

BULETINUL INSTITUTULUI POLITEHNIC DIN IAȘI
Publicat de
Universitatea Tehnică „Gheorghe Asachi” din Iași
Volumul 70 (74), Numărul 2, 2024
Secția
ELECTROTEHNICĂ. ENERGETICĂ. ELECTRONICĂ
DOI: 10.2478/bipic-2024-0008

PUPIL DETECTION SYSTEM IMPLEMENTATION FOR LOW-RESOLUTION EYE IMAGES BASED ON A FULLY-CONNECTED NEURAL NETWORK CLASSIFIER

BY

**GABRIEL BONTEANU^{1,*}, PETRONELA BONTEANU², NICOLAE PATACHE¹,
ARCADIE CRACAN¹ and RADU GABRIEL BOZOMITU²**

¹“Gheorghe Asachi” Technical University of Iași, Romania,
Fundamentals of Electronics Department

²“Gheorghe Asachi” Technical University of Iași, Romania,
Telecommunication and IT Department

Received: April 18, 2024

Accepted for publication: July 8, 2025

Abstract. This paper presents an implementation of a pupil center detection system that uses a classifier based on neural networks. Being intended for the processing of low-resolution eye images, the implementation is done with a fully connected neural network. The classes of the classifier are defined according to the possible positions of the pupil center. The neural network was trained, tested and validated using 39,000 eye images from 18 databases, images obtained from different human subjects and under variable and non-uniform lighting conditions. The solution is distinguished by the fact that it directly provides the coordinates of the center of the pupil and not just the area where it is located. The results show both high accuracy and high processing speed, the proposed solution being suitable for low-budget implementations of real-time gaze detection systems.

Keywords: eye tracking, neural network, gaze detection, machine learning, pupil detection.

*Corresponding author; *e-mail*: gbonteanu@etti.tuiasi.ro

© 2024 Gabriel Bonteanu et al.

This is an open access article licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (CC BY-NC-ND 4.0).

1. Introduction

In our digital age, where the world revolves around communication, technology has become the bridge that connects us to a vast universe of information and opportunities. Human-computer interaction is now more intimate, faster and more complex than ever before, and the efficiency of this communication is a key element in the progress of our society. In this context, the development of an eye tracking system, capable of accurately tracking and interpreting human eye movements, is revealed as an essential turning point for the creation of an effective human-computer interface.

Our eyes are not only mirrors of our inner states, but also powerful communication tools. While written and spoken language remain fundamental means of communication, many of our deepest meanings and subtle intentions are expressed through our movements and the direction of our gaze. By tracking and interpreting these signals, an eye tracking system can reveal human needs, intentions, and reactions in a way never recorded before. This deep understanding of human behaviour holds the promise of revolutionizing the human-computer interface, facilitating more natural, efficient and personalized communication.

This work focuses on exploring eye tracking technologies and developing an advanced system that opens the door to a world where our devices can respond to our desires by simply directing our gaze. Through research and innovation, we aspire to make this vision a reality, ushering in a new era of human-computer interaction. Through the detailed study of eye movements and their deep meanings, this work contributes to the development of a human-computer interface that is more responsive, intuitive and adaptable than ever before. Thus, it offers a provocative perspective on how eye tracking technology can shape the future of human communication and open new horizons in the development of devices and human-computer interaction.

Solutions for gaze detection based on both classical image processing and artificial intelligence techniques have been reported in the literature. Among the most remarkable achievements are the Starburst algorithm (Li *et al.*, 2005), the algorithm based on projections (Xu *et al.*, 2005), the algorithm based on contour analysis (Wang *et al.*, 2004), the one based on the method of integral projections (Zhou *et al.*, 2004), the one based on the Gaussian model (Kociejko *et al.*, 2008), algorithms based on circular Hough transforms (Cherabit *et al.*, 2012) and elliptic (Bozomitu *et al.*, 2019), algorithms based on fitting an ellipse by the method of least squares (Fitzgibbon *et al.*, 1996) but also complex algorithms with very good performance in real conditions such as ExCuSe (Fuhl *et al.*, 2015), ElSe (Fuhl *et al.*, 2016) and PuReST (Santini *et al.*, 2018). But the major disadvantage of solutions based on classic image processing is the fact that their performance drops dramatically in non-uniform conditions and variable lighting, as is characteristic of real-time applications.

Recently, with the development of machine learning techniques, the implementation of gaze tracking systems based on artificial intelligence has become increasingly preferred. For example, the deep neural network technique has demonstrated remarkable performance in understanding and interpreting human behavior (Klaib *et al.*, 2021). Solutions such as PupilNet v2.0 (Fuhl *et al.*, 2017) combines convolutional neural network architectures and have additional refinement methods to increase detection accuracy. While most implementations using AI try to detect only the area where the pupil is from a limited number of possibilities (Xiong *et al.*, 2024), this paper proposes a technique by which the center of the pupil is detected directly. A suitable solution for low-resolution images, 20px x 15px, which lends itself to be implemented with fully connected neurons, is discussed here, while a more elaborate solution, suitable for high-resolution images, was published in (Bonteanu *et al.*, 2024).

The proposed method is based on a classifier implemented with neural networks, a technology that has proven its usefulness in many applications: establishing medical diagnoses (Dampage *et al.*, 2021), natural language recognition (De Coster *et al.*, 2020), fraud detection (Thennakoon *et al.*, 2019), human emotion interpretation (Liu *et al.*, 2022), self-driving assistance (Bojarski *et al.*, 2016), image classification (Krizhevsky *et al.*, 2017) and signal segmentation (Farshad *et al.*, 2021).

2. The proposed pupil center detection system

An application is further assumed in which a remote-type device captures images of the human subject, and following an eye detection stage, low-resolution eye images result, such as 20 pixels by 15 pixels. The implementation of the classifier on such low-resolution images will be exemplified for simplicity. Thus, the justification for the low resolution of the processed images is given by a possible application in which a low-cost camera captures images of the human subject and an eye recognition algorithm segments the image of the eye. The work does not focus on these aspects but on the further processing of the low-resolution image of the eye.

The images used in training, testing and validating the proposed solution come from 2 different sources: 600 eye images used by the authors of the Swirski algorithm (Swirski *et al.*, 2012) as well as 38,401 eye images (DSI-DSXVII) used in testing the ExCuSe algorithm (Fuhl *et al.*, 2015). All the images were labelled manually, so that in the database the image is accompanied by the coordinates of the pupil center. Regardless of their origin and respect to their native resolution, the eye images have been rescaled and cropped so that they are brought to a resolution of 20 pixels horizontally and 15 pixels vertically, as Fig. 1 illustrates. This was the pre-processing step.

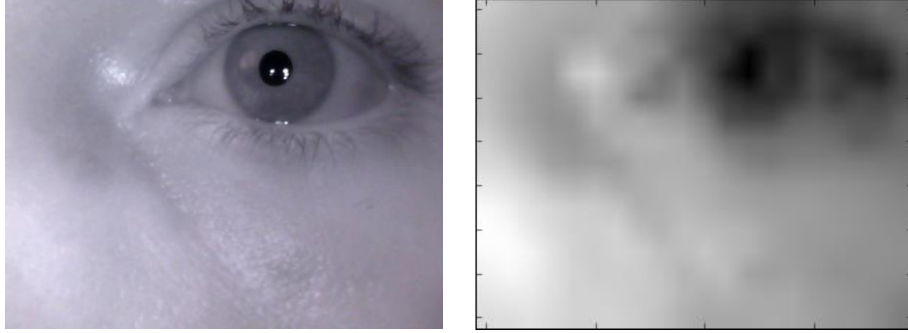


Fig. 1 – Eye image: original versus down-scaled.

The classes of the classifier were defined according to the possible positions of the pupil center. It was considered that the pupil can be in any of the 300 pixels of the image, thus resulting in 300 possible classes. Some of these classes are exemplified in Fig. 2, where the desired states at the output of the classifier are illustrated. For example, if the center of the pupil is found in the first pixel of the image, line 1 and column 1, then the neuron with index 1 will be activated on the output layer of the classifier. If the center of the pupil is found on line 1 and column 2 of the image, then the neuron with index 2 on the output layer will be activated. If the center of the pupil is in the center of the image, the neuron with index 150 in the output layer will be activated. The last class corresponds to those images in which the center of the pupil is found on the last line and the last column, so in the lower right corner of the image, the desired behaviour of the classifier being that of activating the last neuron on the last layer, so the one with index 300.

In general, the index of the neuron on the last layer that will be activated shall depend on the x and y coordinates of the pupil center according to relation (1).

$$n = x + 20 \cdot y \quad (1)$$

The implementation of the classifier was done in python with ReLU type neurons for the first 3 layers and SOFTMAX neurons for the output layer. The architecture of the network is illustrated in Fig. 3. The input layer encodes the normalized grayscale intensities of the pixels in the image to be processed. Given that the number of classes is equal to the number of pixels of the image, 300 neurons were provided on the output layer. For the output neuron layer, the SOFTMAX type was preferred because it transforms the activations into a distribution of probabilities, which are very easy to interpret as class probabilities. The architecture also includes two hidden layers of 100 neurons each.

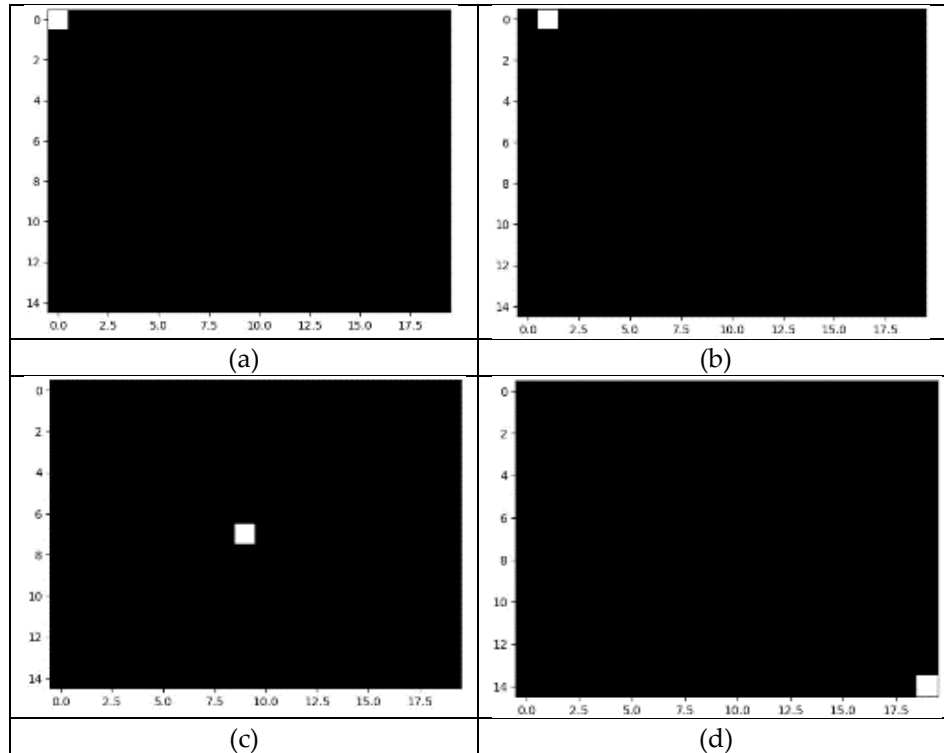


Fig. 2 – Graphical representation of classifier outputs for different classes. The 300 neurons were organized in the form of a matrix of similar size as the input image. The activated pixel is white, the deactivated pixels are black. There are represented: (a) Class 1, (b) Class 2, (c) Class 150 and (d) Class 300.

To calculate the total number of parameters (weights and biases) of a neural network with the specified architecture, we need to consider the connections between neurons and each parameter associated with them:

- **Layer 1** (input): This layer has no parameters of its own, as it only represents the input data, so it does not add any parameters.
- **Layer 2** (fully connected with 100 neurons): For each neuron in this layer, we must assign the weights for all 300 neurons in the previous layer, plus a bias for each neuron in this layer. So, for Layer 2, we have:
Weights: $300 \cdot 100 = 30,000$ parameters
Biases: 100 parameters
- **Layer 3** (fully connected with 100 neurons): As for Layer 2, for each neuron in this layer, we must assign the weights for all 100 neurons in the previous layer, plus a bi-as for each neuron in this layer. So, for Layer 3, we have:
Weights: $100 \cdot 100 = 10,000$ parameters
Biases: 100 parameters

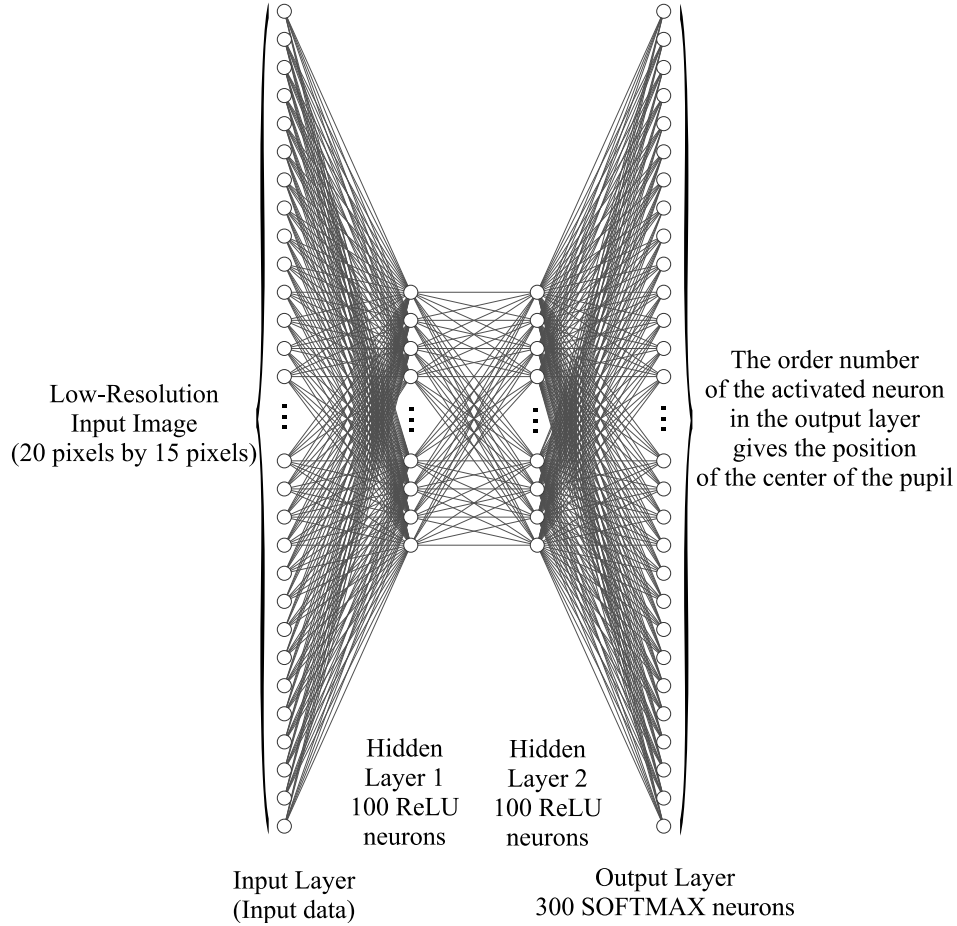


Fig. 3 – Architecture of the proposed neural network classifier.

• **Output layer** (fully connected with 300 neurons): This is the final layer, which generates the output. As in the previous cases, for each neuron in this layer, we must assign the weights for all 100 neurons in the previous layer, plus a bias for each neuron in this layer. So, for the output layer, we have:

Weights: $100 \cdot 300 = 30,000$ parameters

Biases: 300 parameters

In order to calculate the total number of network parameters, all the parameters from each layer shall be summed:

Total Parameters = Layer 2 Weights + Layer 2 Biases + Layer 3 Weights + Layer 3 Biases + Output Layer Weights + Output Layer Biases

Total parameters = $30,000 + 100 + 10,000 + 100 + 30,000 + 300 = 70,500$ parameters. Thus, the specified neural network has a total of **70,500** parameters (weights and biases).

The relationship between the amount of training data and the number of parameters of a neural network is essential to achieve good performance and to avoid problems such as overfitting or underfitting. In the case of a small amount of training data and a neural network with a large number of parameters, there is an increased risk of overtraining. The network might learn the training data in detail and not generalize correctly to new, unseen data. To avoid overtraining in this situation, the regularization technique was applied.

The training of the neural network was done using Stochastic Gradient Descent. The available data were organized as follows: 80% of the eye images constituted the training set, 10% the validation set and the remaining 10% the test set. The hyperparameter adjustment strategy was defined by a regularization parameter $\lambda = 0.5$, a minibatch size = 10, a learning rate of $\eta = 0.1$ for epochs 1-20, for epochs 21-40 a learning rate of $\eta = 0.01$, $\eta = 0.001$ for the next 20 epochs and $\eta = 0.0001$ for the last 20 epochs. Figure 4 shows how the accuracy of testing and validation evolved according to the epoch index.

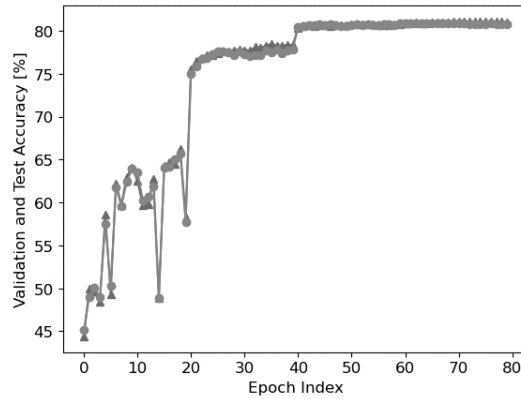


Fig. 4 – Accuracy on validation data (▲) and test data (●) during the training.

In order to estimate the accuracy of the proposed system, it was resorted to the inference of each of the approximately 39,000 images from the 18 databases through the neural network. The results obtained on each of the databases are presented below. The accuracy of the detection was estimated by the Euclidean distance between the manually estimated pupil center and the automatically detected one:

$$ED = |c_{ideal} - c_{detected}| \quad (2)$$

as well as by the detection rate, defined as the ratio between the number of images detected with a Euclidean error below a certain threshold (k) and the total number of images in the respective database:

$$DR_k = \frac{N_{k\text{ px}}}{N_{total}} \cdot 100 (\%) \quad (3)$$

The DSI database (Fuhl *et al.*, 2015) includes 6,554 eye images that contain reflections. The inference results illustrated in Fig. 5 indicate a detection rate at 0 pixels of $DR_0 = 81.85\%$ and a detection rate at 2 pixels of $DR_2 = 99.86\%$. The standard deviation of the Euclidean distance is $\sigma = 0.41$ pixels.

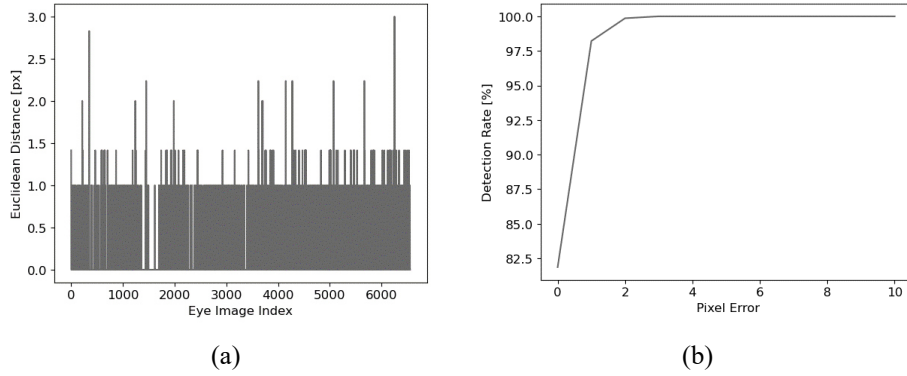


Fig. 5 – (a) Euclidean distance for each of the eye images in DSI. (b) Detection Rates for all the eye images in DSI.

The DSII database (Fuhl *et al.*, 2015) includes 505 eye images captured in bad illumination conditions and including reflections. The inference results shown in Fig. 6 indicate $DR_0 = 82.17\%$, $DR_2 = 99.60\%$ and $\sigma = 0.44$ pixels.

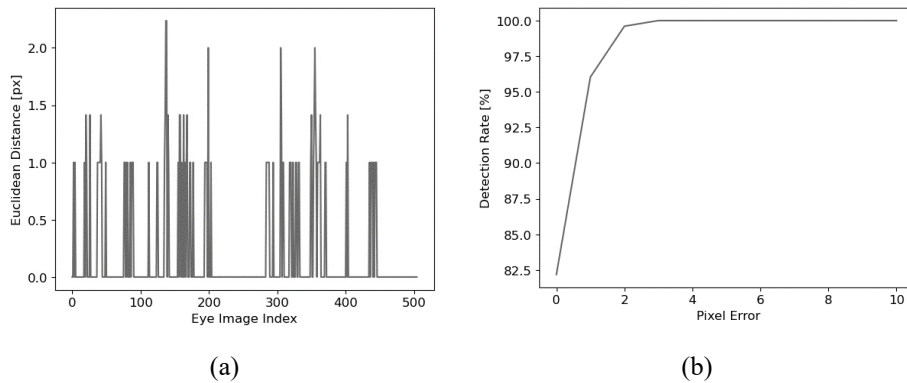


Fig. 6 – (a) Euclidean distance for each of the eye images in DSII. (b) Detection Rates for all the eye images in DSII.

The DSIII database (Fuhl *et al.*, 2015) includes 9,799 eye images captured in bad illumination conditions and including reflections and recording errors. The inference results in Fig. 7 indicate $DR_0 = 86.60\%$, $DR_2 = 99.83\%$ and $\sigma = 0.40$ pixels.

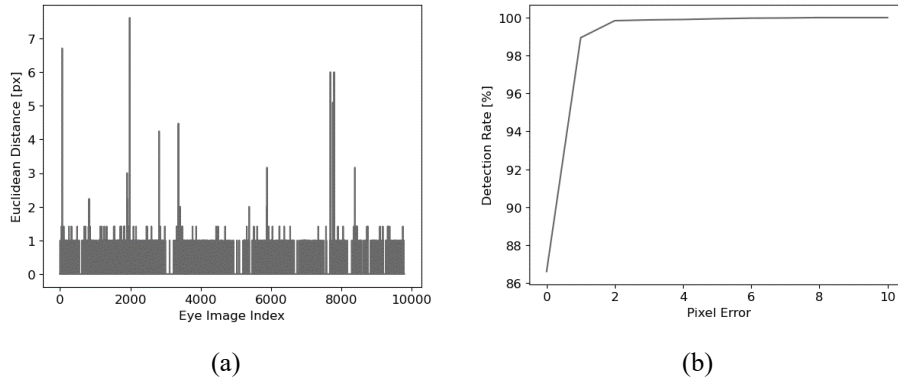


Fig. 7 – (a) Euclidean distance for each of the eye images in DSIII. (b) Detection Rates for all the eye images in DSIII.

The DSIV database (Fuhl *et al.*, 2015) includes 2,655 eye images of human subjects wearing contact lenses captured in bad illumination conditions. An examination of Fig. 8 inference results indicates $DR_0 = 88.51\%$, $DR_2 = 99.88\%$ and $\sigma = 0.34$ pixels.

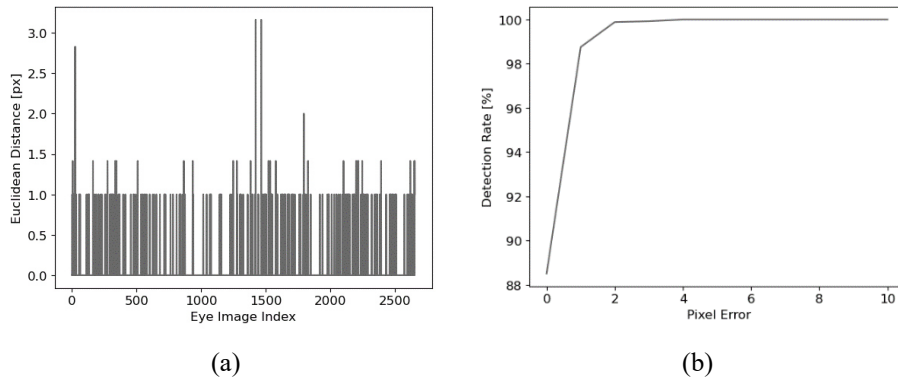


Fig. 8 – (a) Euclidean distance for each of the eye images in DSIV. (b) Detection Rates for all the eye images in DSIV.

The DSV database (Fuhl *et al.*, 2015) includes 2,135 eye images with shifted contact lenses. Figure 9 inference results further supports the observation that $DR_0 = 86.79\%$, $DR_2 = 99.90\%$ and $\sigma = 0.35$ pixels.

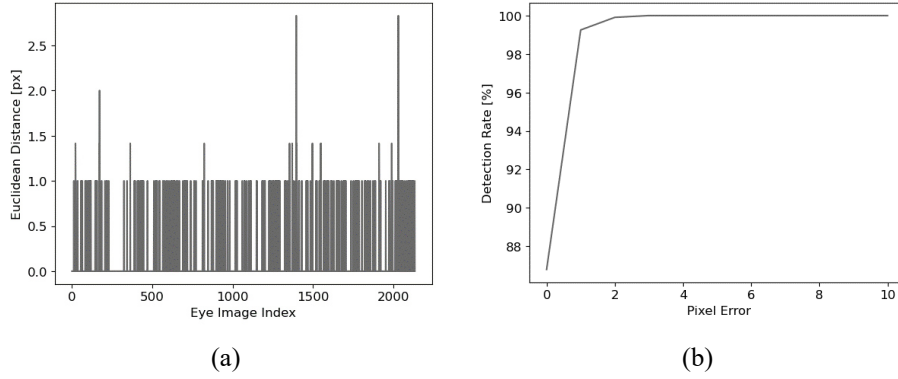


Fig. 9 – (a) Euclidean distance for each of the eye images in DSV. (b) Detection Rates for all the eye images in DSV.

The DSVI database (Fuhl *et al.*, 2015) includes 4,400 eye images of human subjects wearing mascara captured in bad illumination conditions. Inference results depicted in Fig. 10 indicate $DR_0 = 88.00\%$, $DR_2 = 99.81\%$ and $\sigma = 0.37$ pixels.

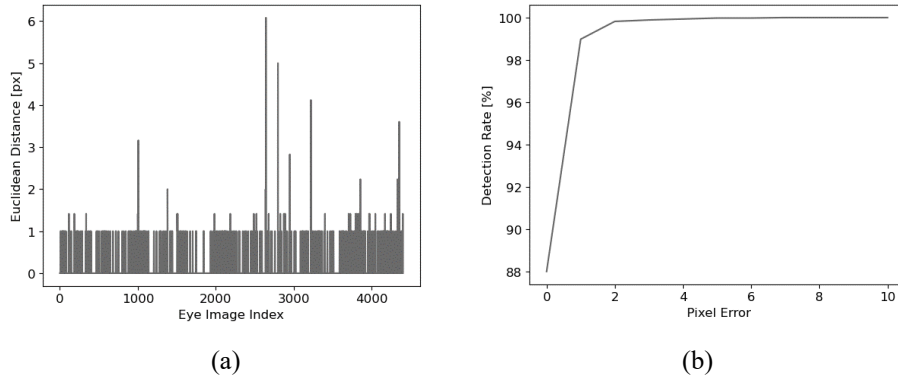


Fig. 10 – (a) Euclidean distance for each of the eye images in DSVI. (b) Detection Rates for all the eye images in DSVI.

The DSVII database (Fuhl *et al.*, 2015) includes 4,890 eye images of human subjects wearing mascara captured in bad illumination conditions and with eye shadow. Inference results shown in Fig. 11 indicate $DR_0 = 83.47\%$, $DR_2 = 99.73\%$ and $\sigma = 0.43$ pixels.

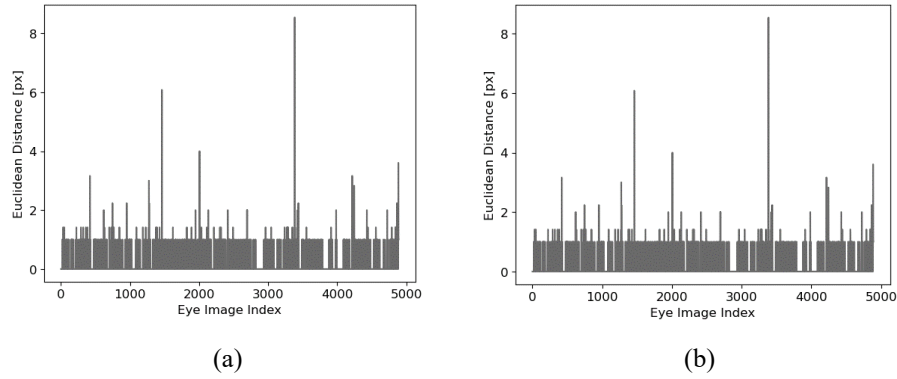


Fig. 11 – (a) Euclidean distance for each of the eye images in DSVII. (b) Detection Rates for all the eye images in DSVII.

The DSVIII (Fuhl *et al.*, 2015) database includes 630 eye images of human subjects wearing mascara captured in bad illumination conditions. Inference results observed in Fig. 12 indicate $DR_0 = 81.58\%$, $DR_2 = 93.80\%$ and $\sigma = 1.08$ pixels.

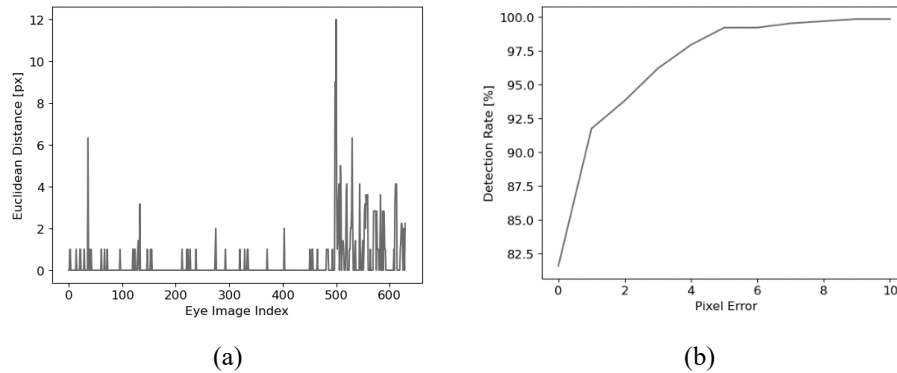


Fig. 12 – (a) Euclidean distance for each of the eye images in DSVIII. (b) Detection Rates for all the eye images in DSVIII.

The DSIX database (Fuhl *et al.*, 2015) includes 2,831 eye images captured in bad illumination conditions with additional black dot. The corresponding inference results are provided in Fig. 13 and indicate $DR_0 = 83.43\%$, $DR_2 = 99.61\%$ and $\sigma = 0.44$ pixels.

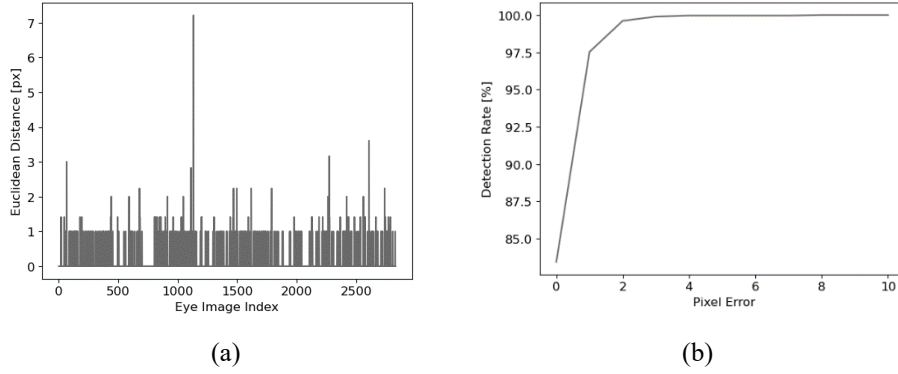


Fig. 13 – (a) Euclidean distance for each of the eye images in DSIX. (b) Detection Rates for all the eye images in DSIX.

The DSX database (Fuhl *et al.*, 2015) includes 2,831 eye images captured in bad illumination conditions with pupil at image border. Figure 14 offers a visual representation of inference results and indicates $DR_0 = 86.78\%$, $DR_2 = 99.87\%$ and $\sigma = 0.47$ pixels.

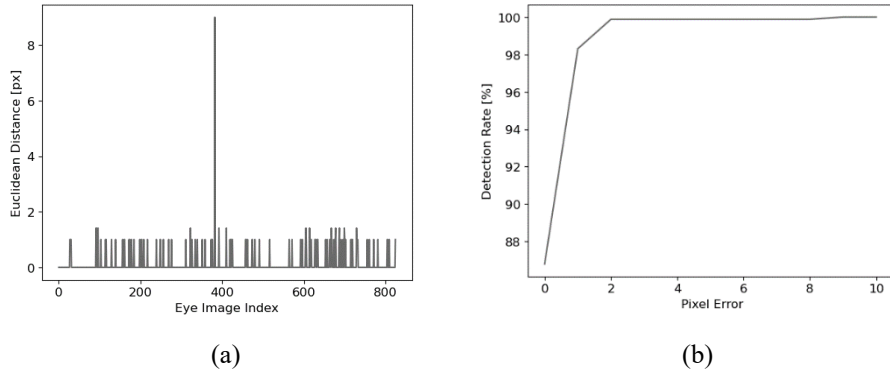


Fig. 14 – (a) Euclidean distance for each of the eye images in DSX. (b) Detection Rates for all the eye images in DSX.

The DSXI database (Fuhl *et al.*, 2015) includes 655 eye images captured in bad illumination conditions with additional black dots and reflections. Inference results in Fig. 15 indicate $DR_0 = 88.24\%$, $DR_2 = 100\%$ and $\sigma = 0.33$ pixels.

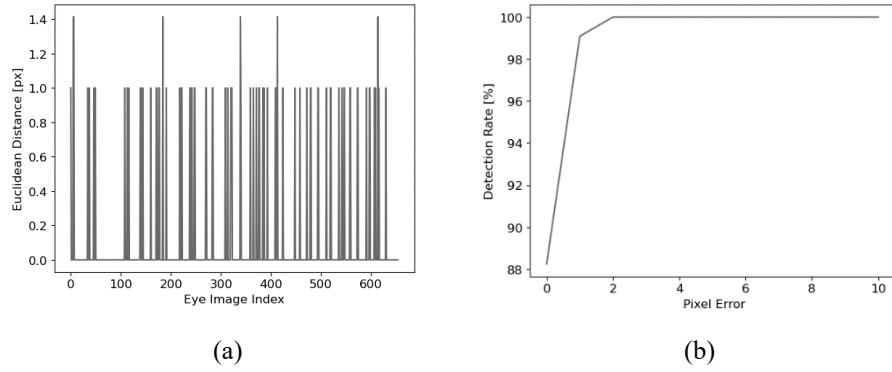


Fig. 15 – (a) Euclidean distance for each of the eye images in DSXI. (b) Detection Rates for all the eye images in DSXI.

The DSXII database (Fuhl *et al.*, 2015) includes 524 eye images captured in bad illumination conditions. The Fig. 16 inference results indicate $DR_0 = 82.44\%$, $DR_2 = 100\%$ and $\sigma = 0.41$ pixels.

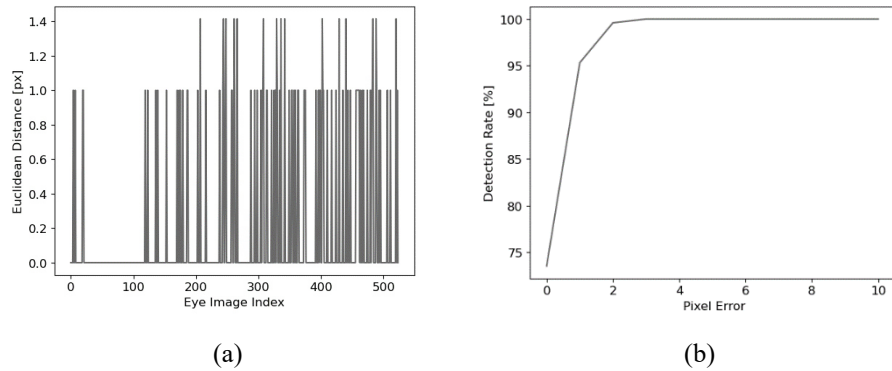


Fig. 16 – (a) Euclidean distance for each of the eye images in DSXII. (b) Detection Rates for all the eye images in DSXII.

The DSXIII database (Fuhl *et al.*, 2015) includes 491 eye images captured in bad illumination conditions and including eyelashes. Inference results in Fig. 17 indicate $DR_0 = 73.52\%$, $DR_2 = 99.59\%$ and $\sigma = 0.49$ pixels.

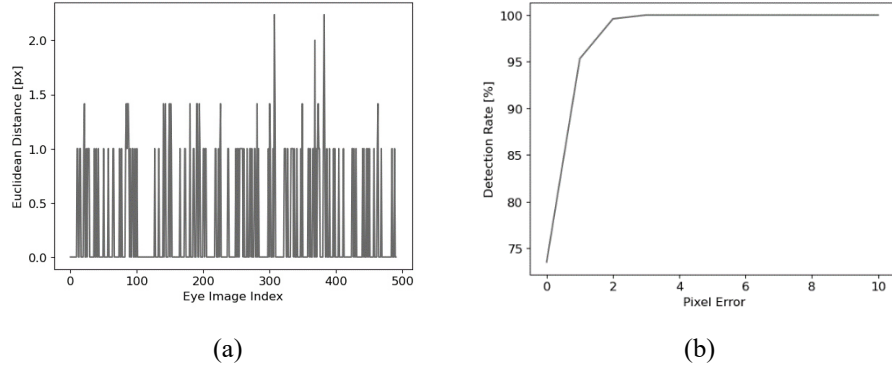


Fig. 17 – (a) Euclidean distance for each of the eye images in DSXIII. (b) Detection Rates for all the eye images in DSXIII.

The DSXIV database (Fuhl *et al.*, 2015) includes 469 eye images captured in bad illumination conditions. Inference results displayed in Fig. 18 indicate $DR_0 = 78.25\%$, $DR_2 = 99.78\%$ and $\sigma = 0.48$ pixels.

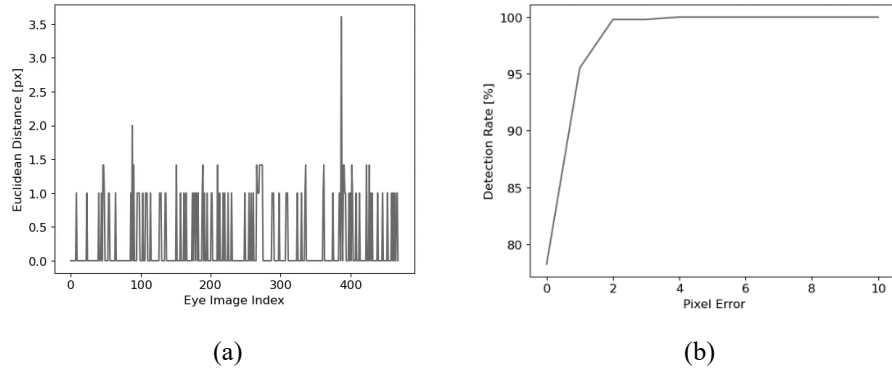


Fig. 18 – (a) Euclidean distance for each of the eye images in DSXIV. (b) Detection Rates for all the eye images in DSXIV.

The DSXV database (Fuhl *et al.*, 2015) includes 363 eye images with shifted contact lenses. Inference results in Fig. 19 indicate $DR_0 = 82.64\%$, $DR_2 = 98.89\%$ and $\sigma = 0.45$ pixels.

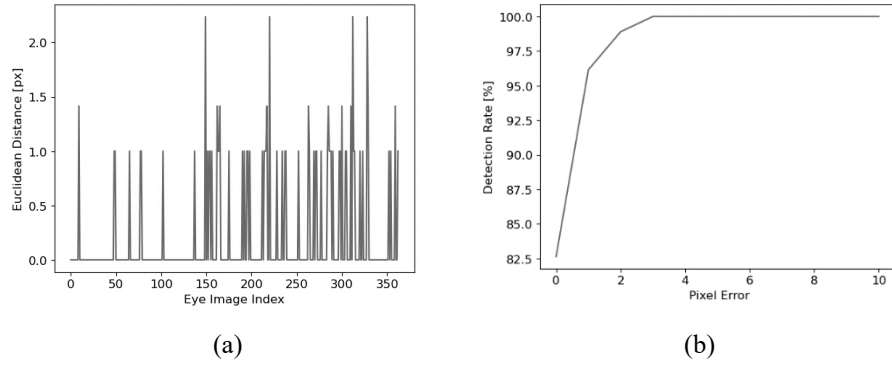


Fig. 19 – (a) Euclidean distance for each of the eye images in DSXV. (b) Detection Rates for all the eye images in DSXV.

The DSXVI (Fuhl *et al.*, 2015) database includes 392 eye images of human subjects wearing mascara and including eyelashes. An examination of Fig. 20 inference results indicates $DR_0 = 77.55\%$, $DR_2 = 100\%$ and $\sigma = 0.44$ pixels.

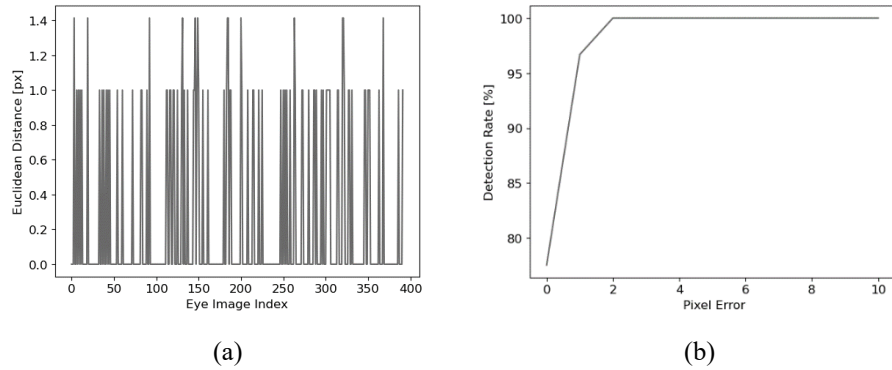


Fig. 20 – (a) Euclidean distance for each of the eye images in DSXVI. (b) Detection Rates for all the eye images in DSXVI.

The DSXVII database (Fuhl *et al.*, 2015) includes 268 eye images captured in bad illumination conditions and including eyelashes. Inference results in Fig. 21 indicate $DR_0 = 84.70\%$, $DR_2 = 100\%$ and $\sigma = 0.39$ pixels.

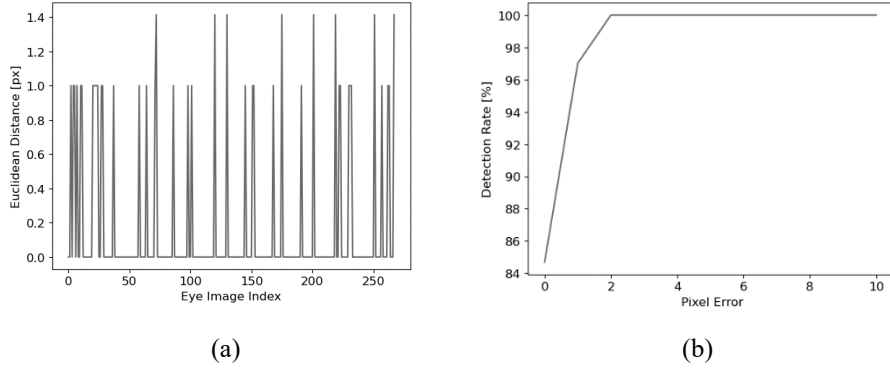


Fig. 21 – (a) Euclidean distance for each of the eye images in DSXVII. (b) Detection Rates for all the eye images in DSXVII.

The Swirski database (Swirski *et al.*, 2012) includes 600 highly off axis eye images and including eyelashes. Inference results shown in Fig. 22 indicate $DR_0 = 89.83\%$, $DR_2 = 99.5\%$ and $\sigma = 0.39$ pixels.

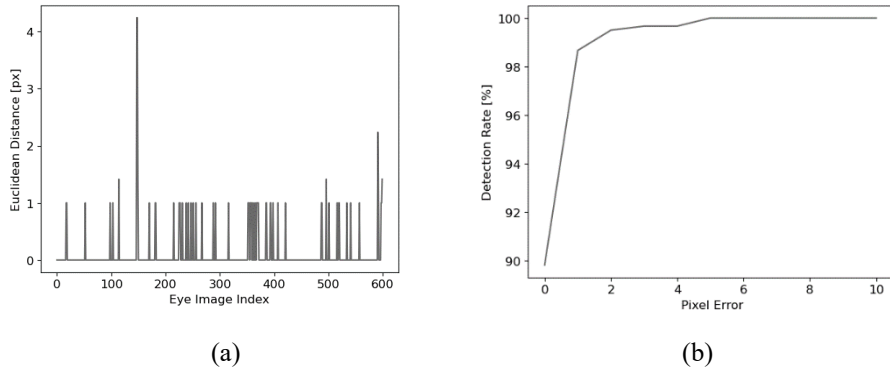


Fig. 22 – (a) Euclidean distance for each of the eye images in Swirski database. (b) Detection Rates for all the eye images in Swirski DB.

The Euclidean distance for each of the eye images in the 18 databases and the Detection Rate for all the eye images in the 18 databases are represented in Fig. 23. The inference results show $DR_0 = 84.96\%$, $DR_2 = 99.70\%$ and $\sigma = 0.42$ pixels.

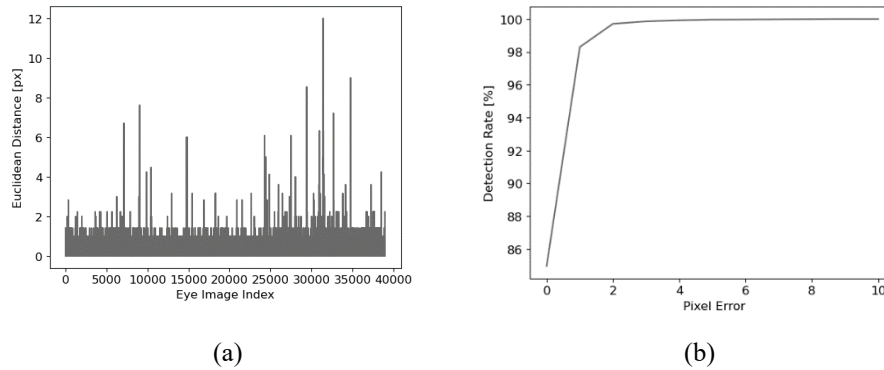


Fig. 23 – (a) Euclidean distance for each of the eye images in all databases. (b) Detection Rates for all the eye images in all databases.

The inference results are also presented in Table 1. The observed processing speed is higher than 100 frames per second.

Table 1
Inference results

Data base	Frames number	DR ₀	DR ₂	σ	Description
DSI	6,554	81.85	99.86	0.41	Reflections
DSII	505	82.17	99.60	0.44	Reflections, bad illumination
DSIII	9,799	86.60	99.83	0.40	Reflections, recording errors, bad illumination
DSIV	2,655	88.51	99.88	0.34	Contact lenses, bad illumination
DSV	2,135	86.79	99.90	0.35	Shifted contact lenses
DSVI	4,400	88.00	99.81	0.37	Bad illumination, mascara
DSVII	4,890	83.47	99.73	0.43	Bad illumination, mascara, eyeshadow
DSVIII	630	81.58	93.80	1.08	Bad illumination, eyelashes
DSIX	2,831	83.43	99.61	0.44	Reflections, additional black dot
DSX	840	86.78	99.87	0.47	Bad illumination, pupil at image boarder
DSXI	655	88.24	100	0.33	Reflections, bad illumination, additional black dot
DSXII	524	82.44	100	0.41	Bad illumination
DSXIII	491	73.52	99.59	0.49	Bad illumination, eyelashes
DSXIV	469	78.25	99.78	0.48	Bad illumination
DSXV	363	82.64	98.89	0.45	Shifted contact lenses
DSXVI	392	77.55	100	0.44	Mascara, eyelashes
DSXVII	268	84.70	100	0.39	Bad illumination, eyelashes
SWIRSKI	600	89.83	99.5	0.39	Highly off axis, eyelashes
ALL	39,001	84.96	99.70	0.42	All databases

The results of the feedforward process of the 39,000 images through the described neural network indicate encouraging results in terms of the detection rate at 0 pixels, with values between 73.52% and 89.93% for each of the considered databases, the average weighted value being 84.96%. Regarding the detection rate at 2 pixels, the values obtained for each of the 18 databases are higher than 93.80%, with a weighted average of 99.70%, indicating a sufficient accuracy for a possible rough control of the cursor in a low budget human machine interface. The low values of the dispersion of the Euclidean distance indicate a good stability of the controlled cursor in such an application.

The fact that the images used in training, testing and validation come from different sources, were captured on different human subjects in variable lighting conditions, including different obstructing elements and reflections, ensures a good reliability of the proposed system.

3. Conclusions

In this paper, a pupil detection system that uses a neural network-based classifier has been presented. The solution is suitable for low-resolution eye images, the implementation being done with a fully connected neural network that was trained, tested and validated using 39,000 eye images from 18 databases. The eye images are obtained from different human subjects and under variable and non-uniform lighting conditions. The solution is distinguished by the fact that it directly provides the coordinates of the pupil center and not just the area where it is located. The results show both high accuracy and high processing speed, the proposed solution being suitable for low-budget implementations of real-time gaze detection systems.

REFERENCES

- Bojarski M., Del Testa D., Dworakowski D., Firner B., Flepp B., Goyal P., Jackel L.D., Monfort M., Muller U., Zhang J., Zhang X., *End to End Learning for Self-Driving Cars*, In Proceedings of the International Conference on Robotics and Auto-mation, 15(4), 2016; pp. 123-137, doi.org/10.48550/arXiv.1604.07316.
- Bonteanu G., Bonteanu P., Cracan A., Bozomitu R.G., *Implementation of a High-Accuracy Neural Network-Based Pupil Detection System for Real-Time and Real-World Applications*, Sensors 2024, 24, 2548, <https://doi.org/10.3390/s24082548>.
- Bozomitu R.G., Păsărică A., Tărniceriu D., Rotariu C., *Development of an Eye Tracking-Based Human-Computer Interface for Real-Time Applications*, Sensors 2019, Vol. 19, Issue 16, Article Number 3630.
- Cherabit N., Djeradi A., Chelali F.Z., *Circular Hough transform for iris localization*, Science and Technology 2012, vol. 2, no. 5, pp. 114-121.

- Dampage U., Balasuriya C., Thilakarathna S., Rathnayaka D., Kalubowila L., *AI-Based Heart Monitoring System*, In Proceedings of the IEEE 4th International Conference on Computing, Power and Communication Technologies (GUCON), Kuala Lumpur, Malaysia, 2021, pp. 1-6, 10.1109/GUCON50781.2021.9573888.
- De Coster M., Van Herreweghe M., Dambre J., *Sign language recognition with transformer networks*, In Proceedings of the 12th international conference on language resources and evaluation (LREC 2020), 2020, pp. 6018-6024.
- Farshad M., Karimi M., *A Signal Segmentation Approach to Identify Incident/Reflected Traveling Waves for Fault Location in Half-Bridge MMC-HVdc Grids*, IEEE Transactions on Instrumentation and Measurement 2021, vol. 71, pp. 1-9, 10.1109/TIM.2021.3139688.
- Fitzgibbon A.W., Pilu M., Fisher R.B., *Direct least squares fitting of ellipses*, In Proceedings of 13th International Conference on Pattern Recognition, Vienna, Austria, 25-29 August 1996, Volume 1, pp. 253-257.
- Fuhl W., Kübler T., Sippel K., Rosenstiel W., Kasneci E., *Excuse: Robust pupil detection in real-world scenarios*, In Proceedings of the 16th International Conference on Computer Analysis of Images and Patterns, Valletta, Malta, 2-4 September 2015, pp. 39-51.
- Fuhl W., Santini T., Kasneci G., Rosenstiel W., Kasneci E., *PupilNet v2.0: Convolutional Neural Networks for CPU based real time Robust Pupil Detection*, arXiv preprint arXiv:1711.00112, 2017.
- Fuhl W., Santini T.C., Kübler T., Kasneci E., *Else: Ellipse selection for robust pupil detection in real-world environments*, In Proceedings of the 9th Biennial ACM Symposium on Eye Tracking Research & Applications, Charleston, SC, USA, 14-17 March 2016, pp. 123-130.
- Klaib A., Alsrchin N., Melhem W., Bashtawi H., Magableh A., *Eye tracking algorithms, techniques, tools, and applications with an emphasis on machine learning and Internet of Things technologies*, Expert Systems with Applications 2021, Vol. 166, doi.org/10.1016/j.eswa.2020.114037.
- Kocejko T., Bujnowski A., Wtorek J., *Eye mouse for disabled*, In 2008 Conference on Human System Interactions, 2008, pp. 199-202.
- Krizhevsky A., Sutskever I., Hinton G.E., *ImageNet Classification with Deep Convolutional Neural Networks*, Communications of the ACM 2017, Vol. 60, no. 6, pp. 84-90, doi.org/10.1145/3065386.
- Li D., Parkhurst J.D., *Starburst: A robust algorithm for video-based eye tracking*, In Proceedings of the IEEE Vision for Human-Computer Interaction Workshop, Iowa State University, Ames, IA, USA, 5 September 2005.
- Liu B., *Sentiment Analysis and Opinion Mining*, Springer Cham, Springer Nature Switzerland AG 2012, 2022, doi.org/10.1007/978-3-031-02145-9.
- Santini T., Fuhl W., Kasneci E., *PuReST: Robust pupil tracking for real-time pervasive eye tracking*, In Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications, Warsaw, Poland, 14-17 June 2018, p. 61.
- Swirski L., Bulling A., Dodgson N.A., *Robust Real-Time Pupil Tracking in Highly off-Axis Images*, In Proceedings of the Symposium on Eye Tracking Research and Applications, Santa Barbara, CA, USA, 28-30 March 2012, pp. 173-176.

- Thennakoon A., Bhagyan C., Premadasa S., Mihiranga S., Kuruwitaarachchi N., *Real-time Credit Card Fraud Detection Using Machine Learning*, In Proceedings of the 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence), 2019, pp. 488-493, 10.1109/CONFLUENCE.2019.8776942.
- Wang Q., Yang J., *Eye detection in facial images with unconstrained background*, Journal of pattern recognition research, 1(1), 2006, 55-62.
- Xiong J., Zhang Z., Wang C., Cen J., Wang Q., Nie J., *Pupil localization algorithm based on lightweight convolutional neural network*, The Visual Computer 2024, pp. 1-17.
- Xu G., Zhang Z., Ma Y., *A Novel and Efficient Method for Iris Automatic Location*, Journal of China University of Mining and Technology. Volume 17, Issue 3, 2007, pp. 441-446, doi.org/10.1016/S1006-1266(07)60122-9.
- Zhou Z.-H., Geng X., *Projection functions for eye detection*, Pattern recognition 2004, vol. 37, no. 5; pp. 1049-1056.

IMPLEMENTAREA UNUI SISTEM DE DETECȚIE A PUPILEI
PENTRU IMAGINI ALE OCHILOR DE REZOLUȚIE JOASĂ BAZAT PE UN
CLASIFICATOR CU REȚEA NEURONALĂ COMPLET CONECTATĂ

(Rezumat)

Această lucrare prezintă o implementare a unui sistem de detectare a centrului pupilei care utilizează un clasificator bazat pe rețele neuronale. Fiind destinată procesării imaginilor oculare de joasă rezoluție, implementarea se face cu o rețea neuronală complet conectată. Clasele clasificatorului sunt definite în funcție de pozițiile posibile ale centrului pupilei. Rețeaua neuronală a fost antrenată, testată și validată folosind 39.000 de imagini oculare din 18 baze de date, imagini obținute de la diferiți subiecți umani și în condiții de iluminare variabile și neuniforme. Soluția se distinge prin faptul că furnizează direct coordonatele centrului pupilei și nu doar zona în care aceasta se află. Rezultatele arată atât acuratețe ridicată, cât și viteză mare de procesare, soluția propusă fiind potrivită pentru implementări de buget redus ale sistemelor de detectare a privirii în timp real.