ISSN 2348-1196 (print)
International Journal of Computer Science and Information Technology Research ISSN 2348-120X (online)
Vol. 11, Issue 4, pp: (72-79), Month: October - December 2023, Available at: www.researchpublish.com

A Deep Learning Approach to Classifying Schizophrenia Based on EfficientNet Convolutional Neural Network Models

Michael Ayeni[1], Prof. S.O Adewale, Dr. Ibam[3]

[1,2]Federal University of Technology, Akure, Ondo State, Nigeria

DOI: https://doi.org/10.5281/zenodo.10390109

Published Date: 15-December-2023

Abstract: **The accurate detection and classification of schizophrenia are vital for early intervention and treatment. The realm of medical image analysis encounters a hurdle due to the scarcity of publicly accessible data, often compelling researchers to grapple with small and imbalanced datasets. To address this issue, transfer learning techniques come to the rescue by allowing the utilization of general features from smaller target datasets. This research investigates the application of the EfficientNets for the classification of Schizophrenia rs-fmri images obtained from schizconnect. The dataset is split into three sets with 60% for training, 20% for testing and the remaining 20% for evaluation The results indicate promising classification performance for different EfficientNet convnet variants, with high precision, recall, F1-score, and accuracy. The B3 architecture, in particular, demonstrates exceptional performance, achieving 99.7% accuracy. The findings of this research provide valuable insights into the classification of schizophrenia using neuroimaging data and state-of-the-art neural network models.**

Keywords: **EfficientNet, Schizophrenia, ConvNets. Transfer Learning.**

I. INTRODUCTION

It has been observed the last couple of years, the progression of artificial intelligence (AI) and Machine Learning (ML) within the medical field has been particularly noteworthy ever since modern computing became more present. This development is due to advancements in computational capabilities and the growing intricacies of medicine. The convergence of AI and medicine has given rise to increased collaboration between these two communities, offering substantial untapped potential [1]. Artificial Intelligence (AI) and Machine Learning (ML) technologies have been thoroughly investigated for developing predictive models, demonstrating diverse applications across multiple medical and healthcare scenarios. [2]. Applications encompass tasks such as the extraction and categorization of medical data [3], real-time analysis of medical imaging [4], the potential for medical condition diagnoses [5], and the automation of medical processes, including detection and classification [6].

Psychological disorders can be broadly categorized into two major groups: psychological and neurological disorders. These conditions have a significant impact on an individual's cognitive and behavioral capabilities [7][8]. One such severe psychiatric disorder is schizophrenia, which presents as a complex set of cognitive and behavioural symptoms. It is often associated with disruptions in brain development caused by genetic or environmental factors [9]. Historically, schizophrenia diagnosis has relied on clinical assessments conducted by psychiatrists, a procedure that can be time-intensive. Nevertheless, there is an increasing curiosity about the capability of Machine Learning algorithms to aid in the prompt identification of schizophrenia. [10].

It's been observed that deep learning, a component of the broader field of machine learning, has gained prominence in the detection and classification of health conditions. It has demonstrated efficacy for integrating, analysing, and predicting data from diverse sources. When it comes to image classification, one of the most accurate techniques is the convolutional neural network (CNN or ConvNet), which employs filters to identify specific features in images [11]

While previous methods for classifying schizophrenia have demonstrated relative success, they tend to be computationally intensive. In this context, the research delves into the utilization of seven compact architectures belonging to the EfficientNet family [12]. These EfficientNet models were explicitly crafted to enhance resource efficiency without compromising high accuracy. The objective is to attain performance comparable to prevailing deep learning approaches, all while reducing resource demands and training duration. Transfer learning methods, notably fine-tuning, have gained popularity as valuable enhancements in deep learning solutions designed for classification purposes.

Various diagnostic techniques are employed to assess schizophrenia. Medical professionals, including neurologists and geriatricians, utilize diverse methods and instruments for diagnosis. The diagnostic process for schizophrenia can encompass the following:

- **Clinical examination**: This assessment is conducted to rule out alternative factors presenting similar symptoms and to identify any associated complications.

- **Tests and screenings**: These evaluations help exclude conditions with symptoms resembling schizophrenia, including screenings for alcohol and drug use. Additional assessments may involve imaging studies such as MRIs or CT scans.

- **Mental health evaluation**: A healthcare provider or mental health specialist assesses the patient's mental state by observing physical appearance, behavior, and inquiring about thoughts, moods, delusions, hallucinations, substance use, and the potential for violence or self-harm. It also involves discussions about the patient's family and personal history.

- **Employing diagnostic criteria for schizophrenia:** Healthcare practitioners may utilize the criteria specified in the Diagnostic and Statistical Manual of Mental Disorders (DSM-5), published by the American Psychiatric Association.

II. AIM AND OBJECTIVES OF THE DISSERTATION

The aim of this research is to use EfficientNet models for schizophrenia classification based on fMRI data.

Objectives are to:

- Design a deep learning model for classification of schizophrenia,
- Implement the model; and
- Evaluate the performance of the model

III. METHODOLOGY

Introduction

As we seek to detect and classify schizophrenia, different techniques and methods are employed. In this chapter, the methods and techniques used for classifying schizophrenia in this dissertation are discussed.

Proposed Schizophrenia Classification Model

The model proposed begins with obtaining dataset from Schizconnect, in this case, the COBRE rs-fMRI dataset is used. The COBRE dataset contains images of individuals with varying gender and ages, hence, we are able to test our model on data with different age ranges and gender. The diagrammatic framework of the proposed model is shown below.

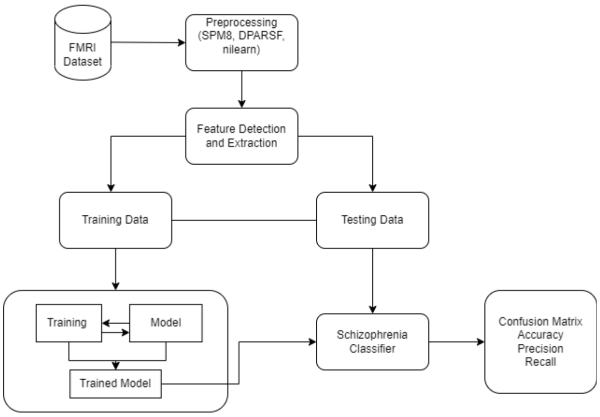


Figure 1: Diagrammatic Framework of the proposed model

ISSN 2348-1196 (print)
International Journal of Computer Science and Information Technology Research ISSN 2348-120X (online)
Vol. 11, Issue 4, pp: (72-79), Month: October - December 2023, Available at: www.researchpublish.com

Participants

The data used in this project is hosted on schizconnect.org. The Center for Biomedical Research Excellence (COBRE) had shared raw functional MRI dataset from a total of 78 individuals diagnosed with schizophrenia. Both groups have participants ranging in age from 18 to 65 years. To ensure data reliability, all participants underwent a thorough screening process. Exclusion criteria included a history of mental retardation, severe head trauma and neurological disorders leading to loss of consciousness exceeding 5 minutes, or substance dependence over the last 12 months. The participants diagnostic information was gathered from 78 patients with schizophrenia, comprising 68 males and 14 females, also information of 91 Healthy with no known disorder were retrieved.

Table 1: Characteristics of the participants involved in this study

Statistics/DX	Patients (SCZ)	Healthy Control
Number of Subjects	78	91
Gender (m/f)	64/14	65/26
Age (years) (mean ± s.d.)	37.8 ± 14.2	38.5 ± 11.7
Age range	19.0-66.0	19.0-66.0

Image Acquisition

These details are from the contributors of the fmri images on schizconnect. The rs-fMRI were gathered using a single-shot full k-space echo-planar imaging (EPI) technique with ramp sampling correction, utilizing the intercommissural line as a reference. The parameters for the scan were as follows:

150 volumes, 33 slices, a repetition time (TR) of 2000 ms, an echo time (TE) of 29 ms, a field of view (FOV) measuring 256 x 256, an acquisition matrix of 64 x 64, and a voxel resolution of 3 x 3 x 4 mm^3.

In Addition, high-resolution T1-weighted image of the brain were gathered with the aid of a multi-echo magnetization-prepared rapid gradient echo sequence with the following settings:

- TR of 2530 ms,
- TE values of [1.64, 3.5, 5.36, 7.22, 9.08] ms,
- Flip angle of 7°, a slab thickness of 176 mm
- FOV of 256 x 256 mm, an acquisition matrix of 256 x 256, and a voxel resolution of 1 x 1 x 1 mm^3.

Image Preprocessing

We conducted pre-processing and feature engineering using MATLAB, with a focus on utilizing the toolboxes, including Data Processing Assistant for Resting State fMRI (DPARSF). DPARSF is a practical plug-in software integrated into Data Processing and Analysis for Brain Imaging (DPABI). It is built upon the foundation of Statistical Parametric Mapping (SPM). We harnessed DPARSF to perform various preprocessing tasks, encompassing slice timing correction, realignment, normalization, and smoothing, along with generating data related to functional connectivity, ReHo (Regional Homogeneity), ALFF/fALFF (Amplitude of Low-Frequency Fluctuations/fractional ALFF), degree centrality, and voxel-mirrored homotopic connectivity (VMHC) results.

Rgb Transformation.

The next stage after preprocessing is converting the images to RGB by stacking three images on top of each other using numpy's dstack function. The idea behind this is that looking at 33 slices, it appears 3 are replica of each other as shown in the visualization below.

ISSN 2348-1196 (print)
International Journal of Computer Science and Information Technology Research ISSN 2348-120X (online)
Vol. 11, Issue 4, pp: (72-79), Month: October - December 2023, Available at: www.researchpublish.com

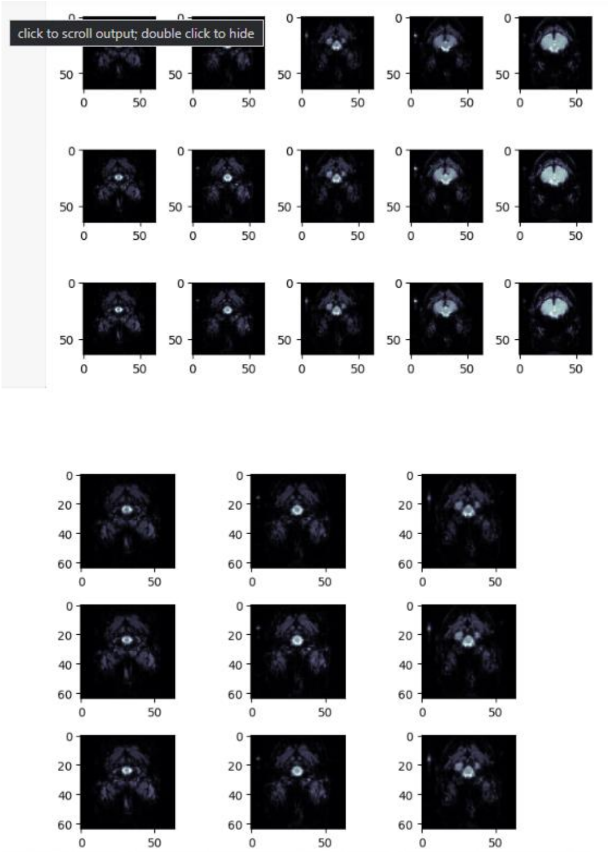


Figure 2: Visualization of different time slices to depict the similarity

Because of this fact, three images that follow each other in the plot above are made into a single image to form an rgb. The final RGB for the first three images now looks like this.

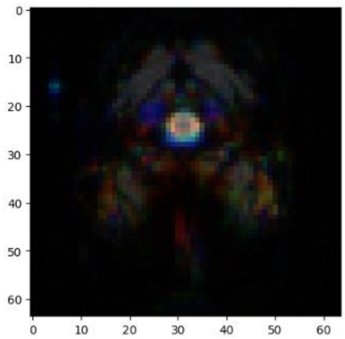


Figure 3: RGB conversion output

However, the last 3 slices upon converting to RGB doesn't seem to hold much information, so they were discarded. This means around 5 time-slices chosen from each image multiplied by 10 RGB images from each slice to give approximately 50 images per subject.

EfficientNets

The expansion process of convolutional neural networks is not fully comprehended and is often carried out in a somewhat arbitrary manner until a satisfactory result is attained. This approach can be time-consuming due to the necessity for manual adjustment of relevant parameters [12]. Previous strategies for scaling networks have involved increasing model depth [13], width [14], and image resolution [15]. Tan et al. conducted a study to investigate the effects of these scaling techniques, aiming to establish a more systematic approach to expanding network architecture. Their research reveals two primary observations: Firstly, enhancing any one aspect of network resolution, depth, or width improves accuracy; however, this accuracy gain diminishes as models grow larger. Secondly, achieving superior accuracy and efficiency requires a balanced adjustment of a network's depth, width, and resolution, rather than concentrating solely on one of these aspects. In light of these findings, the authors introduced an innovative scaling approach that employs a robust compound coefficient, ϕ, for a more organized expansion of networks. Equation 1 illustrates the authors' suggested method for scaling depth, width, and resolution in relation to ∅.

$$d = \alpha^{\emptyset}$$

$$w = \beta^{\emptyset}$$

$$r = \gamma^{\emptyset}$$

$$s.t \quad \alpha.\beta^2.\gamma^2 \approx 2$$

$$\alpha \geq 1, \qquad \beta \geq 1, \qquad \gamma \geq 1$$

Here, the authors introduce constants α, β, and γ, which can be determined through a limited grid search. In simple terms, φ represents a user-defined coefficient controlling the allocation of additional resources for scaling the model. Meanwhile, α, β, and γ specify how these extra resources are distributed to extend the network's width, depth, and resolution. It's important to note that the computational cost of a standard convolution operation is directly linked to d, w^2, and r^2. This implies that doubling the network's depth results in a twofold increase in computational cost, while doubling the width or resolution leads to a fourfold increase. Given that convolution operations are typically the most resource-intensive in ConvNets, employing the provided equation to scale a ConvNet will approximately increase the total computational cost by $(\alpha \cdot \beta^2 \cdot \gamma^2)\varphi$.

The model training process involves the use of the 'binary cross-entropy' loss function sourced from Keras. This particular loss function, alternatively known as logarithmic loss or log loss, operates as a metric within the model, overseeing the misclassification of data class labels. A reduced log loss value signifies greater accuracy. The formulation of the loss function is expressed:

$$L(y,p) = -\frac{1}{N}\sum_{i=1}^{N}[yi \cdot log(pi) + (1 - yi) \cdot log(1 - pi)]]$$

The chosen activation function is the sigmoid, specifically selected for binary classification purposes in the model. Widely known as the logistic function in neural networks, the sigmoid activation function possesses the capability to transform any real-valued number into a value within the range of 0 to 1. It's formula is as described below:

$$f(x) = \frac{1}{1 + e^{-x}}$$

Experimental Settings

In the training phase, a dataset comprising 8620 images was employed, with a distribution of 60% for training, 20% for testing, and the remaining 20% for evaluation. During the data split, a stratified approach was adopted to ensure equal representation of classes in the training, testing, and validation sets. The implementation of this method involved leveraging the Keras package, a component of the TensorFlow Python library. Specifically, we utilized the Keras ImageDataGenerator to create augmentation generators for both the training and validation datasets. Additionally, publicly available pre-trained EfficientNet models [16] were incorporated into the process.

IV. RESULTS AND DISCUSSION

Evaluation Criteria

Assessing the performance of each model included calculating precision, recall, f1-score, and accuracy. The ensuing equations depict the methodologies employed to compute these metrics.

$$Accuracy = \frac{\mathrm{TP + TN}}{\mathrm{TP + TN + FP + FN}}$$

$$Precisions = \frac{\mathrm{TP}}{\mathrm{TP + FP}}\ x\ 100$$

ISSN 2348-1196 (print)
International Journal of Computer Science and Information Technology Research ISSN 2348-120X (online)
Vol. 11, Issue 4, pp: (72-79), Month: October - December 2023, Available at: www.researchpublish.com

$$Recall = \frac{TP}{TP + FN} X\ 100$$

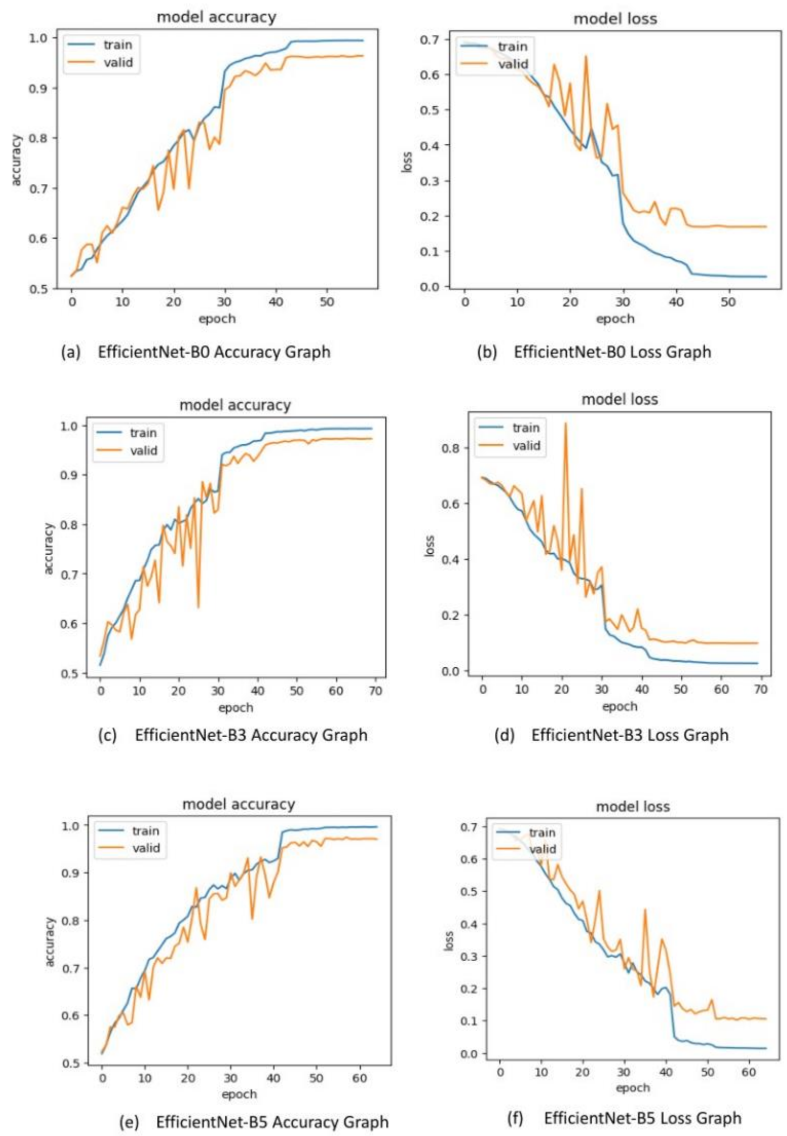
$$f1 - score = 2\ x\ \frac{Precision\ x\ Recall}{Precision + Recall}$$

where TP = true positives, TN = true negatives, FP = false positives, and FN = false negatives.

Analysis of Results

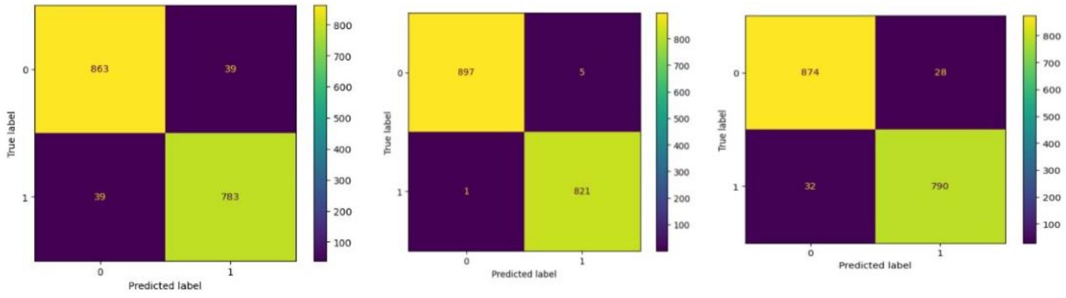
Table 2: Classification Results for the EfficientNet Architectures

EfficientNet	Epoch	Precision	Recall	F1-score	Accuracy
B0	58	95%	95%	95%	95.5%
B1	76	96%	96%	96%	96.3%
B3	61	100%	100%	100%	99.7%
B5	65	97%	97%	97%	96.5%
B7	84	97%	97%	97%	96.64%



(a) EfficientNet-B0 Accuracy Graph
(b) EfficientNet-B0 Loss Graph
(c) EfficientNet-B3 Accuracy Graph
(d) EfficientNet-B3 Loss Graph
(e) EfficientNet-B5 Accuracy Graph
(f) EfficientNet-B5 Loss Graph

Figure 4: Accuracy and Loss graphs



(a) EfficientNet-B0 Confusion Matrix (b) EfficientNet-B3 Confusion Matrix (c) EfficientNet-B5 Confusion Matrix

Figure 5: Confusion Matrix

Limitations

The main challenge encountered in these experiments centered on the issue of overfitting. Because the size of the dataset is a bit substantial, overfitting became unavoidable when utilizing large neural network architectures. To address this, various regularization methods were explored beyond conventional data augmentation techniques. These methods included the integration of dropout layers, early stopping, and model checkpointing. Identifying the appropriate dropout rate required a grid search, which also involved determining the optimal batch size and the number of dense layers. Additionally, alternative activation and loss functions were experimented with, but they tend to exacerbate the issue of overfitting and resulted in unstable learning.

V. CONCLUSION

The EfficientNet family represents a set of advanced convolutional neural networks that have demonstrated exceptional performance on image data by effectively optimizing vital network measurements, achieving a balance between accuracy and efficiency. In the realm of medical image analysis, the challenge arises from the limited availability of publicly accessible data, often necessitating researchers to work with small and imbalanced datasets. To overcome this obstacle, transfer learning techniques prove beneficial by enabling the utilization of general features from smaller target datasets. Among the examined models, EfficientNet-B3 architecture yielded the most favourable outcomes, achieving an accuracy of 99.7%. This study can serve as a valuable reference for future works in the field of medical diagnosis. While the presented models attained a notable level of accuracy, they can be fine-tuned for application in diagnosing various psychiatric disorders.

REFERENCES

[1] Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., ... Wang, Y. (2017). Artificial intelligence in healthcare: Past, present and future. Stroke and Vascular Neurology, 2(4), 230–243. https://svn.bmj.com/content/2/4/230.full.pdf

[2] Wang, G., Zuluaga, M. A., Li, W., Pratt, R., Patel, P. A., Aertsen, M., ... Ourselin, S. (2018). DeepIGeoS: A deep interactive geodesic framework for medical image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 41(7), 1559–1572.

[3] Clough, J., Balfour, D. R., Da Cruz, G. L., Marsden, P., Prieto, C., Reader, A., & King, A. (2019). Weighted Manifold Alignment using Wave Kernel Signatures for Aligning Medical Image Datasets. IEEE Transactions on Pattern Analysis and Machine Intelligence, 988–997.

[4] Ghesu, F. C., Georgescu, B., Zheng, Y., Grbic, S., Maier, A., Hornegger, J., & Comaniciu, D. (2017). Multi-scale deep reinforcement learning for real-time 3D-landmark detection in CT scans. IEEE Transactions on Pattern Analysis and Machine Intelligence, 41(1), 176–189.

[5] Panicker, S. S., & Gayathri, P. (2019). A survey of machine learning techniques in physiology-based mental stress detection systems. Biocybernetics and Biomedical Engineering, 39(2), 444–469. doi: 10.1016/j.bbe.2019.01.004

[6] Li, H.; Zhang, B.; Zhang, Y.; Liu, W.; Mao, Y.; Huang, J.; Wei, L (2020). A semi-automated annotation algorithm based on weakly supervised learning for medical images. Biocybern. Biomed. Eng., 40, 787–802

[7] Kaur, P., & Sharma, M. (2019). Diagnosis of human psychological disorders using supervised learning and nature-inspired computing techniques: A meta-analysis. Journal of Medical Systems, 43(7), 1–30.

[8] Gautam, R., & Sharma, M. (2020). Prevalence and diagnosis of neurological disorders using different deep learning techniques: A meta-analysis. Journal of Medical Systems, 44(2), 1–24.

[9] Owen, M. J., Sawa, A., & Mortensen, P. B. (2016). Schizophrenia. The Lancet, 388, 86–97.

[10] Rezaii, N., Walker, E., & Wolff, P. A. (2019). Machine learning approach to predicting psychosis using semantic density and latent content analysis. NPJ Schizophrenia, 5(1), 203.

[11] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. In Proceedings of the 25th International Conference on Neural Information Processing Systems, Volume 1 (pp. 1097–1105). USA.

[12] Tan, M., & Le, Q. V. (2019). Efficientnet: Rethinking model scaling for convolutional neural networks. CoRR, abs/1905.11946.

[13] He, K., Zhang, X., Ren, S., & Sun, J. (2015). Deep residual learning for image recognition. https://arxiv.org/abs/1512.03385

[14] Sergey Zagoruyko and Nikos Komodakis. (2016). Wide residual networks. https://arxiv.org/abs/1605.07146

[15] Huang, Y., Cheng, Y., Bapna, A., Firat, O., Chen, M. X., Chen, D., ... Wu, Y. (2018). Gpipe: Efficient training of giant neural networks using pipeline parallelism. https://arxiv.org/abs/1811.06965

[16] Yakubovskiy, P. (2020, September). qubvel/efficientnet. [Computer Software]. GitHub. https://github.com/qubvel/efficientnet