Improved Medical Image Classification Accuracy on Heterogeneous and Imbalanced Data using Multiple Streams Network

Mumtaz Ali[1]

School of Computer Science,
Huazhong University of Science and Technology,
Wuhan & Department of Computer Systems Engineering,
Sukkur IBA University, Pakistan

Riaz Ali[2]

Department of Computer Science
Sukkur IBA University
Pakistan

Nazim Hussain[3]

Department of Mathematics
Sukkur IBA University
Pakistan

Abstract—Small and massively imbalanced datasets are long-standing problems on medical image classification. Traditionally, researchers use pre-trained models to solve these problems, however, pre-trained models typically have a huge number of trainable parameters. Small datasets are challenging for them to train a model adequately and imbalanced datasets easily lead to overfitting on the classes with more samples. Multiple-stream networks that learn a variety of features have recently gained popularity. Therefore, in this work, a quad-stream hybrid model called QuadSNet using conventional as well as separable convolutional neural networks is proposed to achieve better performance on small and imbalanced datasets without using any pre-trained model. The designed model extracts hybrid features and the fusion of such features makes the model more robust on heterogeneous data. Besides, a weighted margin loss is used to handle the problem of class imbalance. The QuadSNet is trained and tested on seven different classification datasets. To evaluate the advantages of QuadSNet on small and massively imbalanced data, it is compared with six state-of-the-art pre-trained models on three benchmark datasets based on Pneumonia, COVID-19, and Cancer classification. To assess the performance of QuadSNet on general classification datasets, it is compareed with the best model on each of the remaining four datasets, which contain larger, balanced, grayscale, color or non-medical image data. The results show that QuadSNet handles the class imbalance and overfitting better than existing pre-trained models with much fewer parameters on small datasets. Meanwhile, QuadSNet has competitive performance in general datasets.

Keywords—*Medical image classification; convolutional neural networks; class imbalance; small dataset; margin loss*

I. Introduction

Typically, a huge amount of data is needed to train the neural networks for natural and medical image classification. However, along with the scarcity of sufficient samples, generally, the medical image datasets are massively imbalanced and they possess very limited positive cases. Thus, to obtain high-performance results by incorporating small and imbalanced datasets is a very perplexing task. In recent years, researchers have tried various algorithm-level and data-level methodologies to handle such challenges.

The algorithm level approaches have evolved since the reemergence of deep learning. Transfer learning [5], few-shot learning [18], zero-shot learning [6], Siamese networks [8],

network ensembles [4] and most recent algorithms based on generative adversarial networks [2] have been applied to small and imbalanced datasets. The algorithm level approaches commonly rely on pre-trained models. Despite decent signs of progress, the pre-trained models orthodoxly use millions of parameters and complex architecture to achieve competitive results, which appear to be enormous for a small dataset to train a good model.

Along with the issue of a small dataset, the class imbalance is another challenge. Overmuch parameters of pre-trained models typically lead to overfitting if the classes are imbalanced. Weighting class labels is a very prevalent technique in handling class imbalance. The weighted class label approach produces better results when the model is trained with a suitable loss function and optimization method.

The data level approaches are mainly based on data augmentation, oversampling and undersampling. In medical image processing, conventional data augmentation techniques are sometimes problematic. For instance, a chest X-ray image traditionally depicts the heart at the lower left side, but when the image is augmented with a mirror effect, it will depict the heart at the right side of the chest. Similarly, effects of filliping, zooming, or rotation may change the meaning of a medical image completely. Hence, efficiently designed deep learning models and algorithms may be a superior option for medical image processing.

However, despite good performance, data and algorithm level approaches tend to use very complex models with a large number of parameters. According to [3], recent hybrid models using multiple streams and feature fusion have produced competitive results for small and highly imbalanced datasets. In this work, the algorithm level approaches are followed and a quad-stream hybrid model, called QuadSNet, with separable as well as conventional convolutional neural networks is proposed to classify medical images on small and imbalanced datasets.

The performance of QuadSNet is evaluated together with the state-of-the-art pre-trained models, including DenseNet121 (DN) [12], InceptionV3 (IN) [25], MobileNet (MN) [11], ResNet50 (RN) [9], VGG16 (VG) [24] and Xception (XC) [7]. The proposed model QuadSnet is trained and tested on six medical image datasets and one famous MNIST handwritten digits dataset. QuadSNet is

intended to handle the problem of class imbalance with weighted margin loss on small datasets. Consequently, out of these seven datasets, the three smallest and most imbalanced medical datasets are used as benchmarks to compare the performance of QuadSNet with pre-trained models. For comprehensiveness, two of the three benchmark datasets are grayscale and one is colored. Other than the three benchmark datasets, the remaining four datasets are used to analyze the performance of QuadSNet on the larger, balanced, grayscale, color or non-medical image data.

The proposed technique could well handle the challenges of small datasets and class imbalance using a less complex model. Competitive results are obtained by using a fewer number of trainable parameters than most state-of-the-art pre-trained models. The feature fusion technique makes QuadSNet robust on a range of data. The training data size, type of images, number of classes and number of samples in each class are controlled with QuadSNet at an agreeable level.

The rest of the paper is organized as follows. In Section II, the proposed QuadSNet model and its architecture in detail is discussed. Section III presents a clear demonstration of the datasets, training procedure and experimental results. The best results in all the tables are in **bold**. Discussion and conclusion are in the end.

II. The Proposed QuadSNet

The proposed QuadSNet model uses a quad-stream approach, as illustrated in Fig. 1. Out of the four streams, two streams use conventional convolutional neural networks [16], denoted as $CS1$ and $CS2$, and two streams use separable convolutional neural networks [7], denoted as $SS1$ and $SS2$, respectively. Each stream is based on blocks, whose detailed design is depicted in Fig. 1.

Each of the streams, $CS1$, $CS2$, $SS1$ and $SS2$, consists of four blocks. The design of the streams helps them to extract different features because of dissimilar kernel sizes and different forms of convolutional mode. Due to such a technique, the model becomes wider rather than deeper, and a wider approach helps to reduce the number of trainable parameters without losing accuracy. Features extracted from each stream are concatenated to make a fusion of the learned features. The model learns various features simultaneously. Therefore the training time is drastically reduced.

Separable convolutional neural networks are faster than conventional convolutional neural networks because of their depth-wise and point-wise feature extraction mechanism [7]. They perform fewer multiplications during operation than conventional convolutional neural networks. Therefore, both types of convolutional neural networks in are used in the model to make it faster and more robust. Model functions in a conventional manner as any convolutional neural network may operate. Thus, the quad-stream approach helps the model to learn features simultaneously on a single input image.

A. Feature Representation and Fusion

Formally, $\mathbf{F}_{CS1}$, $\mathbf{F}_{CS2}$, $\mathbf{F}_{SS1}$ and $\mathbf{F}_{SS2}$ are represented as the extracted features of the streams $CS1$, $CS2$, $SS1$ and $SS2$ respectively. The features $\mathbf{F}_{CS1}$ and $\mathbf{F}_{CS2}$ are generated

through conventional convolutional neural networks, whereas $\mathbf{F}_{SS1}$ and $\mathbf{F}_{SS2}$ are the features generated through the streams of separable convolutional neural networks. The $CS1$ and $SS1$ streams use the kernel size of 3×3 for conventional convolutional layers and separable convolutional layers respectively, and use 2×2 for maxpooling layers. Similarly the $CS2$ and $SS2$ streams use the kernel size of 5×5 for conventional convolutional layers and separable convolutional layers, respectively, and 2×2 for maxpooling layers.

The features generated by $CS1$ and $CS2$ are concatenated into $\mathbf{F}_{cs}$. Likewise the features obtained through the streams $SS1$ and $SS2$ are concatenated into $\mathbf{F}_{ss}$. Finally the features $\mathbf{F}_{cs}$ and $\mathbf{F}_{ss}$ are concatenated into $\mathbf{F}_{total}$, as shown in Equation 1.

$$\mathbf{F}_{cs} = \begin{pmatrix} \mathbf{F}_{CS1} \\ \mathbf{F}_{CS2} \end{pmatrix}, \mathbf{F}_{ss} = \begin{pmatrix} \mathbf{F}_{SS1} \\ \mathbf{F}_{SS2} \end{pmatrix}, \mathbf{F}_{total} = \begin{pmatrix} \mathbf{F}_{cs} \\ \mathbf{F}_{ss} \end{pmatrix} \quad (1)$$

The feature fusion in $\mathbf{F}_{total}$ is eventually average-pooled to retain effective features and shrink the size of feature vectors.

B. Micromanagement of Over-fitting

The model handles over-fitting at the micro-level on each block of the model. Each block includes a dropout layer and a batch normalization layer. The dropout layer uses a fixed value of 0.3. The batch normalization layer depends upon two important parameters, momentum and epsilon. The momentum parameter is set to 0.99 and the ϵ is set to 0.0001.

Other than the block-level dropout layers, there is a dropout layer after the average-pooling layer as well. Customarily the dropout layers using a value more than 0.5 in medical image classification are not considered in best practices. Therefore, it is purposefully kept at 0.3.

C. Margin Loss

The margin loss [28] with dynamic weights is adopted to handle the class imbalance and unnecessarily instant overfitting of the model due to a smaller training set. Preliminary experiments showed that the margin loss performs better than the cross-entropy loss. Therefore, the margin loss is opted to be used, as represented in Equation (2).

$$\begin{aligned} \Gamma_\kappa = T_\kappa max(0, m^+ - \|\nu_\kappa\|^2) \quad + \\ \lambda(1 - T_\kappa)max(0, \|\nu_\kappa\|^2 - m^-) \end{aligned} \quad (2)$$

In Equation (2) $\kappa = 1$ if κ class is present, $m^+ = 0.9$, $m^- = 0.1$ and $\lambda = 0.5$ is down-weighting. Here ν_κ represents the feature vector with κ number of classes. The proposed model can be used not only for binary classification but also for multi-class classification. Therefore κ can be any finite discrete number. The values of m^+, m^- and λ dynamically handle the value of total loss Γ_κ. Ultimately the loss is minimized gradually by controlling the imbalance among the κ classes.
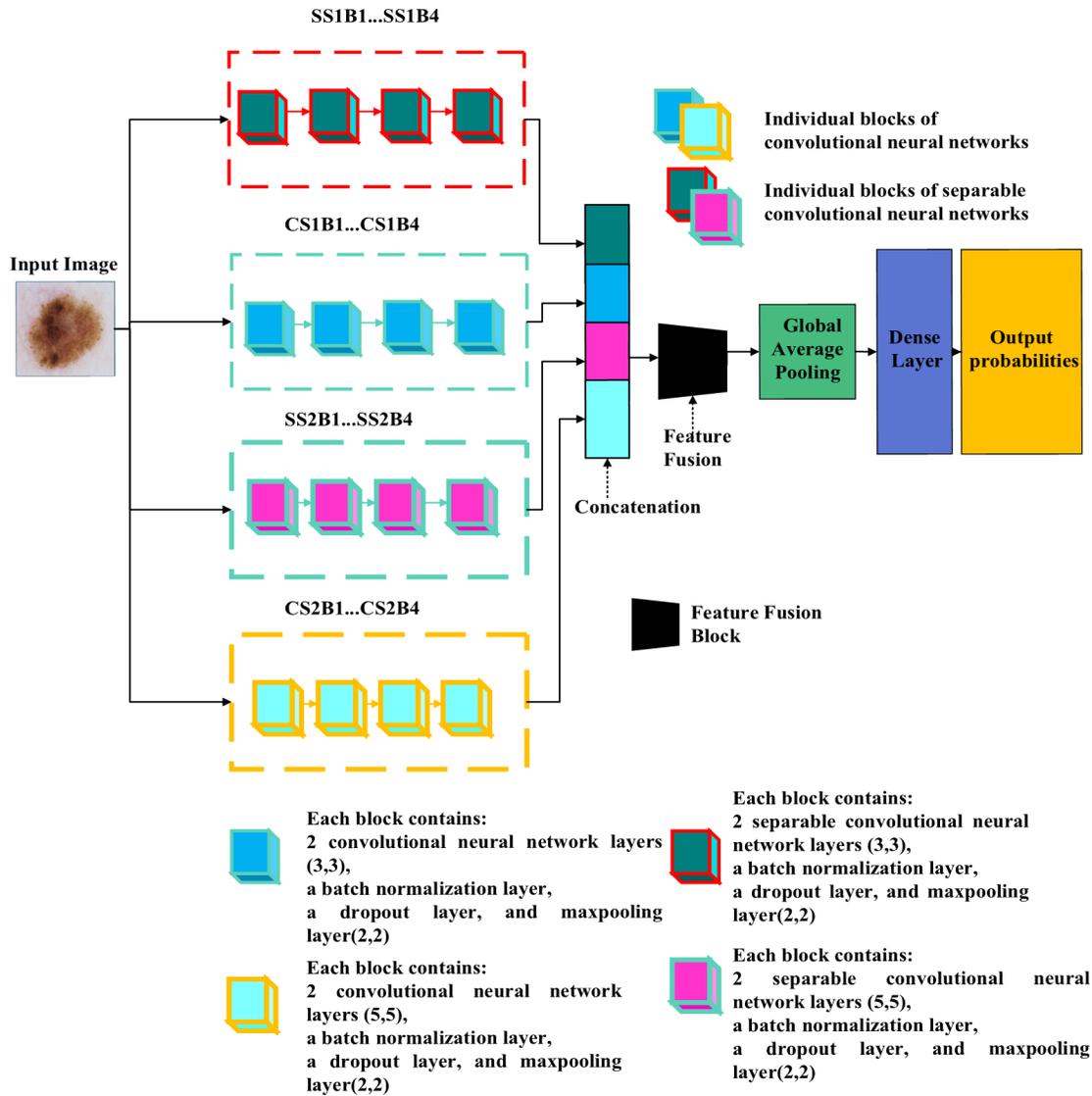
Fig. 1. Block Diagram of the QuadSNet Architecture. There are Four Streams and Each Stream Contains four Individual Blocks.

III. Experiments

A. Datasets

The proposed QuadSNet model and six state-of-the-art pre-trained models are trained and tested on six different types of medical datasets and one non-medical image dataset. For convenience, the seven datasets are symbolically, named, as Θ_1, Θ_2, ..., up to Θ_7 and the six pre-trained models as DN, IN, MN, RN, VG and XC, respectively.

1) Brief description of datasets: Each of the datasets in the list serves a unique purpose because of images' types and classes' degree of balance. Most datasets are imbalanced other than Θ_7, the MNIST dataset. Table I illustrates details of the datasets. The Θ_1, Θ_2, and Θ_3 are used as the benchmark datasets. Dataset Θ_1 consists of chest X-ray images in grayscale. Similarly, Θ_2 consists of CT-scan images in grayscale. Dataset Θ_3 is a colored dataset because of the dermoscopic images. As two of the benchmark datasets are

TABLE I. Brief Description of Datasets. DR: Diabetic Retinopathy

Symbol	Dataset	#Samples
Θ_1	Pneumonia Classification	10855
Θ_2	COVID-19 Classification	746
Θ_3	Skin Lesion Classification	3297
Θ_4	DR Classification on Fundus Photographs	18632
Θ_5	DR Classification on OCT Data	84484
Θ_6	Malaria Cell Classification	27500
Θ_7	MNIST Handwritten Digit Classification	60000

grayscale and one is colored, the models have a fair chance of depicting efficacy. The QuadSNet and each of the six pre-trained models are trained on these datasets. The obtained results are then compared and analyzed comprehensively. Three benchmark datasets are described as follows:

1) Pneumonia (Θ_1): Θ_1 is a chest X-ray dataset for pneumonia classification [14]. This dataset is officially available at

https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia/.

2) COVID-19 (Θ_2): Θ_2 is about COVID-19 (the data is from CT-scans) [29] and it is officially presented at https://covid-ct.grand-challenge.org.

3) Cancer (Θ_3): Dataset Θ_3 is RGB dataset based on dermoscopic images for skin lesion classification [26]. The dataset is a random subset of the dataset present at https://challenge2019.isic-archive.com/data.html.

Other than benchmark datasets, QuadSNet is also trained on Θ_4, Θ_5, Θ_6, Θ_7 to test its capabilities on a variety of image data. A brief description of the remaining datasets is presented as follows:

1) DR Classification on fundus photographs (Θ_4): Θ_4 is a relatively larger dataset than Θ_1, Θ_2 and Θ_3; it is based on color fundus photographs. The data is officially presented at
https://www.kaggle.com/c/diabetic-retinopathy-detection/data.

2) DR Classification on OCT images (Θ_5): Θ_5 is the largest dataset among all the datasets and it is massively imbalanced due to the very high number of normal samples and the smaller number of diseased samples [14]. It is a multi-class classification data based on OCT (Optical Coherence Tomography) images of the retina of diabetic patients. This dataset contains four classes, called Choroidal Neovascularization (CNV), Diabetic Macular Edema (DME), DRUSEN, and Normal. This dataset is officially hosted at https://www.kaggle.com/paultimothymooney/kermany2018.

3) Malarial Cell Classification (Θ_6): Θ_6 is for Malarial cell classification. It is a balanced dataset [23]. This dataset is officially available at https://lhncbc.nlm.nih.gov/publication/pub9932. This dataset contains two classes based on Malarial and Non-Malarial cell images.

4) Hand Written Digit Classification MNIST (Θ_7): This dataset is based handwritten digits [17] from 0 to 9. The dataset Θ_7 is used to investigate the capabilities of QuadSNet on non-medical image datasets. Officially this dataset is present at http://yann.lecun.com/exdb/mnist/.

The main objective of using the seven different datasets is to analyze the diversity of the QuadSNet on a variety of image data. A few of the datasets are smaller and imbalanced except MNIST. The selected datasets are based on colored or grayscale images. Hence QuadSNet is tested accordingly.

B. Training and Testing

The QuadSNet and pre-trained models are trained and tested on a Windows 10 PC equipped with NVidia Gforce GTX 1060, having 16 GB of RAM, Intel Ci7 64 bit processor. All the simulations are performed on Keras with Tensorflow at the backend. The pre-trained models are individually trained on each of the three benchmark datasets.

Essential supplements and parameters, for instance, the learning rate $\gamma = 0.0001$, the optimizer, batch size $\beta = 32$,

TABLE II. COMPARISON OF QUADSNET WITH STATE-OF-THE-ART PRE-TRAINED MODELS IN TERMS OF THE TRAINABLE PARAMETERS

Abbr.	Model	#Trainable parameters
DN	DenseNet121	8,062,504
IN	InceptionV3	23,851,784
MN	MobileNet	**4,253,864**
RN	ResNet50	25,636,712
VG	VGG16	138,357,544
XC	Xception	22,910,480
QN	QuadSNet	**5,235,944**
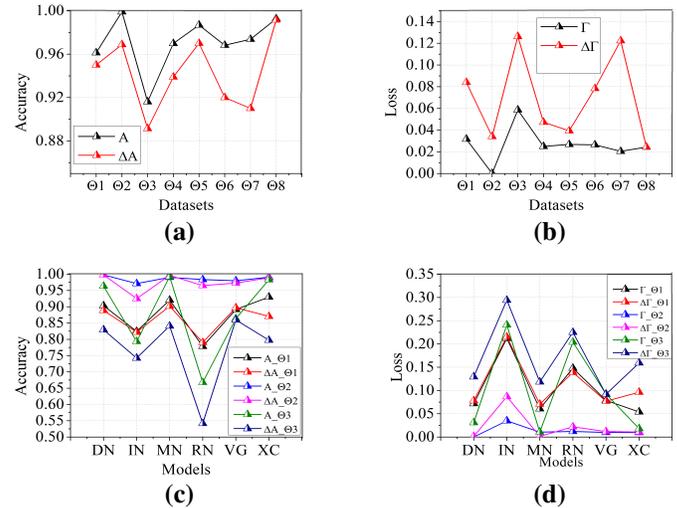
Fig. 2. (a) The Final Training Accuracy (A) and Validation Accuracy (ΔA) of QuadSNet on all Datasets, (b) The Final Training Loss (Γ) and Validation Loss ($\Delta\Gamma$) of QuadSNet on all Datasets, (c) The Final Training Accuracy (A) and Validation Accuracy (ΔA) Curves for Θ_1, Θ_2 and Θ_3 of Different Models and (d) The Final Training Loss (Γ) and Validation Loss ($\Delta\Gamma$) Curves for Θ_1, Θ_2, and Θ_3 of Different Models.

momentum $\mu = 0.009$, and input image size has been kept the same for all the models, including QuadSNet. Adam [15] optimizer has been used throughout the training. The models have been trained up to their maximum potential to nullify any unfairness.

Fig. 2 shows the training accuracy (A), the validation accuracy (ΔA), the final training loss (Γ), and the validation loss ($\Delta\Gamma$) of the QuadSNet model on all the datasets. The distribution of individual datasets into train, validation, and test sets has been kept identical for all the models. The QuadSNet model has been trained on all the datasets, whereas the pre-trained models have been trained only on the three benchmark datasets.

IV. DISCUSSION

The training accuracy A and the validation accuracy ΔA of the pre-trained models and QuadSNet, as presented in Figure 2, exhibit the upshot of training with margin loss. As in this paper the weighted margin loss is used, it noticeably helps to control the effect of imbalance and scarceness of data with the help of QuadSNet. QuadSNet is trained from scratch. Therefore the micromanagement of over-fitting at the block level assists the model to handle it instantaneously.

The pre-trained models traditionally are trained on Ima-

geNet [10], which is mostly based on natural images. In the transfer learning approach, the intended dataset and type of images play a vital role in a model's performance. If the results produced by pre-trained models are studied, datasets Θ_1 and Θ_2 were much easier because of precise edges and object-like structures. Whereas dataset Θ_3 looks harder for the models to extract features due to the lack of natural objects like features.

The same conditions of pre-trained models apply to Quad-SNet as well. Therefore it yields higher accuracy on datasets Θ_1 and Θ_2 than on dataset Θ_3. The massive advantage of QuadSNet over pre-trained models in feature extraction is the fusion of multiple features obtained from various streams. Such features are rich due to different convolution techniques and different sizes of the kernels. The streams in QuadSNet are based on both separable and conventional convolutional neural networks; therefore, the features can be complementary with each other and contribute to the model's accuracy on diverse data. Therefore, QuadSnet achieves a reasonable accuracy on heterogeneous image data.

A. Experimental Results

The systematic analysis of the experimental results of QuadSNet and the pre-trained models after the training reveals a massive difference between the number of trainable parameters. Table II shows the details about the number of trainable parameters for each of the models. QuadSNet and MobileNet have the fewest trainable parameters, which are about one-fifth of the average parameters.

Traditionally accuracy is not considered as a better performance metric for medical image classification on imbalanced data, therefore, sensitivity, specificity and F_1 score are used.

All the datasets are distributed in three portions (Training, Validation and Test). The main objective is to analyze the performance of QuadSNet in comparison with pre-trained models on imbalanced and smaller datasets. Therefore, the chosen three benchmark datasets, Θ_1, Θ_2 and Θ_3, are the smallest and most imbalanced among the seven datasets.

The performances of pre-trained models and QuadSNet on Θ_1, Θ_2 and Θ_3 are depicted in Table III. It can be concluded that QuadSNet achieves the best performance in all the benchmark datasets at almost all the metrics.

Except for the six state-of-the-art pre-trained models, QuadSNet is compred with some of the latest publications on each of the datasets which have better performance on the particular dataset. The results can be observed from Table IV. Each of the compared methods is designed for a particular disease. QuadSNet exhibits competitive accuracy, specificity and sensitivity in comparison with existing works. Most existing works listed in Table IV utilize pre-trained models having a huge number of parameters and use data augmentation techniques, whereas QuadSNet produces comparable results with a few trainable parameters and without any data augmentation. The dataset Θ_7 is a non-medical image dataset. Therefore the results based on accuracy are compred only.

B. Limitations of the Study and Future Work

The proposed model has been thoroughly tested on the datasets described in Section III-A1. However, the proposed

TABLE III. THE PERFORMANCE OF PRE-TRAINED MODELS AND QUADSNET ON Θ_1, Θ_2 AND Θ_3 DATASETS

Dataset	Model	Sensitivity	Specificity	Accuracy	F_1 Score
Θ_1	DN	0.978	0.875	0.920	0.915
	IN	0.987	0.792	0.865	0.846
	MN	1.000	0.926	0.960	0.958
	RN	0.924	0.771	0.830	0.809
	VG	0.990	0.952	0.970	0.969
	XC	0.929	0.912	0.920	0.919
	QN	**0.990**	**0.971**	**0.980**	**0.980**
Θ_2	DN	0.877	1.000	0.891	0.934
	IN	0.956	1.000	0.964	0.977
	MN	0.926	1.000	0.938	0.961
	RN	0.867	1.000	0.881	0.928
	VG	0.961	0.949	0.959	0.974
	XC	**0.987**	0.892	0.964	0.976
	QN	0.962	**1.000**	**0.969**	**0.980**
Θ_3	DN	0.932	0.734	0.817	0.810
	IN	0.711	0.796	0.738	0.785
	MN	0.867	0.807	0.838	0.848
	RN	**0.985**	0.504	0.552	0.305
	VG	0.887	0.804	0.846	0.853
	XC	0.809	0.793	0.802	0.821
	QN	0.937	**0.917**	**0.928**	**0.945**

TABLE IV. COMPARISON OF QUADSNET WITH SOME OF THE LATEST STATE-OF-THE-ART METHODS ON Θ_1, Θ_2 AND Θ_3, Θ_4 , Θ_5, Θ_6 AND Θ_7 DATASETS. SEN: SENSITIVITY, ACC: ACCURACY

Datasets	Author	Sen	Spe	Acc
Θ_1	Rahman et al. [22]	0.990	0.970	0.980
	QN	**0.999**	0.970	**0.981**
Θ_2	Li et al. [20]	0.826	**1.000**	0.976
	QN	**0.960**	1.000	**0.977**
Θ_3	Kassem et al. [13]	0.798	**0.970**	0.943
	QN	**0.937**	0.910	**0.945**
Θ_4	Wan et al. [27]	0.880	**0.950**	0.900
	QN	**0.995**	0.858	**0.915**
Θ_5	Li et al. [19]	0.960	0.980	0.970
	QN	**0.960**	**1.000**	**0.990**
Θ_6	Masud et al. [21]	**0.970**	0.940	**0.970**
	QN	0.959	0.923	0.940
Θ_7	Ali et al. [1]	—	—	0.990
	QN	—	—	**0.999**

model has not been tested on 3-D medical image data, which may impact the performance. Similarly, the proposed model has not been tested in multi-input or multi-output settings where multiple tasks are performed simultaneously. In the future, the QuadSNet's efficacy may be tested on multimodal or 3-D medical image data.

V. CONCLUSION

Compared with popular image datasets such as ImageNet for classification, medical datasets are usually small and class imbalance. This work introduces a new model called QuadSNet that can effectively handle over-fitting and class imbalance on small datasets without using transfer learning and data augmentation to handle such issues for image data. In general, QuadSNet outperforms the pre-trained models in terms of sensitivity, specificity, accuracy, and F_1 score on three medical benchmark datasets. Additionally, QuadSNet has competitive results compared to several state-of-the-art works focusing on a certain disease, demonstrating the effectiveness of the model. QuadSNet exhibits efficacy on a variety of datasets, including grayscale and color images. Due to the ability to handle diverse data, QuadSNet has the potential of

becoming a universal model for medical image classification on small datasets. This paper uses margin loss only, but QuadSNet can be trained and tested using other loss functions to analyze its effectiveness.

REFERENCES

[1] Ali S, Li J, Pei Y, Aslam MS, Shaukat Z, Azeem M (2020) An effective and improved cnn-elm classifier for handwritten digits recognition and classification. Symmetry 12(10):1742

[2] Ali-Gombe A, Elyan E (2019) Mfc-gan: class-imbalanced dataset classification using multiple fake class generative adversarial network. Neurocomputing 361:212–221

[3] Amin SU, Alsulaiman M, Muhammad G, Mekhtiche MA, Hossain MS (2019) Deep learning for eeg motor imagery classification based on multi-layer cnns feature fusion. Future Generation computer systems 101:542–554

[4] An N, Ding H, Yang J, Au R, Ang TF (2020) Deep ensemble learning for alzheimer's disease classification. Journal of Biomedical Informatics p 103411

[5] Apostolopoulos ID, Mpesiana TA (2020) Covid-19: automatic detection from x-ray images utilizing transfer learning with convolutional neural networks. Physical and Engineering Sciences in Medicine p 1

[6] Chen YJ, Chang YJ, Wen SC, Shi Y, Xu X, Ho TY, Jia Q, Huang M, Zhuang J (2020) Zero-shot medical image artifact reduction. In: 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI), IEEE, pp 862–866

[7] Chollet F (2017) Xception: Deep learning with depthwise separable convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 1251–1258

[8] Dunnhofer M, Antico M, Sasazawa F, Takeda Y, Camps S, Martinel N, Micheloni C, Carneiro G, Fontanarosa D (2020) Siam-u-net: encoder-decoder siamese network for knee cartilage tracking in ultrasound images. Medical Image Analysis 60:101631

[9] 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, pp 770–778

[10] Hinton GE, Krizhevsky A, Sutskever I (2012) Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems 25:1106–1114

[11] Howard AG, Zhu M, Chen B, Kalenichenko D, Wang W, Weyand T, Andreetto M, Adam H (2017) Mobilenets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:170404861

[12] Iandola F, Moskewicz M, Karayev S, Girshick R, Darrell T, Keutzer K (2014) Densenet: Implementing efficient convnet descriptor pyramids. arXiv preprint arXiv:14041869

[13] Kassem MA, Hosny KM, Fouad MM (2020) Skin lesions classification into eight classes for isic 2019 using deep convolutional neural network and transfer learning. IEEE Access 8:114822–114832

[14] Kermany DS, Goldbaum M, Cai W, Valentim CC, Liang H, Baxter SL, McKeown A, Yang G, Wu X, Yan F, et al. (2018) Identifying medical diagnoses and treatable diseases by image-based deep learning. Cell 172(5):1122–1131

[15] Kingma J (2014) Adam: A method for stochastic optimization. arXiv preprint arXiv:14126980

[16] LeCun Y, Bengio Y, et al. (1995) Convolutional networks for images, speech, and time series. The handbook of brain theory and neural networks 3361(10):1995

[17] Lecun Y, Cortes C, Burges C (1999) The mnist dataset of handwritten digits(images)

[18] Li C, Zhang D, Tian Z, Du S, Qu Y (2020) Few-shot learning with deformable convolution for multiscale lesion detection in mammography. Medical Physics

[19] Li F, Chen H, Liu Z, Zhang Xd, Jiang Ms, Wu Zz, Zhou Kq (2019) Deep learning-based automated detection of retinal diseases using optical coherence tomography images. Biomedical Optics Express 10(12):6204–6226

[20] Li K, Fang Y, Li W, Pan C, Qin P, Zhong Y, Liu X, Huang M, Liao Y, Li S (2020) Ct image visual quantitative evaluation and clinical classification of coronavirus disease (covid-19). European radiology pp 1–10

[21] Masud M, Alhumyani H, Alshamrani SS, Cheikhrouhou O, Ibrahim S, Muhammad G, Hossain MS, Shorfuzzaman M (2020) Leveraging deep learning techniques for malaria parasite detection using mobile application. Wireless Communications and Mobile Computing 2020

[22] Rahman T, Chowdhury ME, Khandakar A, Islam KR, Islam KF, Mahbub ZB, Kadir MA, Kashem S (2020) Transfer learning with deep convolutional neural network (cnn) for pneumonia detection using chest x-ray. Applied Sciences 10(9):3233

[23] Rajaraman S, Antani SK, Poostchi M, Silamut K, Hossain MA, Maude RJ, Jaeger S, Thoma GR (2018) Pre-trained convolutional neural networks as feature extractors toward improved malaria parasite detection in thin blood smear images. PeerJ 6:e4568

[24] Simonyan A Karen (2014) Very deep convolutional networks for large scale image recognition. arXiv preprint arXiv:14091556

[25] Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A (2015) Going deeper with convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 1–9

[26] Tschandl P, Rosendahl C, Kittler H (2018) The ham10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. Scientific data 5:180161

[27] Wan S, Liang Y, Zhang Y (2018) Deep convolutional neural networks for diabetic retinopathy detection by image classification. Computers & Electrical Engineering 72:274–282

[28] Xiao Q, Luo H, Zhang C (2017) Margin sample mining loss: A deep learning based method for person re-identification. arXiv preprint arXiv:171000478

[29] Zhao J, Zhang Y, He X, Xie P (2020) Covid-ct-dataset: a ct scan dataset about covid-19. arXiv preprint arXiv:200313865