Deep convolutional generative adversarial networks in retinitis pigmentosa disease images augmentation and detection

Pawel Powroznik[1*] , Maria Skublewska-Paszkowska[1] , Katarzyna Nowomiejska[2] ,
Andreas Aristidou[3,4] , Andreas Panayides[4] , Robert Rejdak[2]

[1] Department of Computer Science, Lublin University of Technology, Nadbystrzycka 36B, 20-618 Lublin, Poland
[2] Chair and Department of General and Pediatric Ophthalmology, Medical University of Lublin, Chmielna 1, 20-400 Lublin, Poland
[3] University of Cyprus, 75 Kallipoleos Str., Nicosia 1678, Cyprus
[4] CYENS Centre of Excellence, 23 Dimarchias Square, Nicosia 1016, Cyprus
* Corresponding author's e-mail: p.powroznik@pollub.pl

ABSTRACT

Large medical datasets are crucial for advancing contemporary medical practices that incorporate computer vision and machine learning techniques. These records serve as indispensable resources for identifying patterns that assist healthcare professionals in diagnosing rare diseases and enhancing patient outcomes. Moreover, these datasets drive research into the causes and progression of such diseases, potentially leading to innovative therapeutic strategies. However, the acquisition of such data poses significant challenges due to privacy and ethical concerns, as well as the rarity of certain conditions. Therefore, it is imperative to both collect new medical data and develop tools that facilitate the enhancement of existing datasets while preserving the accurate characteristics of the diseases. This study focuses on leveraging deep convolutional generative adversarial networks (DCGAN) to expand a dataset containing images of retinitis pigmentosa, a rare eye condition affecting the retina. Our study showcases that integrating xtreme gradient boosting (XGBoost) within the DCGAN framework enhances the clarity and quality of these augmented images. By employing hybrid VGG16 alongside XGBoost techniques during training, we observe improvements in detection accuracy. The outcomes of the proposed method are highly promising, with the model achieving all key performance metrics surpassing the 90% threshold as well as improving baseline classification accuracy by almost 19%.

Keywords: generative adversarial networks, medical image augmentation, Rare eye disease classification, retinitis pigmentosa.

INTRODUCTION

In recent years, computer vision has seen rapid advancement, particularly driven by deep learning techniques. A noticeable trend is the increasing depth and complexity of models, with a focus on achieving state-of-the-art performance [1, 2]. This trend can be attributed, in part, to the significance of the dynamic development of artificial intelligence methods, where new deep neural networks undergo rigorous testing on a massive dataset comprising millions of images. However, it is worth noting that while accomplishing an excellent result is essential in laboratory environment, real-world applications often differ significantly, due to variations in dataset sizes, integrity, completeness, and composition. Medical datasets containing images of rare diseases are crucial for advancing diagnostic accuracy, treatment strategies, and studies about atypical disorders [3]. They enable machine learning algorithms to recognize patterns and assist healthcare professionals in identifying diseases that they might encounter infrequently, thereby improving patient outcomes. Moreover, these datasets can help in developing studies on causes and progression of rare diseases, potentially leading to new approaches supporting therapeutic solutions. However, obtaining such data poses significant challenges. Rare diseases, by definition, affect a small number of individuals, which means that

assembling a comprehensive dataset is inherently difficult. Privacy concerns and ethical considerations further complicate data collection, as patient consent and anonymity must be maintained. Additionally, there is often a lack of funding and incentives for research on conditions that affect a limited patient population, leading to scarcity of resources to systematically collect and curate these valuable datasets. Collaboration across institutions and countries, along with strong ethical frameworks, are essential to overcome these barriers and build robust medical datasets for rare diseases. Achievements in the field of machine learning and computer vision may also be helpful here, due to its possibility of expansion of existing medical data sets.

Motivation of study

For many years, medical imaging has had enormous potential in clinical diagnosis and patient assessment. Optical coherence tomography (OCT) images are often the basis for diagnosing eye diseases, including rare ones. The studies [3, 4] presented selected methods of classification, detection and analysis of images illustrating retinitis pigmentosa (RP) diseases. For this purpose, solutions such as residual attention network, RP segmentation network and simple convolutional networks were applied. However, many authors emphasize the huge lack of data they had to deal with while developing tools supporting diagnostics. Creating high-quality synthetic images is crucial for increasing the performance of machine learning models [5]. These data will also lead to better understanding of rare eye diseases. Therefore, this study presents methods that allow for augmentation of image data using both typical methods and, above all, generative neural networks.

Contribution of the proposed work

This study's significant contributions are concisely summarized as follows:
1. The adoption of a deep convolutional generative adversarial network (DCGAN) for data augmentation, which has improved the quality of the dataset.
2. The incorporation of data obtained for consistency using classic augmentation methods found in computer vision issues.
3. The deployment of alteration of VGG16 architecture that effectively integrates the augmented DCGAN-generated dataset with the

initial dataset, resulting in a notable increase in accuracy over past studies. Furthermore, the Xtreme gradient boost (XGBoost) algorithm is employed to refine feature representation.

The primary goal of the article is to examine the impact of small medical datasets augmentation on the quality of classification. The study placed particular emphasis on the detection of rare eye diseases, in particular retinitis pigmentosa.

The further structure of this paper is as follows: Section 2. describes current state of-the-art. Section 3. presents a detail methodology of the proposed approach in this study. Section 4. contains experimental findings with in-depth analysis. Section 5. draws conclusions from the research findings as well as future direction of study.

RELATED WORKS

Generating synthetic images for improving detection performance

Various supervised machine learning (ML) approaches have been widely applied for retinal diseases with great success [6, 7]. Automatic methods have achieved better performance in accordance to traditional ones, by automatically extracting the relevant features [8]. However, in order to obtain high classification accuracy results a well-balanced dataset for training should be gathered, which, in practice, is a great challenge. Especially, this situation may occur for rare eye retinal disease recognition [9, 10]. These difficulties have been overcome by applying generative adversarial network (GAN) that may produce additional samples based on random noise [11]. These synthetic images are very similar to the original ones and thus the training set may be balanced. In supervised learning methods the GAN models are utilizing as a pre-processing step to enlarge the training dataset [10, 12]. In recent studies, various retinal diseases have been detected utilizing GAN models. For early recognition of glaucoma, diabetic retinopathy, cataract, macular edema, and myopia the modified generative adversarial-based crossover salp grasshopper (MGA-CSG) was proposed for fundus images [13]. The GAN model was used for feature vector augmentation. This algorithm resulted in high obtained classifier accuracy with a small sample of input data. Due to fact that classification is based on labelled samples while the GAN is unsupervised method, the semi-supervised feature discrimination and distribution was applied to create hybrid classifier system. As a result,

labelled and unlabelled data were predicted and classified achieving high accuracy of 98.7%. The GAN models were also utilized for increasing detection of age-related macular degeneration (AMD). Ten various GAN architectures were compared for generating images with and without AMD, such as: deep convolutional GAN (DCGAN), least squares generative adversarial networks (LSGAN), wasserstein GAN, wasserstein GAN with gradient penalty, deep regret analytic generative adversarial networks, energy-based generative adversarial network, boundary equilibrium generative adversarial networks, conditional GAN (CGAN), and auxiliary classifier GAN [14]. The synthetic and real AMD and non-AMD images were assessed by experts and three deep learning methods: SqueezeNet, AlexNet and ResNet-18. The machine learning algorithms appeared to be as good as medical experts. Additionally, the web system was developed.

The StyleGAN2-ADA model was applied for improving the classification performance of inherited retinal diseases using fundus autofluorescence images in [15]. The synthetic images, generated utilizing GAN model, were evaluated by four clinical experts indicating whether they were real, fake or it was difficult to decide. Additionally, ten various deep learning models were trained with various proportion of synthetic and real images. The authors stated that classification performance obtained after training on synthetic data was comparable to performance achieved using training on all real data. Improving AMD detection utilizing another GAN method for super-resolution (SR) of the OCT B-scan images is described in [16]. The SRGAN proposed architecture allowed for obtaining SR without combining images of low and high resolutions. The quality of the OCT B-scan images was improved. The classification accuracy for single scale CNN method achieved very high results up to 96.54%. The SRGAN was also applied for improving resolution of fundus images that were further utilized for detecting lesions in retina [17]. Mask-RCNN with three variants, R-50, R-101, and X-101, was used. The average precision was the highest for X-101 model, achieving up to 75.20% for exudates and 67.20% for microaneurysms. A few-sample generator, created based on DCGAN and SRGAN, was introduced in [18] for generating synthetic images for branch retinal arterial occlusion and central retinal arterial occlusion. The obtained data were further applied in DeepDrRVOtraining for detecting the ocular diseases. The proposed framework achieved accuracy, ranging from 86.30% to 100%.

Differentiating between real and artificially generated images

High resolution synthetic images (512×512 pixels), referable AMD as well as nonreferable AMD datasets were obtained using Progressively grown Generative Adversarial Networks (ProGAN) methods [19]. The generated images were distinguished by certified experts. Moreover, deep convolutional neural networks were applied that were trained on separate types of images, real and synthetic. The network learned using real images achieved higher ability to recognize referable AMD and nonreferable AMD at the level of 91.12%.

The ability to distinguish synthetic fundus images, generated by the DCGAN, from the real ones were investigated in [20]. The obtained F-measure results reaching up to 95% proved that the great number of fake images could be recognized. Another study concerning generating synthetic diabetic retinopathy was presented in [21]. Two types of GAN models were used: the DCGAN and the LSGAN. As a result, the high-quality and high-resolution images were obtained. The created data were assessed for visual acuity (inception score). The Frechet inception distance was applied for measuring GANs performance. The study indicated that the DCGAN is the most appropriate tool to generate diabetic retinopathy fundus images.

The counterfactual GAN model was also applied for generating high-quality OCT images based on real images and additional information, such as: age and sex [22]. The synthetic retinal aging images were visually assessed by five experts. Moreover, three ResNet50-based neural network approaches were utilized for sex and age classification purposes. The tests were performed on real OCT images with success, achieving 79.5% accuracy for sex and 65.8% for age predictions.

The deep convolutional GAN models were developed for generating retinal synthetic images (DCGAN) and for both generating images and semi-supervised learning (SS-DCGAN) [23]. Although SS-DCGAN was stated to be insufficient as a synthesizer, it appeared to be a proper method as a classifier. The glaucoma assessment was performed using a small labelled dataset and a large unlabelled dataset. The ability to detect glaucoma from cropped fundus images using the SS-DCGAN achieved high results, up to 84.29% F1-score.

The hybrid deep learning method was applied to detect an early stage of glaucoma diseases using fundus images of patients suffering from diabetes

[24]. This approach allowed to recognize two types of retinal diseases in the same time. The GAN model was applied for enlarging the data by creating synthetic images. They were used for classification training and for improving its performance. The proposed Mayfly optimized deep convolutional neural networks included the optimization algorithm for eluding the classifier overfitting. The presented approach achieved high accuracy, up to 99%. The detection of healthy retina, age-related macular degeneration, diabetic retinopathy as well as crystalline retinopathy using augmentation methods was described in [25]. The CutMix method was applied together with CycleGAN for datasets enlargement with high-quality synthetic images. In order to overcome the issue of synthetic data distribution, the multistage GAN training was used. The ResNet50 sensitivity achieved 95.5%.

In diabetic retinopathy detection the African Buffalo Optimization was also applied with great success [26]. That approach involved the hybrid GAN and Bidirectional Gated Recurrent Unit (BiGRU). The GAN was used to generate synthetic images but also for extract latent variables that were further represented as multifaceted vector spaces. The BiGRU was used for diabetic retinopathy detection utilizing temporal relationships between extracted features. The proposed method achieved 98.5% accuracy.

The hybrid Region-based convolutional neural networks together with GAN (RCNN-GAN) was proposed to identify various stages of diabetic retinopathy [27]. The mild, moderate, severe, and nonproliferative stages of the disease was detected. In that study the GAN was used for extracting features while the RCNN acted as a discriminator. The Archimedes optimization Algorithm was applied to optimize both GAN and RCNN parameters. The proposed solution achieved up to 99.4% accuracy.

Another GAN application is denoising images. This method for adaptive optics images has gained great success [28]. Thirty-two high-quality retinal images were blurred using Gaussian kernel convolutions and image motion blur. The proposed denoiseGAN include CGAN model that utilizes high-quality and synthetic low-quality images. While training process of the denoiseGAN the denoised image is predicted from the synthetic one. Applying the minimalization of the diversities between denoised image and the original one the denoiseGAN acquires the skills to remove added noise from images. This approach achieved 99.51% precision and 85.71% recall.

The GAN models seem to be promising tool for enlarging the datasets, especially for rare diseases. They are suitable methods for enhancing image quality. They also have applications in improving detection performance.

MATERIALS AND METHODS

This section provides an overview of the dataset, classical augmentation methods, DCGAN, XGBoost, and evaluation metrics and measures. This paper aims to enhance retinitis pigmentosa image outcomes by integrating DCGAN with transfer learning. For classification purposes the architecture of the VGG16 model was applied, while the entire proposed methodology is outlined in Figure 1.
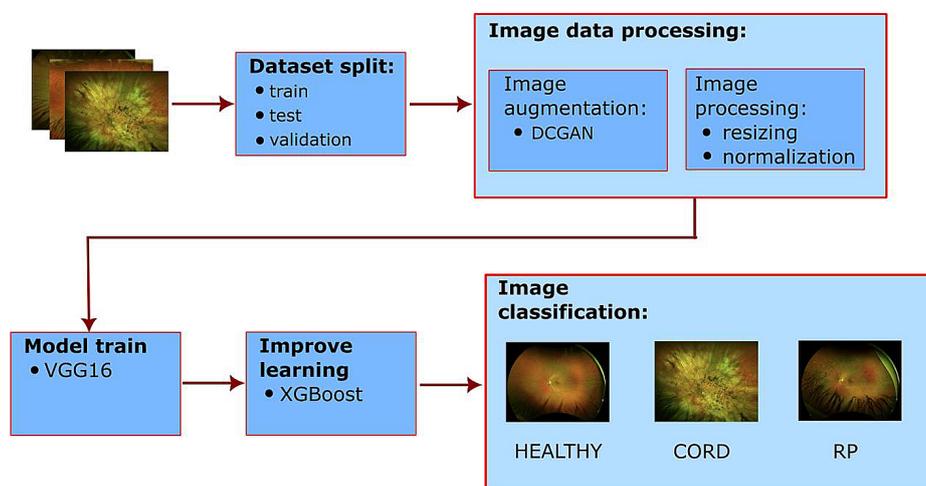
Figure 1. Proposed methodology for DCGAN medical images augmentations and VGG16+XGBoost classification

Datasets

RP is a group of retinal diseases suffering from the progressive dysfunction of rod and cone photoreceptors in the retina. Analysing the type of cell primarily affected, RP may be divided into two categories: rod-dominant (classical RP) and conedominant cone–rod dystrophy (CORD) [3, 6]. This study employs ultra-widefield fundus photography (UWFP) and ultra-widefield fundus autofluorescence (UWFAF) images gathered using an Optos 200TX device (Optos PLC). The created dataset contains healthy patients and patients suffering from cone–rod (CORD) and cone–rod dystrophies (RP). Only cases with clinical characterisation of the retinal dystrophy were selected. UWFP and UWFAF photography were performed after pupil dilation with topical 0.5% tropicamide.

The data were collected from 75 patients, including 25 healthy people, at the chair and department of General and Pediatric Ophthalmology of the Medical University of Lublin in Poland. Totally, 230 optomap retinal photography were gathered. They were grouped into: 132 images with RP located at the periphery of the retina, 48 CORD images and 50 images representing healthy cases. Approval of the Ethic Committee of the Medical University of Lublin has been obtained (no. KE-0254/260/12/2022).

Deep convolutional generative adversarial network

A deep convolutional generative adversarial network is a class of artificial neural network that is used in unsupervised machine learning tasks. It is a variant of the generative adversarial network architecture, which was introduced in [11]. DCGANs were first described in [29]. They are particularly designed to generate new content, like images, that mimic the distribution of real data. Generative adversarial networks consist of two competing networks: a generator and a discriminator. The generator takes in a noise vector from Latent space and generates synthetic data. These fake samples, along with real ones, are inputted into the discriminator, which aims to distinguish between these two distributions. Conversely, the generator aims to mimic the real distribution without directly observing it, making its outputs indistinguishable from genuine samples. Both networks are trained concurrently and adversarially until they reach a balanced state. To address instability during training, the Wasserstein distance is employed. This choice is motivated by its ability to converge across a wider range of distributions and its direct correlation with the quality of the generated data [30]. The resulting formulation of this adversarial competition between generator (G) and discriminator (D) is presented in Eq. 1. The whole structure of DCGAN is presented in Figure 2.
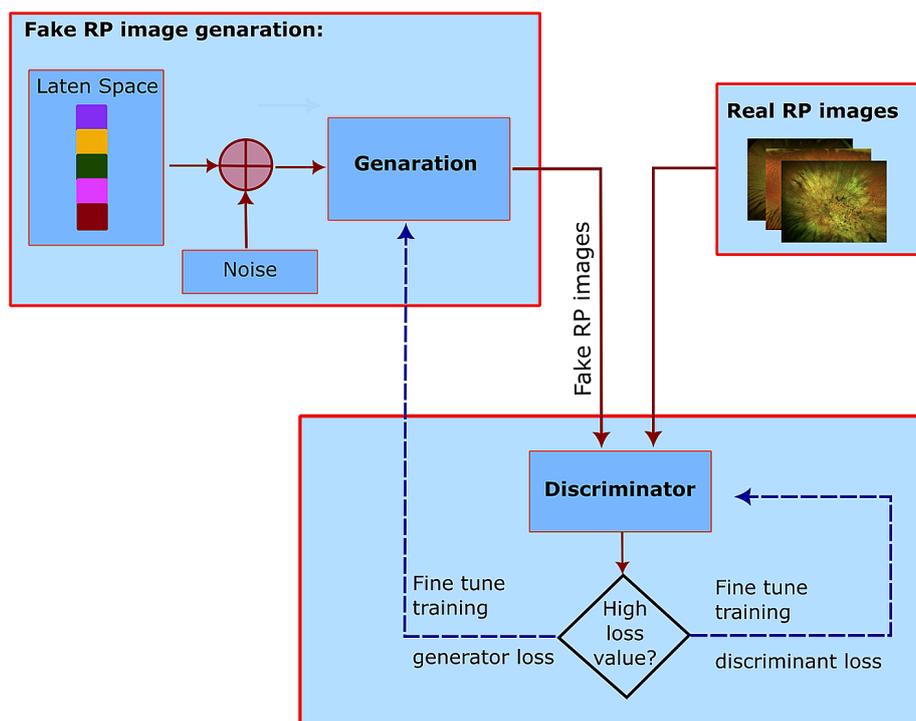
Figure 2. The DCGAN workflow for retinitis pigmentosa data augmentation

$$\min_{G} \max_{D \in \mathcal{D}} \mathbb{E}_{x \sim \mathbb{P}_x}[D(x)] - \mathbb{E}_{z \sim \mathbb{P}_z}[D(G(z))] \quad (1)$$

where: $\mathbb{E}_{x \sim \mathbb{P}_x}$, is binary cross entropy of *G* and *D*, respectively, *z* indicates generated data, $\mathbb{P}_x$ is a probability distribution of the real data *x*, while $\mathbb{P}_z$ corresponds to the prior probability distribution of the noise vector *z*.

Generator

In a GAN the generator is one of the two neural networks that engage in a zero-sum game, where it competes against the discriminator. The role of the generator is to learn how to produce data that is indistinguishable from genuine data provided during the training phase. The generator network takes a random noise vector (latent space vector) as input and generates data. The goal of the generator is to produce data that is indistinguishable from real data [31, 32].

The network's generator is constructed with an 11-layer architecture, encompassing over 15 million adaptable parameters (Fig. 3). This generator begins by accepting a 128-element vector comprised of random numbers uniformly distributed between 0 and 1 (exclusive). The process starts with these values passing through a fully connected (FC) layer, equipped with 15,360 neurons. The output from this layer is then reshaped into a three-dimensional structure, analogous to a 6×5 pixel image, albeit with an

extensive depth of 512 channels. Following this, the network employs alternating layers of standard 2D convolutions (Conv) and 2D transposed convolutions (often called Conv trans up or "deconvolutions"). These layers utilize a *5x5* kernel size and employ 'same' padding to maintain dimensionality, while the transposed convolutions also apply a stride of 2, effectively doubling the spatial dimensions of their input.

Excluding the final layer, each layer's activation is facilitated by a Leaky ReLU function, which introduces non-linearity. The terminal layer, however, employs a hyperbolic tangent (tanh) activation function. This choice is strategic, as the tanh function's output range matches the desired bound [-1, 1] for image generation and offers cantered zero output, which is advantageous for training.

Through five cycles of alternating convolution and transposed convolution layers, with each transposed layer expanding the input's size, the architecture ultimately yields an image with a resolution of 192×160 pixels, containing a single-color channel.

Discriminator

The discriminator employed in this study follows a conventional CNN design tailored for binary classification tasks [31, 32]. Specifically, it comprises 11 layers, in total approximately 9.5 million trainable parameters (Fig. 3). For input, a single-channel image of dimensions 192×160
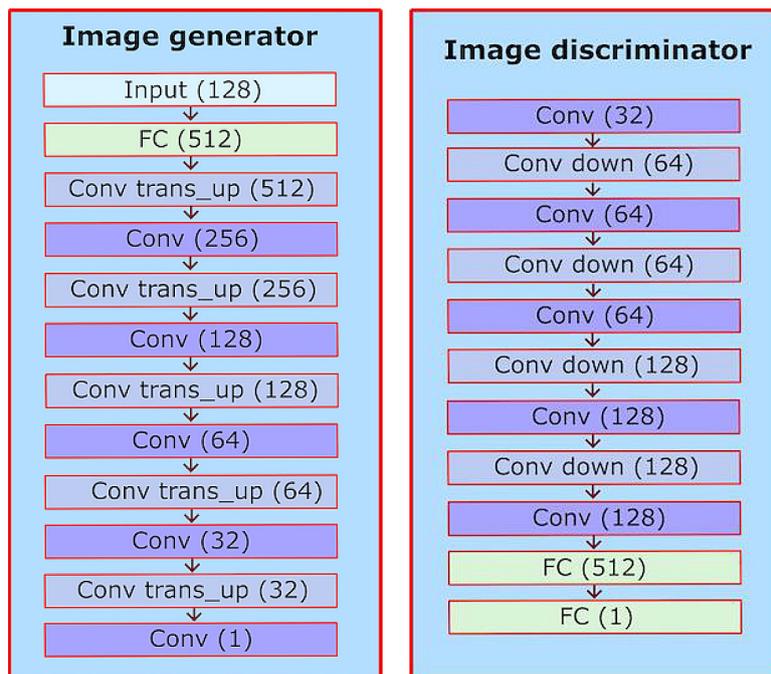
Figure 3. The architecture of deep convolution image generator (left) and discriminator (right)

is utilized. This input undergoes five iterations through convolutional layers, with alternating strides of 1 and 2. The stride of 2 facilitates subsampling since no pooling layers are integrated into the architecture. The discriminator culminates with two fully connected (FC) layers. Throughout the network, all layers employ a Leaky ReLU activation function, with the exception of the final layer, which lacks activation.

Xtreme gradient boost

For the purpose of this study the classical XGBoost algorithm introduced by Chen [33], was slightly modified. XGBoost was intend for datasets with *n* examples and *m* features, as $N\{x_i, y_i\}(|N| = n, x_i, \in R^m, y_i \in R)$. Application the *S* additive function, demonstrated in Equation. 2, enables to forecast the outcome:

$$y_{ii} = \sum_{s=1}^{S} f_s(x_i), f_s \in \mathcal{F} \quad (2)$$

where: $\mathcal{F} = \{f(x) = \vartheta_{q(x)}\}(q: \mathbb{R}^m \to T, \vartheta \in \mathbb{R}^T)$ denotes the spatial regression of trees, *q* is the tree structures, and *T* represents the number of leaves on the tree. The cumulative score is determined by ϑ, aggregating the respective leaves. The regularization objective for acquiring a set of functions is defined by Equation 3 and Equation 4 [33].

$$\mathcal{J} = \sum_{i} l(y_{ii} - y_i) + \sum_{j} \phi(f_j) \quad (3)$$

$$\phi(f) = \gamma T + \frac{1}{2}\lambda\|\vartheta\|^2 \quad (4)$$

where: *l* is the loss function – the difference between the predicted value expressed by y_{ii} and the target value y_i. The ϕ indicates complexity of the model. The extra regularization term λ helps to smooth the final learned weights, reducing the risk of over-fitting. Intuitively, the regularized objective tends to favour models that use simple and predictive functions. Consequently, [33] employed an additive approach, demonstrated in Equation 5. Here, $y_i(t)$ signifies the prediction of the *i-th* instance at the *t-th* iteration.

$$\mathcal{J}^{(t)} = \sum_{i=1} l\left(y_i, y_{ii}^{(t-1)} + f_t(x_i)\right) + \vartheta(f_t) \quad (5)$$

Equation 5. can be improved further with a second-order approximation, as depicted in Equation 6 [32].

$$\mathcal{J}^{(t)} \simeq \sum_{i=1}^{n} \begin{matrix} l(y_i, y_{ii}^{(t-1)}) + g_i f_t(x_i) + \\ + \frac{1}{2}h_i f_t^2(x_i) + \vartheta(f_t) \end{matrix} \quad (6)$$

where: $g_i = \partial_{y_{ii}^{(t-1)}} l(y_i, y_{ii}^{(t-1)})$ and $h_i = \partial^2_{y_{ii}^{(t-1)}} l(y_i, y_{ii}^{(t-1)})$ both g_i and h_i serving as loss functions for the first and second gradient statistics at step *t*, respectively. Equation. 6 can be further transformed by eliminating the constant term, resulting in Equation 7 [33].

$$\widetilde{\mathcal{J}^{(t)}} = \sum_{i=1}^{n} \left[g_i f_t(x_i) + \frac{1}{2}h_i f^2_t(x_i)\right] + \vartheta(f_t) \quad (7)$$

Furthermore, if $S_j = \{i|q(x_i) = j\}$ will be defined as a leaf *j*, the Equation 7. can be expressed in a form of Equation 8 and Equation 9 by expanding ϑ [31].

$$\widetilde{\mathcal{J}^{(t)}} = \sum_{i=1}^{n} \left[g_i f_t(x_i) + \frac{1}{2}h_i f^2_t(x_i)\right] + \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} \omega^2_j$$
$$= \sum_{j=1}^{T}\left[\left(\sum_{i \in S_j} g_i\right)\omega_j + \frac{1}{2}\left(\sum_{i \in S_j} h_i + \lambda\right)\omega^2_j\right] + \gamma T \quad (8)$$

The optimal value ω of leaf *j* for the static function *q* is expressed by Equation 9 and Equation 10 [31].

$$\omega^*_j(q) = -\frac{\sum_{i \in S_j} g_i}{\sum_{i \in S_j} h_i + \lambda} \quad (9)$$

$$\widetilde{\mathcal{J}^{(t)}}(q) = -\frac{1}{2}\sum_{j=1}^{T} \frac{\left(\sum_{i \in S_j} g_i\right)^2}{\sum_{i \in S_j} h_i + \lambda} \quad (10)$$

The whole data augmentation procedure with XGBoost module is described by Figure 4.

Classification model

Concerning the classification task, the well-known VGG16 architecture was utilized. It consists of 16 weight layers. The architecture of VGG16 is distinguished by its clarity and homogeneity, comprising with an input layer, a series of convolutional layers followed by max-pooling layers, ending with fully connected layers (FC). The last FC layer is activated by Softmax function [34]. One parameter we experimented on was the use of dropout [35], which was applied on FC layers. Three different dropout rates were examined: 0% (which is equivalent to no dropout), 25% and 50%. Every model was trained by ImageNet dataset and

---

Algorithm 1 Data augmentation with VGG16 and XGBoost classification

---

Require: Input: An eye image $I(x, y)$

    Step 1: Converting image into gray scale and applying normalization

$$I : \{\mathbb{X} \subseteq \mathbb{R}^n\} \to \{Min, ..., Max\} \text{ change into}$$
$$I_N : \{\mathbb{X} \subseteq \mathbb{R}^n\} \to \{newMin, ..., newMax\}$$
$$I_N = (newMax - newMin)\frac{1}{1+e^{-\frac{I-\beta}{\alpha}}} + newMin$$

    Step 2: Basic image augmentation (rotation, scaling, flipping)

$$Rotation\ I_{new} = (x_r, y_r)$$
$$x_r = (x - center_x) * cos(angle) - (y - center_y) * sin(angle) + center_x$$
$$y_r = (x - center_x) * sin(angle) - (y - center_y) * cos(angle) + center_y$$

$$Scale\ I_{new} = (x_{ratio}, y_{ratio})$$
$$x_{ratio} = \frac{old\_image_x}{new\_image_x}$$
$$y_{ratio} = \frac{old\_image_y}{new\_image_y}$$
$$I_{new}(floor(x * x_{ratio}), floor(y * y_{ratio}))$$
$$Horizontal\ flip\ I_{new} = (width - x - 1, y)$$

    Step 3: Image augmentation using DCGAN

    Step 4: Feature extraction by VGG16 network

    Step 5: Feature mapping with XGBoost

        Adding output to the XGBoost tree by predicting the weight of leaf j

(Eq. (7))

Output: $I_{new} = (x_r, y_r)$

Figure 4. Algorithm 1

fine-tuned on the available images after dataset normalization was applied. The models were trained for a total of 100 epochs with the Adam optimization algorithm [36, 37].

Implementation details

The DCGAN implemented in this study aims to augment the dataset of retinitis pigmentosa images by generating realistic synthetic images. The DCGAN comprises two main components: the Generator and the Discriminator.

The Generator starts with a 128-dimensional noise vector passes through a fully connected layer with 15360 neurons and is reshaped into a 6×5 pixel feature map with 512 channels. The network then uses a series of transposed convolutional layers to progressively upscale the feature maps. Each transposed convolution has a 5×5 kernel size, 'same' padding to maintain spatial dimensions, and a stride of 2 to double the spatial size. Leaky ReLU activation functions with a negative slope coefficient of 0.2 are applied to all layers except the last one. The final layer uses a hyperbolic tangent (tanh) activation function to produce output values in the range [-1, 1], matching the normalized pixel values of the images. After five upsampling stages, the Generator outputs a single-channel image of dimensions 192×160 pixels.

The Discriminator is designed to classify images as real or fake. It accepts an input image of size 192×160 pixels and processes it through five convolutional layers with a *5x5* kernel size, 'same' padding, and alternating strides of 1 and 2 to reduce spatial dimensions while increasing depth. Leaky ReLU activations with a negative slope of 0.2 are used. Two fully connected layers follow the convolutional layers, culminating in a single scalar output representing the probability that the input image is real. The absence of an activation function in the final layer allows for a linear output suitable for the Wasserstein loss computation.

The DCGAN is trained using the Wasserstein loss function with a gradient penalty of 10 to ensure Lipschitz continuity, stabilizing the training process and preventing mode collapse. The Adam optimizer is employed with a learning rate of 0.0002 and decay rate for momentum (β_1) set to 0.5. The batch size during training is set to 32 and continues for approximately 180 epochs or until the Discriminator's loss converges to almost zero.

The VGG16 model is utilized for classifying retinal images into retinitis pigmentosa, cone–rod dystrophies, and healthy cases. VGG16 consists of 13 convolutional and 3 fully connected layers, applying small 3×3 convolutional filters. To adapt VGG16 for this specific classification task, the final fully connected layer is modified to output three

classes. A Softmax activation function is applied to this layer to produce probabilities for each class. The model is initialized with weights pretrained on the ImageNet dataset to leverage transfer learning, allowing it to benefit from features learned from a large and diverse set of images. The input images are resized and normalized to match the expected input format of the network. Training is conducted for 100 epochs using the Adam optimizer. The learning rate is set to 0.0001. To prevent overfitting, dropout regularization is tested on the fully connected layers with rates of 0%, 25%, and 50%, respectively. The loss function used is Categorical Cross-Entropy, appropriate for multi-class classification problems.

To enhance the classification performance further, features from the penultimate layer of VGG16 are extracted and fed into an XGBoost classifier. This combination leverages the deep feature extraction capabilities of the convolutional neural network and the powerful classification abilities of gradient boosting trees, aiming to capture complex nonlinear relationships and interactions between features that may not be fully exploited by the neural network alone.

High-level abstract features extracted from the penultimate layer of the VGG16 model, which capture essential patterns and structures in the retinal images become as input features.

The following values of characteristic parameters were implemented in the model: maximum depth = 6, learning rate (*eta*) = 0.1, and number from estimators is set to 100. As an objective function the multi:softprob is used, while for evaluation multiclass logarithmic loss is chosen.

Traditional augmentation techniques

Data augmentation is a vital technique in image classification that enhances the performance and generalization capabilities of machine learning models, especially deep convolutional neural networks. This method involves creating new training samples by applying various transformations to the existing dataset, thereby increasing its diversity and reducing the likelihood of overfitting. By augmenting the data, models become more robust to variations in lighting, scale, orientation, and other factors commonly encountered in real-world scenarios. Additionally, data augmentation effectively enlarges the training dataset without the need for extra labelled samples, mitigating overfitting risks and improving the model's ability to generalize to new, unseen data. It plays a crucial role in training deep learning models, enabling them to learn features that are more adaptable and responsive to the diverse variations present in real-world images [3].

The constraints of dataset necessitated a limited image modification. Specifically, the following sequences of transformations were applied:
- a horizontal flip – performed with a certain probability,
- brightness – adjusted to a range of 85–115% of the original intensity at the same probability,
- resizing – to 85–115% of its original dimensions, it was shifted by –7% to +7% along each axis, and rotated between –7 and +7 degrees, all contingent on the same probability.

More pronounced alterations to each image were the result of a higher probability value. This probability was denoted by the symbol p, which indicated the augmentation's intensity level. Pixels that were erased during these transformations were filled with a zero value, corresponding to the background intensity. This procedure is explicated in Figure 5. Probability values p of 0.25 and 0.5 were tested.

Algorithm 2 Traditional Augmentation Technique

Require: Input: An eye image $I(x, y)$ and probability p
    Step 1:
        $I_{new_1}(x, y) = Horizontal\ flip(I)$ with probability p
    Step 2:
        $I_{new_2}(x, y) = Adjust\ brightness(I_{new_1}(0.85, 1.15))$ with probability p
    Step 3:
        $I_{new_3}(x, y) = Scale(I_{new_2}(0.85, 1.15))$ with probability p
        $I_{new_4}(x, y) = Translate(I_{new_3}(-7, +7))$ with probability p
        $I_{new_5}(x, y) = Rotate(I_{new_4}(-7, +7))$ with probability p
Output: I_{new_5}

Figure 5. Algorithm 2

EXPERIMENTS AND RESULTS

To assess the effectiveness of the proposed methodology, an experiment was devised as follows: Firstly, a deep neural network architecture, capable of classifying three categories with satisfactory performance, was chosen. The dataset described earlier was used to train this network, employing traditional augmentation techniques. Subsequently, artificially generated images were incorporated into the training dataset, creating a composite dataset. This enhanced dataset was then utilized to train the same network again. The study aimed to demonstrate that the model, when trained with data augmented by the use of GANs, surpasses the performance of the model trained without such augmentation. Additionally, the effectiveness of GANs for data augmentation was compared with that of traditional augmentation techniques, and the impact of both augmentation methods was scrutinized.

Evaluation metrics

The classifier quality assessment relied on standard accuracy (Eq. 11) metric.

$$Accuracy = \frac{Number\ of\ correct\ classifications}{Total\ number\ of\ classifications} \quad (11)$$

where: *TP*, *FP* and *FN* are true-positive, false-positive, false-negative values, correspondingly [38].

It is important to emphasize that all tests were performed using only real patient data. Synthetic data were used only for training the network. For all tests, to verify the accuracy of the developed model, leave-one-patient-out cross-validation (LOPOCV) was conducted. It involved training the model without data from one patient, which was then used for testing. This algorithm was repeated for each patient. Although this procedure is computationally intensive, it provides precise and unbiased insights into the model's performance. Through LOPOCV, the root mean squared error (RMSE) was calculated for *n* tests [39]:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \widehat{y_i})^2 \quad (12)$$

where: *n* – indicates number of tests, y_i – is a true value, $\widehat{y_i}$ denotes predicted value.

$$RMSE = \sqrt{MSE} \quad (13)$$

DCGAN augmentation

Proposed DCGAN models, employed a batch size of 32, steps per epochs 375, utilized the Adam optimizer, applied the Wasserstein loss function and included a gradient penalty equals 10. Training was ended when the discriminator's loss reached 0, indicating its inability to differentiate between real and generated images. This criterion was achieved after approximately 180 epochs, as shown in Fig. 6, illustrating the training losses as well as density distribution of the DCGAN trained on a subset of retinitis pigmentosa peripheral changes subjects (referred to as *RP*). Similar training trends were observed in another subset.

The process initialized with the generation of fake images using the DCGAN. Initially, the image resembles a blank canvas. By the 20th iteration, faint shadows began to emerge on the retinitis pigmentosa images. By the 50th iteration, the eye region became more defined. Upon reaching the 100th iteration, the generated image displayed a clear eye image. Subsequently, details gradually emerged throughout subsequent iterations. Ultimately, beyond iteration 180, the generated image exhibited a compelling outcome, as depicted in Figure 7.

Once fully trained, the GANs were assigned the task of generating synthetic images for each class equals 200% of the number of images, in relation to the number of each subset of the dataset. These generated images were then merged with the images from the original dataset to create eight distinct composite datasets, each characterized by a varying ratio of synthetic to authentic images (i.e., 25%, 50%, 75%, 100%, 125%, 150%, 175%, and 200%).

Classification results

The objective of this experiment was to set a basic classification accuracy for future studies. Three dropout probabilities were examined: 0%, 25%, and 50%, aligning with the probabilities detailed in the traditional augmentation section. The outcomes are illustrated in Figure 8. The top-performing model (i.e., without dropout) obtained an accuracy of 72.20%, serving as the "baseline" for future iterations. Introducing dropout seemed to degrade the model's performance by approximately 4%.

Upon setting the initial performance benchmark, the subsequent phase entailed examining the impact of integrating conventional augmentation methods on the efficacy of the models. Used
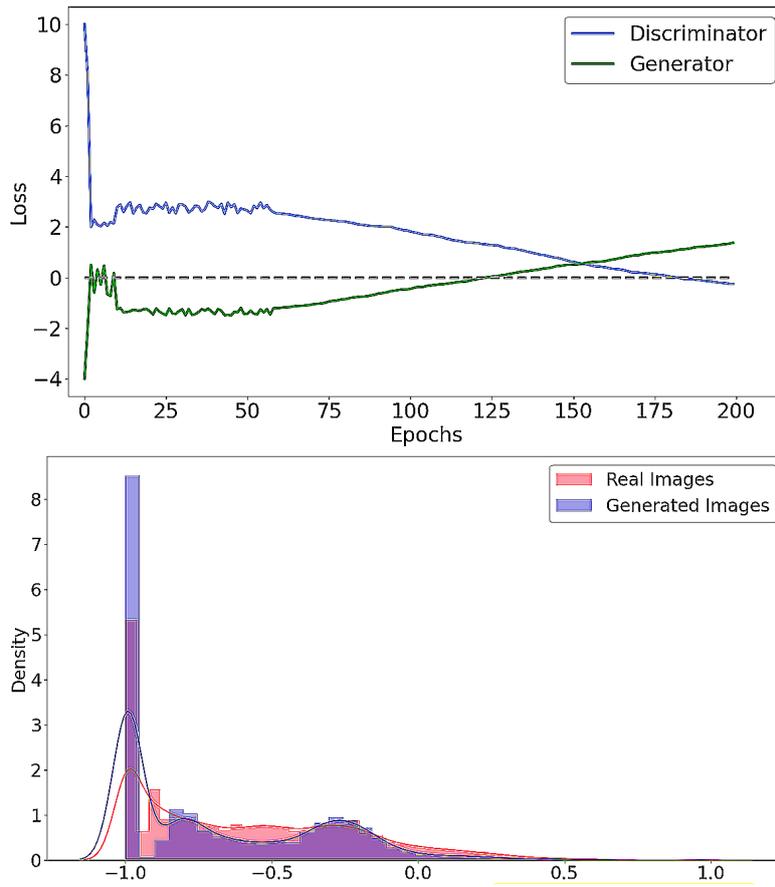
Figure 6. DCGAN loss (top) and density (bottom) on retinitis pigmentosa dataset
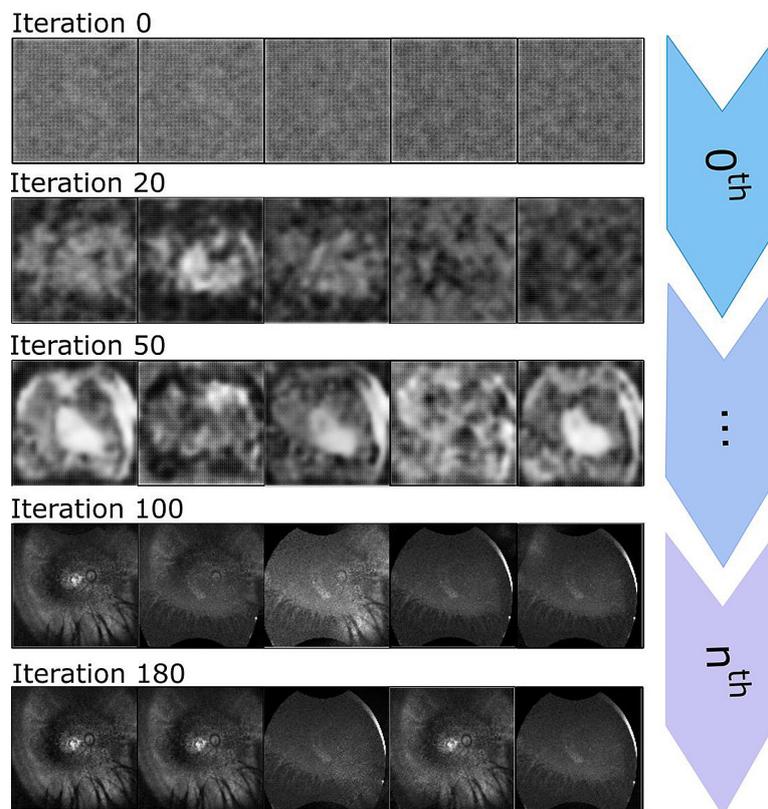
Figure 7. GAN iterations
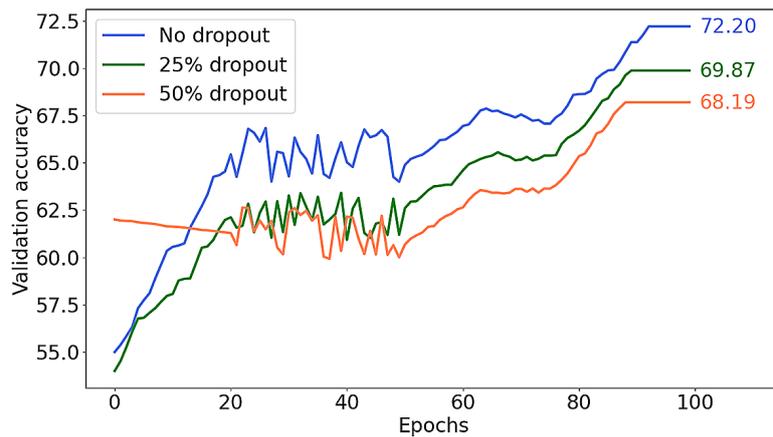
Figure 8. VGG16 with XGBoost trained without augmentation methods. Obtained results will be used as a reference point for further studies

methods are described is section related to traditional augmentation techniques and Figure 5. As illustrated in Figure 9, the comparative performance of two distinct models, one incorporating a 25% dropout rate and the other devoid of dropout, was analysed using the enhanced dataset. The non-dropout model demonstrated higher performance than the others, however, its application may not be as wide due to the potential advantages of regularization in scenarios involving imperfect data.

Contrary to the expectation that augmentation often increases model capabilities, the current study observed that only when a small amount (25%) of the augmentation data and no dropout was used, the baseline was exceeded. This can be attributed to the rigorous nature of OCT image formatting, where even minimal magnification interventions can have a detrimental effect on the classification process. The efficiency of the best model in this part of the

research was 72.89%. Moreover, the augmentation process seemingly worsens the model's ability to converge consistently, as reflected by the pronounced fluctuations in the performance curves depicted in Figure 9. The primary objective of this study was to assess the performance of a VGG16 with XGBoost classifier trained on a dataset enhanced with GAN-generated images. To identify the optimal quantity of synthetic images to include in the original dataset, eight experiments were conducted, as detailed in Section about GAN augmentation. The results of these experiments are illustrated in Figure 10.

Every run outperforms the baseline, demonstrating that GANs can serve as effective data augmentors, even in scenarios where traditional augmentation techniques fall short. The optimal ratio of 75% achieves a score almost 19% higher than the baseline. Additionally, using GAN-generated images for data augmentation does not
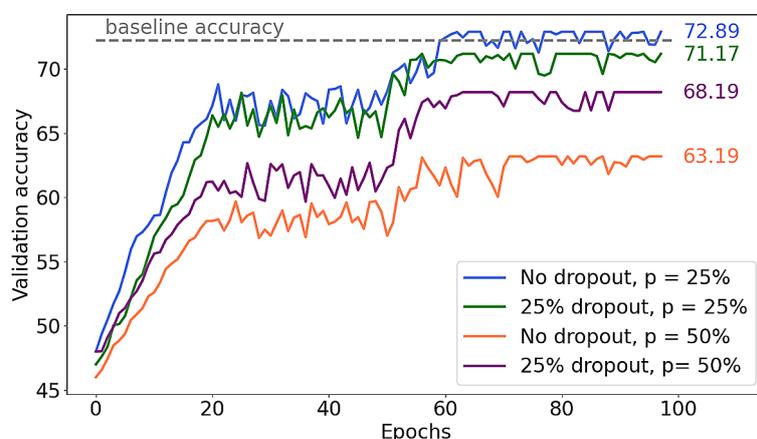
Figure 9. VGG16 with XGBoost trained with traditional augmentation techniques. Two distinct network architectures were analysed across two varying values of *p*. The dash, grey line in the graph signifies the baseline (72.20%) established in the prior experiment
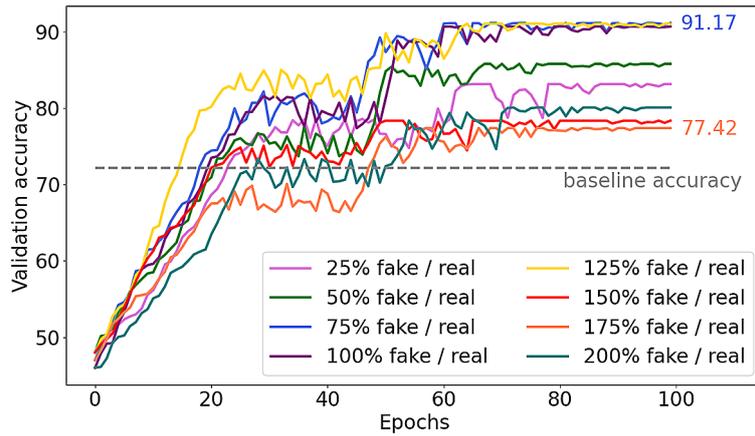
Figure 10. VGG16 with XGBoost trained with DCGAN augmentation techniques. Eight distinct datasets were analysed with different level of fake images and no dropout function. The dash, grey line in the graph signifies the baseline (72.20%)

significantly impact on model convergence, unlike traditional augmentation techniques.

Despite the increased dataset size, nearly all models converged by epoch 80. In contrast, most traditional data augmentation runs (Fig. 11) failed to converge even after 100 epochs. The ideal ratio appears to be between 75% and 125%, with the highest score of 91.17% was obtained by the model trained with a 75% fake/real ratio.

Finally, the combination of traditional and DCGAN-based augmentation methods was evaluated. The optimal model in this instance featured a 25% dropout rate and a 125% fake-to-real ratio, achieved a score of 88.34%. Unlike previous experiments, dropout proved beneficial in this scenario. Figure 11 illustrates the models trained with a 25% dropout rate, $p = 25\%$, across eight different fake-to-real ratios. As shown in Figure 12, the models

incorporating a 25% dropout rate performed marginally better than those without dropout.

Table 1 shows a comparison of classification results when the training set is augmented with images generated using DCGAN and DCGAN combined with traditional augmentation techniques. Eight distinct datasets were analysed, each with varying levels of synthetic images and a 25% dropout function applied during training.

Comparison with the state-of-the-art

GANs are increasingly used in the process of eye disease detection. They are particularly used to enlarge data sets in order to generate synthetic images that reliably represent the disease and its stages. These data are then utilized to train classifiers to improve their efficiency. Selected types
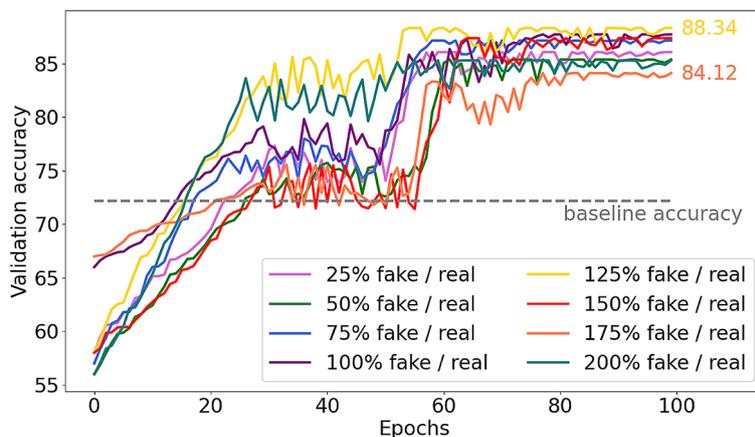
Figure 11. VGG16 with XGBoost trained with DCGAN and traditional augmentation techniques. Eight distinct datasets were analysed with different level of fake images and 25% dropout function and *p*=0.25. The dash, grey line in the graph signifies the baseline (72.20%)
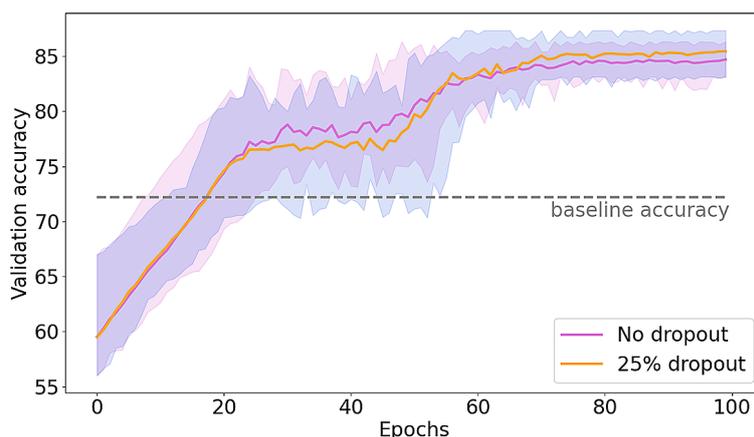
Figure 12. VGG16 with XGBoost trained with DCGAN and traditional augmentation techniques. Two tests for no dropout and 25% dropout with *p*=0.25 were conducted for 8 datasets represent different fake to real images ratio. The thick line represents the mean value, while shaded area was made by the minimum and maximum values achieved in each epoch. The dash, grey line in the graph signifies the baseline (72.20%)

Table 1. Comparison of classification results (accuracy in %) while augmenting the training set with images acquired using DCGAN and DCGAN with traditional augmentation techniques (DCGAN with TA). Eight distinct datasets were analysed with different levels of fake images and 25% dropout function.

Augmentation method	Percentage of artificially generated data							
	25%	50%	75%	100%	125%	150%	175%	200%
DCGAN	83.19	85.83	91.17	90.74	91.09	78.39	77.42	80.11
DCGAN with TA	86.08	85.39	87.17	87.74	88.34	87.39	84.12	85.31

of GANs and their applications are summarized in Table 2. A great majority of studies have been conducted on fundus images from various data sets. Depending on their type, the kinds of GAN used and ML methods, the efficiency of over 78% was achieved. These studies gathered in Table 2 contain different ratios of original to synthetic images. For fundus images, the best detection result was obtained for identify various stages of diabetic retinopathy using GAN-RCNN, achieving an accuracy of 99%. Similarly, in the case of detecting retinal vascular occlusions, using Few-Sample Generator, the efficiency was obtained above 86.30%-100% depending on the dataset used. In the case of studies on OCT images, the efficiency was obtained above 79% for counterfactual GAN.

In this paper, the authors proposed the DCGAN method to augment the fundus images dataset. As a classification method, the VGG16 with XGBoost was proposed, obtaining high accuracy, 91.17%, for detecting RP, CORD and healthy cases. This is the only study comparing to these presented in Table 2, which concerns this rare disease. In this case, GAN methods are highly effective in generating synthetic images

and improving the detection performance. The authors performed an extensive analysis, which proved that GAN methods are more effective in classifying the disease than traditional augmentation methods. Moreover, the effect of dropout on accuracy was investigated. To the best of the authors' knowledge, the ML method in the form of VGG with XGBoost was used for the first time for this type of disease.

DISCUSSION AND CONCLUSIONS

There are many methods of preparing image data for classification purposes, from simple methods that remove distortions [61], to processing using the Fourier transform [62] or the wavelet transform [36, 65].

The article discusses the rapid advancements in computer vision, particularly driven by deep learning techniques, with a focus on achieving state-of-the-art performance through increasingly deep and complex models. Moreover, the importance of medical datasets containing images of rare diseases is highlighted as well as

Table 2. State-of-the art of GAN augmentation for medical images datasets (AF – affine, PX – pixel-level)

Augmentation method	Data type	Improvement [%]	Classification accuracy [%]	Purpose	References
StyleGAN2-ADA	Fundus images	–	78.00–81.00	Enhance AMD detection	[14]
ProGAN	Fundus images	–	82.92–91.12	Enlarge dataset	[19]
GAN-BIGRU	Fundus images	–	98.50	Diabetic retinopathy detection	[26]
GAN-RCNN	Fundus images	–	99.00	Identify various stages of diabetic retinopathy	[27]
Few-Sample Generator	Fundus images	–	86.30–100	Detect retinal vascular occlusions	[18]
Counterfactual GAN	OCT images	–	65.80–79.70	Visualize the individual course of retinal aging	[22]
GAN	PET-MR brain images	10.45	74.00	Alzheimer disease classification	[40]
GAN	MR brain images	7.07	86.30	Alzheimer disease classification	[41]
GAN	MR brain images	11.60	83.49	Alzheimer disease classification	[32]
GAN	MR brain images	3.20	92.00	Cancer classification	[42]
GAN	Coronary angioscopy images	5.19	81.00	Heart disease classification	[43]
DCGAN	Ultrasound scan image	21.43	85.00	Heart disease classification	[44]
GAN-PX-AF	X-ray lung images	14.47	91.00	Pneumothoax classification	[45]
AF-GAN	Computed Tomography and X-ray lung images	8.79	99.00	Covid-19 classification	[46]
AF-PX-GAN	Computed Tomography lung images	8.41	82.50	Cancer classification	[47]
AF-GAN	Computed Tomography lung images	22.87	81.20	Cancer classification	[48]
Custom GAN	Computed Tomography lung images	2.93	95.00	Nodule classification	[49]
Custom GAN	Computed Tomography lung images	5.48	68.71	Cancer classification	[50]
GAN	Computed Tomography lung images	8.91	92.10	Nodule classification	[51]
GAN - AF	X-ray lung images	3.11	83.00	Pneumothoax classification	[52]
GAN	X-ray lung images	2.23	94.10	Pneumothorax and Covid-19 classification	[53]
GAN	Computed Tomography lung images	11.01	68.71	Cancer classification	[54]
GAN	Mammography images	11.25	87.00	Cancer classification	[55]
AF-GAN	Mammography images	6.81	94.00	Cancer classification	[56]
GAN	Computed Tomography breast images	17.27	86.90	Cancer classification	[57]
Custom GAN	Ultrasound scan breast images	1.92	90.40	Cancer classification	[58]
GAN - AF	Mammography images	6.49	88.90	Cancer classification	[59]
Custom GAN	Eye fundus images	4.00	78.00	Eye diseases classification	[60]
DiaGAN-CNN	Eye fundus images	1.0-2.0	98.00	Eye diseases classification	[63]
DC-GAN	Eye fundus images	2.0-6.86	93.80-98.66	Eye diabetic retinopathy	[64]
Custom GAN	Eye fundus images	18.97	91.17	Retinitis pigmentosa classification	Own

the augmentation of medical images using deep convolutional generative adversarial networks (DCGAN) to improve dataset quality and classification accuracy. In studies taking into account enlarged datasets, a noticeable increase in classification accuracy is achieved compared to the initial studies. The top-performing model, without dropout, achieved an accuracy of 72.20%, serving as the baseline for future iterations. Introducing dropout in the models seemed to degrade performance by approximately 4%. The study observed that the optimal ratio of synthetic images to real images was 75%, resulting in a score almost 19% higher than the baseline accuracy, reaching 91.17%. Using GAN-generated images for data augmentation did not significantly impact model convergence, unlike traditional augmentation techniques. The combined augmentation method of traditional and DCGAN-based techniques, with a 25% dropout rate and a 125% fake-to-real ratio, achieved a score of 88.34%, demonstrating the synergistic effect of both augmentation approaches. Models trained with the proposed GAN augmentation methodology outperformed those trained with traditional augmentation techniques by a large margin, showcasing the effectiveness of GAN-based augmentation for improving classification accuracy in medical image datasets. These results highlight the success and implications of the study's methodology in enhancing the performance of machine learning models for medical image analysis, particularly in the context of rare eye diseases. The study faced the challenge of working with a small real dataset due to the nature of rare diseases, in contrast to other datasets focusing on retinal diseases. Despite the limited size of the real dataset, the augmentation techniques, including the use of DCGAN and traditional augmentation methods, were effective in improving classification accuracy and model performance. In the rapidly evolving landscape of healthcare, the significance of data-driven decision-making cannot be overstated. Medical data classification, a critical aspect of modern healthcare analytics, hinges on the availability of high-quality, diverse datasets. These datasets empower machine learning models to identify patterns, predict outcomes, and ultimately, enhance patient care. However, obtaining sufficient medical data for training these models presents formidable challenges, including issues of privacy, data scarcity, and the need

for labelled data. In response to these challenges, data generation techniques have emerged as a pivotal solution. This article delves into the methodologies, applications, and implications of data generation for medical data classification, exploring how synthetic data are transforming the healthcare industry by bridging the gap between data scarcity and the growing demand for accurate, reliable medical insights especially in rare eye diseases classification. The Table 2 presents a brief summary of the article related to the generation of medical data, along with information on the classification tools and the obtained improvement in results.

Based on the results presented in Table 2, it can be concluded that the proposed method significantly enhances the efficiency of retinitis pigmentosa (RP) classification. The developed solution is among the leading approaches in terms of both improvement and accuracy. It is important to acknowledge that retinitis pigmentosa is a rare disease, which limited the dataset available for this study, unlike the larger datasets used in studies such as [44, 48]. Despite this, only marginally better results were obtained in those studies. It is reasonable to assume that expanding the dataset of RP images will lead to further improvements in image quality and, consequently, in classification accuracy. It should be noted that the commonly available VGG-16 model was employed for classification, with adaptations specifically for retinitis pigmentosa images. The development of a dedicated model tailored to this task is likely to yield even greater improvements in classification quality.

Potential limitations

One of the observed limitations is the potential risk of overfitting due to the use of synthetic data generated by the DCGAN. Overfitting occurs when a model learns patterns specific to the training data, including noise and artifacts, which may not generalize well to new, unseen data. The high accuracy achieved might partly reflect the model's proficiency in distinguishing between real and synthetic images rather than capturing the underlying features of retinitis pigmentosa.

Moreover, the synthetic images produced by the DCGAN, while visually similar to real images, may not encompass the full variability present in real-world clinical data. This could

introduce biases, as the GAN might replicate existing biases in the original dataset, such as underrepresentation of disease variations. Consequently, the model's performance might not be consistent across diverse patient populations, potentially limiting its generalizability.

Therefore, it would be necessary to expand the existing set of OCT classes and ensure the effectiveness of the model.

While we report a significant improvement in classification accuracy – surpassing the 90% threshold and improving baseline accuracy by almost 19% – the clinical relevance of this enhancement requires further exploration. In medical diagnostics, even small improvements in accuracy can have substantial impacts on patient outcomes. However, it is crucial to assess whether this improvement translates into better diagnostic decisions, earlier detection, or improved patient management for retinitis pigmentosa. Additionally, factors such as the false-positive and false-negative rates need to be evaluated, as they have direct implication.

The use of synthetic data in medical imaging raises ethical and regulatory considerations. Ensuring patient privacy is essential to maintained, but it is equally important to verify that models trained on synthetic data do not inadvertently incorporate biases or inaccuracies that could affect patient care. Regulatory bodies may require thorough validation of such models before they can be implemented in clinical settings.

Despite the above-mentioned limitations, it should be emphasized that models trained with DCGAN-augmented data exhibited more stable convergence during training compared to those trained with traditional augmentation methods. This stability suggests that the synthetic data generated by the DCGAN aligns well with the underlying data distribution, facilitating more efficient learning. The convergence was generally achieved around the 80th epoch, indicating that the model could effectively learn from the augmented dataset without requiring excessive training time.

The application of dropout did not consistently improve model performance. In some cases, models without dropout achieved higher accuracy. This observation may indicate that, with the enriched dataset provided by DCGAN augmentation, the risk of overfitting was reduced, and the regularization effect of dropout became less critical. Alternatively, dropout may have impeded the model's ability to learn from the complex patterns present in the augmented data.

The substantial improvement in classification accuracy has important implications for the detection of retinitis pigmentosa. Improved diagnostic tools can aid ophthalmologists in identifying this rare disease earlier and more accurately, potentially leading to better patient counselling and management. Additionally, the methodology demonstrated in this study can be applied to other rare diseases where data scarcity is a significant hurdle.

Future directions

To address limitations mentioned before, future research should focus on several aspects:
- External validation: Testing the model on independent datasets from different populations to evaluate its generalizability and robustness.
- Bias mitigation: Implementing techniques to detect and mitigate biases in synthetic data generation, such as incorporating fairness constraints in the GAN training process.
- Clinical impact assessment: Collaborating with clinicians to assess the practical utility of the model, including its integration into diagnostic workflows and its impact on patient outcomes.
- Comparison with other methods: Benchmarking the proposed approach against other advanced models, such as deep ensemble networks, attention mechanisms, or transformer-based architectures.

Moreover, a future study may include experiments with architectures more sophisticated GAN frameworks such as auxiliary classifier GANs, CycleGAN or Progressive Growing GANs to further improve data quality and experimental performance. Subsequent tests may include comparative studies to evaluate the performance of different augmentation techniques on various types of medical image datasets, exploring the strengths and limitations of each approach in different clinical scenarios. Moreover, future work should aim to increase the number of collected retinitis pigmentosa images from real patients obtained from other collaborating research centres.

Overall, the article demonstrates the potential of GAN-based augmentation in improving the quality and accuracy of classification of medical image datasets for rare diseases, paving the way for future research in enhancing machine learning algorithms for healthcare applications.

Acknowledgements

REFERENCES

1. Panayides A.S., Amini A., Filipovic N.D., Sharma A., Tsaftaris S.A., Young A., Foran D., Do N., Golemati S., Kurc T., Huang K., Nikita K. S., Veasey B.P., Zervakis M., Saltz J.H., Pattichis C.S. AI in medical imaging informatics: current challenges and future directions. IEEE J. Biomed. Health 2020, 24.7, 1837–1857. 10.1109/JBHI.2020.2991043.

2. Skublewska-Paszkowska, M., Powroznik, P., Barszcz, M., Dziedzic, K. Identifying and Animating Movement of Zeibekiko Sequences by Spatial Temporal Graph Convolutional Network with Multi Attention Modules. Advances in Science and Technology. Research Journal, 2024, 18(8), 217–227.

3. Nowomiejska K., Powroznik P., Skublewska-Paszkowska M., Adamczyk K., Concilio M., Sereikaite L., Zemaitiene R., Toro M. D., Rejdak R. Residual Attention Network for distinction between visible optic disc drusen and healthy optic discs. Opt. Lasers Eng. 2024, 176, 108056. https://doi.org/10.1016/j.optlaseng.2024.108056.

4. Chen T.C., Lim W.S., Wang V.V., Ko M.L., Chiu S.I., Huang Y.S., Yang M. C., Hu F., Jang J. R., Yang C. H. Artificial intelligence–assisted early detection of retinitis pigmentosa – the most common inherited retinal degeneration. J. Digit. Imaging. 2021, 34, 948–958. 10.1007/s10278-021-00479-6.

5. Li T., Bo W., Hu C., Kang H., Liu H., Wang K., Fu, H. Applicationsof deep learning in fundus images: A review. Med. Image Anal. 2021, 69, 101971.

6. Skublewska-Paszkowska M., Powroznik P., Rejdak R., Nowomiejska K. Application of convolutional gated recurrent units U-net for distin-guishing between retinitis pigmentosa and cone-rod dystrophy. Acta Mech. Autom. 2024, 18, 3.

7. Sun G., Wang X., Xu L., Li C., Wang W., Yi Z., Luo H., Su Y., Zheng J., Li Z., Chen Z., Zheng H., Chen C. Deep learning for the detection of multiple fundus diseases using ultra-widefield images. Ophthalmol. Ther. 2023, 12(2), 895-907.

8. Masumot H., Tabuchi H., Nakakura S., Ohsugi H., Enno H., Ishitobi N., Ohsugi E., Mitamura Y. Accuracy of a deep convolutional neural network in detection of retinitis pigmentosa on ultrawide-field images. PeerJ., 2019, 7:e6900.

9. Jain L., Murthy H. S., Patel C., Bansal D. Retinal eye disease detection using deep learning. In: Proc - Fourteenth International Conference on Information Processing (ICINPRO). IEEE., 2018, 1–6.

10. Li B., Cheng F., Cai H., Zhang X., Cai W. A semi-supervised approach to fault detection and diagnosis for building HVAC systems based on the modified generative adversarial network. Energy Build. 2021, 246, 111044.

11. Goodfellow I., Pouget-Abadie J., Mirza M., Xu B., Warde-Farley D., Ozair S., Courville A. Bengio Y. Generative adversarial nets. Adv. Neural Inf. Process. Syst. 2014, 27.

12. Yan K., Huang J., Shen W., Ji Z. Unsupervised learning for fault detection and diagnosis of air handling units. Energy Build. 2020, 210, 109689.

13. Navaneethan R. and Devarajan H. Enhancingdiabeticretinopathydetection through preprocessing and feature extraction with MGA-CSG algorithm. Expert Syst. Appl. 2024, 249, 123418.

14. Oliveira G. C., Rosa G. H., Pedronette D. C., Papa J. P., Kumar H., L. A. Passos L. A, Kumar D. Robust deep learning for eye fundus images: Bridging real and synthetic data for enhancing generalization. Biomed. Signal Process. Control. 2024, 94, 106263.

15. Veturi Y. A., Woof W., Lazebnik T., Moghul I., Woodward-Court P., Wagner S. K., Cabral de Guimarães T. A., Varela M. D., Liefers B., Patel P. J., Beck S., Webster A. R., Mahroo O., Keane P. A., Michaelides M., Balaskas K., Pontikos N. SynthEye: investigating the impact of synthetic data on artificial intelligence-assisted gene diagnosis of inherited retinal disease. Ophthalmol. Sci. 2023, 3(2), 100258.

16. Das V., Dandapat S., Bora P. K. Unsupervised super-resolution of OCT images using generative adversarial network for improved age-related macular degeneration diagnosis. IEEE Sens. J. 2020, 20(15), 8746–8756.

17. Chaudhuri R. and Deb S. Precise lesion analysis to detect diabetic retinopathy using Generative Adversarial Network (GAN) and Mask- RCNN. Procedia Comput. Sci. 2024, 235, 520–529.

18. Yang, Z. Zhang Y., Xu K., Sun J., Wu Y., Zhou M. DeepDrRVO: A GAN-auxiliary two-step masked transformer framework benefits early recognition and differential diagnosis of retinal vascular occlusion from color fundus photographs," Comput. Biol. Med. 2023, 163, 107148.

19. Burlina P.M., Joshi N., Pacheco K.D., Liu T.A., Bressler, N.M. Assessment of deep generative models for high-resolution synthetic retinal image generation of age-related macular degeneration. JAMA Ophthalmol., 2019, 137(3), 258–264.

20. Mercaldo F., Brunese L., Martinelli F., Santone A.,

Cesarelli M. Gen- erative Adversarial Networks in Retinal Image Classification. Appl. Sci., 2023, 13(18), 10433.

21. Atas I. Comparison of deep convolution and least squares GANs for dia- betic retinopathy image synthesis. Neural. Comput. Appl. 2023, 35(19), 14431–14448.

22. Menten M. J., Holland R., Leingang O., Bogunovic H., Hagag A. M., Kaye R., Riedl S., Traber G. L., Hassan O. N., Pawlowski N., Glocker B., Fritsche L. G., Scholl H. P. N., Sivaprasad S., Schmidt-Erfurth U., Rueckert D., Lotery A. J., PINNACLE Consortium. Exploring healthy retinal aging with deep learning. Ophthalmol Sci. 2023, 3(3), 100294.

23. Diaz-Pinto A., Colomer A., Naranjo V., Morales S., Xu Y., Frangi A.F. Retinal image synthesis and semi-supervised learning for glaucoma assessment. IEEE Trans. Image Process. 2019, 38(9), 2211–2218.

24. Joseph J. and Sreela S. R. MODCN:Fine-tuneddeep-convolutionalneural network with GAN deployed to forecast diabetic eye damage in fundus retinal images. Int. J. Image Graph. 2024, 24(03), 2450029.

25. Choi E. Y., Han S. H., Ryu I. H., Kim J. K., Lee I. S., Han E., Kimf H., Yoo J. Y. C., Yoo T. K. Automated detection of crystalline retinopathy via fundus photography using multistage generative adversarial networks. Biocybern. Biomed. Eng. 2023, 43(4), 725–735.

26. Dohare S., Al Ansari M. S., Naga Ramesh J. V., El-Ebiary Y. A. B., Thenmozhi E. A hybrid GAN-BiGRU model enhanced by african buffalo optimization for diabetic retinopathy detection. Int. J. Adv. Compput. Sc. 2024, 15(1).

27. Krishnamoorthy S., Weifeng Y., Luo J., Kadry S. AO-HRCNN: archimedes optimization and hybrid region-based convolutional neural network for detection and classification of diabetic retinopathy. Artif. Intell. Rev. 2023, 56(Suppl 1), 483–511.

28. Wang S., Li K., Yin Q., Ren J., Zhang J. Semi-supervisedgenerative adversarial learning for denoising adaptive optics retinal images. Biomed. Opt. Express. 2024, 15(3), 1437–1452.

29. Radford A., Metz L., Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv preprint. 2015.

30. Arjovsky M., Chintala S., Bottou L. Wasserstein generative adver- sarial networks. In: International conference on machine learning. 2017, 214–223, PMLR.

31. Basori A.H., Malebary S.J., Alesawi S. Hybrid-DeepConvolutional Generative Adversarial Network (DCGAN) and Xtreme Gradient Boost for X-ray Image Augmentation and Detection. Appl. Sci. 2023, vol. 13(23), 12725.

32. Konidaris F., Tagaris T., Sdraka M., Stafylopatis A. Generative Adversarial Networks as an Advanced Data Augmentation Technique for MRI Data. In: VISIGRAPP (5: VISAPP), 2019, 48–59.

33. Chen T. and Guestrin C. Xgboost: A scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining. 2016, 785–794.

34. Simonyan K. and Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv preprint. 2014.

35. Srivastava N., Hinton G., Krizhevsky A., Sutskever I., Salakhutdinov R. Dropout: a simple way to prevent neural networks from overfitting. J. Mach. Learn. Res. 2014, 15(1), 1929–1958.

36. Kingma D. P. and Ba J. Adam: A method for stochastic optimization. arXiv preprint. 2014.

37. Powroznik P., Wojcicki P., Przylucki S. W. Scalogram as a representation of emotional speech. IEEE Access. 2021, 9, 154044–154057.

38. Plaksyvyi, A., Skublewska-Paszkowska, M., Powroznik, P. A comparative analysis of image segmentation using classical and deep learning approach. Advances in Science and Technology Research Journal, 2023, 17(6).

39. Skublewska-Paszkowska M., Powroznik P. Temporal pattern attention for multivariate time series of tennis strokes classification. Sensors. 2023, 23(5), 2422.

40. Hu S., Yu W., Chen Z., Wang S. Medical image reconstruction using generative adversarial network for Alzheimer disease assessment with class-imbalance problem. In: 2020 IEEE 6th International Conference On Computer And Communications (ICCC). 2020, 1323–1327.

41. Chadebec C., Thibeau-Sutre E., Burgos N., Allassonnière S. Data augmentation in high dimensional low sample size setting using a geometry-based variational autoencoder. IEEE Trans. Pattern Anal. Mach. Intell. 2022, 45, 2879–2896.

42. Anaya-Isaza A., Mera-Jiménez L. Data augmentation and transfer learning for brain tumor detection in magnetic resonance imaging. IEEE Access. 2022, 10, 23217–23233.

43. Miyoshi T., Higaki A., Kawakami H., Yamaguchi O. Automated interpretation of the coronary angioscopy with deep convolutional neural networks. Open Heart. 2020, 7, e001177.

44. Gong Y., Zhang Y., Zhu H., Lv J., Cheng Q., Zhang H., He Y., Wang S. Fetal congenital heart disease echocardiogram screening based on DGACNN: adversarial one-class classification combined with video transfer learning. IEEE T. Med. Imaging. 2–19, 39, 1206–1222.

45. Bhagat V. and Bhaumik S. Data augmentation using generative adversarial networks for pneumonia classification in chest xrays. In: 2019 Fifth International Conference On Image Information Processing (ICIIP). 2019, 574-579.

46. Sedik A., Iliyasu A., Abd El-Rahiem B., Abdel Samea M., Abdel- Raheem A., Hammad M., Peng J., Abd El-Samie F., Abd El-Latif A, Deploying

machine and deep learning models for efficient data-augmented detection of COVID-19 infections. Viruses. 2020, 12, 769.

47. Onishi Y., Teramoto A., Tsujimoto M., Tsukamoto T., Saito K., Toyama H., Imaizumi K., Fujita H. Investigation of pulmonary nodule classification using multi-scale residual network enhanced with 3DGAN- synthesized volumes. Radiol. Phys. Technol. 2020, 13, 160–169.

48. Onishi Y., Teramoto A., Tsujimoto M., Tsukamoto T., Saito K., Toyama H., Imaizumi K.,, Fujita H. Multiplanar analysis for pulmonary nodule classification in CT images using deep convolutional neural network and generative adversarial networks. Int. J. Comput. Assist. Radiol. Surg. 2020, 15, 173–178. https://doi.org/10.1007/s11548-019-02092-z.

49. Tekchandani H., Verma S., Londhe N., Performance improvement of mediastinal lymph node severity detection using GAN and Inception network, Comput. Methods Programs Biomed., 2020, 194, 105478, 202. https://doi.org/10.1016/j.cmpb.2020.105478.

50. Toda R., Teramoto A., Tsujimoto M., Toyama H., Imaizumi K., Saito K., Fujita H., Synthetic CT image generation of shape-controlled lung cancer using semi-conditional InfoGAN and its applicability for type classification, Int. J. Comput. Assist. Radiol. Surg., 2021, 16, 241–251. https://doi.org/10.1007/s11548-021-02308-1.

51. Apostolopoulos I. D., Papathanasiou N. D., Panayiotakis G. S., Classification of lung nodule malignancy in computed tomography imaging utilising generative adversarial networks and semi-supervised transfer learning, Biocybern. Biomed. Eng., 2021, 41, 1243–1257. https://doi.org/10.1016/j.bbe.2021.08.006.

52. Zhang X., Angelini E.D., Haghpanah F.S., Laine A.F., Sun Y., Hiura G.T., Dashnaw S.M., Prince S. W. Hoffman E.A., Ambale-Venkatesh B., Lima J. A., Wild J.M., Hughes E.W., Barr R.G., Shen W., Quantification of lung ventilation defects on hyperpolarized MRI: The Multi-Ethnic Study of Atherosclerosis (MESA) COPD study, Magn. Reson. Imaging, 2022, 92, 140–149. https://doi.org/10.1016/j.mri.2022.06.016.

53. Motamed S., Rogalla P., Khalvati F., Data augmentation using Generative Adversarial Networks (GANs) for GAN-based detection of Pneumonia and COVID-19 in chest X-ray images, Inform. Med. Unlocked, 2021, 27, 100779. https://doi.org/10.1016/j.imu.2021.100779.

54. Toda R., Teramoto A., Kondo M., Imaizumi K., Saito K., Fujita H., Lung cancer CT image generation from a free-form sketch using style-based pix2pix for data augmentation. Sci. Rep., 2022, 12, 12867. 10.1038/s41598-022-16861-5.

55. Desai S.D., Giraddi S., Verma N., Gupta P., Ramya S., Breast cancer detection using gan for limited labeled dataset, In 12th International Conference on Computational Intelligence and Communication Networks (CICN), 2020, 34–39. 10.1109/CICN49253.2020.9242551.

56. KÜÇÜK E. N., UĞur A., Effects of Data Augmentation Techniques on Classification Performance in Knee MRIs, In 5th International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA), 2023, 1–6. 10.1109/HORA58378.2023.10155785.

57. Muramatsu C., Nishio M., Goto T., Oiwa M., Morita T., Yakami M., Kubo T., Togashi K., Fujita, H., Improving breast mass classification by shared data with domain transformation using a generative adversarial network, Comput. Biol. Med., 2020, 119, 103698. https://doi.org/10.1016/j.compbiomed.2020.103698.

58. Pang T., Wong J.H.D., Ng W.L., Chan C.S., Semi-supervised GAN-based radiomics model for data augmentation in breast ultrasound mass classification, Comput. Methods Programs Biomed., 2021, 203, 106018.

59. Guan,Q., Chen Y., Wei Z., Heidari A. A., Hu H., Yang X.H., Zheng J., Zhou Q., Chen H., Chen, F. Medical image augmentation for lesion detection using a texture-constrained multichannel progressive GAN, Comput. Biol. Med., 2022, 145, 105444. https://doi.org/10.1016/j.compbiomed.2022.105444.

60. Araújo T., Aresta G., Mendonça L., Penas S., Maia C., Carneiro A., Mendonça A.M., Campilho, A., Data augmentation for improving proliferative diabetic retinopathy detection in eye fundus images, IEEE Access, 2020, 8, 182462–182474. 10.1109/ACCESS.2020.3028960

61. Prabha D.S., Kumar J.S., Performance analysis of image smoothing methods for low level of distortion, In IEEE International Conference on Advances in Computer Applications (ICACA), 2016, 372–376. 10.1109/ICACA.2016.7887983.

62. Koziel G., Simplified steganographic algorithm based on Fourier transform. Adv. Sci. Lett., 2024, 20(2), 505–509. https://doi.org/10.1166/asl.2014.5322.

63. Shoaib M.R., Emara H.M., Mubarak A.S., Omer O.A., Abd El-Samie F.E., Esmaiel, H. Revolutionizing diabetic retinopathy diagnosis through advanced deep learning techniques: Harnessing the power of GAN model with transfer learning and the DiaGAN-CNN model, Biomedical Signal Processing and Control, 2025, 99, 106790.

64. Devi Y. S., Kumar S. P. Diabetic Retinopathy (DR) Image Synthesis Using DCGAN and Classification of DR Using Transfer Learning Approaches. International Journal of Image and Graphics, 2024, 24(05), 2340009.

65. Powroźnik, P., Czerwiński, D. Spectral methods in Polish emotional speech recognition. Advances in Science and Technology. Research Journal, 2016, 10(32).