Vol. 10, No. 1, March, 2023, pp. 14-19

Journal homepage: http://iieta.org/journals/rces

Pokepedia: Pokemon Image Classification Using Transfer Learning

Venkata Rami Reddy Chirra^{1*}, Afifa Syeda¹⁰, Nuthana Kolla¹⁰, Neha Ghanta¹⁰, Suneetha Muvva²⁰



¹ School of Computer Science & Engineering, VIT-AP University, Amaravati 522237, India ² Independent Researcher, Amaravati 522237, India

Corresponding Author Email: venkataramireddy.chirra@vitap.ac.in

https://doi.org/10.18280/rces.100103

ABSTRACT

Received: 6 January 2023 Accepted: 25 February 2023

Keywords:

computer vision, MobileNetV2, EfficientNetB7, EfficientNetV2L, DenseNet201, ResNet101, VGG19 and VGG16 Identifying images of various objects, living creatures, food, etc., and classifying them using machine learning has become a common task in computer vision. Humans may not identify every object they see, here comes machine learning that eases the life of human beings by identifying the object for the human. Pokémon is a cartoon that is widely watched by the majority of the younger generation around the world. The aim of this work to predict and classify Pokémon images using pre-trained models. In the proposed work, seven pre-trained models namely MobileNetV2, EfficientNetB7, EfficientNetV2L, DenseNet201, ResNet101, VGG19 and VGG16 were utilised to classify ten Pokémon characters which includes Pikachu, Raichu, Charmander, Bulbasaur, Squirtle, Eevee, Piplup, Snorlax, Jigglypuff, and Psyduck. The performance of the pre-trained models were evaluated on a dataset collected from the internet. The ResNet101 pre-trained model produces the highest accuracy of 95.60% when compared with the other models.

1. INTRODUCTION

In recent years, deep learning techniques have gained significant attention in various fields such as Pattern Recognition [1-9], Medical Imaging, Video Analysis, driver drowsiness detection [10, 11], video analysis, Spam detection [12], Healthcare, Clustering [13] and many more. One of the popular techniques is transfer learning, which allows the pre-trained models to be used for a new set of tasks with minimal training data. Transfer learning has shown promising results in image classification tasks, and it can significantly reduce training time and improve model performance.

The classification of Pokémon images can help in building a robust dataset for various applications such as gaming, educational purposes, and research. Due to the complex features and variations in Pokémon images, it is challenging to achieve high accuracy with traditional image classification techniques. Therefore, this research problem can be addressed using transfer learning techniques.

Existing studies on Pokemon image classification using transfer learning have shown promising results in accurately identifying Pokemon species in various scenarios and backgrounds. These studies can be categorized into different areas, such as data augmentation techniques, model architectures, and hyperparameter tuning. Despite the promising results, previous studies have some limitations, such as low accuracy, overfitting, and insufficient dataset size. To overcome these limitations, this report proposes a novel approach for the classification of Pokémon images using transfer learning and data augmentation techniques.

To solve the defects identified in the existing studies, this paper establishes a model based on transfer learning and data augmentation techniques and applies it to classify a dataset of Pokémon images. The aim of this study is twofold: first, it proposes a novel approach that achieves high accuracy in classifying Pokémon images, and second, it demonstrates the effectiveness of transfer learning and data augmentation techniques in improving model performance. The findings shed new light on the effectiveness of transfer learning and data augmentation in improving the performance of deep learning models for image classification tasks, specifically in the context of Pokémon images.

Section 1 presents the research background and motivation followed by a comprehensive review of the existing studies on transfer learning in Section 2. Section 3 describes the methodology utilized in this study, including the datasets, pretrained models, and evaluation metrics. Section 4 presents the results of the experiments and provides a comparative analysis of the performance of the pre-trained models. Section 5 conclusion finally, discuss the potential for future work in Section 6.

2. RELATED WORK

Performances of various pre-trained models were discussed in Table 1.

Author	Pre-trained model	Dataset	Accuracy (%)
Xiang et al.	MobileNetV2	Fruit image	85.12
[14]	WIODIICI VCI V 2	dataset	05.12
Zhang et al.	DansaNat160	NWNU-	87
[15]	Denselvet109	TRASH	02
Palakodati et	VGG16	Fruit image	80.67
al. [6]	V0010	dataset	89.02
Rajayogi et al.	Incontion V3	The Indian	87.0
[16]	nicception v 5	food dataset	07.9
Wang et al.	VCC16	Plant Village	00.4
[17]	10010	dataset	90.4

Bansal et al. [18]	VGG19	Caltech-101	93.73
Ahmad et al. [19]	ResNet	Breast Histology dataset	85

3. MATERIALS AND METHODS

3.1 Dataset augmentation

In this work, we have created the Pokémon dataset with 600 images collected from internet sources. Pokémon dataset consists of ten types of Pokémon characters that are Pikachu, Raichu, Charmander, Bulbasaur, Squirtle, Eevee, Piplup, Snorlax, Jigglypuff, Psyduck. The sample images in the dataset are shown in Figure 1. Data Augmentation was used to increase the images to 7386 images. In this work, we used the techniques like Scaling, Padding, Cropping, Flipping, Rotation, Translation, and Affine transformation for Data Augmentation.



Figure 1. Sample images of dataset





Figure 2. Framework of proposed work

In transfer learning, the knowledge of an already trained model is applied to different but related problems. In this case, the model's output from a layer prior to the output layer is used as input to a new classifier model, which operates the first and middle layers and retrains the following layers. This finetuning can have built a solid machine learning model with comparably little data and this method can reduce the training time for a neural network model. Here we used MobileNetV2, EfficientNetB7, EfficientNetV2L, DenseNet201, ResNet101, VGG19, VGG16 pre-trained models to predict and classify the Pokémon images. The framework of the proposed work givenin Figure 2.

3.2.1 VGG 16 and VGG19

Simonyan and Zisserman [20] created VGG16 for the ILSVRC 2014 competition. It includes sixteen convolutional layers with the most effective 3x3 kernels. The layout opted through authors is just like AlexNet i.e., growth in the quantity of the capabilities map or convolution because of the intensity of the community increases. The community incorporates 138 million parameters. In work, we changed the 1000 classes with our 10 classes. Accuracy is obtained using the Adam optimizer. Similarly, by pushing the intensity to 19 layers VGG19 structure is defined.

3.2.2 MobileNetV2

MobileNet-v2 [21] is a family of computer vision models for mobile devices that can be used as a base for many visual recognition tasks. MobileNetV2 is very similar to the original MobileNet, with the exception of the use of inverted residual blocks with bottlenecking features. Lightweight depth-wise convolutions are used as a source of nonlinearity in the intermediate expansion layer to filter features. The architecture of MobileNetV2 includes a 32-filter initial fully convolution layer as well as 19 additional bottleneck layers. It is employed for classification, detection, and alternative common tasks of CNNs that are 53 layers deep.

3.2.3 EfficientNetB7

As a convolution neural network architecture and scaling method EfficientNet [22] scales all the dimensions of depth/width/resolution uniformly using a compound coefficient. By applying compound scaling to the input image, the network can capture more fine-grained patterns on the bigger image by adding more layers and channels. Compared to the previous Gpipe, EfficientNet-B7 achieves 84.4% top-1 / 97.1% top-5 accuracy, and it is 8.4x smaller and 6.1x faster on CPU inference.

3.2.4 DenseNet201

DenseNet-201 [23] is a 201-layer convolutional neural network with feed-forward connections between each layer. DenseNets have several compelling advantages, including the ability to solve the vanishing-gradient problem, improve feature propagation, encourage feature reuse, and significantly reduce the number of parameters. On the majority of them, DenseNets outperform the state-of-the-art while requiring less memory and computation to achieve high performance.

3.2.5 ResNet101

The ResNet [24] achieved outstanding results in the ImageNet and MS-COCO competitions [6]. A convolutional neural network with 101 layers is called ResNet-101. They noticed relative improvements of 28% when ResNet-101 was

used to replace VGG-16 layers in Faster R-CNN. Residual Networks, often known as ResNets, is an extremely ingenious architecture that combines the input of every CNN block with its output. The foundation of residual blocks is a connection known as a skip connection.

3.2.6 EfficientNetV2L

Tan and Le [25] have come up again with EfficientNetv2 which have faster training speed and better parameter efficiency compared to previous convolutional neural networks [7] and their main idea is to train on smaller image sizes and test on larger image sizes. Training-aware neural architecture search and scaling are used by the authors to develop these models in order to jointly optimize training speed. As a result of the smaller kernel size of 3x3, EfficientNetV2 adds more layers to compensate for the reduced receptive field. The EfficientNetV2 models expect float tensors of pixels with values between [-1, 1] as inputs in the absence of preprocessing. Similarly, by modifying this architecture by

EfficientNetV2L [8] and with the number of classes in our model the structure is defined.

4. RESULTS AND DISCUSSIONS

The accuracies of the various pre-trained models were given in Table 2. From the Table 2 we can observe that ResNet101 pre-trained model produces the highest accuracy of 95.60% when compared with the other pre-trained models. The accuracies of pre-trained models were compared in Figure 3. Confusion Matrix of VGG16 was given in Table 3. Confusion Matrix of VGG19 was given in Table 4. The Confusion Matrix of ResNet101 was given in Table 5. Confusion Matrix of MobileNetV2 was given in Table 6. Confusion Matrix for DenseNet201 was given in Table 7. The Confusion Matrix of EfficientNetB7 was given in Table 8. The Confusion Matrix of EfficientNetV2L was given in Table 9.

Table 2. Accuracies of various pre-trained models (%)

Model	Precision	Recall	F1-Score	Accuracy
VGG16	90	90	90	89.72
VGG19	89	89	89	88.77
ResNet101	96	96	96	95.60
MobileNetV2	81	81	81	80.65
DenseNet201	91	91	91	90.6
EfficientNetB7	95	95	95	94.52
EfficientNetV2L	93	93	93	93.03



Figure 3. Comparison of pre-trained models accuracies

Table 3. VGG16 confusion matrix (%)

Classes	Raichu	Pikachu	Charmander	Bulbasaur	Squirtle	Eevee	Piplup	Snorlax	Jigglypuff	Psyduck
Raichu	82.9	1.8	3.7	3.7	1.2	0.6	1.8	1.8	1.2	0.6
Pikachu	0	93.08	1.88	1.25	0	0	1.25	1.88	0.62	0
Charmander	0	6.62	85.43	1.98	0.66	0.66	1.32	0	0.66	2.64
Bulbasaur	0	1.02	2.04	90.30	5.61	0	0	0.51	0	0.51
Squirtle	0	0	1.39	9.09	86.01	2.79	0.7	0	0	0
Eevee	0	0	0	0	4.7	95.29	0	0	0	0
Piplup	0	0	0.7	0.7	0	2.12	89.36	7.09	0	0
Snorlax	0	3.77	0	1.88	0.94	2.83	10.37	79.24	0	0.94
Jigglypuff	0	0	0	1.23	0.61	0	0	0	98.14	0
Psyduck	0	1.13	0	2.82	0.56	0	0.56	0	0	94.91

Table 4. VGG19 confusion matrix (%)

Classes	Raichu	Pikachu	Charmander	Bulbasaur	Squirtle	Eevee	Piplup	Snorlax	Jigglypuff	Psyduck
Raichu	89.24	0	4.43	0.63	2.53	0	1.26	0.63	0.63	0.63
Pikachu	1.25	89.3	3.77	0.62	1.88	0	1.88	0	1.25	0
Charmander	0	5.96	84.76	1.32	0.66	0.66	5.96	0	0	0.66
Bulbasaur	0.51	0.51	6.63	83.67	7.14	0	0.51	1.02	0	0
Squirtle	0	0	0	5.59	91.6	0	1.39	0	0.7	0.7
Eevee	0	0	0	0	10.58	84.7	4.7	0	0	0
Piplup	0	0	0	0	0.7	4.25	88.65	6.38	0	0
Snorlax	0	0.94	1.88	0.94	5.66	0	11.32	78.3	0.94	0
Jigglypuff	0	0	0	0	0.61	0	0.61	0	98.76	0
Psyduck	0.56	0	0.56	0.56	1.69	0.56	1.12	0	1.12	93.78

 Table 5. ResNet101 confusion matrix (%)

Classes	Raichu	Pikachu	Charmander	Bulbasaur	Squirtle	Eevee	Piplup	Snorlax	Jigglypuff	Psyduck
Raichu	93.6	0	3.7	1.8	0.6	0	0	0	0	0
Pikachu	0	88.6	7.5	0.6	0	0	2.5	0	0.6	0
Charmander	0	3.9	94.7	0	0	0.6	0.6	0	0	0
Bulbasaur	0	0.5	3.5	83.6	8.1	0.5	2	1.5	0	0
Squirtle	0	0	1.3	0.6	95.1	2	0.6	0	0	0
Eevee	0	0	0	0	7	90.5	2.3	0	0	0
Piplup	0	0	0	0	0	1.4	92.9	5.6	0	0
Snorlax	0	0	0	0	0.9	0.9	4.7	92.4	0.9	0
Jigglypuff	0	0	0	0	0	0	0	0	100	0
Psyduck	0	0	1.1	0.5	1.1	0.5	0	0	0	96.6

 Table 6. MobileNetV2 confusion matrix (%)

Classes	Raichu	Pikachu	Charmander	Bulbasaur	Squirtle	Eevee	Piplup	Snorlax	Jigglypuff	Psyduck
Raichu	88.6	1.2	0.6	1.89	0	0	2.5	1.89	0.6	2.5
Pikachu	0	93	0.6	0	0	0	1.8	3.1	0.6	0.6
Charmander	4.6	10.5	72.8	0.6	1.9	0	3.9	0	1.3	3.9
Bulbasaur	5.6	3	5.1	73.9	4.5	0	2.5	1	2	2
Squirtle	3.4	4.8	6.2	13.2	59.4	2	4.8	0	1.3	4.1
Eevee	4.7	0	2.3	0	11.7	71.7	2.3	1.1	2.3	3.5
Piplup	0.7	1.4	2.1	0	0	1.4	82.2	9.9	0	2.1
Snorlax	0.9	3.7	4.7	0	1.8	0	10.3	74.5	3.7	0
Jigglypuff	3.7	0	0.6	1.2	0.6	0	0	0	91.9	1.8
Psyduck	2.2	2.2	0	0	0	0	5	0	0.5	89.8

 Table 7. DenseNet201 confusion matrix (%)

Classes	Raichu	Pikachu	Charmander	Bulbasaur	Squirtle	Eevee	Piplup	Snorlax	Jigglypuff	Psyduck
Raichu	97.5	0.63	1.26	0	0	0	0	0	0	0.63
Pikachu	0	96.85	1.25	0	0	0	0.62	1.25	0	0
Charmander	0	13.24	81.45	1.32	0	0	1.32	1.32	0	1.32
Bulbasaur	0.51	0	4.59	86.73	5.1	0	0	1.02	0	2.04
Squirtle	0	1.4	0	6.3	81.1	1.4	2.8	0.7	0	6.3
Eevee	0	3.52	0	0	4.7	90.6	0	1.2	0	0
Piplup	0	2.12	0	0.7	0	4.25	86.52	5.67	0	0.7
Snorlax	0	4.71	0	0.94	0	0	4.71	89.62	0	0
Jigglypuff	0	0.61	0.61	0	0	0.61	2.46	0	95.1	0.61
Psyduck	0	1.12	0	0	0.56	0	0	0	0	98.3

 Table 8. EfficientNetB7 confusion matrix (%)

Classes	Raichu	Pikachu	Charmander	Bulbasaur	Squirtle	Eevee	Piplup	Snorlax	Jigglypuff	Psyduck
Raichu	98.10	1.27	0	0	0	0	0	0	0	0.63
Pikachu	0.63	96.85	1.26	0.63	0	0	0	0	0.63	0
Charmander	0.66	6.62	90.74	0.66	0.66	0	0	0	0	0.66
Bulbasaur	1.02	0	2.55	93.88	2.55	0	0	0	0	0
Squirtle	0	0.7	0.7	8.4	89.5	0.7	0	0	0	0
Eevee	0	0	0	0	11.77	87.05	0	0	0	1.18
Piplup	0	0	0.71	0	0	1.42	92.2	5.67	0	0

Snorlax	0	0	0	0	0.94	0	5.66	93.4	0	0
Jigglypuff	0	1.85	0	0	0	0	0	0	98.15	0
Psyduck	0	0	0	0	0	0	0	0	0	100

Table 9. EfficientNetV2L confusion matrix (%)

Classes	Raichu	Pikachu	Charmander	Bulbasaur	Squirtle	Eevee	Piplup	Snorlax	Jigglypuff	Psyduck
Raichu	98.74	0.63	0	0	0.63	0	0	0	0	0
Pikachu	0.63	91.2	2.5	1.26	1.26	0	1.26	0.63	0.63	0.63
Charmander	0	6	84.1	2.65	1.2	0.7	2.65	1.2	0	0
Bulbasaur	0	1.02	1.02	86.2	8.7	1.02	2.04	0	0	0
Squirtle	0	0	0	1.4	95.8	2.1	0	0	0	0.7
Eevee	0	0	0	0	2.3	96.5	1.2	0	0	0
Piplup	0	0	0	0	0.71	1.42	90.07	7.8	0	0
Snorlax	0	0	0	0	1.9	0	6.6	91.5	0	0
Jigglypuff	0	0	0	0	0.6	0	0	0	99.4	0
Psyduck	0	0	0	0	0.56	0.56	0	0.56	0	98.3

5. CONCLUSIONS

The Pokémon character classification is useful for Pokémon lovers who are eager to know their character names. We developed our model using different pre-trained models for the classification of Pokémon characters. MobileNetV2, EfficientNetB7, EfficientNetV2L, DenseNet201, ResNet101, VGG19, and VGG16 Pre-trained models were used in this work and compared on their performance based on accuracy in classification. Based on the results it showed that ResNet101 will be the best model among the other transfer learning models taken in this work. Thus, this Transfer learning model can automate the classification of Pokémon characters by displaying its name. Future studies can focus on exploring the performance of transfer learning techniques on other types of visual data, such as videos or 3D images. Additionally, further investigations can be conducted to compare the performance of different transfer learning models on various image classification tasks beyond Pokémon characters.

5. FUTURE SCOPE

The future extent of this model includes deploying it into a mobile application that will be easy to use by children who love to know the characters of Pokémon, they can just scan the Pokémon character image with that app and can know its name. And also, it can further expand the dataset to include more Pokémon characters and images, as well as exploring the use of transfer learning techniques for other computer vision tasks beyond image classification. The findings from this study can serve as a foundation for further research in the field of transfer learning and its applications in computer vision.

REFERENCES

- Reddy, C.V.R., Reddy, U.S., Kishore, K.V.K. (2019).
 Facial emotion recognition using NLPCA and SVM. Traitement du Signal, 36(1): 13-22. https://doi.org/10.18280/ts.360102
- [2] VenkataRamiReddy, C., Kishore, K.K., Bhattacharyya, D., Kim, T.H. (2014). Multi-feature fusion-based facial expression classification using DLBP and DCT.

International Journal of Software Engineering and Its Applications, 8(9): 55-68. https://doi.org/10.14257/ijseia.2014.8.9.05

- [3] Ramireddy, C.V., Kishore, K.K. (2013). Facial expression classification using Kernel-based PCA with fused DCT and GWT features. In 2013 IEEE International Conference on Computational Intelligence and Computing Research, IEEE, pp. 1-6. https://doi.org/10.1109/ICCIC.2013.6724211
- [4] Reddy, C.V.R., Kishore, K.K., Reddy, U.S., Suneetha, M. (2016). Person identification system using feature level fusion of multi-biometrics. In 2016 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), IEEE, pp. 1-6. https://doi.org/10.1109/ICCIC.2016.7919672
- [5] Meena, K., Veni, N.K., Deepapriya, B.S., Vardhini, P.H., Kalyani, B.J.D., Sharmila, L. (2022). A novel method for prediction of skin disease through supervised classification techniques. Soft Computing, 26(19): 10527-10533. https://doi.org/10.1007/s00500-022-07435-8
- [6] Palakodati, S.S.S., Chirra, V.R.R., Yakobu, D., Bulla, S. (2020). Fresh and rotten fruits classification using CNN and transfer learning. Revue d'Intelligence Artificielle, 34(5): 617-622. https://doi.org/10.18280/ria.340512
- [7] Babu, Y.M.M. (2021). Optimized performance and utilization analysis of real-time multi spectral data/image categorization algorithms for computer vision applications. Turkish Journal of Computer and Mathematics Education (TURCOMAT), 12(9): 2212-2227. https://doi.org/10.17762/turcomat.v12i9.3696
- [8] Chirra, V.R.R., Uyyala, S.R., Kolli, V.K.K. (2021). Virtual facial expression recognition using deep CNN with ensemble learning. Journal of Ambient Intelligence and Humanized Computing, 12: 10581–10599. https://doi.org/10.1007/s12652-020-02866-3
- [9] Banothu, B., Murthy, T.S., Reddy, C.V.R., Yakobu, D. (2020). High-order total bounded variation approach for gaussian noise and blur removal. International Journal of Advanced Science and Technology, 29(3): 10152-10161.
- [10] Chirra, V.R.R., Uyyala, S.R., Kolli, V.K.K. (2019). Deep CNN: A machine learning approach for driver drowsiness detection based on eye state. Revue d'Intelligence Artificielle, 33(6): 461-466. https://doi.org/10.18280/ria.330609

- [11] Kumar, B.V., Srinivas, K.K., Anudeep, P., Yadav, N.S., Kumar, G.V., Vardhini, P.H. (2021). Artificial intelligence based algorithms for driver distraction detection: A review. In 2021 6th International Conference on Signal Processing, Computing and Control (ISPCC), IEEE, pp. 383-386.
- [12] Chirra, V.R., Maddiboyina, H.D., Dasari, Y., Aluru, R. (2020). Performance evaluation of email spam text classification using deep neural networks. Review of Computer Engineering Studies, 7(4): 91-95. https://doi.org/10.18280/rces.070403
- [13] Reddy, V.R., Yakobu, D., Prasad, S.S., Vardhini, P.H. (2022). Clustering student learners based on performance using K-Means algorithm. In 2022 International Mobile and Embedded Technology Conference (MECON), IEEE, pp. 302-306. https://doi.org/10.1109/MECON53876.2022.9752165
- [14] Xiang, Q., Wang, X.D., Li, R., Zhang, G.L., Lai, J., Hu, Q.S. (2019). Fruit image classification based on Mobilenetv2 with transfer learning technique. In Proceedings of the 3rd International Conference On Computer Science And Application Engineering, pp. 1-7. https://doi.org/10.1145/3331453.3361658
- [15] Zhang, Q., Yang, Q.F., Zhang, X.J., Bao, Q., Su, J.Q., Liu, X.Y. (2021). Waste image classification based on transfer learning and convolutional neural network. Waste Management, 135: 150-157. https://doi.org/10.1016/j.wasman.2021.08.038
- [16] Rajayogi, J.R., Manjunath, G., Shobha, G. (2019). Indian food image classification with transfer learning. In 2019 4th International Conference on Computational Systems and Information Technology for Sustainable Solution (CSITSS), IEEE, 4: 1-4. https://doi.org/10.1109/CSITSS47250.2019.9031051
- [17] Wang, G., Sun, Y., Wang, J.X. (2017). Automatic imagebased plant disease severity estimation using deep learning. Computational Intelligence and Neuroscience,

Article ID: https://doi.org/10.1155/2017/2917536

[18] Bansal, M., Kumar, M., Sachdeva, M., Mittal, A. (2021). Transfer learning for image classification using VGG19: Caltech-101 image data set. Journal of Ambient Intelligence and Humanized Computing, 14: 3609-3620. https://doi.org/10.1007/s12652-021-03488-z

2917536.

- [19] Ahmad, H.M., Ghuffar, S., Khurshid, K. (2019). Classification of breast cancer histology images using transfer learning. In 2019 16th International Bhurban Conference on Applied Sciences and Technology (IBCAST), IEEE, pp. 328-332. https://doi.org/10.1109/IBCAST.2019.8667221
- [20] Simonyan, K., Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition, arXiv: 1409. 1556. https://doi.org/10.48550/arXiv.1409.1556
- [21] Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., Chen, L.C. (2018). Mobilenetv2: Inverted residuals and linear bottlenecks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, IEEE, pp. 4510-4520.
- [22] Tan, M., Le, Q. (2019). Efficientnet: Rethinking model scaling for convolutional neural networks. In International Conference on Machine Learning, PMLR, 97: 6105-6114. https://doi.org/10.48550/arXiv.1905.11946
- [23] Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K.Q. (2017). Densely connected convolutional networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, IEEE, pp. 4700-4708.
- [24] He, K., Zhang, X., Ren, S., Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, IEEE, pp. 770-778.
- [25] Tan, M., Le, Q. (2021). Efficientnetv2: Smaller models and faster training. In International Conference on Machine Learning, PMLR, pp. 10096-10106.