Medical Image Classification Using Transfer Learning and Network Pruning Algorithms

Luca Saleh
Department of Computer Science
Royal Holloway, University of London
Surrey, UK
Luca.Saleh.2018@live.rhul.ac.uk

Li Zhang
Department of Computer Science
Royal Holloway, University of London
Surrey, UK
li.zhang@rhul.ac.uk

***Abstract*—Deep neural networks show great advancement in recent decades in classifying medical images (such as CT-scans) with high precision to aid disease diagnosis. However, the training of deep neural networks requires significant sample sizes for learning enriched discriminative spatial features. Building a high quality dataset large enough to satisfy model training requirement is a challenging task due to limited disease sample cases, and various data privacy constraints. Therefore in this research, we perform medical image classification using transfer learning based on several well-known deep networks, i.e. GoogLeNet, Resnet and EfficientNet. To tackle data sparsity issues, a Wasserstein Generative Adversarial Network (WGAN) is used to generate new medical image samples to increase the numbers of training instances of the minority classes. The transfer learning process itself also allows the building of strong classifiers by transferring knowledge from the pre-trained image domain to a new medical domain using a small sample size. Moreover, the lottery ticket hypothesis is also used to prune each transfer learning network trained using the new target image data sets. Specifically, the L1 norm unstructured pruning technique is used for network reduction. Hyper-parameter fine-tuning is also performed to identify optimal settings of key network hyper-parameters such as learning rate, batch size and weight decay. A total of 20 trials are used for optimal hyper-parameter selection. Evaluated using multi-class lung X-ray images for pneumonia conditions and brain tumor CT-scans, the fine-tuned EfficientNet model obtains the best brain tumor classification accuracy rate of 96% and a fine-tuned GoogLeNet model with pruning has the highest pneumonia classification accuracy rate of 81.5%.**

***Index Terms*—Neural Networks, Medical Imaging, Transfer Learning, Network Pruning, Computer Aided Diagnosis**

I. Introduction

This research studies a range of distinctive neural network architectures and makes attempt to apply a variety of modifications to the network structures as well as training configurations. Traditional classifiers for computer aided diagnosis rely on a large number of training samples to achieve reliable and accurate predictions, e.g. with respect to brain tumor or lung disease classification [1] [2] [3] [4] [5]. In this research, we aim to maximise model performance and generalisation capabilities in a domain with limited training data. In this regard, transfer learning for computer vision has shown to be very advantageous in the medical domain, producing effective classifiers of through re-training the networks with various different types of medical data [6] [7] [8] [9]. Such a learning process requires a small number of new samples in the new target domain, which significantly increases computational efficiency while attaining reasonable classification performance [10].

In this work, two core transfer learning techniques will be explored, i.e. feature extraction (re-training only the last layer) and fine-tuning (re-training the overall network based on the pre-trained weights) [11] [12] [13] [14] [15]. The research in [16] details feature extraction to be more useful when less data are available and the problem the network was originally trained for is similar to the new target domain. On the other hand, fine-tuning is required when the original and target problems show large variations [17] [18] [19].

Models that are used for transfer learning will typically have weights that were trained on the IMAGENET1K dataset, allowing high-level image features the networks have learnt to be transferred and applied for classification in a new and perhaps very different domain [20] [21] [22] [23].

In this research, we apply transfer learning to GoogLeNet, Resnet-18 and EfficientNet architectures. In order to identify the most optimal training optimiser effectively, Adam, RProp and SGD, are used for model training. Hyper-parameter fine-tuning is also conducted to identify the optimal batch size, learning and weight decay (i.e. L2 regularisation) settings [16] [24] [25]. Additionally, the pruning algorithm, i.e. the lottery ticket hypothesis [26], is also applied for network architecture reduction, in an attempt to improve their generalisation capabilities and reduce prediction loss.

Once a strong classifier has been produced, oversampling techniques e.g. Generative Adversarial Networks (GANs), are explored to generate new images to increase sample sizes of minority classes. Specifically, the Wasserstein GAN with Gradient Penalty (WGAN-GP) architecture is used for image generation. Comprehensive experiments have been carried out in order to identify if generalisation test accuracy can be improved by training the networks with generated samples. The use of WGAN-GP allows effective loss minimisation of the generator and critic networks whilst improving stability of network training. It is less likely to result in mode collapse as the case of the original GAN model [27]. The results in [27] also show high quality eye sample images of sufficient

variations between healthy or strabismus conditions (where the eyes are not aligned) were generated using WGAN. Variation in generated images is also desirable in our research, otherwise the generated images with high resemblance may cause overfitting during training.

Two datasets in the medical domain are used in our study to test transfer learning, hyper-parameter tuning and network pruning methods. The first dataset is a collection of brain CT scans representing four classes, i.e. no tumor, glioma tumor, meningioma tumor and pituitary tumor. The scans in this dataset are all from different angles of the brain, thus finding very tumor specific features is a key for a neural network to correctly classify a given sample. Varying angles as shown in Figure 1 are used for CT-scan image capture to better observe patients' conditions.

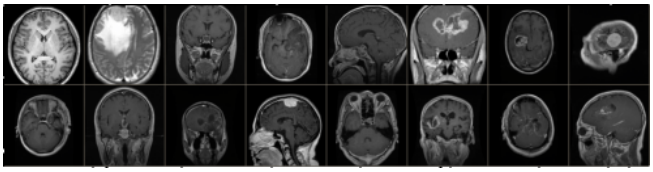


Fig. 1. Collection of 16 CT-scan samples.

The second dataset is a collection of chest X-rays of patients who are either healthy or with pneumonia disease conditions. It consists of three classes, i.e. normal, viral pneumonia and bacterial pneumonia. All X-ray images in this dataset are taken from very similar if not the same angle to a patient. With all the images being taken from the front of the patient in laying down positions, with slight variances in the angle.

II. Related Work

Oversampling with GANs is conducted when real-life disease cases or examples of minority classes are difficult or costly to obtain. Many GANs and their variants are used for image generation. E.g., [28] uses a WGAN model to generate URLs for potentially malicious phishing sites. Naturally there will be far more available examples of benign URLs, so the objective is to use a text GAN to generate the required number of malicious URLs to balance the classes in the dataset before it is used to train a neural network classifier for malicious URL classification.

Class imbalance is an issue that exists in any domain of decision making. The work of [29] discusses addressing data imbalance issues itself, rather than designing an oversampling method to balance the class distribution. Comparing oversampling methods such as SMOTE and WGAN. It was found that GAN-based oversampling produced better performance with improved recall, F1-score and AUC results.

The study of [30] employs oversampling methods to assist fine-grained classification of lung nodules. Between GAN, DCGAN, WGAN as well as random rotation of a sample from 0 to 359 degrees, oversampling through WGAN image generation was found to balance the dataset with the best samples, resulting in improved CNN classifier performance.

Pneumonia classification using Convolutional Neural Networks (CNNs) with X-ray scans has been previously explored by [31]. A total of 5863 data samples were used to train four different models to classify healthy, viral pneumonia and bacterial pneumonia, as well as a binary classifier for healthy and pneumonia. Each model had different augmentations applied to the datasets beforehand. The augmentations including rotating the images and increasing image contrast. Their work achieved a 3-class accuracy rate of 85% and a binary classification result of 90.5%.

Similarly, [32] focuses on classification in the medical domain. A binary classification is performed using MRI scans of the brains to identify brain tumor or healthy cases. Transfer learning is used based on pre-trained ResNet, Xception and MobilNet-V2 networks. These models were pre-trained on the IMAGENET1K dataset. A total of 928 tumor samples and 588 tumor-free samples are used for training. Their results indicated that MobileNet-V2 achieves the highest accuracy rate of 98.25% and a F1 score of 98.42%.

The study in [33] shows the use of Resnet50 and VGG16 as binary classifiers for predicting healthy or pneumonia cases from CT-scan images. The models are trained with 5216 original images over 20 epochs. The VGG16 model obtained the best accuracy rate of 85.58% with Resnet50 having an accuracy rate of 82.21%, when tested with 624 images.

III. The Proposed Methodology

The aim of this work is to produce an optimal classifier with limited training samples. To achieve this, each network will be trained with each of the chosen optimisers. The best network and optimiser will be used for transfer learning tasks where hyper-parameters are fine-tuned. The lottery ticket hypothesis is then applied to the resulting transfer learning model for network pruning. A total of 25 epochs will be used as the standard number of training epochs, with training over reduced epochs also explored to reduce network overfitting.

For the brain tumor dataset [34], 640, 160 and 200 images are used for training, validation and test respectively. For the chest X-ray [35] images, there are 640, 160, 200 samples provided for training, validation and test respectively.

In the case of pneumonia classification, to evaluate the benefits of image generation using WGAN, fine-tuned transfer learning model will be trained on additional WGAN-generated samples.

A. Oversampling with GANs

This research uses a conditional WGAN-GP architecture for image generation, where its training method is provided in [36]. The work of [36] builds on the model presented in the original WGAN paper [37], using a gradient penalty driven through an interpolation of generated and real images. This gradient penalty enforces the Lipschitz constraint as shown in Equation (1) more effectively which governs the image generation process in WGAN.

$$\max_{||\mathbf{f}||_{\mathbf{L}}\leq\mathbf{1}} \mathbb{E}_{x\sim p_r}[f(x)] - \mathbb{E}_{x\sim p_\theta}[f(x)] \tag{1}$$

The WGAN-GP generated images also possess sufficient variations which is the key to avoid overfitting. The WGAN

along with gradient penalty can also effectively prevent mode collapse. Therefore in the case of chest X-ray image generation, our generated X-rays appear to be from slightly different angles. The rib cages and body shapes were also noticeably different.

Oversampling is also performed for the pneumonia dataset. Since these samples are from the same angle, WGAN shows more impressive performance for image generation.

B. Data Pre-Processing

All data used for training, validation and testing requires pre-processing before they are used for network training and test. Each image is resized to have the smaller side of a length of 225, with the other side resized to maintain the original aspect ratio. The image is then cropped around the center to produce an image of size 224x224. The cropped image is then normalised so that each colour channel has a mean of 0.485, 0.456 and 0.406, and a standard deviation of 0.229, 0.224 and 0.225, respectively. Each value represents the red, blue and green channels respectively. Any images produced by a GAN are pre-processed in this way before training.

The above color channel values are adopted as they match the pre-processing steps used on IMAGENET1K samples, which the models used for transfer learning were originally trained on. It is important that the normalisation pre-processing steps are matched as otherwise the networks will fail to successfully extract features from given samples.

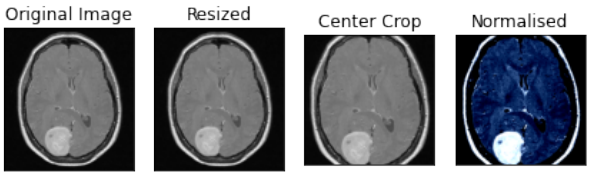


Fig. 2. Brain CT-scan pre-processing steps.

It is clear from Figure 2 how these normalisation steps enhance key features of the brain scan image.

C. The Proposed Models

We perform transfer learning using the following networks, i.e. Resnet18, GoogLeNet and EfficientNet. Using these pre-trained models as feature extractors means that only the final classifying layer will be re-trained using the new dataset. The gradients of all layers of these networks will be frozen apart from the final layer, which will be initialised with random weights and set to produce the correct number of outputs for how many classes it is trained with.

The process of fine tuning described by [16] is to fine-tune the whole entire network with a small learning rate as to not change the original weight values of the network too much. This will be applied by resetting the final classifying layer of the network as described before, but not freezing any layers of the network, so all weights will be adjusted by the chosen optimiser.

When training the models for the transfer learning baseline experiments, the hyper-parameters used were as follows in Table I. Note that no regularisation is used in the baseline experiments.

TABLE I
HYPER-PARAMETERS USED WHEN TRAINING MODELS FOR THE BASELINE EXPERIMENTS.

Learning rate	0.0001
Batch size	4
Epochs	25
Optimisers	Adam, SGD, RProp
Loss function	Cross-entropy loss

D. Lottery Ticket Hypothesis

The lottery ticket hypothesis is a network architecture pruning technique [26]. When applying the lottery ticket hypothesis, pruning will be applied to model weights to remove 50% of the weights using L1 norm unstructured pruning. This pruning will be applied to all layers of the target network once it has been trained on the new target domain. The pruned weights are re-initialised to the original IMAGENET1K weights that the network held before being trained on the net target domain, and then the model is trained again with the goal to increase generalisation capability and reduce the loss in the target domain.

E. Hyper-Parameter Tuning

The hyper-parameters that will be optimised include the learning rate, batch size and weight decay (i.e. L2 regularisation). The range of values that will be explored are shown in Table II.

TABLE II
EXPLORED HYPER-PARAMETER RANGES.

Learning Rate	Log-Uniform range 1e-4 to 1e-1
Batch Size	16, 32, 64, 128
Weight Decay	Choice of 0 or Log-Uniform range 1e-8 to 1e-1

20 trials are performed where parameters are randomly sampled from this variable space.

F. Sample Generation with GANs

To generate high quality chest X-ray images, a conditional WGAN-GP was used. This network was trained over 100 epochs using the Adam optimiser with the detailed parameters shown in Table III.

Using the conditional WGAN-GP model, 192 bacterial pneumonia, 192 healthy and 192 viral pneumonia samples are generated. Therefore a total of 576 samples are generated, which are then divided into a 80-20 training-validation split, with 461 training and 115 validation images.

IV. EVALUATION

A. Classification of brain tumors through CT scan images

1) GoogLeNet: The transfer learning using GoogLeNet shows great efficiency. But there are also cases that the network is prone to over-fitting. Pruning the GoogLeNet transfer learning model showed a decrease in training loss

TABLE III
HYPER-PARAMETERS USED TO TRAIN THE WGAN-GP.

Learning rate	0.0001
Batch Size	64
Latent Noise size	100
Label embedding size	100
Generator Latent Features	128
Critic Latent Features	128
Epochs	100
Optimiser	Adam
Optimiser Betas	$(0, 0.9)$
Critic iterations	5
Gradient Penalty Lambda	10

from 0.223207 to 0.187564 and an increase in testing accuracy from 91% to 94%. The identified optimal hyper-parameters are also shown in Table IV. Figure 3 shows performance of GoogLeNet with Adam optimiser and fine-tuned parameters, which was the best performer.

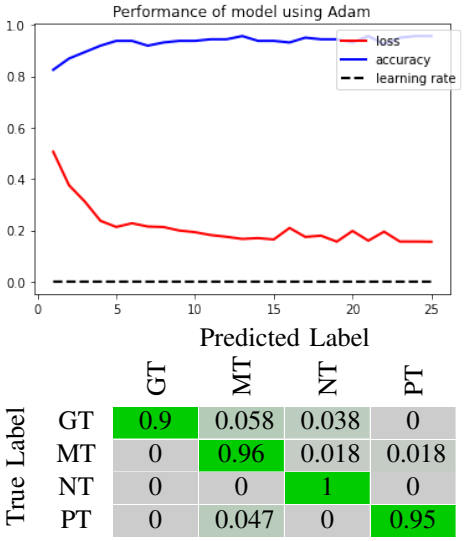


Fig. 3. Validation and testing performance of GoogLeNet using Adam and fine-tuned parameters for tumor classification, where GT is Glioma Tumor, MT is Meningioma Tumor, NT is No Tumor and PT is Pituitary Tumor.

TABLE IV
OPTIMAL HYPER-PARAMETERS FOUND FOR THE FINE-TUNED GOOGLENET.

Learning rate	0.0009747398181460825
Batch Size	128
Weight Decay	1.2076289887829432e-07

2) Resnet18: Optimal hyper-parameter identification is also performed for Resnet18, with detailed settings illustrated in Table V. This optimized model produced a testing accuracy rate of 88%. The lottery ticket pruning method caused a validation loss being larger than that of the original model, with it increasing from 0.314897 to 0.361918. Figure 4 shows performance of the optimized Resnet18 model with Adam optimizer.

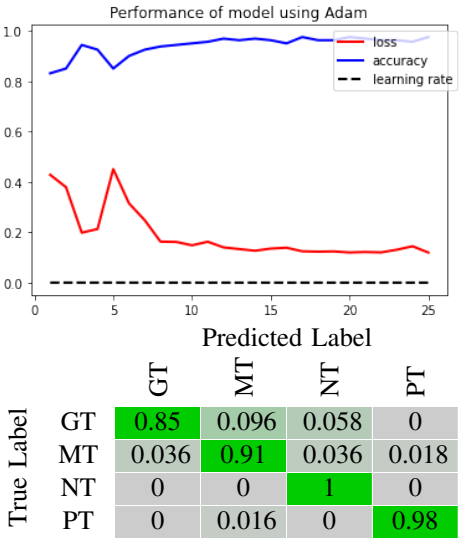


Fig. 4. Validation and testing performance of Resnet18 using Adam and fine-tuned parameters for tumor classification.

TABLE V
OPTIMAL HYPER-PARAMETERS FOUND FOR THE FINE TUNED RESNET18.

Learning rate	0.0001061193920871678
Batch Size	64
Weight Decay	1.6475858424935896e-06

3) EfficientNet: EfficientNet was only comparatively more effective with the Adam optimiser and fine-tuned parameters, with the detailed performance displayed in Figure 5.

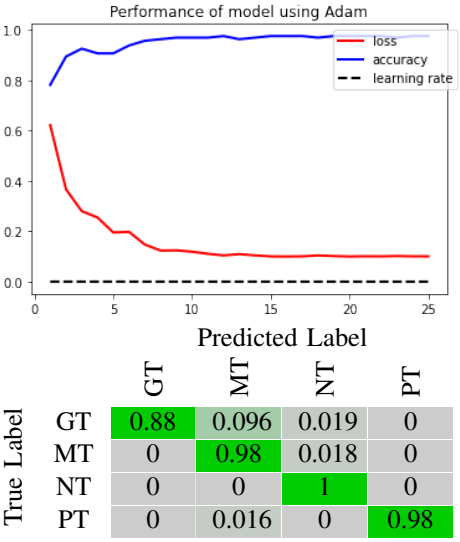


Fig. 5. Validation and testing performance of EfficientNet using Adam and fine-tuned parameters for tumor classification.

B. Pneumonia classification through chest X-rays

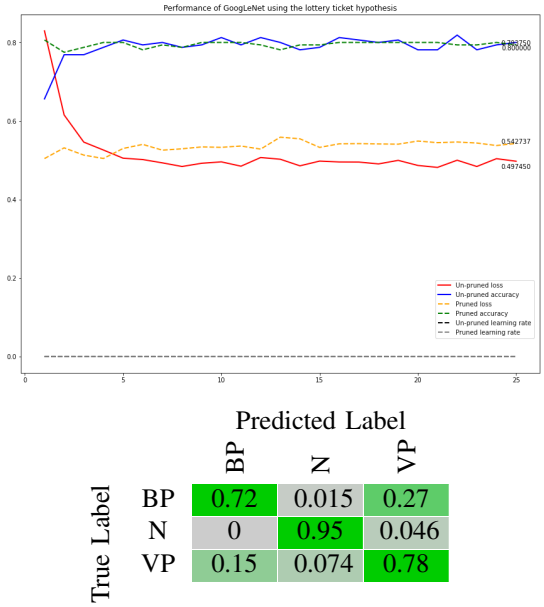
1) GoogLeNet: The transfer learning process using GoogLeNet was comparatively more effective for the chest X-ray dataset for lung condition classification, with the Adam and RProp optimisers producing the test accuracy rates of 77% and 77.5%, respectively.

Using Adam with fine-tuned hyper-parameters showed much stable training. The optimal parameters identified are provided in Table VI.

TABLE VI
OPTIMAL HYPER-PARAMETERS FOUND FOR THE GOOGLENET MODEL.

Learning rate	0.0002983176352303694
Weight Decay	0.0008982934181184593
Batch Size	16

The parameters in Table VI achieved a validation accuracy rate of 81.88% and a testing accuracy rate of 80%. A decrease in loss in the baseline was observed from 0.76 to 0.536. Pruning for this dataset was slightly effective, with the test score reaching 81.5%. Figure 6 shows the results of the pruned GoogLeNet with optimal hyper-parameters.



Predicted Label

True Label	BP	N	VP
BP	0.72	0.015	0.27
N	0	0.95	0.046
VP	0.15	0.074	0.78

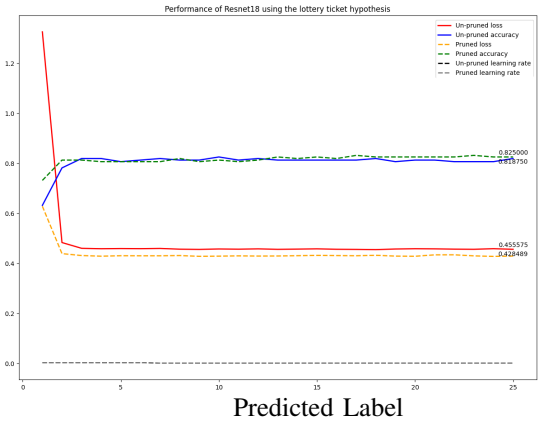
Fig. 6. Pruned GoogLeNet performance with the optimal hyper-parameters found, where the confusion matrix shows classification accuracy for bacterial pneumonia as BP, normal as N and viral pneumonia as VP.

2) Resnet18: We use both Adam and Rprop in hyper-parameter tuning for Resnet18. The identified parameters are provided in Table VII. The experiments indicate that using the Rprop optimiser obtains the highest test accuracy. Pruning these models was more effective for the RProp model, with the test score rising to 75.5%. The validation loss was reduced from 0.455575 to 0.428489.

In contrast, the pruning had little to no effect on the model produced with Adam, with test accuracy remaining at 0.78 and changes to validation loss and accuracy being negligible. Figure 7 shows the performance comparison between the fine-tuned and pruned Resnet18 models, both with optimized hyper-parameters.

TABLE VII
OPTIMAL HYPER-PARAMETERS FOUND FOR THE RESNET18 MODEL.

	Adam	RProp
Learning rate	0.0024808160755579877	0.00200901124909481
Weight Decay	0.00010424329777131112	x
Batch Size	64	32

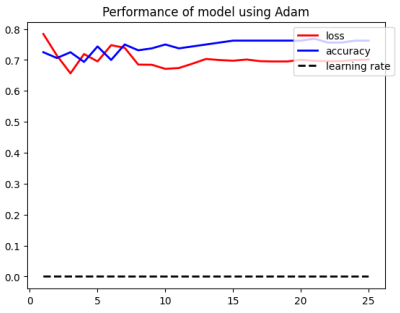


Predicted Label

True Label	BP	N	VP
BP	0.77	0.082	0.15
N	0.018	0.93	0.055
VP	0.29	0.097	0.61

Fig. 7. Validation performance of the Resnet18 model with identified optimal parameters compared to those of the same model with pruning applied. Testing performance of the pruned model presented in the confusion matrix.

3) EfficientNet: Fine tuning using the EfficientNet model with Adam optimizer showed impressive results on its own, with detailed performance presented in Figure 8.



Predicted Label

True Label	BP	N	VP
BP	0.74	0.028	0.24
N	0	0.91	0.09
VP	0.23	0.049	0.72

Fig. 8. Performance of the EfficientNet model with parameter fine-tuning and Adam optimizer for pneumonia classification.

Overall, all the transfer learning models show reasonable performance for both datasets with the detailed results shown in Tables VIII and IX. The best results for each model are highlighted using bold text.

4) Training with GAN generated images: In transfer learning using a pre-trained GoogLeNet for 3-class lung condition identification, the model was trained over 10 epochs with 1000 real data samples. This achieved a best validation accuracy rate of 81.25% and a test accuracy rate of 71.5%. A final loss of 0.545 was also obtained.

TABLE VIII
OVERVIEW OF BRAIN TUMOR CLASSIFICATION TEST ACCURACY RESULTS.

Model	Method	Adam	SGD	RProp	Tuning	LTH
GoogLeNet	Fine Tune (last layer)	69%	38%	39%	-	-
	Fine Tune (all layers)	**95**%	54%	77.5%	91%	94%
Resnet18	Fine Tune (last layer)	66.5%	44.5%	34%	-	-
	Fine Tune (all layers)	**93**%	67.5%	86.5%	88%	-
EfficientNet	Fine Tune (last layer)	74%	40.5%	55%	-	-
	Fine Tune (all layers)	**96**%	50.5%	82%	-	-

TABLE IX
OVERVIEW OF PNEUMONIA CLASSIFICATION TEST ACCURACY RESULTS.

Model	Method	Adam	SGD	RProp	Tuning w/ RProp	Tuning w/ Adam	LTH w/ Rprop	LTH w/ Adam
GoogLeNet	Fine Tune (last layer)	70.5%	45%	45.5%	-	-	-	-
	Fine Tune (all layers)	77%	63.5%	77.5%	-	80%	-	**81.5%**
Resnet18	Fine Tune (last layer)	69%	62.5%	53%	-	-	-	-
	Fine Tune (all layers)	72.5%	72.5%	75%	77.5%	**78**%	77.5%	**78**%
EfficientNet	Fine Tune (last layer)	76%	58.5%	62.5%	-	-	-	-
	Fine Tune (all layers)	**79**%	56.5%	77%	-	-	-	-

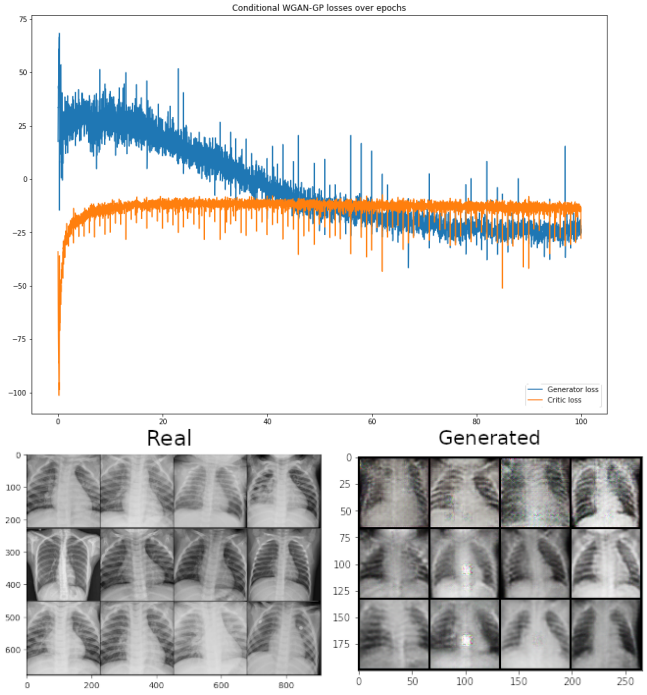


Fig. 9. WGAN-GP training loss and samples generated by the conditional WGAN-GP, with bacterial pneumonia (top), normal (middle) and viral pneumonia (bottom).

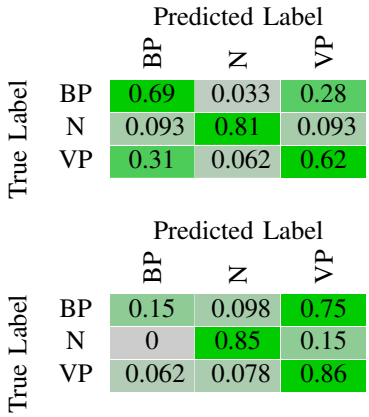


Fig. 10. Comparison of classifiers trained on real data only (top) and combination of real and generated data (bottom).

Then training this classifier with generated data, the results showed over-fitting, with a validation accuracy of 97%, final loss of 0.1146 but a testing accuracy of 51.5% for the 3-class classification task.

However, the empirical results in Figure 10 indicated a success in model improvement as a binary classifier, with a very large increase in viral pneumonia classification accuracy from 62% to 86%, and a slight increase in normal/healthy classification from 81% to 85%. Figure 9 shows the training loss comparison of the generator and critic, as well as generated example images using WGAN. Figure 10 shows the detailed confusion matrix result comparison.

V. CONCLUSION

In this research, we explore transfer learning, network pruning and hyper-parameter fine-tuning. In addition, WGAN is used for image generation to tackle data sparsity issues. As indicated in Tables VIII and IX, our transfer learning and some pruned networks achieved much improved performance. The empirical results indicate the effectiveness of the transfer learning methods and pruning algorithms. Adjusting all the weights of the entire network to fit the new target domain is comparatively more effective, than simply re-training the last network layer as in traditional transfer learning. In addition, pushing the models further using the network pruning and hyper-parameter fine-tuning results in increases in generalisation accuracy.

The use of GAN-generated images is proven to be effective at improving model's generalisation capabilities, but distinguishing more complex class cases requires the WGAN-GP to be more optimised, as multi-class classification was more challenging [25] [12] [38] [15], with the features of

the different pneumonia types being more complex than what the WGAN-GP model learnt. Future work would involve improving the performance of the WGAN-GP model, e.g. by training it over a greater number of epochs to generate higher quality X-ray samples that allow better differentiation between viral and bacterial pneumonia cases.

REFERENCES

[1] X. Luo, H. Wu, Z. Wang, J. Wang, and D. Meng, "A novel approach to large-scale dynamically weighted directed network representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 12, pp. 9756–9773, 2021.

[2] C. Wall, L. Zhang, Y. Yu, and K. Mistry, "Deep recurrent neural networks with attention mechanisms for respiratory anomaly classification," in *2021 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–8, IEEE, 2021.

[3] L. Zhang, C. P. Lim, Y. Yu, and M. Jiang, "Sound classification using evolving ensemble models and particle swarm optimization," *Applied soft computing*, vol. 116, p. 108322, 2022.

[4] C. Wall, L. Zhang, Y. Yu, A. Kumar, and R. Gao, "A deep ensemble neural network with attention mechanisms for lung abnormality classification using audio inputs," *Sensors*, vol. 22, no. 15, p. 5566, 2022.

[5] L. Zhang, K. Mistry, C. P. Lim, and S. C. Neoh, "Feature selection using firefly optimization for classification and regression models," *Decision Support Systems*, vol. 106, pp. 64–85, 2018.

[6] B. Sowan, M. Eshtay, K. Dahal, H. Qattous, and L. Zhang, "Hybrid pso feature selection-based association classification approach for breast cancer detection," *Neural Computing and Applications*, vol. 35, no. 7, pp. 5291–5317, 2023.

[7] T. Y. Tan, L. Zhang, and C. P. Lim, "Adaptive melanoma diagnosis using evolving clustering, ensemble and deep neural networks," *Knowledge-Based Systems*, vol. 187, p. 104807, 2020.

[8] T. Y. Tan, L. Zhang, S. C. Neoh, and C. P. Lim, "Intelligent skin cancer detection using enhanced particle swarm optimization," *Knowledge-based systems*, vol. 158, pp. 118–135, 2018.

[9] T. Y. Tan, L. Zhang, C. P. Lim, B. Fielding, Y. Yu, and E. Anderson, "Evolving ensemble models for image segmentation using enhanced particle swarm optimization," *IEEE access*, vol. 7, pp. 34004–34019, 2019.

[10] T. Lawrence, L. Zhang, K. Rogage, and C. P. Lim, "Evolving deep architecture generation with residual connections for image classification using particle swarm optimization," *sensors*, vol. 21, no. 23, p. 7936, 2021.

[11] H. Xie, L. Zhang, C. P. Lim, Y. Yu, and H. Liu, "Feature selection using enhanced particle swarm optimisation for classification models," *Sensors*, vol. 21, no. 5, p. 1816, 2021.

[12] S. Slade, L. Zhang, H. Huang, H. Asadi, C. P. Lim, Y. Yu, D. Zhao, H. Lin, and R. Gao, "Neural inference search for multiloss segmentation models," *IEEE Transactions on Neural Networks and Learning Systems*, vol. In Press, 2023.

[13] K. Mistry, L. Zhang, S. C. Neoh, C. P. Lim, and B. Fielding, "A micro-ga embedded pso feature selection approach to intelligent facial emotion recognition," *IEEE transactions on cybernetics*, vol. 47, no. 6, pp. 1496–1509, 2016.

[14] K. Wang, "An overview of deep learning based small sample medical imaging classification," in *2021 International Conference on Signal Processing and Machine Learning (CONF-SPML)*, pp. 278–281, 2021.

[15] P. Kinghorn, L. Zhang, and L. Shao, "A hierarchical and regional deep learning architecture for image description generation," *Pattern Recognition Letters*, vol. 119, pp. 77–85, 2019.

[16] S. Slade, L. Zhang, Y. Yu, and C. P. Lim, "An evolving ensemble model of multi-stream convolutional neural networks for human action recognition in still images," *Neural computing and applications*, vol. 34, no. 11, pp. 9205–9231, 2022.

[17] L. Zhang, C. P. Lim, and Y. Yu, "Intelligent human action recognition using an ensemble model of evolving deep networks with swarm-based optimization," *Knowledge-based systems*, vol. 220, p. 106918, 2021.

[18] T. Y. Tan, L. Zhang, and C. P. Lim, "Intelligent skin cancer diagnosis using improved particle swarm optimization and deep learning models," *Applied Soft Computing*, vol. 84, p. 105725, 2019.

[19] S. C. Neoh, L. Zhang, K. Mistry, M. A. Hossain, C. P. Lim, N. Aslam, and P. Kinghorn, "Intelligent facial emotion recognition using a layered encoding cascade optimization model," *Applied soft computing*, vol. 34, pp. 72–93, 2015.

[20] H. Xie, L. Zhang, C. P. Lim, Y. Yu, C. Liu, H. Liu, and J. Walters, "Improving k-means clustering with enhanced firefly algorithms," *Applied Soft Computing*, vol. 84, p. 105763, 2019.

[21] B. Fielding and L. Zhang, "Evolving image classification architectures with enhanced particle swarm optimisation," *IEEE access*, vol. 6, pp. 68560–68575, 2018.

[22] L. Zhang and C. P. Lim, "Intelligent optic disc segmentation using improved particle swarm optimization and evolving ensemble models," *Applied Soft Computing*, vol. 92, p. 106328, 2020.

[23] L. Zhang, C. P. Lim, and C. Liu, "Enhanced bare-bones particle swarm optimization based evolving deep neural networks," *Expert Systems with Applications*, p. 120642, 2023.

[24] H. Xie, L. Zhang, and C. P. Lim, "Evolving cnn-lstm models for time series prediction using enhanced grey wolf optimizer," *IEEE access*, vol. 8, pp. 161519–161541, 2020.

[25] W. Lu, D. Zhao, C. Premebida, L. Zhang, W. Zhao, and D. Tian, "Improving 3d vulnerable road user detection with point augmentation," *IEEE Transactions on Intelligent Vehicles*, 2023.

[26] J. Frankle and M. Carbin, "The lottery ticket hypothesis: Finding sparse, trainable neural networks," 2018.

[27] D. Kim, J. Joo, and S. C. Kim, "Fake data generation for medical image augmentation using gans," in *2022 International Conference on Artificial Intelligence in Information and Communication (ICAIIC)*, pp. 197–199, 2022.

[28] A. Anand, K. Gorde, J. R. Antony Moniz, N. Park, T. Chakraborty, and B.-T. Chu, "Phishing url detection with oversampling based on text generative adversarial networks," in *2018 IEEE International Conference on Big Data (Big Data)*, pp. 1168–1177, 2018.

[29] N. M. Nafi and W. H. Hsu, "Addressing class imbalance in image-based plant disease detection: Deep generative vs. sampling-based approaches," in *2020 International Conference on Systems, Signals and Image Processing (IWSSIP)*, pp. 243–248, 2020.

[30] Q. Wang, X. Zhou, C. Wang, Z. Liu, J. Huang, Y. Zhou, C. Li, H. Zhuang, and J.-Z. Cheng, "Wgan-based synthetic minority over-sampling technique: Improving semantic fine-grained classification for lung nodules in ct images," *IEEE Access*, vol. 7, pp. 18450–18463, 2019.

[31] J. Garstka and M. Strzelecki, "Pneumonia detection in x-ray chest images based on convolutional neural networks and data augmentation methods," in *2020 Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA)*, pp. 18–23, 2020.

[32] M. Arbane, R. Benlamri, Y. Brik, and M. Djerioui, "Transfer learning for automatic brain tumor classification using mri images," in *2020 2nd International Workshop on Human-Centric Smart Environments for Health and Well-being (IHSH)*, pp. 210–214, 2021.

[33] Y. Yang, G. Mei, and F. Piccialli, "Explainable deep learning models on the diagnosis of pneumonia," in *2021 IEEE/ACM Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE)*, pp. 134–138, 2021.

[34] S. Bhuvaji, A. Kadam, P. Bhumkar, S. Dedge, and S. Kanchan, "Brain tumor classification (mri)," 2020.

[35] D. Kermany, K. Zhang, and M. Goldbaum, "Labeled optical coherence tomography (oct) and chest x-ray images for classification," 2018.

[36] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. Courville, "Improved training of wasserstein gans," 2017.

[37] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein gan," 2017.

[38] P. Kinghorn, L. Zhang, and L. Shao, "A region-based image caption generator with refined descriptions," *Neurocomputing*, vol. 272, pp. 416–424, 2018.