The Korean Institute of Communications and Information Sciences

Available online at www.sciencedirect.com

ScienceDirect

ICT Express 8 (2022) 271–275

www.elsevier.com/locate/icte

DFU_SPNet: A stacked parallel convolution layers based CNN to improve Diabetic Foot Ulcer classification

Sujit Kumar Das*, Pinki Roy, Arnab Kumar Mishra

Department of CSE, National Institute of Technology, Silchar 788010, Assam, India

Received 2 June 2021; received in revised form 12 August 2021; accepted 27 August 2021
Available online 16 September 2021

Abstract

Diabetic Foot Ulcer (DFU) is a complication of diabetes that causes lower limb amputation. In this work, a unique stacked parallel convolution layers-based network (DFU_SPNet) is proposed to perform DFU vs. normal skin classification. The main objective of this work is to design an effective CNN-based classification model, along with proper fine-tuning of optimizer settings. DFU_SPNet consists of 3 blocks of parallel convolution layers with multiple kernel sizes, for local and global feature abstractions. The proposed DFU_SPNet, trained using SGD (with momentum) optimizer with $1e-2$ learning rate on the DFUNet dataset, outperformed the current state-of-the-art results with an AUC of 0.974.

Keywords: Diabetic Foot Ulcer; Parallel convolution; Convolutional neural network; Classification

1. Introduction

Diabetic patients have several complications, and DFU is one of them. The DFU results in frequent hospitalization, lower feet or limb amputation, low quality of life, and costly treatment. The data indicates that 15% of diabetic patients suffer from DFU [1]. Due to failure in early diagnosis and lack of proper treatments on time, more than a million diabetic patients lose part of their legs every year [2]. The traditional medical diagnosis of DFU and treatment requires several expensive and time-consuming clinical tests like MRI, CT scans, X-rays, blood tests etc. In recent years the major developments of Machine Learning (ML), Computer Vision (CV), and Deep Learning (DL) algorithms in the medical imaging field [3–5] suggest the potential of such systems in automatic DFU identification.

In recent times, multiple approaches are introduced to classify normal and abnormal DFU skin. Goyal et al. [6] have proposed a novel CNN architecture and named it as DFUNet. DFUNet consisted of parallel convolution blocks to concatenate features from convolution operations and traditional single convolutional layers. Further, to create baseline results on DFU identification, Linde–Buzo–Gray (LBP) [7] based low-level image processing features were extracted. Also, standard LeNet [8], AlexNet, and GoogleNet [9] have been used to extract high-level features. These features are fed to ML classifiers for DFU identification. One potential limitation of the proposed parallel convolutional block based DFUNet is the lack of adequate transition layers between parallel convolutional blocks. As demonstrated by [10], appropriate usage of transition layers can help learn better abstractions from multi-layer feature map concatenations. In another work, Alzubaidi et al. [11] have proposed a novel CNN architecture DFU_QUTNet to extract multi-level features, and these features were finally fed to Support Vector Machine (SVM) and K-Nearest Neighbors (KNN) classifiers. The novel DFU_QUTNet architecture was designed by increasing the network width without compromising computation cost. The classification results of DFU_QUTNet based features and SVM classifier outperformed the standard CNN models (GoogleNet, AlexNet, and VGG16). Even though impressive performances were obtained by the DFU_QUTNet model, a potential limitation of this approach is due to the use of same kernel sizes (3×3) for the parallel convolutional layers. Instead of such same sized convolutional kernels, use of heterogeneous types of kernels could help

* Corresponding author.
E-mail addresses: dassujit88@gmail.com (S.K. Das), pinkiroy2405@gmail.com (P. Roy), arnab.mishra.1992@gmail.com (A.K. Mishra).

Peer review under responsibility of The Korean Institute of Communications and Information Sciences (KICS).

https://doi.org/10.1016/j.icte.2021.08.022

improve the receptive field of the entire network. In the DFU Grand Challenge (DFUC), 2020 Cassidy et al. [12] have used Faster-RCNN [13] and Inception-v2-Resnet101, FRCNN Inception-v2-ResNet101, YOLOv5, [14] and EfficientDet [15] to identify DFU cases. They have observed that these networks are significant in providing promising results. In a DFU localization task, Goyal et al. [16] have used Faster R-CNN [13] with InceptionV2 [9] architecture and a two-stage transfer learning approach. Yap et al. [17] used Faster R-CNN [13], YOLOv5 [14], and EfficientDet [15] for DFU identification. Faster R-CNN focuses on reducing false positives and improving detection performance. Also, the Monte Carlo cross-validation methodology has been used during training. However, for the classification of normal vs. abnormal (DFU) skin patches, the current literature lacks appropriate exploration of efficient techniques. Moreover, adequate fine-tuning of DL network optimizer settings is not studied in the current literature.

Motivated by the limitations of the current literature in this area, the objectives of the current work are identified as shown below-

- To design a new CNN architecture for improving the classification performance of normal vs. abnormal DFU cases, with the help of parallel convolutional layers with appropriate transition layers for abstraction.
- To study optimizer setting based fine-tuning of the proposed CNN model, by exploring multiple optimizers and learning rates, so as to enhance its performance.

2. Material and methods

2.1. Dataset description

The DFU dataset is not publicly available, but can be accessed through the proper procedure. We have acquired it following the release agreement provided by the dataset principal investigator.[1] The samples are collected by the DFU dataset creator utilizing a database consisting of 292 Ulcer foot images from Lancashire Teaching Hospitals, UK. Further, for normal samples, images of 105 healthy foot images are captured from the same hospital. All the samples are acquired with patient consent and approval from the NHS Research Ethics Committee to use these samples for research purposes. The natural color characteristics of the samples are ensured by being captured in room light conditions with Nikon D3300 capturing device. Also, images are captured in close-up range in parallel orientation from a distance of 30–40 cm from the ulcer area using the additional lens. After collecting the full foot image samples, Region Of Interest (ROIs) are extracted, and annotation is done as normal and abnormal classes by medical experts. Finally, the dataset consisted of 1679 colored image patches (641 normal and 1038 abnormal classes). A few examples of both normal and abnormal classes of DFU datasets are shown in Fig. 1. The image patches were not uniform in terms of size. Therefore, for the current work, image patches are rescaled to 128 × 128 for computational simplicity.

[1] http://www2.docm.mmu.ac.uk/STAFF/M.Yap/dataset.php

Fig. 1. Examples of image patches from the dataset.

2.2. Method

The proposed DFU_SPNet shown in Fig. 2 consists of three stacked parallel convolution layers of kernel size 1 × 1, 3 × 3, and 5 × 5. The number of the kernels are in increasing order (32, 64, and 128). These diverse set of convolutional Kernels helped in learning both local as well as global feature abstractions effectively. In the stacked parallel convolution block, 1 × 1 strides with ReLU activation have been used for each convolution layer. The preceding layer of the first stacked parallel block is a single convolution layer of 7 × 7 with strides 2 × 2 followed by a transition layer. The transition layer consists of BatchNormalization (BN) and LeakyReLU activation. It helps in standardizing the inputs to accelerate the training. The concatenate output of each stacked parallel convolution block is passed through a sequence of 1 × 1 convolution, transition, and 2 × 2 MaxPooling layer. Finally, the flatten output of three such stacked parallel convolution blocks and its successive sequence layers are passed through a fully connected layer of 32 nodes with a dropout layer. The output layer consists of a sigmoid activation function to get a binary predicted value of either 0 (Normal) or 1 (Abnormal) class. The proposed DFU_SPNet with multiple different sized kernels is proved to be a significant architecture to learn important features from the input images. The first convolution layers' activation maps, which are the input of the first stacked parallel convolution layer of DFU_SPNet, are shown in Fig. 3.

3. Results

3.1. Experimental setup

All experiments here, are accomplished using Tensorflow [18] python package in a freely available Kaggle GPU environment. At initial stage, the dataset is divided into an 80:20 random split where 80% of data are used for training and 20% of data kept for testing. Further, 20% of the training data are kept as a validation set to overcome the overfitting problem. These resulted in 1074 training, 336 for testing, and 269 for validation. Further, to increase the number of training samples,

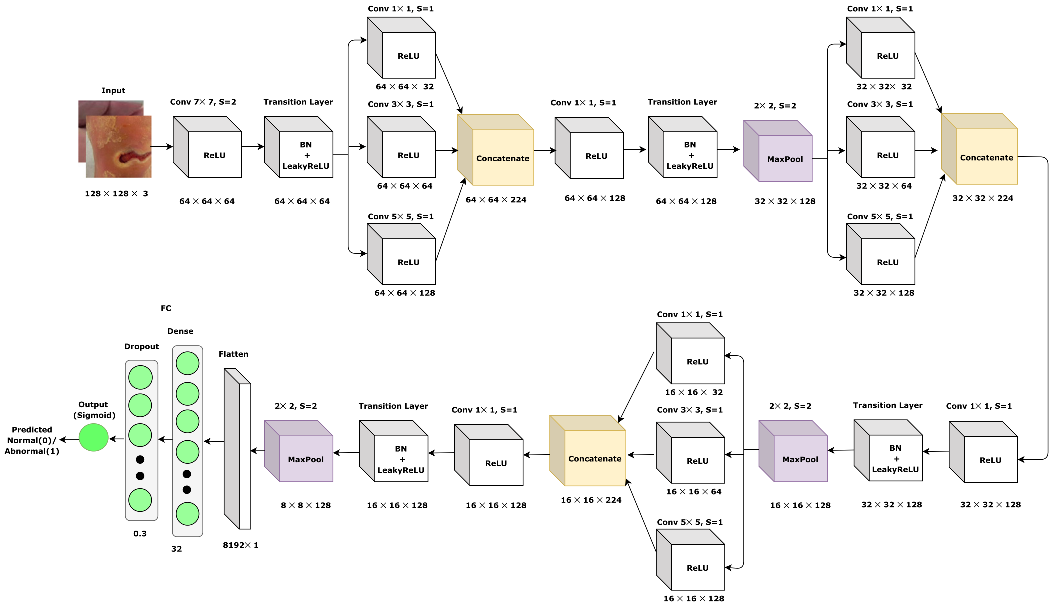

Fig. 2. The proposed model diagram (DFU_SPNet).

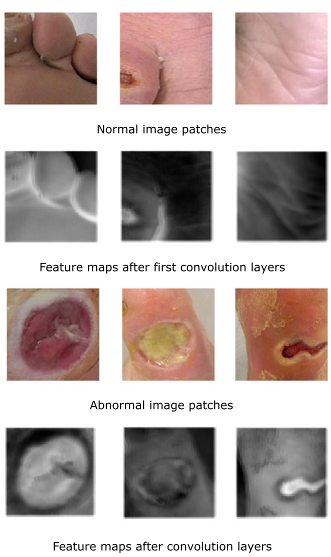

Fig. 3. Features maps of DFU_SPNet after first convolution.

natural data augmentation techniques are used. The horizontal and vertical flips of the training images increased the training set to 3222 image patches. DFU_SPNet is trained on multiple optimizer and learning rate setups. The popularly used are Adam, SGD (with momentum), and Adagrad optimizers with learning rates $1e-2$ and $1e-3$. It made a complete six sets of experiments. The performance of the proposed DFU_SPNet is measured in terms of various important metrics like accuracy, sensitivity, specificity, precision, F1-Scores, AUC, and Error rate. To observe the average behavior of the model's performance, results are collected from 10 different random splits on a run of 100 epochs and patience of 50 epochs (as early stopping). The 6 sets of experiments and 10 runs for each set resulted in 60 times training and testing of the proposed DFU_SPNet.

Table 1
Average Accuracy, F1-Scores and AUC of the proposed DFU_SPNet.

Opt_Lr	Accuracy	F1-Scores	AUC
Adam_1e-2	0.944	0.928	0.968
Adam_1e-3	0.922	0.902	0.950
SGD_1e-2	**0.964**	**0.954**	**0.974**
SGD_1e-3	0.952	0.939	0.972
Adagrad_1e-2	0.934	0.917	0.970
Adagrad_1e-3	0.925	0.907	0.968

3.2. Experimental results

The accuracy, F1-Score, and AUC are the most commonly used performance metrics to evaluate such DL models. In Table 1, multiple optimizer and learning rate settings (Opt_Lr), based average accuracy, F-Scores, and AUC results of DFU_SPNet on testing samples are shown. The results are quite promising among all sets, but the best results are achieved when DFU_SPNet is trained using SDG optimizer and $1e-2$ learning rate. The sensitivity and specificity measures give the idea of how accurately a model predicts positive and negative classes separately. In an automatic disease prediction system, these metrics play a very important role. The bar graph shown in Fig. 4 represents the model performance with respect to sensitivity and specificity. The proposed DFU_SPNet performed very well with a good sensitivity score. The highest sensitivity achieved by DFU_SPNet is 0.984. The specificity is comparatively lower than sensitivity by a difference of 0.032, with the value of 0.952. The proposed model is evaluated on precision, recall, and error rate evaluation metrics, and the results are presented in Table 2. The proposed DFU_SPNet has shown significant results in all of these evaluation metrics under consideration.

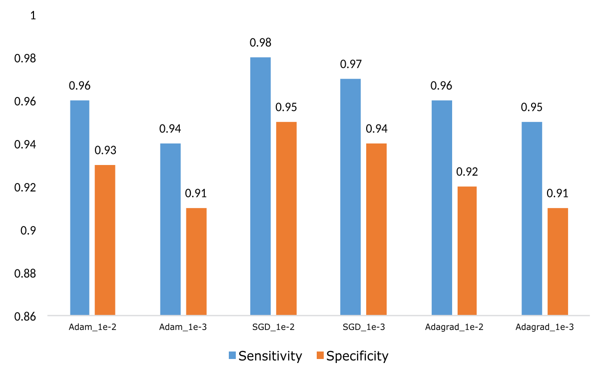

Fig. 4. DFU_SPNet Sensitivity and Specificity bar graph.

Table 2
Average Precision, Recall and Error Rate of the proposed DFU_SPNet.

Opt_Lr	Precision	Recall	Error rate
Adam_1e-2	0.897	0.961	0.056
Adam_1e-3	0.869	0.937	0.077
SGD_1e-2	**0.926**	**0.984**	**0.036**
SGD_1e-3	0.912	0.968	0.048
Adagrad_1e-2	0.878	0.961	0.066
Adagrad_1e-3	0.865	0.953	0.075

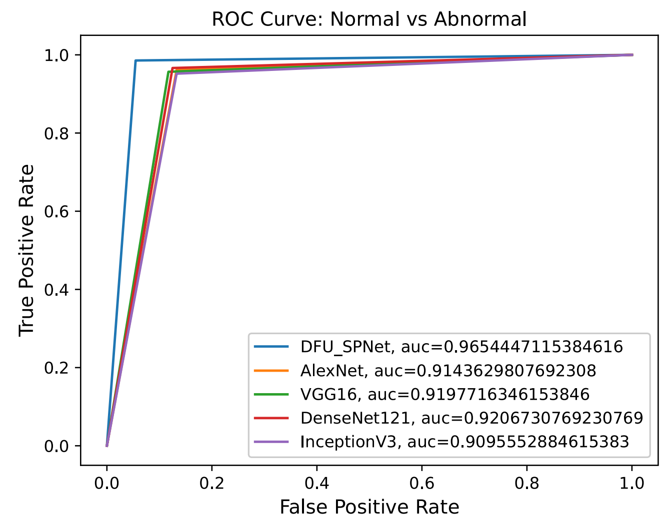

Fig. 5. ROC Curve comparison with standard CNNs.

Table 3
Comparison table with SOTA approaches.

Results	DFUNet[6]	DFU_QUTNet[11]	DFU_SPNet
Accuracy	0.925	–	**0.964**
Sensitivity	0.934	–	**0.984**
Specificity	0.911	–	**0.951**
Precision	0.945	**0.954**	0.926
Recall	–	0.936	**0.984**
F1-Scores	0.939	0.945	**0.954**
AUC	0.961	–	**0.974**

4. Discussion

In this work, classification of normal and abnormal DFU skin patches is performed. The proposed stacked parallel convolution layer-based CNN architecture with an intermediate sequence of convolution, transition, and maxpooling layers is proved to be a unique solution for the task. The effectiveness of the proposed DFU_SPNet in normal vs. abnormal classification is measured with various optimizer and learning rate settings. Among all the settings under consideration, SGD (with momentum) optimizer on $1e-2$ learning rate achieved the highest results in terms of every evaluation metric. The possible reason of such results is that the classes in DFU dataset are linearly separable, except for a few examples. In such situation SGD (with momentum) has more generalization power and the network requires less number of epochs to converge [19]. The superiority of the proposed approach is due to multiple reasons, which are mentioned below:

- The use of stacked parallel convolutional layers helps in better processing regions of interests that are of varying sizes, as such an approach allows for the network to have a much better receptive field.
- The use of multiple different kernel sizes further improves the detection of the disease at multiple scales.
- Moreover, fine-tuning of the optimizer settings for the particular DFU classification task empowers the predictive capabilities of the DFU_SPNet significantly.

Further, the ROC curve in Fig. 5, is a comparison of the proposed DFU_SPNet with a few most commonly used CNN architectures (AlexNet, VGG16, DenseNet121, InceptionV3) on the same sample split. It shows that the DFU_SPNet outperformed the standard CNN architectures in the DFU normal vs. abnormal classification task. The results of DFU_SPNet are compared with the SOTA results in Table 3. The results from Table 3 show that the proposed approach outperformed in case of almost every evaluation metric, except precision. The reason for such precision score in DFU_SPNet is the slightly higher false positive (FP) value. A closer look at the comparison table shows a very high recall value in DFU_SPNet. Therefore, the true positive value is very promising, resulting in a significantly powerful prediction model for DFU abnormal cases. A model with a significant recall or sensitivity score (0.984) with a good precision value is generally considered as an efficient automatic prediction system. Further, in Fig. 6 a few examples of correctly and mis-classified with their predictions from both the classes are shown. The high sensitivity and specificity result, already shown that the proposed DFU_SPNet performed very well in predicting normal and abnormal classes correctly. However, the observations from the mis-classified samples of both classes make it clear that those samples are hard to identify correctly, even for any DFU expert. Further, most of the mis-classified samples have an inter-class similarity problem.

5. Conclusions

In the current work, a stacked parallel convolution layer with a suitable intermediate combination of convolution, transition, and maxpooling layers based CNN network is proposed to identify DFU by classifying normal and abnormal DFU skin

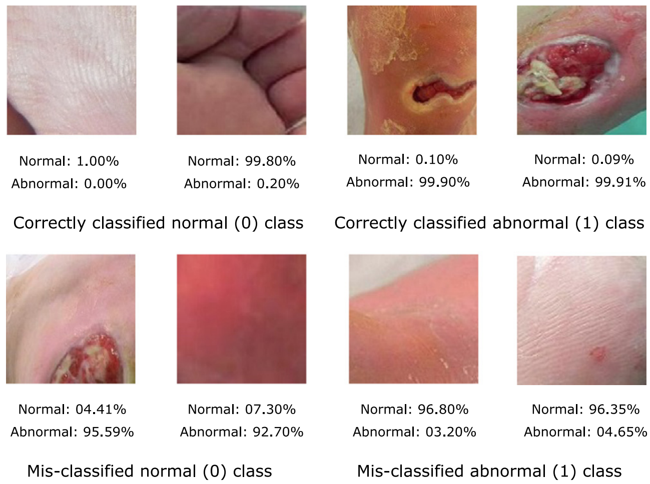

Fig. 6. DFU_SPNet based prediction examples.

patches. The various filter sizes in parallel convolution layers and intermediate layers among each parallel block helped in extracting important features from the input images. Further, natural data augmentation to increase training images helped in solving limited training data related problems. The proposed DFU_SPNet, trained using multiple optimizer and learning rate settings, successfully outperformed the most commonly use standard CNN architectures. The promising mean average accuracy, sensitivity, F1-score, and AUC achieved by DFU_SPNet also show improvements over the SOTA works on the same dataset. Such results suggest that the DFU_SPNet can provide a helping hand to DFU experts in making a faster decision.

CRediT authorship contribution statement

Sujit Kumar Das: Conceptualization, Methodology, Software, Investigation, Writing – original draft, Writing – review & editing. **Pinki Roy:** Resources, Supervision. **Arnab Kumar Mishra:** Validation, Writing – review & editing, Visualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Sebastiano Leone, Renato Pascale, Mario Vitale, Silvano Esposito, Epidemiology of diabetic foot, Infez. Med. 20 (Suppl. 1) (2012) 8–13.
[2] Scott D. Ramsey, Katherine Newton, David Blough, David K. Mccul-loch, Nirmala Sandhu, Gayle E. Reiber, Edward H. Wagner, Incidence, outcomes, and cost of foot ulcers in patients with diabetes, Diabetes Care 22 (3) (1999) 382–387.
[3] Arnab Kumar Mishra, Sujit Kumar Das, Pinki Roy, Sivaji Bandyopadhyay, Identifying COVID19 from chest CT images: a deep convolutional neural networks based approach, J. Healthc. Eng. 2020 (2020).
[4] Arnab K. Mishra, Pinki Roy, Sivaji Bandyopadhyay, Sujit K. Das, Breast ultrasound tumour classification: A Machine Learning—Radiomics based approach, Expert Syst. (2021) e12713.
[5] Sujit Kumar Das, Pinki Roy, Arnab Kumar Mishra, Recognition of ischaemia and infection in diabetic foot ulcer: A deep convolutional neural network based approach, Int. J. Imaging Syst. Technol. (2021).
[6] Manu Goyal, Neil D. Reeves, Adrian K. Davison, Satyan Rajbhandari, Jennifer Spragg, Moi Hoon Yap, Dfunet: Convolutional neural networks for diabetic foot ulcer classification, IEEE Trans. Emerg. Top. Comput. Intell. 4 (5) (2018) 728–739.
[7] Zhenhua Guo, Lei Zhang, David Zhang, A completed modeling of local binary pattern operator for texture classification, IEEE Trans. Image Process. 19 (6) (2010) 1657–1663.
[8] Yann LeCun, Yoshua Bengio, et al., Convolutional networks for images, speech, and time series, Handb. Brain Theory Neural Netw. 3361 (10) (1995) 1995.
[9] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, Andrew Rabinovich, Going deeper with convolutions, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 1–9.
[10] Gao Huang, Zhuang Liu, Laurens Van Der Maaten, Kilian Q. Weinberger, Densely connected convolutional networks, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 4700–4708.
[11] Laith Alzubaidi, Mohammed A. Fadhel, Sameer R. Oleiwi, Omran Al-Shamma, Jinglan Zhang, DFU_QUTNet: diabetic foot ulcer classification using novel deep convolutional neural network, Multimedia Tools Appl. 79 (21) (2020) 15655–15677.
[12] Bill Cassidy, Neil D. Reeves, Pappachan Joseph, David Gillespie, Claire O'Shea, Satyan Rajbhandari, Arun G. Maiya, Eibe Frank, Andrew Boulton, David Armstrong, et al., Dfuc2020: Analysis towards diabetic foot ulcer detection, 2020, arXiv preprint arXiv:2004.11853.
[13] Shaoqing Ren, Kaiming He, Ross Girshick, Jian Sun, Faster r-cnn: Towards real-time object detection with region proposal networks, 2015, arXiv preprint arXiv:1506.01497.
[14] Glenn Jocher, Alex Stoken, Jirka Borovec, Liu Changyu, Adam Hogan, ultralytics/yolov5: v3. 0, 2020, Zenodo.
[15] Mingxing Tan, Ruoming Pang, Quoc V. Le, Efficientdet: Scalable and efficient object detection, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, pp. 10781–10790.
[16] Manu Goyal, Neil D. Reeves, Satyan Rajbhandari, Moi Hoon Yap, Robust methods for real-time diabetic foot ulcer detection and localization on mobile devices, IEEE J. Biomed. Health Inf. 23 (4) (2018) 1730–1741.
[17] Moi Hoon Yap, Ryo Hachiuma, Azadeh Alavi, Raphael Brungel, Manu Goyal, Hongtao Zhu, Bill Cassidy, Johannes Ruckert, Moshe Olshansky, Xiao Huang, et al., Deep learning in diabetic foot ulcers detection: A comprehensive evaluation, 2020, arXiv preprint arXiv:2010.03341.
[18] Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, et al., Tensorflow: Large-scale machine learning on heterogeneous distributed systems, 2016, arXiv preprint arXiv:1603.04467.
[19] Ashia C. Wilson, Rebecca Roelofs, Mitchell Stern, Nathan Srebro, Benjamin Recht, The marginal value of adaptive gradient methods in machine learning, 2017, arXiv preprint arXiv:1705.08292.