Generative Adversarial Networks (GANs) in Medical Imagining

Alexandra-Georgiana Andrei
AI Multimedia Lab
National University of Science and Technology
POLITEHNICA Bucharest
Bucharest, Romania
alexandra.andrei@upb.ro

Bogdan Ionescu
AI Multimedia Lab
National University of Science and Technology
POLITEHNICA Bucharest
Bucharest, Romania
bogdan.ionescu@upb.ro

***Abstract*—Generative Adversarial Networks (GANs) have shown immense potential in medical imaging, particularly in generating synthetic data to address the challenges of limited annotated datasets. This paper provides a concise overview of key applications of GANs in medical image augmentation, focusing on their role in enhancing training data for improved diagnostic accuracy. We discuss the technical challenges and limitations associated with GANs, including training instability, mode collapse, and privacy concerns in sensitive medical data. Our work contributes specifically to the augmentation of histopathology images and CT slices using various GAN architectures, such as DCGAN, cGAN, and StyleGAN. Through extensive experimentation on hyperparameters and model architectures, we demonstrate that GAN–augmented datasets can significantly improve classifier performance, while carefully addressing privacy risks by evaluating potential "fingerprints" in synthetic images. These findings support the clinical utility of GAN–based augmentation in medical imaging and establish a foundation for future research on privacy-preserving synthetic data generation.**

***Index Terms*—Generative Adversarial Networks (GANs), data augmentation, medical imaging, histopathology image synthesis, CT image generation, privacy in synthetic data.**

I. Introduction

Biomedical imaging has andvanced significantly in recent years, largely due to the integration of machine learning (ML) and artificial intelligence (AI) technologies. Among these advancements, generative models — especially Generative Adversarial Networks (GANs) — have proven to be powerful tools for generating synthetic images that resemble real biomedical images. The progress in GAN development has unlocked new opportunities for both research and practical applications, as high-quality synthetic images are now being produced across various fields using these models.

Large annotated datasets are often essential for achieving strong performance with deep learning models, yet these datasets are not always readily available in healthcare. Barriers include medical data confidentiality, legal restrictions, data complexity, and the high cost of data labeling. GAN-based image synthesis methods have been explored as a way to address this challenge across different fields, including medicine. Based on the type of translation between the input and output domains, GAN-based image synthesis can follow two main approaches: Latent-to-Image (Z2I), which generates images from a latent vector or random noise, and Image-to-Image (I2I), which translates images from one domain to another.

In recent years, GANs have been widely applied in healthcare for data augmentation, especially in medical imaging, targeting various data types. Nie et al. [1] proposed context-aware GANs to generate computed tomography (CT) images from magnetic resonance images (MRIs). Similarly, Yang et al. [2] developed a method using cGANs for MRI image generation, while Salehinejad et al. [3] applied DCGANs to create synthetic chest X-ray images. Frid-Adar et al. [4] focused on generating synthetic CT images to support liver lesion classification, and Madani et al. [5] showed that GAN-based data augmentation improved accuracy over traditional methods for chest X-ray analysis. Additionally, Ho et al. [6] demonstrated that diffusion probabilistic models could produce high-fidelity images on par with GANs, as illustrated by Puria Azadi et al. [7], who used these models to synthesize high-quality histopathology images of brain cancer.

Various GAN architectures have been adapted for specialized tasks in histopathology: InfoGAN for stain normalization [8], CycleGAN for ink mark removal [9], the Pix2Pix model for segmentation [10], and GAN models for generating label masks to create histological samples [11]. Super-resolution GANs have also been applied to enhance image resolution.

For data augmentation, Quiros et al. [12] introduced PathologyGAN, a framework that synthesizes H&E-stained colorectal and breast cancer tissues through a structured latent space, combining BigGAN, StyleGAN, and the Relativistic Average Discriminator, with performance assessed using FID scores for varying dataset sizes. Jerry Wei et al. [13] employed CycleGANs to generate synthetic colorectal polyp images from normal mucosa, with Turing tests by four gastrointestinal pathologists revealing that at least two could not statistically distinguish synthetic images from real ones. Liu et al. [14] used GANs to augment histological glioma specimens to improve IDH mutation status prediction from H&E slides. Boyd et al. [15] extended the visual field of histopathology tiles from lymph node whole-slide images, achieving FID scores around 21 for the CAMELYON17 (breast cancer) and 37 for the CRC (colorectal cancer) datasets.

This paper presents a short overview of the main applica-

1st Workshop on Artificial Intelligence for Multimedia,
AI Multimedia Lab, CAMPUS Research Institute, POLITEHNICA Bucharest,
Bucharest, Romania, 8 November, 2024

tions of GANs in medical imaging, examines key challenges and limitations, and highlights our contributions to the field, specifically in the synthesis of histopathology images and CT slices.

II. Variants of GANs

Generative Adversarial Networks [16] are used to generate new data by learning an implicit representation of the dataset distribution. GANs consist in a pair of networks: generator (*G*) and discriminator (*D*). The *G* is responsible for crating new data samples by learning to mimic the underlying distribution of the original dataset. The *D* is a classifier that evaluates these generated images alogside the real images, determining whether each sample is "real" or "fake". During training, the *G* continuously tries to improve its outputs to "fool" the discriminator, while the discriminator refines its ability to detect fake images. This adversarial process drives both networks to improve iteratively, resulting in the generator producing increasingly realistic images as training progresses.

In this section, we focus on the specific types of GANs that we've employed in our work so far: Deep Convolutional GAN (DCGAN) [17], Conditional GAN (cGAN) [18], and StyleGAN [19]. These variants were selected for their unique capabilities in generating high-quality synthetic images tailored to medical imaging applications. DCGAN leverage deep convolutional layers to enhance image quality and stability, which is valuable in generating realistic medical images with preserved structural features. cGANs allow for controlled image generation by conditioning on specific class labels, making them particularly useful for producing distinct tissue types in histopathology. Figure 1 presents the comparison of the original GAN method and the cGAN method. StyleGAN, known for its innovative style-based architecture, provides fine-grained control over generated image features, facilitating the synthesis of highly detailed images that maintain the complexity of medical images. Each of these GAN types contributes specific strengths to our work, enabling us to create diverse, clinically relevant synthetic images that enhance dataset quality and support reliable model training.

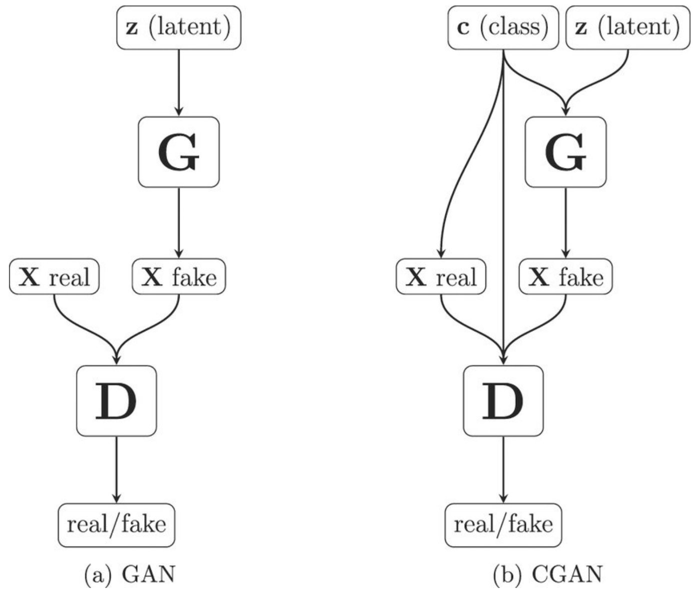


Fig. 1: Comparison of original GAN and cGAN methods. *G* and *D* denote the generator and discriminator networks, with *X real* and *X fake* as real and generated samples, respectively. In cGANs, class labels are also provided to the generator, enabling it to generate samples conditioned on these labels.

III. Challenges and Limitations of GANs in Medical Imaging

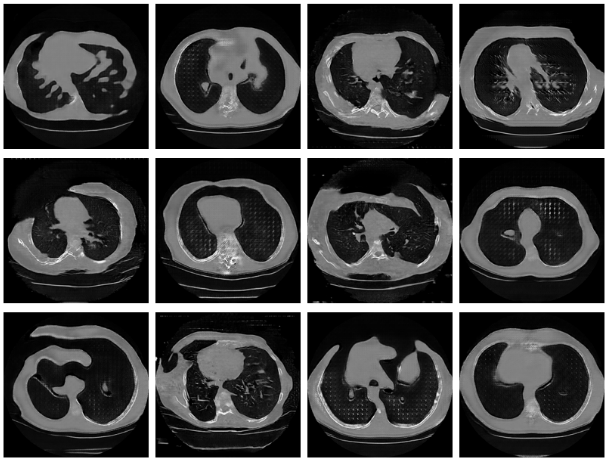
While GANs offer significant promise in medical imaging, their practical application is not without challenges. The unique complexities of medical data, combined with the technical limitations of GANs, present several obstacles that can impact the accuracy, reliability, and clinical applicability of these models.

Training Instability: The training process needs careful balancing between the *G* and the *D*. Without this syncronization, it is difficult to obtain stable training results, leading to poor convergence and difficulty in achieving the desired image quality and realism. This instability is particularly problematic in medical imaging, where accurate and reliable image synthesis is crucial. Additionally, as GANs are based on deep neural networks, they inherit the interpretability challenges commonly associated with neural networks, which can make it difficult for researchers and clinicians to understand the underlying processes that drive the generation of synthetic images [20].

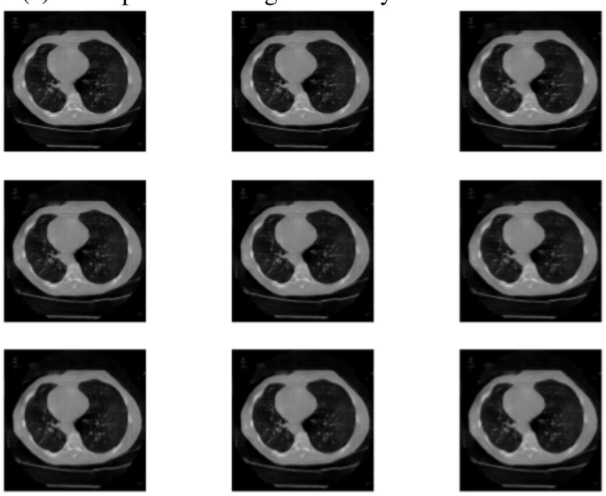
Mode Collapse: is a common issue in GANs, where the *G* produces a limited variety of images that lack diversity, often leading to images with similar features. This issue can be particularly problematic in medical imaging applications where capturing the full diversity of clinical presentations is essential. For instance, generating synthetic images that fail to represent rare cases or atypical manifestations can lead to biased models that overlook critical variations in patient data. This lack of diversity in synthetic images can skew the training process, resulting in models that may perform well on common cases but struggle with uncommon or complex cases [20].

Evaluation Metrics: One of the biggest challenges in GAN-based image synthesis for medical imaging is the lack of standardized, clinically relevant evaluation metrics. Traditional metrics, such as Fréchet Inception Distance (FID) and Inception Score (IS), measure image quality and diversity but may not reflect clinical utility or relevance. In the medical field, the synthesized images need to be not only visually realistic but also clinically accurate. This discrepancy complicates the assessment of GANs in medical applications, where human evaluation by clinicians is often required to confirm the quality and usefulness of generated images, adding time and cost to the validation process [21].

Privacy and Confidentiality Ethical Concerns: The use of GANs in medical imaging raises significant concerns related to data privacy, confidentiality, and regulatory compliance.



(a) Example of training instability in GAN models.



(b) Example of mode collapse in GAN models.

Fig. 2: Examples of GAN model failures in data generation.

Regulations such as the General Data Protection Regulation (GDPR) [22] in Europe and the Health Insurance Portability and Accountability Act (HIPAA) [23] in the United States enforce strict guidelines on the handling of personal and medical data. While synthetic data generated by GANs is often seen as a solution to the scarcity of annotated datasets in healthcare, questions remain about whether this data can be truly “anonymized”. A primary concern is the potential for synthetic images to contain identifiable features, or “fingerprints” from the original training data [24], [25]. This poses a risk that generated images might retain subtle yet identifiable features of the source data, which could compromise patient privacy if the images are traced back to specific individuals.

Figure 2 illustrates the two presented challenges: training instability, which results in distorted CT slices that do not adhere to anatomical structures, and mode collapse, where the generator repeatedly produces the same high-quality sample capable of fooling the discriminator.

IV. Our Contributions

In this section, we outline our contributions to advancing GAN-based image augmentaion in medical images, with focus on histopathology and CT images. By leveraging different GAN architectures, we address challenges in generating synthetic images that are diverse, realistic and privacy-preserving.

A. Synthetic Colorectal Cancer Histopathology Image Augmentation Using Conditional Generative Adversarial Networks

We propose a latent-to-image approach for generating synthetic images by applying a Conditional Deep Convolutional Generative Adversarial Network for generating images of human colorectal cancer and healthy tissue. We generate high-quality images of various tissue types that preserve the general structure and features of the source classes. The quality of these images is evaluated through perceptual experiments with pathologists and the Fréchet Inception Distance (FID) metric. Using the generated data to train classifiers significantly improved MobileNet’s [26] accuracy by 22.1%, and also enhanced the accuracies of DenseNet [27], ResNet [28], and EfficientNet [29]. We further validated the robustness and versatility of our model on a different dataset, yielding promising results. Additionally, we make a novel contribution by addressing security and privacy concerns in personal medical image data, ensuring that training medical images “fingerprints” are not contained in the synthetic images generated with the model

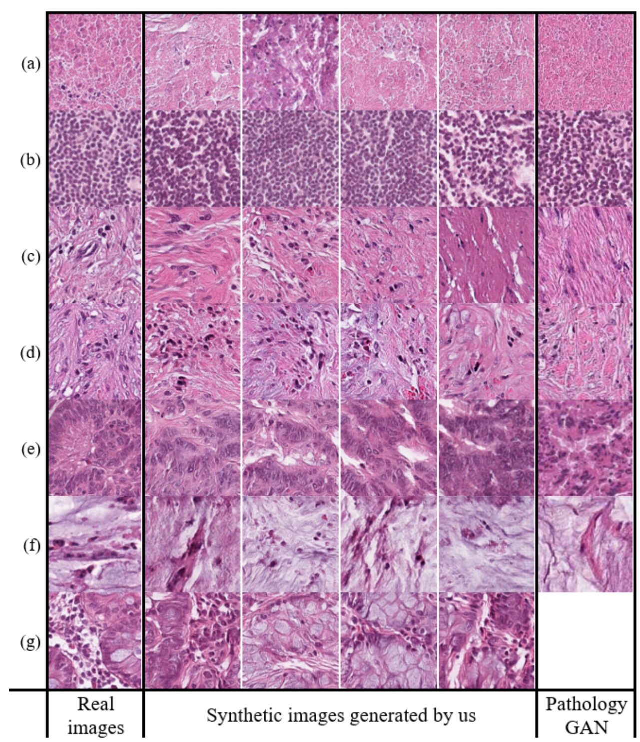


Fig. 3: Examples of real and synthetic images for NCT-CRC-HE-100K colorectal cancer dataset: (a) DEB, (b) LYM, (c) MUS, (d) STR, (e) TUM, (f) MUC, and (g) NORM.

we propose. Figure 3 allows a visual comparison by showing tissue samples generated by our method alongside a sample produces by PathologyGAN [12], except for the NORM tissue type, which they did not generate in their work. The generated images preserve the general structure and features of the original tissue types, illustrating the effectiveness of our synthetic data generation process.

Further details regarding our methodologies, experimental setup, and comprehensive results will be presented in an upcoming publication. In the interest of this workshop, we have included only a summary of key findings.

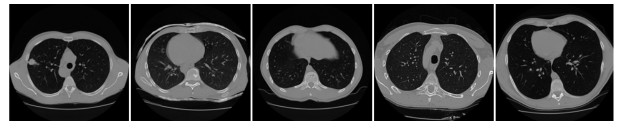
B. CTs Slices Augmentation using different GANs

For CT images, we experimented with various GAN architectures, including DCGAN, cGAN, and StyleGAN, to identify models that effectively address training instability and mode collapse. By optimizing model architectures and training protocols, we aimed to improve the reliability and quality of CT image synthesis. In our work on augmenting CT slices using various GAN models, we experimented with a range of parameters for both G and the D networks to optimize performance and avoid the challenges presented in Section III. Specifically, we tested different values for the learning rate (lr) to find the optimal balance between stability and convergence speed. Additionally, we explored variations in the $\beta 1$ and $\beta 2$ parameters of the Adam optimizer, which control the momentum and adaptivity of the learning process, affecting how quickly and effectively the networks learn. Adjustments to the latent dimension were also investigated to determine how the complexity of the generated samples could be enhanced while maintaining training efficiency. Table I present the hyperparameters tested in our studies so far. Through these parameter explorations, we aimed to improve the quality and diversity of generated CT slices, ensuring realistic augmentations suitable for medical imaging applications. Further details, including an in-depth analysis of the generative capabilities of different GAN models, experimental findings, and evaluations of the generated data, will be presented in an upcoming publication. Figure 4 presents some of the results obtained using different variations of the architectures presented in Section II.

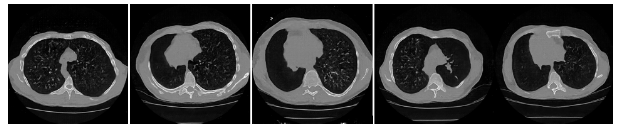
One step further, we examine privacy risks in synthetic CT images generated by each GAN model through their use in the ImageCLEF benchmarking campaign for our GANs task [24], [25]. In this evaluation, we assess the potential presence of "fingerprints" or residual identifiable patterns in the synthetic images that might inadvertently disclose details of the original data, thereby addressing important privacy concerns.

TABLE I: Hypermeter values and latent dimensions used in our studies.

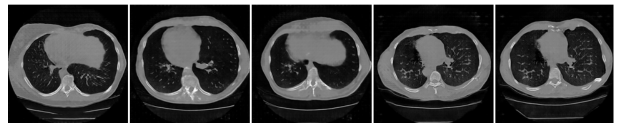
Parameter	Value
Learning rate	{0.0001, 0.0001, 0.0002, 0.001}
$\beta 1$	{0.5, 0.9}
$\beta 2$	{0.9, 0.99, 0.999}
Latent dimension	{50, 100, 500, 1000}



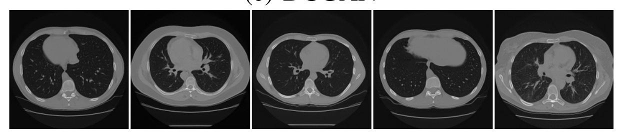
(a) Real images



(b) cGAN



(c) DCGAN



(d) StyleGAN

Fig. 4: Examples of images generated with different GANs architectures.

V. Discussion and Conclusions

This paper highlights the application of Generative Adversarial Networks (GANs) in medical image augmentation, with a focus on histopathology images and CT slices. By employing different GAN architectures, such as DCGAN, cGAN, and StyleGAN, we demonstrated the potential of synthetic data generation to enhance training datasets.

Future work should focus on refining GAN architectures to further enhance image quality, stability, and interpretability, as well as developing robust privacy-preserving techniques to ensure synthetic images remain fully anonymized. Additionally, improving evaluation metrics tailored specifically to clinical relevance will be essential for validating GAN applications in healthcare.

References

[1] C. P. S. R. D. S. D. Nie, R.Trullo, "Medical image synthesis with context-aware generative adversarial networks," in *MICCAI*. Springer, 2017, pp. 417–425.

[2] Q. Y. et al., "Mri cross-modality image-to-image translation," *Scientific reports*, vol. 10, no. 1, pp. 1–18, 2020.

[3] T. D. E. C. J. B. H. Salehinejad, S. Valaee, "Generalization of deep neural networks for chest pathology classification in x-rays using generative adversarial networks," in *2018 IEEE international conference on acoustics, speech and signal processing (ICASSP)*. IEEE, 2018, pp. 990–994.

[4] M. F.-A. et al., "Gan-based synthetic medical image augmentation for increased cnn performance in liver lesion classification," *Neurocomputing*, vol. 321, pp. 321–331, 2018.

[5] A. K. T. S.-M. M. Moradi, A. Mandani, "Chest x-ray generation and data augmentation for cardiovascular abnormality classification," in *Medical imaging 2018: Image processing*, vol. 10574. SPIE, 2018, pp. 415–420.

[6] P. A. J. Ho, A. Jain, "Denoising diffusion probabilistic models," in *Advances in Neural Information Processing Systems*, vol. 33. Curran Associates, Inc., 2020, pp. 6840–6851.

[7] P. A. M. et al., “A morphology focused diffusion probabilistic model for synthesis of histopathology images,” *arXiv e-prints*, pp. arXiv–2209, 2022.

[8] B. B. J. A. d. L. P. H. d. W. F. Zanjani, S. Zinger, “Stain normalization of histopathology images using generative adversarial networks,” in *2018 IEEE 15th International symposium on biomedical imaging (ISBI 2018)*. IEEE, 2018, pp. 573–577.

[9] C. V. J. R. S. Ali, N. K. Alham, “Ink removal from histopathology whole slide images by combining classification, detection and image generation models,” in *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*. IEEE, 2019, pp. 928–932.

[10] K. W. X. G. D. Wang, C. Gu, “Adversarial neural networks for basal membrane segmentation of microinvasive cervix carcinoma in histopathology images,” in *2017 International conference on machine learning and cybernetics (ICMLC)*, vol. 2. IEEE, 2017, pp. 385–389.

[11] A. C. S. G. B. Venkatesh, T. Shah, “Restoration of marker occluded hematoxylin and eosin stained whole slide histology images using generative adversarial networks,” in *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*, 2020, pp. 591–595.

[12] A. C. Quiros, R. Murray-Smith, and K. Yuan, “Pathologygan: Learning deep representations of cancer tissue,” *arXiv preprint arXiv:1907.02644*, 2019.

[13] J. W. et al., “Generative image translation for data augmentation in colorectal histopathology images,” *Proceedings of machine learning research*, vol. 116, p. 10, 2019.

[14] S. L. et al., “Isocitrate dehydrogenase (idh) status prediction in histopathology images of gliomas using deep learning,” *Scientific reports*, vol. 10, no. 1, pp. 1–11, 2020.

[15] J. B. et al., “Self-supervised representation learning using visual field expansion on digital pathology,” in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021, pp. 639–647.

[16] I. J. G. et al., “Generative adversarial networks,” *Communications of the ACM*, vol. 63, no. 11, pp. 139–144, 2020.

[17] M. S. C. A. Radford, L, “Unsupervised representation learning with deep convolutional generative adversarial networks,” *arXiv preprint arXiv:1511.06434*, 2015.

[18] M. Mirza and S. Osindero, “Conditional generative adversarial nets,” *arXiv preprint arXiv:1411.1784*, 2014.

[19] T. Karras, S. Laine, and T. Aila, “A style-based generator architecture for generative adversarial networks,” *CoRR*, vol. abs/1812.04948, 2018. [Online]. Available: http://arxiv.org/abs/1812.04948

[20] Z. Ahmad, Z. u. A. Jaffri, M. Chen, and S. Bao, “Understanding gans: fundamentals, variants, training challenges, applications, and open problems,” *Multimedia Tools and Applications*, pp. 1–77, 2024.

[21] P. Salehi, A. Chalechale, and M. Taghizadeh, “Generative adversarial networks (gans): An overview of theoretical model, evaluation metrics, and recent developments,” *arXiv preprint arXiv:2005.13178*, 2020.

[22] European Parliament and Council of the European Union. Regulation (EU) 2016/679 of the European Parliament and of the Council. [Online]. Available: https://data.europa.eu/eli/reg/2016/679/oj

[23] Centers for Medicare & Medicaid Services, “The Health Insurance Portability and Accountability Act of 1996 (HIPAA),” Online at http://www.cms.hhs.gov/hipaa/, 1996.

[24] A. Andrei, A. Radzhabov, D. Karpenka, Y. Prokopchuk, V. Kovalev, B. Ionescu, and H. Müller, “Overview of 2024 imageclefmedical gans task–investigating generative models’ impact on biomedical synthetic images,” in *CLEF2024 Working Notes, CEUR Workshop Proceedings, CEUR-WS. org, Grenoble, France*, 2024.

[25] ——, “Overview of 2024 imageclefmedical gans task–investigating generative models’ impact on biomedical synthetic images,” in *CLEF2024 Working Notes, CEUR Workshop Proceedings, CEUR-WS. org, Grenoble, France*, 2024.

[26] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, “Mobilenetv2: Inverted residuals and linear bottlenecks,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 4510–4520.

[27] F. Iandola, M. Moskewicz, S. Karayev, R. Girshick, T. Darrell, and K. Keutzer, “Densenet: Implementing efficient convnet descriptor pyramids,” *arXiv preprint arXiv:1404.1869*, 2014.

[28] B. Koonce and B. Koonce, “Resnet 50,” *Convolutional neural networks with swift for tensorflow: image recognition and dataset categorization*, pp. 63–72, 2021.

[29] ——, “Efficientnet,” *Convolutional neural networks with swift for Tensorflow: image recognition and dataset categorization*, pp. 109–123, 2021.