

Article

Not peer-reviewed version

Webcam-Based Pain Measurement Using Pupillary Diameter

[Natalia Shamantseva](#) , Arseniy Polyakov , [Vsevolod Lyakhovetski](#) , Margarita Bystrova , [Ivan Sakun](#) , [Sergey Ananyev](#) , [Yury Gerasimenko](#) , [Tatiana Moshonkina](#) *

Posted Date: 11 March 2026

doi: 10.20944/preprints202603.0782.v1

Keywords: video pupillometry; pupillary diameter; pain measurement; neural network; transcutaneous spinal cord stimulation



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Article

Webcam-Based Pain Measurement Using Pupillary Diameter

Natalia Shamantseva, Arseniy Polyakov, Vsevolod Lyakhovetskii, Margarita Bystrova, Ivan Sakun, Sergey Ananyev, Yuri Gerasimenko and Tatiana Moshonkina *

Pavlov Institute of Physiology, Russian Academy of Sciences, 199034 St. Petersburg, Russia

* Correspondence: moshonkina@infran.ru

Abstract

Pupillometry is a reliable method of pain control. There are experimental conditions under which standard pupillometry equipment cannot be used. Studying effects of different pulse forms used for transcutaneous spinal cord stimulation (tSCS) is one such task. The aim was to create a system for recording pupil diameter based on a web camera because it can be synchronized with external equipments, which allows the diameter to be recorded simultaneously with other physiological signals. A markerless system for recording and analyzing pupil diameter using deep neural networks was developed based on a commercially available web camera. The accuracy of this system was compared with the accuracy of measurements using manual analysis with the ImageJ. The system was tested in a study of the dependence of tolerance to tSCS on the shape of stimulating pulses, which involved the participation of volunteers (n=12). The results of the developed pupillometry were compared with the pain rating scale traditionally used in such studies. The developed system is accurate in determining the pupil diameter, comparable to human accuracy. The pupillometry results reproduced those obtained using a subjective pain scale. The developed method was found to be a reliable method for recording pain in electrophysiological studies.

Keywords: video pupillometry; pupillary diameter; pain measurement; neural network; transcutaneous spinal cord stimulation

1. Introduction

In the context of the prevalence of transcutaneous spinal cord stimulation (tSCS) for spinal network modulation in neurophysiological movement studies in humans and motor rehabilitation, the significance of kilohertz-modulated pulses for this stimulation has been discussed [1]. Studies have shown that tSCS with a carrier frequency of 5–10 kHz results is better tolerated than conventional rectangular pulses during both single-pulse and train stimulation [2,3]. In these studies, a pain rating scale was used to assess the pain experienced during stimulation. Participants ranked their pain on a scale from 0 to 10, where 0 indicated no pain and 10 indicated the most painful sensation imaginable. A comparative functional study of tSCS using kHz-modulated and conventional pulses is therefore highly relevant [1]. In such a study, it is important to synchronize and monitor both the parameters of movements induced by tSCS and the relationship between motor responses and any discomfort or pain associated with the stimulation. The pain rating scale is subjective and not instrumental. It cannot be used to synchronize with stimulation and recording devices.

Pupillary diameter monitoring is a recognised method of measuring pain during anesthesia and analgesia [4]. Infrared pupillometers are used in clinical practice and research studies [5]. Ultrasound technology is also beginning to be used for this purpose [6]. However, none of the commercially available devices are suited during the study of motor function modulation through tSCS, since pain measurement must be performed during movement and/or in a specific experimental position. Classical studies of tSCS effects on locomotor networks are typically conducted when individuals are

lying on their side, with each leg suspended on an independent swing-like support, allowing horizontal-plane movements of the legs [7–9]. In this configuration, the legs are unloaded, making the initiation of locomotor movements by tSCS easier compared with other body and leg positions. This setup is preferable for a comparative functional study of spinal stimulation using kHz-modulated and conventional pulses.

The aim of this study was to develop a system for recording and analysing pupillary diameter in the aforementioned experimental setup, in order to monitor the onset and progression of pain during movement investigations by the spinal stimulation. This system should digitise the pupillary diameter during the movement experiment and synchronise with the other devices in the experimental setup.

We selected webcam video recording of the eye because we use the optical motion capture technology to record and visualise movements. Usually, the optical motion capture software synchronises devices for muscle activity recording and the stimulator, and it also allows synchronisation of a webcam.

The proposed pupillary recording system is based on video recording of the eye in the visible light range, recognition of pupil boundaries and a calibration reference within the video frames using machine learning algorithms and reconstruction of pupil diameter from the detected boundaries and the reference. We tested the developed technology through a comparative study examining tolerance to tSCS with both kHz-modulated and conventional pulses.

2. Materials and Methods

2.1. Subjects

Twelve healthy volunteers (one female) participated in the study, with a mean age of 25 ± 4 years and body mass index of 23.1 ± 1.6 kg/m². Exclusion criteria included any neurological and musculoskeletal disorder, any acute or chronic disease symptoms, medical or surgical procedures or trauma within four weeks prior to the study, and pregnancy. All participants self-assessed as healthy on the day of the study. Ethical approval was obtained from the Ethics Committee of the Pavlov Institute of Physiology, Russian Academy of Sciences (Minutes No. 25-07, dated 1 August 2025). The study was conducted in accordance with the Declaration of Helsinki (World Medical Association, 2013). Prior to participation, all participants received detailed information about the study objectives and procedures and provided written informed consent.

2.2. Protocol

Participants were examined in a lateral position on their left side on an examination couch. Both lower limbs were placed on independent horizontal suspension supports, each attached to the ceiling with flexible cables in a swing-like configuration, which allowed for stepping-like movements [7,8]. For each participant, two recording sessions were conducted on the same day. During sessions participants were instructed to look at the webcam lens positioned in front of their face. Each session began with a 30-second baseline period during which no stimulation was delivered and participants were instructed to stay relaxed. This baseline served as a control segment for subsequent analysis of pupil area dynamics. Following the baseline period, the spinal stimulation started.

Two stimulation paradigms were tested: (1) with single rectangular monopolar pulses not modulated at kHz frequency (MN) and (2) with single rectangular monopolar pulses modulated at a 5 kHz frequency (MM). The duration of the MN and MM pulses was 1 msec. The pulses followed one after another with an interval of about 3 sec or longer, and their intensity increased. At each intensity level, three pulses were delivered. This is a procedure for motor evoked potential recording during spinal stimulation that also repeats the protocol from the study of tSCS tolerance that was used in the early researches [2,3]. Each stimulation paradigm lasted up to 6 minutes. Stimulation intensity began at 10 mA and was increased in 10 mA incremental steps until either the participant requested termination or a maximum current intensity of 250 mA was reached. Participants were

instructed to verbally indicate when they reached their maximum tolerable intensity (i.e., when they could not tolerate further stimulation). At this point, they reported their perceived pain level using a Visual Analogue Scale (VAS) ranging from 0 to 10, where 0 represented "no pain or discomfort at all" and 10 represented "the worst pain" [10]. The order of the MN and MM paradigms differed among participants, with an interval of at least five minutes between paradigms.

2.3. Eye Video Recording

A USB webcam (Logitech C922 Pro Stream, Logitech, Switzerland) was mounted on a microphone stand and positioned approximately 10–15 cm from the anterior surface of the right eye, perpendicular to the pupil plane. Video acquisition was performed at 30 frames per second throughout all experimental sessions. Illumination was provided by a Ulanzi VL49 RGB light (Ulanzi, Shenzhen, China), directed toward the right eye of the participant. The color temperature was set to 2500 K to maintain consistent warm illumination and minimise spectral variations across sessions. Luminance intensity was monitored using a calibrated lux meter (RADEX LUPIN, RADEX, Russia). The illuminance was maintained at 20–40 lux throughout all recordings. Illuminance measurements were performed immediately before each recording session by positioning the lux meter sensor directly adjacent to the lateral canthus of the eye, ensuring that light from the lamp struck the sensor perpendicularly. The illuminance varied by less than 4 lux during testing for each participant.

To account for variations in camera scale caused by head movements or camera displacement during the experiment, a reference was affixed to the skin adjacent to the eye (1–2 cm below the lower eyelid). The reference consisted of a white paper sticker with two parallel red line segments oriented approximately horizontally, with its plane aligned to the anterior ocular surface.

2.4. Transcutaneous Spinal Cord Stimulation

Stimulation was delivered using Neostim-5 (LTD Cosyma, Moscow, Russia). The cathode (diameter 2.5 cm, ValuTrobe® Axelgaard Manufacturing Co., Fallbrook, CA, USA) was positioned midline between the spinous processes of the T11–T12 vertebrae. Two anodes (5 × 10 cm², ValuTrobe® Axelgaard Manufacturing Co., Fallbrook, CA, USA) were placed symmetrically over the iliac crests.

The Neostim-5 and webcam were synchronized by Qualisys motion capture system (Qualisys, Gothenburg, Sweden).

2.5. Pupil Boundary Detection with DeepLabCut

2.5.1. Training Dataset Preparation

Pupil boundary keypoints were detected using DeepLabCut (version 3.0.0rc9), an open-source deep learning framework for markerless pose estimation [11]. A single-animal project was created.

Nine keypoints were manually annotated to define the pupil boundary: eight points distributed at 45° intervals around the pupil perimeter (P1–P8) and one at the pupil center (P9), which was not used in further analysis. Additionally, four keypoints were labelled on a reference visible in the video frame (S1–S4) to serve for subsequent normalisation (Figure 1). The distance between S2 and S4 was 8 mm.

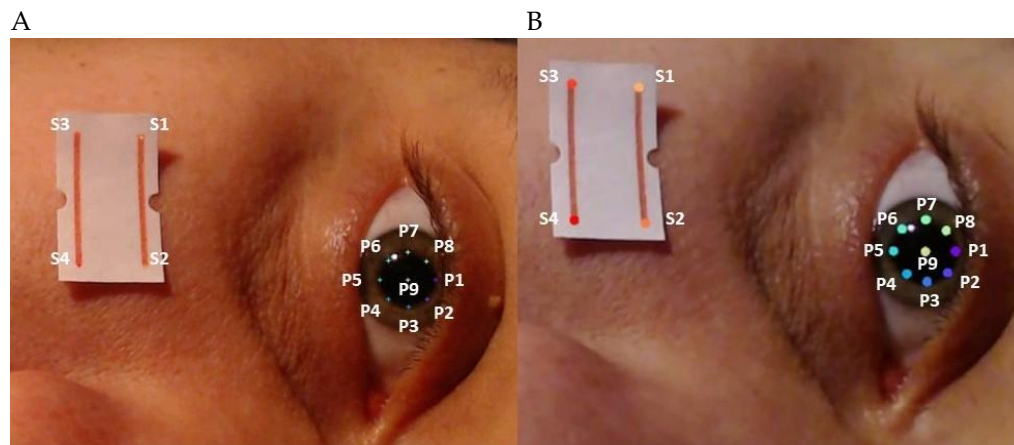


Figure 1. Comparison of manually annotated and neural network-predicted keypoints. A: manually arranged keypoints. B: keypoints arranged by neural network. P1–P8 denote the pupil boundary, P9 indicates the pupil center, and S1–S4 mark the endpoints of the reference marker (red lines).

Training frames were extracted from video recordings with k-means clustering to ensure representation of diverse pupil states and lighting conditions. A total of 701 frames (consisting of 92 images from an open dataset and frames extracted across 12 participants) were manually labelled using Napari-DeepLabCut graphical interface. Labels placement was performed manually by a single rater based on visual inspection. Label P1 was placed at the uppermost point along the pupil boundary. Labels P2–P8 were subsequently placed in a clockwise direction along the pupil contour. Label P9 was placed at the center of the pupil. Approximately equal spacing was maintained between all boundary labels. Frames in which keypoints were occluded or not visible were left unlabelled, following established best practices.

Following initial labelling, the training dataset was created with a training fraction of 0.95 (95% training, 5% validation) and shuffle number 1. The dataset was configured with uniform cropping and image augmentation parameters including affine transformations (rotation $\pm 30^\circ$, scaling 0.5–1.25), Gaussian noise ($\sigma = 12.75$), and motion blur to enhance model robustness.

2.5.2. Network Architecture and Training

The neural network backbone was a ResNet-50 architecture pretrained on ImageNet, leveraging transfer learning to improve generalisation performance with limited training data. The network consisted of the ResNet-50 feature extractor followed by custom deconvolutional layers to predict heatmaps for keypoint localisation and location refinement fields for sub-pixel accuracy.

Training was performed using the PyTorch engine in Google Colab with the following hyperparameters: batch size of 8200 training epochs (with model snapshots saved every 5 epochs), AdamW optimizer with initial learning rate of 0.0005, and a learning rate schedule with step decay at epochs 90 and 120 to 0.0001 and 0.00001, respectively [12]. Image crops of 448×448 pixels were randomly sampled during training with a maximum shift of 10% to improve spatial invariance. The colour mode was set to RGB, and images were normalised prior to network input. The key metric to choose the best training epoch was the highest value of test Mean Average Precision.

2.5.3. Model Evaluation and Refinement

The model performance was evaluated by the built-in DeepLabCut algorithm after every training session. This algorithm found the average root mean square error between the manual (test dataset) and predicted labels [13]. The pCutoff likelihood filter for the predicted labels was set to 0.2.

Early versions of the model were not accurate enough for the pupil border detection task. To optimise detection accuracy, we labelled more variable frames and added them to the training and validation datasets. Furthermore, an iterative active refinement strategy was employed. During this

process, frames with visibly incorrect marker placements were manually identified and relabelled. These corrected frames were then merged into the training dataset and the network was retrained.

2.5.4. Video Analysis

Following training, the network was applied to all experimental videos to predict keypoint coordinates for every frame. For each detected keypoint, the software outputted x- and y-pixel coordinates alongside a likelihood score (range 0–1) indicating prediction confidence. Results were exported as CSV files containing frame-by-frame coordinates and likelihood values. Additionally, labelled videos with overlaid keypoint predictions were generated to enable visual quality control of the tracking results.

2.6. Post-Processing Pipeline in MATLAB

All post-processing analyses were performed in MATLAB (R2023a, MathWorks, Natick, MA, USA). The pipeline consisted of four sequential stages: (1) ellipse fitting to the detected pupil boundary keypoints and area calculation, (2) outlier detection and temporal interpolation, (3) the reference length extraction, and (4) normalisation of pupil metrics by the reference.

2.6.1. Pupil Area Calculation and Ellipse Fitting

The DeepLabCut-generated CSV files containing coordinates and likelihood scores for the nine pupil keypoints (P1–P8: boundary markers; P9: center) were imported using a custom MATLAB script.

Temporal smoothing. To minimize marker instability caused by camera or head movements, a sliding median filter (window = 3 frames, step = 1 frame) was applied to the x- and y-coordinates of each keypoint. This approach allowed for robust boundary reconstruction by utilizing a larger set of markers across consecutive frames (up to 24 markers for a 3-frame window).

Ellipse fitting algorithm. For each video frame, an ellipse was fitted to the eight pupil boundary points using a two-stage robust fitting procedure:

Initial fit: A conic section was fitted to all available valid keypoints using the least squares method via singular value decomposition, which solves the general conic equation: $ax^2 + bxy + cy^2 + dx + ey + f = 0$.

Outlier removal: The perpendicular distance from each keypoint to the fitted conic was calculated. Outliers were identified using a robust statistical threshold: points lying beyond $\text{median}(\text{distance}) + k \times \text{interquartile range (IQR)}$ were excluded. Additionally, an absolute distance threshold of $\alpha \times a_0$ was applied, where a_0 is the semi-major axis length from the initial fit and $\alpha = 0.1$.

Refined fit: The ellipse was re-fitted using only the inlier keypoints. The fitted conic parameters were converted to geometric ellipse parameters (center coordinates x_c, y_c ; semi-major axis a ; semi-minor axis b ; orientation angle φ).

Pupil area. For all frames where a valid ellipse was fitted (requiring a minimum of 5 keypoints), the pupil area was calculated using the derived semi-axes as $A = \pi ab$.

2.6.2. Outlier Detection and Temporal Interpolation

To remove physiologically implausible values resulting from tracking errors or blinks, an IQR-based outlier detection algorithm was applied to the pupil area time series. Area values falling outside the range $[Q_1 - 1.5 \times \text{IQR}, Q_3 + 3.0 \times \text{IQR}]$ were replaced with NaN. Missing values (from outlier removal, low-likelihood frames, or failed ellipse fits) were interpolated using linear interpolation with a maximum gap threshold of 30 frames (~1 second). Gaps exceeding this duration were left as NaN, as longer interruptions typically indicated sustained blinks or pupil occlusions that could not be reliably interpolated.

2.6.3. Reference Length Calculation

To normalise pupil area across participants and account for inter-individual variability in camera-to-eye distance, a visible reference marker was used. The reference consisted of a red line segment captured within the video frame, defined by two endpoints (keypoints S1 and S2 for one endpoint pair, S3 and S4 for another pair).

Reference coordinates were extracted from the DeepLabCut CSV output using a separate MATLAB script. The Euclidean distance L between the most stable marker end-points was calculated frame-by-frame as:

$$L = \sqrt{(X_i - X_j)^2 + (Y_i - Y_j)^2}$$

where (X_i, Y_i) and (X_j, Y_j) are the pixel coordinates of the two endpoints.

Normalizing reference lengths were processed by: (1) excluding values outside manually set minimum and maximum thresholds; (2) removing statistical outliers using IQR range; and (3) applying linear interpolation to data gaps shorter than a specified duration threshold.

2.6.4. Normalisation Procedure

Pupil area was normalised by the reference length using a custom MATLAB script to yield dimensionless metrics invariant to camera positioning:

$$A_{\text{norm}} = \frac{A}{L^2}$$

where A is pupil area (px^2) and L is the reference length (px).

Normalisation was performed only for frames in which both the pupil metric and the reference length were available (non-NaN) and the reference length exceeded a minimum threshold of 1×10^{-6} pixels to avoid division by near-zero values. The normalised time series were saved for subsequent statistical analysis.

2.7. Validation of Neural Network Predictions

To validate the accuracy of pupil area measurements obtained from the neural network, manual measurements were performed in ImageJ (National Institutes of Health, Bethesda, MD, USA) using ellipse fitting on a subset of frames. 120 frames were randomly selected from the first 30 seconds (baseline period) of the initial recording session across all 12 participants (10 frames per participant). For each frame, three measurements were independently calculated using both ImageJ manual analysis and the neural network: (1) pupil area in pixels, (2) distance between the two most stable reference keypoints (S2 and S4) in pixels, and (3) pupil diameter converted to millimeters using the calibrated reference length (8 mm).

Normality of the data was assessed using the Shapiro–Wilk test. Since the data did not follow a normal distribution, agreement between manual and automated measurements was evaluated using Bland–Altman analysis [14] for repeated measures with non-parametric summary statistics. Bias was expressed as the median difference with interquartile range [Q1; Q3], and limits of agreement (LOA) were calculated as the 2.5th and 97.5th percentiles of the differences. Additionally, Pearson correlation coefficients for repeated measurements were computed to assess the strength of linear association between the two methods. All statistical analyses were performed in RStudio (version 4.3.1, R Foundation for Statistical Computing, Vienna, Austria).

2.8. Analysis of the Relationship Between Stimulation Intensity, Pain, and Pupil Diameter

For each tSCS paradigm (MM and MN), two time intervals were extracted from the continuous tracking data for comparison: (1) a 30-second baseline period (900 frames) recorded prior to the stimulation, and (2) a 9-second period (270 frames) corresponding to the maximum tolerable stimulation intensity. Three participants who tolerated the maximum stimulation intensity of 250 mA without requesting termination were excluded from this analysis, as their data did not contain a

clearly defined tolerance threshold. The remaining nine participants were included in the statistical comparison. No likelihood filtering was applied to the tracking data.

To process the extracted frames, the normalized pupil area (A) was averaged across all frames within each respective time interval. This averaged normalized area was then converted into an absolute pupil diameter (D , in mm) using the formula $D = 8 \cdot \sqrt{A/\pi}$, where 8 represents the physical reference distance (in mm) between the S4 and S2 markers.

The pupillary response, defined as the absolute change in pupil diameter (in mm), was calculated for each stimulation paradigm by subtracting the average baseline diameter from the average diameter at the maximum tolerable intensity.

Since the data were non-normally distributed, the Wilcoxon signed-rank test was used to evaluate whether pupil diameter during stimulation differed significantly from baseline. A p -value < 0.05 was considered statistically significant.

3. Results

3.1. Network Performance Metrics

The final training session consisted of 186 epochs. At the 180th training epoch the ResNet-50-based neural network achieved robust detection performance on the validation set. Quantitative evaluation metrics are summarized in Table 1.

Table 1. DeepLabCut neural network performance metrics on training and validation datasets.

Metric	Training Set	Validation Set
Mean Average Precision (mAP), %	90.81	87.54
Mean Average Recall (mAR), %	92.76	89.44
Root Mean Squared Error (RMSE), pixels	7.34	8.14
RMSE at pCutoff = 0.2, pixels	6.77	7.42

The network achieved Mean Average Precision (mAP) of 87.54% and Mean Average Recall (mAR) of 89.44% on the validation set, indicating high accuracy in keypoint prediction. The RMSE was 8.14 pixels for the test set.

3.2. Agreement with ImageJ Manual Calculations

A Bland–Altman analysis comparing the pupil area determined by a neural network and manual measurements demonstrated a median bias of 15.5 [−83.1; 200.3] square pixels and limits of agreements (LOA) of −626 and 731 square pixels. The two methods were strongly correlated ($r = 0.51$, $p < 0.001$) (Figure 2).

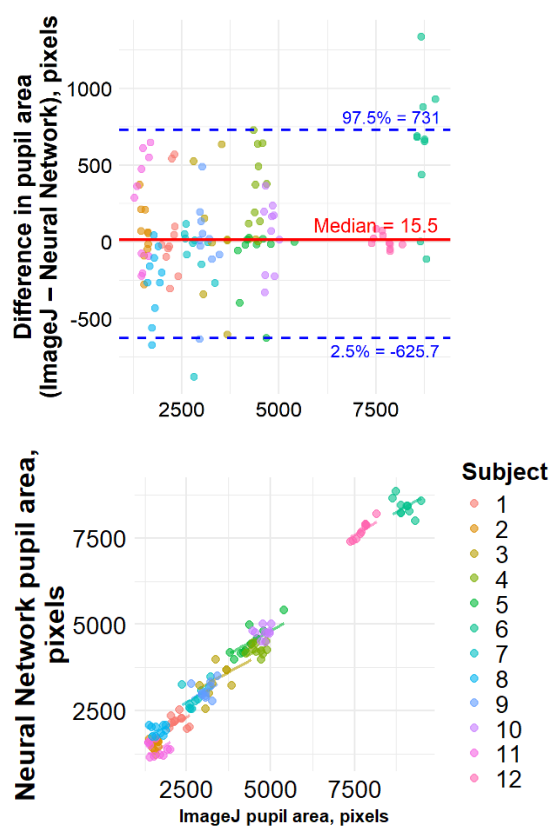


Figure 2. Results of non-parametric Bland–Altman (left) and Spearman’s correlation (right) analyses for paired measurements comparing pupil area (square pixels) obtained using the neural network and ImageJ.

The bias in a reference length ($S_4 - S_2$) was 0.6 [-0.2; 1.1] pixels, with LOA from -1.6 to 6.9 pixels and a correlation of 0.97 ($p < 0.001$) (Figure 3).

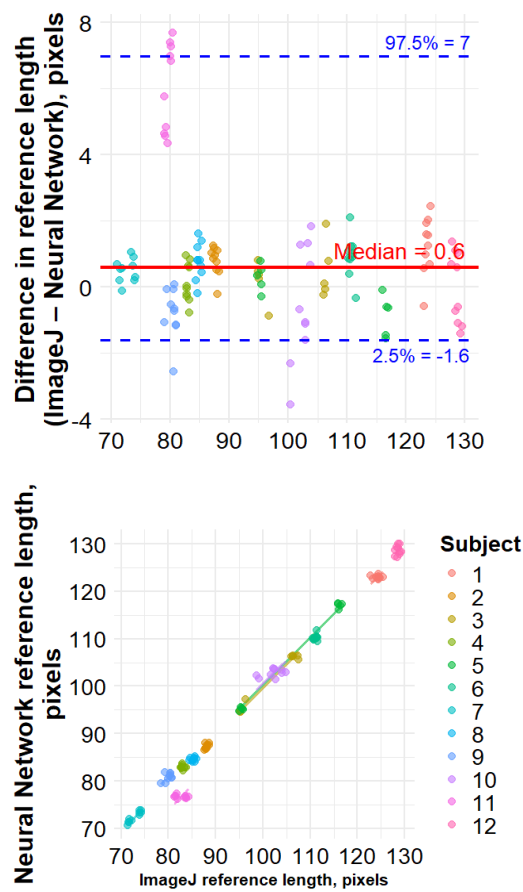


Figure 3. Results of non-parametric Bland–Altman (left) and Spearman’s correlation (right) analyses for paired measurements comparing reference length (pixels) obtained using the neural network and ImageJ].

The bias in pupil diameter was 0.016 [-0.08; 0.176] mm, with LOA from -0.62 to 0.52 mm and a correlation of 0.53 ($p < 0.001$) (Figure 4).

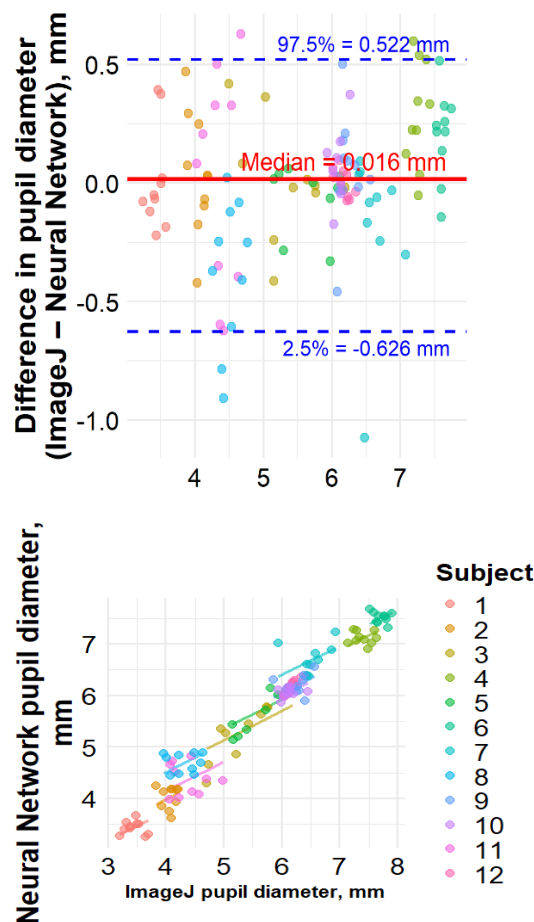


Figure 4. Results of non-parametric Bland–Altman (left) and Spearman’s correlation (right) analyses for paired measurements comparing pupil diameter (mm) obtained using the neural network and ImageJ.

3.3. Maximum Tolerable Intensity

Maximum tolerable intensities were 240 [215; 250] mA for MM and 235 [120; 240] mA for MN pulses. Maximum tolerable intensity for the MM pulses tended to be higher than for the MN pulse ($p = 0.06$).

3.4. Pupil Diameter versus Stimulation Paradigm

The pupil diameter before the MM paradigm was 5.9 [4.3; 6.2] mm, and before the MN paradigm was 5.7 [4.3; 6.1] mm. No significant differences were obtained between these control values ($p > 0.05$). At the maximum tolerable intensity, the pupil diameter was 6.2 [4.3; 6.6] mm for the MM paradigm, with a pupillary response of 0.2 [-0.1; 0.3] mm, which was not significantly different compared to the control value (Figure 5). The MN paradigm was found to significantly dilate the pupil by up to 6.4 [5.4; 7.1] mm ($p = 0.03$), producing a pupillary response of 0.9 [0.2; 1.4] mm. The pupillary response for the MN paradigm was significantly greater than for the MM paradigm ($p = 0.01$).

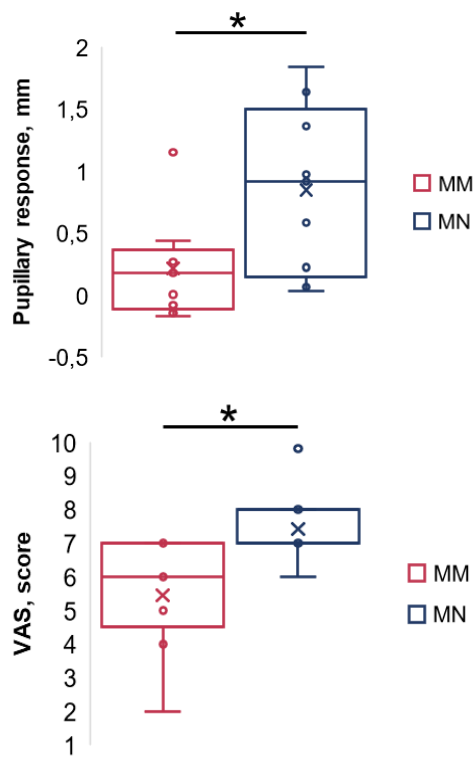


Figure 5. Pupillary responses (left) and Visual Analogue Scale scores (right) for the MM and MN stimulation paradigm, number of participants = 9; * $p < 0.05$.

3.5. Visual Analogue Scale Score versus Stimulation Paradigm

VAS scores were significantly higher for the MN pulses and were equal to 7 [7; 8] compare to 6 [4; 7] for the MM pulses ($p < 0.01$) (Figure 5).

4. Discussion

This study shows that webcam-based video pupillometry combined with DeepLabCut provides a practical and sensitive approach for monitoring pupillary diameter during noninvasive spinal stimulation. The approach enabled continuous, synchronised assessment of pupil dynamics and demonstrated that pupil diameter differentiates between tSCS paradigms at maximum tolerable intensity.

4.1. Comparison with Existing Methods

In contrast to infrared pupillometers and ultrasound-based approaches, which are often limited to static clinical settings [4–6], the proposed method operating with a standard RGB webcam and flexible mounting, makes it compatible with side-lying, leg-suspended locomotor setup [7–9]. While clinical devices typically provide standardised optics and illumination, they are less suited for dynamic tSCS protocols requiring integration with motion capture system and electromyography. In this study, the use of DeepLabCut for markerless pupil boundary detection and reference lines allowed or the reconstruction of pupil diameter in millimeters under a common spinal stimulation protocol [2,3,8,15].

Several deep learning approaches for pupil boundary detection have been proposed [16–19] Convolutional neural networks for visible-light pupillometry have been developed to estimate elliptical pupil parameters directly from images, eliminating the need for a separate ellipse-fitting step. DeepLabCut has been successfully applied to detect pupil boundaries in rodent eye-tracking studies, and dedicated repositories provide pre-trained models for rodent pupillometry. The use of Raspberry Pi computers and cameras is a key component of this technique, which “require extensive

technical and experimental expertise and may not be available to the majority of laboratories" [16]. Other work describes an algorithm for obtaining a pupil mask and extracting the parameters of an ellipse that represents the boundary of an elliptical pupil recorded in visible light [17]. This algorithm has a limitation in that it uses a training dataset containing images that differ from the experimentally acquired images. The experimental images were captured with a commercially available Android smartphone. A webcam-based deep learning eye-tracking system has also been explored for online behavioral experiments [18]. This system records eye fixations, eye movements, and blinks, but not pupil diameter. However, none of these implementations address the specific requirements of our experimental paradigm.

Traditional computer vision methods for pupil detection, such as the circular Hough transform, ellipse fitting, Daugman's integro-differential operator, and the radial symmetry transform, rely on handcrafted features and perform well under controlled infrared illumination with minimal occlusion [20]. However, these algorithms are sensitive to specular reflections, partial eyelid occlusion, and motion blur, which are unavoidable in a visible-light setup where stimulation-evoked movements shift the head and face within the camera field of view. In this study, the use of DeepLabCut's ResNet-50 backbone with transfer learning and iterative active refinement allowed robust keypoint detection in visible light across diverse pupil states, lighting conditions, and head positions, achieving a mAP of 87.54% and mAR of 89.44% on the validation set despite the challenging recording environment [21].

4.2. Pupillary Response to Transcutaneous Spinal Cord Stimulation

In our study, the MN stimulation paradigm was less tolerable than the MM paradigm. Three results demonstrate this effect: the tolerable intensity was lower in the MN paradigm than in the MM paradigm, the VAS score was higher in the MN paradigm, and the pupil diameter and its change were higher in the MN paradigm. These results are consistent with previous studies that found tolerance to kHz-modulated pulses was higher than tolerance to conventional pulses [2,3,15]. In previous studies, researchers tested tolerance to tSCS using a pain rating scale or stimulation until participants reported it was too intense.

4.3. Future Directions

Taken together, our results demonstrate that the developed system for recording and analyzing pupil diameter can be used in further studies of the functional effects of tSCS with conventional and modulated pulses and the relationship between functional effect and tolerance to stimulation. In such experiments, tSCS is delivered continuously at 15-30 Hz for several minutes while the participant lies in the side-lying position with legs suspended, producing rhythmic stepping movements [8]. Because the stimulation is ongoing and movements are continuous, pain and discomfort cannot be assessed using single-point subjective scales; instead, a dynamic objective measure is required. The webcam-based pupillometry system presented here meets this requirement by providing frame-by-frame pupil diameter values that are synchronised with electromyography and motion capture data. Future work will extend the current single-pulse protocol to continuous tSCS paradigms and will evaluate whether differences in pupillary responses across waveforms observed during single pulse stimulation are replicated during sustained locomotor activity.

5. Limitations

We opted not to apply likelihood-based filtering, as the combination of long recording sessions (up to 6 min) and stimulation-induced movement artifacts made standard thresholds impractical. Tracking quality was therefore verified through visual inspection.

Another limitation of the method is the difficulty of controlling illumination during video acquisition. When the optical contrast between the pupil and the iris is low, a high light intensity is

required. Additionally, controlled (directional) illumination is necessary because diffuse or irregular lighting produces multiple reflections that obscure the pupil-iris boundary.

Another limitation of the method is the necessity of a reference that compensates only for uniform changes in imaging scale, specifically situations in which the pupil moves closer to or farther from the camera. In cases of non-uniform scale changes, such as rotation of the plane containing both the pupil and the reference (e.g., due to head tilt), normalization would require a two-dimensional reference object with a known area (e.g., a triangle, square, or ellipse), rather than a line segment. Accordingly, the DeepLabCut model would need additional training to detect at least three points of the normalization marker to reconstruct the triangular boundary and compute its area.

The final limitation of the method is that the current model is only trained to recognize a specific reference. If a researcher uses a different object as a reference, the model will require additional training.

6. Conclusions

We have developed a system for recording and analyzing pupillary diameter in a human movement study protocol to control pain during spinal stimulation. The system uses a webcam and works in standard room lighting. It is synchronized with electrophysiological equipment. It can digitize pupillary diameter during movement experiments.

Author Contributions: Conceptualization, Y.G. and T.M.; methodology, N.S. and T.M.; software, A.P. and V.L.; validation, N.S. and M.B.; formal analysis, A.P., V.L. and N.S.; investigation, A.P., I.S., M.B. and S.A.; data curation, N.S.; writing—original draft preparation, A.P. and N.S.; writing—review and editing, T.M.; visualization, V.L. and N.S.; supervision, Y.G.; project administration, T.M.; funding acquisition, T.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by state funding allocated to the Pavlov Institute of Physiology, Russian Academy of Sciences (No. 1021062411782-5-3.1.8), as well as by Cosyma LLC fundings for the Pavlov Institute of Physiology, Russian Academy of Sciences' study on the theme of 'Searching for optimal methods for analysing movements recorded during transcutaneous electrical stimulation of the human spinal cord' (No. 2025-10-01) and of 'A comparative study of the effect of stimulation frequency during electrical transcutaneous spinal cord stimulation on locomotor function' (No. 2025-06-17).

Institutional Review Board Statement: Ethical approval was obtained from the Ethics Committee of the Pavlov Institute of Physiology, Russian Academy of Sciences (Minutes No. 25-07, dated 1 August 2025). The study was conducted in accordance with the Declaration of Helsinki (World Medical Association, 2013).

Informed Consent Statement: Prior to participation, all participants received detailed information about the study objectives and procedures and provided written informed consent.

Data Availability Statement: The data supporting the findings of this study are available upon request from the corresponding author.

Acknowledgments: We extend our gratitude to the research participants and Elena Prazdnikova for her assistance with participant recruitment and to Dr. Roman Malashin from Pavlov Institute of Physiology for critically reading the manuscript.

Conflicts of Interest: Yury Gerasimenko and Tatiana Moshonkina, who are researchers on the study team, hold a shareholder interest in Cosyma LLC (Moscow, Russia). Cosyma LLC produced the spinal cord stimulation device used in this study. They have certain inventorship rights to intellectual property licensed by Cosyma LLC. The remaining authors declare no conflicts of interest.

References

1. Shamantseva, N.; Moshonkina, T. On Carrier Frequency in Transcutaneous Spinal Cord Electrical Stimulation: A Narrative Review. *J. Neural Eng.* **2025**, *22*, 051002, doi:10.1088/1741-2552/ae08e8.
2. Manson, G.A.; Calvert, J.S.; Ling, J.; Tychon, B.; Ali, A.; Sayenko, D.G. The Relationship between Maximum Tolerance and Motor Activation during Transcutaneous Spinal Stimulation Is Unaffected by the Carrier Frequency or Vibration. *Physiological Reports* **2020**, *8*, e14397, doi:10.14814/phy2.14397.
3. Dalrymple, A.N.; Hooper, C.A.; Kuriakose, M.G.; Capogrosso, M.; Weber, D.J. Using a High-Frequency Carrier Does Not Improve Comfort of Transcutaneous Spinal Cord Stimulation. *J. Neural Eng.* **2023**, *20*, 016016, doi:10.1088/1741-2552/acabe8.
4. Packiasabapathy, S.; Rangasamy, V.; Sadhasivam, S. Pupillometry in Perioperative Medicine: A Narrative Review. *Can J Anesth/J Can Anesth* **2021**, *68*, 566–578, doi:10.1007/s12630-020-01905-z.
5. Ledowski, T. Objective Monitoring of Nociception: A Review of Current Commercial Solutions. *British Journal of Anaesthesia* **2019**, *123*, e312–e321, doi:10.1016/j.bja.2019.03.024.
6. Rajendran, G.; Mahalingam, S.; Ramkumar, A.; Krishnamoorthy, Y.; Kumares, P.T.; Vijayan, V.; Elanjaeran, R.; Kannan, R.; Prakasam, S.; Salih, A. Ultrasound for Pupillary Assessment - A Systematic Review. *JEM Reports* **2025**, *4*, 100158, doi:10.1016/j.jemrpt.2025.100158.
7. Selionov, V.A.; Ivanenko, Y.P.; Solopova, I.A.; Gurfinkel, V.S. Tonic Central and Sensory Stimuli Facilitate Involuntary Air-Stepping in Humans. *Journal of Neurophysiology* **2009**, *101*, 2847–2858, doi:10.1152/jn.90895.2008.
8. Gorodnichev, R.M.; Pivovarova, E.A.; Puhov, A.; Moiseev, S.A.; Gerasimenko, Y.P.; Savochin, A.A.; Moshonkina, T.R.; Chsherbakova, N.A.; Kilimnik, V.A.; Selionov, V.A.; et al. Transcutaneous Electrical Stimulation of the Spinal Cord: A Noninvasive Tool for the Activation of Stepping Pattern Generators in Humans. *Human Physiology* **2012**, *38*, 158–167, doi:10.1134/S0362119712020065.
9. Ivanenko, Y.P.; Gurfinkel, V.S.; Selionov, V.; Solopova, I.A.; Sylos-Labini, F.; Guertin, P.A.; Lacquaniti, F. Tonic and Rhythmic Spinal Activity Underlying Locomotion. *CPD* **2017**, *23*, 1753–1763, doi:10.2174/1381612823666170125152246.
10. Delgado, D.A.; Lambert, B.S.; Boutris, N.; McCulloch, P.C.; Robbins, A.B.; Moreno, M.R.; Harris, J.D. Validation of Digital Visual Analog Scale Pain Scoring With a Traditional Paper-Based Visual Analog Scale in Adults. *JAAOS Global Research & Reviews* **2018**, *2*, e088, doi:10.5435/JAAOSGlobal-D-17-00088.
11. Mathis, A.; Mamidanna, P.; Cury, K.M.; Abe, T.; Murthy, V.N.; Mathis, M.W.; Bethge, M. DeepLabCut: Markerless Pose Estimation of User-Defined Body Parts with Deep Learning. *Nat Neurosci* **2018**, *21*, 1281–1289, doi:10.1038/s41593-018-0209-y.
12. DeepLabCut DeepLabCut/Examples/COLAB/COLAB_YOURDATA_TrainNetwork_VideoAnalysis.Ipynb at Main · DeepLabCut/DeepLabCut Available online: https://github.com/DeepLabCut/DeepLabCut/blob/main/examples/COLAB/COLAB_YOURDATA_TrainNetwork_VideoAnalysis.ipynb (accessed on 15 February 2026).
13. DeepLabCut DeepLabCut/Docs/standardDeepLabCut_UserGuide.Md at Main · DeepLabCut/DeepLabCut Available online: https://github.com/DeepLabCut/DeepLabCut/blob/main/docs/standardDeepLabCut_UserGuide.md (accessed on 28 February 2026).
14. Bland, J.M.; Altman, D.G. Statistical Methods for Assessing Agreement between Two Methods of Clinical Measurement. *International Journal of Nursing Studies* **2010**, *47*, 931–936, doi:10.1016/j.ijnurstu.2009.10.001.
15. Yang, C.; Veit, N.C.; McKenzie, K.A.; Aalla, S.; Kishta, A.; Embry, K.; Roth, E.J.; Lieber, R.L.; Jayaraman, A. The Effects of Stimulation Waveform and Carrier Frequency on Tolerance and Motor Thresholds Elicited by Transcutaneous Spinal Cord Stimulation in Stroke. *Clinical Neurophysiology Practice* **2025**, *10*, 150–158, doi:10.1016/j.cnp.2025.04.001.
16. Privitera, M.; Ferrari, K.D.; von Ziegler, L.M.; Sturman, O.; Duss, S.N.; Floriou-Servou, A.; Germain, P.-L.; Vermeiren, Y.; Wyss, M.T.; De Deyn, P.P.; et al. A Complete Pupillometry Toolbox for Real-Time Monitoring of Locus Coeruleus Activity in Rodents. *Nat Protoc* **2020**, *15*, 2301–2320, doi:10.1038/s41596-020-0324-6.

17. Phatak, A.; Mandal, A.C.; Balaji, J.J.; Lakshminarayanan, V. Direct Estimation of Pupil Parameters Using Deep Learning for Visible Light Pupillometry 2023.
18. Saxena, S.; Fink, L.K.; Lange, E.B. Deep Learning Models for Webcam Eye Tracking in Online Experiments. *Behav Res* **2024**, *56*, 3487–3503, doi:10.3758/s13428-023-02190-6.
19. Schroeder-Lab/EyeVideoAnalysis 2025.
20. Vázquez Romaguera, T.; Vázquez Romaguera, L.; Castro Piñol, D.; Vázquez Seisdedos, C.R.; Vázquez Romaguera, T.; Vázquez Romaguera, L.; Castro Piñol, D.; Vázquez Seisdedos, C.R. Pupil Center Detection Approaches: A Comparative Analysis. *Computación y Sistemas* **2021**, *25*, 67–81, doi:10.13053/cys-25-1-3385.
21. Mathis, A.; Mamidanna, P.; Cury, K.M.; Abe, T.; Murthy, V.N.; Mathis, M.W.; Bethge, M. DeepLabCut: Markerless Pose Estimation of User-Defined Body Parts with Deep Learning. *Nat Neurosci* **2018**, *21*, 1281–1289, doi:10.1038/s41593-018-0209-y.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.