



# Label-Preserving Data Augmentation in Latent Space for Diabetic Retinopathy Recognition

Zhihao Zhao<sup>1,2(✉)</sup>, Junjie Yang<sup>1,2</sup>, Shahrooz Faghihroohi<sup>1</sup>, Kai Huang<sup>3</sup>,  
Mathias Maier<sup>2</sup>, Nassir Navab<sup>1</sup>, and M. Ali Nasseri<sup>2(✉)</sup>

<sup>1</sup> TUM School of Computation, Information and Technology, Technical University of  
Munich, Arcisstrasse 21, Munich 80333, Germany

[zhihao.zhao@tum.de](mailto:zhihao.zhao@tum.de)

<sup>2</sup> Klinik und Poliklinik für Augenheilkunde, Technische Universität München,  
IsmaningerStr. 22, München 81675, Germany

[nasseri@in.tum.de](mailto:nasseri@in.tum.de)

<sup>3</sup> School of Computer Science and Engineering, Sun Yat-Sen University,  
Panyu District 132, Guangzhou 510006, China

**Abstract.** AI based methods have achieved considerable performance in screening for common retinal diseases using fundus images, particularly in the detection of Diabetic Retinopathy (DR). However, these methods rely heavily on large amounts of data, which is challenging to obtain due to limited access to medical data that complies with medical data protection legislation. One of the crucial aspects to improve performance of the AI model is using data augmentation strategy on public datasets. However, standard data augmentation methods do not keep the labels. This paper presents a label-preserving data augmentation method for DR detection using latent space manipulation. The proposed approach involves computing the contribution score of each latent code to the lesions in fundus images, and manipulating the lesion of real fundus images based on the latent code with the highest contribution score. This allows for a more targeted and effective label-preserving data augmentation approach for DR detection tasks, which is especially useful given the imbalanced classes and limited available data. The experiments in our study include two tasks, DR classification and DR grading, with 4000 and 2000 labeled images in their training sets, respectively. The results of our experiments demonstrate that our data augmentation method was able to achieve a 6% increase in accuracy for the DR classification task, and a 4% increase in accuracy for the DR grading task without any further optimization of the model architectures.

**Keywords:** Diabetic Retinopathy · Latent Space · Data  
Augmentation

**Supplementary Information** The online version contains supplementary material available at [https://doi.org/10.1007/978-3-031-43898-1\\_28](https://doi.org/10.1007/978-3-031-43898-1_28).

# 1 Introduction

Retinal fundus images are widely used in diabetic retinopathy (DR) detection through a data-driven approach [4]. Specifically, lesions in fundus images play a major role in DR recognition. Some of most prominent lesions associated with DR include microaneurysms, hard exudates, soft exudates, hemorrhages, intraretinal microvascular abnormalities, neovascularization [15]. Hard exudates and soft exudates [3, 21, 23] are commonly observed in the early stages of diabetic retinopathy. Therefore, it is important using data augmentation on the lesions of DR images in AI-based models, especially exudates. However, the size and quantity of exudates can potentially serve as discriminatory indicators of the severity of diabetic retinopathy [18]. As a result, *the most crucial issue is that we need augment the lesions without changing the severity level of DR, especially in the region of exudates.*

Classic image processing methods for data augmentation include random rotation, vertical and horizontal flipping, cropping, random erasing and so on. In some cases, standard data augmentation methods can increase the size of dataset and prevent overfitting of the neural network [8, 22]. But in medical field, conventional data augmentation may damage the semantic content of the image. In ophthalmology diagnosis, the eyes sometimes need to be divided into left and right, and each eye has its own features that are important for image recognition. In this case, we cannot augment the image by flipping method. The same is true for random clipping as the clipped area may not contain any pathological information, which may cause labeling errors. Therefore, the standard methods may damage the semantics of the medical image and also cannot add pathological diversity in the lesion region. Synthetic images generated by GANs can handle the problem of pathological diversity [7, 11], but the generated images are unlabeled. Even if a GAN model is trained separately for different categories of images [2, 5], there is still a possibility of labeling errors in the generated images. Because the generated images are only a sample of the distribution of a given dataset, and the label of generated images are not controlled. Moreover, it is time consuming to train a new model for each category separately. Therefore, *it is important to find a way to use the GAN model to increase the size and diversity of the data while keeping the labels in data augmentation.*

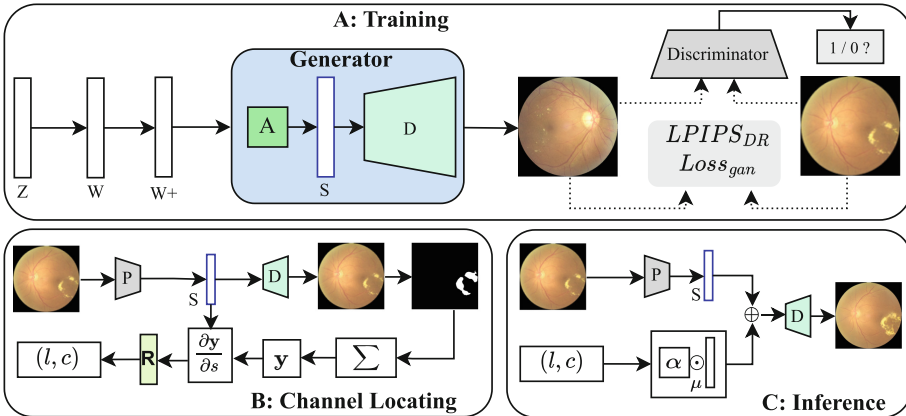
For the aforementioned problems, this paper proposes a data augmentation method for dealing with the class imbalanced problem by manipulating the lesions in DR images to increase the size and diversity of the pathologies while preserving the labels of the data. In this paper, we train a DR image generator based on StyleGAN3 [14] to perform manipulation on lesions in latent space. To make the fundus images after manipulation more realistic in terms of DR pathologies, we add an LPIPS [26] loss based on DR detection to the network. In a well-disentangled latent space, one latent code can control the generation of a specific attribute [6]. In this paper, we mainly focus on exudates augmentation, our purpose is to find the latent code that individually controls the generation of exudates region. Different from StyleSpace [25], we do not need to detect the position of latent code in latent space by calculating the jacobian

matrix of the generated image with respect to latent space. Instead, we set the mask of exudates as the contribution map of generated image to DR pathologies, and apply backpropagation to obtain the contribution map of DR lesion with respect to latent space. Then we analyze the contribution score of each latent code, the position of latent code that controls the exudates region is determined by highest contribution score.

The major contributions of this work can be summarized as follows. Firstly, we introduce LPIPS loss based on DR detection in the training phase of StyleGAN3 to make the reconstructed images have more realistic lesions. Secondly we apply channel locating method with gradient and backpropagation algorithms to identify the position of latent code with the highest contribution score to lesions in DR images. Finally, we perform manipulation on lesions for data augmentation. We also show the performance of proposed method for DR recognition, and explore the correlation between label preservation and editing strength.

## 2 Methods

Figure 1 illustrates our overall structure of the data augmentation method. In part A, StyleGAN3 model is trained to generate DR images. The network maps the noise vector  $Z$  to  $W$  and  $W+$  space through two linear connection layers, and then maps it to  $S$  space via affine transformation layer  $A$ . LPIPS loss based on DR detection was added in training phase. In part B, we use projector  $P$  to embed real image to  $S$  space, then decoder  $D$  trained in part A is used for reconstruction. We also define  $y$  as the importance score of DR pathology (exudates), and then the lesion map can be used as the contribution map of  $y$



**Fig. 1.** The flowchart of the proposed data augmentation method. Part A is the training phase of StyleGAN3 which include additional LPIPS loss based on DR detection. In part B, we detect the position  $(l, c)$  of latent code with highest contribution score to DR pathologies. Part C is the inference phase of image manipulation on  $S$  space of real image.

with respect to output image. The meaning of lesion map is the contribution score matrix of each pixel to DR pathology in output image, because 0 in the lesion map means non-lesion, 1 means lesion. If we want to get the contribution score of each latent code to DR pathology in  $S$  space, we can compute the gradient of  $\mathbf{y}$  with respect to  $S$  space by applying back-propagation. The gradient list of all latent codes is denoted as  $R$  which can be used as a list of contribution score for latent space because of the meaning of lesion map. In part C, we perform manipulation on real image by fusing the latent code of position  $(l, c)$  in  $S$  space and average space  $\mu$  with editing strength  $\alpha$ .

## 2.1 Part A: Training Phase of StyleGAN3

StyleGAN3 has the ability to finely control the style of generated images. It achieves style variations by manipulating different directions in the latent vector space. In latent space of StyleGAN3, the latent code  $z \in Z$  is a stochastic normally distributed vector, always called  $Z$  space. After the fully connected mapping layers,  $z$  is transformed into a new latent space  $W$  [1]. In affine transformation layers of StyleGAN3, the input vectors are denoted as  $W+$  space, and the output vectors are denoted as  $S$  space. In [25], the experiments demonstrate that the **style code**  $s \in S$  can better reflect the disentangled nature of the learned distribution than  $W$  and  $W+$  spaces.

**LPIPS Loss Based on DR Detection.** Due to the good performance of LPIPS in generation tasks [26], we use additional LPIPS loss in generator of StyleGAN3. But the original LPIPS loss is based on the Vgg and Alex network trained on ImageNet, which don't include fundus images. As we want to make the fundus images after manipulation have more realistic DR pathology, we customized the LPIPS loss based on DR detection tasks. Firstly, we train a DR detection network based on VGG16 [24] and select the 5 feature maps after max pooling layers as input for LPIPS. In this way, the loss function will make the generator focus more on lesion of fundus images. At last, we add original StyleGAN loss and LPIPS loss together when training.

$$\mathcal{Loss} = \lambda_{lips} \mathcal{L}_{lips} + \lambda_{gan} \mathcal{L}_{gan} \quad (1)$$

## 2.2 Part B: Detecting the Channel of Style Code Having Highest Contribution to Exudates

As shown in Fig. 1, we define  $\mathbf{y} = f(x)$  as the function to compute the importance score of DR pathology for  $x$ , so  $\mathbf{y}$  represents the importance score. The meaning of lesion map is the contribution score matrix of each pixel to DR pathology in output image. In the paper, because the binary segmentation mask only values 1 and 0. the value 1 can give the meaning of contribution, 0 means no contribution. So we can regard binary semantic segmentation mask as a contribution matrix. Due to these prerequisites, the binary semantic mask has same meaning with

gradient map. Because of these characteristic that we can consider set the mask as the contribution matrix of  $\mathbf{y}$  with respect to the output image.

$$\begin{cases} mask = \mathbf{y}' = \frac{\partial f}{\partial x} \\ img = x \end{cases} \quad (2)$$

When we get output image from the generator of a well-trained StyleGAN3, we apply a pre-trained exudates segmentation network based on Unet [20] to obtain the mask of the exudates. As Eq. 3 shown, we can compute the importance score  $\mathbf{y}$  by integral calculus. But images and vectors in middle of StyleGAN3 are not continuous. We also do not know how function  $f$  works, the things we know are  $x$  is the output image, and lesion mask means the contribution matrix. And if a pixel value in gradient map of output image is 0, then the pixel in output image will also contribute 0 to the lesion because of forward propagation in network. So we can treat lesion map as gradient matrix of  $y$ . Since the lesion map has the property that 1 means contribution and 0 means no contribution, and we want to compute the importance score of DR pathology for  $x$ . We can rewrite the integral calculus as computing the sum of all pixels in the image based on the mask. Here,  $\odot$  denote the pixel-wise dot product.  $M$  is the number of pixels in image.

$$\mathbf{y} = \int_0^{M-1} x dx = \sum_{n=0}^{M-1} img \odot mask \quad (3)$$

After  $\mathbf{y}$ , the importance score of DR pathology is computed. We can calculate the contribution of each latent code to DR pathology in  $S$  space by computing the gradient of  $\mathbf{y}$  with respect to each latent code in  $S$  space, based on back-propagation algorithm in neural network as shown in the Eq. 4.

$$\mathbf{R}_k = \mathbf{y}^{L-l} = \frac{\partial^{L-l} f}{\partial S_k^{L-l}} \quad (4)$$

where  $\mathbf{R}$  is the contribution score list of latent codes in  $S$  space,  $k \in K$  is the dimension of  $S$  space,  $\mathbf{R}_k$  is the contribution score of  $k$ th-dimension of  $S$  space. In our paper,  $S$  space has  $K$  dimensions in total including 6048 dimensions of feature maps and 3040 dimensions of tRGB blocks. So, in StyleGAN3, depending on the structure of the network, the  $K$  dimensions can be correspond to different layers  $l$  and channels  $c$  depending on the number of channels in each layers.  $L$  is the number of layers in StyleGAN3. Since we want to do image editing without affecting other features of the image as much as possible to avoid changes in image label. From the experiments in [6], a well disentanglement latent space, one latent code can control the generation of a specific attribute. So when editing the image, it is not necessary to manipulate on whole  $S$  space. We can only modify the style code with highest contribution score.

To ensure that the collected style codes are not impacted by individual noise and are consistent across all images, we calculate the average of 1K images. Afterwards, the collected contribution list  $\mathbf{R}$  are sorted from high to low based

on contribution score of each latent code. To make it easier to understand we can convert  $k$  to  $(l, c)$  based on the layers and channels in StyleGAN3 network structure.

$$(l, c) = k = \arg \max_{k \in K} \sum_{i=0}^{1K} R_k^i \quad (5)$$

### 2.3 Part C: Manipulation on Real Images in Inference Phase

We project the real images to style spaces using projector provided by Karras et al. [14]. Based on the position  $(l, c)$  of contribution score in  $\mathbf{R}$ , we perform manipulation on the layer and channel of highest contribution score when the data flow forward in decoder of StyleGAN3  $D(\cdot)$ .

$$\begin{aligned} S_{new} &= S \odot (1 - m) + [S + \alpha\mu] \odot m \\ I^{aug} &= G(S_{new}) \end{aligned} \quad (6)$$

where  $S_{new}$  denotes the  $S$  space after manipulation,  $I^{aug}$  denotes the generated image after manipulation.  $\alpha$  is the editing strength.  $\mu$  is the mean vector of style space, which is obtained by projecting lots of real images to  $S$  space.  $m$  is a special vector with same dimension with  $S$ , the value of all elements in the vector is 0 except for the element at the  $(l, c)$  position, which is equal to 1. It means which layer and channel is selected.

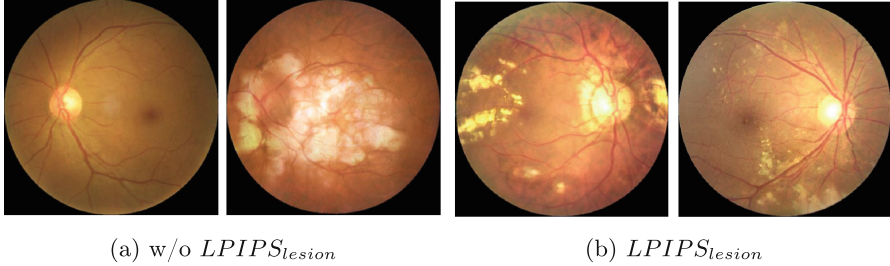
## 3 Experiments

### 3.1 Setup Details

**Dataset.** We carry out all experiments on publicly available datasets. We train StyleGAN3 on EyePACS [9] dataset. We use EyeQ [10] to grade the quality of EyePACS and select 12K images for training and 8K images for testing from the dataset that has been classified as “Good” or “Usable”. For evaluation of our data augmentation method. ODIR-5K (Ocular Disease Intelligent Recognition) [17] dataset was used for DR classification tasks. It include 6392 images with different ocular diseases of which 4000 are training set and 2392 are testing set. APTOS 2019 [13] dataset was used for grading evaluation of diabetic retinopathy as no DR, mild, moderate, severe and proliferative DR. There are 3662 image with labels in total, we split 2000 as training set and 1662 as testing set.

**Evaluation Metric.** We employ Frechet Inception Distance (FID) [12] and Learned Perceptual Image Patch Similarity (LPIPS) [26] to measure the distance between synthetic and real data distribution. FID and LPIPS can evaluate the quality of images created by a generative model. To measure the performance of human perception score, we apply Precision, Recall, Accuracy and F1 score as the evaluation metric.

**Experimental Settings.** All experiments are conducted on a single NVIDIA RTX A5000 GPU with 24GB memory using PyTorch implementation. For StyleGAN3, we initialize the learning rate with 0.0025 for  $G(\cdot)$  and 0.002 for  $D(\cdot)$ . For optimization, we use the Adam optimizer with  $\beta_1 = 0.0$ ,  $\beta_2 = 0.99$ . We empirically set  $\lambda_{l_{lips}} = 0.6$ ,  $\lambda_{gan} = 0.4$ . For evaluation of classification and detection tasks, we both initialize the learning rate with 0.001 for the training of VGG16.



**Fig. 2.** Visual results of StyleGAN3.

### 3.2 Evaluation and Results

**Results of Synthetic Images.** For visual evaluation of generation of StyleGAN3, we show some synthetic images in Fig. 2. Generated images in Fig. 2a show the image without LPIPS loss in training phase. The exudates are sometimes not in the reasonable region, or even not generated. But the synthetic images are visually comparable to real fundus images, in Fig. 2b. For subjective quality evaluation, we also conduct a user study by randomly choosing 200 images which consists of 100 synthetic images and 100 real images, then professional ophthalmologists are asked to determine which images are synthetic. We evaluate the human perception score by Accuracy, Precision, Recall and F1 score in Table 1. Evidently, the proposed loss function based on LPIPS significantly improves the quality of generated images. And from the human perception score, we can also know that even for professional doctors there is only a 50% possibility to predict which one is real and which one is synthetic. This means that synthetic images might be highly confusing even for experienced clinicians.

**Table 1.** Image Quality Assessment

	Objective Quality Metrics		Human Perception Score			
	FID(↓)	LPIPS(↓)	Acc	Recall	Precision	F1
w/o LPIPS	9.329	0.457	0.595	0.760	0.571	0.652
LPIPS	6.459	0.382	0.460	0.462	0.490	0.476

**Results of Data Augmentation.** In our work, we perform the data augmentation method in two tasks. One is the classification of DR images based on ODIR-5K dataset, another is the grading of the DR severity in APTOS-2019. The distribution of the images in ODIR-5K dataset is  $\{34.2, 65.8\}\%$  for  $\{\text{DR}, \text{NDR}\}$  in training set and  $\{31.5, 68.5\}\%$  for  $\{\text{DR}, \text{NDR}\}$  in testing set. The distribution of the images per grade in APTOS-2019 is  $\{47.6, 10.8, 27.8, 5.3, 8.5\}\%$  for  $\{0:\text{No DR}, 1:\text{Mild}, 2:\text{Moderate}, 3:\text{Severe}, 4:\text{PDR}\}$  in training set, and  $\{51.4, 9.4, 26.5, 5.2, 7.5\}\%$  in testing set, respectively.

In ODIR-5k dataset, we perform the manipulation on DR image in training dataset to make DR images have same number with NDR images [16]. And then we train a classification network based on VGG16 with training dataset. We also perform manipulation on APTOS-2019 dataset for level 2–4. It is important to note that, level 1 in APTOS-2019 means mild severity of NPDR, the main lesion is microaneurysms in this stage. Therefore we will not edit the image at this stage because the exudation has not yet occurred. We compare our data augmentation method with standard method in these two tasks. Quantitative evaluation results are shown in Table 2. The results show that standard data augmentation method will not improve the accuracy too much, even the rotation ( $-30^\circ, 30^\circ$ ) will reduce the accuracy. Patho-GAN [19] can not perform augmentation in detection task, because it is not label-preserving method. In detection task, we calculate the accuracy for each level of severity. In level 2 and 3, the accuracy improvement is significant, especially level 3 is most imbalanced in original dataset. In level 4, the accuracy also have been improved, even though the main lesion in this stage is the growing of new abnormal blood vessels, augmenting the dataset can still improve the accuracy of recognition after changing the exudation area. The accuracy of level 0 and 1 have not changed much, the accuracy of level 1 has even decreased, and the reason for this phenomenon is most likely due to the imbalance of data categories caused by not performing augmentation on level 1. This also shows that data imbalance does have an impact on accuracy.

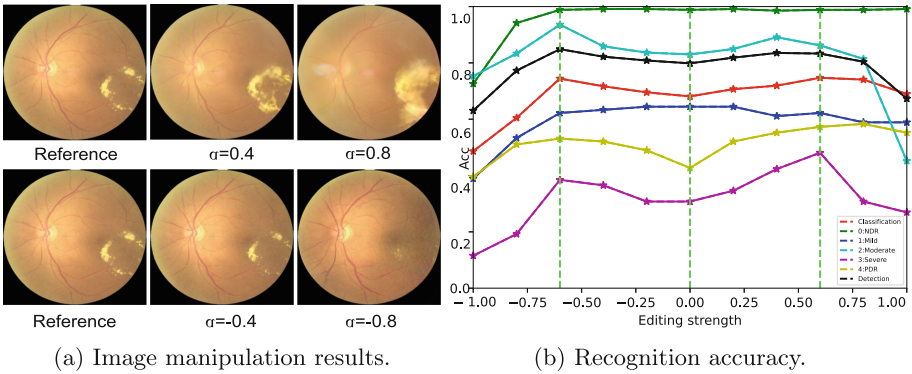
**Table 2.** Comparison with standard data augmentation.

	Classification(ACC)	Detection(level: 0-4 ACC)					
	DR	NoDR	Mild	Moderate	Severe	PDR	ALL
No	0.681	0.988	0.644	0.830	0.307	0.427	0.798
Flipping	0.689	0.991	0.611	0.846	0.153	0.541	0.805
rotation	0.676	0.988	0.688	0.772	0.307	0.468	0.787
Patho-GAN	0.729	-	-	-	-	-	-
Ours	0.743	0.991	0.611	0.891	0.480	0.552	0.841

**Ablation Study.** Since our approach is label-preserving data augmentation, it is more concerned with whether our augmentation method affects the label of the image. In order to investigate the correlation between label preservation



and editing strength, we conduct comprehensive ablation studies to examine the effect of different editing strength  $\alpha$  on the recognition accuracy. As we know, if we change exduates too much, the labels of images will be changed. Such as, level 1 will change to level 2, level 2 will change to level 3. If such a thing happens, our method will not only have no improvement on the accuracy rate, but will also have a negative impact. The Fig. 3a show the manipulation results with different editing strength. We also conduct experiments of effect on the accuracy with different editing strength as shown in Fig. 3b. Results show that when editing strength are in the range of  $(-0.6, 0.6)$ , the accuracy will be improved, otherwise the accuracy is significantly reduced.



**Fig. 3.** Effects of different editing strengths on DR recognition.

## 4 Conclusions

In this paper we propose a label-preserving data augmentation method to deal with the classes imbalanced problem in DR detection. The proposed approach computes the contribution score of latent code, and perform manipulation on the lesion of real images. In this way, the method can augment the dataset in a label-preserving manner. It is a targeted and effective label-preserving data augmentation approach. Although the paper mainly discusses exudates, the proposed method can be applied to other types of lesions as long as we have the lesion mask and the GAN model can generate realistic lesions.

**Data Declaration.** Data underlying the results presented in this paper are available in APTOS [13], ODIR-5K [17], EyePACS [9].

## References

1. Abdal, R., Qin, Y., Wonka, P.: Image2stylegan: how to embed images into the stylegan latent space? In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 4432–4441 (2019)

2. Agustin, T., Utami, E., Al Fatta, H.: Implementation of data augmentation to improve performance CNN method for detecting diabetic retinopathy. In: 2020 3rd International Conference on Information and Communications Technology (ICOIACT), pp. 83–88. IEEE (2020)
3. Akram, U.M., Khan, S.A.: Automated detection of dark and bright lesions in retinal images for early detection of diabetic retinopathy. *J. Med. Syst.* **36**, 3151–3162 (2012)
4. Alyoubi, W.L., Shalash, W.M., Abulkhair, M.F.: Diabetic retinopathy detection through deep learning techniques: a review. *Inf. Med. Unlocked* **20**, 100377 (2020)
5. Araújo, T., et al.: Data augmentation for improving proliferative diabetic retinopathy detection in eye fundus images. *IEEE Access* **8**, 182462–182474 (2020)
6. Bau, D., et al.: Gan dissection: visualizing and understanding generative adversarial networks. *arXiv preprint [arXiv:1811.10597](https://arxiv.org/abs/1811.10597)* (2018)
7. Bellemo, V., Burlina, P., Yong, L., Wong, T.Y., Ting, D.S.W.: Generative Adversarial Networks (GANs) for Retinal Fundus Image Synthesis. In: Carneiro, G., You, S. (eds.) *ACCV 2018*. LNCS, vol. 11367, pp. 289–302. Springer, Cham (2019). [https://doi.org/10.1007/978-3-030-21074-8\\_24](https://doi.org/10.1007/978-3-030-21074-8_24)
8. Elloumi, Y.: Cataract grading method based on deep convolutional neural networks and stacking ensemble learning. *Int. J. Imaging Syst. Technol.* **32**(3), 798–814 (2022)
9. EyePACS: diabetic retinopathy detection. <https://www.kaggle.com/datasets/tanlikesmath/diabetic-retinopathy-resized> (2015), Accessed 20 July 2015
10. Fu, H., et al.: Evaluation of Retinal Image Quality Assessment Networks in Different Color-Spaces. In: Shen, D., et al. (eds.) *MICCAI 2019*. LNCS, vol. 11764, pp. 48–56. Springer, Cham (2019). [https://doi.org/10.1007/978-3-030-32239-7\\_6](https://doi.org/10.1007/978-3-030-32239-7_6)
11. Guo, J., Pang, Z., Yang, F., Shen, J., Zhang, J.: Study on the method of fundus image generation based on improved GAN. *Math. Probl. Eng.* **2020**, 1–13 (2020)
12. Heusel, M., Ramsauer, H., Unterthiner, T., Nessler, B., Hochreiter, S.: GANs trained by a two time-scale update rule converge to a local nash equilibrium. *Adv. Neural Inf. Process. Syst.* **30** (2017)
13. Hospital, A.E.: APTOS 2019 Blindness Detection. <https://www.kaggle.com/competitions/aptos2019-blindness-detection> (2019), Accessed 27 June 2019
14. Karras, T., et al.: Alias-free generative adversarial networks. *Adv. Neural Inf. Process. Syst.* **34**, 852–863 (2021)
15. Lin, L., et al.: The SUSTech-SYSU dataset for automated exudate detection and diabetic retinopathy grading. *Sci. Data* **7**(1), 409 (2020)
16. Mungloo-Dilmohamud, Z., Heenaye-Mamode Khan, M., Jhumka, K., Beedassy, B.N., Mungloo, N.Z., Peña-Reyes, C.: Balancing data through data augmentation improves the generality of transfer learning for diabetic retinopathy classification. *Appl. Sci.* **12**(11), 5363 (2022)
17. NIHDS-PKU: Ocular Disease Intelligent Recognition ODIR-5K. <https://odir2019.grand-challenge.org/> (2019)
18. Niu, S., et al.: Multimodality analysis of hyper-reflective foci and hard exudates in patients with diabetic retinopathy. *Sci. Reports* **7**(1), 1568 (2017)
19. Niu, Y., Gu, L., Zhao, Y., Lu, F.: Explainable diabetic retinopathy detection and retinal image generation. *IEEE J. Biomed. Health Inf.* **26**(1), 44–55 (2021)
20. Ronneberger, O., Fischer, P., Brox, T.: U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) *MICCAI 2015*. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015). [https://doi.org/10.1007/978-3-319-24574-4\\_28](https://doi.org/10.1007/978-3-319-24574-4_28)

21. Santhi, D., Manimegalai, D., Parvathi, S., Karkuzhali, S.: Segmentation and classification of bright lesions to diagnose diabetic retinopathy in retinal images. *Biomed. Eng. (Berl)* **61**(4), 443–453 (2016)
22. Shorten, C., Khoshgoftaar, T.M.: A survey on image data augmentation for deep learning. *J. Big Data* **6**(1), 1–48 (2019)
23. Sidibé, D., Sadek, I., Mériaudeau, F.: Discrimination of retinal images containing bright lesions using sparse coded features and SVM. *Comput. Bio. Med.* **62**, 175–184 (2015)
24. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. *arXiv preprint [arXiv:1409.1556](https://arxiv.org/abs/1409.1556)* (2014)
25. Wu, Z., Lischinski, D., Shechtman, E.: Stylespace analysis: disentangled controls for stylegan image generation, In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp. 12863–12872 (2021)
26. Zhang, R., Isola, P., Efros, A.A., Shechtman, E., Wang, O.: The unreasonable effectiveness of deep features as a perceptual metric, In: *CVPR* (2018)