# Deep learning method for accessible eccentric photorefraction

Mathieu Vu, Emilie Chouzenoux, Jean-Christophe Pesquet, Stéphane Boutinon, Marius Peloux, Philippe Pinault

# DEEP LEARNING METHOD FOR ACCESSIBLE ECCENTRIC PHOTOREFRACTION

*Mathieu Vu, Émilie Chouzenoux *, Jean-Christophe Pesquet*

Université Paris-Saclay, Inria, CentraleSupélec, OPIS, CVN
Paris, France

*Stéphane Boutinon, Marius Peloux, Philippe Pinault*

Essilor International
Créteil, France

## ABSTRACT

Affordable and accessible screening tools are crucial to diagnose and, hence, effectively address vision impairment. We propose a novel pipeline based on photorefraction which complies with these requirements. It includes an innovative, easy-to-use, optical portable device combined with an image processing algorithm to automatically compute refraction errors from twelve jointly captured images. To do so, a supervised deep convolutional model is proposed, trained on synthetic images generated by a physical-aware simulator, as well as on real images carefully collected from a cohort of volunteers. We provide in this article an assessment of our approach accuracy and robustness on both simulated and real datasets. Our promising results pave the way for a practical application of the proposed device for an accurate vision loss diagnostic in a real setting.

***Index Terms*—** deep learning, convolutional network, retinal images, eye refraction, eccentric photorefraction, photoretinoscopy

## 1. INTRODUCTION

Almost 2 billion people worldwide benefit from vision correction while an estimated number of 2.7 billion people suffer from uncorrected refractive error due to a lack of affordability, access, awareness, and in some cases, acceptance [1]. A key element to reduce poor vision is to provide accessible and affordable screening tools. In this work, we introduce an innovative device complying with those criteria allowing non-expert users to screen patients with vision impairment, as well as an embedded deep learning-based algorithm that computes objective eye refraction errors from the eye images captured by the device's integrated camera.

### 1.1. Photorefraction principle

Photorefraction is a non-invasive screening method to measure eye refraction [2]. Its optical principle is summarized in

*M.V. and E.C acknowledge support from the ERC Starting Grant MAJORIS ERC-2019-STG-850925.

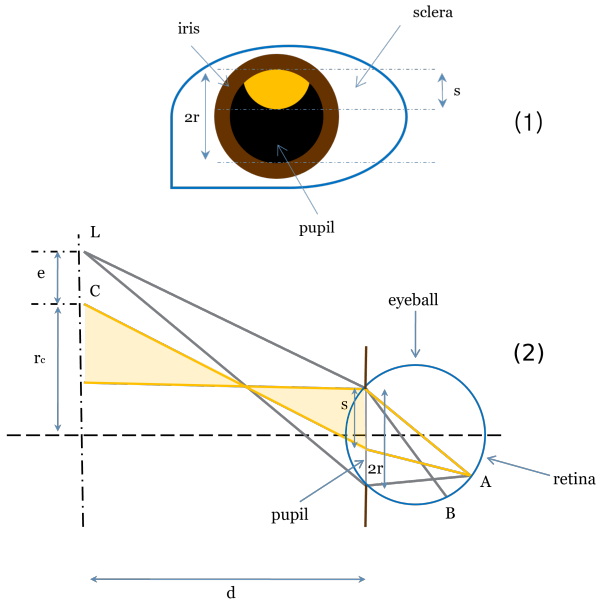


**Fig. 1**. (1) Eye physiology scheme. (2) Photorefraction principle scheme

Figure 1(2), given the schematic description of the eye, in Figure 1(1). Consider a camera with a round aperture of diameter $2r_c$ and center $O$. The user's eye is at a distance $d$ from this aperture, and $C$ corresponds to the upper border of the aperture. A point source $L$ is positioned at a distance $e$ - called eccentricity - from point $C$ along meridian line $(OC)$. Light is emitted by $L$, enters the eye through the pupil and reaches the user's retina, leading to a spot $AB$ whose size and shape depends on the subject's level of ametropia (i.e., abnormal refractive eye condition such as myopia, hyperopia, or astigmatism). The light from this spot is diffusely reflected towards the camera. The latter, focusing on the eye pupil plane (i.e., pupil radius $r$), observes a light pattern whose size $s$, shape, and orientation, are functions of $(i)$ the subject's ametropia, $(ii)$ the light source position in relation to the camera aperture, and $(iii)$ the measurement distance $d$.

Therefore, on a photorefraction device embedding sev-

eral light sources, with controlled distance measurement and light source positions, ametropia can be estimated, and then adressed if needed with a suitable correction (e.g., glasses). Quantitative estimation of ametropia commonly relies on the estimation of three scalar parameters, namely $M$, $J_0$, and $J_{45}$, which are equivalent substitutes of the Sphere $S$, Cylinder $C$, and Axis $\alpha$ parameters commonly used in glasses prescription. Here, we follow Thibos' definition of those parameters [3], with the convention of negative cylinder values:

$$M = S + \frac{C}{2}, \quad J_0 = -\frac{C}{2}\cos 2\alpha, \quad J_{45} = -\frac{C}{2}\sin 2\alpha. \quad (1)$$

Such representation is more practical than using $S$, $C$, and $\alpha$, for numerical and graphical analysis of optometric data, with the advantages of being expressed in a common unit, namely the dioptre, and of expressing quantities that can be analyzed individually from the other parameters values [4].

### 1.2. State-of-the-art

Using infrared light from a light-emitting diode (LED) source, for the photorefraction screening, presents the advantage of preventing pupillary constriction, thus allowing the refraction to be done with a pupil as dilated as possible [5]. Most device designs hence include a camera along with multiple infrared LED with various eccentricities, with the aim to extend the measurable range [5, 2, 6]. Images are then processed by an external computational tool [7] or even, more recently, by miniaturized computers [6], to increase portability. This is the case, for instance, of the 2WIN device (Adaptica, Italy) which performance was assessed in different contexts [8, 9]. Other recent tools perform eccentric photorefraction from a smartphone device [10, 11], yielding an even more accessible screening thanks to their ubiquity. However, the aforementioned approaches have limited measurable range due to the use of a unique and visible light source.

While various device designs were proposed for photorefraction, most share a common methodology to compute eye refraction parameters from light crescents. The standard method consists of computing a power value alongside three meridians, obtained by performing photorefraction with an eccentric light source at three different angles [7, 6]. The brightness profile on each meridian is then used to compute a set of refraction parameters. More recently, machine learning models such as Support Vector Regression models [10] and Deep Convolutional Neural Networks [11], have also been proposed to this task, claiming better performance than traditional image processing methods.

### 1.3. Contribution

In this work, we introduce an end-to-end eccentric photorefraction pipeline, merging an innovative capture device with an original refraction parameters regression method based on deep learning. Our approach infers $(M, J_0, J_{45})$ from a patient eye picture to determine its refraction error. The proposed capture device is designed with the purpose of being a portable and affordable device, without sacrificing its effectiveness. Its design, incorporating infrared LEDs, allows for a measurable range that covers most vision impairments. Our deep-learning based regression model shows accurate results evaluated on both synthetic and real images, and an assessed robustness, increasing confidence in its prediction ability.

The remaining of the paper is organized as follows: the measurement device is introduced in Section 2. The proposed deep learning architecture, and the data generation strategy, is described in Section 3. Numerical experiments are presented in Section 4, and Section 5 concludes the paper.

## 2. PROPOSED PHOTOREFRACTION DEVICE

Our screening tool is designed such that the user can conveniently hold it in hand and point it at a patient. It consists of a camera whose aperture is surrounded by 12 infrared LEDs (cf. Figure 3, left) which serves as eccentric light sources for the photorefraction. At a given distance from the patient (typically greater than $50$ cm), the user triggers the capture process. One after another, the LEDs emit infrared light that is partly reflected by the patient's eyes and synchronously captured by the camera before being sent to a connected smartphone. The capture process lasts in total 400 ms, which allows to minimize perturbations that can occur over time (eg. lighting variation, motion blur etc).

The resulting 12 images (see Figure 2 for an example) are then slightly numerically refined using classic image processing techniques (e.g., denoising) by the software running on the smartphone. After pupils detection, images are cropped so that each eye (left / right) is processed independently and pupils are centered. Each input is thus a standardized size image with 12 channels, each corresponding to a LED.

## 3. PROPOSED DEEP LEARNING REGRESSION PIPELINE

Given a 12 channels input image $x = (x^{(c)})_{1 \leq c \leq 12}$, our model estimates a target among $(M, J_0, J_{45})$, as $f_{\theta_j}(x)$, with parameters $\theta_j$ trained separately for each target $j \in \{M, J_0, J_{45}\}$. We retained the convolutional neural network (CNN) architecture described in Fig. 3 (right), with six convolutional layers followed by six fully connected layers for a total of 12,141,441 parameters, that can be viewed as a VGG-like architecture [12]. Larger architectures such as VGG-16 or Deep Residual Networks (ResNet) [13] led to training instabilities, while smaller models produced larger prediction errors.

Given a training set $\mathcal{D} = \{x_i, y_{i,j}\}_i$ of images paired with their corresponding ground truth parameter values, the CNN weights $\theta_j$ are learned to minimize the mean squared error

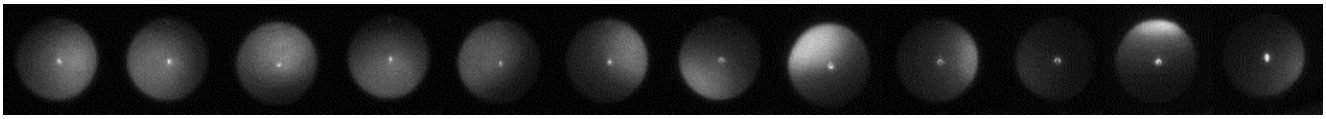

**Fig. 2**. Example of twelve eye pupil images captured sequentially by a different LED of the proposed device.

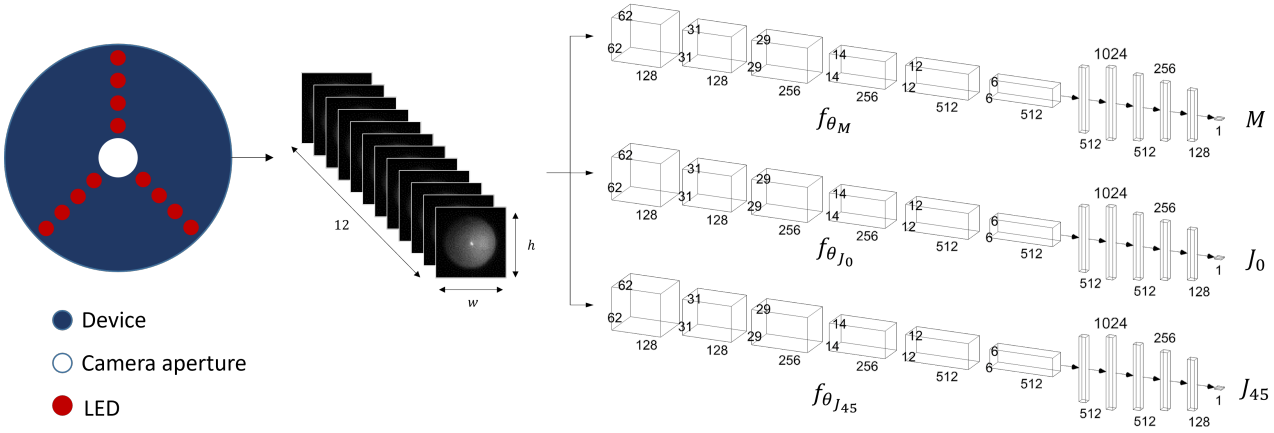


**Fig. 3**. Proposed pipeline: Schematic layout of the 12 LEDs (in red) on the capture device (left). Each LED emits infrared light while the camera synchronously captures an image of the patient's eyes where light is reflected, yielding 12 images for each eye of a patient (see Fig. 2). Our deep learning model (right) has three branches, each taking as input the 12 images, and predicting a given target parameter among $(M, J_0, J_{45})$.

loss $\frac{1}{|\mathcal{D}|}\sum_{i=1}^{|\mathcal{D}|}(f_{\theta_j}(x_i) - y_{i,j})^2$, by backpropagation. In the testing phase, we measure the model prediction error using mean and standard deviation errors.

### 3.1. Data

**Synthetic images.** We developed an internal simulation model inspired from [14, 15] able, given a certain set of parameters, to simulate realistic synthetic images of an eye captured by our device. A total number of 50,000 patients images were synthesized with this simulator, to train/validate our deep learning model.

**Real images.** Volunteers went through the process of image capture with our proposed device along with another eye evaluation using a WAM-700 wavefront analyzer (Essilor, France) providing reference values for objective refraction. The process was executed multiple times for each of the approximately 250 volunteers, producing in total around 3000 samples. The data was anonymized before any processing or storage, in compliance with GDPR [16]. An ethic comity supervised the process and the device design was accredited and approved to be safe, in particular regarding LED intensities.

### 3.2. Neural network training

In all our experiments (i.e., for synthetic as well as real data), the model for each target parameter is trained by Adam optimizer [17] for 200 epochs, with a mini-batch size of 128 samples. The learning rate is originally set to $10^{-4}$ and target parameters are normalized to obtain a reduced and centered distribution. Datasets were randomly split such that 70% were included in the training set and the remaining 30% were in the test set; 10% of the training set is separated to be used as validation set. A 10-fold cross-validation was performed for the real dataset only, as it contains fewer samples.

### 3.3. Neural network robustness

Since the seminal work in [18], it is known that neural networks are sensitive to adversarial noise which can lead to erroneous outputs. Adversarial perturbations can be generated maliciously, but they can also arise accidentally. Evaluating the robustness of our proposed deep learning architecture thus appears mandatory to permit its reliable use. Basically, if $x = (x^{(c)})_{1\leq c\leq 12}$ is a considered 12-channel input image and $f_{\theta_j}(x)$ with $j \in \{M, J_0, J_{45}\}$ is one of the produced output estimates, we are interested in quantifying the effect of an input error $b$ in the acquisition process. By assuming that the input error is small enough and the network $f_{\theta_j}$ is differentiable at $x$, we can use the following first-order approximation:

$$f_{\theta_j}(x+b) \simeq f_{\theta_j}(x) + \big(\nabla f_{\theta_j}(x)\big)^\top b. \tag{2}$$

If the noise $b$ is assumed to be Gaussian with zero-mean and covariance matrix $\Lambda_b$, the output error $e = f_{\theta_j}(x+b) - f_{\theta_j}(x)$ is thus approximately Gaussian with zero-mean and variance

$$\sigma_e^2(x) = \big(\nabla f_{\theta_j}(x)\big)^\top \Lambda_b \nabla f_{\theta_j}(x), \tag{3}$$

where $\nabla f_{\theta_j}(x)$ is the gradient of the network at $x$, easily computed by backpropagation. $\Lambda_b$ can be chosen as a diagonal matrix whose elements corresponding to pixels in the pupil area are $\sigma_b^2$, and those outside are zero. Equation (2) can be used to calculate confidence intervals. In addition, the quantity $\sigma_e(x)/\sigma_b$ defines the so-called weighted local Lispchitz constant of network $f_{\theta_j}$ at point $x$ [19]. A reasonable value of the Lipschitz constant is a strong indicator of the robustness of the network [20].

## 4. EXPERIMENTS

**Benchmarks comparison**. We start our experiments by comparing different image processing methods, on the task of estimating $(M, J_0, J_{45})$ parameters given the output data of our imaging device. Namely, we compare our CNN to a reference method for refraction errors computation in eccentric photorefraction, i.e., the brightness slope method (BSM) [6, 7], as well as to a competing deep learning-based method relying on a fully connected (FC) architecture with 5 layers respectively of size 256 - 512 - 256 - 128 - 1, resulting in a total of 12,879,105 parameters. Table 1 presents statistics of the prediction errors for the three methods. Since a biased model can easily be corrected by offsetting predictions, our main criterion in the evaluation of the models is their error deviation. In particular, the standard deviation metric provides crucial insights on the effectiveness of the methods. Minimum and maximum values of the errors complete the analysis to evaluate the error spreading. To better assess these results, recall that glasses are prescribed with a precision of 0.25 D. Deep learning-based approaches sensibly outperform the BSM method which relies on image processing and optic formulas. The proposed CNN architecture provides the best results in most cases, and, in particular, outperforms both benchmarks in terms of standard deviation and maximum error scores.

| | $M$ | | | $J_0$ | | | $J_{45}$ | | |
|---|---|---|---|---|---|---|---|---|---|
| Error | BSM | FC | CNN | BSM | FC | CNN | BSM | FC | CNN |
| min | -4.66 | -2.31 | **-2.16** | -2.82 | -1.22 | **-0.97** | -2.56 | **-0.83** | -1.17 |
| mean | 0.05 | **0.02** | -0.03 | -0.04 | **0.01** | 0.02 | 0.01 | 0.01 | -0.01 |
| std | 0.75 | 0.60 | **0.56** | 0.43 | 0.28 | **0.26** | 0.32 | 0.23 | **0.22** |
| max | 5.73 | 2.86 | **2.50** | 2.62 | 1.01 | **0.88** | 2.36 | 1.13 | **0.80** |

**Table 1**. Statistic of prediction error wrt. ground truth (in D.).

**Synthetic vs real data.** When trained and tested on synthetic data, our CNN reaches very satisfying scores. As Table 2 shows, the performance of the CNN trained on synthetic data slightly degrades when tested on real data. The results improve greatly when training on real data.

**Comparison with the 2WIN device**. Table 3 displays the comparative performance to another portable device, namely the 2WIN vision screener produced by Adaptica whose performance are reported from other studies [8, 9]. This device also measured objective photorefraction and was compared against retinoscopy under cycloplegia. Our method shows overall much better results particularly in terms of error mean on $M$ and standard deviation. While both devices are similar in terms of usage, our method seems to show better screening accuracy than the 2WIN device.

| Train | Test | $M$ | $J_0$ | $J_{45}$ |
|---|---|---|---|---|
| Synthetic | Synthetic | 0.10 | 0.07 | 0.07 |
| | Real | 0.70 | 0.33 | 0.32 |
| Real | Real | **0.56** | **0.26** | **0.22** |

**Table 2**. CNN prediction error standard deviation wrt. ground truth (in D.) on synthetic and real data.

| | 2WIN [8] | | 2WIN [9] | | Ours | |
|---|---|---|---|---|---|---|
| | mean | std | mean | std | mean | std |
| $M$ | 0.41 | 1.37 | 1.25 | 1.15 | **-0.03** | **0.56** |
| $J_0$ | -0.08 | 0.35 | 0.19 | 0.45 | **0.02** | **0.26** |
| $J_{45}$ | **0.00** | 0.25 | -0.05 | **0.15** | -0.01 | 0.22 |

**Table 3**. Performance comparison to the 2WIN device.

**Robustness** Following the methodology described in Section 3.3, we provide in Figure 4, histograms of the weighted local Lipschitz constant computed at each point of a test set. We observe globally that reasonable values are obtained, demonstrating the robustness of our approach.

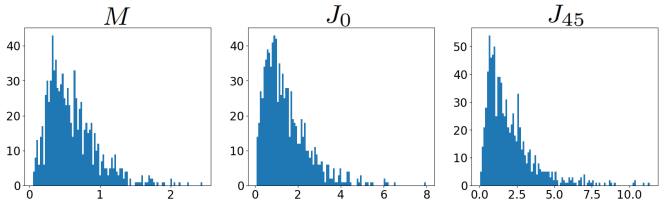


**Fig. 4**. Histograms of the weighted local Lipschitz constant computed for each test sample, for each target.

**Implementation performance**. With an average inference time of only 100 ms per eye for all three targets on a Samsung Galaxy S20 FE equipped with an Exynos 990 processor and only 130 MB of storage required, our models comply to the inherent constraints of an embedded software.

## 5. CONCLUSION

In this paper, we have shown that combining a novel capture device with a deep learning-based algorithm allows us to design an accurate and robust screening tool for eccentric photorefraction. Such a device should respond to the worldwide increasing demand for identifying and correcting vision impairment, particularly in countries where accessibility and affordability are critical requirements. Our future endeavour will be dedicated to validation on additional cohorts by paying attention to cultural and physiological diversity.

## 6. COMPLIANCE WITH ETHICAL STANDARDS

The human data collected in this study were approved by Parkway Independent Ethics Committee in Singapore (IRB Reference number : PIEC/2021/048). The written informed consent was obtained from participants and children's parents.

## 7. REFERENCES


[1] W. H. Organization, "Eye care, vision impairment and blindness," 2023. https://www.who.int/health-topics/blindness-and-vision-loss.

[2] T. D. Cole, "Multimeridian photorefraction: a technique for the detection of visual defects in infants and preverbal children," *Johns Hopkins APL Technical Digest*, vol. 12, no. 2, pp. 166–175, 1991.

[3] L. N. Thibos, W. Wheeler, and D. Horner, "Power vectors: an application of fourier analysis to the description and statistical analysis of refractive error," *Optometry and Vision Science*, vol. 74, no. 6, pp. 367–375, 1997.

[4] H. Starynkevitch, "Réfraction subjective: Une nouvelle méthode vectorielle de détermination du cylindre (1/3)," *Points de Vue*, 2020.

[5] F. Schaeffel, L. Farkas, and H. C. Howland, "Infrared photoretinoscope," *Applied Optics*, vol. 26, no. 8, pp. 1505–1509, 1987.

[6] R. Agarwala, A. Leube, and S. Wahl, "Utilizing mini-computer technology for low-cost photorefraction: a feasibility study," *Biomedical Optics Express*, vol. 11, no. 11, pp. 6108–6121, 2020.

[7] F. Gekeler, F. Schaeffel, H. C. Howland, and J. Wattam-Bell, "Measurement of astigmatism by automated infrared photoretinoscopy," *Optometry and Vision Science*, vol. 74, no. 7, pp. 472–482, 1997.

[8] S. J. Martin, H. E. Htoo, N. Hser, and R. W. Arnold, "Performance of two photoscreeners enhanced by protective cases," *Clinical Ophthalmology*, pp. 1427–1435, 2020.

[9] A. Kurent, "Comparison of the cycloplegic refractive measurements with handheld, table-mounted refractometers and retinoscopy in children," *Ophthalmology Journal*, vol. 7, pp. 200–207, 2022.

[10] Z. Yang, E. Y. Fu, G. Ngai, H. V. Leong, C.-w. Do, and L. Chan, "Screening for refractive error with low-quality smartphone images," in *Proceedings of the 18th International Conference on Advances in Mobile Computing & Multimedia*, pp. 119–128, 2020.

[11] E. Y. Fu, Z. Yang, H. V. Leong, G. Ngai, C.-w. Do, and L. Chan, "Exploiting active learning in novel refractive error detection with smartphones," in *Proceedings of the 28th ACM International Conference on Multimedia*, pp. 2775–2783, 2020.

[12] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[13] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016.

[14] Y.-L. Chen, B. Tan, and J. Lewis, "Simulation of eccentric photorefraction images," *Optics Express*, vol. 11, no. 14, pp. 1628–1642, 2003.

[15] Y. Wu, L. N. Thibos, and T. R. Candy, "Two-dimensional simulation of eccentric photorefraction images for ametropes: factors influencing the measurement," *Ophthalmic and Physiological Optics*, vol. 38, no. 4, pp. 432–446, 2018.

[16] European Parliament and Council of the European Union, "Regulation (EU) 2016/679 of the European Parliament and of the Council," 2023-05-26. https://data.europa.eu/eli/reg/2016/679/oj.

[17] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[18] C. Szegedy, W. Zaremba, I. Sutskever, J. Bruna, D. Erhan, I. Goodfellow, and R. Fergus, "Intriguing properties of neural networks," *ICLR, arXiv preprint arXiv:1312.6199*, 2014.

[19] P. L. Combettes and J.-C. Pesquet, "Lipschitz certificates for neural network structures driven by averaged activation operators," *SIAM Journal on Mathematics of Data Science*, vol. 2, pp. 529–557, 2020.

[20] K. Gupta, F. Kaakai, B. Pesquet-Popescu, J.-C. Pesquet, and F. D. Malliaros, "Multivariate Lipschitz analysis of the stability of neural networks," *Frontiers in Signal Processing*, vol. 2, 2022.