Journal of the Neurological Sciences 463 (2024) 123089

Contents lists available at ScienceDirect

Journal of the Neurological Sciences

journal homepage: www.elsevier.com/locate/jns

Deep learning of Parkinson's movement from video, without human-defined measures

Jiacheng Yang [a], Stefan Williams [b,c,*], David C. Hogg [a], Jane E. Alty [c,d], Samuel D. Relton [b]

[a] School of Computing, University of Leeds, UK
[b] Leeds Institute of Health Sciences, University of Leeds, UK
[c] Leeds Teaching Hospitals NHS Trust, UK
[d] Wicking Dementia Research and Education Centre, University of Tasmania, Australia

ARTICLE INFO

Keywords:
Parkinson's disease
Bradykinesia
Computer vision
Video
Deep learning
Artificial intelligence

ABSTRACT

Background: The core clinical sign of Parkinson's disease (PD) is bradykinesia, for which a standard test is finger tapping: the clinician observes a person repetitively tap finger and thumb together. That requires an expert eye, a scarce resource, and even experts show variability and inaccuracy. Existing applications of technology to finger tapping reduce the tapping signal to one-dimensional measures, with researcher-defined features derived from those measures.
Objectives: (1) To apply a deep learning neural network directly to video of finger tapping, without human-defined measures/features, and determine classification accuracy for idiopathic PD versus controls. (2) To visualise the features learned by the model.
Methods: 152 smartphone videos of 10s finger tapping were collected from 40 people with PD and 37 controls. We down-sampled pixel dimensions and videos were split into 1 s clips. A 3D convolutional neural network was trained on these clips.
Results: For discriminating PD from controls, our model showed training accuracy 0.91, and test accuracy 0.69, with test precision 0.73, test recall 0.76 and test AUROC 0.76. We also report class activation maps for the five most predictive features. These show the spatial and temporal sections of video upon which the network focuses attention to make a prediction, including an apparent dropping thumb movement distinct for the PD group.
Conclusions: A deep learning neural network can be applied directly to standard video of finger tapping, to distinguish PD from controls, without a requirement to extract a one-dimensional signal from the video, or pre-define tapping features.

1. Introduction

Neurological disorders are the leading cause of disability in the world, and the fastest growing of these is Parkinson's disease, with numbers predicted to rise from 6.9 million in 2015 to 14.2 million in 2040 [1]. It is notable that Parkinson's remains a clinical diagnosis that relies largely on visual judgements made by the eye of an experienced physician. Indeed, in 1817, three of the six cases originally described by James Parkinson were people he simply observed in the street [2], and in modern diagnostic criteria the cardinal motor feature of the condition is a visual sign: bradykinesia [3]. This term refers to *slowness of movement and decrement in amplitude or speed (or progressive hesitations / halts) as movement continues* [3]. A common clinical test for bradykinesia is finger tapping, in which the clinician observes the patient repeatedly tapping their index finger and thumb together 'as wide and as quick as possible' (for 10 taps [4] or 10 s [5]), and judges whether the tap speed, amplitude and/or rhythm are impaired [3–5].

Reliance on clinician visual judgement has two limitations. First, it requires an experienced specialist clinician, a resource in short supply globally [6]. Second, visual judgements are inherently subjective, and, even experts show considerable inter-rater variability for visual signs [5,7–10]. This has the potential to contribute to delayed or inaccurate diagnosis, impaired assessment of change over time, and imprecise measurement of treatment outcomes.

In response, there have been numerous attempts to use technology to objectively quantify the finger tapping test for Parkinson's bradykinesia.

* Corresponding author at: Department of Neurology, Leeds Teaching Hospitals NHS Trust, Leeds General Infirmary, Great George Street, Leeds LS1 3EX, UK.
E-mail address: stefanwilliams@doctors.org.uk (S. Williams).

https://doi.org/10.1016/j.jns.2024.123089
Received 20 February 2023; Received in revised form 5 June 2024; Accepted 5 June 2024
Available online 10 June 2024

Two common approaches have been to measure finger tapping with 'wearable' sensors (such as gyroscope and infrared devices) [11–30], or by tapping on a smartphone screen [31–36]. The first of these requires specialist equipment, that has never become widely available. The second cannot record tapping amplitude, and relies on patient motivation to interact with a specific application.

More recently, computer vision techniques have been applied to standard video in attempts to automate the finger tapping test [37–47]. Unlike wearable sensors, video information is much closer to what an expert clinician sees during examination. However, previous publications discard most of this information at the outset, following the same basic approach used in wearable and screen-tapping studies. That is to first convert the tapping signal into a one-dimensional (1D) time series, such as finger-to-thumb distance over time. The resultant 1D time series is usually then processed further to derive specific measurements, such as opening velocity and coefficient of variation, which researchers choose to be representative of elements in the clinical definition of bradykinesia. Such measurements by themselves, or in combination, cannot discriminate the hand of a person with Parkinson's (PwP) from that of a control participants, because the distributions of scores in PwP and controls overlap considerably [11–14,23,24,29,30,32,34,35]. Machine learning, including deep learning, has been applied to finger tapping measures from video [37–48]. However, it is important to note that these studies apply machine learning only *after* the signal has been first reduced and constrained to researcher-defined measures from a 1D signal.

It is possible that there are movement patterns present during finger tapping in Parkinson's that are outside the current definition of bradykinesia. Such patterns could be invisible to clinicians, or only recognisable to clinicians in an automatic, unconscious way. Deep learning using neural networks is a machine learning technique that does not require human-defined features, such as finger to thumb distance or opening velocity; instead, it can learn the features most predictive of a given category, without predefined rules. This provides a way to look for patterns that are characteristic of Parkinson's in video, without restricting patterns to one-dimensional, human-defined measures.

To the best of our knowledge, only one previous paper has applied a deep learning neural network directly to 2D video data [49]. However, this was to disintiguish two grades of bradykinesia (0 and 1, rather than the 5 grades in the standard rating scale [50]). The study did not feature control participants, and did not apply deep learning to the differences between controls and PwP.

In this paper we describe and test a new approach to computer vision of video of the finger tapping test. Specifically, we apply a deep learning neural network *directly* to video of finger tapping. We test the ability of this network to discriminate Parkinson's finger tapping from that of control participants. In addition, we use class activation maps to visualise the most important spatial and temporal regions of the tapping video that the network learns to 'look at' to distinguish Parkinson's from control.

2. Methods

2.1. Ethical review

The study was approved by the United Kingdom Health Research Authority, North of Scotland Research Ethics Committee (IRAS project ID 256116). Informed, written consent was given by all participants.

2.2. Participants and video recording

A total of 152 videos of finger tapping were collected from 77 participants, consisting of 40 people with idiopathic Parkinson's (PwP) and 37 healthy controls. The left and right hands of each participant were recorded separately, with two videos rejected because the hand moved outside the camera frame.

All PwP had previously been diagnosed by a movement disorder specialist neurologist at Leeds Teaching Hospitals NHS Trust United Kingdom, according to Movement Disorder Society clinical diagnostic criteria [3]. At the time of video recording, they were subjectively and objectively in the 'on' state, following their regular medication regime. PwP with postural hand tremor were not excluded. Controls were recruited from the spouses, partners or friends of the PwP, as well as hospital or university staff. The control participants had no history of Parkinson's or other neurological diagnosis.

Participants rested their elbow on a chair arm with their forearm and hand free to move as per the Item 3.4 Finger Tapping protocol of the Movement Disorder Society revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS) [51]. Only the hand and forearm were within the video frame. The distance between camera and hand was approximately 1 m, but not tightly defined. Digits 1 and 2 were closest to the camera. No specific instructions were given for flexion or extension of digits 3 to 5, and participants were free to position them as they preferred, although the researcher gave a brief demonstration in which they were extended.

A smartphone, placed on a tripod lateral to the participant, was used to record standard video with only ambient lighting (1920 × 1080 pixel resolution, 60 frames per second). Participants were asked to tap index finger and thumb together "as quickly and as big as possible" [51] for at least 10 s. Each video was edited to an 11 s clip: 1 s prior to tapping onset and 10 s of finger tapping (the duration used in the MBRS clinical rating scale [52]).

2.3. Pre-processing

The high spatial resolution of the videos and the large proportion of background and forearm captured in the videos would reduce the efficiency of the neural network learning process. Therefore, each frame of the video was assigned a bounding box for the hand using hand tracking software [53]. An overall box – which encapsulates all the frame-specific boxes – was used to crop the video. To have uniform video size, the cropped video was down-sampled to 80 × 60 pixels with Gaussian smoothing and bicubic interpolation (retaining 60 frames per second) using the OpenCV library [54], before conversion to greyscale.

The videos were split into 1 s clips, overlapping one another by 0.5 s, so clips were from 0 to 1 s, 0.5-1.5 s, 1-2 s, etc. This procedure resulted in a total of 3148 clips (1637 patients and 1511 controls), each 60 frames long with 80 × 60 pixels. These clips provided the visual data to train and evaluate the deep learning model.

Prior to data being passed to the neural network, the training data was mean-centred and divided by the overall standard deviation of pixel values (i.e. Z-score normalization). The test data were normalized using the mean and standard deviation of the training data.

2.4. Model design

Since the videos are stored as a 3-tensor with dimensions 60 × 80 × 60, our model is a neural network based primarily on 3D convolutional filters. In particular, we have three layers of convolutions and max-pooling, with batch normalization [55] between the convolutions and max-pooling operations to improve the convergence rate. A final batch normalization occurs after the third max-pooling before a fully connected layer with a ReLU nonlinearity and the output layer with a sigmoid nonlinearity. Parameters in the 3D convolutions and the dense layers had L2 (i.e. Ridge) regularization to prevent overfitting, in addition to a dropout layer (10% dropout) used between the fully connected and output layer. In total this model has 20 million learnable parameters. The network architecture is described graphically in Supplementary Fig. S1.

The model was built using the Keras framework [56] and trained using the Adam optimizer [57] applied to the binary cross-entropy loss function. Default momentum parameters were used for the optimization. The learning rate, number of epochs, and regularization strength were chosen using a grid search combined with 6-fold cross-validation to maximize the F1-score. The learning rate was selected from ten log-spaced values between 0.1 and 0.0001, the epochs from [100, 200, 400] and regularization strength from [0, 0.001, 0.01, 0.05, 0.1].

The model was trained and tested in Keras on an NVIDIA P100 GPU, provided by the Leeds Advanced Research Computing facility.

2.5. Performance of the deep learning neural network

The classification accuracy of the model to discriminate PwP from controls was assessed using five metrics: accuracy, precision, recall, F1-score and Area Under the Receiver Operating Characteristic curve (AUROC).

We used cross-validation to obtain realistic performance estimates. Examples were split into 6 random folds containing all the clips from a given participant. The model, including the normalization of training and test data, was then trained on 4 of these folds, with one used as the validation set and one for testing; repeating this procedure for each combination of folds [58]. This approach ensured that no participant is in the training and test data simultaneously. The mean and standard deviation of performance over the 6 testing folds is reported.

2.6. Visualisation of the learned features

In addition, we examined the features that the deep learning model derived to investigate how it learned to differentiate between the patient and control groups.

We plotted the value of the activation function for selected features of each 3D convolutional layer to examine what the network "sees" at each stage. This visualisation was done for a single PwP and control participant over a range of video frames. To select the features from the final convolutional layer that best distinguish between patient and control, we determined the distribution of the maximal activation for each feature on all patient and control videos in the training set, using a violin plot. That is, for each video we ran the model and captured the maximal activation of each feature on the third convolutional layer. The features that exhibited the largest difference between patients and controls were further investigated as follows.

First saliency mapping, a basic form of attention mechanism, was applied – showing the pixels that the model focuses on to make its classification. Class Activation Maps [59] show the attention of the final convolutional layers over the input, making use of the derivative of the output classification with respect to the input image. Pixels with steeper gradients have more influence over the final decision and therefore are intuitively similar to a notion of attention. This form of attention was applied to a PwP and control video with results plotted at frames 0, 8, and 16, enabling an illustrative view of where each feature is "looking" during the clip.

We further plotted the timeseries of activation values for these features over the second video to determine when the feature is activated. The combination of where and when the features activate enabled us to describe their function.

3. Results

3.1. Participants

40 people with Parkinson's (mean age 68 years, SD 10) and 37 controls (mean age 56 years, SD 19) participated, with a total of 79 Parkinson's hand videos and 73 control hand videos. The median number of years since diagnosis was 4, and the median Modified Hoehn and Yahr Scale was 2 (Table 1).

Table 1
Participant characteristics.

	Parkinson's	Controls
Number of participants	40	37
Hands	79	73
Mean age [SD], years	68 [10]	56 [19]
Male:Female	27:13	13:24
Median years since diagnosis	4	N/A
Median H&Y [IQR]	2 [1, 3]	N/A
H&Y = 1	17	
H&Y = 1.5	1	
H&Y = 2	7	
H&Y = 2.5	2	
H&Y = 3	11	
H&Y = 4	2	
H&Y = 5	0	

Parkinson's: Idiopathic Parkinson's disease. H&Y: Modified Hoehn and Yahr Scale [61]. IQR: Interquartile Range.

3.2. Performance of the neural network to discriminate Parkinson's from controls

Using our cross-validation procedure the model obtained excellent performance on the training data, with some evidence of overfitting (likely due to the limited amount of data available). The test accuracy was 0.69, with test precision 0.73, test recall 0.76, and test AUROC 0.76 (Table 2). This model had a learning rate of 7.2e-4, regularization strength of 0.1, and ran for 400 epochs.

3.3. Visualisation of the learned features

To investigate how the final model learned to differentiate between PwP and controls, three approaches were used.

The first approach is to plot the activation function values for selected convolutional filters on each of the three 3D convolutional neural network layers, shown in Fig. 1. Two filters are chosen at each layer to illustrate the diversity of learned features and are plotted for a PwP and control participant at frames 0, 5, and 10 of the clip. Supplementary Video 1 shows an example of three features in third layer of the network for a PwP and control.

The features selected from the first layer show a movement detector to separate the hand from the background and an edge detector that focuses on the upper edges of the arm, thumb, and fingers. From layer two we show a feature that appears to focus on movement of the forearm and one that follows movement in the lower edge of the image (forearm and thumb). The features selected from the third layer include one which follows the upper edge of the hand and thumb, and one which appears to detect stationary patches of skin: the fingers are visible from the control hand but not in the patient hand, though the forearm is visible in both.

Table 2
Performance of the neural network.

Metric	Training data (SD)	Test data (SD)
Binary cross-entropy loss	0.35 (6e-3)	1.0 (6e-3)
Accuracy	0.91 (5e-3)	0.69 (1e-1)
Precision	0.90 (7e-3)	0.73 (2e-1)
Recall	0.89 (2e-2)	0.76 (1e-1)
F1-score	0.90 (1e-2)	0.71 (1e-1)
AUROC	0.97 (4e-3)	0.76 (8e-2)

Binary cross-entropy loss is the function used to penalize model inaccuracy during the training process. Accuracy is the proportion of correct predictions (true positives and true negatives divided by all predictions). Precision is equivalent to specificity (true positives divided by true positives plus false positives). Recall is equivalent to sensitivity (true positives divided by true positives plus false negatives). F1-score is the harmonic mean of precision and recall. AUROC: Area under the receiver operating characteristic curve.

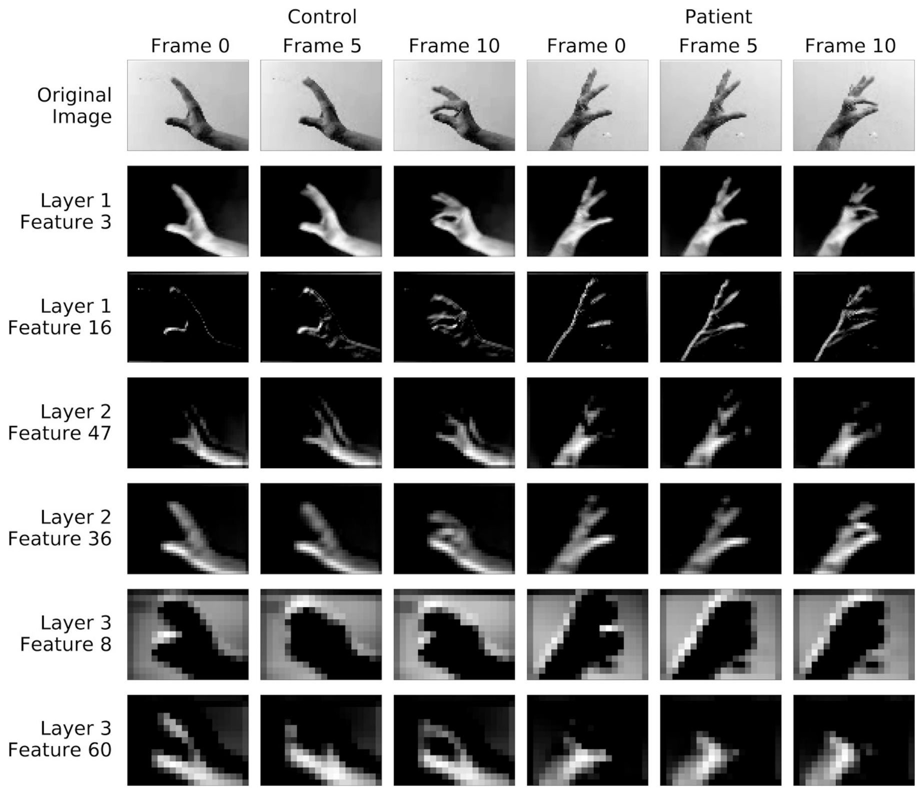

Fig. 1. Activation of selected features on the three convolutional layers for three frames of finger tapping video from a PwP ('patient') and control participant. Features that the network has used to learn to discriminate Parkinson's and controls are represented by brighter pixels.

To investigate which features from the final convolutional layer best differentiate between PwP and control, Supplementary Fig. S2 shows a violin plot of maximal feature activation over the entire training set. Features 17, 19, 12, 38, and 32 provided the strongest differentiation (largest difference in mean value between the PwP and control participants).

In Fig. 2, we show class activation maps (saliency maps) for the five features that provide the strongest differentiation between patient and control participants. Frames 0, 8, and 16 are shown for two videos from the training set. These maps allow us to see the areas where the network is focusing its attention when making a prediction.

In Fig. 3 we plot the maximal activation of each feature over the duration of the video clip. Solid lines correspond to the patient whilst dotted lines are for the control participant. The class activation maps show *where* each feature is focused (Fig. 2), whilst the activation pattern over time shows *when* it activates (Fig. 3). Supplementary Video 2 provides an animated version of Fig. 2, showing the features over time (the attention over the tapping cycle).

Combining these two figures allows us to determine the function of each feature as follows: (Feature 17) thumb movement (an apparent dropping movement in the patient group), (Feature 19) the number of opening and closing cycles, (Feature 12) the section of the image with the fastest movement, (Feature 38) the connection between thumb and wrist, and (Feature 32) a smoothed version of overall hand movement.

4. Discussion

Our results show that a deep learning neural network can be directly applied to smartphone video of finger tapping, without predefined rules or measures, to distinguish people with Parkinson's from control participants. Performance of this neural network was good, with test accuracy 0.69, precision 0.73, recall 0.76, and AUROC 0.76. In addition, we have demonstrated a visual representation of the network's attention within the video, showing what it 'looks at' in the video to classify as Parkinson's or control.

These findings suggest a potential method to assist or augment the discrimination of Parkinson's from controls that does not require special equipment or patient motivation. The products of deep learning models can be run on standard smartphones (which are ubiquitous), and observation of the finger tapping test is already a standard part of routine clinical care. This contrasts with specialist apps that require a patient to tap the screen or keyboard, for example [60]. A video-based method is contactless, compatible with infection control and remote video consultations. Only a ten second clip was recorded. Furthermore, our novel method is not constrained by an attempt to predefine and therefore restrict the parameters for movement patterns prior to machine learning. We did not begin with a one-dimensional time series such as finger-thumb distance, or tell the computer which features (such as speed or rhythm) to look for in finger tapping, or that people with Parkinson's have something called bradykinesia and how that is defined. This provides the opportunity to find new discriminating features of finger tapping movement that push beyond current definitions of bradykinesia, and may not have previously been considered or investigated. Could there be other novel features of finger tapping that we don't yet recognise ourselves? Or, that we unconsciously recognise as typical of Parkinson's without consciously naming them as a feature of bradykinesia?

The neural network, without any prior knowledge, found some features consistent with what we know about bradykinesia. Feature 19 and feature 12 appeared to be related to the speed of finger tapping. This suggests that the network is learning clinically valid features of tapping. However, in addition, one of the most discriminating features identified by the network is a movement feature not previously described. This was feature 17 – the mirrored movement of finger and thumb in Parkinson's tapping. In the Parkinson's group, the thumb and finger both moved apart from each other and then back towards each other, whereas in controls the thumb made an initial extension movement and then remained static so that it was the finger tapping *against* the thumb,

Fig. 2. Class Activation Maps for selected features on the third convolutional layer in three example video frames from a control (left) and patient (right). The brighter, yellow/green regions show where the network is focusing its attention when making a prediction. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

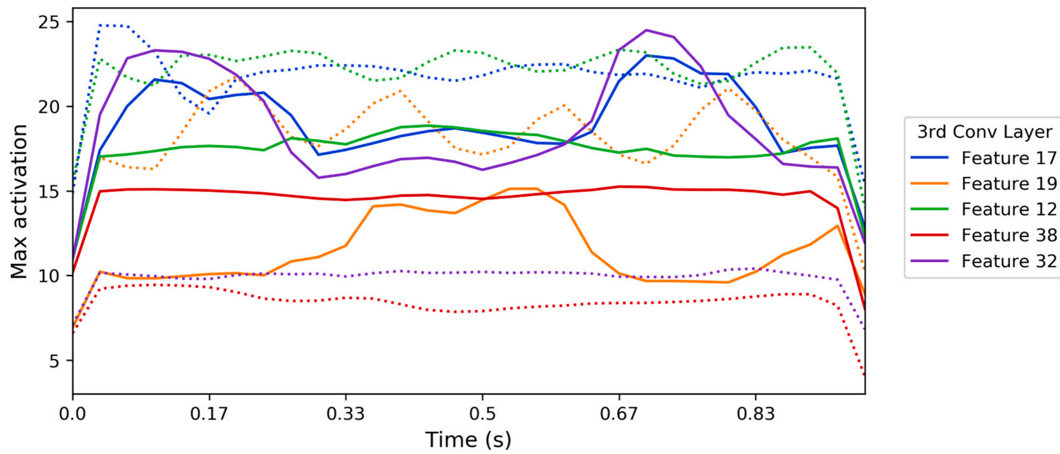

Fig. 3. Maximal activation of the identified features over a 1 s video segment for a PwP (solid lines) and control participant (dotted lines). Class Activation Maps for the same features, participants and 1 s segment are shown in Fig. 2 and Supplementary Video 2.

rather than *with* the thumb. It is unclear whether this thumb movement pattern would help discriminate PwP outside our specific participant group. We are not aware of clinicians previously noticing such a phenomenon.

The results have limitations. The number of videos is small for a deep learning study, even when considering the division into 1 s segments to make 3148 clips. The large number of parameters of the network in comparison to the amount of data available amounts to a significant risk of overfitting. This refers to the tendency of a machine learning technique to learn differences between two categories that are specific to the dataset it is trained on, rather than inherent to the categories in general. The drop in performance figures from training to test data is evidence of a degree of overfitting, meaning that the likely performance in a wider population is unclear at this stage.

Nevertheless, the test data performance results (i.e. the lower figures in Table 2) are similar to, or better than, many other approaches based on hand-coded features mentioned in the introduction [13,15,16,22,28,30,34,37,39,61]. Furthermore, in a different study with a subset of the same participant population from the current study, 21 human movement disorder expert clinician raters showed only moderate agreement for finger tapping MDS-UPDRS, with intraclass correlation coefficient of 0.53, and 24% of control participant videos were rated as bradykinesia by MBRS subscores [10]. When asked to guess Parkinson's or control status in the videos, the clinician raters were correct in 70% of videos - a strikingly similar figure to the test accuracy of our model [10]. That study suggests discriminating PwP from controls in finger tapping videos is not an easy task.

The videos were split into 1 s segments before processing by the neural network, so that any patterns with a longer duration would not be learned. This would largely exclude the phenomenon of ‘decrement’: progressive decrease in amplitude or speed as tapping continues [62]. However, our results suggest that despite this there are movement features within 1 s time windows that differ between PwP and controls. Furthermore, all attempts to quantify decrement in the literature show overlap between PwP and control groups [12,20,23,63], often with no significant group difference [13,20,21,23,24], suggesting that decrement may not always be an important distinguishing feature.

It is important to consider data storage with any video-based technology; currently we need to store videos for training models, but in actual deployment once the models have been developed, it is possible to do the analysis on-device where videos can be deleted immediately afterwards.

Another limitation of our study is that the only video data involved was finger tapping. Human diagnosis of Parkinson's is based on a much more comprehensive examination, together with details of the history of symptoms [3]. It is important to note that the current technology is insufficient to make a diagnosis, as that still relies on clinical assessment. However, there is potential for quantitative assessment of the movement to aid that clinical acumen – as an adjunct to diagnosis - and there is growing interest in technologies to aid earlier detection of neurodegenerative disorders [64,65]. An additional limitation is that the Parkinson's group mean age was higher than controls, and it is possible that some of the difference in tapping patterns relate to changes with aging rather than Parkinson's. The male:female ratio differed between the groups, another possible confounder. We did not have a large enough sample size to allow meaningful analysis of left and right hands in combination. Our participants did not have a broad range of skin colour and tone, limiting generalisability of the method to some extent. We did not record the proportion of hands with postural tremor, also a possible confounder. However, we would expect that only a minority of PwP had postural tremor, because a smaller subset of the same videos used in a previous study involved only 11 of 68 videos showing postural tremor [38]. Our visual representations of the network's attention during finger tapping do not suggest that it is using tremor movement to discriminate PwP from controls.

There are several previous computer vision reports involving standard video of finger tapping in Parkinson's. These have started with specific, human-defined measurements to be extracted from the video, followed by analysis of researcher-chosen features from that one-dimensional signal. A 2014 study used index finger tip coordinates [37] and a support vector machine (SVM) method to combine tapping features chosen by the researchers. They reported an accuracy of 95%, AUROC 90%, for discrimination between healthy controls and PwP, but participant numbers were very small (13 PwP, 6 controls), suggesting overfitting. Another study used a measure of overall hand movement (pixel optical flow) and machine learning such including SVM, to distinguish Parkinson's from controls, with 0.63 test accuracy [39]. A third study converted four video measures of upper limb parkinsonism into a 1D time series of frequency, and with SVM classification, they reported precision accuracy of 0.82 for discrimination of PwP from controls [46]. This is higher than our precision figure of 0.73, perhaps because other upper limb tests were included and their PwP cohort had longer mean disease durations of (8.7 and 6.4 years, compared with our median of 4 years). A 2021 paper reduced webcam video data to a 1D time series of finger to thumb distance. Researcher-selected features with several machine learning techniques to distinguish PwP from control showed sensitivity and specificity ranging from 41% to 73% and 23% to 81%, respectively, depending on the feature [47].

These methods are fundamentally different from ours – they do not apply deep learning to video, instead using simpler machine learning techniques (not deep learning) on a 1D time series of a researcher-selected feature. One published method is similar to our approach. Yin et al applied a deep learning neural network directly to 2D video data [49]. However, this was to disintiguish two grades of bradykinesia (0 and 1, rather than the 5 grades in the standard rating scale) [50]. The study did not feature control participants, and did not apply deep learning to the differences between controls and PwP. So whilst deep learning was applied directly to video, it was for a fundamentally different classification task. Nor did their study provide a visual representation of the network's attention within the video.

Six other video-based studies measured relative finger-thumb distance during tapping, to grade bradykinesia severity and correlate that with clinician ratings, but they did not attempt machine learning classification of videos into PwP versus controls [40–42,44,45,66].

There is a body of literature reporting the measurement of finger tapping in Parkinson's using wearable equipment such as gyroscopes [5], electromagnetic sensors [28,67], and infrared camera markers [22], or tapping on a smartphone screen [32]. Single measures recorded with such devices, such as tapping frequency, opening velocity or the coefficient of variation of tapping peaks, show overlapping patient and control group scores that cannot alone discriminate Parkinson's from health [11–14,23,24,29,30,32,34,35]. This is despite many studies using a protocol in which medication is withheld prior to some or all of the recordings, artificially exaggerating parkinsonism [5,20,21,23,27,29,61]. When single measures are combined with machine learning techniques, the results are comparable with ours. For example, the application of evolutionary algorithms to data from electromagnetic sensors measuring finger to thumb distance could discriminate early-stage Parkinson's from normal controls, with area under the receiver operating characteristic curve of 0.899 [68]. Such approaches require specific equipment, limiting widespread use.

There are several potential future extensions to the work described here: (1) collection of a large dataset to validate and refine the neural network, (2) application of the same method to other clinical signs, and (3) the use of traditional statistical methods to test sensitivity and specificity of tapping features identified by the model, such as feature 17, thumb movement.

In conclusion, a deep learning neural network can be applied directly to standard video of finger tapping, to distinguish Parkinson's from controls, without a requirement to extract a one-dimensional measures or pre-define tapping features. The location and timing of what the neural network's 'sees' in the video, to learn to distinguish Parkinson's tapping from controls, can be visualised.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jns.2024.123089.

Financial disclosures (for the previous 12 months)

Jiacheng Yang has no disclosures to report. Stefan Williams received funding from the National Institute for Health Research. David C Hogg received grant funding from EPSRC, Innovate UK, and the EU; he is director of a doctoral training centre with UKRI funding. Jane E Alty received: grant funding from the National Health and Medical Research Council, and the National Institute for Health Research; royalties from Taylor and Francis Publishing; stock ownership Clearsky Medical Diagnostics Ltd.; Honoraria from Stada and Ipsen. Samuel D Relton received funding from National Institute for Health Research, Harry Summer Kidney Fund and Heart Research UK.

CRediT authorship contribution statement

Jiacheng Yang: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Writing – review & editing. **Stefan Williams:** Conceptualization, Data curation, Investigation, Methodology, Writing – original draft, Writing – review & editing. **David C. Hogg:** Conceptualization, Formal analysis, Investigation, Methodology, Resources, Supervision, Writing – review & editing. **Jane E. Alty:** Conceptualization, Investigation, Methodology, Supervision, Writing – original draft, Writing – review & editing. **Samuel D. Relton:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Supervision, Visualization, Writing – original draft, Writing – review & editing.

Declaration of competing interest

No specific funding was received for this work. The authors declare that there are no conflicts of interest relevant to this work.

References

[1] E.R. Dorsey, B.R. Bloem, The Parkinson pandemic - a call to action, JAMA Neurol. 75 (1) (2018) 9–10, https://doi.org/10.1001/jamaneurol.2017.3299.

[2] J. Parkinson, An Essay on the Shaking Palsy, Published online, 1817.

[3] R.B. Postuma, D. Berg, M. Stern, et al., MDS clinical diagnostic criteria for Parkinson's disease, Mov. Disord. 30 (12) (2015) 1591–1601, https://doi.org/10.1002/mds.26424.

[4] C.G. Goetz, D. Nyenhuis, W. Poewe, et al., Movement Disorder Society-sponsored revision of the unified Parkinson's disease rating scale (MDS-UPDRS): scale presentation and clinimetric testing results, Mov. Disord. 23 (15) (2008) 2129–2170, https://doi.org/10.1002/mds.22340.

[5] D.A. Heldman, J.P. Giuffrida, R. Chen, et al., The modified bradykinesia rating scale for Parkinson's disease: reliability and comparison with kinematic measures, Mov. Disord. 26 (10) (2011) 1859–1863, https://doi.org/10.1002/mds.23740.

[6] A. Burton, How do we fix the shortage of neurologists? Lancet Neurol. 17 (6) (2018) 502–503, https://doi.org/10.1016/S1474-4422(18)30143-1.

[7] C.H. Adler, T.G. Beach, J.G. Hentz, et al., Low clinical diagnostic accuracy of early vs advanced Parkinson disease, Neurology 83 (5) (2014), https://doi.org/10.1212/WNL.0000000000000641, 406 LP - 412.

[8] N.P.S. Bajaj, V. Gontu, J. Birchall, J. Patterson, D.G. Grosset, A.J. Lees, Accuracy of clinical diagnosis in tremulous parkinsonian patients: a blinded video study, J. Neurol. Neurosurg. Psychiatry 81 (11) (2010) 1223–1228.

[9] C.G. Goetz, G.T. Stebbins, Assuring interrater reliability for the UPDRS motor section: utility of the UPDRS teaching tape, Mov. Disord. 19 (12) (2004) 1453–1456, https://doi.org/10.1002/mds.20220.

[10] S. Williams, D. Wong, J.E. Alty, S.D. Relton, Parkinsonian hand or clinician's eye? Finger tap bradykinesia interrater reliability for 21 movement disorder experts, J. Parkinsons Dis. 13 (4) (2023) 525–536, https://doi.org/10.3233/jpd-223256.

[11] J. Costa, H.A. Gonzalez, F. Valldeoriola, C. Gaig, E. Tolosa, J. Valls-Sole, Nonlinear dynamic analysis of oscillatory repetitive movements in Parkinson's disease and essential tremor, Mov. Disord. 25 (15) (2010) 2577–2586, https://doi.org/10.1002/mds.23334.

[12] M.J. Lee, S.L. Kim, C.H. Lyoo, J.O. Rinne, M.S. Lee, Impact of regional striatal dopaminergic function on kinematic parameters of Parkinson's disease, J. Neural Transm. 122 (5) (2015) 669–677, https://doi.org/10.1007/s00702-014-1296-x.

[13] L. di Biase, S. Summa, J. Tosi, et al., Quantitative analysis of bradykinesia and rigidity in Parkinson's disease, Front. Neurol. 9 (MAR) (2018) 1–12, https://doi.org/10.3389/fneur.2018.00121.

[14] J.W. Kim, Y. Kwon, J.S. Yun, et al., Regression models for the quantification of parkinsonian bradykinesia, Biomed. Mater. Eng. 26 (Suppl. 1) (2015) S2249–S2258, https://doi.org/10.3233/BME-151531.

[15] M. Yokoe, R. Okuno, T. Hamasaki, Y. Kurachi, K. Akazawa, S. Sakoda, Opening velocity, a novel parameter, for finger tapping test in patients with Parkinson's disease, Parkinsonism Relat. Disord. 15 (6) (2009) 440–444, https://doi.org/10.1016/j.parkreldis.2008.11.003.

[16] J. Stamatakis, J. Ambroise, J. Cremers, et al., Finger tapping clinimetric score prediction in Parkinson's disease using low-cost accelerometers, Comput. Intell. Neurosci. (2013), https://doi.org/10.1155/2013/717853. Published online.

[17] D.A. Heldman, A.J. Espay, P.A. LeWitt, J.P. Giuffrida, Clinician versus machine: reliability and responsiveness of motor endpoints in Parkinson's disease, Parkinsonism Relat. Disord. 20 (6) (2014) 590–595, https://doi.org/10.1016/j.parkreldis.2014.02.022.

[18] O. Martinez-Manzanera, E. Roosma, M. Beudel, R.W.K. Borgemeester, T. Van Laar, N.M. Maurits, A method for automatic and objective scoring of bradykinesia using orientation sensors and classification algorithms, IEEE Trans. Biomed. Eng. 63 (5) (2016) 1016–1024, https://doi.org/10.1109/TBME.2015.2480242.

[19] R. Agostino, A. Curra, M. Giovannelli, N. Modugno, M. Manfredi, A. Berardelli, Impairment of individual finger movements in Parkinson's disease, Mov. Disord. 18 (5) (2003) 560–565, https://doi.org/10.1002/mds.10313.

[20] H. Ling, L.A. Massey, A.J. Lees, P. Brown, B.L. Day, Hypokinesia without decrement distinguishes progressive supranuclear palsy from Parkinson's disease, Brain 135 (4) (2012) 1141–1153, https://doi.org/10.1093/brain/aws038.

[21] C. Lainscsek, P. Rowat, L. Schettino, et al., Finger tapping movements of Parkinson's disease patients automatically rated using nonlinear delay differential equations, CHAOS 22 (1) (2012), https://doi.org/10.1063/1.3683444.

[22] E. Ruzicka, R. Krupicka, K. Zarubova, J. Rusz, R. Jech, Z. Szabo, Tests of manual dexterity and speed in Parkinson's disease: not all measure the same, Parkinsonism Relat. Disord. 28 (2016) 118–123, https://doi.org/10.1016/j.parkreldis.2016.05.009.

[23] M. Bologna, G. Leodori, P. Stirpe, et al., Bradykinesia in early and advanced Parkinson's disease, J. Neurol. Sci. 369 (2016) 286–291, https://doi.org/10.1016/j.jns.2016.08.028.
[24] Bank PJM, J. Marinus, C.G.M. Meskers, J.H. de Groot, J.J. van Hilten, Optical hand tracking: a novel technique for the assessment of bradykinesia in Parkinson's disease, Mov. Disord. Clin. Pract. 4 (6) (2017) 875–883, https://doi.org/10.1002/mdc3.12536.
[25] Keisuke Shima, Toshio Tsuji, Akihiko Kandori, Masaru Yokoe, Saburo Sakoda, et al., Measurement and evaluation of finger tapping movements using magnetic sensors, Mov. Disord. 26 (6) (2018) 5628–5631, https://doi.org/10.3233/BME-151531.
[26] A. Kandori, M. Yokoe, S. Sakoda, et al., Quantitative magnetic detection of finger movements in patients with Parkinson's disease, Neurosci. Res. 49 (2) (2004) 253–260, https://doi.org/10.1016/j.neures.2004.03.004.
[27] Y. Sano, A. Kandori, K. Shima, et al., Quantifying Parkinson's disease finger-tapping severity by extracting and synthesizing finger motion properties, Med. Biol. Eng. Comput. 54 (6) (2016) 953–965, https://doi.org/10.1007/s11517-016-1467-z.
[28] C. Gao, S. Smith, M. Lones, et al., Objective assessment of bradykinesia in Parkinson's disease using evolutionary algorithms: clinical validation, Transl. Neurodegener. 7 (1) (2018) 1–8, https://doi.org/10.1186/s40035-018-0124-x.
[29] W.P. Teo, J.P. Rodrigues, F.L. Mastaglia, G.W. Thickbroom, Comparing kinematic changes between a finger-tapping task and unconstrained finger flexion-extension task in patients with Parkinson's disease, Exp. Brain Res. 227 (3) (2013) 323–331, https://doi.org/10.1007/s00221-013-3491-7.
[30] D.R. Roalf, P. Rupert, D. Mechanic-Hamilton, et al., Quantitative assessment of finger tapping characteristics in mild cognitive impairment, Alzheimer's disease, and Parkinson's disease, J. Neurol. 265 (6) (2018) 1365–1375, https://doi.org/10.1007/s00415-018-8841-8.
[31] S. Arora, V. Venkataraman, A. Zhan, et al., Detecting and monitoring the symptoms of Parkinson's disease using smartphones: a pilot study, Parkinsonism Relat. Disord. 21 (6) (2015) 650–653, https://doi.org/10.1016/j.parkreldis.2015.02.026.
[32] C.Y. Lee, S.J. Kang, S.K. Hong, H.I. Ma, U. Lee, Y.J. Kim, A validation study of a smartphone-based finger tapping application for quantitative assessment of bradykinesia in Parkinson's disease, PLoS One 11 (7) (2016), https://doi.org/10.1371/journal.pone.0158852.
[33] P. Kassavetis, T.A. Saifee, G. Roussos, et al., Developing a tool for remote digital assessment of Parkinson's disease, Mov. Disord. Clin. Pract. 3 (1) (2016) 59–64, https://doi.org/10.1002/mdc3.12239.
[34] G. Mitsi, E.U. Mendoza, B.D. Wissel, et al., Biometric digital health technology for measuring motor function in Parkinson's disease: results from a feasibility and patient satisfaction study, Front. Neurol. (2017) 8, https://doi.org/10.3389/fneur.2017.00273.
[35] L. Lalvay, M. Lara, A. Mora, et al., Quantitative measurement of akinesia in Parkinson's disease, Mov. Disord. Clin. Pract. 4 (3) (2017) 316–322, https://doi.org/10.1002/mdc3.12410.
[36] J. Prince, S. Arora, M. de Vos, Big data in Parkinson's disease: using smartphones to remotely detect longitudinal disease phenotypes, Physiol. Meas. 39 (4) (2018), https://doi.org/10.1088/1361-6579/aab512.
[37] T. Khan, D. Nyholm, J. Westin, M. Dougherty, A computer vision framework for finger-tapping evaluation in Parkinson's disease, Artif. Intell. Med. 60 (1) (2014) 27–40, https://doi.org/10.1016/j.artmed.2013.11.004.
[38] S. Williams, Z. Zhao, A. Hafeez, et al., The discerning eye of computer vision: can it measure Parkinson's finger tap bradykinesia? J. Neurol. Sci. 416 (February) (2020) 117003 https://doi.org/10.1016/j.jns.2020.117003.
[39] S. Williams, S.D. Relton, H. Fang, et al., Supervised classification of bradykinesia in Parkinson's disease from smartphone videos, Artif. Intell. Med. (2020) 110, https://doi.org/10.1016/j.artmed.2020.101966.
[40] Y. Liu, J. Chen, C. Hu, et al., Vision-based method for automatic quantification of parkinsonian bradykinesia, IEEE Trans. Neural. Syst. Rehabil. Eng. 27 (10) (2019) 1952–1961, https://doi.org/10.1109/TNSRE.2019.2939596.
[41] J.H. Shin, J.N. Ong, R. Kim, et al., Objective measurement of limb bradykinesia using a marker-less tracking algorithm with 2D-video in PD patients, Parkinsonism Relat. Disord. 81 (April) (2020) 129–135, https://doi.org/10.1016/j.parkreldis.2020.09.007.
[42] M. Lu, J.C. Niebles, L. Fei-Fei, et al., Quantifying Parkinson's disease motor severity under uncertainty using MDS-UPDRS videos, Med. Image Anal. 73 (2021) 102179, https://doi.org/10.1016/j.media.2021.102179.
[43] Z. Li, K. Lu, M. Cai, X. Liu, Y. Wang, J. Yang, An automatic evaluation method for Parkinson's dyskinesia using finger tapping video for small samples, J. Med. Biol. Eng. 42 (3) (2022) 351–363, https://doi.org/10.1007/s40846-022-00701-y.
[44] G. Vignoud, C. Desjardins, Q. Salardaine, et al., Video-based automated assessment of movement parameters consistent with MDS-UPDRS III in Parkinson's disease, J. Parkinsons Dis. 12 (7) (2022) 2211–2222, https://doi.org/10.3233/JPD-223445.
[45] G. Morinan, Y. Dushin, G. Sarapata, et al., Computer vision quantification of whole-body parkinsonian bradykinesia using a large multi-site population, NPJ Parkinsons Dis. 9 (1) (2023) 1–12, https://doi.org/10.1038/s41531-023-00454-8.
[46] M.R. Ali, J. Hernandez, E. Ray Dorsey, E. Hoque, D. McDuff, Spatio-temporal attention and magnification for classification of Parkinson's disease from videos collected via the internet, in: Proceedings - 2020 15th IEEE International Conference on Automatic Face and Gesture Recognition, FG 2020, 2020, pp. 207–214, https://doi.org/10.1109/FG47880.2020.00008. Published online.
[47] M.H.G. Monje, S. Domínguez, J. Vera-Olmos, et al., Remote evaluation of Parkinson's disease using a conventional webcam and artificial intelligence, Front. Neurol. 12 (December) (2021) 1–11, https://doi.org/10.3389/fneur.2021.742654.
[48] Z. Guo, W. Zeng, T. Yu, et al., Vision-based finger tapping test in patients with Parkinson's disease via spatial-temporal 3D hand pose estimation, IEEE J. Biomed. Health Inform. 26 (8) (2022) 3848–3859, https://doi.org/10.1109/JBHI.2022.3162386.
[49] Z. Yin, V.J. Geraedts, Z. Wang, M.F. Contarino, DIbeklioglu H, Van Gemert J., Assessment of Parkinson's disease severity from videos using deep architectures, IEEE J. Biomed. Health Inform. 26 (3) (2022) 1164–1176, https://doi.org/10.1109/JBHI.2021.3099816.
[50] C.G. Goetz, S. Fahn, P. Martinez-Martin, et al., MDS-UPDRS: The MDS-Sponsored Revision of the Unified Parkinson's Disease Rating Scale 1, 2008. www.movementdisorders.orginfo@movementdisorders.org (414).
[51] C.G. Goetz, D. Nyenhuis, W. Poewe, et al., Movement Disorder Society-sponsored revision of the unified Parkinson's disease rating scale (MDS-UPDRS): scale presentation and clinimetric testing results, Mov. Disord. 23 (15) (2008) 2129–2170, https://doi.org/10.1002/mds.22340.
[52] D.A. Heldman, J.P. Giuffrida, R. Chen, et al., The modified bradykinesia rating scale for Parkinson's disease: reliability and comparison with kinematic measures, Mov. Disord. 26 (10) (2011) 1859–1863, https://doi.org/10.1002/mds.23740.
[53] Victordibia/Handtracking: Building a Real-time Hand-Detector using Neural Networks (SSD) on Tensorflow, Accessed February 25, https://github.com/victordibia/handtracking, 2021.
[54] OpenCV - OpenCV, Accessed February 25, https://opencv.org/, 2021.
[55] Batch Normalization | Proceedings of the 32nd International Conference on International Conference on Machine Learning - Volume 37. Accessed February 25, 2021. https://doi.org/10.5555/3045118.3045167.
[56] Keras: the Python Deep Learning API. https://keras.io/. (Accessed 25 February 2021).
[57] D.P. Kingma, J.L. Ba, Adam: A method for stochastic optimization, in: *3rd International Conference on Learning Representations, ICLR* 2015 - *Conference Track Proceedings*. International Conference on Learning Representations, ICLR, 2015.
[58] E. Choi, A. Schuetz, W.F. Stewart, J. Sun, Using recurrent neural network models for early detection of heart failure onset, J. Am. Med. Inform. Assoc. 24 (2) (2017) 361–370, https://doi.org/10.1093/jamia/ocw112.
[59] R.R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, D. Batra, Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization, Published online October 7, 2016, https://doi.org/10.1007/s11263-019-01228-7.
[60] B.M. Bot, C. Suver, E.C. Neto, et al., The mPower study, Parkinson disease mobile data collected using ResearchKit, Sci. Data. (2016) 3, https://doi.org/10.1038/sdata.2016.11.
[61] R. Krupicka, Z. Szabo, S. Viteckova, E. Ruzicka, Motion capture system for finger movement measurement in Parkinson disease, Radioengineering 23 (2) (2014) 659–664.
[62] Z. Zhao, H. Fang, S. Williams, et al., Time series clustering to examine presence of decrement in Parkinson's finger-tapping bradykinesia, in: Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS, 2020, pp. 780–783, https://doi.org/10.1109/EMBC44109.2020.9175638, 2020-July.
[63] Z. Zhao, H. Fang, S. Williams, et al., Time series clustering to examine presence of decrement in Parkinson's finger-tapping bradykinesia, in: Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS, 2020, pp. 780–783, https://doi.org/10.1109/EMBC44109.2020.9175638, 2020-July.
[64] I.H.J. Willemse, S. Schootemeijer, R. van den Bergh, H. Dawes, J.H. Nonnekes, B.P. C. van de Warrenburg, Smartphone applications for movement disorders: towards collaboration and re-use, Parkinsonism Relat. Disord. 120 (2024) 105988, https://doi.org/10.1016/j.parkreldis.2023.105988. September 2023.
[65] W. Maetzler, A. Mirelman, A. Pilotto, R. Bhidayasiri, Identifying subtle motor deficits before Parkinson's disease is diagnosed: what to look for? J. Parkinsons Dis. (2024) 1–10, https://doi.org/10.3233/jpd-230350. Published online.
[66] S. Williams, Z. Zhao, A. Hafeez, et al., The discerning eye of computer vision: can it measure Parkinson's finger tap bradykinesia? J. Neurol. Sci. 416 (February) (2020) 117003 https://doi.org/10.1016/j.jns.2020.117003.
[67] A.J. Espay, D.E. Beaton, F. Morgante, C.A. Gunraj, A.E. Lang, R. Chen, Impairments of speed and amplitude of movement in Parkinson's disease: a pilot study, Mov. Disord. 24 (7) (2009) 1001–1008, https://doi.org/10.1002/mds.22480.
[68] C. Gao, P. Zhang, J. Liu, et al., Objective assessment of bradykinesia in Parkinson's disease using evolutionary algorithms: clinical validation, Transl. Neurodegener. 7 (1) (2018) 18, https://doi.org/10.1186/s40035-018-0124-x.