Detecting rare but relevant events in systems neuroscience

Dissertation

zur Erlangung des Grades eines
Doktors der Naturwissenschaften

der Mathematisch-Naturwissenschaftlichen Fakultät
und
der Medizinischen Fakultät
der Eberhard-Karls-Universität Tübingen

vorgelegt von

Joachim Bellet
aus Paris, Frankreich

Juni - 2019

Tag der mündlichen Prüfung:	28 Oktober 2019
Dekan der Math.-Nat. Fakultät:	Prof. Dr. W. Rosenstiel
Dekan der Medizinischen Fakultät:	Prof. Dr. I. B. Autenrieth
1. Berichterstatter:	Prof. Dr. Ziad M. Hafed
2. Berichterstatter:	Prof. Dr. Andreas Bartels
3. Berichterstatter:	Prof. Dr. Martin Giese
Prüfungskommission:	Prof. Dr. Ziad M. Hafed
	Prof. Dr. Andreas Bartels
	Prof. Dr. Martin Giese
	Prof. Dr. Uwe Ilg

Erklärung

Ich erkläre, dass ich die zur Promotion eingereichte Arbeit mit dem Titel:

“Detecting rare but relevant events in systems neuroscience”

selbständig verfasst, nur die angegebenen Quellen und Hilfsmittel benutzt und wörtlich oder inhaltlich übernommene Stellen als solche gekennzeichnet habe. Ich versichere an Eides statt, dass diese Angaben wahr sind und dass ich nichts verschwiegen habe. Mir ist bekannt, dass die falsche Abgabe einer Versicherung an Eides statt mit Freiheitsstrafe bis zu drei Jahren oder mit Geldstrafe bestraft wird.

Tübingen, den 28.10.19

Joachim Bellet

Acknowledgement

I want to thank Ziad Hafed for giving me the opportunity to carry out this thesis. It was a great adventure during which I learned a lot, even when experiments where not successful. Ziad Hafed was a great mentor, always available to answer my questions.

I also want to thank my colleagues for their help and their kindness. Thank you Alla Ignashchenkova, Chih-Yang Chen, Xiaoguang Tian, Konstantin-Friedrich Willeke, Antimo Buonocore, Araceli Ramirez-Cardenas, Tatiana Malevich, Fatemeh Khademi, Amarender Bogadhi, and Matthias-Philipp Baumann.

I want to thank my family and my friends for their love and support.

Finally, I want to thank my wife, Marie, for her love, for our child, for all the happiness in my life, and for collaborating with me on this thesis.

Abstract

Animals actively move their sensory organs, often in a rhythmic manner, to gather information from the external environment. The movements performed to sense the world are often very subtle and hard to detect in recording devices. For instance, in the visual domain, eye movements with amplitudes smaller than a degree of visual angle can occur. These tiny movements, called microsaccades, are at the threshold of the resolution of most recording techniques and one could be tempted to ignore them when studying vision. Yet, they might play an important role in visual processing. My thesis shows that microsaccades should not be ignored, that an algorithm can detect them accurately, and that the same algorithm can be used to detect any other seemingly "petty" events that deserve to be detected among noisy signals.

In the first part, we demonstrated that microsaccades have a long-lasting impact on visual processing. We designed behavioral experiments to probe visual detectability and reaction time for stimuli presented at various moments relative to microsaccade onset. By probing the behavioral performance at multiples time points, we could reconstruct a signal that revealed oscillations occurring during visual processing. These oscillations occurred in the beta and alpha range and were synchronized to microsaccade generation. Moreover, the oscillations were sequential, occurring as two pulses, one in each hemifield, depending on the direction of the microsaccade. We also found that microsaccades are associated with a long-lasting increase in contrast sensitivity for stimuli presented in the same hemifield than their direction. These discoveries were important because they demonstrated that visuomotor processing is almost never exempt from the impact of subtle, seemingly irrelevant, movement behaviors. The results therefore established the need for accurate detection of microsaccades and other potentially significant events in brain activity and behavior. We thus designed, in a second study, a deep neural network that performs human-level eye movements detection even in noisy eye traces. Our algorithm outperformed the state-of-the-art algorithm for eye movement detection as well as many commonly used algorithms. In a third study, we finally showed that our algorithm can be generalized to other types of signals by detecting complex spikes in extracellular recordings of cerebellar Purkinje cells. We demonstrated human-level detection of complex spikes, outperforming commonly used online algorithms. Furthermore, our approach also accurately estimated the duration of complex spikes, which provides important information about the coding of error in the cerebellum.

Putting all of the above together, this thesis argues for a careful control of exploratory movements when studying sensory processing. It also provides the tools necessary to approach a problem that is common in many different fields of neuroscience: the detection of an event of interest in a noisy signal.

Content

Abstract *1*

Part I. Scientific background *4*

Introduction **4**

1 Active perception: the example of microsaccades **5**
1.1 Presaccadic compression of space and time 5
1.2 Presaccadic attention 5
1.3 Saccadic suppression 6
1.4 Open question 7

2 Saccades, oscillations, and attention **8**
2.1 Saccades reset brain oscillations 8
2.2 Dense sampling 8
2.3 What is attention? 10
2.4 Microsaccade generation and its relation to attention 10
2.5 Open question 10

3 Microsaccade detection: eye tracking and algorithms **11**
3.1 The scleral search coil 11
3.2 Purkinje-image eye tracker 11
3.3 Engbert and Mergenthaler saccade detection algorithm 13
3.4 Otero-Millan et al. algorithm 14
3.5 Recent state-of-the-art algorithms 14
3.6 Detecting microsaccades during smooth pursuit 14
3.7 Open question 15

4 Microsaccade detection, a generalizable problem **16**
4.1 Detecting larger saccades and post-saccadic oscillations 16
4.2 Detection in other signals: the example of complex spikes 17
4.3 Open question 17

Part II Main results *18*

Statement of contribution **18**

5. Long term impacts of microsaccades on visual processing **19**
5.1 Microsaccades reset oscillations in the alpha and beta range 19
5.2 Microsaccades increase contrast sensitivity in one hemifield 19

6. U'n'Eye a state-of-the-art algorithm for (micro)saccade detection **20**
6.1 Deep neural network for microsaccade detection 20
6.2 State-of-the-art eye movement detection algorithm 20
6.3 Robustness of U'n'Eye 21

7. Generalization to complex spike detection **22**
7.1 Our algorithm detects complex spikes like a human expert 22
7.2 Accurate detection of complex spike duration 22

Part III Discussion *23*

8 Microsaccades and attention **24**

9. Futures research opportunities **26**

References *28*

Study 1: Sequential hemifield gating of α and β-behavioral performance oscillations after microsaccades *35*

Study 2: Human-level saccade detection performance using deep neural networks *52*

Study 3 : Using deep neural networks to detect complex spikes of cerebellar Purkinje Cells *67*

Part I. Scientific background

Introduction

The brain is a complex organ composed of billions of neurons sharing the information that is used to perform its function. Because current techniques are limited, neuroscientists have to study the brain by measuring either a small subsample of its neurons or the signal originating from the pooled activity of many neurons. Either option results in a noisy measurement of brain activity that can rarely be decoded with a single observation. This problem is often bypassed by averaging the signal over many trials. However, such trial averaging method imposes a necessary requirement to control as many experimental parameters as possible, such that trials are "identical". For this reason, to reduce variability between trials, neural processing has been mainly studied by testing the effect of stimuli on anesthetized, immobilized, or passive animals. Yet, a growing body of evidence indicates that endogenous neural events and spontaneous movements play an important role in the computations that the brain performs. Thus, to understand sensory processing, one needs to take into account the active contribution of the organism that senses the world. In this thesis, I will develop the importance of studying active sensing using the example of microsaccades, which are seemingly spontaneous eye movements occurring during fixation. I will then describe the state-of-the-art algorithm that we designed to detect microsaccades in recordings of eye position. Finally, I will show that our algorithm can be used to detect other spontaneous or rare events and therefore be generalized to approach other problems in systems neuroscience, in which spontaneous or rare events are relevant for brain computations. This first section provides the scientific background that motivated this thesis.

1 Active perception: the example of microsaccades

It is easy to assume that perception is a passive phenomenon. For any modality, the overwhelming majority of studies find correlates of perception in the neural activity elicited by stimuli presented to passive animals (Kandel et al., 2000). Yet, a growing body of evidence shows that movement and spontaneous changes in neural activity play a major role in sensory processing (Schroeder et al., 2010). This is the case for taste (Halpern, 1983; Gutierrez et al., 2010), smell (Kepecs et al., 2006), audition (Guinan J. J., 2006; Garinis et al., 2011), touch (Lederman and Klatzky, 1987; Hughes and Jansson, 1994), and vision (Kowler, 2011). Perhaps, the most studied movements impacting visual perception are saccades (Wurtz, 2008; Kowler, 2011). Saccades are ballistic eye movements that shift the gaze from one location to another in ~50 ms. In my thesis, I explored the impact of a special kind of saccade called microsaccade. Microsaccades are saccades occurring 1 to 3 times per second while maintaining fixation at a certain location. They are of special interest because they still occur in experiments where the paradigm imposes fixation in order to prevent any saccades from happening. They therefore represent a violation of the ideal of "identical" trials that experimentalists aspire towards, as stated above. In the following, I will describe the well-known effects of saccades on visual processing and show that microsaccades have similar impacts to those of larger saccades.

1.1 Presaccadic compression of space and time

Starting about 100 ms prior to saccade onset, a phenomenon called saccadic compression occurs. In this phenomenon, the perception of the position of briefly flashed stimuli is shifted towards the endpoint of a saccade (Matin and Pearce, 1965, 1969; Matin et al., 1970; Morrone et al., 1997; Ross et al., 1997). Similarly, the position of receptive fields in the frontal eye fields (FEF) and the lateral intraparietal cortex (LIP) are also shifted during this brief interval (Colby et al., 1996; Colby and Goldberg, 1999; Kusunoki and Goldberg, 2006; Sommer and Wurtz, 2006). This suggests a brief alteration of sensorimotor processing in the brain around the time of saccades. Such alteration also affects aspects of temporal coding, since perceptual experiments suggest that, within the 100 ms preceding a saccade, the time interval separating two stimuli is perceived to be shorter (Morrone et al., 2005). Microsaccades also induce perisaccadic compression of space (Hafed, 2013) as well as time perception (Yu et al., 2017). Figure 1 A. depicts the compression of space prior to microsaccades.

1.2 Presaccadic attention

At the same time at which presaccadic compression occurs, detection performance also increases for targets presented close to the saccade endpoint, depicted in Figure 1 B. (Hoffman and Subramaniam, 1995; Deubel et al., 1996; McPeek et al., 1999; Godijn and Theeuwes, 2003; Gersch et al., 2004, 2008, 2009; Van der Stigchel and Theeuwes, 2005; Baldauf and Deubel, 2008; Rolfs et al., 2011; Rolfs and Carrasco, 2012). This increase in performance is present when the saccade is instructed but also when the saccade occurs spontaneously. Similarly, microsaccades are associated with an increase in visual

sensitivity prior to their onset (Hafed, 2013), and neural correlates of this phenomenon have been found in the superior colliculus (SC) and FEF (Chen et al., 2015).

1.3 Saccadic suppression

When a saccade occurs, the image projected on the retina moves substantially. Yet, we still see the world as stable. For this reason, many researchers have investigated visual processing at the time of saccade occurrence (Binda and Morrone, 2018). The consensus is that contrast sensitivity thresholds increase for stimuli presented during the saccade and even shortly before. This effect most strongly impacts stimuli with low spatial frequencies. Saccadic suppression has also been found with microsaccades, both at the behavioral and neuronal levels (Hafed and Krauzlis, 2010; Chen and Hafed, 2017). Figure 1 C depicts an example of this, showing a decrease in visual sensitivity in visual neurons of the superior colliculus around the time of microsaccades.

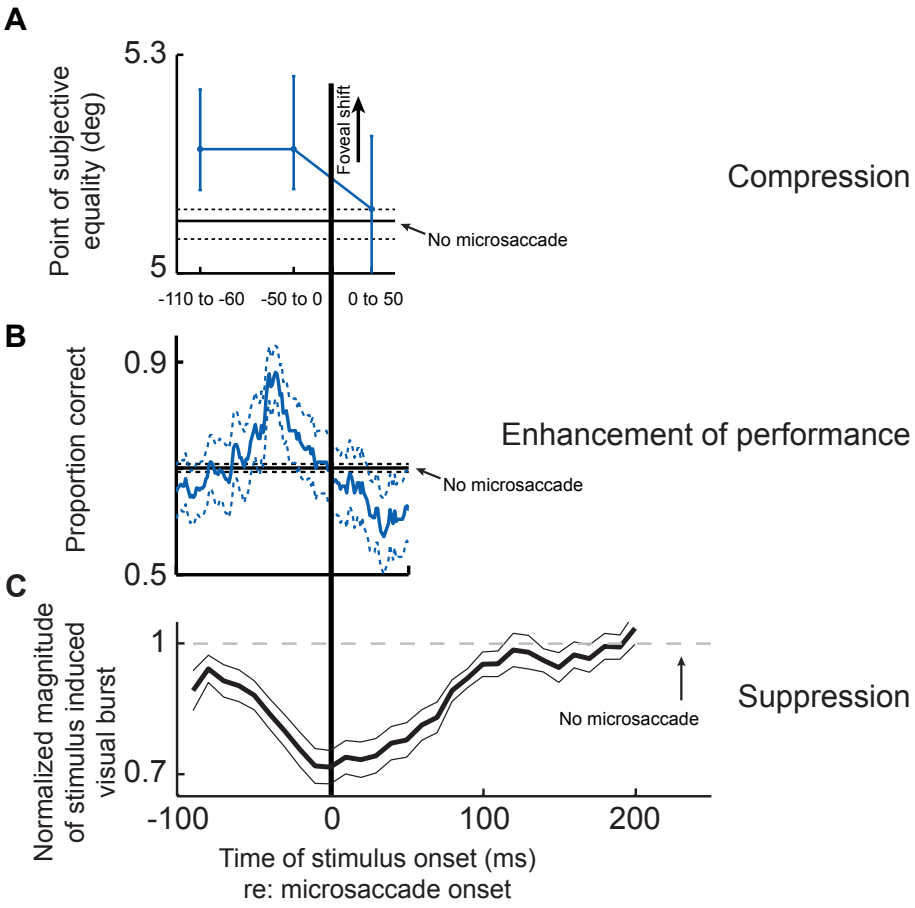


Figure 1. Microsaccade-associated changes in visual processing. **A.** Prior to a microsaccade, the position of a target is estimated closer to the fovea than it really is. **B.** 50 ms prior to the microsaccade, visual performance is increased relative to baseline. **C.** Around the time of microsaccade onset, the burst evoked by a visual stimulus in the superior colliculus is reduced relative to the burst for an identical stimulus presented when no microsaccade occurred. A and B are adapted from (Hafed, 2013) and C is adapted from (Hafed and Krauzlis, 2010)

1.4 Open question

Microsaccades are associated with alteration of visual processing close to the time of their generation. It is, however, unknown whether they impact vision for the remaining period that separates them. If so, this will bring further evidence that vision is an active process even when the eye no longer moves, and it would also provide clear justification for developing algorithms like those presented in this thesis for the service of systems neuroscience.

2 Saccades, oscillations, and attention

As seen above, many experiments have investigated the influences of saccades on the processing of stimuli presented close to the time of their generation. It is therefore commonly believed that saccadic influence on visual processing is short-lived. Yet, many electrophysiological fluctuations lasting several hundreds of milliseconds have been found to be reset by saccades. These fluctuations happen in frequency bands that are also linked to attention. In this section, I will therefore introduce the relationships between saccades, oscillations, and attention.

2.1 Saccades reset brain oscillations

Saccades reset oscillatory activities at various frequency bands in the visual system and hippocampus (Bartlett et al., 2011; Ito et al., 2011b, 2011a; Jutras et al., 2013; Zanos et al., 2015). Even microsaccades reset alpha oscillations (~10 Hz) in the occipital EEG (Gaarder et al., 1964) and induce gamma synchrony in the LFP of macaque early visual cortex (Bosman et al., 2009; Lowet et al., 2015, 2016, 2018a). However, the fluctuations observed in LFP and EEG might be originating from distant areas in the brain and be observed in the visual areas only because of volume conduction (Cohen, 2017). Even if they are originating from visual areas, it is not known whether they have a functional impact on visual processing and behavior. In this thesis, I make use of a technique called dense sampling to show a causal link between oscillations and behavior.

2.2 Dense sampling

Dense sampling consists in probing behavioral performance (e.g. hit rate or reaction time) for sensory targets presented at variable times from a resetting event. A running average of the performance is then computed and a signal is obtained. This signal is then treated as an electrophysiological signal to test for the presence of oscillatory patterns compared to surrogate signals from permutations (Fig. 2). Dense sampling has been recently used to reveal periodicity in behavioral responses after a resetting by an exogenous cue (Fiebelkorn et al., 2011, 2013; Landau and Fries, 2012; Song et al., 2014; Drewes et al., 2015; Dugué et al., 2016; Re et al., 2019). Most studies using dense sampling interpret the observed fluctuations in behavior as a consequence of a rhythmic alternation in the locus of a “spotlight” of attention.

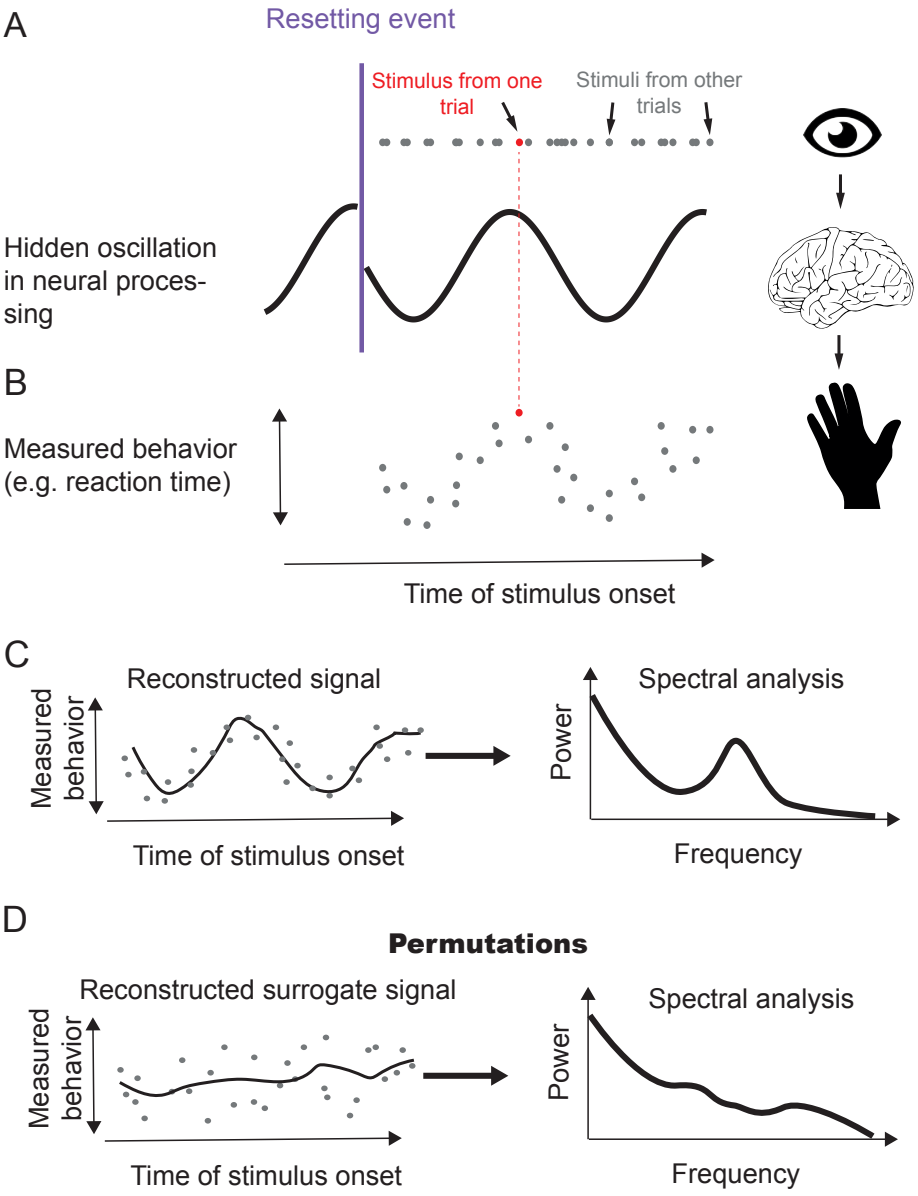


Figure 2: The method of dense sampling and its steps. **A.** A sensory stimulus is probed at various times relative to an event that is hypothesized to reset neural oscillations. **B.** The neural oscillation affects sensory processing such that the behavioral response to the same stimulus depends on the time of stimulus presentation relative to the resetting event. **C**. Left panel: a signal is generated from the many data points as a time course of behavioral performance. Right Panel: the signal is analyzed in the frequency domain (shown) or in the time-frequency domain. **D.** Statistical tests are obtained by shuffling the original data points in time, in order to obtain a surrogate signal and test the probability of observing a peak in power as strong as that observed in the original dataset. Such probability is estimated after making the shuffling operation a great number of times.

2.3 What is attention?

According to Willam James: “Everyone knows what attention is. It is the taking possession by the mind, in clear and vivid form, of one out of what seem several simultaneously possible objects or trains of thought” (James, 1890). In reality, experimenters have had a very hard time making a consensus on the definition of attention (Fernandez-Duque and Johnson, 2002). A majority of scientists consider attention as a “selection operation” occurring at the level of sensory processing; that is, certain stimuli may be selectively processed at the expense of other, simultaneously present, stimuli. Neurobiologically, such selectivity may appear as a change in sensory activity to a given stimulus, even if the stimulus attributes are themselves physically unaltered. For example, some researchers call “attention” any change in neural activity correlated with the instruction to focus on a certain aspect of the sensory environment (Reynolds et al., 2000; Fries et al., 2001). Since a change in neural activity (e.g. increase in firing rate) is not always a synonym of better sensory processing, other researchers prefer to consider only an increase in behavioral performance as a marker for attention. In recent years, an increasing body of literature using dense sampling or studying behavioral performance as a function of electrophysiological measurements suggests that attention is an oscillating process (Buschman and Kastner, 2015). In the next section, I will describe why microsaccades might be a confounding factor when studying oscillatory fluctuations of attention.

2.4 Microsaccade generation and its relation to attention

The overwhelming majority of experiments investigating visual attention uses exogenous cues or central arrows to orient attention (Posner, 1980). Yet, exogenous stimuli as well as asymmetric central cues impact both the probability of microsaccade occurrence and their direction in a time dependent manner (Hafed and Clark, 2002; Engbert and Kliegl, 2003; Laubrock et al., 2005; Pastukhov and Braun, 2010; Hafed and Ignashchenkova, 2013; Pastukhov et al., 2013). This property, combined with the effect of microsaccades on visual processing mentioned earlier, might be a confounding factor when studying attention (Hafed, 2013; Tian et al., 2016). Microsaccades also tend to occur rhythmically (Bosman et al., 2009) in a self-paced manner (Amit et al., 2017) and might, therefore, drive the oscillations observed in the experiments mentioned in the previous section.

2.5 Open question

Covert attention is an actively studied topic relying on experiments in which subjects perform fixation. Yet, microsaccade direction and frequency are systematically affected by the cue onset in such experiments and could be a confounding factor (given the active perception literature alluded to above in Section 1). Would it be sufficient to control for the effects of microsaccades by excluding trials in which the tested stimulus occurred only close to the time of a microsaccade?

3 Microsaccade detection: eye tracking and algorithms

In the previous sections, I provided evidence that a seemingly spontaneously generated event, the microsaccade, is important for visual processing. Yet, many studies ignore this type of eye movement in analyses. One reason for ignoring microsaccades is that they are hard to detect because their amplitude can be close to the noise of the recording equipment for eye tracking. I will present now the two recording techniques that are most commonly used to detect microsaccades. I will also review the popular and state-of-the-art algorithms for microsaccade detection, paving the way for developing my own state-of-the-art benchmark algorithm.

3.1 The scleral search coil

The scleral search coil is considered to be the most accurate technique for measuring eye position. It consists of recording the current induced by a uniform magnetic field on a coil made of thin metal wire physically attached to the eye sclera (Robinson, 1963). By rapidly alternating the strength of the magnetic field along a given axis, one can induce currents in the implanted wire loop, and the amount of induced current will depend in a lawful manner on the angle of the eye (i.e. the angle of the wire loop) relative to the axis of the magnetic field (through well-known physical principles of electromagnetic induction). If the magnetic field is now generated across multiple axes, then changes in different eye movement axes, such horizontal and vertical eye position and also torsional rotations, can all be detected from the same implanted wire loop. This method can provide a very fine estimation of gaze position, down to a resolution 1 min of arc (approximately the spacing between individual foveal photoreceptors). However, the search coil technique is hard to set up. Used in non-human animals, one needs to surgically implant the wire in the sclera. For use in humans, a thick and uncomfortable lens holding the wire has to be placed onto the eye. In this case, the eye often needs to be anesthetized to prevent discomfort, and the experiment has to be performed in the presence of a physician (McCamy et al., 2015). However, because of its superior spatial resolution (Fig. 3), the search coil is the technique of choice for investigating microsaccades, particularly in animal models.

3.2 Purkinje-image eye tracker

The most common way to record eye movements is to illuminate the eye with an infrared source of light and to record the position of the Purkinje images reflected by the various optical elements of the eye (Collewijn, 1999). Purkinje images are reflections of the infrared light formed as the light ray crosses different tissues in its path. The relative position between the first Purkinje reflection, which is the reflection from the front surface of the cornea, and the center of the pupil can be used to estimate the position of gaze, and infrared lights are used because camera sensors can detect the infrared reflections without visually disturbing the subject with a very bright illuminant. This technique can be accurate enough to detect microsaccades but suffers from several sources of noise, like the change in pupil diameter and the asperity of the corneal surface (Nyström et al. 2016). Figure 3 shows an example of an eye trace recorded both with such an eye-tracker in combination with a scleral search coil in a macaque monkey. Some devices, called dual Purkinje eye-

trackers, measure the relative position of the first and the fourth Purkinje reflection. The fourth Purkinje reflection comes from the light ray reflecting at the rear of the eye and refracting at the posterior part of the lens. This technique suffers from less noise as it is not affected by changes in the size of the pupil. However, the dual Purkinje eye tracker requires the constant mechanical alignment of mirrors to the fourth Purkinje reflection by a servomechanism. This makes the system cumbersome to calibrate and use, and most laboratories avoid this eye tracker. Blinks, fast eye movements, or eccentric eye positions often also interrupt the acquisition of the signal.

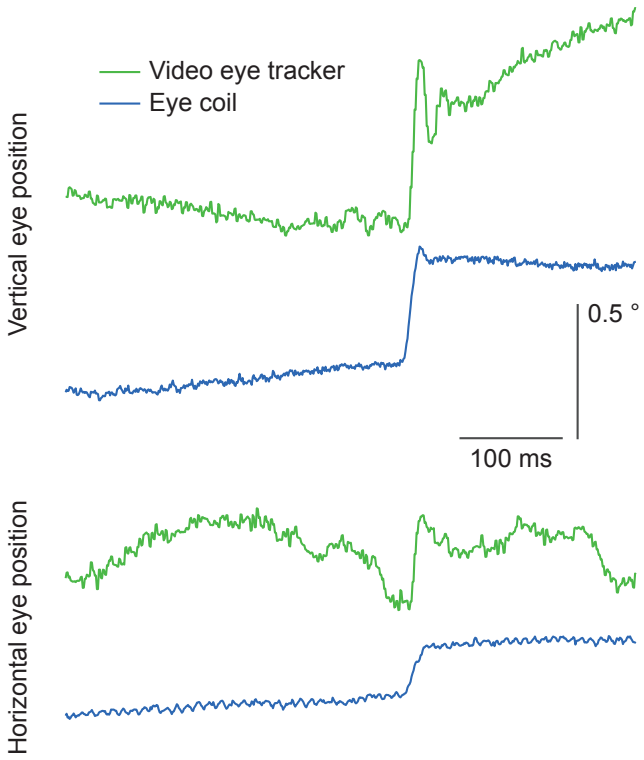


Figure 3: Simultaneous recording of eye position with a scleral search coil and a video eye tracker. The eye tracker uses the first Purkinje image and an estimate of the center of the pupil to calculate gaze position. This results in a noisier estimate of the eye position compared to the eye coil (greater high frequency wiggle in the green traces than in the blue traces). Moreover, the video eye tracker shows substantial variation in eye position before and after the saccade, which is not present in the coil data (e.g. the strong drift lasting >100 ms after the saccade in the upper most trace; it is not present in the blue trace recorded simultaneously with the coil). This is due to artifacts caused by pupil diameter changes in the video eye tracker. Note also that the topmost trace shows a short-lived postsaccadic oscillation in eye position. This is likely due to oscillations of the anterior parts of the eye (Nyström et al., 2013; Bouzat et al., 2018), which the video eye tracker detects. The eye coil does not reveal strong postsaccadic oscillation because it only tracks the position of the eyeball.

3.3 Engbert and Mergenthaler saccade detection algorithm

The most common way to automatically detect saccades is to set a threshold on the instantaneous radial velocity from an eye position trace (Fig. 4). If eye velocity exceeds the threshold, the time point is flagged as being part of a saccade. Because microsaccades are so small, their eye velocities are also low. Therefore, the velocity threshold needs to be as close as possible to the noise of the eye-tracker such that even the smallest microsaccade is detected. Engbert and Mergenthaler (Engbert and Mergenthaler, 2006) proposed a solution that consists of setting the threshold as a multiple of the median radial velocity over an entire trial. Since saccades are rare events, the median is only influenced by the noise in the eye tracker. This makes the threshold flexible to changes in background noise and increases the accuracy of the detection algorithm. Engbert and Mergenthaler's algorithm needs to set a single parameter which is the factor with which the median velocity is multiplied to obtain a threshold. The choice of this parameter can change the performance of the algorithm, leading either to too many missed events (high threshold) or too many false alarms (low threshold) (Fig. 4).

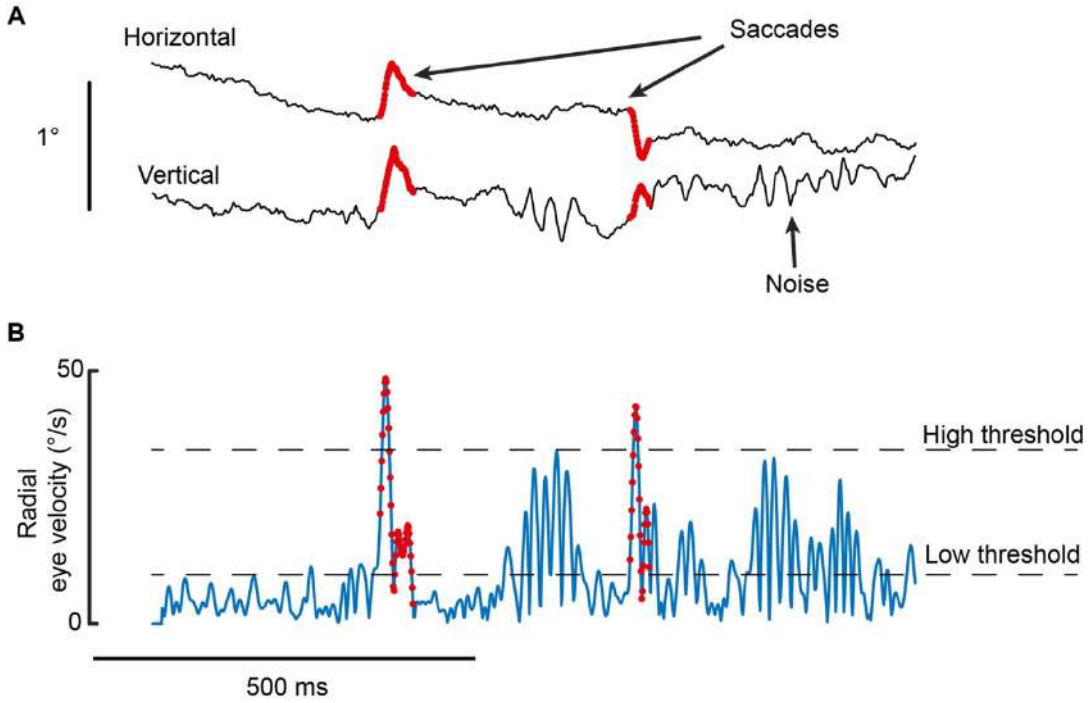


Figure 4: Detection of saccades using a velocity threshold. **A**. Horizontal and vertical eye position recorded with a video-based eye tracker. Two small saccades have been manually labeled. **B.** Corresponding radial eye velocity. The saccades induce peaks in eye velocity and can therefore be detected using a threshold on the eye velocity. A high threshold can avoid labeling the noise as saccades but will underestimate the duration of saccades and might miss smaller saccades. Instead, a lower threshold would miss less saccades but would result in more false positives.

3.4 Otero-Millan et al. algorithm

To avoid the choice of arbitrary parameters, Otero-Millan and colleagues (Otero-Millan et al., 2014) developed an unsupervised algorithm for microsaccade detection. They essentially used the same Engbert and Mergenthaler algorithm, but they added a means to effectively choose the velocity threshold in a different manner. In their approach, the threshold is set such that an average of 5 events per second is detected, which is a bit higher than the normal rate of microsaccades. This first step will select microsaccades and some events emerging from the noise. The rationale is then that microsaccades will be separable from the noise in a space composed of the peak velocity, the first acceleration peak, and the last acceleration peak. The microsaccades are then selected from the noise by k-mean clustering of the detected events in 3-dimensional space.

3.5 Recent state-of-the-art algorithms

More recently, other unsupervised algorithms have been developed that are not commonly used yet. However, as they have shown better performance than the algorithms by Engbert and Mergenthaler and Otero-Millan et al., they might become popular in the near future. Mihali and colleagues (Mihali et al., 2017) developed a Bayesian approach for microsaccade detection based on a semi-Markov model that outputs a probability for each time point to belong to a microsaccade or a drift. The approach developed by Mihali and colleagues outperforms the commonly used algorithms discussed earlier. Finally, Sheynikhovich and colleagues (Sheynikhovich et al., 2018) developed another method that first selects events using the Engbert and Mergenthaler algorithm with a low threshold parameter, and then separates the microsaccade from the noise by unsupervised clustering. The clustering is performed on the absolute velocity of the horizontal and vertical component of the eye trace. The Sheynikhovich algorithm outperforms all algorithms mentioned earlier.

3.6 Detecting microsaccades during smooth pursuit

The detection of microsaccades is even more challenging when the target to fixate is moving. Under such a condition, the eye engages in a movement called smooth pursuit; this means that there is a sustained period of elevated instantaneous eye velocity above the baseline fixation velocity. The microsaccades occurring during smooth pursuit often compensate for the offset between the target position and the eye position and are thus called catch-up saccades. Catch-up saccades are particularly difficult to detect because the smooth pursuit is already a high velocity eye movement and thus reduces the signal gain by microsaccades. Moreover, saccades going into the opposite direction actually reduce the eye velocity. Therefore, a simple velocity threshold would miss these movements entirely. To address this issue, Daye and Optican (Daye and Optican, 2014) developed an algorithm based on a particle filter that cancels out the slow component of eye movement to let saccades emerge in the signal.

3.7 Open question

When manually analyzing the presence of microsaccades in eye traces, it is striking how many mistake automatic algorithms still make. Is it possible to design a better algorithm that labels microsaccades in the same way that a human would do it?

4 Microsaccade detection, a generalizable problem

To correctly detect microsaccades, one would need to implement an algorithm knowing all the rules that discriminate a microsaccade from any noise that the eye-trace could contain. Expert users who label microsaccades manually exploit all such rules, without necessarily verbalizing them explicitly. Therefore, there are two problems in translating expect human performance to machine algorithms: on the one hand, the rules are hard to implement programmatically; on the other hand, we use them without noticing when manually labeling eye traces. Supervised machine learning approaches have often proved to be helpful in such kind of situation (Theis et al., 2016; Mathis et al., 2018). Moreover, machine learning approaches are flexible, such that different rules can be learned by the same method once new training evidence can be provided. In this thesis, I describe an algorithm, based on convolutional neural networks, that proved its efficacy for detecting not only microsaccades but also other types of eye movements and even an electrophysiological event called complex spike.

4.1 Detecting larger saccades and post-saccadic oscillations

In the previous section, I outlined that the detection of microsaccades is a challenging problem. One might think that the detection of larger saccades is not difficult because their velocity reaches higher peaks relative to the background noise. Yet, in the case of saccades, the onset and offset of the movement have velocities close to the noise of the eye-tracker since the eye accelerates from and decelerates towards fixation at the beginning and end of a saccade. Therefore, it is still a challenge is to detect the timing of large saccades accurately. This is of great importance because many psychophysics and electrophysiology experiments analyze their data relative to onsets and offsets of saccades. For instance, saccade onset can be used to accurately measure saccadic reaction time. Saccade offset, on the other hand, is the time at which a novel stable image is projected on the retina. A great number of studies aimed at developing algorithms for saccade detection (Andersson et al., 2017).

At the end of the saccade, the inertia of the movement makes the pupil move relative to the iris (Nyström et al., 2013; Bouzat et al., 2018). This results in so-called post saccadic oscillations (PSO) that last about 30 ms (see figure 3 where PSOs are visible in the video-based eye tracker and not the search coil tracker). PSOs are challenging to detect because their velocity is close to that of the saccade. PSOs might be important for visual processing and have recently become an active topic of research (Tabernero and Artal, 2014; Hooge et al., 2015).

Larsson and colleagues provided a benchmark dataset (with a video eye tracker) that contains saccades and PSOs labeled by human experts (Larsson et al., 2013). This benchmark dataset includes eye movements performed while viewing static pictures or moving objects. It is used by any study attempting to compare the performance of their algorithm to previously published algorithms.

4.2 Detection in other signals: the example of complex spikes

The problem of accurately detecting the onsets and offsets of events in a signal is not specific to eye movements. Similarly, the movement of other body parts or the spontaneous emergence of events in electrophysiological signals can be of interest in systems neuroscience (e.g. hand movement, ponto-geniculo-occipital waves, hippocampal ripples, etc.). In this thesis, we decided to tackle the problem of complex spike detection in extracellular recordings of cerebellar Purkinje cells. Complex spikes are high frequency events occurring upon excitation of the dendrites of Purkinje cells by the climbing fibers from the inferior olive. The shape of complex spikes differs from one cell to another, which makes it challenging to design an automatic algorithm for complex spike detection. Even within one cell, the duration of a complex spike can vary, and this variation has recently been suggested to encode information (Yang and Lisberger, 2014; Herzfeld et al., 2015, 2018; Junker et al., 2018). However, in all of these studies, the duration of complex spikes was detected manually requiring months of work.

4.3 Open question

Given that the problem of detecting microsaccades is encountered in other fields of neuroscience for other events and other signals, is it possible to design a single algorithm that would be used for the detection of any such event?

Part II Main results

Statement of contribution

This thesis comprises 3 publications, which are summarized in the following sections. The individual publications and manuscripts can be found in the annexes.

1. Bellet, J., Chen, C.Y. and Hafed, Z.M., 2017. Sequential hemifield gating of α- and β-behavioral performance oscillations after microsaccades. *Journal of neurophysiology*, *118*(5), pp.2789-2805.

I designed the project together with Z.M. Hafed. I performed the behavioral experiments with the human subjects as well as with one monkey. I analyzed the human behavioral data. I edited the paper together with Z.M. Hafed.

2. Bellet, M.E.*, **Bellet, J.***, Nienborg, H., Hafed, Z.M. and Berens, P., 2018. Human-level saccade detection performance using deep neural networks. *Journal of neurophysiology*, *121*(2), pp.646-661. *(* shared first authorship)*

I designed the study together with M.E. Bellet, Z.M. Hafed and P. Berens. I implemented parts of the saccade detection algorithm. I analyzed the data together with M.E. Bellet. I labeled manually two datasets provided in the paper. I wrote the paper together with M.E. Bellet, Z.M. Hafed and P. Berens.

3. Markanday, A.*, **Bellet, J.***, Bellet, M.E.*, Hafed, Z.M. and Thier, P., 2019. Using deep neural networks to detect complex spikes of cerebellar Purkinje Cells. (under revision, * *shared first authorship*)

I designed the study together with A. Markanday, M.E. Bellet, Z.M. Hafed and P. Thier. I designed the post-processing steps to refine the output of the algorithm. I analyzed the data. I wrote the paper together with A. Markanday, M.E. Bellet, Z.M. Hafed and P. Thier.

5. Long term impacts of microsaccades on visual processing

As I described in section 1, it is well established that microsaccades influence visual processing for stimuli presented close to the time of their appearance. Yet, electrophysiological experiments suggest that some oscillatory activities are phase-locked to microsaccade generation. In my first study, I showed that microsaccades leave an imprint on visual processing that lasts several hundreds of milliseconds after their occurrence. This imprint is dependent on the direction of microsaccades.

5.1 Microsaccades reset oscillations in the alpha and beta range

5.1.1 Global effect

Saccadic reaction time is a reliable measure of the efficiency of visual processing (Hafed and Krauzlis, 2010; Chen et al., 2018). Using the method of dense sampling (described in section 2.2), we measured saccadic reaction times of 10 human subjects for targets presented at various times relative to microsaccade onset. Microsaccades occurred spontaneously during a long fixation period and were not reset by any visual transient. Under these conditions, we observed oscillations in the signal obtained from the reaction times in the alpha/beta range that lasted for ~500 ms after the onset of microsaccades. By visual inspection of every single eye trace, I took care that no microsaccade or blink occurred between the measured microsaccade and the target stimulus appearance. Thus, this effect is not explained by the rhythmic nature of microsaccade generation, which lies anyway in a lower frequency range.

5.1.2 Hemifield specific effect

To dissect this phenomenon, we performed the same analysis as described above but independently for targets presented in the same hemifield or in the opposite hemifield relative to the microsaccade direction. This revealed that the oscillatory activity is composed of two pulses occurring successively in each hemifield. First, from ~100 ms to ~400 ms, a pulse in the beta range (13-20 Hz) occurs in the same hemifield relative to the microsaccade direction. Then, from ~400 ms to ~600 a pulse in the alpha range (8-13 Hz) occurred in the opposite hemifield relative to the saccade direction.

5.2 Microsaccades increase contrast sensitivity in one hemifield

Following the same logic as in the first experiment, we tested the probability to detect low contrast stimuli presented at different times relative to microsaccade onset. We found that in the period between 100 ms and 400 ms following a microsaccade onset, the target was better perceived in the same hemifield as the microsaccade direction. This effect occurs exactly at the same period and in the same hemifield as the beta oscillations revealed by the first experiment.

6. U'n'Eye a state-of-the-art algorithm for (micro)saccade detection

We demonstrated in the first study that microsaccades impact vision for an extended period of time after their occurrence. Microsaccade direction lateralizes the effect they have on visual processing. Because of their systematic alteration by cueing, microsaccades are even more likely to be a confounding factor in attention experiments. Thus, the need for detecting microsaccades in eye traces becomes more important. Yet, carefully detecting microsaccades requires tedious manual labeling as no algorithm to date is fully satisfactory.

6.1 Deep neural network for microsaccade detection

We aimed at designing a machine learning algorithm that would detect microsaccades in eye traces like a human would do. Our constraint was to reach a performance that is at the level of a human expert while minimizing the number of training samples needed to train the algorithm. We developed a convolutional neural network called U'n'Eye inspired by the architecture of the U-Net image segmentation algorithm (Ronneberger et al., 2015). U'n'Eye is adapted for processing signals with few training examples. To this end, we designed it with fewer layers than the original architecture and included batch normalization operations that were not present in U-Net.

U'n'Eye was tested on three different datasets of eye traces, recorded with video eye trackers or search coils, including microsaccades occurring during fixation or smooth pursuit. All dataset where entirely annotated by human experts to train and evaluate the performance of U'n'Eye. For every dataset, U'n'Eye outperformed every other algorithm that we tested. This included the classical algorithm from Engbert and Mergenthaler (Engbert and Mergenthaler, 2006), the unsupervised algorithm from Otero-Millan and colleagues (Otero-Millan et al., 2014), the last state-of-the-art algorithm for microsaccade detection from Sheynikhovich and colleagues (Sheynikhovich et al., 2018), and the most recent state-of-the-art algorithm for saccade detection from Pekkanen and Lappi (Pekkanen and Lappi, 2017). In fact, U'n'Eye detected microsaccades with a performance that was at the level of the agreement between two human experts.

6.2 State-of-the-art eye movement detection algorithm

To further demonstrate the capacities of our algorithm, we also tested its performance on a benchmark dataset for eye movement classification (Larsson et al., 2013) described in section 4.2. U'n'Eye can be trained to classify more than two categories and we, therefore, tested its performance in detecting saccades, post-saccadic oscillations (PSO), and blinks. U'n'Eye outperformed the previous state-of-the-art algorithm designed for the detection of saccades and PSOs. It also labeled blinks accurately although no algorithm could be compared to U'n'Eye for this class of eye movement.

6.3 Robustness of U’n’Eye

We also showed that our algorithm can reach state-of-the-art performance in microsaccade detection with only 50 seconds of eye traces as training samples. It can accurately detect saccades even if some labels are missing in the training set. In a large cohort study, U’n’Eye can be trained with eye trace samples from a few subjects and generalize to detect eye movements from other new subjects.

7. Generalization to complex spike detection

As we found that a single convolutional neural network can be used for the detection of several types of eye movements, we considered extending its use to other types of events in other kinds of signals. We choose to approach the problem of complex spike detection in extracellular recordings of Purkinje cells in the cerebellum.

7.1 Our algorithm detects complex spikes like a human expert

Human experts in complex spike detection use the local field potential trace together with the high-passed signal trace to spot complex spikes. We thus decided to use these signals as input to our network. We also increased the size of the convolutional and max-pooling kernels so that the span of the signal taken into consideration by the network to estimate a time bin was equivalent to what a human expert uses. Finally, because complex spikes have a similar waveform within one cell, we designed a post-processing step than refined the outputs of the network by selecting only events that look alike within one recording. Our algorithm was compared to the standard automatic approach that is online detection by manual selection of bounding boxes to select the complex spike waveform. Our algorithm outperforms, by far, the online sorting approach and detects complex spikes like a human expert would do. The disagreements between our algorithm and the human expert labeling are negligible and more often are because of mistakes from the human expert.

7.2 Accurate detection of complex spike duration

The duration of complex spikes has been suggested to encode the strength of motor learning in the cerebellum (Yang and Lisberger, 2014). To date, measuring the duration of complex spikes has always been done manually since no algorithm has been developed for this purpose (Herzfeld et al., 2015, 2018; Junker et al., 2018). We compared the estimation of complex spike duration by our algorithm with that of a human expert. We found an error in estimation that was typically below 1 millisecond. Moreover, our algorithm tracked the changes in complex spike duration within one cell remarkably well. Thus, our algorithm is suited for studying the code that the duration of complex spikes might carry and will greatly speed up research in this domain.

Part III Discussion

Our studies reveal the necessity and feasibility to detect rare events in different signals used in systems neuroscience. We showed that microsaccades impact vision in a stereotypical manner for several hundreds of milliseconds after their occurrence. This observation revives the debate about the potential influence of microsaccades on the interpretation of many experiments. Therefore, it is all the more necessary to detect microsaccades, and the algorithm that we developed transforms this tedious task into an easy routine check. The tool we provide is versatile such that it can approach other problems where a certain event needs to be detected in a noisy signal. I will now discuss how our results articulate with recent scientific discoveries.

8 Microsaccades and attention

In the first study that I described in section 5, microsaccades occurring in steady-state were found to be associated with rhythmic fluctuations and a long-lasting increase in contrast sensitivity. Because these features are hallmarks of the presence of attention, we will discuss two questions that emerge from these findings.

8.1 Microsaccades and attention: a chicken and egg problem?

On the one hand, one might interpret that microsaccades trigger a lateralized increase in visual performance that would explain shifts in so-called attention observed in cuing experiments. On the other hand, one might believe that spontaneous changes in attention both trigger microsaccades and the increase in visual sensitivity observed in our experiments. I will review a few recent studies that tried to solve this chicken and egg problem.

In a recent paper, the team of Robert Desimone (Lowet et al., 2018b) studied a correlate of attention in areas V4 and IT of the macaque monkey. The correlate is the increase of firing rate for neurons whose receptive fields are at the cued location. They revealed that the increase in firing rate occurred only after the occurrence of a microsaccade in the direction of the receptive field. Moreover, the identity of the object presented in the receptive field can be more readily decoded if a microsaccade directed toward the receptive field is generated before the stimulus presentation. These results suggest that sensory processing is boosted in a cueing task only after a microsaccade has occurred. The author interpretation of the result is that attention triggers the microsaccade, which in turn cause the changes in visual processing. However, the cue used in this study is perfect to attract microsaccades reflexively (Meyberg et al., 2017). Thus, it is more parsimonious to interpret the result without invoking the involvement of attention at all.

Another way to disentangle this chicken and egg problem would be to prove that attention can occur when microsaccades are absent. Poletti and colleagues (Poletti et al., 2017) performed a series of tasks probing the change of performance induced by a cue indicating the most likely position for a stimulus to occur. Under these conditions, the manual reaction time is faster and the sensitivity is higher when the cue is valid than when the cue is invalid or neutral. Using a high-resolution dual Purkinje eye tracker and controlling that no microsaccade occurred during the cue-target interval, they ruled out the involvement of microsaccades in the effect they observed. However, it is also contestable whether the changes in performance measured in this study are caused by a boost in sensory processing. It could be argued, on the contrary, that invalid and neutral cues are distractors that impair performance relative to valid cues. An important control, ignored in this study, would have been a condition where no cue is presented prior to the target appearance. Moreover, the fact that no microsaccade occurred in their specific measurement interval does not eliminate the fact that prior (or upcoming) microsaccades still have an influence on perceptual performance.

8.2 Rhythmic sampling of attention without microsaccade?

I will now describe three recent studies that considered the impact of microsaccades in their results, but ruled it out. A thorough look at their methods reveals that the question is, in fact, still very much open.

Landau and colleagues (Landau et al., 2015) measured by magnetoencephalography the phase-amplitude coupling of gamma oscillations nested in a theta cycle occurring in antiphase between hemifields. They observed that performance in detecting a target is correlated with the phase of this oscillatory activity. As microsaccades tend to occur rhythmically in the theta band, it could be that their results are a byproduct of spontaneous microsaccade generation. They dismiss this possibility by displaying the position of the eye relative to the target for the "hit" and "miss" conditions. As the eye position is not significantly diverging for the two conditions, they conclude that microsaccades are not involved in the observed effect. However, our study shows that the position of the eye is not a good predictor of the performance, whereas the direction of the microsaccade is. Averaging the eye position across trials is thus not a good control for detecting microsaccades.

In another study, Spyropoulos and colleagues (Spyropoulos et al., 2018) observed with electrocorticogramms over the areas V1 and V4 of two monkeys that the cued location is associated with a decrease in theta activity. They observed microsaccades correlated with the theta activity but that the effect remained when excluding microsaccades. Yet, the detection of microsaccades was done by selecting events crossing the mean velocity of the eye traces plus 5 standard deviations. This is a very conservative threshold. While this custom-made heuristic might spot the biggest of the microsaccades, it does not ensure that the totality of microsaccades is detected.

Fiebelkorn and colleagues (Fiebelkorn et al., 2018) observed a variety of oscillatory activities in the frontal eye field and the lateral intraparietal area that correlate with visual performance. Here also microsaccades are claimed to play no role in the effect they observe: "removing trials with microsaccades did not substantially change the pattern of the results or alter the statistical significance". However, no description of the method for microsaccade detection is provided, and we can assume that the authors did not go through every eye trace to detect microsaccades manually.

The conclusion of the three studies described above suffice for Helfrich and colleagues to conclude that "several studies have shown that rhythmic attentional sampling is not a microsaccade artifact" (Helfrich et al., 2018). However, none of these studies used methods that are sensitive enough to detect all microsaccades and thus, such claim cannot be made. We hope that the scientific community will approach the problem more rigorously in future studies perhaps with the help of U'n'Eye.

9. Futures research opportunities

The algorithms we developed revealed an easy and efficient way to detect events in noisy signals. Although they were only tested for detecting eye movements and complex spikes, they could provide great help in many other fields. I will now discuss other problems that are similar to the one approached in this thesis and that could benefit from using our approach. The list is not exhaustive but still reveals how general this problem is.

9.1 Detecting other movements

It is now evident that visual processing is organized around the transient occurrence of saccades. Yet, movements have also been shown to affect sensory processing in other modalities. (Halpern, 1983; Lederman and Klatzky, 1987; Hughes and Jansson, 1994; Guinan J. J., 2006; Kepecs et al., 2006; Gutierrez et al., 2010; Garinis et al., 2011). Studying active sensing is challenging because one needs first to acquire an accurate measurement of the movement involved in sensory sampling. Many recording devices have been developed for recording such movements. In olfaction, for instance, breathing patterns can be acquired using thermocouples that sense change in the temperature of the air flowing in or out of the nostrils (Kepecs et al., 2006). In hearing, the endogenously generated movement of the basilar membrane can be detected by recording the sound that the eardrum produces in reaction to this movement (Gruters et al., 2018). In the tactile domain, one can record movements using electromyograms (Reaz et al., 2006), magnetic coils (Markanday et al., 2018) or video tracking (Mathis et al., 2018). The signal resulting from such measurements is often noisy and detecting movements could be made easy using our algorithm.

9.2 Detecting transient electrophysiological events

Our neural network could also be used for detecting a multitude of electrophysiological events originating in the brain.

Sleep, alone, hosts a collection of spontaneously generated events that are challenging to spot: K-complex, spindles, ponto-geniculo-occipital waves, and hippocampal ripples. These events can be used for scoring the different stages of sleep and are associated with different types of memory consolidation (Diekelmann and Born, 2010). Therefore, their detection can have implication for fundamental and clinical research. The K-complex is a biphasic EEG fluctuation that occurs spontaneously or upon sensory stimulation during phase II of NREM sleep (Colrain, 2005). Sleep spindles also occur during phase II of NREM sleep (De Gennaro and Ferrara, 2003). Spindles are oscillatory events (12-12 Hz) generated in the thalamic reticular nucleus and propagating as a traveling wave through the cortex. Ponto-geniculo-occipital waves are generated during REM sleep in the pontine nucleus and propagate through the lateral geniculate nucleus to the occipital cortex (McCarley et al., 1978). Hippocampal ripples are high-frequency events occurring mostly during NREM sleep. They are causally involved in memory consolidation (Girardeau and Zugaro, 2011). The detection of K-complex, spindles, ponto-geniculo-occipital waves, and ripples are challenging problems (Bremer et al., 1970; Jansen, 1990; Li et al., 2003; Sethi and Kemere, 2014; Warby et al., 2014; Lajnef et al., 2015).

The brain can also generate signals that are important to detect in pathological conditions. For instance, spike and wave discharges occur in the same frequency band as sleep spindle (Beenhakker and Huguenard, 2009; Leresche et al., 2012) but are a hallmark of epileptic seizures. Accurately detecting their occurrence and duration is a necessary step when testing the effect of therapies (Venzi et al., 2016) and when trying to understand their etiology (Polack et al., 2007). The detection of spike and wave discharge is a hard and active field of research (Van Hese et al., 2009; Ovchinnikov et al., 2010; van Luijtelaar et al., 2016; Maksimenko et al., 2017). Even in the periods between seizures, one can detect events, called either "fast ripples" or "high-frequency oscillations", that predict the intensity of seizures. The detection of these events is also difficult, and a plethora of studies attempt to address this problem (Gardner et al., 2007; Bénar et al., 2010; Birot et al., 2013; Chaibi et al., 2013; López-Cuevas et al., 2013; Burnos et al., 2014; Amiri et al., 2016; Fedele et al., 2016).

References

Amiri M, Lina JM, Pizzo F, Gotman J (2016) High Frequency Oscillations and spikes: Separating real HFOs from false oscillations. Clin Neurophysiol 127:187–196 Available at: http://dx.doi.org/10.1016/j.clinph.2015.04.290.

Amit R, Abeles D, Bar-Gad I, Yuval-Greenberg S (2017) Temporal dynamics of saccades explained by a self-paced process. Sci Rep 7:1–15 Available at: http://dx.doi.org/10.1038/s41598-017-00881-7.

Andersson R, Larsson L, Holmqvist K, Stridh M, Nyström M (2017) One algorithm to rule them all? An evaluation and discussion of ten eye movement event-detection algorithms. Behav Res Methods 49:616–637 Available at: http://link.springer.com/10.3758/s13428-016-0738-9.

Baldauf D, Deubel H (2008) Properties of attentional selection during the preparation of sequential saccades. Exp Brain Res 184:411–425.

Bartlett a. M, Ovaysikia S, Logothetis NK, Hoffman KL (2011) Saccades during Object Viewing Modulate Oscillatory Phase in the Superior Temporal Sulcus. J Neurosci 31:18423–18432.

Beenhakker MP, Huguenard JR (2009) Neurons that fire together also conspire together: is normal sleep circuitry hijacked to generate epilepsy? Neuron 62:612–632 Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2748990&tool=pmcentrez&rendertype=abstract [Accessed May 29, 2014].

Bénar CG, Chauvière L, Bartolomei F, Wendling F (2010) Pitfalls of high-pass filtering for detecting epileptic oscillations: A technical note on "false" ripples. Clin Neurophysiol 121:301–310 Available at: http://dx.doi.org/10.1016/j.clinph.2009.10.019.

Binda P, Morrone MC (2018) Annual Review of Vision Science Vision During Saccadic Eye Movements. Available at: https://doi.org/10.1146/annurev-vision-091517-.

Birot G, Kachenoura A, Albera L, Bénar C, Wendling F (2013) Automatic detection of fast ripples. J Neurosci Methods 213:236–249 Available at: http://dx.doi.org/10.1016/j.jneumeth.2012.12.013.

Bosman C a, Womelsdorf T, Desimone R, Fries P (2009) A microsaccadic rhythm modulates gamma-band synchronization and behavior. J Neurosci 29:9471–9480 Available at: http://www.ncbi.nlm.nih.gov/pubmed/19641110 [Accessed January 26, 2014].

Bouzat S, Freije ML, Frapiccini AL, Gasaneo G (2018) Inertial Movements of the Iris as the Origin of Postsaccadic Oscillations. Phys Rev Lett 120:178101 Available at: https://link.aps.org/doi/10.1103/PhysRevLett.120.178101.

Bremer G, Smith JR, Karacan I (1970) Automatic detection of the K-complex in sleep electroencephalograms. IEEE Trans Biomed Eng:314–323.

Burnos S, Hilfiker P, Sürücü O, Scholkmann F, Krayenbühl N, Grunwald T, Sarnthein J (2014) Human intracranial high frequency oscillations (HFOs) detected by automatic time-frequency analysis. PLoS One 9.

Buschman TJ, Kastner S (2015) From Behavior to Neural Dynamics: An Integrated Theory of Attention. Neuron 88:127–144 Available at: http://www.sciencedirect.com/science/article/pii/S0896627315007746 [Accessed October 8, 2015].

Chaibi S, Sakka Z, Lajnef T, Samet M, Kachouri A (2013) Automated detection and classification of high frequency oscillations (HFOs) in human intracereberal EEG. Biomed Signal Process Control 8:927–934 Available at: http://dx.doi.org/10.1016/j.bspc.2013.08.009.

Chen C-Y, Hafed ZM (2017) A neural locus for spatial-frequency specific saccadic suppression in visual-motor neurons of the primate superior colliculus. J Neurophysiol 117:1657–1673.

Chen C, Ignashchenkova A, Thier P, Hafed ZM (2015) Neuronal Response Gain Enhancement

prior to Microsaccades. Curr Biol:1–10 Available at: http://dx.doi.org/10.1016/j.cub.2015.06.022.
Chen CY, Sonnenberg L, Weller S, Witschel T, Hafed ZM (2018) Spatial frequency sensitivity in macaque midbrain. Nat Commun 9.
Cohen MX (2017) Where Does EEG Come From and What Does It Mean? Trends Neurosci 40:208–218 Available at: http://dx.doi.org/10.1016/j.tins.2017.02.004.
Colby CL, Duhamel J-R, Goldberg ME (1996) Visual, presaccadic, and cognitive activation of single neurons in monkey lateral intraparietal area. J Neurophysiol 76:2841–2852.
Colby CL, Goldberg ME (1999) Space and attention in parietal cortex. Annu Rev Neurosci 22:319–349.
Collewijn H (1999) Eye movement recording. Vis Res A Pract Guid to Lab methods:245–285.
Colrain IM (2005) The K-complex: a 7-decade history. Sleep 28:255–273.
Daye PM, Optican LM (2014) Saccade detection using a particle filter. J Neurosci Methods 235:15–168 Available at: http://dx.doi.org/10.1016/j.jneumeth.2014.06.020.
De Gennaro L, Ferrara M (2003) Sleep spindles: an overview. Sleep Med Rev 7:423–440.
Deubel H, Schneider WX, Bridgeman B (1996) Postsaccadic target blanking prevents saccadic suppression of image displacement. Vision Res 36:985–996.
Diekelmann S, Born J (2010) The memory function of sleep. Nat Rev Neurosci 11:114–126 Available at: http://www.ncbi.nlm.nih.gov/pubmed/20046194 [Accessed July 10, 2014].
Drewes J, Zhu W, Wutz A, Melcher D (2015) Dense sampling reveals behavioral oscillations in rapid visual categorization. Sci Rep 5.
Dugué L, Roberts M, Carrasco M (2016) Attention Reorients Periodically. Curr Biol:1–7 Available at: http://linkinghub.elsevier.com/retrieve/pii/S0960982216303992.
Engbert R, Kliegl R (2003) Microsaccades uncover the orientation of covert attention. Vision Res 43:1035–1045 Available at: http://linkinghub.elsevier.com/retrieve/pii/S0042698903000841 [Accessed February 1, 2014].
Engbert R, Mergenthaler K (2006) Microsaccades are triggered by low retinal image slip. Proc Natl Acad Sci.
Fedele T, van ‘t Klooster M, Burnos S, Zweiphenning W, van Klink N, Leijten F, Zijlmans M, Sarnthein J (2016) Automatic detection of high frequency oscillations during epilepsy surgery predicts seizure outcome. Clin Neurophysiol 127:3066–3074 Available at: http://dx.doi.org/10.1016/j.clinph.2016.06.009.
Fernandez-Duque D, Johnson ML (2002) Cause and effect theories of attention: The role of conceptual metaphors. Rev Gen Psychol 6:153–165.
Fiebelkorn IC, Foxe JJ, Butler JS, Mercier MR, Snyder AC, Molholm S (2011) Ready, set, reset: stimulus-locked periodicity in behavioral performance demonstrates the consequences of cross-sensory phase reset. J Neurosci 31:9971–9981 Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3343369&tool=pmcentrez&rendertype=abstract [Accessed September 25, 2014].
Fiebelkorn IC, Pinsk MA, Kastner S (2018) A Dynamic Interplay within the Frontoparietal Network Underlies Rhythmic Spatial Attention. Neuron.
Fiebelkorn IC, Saalmann YB, Kastner S (2013) Rhythmic sampling within and between objects despite sustained attention at a cued location. Curr Biol 23:2553–2558 Available at: http://www.ncbi.nlm.nih.gov/pubmed/24316204 [Accessed May 13, 2014].
Fries P, Reynolds JH, Rorie AE, Desimone R, Fries P, Reynolds JH (2001) Modulation of oscillatory neuronal synchronization by selective visual attention. Science (80-) 291:1560–1563.
Gaarder K, Krauskopf J, Graf V, Kropfl W, Armington JC (1964) Averaged Brain Activity Following Saccadic Eye Movement. Science 146:1481–1483 Available at: http://www.ncbi.nlm.nih.gov/pubmed/14208580.
Gardner AB, Worrell GA, Marsh E, Dlugos D, Litt B (2007) Human and automated detection of

high-frequency oscillations in clinical intracranial EEG recordings. Clin Neurophysiol 118:1134–1143.
Garinis AC, Glattke T, Cone BK (2011) The MOC Reflex During Active Listening to Speech. J Speech, Lang Hear Res 54:1464–1476.
Gersch TM, Kowler E, Dosher B (2004) Dynamic allocation of visual attention during the execution of sequences of saccades. Vision Res 44:1469–1483.
Gersch TM, Kowler E, Schnitzer BS, Dosher BA (2008) Visual memory during pauses between successive saccades. J Vis 8:15.
Gersch TM, Kowler E, Schnitzer BS, Dosher BA (2009) Attention during sequences of saccades along marked and memorized paths. Vision Res 49:1256–1266.
Girardeau G, Zugaro M (2011) Hippocampal ripples and memory consolidation. Curr Opin Neurobiol 21:452–459.
Godijn R, Theeuwes J (2003) Parallel allocation of attention prior to the execution of saccade sequences. J Exp Psychol Hum Percept Perform 29:882.
Gruters KG, Murphy DLK, Jenson CD, Smith DW, Shera CA, Groh JM (2018) The eardrums move when the eyes move: A multisensory effect on the mechanics of hearing. Proc Natl Acad Sci 115:E1309–E1318.
Guinan J. J. J (2006) Olivocochlear efferents: anatomy, physiology, function, and the measurement of efferent effects in humans. Ear Hear 27:589–607.
Gutierrez R, Simon SA, Nicolelis MAL (2010) Licking-Induced Synchrony in the Taste-Reward Circuit Improves Cue Discrimination during Learning. J Neurosci 30:287–303.
Hafed ZM (2013) Alteration of visual perception prior to microsaccades. Neuron 77:775–786 Available at: http://www.ncbi.nlm.nih.gov/pubmed/23439128 [Accessed January 23, 2014].
Hafed ZM, Clark JJ (2002) Microsaccades as an overt measure of covert attention shifts. Vision Res 42:2533–2545 Available at: http://www.ncbi.nlm.nih.gov/pubmed/12445847.
Hafed ZM, Ignashchenkova A (2013) On the dissociation between microsaccade rate and direction after peripheral cues: microsaccadic inhibition revisited. J Neurosci 33:16220–16235 Available at: http://www.ncbi.nlm.nih.gov/pubmed/24107954 [Accessed January 22, 2014].
Hafed ZM, Krauzlis RJ (2010) Microsaccadic suppression of visual bursts in the primate superior colliculus. J Neurosci 30:9542–9547 Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2922969&tool=pmcentrez&rendertype=abstract [Accessed September 9, 2014].
Halpern BP (1983) Tasting and smelling as active, exploratory sensory processes. Am J Otolaryngol Neck Med Surg 4:246–249.
Helfrich RF, Fiebelkorn IC, Szczepanski SM, Lin JJ, Parvizi J, Knight RT, Kastner S (2018) Neural Mechanisms of Sustained Attention Are Rhythmic. Neuron 99:854-865.e5.
Herzfeld DJ, Kojima Y, Soetedjo R, Shadmehr R (2015) Encoding of action by the Purkinje cells of the cerebellum. Nature 526:439.
Herzfeld DJ, Kojima Y, Soetedjo R, Shadmehr R (2018) Encoding of error and learning to correct that error by the Purkinje cells of the cerebellum. Nat Neurosci 21:736.
Hoffman JE, Subramaniam B (1995) The role of visual attention in saccadic eye movements. Percept Psychophys 57:787–795.
Hooge I, Nyström M, Cornelissen T, Holmqvist K (2015) The art of braking: Post saccadic oscillations in the eye tracker signal decrease with increasing saccade size. Vision Res 112:55–67 Available at: http://dx.doi.org/10.1016/j.visres.2015.03.015.
Hughes B, Jansson G (1994) Texture perception via active touch. Hum Mov Sci 13:301–333.
Ito J, Maldonado P, Gruen S (2011a) Cross-frequency coupling of eye-movement related LFP activities of freely viewing monkeys. BMC Neurosci 12:P132.
Ito J, Maldonado P, Singer W, Grün S (2011b) Saccade-related modulations of neuronal excitability support synchrony of visually elicited spikes. Cereb Cortex 21:2482–2497.

James, William. "Attention." The principles of psychology 1 (1890): 402-458.
Jansen BH (1990) Artificial neural nets for K-complex detection. IEEE Eng Med Biol Mag 9:50–52.
Junker M, Endres D, Sun ZP, Dicke PW, Giese M, Thier P (2018) Learning from the past: A reverberation of past errors in the cerebellar climbing fiber signal. PLoS Biol 16:1–24.
Jutras MJ, Fries P, Buffalo E a (2013) Oscillatory activity in the monkey hippocampus during visual exploration and memory formation. Proc Natl Acad Sci 110:13144–13149.
Kandel ER, Schwartz JH, Jessell TM, Jessell D of B and MBT, Siegelbaum S, Hudspeth AJ (2000) Principles of neural science. McGraw-hill New York.
Kepecs A, Uchida N, Mainen ZF (2006) The sniff as a unit of olfactory processing. Chem Senses 31:167–179.
Kowler E (2011) Eye movements: The past 25years. Vision Res 51:1457–1483 Available at: http://dx.doi.org/10.1016/j.visres.2010.12.014.
Kusunoki M, Goldberg ME (2006) The Time Course of Perisaccadic Receptive Field Shifts in the Lateral Intraparietal Area of the Monkey. J Neurophysiol 89:1519–1527.
Lajnef T, Chaibi S, Eichenlaub J-B, Ruby PM, Aguera P-E, Samet M, Kachouri A, Jerbi K (2015) Sleep spindle and K-complex detection using tunable Q-factor wavelet transform and morphological component analysis. Front Hum Neurosci 9:1–17.
Landau AN, Fries P (2012) Attention samples stimuli rhythmically. Curr Biol 22:1000–1004 Available at: http://www.ncbi.nlm.nih.gov/pubmed/22633805 [Accessed July 10, 2014].
Landau AN, Schreyer HM, Van Pelt S, Fries P (2015) Distributed attention is implemented through theta-rhythmic gamma modulation. Curr Biol 25:2332–2337.
Larsson L, Nystrom M, Stridh M (2013) Detection of saccades and postsaccadic oscillations in the presence of smooth pursuit. IEEE Trans Biomed Eng.
Laubrock J, Engbert R, Kliegl R (2005) Microsaccade dynamics during covert attention. 45:721–730.
Lederman SJ, Klatzky RL (1987) Hand Movements: A Window into Haptic Object Recognition. Cogn Psychol 19:342–368.
Leresche N, Lambert RC, Errington AC, Crunelli V (2012) From sleep spindles of natural sleep to spike and wave discharges of typical absence seizures: is the hypothesis still valid? Pflugers Arch 463:201–212 Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3256322&tool=pmcentrez&rendertype=abstract [Accessed October 22, 2013].
Li C, Radulovacki M, Carley DW (2003) An automated pontine-wave detection system. Sleep 26:613–618 Available at: http://www.ncbi.nlm.nih.gov/pubmed/12938817.
López-Cuevas A, Castillo-Toledo B, Medina-Ceja L, Ventura-Mejía C, Pardo-Peña K (2013) An algorithm for on-line detection of high frequency oscillations related to epilepsy. Comput Methods Programs Biomed 110:354–360.
Lowet E, Gips B, Roberts MJ, De Weerd P, Jensen O, van der Eerden J (2018a) Microsaccade-rhythmic modulation of neural synchronization and coding within and across cortical areas V1 and V2 Vanduffel W, ed. PLOS Biol 16:e2004132 Available at: http://dx.plos.org/10.1371/journal.pbio.2004132 [Accessed October 14, 2018].
Lowet E, Gomes B, Srinivasan K, Zhou H, Schafer RJ, Desimone R (2018b) Enhanced Neural Processing by Covert Attention only during Microsaccades Directed toward the Attended Stimulus. Neuron:1–8 Available at: https://doi.org/10.1016/j.neuron.2018.05.041.
Lowet E, Roberts MJ, Bosman CA, Fries P, de Weerd P (2016) Areas V1 and V2 show microsaccade-related 3-4-Hz covariation in gamma power and frequency. Eur J Neurosci 43:1286–1296.
Lowet E, Roberts MJ, Bosman CA, Fries P, De Weerd P (2015) Areas V1 and V2 show microsaccade-related 3-4-Hz covariation in gamma power and frequency. Eur J Neurosci:1–11.

Maksimenko VA, Van Heukelum S, Makarov V V., Kelderhuis J, Lüttjohann A, Koronovskii AA, Hramov AE, Van Luijtelaar G (2017) Absence Seizure Control by a Brain Computer Interface. Sci Rep 7:1–8.
Markanday A, Messner J, Thier P (2018) A loss of a velocity-duration trade-off impairs movement precision in patients with cerebellar degeneration. Eur J Neurosci 48:1976–1989.
Mathis A, Mamidanna P, Cury KM, Abe T, Murthy VN, Mathis MW, Bethge M (2018) DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. Nat Neurosci 21 Available at: http://www.nature.com/articles/s41593-018-0209-y.
Matin E, Pearce DG (1969) Visual perception of direction when voluntary saccades occur. I. Relation of visual direction of a fixation target extinguished before a saccade to a flash presented during the saccade. Percept Psychophys 5:65–80.
Matin L, Matin E, Pola J (1970) Visual perception of direction when voluntary saccades occur: II. Relation of visual direction of a fixation target extinguished before a saccade to a subsequent test flash presented before the saccade. Percept Psychophys 8:9–14.
Matin L, Pearce DG (1965) Visual perception of direction for stimuli flashed during voluntary saccadic eye movements. Science (80-) 148:1485–1488.
McCamy MB, Otero-Millan J, Leigh RJ, King SA, Schneider RM, Macknik SL, Martinez-Conde S (2015) Simultaneous Recordings of Human Microsaccades and Drifts with a Contemporary Video Eye Tracker and the Search Coil Technique. PLoS One 10:e0128428 Available at: http://dx.plos.org/10.1371/journal.pone.0128428.
McCarley RW, Nelson JP, Hobson JA (1978) Ponto-geniculo-occipital (PGO) burst neurons: correlative evidence for neuronal generators of PGO waves. Science (80-) 201:269–272.
McPeek RM, Maljkovic V, Nakayama K (1999) Saccades require focal attention and are facilitated by a short-term memory system. Vision Res 39:1555–1566.
Meyberg S, Sinn P, Engbert R, Sommer W (2017) Revising the link between microsaccades and the spatial cueing of voluntary attention. Vision Res 133:47–60 Available at: http://dx.doi.org/10.1016/j.visres.2017.01.001.
Mihali A, van Opheusden B, Ma WJ (2017) Bayesian microsaccade detection. J Vis 17:13 Available at: http://www.ncbi.nlm.nih.gov/pubmed/28114483%0Ahttp://jov.arvojournals.org/article.aspx?doi=10.1167/17.1.13.
Morrone MC, Ross J, Burr D (2005) Saccadic eye movements cause compression of time as well as space. Nat Neurosci 8:950–954.
Morrone MC, Ross J, C. Burr D, Burr DC (1997) Apparent position of visual targets during real and simulated saccadic eye movements. J Neurosci 17:7941–7953.
Nyström M, Hooge I, Holmqvist K (2013) Post-saccadic oscillations in eye movement data recorded with pupil-based eye trackers reflect motion of the pupil inside the iris. Vision Res 92:59–66.
Nyström, Marcus, Ignace Hooge, and Richard Andersson. "Pupil size influences the eye-tracker signal during saccades." Vision research 121 (2016): 95-103.
Otero-Millan J, Castro JLA, Macknik SL, Martinez-Conde S (2014) Unsupervised clustering method to detect microsaccades. J Vis 14:18–18 Available at: http://jov.arvojournals.org/Article.aspx?doi=10.1167/14.2.18 [Accessed October 8, 2018].
Ovchinnikov A, Lüttjohann A, Hramov A, Van Luijtelaar G (2010) An algorithm for real-time detection of spike-wave discharges in rodents. J Neurosci Methods 194:172–178.
Pastukhov A, Braun J (2010) Rare but precious: Microsaccades are highly informative about attentional allocation. Vision Res 50:1173–1184 Available at: http://dx.doi.org/10.1016/j.visres.2010.04.007.
Pastukhov A, Vonau V, Stonkute S, Braun J (2013) Spatial and temporal attention revealed by microsaccades. Vision Res 85:45–57 Available at: http://dx.doi.org/10.1016/j.visres.2012.11.004.

Pekkanen J, Lappi O (2017) A new and general approach to signal denoising and eye movement classification based on segmented linear regression. Sci Reports 2017 71 7:17726 Available at: http://www.nature.com/articles/s41598-017-17983-x.
Polack P-O, Guillemain I, Hu E, Deransart C, Depaulis A, Charpier S (2007) Deep layer somatosensory cortical neurons initiate spike-and-wave discharges in a genetic model of absence seizures. J Neurosci 27:6590–6599 Available at: http://www.ncbi.nlm.nih.gov/pubmed/17567820 [Accessed October 22, 2013].
Poletti M, Rucci M, Carrasco M (2017) Selective attention within the foveola. Nat Neurosci 24 Available at: http://www.nature.com/doifinder/10.1038/nn.4622.
Posner MA (1980) Orienting of attention. Q J Exp Psychol 32:3–25 Available at: http://www.tandfonline.com/doi/abs/10.1080/00335558008248231#.VZ4m9flVj84.
Re D, Inbar M, Richter CG, Landau AN (2019) Feature-Based Attention Samples Stimuli Rhythmically. Curr Biol:1–7 Available at: https://doi.org/10.1016/j.cub.2019.01.010.
Reaz MBI, Hussain MS, Mohd-Yasin F (2006) Techniques of EMG signal analysis: detection, processing, classification and applications. Biol Proced Online 8:11.
Reynolds JH, Pasternak T, Desimone R (2000) Attention increases sensitivity of V4 neurons. Neuron 26:703–714.
Robinson DA (1963) A method of measuring eye movemnent using a scieral search coil in a magnetic field. IEEE Trans bio-medical Electron 10:137–145.
Rolfs M, Carrasco M (2012) Rapid Simultaneous Enhancement of Visual Sensitivity and Perceived Contrast during Saccade Preparation. J Neurosci 32:13744-13752a Available at: http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.2676-12.2012.
Rolfs M, Jonikaitis D, Deubel H, Cavanagh P (2011) Predictive remapping of attention across eye movements. Nat Neurosci 14:252–258 Available at: http://dx.doi.org/10.1038/nn.2711.
Ronneberger O, Fischer P, Brox T (2015) U-net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical image computing and computer-assisted intervention, pp 234–241. Springer.
Ross J, Morrone MC, Burr DC (1997) Compression of visual space before saccades. Nature 386:598.
Schroeder CE, Wilson D a, Radman T, Scharfman H, Lakatos P (2010) Dynamics of Active Sensing and perceptual selection. Curr Opin Neurobiol 20:172–176 Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2963579&tool=pmcentrez&rendertype=abstract [Accessed March 19, 2014].
Sethi A, Kemere C (2014) Real time algorithms for sharp wave ripple detection. In: 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp 2637–2640. IEEE.
Sheynikhovich D, Bécu M, Wu C, Arleo A (2018) Unsupervised detection of microsaccades in a high-noise regime. J Vis 18:19.
Sommer M a, Wurtz RH (2006) Influence of the thalamus on spatial visual processing in frontal cortex. Nature 444:374–377.
Song K, Meng M, Chen L, Zhou K, Luo H (2014) Behavioral oscillations in attention: rhythmic α pulses mediated through θ band. J Neurosci 34:4837–4844 Available at: http://www.ncbi.nlm.nih.gov/pubmed/24695703 [Accessed September 15, 2014].
Spyropoulos G, Bosman CA, Fries P (2018) A theta rhythm in macaque visual cortex and its attentional modulation. Proc Natl Acad Sci 115:E5614–E5623.
Tabernero J, Artal P (2014) Lens oscillations in the human eye. Implications for post-saccadic suppression of vision. PLoS One 9:e95764.
Theis L, Berens P, Froudarakis E, Reimer J, Rosón MR, Baden T, Euler T, Tolias AS, Bethge M (2016) Benchmarking spike rate inference in population calcium imaging. Neuron 90:471–482.
Tian X, Yoshida M, Hafed ZM (2016) A microsaccadic account of attentional capture and

inhibition of return in Posner cueing. Front Syst Neurosci 10:23.
Van der Stigchel S, Theeuwes J (2005) The influence of attending to multiple locations on eye movements. Vision Res 45:1921–1927.
Van Hese P, Martens JP, Waterschoot L, Boon P, Lemahieu I (2009) Automatic detection of spike and wave discharges in the EEG of genetic absence epilepsy rats from strasbourg. IEEE Trans Biomed Eng 56:706–717.
van Luijtelaar G, Lüttjohann A, Makarov V V, Maksimenko VA, Koronovskii AA, Hramov AE (2016) Methods of automated absence seizure detection, interference by stimulation, and possibilities for prediction in genetic absence models. J Neurosci Methods 260:144–158 Available at: http://dx.doi.org/10.1016/j.jneumeth.2015.07.010.
Venzi M, David F, Bellet J, Cavaccini A, Bombardi C, Crunelli V, Di Giovanni G (2016) Role for serotonin2A (5-HT2A) and 2C (5-HT2C) receptors in experimental absence seizures. Neuropharmacology 108:292–304 Available at: http://www.sciencedirect.com/science/article/pii/S0028390816301563.
Warby SC, Wendt SL, Welinder P, Munk EGS, Carrillo O, Sorensen HBD, Jennum P, Peppard PE, Perona P, Mignot E (2014) Sleep-spindle detection: Crowdsourcing and evaluating performance of experts, non-experts and automated methods. Nat Methods 11:385–392.
Wurtz RH (2008) Neuronal mechanisms of visual stability. Vision Res 48:2070–2089.
Yang Y, Lisberger SG (2014) Purkinje-cell plasticity and cerebellar motor learning are graded by complex-spike duration. Nature 510:529.
Yu G, Yang M, Yu P, Dorris MC (2017) Time compression of visual perception around microsaccades. J Neurophysiol 118:416–424.
Zanos TP, Mineault PJ, Nasiotis KT, Guitton D, Pack CC (2015) A sensorimotor role for traveling waves in primate visual cortex. Neuron 85:615–627 Available at: http://dx.doi.org/10.1016/j.neuron.2014.12.043.

J Neurophysiol 118: 2789–2805, 2017.
First published August 9, 2017; doi:10.1152/jn.00253.2017.

RESEARCH ARTICLE | *Higher Neural Functions and Behavior*

Sequential hemifield gating of α- and β-behavioral performance oscillations after microsaccades

Joachim Bellet,[1,2,3] Chih-Yang Chen,[1,2,3] and Ziad M. Hafed[1,3]

[1]*Werner Reichardt Centre for Integrative Neuroscience, Tuebingen University, Tuebingen, Germany;* [2]*Graduate School of Neural and Behavioural Sciences, International Max Planck Research School, Tuebingen University, Tuebingen, Germany; and* [3]*Hertie Institute for Clinical Brain Research, Tuebingen University, Tuebingen, Germany*

Submitted 5 April 2017; accepted in final form 8 August 2017

Bellet J, Chen CY, Hafed ZM. Sequential hemifield gating of α- and β-behavioral performance oscillations after microsaccades. *J Neurophysiol* 118: 2789–2805, 2017. First published August 9, 2017; doi:10.1152/jn.00253.2017.—Microsaccades are tiny saccades that occur during gaze fixation. Even though visual processing has been shown to be strongly modulated close to the time of microsaccades, both at central and peripheral eccentricities, it is not clear how these eye movements might influence longer term fluctuations in brain activity and behavior. Here we found that visual processing is significantly affected and, in a rhythmic manner, even several hundreds of milliseconds after a microsaccade. Human visual detection efficiency, as measured by reaction time, exhibited coherent rhythmic oscillations in the α- and β-frequency bands for up to ~650–700 ms after a microsaccade. Surprisingly, the oscillations were sequentially pulsed across visual hemifields relative to microsaccade direction, first occurring in the same hemifield as the movement vector for ~400 ms and then the opposite. Such pulsing also affected perceptual detection performance. Our results suggest that visual processing is subject to long-lasting oscillations that are phase locked to microsaccade generation, and that these oscillations are dependent on microsaccade direction.

NEW & NOTEWORTHY We investigated long-term microsaccadic influences on visual processing and found rhythmic oscillations in behavioral performance at α- and β-frequencies (~8–20 Hz). These oscillations were pulsed at a much lower frequency across visual hemifields, first occurring in the same hemifield as the microsaccade direction vector for ~400 ms before switching to the opposite hemifield for a similar interval. Our results suggest that saccades temporally organize visual processing and that such organization can sequentially switch hemifields.

microsaccades; fixational eye movements; α-rhythms; β-rhythms; perceptual oscillations

MICROSACCADES ARE SMALL SACCADES that occur during gaze fixation (Cherici et al. 2012; Hafed et al. 2015; Krauzlis et al. 2017; Poletti and Rucci 2016). These eye movements are governed by similar brain generation mechanisms as larger saccades (Goffart et al. 2012; Hafed 2011; Hafed et al. 2009; Hafed and Krauzlis 2012; Peel et al. 2016; Zuber et al. 1965), suggesting a precise level of control over the metrics and kinematics of these movements (Buonocore et al. 2017; Guerrasio et al. 2010; Havermann et al. 2014; Ko et al. 2010; Poletti et al. 2013; Tian et al. 2016). Interestingly, despite their small size, microsaccades are also associated with substantial modulations in neural activity and behavior and over a wide range of retinal eccentricities all the way from the fovea to the periphery (Bosman et al. 2009; Chen and Hafed 2017; Chen et al. 2015; Hafed 2013; Hafed et al. 2015; Hafed and Krauzlis 2010; Herrington et al. 2009; Tian et al. 2016; Zuber and Stark 1966). Such perimicrosaccadic modulations can be quite strong. For example, suppression of visual sensitivity for stimuli presented around the time of microsaccades can reach levels of ~50% in structures like the superior colliculus (SC) and lasts for ~100 ms after movement onset (Chen and Hafed 2017; Chen et al. 2015; Hafed and Krauzlis 2010). Moreover, perimicrosaccadic changes in vision result in performance changes that are virtually identical to those in a variety of experiments involving "covert" processing without eye movements (Chen et al. 2015; Hafed 2013; Hafed et al. 2015; Tian et al. 2016). Therefore, understanding the relationships between microsaccades and visual performance not only can enrich the microsaccadic literature itself but also can potentially illuminate findings in other related topics in perception and cognition.

The majority of studies of perimicrosaccadic modulations in vision so far have focused on a tight temporal window of ~50–100 ms around microsaccade onset. Since microsaccades are driven by saccade motor control circuitry (Hafed et al. 2009; Hafed and Krauzlis 2012; Peel et al. 2016), and since they also refresh retinal images (Martinez-Conde et al. 2000), it is reasonable to expect that changes in visual representations would occur during such a tight temporal window (Hafed et al. 2015; Krauzlis et al. 2017). However, microsaccades may also be associated with longer term changes in brain state going beyond simple "transient recovery."

In this study, we specifically hypothesized that microsaccades globally reset the visual system, meaning that after "transient recovery," the brain does not return to a random state but to a coherently structured steady-state oscillatory pattern. Such a pattern is so coherent and reliable across individuals that it manifests in alterations of the visual system's efficiency in detecting stimuli even at the behavioral level and even up to several hundreds of milliseconds after microsaccades. In this regard, we were motivated by findings of synchronization of endogenous brain rhythm behavioral manifestations with mo-

Address for reprint requests and other correspondence: Z. M. Hafed, Werner Reichardt Centre for Integrative Neuroscience, Otfried-Mueller Str. 25, Tuebingen 72076, Germany (e-mail: ziad.m.hafed@cin.uni-tuebingen.de).

tor actions. For example, Benedetto and colleagues have recently found a synchronous relationship between ~3.5- and 8-Hz behavioral oscillation rhythms of contrast sensitivity and reaching arm movements (Benedetto et al. 2016; Tomassini et al. 2015). Moreover, large, voluntary saccades are associated with phase locking of ~3- to 4-Hz behavioral oscillations (Benedetto and Morrone 2017; Hogendoorn 2016; Wutz et al. 2016), and microsaccades have been shown to reset physiological α rhythms (~10 Hz) in occipital cortex (Gaarder et al. 1966).

In our experiments, we therefore explored both time-dependent and microsaccade-direction-dependent long-term behavioral fluctuations after the occurrence of microsaccades. Unlike in the large-saccade literature (Benedetto and Morrone 2017; Hogendoorn 2016; Wutz et al. 2016), we observed higher frequency oscillations in the α- and β-frequency ranges. This is interesting as it might relate to findings of higher frequency behavioral oscillations in attentional state (Song et al. 2014), which is known to be correlated with microsaccades (Engbert and Kliegl 2003; Hafed 2013; Hafed and Clark 2002; Hafed et al. 2011; 2013; Kliegl et al. 2009; Laubrock et al. 2005; Peel et al. 2016; Tian et al. 2016).

As we will show, our results demonstrate that visual processing is almost never exempt from saccade-related influences, highlighting the importance of adopting an "active perception" approach to vision science.

MATERIALS AND METHODS

We performed two behavioral experiments on human subjects, and we also analyzed neurophysiological data from the SC of two male rhesus macaque monkeys (*Macaca mulatta*, aged 7 yr). As we describe below, the logic of all experiments was to present stimuli after microsaccade occurrence and to measure either behavioral or neural responses to these stimuli, with the specific goal of probing longer times after microsaccade onset than we had previously explored (Chen and Hafed 2017; Chen et al. 2015; Hafed 2013; Hafed et al. 2015; Hafed and Krauzlis 2010).

Ethics committees at Tuebingen University approved the human experiments, and human subjects provided informed, written consent in accordance with the Declaration of Helsinki. Monkey experiments were approved by the regional governmental offices of the city of Tuebingen.

Behavioral Tasks

Experiment 1: reaction time task. Human subjects sat in a dark room facing a computer display (41 pixels/°; 85 Hz). Head fixation was achieved through a custom-made chin/forehead rest (Hafed 2013), and we measured eye movements using a video-based eye tracker (EyeLink 1000; SR Research). In each trial, a white fixation spot (6-min arc in diameter; 89.1 cd/m^2) appeared on a gray screen (22.3 cd/m^2) (Fig. 1*A*). After a random time, a target (white circle subtending 1° in diameter) appeared at 5° either to the right or left of fixation, and the fixation spot disappeared. Subjects were instructed to execute a saccade as fast as possible to the target. Unbeknownst to the subjects, they continuously made microsaccades during fixation, and our experimental goal was to investigate reaction time (RT) to the target as a function of when this target appeared after a given microsaccade (Fig. 1*A*). We wanted to sample as long an interval as possible between a given microsaccade and target onset, such that we could uncover long-term microsaccadic effects on RT. As a result, and for the first six subjects who took part in the experiment, we programmed the experiment such that the target could appear between 1,000 and 3,000 ms after the initial fixation spot onset. The idea behind using such a long interval was to increase our chances of having trials with a very long delay between a microsaccade and target onset (i.e., to obtain long time courses of RT fluctuations; e.g., see Fig. 3*A* in RESULTS). Specifically, since intermicrosaccadic intervals can be short (e.g., see Fig. 4*B* in RESULTS), we needed to have sufficient trials with long delays to be able to sample enough cases in which no intervening microsaccades occurred for a long time after a given movement until target onset. Having said that, and to accelerate data collection, we found that we could shorten trial lengths for the last four subjects that we recruited; in this case, the target could appear between 300 and 1,300 ms after fixation spot onset. Even in these cases, we ensured that our data analyses included only microsaccades occurring after >300 ms of steady fixation (for reasons clarified below).

We collected data from 10 subjects (6 females; 2 authors; 22–39 yr old). All subjects completed 5 sessions of 1 h each, and we analyzed a total of 10,817 trials in this experiment. As stated above, this large number of trials was needed to allow us to sample enough time points after individual microsaccades. In reality, for any given time bin in our analyses, the number of trials used was much less than the total number of trials collected (e.g., see Fig. 3*A* in RESULTS).

Experiment 2: perceptual detection task. In a second experiment, we collected data from 14 human subjects (9 females; 2 authors; 22–39 yr old), and all subjects completed 3–5 sessions of 1 h each.

Each trial started with the appearance of a black fixation spot subtending a 6-min arc in diameter along with white vertical lines (3-min arc × 4°) that were vertically centered 3° above and below the two probable horizontal locations of an upcoming target (Fig. 1*B*). The target was a dot subtending a 4-min arc in diameter presented horizontally 5° to the right or left of fixation. The target was presented only briefly for ~12 ms. The luminance of the target was darker than the background, and it was continuously adjusted from trial to trial with a staircase procedure (see below) to maintain correct target detection performance of ~50% throughout all sessions. In our design, "correct target detection" meant that subjects reported both a conscious percept of the target and its correct location, because they also always had the option to report "target not seen" (see below). We chose to continuously adjust task difficulty to maintain ~50% correct detections to demonstrate the robustness of our observed modulations in target detection likelihoods at different times after microsaccades (e.g., see Fig. 7 in RESULTS) and also to ensure stable performance in the face of fluctuations in subject arousal during the sessions. This has also allowed us to combine subjects in analyses, as we describe below.

The target was presented at a random time ranging from ~61 to ~1,061 ms after gaze entered a virtual window of 2° radius around the fixation spot. Subjects were instructed to press a button corresponding to the target location (right or left) as soon as it was consciously seen (Fig. 1*B*). Because the target was difficult to see, we presented a question mark at screen center 500–700 ms after target presentation to indicate to subjects that they needed to make a response (Fig. 1*B*). When the question mark was visible, subjects could either press a middle button to indicate that the target was not seen or to press the button corresponding to the right/left target location if it was seen. Subjects were instructed not to guess the location and to only answer when the target was consciously seen. To ensure that the subjects followed the task instructions, 10% of the trials were designated as catch trials with no target appearing at all, and we confirmed that false alarms on these catch trials were much less frequent than correct detections; we also confirmed that erroneous localizations of the target when subjects reported consciously seeing it were rare (e.g., see Fig. 7*A* in RESULTS). Auditory feedback signaled if target localization was correct or not. There was no auditory feedback when the subjects reported not seeing the target.

For each subject, the first session in this experiment was used solely to estimate target contrast needed to achieve ~50% correct target detections. The luminance of the target was first started at background

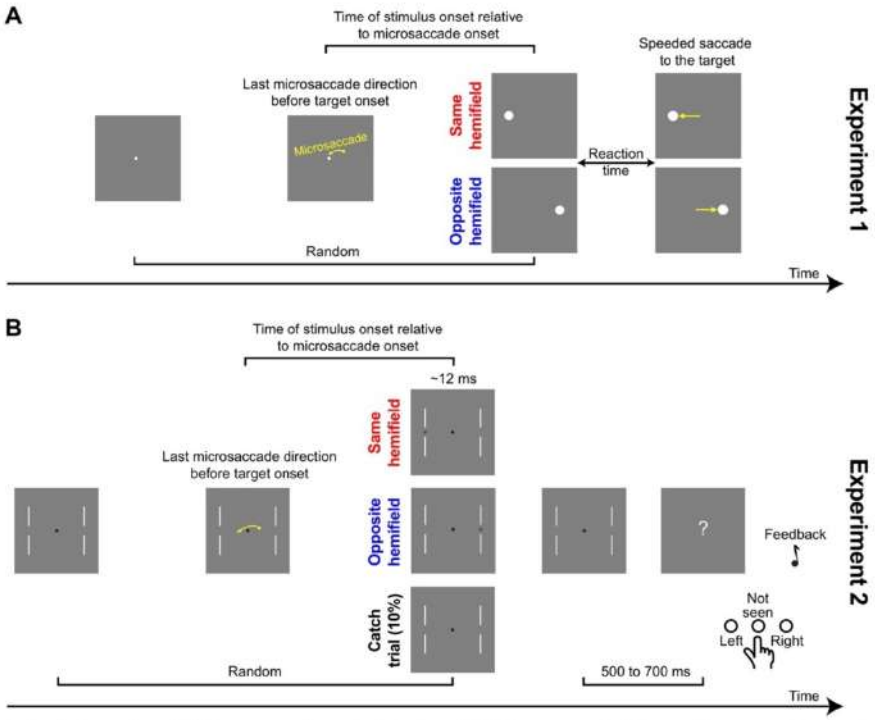


Fig. 1. Behavioral tasks. *A*: *experiment 1*. Subjects generated a target-directed saccade. Before target onset, subjects could generate a microsaccade either toward the same or opposite hemifield as the target. We analyzed reaction time (RT) as a function of target onset time and direction relative to microsaccades, and we ensured that there were no intervening movements (saccades or microsaccades) or blinks between the microsaccade of interest and target onset. *B*: *experiment 2*. Subjects detected a target at threshold by indicating its location (or that they did not see any target). Task difficulty was continuously adjusted to keep conscious target detections at ~50% (see MATERIALS AND METHODS).

contrast, which meant that subjects could not detect the target. After every "target not seen" response by the subjects (Fig. 1*B*), the luminance of the target was stepped by one display-gun value on our eight-bit display driver (i.e., 1 out of 256 levels). As soon as the target began to be detected, we started a procedure of monitoring correct target detections based on the previous 50 trials. Target luminance was increased or decreased by 1 display-gun value if average performance was above 66% or below 34%, respectively, during the previous 50 trials. For the subsequent sessions, target luminance started at the level predicted from the first session, and it was continuously adjusted (always based on performance in the previous 50 trials) using the same incremental procedure. This approach enabled us to track the ~50% threshold well (e.g., see Fig. 8*A* in RESULTS). We analyzed a total of 6,693 trials from this task.

Experiment 3: monkey SC recordings. We performed a novel analysis on data that were part of the set used to describe a previous report (Chen et al. 2015); here, we critically extended our analysis window as long as possible after microsaccade onset to explicitly sample longer time courses of neural modulations of SC visual sensitivity after microsaccades than were analyzed in the previous report. Briefly, two monkeys fixated, and we presented a 2.2 cycles/° vertical sine wave grating in a neuron's visual response field. The grating was tailored in size to the response field size, and we varied the contrast from trial to trial. We considered as "baseline" visual response the response of SC neurons when the grating appeared with no microsaccades occurring within ±150 ms from grating onset. We then measured the visual response of the same neurons when the grating appeared at a given time after a microsaccade. Once again, the difference in the present study from our earlier experiments was that we increased the time window of analysis as much as possible (see Fig. 10 in RESULTS).

Data Analysis

Microsaccades and saccades were detected offline using velocity and acceleration criteria (Chen and Hafed 2013), and microsaccade misdetections were checked manually for all trials. We defined as microsaccades all fixational saccades with amplitudes <1°; in reality, median amplitude was a 14.6-min arc in *experiment 1* and a 17.6-min arc in *experiment 2*.

Since we observed that microsaccade frequency increased substantially in the first 300 ms following fixation spot onset (e.g., see Fig. 4*C* in RESULTS), we only analyzed trials in which the last microsaccade to occur before target onset actually occurred >300 ms after fixation spot onset. This was done to ensure that the phase locking that we observed in this study (see RESULTS) was to microsaccades and not necessarily to another resetting event (i.e., fixation spot onset). In a control analysis, we also analyzed data when target onset time was measured with respect to fixation spot onset and not to microsaccade onset; in this case, it was irrelevant whether or not a microsaccade had occurred between fixation spot onset and target onset.

Experiment 1: RT task. Our goal in this experiment was to analyze the spectrotemporal variation of RT as a function of when a target appeared after a given microsaccade (Fig. 1*A*). We thus performed a series of analysis steps that are graphically illustrated in Fig. 2 on an example artificial data set designed to clarify our procedures as much as possible (real data are presented in all subsequent figures). In what follows, we step through the details of our analyses with the aid of this figure.

We first excluded the 2.5% fastest and 2.5% slowest RTs of each subject from any further analysis. We then demeaned each subject's RT distribution by obtaining RT values expressed as a difference from the overall mean RT, thus reducing intersubject variability. We then pooled trials from all subjects together, similar to what has been done previously (Hogendoorn 2016), and we plotted RT as a function of when a target on a given trial appeared after a microsaccade (also see below for further justification). This resulted in a scatter plot similar to that shown in Fig. 2*A*, *top*. The mean RT value per subject was obtained from all trials regardless of when microsaccades or target onsets occurred. Also, in every trial, the demeaned RT value was plotted as a function of the time of target appearance relative to microsaccade onset for the very last microsaccade to occur before target onset (Fig. 1*A*), such that we ensured that there were no intervening microsaccades (or other eye movements) between the microsaccade of interest and the target onset. This means that if there were multiple microsaccades during the fixation period before target onset (a rarity), the time interval of interest to us was the time between target onset and the final microsaccade to occur before such onset. We also excluded any trials with blinks between the microsaccade of interest and target onset.

To obtain the time course of RT fluctuations, we computed the mean value across all trials using a running window of 50 ms in width in steps of 1 ms (Fig. 2*B*, *top*). Note that such a running window is

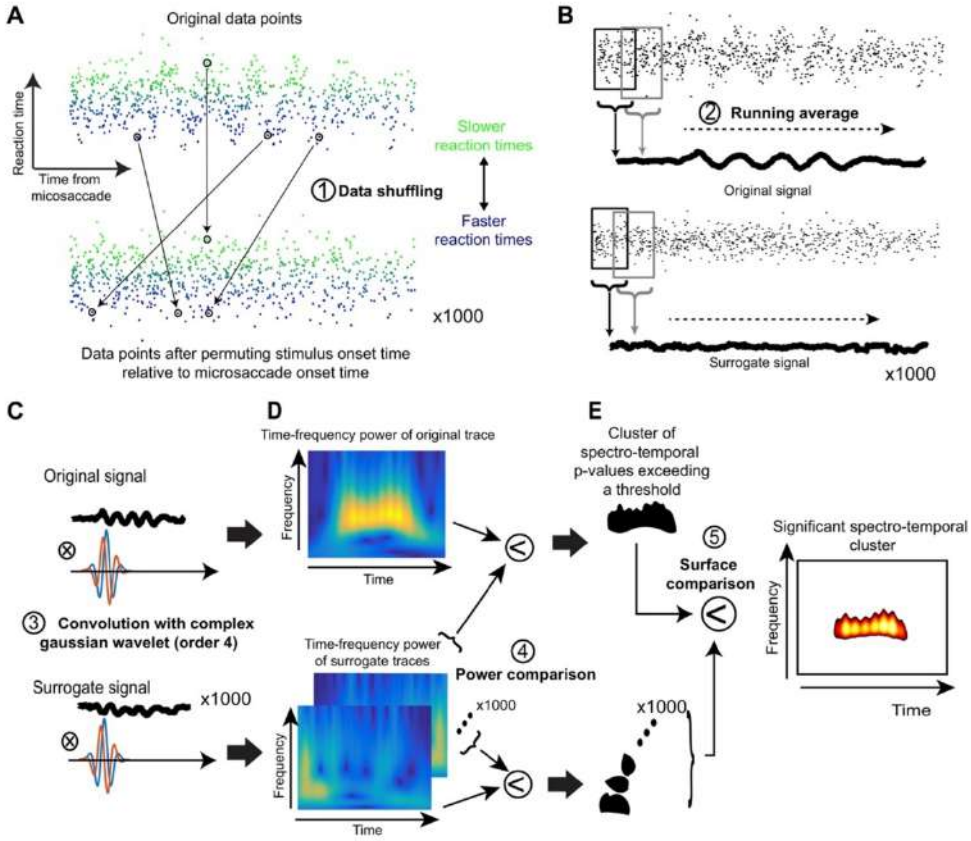


Fig. 2. Steps used for time-frequency analyses in our study. *A*: we collected all trials and plotted each trial's RT (as in the case of *experiment 1*) as a function of target onset time relative to a microsaccade. In this illustrative example, we show artificial data that have an underlying oscillatory pulse embedded in noise. *Bottom*: we shuffled the data in time, and we did this 1,000 times to obtain 1,000 surrogate data sets. *B*: we filtered the data using a running average to obtain a time course of fluctuations in the real data (*top*) and also for each of the 1,000 surrogate data sets. *C*: we then convolved each of the time course traces with a complex Gaussian wavelet of order 4. *D*: this allowed us to obtain time-frequency power spectra. For the real data, high power was observed during the interval in which an oscillation pulse was present in the original signal. The surrogate data sets had high power at random places in the time-frequency plots. *E*: we established statistical significance, with corrections for multiple comparisons, by finding a cluster in 2-dimensional time-frequency space in the original data set that was unlikely to be observed by chance from similar clusters observed in the surrogate data sets. We then plotted this "outlying" cluster with its associated *P* values as the significant cluster of spectrotemporal oscillation.

equivalent to implementing a 20-Hz low-pass filter, and it reduces the amplitude of oscillations. The higher in frequency an oscillation is (and approaching 20 Hz), the more its amplitude will be reduced by the running window analysis. This also means that our analyses could not detect oscillations higher than 20 Hz in frequency; however, to our knowledge, behavioral oscillations are rarely studied above this frequency. For some analyses (e.g., see Fig. 5 in RESULTS), we performed this same RT time-course analysis but only for trials in which the target appeared either in the "same" hemifield as the microsaccade direction vector or in the "opposite" hemifield. In addition, in yet other analyses (e.g., see Figs. 3 and 5 in RESULTS), we also computed the RT time courses for only microsaccades smaller than a 30-min arc in amplitude.

To obtain the spectral content of RT fluctuations, we analyzed the RT time course in the interval 100–800 ms after microsaccade onset (Fig. 2, *A* and *B*). We excluded the first 100 ms after microsaccade onset because this period is dominated by "microsaccadic suppression" effects on RT (Chen and Hafed 2017; Chen et al. 2015; Hafed et al. 2015; Hafed and Krauzlis 2010; Tian et al. 2016), as we detail in RESULTS. We first removed the slow fluctuation trend of the signal (any <2-Hz fluctuation) by using a finite impulse response low-pass filter (Fieldtrip Toolbox; Oostenveld et al. 2011) and subtracting the result from the original curve (Fig. 2*B*, *top*). The detrended signal was then transformed using continuous complex Gaussian wavelets (order 4) for frequencies ranging from 2 to 20 Hz in steps of 0.25 Hz (Fig. 2*C*, *top*). The power of the resulting time-frequency image was defined as the square of the absolute value of the complex transform (Fig. 2*D*, *top*). We now had a time-frequency plot of spectrotemporal variation in RT (Fig. 2*D*, *top*).

To assess statistical significance of given time intervals and/or frequency bands in the spectral analysis, we used the following starting point: the hypothesis in this study was that the oscillations that we observed in RT were caused by the time at which a stimulus was presented relative to a microsaccade (see RESULTS). In other words, such oscillations can hardly ever be observed if we were to compose surrogate signals by randomly arranging the behavioral responses in time (i.e., without regard to the real time of stimulus onset relative to a microsaccade; Fig. 2*A*, *bottom*). We therefore tested the probability that random permutations in the *x*-label values of figures like Fig. 2*A*, *top*, would result in as much as or more spectral power than the original data. To generate a statistic, the following procedure was repeated 1,000 times. We used data points that were collected outside the microsaccadic suppression period (i.e., >100 ms after microsaccade onset), and we arranged them randomly relative to the time of a virtual microsaccade (Fig. 2*A*, *bottom*). Care was taken such that there were as many data points assigned for each time point as in the original measurements, such that the difference in power between original signals and surrogate signals could not be explained by differences in variability (Fig. 2*A*). The randomly arranged data points were then processed using the exact same procedures as the original data: time courses were obtained using running windows (Fig. 2*B*, *bottom*); time courses were detrended; and finally, time courses were wavelet-transformed (Fig. 2, *C* and *D*, *bottom*). For each time-frequency point in the original power spectrum, *P* values were defined as the proportion of power spectra in surrogate data that led to an equal or higher power than the original data (Fig. 2*E*). $P < 0.05$ was considered as significant. To control for multiple comparisons, we additionally used a nonparametric statistical test of the size of the cluster of adjoining significant *P* values (Maris and Oostenveld 2007). More specifically, the multiple comparisons test consisted of determining the probability of obtaining a cluster (defined as a 2-dimensional region in time-frequency space) of adjoining significant *P* values as big as or bigger than the one obtained with the original data. To this end, each of the 1,000 power spectra obtained by random permutations was compared with the 999 other random power spectra (Fig. 2*D*, *bottom*). This resulted in 1,000 sets of *P* values. Clusters of significant *P* values from the original power spectrum passed the multiple comparisons test if their size was >95% of the biggest clusters of adjoining significant *P* value in each randomly drawn set of *P* values (Fig. 2*E*).

Our approach to statistical tests of the oscillations provided us with the most conservative estimates of significant oscillations in our data. This was important because using a running window (i.e., a low-pass filter) can cause ringing in signals that may erroneously appear like oscillations by visual inspection. For example, see Figs. 3*B* and 5, *B* and *E*, in RESULTS for examples of surrogate traces from our analyses after the shuffling step of Fig. 2*A*. In some of these sample shuffled traces, visual inspection might lead one to assume that an oscillation was present. Thus our approach was to err on the conservative side to be certain that our interpretations of oscillations are highly reliable. This is also why we additionally ensured using similar numbers of data points in each time bin in our shuffled signals as in the real data.

To rule out the possibility that RT oscillations may have resulted from an increase of variance in RT after microsaccade onset, for example, as suggested in (Zoefel and Sokoliuk 2014), we investigated whether the phase of RT oscillation time courses was consistent across subjects. To represent the instantaneous phase of oscillations in each subject (e.g., see Fig. 6), we first obtained individual time courses of RT using an averaging running window of 50 ms in the "same hemifield" condition (i.e., when the target appeared in the same hemifield as a microsaccade direction vector). In the "opposite hemifield" condition, we chose a 78-ms averaging window instead to obtain the best estimate of the phase of α-oscillations in each subject (because "opposite" oscillations were lower in frequency than "same" oscillations, as we show in RESULTS). We then band-passed the traces in the frequency bands of interest based on Fig. 5 (13–20 Hz in the same condition and 8–12 Hz in the opposite condition; see RESULTS), and we extracted the instantaneous phase with a Hilbert transform (see Fig. 6 in RESULTS). Even though such a transform, which is well established to evaluate phase (Le Van Quyen et al. 2001), may introduce phase distortions at the edges of a signal interval, such distortions would also occur in permuted data during significance testing (see below); thus, if anything, there may be more false negatives at the edges of an interval being analyzed. To quantify the instantaneous phase consistency between subjects, we calculated the phase-locking value (PLV) at each time point after microsaccade onset:

$$PLV = \frac{1}{n} \cdot \left| \sum_{k=1}^{n} e^{i\theta_k[2\pi]} \right| \quad (1)$$

where n is the number of subjects and θ_k the instantaneous phase from the k^{th} subject. The PLV is between 0 (no coherence) and 1 (complete coherence) (Lachaux et al. 1999).

We assessed the statistical significance of an increased PLV by using random permutations (e.g., Fig. 2*A*) and cluster corrections for multiple comparisons. The PLV was computed for 1,000 sets of 10 surrogate signals (1 per subject). A *P* value for each time bin was obtained by comparing the original PLV traces to the surrogate ones. We also obtained surrogate *P* values by comparing each surrogate PLV trace to the 999 other traces. We then measured the size of the clusters of adjoining *P* values from the original PLV trace that were <0.1. This threshold was meant to reveal long-lasting PLV increases, but the exact choice does not affect the false alarm rate of the statistical test (Maris and Oostenveld 2007). We tested the probability that clusters obtained from surrogate *P* values below the threshold had a size equal to or exceeding the size of the cluster from the original signal. If <5% of the biggest clusters from each of the 1,000 surrogate *P* value traces were bigger than the cluster from the original signal, then the whole period of increased PLV was considered as significant.

Finally, we also performed discrete Fourier transform (DFT) analysis on RT fluctuations in the interval 100–400 ms after microsaccade onset (i.e., the interval in which we observed significant oscillations throughout the different conditions; see RESULTS). We assessed signif-

icance of individual frequency ranges using a similar approach to that described above (including multiple comparison corrections), except that there was now only one dimension to work with (i.e., frequency) as opposed to two (i.e., time and frequency). We also performed such DFT analysis for microsaccades either smaller than or larger than the median amplitude.

It should be noted once again here that we combined all trials from all subjects in most of our analyses, as in Hogendoorn (2016). Given that our subjects performed similar numbers of trials to each other, it was unlikely that our results could be dominated entirely by only a subset of outlying subjects, and so our choice was justified. We also confirmed this by our PLV analyses described above, because such analyses can reveal phase consistency across individuals, and we took measures in all of our statistical analyses to take the varying numbers of observations per time bin into account. If each subject's data were first grouped into a single average trace before averaging across subjects, as performed in some other studies, such information about variability of numbers of observations per time bin would have been lost, which in turn increases the likelihood that outlying subjects may have distorted the measured signal-to-noise ratios in the averages.

Experiment 2: perceptual detection task. We used a similar time-course analysis on detectability as a function of time of target appearance relative to a microsaccade. Specifically, we computed the proportion of detected targets combining data points from all subjects together with a running window of 50 ms in width. We also used the same spectral analyses on these time courses as in the RT task described above and illustrated in Fig. 2. We should emphasize here that our task was not a classic two-alternative forced choice task. Thus, by "proportion of correct target detections," we mean the fraction of noncatch trials in which subjects reported seeing the target and correctly localized it. Thus, if this proportion was 50% (as per task design), then this means that the remaining 50% of the trials could be of two types: *1*) either subjects indicated "target not seen" in their response (which was the great majority of cases; see RESULTS); or *2*) subjects erroneously reported seeing the target in the wrong location. This also means that microsaccadic suppression alluded to above can lead to reductions in "correct target detections" to levels well below 50%, as we show in RESULTS.

To assess the statistical reliability of the difference in perceptual detection performance between two conditions (e.g., "same" vs. "opposite" or "close" vs. "far"; see Fig. 7), we used two-proportion z-tests.

To study the effects of microsaccade directions while canceling out the influence of gaze distance to the target (or equivalently, retinotopic target position relative to the fovea) at the time of target onset, we downsampled the data sets such that the distribution of retinotopic target positions was the same whether the target appeared in the same or opposite hemifield from a microsaccade. Specifically, we split the entire distribution of observed retinotopic target positions into 50 quantiles. Then, within each quantile, we had a distribution of N "same" trials and a distribution of M "opposite" trials. If N was larger than M, we randomly picked M trials from the "same" distribution to match the numbers of trials in the quantile to the "opposite" distribution; similarly, if M was larger than N, we randomly picked N trials from the "opposite" distribution. This resulted in matched distributions of retinotopic target positions between same and opposite trials. From these matched distributions, we compared detection performance on same vs. opposite trials, and we also measured detection performance from the same data after shuffling the "same" and "opposite" labels (to obtain a surrogate data set). Since this downsampling procedure required randomly selecting subsets of data, we repeated this procedure 1,000 times, obtaining 1,000 downsampled distributions and also 1,000 surrogate measurements from shuffled data. The P value was calculated as the likelihood of observing differences between "same" and "opposite" shuffled data sets that were larger in absolute value than the differences in performance between the real "same" and "opposite" conditions observed with overlapping retinotopic target positions.

We also compared detection performance on same vs. opposite microsaccade trials to detection performance on "close" vs. "far" trials. We repeated the same comparison of perceptual detection performance but after labeling trials as either having "close" or "far" target positions relative to the fovea instead of being from "same" or "opposite" microsaccade trials. To obtain "close" and "far" trials, we performed a median split on the Euclidean distance between gaze and the target (see Fig. 7*B*, *right*). In other words, in this analysis, we simply asked whether any changes in perceptual detection performance between same and opposite microsaccade trials were in reality caused by the potential that retinotopic target position relative to the fovea altered visual acuity and therefore detection performance (see RESULTS). Finally, to explore whether there were different effects for different microsaccade sizes, we repeated the comparisons between "same" and "opposite" trials but only for trials in which microsaccades were either less than a 30-min arc in amplitude or when they were classified as being larger or smaller than the median amplitude.

Experiment 3: monkey SC recordings. We repeated the time course analyses described in detail in (Hafed and Krauzlis 2010), but we extended them in time as much as possible given the data set available. We combined trials from 20, 40, and 80% contrast because these trials consistently exhibited robust visual responses. We normalized responses to the "baseline" of each contrast individually to ensure that our plots of neural fluctuations (e.g., see Fig. 10) isolate the influence of microsaccades on visual response strength, independent of stimulus contrast. As stated above, we considered as "baseline" neural responses those responses that were observed when stimuli appeared without any microsaccades occurring within ±150 ms from stimulus onset.

RESULTS

Microsaccades Reset the Phase of α- and β-Frequency Oscillatory RT Fluctuations

In *experiment 1*, we analyzed visual processing efficiency by asking 10 human subjects to perform a speeded RT task. Subjects fixated on a spot, and we then presented a bright target at 5° eccentricity either to the right or left of fixation (Fig. 1*A*, see MATERIALS AND METHODS). Subjects had to look at the target as fast as possible, and we investigated RT modulations as a function of when the target appeared relative to a microsaccade. This task has previously been shown to provide a very sensitive behavioral measure of visual sensitivity changes of SC neurons around microsaccades (Chen and Hafed 2017; Hafed and Krauzlis 2010). The key here was to densely sample many times even long after movement occurrence. We measured RT for stimuli appearing up to 800 ms after a given movement, and we took care to ensure that there were no other microsaccades, saccades, or blinks occurring between the microsaccade of interest and the target (see MATERIALS AND METHODS). We also aimed to identify only common effects across individuals. We thus subtracted each subject's overall mean RT from each trial's measurement, obtaining a differential RT value, and we then combined trials from all subjects in analyses.

When the target appeared immediately after a microsaccade, we observed an expected RT cost (Chen and Hafed 2017; Chen et al. 2015; Hafed and Krauzlis 2010). Specifically, Fig. 3*A*, *top*, shows a plot of differential RT value across subjects as a function of target onset time after a microsaccade. The error bars in Fig. 3 show SE values across all trials from all subjects combined (see MATERIALS AND METHODS), and Fig. 3*A*, *bottom*,

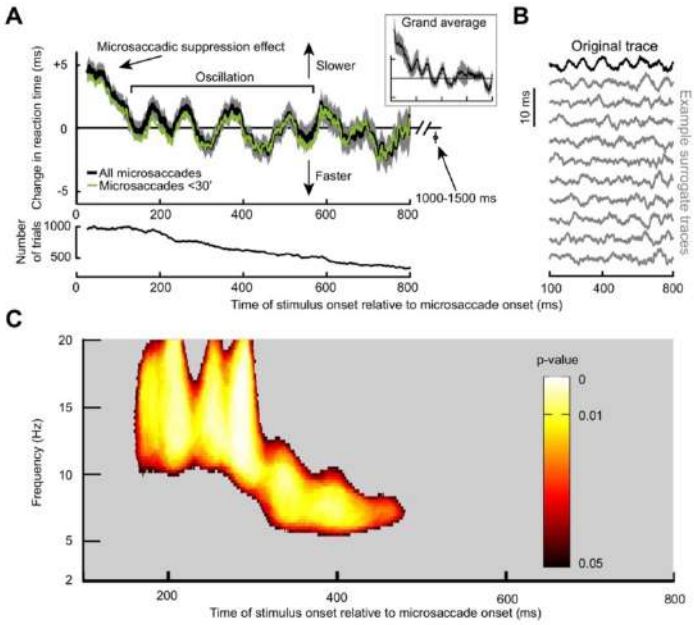


Fig. 3. Long-term microsaccadic influence on RT. *A*: change in mean RT as a function of target onset time relative to a microsaccade. This trace shows the demeaned RT trace before detrending (see MATERIALS AND METHODS). After an initial ~100-ms period of RT costs (i.e., increases) due to "microsaccadic suppression," RT oscillated coherently for almost 600 ms. Error bars denote SE across all trials from all subjects combined (see MATERIALS AND METHODS), and we used a running window of 50 ms (in steps of 1 ms); the *inset* shows that we obtained very similar RT fluctuations when individual subject data were first averaged into a single subject trace and then all the single subject traces averaged together to obtain a grand average. The *bottom* shows the number of trials used per time bin in the *top* and using the same running window procedure. Also, the green trace shows results from only trials with microsaccades less than a 30-min arc in amplitude. Note that the RT oscillation (and subsequent measurement for 1,000–1,500 ms shown as an individual data point) hovered slightly below 0. This is because the total average used to demean RT included all trials even those with target onset <300 ms from fixation spot onset; because these early trials almost always had a microsaccade (e.g., Fig. 4*C*), they had long RTs due to microsaccadic suppression, and this elevated the average RT value slightly. *B*: the *top trace* in black is the original demeaned and detrended RT trace from the same data in *A*. Below this trace are 10 example surrogate traces according to the permutation procedures of Fig. 2*A*. Our goal was to establish statistically whether oscillations in the original trace that are visible by inspection were not ones expected by chance as might happen in some of the surrogate traces. *C*: power spectrum of the trace in *A* during the interval following the initial microsaccadic suppression period of 100 ms. Colored pixels indicate times and frequency bands with a significant oscillation that is not expected by chance from the surrogate traces (Fig. 2; see MATERIALS AND METHODS).

shows the number of trials within each time bin used to obtain the data in Fig. 3*A*, *top*. In Fig. 3*A*, *bottom*, the time course of number of trials was analyzed using the same running window as that used in Fig. 3*A*, *top*. As can be seen from Fig. 3*A*, *top*, for up to ~100 ms, RT was increased relative to the mean, and it gradually decreased back to "baseline," consistent with a short-lived "microsaccadic suppression" effect known to influence both RT and visual sensitivity in SC and other brain structures (Chen and Hafed 2017; Chen et al. 2015; Hafed and Krauzlis 2010). Interestingly, after ~100 ms, the RT recovery was not back toward a "stable" baseline. Instead, RT kept fluctuating in a rhythmic fashion, becoming sequentially either faster or slower than average (e.g., see the period labeled "Oscillation" in Fig. 3*A*). Moreover, this oscillation was long-lasting and persisted for almost the entire 800 ms that we sampled, and it also occurred when we only included trials with microsaccades less than a 30-min arc in amplitude (Fig. 3*A*, thin green curve). The oscillation was also present (Fig. 3*A*, *inset*) when we first averaged all trials of a single subject into a single curve and then averaged all the individual subjects' single traces to obtain a grand average, as is often done in other studies (e.g., Fiebelkorn et al. 2013).

Since microsaccades have an intrinsic rhythm to them (Bosman et al. 2009; Hafed and Ignashchenkova 2013; Tian et al. 2016), we excluded the possibility that the fluctuations in Fig. 3*A* reflected the occurrence of subsequent microsaccades coming in a rhythmic fashion. Specifically, in our analyses, there was only one single microsaccade before target onset in the analyzed intervals, by design, even for stimuli appearing >700 ms after a movement. To confirm this, we inspected eye velocity traces from all of our accepted trials and ensured that there were no spikes in velocity after the spike associated with the microsaccade of interest. For example, the black radial eye velocity trace of Fig. 4*A* shows average eye velocity from all accepted trials, with only a peak at the microsaccade of interest. In contrast, when we plotted average eye velocity after a randomly chosen microsaccade irrespective of whether a subsequent microsaccade had occurred or not, we obtained the dashed gray curve in Fig. 4*A*. As expected, this curve was elevated compared with the black one, with a peak occurring

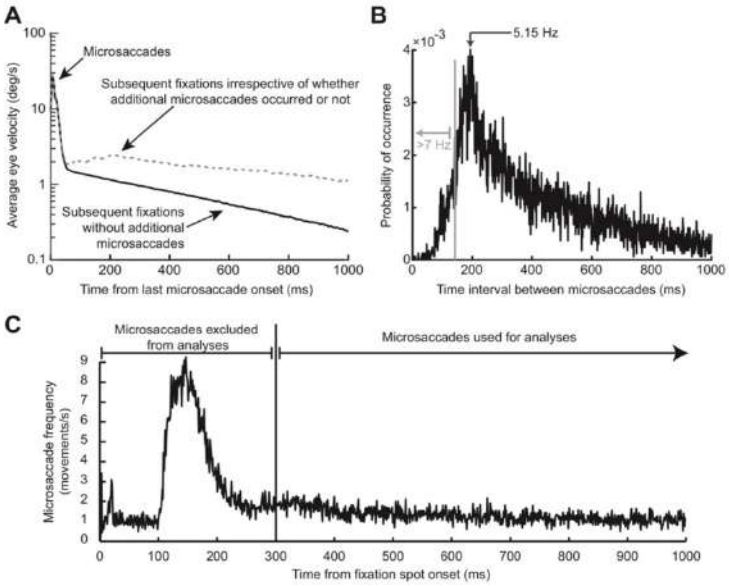


Fig. 4. The RT oscillation in Fig. 3 was not due to either intrinsic microsaccadic rhythms or visual transients associated with the onset of the fixation spot at trial beginning. *A*: the black curve shows mean eye velocity after the last detected microsaccade before target onset from all trials in Fig. 3. For comparison, the dashed gray curve shows mean eye velocity after any given microsaccade irrespective of whether there were subsequent microsaccades or not (from a similar number of trials). The black curve is characterized by an absence of positive-going fluctuations in eye velocity after the first microsaccade (at 0 ms), whereas the gray curve shows elevations peaking at ~200 ms and persisting later, consistent with the likelihood of observing subsequent microsaccades (also see *B*). A similar result was also obtained for *experiment 2*. *B*: distribution of intermicrosaccadic intervals in *experiment 1*. If 2 microsaccades occurred before target onset, the time between the final microsaccade and target onset contributed to our analyses (Fig. 1*A*), but the time between the two microsaccades allowed us to estimate intrinsic microsaccadic rhythms. As can be seen from the histogram, intrinsic microsaccadic rhythms might predict dominant oscillations of ~5 Hz or less, which is lower than any of the oscillations that we observed in our data (> 7 Hz; also see Fig. 5). *C*: microsaccade rate after fixation spot onset. In the interval >300 ms after fixation spot onset, microsaccade rate was constant. We chose to analyze only trials in which the last microsaccade before target onset occurred >300 ms after the beginning of any given trial, to make sure that the effects presented in this study were not phase locked to fixation spot onset. Similar results were obtained from *experiment 2*.

near ~200 ms, reflecting normal intermicrosaccadic intervals and a subsequent plateau reflecting the elevated likelihood of obtaining microsaccades throughout the remaining time interval. Consistent with this, intermicrosaccadic intervals in our data (Fig. 4*B*) peaked at ~200 ms and had a broad tail later. This means that the oscillations in Fig. 3*A* were not due to the occurrence of subsequent microsaccades. This was further confirmed when we considered the details of the intermicrosaccadic interval distribution shown in Fig. 4*B* more closely. According to this distribution, an oscillation dictated by when subsequent microsaccades are triggered would necessarily need to be ~5 Hz (the inverse of ~200 ms) or lower in frequency. However, and as we show in more detail next, our oscillations were significantly higher in frequency, where there would be no microsaccades (see region >7 Hz in Fig. 4*B*). Thus microsaccades have a persistent influence on RT modulations, revealing a coherent long lasting oscillation.

We analyzed the spectral properties of the RT oscillation by computing time-frequency plots in the range of 2–20 Hz. Critically, we performed data shuffling to investigate whether oscillatory rhythms in any given frequency band were expected by chance. Such data shuffling (described in detail in MATERIALS AND METHODS as well as in Fig. 2) gave us a set of surrogate traces, examples of which are shown in Fig. 3*B*, and our goal was to ask whether our original data (Fig. 3, *A* and *B*, *top*) was significantly different from these surrogate traces. In Fig. 3*C*, we only plotted time points after 100 ms (i.e., after the initial microsaccadic suppression interval), and we also only labeled time/frequency ranges that were statistically significant according to conservative criteria (Fig. 2 and see MATERIALS AND METHODS). RT oscillated in the α- and β-frequencies (i.e., within ~8–20 Hz) (Wang 2010); the oscillation gradually decreased in frequency, starting in the low-β/high-α range (i.e., ~10–20 Hz) and then finishing in the low-α range (i.e., ~8–10 Hz).

We also considered the possibility that the RT oscillation was not due to microsaccades but instead to the visual transient associated with fixation spot onset at trial beginning, which could globally reset the visual system. If microsaccades were temporally synchronized to such onset, then the RT oscillation could in reality reflect phase locking to the visual event and not to microsaccades. However, we only analyzed trials with >300 ms of steady fixation. This ensured that we excluded any microsaccades synchronized with fixation spot onset. Indeed, we did find that the likelihood of microsaccades was elevated in the first 300 ms after fixation spot onset (Fig. 4*C*), so

removing these trials was necessary to avoid possible ambiguities about whether phase locking was to the fixation spot onset or not. We also plotted RT time courses as a function of time of target onset from fixation spot onset and irrespective of microsaccades. We found no oscillation but instead a strong RT cost in the first ~500 ms followed by a return to baseline (data not shown); such early RT cost is to be expected given the increased microsaccade likelihood in this early period, since microsaccadic suppression is associated with RT increases.

We also performed DFT analysis on the RT time course of Fig. 3*A* in the period between 100 and 400 ms after microsaccade onset. We found significant power ($P < 0.05$; see MATERIALS AND METHODS) in the range of ~10–16 Hz consistent with Fig. 3*C*. Such significant power also existed when performing the DFT analysis for only trials with microsaccades less than the median amplitude (14.6-min arc) but not for microsaccades larger than the median amplitude (although a nonsignificant peak in a similar frequency range was still evident).

To summarize, our results so far suggest that microsaccades influence RT in a rhythmic fashion for several hundreds of milliseconds (Fig. 3), that this effect also holds for the smallest movement amplitudes, and that this effect cannot be explained by either microsaccade frequency or by visual transients associated with trial onsets (Fig. 4).

Sequential Hemifield Pulses of α- and β-Frequency RT Oscillations After Microsaccades

Even if visual transients associated with trial onsets were not responsible for our results, it could still be the case that the RT oscillation was due to visual transients associated with microsaccades themselves, since microsaccades shift and refresh retinal images. Additionally, a putative extra-retinal influence associated with movement triggering might contribute. In this case, microsaccade direction might have differential effects, since microsaccade generation would necessarily cause lateralized activation of the oculomotor system (Hafed 2011; Hafed et al. 2009, 2015; Hafed et al. 2009; Hafed and Krauzlis 2012; Krauzlis et al. 2017). We thus separated microsaccades according to the visual hemifield to which they were directed. For example, if a microsaccade was directed to the right hemifield, we asked how RT was modulated for rightward targets (presented in the "same hemifield" as the microsaccade) as opposed to leftward ("opposite") ones. Immediately after a microsaccade (and for up to ~100 ms), both "same" and "opposite" targets experienced increased RT as expected from microsaccadic suppression (Fig. 5, *A* and *D*). However, surprisingly, separating movement directions revealed that the coherent oscillation in Fig. 3 consisted of two separate oscillating "pulses" appearing sequentially in each hemifield, and these observations persisted even when considering only microsaccades less than a 30-min arc in amplitude (green traces in Fig. 5, *A* and *D*). From ~160 to 360 ms, a low-β/high-α-oscillation pulse appeared in the same hemifield as the microsaccade (Fig. 5*A*); for target onsets ~380–635 ms after microsaccades, a lower α-pulse appeared in the opposite hemifield (Fig. 5*D*). We obtained these specific frequency ranges by repeating the analyses highlighted in Fig. 2 for the present data subsets to assess statistical significance in the spectrotemporal domain. Specifically, we obtained surrogate traces, examples of which

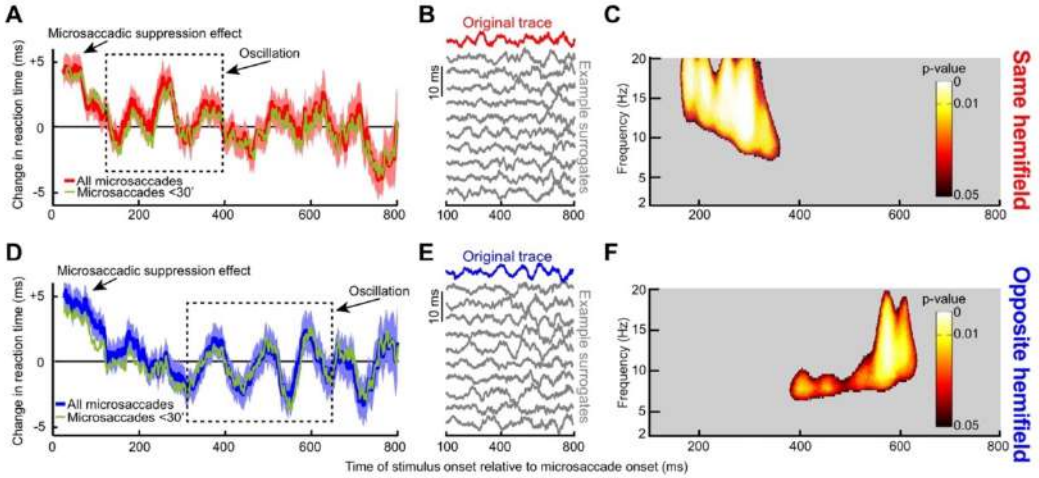


Fig. 5. Dependence of RT oscillations on microsaccade direction. *A*: same analysis as in Fig. 3*A* but only for trials with the target appearing in the same hemifield as a microsaccade. There was an initial microsaccadic suppression effect, followed by an oscillation. However, the oscillation only lasted for up to ~400 ms as indicated by the dashed rectangle, which highlights a pulse of statistically significant RT oscillation (see *C*). *B*: original and example surrogate traces from the data in *A*, as in Fig. 3*B*. *C*: same analysis as in Fig. 3*C* illustrating that the statistically significant RT oscillation pulse was restricted to only the first part of our sampled period after microsaccades. *D*–*F*: same as *A*–*C* but for targets appearing in the opposite hemifield. In this case, the RT oscillation pulse was delayed relative to *A*–*C* (dashed rectangle in *D*; also see *F* for the statistically significant times and frequencies). Thus, when separating microsaccade directions, we found that the long-term RT oscillation in Fig. 3 reflected sequential hemifield gating of RT oscillatory pulses first in the same hemifield as a microsaccade and then in the opposite hemifield. Moreover, the effect was consistent across individual subjects (Fig. 6). Green traces show the RT time courses for microsaccades less than a 30-min arc in amplitude.

are shown in Fig. 5, *B* and *E*, and we then computed significant time-frequency clusters shown in Fig. 5, *C* and *F*). These clusters confirm the sequential nature of the oscillations across hemifield that were evident in Fig. 5, *A* and *D*.

We next assessed the consistency of same- and opposite-hemifield RT oscillation pulses across individual subjects. In Fig. 6*A*, we plotted the phase of 13- to 20-Hz oscillations for an RT curve like that shown in Fig. 5*A* but now obtained from each subject individually. We found that the phase was largely consistent across subjects during the initial same-hemifield RT pulse of Fig. 5*A*, and it got scrambled later (Fig. 6*A*). This was further supported by plotting the PLV (see MATERIALS AND METHODS) of the phase traces shown in Fig. 6*A* (Fig. 6*B*). The shaded gray region in this case (Fig. 6*B*) indicates the time interval during which the PLV was statistically significant according to permutation tests with 1,000 surrogate analyses ($P < 0.05$; see MATERIALS AND METHODS). In other words, during the early same-hemifield RT pulse shown in Fig. 5, *A* and *C*, PLV in our real data from individual subjects consistently deviated from chance expectations obtained from the surrogate traces. We also performed similar analyses for the opposite-hemifield trials, this time within a frequency band (8–12 Hz) consistent with the opposite-hemifield pulse frequency range in Fig. 5, *D* and *F*. Once again, phase was consistent across individuals, but only in the late period in which the opposite-hemifield RT oscillation pulse occurred in Fig. 5, *D* and *F* (Fig. 6, *C* and *D*). Thus the opposite-hemifield RT oscillation pulse in Fig. 5, *D* and *F*, was consistently observed across individual subjects. These results, combined, reveal that microsaccade-phase-locked RT oscillations are dependent on movement direction.

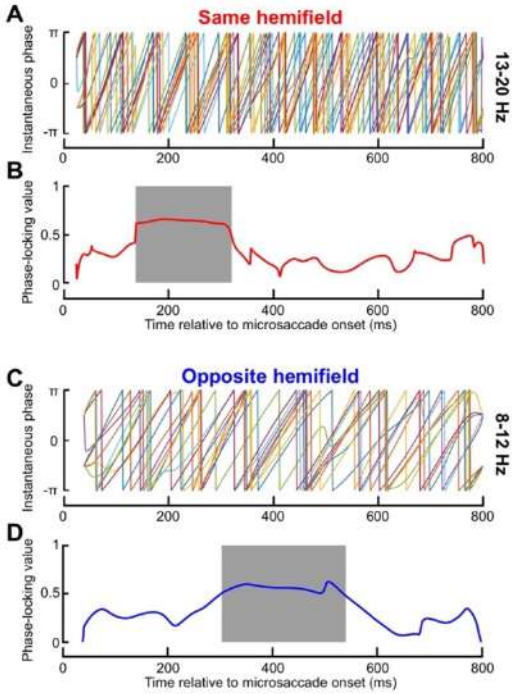


Fig. 6. Consistency of same- and opposite-hemifield RT oscillation pulses across individual subjects. *A*: phase of 13- to 20-Hz oscillations for each subject. Phase was largely consistent across subjects during the initial same-hemifield RT pulse of Fig. 5A, C, but it was inconsistent later. *B*: phase-locking value (PLV; see MATERIALS AND METHODS) of the phase traces shown in *A*. The shaded region shows the interval with significant PLV ($P < 0.05$; see MATERIALS AND METHODS), which was consistent with the early pulse in Fig. 5, *A* and *C*. Thus the results in Fig. 5, *A* and *C*, were reliable across individual subjects. *C*: similar analyses but for opposite hemifield trials. Once again, phase was consistent across individuals but only in the late period in which the opposite-hemifield RT oscillation pulse occurred in Fig. 5, *D* and *F*; also see *D*. Thus the opposite-hemifield RT oscillation pulse in Fig. 5, *D* and *F*, was consistently observed across individual subjects. *D*: same as in *B* but for the data in *C*.

Same-Hemifield Advantage in Perceptual Detection During the First α- and β-frequency Oscillatory RT Pulse

In *experiment 2* (Fig. 1*B*), we asked whether visual detection capabilities can also be affected long after microsaccades. We performed this experiment on 14 subjects who had to detect a small, briefly flashed target (see MATERIALS AND METHODS). Target contrast was continuously adjusted to maintain an average correct detection performance of ~50%. This means that subjects reported consciously seeing the target, and correctly localized it, on ~50% of the trials (see MATERIALS AND METHODS). Because this meant that the task was difficult, we placed visual placeholders to aid subjects in estimating the two possible target locations (Fig. 1*B*).

In Fig. 7*A*, *top*, we plotted the proportion of correct target detections as a function of time after microsaccades. Immediately after microsaccade onset, and for ~100 ms, subjects had immense difficulty in seeing the target; their likelihood of reporting that they saw the target and with correct localization was lower than the ~50% goal of this experiment's design (this means that they pressed "target not seen" more often than they pressed a seen target location). This is consistent with microsaccadic suppression. Interestingly, after such suppression, and specifically during the period in which the same-hemifield RT oscillation pulse occurred in Fig. 5, *A* and *C*, perceptual detection in the present experiment was also higher for the target appearing in the same vs. opposite hemifield as the microsaccade (dashed rectangle in Fig. 7*A*, *top*; $P = 0.0033$ in the interval 100–400 ms; 2-proportion *z*-test; also see Fig. 7*C*). This effect is remarkable because it emerged even though task difficulty was continuously titrated to maintain ~50% performance. This effect is also reminiscent of other observations made on shorter analysis time intervals than ours (Yuval-Greenberg et al. 2014), although it is not clear whether these authors analyzed performance during microsaccadic suppression or after it (Tian and Chen 2015). For later times, performance in both hemifields was similar (Fig. 7*A*, *top*; $P = 0.8035$ in the interval 400–800 ms; 2-proportion *z*-test). These results suggest that during the first RT oscillation pulse in the same hemifield observed in Fig. 5, perceptual detection capabilities were also improved (Fig. 7*A*). However, we did not find any statistically significant high-frequency oscillations in performance as we did for RT.

The early same-hemifield perceptual detection improvement (Fig. 7*A*) was also consistent across individuals (Fig. 8*A*), further suggesting that this effect was a robust property of

long-term microsaccadic influences. However, we have to emphasize here that we were unable to realize statistical significance on a per-subject basis due to the low numbers of trials per subject; nonetheless, all but three subjects showed trends consistent with our pooled observations (Fig. 8*A*).

We also ensured faithful task performance by the subjects despite task difficulty. For example, incorrect reports of target location were much less frequent than correct responses (Fig. 7*A*, *bottom*), and only 8% of catch trials had false alarms. This means that subjects followed the instruction to only report targets when they were consciously seen. Similarly, we ensured that target contrast was not different between same and opposite trials when performance was different. Specifically, in Fig. 8*B*, we compared target luminance contrast between the two sets of trials and confirmed that target contrast could not explain the early same-hemifield performance advantage observed in Fig. 7*A*.

It could also be suggested that the improved perceptual performance in Fig. 7*A* (dashed rectangle) was due to microsaccades bringing the fovea slightly closer to the peripheral target than with oppositely directed movements. Indeed, when we measured the position of the peripheral target relative to the fovea at the time of target onset in trials with improved perceptual performance (i.e., with targets 100–400 ms after microsaccades), we found that the target was slightly closer to the fovea on same than opposite trials (Fig. 7*B*, *left*), which was due to the eye becoming better aligned to the fixation spot (see Fig. 9). To ask whether this slight difference in retinotopic position was sufficient to explain the same-hemifield advantage, we performed two additional analyses. First, we only took trials in which target position overlapped between same and opposite trials (see MATERIALS AND METHODS), and we compared detection performance. The same-hemifield advantage was still present ($P = 0.034$ permutation test; see MATERIALS AND METHODS). Second, we reanalyzed the original data by now classifying trials according to the position of the target relative to the fovea rather than according to microsaccade direction (Fig. 7*B*, *right*); we performed a median split on Euclidean distance between target and fovea, and we compared "close" vs. "far" target positions, regardless of microsaccade direction. We no longer observed a difference in performance (Fig. 7*C*), suggesting that our results could not be explained by the target being closer to the fovea on same vs. opposite microsaccade trials. Consistent with this, we still observed the same-hemifield advantage when restricting our analyses to only trials with microsaccades less than a 30-min arc in amplitude and even those with microsaccades less than the median amplitude of a

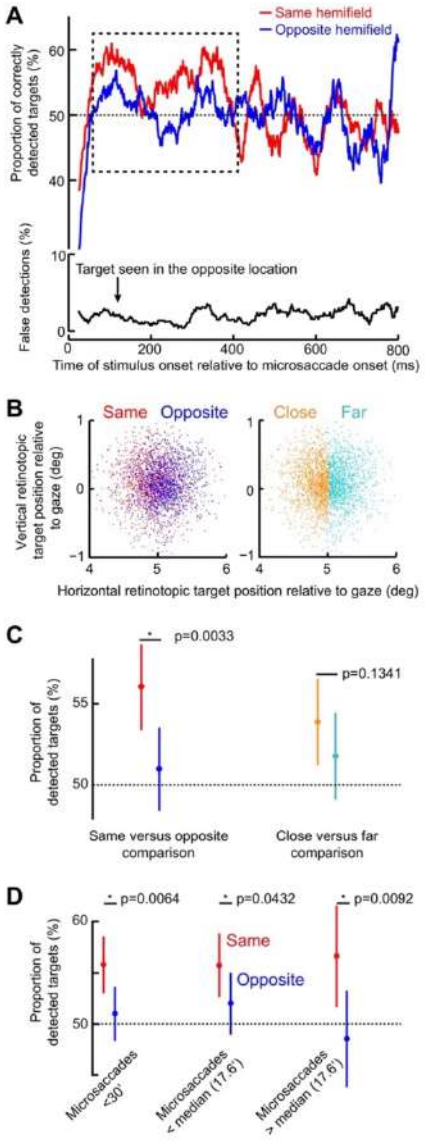


Fig. 7. Same-hemifield advantage in detectability during the period of the initial RT pulse. *A*, *top*: mean detectability (i.e., correct responses) as a function of time after microsaccade onset. Immediately after microsaccades, detectability was impaired, consistent with microsaccadic suppression effects (see MATERIALS AND METHODS). However, in the dashed rectangle, which corresponds closely with the same-hemifield RT pulse in Fig. 5, *A* and *C*, detectability was significantly higher for targets appearing in the same hemifield as a microsaccade than for opposite targets ($P = 0.0033$; 2-proportion *z*-test). *A*, *bottom*: incorrect target localizations, which were rare and unmodulated by microsaccadic suppression, suggesting that subjects followed task instructions (also see RESULTS). *B*: retinotopic target position at the time of target onset for trials in the early (100–400 ms) interval after microsaccade onset (each dot is a trial). Positions are colored based on the direction of the last microsaccade relative to the target (*left*) or whether the target was closer or farther than the median Euclidean distance observed. *C*: mean detectability during the early (100–400 ms) period. The *left pair* compares conditions when the target was presented in the same or opposite hemifield relative to a microsaccade. The *right pair* compares closer or farther targets from the fovea. Detectability was only enhanced in the same hemifield condition (*left*) and did not depend on retinotopic target position (*right*). *D*: mean detectability as in *C* but now only when microsaccades were less than a 30-min arc in amplitude or when they were smaller or bigger than the median. In all cases, the same-hemifield advantage of *A* was observed. Error bars denote 95% confidence intervals.

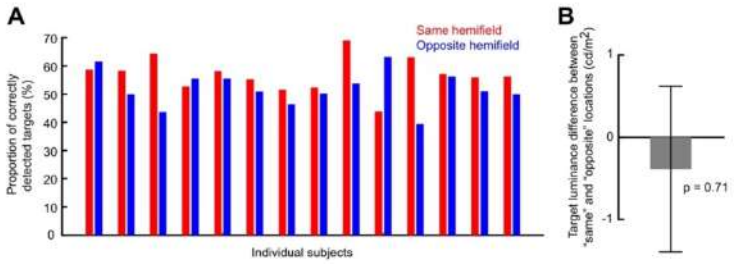


Fig. 8. Individual subject data and same vs. opposite target contrasts in *experiment 2*. *A*: mean detectability for each subject and each target hemifield location relative to microsaccade direction in the perceptual detection task. This analysis is restricted to the early period (100–400 ms) after microsaccade onset, which is the period in which we observed differential detection performance across hemifields (Fig. 7). Note that 11 out of the 14 subjects detected the target better when it was presented in the same hemifield as the microsaccade, consistent with the overall population average results in Fig. 7. Thus no outlier subject minority was responsible for the differences observed in Fig. 7 when the data were averaged across all participants. *B*: difference in target contrast between same and opposite trials in the early period. There was no statistically significant difference ($P = 0.71$; *t*-test), suggesting that the performance differences in Fig. 7 were not due to differences in target contrast. Error bar denotes SE.

17.6-min arc (Fig. 7*D*). For these trials, retinotopic target position was very similar between same and opposite trials.

Finally, we wondered whether our results could reflect microsaccades being pulled peripherally by a rhythmic process of covert peripheral visual attention, as suggested by recent evidence (Busch et al. 2009; Busch and VanRullen 2010; Dugué et al. 2015; Fiebelkorn et al. 2013; Landau and Fries 2012; Landau et al. 2015; Song et al. 2014). We analyzed microsaccade directions relative to the fixation spot (Fig. 9) and confirmed previous findings that microsaccades redirect gaze to the spot and not the periphery (Buonocore et al. 2017; Guerrasio et al. 2010; Ko et al. 2010; Tian et al. 2016). That is, eye position was closer to the fixation spot after microsaccades than before them (Fig. 9). Thus our results are likely due to microsaccade generation itself. Of course, some relation to peripheral covert attention might still be expected (Hafed 2013; Tian et al. 2016) given our results, as we discuss in more detail in DISCUSSION.

Postsuppression Enhancement in SC Visual Sensitivity After Microsaccades

Finally, because our *experiment 1* was motivated by our recent SC studies relating perimicrosaccadic visual sensitivity in this brain structure to RT (Chen and Hafed 2017; Hafed and Krauzlis 2010), we revisited these studies from the perspective of what we have learned so far from the present results. We specifically hypothesized that RT oscillations should correlate with visual sensitivity even in the absence of any overt response. In our previous SC studies, visual response strength was the only assay of microsaccadic suppression, and monkeys did not perform any task other than fixation (Chen and Hafed 2017; Hafed and Krauzlis 2010). SC visual response strength in these studies correlated remarkably well with RT effects obtained from completely different behavioral sessions (Chen and Hafed 2017). This means that if we were to record SC visual responses in the absence of any behavioral task, then a faster RT following the initial microsaccadic suppression effect (e.g., Fig. 3) should be mirrored by stronger visual responses than "baseline" (with "baseline" defined as response strength in the absence of any nearby microsaccades; see MATERIALS AND METHODS). We thus analyzed visual response strength after microsaccades in two monkeys, using data from Chen et al. (2015), by extending the time course of analysis beyond the initial 100 ms that we typically used in our earlier studies. For up to ~100 ms after microsaccade onset, visual response strength in both visual and visual-motor SC neurons was suppressed, as expected (Fig. 10). Specifically, the curves in Fig. 10 were below the normalized "baseline" response value of 1 for all time points less than that indicated by the dashed vertical line (error bars denote 95% confidence intervals).

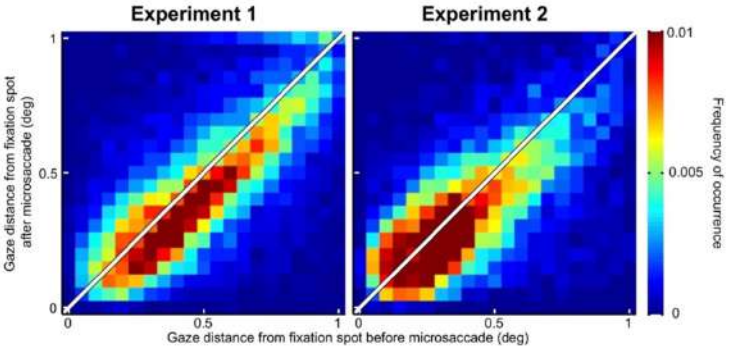


Fig. 9. Microsaccades acted to primarily correct for foveal motor errors during gaze fixation. We plotted gaze distance from the fixation spot before and after microsaccades, for all microsaccades that were used in the analyses of Figs. 3, 5, and 7. During both experiments, microsaccades brought gaze closer to the fixation spot than before the movements ($P = 1.2111 \times 10^{-54}$ in *experiment 1* and $P = 6.9096 \times 10^{-29}$ in *experiment 2*; Wilcoxon rank sum test), consistent with recent observations in previous studies. This suggests that our microsaccades were not necessarily reflecting endogenous attention shifts towards the periphery but were instead part of a deliberate oculomotor strategy to optimize eye position on the fixated target.

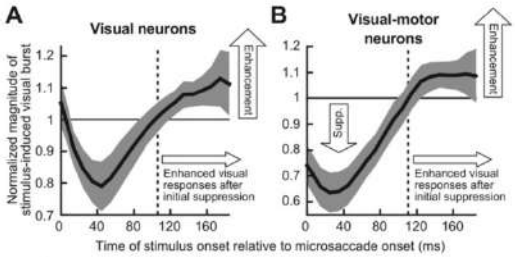


Fig. 10. Neural implications of an RT oscillation after initial microsaccadic suppression. *A*: we performed time course analyses of visual response strength in superior colliculus (SC) purely visual neurons. The neurons were the same as those used in Chen et al. (2015), and the analysis was identical to that performed in Hafed and Krauzlis (2010), except that we extended the analysis window beyond the typical 100 ms after microsaccade onset that we had used earlier (see MATERIALS AND METHODS). This allowed us to explore whether recovery from neural microsaccadic suppression is towards a constant baseline or not, as we did in the behavioral experiment of Fig. 1. Error bars denote 95% confidence intervals. As can be seen, after the initial ~100 ms of neural suppression, visual response strength was enhanced (i.e., the curve went above 1), rather than being equal to "baseline," suggesting that RT oscillations like in Fig. 3 can reflect oscillations in visual neural sensitivity. *B*: a similar observation was made for visual sensitivity of visual-motor SC neurons, which are better correlated to RT (Chen and Hafed 2017; Hafed and Krauzlis 2010). In *A* and *B*, the "baseline" response (normalized to 1) was obtained from trials in which there were no microsaccades within ±150 ms from stimulus onset. Note that this data set did not allow us to sample neural sensitivity for longer periods after a microsaccade (as in Fig. 3), but the postsuppression enhancement nonetheless suggests that visual neural fluctuations can also go above baseline, consistent with the Fig. 3 RT oscillation.

Interestingly, visual response strength in both visual and visual-motor SC neurons indeed increased above baseline after this initial microsaccadic suppression (Fig. 10; curves lying above 1 in the interval after the vertical dashed line). Thus, like RT costs, recovery from neural microsaccadic suppression was also not back to a single "baseline," but there was higher visual sensitivity after the initial suppression. These neural analyses did not allow us to extend the time course for hundreds of milliseconds like in our behavioral data, because our earlier neural experiments did not sample long times after microsaccades (Chen et al. 2015); however, they nonetheless demonstrate that our RT effects in Fig. 3 may be related to visual sensitivity modulations after microsaccades, as suggested earlier (Chen and Hafed 2017; Hafed et al. 2015; Hafed and Krauzlis 2010; Tian et al. 2016). This is also similar to behavioral results with large saccades showing enhancement after the suppression (Burr et al. 1994; Diamond et al. 2000; Knöll et al. 2011). We also saw hints of direction dependence in the post-suppression enhancement of neural activity in Fig. 10, similar to Fig. 5, but the data (constrained by prior neural recordings) were not sufficient to allow us to establish statistical significance.

DISCUSSION

We found that microsaccade occurrence has a profound long-term influence on brain state, resulting in coherent pulsed α/β-RT oscillations first in the same visual hemifield as the movement vector and then in the opposite hemifield. We also found long-term changes in perceptual detection and SC neural activity. In what follows, we discuss these observations and relate them to observations of physiological rhythms in the brain, attentional fluctuations, and motor control in general. We also discuss our methodological choices and their potential limitations.

We think that the lateralization of the two oscillatory pulses that we observed in RT (Fig. 5) is likely jumpstarted by microsaccade movement commands, which necessarily result in lateralized action-potential bursts in spatially organized visual-motor structures like the SC (Hafed et al. 2009; Hafed and Krauzlis 2012). Interestingly, the switch to the opposite hemifield RT oscillation that we observed in Fig. 5, *D–F*, came at approximately the time at which one would normally make a second eye movement during natural scene viewing and also during fixation in the case of microsaccades (e.g., Fig. 4*B*). It would be interesting to investigate the implications of such hemifield switching under more natural conditions. For example, there is evidence for "facilitation of return" in saccadic scanning of natural scenes (Wilming et al. 2013), in which "return saccades" are frequently observed (i.e., a saccade occurs opposite in direction to a previous one). It might be the case that the brain mechanisms underlying the sequential hemifield switches that we have observed could be related to such increases in the propensity to make return saccades. This can also apply in the case of scanning objects in far environments, in which microsaccades would be necessary given the small images projected by far visual features onto the retina. Thus the results that we obtained in a laboratory setting with a fixation marker can potentially be extended to more naturalistic scenarios.

We also think that our results may be related to observations of neural oscillatory patterns in a variety of visual areas. For example, it was previously shown that microsaccades cause broadband modulation of visual areas immediately after movements (Bosman et al. 2009; Lowet et al. 2016). This is consistent with image refreshing (Martinez-Conde et al. 2000), but there could also be longer term oscillations in neural activity, and future work should explore the links between neural and behavioral oscillations in relation to microsaccades. Currently, such links may not always be obvious, especially because during parts of the longer term intervals that we have focused on here, some visual areas were shown previously to exhibit low frequency oscillations and not higher frequency ones like we saw. For example, the neural oscillations that were reported previously were primarily consistent with microsaccade frequency (Bosman et al. 2009) and not necessarily with the higher frequencies that we observed in our data. Similarly, if neural oscillation results were indeed to turn out to be linked to our behavioral observations, then we think that longer term analyses of neural data (say, at 600 ms or more after a movement) need to take microsaccade direction into account. For example, the neurophysiological experiments of Bosman et al. (2009) did not separate different movement directions, and this could be why oscillatory patterns may not have lasted for too long (if there is sequential hemifield pulsing but all movements are combined, then such pulsing might be masked). Indeed, even at the neural level, movement direction does seem to matter a great deal for large saccades, which are associated with direction-dependent β-frequency waves in area V4 (Zanos et al. 2015). α-Oscillations have also been observed in V4 after large saccades (Zanos et al. 2016), but these

oscillations were once again short-lived because potential later oscillations in the opposite hemifield were not investigated.

Related to the above, we observed α- and β-behavioral oscillations in RT (Figs. 3 and 5). The β-oscillations that we observed might be a lingering component of β-rhythms associated with microsaccade generation, since β-rhythms can be movement-related (Wang 2010). On the other hand, the α-oscillations may be related to occipital α-rhythms, which synchronize to microsaccades (Gaarder et al. 1966). Of course, this requires an assumption that occipital modulations can manifest in RT modulations, but this is quite reasonable. In fact, even manual RT is a highly sensitive measure of properties of the early visual system, like magno- and parvo-cellular pathways (Breitmeyer 1975). Similarly, the early evoked visual response of SC neurons is highly predictive of RT's collected from separate sessions (Chen and Hafed 2017). In this regard, such sensitivity of RT could be why it was easier to see significant oscillations in *experiment 1* than in *experiment 2*. Indeed, maybe more data would have revealed significant oscillations also in *experiment 2*, but this itself would indicate that the effect is less robust. The visual stimuli were also very different in both experiments, so it could be that our task in *experiment 2* was not sensitive enough to behaviorally manifest the underlying brain oscillations. It is nonetheless still interesting that the same-hemifield advantage in Fig. 7 emerged at the same time as the same-hemifield RT oscillation pulse in Fig. 5.

The above statements suggest that the implications of our findings are not necessarily that RT or detection effects in our present scenarios are the ultimate goal of microsaccade-related synchronization of endogenous brain rhythms. Rather, we think that our behavioral assays were sensitive measures of underlying oscillations, which could be useful in a variety of other ways. For example, oscillations could increase the efficacy of transmitting information between areas (for example, from the lateral geniculate nucleus to primary visual cortex) (Bastos et al. 2014). This could happen, say, through the so-called idea of "communication through coherence" (Fries 2005). Additionally, the phase of oscillations slower than 20 Hz could be used to potentially increase the information conveyed by each spike in V1 (Montemurro et al. 2008), and α-oscillations could order and prioritize salient information from the unattended visual field (Jensen et al. 2012). In this regard, such oscillations could modulate when and how neural assemblies might inhibit other assemblies. Overall, this could help in allocating resources to "sample" different regions of an image. At the circuit level, oscillations can also emerge out of known microcircuit architecture in cortex with local inhibitory loops, so they also represent structure-function relationships in brain operation.

In our experiments, we also observed higher oscillation frequencies than those predicted by attentional sampling at the behavioral level (Fiebelkorn et al. 2013; Landau and Fries 2012). However, our results are still consistent with hypotheses about perceptual rhythms in the visual system (Busch et al. 2009; Busch and VanRullen 2010; VanRullen 2016), and we think that it is quite reasonable that these rhythms would be related to saccades and microsaccades, especially given how active visual perception is under natural conditions.

With respect to attention, it could additionally be argued that covert attention might have been the source of the RT oscillations and same-hemifield detection advantages that we have observed, independently of microsaccades. Indeed, attention samples locations rhythmically (Busch et al. 2009; Busch and VanRullen 2010; Dugué et al. 2015; Fiebelkorn et al. 2013; Landau and Fries 2012; Landau et al. 2015; Song et al. 2014), and microsaccades are correlated with attention (Engbert and Kliegl 2003; Hafed and Clark 2002; Hafed et al. 2011). However, time locking of our effects to microsaccades would require that the time and direction of any given movement would have to perfectly match the time and direction of every single attention shift or at least within a small window associated with β-rhythms (<50 ms). Moreover, we found that microsaccades correct for eye position error (Fig. 9), consistent with previous findings (Guerrasio et al. 2010; Ko et al. 2010; Tian et al. 2016), rather than increase such error, which is what would happen if they were pulled peripherally by oscillations in the locus of peripheral covert attention. Instead, we think that our results are more likely related to well-known perimicrosaccadic influences on visual sensitivity, as in the case of microsaccadic (Chen et al. 2015; Hafed et al. 2015; Hafed and Krauzlis 2010; Zuber and Stark 1966) and saccadic (Benedetto and Morrone 2017) suppression. However, even with this view, a relation to attention can still be expected (Hafed 2013; Hafed et al. 2015; Tian et al. 2016). For example, since microsaccades are reflexively reset by sensory cues (Hafed and Ignashchenkova 2013; Rolfs et al. 2008), it may be possible based on our current results that some behavioral effects in attention tasks may be partially influenced by microsaccades occurring several hundreds of milliseconds before a target.

It was also recently found that large voluntary saccades are associated with behavioral performance oscillations like in our results (Benedetto and Morrone 2017; Hogendoorn 2016; Wutz et al. 2016). However, an intriguing difference between these observations and ours is that our oscillations were higher in frequency than in these studies. Indeed, the lower frequency oscillations associated with large voluntary saccades may be related either to attentional rhythms (Hogendoorn 2016) or the intrinsic saccadic/microsaccadic rhythms of the oculomotor system (Benedetto and Morrone 2017), which are both slower than the α- and β-frequencies that we observed. Since microsaccades cannot be generally made at such high frequencies (e.g., Fig. 4*B*), this leads us to hypothesize that microsaccades may "ride" on slow frequency rhythms, as we have recently suggested (Tian et al. 2016), and as is the case with the large saccade results of Benedetto and Morrone (2017) and Hogendoorn (2016), but that they reset the phase of higher frequency oscillations. Such resetting could serve the purpose of regularizing visual processing in between two successive movements (Chen and Hafed 2017), particularly with the sequential hemifield pulsing that we observed. Interestingly, our DFT analysis did not reveal significant ~10- to 16-Hz oscillations for microsaccades larger than the median amplitude, but there was also a secondary peak at ~4–5 Hz (although still non-significant) emerging for these larger microsaccades. It would be interesting to investigate the interactions between multiple frequency bands and saccade/microsaccade sizes at the mechanistic level in future research.

Finally, our observation of coherent effects when pooling multiple subjects' data suggests that our observations were common across individuals, even if there may have been individual differences in more subtle ways (in, say, the exact

values of α- or β-oscillation frequencies that each individual might exhibit). We find both this observation and the fact that the coherent oscillations appeared at the behavioral level (i.e., at the overall output of the entire brain processing cascade) intriguing. In this sense, a small effect size (e.g., for the absolute RT difference between peak and trough in Fig. 3*A*) is to be perfectly expected after pooling all subjects together; what is more interesting is that such a coherent effect occurs at all in the first place. Also, amplitudes of behavioral oscillations at the high frequencies that we studied are generally mild even in earlier studies (Song et al. 2014). Thus it is the phase consistency after pooling that is interesting here, rather than the effect size itself. Having said that, it may still be argued that an across-subject average oscillation can emerge even if individual subjects had no oscillations at all, but if they instead each somehow had a "strategically timed" single peak in RT (to give rise to the highly specific results in Figs. 3 and 5 after pooling). This is both physiologically and statistically unlikely, but we explored it further by running one of our monkeys on multiple sessions of a version of *experiment 1* to obtain reliable within-subject RT time courses. The results that we obtained (Fig. 11), which are part of an ongoing follow-up investigation of neural mechanisms, confirmed that multiple RT peaks can indeed appear within an individual subject.

The above points touch on the importance of statistical analyses in studies like ours. In our approach, we pooled subjects but this was because of the daunting task of data collection (e.g., the number of trials per time bin in Fig. 3*A* was an order of magnitude smaller than the total number of trials collected because of factors related to Fig. 4). However, we were careful to correct for multiple comparisons and other factors in analyses, especially because running average operations can cause ringing in traces that may appear to be an oscillation visually even when there is no oscillation (e.g., see some surrogate traces in Figs. 3*B* and 5, *B* and *E*). We find it unlikely that our results were entirely explained by ringing, especially given the sequential hemifield pulsing that we observed in Fig. 5 (it is not clear how filtering-related ringing artifacts can give rise to such a finding). However, we acknowledge that our filtering choices restricted our analyses to frequencies <20 Hz. It may be possible that higher frequencies are present, but this remains to be seen in future experiments.

Overall, our results bridge action, perception, and rhythmic brain fluctuations in a manner that should helpfully inform many exciting future neurophysiological experiments in vision science.

ACKNOWLEDGMENTS

We are grateful for helpful comments on the manuscript from three peer reviewers.

GRANTS

We were funded by the Werner Reichardt Centre for Integrative Neuroscience (CIN), an Excellence Cluster (EXC307) funded by the Deutsche Forschungsgemeinschaft (DFG). We were also supported by the Hertie Institute for Clinical Brain Research.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the authors.

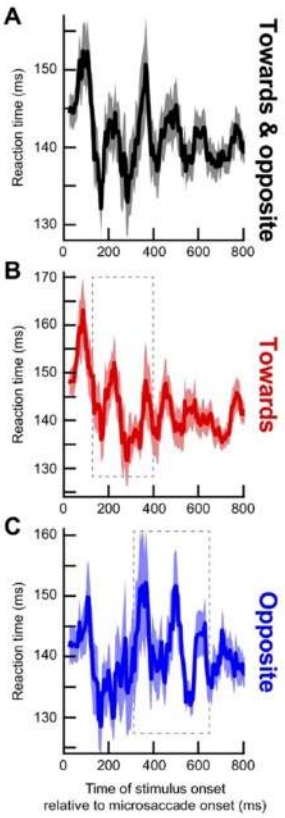


Fig. 11. RT fluctuations within an individual monkey subject. We ran *monkey N* on 11 sessions of a version of *experiment 1*. A target like in *experiment 1* could appear at random times after microsaccades, but it could appear at one of eight different equally spaced directions. We defined microsaccades as being towards or opposite the target if their directions were within ±22.5° from the direction of the target or the direction diametrically opposite it. *A*: an analysis like that in Fig. 3 (with a 50-ms running window in 5-ms steps) revealed RT fluctuations similar to those in Fig. 3, with multiple peaks at different times; thus multiple peaks of RT variability can indeed happen within an individual subject. Discrete Fourier transform analysis on the shown interval revealed a spectral peak at ~7.7 Hz. *B* and *C*: we repeated the analysis for cases in which the target appeared in a direction congruent with microsaccade direction ("Towards") or opposite it ("Opposite"). The highest amplitude RT fluctuations occurred early for "Towards" microsaccades and later for "Opposite" movements, just like in Fig. 5. As a reference, the thin gray rectangles show the time intervals in which our average analyses in humans showed significant oscillatory pulses in similar conditions (Fig. 5). This monkey showed similar patterns to the pooled human data. Error bars denote SE across trials, and the figure is formatted similarly to Figs. 3 and 5.

AUTHOR CONTRIBUTIONS

J.B. and Z.M.H. conceived and designed research; J.B. and Z.M.H. performed experiments; J.B., C.-Y.C., and Z.M.H. analyzed data; J.B. and Z.M.H. interpreted results of experiments; J.B. and Z.M.H. prepared figures; J.B., C.-Y.C., and Z.M.H. edited and revised manuscript; J.B., C.-Y.C., and Z.M.H. approved final version of manuscript; Z.M.H. drafted manuscript.

REFERENCES

Bastos AM, Briggs F, Alitto HJ, Mangun GR, Usrey WM. Simultaneous recordings from the primary visual cortex and lateral geniculate nucleus reveal rhythmic interactions and a cortical source for γ-band oscillations. *J Neurosci* 34: 7639–7644, 2014. doi:10.1523/JNEUROSCI.4216-13.2014.

Benedetto A, Morrone MC. Saccadic suppression is embedded within extended oscillatory modulation of sensitivity. *J Neurosci* 37: 3661–3670, 2017. doi:10.1523/JNEUROSCI.2390-16.2016.

Benedetto A, Spinelli D, Morrone MC. Rhythmic modulation of visual contrast discrimination triggered by action. *Proc Biol Sci* 283: 283, 2016. doi:10.1098/rspb.2016.0692.

Bosman CA, Womelsdorf T, Desimone R, Fries P. A microsaccadic rhythm modulates gamma-band synchronization and behavior. *J Neurosci* 29: 9471–9480, 2009. doi:10.1523/JNEUROSCI.1193-09.2009.

Breitmeyer BG. Simple reaction time as a measure of the temporal response properties of transient and sustained channels. *Vision Res* 15: 1411–1412, 1975. doi:10.1016/0042-6989(75)90200-X.

Buonocore A, Chen CY, Tian X, Idrees S, Muench T, Hafed ZM. Alteration of the microsaccadic velocity-amplitude main sequence relationship after visual transients: implications for models of saccade control. *J Neurophysiol* 117: 1894–1910, 2017. doi:10.1152/jn.00811.2016.

Burr DC, Morrone MC, Ross J. Selective suppression of the magnocellular visual pathway during saccadic eye movements. *Nature* 371: 511–513, 1994. doi:10.1038/371511a0.

Busch NA, Dubois J, VanRullen R. The phase of ongoing EEG oscillations predicts visual perception. *J Neurosci* 29: 7869–7876, 2009. doi:10.1523/JNEUROSCI.0113-09.2009.

Busch NA, VanRullen R. Spontaneous EEG oscillations reveal periodic sampling of visual attention. *Proc Natl Acad Sci USA* 107: 16048–16053, 2010. doi:10.1073/pnas.1004801107.

Chen CY, Hafed ZM. Postmicrosaccadic enhancement of slow eye movements. *J Neurosci* 33: 5375–5386, 2013. doi:10.1523/JNEUROSCI.3703-12.2013.

Chen CY, Hafed ZM. A neural locus for spatial-frequency specific saccadic suppression in visual-motor neurons of the primate superior colliculus. *J Neurophysiol* 117: 1657–1673, 2017. doi:10.1152/jn.00911.2016.

Chen CY, Ignashchenkova A, Thier P, Hafed ZM. Neuronal response gain enhancement prior to microsaccades. *Curr Biol* 25: 2065–2074, 2015. doi:10.1016/j.cub.2015.06.022.

Cherici C, Kuang X, Poletti M, Rucci M. Precision of sustained fixation in trained and untrained observers. *J Vis* 12: 12, 2012. doi:10.1167/12.6.31.

Diamond MR, Ross J, Morrone MC. Extraretinal control of saccadic suppression. *J Neurosci* 20: 3449–3455, 2000.

Dugué L, McLelland D, Lajous M, VanRullen R. Attention searches nonuniformly in space and in time. *Proc Natl Acad Sci USA* 112: 15214–15219, 2015. doi:10.1073/pnas.1511331112.

Engbert R, Kliegl R. Microsaccades uncover the orientation of covert attention. *Vision Res* 43: 1035–1045, 2003. doi:10.1016/S0042-6989(03)00084-1.

Fiebelkorn IC, Saalmann YB, Kastner S. Rhythmic sampling within and between objects despite sustained attention at a cued location. *Curr Biol* 23: 2553–2558, 2013. doi:10.1016/j.cub.2013.10.063.

Fries P. A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends Cogn Sci* 9: 474–480, 2005. doi:10.1016/j.tics.2005.08.011.

Gaarder K, Koresko R, Kropfl W. The phasic relation of a component of alpha rhythm to fixation saccadic eye movements. *Electroencephalogr Clin Neurophysiol* 21: 544–551, 1966. doi:10.1016/0013-4694(66)90173-8.

Goffart L, Hafed ZM, Krauzlis RJ. Visual fixation as equilibrium: evidence from superior colliculus inactivation. *J Neurosci* 32: 10627–10636, 2012. doi:10.1523/JNEUROSCI.0696-12.2012.

Guerrasio L, Quinet J, Büttner U, Goffart L. Fastigial oculomotor region and the control of foveation during fixation. *J Neurophysiol* 103: 1988–2001, 2010. doi:10.1152/jn.00771.2009.

Hafed ZM. Mechanisms for generating and compensating for the smallest possible saccades. *Eur J Neurosci* 33: 2101–2113, 2011. doi:10.1111/j.1460-9568.2011.07694.x.

Hafed ZM. Alteration of visual perception prior to microsaccades. *Neuron* 77: 775–786, 2013. doi:10.1016/j.neuron.2012.12.014.

Hafed ZM, Chen CY, Tian X. Vision, perception, and attention through the lens of microsaccades: mechanisms and implications. *Front Syst Neurosci* 9: 167, 2015. doi:10.3389/fnsys.2015.00167.

Hafed ZM, Clark JJ. Microsaccades as an overt measure of covert attention shifts. *Vision Res* 42: 2533–2545, 2002. doi:10.1016/S0042-6989(02)00263-8.

Hafed ZM, Goffart L, Krauzlis RJ. A neural mechanism for microsaccade generation in the primate superior colliculus. *Science* 323: 940–943, 2009. doi:10.1126/science.1166112.

Hafed ZM, Ignashchenkova A. On the dissociation between microsaccade rate and direction after peripheral cues: microsaccadic inhibition revisited. *J Neurosci* 33: 16220–16235, 2013. doi:10.1523/JNEUROSCI.2240-13.2013.

Hafed ZM, Krauzlis RJ. Microsaccadic suppression of visual bursts in the primate superior colliculus. *J Neurosci* 30: 9542–9547, 2010. doi:10.1523/JNEUROSCI.1137-10.2010.

Hafed ZM, Krauzlis RJ. Similarity of superior colliculus involvement in microsaccade and saccade generation. *J Neurophysiol* 107: 1904–1916, 2012. doi:10.1152/jn.01125.2011.

Hafed ZM, Lovejoy LP, Krauzlis RJ. Modulation of microsaccades in monkey during a covert visual attention task. *J Neurosci* 31: 15219–15230, 2011. doi:10.1523/JNEUROSCI.3106-11.2011.

Hafed ZM, Lovejoy LP, Krauzlis RJ. Superior colliculus inactivation alters the relationship between covert visual attention and microsaccades. *Eur J Neurosci* 37: 1169–1181, 2013. doi:10.1111/ejn.12127.

Havermann K, Cherici C, Rucci M, Lappe M. Fine-scale plasticity of microscopic saccades. *J Neurosci* 34: 11665–11672, 2014. doi:10.1523/JNEUROSCI.5277-13.2014.

Herrington TM, Masse NY, Hachmeh KJ, Smith JE, Assad JA, Cook EP. The effect of microsaccades on the correlation between neural activity and behavior in middle temporal, ventral intraparietal, and lateral intraparietal areas. *J Neurosci* 29: 5793–5805, 2009. doi:10.1523/JNEUROSCI.4412-08.2009.

Hogendoorn H. Voluntary saccadic eye movements ride the attentional rhythm. *J Cogn Neurosci* 28: 1625–1635, 2016. doi:10.1162/jocn_a_00986.

Jensen O, Bonnefond M, VanRullen R. An oscillatory mechanism for prioritizing salient unattended stimuli. *Trends Cogn Sci* 16: 200–206, 2012. doi:10.1016/j.tics.2012.03.002.

Kliegl R, Rolfs M, Laubrock J, Engbert R. Microsaccadic modulation of response times in spatial attention tasks. *Psychol Res* 73: 136–146, 2009. doi:10.1007/s00426-008-0202-2.

Knöll J, Binda P, Morrone MC, Bremmer F. Spatiotemporal profile of peri-saccadic contrast sensitivity. *J Vis* 11: 15, 2011. doi:10.1167/11.14.15.

Ko HK, Poletti M, Rucci M. Microsaccades precisely relocate gaze in a high visual acuity task. *Nat Neurosci* 13: 1549–1553, 2010. doi:10.1038/nn.2663.

Krauzlis RJ, Goffart L, Hafed ZM. Neuronal control of fixation and fixational eye movements. *Philos Trans R Soc Lond B Biol Sci* 372: 20160205, 2017. doi:10.1098/rstb.2016.0205.

Lachaux JP, Rodriguez E, Martinerie J, Varela FJ. Measuring phase synchrony in brain signals. *Hum Brain Mapp* 8: 194–208, 1999. doi:10.1002/(SICI)1097-0193(1999)8:4<194::AID-HBM4>3.0.CO;2-C.

Landau AN, Fries P. Attention samples stimuli rhythmically. *Curr Biol* 22: 1000–1004, 2012. doi:10.1016/j.cub.2012.03.054.

Landau AN, Schreyer HM, van Pelt S, Fries P. Distributed attention is implemented through theta-rhythmic gamma modulation. *Curr Biol* 25: 2332–2337, 2015. doi:10.1016/j.cub.2015.07.048.

Laubrock J, Engbert R, Kliegl R. Microsaccade dynamics during covert attention. *Vision Res* 45: 721–730, 2005. doi:10.1016/j.visres.2004.09.029.

Le Van Quyen M, Foucher J, Lachaux J, Rodriguez E, Lutz A, Martinerie J, Varela FJ. Comparison of Hilbert transform and wavelet methods for the analysis of neuronal synchrony. *J Neurosci Methods* 111: 83–98, 2001. doi:10.1016/S0165-0270(01)00372-7.

Lowet E, Roberts MJ, Bosman CA, Fries P, De Weerd P. Areas V1 and V2 show microsaccade-related 3-4-Hz covariation in gamma power and frequency. *Eur J Neurosci* 43: 1286–1296, 2016. doi:10.1111/ejn.13126.

Maris E, Oostenveld R. Nonparametric statistical testing of EEG- and MEG-data. *J Neurosci Methods* 164: 177–190, 2007. doi:10.1016/j.jneumeth.2007.03.024.

Martinez-Conde S, Macknik SL, Hubel DH. Microsaccadic eye movements and firing of single cells in the striate cortex of macaque monkeys. *Nat Neurosci* 3: 251–258, 2000. doi:10.1038/72961.

Montemurro MA, Rasch MJ, Murayama Y, Logothetis NK, Panzeri S. Phase-of-firing coding of natural visual stimuli in primary visual cortex. *Curr Biol* 18: 375–380, 2008. doi:10.1016/j.cub.2008.02.023.

Oostenveld R, Fries P, Maris E, Schoffelen JM. FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput Intell Neurosci* 2011: 156869, 2011. doi:10.1155/2011/156869.

Peel TR, Hafed ZM, Dash S, Lomber SG, Corneil BD. A causal role for the cortical frontal eye fields in microsaccade deployment. *PLoS Biol* 14: e1002531, 2016. doi:10.1371/journal.pbio.1002531.

Poletti M, Listorti C, Rucci M. Microscopic eye movements compensate for nonhomogeneous vision within the fovea. *Curr Biol* 23: 1691–1695, 2013. doi:10.1016/j.cub.2013.07.007.

Poletti M, Rucci M. A compact field guide to the study of microsaccades: Challenges and functions. *Vision Res* 118: 83–97, 2016. doi:10.1016/j.visres.2015.01.018.

Rolfs M, Kliegl R, Engbert R. Toward a model of microsaccade generation: the case of microsaccadic inhibition. *J Vis* 8: 1–23, 2008. doi:10.1167/8.11.5.

Song K, Meng M, Chen L, Zhou K, Luo H. Behavioral oscillations in attention: rhythmic α pulses mediated through θ band. *J Neurosci* 34: 4837–4844, 2014. doi:10.1523/JNEUROSCI.4856-13.2014.

Tian X, Chen CY. Probing perceptual performance after microsaccades. *J Neurosci* 35: 2842–2844, 2015. doi:10.1523/JNEUROSCI.4983-14.2015.

Tian X, Yoshida M, Hafed ZM. A microsaccadic account of attentional capture and inhibition of return in Posner cueing. *Front Syst Neurosci* 10: 23, 2016. doi:10.3389/fnsys.2016.00023.

Tomassini A, Spinelli D, Jacono M, Sandini G, Morrone MC. Rhythmic oscillations of visual contrast sensitivity synchronized with action. *J Neurosci* 35: 7019–7029, 2015. doi:10.1523/JNEUROSCI.4568-14.2015.

VanRullen R. Perceptual cycles. *Trends Cogn Sci* 20: 723–735, 2016. doi:10.1016/j.tics.2016.07.006.

Wang XJ. Neurophysiological and computational principles of cortical rhythms in cognition. *Physiol Rev* 90: 1195–1268, 2010. doi:10.1152/physrev.00035.2008.

Wilming N, Harst S, Schmidt N, König P. Saccadic momentum and facilitation of return saccades contribute to an optimal foraging strategy. *PLOS Comput Biol* 9: e1002871, 2013. doi:10.1371/journal.pcbi.1002871.

Wutz A, Muschter E, van Koningsbruggen MG, Weisz N, Melcher D. Temporal integration windows in neural processing and perception aligned to saccadic eye movements. *Curr Biol* 26: 1659–1668, 2016. doi:10.1016/j.cub.2016.04.070.

Yuval-Greenberg S, Merriam EP, Heeger DJ. Spontaneous microsaccades reflect shifts in covert attention. *J Neurosci* 34: 13693–13700, 2014. doi:10.1523/JNEUROSCI.0582-14.2014.

Zanos TP, Mineault PJ, Guitton D, Pack CC. Mechanisms of saccadic suppression in primate cortical area V4. *J Neurosci* 36: 9227–9239, 2016. doi:10.1523/JNEUROSCI.1015-16.2016.

Zanos TP, Mineault PJ, Nasiotis KT, Guitton D, Pack CC. A sensorimotor role for traveling waves in primate visual cortex. *Neuron* 85: 615–627, 2015. doi:10.1016/j.neuron.2014.12.043.

Zoefel B, Sokoliuk R. Investigating the rhythm of attention on a fine-grained scale: evidence from reaction times. *J Neurosci* 34: 12619–12621, 2014. doi:10.1523/JNEUROSCI.2134-14.2014.

Zuber BL, Stark L. Saccadic suppression: elevation of visual threshold associated with saccadic eye movements. *Exp Neurol* 16: 65–79, 1966. doi:10.1016/0014-4886(66)90087-2.

Zuber BL, Stark L, Cook G. Microsaccades and the velocity-amplitude relationship for saccadic eye movements. *Science* 150: 1459–1460, 1965. doi:10.1126/science.150.3702.1459.

J Neurophysiol 121: 646–661, 2019.
First published December 19, 2018; doi:10.1152/jn.00601.2018.

INNOVATIVE METHODOLOGY | *Sensory Processing*

Human-level saccade detection performance using deep neural networks

Marie E. Bellet,[1]* Joachim Bellet,[2,3,4]* Hendrikje Nienborg,[2] Ziad M. Hafed,[2,4]* and Philipp Berens[1,2,5]*

[1]*Institute for Ophthalmic Research, University of Tübingen, Tübingen, Germany;* [2]*Werner Reichardt Centre for Integrative Neuroscience, University of Tübingen, Tübingen, Germany;* [3]*International Max Planck Research School for Cognitive and Systems Neuroscience, Tübingen, Germany;* [4]*Hertie Institute for Clinical Brain Research, University of Tübingen, Tübingen, Germany; and* [5]*Bernstein Center for Computational Neuroscience, Tübingen, Germany*

Submitted 5 September 2018; accepted in final form 17 December 2018

Bellet ME, Bellet J, Nienborg H, Hafed ZM, Berens P. Human-level saccade detection performance using deep neural networks. *J Neurophysiol* 121: 646–661, 2019. First published December 19, 2018; doi:10.1152/jn.00601.2018.—Saccades are ballistic eye movements that rapidly shift gaze from one location of visual space to another. Detecting saccades in eye movement recordings is important not only for studying the neural mechanisms underlying sensory, motor, and cognitive processes, but also as a clinical and diagnostic tool. However, automatically detecting saccades can be difficult, particularly when such saccades are generated in coordination with other tracking eye movements, like smooth pursuits, or when the saccade amplitude is close to eye tracker noise levels, like with microsaccades. In such cases, labeling by human experts is required, but this is a tedious task prone to variability and error. We developed a convolutional neural network to automatically detect saccades at human-level accuracy and with minimal training examples. Our algorithm surpasses state of the art according to common performance metrics and could facilitate studies of neurophysiological processes underlying saccade generation and visual processing.

NEW & NOTEWORTHY Detecting saccades in eye movement recordings can be a difficult task, but it is a necessary first step in many applications. We present a convolutional neural network that can automatically identify saccades with human-level accuracy and with minimal training examples. We show that our algorithm performs better than other available algorithms, by comparing performance on a wide range of data sets. We offer an open-source implementation of the algorithm as well as a web service.

algorithm; deep neural network; eye movements; microsaccade; saccade

INTRODUCTION

Eye tracking is widely used in both animals and humans to study the mechanisms underlying perception, cognition, and action, and it is useful for investigating neurological and neurodegenerative diseases in human patients (Carpenter 1988; Kowler 2011; Leigh and Kennard 2004; Leigh and Zee 2015; MacAskill and Anderson 2016). This is in part due to practical reasons: recording eye movements is relatively easy (Duchowski 2007), while, at the same time, eye movements can be highly informative about brain state (Borji and Itti 2014; Haji-Abolhassani and Clark 2014).

The most prominent type of eye movement, in terms of eyeball rotation speed, is a ballistic shift in gaze position, called saccade. This type of eye movement occurs 3–5 times per second, and it can realign the fovea with interesting scene locations within only ~50 ms. Naturally, saccades cause dramatic changes in visual input when they occur, and they therefore impact neural processing in different visual areas and also in a variety of ways (Burr et al. 1994; Crevecoeur and Kording 2017; Duhamel et al. 1992; Golan et al. 2017; Ross et al. 1997; Reppas et al. 2002; Sommer and Wurtz 2008; Yao et al. 2018; Zirnsak et al. 2014). This even happens for the tiniest of saccades, called microsaccades, that occur when gaze is fixed (Bellet et al. 2017; Bosman et al. 2009; Chen and Hafed 2017; Hafed 2011; Hafed et al. 2015; Hass and Horwitz 2011; Herrington et al. 2009; Gur et al. 1997; Leopold and Logothetis 1998; Yu et al. 2017). Therefore, studies not quantitatively analyzing microsaccades can miss important behavioral and neural modulations in experiments (Hafed 2013). Saccades and microsaccades are, additionally, key discrete events in eye tracking traces that can be useful for parsing other eye movement epochs (e.g., smooth pursuits, ocular drifts, ocular tremors) for further analysis. Therefore, detecting saccades is typically the first step in any quantitative analysis of behavior or neural activity that might be impacted by these eye movements.

Several algorithms have been proposed for automating the task of saccade detection (reviewed in Andersson et al. 2017). For example, Engbert and Mergenthaler (2006) developed a method for classifying saccades and microsaccades based on an adaptive threshold. This algorithm (which we refer to here as EM) is particularly popular because of its simple implementation and ease of use, as well as its ability to detect even microsaccades. However, this algorithm, like others, may still mislabel some microsaccades due to high eye tracker noise (as is typical with video-based eye trackers) as well as small catch-up saccades occurring during smooth pursuit. Other existing algorithms (Larsson et al. 2013; Pekkanen and Lappi 2017) have the added advantage of providing additional labels for fixations and postsaccadic oscillations (PSO) in eye position.

* M.E.B. and J.B. contributed equally to this work; Z. M. Hafed and P. Berens are co-senior authors.

Address for reprint requests and other correspondence: Z. M. Hafed, Werner Reichardt Centre for Integrative Neuroscience, Tübingen, Germany (e-mail: ziad.m.hafed@cin.uni-tuebingen.de).

Despite their success, several shortcomings still render the use of existing algorithms either less reliable than desired or, at the very least, cumbersome. While the performance of many published algorithms is promising (Andersson et al. 2017; Pekkanen and Lappi 2017), it does not reach the level of trained human experts. Also, none of the existing algorithms show convincing performance for all eye movement-related events that may need to be analyzed (e.g., fixations, saccades, PSO, blinks, smooth pursuits). In addition, equipment-dependent hyperparameters, such as thresholds, need to be chosen for most algorithms, a fact that renders broad usability difficult. For example, even simple changes in eye tracking hardware, involving changes in sampling frequency or measurement noise, require retuning of such parameters. Retuning is also needed when the ranges of eye movement amplitudes being studied are modified (e.g., microsaccades vs. larger saccades). Perhaps most importantly, objective parameter estimation in existing algorithms is currently a challenging task because of a limited amount of available reliably labeled data. Finally, in many cases, applying available online resources is not straightforward. As a result of all of the above shortcomings, current laboratory practice often still involves experimenters spending substantial amounts of time to carefully relabel at least parts of their data after automatic saccade detection.

Here we propose a convolutional neural network (CNN) for classifying eye movements. The architecture of the network is inspired by U-Net, which has successfully been used for image segmentation (Ronneberger et al. 2015). We evaluated our network (U'n'Eye) on four challenging data sets containing small saccades occurring during fixations or smooth pursuits. On these data sets, U'n'Eye reached the performance level of human experts in labeling saccades and microsaccades, while being much faster. The network also beat state-of-the-art algorithms on a benchmark data set not just for saccade detection, but also for PSO. As we show here, our network can be trained quickly, even on a standard laptop, and with minimal amounts of training data. More importantly, our network's adaptability to different data sets makes U'n'Eye the novel state-of-the-art eye movement detection algorithm. We provide an easily accessible web service for running U'n'Eye (http://uneye.berenslab.org), as well as an open source implementation (https://github.com/berenslab/uneye). Our labeled data sets will also be freely available upon publication.

METHODS

Data sets. All experiments used for collecting the data sets were approved by ethics committees at T bingen University. Human subjects provided informed, written consent in accordance with the Declaration of Helsinki. Monkey experiments were approved by the regional governmental offices of the city of T bingen.

Data set 1 was collected from human subjects using the Eyelink 1000 video-based eye tracker (SR Research) sampling eye position at 1 kHz. The data set contains mostly microsaccades and small-amplitude memory-guided saccades. It contains 2,000 trials of 1 s. Out of these 2,000 trials, 1,000 were selected to compare U'n'Eye to other algorithms via cross-validation (Fig. 4). We named these trials "set1A." When testing for the impact of missing labels on performance (Fig. 7*B*), we used the other 1,000 trials, "set1B," to train networks and tested them on set1A.

Data set 2 was collected from three male rhesus macaque monkeys implanted with scleral search coils (in one eye for each of the monkeys). The data set contains catch-up saccades generated during smooth pursuit. Eye position was again sampled at 1 kHz. For the trials containing smooth pursuit of sinusoidal target motion trajectories in this data set, the data were obtained from the experiments described in Hafed et al. (2008) and Hafed and Krauzlis (2008). For the trials containing pursuit of constant speed, the experimental conditions are described in Buonocore et al. (2018). Eye movement calibration for search coil data was done according to the procedures in Tian et al. (2016). The overall data set consists of 2,000 segments of 1 s of eye traces. Like in the case of data set 1, we split the set into two sets of 1,000 segments each, "set2A" and "set2B." set2A was used to compare U'n'Eye to Daye and Optican's (2014) algorithm (Fig. 4).

Data set 3 was collected from a single male macaque monkey using the Eyelink 1000 video-based system sampling eye position at 500 Hz. The data set contains microsaccades generated during fixation. The data were obtained from experiments described in Kawaguchi et al. (2018). It consists of 403 segments of 1.438 s. Similarly to data sets 1 and 2, we split the data in two subsets, "set3A" and "set3B." Set3A contains 350 segments and set3B 53 segments. Set3A was used for comparing U'n'Eye's performance to other that of algorithms (Fig. 4).

For the results shown in Fig. 7*D*, we used setA of all data sets for training and the respective setB for testing.

Data set 4 was collected from the same eye tracker as data set 1 but with different sets of subjects. It comes from a recently published study (Bellet et al. 2017) in which subjects had to keep fixation at the center of the screen before a peripheral target appearance. We selected 630 segments of 750 ms from each of 10 subjects (4,725 s in total). The data set includes not only successful trials, in which subjects maintained fixation, but also trials containing blinks or saccades outside of the fixation window. Again, we split the data set into two subsets. Set4A contained 330 segments per subject and was used to train networks. Set4B contained 300 segments per subject and was used to test the performance of the networks.

In all data sets, we manually detected saccades using a custom-made graphical user interface (GUI) in MATLAB. The GUI displayed horizontal and vertical eye position traces, as well as filtered radial eye velocity. The GUI internally estimated saccade onset and end times using a combination of velocity and acceleration thresholds (Chen and Hafed 2013). The user then manually interacted with the GUI to delete false alarms, correct false negatives, and adjust estimation of onset and offset timing.

Simulated saccades. To test the performance of our network on noisy labeled data, we designed artificial eye traces for which we knew the ground truth. Saccades ranging from 0.5 to 60° were simulated using an adaptation of a model for saccade waveforms (Dai et al. 2016). The model is a sum of soft ramp functions, which follows the relationship between amplitude and peak velocity observed in real saccades (Dai et al. 2016). Since the model is originally one dimensional, we adapted it so that it generates two-dimensional trajectories. Saccade generation in time was made to follow a Poisson process with λ equal to 3 saccades/s. Simulated blinks were also added by inducing sharp transients in the eye traces. Finally, a Gaussian white noise with a standard deviation of 0.02° was added to the trace. Then, as described in RESULTS, we trained U'n'Eye under a variety of conditions in which we intentionally removed a subset of saccade labels during training, to explore robustness to missing labels (Fig. 7).

U'n'Eye: our convolutional neural network. The architecture of CNN was inspired by U-Net, a CNN first used for image segmentation (Ronneberger et al. 2015). Here we modified U-Net to meet the requirements of an eye movement classifier. The network was built of seven convolutional layers with kernel size 5, each followed by a linear-rectifying unit (ReLU) and a BatchNorm layer, both described in detail in RESULTS. Batches consisted of samples of the same duration. The input to the network was eye velocity which was computed as the first-order difference of the eye position signal. The input was of dimension $N \times T \times 2$, where N is the batch size, T the number of time points, and 2 the number of coordinates (horizontal

and vertical eye velocity). The number of input time points could be variable but had to be a multiple of 25 bins due to the max pooling operations. The output of the network was a matrix of dimension $N \times K \times T$, where K was the user-defined number of classes. For example, we could have a "saccade" and "fixation" class in the networks of Fig. 3 and we could also add other classes like "PSO" in the network of Fig. 6.

We applied a softmax (Bishop 2016) activation function to the output of the last convolutional layer x:

$$Softmax(x_i) = \frac{e^{x_i}}{\sum_{j=1}^{k} e^{x_j}} \quad (1)$$

where x_i is the layer corresponding to *class i*. Thus, the network's output y represented the sample-by-sample conditional probability of each class (e.g., "fixation" or "saccade") given the eye-velocity x and the network weights w:

$$Y_k = p(k = 1|x, w) \quad (2)$$

The final prediction of the algorithm represented the class that maximized this conditional probability:

$$\hat{k} = argmax_k p(k = 1|x, w) \quad (3)$$

We chose the kernel sizes of the convolutional and max pooling operations in a way to capture a relevant signal range around each time point. Based on the given kernel sizes of the network, it can be shown that the prediction of one time bin is influenced by the preceding and following 89 time bins of the velocity signal (Fig. 2*B*, red color).

Network training. We trained the network with minibatches whose size depended on the total number of training samples. We performed 10 training iterations in each epoch. Overfitting on the training set was prevented by computing the loss on a validation set and stopping training when the validation loss increased for three successive epochs. We used a multiclass error function, which, for two classes, equals the cross entropy loss. Weight-regularization was done with L2-penalty (Bishop 2016), which corresponds to a Gaussian prior with zero mean over the network weights. The optimal parameter λ was determined to be 0.01. The loss function was thus defined as:

$$L = -\sum_{n=1}^{N}\sum_{k=1}^{K} t_{nk} log y(x_n, w) + \lambda \|w\|_2^2 \quad (4)$$

where N is the number of time points and K the number of classes. The ground truth label t_{nk} equals 1 if the *time point n* belongs to *class k*. Gradient computation was done with PyTorch autograd method.

We used the Adam optimizer (Kingma and Ba 2014) with an initial learning rate of 0.001. Adam is a stochastic gradient-based optimizer that uses adaptive learning rates for different weights of the network. An additional step decay by a factor of 2 was applied to the current learning rates when the loss on the validation set increased during one epoch.

Postprocessing. In the case of binary prediction into the classes fixation and saccade, we provided the possibility to define thresholds for minimum saccade duration and minimum saccade distance. If thresholds were given, saccades closer than the minimum distance were merged and saccades shorter than the minimum duration were removed. We obtained the results reported here with a minimum saccade distance threshold of 10 ms for data set 1 and of 3, 4, and 5 ms for data set 2, because we previously observed that some saccades occurred very close in time in this data set. For data sets 1–4, we used a minimum saccade duration threshold of 6 ms. The same thresholds were used for the algorithm Engbert and Mergenthaler (2006).

Data augmentation. U'n'Eye performs better with a bigger training set. However, we aimed to reduce the amount of saccades that a user should provide to train U'n'Eye. In this study, to increase the number of training samples, the input eye positions were rotated and added to the original training samples:

$$x_2 = xcos(\theta) + ysin(\theta) \quad (5)$$

$$y_2 = -xsin(\theta) + ycos(\theta) \quad (6)$$

where x and y are the horizontal and vertical eye positions. We used $\theta - (1/4\pi, 3/4\pi, 5/4\pi, 7/4\pi)$ radians. Thus, we could increase by fivefold the size of our training set without causing overfitting.

Performance measures. To evaluate the eye movement detection performance of our network, we used the following metrics: Cohen's kappa, F1 score, and onset and offset time differences.

Cohen's kappa is a sample-based statistic. It reflects how much two coders agree on the class that each time bin belongs to, while controlling for chance agreement of the two coders. It is given by

$$\kappa = \frac{p_0 - p_e}{1 - p_e} \quad (7)$$

where p_0 is the proportion of time bins for which two coders agree, and p_e is the proportion of time bins for which agreement can be expected by chance.

For a binary classification of fixation vs. saccades, the Cohen's kappa value p_e is given by

$$P_e = \frac{1}{N^2} \times \sum_{k=1}^{K} nk_{coder1} \times nk_{coder2} \quad (8)$$

where nk_{coderX} is the number of time bins coder X assigned to *class k*.

The F1 score is a measure of classification accuracy that combines precision and recall of a predictor. Precision is defined as the proportion of correctly classified saccades over all predicted saccades. Recall is defined as the proportion of correctly classified saccades over all saccades in the ground truth. The F1 score is the harmonic mean of these two measures. It is given by

$$F1 = 2 \times \frac{TP}{2 * TP + FN + FP} \quad (9)$$

where TP is the number of true positives, FN the number of false negatives, and FP the number of false negatives. For all true positive saccades, we compared saccade timing between the ground truth and prediction by calculating the absolute time differences between true and predicted saccade onsets and offsets.

Evaluation on a benchmark data set. We evaluated U'n'Eye performance on a benchmark data set by Andersson et al. (2017). This data set comprises 500 Hz eye-tracking data from humans looking at images, movies, or moving dots. It contains human labels for the events fixations, smooth pursuits, saccades, PSO, and blinks. Events that the human experts did not assign to any of these classes were labeled as "others." For some trials, the data set contained labels from two different human coders. For other trials, only one label was available. We trained 20 independent networks with different random initializations on the data with labels from one human coder (coder RA). Performance was then tested on the trials with labels from two coders, which makes our result comparable with previously reported results (Pekkanen and Lappi 2017). Note that we were not able to reproduce the interrater measures reported by Andersson et al. (2017) in line with the results of Pekkanen and Lappi (2017). For comparability with the NSLR-HMM algorithm (Pekkanen and Lappi 2017), we excluded the event labels "other" for the calculation of Cohen's kappa scores. The performance on the class "blinks" was not compared with other algorithms since it was not reported.

Evaluation of other algorithms. We compared U'n'Eye performance on our data sets to several already published algorithms. For data sets 1 and 3, which contain microsaccades occurring during fixation of a static target, we evaluated the performance of three algorithms designed for microsaccade detection (Engbert and Mer-

genthaler 2006; Otero-Millan et al. 2014; Sheynikhovich et al. 2018) and one algorithm designed for saccade detection in a high-noise regime (Pekkanen and Lappi 2017).

The algorithm by Engbert and Mergenthaler (2006) is commonly used as an unsupervised method to detect microsaccades. It selects saccades based on a threshold that depends on the level of the noise in the velocity. One parameter, called λ, can be also be fit to the data to obtain better results. This λ is multiplied with the velocity noise to determine a threshold for saccade selection. Here we chose λ values that maximized the metric of interest on the training data from our cross-validations. This was done to give this algorithm the benefit of the doubt in our comparisons. Importantly, before saccade detection, we smoothed the eye traces using a five-point average independently of the sampling frequency, as described by Engbert and Mergenthaler (2006).

The approach from Otero-Millan et al. (2014) is an unsupervised method. It gives an estimate of saccade onset and offset timing and thus can be compared in terms of both the Cohen's kappa and F1 metrics.

Another unsupervised method has been developed by Sheynikhovich et al. (2018). This algorithm only gives an estimate of microsaccade occurrence at one point in time without determining onset and offset. We thus compared only the performance in terms of F1 score for this algorithm. We considered as true positive any saccade detected ± 10 ms away from a ground truth saccade.

For data set 2, we evaluated the performance of the method by Daye and Optican (2014), which uses particle filters to detect saccades embedded in high-velocity eye movements. The algorithm was kindly provided by the authors. To increase performance, we detected saccades independently in the horizontal and vertical channel and then merged the predictions. This is because the Daye and Optican algorithm only considers as a saccade an event crossing a threshold both in horizontal and vertical components at the same time, which increases the number of false negatives. To increase performance, the parameters were tuned differently for trials containing sinusoidal pursuit than for those containing linear pursuit, again to give the algorithm the benefit of the doubt when comparing to U'n'Eye. To detect saccades in sinusoidal pursuit, the parameters were set to $\Omega = 10^{-4}$, $\xi = 3.\ 10^{-4}$, $N = 100$, $m = 20$, $\lambda = 5.\ 10^{-3}$, $\psi = 2.\ 10^{-3}$. To detect saccades in linear pursuit, the parameters were set to $\Omega = 10^{-3}$, $\xi = 3.\ 10^{-3}$, $N = 100$, $m = 20$, $\lambda = 5.\ 10^{-3}$, $\psi = 10^{-4}$.

For all unsupervised algorithms, the 10 testing subsets from the cross-validation data were evaluated at once to yield better clustering estimates.

The results of the algorithm by Pekkanen and Lappi (2017) on data sets 1 and 3 were obtained by estimating the model's parameters via cross-validation using the same training folds as for U'n'Eye. The estimated parameters were kindly provided by the authors.

Compute time. The computation times of our algorithm reported here were achieved on a personal computer with a 2 GHz Intel Core i5 processor at 8 GB RAM running on Mac OS X 10.11.6.

Code and data availability. A web service for running the algorithm is available at http://uneye.berenslab.org. All code is available from https://github.com/berenslab/uneye. Data will be available upon publication.

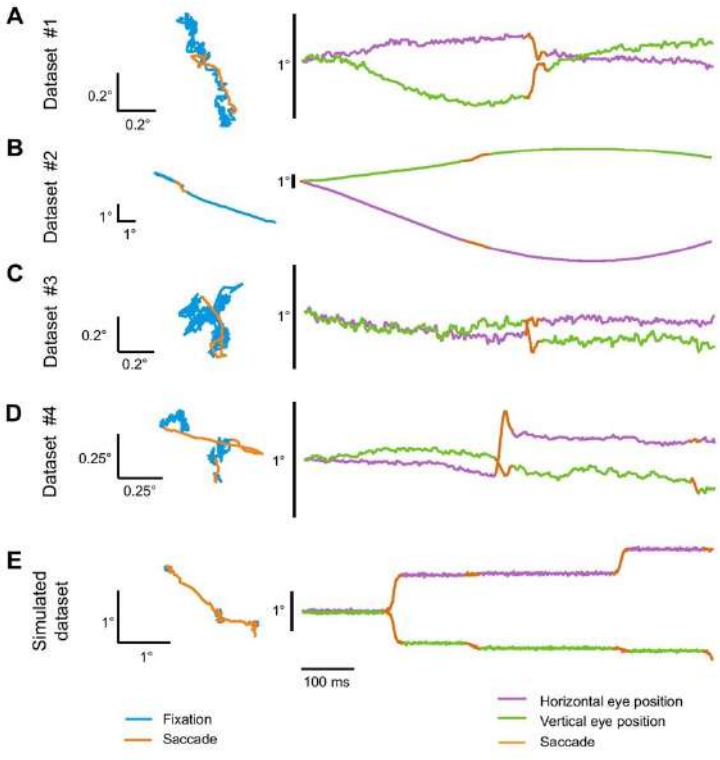


Fig. 1. Examples of eye traces containing saccades for detection. *A*: microsaccades during fixation recorded with a video-based eye tracker. *B*: catch-up saccades during smooth tracking recorded with scleral search coils. *C*: microsaccades during fixation recorded with a video-based eye tracker. *D*: microsaccades during fixation recorded with the same video-based eye tracker as in *A* but for different sets of subjects. *E*: simulated saccades. In all panels, 2D plots on the left are the 2D representation of the eye trajectory over 1 s of recording, and to the right of them are the horizontal and vertical components of the corresponding traces presented as a function of time; in this case, an upward deflection in the shown traces corresponds to a rightward or upward eye movement for the horizontal and vertical components, respectively. Note that, in *B*, we refer to the nonsaccadic smooth change in eye position as "fixation" for simplicity, since the primary goal of our algorithm was to detect saccades, irrespective of whether they happened during fixation or embedded in smooth pursuit eye movements.

RESULTS

Design of a convolutional neural network for eye movement classification. We developed a CNN that predicts the state of the eye for each time point of an eye trace. The aim of the network was to segment eye movement recordings (Fig. 1) into epochs containing saccades/microsaccades (orange highlights) vs. epochs not containing these eye movements (but see also *U'n'Eye: new state-of-the-art eye movement classifier* below for additionally classifying PSO using our network). Our primary goal was to have a network that can seamlessly handle the challenging scenarios of tiny microsaccades during fixation (Fig. 1*A*), small catch-up saccades embedded in relatively high smooth pursuit eye velocities (Fig. 1*B*), and microsaccades and saccades occurring in recordings with higher noise levels associated with video-based eye trackers when compared with, say, scleral search coil techniques (Fuchs and Robinson 1966; Judge et al. 1980) (Fig. 1*C*). We therefore trained and tested the network on three different challenging data sets (see METHODS and Table 1), which contain labels for fixations and saccades manually determined by human experts. To test the ability of our network to generalize across eye movement traces recorded from different individuals, we also included a fourth data set (Fig. 1*D*), which was obtained from 10 different subjects using the same eye tracker as in data set 1 (see METHODS). Testing generalizability was also achieved using with a fifth and final data set containing artificially generated noisy eye movement traces, in which the ground truth for saccade times was known (see METHODS) (Fig. 1*E*). Finally, we compared our network's performance to different existing algorithms, both on our data sets and also on a publicly available benchmark data set (Larsson et al. 2013) (http://dev.humlab.lu.se/www-transfer/people/marcus-nystrom/annotated_data.zip).

The network operates on the eye velocity signal and requires no other preprocessing. Eye velocity is computed as the differential of eye position (see METHODS), and chunks of eye velocity signals are then input to the network. Briefly, the network's architecture is based on the U-Net, a CNN for pixel-by-pixel image segmentation (Ronneberger et al. 2015), which we modified to process one-dimensional signals and output a predictive probability for each eye movement class at every time point (Fig. 2*A*). A major change compared with the original U-Net architecture is that we introduced batch normalization (BatchNorm) layers (Klibisz et al. 2017). BatchNorm layers subtract a mean from their input and divide it by a standard deviation. Both of these parameters are estimated for each layer over minibatches of training samples during learning. This method normalizes the distribution of activations across the network layers, allowing for higher learning rates and reducing overfitting (Ioffe and Szegedy 2015). We also applied a rectified linear unit (ReLu) function between each convolutional and batch normalization layer. The ReLu function, or heaviside step function, introduces nonlinearities in the network, allowing it to apply arbitrary-shaped functions to the input data. Finally, the U-shaped architecture of the network leads to temporal downsampling and upsampling in the hidden layer representations (Fig. 2). Downsampling is achieved by max pooling (MaxPool) operations that reduce the dimensionality of the network content, extracting relevant features. Upsampling is realized by transposed convolution. Convolutional kernels and max pooling operations together lead to the integration of information over time. Due to the network design, the probability assigned to each time bin can be influenced by ±89 preceding and following time bins (Fig. 2*B*). Thus, U'n'Eye takes into account a large enough signal to make point predictions of the correct eye movement class.

U'n'Eye achieves human-level performance. Our network achieved human-level performance after training on our data sets. We first illustrate this with three example scenarios for detecting saccades (Fig. 3). For illustrative purposes, we also show how the commonly used EM algorithm might perform for the examples; we later provide an exhaustive quantitative

Table 1. *Data set characteristics*

Data Set	*1*	*2*	*3*	*4*
Subjects	Humans	Monkeys	Monkeys	Humans
Eye tracker	Eyelink 1000	Search coil	Eyelink 1000	Eyelink 1000
Sampling frequency, Hz	1,000	1,000	500	1,000
Saccade type	Microsaccades and memory saccades	Saccades during smooth pursuit	Microsaccades	Microsaccades and saccades
Duration				
Mean ± SD, ms	44.58 ± 15.42	37.51 ± 8.81	23.12 ± 6.52	31.66 ± 8.93
Median, ms	42	36	22	31
Minimum, ms	11	18	8	8
Maximum, ms	169	97	54	110
Amplitude				
Mean ± SD, °	0.69 ± 0.93	1.07 ± 0.70	0.22 ± 0.13	0.33 ± 0.28
Median, °	0.43	0.96	0.20	0.24
Minimum, °	0.02	0.04	0.010	0.008
Maximum, °	11.34	7.03	1.27	2.66
Velocity				
Mean peak ± SD, °/s	102.46 ± 68.82	68.23 ± 42.98	208.41 ± 65.95	61.93 ± 35.18
Median peak, °/s	81.91	56.59	198.86	52.96
Minimum peak, °/s	17.81	11.49	85.28	15.32
Maximum peak, °/s	547.72	450.44	560.28	423.18
Median instantaneous, °/s	5.63	15.70	10.20	5.63

All statistics refer to saccades. Note that minimum saccade amplitude may appear very low due to the existence of some saccades that had very strong dynamic overshoot (a substantial saccadic movement followed by one lobe of a postsaccadic oscillation almost to the original eye position before saccade onset). The statistics of the simulated data set are described in METHODS.

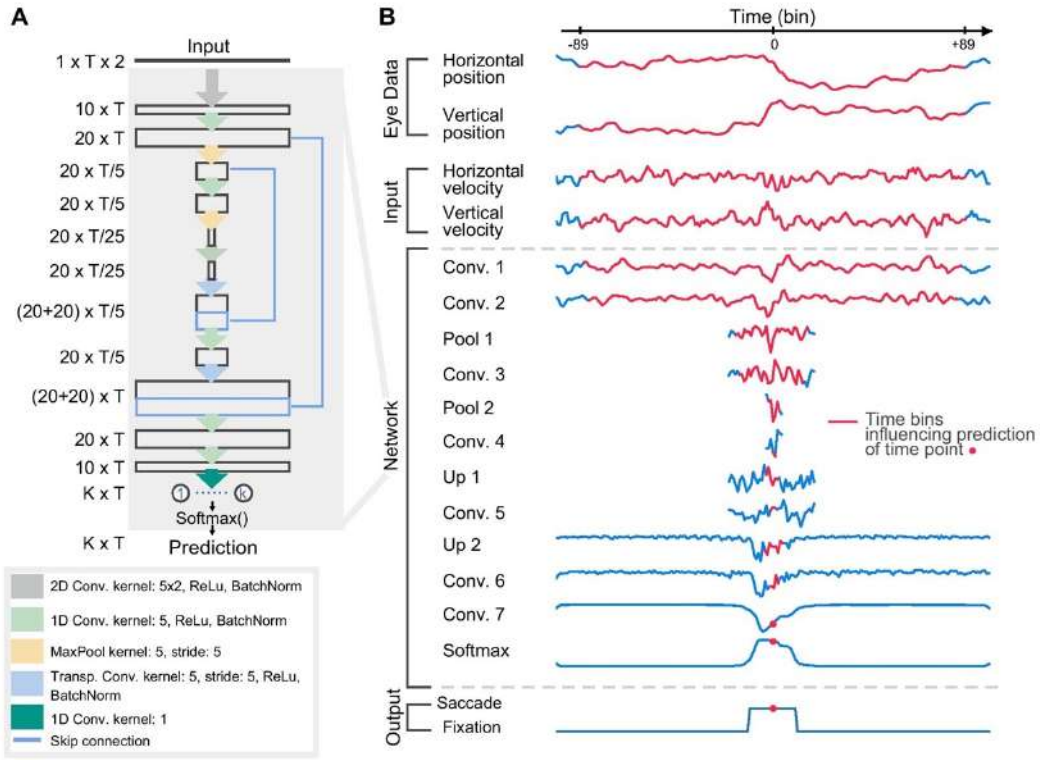


Fig. 2. U'n'Eye. *A*: network architecture. The input matrix contains horizontal and vertical eye velocity. *T* is the number of input time points (see *B*), and *K* is the user-defined number of eye movement classes (e.g., "fixation" vs. "saccade" in a binary classifier). The different network layers are described in *Design of a convolutional neural network for eye movement classification* under RESULTS. *B*: the output probability of one time bin is influenced by 89 time samples before and after this time bin. For each layer of the network, the red color indicates the range of influence of the time bin indicated by the red dot in the output. Traces show the projection of the layer's output onto its first principal component. The outputs of convolutional (Conv) layers 6 and 7 resemble the final classifier's output probability (Softmax), whereas early convolutional layers 1 and 2 seem to perform noise reduction.

comparison with several more algorithms (Fig. 4). In the first example, a small microsaccade occurred with substantial oscillation in eye position toward movement end, and with the amplitude of the movement being near the eye tracker noise level (Fig. 3*A*). Human coder 1 considered the postsaccadic oscillation as part of the saccade, and so did our network trained on his training set (compare the binary classification output of the coder and network 1 below the eye movement traces in Fig. 3*A*). On the other hand, coder 2 determined that the saccade ended earlier, and our network trained on his training set did the same (again, compare the classification output for human coder 2 and network 2). Thus, our network could match the criterion used by an individual human coder very well. Moreover, our network successfully avoided a false detection by the EM algorithm on these traces. In the second example, the EM algorithm missed all three saccades, which is perfectly reasonable since this algorithm was never designed to work in association with smooth pursuit eye movements, but our network successfully flagged them (Fig. 3*B*). Finally, the eye movement in the third example was collected with a video-based eye tracker having substantially more noise (Fig. 3*C*). In this case, one false detection made by the EM algorithm was successfully excluded by our network.

To present more quantitative performance measures, we first tested our network on our in-house data sets (Fig. 1) and compared its performance to that of commonly used or recently published algorithms. For our network, we performed 10-fold cross-validation separately for data sets 1–3. In each cross-validation round, 90% of the data were used for training the network, and the remaining 10% were used to test performance. A separate validation set from each data set was used to detect overfitting of the network. To prevent such overfitting, we regularized the weights of the network using the L2 penalty (Bishop 2016) (see METHODS), preventing the parameters of the network from deviating excessively from zero. Furthermore, we made use of early stopping. For this, a separate validation set was used, and the validation set error was computed in each epoch. Training was stopped at the point of smallest validation set error. For data sets 1 and 2, 950 s of eye traces were used for cross-validation and 50 s for valida-

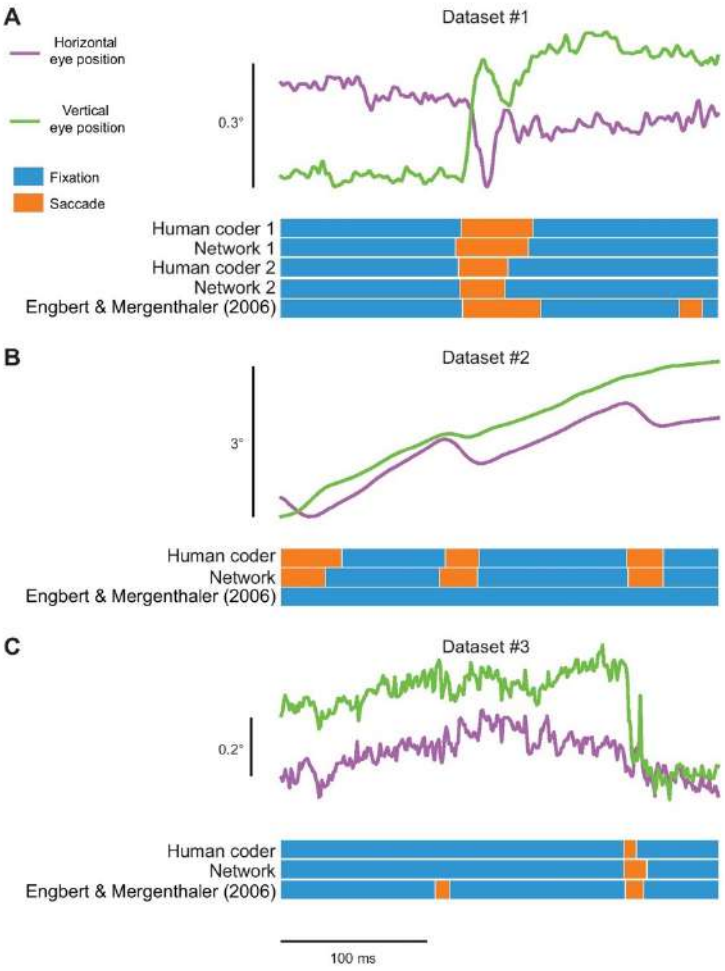


Fig. 3. Examples of eye traces from our first three data sets. Saccades are labeled by either human coders, different instances of U'n'Eye, or a popular algorithm from the literature included here for illustrative purposes (also see Fig. 4 for detailed performance comparisons to several algorithms). *A*: an example microsaccade exhibiting substantial postsaccadic oscillation (PSO). The top two traces show eye position as a function of time in an identical format to Fig. 1. Below the eye position traces, we show labels for "fixation" or "saccade" made by two human experts (human coder 1 and human coder 2) as well as predictions of two separate networks. Network 1 was trained on labels from human coder 1, and network 2 was trained on labels from human coder 2. Note how each network matched the performance of its corresponding human coder. The very bottom row shows the performance of the Engbert and Mergenthaler (2006) algorithm (EM), which suffered from a false alarm later in the trace due to eye tracker noise. *B*: saccades embedded in smooth pursuit eye movements. Here, our network successfully detected three catch-up saccades, all of which were missed by the EM algorithm, which was not designed to work with eye movement records containing smooth pursuit. The reason that these saccades were missed is that the saccades were directed opposite to the ongoing pursuit, resulting in momentary reductions in eye speed, as opposed to increases. *C*: an example microsaccade embedded in high eye tracker noise. Once again, the EM algorithm suffered from a false alarm due to eye tracker noise.

tion. Thus, each training set contained 855 s of data. For data set 3, 330 s were used for cross-validation and 23 s for validation, resulting in 297 s of data in each training set. For the other algorithms that we tested, we used the same cross-validation approach in the case of supervised algorithms [EM (Engbert and Mergenthaler 2006), Pekkanen and Lappi (2017)]. Note that we used the EM algorithm as a supervised method since we fitted its single parameter on our training data (see METHODS). For unsupervised methods (Daye and Optican 2014; Otero-Millan et al. 2014; Sheynikhovich et al. 2018), the identical 10 test sets were evaluated without using the training set (see METHODS).

Finally, similarity of the algorithms' predictions to human labels was evaluated using three metrics. First, we calculated Cohen's kappa, which is a sample-by-sample similarity measure that takes chance agreement of two predictors into account (Cohen 1960). Second, we calculated the F1 score, which is an accuracy measure that considers precision and recall of a classifier. Recall corresponds to the number of correctly detected saccades divided by the number of saccades that were labeled by the human expert. Precision, on the other hand, is the number of correctly classified saccades divided by the total number of saccades detected by the classifier (see METHODS). The F1 score is defined as the harmonic mean of both, and it thus only measures how accurate saccades were detected without taking into account their timing (i.e., exact saccade onset and offset times). Correctly labeling saccade onset and offset can be crucial for further analyses. Therefore, for our third and final metric, we additionally computed the absolute time difference in onset and offset of correctly predicted saccades and of saccades labeled by the human experts. This measure reflects how well an algorithm agrees with the human coder in terms of saccade start and end.

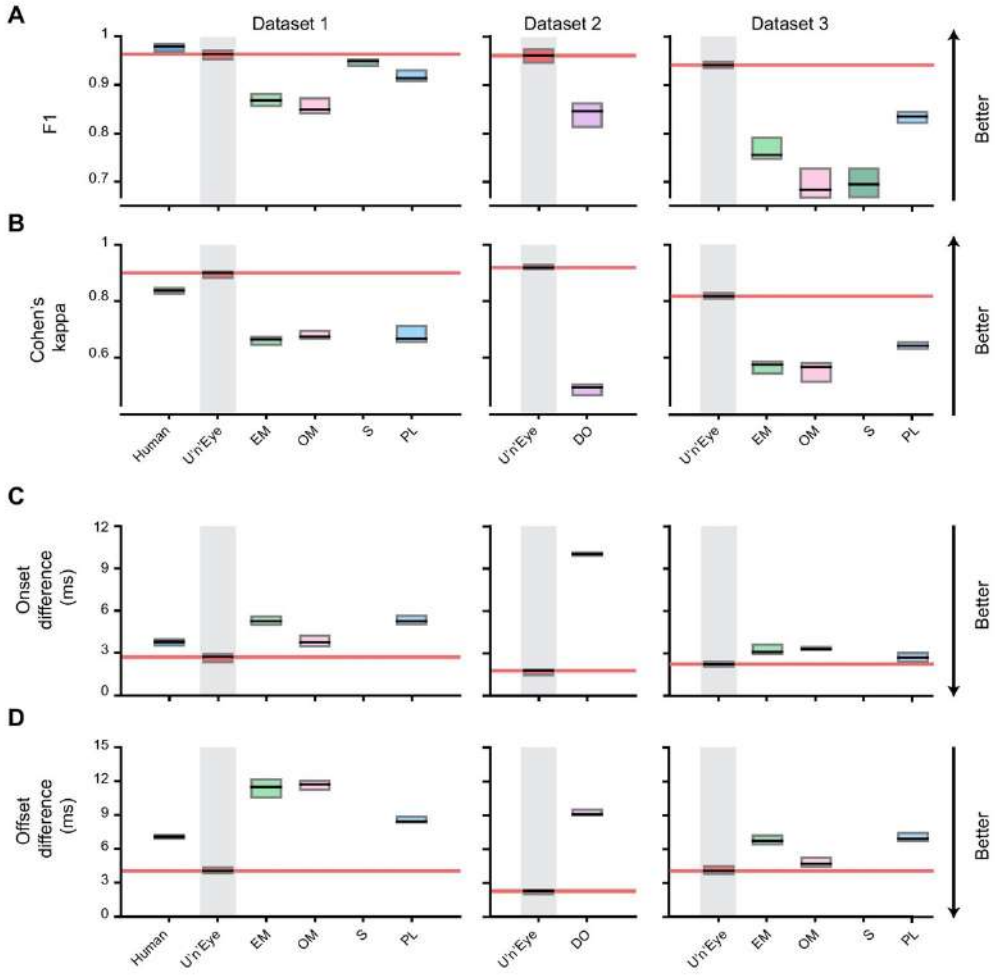


Fig. 4. High performance of U'n'Eye. Each panel shows results from one performance metric described in the text, and on each of our first three data sets. For each metric, we show the median across 10 different cross-validation runs. The boxes show two quartiles of the distributions. *A*: F1 score (See *Eq. 9* and the surrounding text) summarizing precision and recall performance between two predictors. The first predictor was always a human coder (considered as ground truth). Therefore, the first column indicates agreement between a second coder (labeled Human in the figure) to the original coder used to train our network. *B*: Cohen's kappa measuring sample to sample agreement. *C*: average absolute difference in the timing of saccade onset times. *D*: same as *C* but for saccade offset times. In all cases, our network (highlighted by gray rectangles) demonstrated superior performance (the arrows on the far right side indicate the direction of superior performance for each metric). EM, Engbert and Mergenthaler (2006); OM, Otero-Millan et al. (2014); S, Sheynikhovich et al. (2018); PL, Pekkanen and Lappi (2017); DO, Daye and Optican (2014). Note that we only tested data set 2 on DO because only this algorithm was explicitly designed to deal with smooth pursuit eye movements.

U'n'Eye reached high similarity to the human coder (Fig. 4, *A* and *B*) and outperformed all the other compared algorithms (Fig. 4, *A* and *B*). U'n'Eye also detected saccade onset and offset in high agreement with the human labels. On average, saccade onset differences to human labels were smaller than 3 ms, and saccade offset differences were smaller than 4 ms. Saccade onset and offset labels by the other algorithms deviated more strongly from the human-labeled saccades (Fig. 4, *C* and *D*; Table 2). This indicates that U'n'Eye's saccade predictions were more humanlike.

In the more challenging data set 2, in which saccades occurred during smooth pursuit eye movements, U'n'Eye outperformed the algorithm by Daye and Optican (2014), which was designed to overcome this difficulty. Here, saccade peak velocity was close to the instantaneous velocity of the ongoing smooth pursuit movements. In fact, the minimum saccade peak

Table 2. *Comparison of U'n'Eye performance to other algorithms on data sets 1, 2, and 3*

Data Set	Algorithm	F1	Cohen's Kappa	Δ Onset, ms	Δ Offset, ms
1	U'n'Eye	**0.96** ± 0.01	**0.89** ± 0.02	**2.66** ± 0.34	**4.11** ± 0.41
	EM	0.87 ± 0.03	0.66 ± 0.02	5.39 ± 0.49	11.28 ± 1.00
	OM	0.85 ± 0.03	0.68 ± 0.03	3.80 ± 0.48	11.50 ± 0.77
	S	0.95 ± 0.02			
	PL	0.92 ± 0.02	0.68 ± 0.03	5.51 ± 0.88	8.53 ± 0.60
2	U'n'Eye	**0.96** ± 0.01	**0.92** ± 0.01	**1.70** ± 0.29	**2.19** ± 0.37
	DO	0.84 ± 0.03	0.49 ± 0.03	9.99 ± 0.19	9.22 ± 0.35
3	U'n'Eye	**0.94** ± 0.01	**0.82** ± 0.02	**2.23** ± 0.22	**3.99** ± 0.60
	EM	0.77 ± 0.04	0.58 ± 0.04	3.22 ± 0.53	6.87 ± 0.66
	OM	0.68 ± 0.07	0.55 ± 0.06	3.48 ± 0.55	4.90 ± 0.64
	S	0.70 ± 0.04			
	PL	0.83 ± 0.02	0.64 ± 0.03	2.74 ± 0.39	6.94 ± 0.50

In bold are the best performances for each data set. In all cases, U'n'Eye achieved highest performance. Values report mean and standard deviation across validation sets. EM, Engbert and Mergenthaler (2006); OM, Otero-Millan et al. (2014); PL, Pekkanen and Lappi (2017); S, Sheynikhovich et al. (2018).

velocity in this data set was smaller than the median instantaneous velocity during pursuit (Table 1). Yet U'n'Eye succeeded in detecting such saccades. This was because the network architecture utilized a substantial time window (Fig. 2), allowing it to infer changes in the state of the eye even if the instantaneous velocity is low compared with the surrounding eye trace.

We next addressed the question of whether U'n'Eye can achieve a similar level of interhuman agreement when multiple human experts analyze the same data. For this, we used data set 1 because, among the four data sets, it contained saccades with the widest range of amplitudes (from as small as 0.02° up to a size of 11°; see Table 1 for a reason why saccades as small as 0.02° were possible). We could thus assess interrater agreement for a broad range of saccades. Data set 1 was labeled by a second independent human coder (Fig. 3, *top panel*; Fig. 4, data set 1). Coder 1 estimated saccade timing based on a combination of the raw eye traces and the smoothed radial velocity, whereas coder 2 used the raw eye traces only. We trained independent networks either with labels from coder 1 or coder 2 (network 1 and network 2, respectively), and we tested the networks' performance on the 10 test sets from the 10-fold cross-validation routine described above, both against ground truth labels from coder 1 or coder 2. U'n'Eye's saccade labels were as similar to both human coders as the human labels were to each other (Table 3). In terms of the F1 score, the interhuman agreement was not significantly different from the network-human agreement (Table 4). Interestingly, network 1 showed higher similarity scores than coder 2 when both were compared with labels of coder 1 in the test sets, and vice versa for network 2 and coder 2. This is reflected by larger Cohen's kappa scores and smaller onset and offset differences (Table 4, all $P<5\times10^{-5}$ after Bonferroni correction for multiple comparisons, Student's paired samples *t*-test for Cohen's kappa and F1 scores, and independent samples *t*-test for on- and offset differences). This indicates that U'n'Eye's saccade estimation surpasses interrater consistency.

U'n'Eye misses only a small fraction of microsaccades. We then analyzed the patterns of agreement and disagreement between U'n'Eye and human labeling. For true positive saccades, the two dimensional histogram of detected movements reflected the typical main sequence relationship between peak velocity and amplitude of saccades (Fig. 5, *A*, *D*, and *G*) (Zuber et al. 1965). A few false positives were present within the range of the main sequence, suggesting that the human coder forgot to label some saccades (for example, see the movement in the *inset* in Fig. 5*B*). Concerning the rare false negatives that occurred, some of them had fairly large amplitudes (beyond eye tracker noise). Closer inspection revealed that there were pairs of successive saccades that had very short intersaccadic intervals. The network lumped them into one movement, whereas the human coders separated them. Most remaining disagreements between the human and the network were associated with the smallest microsaccades, closest to eye tracker noise levels.

Table 3. *Interrater comparison*

	Cohen's kappa	F1	Δ Onset, ms	Δ Offset, ms
Coder 1 vs. coder 2	0.83 ± 0.02	**0.98** ± 0.01	3.72 ± 0.39	7.10 ± 0.34
Network 1 vs. coder 1	**0.89** ± 0.02	0.96 ± 0.01	**2.65** ± 0.34	**4.11** ± 0.41
Network 2 vs. coder 2	**0.89** ± 0.01	0.96 ± 0.01	**2.00** ± 0.11	**4.81** ± 0.33
Network 2 vs. coder 1	0.85 ± 0.01	0.96 ± 0.01	3.34 ± 0.34	5.58 ± 0.33
Network 1 vs. coder 2	0.86 ± 0.01	0.96 ± 0.01	2.82 ± 0.32	6.57 ± 0.53

The first row shows the similarity measures between labels from two human experts (coder 1 and coder 2). Network 1 was trained on labels from coder 1, and network 2 was trained on labels from coder 2. In bold are comparisons leading to best performances. Values report mean and standard deviation across cross-validations. Intercoder agreement was evaluated on the 10 test samples from cross-validation.

U'n'Eye: new state-of-the-art eye movement classifier. To compare our algorithm to state-of-the-art methods for eye movement classification, we next evaluated its performance on a benchmark data set (Larsson et al. 2013), which has previously been used for the comparison of 12 eye movement classifiers (Andersson et al. 2017; Pekkanen and Lappi 2017). The data set comprises 500-Hz eye tracking recordings from humans watching videos, images, or moving dots, and it contains human labels for fixations, smooth pursuits, saccades,

Table 4. *Statistical tests in interrater comparison*

Metric	Comparison	Test	*t*-Value	*P* Value
Kappa to C1	N1 vs. C2	paired *t*-test	18.38	$2.98\cdot10^{-7}$
Kappa relative to C2	N2 vs. C1	paired *t*-test	10.88	$3.69\cdot10^{-10}$
F1 relative to C1	N1 vs. C2	paired *t*-test	−3.52	$5.08\cdot10^{-2}$
F1 relative to C2	N2 vs. C1	paired *t*-test	−3.7	$5.19\cdot10^{-2}$
Onset distance relative to C1	N1 vs. C2	independent *t*-test	−6.6	$2.98\cdot10^{-5}$
Onset distance relative to C2	N2 vs. C1	independent *t*-test	−13.6	$5.26\cdot10^{-10}$
Offset distance relative to C1	N1 vs. C2	independent *t*-test	−17.9	$5.28\cdot10^{-12}$
Offset distance relative to C2	N2 vs. C1	independent *t*-test	−15.3	$7.33\cdot10^{-11}$

Network 1 (N1) was trained on labels from coder 1 (C1), and network 2 (N2) was trained on labels from coder 2 (C2). All *P* values were Bonferroni corrected for multiple comparisons.

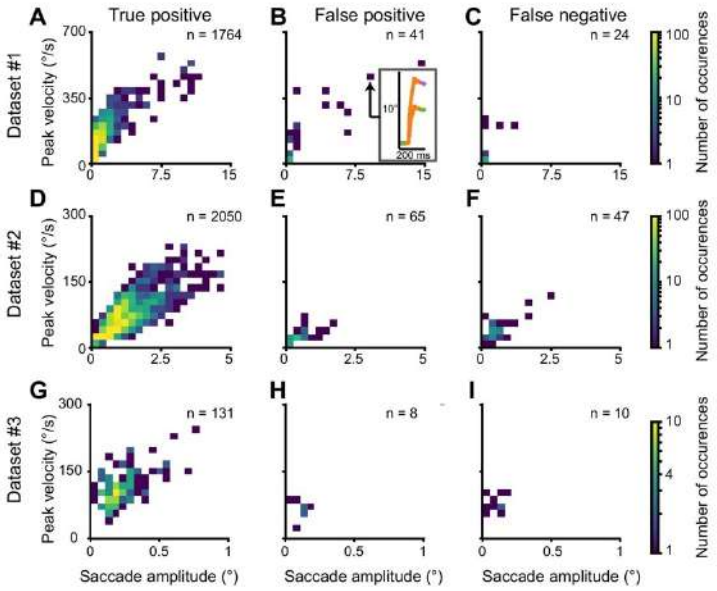


Fig. 5. Location on the main sequence of detected and undetected saccades. *A*, *D*, and *G*: saccades that were detected both by a human expert and U'n'Eye. The detected saccades expectedly followed the main sequence relationship between peak velocity and movement amplitude. *B*, *E*, and *H*: saccades that were detected only by U'n'Eye. Most saccades were small and close to the eye tracker noise, likely being cautiously unlabeled by human coders. In the inset, a large saccade was detected by U'n'Eye but not by the human coder, suggesting a possible lapse by the latter. *C*, *F*, and *I*: saccades missed by U'n'Eye. Most of these were very small.

PSO (Fig. 6*A*), and blinks. We therefore used U'n'Eye as a multiclass classifier to predict saccades, PSOs, and blinks (Fig. 6*B*). Fixations and smooth-pursuit eye movements were both assigned to the fixation class. U'n'Eye output a predictive probability for each class (Fig. 6*D*), with the prediction value corresponding to the class that maximized this predictive probability (Fig. 6*C*). We trained U'n'Eye on one part of the data and evaluated its performance on the test trials listed in Andersson et al. (2017; their Table 11). When considering the whole benchmark data set, U'n'Eye outperformed the state-of-

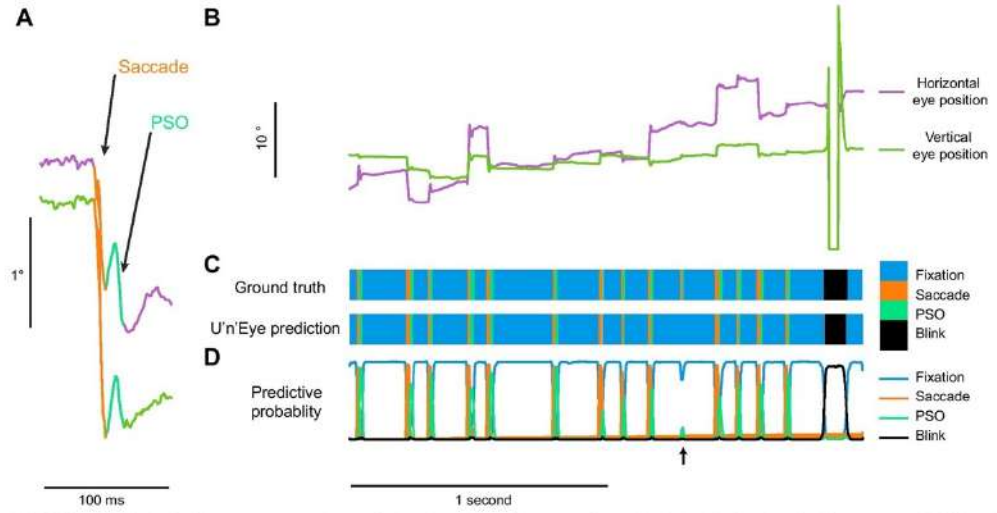


Fig. 6. Multiclass labeling by U'n'Eye. *A*: an example saccade showing substantial postsaccadic oscillation (PSO) from the data in Larsson et al. (2013). *B*: an example full trace from the same data set showing sequences of saccades, PSOs, and blinks. *C*: for the trace in *B*, ground truth labels are shown, in addition to labels by U'n'Eye. The latter successfully classified all ground truth labels, except for one instance marked by a black vertical arrow. *D*: nonetheless, the predictive probability of the network still showed a transient for the missed microsaccade (upward black arrow), suggesting that additional postprocessing may be used to improve the performance of U'n'Eye even more. For example, the user could manually inspect significant transients in predictive probability.

the-art classifiers for saccades and PSOs (Table 5). Moreover, U'n'Eye's performance lay within the range of the intercoder agreement of the two human experts who labeled the data set (Table 5). This result indicates that U'n'Eye is very well suited for multiclass eye movement classification.

Practical considerations for U'n'Eye usage. To better understand the practical aspects of using our approach, we additionally assessed how U'n'Eye performs under different training scenarios. The results of this section can be used as good practice guidelines by the users in their own applications that employ our algorithm.

First, we studied how the amount of training data impacts saccade detection performance. In practice, the available number of annotated training samples might be limited. To achieve good performance of U'n'Eye, small training sets were sufficient (Fig. 7*A*). Even with only 50 s of labeled data, our network outperformed other algorithms. Using more training samples led to a further increase of performance. Training time was also no limiting factor, since training a new network even on a CPU took only ~2 min for every minute of training data (Fig. 7*A*).

In machine learning, the quality of the training data is also crucial for the performance of a classifier, since the latter directly learns from the human ground truth labels. Human labeling, however, is prone to mistakes and lapses: saccades might be missed by the human coder, leading to noisy labels. We therefore assessed how U'n'Eye's performance was influenced by noise-corrupted labels. We evaluated the network's performance when trained on real data (data set 1) from which we artificially removed a fixed fraction of saccade labels. U'n'Eye was robust to the presence of noisy labels in the training data: even with 20% of missing labels, our network outperformed other algorithms (Fig. 7*B*). We also trained the network on simulated data (Fig. 1*E*) for which we knew the ground truth. While noise-corrupted labels in the training data impaired saccade detection performance as expected, this effect could be compensated for by using a larger amount of training data (Fig. 7*C*). This indicates that U'n'Eye can achieve good performance even if the human coder misses some saccades in the training set.

Table 5. *Performance of U'n'Eye compared with state-of-the-art algorithms*

Event		Coder MN	U'n'Eye	NSLR-HMM	LNS
Saccades	Image	0.91	0.89		0.81
	Dot	0.80	0.79		0.75
	Video	0.88	0.89		0.81
	All	0.89	**0.88**	0.82	0.81
PSOs	Image	0.76	0.72		0.64
	Dot	0.59	0.59		0.53
	Video	0.73	0.68		0.63
	All	0.73	**0.70**	0.53	0.64
Blinks	Image	0.92	0.84		
	Dot	0.77	0.71		
	Video	0.82	0.84		
	All	0.91	0.83		

Naive Segmented Linear Regression-Hidden Markov models (NSLR-HMM) and Larsson, Nystr m, and Stridh (LNS) values were taken from Pekkanen and Lappi (2017) and Andersson et al. (2017), respectively. For U'n'Eye, values are the median across 20 independent networks. In bold are the values reached by the best performing algorithm. MN, initial of the expert labeling the dataset (anonymous); PSO, postsaccadic oscillation.

Next, we studied how well an already-trained network can be applied to label saccades in new data, for which no training labels are available. Our results show that this is possible if the new data has broadly the same signal characteristics as the data used for training the network (i.e., if it was sampled with the same eye tracker during a sufficiently similar task). To illustrate this, we trained networks on our first three data sets and evaluated their performance on each data set plus an additional data set 4 on which none of the networks was trained. Data set 4 was similar to data set 1 in that it was recorded with the same eye tracker in human subjects performing fixations (Table 1). Therefore, a network trained on data set 1 performed very well not only in detecting saccades in the same data set, but also in data set 4 (Fig. 7*D*). Overall, good performance was guaranteed when the test set exhibited similar statistics as the training set or was exposed in training to a sufficiently wide variety of training samples (Fig. 7*D*).

Likewise, our network extrapolated well over subjects, for example in large cohort studies with many different observers (as is often the case in clinical investigations of neurological diseases). We studied whether a network trained on data from one subject was able to detect saccades well in data from another subject. To this end, we trained separate networks on data from 10 individual human subjects in data set 4 and applied them to all other subjects. Overall, performance on data that came from the same subject as the training data were only marginally higher than performance on data that came from a different subject (F1 mean and SD: 0.96 and 0.01 vs. 0.92 and 0.08, Fig. 7*E*). The higher standard deviation of intersubject performance was due to the apparent difference between data from certain subjects (Fig. 7*E*). We therefore advise users to combine training data from a few subjects to obtain a network that is able to deal with different signal statistics (Fig. 7*E*, network trained on all). Note that for the network trained on a combination of subjects, we made sure to keep the number of training samples the same as for networks trained on individual subjects. Thus, the better performance was a result of having more variable samples in the training set and not of more training examples being available.

Eye movement representation becomes disentangled along network layers. We finally had a closer look at how the network achieves the separation of two eye states (e.g., fixations and saccades; Fig. 8*A*). In the velocity domain, saccades and fixations can show highly overlapping distributions (Fig. 8*B*). This explains why velocity threshold-based algorithms can fail to distinguish fixations from saccades (Fig. 4). Here, we showed that U'n'Eye can differentiate between fixations and saccades with high accuracy (Fig. 4). The classification was based on the output layer of the network. To illustrate how this decision arises throughout the hidden layers, we performed principal component analysis (PCA) on the features of each convolutional layer. The fraction of explained variance by the first two principal components (PCs) reflects the U-shaped architecture of the network (Fig. 8*C*): in the middle layers, information is distributed across more components than in early and late layers. We projected the hidden layer activations onto the PC space and labeled time bins according to their ground truth labels (fixation or saccade, Fig. 8*D*). We observed in higher layers that the two classes were better separated (Fig. 8*D*). Finally, in the output layer, fixations and saccades became linearly separable (Fig. 8*E*). Thus, through training, the net-

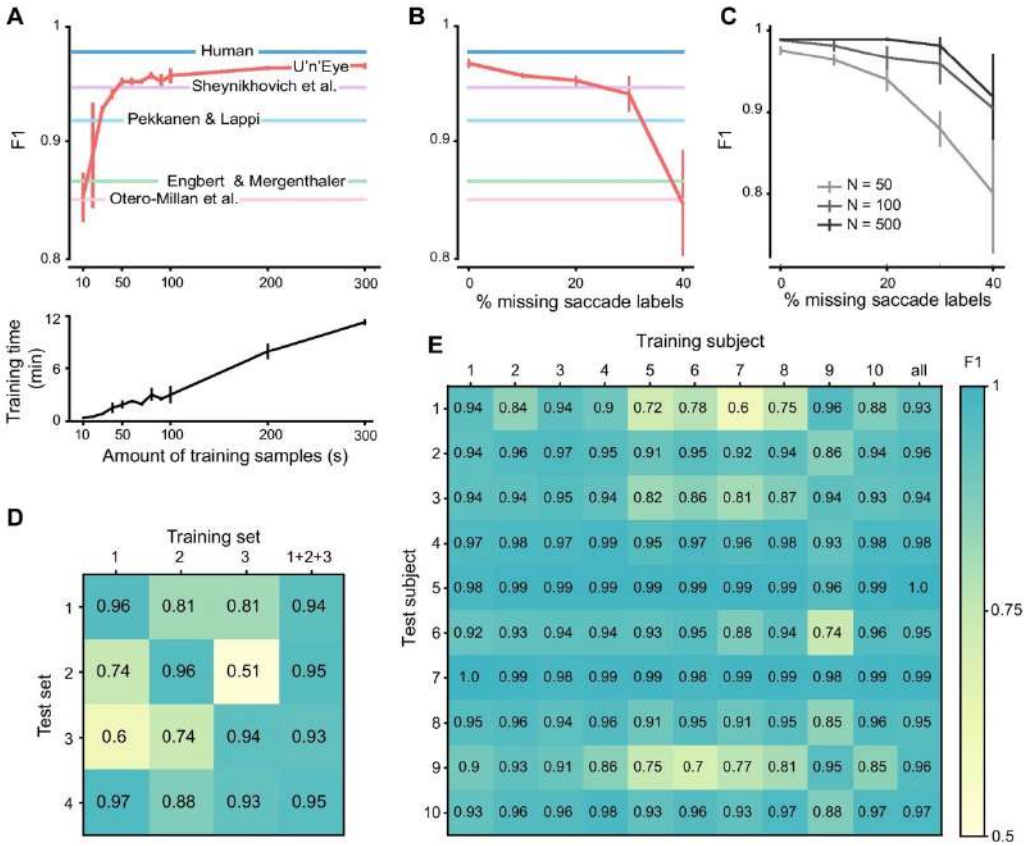


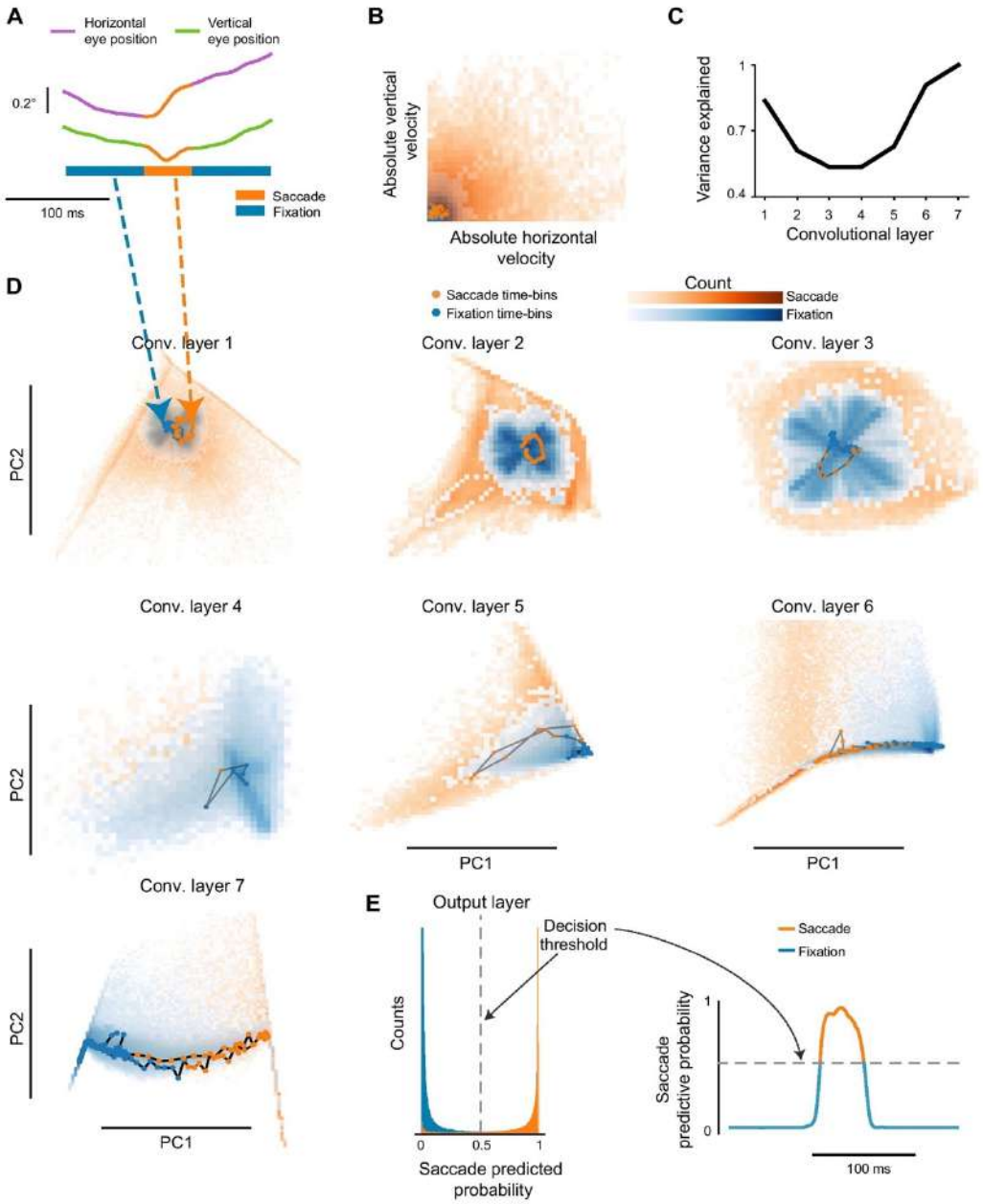
Fig. 7. Robustness of U'n'Eye performance under a variety of training regimes. *A*, *top*: U'n'Eye saccade detection performance (red) as a function of amount of training samples. *Bottom*: linear increase of training time with the number of training samples on a CPU. Data shows mean ± standard deviation across the same training epochs as in the *top* panel. The colored horizontal lines with labels refer to the performance of other algorithms from the literature that we tested. *B*: U'n'Eye saccade detection performance as a function of the fraction of saccade labels missing in 300 s of training data. The colored horizontal lines refer to the same algorithms as in *A*. Also, the red line in *A* and *B* shows mean ± standard deviation across networks trained on three different subsets of data set 1. Performance of U'n'Eye and other algorithms was evaluated on 1,000 s of test data from data set 1. *C*: U'n'Eye saccade detection performance in simulated data with missing labels for different amounts of training samples *N* in seconds. *D*: U'n'Eye saccade detection performance for different combinations of training and test sets. Each number in a square (and its associated color code) indicates the F1 score for training on one data set and testing on another. The column labeled 1+2+3 on the far right shows results when the network was trained on all three data sets simultaneously (but ensuring the same amount of training data as in the other columns of the figure). *E*: U'n'Eye saccade detection performance for combinations of different human subjects in training and test data from data set 4. Each column shows the results of training on a single subject from the data set, in a similar format to *D*. The final column on the right indicates that training the network on a (small) population of subjects yields best performance, and the rest of the figure indicates that additional subjects can then be tested with the pretrained network without much loss in performance. The same color scale was used for *D* and *E*.

work effectively learns to extract relevant features and to project those onto a plane where the two eye movement classes are linearly separable.

DISCUSSION

In this study, we presented U'n'Eye, a convolutional neural network for eye movement classification. We demonstrated that U'n'Eye achieved human-level performance in the detection of saccades and microsaccades. In addition, the network was able to predict other classes of eye movements, which we exemplified with the detection of blinks and PSOs in a benchmark data set.

Furthermore, we showed that U'n'Eye achieved excellent performance both when trained on a single type of data with labels from one coder and when trained on different data sets with labels from two coders. While data sets 1 and 3 used in this study contained data with only one type of visual task and labels from one coder each, data set 2 was composed of two different pursuit tasks and contained labels from two different human coders. Moreover, data set 4 allowed us to conclude that



J Neurophysiol • doi:10.1152/jn.00601.2018 • www.jn.org

our algorithm has generalization properties and can therefore be used when training on a subset of individuals and then testing with large cohorts of subsequent subjects measured with similar eye tracking technology. The data set by Andersson et al. (2017) also contained greatly varying types of saccades and other eye movements. Still, U'n'Eye achieved good performance when trained and tested on this data set. Note that the network might fail to detect eye movements when tested on data that show a very different distribution than the data it was trained on. We therefore recommend to either train a network with a variety of data or to train separate specialized networks for each task.

In this regard, our approach falls in the class of supervised learning algorithms, as opposed to methods not requiring parameter estimation based on annotated data (Engbert and Mergenthaler 2006; Otero-Millan et al. 2014; Sheynikhovich et al. 2018). However, we typically see in different scenarios that casting an algorithmic issue as a supervised problem helps in terms of performance. For example, we recently showed that supervised techniques perform as well as, or better than, unsupervised ones for spike inference from calcium imaging data (Berens et al. 2018; Theis et al. 2016). Similarly, Mathis et al. (2018) recently showed that supervised learning provides superior animal tracking with few annotated samples. We showed here that the situation is similar for eye movement detection. Importantly, we showed that performance generalizes to new unseen data sets and subjects, yielding better performance than any of the unsupervised algorithms. Of course, there is some manual work involved in preparing the training samples for our network, but we posit here that this amount of manual work is significantly less intensive than the manual postprocessing that we typically perform with other saccade-detection algorithms.

U'n'Eye is publicly available and provides a user friendly interface as well as a web service in which users can upload their data and receive classification outputs (see METHODS). No parameter tuning is needed even for training (e.g., learning rate, and so on) since the standard settings were found to work well across data sets. Instead, an experimenter just needs to provide a few hundred seconds of labeled data to train the network once. Even if some labels are missing in the training data, U'n'Eye can still reach high performance. We recommend, however, to use only carefully annotated data for training, as this should improve results.

Of the few algorithms that are capable of detecting saccades as well as PSO (Larsson et al. 2013; Pekkanen and Lappi 2017; Zemblys et al. 2018b), U'n'Eye achieves highest performance. Note that Zemblys et al. (2018a) also recently proposed a deep learning method for eye movement detection. Their approach consists of generating a large training set out of a small human-labeled data set using a generative neural network. A second network is then trained on this data to classify eye movements. This method reports performance similar to that of U'n'Eye in a subset of the benchmark data set by Andersson et al. (2017), but it remains to be seen how this algorithm performs on more exhaustive tasks like the ones that we reported here. For example, the applicability to data containing smooth pursuit has not been demonstrated. Conversely Startsev et al. (2018) recently published a deep learning approach showing reasonable performance, but again they tested only on a subset of the benchmark data set containing smooth pursuit.

Recently, a Bayesian approach for the detection of microsaccades based on a generative model has been proposed (Mihali et al. 2017). Inherently, Bayesian methods provide estimates of uncertainty, in addition to estimates of the quantity of interest. Indeed, it is an interesting future perspective to combine U'n'Eye with Bayesian Deep Learning techniques to provide uncertainty estimates for the detected eye movements (Gal and Ghahramani 2015).

Future work should include combining data sets with different characteristics, such as different sampling frequencies, to obtain a network that can generalize on a large range of data. Such a network could be used by a large part of the scientific community, which would allow for reproducibility of scientific results. We recommend that anyone who uses our algorithm to publish the weights of the trained network so that eye movement detection can be reproduced. For our own trained networks, all weights have been published online (https://github.com/berenslab/uneye) along with the code of the network. This has the advantage that users with similar data characteristics to one of our three data sets (e.g., microsaccades during fixation with a video-based eye tracker as in data set 3) can directly use our weights from the proper data set without having to retrain their own network. We also intend to make all three data sets publicly available, facilitating the further development for eye movement detection algorithms.

Of course, it should be noted that some prediction errors may still occur with U'n'Eye. However, such errors fall within the range of interrater variability across humans anyway. Also, even when U'n'Eye does make mistakes, the predictive probability that it outputs can be used to retrieve missed events (e.g., see the upward black arrow in the bottom of Fig. 6*D*). For example, detecting peaks in the predictive probability output that did not cross the threshold can accelerate eventual manual postprocessing.

Finally, U'n'Eye's capacity to learn nonlinear relationships between an eye trace and some annotated labels opens new horizons in neuroscience: the network could be used to understand the properties of neural activities related in a complex manner to eye movements. For example, the disentanglement in later layers (Fig. 8) could be used to quantitatively analyze the activity patterns of premotor neurons in the brain stem, which themselves ultimately transform brain processing into individual ocular muscle innervations. Furthermore, U'n'Eye could be turned into a generative model for eye movements, as was shown for neural networks that are used for image clas-

Fig. 8. Disentanglement of fixations and saccades throughout the network. *A*: example eye trace with a microsaccade. *B*: distribution of data set 2 in the velocity domain. Fixations and saccades (shown in bluish and orangish colors, respectively) showed overlapping distributions. *C*: fraction of explained variance by the two first principal components (PCs) of the network's convolutional (Conv.) layers. There was a reduction in the middle layers followed by a peak at the final seventh layer. *D*: projection of hidden layer activations by eye traces of data set 2 onto the first two principal components. Fixation and saccade classes became better separated throughout the hidden layers. *B–D*: Dots indicate the time points of the example eye trace in *A*, and the rest of the background data show the entire data set time samples. *E*: the probability output allowed for a linear separation of the two classes. Time points with a saccade predictive probability above 0.5 were classified as a saccade.

sification (Gatys et al. 2015). The information about eye movements that is contained in the network architecture might in the future be used to identify variations in eye movement characteristics that could hint at underlying pathologies.

ACKNOWLEDGMENTS

We thank Konstantin-Friedrich Willeke and Antimo Buonocore for providing help with labeling saccade data, Murat Ayhan for input on deep neural networks, and Adam von Daranyi for setting up the web service.

GRANTS

This work was funded by the German Ministry of Education and Research (FKZ 01GQ1601) and the German Research Foundation (EXC307, EXC2064/1–Project ID 390727645, SFB 1233–Project ID 276693517, BE5601/4-1).

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the authors.

AUTHOR CONTRIBUTIONS

M.E.B., J.B., Z.M.H., and P.B. conceived and designed research; M.E.B. and J.B. analyzed data; M.E.B., J.B., Z.M.H., and P.B. interpreted results of experiments; M.E.B. prepared figures; M.E.B. and J.B. drafted manuscript; M.E.B., J.B., Z.M.H., and P.B. edited and revised manuscript; M.E.B., J.B., H.N., Z.M.H., and P.B. approved final version of manuscript; J.B. and H.N. performed experiments.

ENDNOTE

At the request of the authors, readers are herein alerted to the fact that additional materials related to this manuscript may be found at the institutional website of one of the authors, which at the time of publication they indicate is: at http://uneye.berenslab.org. All code is available from https://github.com/berenslab/uneye. These materials are not a part of this manuscript and have not undergone peer review by the American Physiological Society (APS). APS and the journal editors take no responsibility for these materials, for the website address, or for any links to or from it.

REFERENCES

Andersson R, Larsson L, Holmqvist K, Stridh M, Nystr m M. One algorithm to rule them all? An evaluation and discussion of ten eye movement event-detection algorithms. *Behav Res Methods* 49: 616–637, 2017. doi:10.3758/s13428-016-0738-9.

Bellet J, Chen CY, Hafed ZM. Sequential hemifield gating of α- and β-behavioral performance oscillations after microsaccades. *J Neurophysiol* 118: 2789–2805, 2017. doi:10.1152/jn.00253.2017.

Berens P, Freeman J, Deneux T, Chenkov N, McColgan T, Speiser A, Macke JH, Turaga SC, Mineault P, Rupprecht P, Gerhard S, Friedrich RW, Friedrich J, Paninski L, Pachitariu M, Harris KD, Bolte B, Machado TA, Ringach D, Stone J, Rogerson LE, Sofroniew NJ, Reimer J, Froudarakis E, Euler T, Rom n Ros n M, Theis L, Tolias AS, Bethge M. Community-based benchmarking improves spike rate inference from two-photon calcium imaging data. *PLOS Comput Biol* 14: e1006157, 2018. doi:10.1371/journal.pcbi.1006157.

Bishop CM. *Pattern Recognition and Machine Learning* New York: Springer, 2016.

Borji A, Itti L. Defending Yarbus: eye movements reveal observers' task. *J Vis* 14: 29, 2014. doi:10.1167/14.3.29.

Bosman CA, Womelsdorf T, Desimone R, Fries P. A microsaccadic rhythm modulates gamma-band synchronization and behavior. *J Neurosci* 29: 9471–9480, 2009. doi:10.1523/JNEUROSCI.1193-09.2009.

Buonocore A, Skinner J, Hafed ZM. Eye-position error influence over "open-loop" smooth pursuit initiation. *bioRxiv*: 404491, 2018. doi:10.1101/404491.

Burr DC, Morrone MC, Ross J. Selective suppression of the magnocellular visual pathway during saccadic eye movements. *Nature* 371: 511–513, 1994. doi:10.1038/371511a0.

Carpenter RH. *Movements of the Eyes* (2nd rev. ed.) London: Pion, 1988.

Chen CY, Hafed ZM. Postmicrosaccadic enhancement of slow eye movements. *J Neurosci* 33: 5375–5386, 2013. doi:10.1523/JNEUROSCI.3703-12.2013.

Chen CY, Hafed ZM. A neural locus for spatial-frequency specific saccadic suppression in visual-motor neurons of the primate superior colliculus. *J Neurophysiol* 117: 1657–1673, 2017. doi:10.1152/jn.00911.2016.

Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas* 20: 37–46, 1960. doi:10.1177/001316446002.000104.

Crevecoeur F, Kording KP. Saccadic suppression as a perceptual consequence of efficient sensorimotor estimation. *eLife* 6: e25073, 2017. doi:10.7554/eLife.25073.

Dai W, Selesnick I, Rizzo JR, Rucker J, Hudson T. A parametric model for saccadic eye movement. Signal Processing in Medicine and Biology Symposium (SPMB). Philadelphia, PA, December 3, 2016. doi:10.1109/SPMB.2016.7846860.

Daye PM, Optican LM. Saccade detection using a particle filter. *J Neurosci Methods* 235: 157–168, 2014. doi:10.1016/j.jneumeth.2014.06.020.

Duchowski AT. *Eye Tracking Methodology: Theory and Practice* (2nd ed.). London: Springer, 2007.

Duhamel JR, Colby CL, Goldberg ME. The updating of the representation of visual space in parietal cortex by intended eye movements. *Science* 255: 90–92, 1992. doi:10.1126/science.1553535.

Engbert R, Mergenthaler K. Microsaccades are triggered by low retinal image slip. *Proc Natl Acad Sci USA* 103: 7192–7197, 2006. doi:10.1073/pnas.0509557103.

Fuchs AF, Robinson DA. A method for measuring horizontal and vertical eye movement chronically in the monkey. *J Appl Physiol* 21: 1068–1070, 1966. doi:10.1152/jappl.1966.21.3.1068.

Gal Y, Ghahramani Z. Bayesian convolutional neural networks with Bernoulli approximate variational inference (Preprint). *arXiv* 1506.02158, 2015.

Gatys LA, Ecker AS, Bethge M. A neural algorithm of artistic style (Preprint). *arXiv* 1508.06576, 2015.

Golan T, Davidesco I, Meshulam M, Groppe DM, M gevand P, Yeagle EM, Goldfinger MS, Harel M, Melloni L, Schroeder CE, Deouell LY, Mehta AD, Malach R. Increasing suppression of saccade-related transients along the human visual hierarchy. *eLife* 6: e27819, 2017. doi:10.7554/eLife.27819.

Gur M, Beylin A, Snodderly DM. Response variability of neurons in primary visual cortex (V1) of alert monkeys. *J Neurosci* 17: 2914–2920, 1997. doi:10.1523/JNEUROSCI.17-08-02914.1997.

Hafed ZM. Mechanisms for generating and compensating for the smallest possible saccades. *Eur J Neurosci* 33: 2101–2113, 2011. doi:10.1111/j.1460-9568.2011.07694.x.

Hafed ZM. Alteration of visual perception prior to microsaccades. *Neuron* 77: 775–786, 2013. doi:10.1016/j.neuron.2012.12.014.

Hafed ZM, Chen CY, Tian X. Vision, perception, and attention through the lens of microsaccades: mechanisms and implications. *Front Syst Neurosci* 9: 167, 2015. doi:10.3389/fnsys.2015.00167.

Hafed ZM, Goffart L, Krauzlis RJ. Superior colliculus inactivation causes stable offsets in eye position during tracking. *J Neurosci* 28: 8124–8137, 2008. doi:10.1523/JNEUROSCI.1317-08.2008.

Hafed ZM, Krauzlis RJ. Goal representations dominate superior colliculus activity during extrafoveal tracking. *J Neurosci* 28: 9426–9439, 2008. doi:10.1523/JNEUROSCI.1313-08.2008.

Haji-Abolhassani A, Clark JJ. An inverse Yarbus process: predicting observers' task from eye movement patterns. *Vision Res* 103: 127–142, 2014. doi:10.1016/j.visres.2014.08.014.

Hass CA, Horwitz GD. Effects of microsaccades on contrast detection and V1 responses in macaques. *J Vis* 11: 3, 2011. doi:10.1167/11.3.3.

Herrington TM, Masse NY, Hachmeh KJ, Smith JE, Assad JA, Cook EP. The effect of microsaccades on the correlation between neural activity and behavior in middle temporal, ventral intraparietal, and lateral intraparietal areas. *J Neurosci* 29: 5793–5805, 2009. doi:10.1523/JNEUROSCI.4412-08.2009.

Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift (Preprint). *arXiv* 1502.03167, 2015.

Judge SJ, Richmond BJ, Chu FC. Implantation of magnetic search coils for measurement of eye position: an improved method. *Vision Res* 20: 535–538, 1980. doi:10.1016/0042-6989(80)90128-5.

Kawaguchi K, Clery S, Pourriahi P, Seillier L, Haefner RM, Nienborg H. Differentiating between models of perceptual decision making using pupil size inferred confidence. *J Neurosci* 38: 8874–8888, 2018. doi:10.1523/JNEUROSCI.0735-18.2018.

Kingma DP, Ba J. Adam: a method for stochastic optimization (Preprint). *arXiv* 1412.6980, 2014.

Klibisz A, Rose D, Eicholtz M, Blundon J, Zakharenko S. Fast, simple calcium imaging segmentation with fully convolutional networks. In: *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, edited by Cardoso MJ, Arbel T. Cham, Switzerland: Springer, 2017, p. 285–293. doi:10.1007/978-3-319-67558-9_33.

Kowler E. Eye movements: the past 25 years. *Vision Res* 51: 1457–1483, 2011. doi:10.1016/j.visres.2010.12.014.

Larsson L, Nystr m M, Stridh M. Detection of saccades and postsaccadic oscillations in the presence of smooth pursuit. *IEEE Trans Biomed Eng* 60: 2484–2493, 2013. doi:10.1109/TBME.2013.2258918.

Leigh RJ, Kennard C. Using saccades as a research tool in the clinical neurosciences. *Brain* 127: 460–477, 2004. doi:10.1093/brain/awh035.

Leigh RJ, Zee DS. *The Neurology of Eye Movements*. New York: Oxford University Press, 2015. Contemporary Neurology Series 90.

Leopold DA, Logothetis NK. Microsaccades differentially modulate neural activity in the striate and extrastriate visual cortex. *Exp Brain Res* 123: 341–345, 1998. doi:10.1007/s002210050577.

MacAskill MR, Anderson TJ. Eye movements in neurodegenerative diseases. *Curr Opin Neurol* 29: 61–68, 2016. doi:10.1097/WCO.0000000000000274.

Mathis A, Mamidanna P, Cury KM, Abe T, Murthy VN, Mathis MW, Bethge M. Deeplabcut: markerless pose estimation of user-defined body parts with deep learning. *Nat Neurosci* 21: 1281–1289, 2018. doi:10.1038/s41593-018-0209-y.

Mihali A, van Opheusden B, Ma WJ. Bayesian microsaccade detection. *J Vis* 17: 13, 2017. doi:10.1167/17.1.13.

Otero-Millan J, Castro JLA, Macknik SL, Martinez-Conde S. Unsupervised clustering method to detect microsaccades. *J Vis* 14: 18, 2014. doi:10.1167/14.2.18.

Pekkanen J, Lappi O. A new and general approach to signal denoising and eye movement classification based on segmented linear regression. *Sci Rep* 7: 17726, 2017. doi:10.1038/s41598-017-17983-x.

Reppas JB, Usrey WM, Reid RC. Saccadic eye movements modulate visual responses in the lateral geniculate nucleus. *Neuron* 35: 961–974, 2002. doi:10.1016/S0896-6273(02)00823-1.

Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*, edited by Navab N, Hornegger J, Wells W, Frangi A. Cham, Switzerland: Springer, 2015, p. 234–241. Lecture Notes in Computer Science 9351.

Ross J, Morrone MC, Burr DC. Compression of visual space before saccades. *Nature* 386: 598–601, 1997. doi:10.1038/386598a0.

Sheynikhovich D, B cu M, Wu C, Arleo A. Unsupervised detection of microsaccades in a high-noise regime. *J Vis* 18: 19, 2018. doi:10.1167/18.6.19.

Sommer MA, Wurtz RH. Brain circuits for the internal monitoring of movements. *Annu Rev Neurosci* 31: 317–338, 2008. doi:10.1146/annurev.neuro.31.060407.125627.

Startsev M, Agtzidis I, Dorr M. 1D CNN with BLSTM for automated classification of fixations, saccades, and smooth pursuits. *Behav Res Methods*, 2018. doi:10.3758/s13428-018-1144-2.

Theis L, Berens P, Froudarakis E, Reimer J, Rom n Ros n M, Baden T, Euler T, Tolias AS, Bethge M. Benchmarking spike rate inference in population calcium imaging. *Neuron* 90: 471–482, 2016. doi:10.1016/j.neuron.2016.04.014.

Tian X, Yoshida M, Hafed ZM. A microsaccadic account of attentional capture and inhibition of return in posner cueing. *Front Syst Neurosci* 10: 23, 2016. doi:10.3389/fnsys.2016.00023.

Yao T, Treue S, Krishna BS. Saccade-synchronized rapid attention shifts in macaque visual cortical area MT. *Nat Commun* 9: 958, 2018. doi:10.1038/s41467-018-03398-3.

Yu G, Yang M, Yu P, Dorris MC. Time compression of visual perception around microsaccades. *J Neurophysiol* 118: 416–424, 2017. doi:10.1152/jn.00029.2017.

Zemblys R, Niehorster DC, Holmqvist K. gazeNet: End-to-end eye-movement event detection with deep neural networks. *Behav Res Methods*, 2018a. doi:10.3758/s13428-018-1133-5.

Zemblys R, Niehorster DC, Komogortsev O, Holmqvist K. Using machine learning to detect events in eye-tracking data. *Behav Res Methods* 50: 160–181, 2018b. doi:10.3758/s13428-017-0860-3.

Zirnsak M, Steinmetz NA, Noudoost B, Xu KZ, Moore T. Visual space is compressed in prefrontal cortex before eye movements. *Nature* 507: 504–507, 2014. doi:10.1038/nature13149.

Zuber BL, Stark L, Cook G. Microsaccades and the velocity-amplitude relationship for saccadic eye movements. *Science* 150: 1459–1460, 1965. doi:10.1126/science.150.3702.1459.

Using deep neural networks to detect complex spikes of cerebellar Purkinje Cells

Abbreviated title: Detecting complex spikes using deep learning

Akshay Markanday[1, 2]†, Joachim Bellet[1-3]†, Marie E. Bellet[3]†, Ziad M. Hafed[1, 3*], Peter Thier[1, 3*]

[1]Hertie Institute for Clinical Brain Research, Tübingen, Germany
[2]Graduate School of Neural and Behavioral Sciences, International Max Planck Research School, Tübingen University, Tübingen, Germany
[3]Werner Reichardt Centre for Integrative Neuroscience (CIN), Tübingen, Germany

†A.M, J.B and M.E.B contributed equally to this work
*Z.M.H. and P.T. contributed equally to this work

Correspondence
Peter Thier and Ziad M. Hafed,
Department of Cognitive Neurology,
Hertie Institute for Clinical Brain Research,
Hoppe-Seyler-Str. 3, 72076
Tübingen, Germany.
Email: thier@uni-tuebingen.de and ziad.m.hafed@cin.uni-tuebingen.de

Number of pages: 40

Number of figures: 9

Number of words (abstract/introduction/discussion):237/ 745/ 1759

Conflict of interest: The authors declare no competing financial interests

Acknowledgements: Supported by DFG Research Unit 1847 "The Physiology of distributed computing underlying higher brain functions in non-human primates".

Abstract

One of the most powerful excitatory synapses in the entire brain is formed by cerebellar climbing fibers, originating from neurons in the inferior olive, that wrap around the proximal dendrites of cerebellar Purkinje cells. The activation of a single olivary neuron is capable of generating a large electrical event, called "complex spike", at the level of the postsynaptic Purkinje cell, comprising of a fast initial spike of large amplitude followed by a slow polyphasic tail of small amplitude spikelets. Several ideas discussing the role of the cerebellum in motor control are centered on these complex spike events. However, these events are extremely rare, only occurring 1-2 times per second. As a result, drawing conclusions about their functional role has been very challenging, as even few errors in their detection may change the result. Since standard spike sorting approaches cannot fully handle the polyphasic shape of complex spike waveforms, the only safe way to avoid omissions and false detections has been to rely on visual inspection of long traces of Purkinje cell recordings by experts. Here we present a supervised deep learning algorithm for rapidly and reliably detecting complex spikes as an alternative to tedious visual inspection. Our algorithm, utilizing both action potential and local field potential signals, not only detects complex spike events much faster than human experts, but it also excavates key features of complex spike morphology with a performance comparable to that of such experts.

Key words: Convolutional neural network**,** complex spike, simple spike, LFP, action potentials, cerebellum

65 **Significance statement**

66 Climbing fiber driven "complex spikes", fired at perplexingly low rates, are known to play a
67 crucial role in cerebellum-based motor control. Careful interpretations of these spikes require
68 researchers to manually detect them, since conventional online or offline spike sorting
69 algorithms (optimized for analyzing the much more frequent "simple spikes") cannot be fully
70 trusted. Here, we present a deep learning approach for identifying complex spikes, which is
71 trained on local field and action potential recordings from cerebellar Purkinje cells. Our
72 algorithm successfully identifies complex spikes, along with additional relevant
73 neurophysiological features, with an accuracy level matching that of human experts, yet with
74 very little time expenditure.

75

76

77

78

79

80

81

82

83

84

85

Introduction

The Purkinje cell (PC) output, the sole output of the cerebellar cortex, is characterized by two distinct types of responses (Fig. 1A, bottom), the simple spike (SS) and the complex spike (CS) (Thach, 1968). SSs are ordinary sodium-potassium spikes with a simple bi- or tri-phasic shape in extracellular recordings (Fig. 1B). These spikes, lasting only a fraction of a millisecond and firing up to several hundred times per second, reflect the concerted impact of mossy fiber input, mediated via the granule cell-parallel fiber system, as well as inhibitory interneurons. On the other hand, an individual CS (Fig. 1C), elicited by a single climbing fiber originating from the inferior olivary nucleus and pervading the proximal dendrites of a PC, is characterized by a polyphasic somatic spike consisting of a first back propagated axonal spike component followed by a series of spikelets riding on a long-lasting, calcium dependent depolarization (Eccles et al., 1967; Fujita, 1968; Thach, 1968; Llinas and Sugimori, 1980; Stuart and Häusser, 1994; Davie et al., 2008). In addition to an exceptional morphology, CSs also exhibit an unusual, perplexingly low firing rate of at most two spikes per second (Fig. 1A, bottom). What could these infrequent, yet unique events possibly tell us about their purpose, and what might be the best statistical tool allowing us to unravel the full extent of information carried by them? These are questions that have kept researchers busy until today.

Thinking about the role of CSs has been guided by two, not necessarily incompatible, ideas: motor timing and motor learning. The first idea, championed by Llinás and his coworkers, was prompted by the characteristic 8-10 Hz rhythmicity and synchronicity of inferior olivary neurons, a pattern that seemed to reflect the temporal structure of many forms of motor behavior, as well as physiological and pathological tremor (Llinas, 1974; Leznik and Llinás, 2005). The second idea emphasized the role of performance errors in driving motor learning. On experiencing an error, the climbing fiber system is assumed to produce a CS, which helps

111 to predictively correct future manifestations of the same motor behavior by modifying the
112 impact of parallel fibers on targeting PCs (Marr, 1969; Albus, 1971; Ito, 1972). This concept
113 has indeed received support from a number of experimental studies (Oscarsson, 1980;
114 Kitazawa et al., 1998; Medina and Lisberger, 2008; Herzfeld et al., 2015, 2018). However,
115 not all findings have been fully compatible with this so-called Marr-Albus-Ito hypothesis, at
116 least not in its original form. For instance, recent work on oculomotor learning has suggested
117 that CS discharge is not only influenced by a current error, but also by a memory of past
118 errors suitable to stabilize behavioral adaptations (Catz et al., 2005; Dash et al., 2010; Junker
119 et al., 2018). An analogous influence of past errors on CS discharge has also been noted in
120 recent studies of eye-blink conditioning (Ohmae and Medina, 2015). Finally, others have
121 advocated that CSs may not be confined to encoding unexpected errors, but to also offer a
122 prediction of the multiple kinematic parameters of the upcoming movement (Streng et al.,
123 2017).

124 Reaching consensus on the diverse views of CS functions would be substantially facilitated
125 by more data on these sparse neural events, collected in conjunction with advanced
126 behavioral paradigms. Yet, it is exactly their unique properties of rarity and complex and
127 highly idiosyncratic spike morphology that have hampered progress. In fact, CS spike
128 morphology not only differs between individual PCs, but it also often changes over the
129 course of a single recording from the same PC. This is why using standard spike sorting
130 software to detect CSs has turned out to be error prone. Critically, given the rarity of CSs,
131 even a few missing or erroneously detected CS events will have profound impacts on
132 conclusions drawn about their functional role. Consequently, researchers are compelled to
133 meticulously label CSs manually, or at least to visually control the CSs detected by
134 conventional spike sorting approaches, an exhausting approach that constrains the amount of
135 experimental data that can be processed.

In this paper, we exploited a state-of-the-art convolutional neural network (CNN) approach to dramatically reduce the burden of investigators in identifying CSs. We show that our network is able to learn fast and that it easily matches the performance of an experienced human expert in detecting CSs. Our algorithm also extracts a number of key parameters on CS timing and morphology, in a regularized and systematic manner, which we believe is particularly important for understanding the functional role of CSs.

Materials and Methods

Animals, preparation, surgical procedures, and recording methods

Two adult male rhesus macaques (*Macaca mulatta*) of age 10 (monkey K) and 8 (monkey E) years, purchased from the German Primate Center, Göttingen, were subjects in this study. Initial training of all animals required them to voluntarily enter an individually customized primate chair and get accustomed to the setup environment, a procedure that could last for up to three months. Following initial training, they underwent the first major surgical procedure in which foundations of all implants were fixed to the skull using titanium bone screws, and then allowed to rest for a period of approximately 3-4 months to improve the long-term stability of the implant foundations. Then, a titanium-based hexagonal tube-shaped head post was attached to the implanted head holder base to painlessly immobilize the head during experiments, and scleral search coils were implanted to record eye positions using electromagnetic induction (Judge et al., 1980; Bechert and Koenig, 1996). Within 2-3 weeks of recovery from the eye-coil implantation procedure, monkeys quickly recapitulated the already learned chair-training protocol, and were trained further on their respective behavioral paradigms. Once fully trained, a cylindrical titanium recording chamber, whose position and orientation were carefully planned based on pre-surgical MRI and later

confirmed by post-surgical MRI, was finally mounted on the implanted chamber base, tilting backwards by an angle of 30° with respect to the frontal plane, right above the midline of the cerebellum. A part of the skull within the chamber was removed to allow precise electrode access to our region of interest, the oculomotor vermis (OMV, lobuli VIc/VIIa), for electrophysiological recordings. All surgical procedures were carried out under aseptic conditions using general anesthesia, and post-surgical analgesics were delivered until full recovery. See Prsa et al. (2009) for full details. All experiments and surgical procedures were approved by the local animal care authority (Regierungspräsidium Tübingen) and complied with German and European law as well as the National Institutes of Health's *Guide for the Care and Use of Laboratory Animals*. All procedures were carefully monitored by the veterinary service of Tübingen University.

Behavioral tasks

In-house software (NREC), running on a Linux PC (http://nrec.neurologie.uni-tuebingen.de), was used for data collection, stimulus presentation, and operations control. The two monkeys were trained on a fatigue inducing repetitive fast eye movements (saccades) task (Fig. 1A, top; Prsa et al., 2010). A trial started with a red fixation dot (diameter: 0.2°) displayed at the center of a CRT monitor placed 38 cm in front of the monkey. After a short and variable fixation period (400-600 ms from trial onset), the fixation dot disappeared and at the same time, a target, having the same features as the fixation dot, appeared on the horizontal axis at an eccentricity of 15°. In a given session, the target was presented consistently either on the left or right of the central fixation dot. The maximum number of trials (>200) per session depended on the willingness of the monkey to cooperate and on the duration for which a PC could be kept well isolated. Each trial lasted for 1200 ms, and inter-trial intervals were kept

very short (100 ms) to maximize the induction of fatigue. At the end of every correct trial, monkeys were rewarded with a drop of water.

Electrophysiological recordings

Extracellular recordings with commercially available glass-coated tungsten microelectrodes (impedance: 1-2 MΩ; Alpha Omega Engineering, Nazareth, Israel) were performed using a modular multi-electrode manipulator (Electrode Positioning System and Multi-Channel Processor, Alpha Omega Engineering) whose position was estimated, based on the position and orientation of the chamber relative to the brain, using a stereotactic apparatus and later confirmed by post-surgical MRI scans. Saccade-related modulation of an intense background activity, reflecting multi-unit granule cell activity, paralleled by saccade-related modulation in the local field potential record (LFP, <150 Hz bandwidth) served as electrophysiological criteria for identifying the OMV (Fig. 1A, middle). Extracellular potentials, sampled at 25 KHz, were high band-pass (300 Hz - 3 KHz) and low-pass filtered (<150 Hz) to differentiate PC action potentials and LFP signals, respectively (Fig. 1A, bottom).

Multi Spike Detector: the online spike sorting algorithm

Single PC units were identified online by the presence of a high-frequency SS discharge accompanied by the signatory, low-frequency CS discharge using a real-time spike sorter, the Alpha Omega Engineering Multi Spike Detector (MSD). The MSD, designed for detecting sharp waveforms uses a template matching algorithm developed by Wörgötter et al. (1986), sorts waveforms according to their shape. The algorithm employs a continuous comparison of the electrode signal against an 8-point template defined by the experimenter to approximate the shape of the spike of interest. The sum of squares of the difference between template and electrode signal is used as a statistical criterion for the goodness of fit. Whenever the

goodness of fit crosses a threshold, the detection of a spike is reported. The 8-point template can be adjusted manually or alternatively, run in an adaptive mode that allows it to keep track of waveforms that may gradually change over time.

Identification of simple spikes and complex spikes in Purkinje cells

As opposed to short duration SSs (Fig. 1B), characterized by short median inter-spike intervals (Fig. 1E), the long duration CSs (Fig. 1C) were much more rare. In addition to the 10-20 msec long pause triggered by a CS in the SS firing (e.g. Fig. 1F, Bell and Grimm, 1969; Latham and Paul, 1971; McDevitt et al., 1982), the presence of a CS was also indicated by a massive deflection of the LFP signal, lasting for the whole duration of a CS (Fig. 1D). While the MSD-based detection of abundantly available SS events can be trusted most of the time, since the consequences of erroneously including or missing a few SSs are less problematic, MSD-based detection of much rarer CS events is error prone, the costs of which cannot be neglected. Consequently, thorough analysis of PC data often requires experimenters to visually control the quality of MSD-based detections post-hoc, and many times, to even manually identify CS events.

Convolutional neural network

We used the architecture of a CNN that was originally designed to segment images ("U-Net", Ronneberger et al., 2015) and later successfully adapted for the detection of saccades in eye position recordings ("U'n'Eye"; see Bellet et al. (2018) for details). For CS detection, we input the LFP and action potential signals, sampled at the same frequency of 25 KHz, to the network (Fig. 2A, top). The output was a bin-wise predictive probability of CS occurrence (Fig. 2A, bottom).

The network consists of convolutional and max-pooling layers. Max-pooling is an operation that down-samples the input in order to reduce the dimensionality of its representation in the network. It filters the input with a certain window size and extracts only the maximum value. It then steps further on the input, repeating the same operation on the next time window. Convolutional layers extract relevant features of the input signal by learning the parameters of its convolutional kernel during training. We chose the size of the max-pooling (mp) and convolutional kernels (c) as 7 and 9 bins, respectively. These influence the signal interval (SI) taken into account for labeling one time bin in the output, as described by the formula,

$$SI = \frac{mp^2 + (mp^2 \times c) + (mp \times c) - mp + 2 \times c - 2)}{2}$$

In our case, the SI corresponds to 281 time bins before and after each classified bin.

Training and testing procedures

We recorded a total of 160 PCs, out of which 119 PCs were selected, based on careful visual assessment of MSD-based CS detection by a human expert (author AM), for in-depth statistical analysis. These PCs remained stable throughout the recording session with clearly isolated CSs and associated signatory SS pauses and LFP deflections. The remaining 41 PCs, for which it was deemed that MSD-based analysis might have led to spurious detections of SSs and CSs, were excluded from analysis.

To prepare the training set, we asked our human expert, who is experienced in electrophysiological recordings from PCs, to visually identify CS events and manually label their start and end points. The expert used small segments of action potential and LFP recordings during labeling, without access to eye movement data. For each PC, 24 segments, each 250 ms long, were manually labeled. To avoid having segments in which a part of a CS may have been truncated (at the beginning or end of a segment), we excluded the first and last 9 ms of each segment during training, thereby reducing its size to 232 ms. Since the

network was trained on the manually labeled data, recording segments from the excluded set of 41 PCs, for which the MSD-based CS detection was poor but the human expert-based visual identification was still feasible, were also included for training the network in addition to the selected set of 119 PCs. The number of recording segments for a given PC included in training naturally varied with the number of CSs found in the particular cell, but we ensured including recording segments from all 160 PCs in training.

Since the MSD-based CS detection in 41 PCs was already unsatisfactory, as stated above, a comparison based on the performance of our algorithm and the MSD on these particular PCs would have been too biased in favor of our algorithm. Therefore, to fully test our algorithm's performance while still giving the MSD-based approach the benefit of the doubt, we used cross-validation on recordings from only the selected pool of 119 PCs. For every PC tested for CS detection, we trained a separate network excluding the currently tested PC from the training set. This allowed us to test how well the network generalized to new data sets, on which it had not been trained, and it also allowed us to have multiple performance tests on our algorithm. Therefore, the training set always comprised the remaining 159 PCs not being currently tested. The total number of recording segments used in any given training set was 970-988, depending on the PC under test. Other parameters of network training such as loss function, learning rate, batch size, and early stopping criterion, were chosen as described in Bellet et al. 2018 for U'n'Eye.

We also performed one more performance test of our algorithm, which was concerned with establishing consistency with expert labeling. For 7 PCs (out of our 119 selected ones described above), we asked our human expert to manually label CSs in the entire records, and not just a small training subset within each of them. This allowed us to directly compare the labeling of the entire records of these 7 PCs by both our algorithm and the human expert. Our

algorithm in this case was based on training the network on segments from the remaining 159 PCs (other than the currently tested one), as described above.

Post-processing

We implemented three post-processing steps to enhance the quality of CSs detected by our algorithm. First, time shifts between the detected start points of all CSs fired by a particular PC were corrected by re-aligning them. To this end, we computed the average waveform from the first estimation of start times of all detected CSs. This average-waveform template was then used as a reference to realign each waveform within a ±2 ms window around CS start so that the cross-correlation was maximized (Fig. 2B). Second, action potential and LFP waveforms, occurring within 2 ms after CS start, were projected onto a two-dimensional plane (Fig. 2C) using the UMAP dimensionality reduction technique (McInnes et al., 2018). This allowed us to use the third post-processing step to cluster waveforms into suitable CSs and unsuitable ones. In this third step, groups of waveforms were identified (Fig. 2D) using HDBSCAN, a hierarchical clustering algorithm (Campello et al., 2013) that builds a tree to describe the distance between data points. The algorithm minimizes the spanning size of the tree and further reduces the complexity of the tree to end up with a minimum number of leaf nodes, corresponding to the clusters. We used the default parameters for HDBSCAN with the option to find only one cluster. Waveforms were excluded if they belonged to a cluster for which the average predictive probability output from the network remained below 0.5 for more than 3 ms (Fig. 2E).

Quality metrics

We evaluated the performance of our algorithm in detecting CSs using the so-called F1 score (Dice, 1945; Sørensen, 1948), which compares the consistency of CS labels predicted by the

algorithm, to "ground-truth" labels provided by the human expert. The F1 score is the harmonic mean of recall (the ratio of true positive detections and all true CS labels) and precision (the ratio of true positive detections and all CS labels predicted by the algorithm), as given by the following equation

$$F1 = \frac{2 \times recall \times precision}{recall + precision}$$

In our case, an F1 score of 1 would suggest that the CSs predicted by our algorithm perfectly matched the "ground-truth" labels provided by the human expert. However, a lower F1 score may suggest that CSs were either erroneously missed or falsely detected. For quality assessment, we also computed the post-CS firing rate of SSs, a signatory feature immune to labels detected by the human expert, which served as a reliable and objective criterion for the identification of a CS. Finally, the resulting CS waveforms were scrutinized by visual inspection.

Results

CNN-based algorithm reliably detects complex spikes

The main idea of our approach was to train a classifier to extract relevant features from electrophysiological recordings of PCs and to identify CSs. This was realized with the help of a CNN that uses the LFP and action potential signals as inputs (Fig. 2A, top). We chose these two inputs because human experts achieve consensus on the presence or absence of a CS, more easily and reliably, if both action potentials and LFPs are simultaneously available. Our network uses convolutional and max pooling operations to extract the temporal features relevant for distinguishing CSs from the surrounding signal. In the end, the network predicts the probability of the presence of a CS for each time bin. Time bins for which the predictive probability exceeded the threshold of 0.5 are classified as CSs (Fig. 2A, bottom). The

prediction for each time bin depends on an interval in the input signal whose size is determined by the size of the max-pooling and convolutional kernels of the CNN (Methods). Our analysis considered an interval of 281 time bins before and after the time bin containing a predicted CS event. As our sampling rate was 25 kHz, a 10 ms duration CS would span 250 time bins. This means that the network was often using information surrounding CS events (281 versus 250 time bins) to classify CSs.

One of the key requirements for correct CS classification is the quality of the recorded PC signal, which may naturally depend on several factors. For example, subtle drifts between electrode tip and the cell body during a recording session can lead to sudden or gradual changes in the signal-to-noise ratio of the PC signal, and potentially change the morphology of the CS waveform. Also, several SSs firing in close proximity to each other might lead to complex waveforms that may erroneously be detected as CS events. Furthermore, there is also a possibility of CS waveforms being modified by the presence of preceding SSs (Servais et al., 2004; Zang et al., 2018). In order to make our algorithm more resilient to such influences, we added automatic post-processing steps at the output of the CNN. We first fine-tuned the CS start points (Fig. 2B, Methods), and we then differentiated between candidate waveforms using a clustering algorithm in a dimensionally-reduced space (Fig. 2C, Methods). The waveform clusters after dimensionality reduction represented potential candidates for CSs of the recorded PC. Some of these candidates needed to be excluded. For example, if the network in the first step mistakenly classified non-CS events as CSs, then the clustering method would help to refine the classification and exclude these events post-hoc: amongst the CS events erroneously detected by the network might be SSs that are revealed by a separate cluster in the two-dimensional space (Fig, 2C and D, black vs. orange and blue). These false positive events were removed by applying a threshold to the average predictive probability output of the network of the respective cluster (Fig. 2E). Not only non-CS events

might have contributed to a distinct cluster separated from the main CS cluster, but true CSs with slightly deviant waveforms (Fig. 2D orange vs. blue) might also have led to separate clusters in the two-dimensional space (Fig 2C orange vs. blue). For all CS clusters that met the defined threshold criterion on predictive probability (Fig. 2E, cluster 1 and 2), the output of our algorithm, CS timing and corresponding cluster IDs, allowed the user to carefully inspect each cluster and decide whether to include clusters with deviant, yet true, CSs or not.

Objective quality measure confirms identity of complex spikes

It is well-established that SS firing rate decreases during 10-20 ms after the emission of a CS (Bell and Grimm, 1969; Latham and Paul, 1971; McDevitt et al., 1982, Fig. 1F). This physiological feature, independent of the subjective assessment of the human expert, provided us with an additional means for objectively measuring the CS labeling quality of our algorithm. For 119 PCs, we evaluated SS firing rates before and after the occurrence of CSs detected by our algorithm. As depicted in Fig. 3, CSs identified by the algorithm were followed by a clear and significant decrease in the neurons' SS firing rates by 96% on average (Fig. 3A). In the pre-CS period of 3 to 8 ms, median SS firing rate of the 119 PCs was 58.7 spikes/s; this dropped to 10.5 spikes/s in the post-CS period of 10-15 ms (Fig. 3B, Wilcoxon signed-rank test: $p = 2.18 \times 10^{-20}$). This indicates a very low probability of false positive CS detections, since such false positives would increase the apparent post-CS firing rate of SSs.

The new algorithm outperforms a widely-used online sorter

The spike sorting application MSD, based on a template matching algorithm suggested by Wörgötter et al. (1986) for online CS detection, has been widely used by several laboratories

379 as an aid in supporting the visual inspection of PC records (e.g. Catz et al., 2005). This is
380 why we compared the performance of our CNN-based approach to that of the MSD for the
381 same 119 PCs used to test the performance of the algorithm in the previous section. Overall,
382 our algorithm detected 23% more CS events than the MSD ($p = 1.4 \times 10^{-25}$, Wilcoxon signed-
383 rank test; Fig. 4A). In order to objectively quantify the difference in CS detection by our
384 algorithm and the MSD, and to verify that the additionally detected events were indeed CSs,
385 we again evaluated the decrease of post-CS SS firing rate. The median decrease of SS firing
386 rate after CSs detected only by our algorithm and not by the MSD was significantly stronger
387 than the decrease induced by CSs detected only by the MSD and not by our algorithm ($p =$
388 1.4×10^{-5}, Wilcoxon signed-rank test; Fig. 4B). This indicates that the CSs detected by our
389 algorithm and missed by the MSD were veridical, whereas CSs only detected by the MSD
390 and not by our algorithm were probably erroneous detections (false positives). This view is
391 also supported by a consideration of the time course of SS firing rate aligned to the start time
392 of detected CSs. SS firing rate for CSs only detected by our algorithm and not by the MSD
393 revealed a peak, approximately 3 ms earlier than in the case of CSs that were detected only
394 by the MSD (Fig. 4C). This suggests that SSs occurring shortly before a CS altered the
395 waveform of the latter (Servais et al., 2004) (also see Fig. 2D showing how the amplitude of
396 the average CS waveform of cluster 2 was reduced), therefore impeding its detection by the
397 MSD.

398 We also found that CS waveforms for CSs only detected by our algorithm and not by the
399 MSD were similar in shape to the CSs detected by both our algorithm and the MSD (Fig. 5,
400 middle column vs. left). CSs labeled only by the MSD, on the other hand, deviated from this
401 waveform shape (Fig. 5 right vs. left). This impression clearly also concurs with the weaker
402 post-CS depression of SS firing rate seen in the pool of CS events detected only by the MSD

(Fig. 4C). In summary, our algorithm is both more sensitive and less error prone than the MSD-based detection.

We also evaluated to what extent the predictions from both approaches agreed with labels from a human expert. To this end, we computed the F1 score (see Methods) on short recording segments from the same 119 neurons as in the previous section for which we had "ground-truth" labels from the human expert. The F1 score is a measure of consistency in performance between an algorithm and the human expert. As shown in Fig. 6, our algorithm achieved overall higher F1 scores than the MSD, and it also showed much less variability between the different PC records (Fig. 6A). In fact, for the majority of recorded PCs, our algorithm agreed with the human expert on all CS labels, reflected by an F1 score of 1. This indicates that the predictions by our approach are more "human-like" than the ones labeled by the MSD. To achieve good performance in terms of F1 score, our algorithm also did not need a lot of training data. With only 50 training records of 232 ms of data each (sampled at 25 kHz), our algorithm outperformed the MSD algorithm (Fig. 6B). Larger training sets, of course, yielded even higher performance (Fig. 6B).

CNN approach reaches human expert-level performance

Finally, for 7 PCs, we asked our human expert to fully label the entire recorded data for each neuron, instead of only a tiny training set (Methods). We then compared the CS labels of our algorithm to the ones placed by the human expert on the entire records of the neurons (spanning a time range of approximately 8-14 minutes of neural recording). Overall, the predictions of our algorithm agreed very well with the human labeling (Fig. 7A). A few events were identified as CSs by our algorithm but not by the human expert. However, also the waveforms of these events matched the waveforms of CSs that were labeled by the

427 human expert (Fig 7A, cells 3, 5, and 6), indicating that the CSs ignored by the expert were
428 indeed genuine CSs. For one of the PCs, the waveforms of additionally detected events
429 indicated that our algorithm mistakenly labeled some SSs as CSs (Fig. 7, cell 7). These false
430 positive detections, whose average predictive probability remained above the threshold (0.5)
431 for more than 3 ms and were not removed during automatic post-processing, however, would
432 appear as isolated clusters after dimensionality reduction (Fig. 2C). Hence, such false
433 detections could be easily removed post-hoc by inspecting the properties of the CSs in the
434 respective isolated cluster. For false positive labels, the average duration of pause in SS firing
435 after these events would also be reduced to the average refractory period of SSs in this
436 recording.

437 The comparison with human labels further showed that our algorithm reliably identified the
438 ends of CSs and, considering knowledge of CS start, provided a quantitative estimate of CS
439 duration. For the recording segments from the 119 PCs, we compared the end times of all
440 CSs that were detected by both our algorithm and the human expert. As shown in Fig. 8A, the
441 estimate of CS end times provided by our algorithm and the human expert differed only very
442 slightly. Correspondingly, average CS durations per neuron predicted by our algorithm and
443 the human expert were highly correlated ($\rho = 0.78$, $p = 1.12 \times 10^{-22}$, Spearman correlation;
444 Fig. 8B). In light of a possible CS duration code supplementing a CS rate code (Yang and
445 Lisberger, 2014; Herzfeld et al., 2015; Warnaar et al., 2015; Herzfeld et al., 2018; Junker et
446 al., 2018), it is important to precisely identify the end times of CSs and to track changes in
447 CS duration in conjunction with behavioral changes even within individual PCs. Our
448 algorithm was indeed capable of identifying small variations in CS duration similar to the
449 expert. This is indicated by a strong correlation ($\rho = 0.62$, $p = 6.81 \times 10^{-92}$, Spearman
450 correlation) of the residuals of human-labeled and algorithm-labeled CS end times of the

451 selected 119 PCs, obtained by subtracting the mean CS duration of the respective PC (Fig.
452 8C).

453

454 **Discussion**

455 This study proposes a largely automated approach to CS detection as a sensitive and reliable
456 alternative to tedious and experience-dependent manual labeling. The approach is based on a
457 CNN, trained on two input vectors (Fig. 9A), a high frequency band pass signal for the
458 extraction of action potentials and a simultaneously sampled lower-frequency band pass
459 signal reflecting LFPs. After training with surprisingly little data, our algorithm outperformed
460 a widely used spike sorter deploying a user defined template. Moreover, our algorithm also
461 easily caught up with the performance of an experienced human expert. Searching manually
462 for rare events like CSs, amidst a sea of high-frequency SS signals, not only requires several
463 weeks of tedious effort, but, as demonstrated by research on visual search (Wolfe et al., 2005;
464 Evans et al., 2011), is also error prone, even among experts. Our network renders CS
465 detection not just feasible, but also, more objective and systematic. Steps describing the
466 general workflow of our algorithm are summarized in Fig. 9.

467

468 *Limitations of conventional spike sorting algorithms*

469 The major challenge that any approach for detecting CSs meets is the polymorphic
470 complexity of these neural events (Warnaar et al., 2015). The MSD spike sorter relies on user
471 defined templates to identify distinct spike waveforms. However, no matter how well isolated
472 a PC neuron may be, spike waveforms may change for internal reasons or because the
473 position of the neuron relative to the electrode may drift over time. The MSD, like other

automatic online or offline sorting approaches, tries to accommodate these changes by adapting the original template. The principal virtue of template adaptation notwithstanding, it may not be sufficient to keep track of a changing CS or, alternatively, may gradually render the template indistinguishable from the waveforms of unrelated neural activity (including the much more frequent SSs in the signal). Hence, the sorter may miss a true CS or falsely qualify other waveforms as CSs because of similar morphological features. To avoid erroneous detections and omissions, most analysts resort to manual detection. Experienced human experts may in principle reach a high level of agreement by using visual search to identify CS events. However, this approach is very tedious and therefore inevitably associated with fluctuations of attention, which jeopardizes the analyst´s performance (Wolfe et al., 2005). The tediousness of the manual detection approach is increased even further if attempts are made to pinpoint the times of CS start and end or to identify distinct features of the CS morphology such as its spikelet architecture (Warnaar et al., 2015). Conventional spike sorters based on template matching (Catz et al., 2005; Dash et al., 2010; Herzfeld et al., 2015, 2018; Junker et al., 2018) or even simpler voltage-threshold crossings can be useful to facilitate visual inspection. However, the need to double check detected CS events will forestall gains in investments of time and effort only minimally.

Our algorithm is more sensitive and performs better than the online sorter

Our CNN-based algorithm, trained on action potential and LFP signals, clearly outperformed the MSD. Not only was it more sensitive in detecting more CSs, but it also rejected many false CSs, as compared to the MSD. This can best be seen in the example of Fig. 2C. In this figure, the Cluster 1 waveforms, despite sharing a similar shape of the initial spike component with the genuine CSs in Cluster 3, appeared as a clearly separated group in our

498 dimensionally reduced space. These erroneous waveforms were therefore safely rejected. On
499 the other hand, waveforms belonging to Cluster 2, neighboring the main Cluster 3, were still
500 accepted due to close resemblance of their features to the genuine ones.

501 It is likely that there can be interactions between SS occurrence and CS waveform
502 appearance. Specifically, a study on PCs in non-anaesthetized mice has demonstrated that the
503 shape of the CS waveform can be altered by preceding SSs (Servais et al., 2004).
504 Furthermore, recently conducted experiments on climbing fiber responses in PCs have
505 revealed that the potassium currents, by means of voltage gating in a branch-specific manner,
506 can regulate the climbing fiber driven calcium ion influx leading to changes in CS waveform
507 amplitude (Zang et al., 2018). This may explain why the additional CSs detected by our
508 algorithm might have potentially deceived the online sorter. The genuine nature of the
509 additional CSs detected by our algorithm was confirmed with the help of another prominent
510 physiological marker-a pause in spontaneous firing activity of SSs 10-20 ms right after the
511 occurrence of a CS. The additional CSs that were detected by the online sorter and not by our
512 algorithm did not show a clear suppression of SS firing.

513 A major factor, contributing to unsatisfactory performance of conventional sorters, is the fact
514 that they typically rely only on information from the action potential record, rather than using
515 complementary information from time synchronized LFP recordings, which is what human
516 experts would do when searching PC recordings for CSs. In accordance with a very recent
517 Principal Component Analysis (PCA) based approach (Zur and Joshua, 2019), demonstrating
518 improved CS sorting by exploiting LFP frequency bands, the high performance of our
519 algorithm in detecting CSs also critically relies on the use of LFP signals. The virtue of the
520 PCA-based approach notwithstanding, it is clearly outperformed by our network. First, our
521 approach gives a good estimate of CS occurrence without requiring a subsequent manual
522 selection of the cluster in a principal component space. Second, as compared to the PCA, the

523 UMAP dimensionality reduction technique is more resistant to changes in waveform shape,
524 such as reductions in waveform amplitude due to relative shifts in position between electrode
525 tips and cell bodies. Third, the performance of our algorithm is indifferent to occasional
526 oscillations that may occur in the LFP signal that may impede the performance of the PCA-
527 based approach, which relies on threshold crossings for event detection. Finally, as discussed
528 further below, the CNN, but not the PCA, offers precise information on timing, enabling us to
529 study CS durations much more systematically and objectively.

530 It is well established (Eccles et al., 1967) that each PC receives input from only one climbing
531 fiber. Therefore, it is very unlikely to find a second CS with completely different properties
532 in addition to the first CS in a PC record. Surprisingly, we found two PCs (see Fig. 9C for an
533 example) for which the CNN delineated a completely separate, large cluster of CSs in
534 addition to the main cluster. At first glance, this might have suggested a violation of the
535 aforementioned architectural principle. However, the CSs found in the respective second
536 clusters could be easily discarded post-hoc because of the insufficient suppression that they
537 induced in SS firing as compared to the genuine CSs. Therefore, although rare, even if
538 genuine CSs that belonged to a neighboring PC (Fig. 9C, seen as much smaller amplitude
539 waveforms in Cluster 2) were captured by the electrode tip, these CSs could easily be
540 identified based on their cluster IDs and scrutinized for selection.

541 To test whether our algorithm could really take over the burden of labeling CSs manually, we
542 made a one to one comparison of the performance of the CNN and the human expert on
543 records of 7 PCs for which all CSs had been labeled manually. Indeed, our algorithm's
544 performance matched the human-level expertise in detecting CSs in all PCs, except for one in
545 which additional CSs were detected by our algorithm (Fig. 7, Cell 7). The location of these
546 CSs in a distinct cluster in two dimensional feature space allowed the experimenter to easily

evaluate the validity of the identification of the waveform as CS and, in this case, to conclude that it was spurious.

Our algorithm detects start and end points of CSs with human-level performance

The prevailing idea of CSs serving as the "teaching-signal" for post-synaptic PCs (Marr, 1969; Albus, 1971; Ito, 1972), for which the occurrence of each CS event might be the only source of relevant information (Rushmer et al., 1976; Gellman et al., 1985), has been challenged by studies that demonstrated that the duration of action potential bursts fired by olivary neurons may vary and that this may be reflected by changes in the duration and the spikelet architecture of CSs (Llinás and Yarom, 1981; Ruigrok and Voogd, 1995; Maruta et al., 2007; Mathy et al., 2009; Bazzigaluppi et al., 2012; De Gruijl et al., 2012; Rasmussen et al., 2013; Zang et al., 2018). These observations have suggested that not only the occurrence of a CS, but also its duration may be relevant for motor learning. Addressing this possibility requires experimenters to invest even more time to manually label the start and end times of CS waveforms in addition to just detecting the events themselves. Not surprisingly, given the amount of time and effort involved, only a handful of attempts have been made to test this idea (Yang and Lisberger, 2014; Herzfeld et al., 2015, 2018; Junker et al., 2018) with inconsistent results. In order to achieve consensus, larger data sets collected under more diverse conditions would have to be explored, a necessity researchers have been reluctant to meet because of the hassles of the manual timing analysis. Since our CNN-based approach is able to effortlessly follow the performance of the human expert in detecting the start and end of the CS waveforms, by applying the expert's "mental rules" learned during training, quantifying task related changes in the architecture of CSs collected at different times in an experiment will become much more feasible in the future.

571 *Deep learning as a research tool*

572 More broadly, deep learning allows modeling non-linear relationships between input and
573 output for which no analytical solutions may exist. It is exactly this property of deep learning
574 that explains why this machine learning approach has recently emerged as a potentially
575 powerful research tool, which can tremendously reduce the workload of scientists (Ciregan et
576 al., 2012; Havaei et al., 2017; Oztel et al., 2017; Bellet et al., 2018). In light of recent
577 developments, in which deep learning has been successfully utilized to not only design
578 stimuli with controlled higher order statistics (Gatys et al., 2015), but also to model non-
579 linear relationships in neural data (Ecker et al., 2018), it is not hard to imagine that the full
580 potential of deep learning will significantly boost the pace of neuroscientific research in the
581 coming years. Certainly, in the case of cerebellar neurophysiology, we believe that our use of
582 deep learning to detect the rare, but relevant, CS events will allow much renewed
583 investigation of the contentious functional roles of these events in motor control and beyond.

584

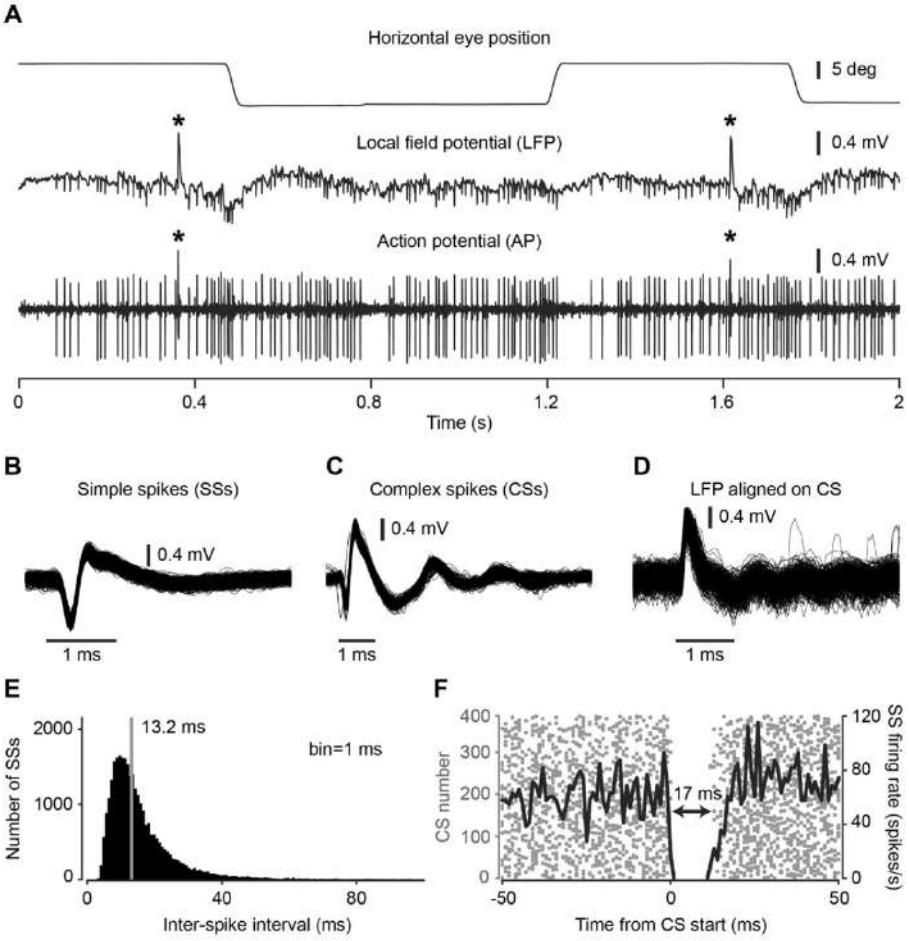
585 *Conclusion*

586 So far, all analysis involving CSs has been based on extremely laborious, manual, or semi-
587 automated methods lasting up to several weeks. This enormously slows down the pace of
588 developments in the field. On the other hand, our deep learning approach can reverse this
589 reality. For example, for a database like ours (160 PCs), our approach requires the human
590 expert to invest only 2-3 hours of CS labeling for training purposes and another 3-4 hours to
591 later verify the results. Given that it takes 3-4 hours to manually label all CSs found in
592 recordings of just one PC, this investment in time is negligible compared to the alternative of
593 manually labeling all recorded PCs. Moreover, our automated algorithm performs this task at
594 par with human experts, and it renders more systematic valuable information about the timing

595 and morphology of CS waveforms. The algorithm will be made available for use via an open
596 source implementation https://github.com/jobellet/detect_CS with provisions for retraining
597 the network to new users' own measurements. We strongly believe that the gains in time and
598 reliability that our tool offers may substantially facilitate the quest for a better understanding
599 of the roles of the still largely mysterious CSs.

600

601 **Figures**

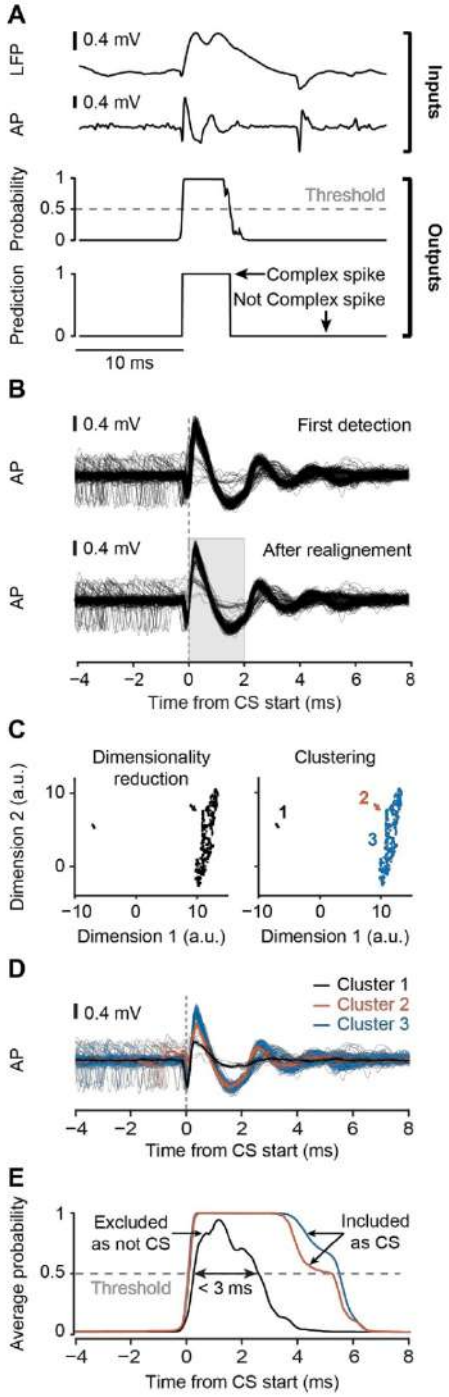


602

603 **Figure 1. Characteristics of an exemplary Purkinje cell.** (A) Local field potential (LFP,
604 low passed, <150 Hz, middle panel) and action potential (AP, high band-passed, 300 Hz - 3
605 KHz, bottom panel) activity in relation to horizontal eye movements (top panel). CSs are
606 marked by asterisks. (B) Isolated SS waveforms aligned on SS start. (C) Isolated CS
607 waveforms aligned on CS start. (D) LFP responses aligned to CS start. (E) Histogram of
608 inter-spike intervals of SSs. Solid gray line depicts the median value (13.2 ms). (F) Raster

609 plot showing a 17 ms pause in SS activity caused by the occurrence of a CS. Solid black line

610 represents the mean SS firing rate aligned to CS start.



611

612 **Figure 2. Pipeline for complex spike detection.** (A) Input to the network (LFP and action
613 potential signal, labels as AP) as well as its output (bin-wise predictive probability for CS
614 occurrence and binary CS classification). (B) Waveforms aligned to the first estimation of
615 start times of all CSs detected by the network (upper panel) used for computing an average
616 waveform that served as a template for realigning the waveforms of all detected CS events
617 (lower panel). (C) Projection of the waveforms during the time interval shaded in gray in B
618 onto a two-dimensional plane and identification of clusters in this space. Different colors
619 indicate distinct clusters. (D) Waveforms of the clusters in (C). Note that Cluster 1 clearly
620 violates well-known CS waveform shapes. (E) Average predictive probability output of the
621 network for the events in each cluster. Clusters, whose probability output exceeds the
622 classification threshold of 0.5 (dashed gray line) for less than 3 ms, are excluded as not
623 representing CSs (Cluster 1).

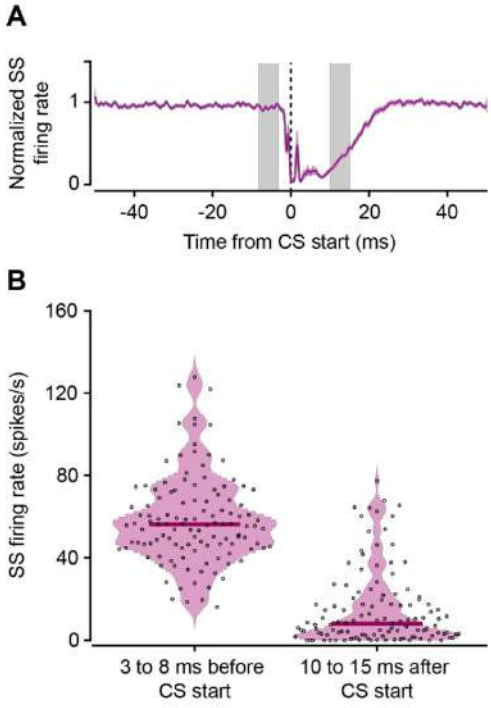
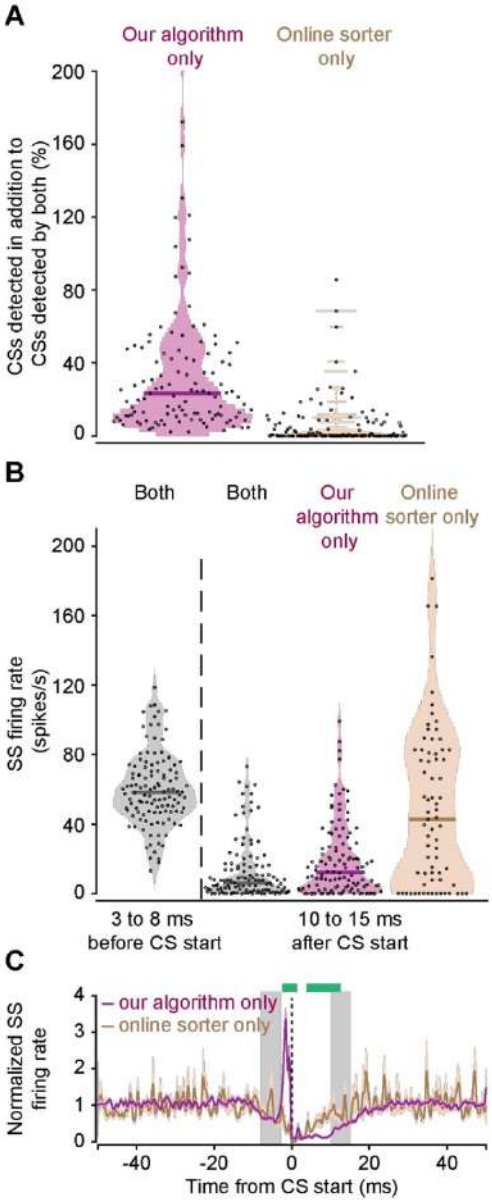


Figure 3. Decrease of SS rate after CSs. (A) Baseline-normalized mean SS firing rate aligned to the start of CSs detected by our algorithm. Data shows mean ± SEM over 119 PCs. Note that the small sharp peak in the SS response, seen immediately after CS start (vertical dashed line in black), is a result of the detection of initial large components of CSs in some PCs where these initial components resembled the shape of SSs and were most probably falsely detected as SSs by the online sorter. (B) Violin plots showing SS firing rate -8 to -3 ms before and 10 to 15 ms after CS start. Each dot represents the average SS firing rate aligned to start time of all CSs in one PC predicted by our algorithm. Thick lines indicate the median SS firing rate of all PCs.



636

637

638 **Figure 4. Comparison of CS detection by our algorithm and by the online sorter**

639 **application, MSD.** (A) Violin plots showing percentage of CSs detected exclusively by our

algorithm and the online sorter. 100% corresponds to the number of CSs detected by both methods. Our algorithm detected significantly more CSs than the MSD. (B) Violin plots showing SS firing rate aligned to the start of the CSs predicted by both algorithms (gray) or of the events additionally labeled as CSs by either our algorithm (pink) or the online sorter (beige). The decrease in SS firing after CSs predicted by our algorithm but not by the online sorter indicates a higher sensitivity of our algorithm. (A and B) Each dot represents the average SS firing rate aligned to all CSs for the recording of one neuron. Thick lines indicate the median. (C) Pause in averaged SS firing rate following a CS. Gray shaded region represents the period of 3-8 ms before and 10-15 ms after CS start used for comparing SS firing rates in panel B. The sharp increase in SS firing rate approximately 3 ms prior to CS start (vertical dashed line in black), observed only for CSs detected by our algorithm (pink), and not the MSD (beige), suggests that these SSs occurring shortly before the start of CSs might have altered their waveform. Only our algorithm was sensitive enough to detect such CSs with altered waveforms. Green bars on top show intervals with a significant difference between the two traces (random permutations cluster-corrected for multiple comparisons).

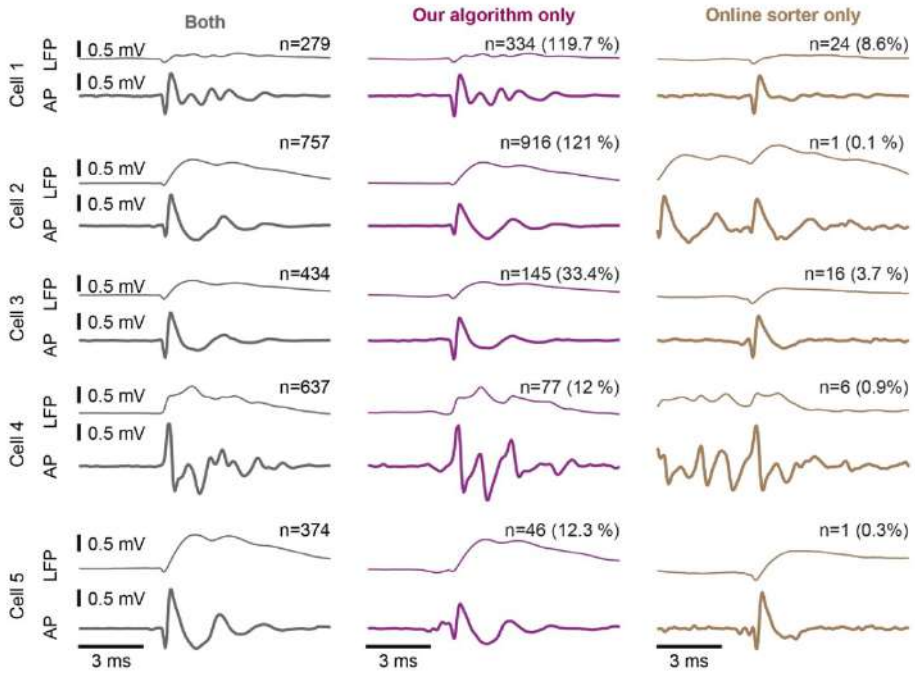


Figure 5. Waveforms of events labeled as CSs by our algorithm and the online sorter application MSD. Examples from seven neurons showing the average waveform in the LFP and action potentials of CSs detected by both methods (left), by our algorithm only (middle) or by the online sorter only (right).

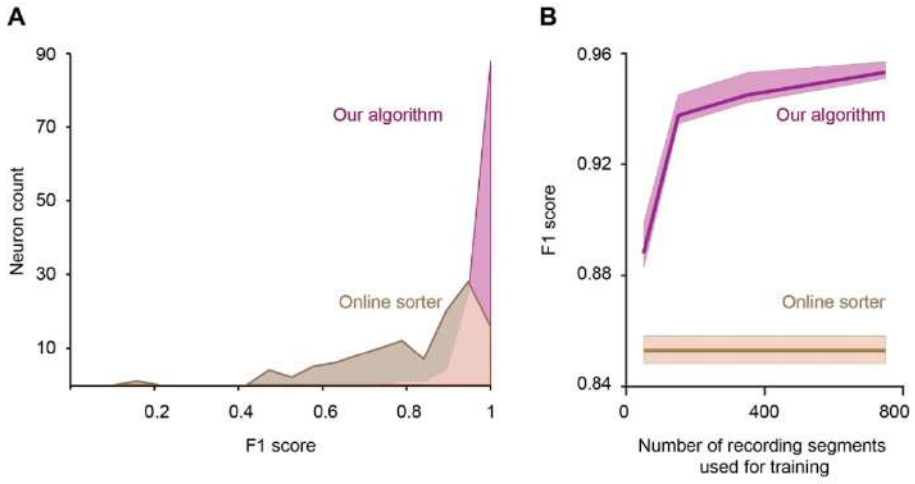
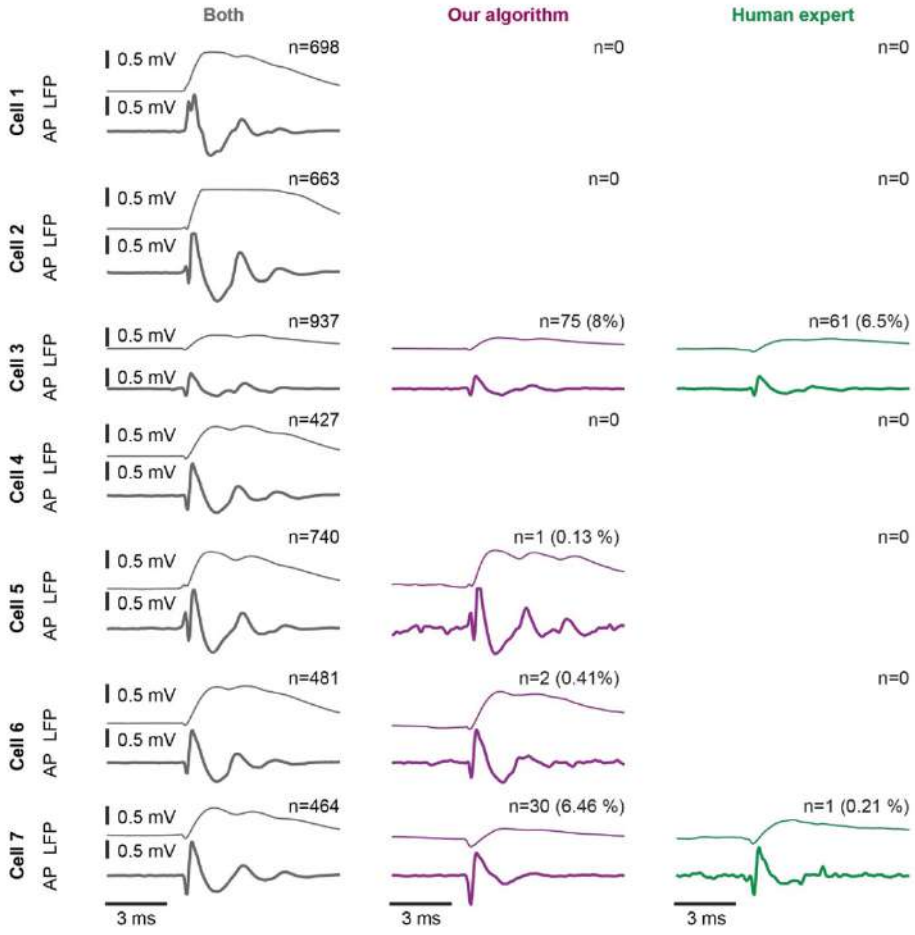


Figure 6. Classification agreement of our algorithm and the online sorter application MSD with a human expert. (A) Distribution of F1 scores of our algorithm and the online sorter computed by comparing CS labels with the human expert. Data from 119 neurons. (B) F1 score of our algorithm as a function of the number of recording segments used for training (pink) and F1 score achieved by the online sorter (beige). Think lines indicate the mean and the shaded area represents 95% confidence interval of the mean obtained by bootstrapping.



671

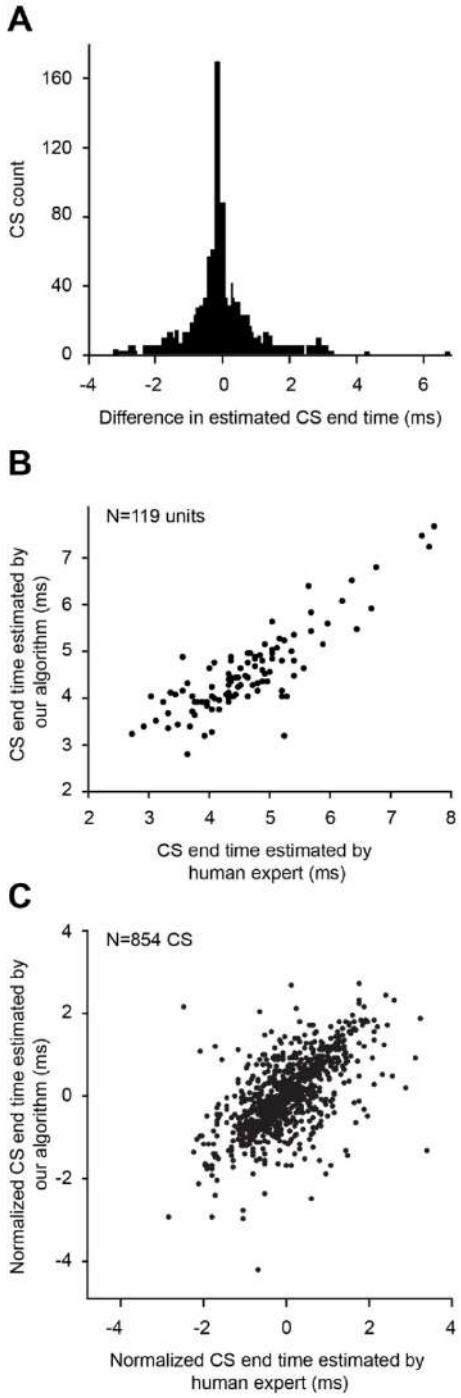
672

673 **Figure 7. Waveforms of events labeled as CSs by our algorithm and the human expert.**

674 Examples from seven neurons showing the average waveform in the LFP and action

675 potentials of CSs detected by both the human expert and our algorithm (left), by our

676 algorithm only (middle) or by the human expert only (right).



677

678 **Figure 8. Comparison of CS end times estimated by our algorithm and by the human**
679 **expert.** (A) Distribution of difference in CS end times labeled by our algorithm and by the
680 human expert. Data shows all CSs detected by both our algorithm and the human expert in
681 short recording segments from 119 neurons. (B) Correlation of CS end times estimated by
682 our algorithm (network) and the human expert. Each dot shows the average end time of all
683 CSs from one neuron. (C) Correlation of all CS end times pooled across the 119 neurons. The
684 end time of each CS was normalized by subtracting the average end time of the respective
685 neuron.
686
687
688
689
690
691
692
693
694
695

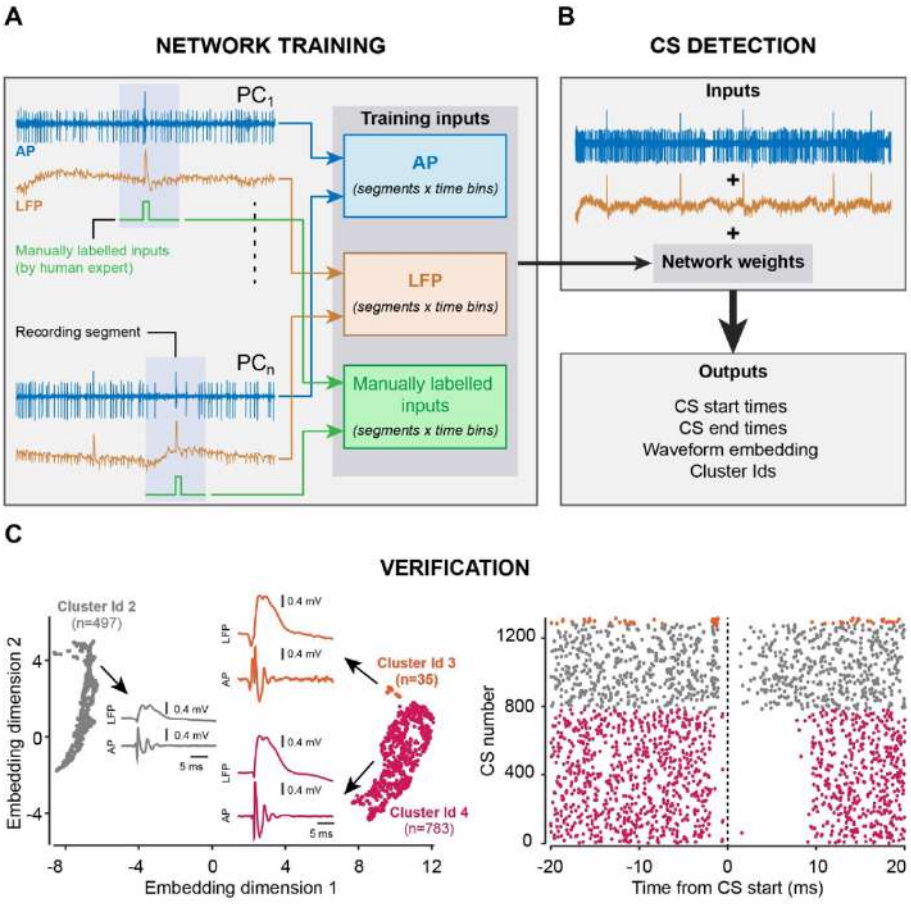


Figure 9. Workflow for using our algorithm. (A) The experimenter selects small segments of signal containing at least one CS each. Each segment is fed into the neural network in the form of three matrices containing the action potentials, the LFPs, and the labels separately. After training, the network outputs a set of weights. (B) The weights are used for evaluating new signals. (C) The output of the algorithm contains information about waveform shape that can be grouped in a dimensionality reduced space. This helps manual verifications, for example by inspecting the pause in SS firing rate after CS events in each cluster.

706 **References**
707
708 Albus JS (1971) A theory of cerebellar function. Mathematical Biosciences 10:25-61.
709 Bazzigaluppi P, De Gruijl JR, van der Giessen RS, Khosrovani S, De Zeeuw CI, de Jeu MTG
710 (2012) Olivary subthreshold oscillations and burst activity revisited. Frontiers in
711 neural circuits 6:91-91.
712 Bechert K, Koenig E (1996) A search coil system with automatic field stabilization,
713 calibration, and geometric processing for eye movement recording in humans.
714 Bell C, Grimm R (1969) Discharge properties of Purkinje cells recorded on single and double
715 microelectrodes. Journal of Neurophysiology 32:1044-1055.
716 Bellet ME, Bellet J, Nienborg H, Hafed ZM, Berens P (2018) Human-level saccade detection
717 performance using deep neural networks. Journal of neurophysiology.
718 Campello RJ, Moulavi D, Sander J (2013) Density-based clustering based on hierarchical
719 density estimates. In: Pacific-Asia conference on knowledge discovery and data
720 mining, pp 160-172: Springer.
721 Catz N, Dicke PW, Thier P (2005) Cerebellar complex spike firing is suitable to induce as
722 well as to stabilize motor learning. Current Biology 15:2179-2189.
723 Ciregan D, Meier U, Schmidhuber J (2012) Multi-column deep neural networks for image
724 classification. In: 2012 IEEE Conference on Computer Vision and Pattern
725 Recognition, pp 3642-3649.
726 Dash S, Catz N, Dicke PW, Thier P (2010) Specific vermal complex spike responses build
727 up during the course of smooth-pursuit adaptation, paralleling the decrease of
728 performance error. Experimental brain research 205:41-55.
729 Davie JT, Clark BA, Häusser M (2008) The origin of the complex spike in cerebellar Purkinje
730 cells. Journal of Neuroscience 28:7599-7609.
731 De Gruijl JR, Bazzigaluppi P, de Jeu MTG, De Zeeuw CI (2012) Climbing fiber burst size
732 and olivary sub-threshold oscillations in a network setting. PLoS computational
733 biology 8:e1002814-e1002814.
734 Dice LR (1945) Measures of the amount of ecologic association between species. Ecology
735 26:297-302.
736 Eccles JC, Ito M, Szentágothai J (1967) The cerebellum as a neuronal machine. Oxford,
737 England: Springer-Verlag.
738 Ecker A, H. Sinz F, Froudarakis E, Fahey P, Cadena S, Y. Walker E, Cobos E, Reimer J,
739 Tolias A, Bethge M (2018) A rotation-equivariant convolutional neural network model
740 of primary visual cortex.
741 Evans KK, Cohen MA, Tambouret R, Horowitz T, Kreindel E, Wolfe JM (2011) Does visual
742 expertise improve visual recognition memory? Attention, Perception, &
743 Psychophysics 73:30-35.
744 Fujita Y (1968) Activity of dendrites of single Purkinje cells and its relationship to so-called
745 inactivation response in rabbit cerebellum. Journal of Neurophysiology 31:131-141.
746 Gatys L, Ecker A, Bethge M (2015) A Neural Algorithm of Artistic Style.
747 Gellman R, Gibson AR, Houk JC (1985) Inferior olivary neurons in the awake cat: Detection
748 of contact and passive body displacement. Journal of Neurophysiology 54:40-60.
749 Havaei M, Davy A, Warde-Farley D, Biard A, Courville A, Bengio Y, Pal C, Jodoin P-M,
750 Larochelle H (2017) Brain tumor segmentation with Deep Neural Networks. Medical
751 Image Analysis 35:18-31.
752 Herzfeld DJ, Kojima Y, Soetedjo R, Shadmehr R (2015) Encoding of action by the Purkinje
753 cells of the cerebellum. Nature 526:439.
754 Herzfeld DJ, Kojima Y, Soetedjo R, Shadmehr R (2018) Encoding of error and learning to
755 correct that error by the Purkinje cells of the cerebellum. Nature neuroscience
756 21:736.
757 Ito M (1972) Neural design of the cerebellar motor control system. Brain research 40:81-84.
758 Judge SJ, Richmond BJ, Chu FC (1980) Implantation of magnetic search coils for
759 measurement of eye position: An improved method. Vision Research 20:535-538.

760 Junker M, Endres D, Sun ZP, Dicke PW, Giese M, Thier P (2018) Learning from the past: A
761 reverberation of past errors in the cerebellar climbing fiber signal. PLoS biology
762 16:e2004344.
763 Kitazawa S, Kimura T, Yin P-B (1998) Cerebellar complex spikes encode both destinations
764 and errors in arm movements. Nature 392:494.
765 Latham A, Paul D (1971) Spontaneous activity of cerebellar Purkinje cells and their
766 responses to impulses in climbing fibres. The Journal of Physiology 213:135-156.
767 Leznik E, Llinás R (2005) Role of gap junctions in synchronized neuronal oscillations in the
768 inferior olive. Journal of neurophysiology.
769 Llinas R (1974) Motor aspects of cerebellar control. Eighteenth Bowditch lecture.
770 Physiologist 17:19-46.
771 Llinas R, Sugimori M (1980) Electrophysiological properties of in vitro Purkinje cell dendrites
772 in mammalian cerebellar slices. The Journal of physiology 305:197-213.
773 Llinás R, Yarom Y (1981) Electrophysiology of mammalian inferior olivary neurones in vitro.
774 Different types of voltage-dependent ionic conductances. The Journal of Physiology
775 315:549-567.
776 Marr D (1969) A theory of cerebellar cortex. The Journal of physiology 202:437-470.
777 Maruta J, Hensbroek RA, Simpson JI (2007) Intraburst and Interburst Signaling by Climbing
778 Fibers. The Journal of Neuroscience 27:11263-11270.
779 Mathy A, Ho SSN, Davie JT, Duguid IC, Clark BA, Häusser M (2009) Encoding of
780 Oscillations by Axonal Bursts in Inferior Olive Neurons. Neuron 62:388-399.
781 McDevitt CJ, Ebner TJ, Bloedel JR (1982) The changes in Purkinje cell simple spike activity
782 following spontaneous climbing fiber inputs. Brain research 237:484-491.
783 McInnes L, Healy J, Melville J (2018) Umap: Uniform manifold approximation and projection
784 for dimension reduction. arXiv preprint arXiv:180203426.
785 Medina JF, Lisberger SG (2008) Links from complex spikes to local plasticity and motor
786 learning in the cerebellum of awake-behaving monkeys. Nature neuroscience
787 11:1185.
788 Ohmae S, Medina JF (2015) Climbing fibers encode a temporal-difference prediction error
789 during cerebellar learning in mice. Nature neuroscience 18:1798.
790 Oscarsson O (1980) Functional organization of olivary projection to the cerebellar anterior
791 lobe. The Inferior Olivary Nucleus:279-290.
792 Oztel I, Yolcu G, Ersoy I, White T, Bunyak F (2017) Mitochondria segmentation in electron
793 microscopy volumes using deep convolutional neural network. In: 2017 IEEE
794 International Conference on Bioinformatics and Biomedicine (BIBM), pp 1195-1200.
795 Prsa M, Dicke PW, Thier P (2010) The absence of eye muscle fatigue indicates that the
796 nervous system compensates for non-motor disturbances of oculomotor function.
797 Journal of Neuroscience 30:15834-15842.
798 Prsa M, Dash S, Catz N, Dicke PW, Thier P (2009) Characteristics of responses of Golgi
799 cells and mossy fibers to eye saccades and saccadic adaptation recorded from the
800 posterior vermis of the cerebellum. Journal of Neuroscience 29:250-262.
801 Rasmussen A, Jirenhed D-A, Zucca R, Johansson F, Svensson P, Hesslow G (2013)
802 Number of Spikes in Climbing Fibers Determines the Direction of Cerebellar
803 Learning. The Journal of Neuroscience 33:13436-13440.
804 Ronneberger O, Fischer P, Brox T (2015) U-net: Convolutional networks for biomedical
805 image segmentation. In: International Conference on Medical image computing and
806 computer-assisted intervention, pp 234-241: Springer.
807 Ruigrok TJH, Voogd J (1995) Cerebellar Influence on Olivary Excitability in the Cat.
808 European Journal of Neuroscience 7:679-693.
809 Rushmer DS, Roberts WJ, Augter GK (1976) Climbing fiber responses of cerebellar Purkinje
810 cells to passive movement of the cat forepaw. Brain Research 106:1-20.
811 Servais L, Bearzatto B, Hourez R, Dan B, Schiffmann SN, Cheron G (2004) Effect of simple
812 spike firing mode on complex spike firing rate and waveform in cerebellar Purkinje
813 cells in non-anesthetized mice. Neuroscience letters 367:171-176.

814 Sørensen T (1948) A method of establishing groups of equal amplitude in plant sociology
815 based on similarity of species and its application to analyses of the vegetation on
816 Danish commons. Biol Skr 5:1-34.
817 Streng ML, Popa LS, Ebner TJ (2017) Climbing fibers control Purkinje cell representations of
818 behavior. Journal of Neuroscience 37:1997-2009.
819 Stuart G, Häusser M (1994) Initiation and spread of sodium action potentials in cerebellar
820 Purkinje cells. Neuron 13:703-712.
821 Thach W (1968) Discharge of Purkinje and cerebellar nuclear neurons during rapidly
822 alternating arm movements in the monkey. Journal of neurophysiology 31:785-797.
823 Warnaar P, Couto J, Negrello M, Junker M, Smilgin A, Ignashchenkova A, Giugliano M,
824 Thier P, De Schutter E (2015) Duration of Purkinje cell complex spikes increases with
825 their firing frequency. Frontiers in cellular neuroscience 9:122.
826 Wolfe JM, Horowitz TS, Kenner NM (2005) Rare items often missed in visual searches.
827 Nature 435:439-440.
828 Wörgötter F, Daunicht WJ, Eckmiller R (1986) An on-line spike form discriminator for
829 extracellular recordings based on an analog correlation technique. Journal of
830 neuroscience methods 17:141-151.
831 Yang Y, Lisberger SG (2014) Purkinje-cell plasticity and cerebellar motor learning are
832 graded by complex-spike duration. Nature 510:529.
833 Zang Y, Dieudonné S, De Schutter E (2018) Voltage-and Branch-Specific Climbing Fiber
834 Responses in Purkinje Cells. Cell reports 24:1536-1549.
835 Zur G, Joshua M (2019) Using extracellular low frequency signals to improve the spike
836 sorting of cerebellar complex spikes. bioRxiv:556985.
837