On the Detection of Retinal Image Synthesis Obtained Through Generative Adversarial Network

Marcello Di Giammarco[1,2], Antonella Santone[3],
Mario Cesarelli[4], Fabio Martinelli[5] and Francesco Mercaldo[1,3]
[1]*Institute for Informatics and Telematics (IIT), National Research Council of Italy (CNR), Pisa, Italy*
[2]*Department of Information Engineering, University of Pisa, Pisa, Italy*
[3]*Department of Medicine and Health Sciences "Vincenzo Tiberio", University of Molise, Campobasso, Italy*
[4]*Department of Engineering, University of Sannio, Benevento, Italy*
[5]*Institute for High Performance Computing and Networking (ICAR), National Research Council of Italy (CNR), Rende (CS), Italy*
{*marcello.digiammarco, francesco.mercaldo*}*@iit.cnr.it*, {*antonella.santone, francesco mercaldo*}*@unimol.it,*

Keywords: Adversarial Machine Learning, Generative Adversarial Networks, Deep Learning, Retinal Imaging, Robustness, Security.

Abstract: Adversarial machine learning on medical imaging is one of the many applications for which the evaluation of Generative Adversarial Networks in the medical field has demonstrated remarkable interest. This paper proposes a method in which Convolutional neural Networks are trained and tested on the binary classification of real and fake images, generated through generative adversarial networks. In this paper, the considered experiments are on the RGB fundus retina images of the human eye. Results highlight networks with optimal performance, and completely recognize real/fake classification; however, on the other hand, other networks mis-classify the images, enhancing security and reliability problems.

1 INTRODUCTION

Generative Adversaries Networks (GANs) are an exciting recent innovation in machine learning. GANs are generative models: they create new instances of data similar to your training data. The number of novel GAN concepts, methods, and applications has increased significantly the interest to these systems This huge success is attributed to the high similarity of the generated images to the real ones (Goodfellow et al., 2020). For the diagnosis and identification of diseases, medical imaging is crucial (Pan and Xin, 2024), (Huang et al., 2024), (Zhou et al., 2023), (Brunese et al., 2022b). Furthermore, research on artificial intelligence using such images could potentially have a beneficial effect on improvements in healthcare. One of the challenges of machine learning in biomedical imaging regards the discrimination ability between real and fake images, which can alter the performance and invalidate the results. To address this challenge, the fake images can be generated with GANs. Moreover, GANs reveal their features in many applications in healthcare and biomedical domains, from data augmentation to anomaly detection passing through medical image synthesis.

In this paper we propose a method able to recognize the real/fake classification. Fake images are generated at different epochs of the GAN considering RGB fundus retina image datasets. In the experiments, resistance and mis-classifications of the models in the real/fake discriminations are obtained. Moreover, the paper analyzed also trends of the output model increasing the epochs of the GAN, showing interesting results.

In the next section, is reported the related works; while a brief overview of the GAN and its working principles is provided in Section 3. The proposed method, applied in the RGB retinal images, is presented in Section 4, followed by the experimental analysis and results in Section 5. Lastly, a conclusion and future research goals are shown in the last section.

Di Giammarco, M., Santone, A., Cesarelli, M., Martinelli, F. and Mercaldo, F.
On the Detection of Retinal Image Synthesis Obtained Through Generative Adversarial Network.
DOI: 10.5220/0013228300003911

In *Proceedings of the 18th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2025) - Volume 1*, pages 611-618
ISBN: 978-989-758-731-3; ISSN: 2184-4305

2 RELATED WORKS

A review of the literature on GAN use in medical field is taken advantage of in this section.

The review of Singh et al.(Singh and Raza, 2021) reports several GAN applications in the medical field. The most recent developments in GAN-based clinical applications for cross-modality synthesis and medical image production are presented in this chapter. Deep convolutional GAN (DCGAN), Laplacian GAN (LAPGAN), pix2pix, CycleGAN, and unsupervised image-to-image translation model (UNIT) are among the GAN frameworks that have become popular for medical image interpretation. They continue to enhance their performance by adding more hybrid architecture.

Recent work is represented by Sai Khal et al. (Akhil et al., 2024). In their work, authors investigated the synthesis of realistic and superior chest X-ray pictures using DCGAN. Their DCGAN model generated synthetic images that exhibited remarkable visual resemblances to real chest X-ray images, including anatomical characteristics, radiographic noise, and disease patterns. They also employed a range of evaluation metrics to statistically examine the diversity and realism of the generated images, demonstrating that their approach produced incredibly realistic synthetic data.

Looking for the retinal imaging domain, in (Diaz-Pinto et al., 2019) authors proposed a novel retinal image synthesizer and a semi-supervised learning technique for glaucoma evaluation based on deep convolutional GANs was innovative. Furthermore, their system is trained on an unparalleled quantity of publicly accessible photos (86926 images), as far as we are aware. As a result, their technology can automatically offer labels in addition to producing synthetic visuals.

In terms of security, the research of Mirsky et al. (Mirsky et al., 2019) represents a great example of the DCGAN application. In their article, they demonstrated how an attacker can modify or delete medical condition evidence from volumetric (3D) medical scans using deep learning. This could be done by an attacker to thwart a political candidate, ruin research, perpetrate insurance fraud, carry out terrorism, or even kill someone. The authors demonstrated how to automate the framework (CT-GAN) and used a 3D conditional GAN to execute the assault. They concentrated on injecting and extracting lung cancer from CT scans to assess the attack. Additionally, they investigated the attack surface of a contemporary radiology network and illustrated one attack vector: they used a clandestine penetration test to intercept and alter CT scans on an operational hospital network.

In the work of Nagaraju et al. (Nagaraju and Stamp, 2022), authors presented a method that employed GANs to create fake malware images and evaluate how well different classification methods worked for these images. Their results demonstrated that while the resulting multiclass classification problem is challenging, they were able to obtain convincing results when they restricted the problem to differentiating between authentic and fraudulent samples. The paper's main finding is that, from a deep learning standpoint, GAN-generated images fall short of deep fake malware images, even if they may resemble real malware images quite a bit.

3 GAN BACKGROUND

Since the proposed method focuses on the GAN, Figure 1 reports a block schematization of the GAN operation.

A GAN main working principle is based on a framework for competitive learning that combines two neural networks: the discriminator and the generator. By learning to map the latent space to the data space, the generator creates synthetic data samples that resemble the real data distribution given random noise (latent vectors) as input. It produces low-quality samples at first, but with training, it improves its ability to produce more realistic samples. In order to differentiate between actual and GAN-generated data samples, the discriminator works as a binary classifier. It learns to give real samples high probability and fraudulent samples low probabilities. The discriminator may function randomly at first, but it learns to distinguish between real and fraudulent samples with practice. The discriminator and generator are simultaneously and competitively trained during the training procedure. The generator's main objective is to trick the discriminator by producing samples that are indistinguishable from actual data. The discriminator, on the other hand, seeks to distinguish between authentic and fraudulent samples correctly. Both networks gradually improve during the training: the discriminator gets better at distinguishing real samples from fake ones, and the generator enhances at producing real examples. The discriminator and generator engage in an adversarial interaction in which the discriminator aims to maximize its ability to distinguish real from fake samples, while the generator tries to minimize the discriminator's ability to distinguish real and synthetic samples by producing increasingly authentic samples. When the generator generates samples that are statistically comparable to actual data and the

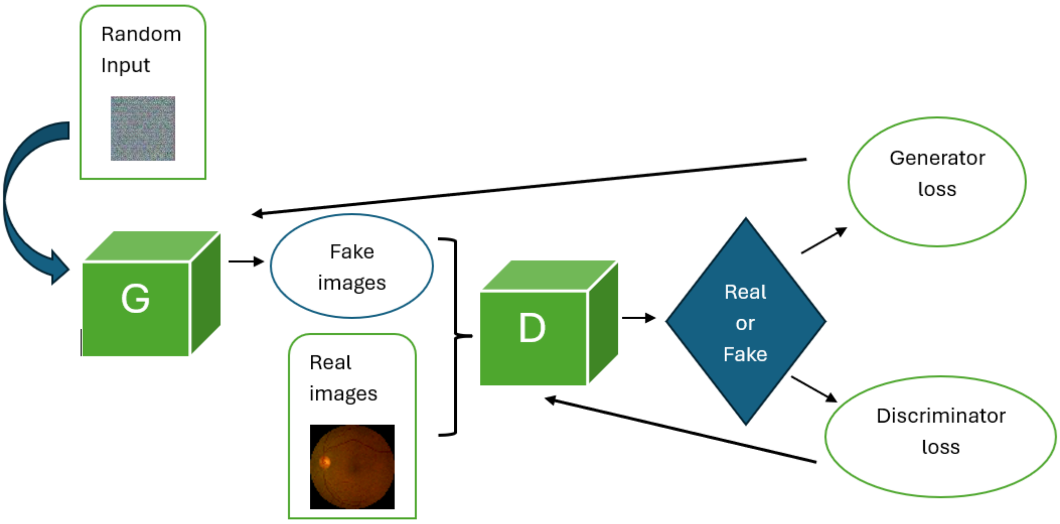

Figure 1: Block schematization of GAN.

discriminator finds it difficult to consistently distinguish between the original and GAN-generated samples, this adversarial process results in a Nash equilibrium. When the discriminator is unable to differentiate between synthetic and genuine samples more accurately than chance, and the generator produces samples that are exactly like real data, the GAN has ideally reached convergence. Convergence can be difficult to achieve and requires careful tweaking of training procedures, network architectures, and hyperparameters. The adversarial training dynamics between the discriminator and generator, which result in the creation of high-quality synthetic data, are essentially the fundamental working principle of a GAN (You et al., 2022; Wang et al., 2023; Singh and Raza, 2020).

This behavior determines a metric for how well the generator is producing accurate data with generator loss. The difference between the generated and actual data is used to calculate this loss. Rather, discriminator loss is a metric that quantifies the discriminator's ability to distinguish between created and genuine data. It is calculated using the discriminator's accuracy in identifying authentic or fraudulent samples. (Mercaldo et al., 2024b)

Following this broad overview, we consider a particular kind of GAN known as DCGAN (Deep Convolutional Generative Adversarial Network (Fang et al., 2018; Liu et al., 2022)) in our work. By adding deep convolutional neural networks, DCGAN is an expansion of the GAN concept, although the working principles are the same as previously shown. Compared to previous GAN architectures, DCGANs can generate high-quality images, which has led to their widespread application for tasks like image production, super-resolution, and style transfer. By utilizing convolutional layers (Brunese et al., 2022a), (Martinelli et al., 2022),(Mercaldo et al., 2024a), (Di Giammarco et al., 2024a), (Mercaldo et al., 2022), hierarchical feature learning, and architectural improvements like batch normalization and Leaky ReLU, the DCGAN architecture strongly improves the quality of generated images.

4 THE METHOD

In this section, our method is presented and exploited. The method is based on the discrimination between GAN-generated images and real ones in the biomedical environment and aims to obtain an output model able to distinguish those biomedical images.

RGB images of the fundus retina represent the case study, but the concept behind is available for any biomedical imaging. The main steps are shown in Figure 2.

The approach starts from the dataset. The dataset consists of an RGB fundus retina and is called the original dataset. This dataset is the ”real dataset” cited in Figure 1 and several generations of fake images, varying the number of epochs, are taken into account for the evaluation goal of the CNNs. In Figure 2 are reported the epochs, from 250 to 500; and the GAN-generated fake images samples more and more similar to the original ones as the epochs increase. 2,000 images are obtained at the end of the adversarial image generation, beginning with the first epoch where the distortion is fully superimposed on the input image and ending with the last epoch where the distortion is undetectable to medics. In this sense,

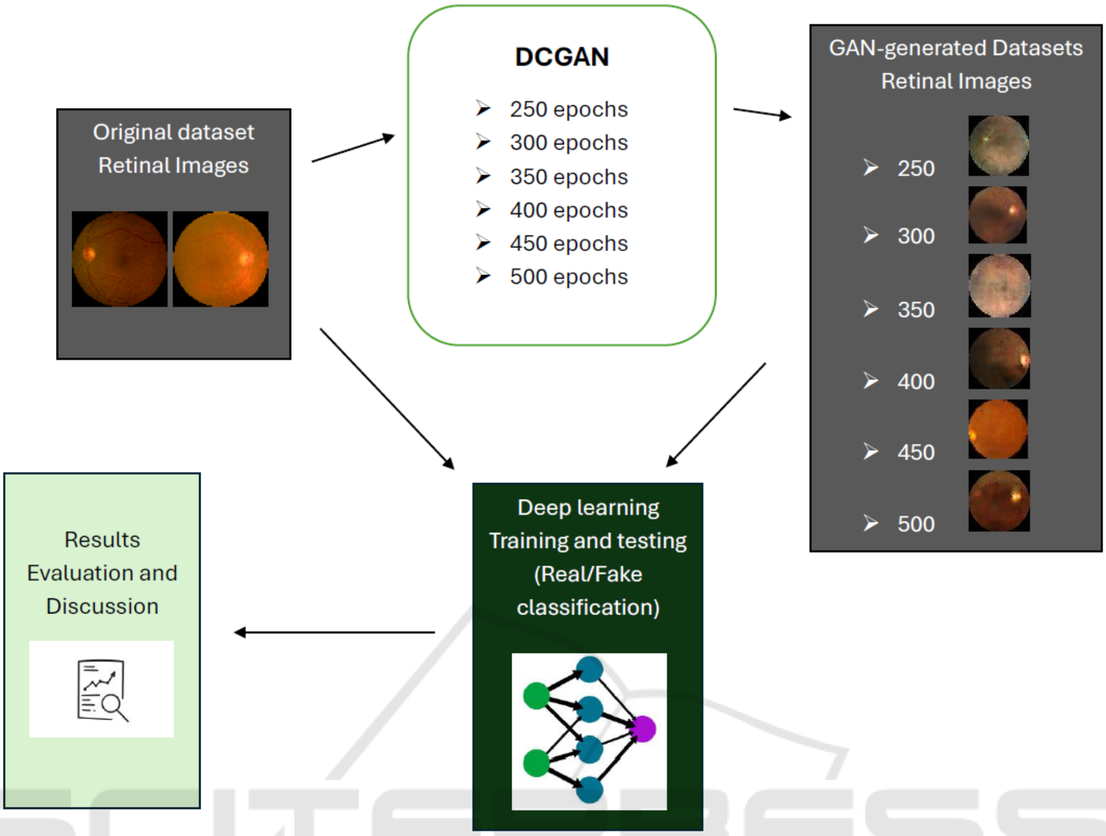

Figure 2: Main steps of the case study on RGB retinal images.

the main focus of our study is on the results of the DL network classification, specifically how well these networks discriminate between real and fake images. The selected epochs for the CNNs evaluation are: 250, 300, 350, 400, 450 and 500 Each of these sets of fake images and the original samples are the inputs of the CNNs, which aim to distinguish real and fake retinal images. The networks under analysis are the following: Standard_CNN (Di Giammarco et al., 2022a), (Di Giammarco et al., 2022b), InceptionV3 (Xia et al., 2017), Resnet50 (He et al., 2016), and the MobileNet (Howard et al., 2017). The networks are evaluated on the main metrics: accuracy, precision, recall, and loss. The main attention regards the accuracy trends over the GAN-epochs. In other words, for the previous networks, the CNNs evaluation concerns whether the accuracy was reduced or not after the classifications of the set of fake images at different GAN-epochs.

5 RESULTS AND DISCUSSION

In this section, the cited experiments on the dataset are conducted and reported.

The dataset under analysis is the RetinaMNIST reported at the following link: [1]. Going into detail, this dataset consists of 2.000 images of RGB fundus retina.

The mentioned networks are trained and evaluated in a binary classification throughout several DCGAN epochs generation in the experimental analysis section to differentiate real from fake retinal images.

For the training, validation, and testing sets, the two classes—the real and the fake—are split at an 80-10-10 rate following each GAN image generation. The following hyper-parameter combinations are used to train and test the datasets on the four networks listed above 50 epochs, 32 batch size, 0.0001 learning rate, and the input size as an image, or 64x64x3. This combination is the average combination for all networks, based on multiple testing and outcomes. Tables 1, 2, 3 and 4 report the evaluation

[1] https://medmnist.com/

of the metrics in terms of accuracy precision, recall, F-Measure and Area Under the Curve (AUC) for the four CNNs, respectevely.

A list of considerations through the evaluation of the results is reported below:

- Standard_CNN, network developed by the authors (Di Giammarco et al., 2022b) shows the best result. Metrics such as accuracy, precision and recall maintain the maximum values in all the DCGAN epochs, meaning that this network completely recognizes the real and the fake images. Also in the RGB case, the Standard_CNN obtains the most reliable output model, such as the previous work with the grayscale retinal images (Di Giammarco et al., 2024b).
- Different behavior is for the MobileNet. The network indeed distinguishes between the real and the GAN-generated images, but it is possible to observe that the loss gradually enhances according to the number of DCGAN epochs. This behavior will produce misclassification when the DCGAN epochs are enhanced. However, in our study (until 500 DCGAN epochs), the MobileNet model correctly recognized real and fake images.
- Finally, the ResNet50 and the InceptionV3 strongly decrease their performance as the DCGAN epochs increase. In particular, these decreasing trends are evident from 400 to 500 epochs, within this latter situation the networks fail the classification, and are not able to distinguish real and fake images.

A better view of these two trends for the ResNet50 and InceptionV3 is reported in Figures 3 and 4.

The two figures show the accuracy-epochs trend of the ResNet50 and the InceptionV3. The plot 3 reveals that the accuracy quickly drops from 450 to 500 DCGAN epochs, while the int the InceptionV3 plot, shows in Figure 4 the main decreasing is between 400 and 450, and report the 0,5 at 500 DCGAN epochs.

In the output, networks such as Standard_CNN and MobileNet correctly recognize the images, impossible to distinguish with a human eye analysis. In this way, the output model can be applied to medical imaging instruments with deep learning prediction implementation. These models represent, for any disease retinal image classification, the best solution to prevent possible attacks and to improve reliability and credibility. On the other hand, CNNs like ResNet50 and InceptionV3 with high values of DCGAN epochs, which means fake images are very similar to the original ones, not recognize the real and the fake images. This behavior does not guarantee reliability on possible predictions, with high risk on the diagnosis procedure and, consequentially, on the patient's health.

The optimal response of the Standard_CNN with the retinal images is confirmed also in the diabetic retinopathy classifications reported in the paper (Mercaldo et al., 2023).

6 CONCLUSIONS AND FUTURE WORKS

In conclusion, in this work, we present a general method to discriminate between real and fake images generated with the DCGAN to enhance data security and reliability. Being extendable on each biomedical imaging technique, the authors decide to analyze the RGB retinal images case. From the four CNNs in the experiment, two of these i.e. Standard_CNN and MobileNet, distinguish the real/fake task, and generate an output model optimal for automated AI-based imaging devices. The other two networks represent a bad choice for the same imaging devices, with the possibility of altering results including GAN-generated images in the testing folder, drastically reducing the reliability.

For future works, authors will exploit additional GANs for testing the CNNs. Another future work is related to the application of federated machine learning for data privacy reasons.

ACKNOWLEDGEMENTS

This work has been partially supported by EU DUCA, EU CyberSecPro, SYNAPSE, PTR 22-24 P2.01 (Cybersecurity) and SERICS (PE00000014) under the MUR National Recovery and Resilience Plan funded by the EU - NextGenerationEU projects, by MUR - REASONING: foRmal mEthods for computAtional analySis for diagnOsis and progNosis in imagING - PRIN, e-DAI (Digital ecosystem for integrated analysis of heterogeneous health data related to high-impact diseases: innovative model of care and research), Health Operational Plan, FSC 2014-2020, PRIN-MUR-Ministry of Health, the National Plan for NRRP Complementary Investments D^3 4 Health: Digital Driven Diagnostics, prognostics and therapeutics for sustainable Health care, Progetto MolisCTe, Ministero delle Imprese e del Made in Italy, Italy, CUP: D33B22000060001, FORESEEN: FORmal mEthodS for attack dEtEction in autonomous driviNg systems CUP N.P2022WYAEW and ALOHA: a framework for monitoring the physical and psychological health status of the Worker through Object de-

Table 1: Metrics evaluation for the Standard_CNN.

Metrics	250 Epochs	300 Epochs	350 Epochs	400 Epochs	450 Epochs	500 Epochs
Accuracy	1.0	1.0	1.0	1.0	1.0	1.0
Precision	1.0	1.0	1.0	1.0	1.0	1.0
Recall	1.0	1.0	1.0	1.0	1.0	1.0
F-Measure	1.0	1.0	1.0	1.0	1.0	1.0
AUC	1.0	1.0	1.0	1.0	1.0	1.0
Loss	0.0	0.0	0.0	0.0	0.0	0.0

Table 2: Metrics evaluation for the MobileNet.

Metrics	250 Epochs	300 Epochs	350 Epochs	400 Epochs	450 Epochs	500 Epochs
Accuracy	1.0	1.0	1.0	1.0	1.0	1.0
Precision	1.0	1.0	1.0	1.0	1.0	1.0
Recall	1.0	1.0	1.0	1.0	1.0	1.0
F-Measure	1.0	1.0	1.0	1.0	1.0	1.0
AUC	1.0	1.0	1.0	1.0	1.0	1.0
Loss	1.24×10^{-7}	4.91×10^{-7}	1.18×10^{-6}	5.43×10^{-6}	9.85×10^{-6}	2.23×10^{-5}

Table 3: Metrics evaluation for the ResNet50.

Metrics	250 Epochs	300 Epochs	350 Epochs	400 Epochs	450 Epochs	500 Epochs
Accuracy	1.0	1.0	1.0	0.970	0.997	0.5
Precision	1.0	1.0	1.0	0.970	0.997	0.5
Recall	1.0	1.0	1.0	0.970	0.997	0.5
F-Measure	1.0	1.0	1.0	0.970	0.997	0.5
AUC	1.0	1.0	1.0	0.982	0.999	0.497
Loss	1.24×10^{-7}	1.35×10^{-5}	6.0×10^{-4}	3.74×10^{-3}	2.358	4.389

Table 4: Metrics evaluation for the InceptionV3.

Metrics	250 Epochs	300 Epochs	350 Epochs	400 Epochs	450 Epochs	500 Epochs
Accuracy	1.0	1.0	0.995	0.950	0.600	0.5
Precision	1.0	1.0	0.995	0.950	0.600	0.5
Recall	1.0	1.0	0.995	0.950	0.600	0.5
F-Measure	1.0	1.0	0.995	0.950	0.600	0.5
AUC	1.0	1.0	0.999	0.965	0.624	0.5
Loss	7.28×10^{-3}	1.02×10^{-2}	4.22×10^{-2}	7.81×10^{-2}	0.673	0.693

tection and federated machine learning, Call for Collaborative Research BRiC -2024, INAIL.

REFERENCES

Akhil, M. S., Sharma, B. S., Kodipalli, A., and Rao, T. (2024). Medical image synthesis using dcgan for chest x-ray images. In *2024 International Conference on Knowledge Engineering and Communication Systems (ICKECS)*, volume 1, pages 1–8. IEEE.

Brunese, L., Brunese, M. C., Carbone, M., Ciccone, V., Mercaldo, F., and Santone, A. (2022a). Automatic pi-rads assignment by means of formal methods. *La radiologia medica*, pages 1–7.

Brunese, L., Mercaldo, F., Reginelli, A., and Santone, A. (2022b). A neural network-based method for respiratory sound analysis and lung disease detection. *Applied Sciences*, 12(8):3877.

Di Giammarco, M., Dukic, B., Martinelli, F., Cesarelli, M., Ravelli, F., Santone, A., and Mercaldo, F. (2024a). Reliable leukemia diagnosis and localization through explainable deep learning. In *2024 Fifth International Conference on Intelligent Data Science Technologies and Applications (IDSTA)*, pages 68–75. IEEE.

Di Giammarco, M., Iadarola, G., Martinelli, F., Mercaldo,

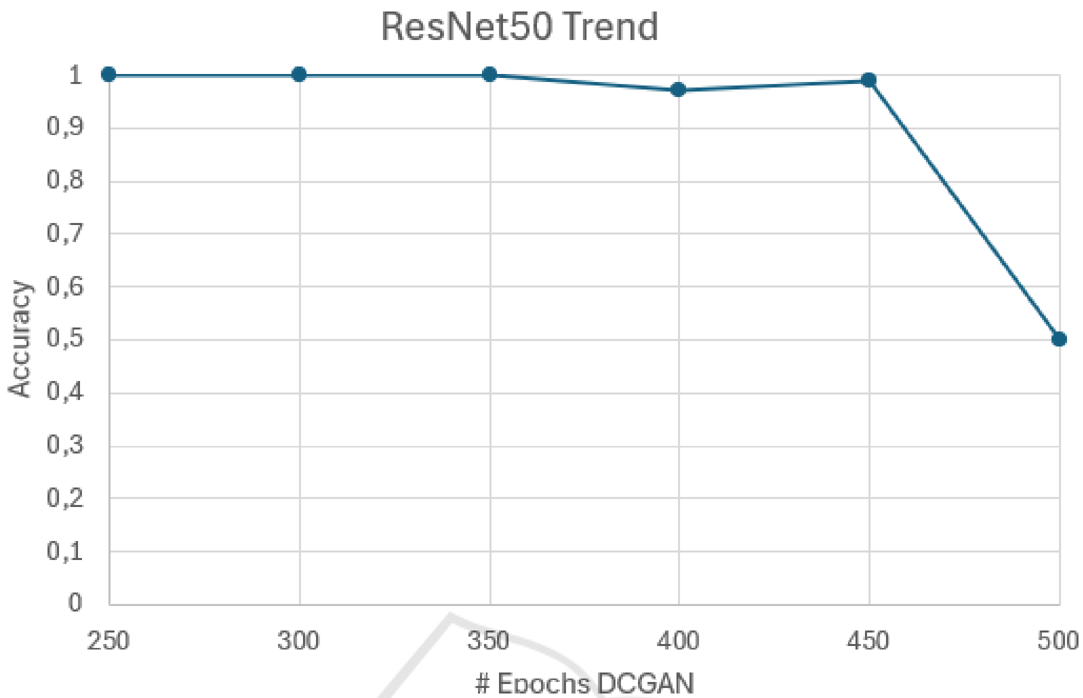

Figure 3: Accuracy-epochs trends of the ResNet50.

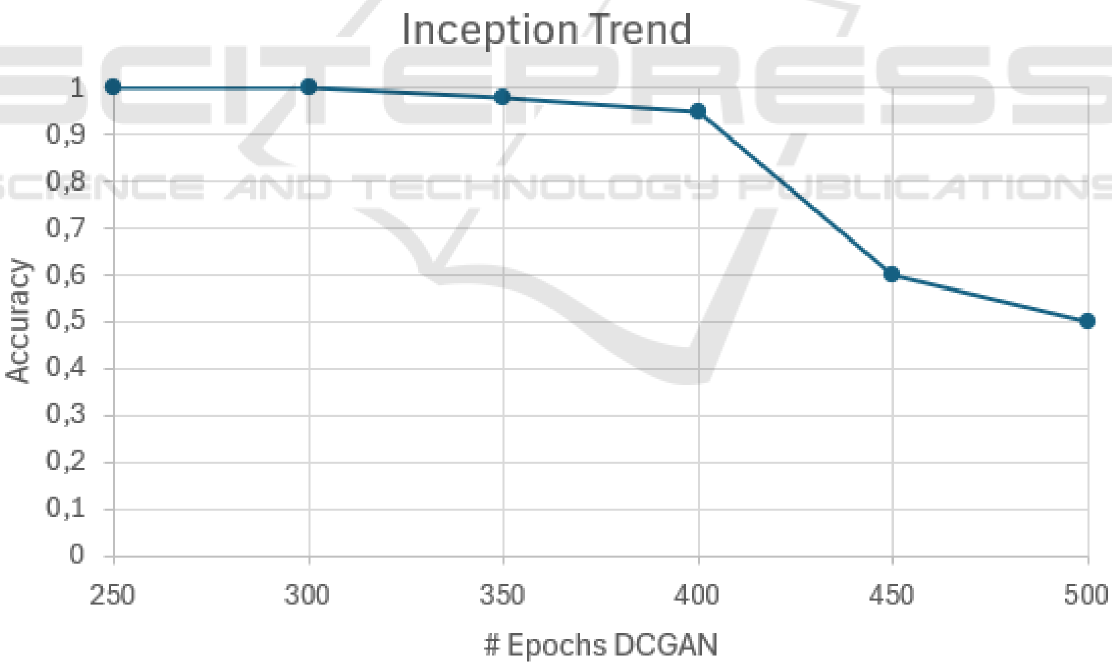

Figure 4: Accuracy-epochs trends of the InceptionV3.

F., Ravelli, F., and Santone, A. (2022a). Explainable deep learning for alzheimer disease classification and localisation. In *International Conference on Applied Intelligence and Informatics, 1-3 September, Reggio Calabria, Italy*. in press.

Di Giammarco, M., Iadarola, G., Martinelli, F., Mercaldo, F., and Santone, A. (2022b). Explainable retinopathy diagnosis and localisation by means of class activation mapping. In *2022 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8. IEEE.

Di Giammarco, M., Santone, A., Cesarelli, M., Martinelli, F., and Mercaldo, F. (2024b). Evaluating deep learning resilience in retinal fundus classification with generative adversarial networks generated images. *Electronics*, 13(13):2631.

Diaz-Pinto, A., Colomer, A., Naranjo, V., Morales, S., Xu, Y., and Frangi, A. F. (2019). Retinal image synthesis and semi-supervised learning for glaucoma assessment. *IEEE transactions on medical imaging*, 38(9):2211–2218.

Fang, W., Zhang, F., Sheng, V. S., and Ding, Y. (2018). A method for improving cnn-based image recognition using dcgan. *Computers, Materials & Continua*, 57(1).

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y. (2020). Generative adversarial networks. *Communications of the ACM*, 63(11):139–144.

He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778.

Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., and Adam, H. (2017). Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv:1704.04861*.

Huang, P., Xiao, H., He, P., Li, C., Guo, X., Tian, S., Feng, P., Chen, H., Sun, Y., Mercaldo, F., et al. (2024). La-vit: A network with transformers constrained by learned-parameter-free attention for interpretable grading in a new laryngeal histopathology image dataset. *IEEE Journal of Biomedical and Health Informatics*.

Liu, B., Lv, J., Fan, X., Luo, J., and Zou, T. (2022). Application of an improved dcgan for image generation. *Mobile information systems*, 2022(1):9005552.

Martinelli, F., Mercaldo, F., and Santone, A. (2022). Water meter reading for smart grid monitoring. *Sensors*, 23(1):75.

Mercaldo, F., Ciaramella, G., Iadarola, G., Storto, M., Martinelli, F., and Santone, A. (2022). Towards explainable quantum machine learning for mobile malware detection and classification. *Applied Sciences*, 12(23):12025.

Mercaldo, F., Di Giammarco, M., Apicella, A., Di Iadarola, G., Cesarelli, M., Martinelli, F., and Santone, A. (2023). Diabetic retinopathy detection and diagnosis by means of robust and explainable convolutional neural networks. *Neural Computing and Applications*, 35(23):17429–17441.

Mercaldo, F., Di Giammarco, M., Ravelli, F., Martinelli, F., Santone, A., and Cesarelli, M. (2024a). Alzheimer's disease evaluation through visual explainability by means of convolutional neural networks. *International Journal of Neural Systems*, 34(2):2450007–2450007.

Mercaldo, F., Martinelli, F., and Santone, A. (2024b). Deep convolutional generative adversarial networks in image-based android malware detection. *Computers*, 13(6).

Mirsky, Y., Mahler, T., Shelef, I., and Elovici, Y. (2019). {CT-GAN}: Malicious tampering of 3d medical imagery using deep learning. In *28th USENIX Security Symposium (USENIX Security 19)*, pages 461–478.

Nagaraju, R. and Stamp, M. (2022). Auxiliary-classifier gan for malware analysis. In *Artificial Intelligence for Cybersecurity*, pages 27–68. Springer.

Pan, H. and Xin, L. (2024). Fdts: A feature disentangled transformer for interpretable squamous cell carcinoma grading. *IEEE/CAA Journal of Automatica Sinica*, 12(JAS-2024-1027).

Singh, N. K. and Raza, K. (2020). Medical image generation using generative adversarial networks. *arXiv preprint arXiv:2005.10687*.

Singh, N. K. and Raza, K. (2021). Medical image generation using generative adversarial networks: A review. *Health informatics: A computational perspective in healthcare*, pages 77–96.

Wang, X., Guo, H., Hu, S., Chang, M.-C., and Lyu, S. (2023). Gan-generated faces detection: A survey and new perspectives. *ECAI 2023*, pages 2533–2542.

Xia, X., Xu, C., and Nan, B. (2017). Inception-v3 for flower classification. In *2017 2nd international conference on image, vision and computing (ICIVC)*, pages 783–787. IEEE.

You, A., Kim, J. K., Ryu, I. H., and Yoo, T. K. (2022). Application of generative adversarial networks (gan) for ophthalmology image domains: a survey. *Eye and Vision*, 9(1):6.

Zhou, X., Tang, C., Huang, P., Tian, S., Mercaldo, F., and Santone, A. (2023). Asi-dbnet: an adaptive sparse interactive resnet-vision transformer dual-branch network for the grading of brain cancer histopathological images. *Interdisciplinary Sciences: Computational Life Sciences*, 15(1):15–31.