

Human Eye Aberrometry Data Generation Using Generative Adversarial Neural Network

M. O. Yaroshenko,  [0000-0002-3092-3856](https://orcid.org/0000-0002-3092-3856)

National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute"  [00syn5v21](https://www.researchgate.net/profile/M-O-Yaroshenko)
Kyiv, Ukraine

Abstract—It's obvious that for development and improvement of methods and apparatus for diagnosis and treatment of optical flaws of human eye at the modelling stage, it's necessary to have sets of real measurements. However, data requests to clinics are accompanied by substantial amount of bureaucracy procedures and, at the same time, acquired dataset may be too small, which can be critical, for example, for training of neural networks. According to the analysis of existing publications, publicly available datasets of aberrometry data (sets of eye's refractive flaws) are rare and consist of relatively low number of measurements. But, due to current development state of neural networks, it is possible to generate data based on real measurements. The most common solutions are methods based on the usage of the Generative Adversarial Networks (GAN). This tendency is also relevant for the modern ophthalmology, but no publications aimed at aberrometry data synthesis were found. For this reason, objective of this work is development of solution for generation of sets of human eye's refractive errors using neural networks. Proposed solution includes generator and critic networks trained according to the Wasserstein GAN with Gradient Penalty (WGAN GP) algorithm. In order to improve training, the method of data augmentation called Data Augmentation Optimized for GAN (DAG) was used, moreover, the possibility of augmentation of aberrometry data in two forms was implemented — for both Zernike coefficient vectors and wavefront pixel images. According to the result's evaluation, generated data has the distribution close to the real sample (Fréchet distance equals 0.7) and, at the same time, it is neither a copy of real measurements (92% creativity rate) nor duplication of a few aberration sets (diversity metric equals 3.64 which is close to the optimal 3.83). The direction of further improvement includes enhancement of existing architectures of generator and critic, search or creation of bigger training dataset and refinement of data augmentation technics.

Keywords: *Ophthalmology; Generative adversarial networks; Data augmentation.*

I. INTRODUCTION

Modern treatment and correction of eye's optical flaws uses variety of different methods and equipment ranging from individually adjusted contact lenses and glasses to laser surgery. Precise diagnosis of aberrations (optical flaws of human eye) is the pledge of successful choice of treatment approaches. It is reasonable that at the modelling stage of new treatment and diagnosis method's development, high amount of aberration examples is needed. Taking into consideration capabilities and efficiency of neural networks in the sphere of diagnosis, the availability of aberration datasets becomes highly relevant. But, nevertheless, there are some obstacles on the way of acquiring such datasets for research. Firstly, the nature of the data — sets of aberrations of human's eye is the medical confidential information which is, moreover, has high degree of uniqueness, that can be used for identification of persons. This fact restricts the data's accessibility. Secondly, detailed measurement of high order aberrations is quite uncommon. For this reason, not many people undergo such diagnosis. The latter case is crucial for development of diagnosis methods based on neural networks because

small size of a dataset complicates development of network's architecture and leads to a risk of overfitting.

Thus, the solution to the problem of small amount of aberration data is relevant. Extensive increase of such datasets using the mass diagnosis campaigns and programs requires administrative efforts and high costs. That's why usage of technical methods is highly actual and effective.

II. LITERATURE REVIEW

A. Datasets of Ophthalmological Measurements

Accessibility of representative datasets is crucial for statistical data processing methods. Such sets of measurements are highly important for artificial neural networks. Thus, it is reasonable to conduct a search of dataset expansion methods among existing solutions for neural networks, or among solutions which directly use neural networks for it.

As it was mentioned before, ophthalmological datasets are quite rare. It can be observed from the fact that only one survey [1] of such datasets was found



during the preparation of this work. Despite the completeness of this survey, it lacks the information of human eye’s aberrations — all the investigated datasets consist of graphical information, from which no data on refractive errors can be directly acquired. The work [2] suggests another survey of methods for artificial ophthalmological data generation using GANs, first proposed by Goodfellow I. J. et al. in [3]. Similar to [1], the aberration-related information is not provided by [2]. One of the biggest publicly available datasets of human eye’s aberrations includes 50 measurements for the research [4]. This dataset would be used in current work for preparation and evaluation of proposed solution, but, obviously, generated sample may have low level of similarity to the real data because of dataset size insufficient for training high-performance GANs (usually it requires thousands of vectors). So, from [1] and [2] it can be concluded, that, firstly, among the few ophthalmological datasets the graphical ones (photos of fundus, iris etc.) are of research community interest, and, secondly, usage of GANs for ophthalmological data generation already

established itself as an effective solution. That is why it’s reasonable and actual to expand existing datasets of aberrations namely through the involvement of GANs taking into account low number of measurements provided.

B. Human Eye’s Aberrations

Aberrations can be divided into two types: chromatic and monochromatic. Chromatic aberrations are caused by difference in wave distribution for rays with different wavelength. Monochromatic aberrations emerge during the distribution of light emission with single wavelength. Generation of information, which describes monochromatic aberrations, is the object of research for this work.

Aberration maps — surfaces of wavefronts — are the main source of refraction error data. The wavefronts themselves are usually defined by the weighted sum of surfaces described by the Zernike modes, which allow to specify any surface determined in the unit circle with required accuracy [5]:

$$W(\rho, \varphi) = \sum_n \sum_m N_n^m R_n^m(\rho) \left[C_n^m \cos(|m|\varphi) + C_n^{-m} \sin(|m|\varphi) \right]$$

where $W(\rho, \varphi)$ – value of wavefront for the point inside the unit circle with polar coordinates (ρ, φ) , N_n^m – value of the norming factor for n -th radial order and m -th angular frequency, $R_n^m(\rho)$ – value of Zernike polynomial, C_n^m – Zernike coefficient which equals standard deviation (Root Mean Square) of the mode. The norming factor N_n^m in its turn is defined as:

$$N_n^m = \sqrt{\frac{2(n+1)}{1+\delta}}, \quad \delta = \begin{cases} 1, & m = 0; \\ 0, & m \neq 0. \end{cases}$$

And $R_n^m(\rho)$ is described as:

$$R_n^m(\rho) = \sum_{k=0}^{\frac{n-|m|}{2}} \frac{(-1)^k (n-k)!}{k! \left(\frac{n+|m|}{2} - k\right)! \left(\frac{n-|m|}{2} - k\right)!} \rho^{n-2k}.$$

Allocations of radial orders and angular frequencies is shown at Fig. 1.

C. Generative Adversarial Neural (GAN) Networks

Training of the GAN is the competition between two neural networks: the generator G and discriminator D . During the training, the discriminator’s objective is to accurately distinguish, i.e., classify, real sample x and artificial $G(z)$ ones synthesized by the generator used

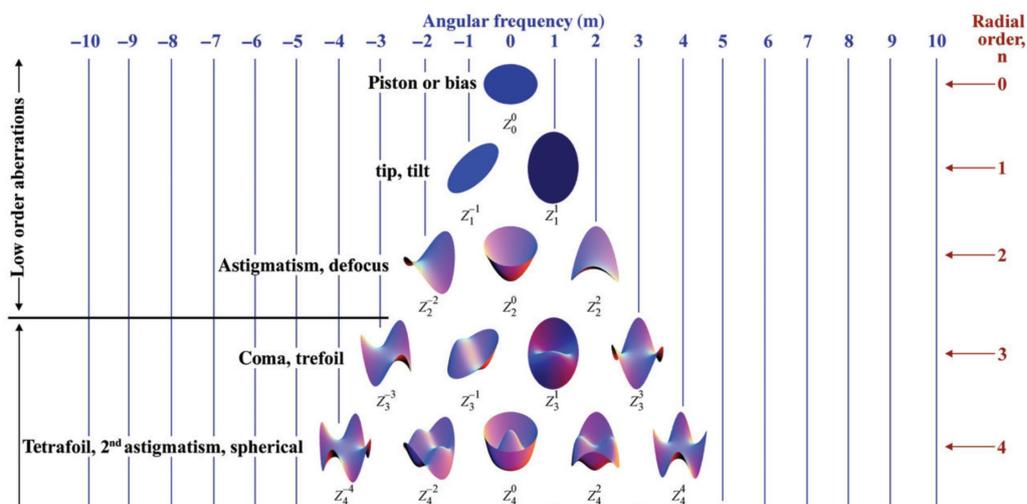


Fig. 1 Visualization of Zernike polynomials from 0-th to 4-th radial orders [6]

random values z . Noise as a generator’s input provides generated data’s diversity and, according to assumptions, represents the sets of latent variables which implicitly characterize data sample. The head’s rotation angle on photo can be considered as an illustrative example of latent variable. Commonly, the discriminator’s output on real data should be equal 1, on generated – 0.

Formally, the training of the GAN is described as a minimax game with value function $V(D,G)$ [3]:

$$\min_G \max_D V(D, G) = E_{x \sim P_d} [\log(D(x))] + E_{z \sim P_z} [\log(1 - D(G(z)))]$$

where P_d is the real data distribution, P_z is the noise distribution.

Training and picking the right architecture for the GAN is not an easy task, because, unlike other types of neural networks, loss functions of generator and discriminator depend on each other and change in the process of training. Especially relevant it becomes for the medical data which is usually composed in small datasets increasing the risk of overfitting. In order to stabilize the training, the Wasserstein GAN was proposed in [7], changing the approach of sample evaluation by critic (analog of discriminator of the common GAN), including, usage of value of unrestricted range as a critic’s metric for the realism of sample, application of weight clipping, critic’s L_D and generator’s L_G loss functions change to the Wasserstein distance-based:

$$L_D = E[D(G(z))] - E[D(x)],$$

$$L_G = -E[D(G(z))].$$

Despite the better stability during the training, WGAN also has drawbacks. In [8] the negative effect of weight clipping is described, namely the critic’s tendency to learn simple functions and necessity of the clipping value’s fine-tuning, otherwise the risk of gradient explosion or vanishing significantly increases. To address these issues in [8] addition of gradient penalty to critic’s loss function was introduced:

$$L_D = E[D(G(z))] - E[D(x)] + \lambda E \left[\left(\|\nabla_{\bar{x}} D(\bar{x})\|_2 - 1 \right)^2 \right],$$

where $\bar{x} = tx + (1-t)G(z)$, t is a random value in range $[0,1]$, λ is a tunable coefficient which usually equals 10. Experimental evaluation of the method proved better convergence of the training and more realistic images to be generated.

In case of small datasets, researchers often use data augmentation in order to expand dataset adding simple processing of training sample, e.g. rotation, cropping, translation for image data. Indeed, augmentation usage cannot properly substitute filling the dataset with higher

amount real measurements, but it is able to significantly improve neural network training. However, involvement of augmentation into the GAN training can bring new risks — generator can integrate augmentations into synthesized images. Harming influence of straightforward augmentation usage for GANs was experimentally demonstrated in [9]. It caused increase of the distance between real and generated images’ distributions measured by the Fréchet Inception Distance (FID) [10] metric from 6.8 to 47.3 for the MNIST dataset [11]. Authors of [9] proposed to use invertible transformations for both real and generated images, whereby the separate discriminator D_k is assigned to each invertible transformation T_k . For the sake of regularization, all the discriminators share all the layers except the last ones. It was theoretically and experimentally proved that the proposed method — Data Augmentation Optimized for GAN (DAG) — doesn’t distort distribution of generated images. Thus, generator safely receives more feedback information from few discriminators D_k . Proposed in [9] minimax game for GAN is described as:

$$\max_{D, \{D_k\}} V(D, \{D_k\}, G) = V(D, G) + \frac{\lambda_u}{K-1} \sum_{k=2}^K V(D_k, G),$$

$$\min_G V(D, \{D_k\}, G) = V(D, G) + \frac{\lambda_v}{K-1} \sum_{k=2}^K V(D_k, G),$$

where λ_u and λ_v – coefficients for training configuration, K – number of transformations, wherein the transformation of $k=1$ is considered as an absence of transformations. The diagram of the method is shown in Fig. 2. According to experiments, the usage of DAG for the SS-GAN [12] training with the CIFAR-10 [13] (only 25% of images were used) dataset improved the FID reducing it from 46.2 to 30.3-35.2 depending on the set of transformations T_k . Hence, the DAG can be considered as an effective method for data augmentation in terms of quality of generated images.

Unfortunately, metric, which can comprehensively evaluate GAN-generated data, doesn’t exist. According to [14], optimal GAN have to generate data which has the distribution similar to real sample’s one, but at

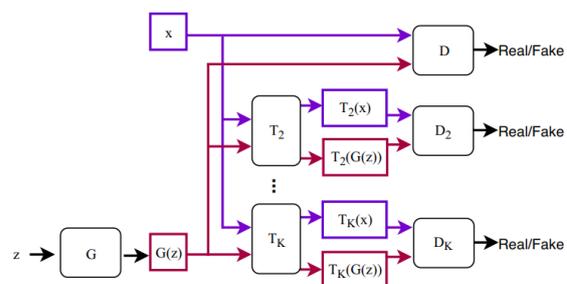


Fig. 2 Diagram of GAN training with DAG. Violet color displays the stream of real data, red color — the stream of generated one [9].

the same time doesn't copy neither real data nor its own specimen. Thus, in [14], GAN-generated data is proposed to evaluate using metrics of inheritance, creativity and diversity.

For the inheritance calculation of graphical information, the FID metric is often used, which involves well-known network for image classification Inception [15] and shows the similarity between real and generated image sets based on the Inception's response on the both sets. Usually, the last layer of the Inception is removed for the FID calculation, and outputs of the remaining network is used for computation of the Fréchet distance [16] between measurements and synthesized data:

$$FID(X, Y) = \|\mu_X - \mu_Y\|^2 + \text{Tr}(\Sigma_X + \Sigma_Y - 2\sqrt{\Sigma_X \Sigma_Y}),$$

where X and Y — responses of penultimate Inception's layer to real and synthesized data respectively; μ_X and μ_Y — average values of X and Y respectively; Σ_X and Σ_Y — covariance matrices of X and Y respectively; $\text{Tr}(\bullet)$ — matrix's trace (sum of all the matrix's diagonal elements).

Creativity of generated data is defined as a ratio of copies of real specimen to the total size of generated sample. According to [14], the Structural Similarity Index Measure (SSIM) [17] between all the pairs of real and synthesized images has to be calculated and, in case it equals 0.8, synthesized image is considered to be a copy of a real one. Calculation of the SSIM between two images x and y is defined as:

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)},$$

where μ_x — average value of x , μ_y — average value of y , σ_x^2 — variance of x , σ_y^2 — variance of y , σ_{xy} — covariance of x and y , $c_1 = (k_1L)^2$ and $c_2 = (k_2L)^2$ — variables for stabilization of small denominator's value, coefficients k_1 and k_2 are usually equal 0.01 and 0.03 respectively, L — dynamic range of the signals.

For synthesized image's diversity evaluation, the SSIM is also used — consequential clustering based on the pairwise SSIM is done after exclusion of the copies of real specimen inside the set of generated images. In that way synthesized duplicates are removed. The SSIM threshold for considering image as a copy also equals 0.8.

Despite the aforementioned image-aimed GANs, data augmentation and evaluation, these methods are applicable to generation of human eye's aberrations described by vectors of Zernike coefficients. It is possible because of two reasons. Firstly, GANs are also capable of non-image data generation depending on the training

datasets and architecture. Secondly, wavefront $W(\rho, \varphi)$ can be represented as a monochrome image with pixels' values of $W(\rho, \varphi)$. In general, transformation "Zernike coefficients — Wavefront image" can be considered as invertible, because vector of Zernike coefficients can be restored from the image of a wavefront. Thus, it is possible to integrate this transformation in the DAG without additional risks of generated data distortion.

III. OBJECTIVE OF THE WORK

Based on the foregoing, WGAN with gradient penalty (WGAN GP) is one of the best solutions among the generative neural networks in terms of training stability and quality of the result. These advantages are useful for medical data generation, given the restricted accessibility and small sizes of the datasets. It is reasonable to use data augmentation for training improvement, but its straightforward usage can distort generated data. This issue is addressed by GAN-specific augmentation methods. DAG is one of them, and its effectiveness is proven both theoretically and practically. Evaluation of generated data using inheritance, creativity and diversity metrics allows better understanding of the result and, therefore, further improvement of GAN training. Thus, the objective of this work is to develop the solution for aberrometry data generation based on WGAN GP with DAG for sample augmentation, and inheritance, creativity and diversity metrics for evaluation of synthesized data.

IV. THE PROPOSED SOLUTION

The solution for aberrometry data generation, developed for this work, includes usage of the WGAN GP with DAG. Zernike modes with radial order from 1 to 6 (as commonly used) with coefficients from the [4] will be used for training and evaluation. For the result assessment the metrics of inheritance, creativity and diversity are involved.

A. Training of the proposed network

To speed up training and exclude complicated patterns for the GAN to learn, forward $W(C)$ and inverse $W^{-1}(w)$ translations from the vector of coefficients C to the waveform image w and vice versa are used for data augmentation methods. In other cases, data is represented as a vector of Zernike coefficients scaled to [0;1] (Fig. 3).

Transformation T_1 means passing the coefficients intact, transformations from T_2 to T_M require application of $W(C)$ and $W^{-1}(w)$ to include the classical techniques of image augmentation such as mirroring, rotation by 90° etc., transformations from T_{M+1} to T_N are for vector data, e.g. shuffle of vector elements, scaling, etc. Full list of transformations shown in Table 1.



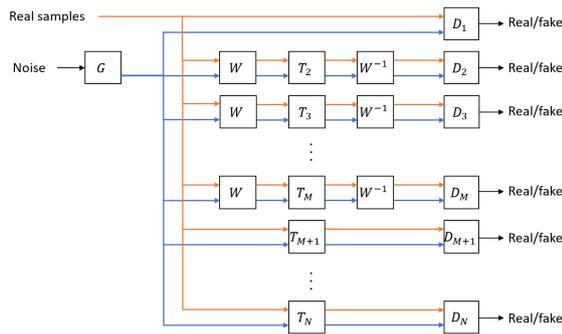


Fig. 3 Diagram of forward data processing during the proposed WGAN GP's training.

TABLE 1 TRANSFORMS FOR DATA AUGMENTATION

No, k	Data type	Transformation, T_k
1	Vector of Zernike coefficients	Intact values
2	Wavefront image	Clockwise rotation by 180°
3	Wavefront image	Mirroring along the X axis (top-down)
4	Wavefront image	Clockwise rotation by 90° , with consequent mirroring along the Y axis (left-right)
5	Vector of Zernike coefficients	Swap of vector halves
6	Vector of Zernike coefficients	Arranging of vector's elements in reverse order

According to the WGAN GP definition, critic's loss functions are defined as:

$$L_{D_k} = E[D_k(T_k(x))] - E[D_k(T_k(G(z)))] + \lambda E\left[\left(\|\nabla_{\hat{x}} D_k(\hat{x})\|_2 - 1\right)^2\right],$$

where $\hat{x} = tT_k(x) + (1-t)T_k(G(z))$, t is a random value in range $[0,1]$, λ is a tunable coefficient. Whereas, according to the DAG, the critic is a single network with $N=6$ output branches, each consists of one classification layer, the total loss function is calculated as:

$$L_D = L_{D_1} + \frac{\lambda_u}{K-1} \sum_{k=2}^K L_{D_k}, \quad \lambda_u = 0.4.$$

For the generator loss function takes into account outputs from all critics:

$$L_G = L_{G_1} + \frac{\lambda_v}{K-1} \sum_{k=2}^K L_{G_k}, \quad \lambda_v = 0.4,$$

where $L_{G_k} = -E[D_k(T_k(G(z)))]$.

B. Architecture of Networks

Both generator and critic have simple structure of multilayer perceptron (Fig. 4) with gradual increase (for generator) and decrease (for critic) of nodes on each layer.

According to the DAG method, critic network has one input and few outputs — one for each transformation T_k . As will be seen later, architectures' configuration is appropriate but application of more contemporary networks can lead to better results, however, may require more effort on fine-tuning of hyperparameters and learning settings.

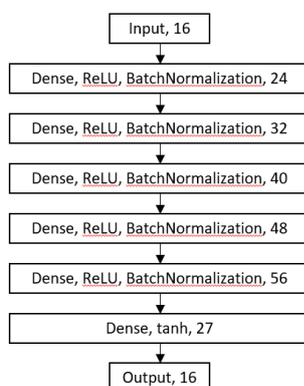
C. Training Parameters

WGAN GP training requires few iterations of critic weight's optimization before each update of the generator's ones. For current work this amount equals 8. Mini-batch also chosen as 8. Adam optimizer is used [18] with its learning rates which equal $1 \cdot 10^{-4}$ for the generator and $2 \cdot 10^{-4}$ for the critic. Coefficient λ for critic's loss function equals 10. Number of epochs is chosen 1000. Coefficients k_1 and k_2 for the SSIM calculation are both 0.01.

D. Results

Data preparation, training and result evaluation is done in the Kaggle [19] environment using Python with TensorFlow [20], SciPy [21] and Matplotlib [22] libraries.

Generator Network:



Critic Network:

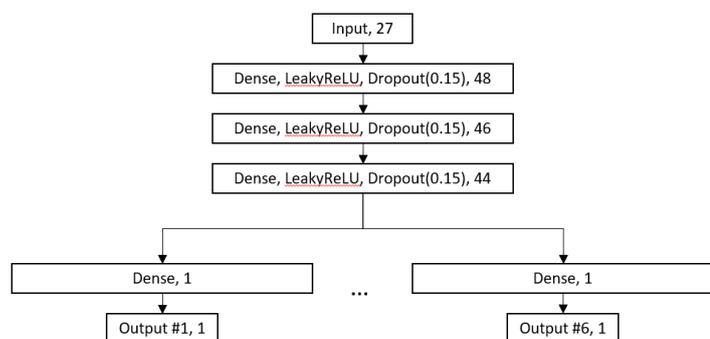


Fig. 4 Architectures of generator and critic networks

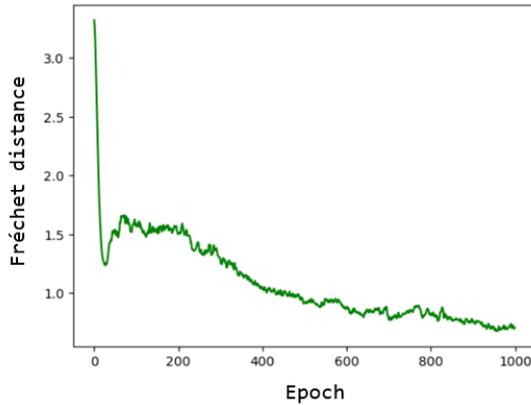
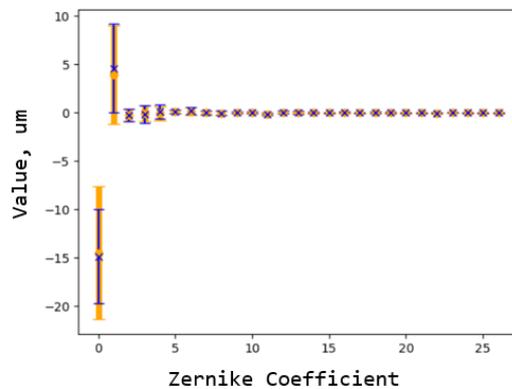
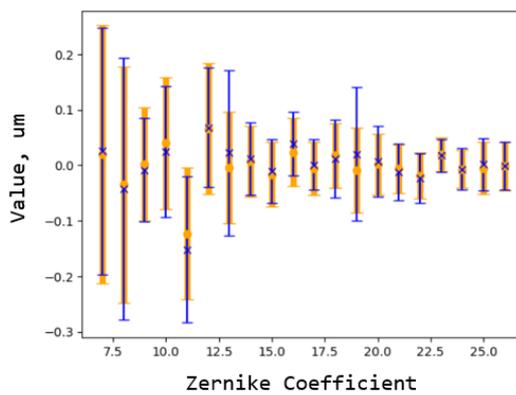


Fig. 5 Fréchet distance between generated and real data throughout the training



a)



b)

Fig. 6 Distributions of generated (thin bars – ± 1 standard deviation, cross – mean value) and real (thick bars – ± 1 standard deviation, circle – mean value) Zernike coefficients: a) all the coefficients with $n = 1..6$; b) coefficients from C_3^1 to C_6^6 .

Fréchet distance between generated 50 Zernike coefficient vectors and 50 real measurements is used for generated data quality monitoring (Fig. 5). We don't need data specimen reduction with other neural network, as it commonly used for FID, because chosen number of Zernike modes (27) doesn't compose vectors of large volume.

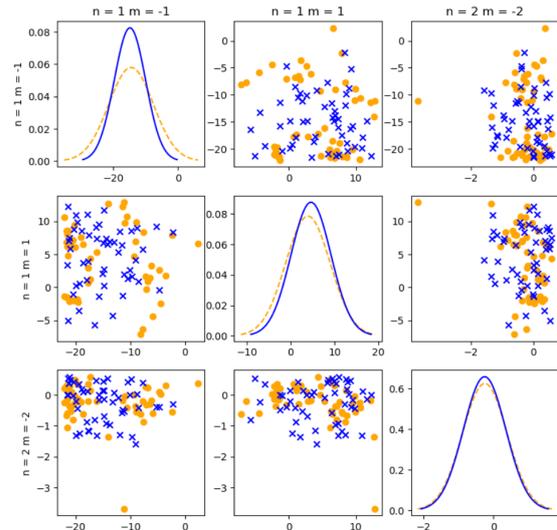


Fig. 7 Demonstration of pairwise relations between generated (cross) and real (circle) Zernike coefficients (in μm). Distributions of generated (solid line) and real (dashed line) are situated on the diagonal (values in μm)

As it can be seen from the Fig. 5, usage of DAG with WGAN GP makes training stable and provides its convergence. Acquired value of 0.7 of Fréchet distance is considered as an inheritance metric in this work. This value can be lowered further by training for more epochs but for this work this distance can be considered as sufficient for demonstration of effectiveness of the proposed solution. Visually training results can be observed at Fig. 6 and Fig. 7.

0.92 as a creativity metrics value proves the fact that the majority of generated vectors are not copies of real ones. Diversity level of 3.64 for synthesized data is close to the optimal value of 3.83 (the case when each cluster has one vector).

Thus, proposed solution is suitable for generation of artificial datasets of human eye's aberrations in the form of Zernike coefficient vectors. Generated data has distribution which can be considered as close to the real one, and, at the same time, it doesn't copy neither real data nor itself.

CONCLUSION

In this work the solution based on usage of WGAN GP combined with DAG was firstly proposed for generation of human eye aberrations in the form of Zernike coefficient vectors. Also, it was implemented using combination of augmentation methods for both ways of aberration representation — vector and wavefront image — for better training.

According to the results of network's training, the proposed solution is capable of generation of the data which distribution can be considered as close to the real measurement's (Fréchet distance equals 0.7) and, at the same time, synthesized vectors neither copies

of real ones (92% creativity rate) nor copies of themselves (diversity metric equals 3.64 which is close to the optimal 3.83). It should be noted that the result was achieved by using relatively small training dataset of 50 measurements. Obviously, the training would lead to more optimal networks' configuration in case of higher number of vectors.

Despite the successful metric values, it is important to take into account the synthetic nature of generated data applying it to real-world solutions. That fact can impose limitations on its usage, such as different degree of similarity between coefficients' pairwise relations, low variance of some generated Zernike coefficients, etc.

These kinds of problems are typical for synthetic data and the proposed solution objectively cannot avoid them due to inability of neural networks to learn all dependencies between data features perfectly, let alone the most subtle ones.

Further research can be aimed at usage of more contemporary network architectures, such as SAGAN [23], with more effective methods of data augmentation. Probably, all the applied and newly designed methods should take into account limited accessibility of datasets with human eye aberration — it can be considered as the main factor which impedes application of common approach to data generation with GANs.

REFERENCES

- [1] S. M. Khan *et al.*, "A global review of publicly available datasets for ophthalmological imaging: barriers to access, usability, and generalizability," *Lancet Digit Health*, vol. 3, no. 1, pp. e51–e66, Jan. 2021, DOI: [10.1016/S2589-7500\(20\)30240-5](https://doi.org/10.1016/S2589-7500(20)30240-5).
- [2] A. You, J. K. Kim, I. H. Ryu, and T. K. Yoo, "Application of generative adversarial networks (GAN) for ophthalmology image domains: a survey," *Eye and Vision*, vol. 9, no. 1, p. 6, Dec. 2022, DOI: [10.1186/s40662-022-00277-3](https://doi.org/10.1186/s40662-022-00277-3).
- [3] I. J. Goodfellow *et al.*, "Generative Adversarial Networks," Jun. 2014, URL: <https://arxiv.org/abs/1406.2661>.
- [4] J. Jarosz, P. Mecé, J.-M. Conan, C. Petit, M. Paques, and S. Meimon, "High temporal resolution aberrometry in a 50-eye population and implications for adaptive optics error budget," *Biomed Opt Express*, vol. 8, no. 4, p. 2088, Apr. 2017, DOI: [10.1364/BOE.8.002088](https://doi.org/10.1364/BOE.8.002088).
- [5] I. Chyzh, G. Tymchyk, and T. Shisha, *Aberometriya optychnoyi systemy oka lyudyny: Monohrafiya [Aberrometry of the optical system of the human eye: Monograph]*. Kyiv, Ukraine: National Technical University of Ukraine "Kyiv Polytechnic Institute," 2013.
- [6] V. Lakshminarayanan and A. Fleck, "Zernike polynomials: a guide," *J Mod Opt*, vol. 58, no. 7, pp. 545–561, Apr. 2011, DOI: [10.1080/09500340.2011.554896](https://doi.org/10.1080/09500340.2011.554896).
- [7] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein GAN," Jan. 2017, URL: <https://arxiv.org/abs/1701.07875>.
- [8] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. Courville, "Improved Training of Wasserstein GANs," Mar. 2017, URL: <https://arxiv.org/abs/1704.00028>.
- [9] N.-T. Tran, V.-H. Tran, N.-B. Nguyen, T.-K. Nguyen, and N.-M. Cheung, "On Data Augmentation for GAN Training," *IEEE Transactions on Image Processing*, vol. 30, pp. 1882–1897, 2021, DOI: [10.1109/TIP.2021.3049346](https://doi.org/10.1109/TIP.2021.3049346).
- [10] M. Heusel, H. Ramsauer, T. Unterthiner, B. Nessler, and S. Hochreiter, "GANs Trained by a Two Time-Scale Update Rule Converge to a Local Nash Equilibrium," Jun. 2017, URL: <https://arxiv.org/abs/1706.08500>.
- [11] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998, DOI: [10.1109/5.726791](https://doi.org/10.1109/5.726791).
- [12] T. Chen, X. Zhai, M. Ritter, M. Lucic, and N. Houlsby, "Self-Supervised GANs via Auxiliary Rotation Loss," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019, pp. 12146–12155, DOI: [10.1109/CVPR.2019.01243](https://doi.org/10.1109/CVPR.2019.01243).
- [13] "CIFAR-10 and CIFAR-100 datasets." [Online]. Available: <https://www.cs.toronto.edu/~kriz/cifar.html>. [Accessed: 16-Jul-2023].
- [14] S. Guan and M. Loew, "Evaluation of Generative Adversarial Network Performance Based on Direct Analysis of Generated Images," in *2019 IEEE Applied Imagery Pattern Recognition Workshop (AIPR)*, 2019, pp. 1–5, DOI: [10.1109/AIPR47015.2019.9174595](https://doi.org/10.1109/AIPR47015.2019.9174595).
- [15] C. Szegedy *et al.*, "Going deeper with convolutions," in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 1–9, DOI: [10.1109/CVPR.2015.7298594](https://doi.org/10.1109/CVPR.2015.7298594).
- [16] D. C. Dowson and B. V. Landau, "The Fréchet distance between multivariate normal distributions," *J Multivar Anal*, vol. 12, no. 3, pp. 450–455, Sep. 1982, DOI: [10.1016/0047-259X\(82\)90077-X](https://doi.org/10.1016/0047-259X(82)90077-X).
- [17] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multiscale structural similarity for image quality assessment," in *The Thirty-Seventh Asilomar Conference on Signals, Systems & Computers, 2003*, pp. 1398–1402, DOI: [10.1109/ACSSC.2003.1292216](https://doi.org/10.1109/ACSSC.2003.1292216).
- [18] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," Dec. 2014, URL: <https://arxiv.org/abs/1412.6980>.
- [19] "Kaggle: Your Machine Learning and Data Science Community." [Online]. Available: <https://www.kaggle.com/>. [Accessed: 04-Aug-2023].
- [20] "TensorFlow." [Online]. Available: <https://www.tensorflow.org/>. [Accessed: 04-Aug-2023].
- [21] "SciPy." [Online]. Available: <https://scipy.org/>. [Accessed: 04-Aug-2023].
- [22] "Matplotlib — Visualization with Python." [Online]. Available: <https://matplotlib.org/>. [Accessed: 04-Aug-2023].
- [23] H. Zhang, I. Goodfellow, D. Metaxas, and A. Odena, "Self-Attention Generative Adversarial Networks," May 2018, URL: <https://arxiv.org/abs/1805.08318>.

Надійшла до редакції 26 серпня 2023 року

Прийнята до друку 19 грудня 2023 року



Генерування аберометричних даних шляхом застосування генеративно-змагальної нейронної мережі

М. О. Ярошенко,  [0000-0002-3092-3856](https://orcid.org/0000-0002-3092-3856)

Національний технічний університет України

«Київський політехнічний інститут імені Ігоря Сікорського»  [00syn5v21](https://ror.org/00syn5v21)

Київ, Україна

Анотація—Отримання медичних даних для статистичних досліджень, розробки нових методів лікування та відповідного обладнання є процесом, який супроводжується великою кількістю бюрократичних процедур, а обсяг отриманої вибірки може виявитись недостатнім. Остання проблема особливо актуальна для розробки методів на основі штучних нейронних мереж. Анонізовані вибірки медичних даних у відкритому доступі є нечисленними, причому серед них зазвичай не представлені певні специфічні дослідження. Ці фактори також є релевантними для аберацій — оптичних похибок людського ока. Дійсно, аналіз існуючих публікацій демонструє вкрай малу кількість датасетів з аберометричною інформацією, в той час як більший інтерес для наукової спільноти становить обробка офтальмологічних зображень. Діагностику для визначення аберацій високих порядків роблять нечасто, тому для отримання великих обсягів даних необхідно запроваджувати кампанії для діагностики населення, що може бути затратним з точки зору часу та коштів. Іншим способом є використання існуючих методів генерації даних, таких як генеративні змагальні нейронні мережі (Generative Adversarial Neural Networks, GAN). Втім, їхнє навчання є нестабільним і, за малих обсягів даних, виникає ризик перенавчання. Більш стабільний вид GAN — Wasserstein GAN (WGAN) — використовує інший підхід до визначення функцій втрат та жорстке обмеження ваг під час оптимізації. Однак він також має недоліки: наприклад, обмеження ваг вимагає додаткових зусиль на підбір порогового значення, бо в іншому випадку існує ризик вибуху або зникнення градієнтів. Недоліки WGAN усунуто додаванням градієнтного штрафу (Gradient Penalty, GP). Незважаючи на високу стабільність навчання WGAN GP, розмір навчаючої вибірки також грає важливу роль в підготовці мережі. З метою його нарощування, що є актуальним для нечисельних навчаючих вимірювань, використовуються методи аугментації даних — утворення нових примірників шляхом застосування до них нескладних перетворень. Однак звичайне застосування аугментації даних при навчанні GAN не є припустимим через інтеграцію цих перетворень у згенеровані примірники. Одним з методів навчання GAN, які дозволяють використання аугментації даних, є Data Augmentation Optimized for GAN (DAG). Незважаючи на те, що більшість архітектур GAN та супутніх методів навчання та нарощування даних описані для роботи з інформацією у вигляді зображень, це не є перепорою у їхньому застосуванні до вирішення задачі генерації аберометричних даних, адже така інформація може бути представлена у двох формах — вектори коефіцієнтів та піксельні зображення хвильових фронтів. Таким чином, задачею даної роботи є розробка методу генерації аберометричних даних на базі WGAN GP із застосуванням DAG. Запропоноване рішення є WGAN GP оригінальної архітектури, для навчання якої використовувались методи нарощування даних як для графічної форми хвильових фронтів, так і для векторів коефіцієнтів Церніке. Аналіз результату генерації за спеціалізованими метриками спадковості, творчості та різноманіття показав, що запропоноване рішення здатне синтезувати дані, що є схожими на реальні (відстань Фреше дорівнює 0.7), і, які, водночас, не копіюють реальні вимірювання (метрика креативності на рівні 92%), та не мають великої кількості самоповторів (значення метрики різноманіття має значення 3.64, що близько до оптимально 3.83). Подальші дослідження можуть бути направлені на використання більш досконалих архітектур штучних нейронних мереж, засобів аугментації даних для GAN та пошук або створення більших навчаючих вибірок.

Ключові слова — офтальмологія; генеративні змагальні мережі; аугментація даних.

