Technical Journal, University of Engineering and Technology (UET) Taxila, Pakistan Vol. 29 No. 1-2024
ISSN:1813-1786 (Print) 2313-7770 (Online)

Deep Learning Based Multi-Class Eye Disease Classification: Enhancing Vision Health Diagnosis

J. Aslam[1], M. A. Arshed[2], S. Iqbal[3], H. M. Hasnain[4]

[1,2,3]*School of Systems and Technology, University of Management and Technology, Lahore, Pakistan*
[4]*Department of Computer Science, The Islamia University of Bahawalpur, Bahawalpur, Pakistan*

[2]muhammadasadarshed@gmail.com

Abstract- Retinal abnormalities impact millions of people globally. Timely detection and treatment of these abnormalities could prevent further progression, potentially saving countless individuals from preventable blindness. However, manual disease detection is a slow, laborious process and lacks consistency in results. This study uses convolutional neural networks to categorize eye disease using a publicly available dataset. Five different pre-trained models based on convolutional neural networks (CNNs), including VGG-16, VGG-19, ResNet-50, ResNet-152, and DenseNet-121, were used in this study. We were able to detect eye diseases at the cutting edge using the refined VGG-19. With testing accuracy of 95% on the dataset, this model accurately predicted eye diseases due to the effective and same weighted precision, recall, and F1 score of 95%. The model also significantly reduces training loss while improving accuracy.

Keywords- Eye Disease, Pre-Trained, Fine-Tuning, Convolutional Neural Networks

I. INTRODUCTION

The eye is a remarkable organ that grants us the gift of vision, playing a vital role in our daily lives. However, ocular diseases and conditions can significantly impact a person's vision and overall quality of life. If these eye conditions are not addressed promptly, they can lead to partial or complete blindness. Visual issues are common and affect almost everyone at some point in their lives. While some problems can be treated at home, others require the expertise of a specialist.

Among the various retinal diseases are Choroidal Neovascularization (CNV), Age-Related Macular Degeneration (AMD), Diabetic Macular Edema (DME), glaucoma, cataract, Drusen, and diabetic retinopathy (DR). Identifying these eye conditions early on is crucial in preventing vision loss. However, early detection can be challenging due to the subtlety and diverse characteristics of initial symptoms. Ophthalmologists often use retinal images obtained through fundus lenses or Optical Coherence Tomography (OCT) to diagnose retinal diseases. Nevertheless, this manual process is time-consuming and can be difficult to achieve accurate diagnoses.

Diagnosing ocular diseases without specialized knowledge is challenging because these diseases may manifest only minor changes in the appearance of the eye or its surroundings. As a result, symptoms may be easily overlooked, misdiagnosed, or mistaken for unrelated or typical variations. Overcoming this challenge requires advanced detection techniques capable of recognizing even the most subtle visual cues associated with various eye disorders.

The incorporation of artificial intelligence (AI) and deep learning technologies, particularly Convolutional Neural Networks (CNNs), has significantly advanced medical image processing in recent years. The detection of eye diseases has gained significantly from these technologies. Medical experts, especially ophthalmologists, can use these algorithms to diagnose and classify eye diseases at an early stage, resulting in more effective and specific treatment strategies.

There are several crucial phases involved in using retinal images to diagnose eye diseases, including feature extraction, categorization, and image pre-processing. Accurate diagnoses can be made with the use of deep learning algorithms used in conjunction with machine learning and image processing techniques.

Globally, there are regional variations in the occurrence of eye issues, which are determined by factors like age, gender, occupation, lifestyle, economic status, habits, and cultural standards. For instance, tropical populations may experience higher rates of eye infections due to environmental factors like dust, humidity, and sunshine. Both industrialized and developing countries experience a high prevalence of eye disorders, with large optical morbidity rates found in several Asian countries. However, these diseases are still underdiagnosed in some areas and are not adequately treated.

According to the World Health Organization (WHO), approximately 285 million people worldwide experience visual problems, with 246 million having poor eyesight and 39 million being blind[1-2]. It is crucial to offer affordable or free comprehensive eye care services to residents of underprivileged neighbourhoods and slums. Moderate-to-severe distance nearsightedness or blindness can result from a variety of eye conditions, including untreated presbyopia, unresolved refractive errors, cataracts, glaucoma, corneal opacities, diabetic retinopathy, and trachoma. The facts indicate that these diseases impact roughly 1 billion people worldwide. It is believed that around half of these cases could have been avoided or properly treated if people had better access to eye care and treatment.

II. Literature Review

Deep learning has found application in numerous aspects of cataract management, encompassing both clinical and surgical domains. Its applications range from the diagnosis of cataracts to enhancing biometry for precise calculation of intraocular lens (IOL) power.

In order to identify and categorize eye diseases, Xu [3] used the AlexNet and Visual-DN CNN-based algorithms. They achieved an accuracy of 86.2% using 8,030 fundus images for eye cataracts identification.

Another study conducted by Zhang [4] also focused on cataract detection and grading. They reported a significant improvement in accuracy, achieving an impressive 93%. Their analysis was based on 1,352 fundus images.

In the study proposed by Gosh [5], they utilized a CNN model and a diverse dataset that included glaucoma, retinal disease, and normal eye cataracts. Unfortunately, further details about their findings were not provided in the current text. Their CNN achieved an accuracy of 82%, which is considered acceptable based on CNN standards.

In a recent study conducted by Ahmed they focused on cataract analysis using a CNN model with VGG-19 architecture. Impressively, their research achieved an outstanding overall accuracy of 97.47%, with a precision rate of 97.47% and a relatively low loss of 5.27% [6].

The performance of the pretrained models cannot be neglected due to the effective performance in vision problems[7-9], and [10]. This paper's primary focus lies in detecting cataract disease through transfer learning. It also presents a comparative analysis of five different deep learning models, namely DenseNet121, ResNet50, ResNet152, VGG-16 and VGG-19 InceptionV3, and InceptionResNetV2 for cataract disease detection. The research explores various approaches and hyperparameters used to implement these models to achieve accurate identification of eight different types of eye diseases.

An advanced and automated image processing algorithm was suggested for diagnosing glaucoma from fundus images in the study [11]. This method employs bend point detection and tracks blood vessels to enhance accuracy and reliability in the diagnosis process. In the study [12], a tailored threshold-based algorithm was crafted to identify red lesions associated with diabetic retinopathy (DR) in fundus images. Each image was processed independently in this approach.

Diabetic retinopathy (DR) leads to diverse retinal abnormalities, such as hard exudates, hemorrhages, microaneurysms, and various symptoms. Various machine-learning methods have been created to detect DR and other diseases. These methods employ different image processing and computer vision approaches for analysis and feature extraction, often utilizing limited and small-scale datasets [13].

A network named ReLayNet [14], structured as an encoder-decoder, has been developed to segment various layers of the retina, including the identification of accumulated fluid in images.

The efficient utilization of deep neural networks for screening age-related macular degeneration (ARMD) has been achieved through the analysis of color fundus images. Experiments were conducted in the study [15] using a customized VGG-16 architecture that incorporates batch normalization.

A proficient convolutional neural network (CNN) was crafted for analyzing the optic disc through digital fundus images. This involved a swift and automated deep learning approach for detecting glaucoma [16].

Various traits of abnormal eyes were identified by utilizing pre-trained convolutional neural networks (CNN) and other deep learning-based neural networks. These models were employed for the classification of diseases and the early detection of abnormalities in their initial stages [17-18]. The major contributions are enlisted below.

- Prior research has predominantly focused on 2-3 classes, whereas our study comprehensively addresses 8 classes.
- We employ a transfer learning approach to enhance efficiency and reduce training time.
- This study incorporates a fine-tuning approach to further optimize model performance.
- The suggested model outperforms current methods, demonstrating superior effectiveness.

III. Methodology

A. Dataset

In this study, we utilized the publicly available Ocular Disease Intelligent Recognition (ODIR) dataset from Kaggle (see Figure 1). This dataset consists of colored fundus images from a diverse group of individuals, encompassing 8 distinct ocular disease diagnosis categories. The ODIR dataset comprises comprehensive ophthalmic data from 5,000 patients, including age, color fundus images of both left and right eyes, and diagnostic keywords provided by medical professionals. The dataset aims to present a comprehensive and authentic representation of patient data gathered from multiple hospitals and medical centers across China by Shang Gong Medical Technology Co., Ltd[19].

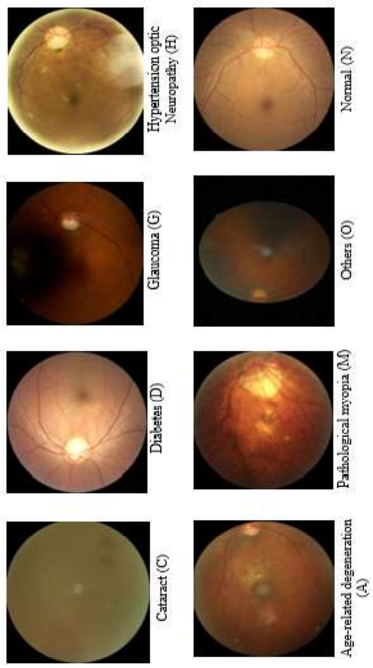

Figure 1. ODIR Dataset Samples

B. Transfer Learning

In this study, we have considered the transfer learningtechnique[20] where a pre-existing model, trained on one task, is utilized to tackle a different but related task. Instead of training a model from scratch, transfer learning takes advantage of the knowledge and learned representations already present in the pre-trained model to accelerate and improve learning on the new target task.

Deep neural networks trained on extensive datasets, like ImageNet for image recognition or BERT for natural language processing, capture generic features and patterns that can be transferred to various tasks. These models learn hierarchical representations that can be repurposed for different problems. By taking a pre-trained model, initially designed for a complex task, and applying it to a related yet distinct task or a smaller dataset, significant time and computational resources can be saved, all while enhancing the model's performance on the new task. In this study, we leverage pre-trained models that were trained on large datasets and then fine-tune them for specific tasks defined within our target dataset. By doing so, we aim to leverage the insights and learned representations captured by these pre-trained models to enhance performance on our specific task.

In this study, we have used all the ImageNet weights for the eye disease classification.

C. Pretrained Model Architectures

Pre-trained models are models that have been trained on large datasets, such as ImageNet, which contains millions of records and images. These models are already created and trained by others, that is why referred to as pre-trained models. For instance, in the context of eye disease classification in this study, we utilize pre-trained models like VGG16, RESNET50, and others, which have been trained on extensive datasets.

Among the deep CNN architectures used for eye disease classification in this research are:

- ResNet50: It is a well-known convolutional neural network with 50 layers and approximately 23.5 million trainable parameters. Its design, incorporating residual blocks, has led to its widespread adoption as a preferred model for various computer vision tasks [21].
- ResNet152: This model consists of 152 layers and around 60.2 million trainable parameters. Its use of residual blocks enhances its capacity to learn and represent complex visual patterns, making it a powerful model for diverse computer vision applications [21].
- VGG16: It is a 16-layer variation of VGG models, comprising 13 convolutional layers and three fully connected layers. VGG16 has been widely used in various computer vision tasks for its effectiveness and simplicity [22].
- VGG-19: It is a variation of the VGG model with 19 layers, comprising 16 convolution layers, 3 fully connected layers, 5 MaxPooling layers, and 1 SoftMax layer. VGG-16, on the other hand, is a 16-layer version of VGG models, containing 13 convolutional layers and three fully connected layers [22].
- DenseNet-201: With 201 layers and approximately 20.6 million trainable parameters, DenseNet-201's dense connection pattern makes it an excellent choice for various computer vision tasks. It allows for optimal feature reuse and smooth gradient flow across the network [23].

D. Hyper Parameters

To obtain optimal performance, we carefully set the hyperparameters for training the deep CNN models employed in this work. Iterative experimentation was used to fine-tune the hyperparameters in order to strike the correct balance between model complexity and training efficiency.

We fixed the number of epochs for all models to 200 to ensure that the models get enough iterations over the training data without overfitting. A learning rate of 0.001 was used to achieve constant convergence during gradient updates, avoiding abrupt oscillations that could impede training.

To further regularize the models and prevent overfitting, we employed a batch size of 32, balancing computational efficiency with the ability to capture meaningful gradients from each mini-batch. Additionally, we utilized the Adam optimizer due to its adaptive learning rate and momentum features, which help in efficient weight updates during training, see Table 1 for hyper-parameters values.

Table 1. Experimental Hyper-Parameters

Parameter	**Values**
Batch Size	32
Number of Epochs	200
Activation	ReLU and SoftMax
Optimizer	Adam
Learning Rate	0.0001
Freeze Layers	Yes

Figure 2 depicts the abstract process of the proposed study.

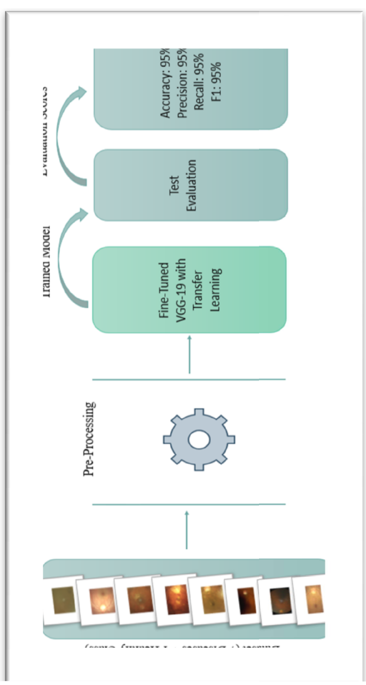

Figure 2. Abstract Diagram of the Proposed Study

IV. RESULTS

The evaluation of our eye disease detection model on the test set has provided us with valuable insights into its performance, offering implications for our project's objectives. Remarkably, the model achieved an accuracy of 95%, showcasing its competence in precisely classifying various eye diseases and bolstering its relevance in real clinical settings.

A deeper analysis of precision, recall, and F1 scores for each eye disease category sheds light on the model's effectiveness in identifying specific conditions. In the training process, accuracy is increasing while loss is decreasing, see Figure 3.

The confusion matrix further enhances our understanding of the model's predictive behavior, revealing potential misclassifications or biases towards certain classes (see Figure 4). The proposed model tested for approximately 973 images. By scrutinizing the distribution of predicted labels in comparison to the ground truth labels, we can pinpoint areas where the model might benefit from further refinement or targeted improvement. These findings empower us to refine and optimize the model, ensuring it delivers accurate and reliable results in the challenging realm of eye disease detection.

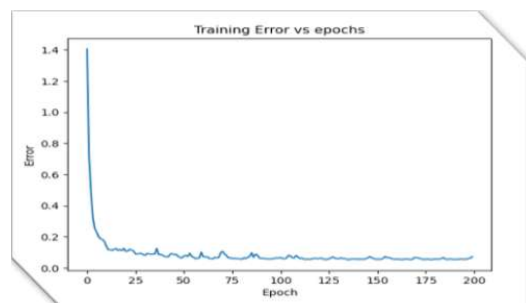

Figure 3.VGG-19 Training Error

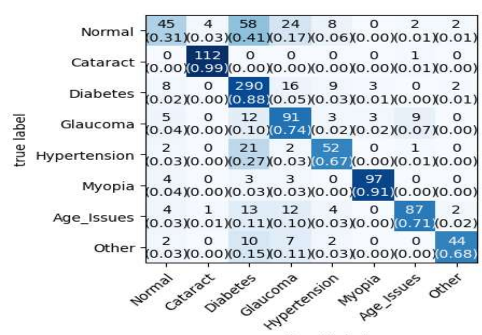

Figure 4.Confusion Matrix of VGG-19

For the efficacy of the proposed model, we have trained the different models with same parameters for better comparison, see Table 1.

VGG-19 outperformed than VGG-16, Resnet-50, ResNet152 and DenseNet-121 pretrained models with same configurations.

Table 1. CNN-Based Pretrained Model Comparison

Model	Accuracy	Precision	Recall	F1
VGG-19 Proposed	0.95	0.95	0.95	0.95
VGG-16	0.94	0.94	0.94	0.94
ResNet-50	0.93	0.93	0.93	0.93
ResNet-152	0.94	0.94	0.94	0.94
DenseNet-121	0.92	0.92	0.92	0.92

V. CONCLUSION

In this study, we explore the performance of various pre-trained classification models in the context of classifying eye diseases from a multiclass perspective. Detecting and classifying eye problems automatically remains a challenging task, particularly for early diagnosis. However, by leveraging transfer learning of existing models, we were able to achieve high accuracy while reducing the workload on the new model.

We evaluated the performance of five pre-trained models: VGG-16, VGG-19, ResNet-50, ResNet-152, and DenseNet-121, on the task of classifying eye diseases. To assess the classification process, we used several evaluation metrics, including recall, precision, accuracy, and F1-score, across all five models. Among the models tested, VGG-19 emerged as the top-performing one, achieving an impressive 95% accuracy, recall, precision, and F1-score. The results shown that transfer learning is effective to gain effective results with limited data and limited training process. The computerized classification of eye diseases has advanced significantly as a result. However, in order to improve the methods and carry out a more thorough examination.The ultimate goal is to improve diagnostic precision and help identify and treat eye diseases more effectively.

The future direction could involve implementing advanced classification methods with extensive datasets. This approach can be applied to address other challenges in medical imaging, aiming for prompt and dependable results.

REFERENCES

[1] S. Resnikoff "Global data on visual impairment in the year 2002," *SciELO Public HealthS Resnikoff, D Pascolini, D Etya'Ale, I Kocur, R Pararajasegaram, GP Pokharel, SP MariottiBulletin of the world health organization, 2004•SciELO Public Health*, vol. 82, no. 11, 2004, Accessed: Dec. 26, 2023. [Online]. Available: https://www.scielosp.org/pdf/bwho/v82n11/v82n11a09.pdf

[2] D. Pascolini, S. Mariotti, … G. P.-O., and undefined 2004, "2002 global update of available data on visual impairment: a compilation of population-based prevalence studies," *Taylor & FrancisD Pascolini, SP Mariotti, GP Pokharel, R Pararajasegaram, D Etya'Ale, AD NégrelOphthalmic epidemiology, 2004•Taylor & Francis*, vol. 11, no. 2, pp. 67–115, Apr. 2004, doi: 10.1076/opep.11.2.67.28158.

[3] X. Xu, L. Zhang, J. Li, … Y. G.-I. journal of biomedical, and undefined 2019, "A hybrid global-local representation CNN model for automatic cataract grading," *ieeexplore.ieee.orgX Xu, L Zhang, J Li, Y Guan, L ZhangIEEE journal of biomedical and health informatics, 2019•ieeexplore.ieee.org*, Accessed: December. 26, 2023. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/8705272/

[4] H. Zhang, K. Niu, Y. Xiong, W. Yang, … Z. H.-C. methods and, and undefined 2019, "Automatic cataract grading methods based on deep learning," *Elsevier*, Accessed: December. 26, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0169260719307163

[5] C. L. Dondapati*, A. Ghosh, and Dr. TYJ. N. Malleswari, "Classification of Eye Disorders based on Deep Convolutional Neural Network," *International Journal of Innovative Technology and Exploring Engineering*, vol. 9, no. 6, pp. 1388–1393, Apr. 2020, doi: 10.35940/IJITEE.F4209.049620.

[6] M. Khan, M. Ahmed, … R. R.-2021 I. W. A. I., and undefined 2021, "Cataract detection using convolutional neural network with VGG-19 model," *ieeexplore.ieee.orgMSM Khan, M Ahmed, RZ Rasel, MM Khan2021 IEEE World AI IoT Congress (AIIoT), 2021•ieeexplore.ieee.org*, Accessed: Dec. 26, 2023. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9454244/

[7] M. Mubeen, M. A. Arshed, and H. A. Rehman, "DeepFireNet - A Light-Weight Neural Network for Fire-Smoke Detection," pp. 171–181, 2022, doi: 10.1007/978-3-031-10525-8_14.

[8] "Multiclass Brain Tumor Classification from MRI Images using Pre-Trained CNN Model," *VFAST Transactions on Software Engineering*, doi: 10.21015/VTSE.V10I4.1182.

[9] M. A. Arshed *et al.*, "Pneumonia Classification from Chest X-ray Images Using Pre-Trained Network Architectures," *VAWKUM Transactions on Computer Sciences*, 2022, Accessed: Dec. 05, 2023. [Online]. Available: https://www.researchgate.net/profile/Muham

mad-Asad-Arshed/publication/366876512_Pneumonia_Classification_from_Chest_X-ray_Images_Using_Pre-Trained_Network_Architectures/links/63b66a0e097c7832ca8f2bad/Pneumonia-Classification-from-Chest-X-ray-Images-Using-Pre-Trained-Network-Architectures.pdf?_sg%5B0%5D=started_experiment_milestone&origin=journalDetail&_rtd=e30%3D

[10] M. Ubaid, M. Khan, … M. R.-2021 I., and undefined 2021, "COVID-19 SOP's violations detection in terms of face mask using deep learning," *ieeexplore.ieee.org*, Accessed: Dec. 26, 2023. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9692999/

[11] Issac SMA, Dutta MK (2018) An automated and robust image processing algorithm for glaucoma diagnosis from fundus images using novel blood vessel tracking and bend point detection. *Int J Med Inform* 110:52–70. https://doi.org/10.1016/j.ijmedinf.2017.11.015.

[12] Ganguly S, Ganguly S, Srivastava K, Dutta MK, Parthasarathi M, Burget R, Riha K (2014) An adaptive threshold based algorithm for detection of red lesions of diabetic retinopathy in a fundus image. In: *2014 International conference on medical imaging, m-health and emerging communication systems (MedCom)* pp 91–94. https://doi.org/10.1109/MedCom.2014.7005982.

[13] Long, S., Chen, J., Hu, A., Liu, H., Chen, Z., & Zheng, D. (2020). Microaneurysms detection in color fundus images using machine learning based on directional local contrast. *BioMedical Engineering Online*, 19(1), 1–23. https://doi.org/10.1186/S12938-020-00766-3/FIGURES/13.

[14] Roy, A., Conjeti, S., Karri, S., … D. S.-B. optics, & 2017, undefined. (n.d.). ReLayNet: retinal layer and fluid segmentation of macular optical coherence tomography using fully convolutional networks. Opg.Optica.OrgAG Roy, S Conjeti, SPK Karri, D Sheet, A Katouzian, C Wachinger, N NavabBiomedical Optics Express, 2017•opg.Optica.Org. Retrieved Dec. 28, 2023, from https://opg.optica.org/abstract.cfm?uri=boe-8-8-3627.

[15] A. Govindaiah, M. A. Hussain, R. T. Smith and A. Bhuiyan, "Deep convolutional neural network based screening and assessment of age-related macular degeneration from fundus images," *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, Washington, DC, USA, 2018, pp. 1525-1528, doi: 10.1109/ISBI.2018.8363863.

[16] R. C. Joshi, M. K. Dutta, P. Sikora and M. Kiac, "Efficient Convolutional Neural Network Based Optic Disc Analysis Using Digital Fundus Images," *2020 43rd International Conference on Telecommunications and Signal Processing (TSP), Milan, Italy, 2020*, pp. 533-536, doi: 10.1109/TSP49548.2020.9163560.

[17] Wu, T., Liu, L., Zhang, T., & Wu, X. (2022). Deep learning-based risk classification and auxiliary diagnosis of macular edema. *Intelligence-Based Medicine*, 6, 100053. https://doi.org/10.1016/J.IBMED.2022.100053.

[18] Kumar, S., & Kumar, B. (2022). Automatic early glaucoma detection by extracting parapapillary atrophy and optic disc from fundus image using SVM. *Multimedia Tools and Applications*, 81(10), 13513–13535. https://doi.org/10.1007/S11042-021-11023-7.

[19] "Ocular Disease Recognition | Kaggle." https://www.kaggle.com/datasets/andrewmvd/ocular-disease-recognition-odir5k (accessed Dec. 26, 2023).

[20] A. Hosna, E. Merry, J. Gyalmo, Z. Alom, Z. Aung, and M. A. Azim, "Transfer learning: a friendly introduction," *J Big Data*, vol. 9, no. 1, pp. 1–19, Dec. 2022, doi: 10.1186/S40537-022-00652-W/FIGURES/6.

[21] Huang et al., Deep residual learning for image recognition. Openaccess.Thecvf.Com. Retrieved Dec. 26, 2023, from http://openaccess.thecvf.com/content_cvpr_2016/html/He_Deep_Residual_Learning_CVPR_2016_paper.html.

[22] Simonyan, K., & Zisserman, A. (2014). Very Deep Convolutional Networks for Large-Scale Image Recognition. *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*. https://arxiv.org/abs/1409.1556v6.

[23] Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2016). Densely Connected Convolutional Networks. *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, 2017-January, 2261–2269. https://doi.org/10.1109/CVPR.2017.243