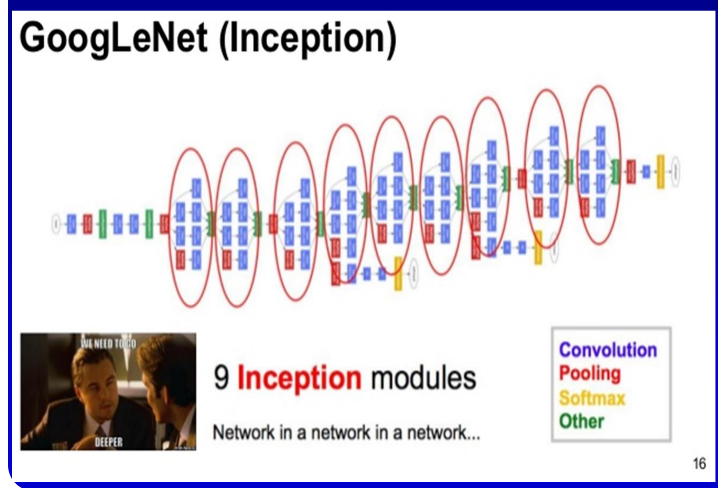
THE EFFECTIVENESS OF TRANSFER LEARNING USING DINET ON MEDICAL IMAGE CLASSIFICATION



James Boit[1]; David Zeng, Phd[2];

[1]Dakota State University; [2]Dakota State University

Introduction

- Background: In the last decade, convolutional neural networks (CNN or ConvNets) have found great success in solving a variety of computer vision problems such as detection, segmentation and recognition of objects. However, this was not the case until the emergence of ImageNet [1], a large-scale image dataset with annotated data used for visual recognition tasks.
- Research Questions (RQ): 1. What is the effectiveness of transfer learning techniques using our novel architecture on medical image classification? 2. What is the optimal cut-off point that gives the best model performance?
- Proposed Architecture: Combines architectural functionality of DENSENET-121 and Inception networks by using DINET modules at different positions of the network.
- Model: We use our novel architecture referred to as DINET and perform a classification problem using a baseline dataset (CIFAR-10).
- Preliminary results: DINET v1 shows promising results with accuracy of about 84.5%.

Literature Review

Why DL?

- Recent advancements in Next-generation computing architectures, for example, Field Programmable Gate Arrays (FPGA); bare metal servers optimized for single-server tenants; modern innovations in DL algorithms, such as Generative Adversarial Networks (GAN); and the rise of powerful and efficient Graphical Processing Units (GPUs) have revolutionized industry verticals such as finance, healthcare, manufacturing etc.
- Deep learning structures such as AlexNet [1] VGGNet [2], RESNet [3], Inception [4] and DenseNet [5] have been successfully used for computer vision problems
- SE-inception-DenseNet [6] improved multi-scale feature learning for unconstrained face recognition tasks.
- In healthcare industries, deep learning approaches have shown great promise in medical diagnosis[7] and treatment.
- DenseNet-121: Big on feature reuse, ensures maximum information flow. Concatenation of dense blocks provides parameter and model performance.
- Inception Network: Class of complex and heavily engineered CNN developed by Google. Comprises 22 Layers to handle classification tasks. Use of residual connections helps mitigate vanishing gradient problem.
- From Literature, we are motivated to introduce a novel architecture to effectively learn multi-scale features and investigate the behavioral performance of our DINET model on various medical image classification tasks, for example ChestX-ray 14 dataset [8].

Methodology

- Method: We apply nonlinearity functions and train the model to evaluate a general image dataset.
- Dataset: Publicly available CIFAR-10 dataset [9]. Comprises of 60,000 images (90% training and 10% test data). CIFAR-10 dataset consists of 10 different classes of images namely: airplanes, cars, birds, cats, deer, dogs, frogs, horses, ships, and trucks.
- Image Size: Each color image is of size 32×32 pixels.
- Baseline Model: DINET architecture as our base model.
- Activation Functions: Examples, ReLU and Sigmoid activation functions.
- Technical implementation: Mini-batch, stochastic gradient descent with momentum, dropout and L2 regularization, batch normalization, optimization of epochs.
- Optimization: We train the model using several Optimizers e.g. Adam, SGD, RMSprop e.t.c
- Framework: Keras with TensorFlow as backend.
- Environment: A GPU environment with a dual TITAN V 12GB memory cards.

Experiment

- Auxiliary classifiers: We disabled the auxiliary classifiers by inserting the DINET modules at different positions (top, middle, bottom) of our novel network.
- Image Size: We resized the images to 112×112pixels.
- Training: We used 100 epochs, minibatch =32

Architecture	Parameters	Accuracy
Inception	6,401, 870	83.86%
DINET V1	6,054,250	84.45%
DINET V2	5,856,098	86.12%
DINET V3	6,026,490	85.94%

Table 1: Evaluation performance of DINET Network

Discussion

- Task: We empirically evaluate our novel architecture DINET using CIFAR-10 Dataset.
- Observation 1: We find that the exclusion of auxiliary classifiers doesn't affect the network from the problem of vanishing gradient.
- Observation 2: Improved and increased interaction among learned feature maps i.e. tightly coupled feature maps leads to performance gains generated by our novel architecture.

Preliminary Results

The performance evaluation of our proposed architecture is presented in the following diagrams:

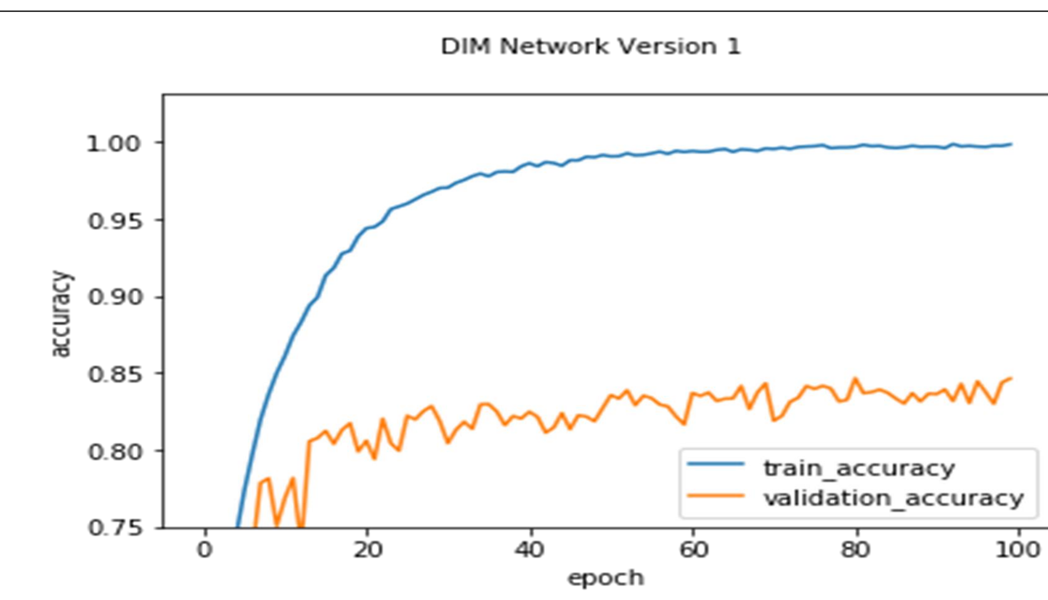


Figure 1: DINET Network version 1

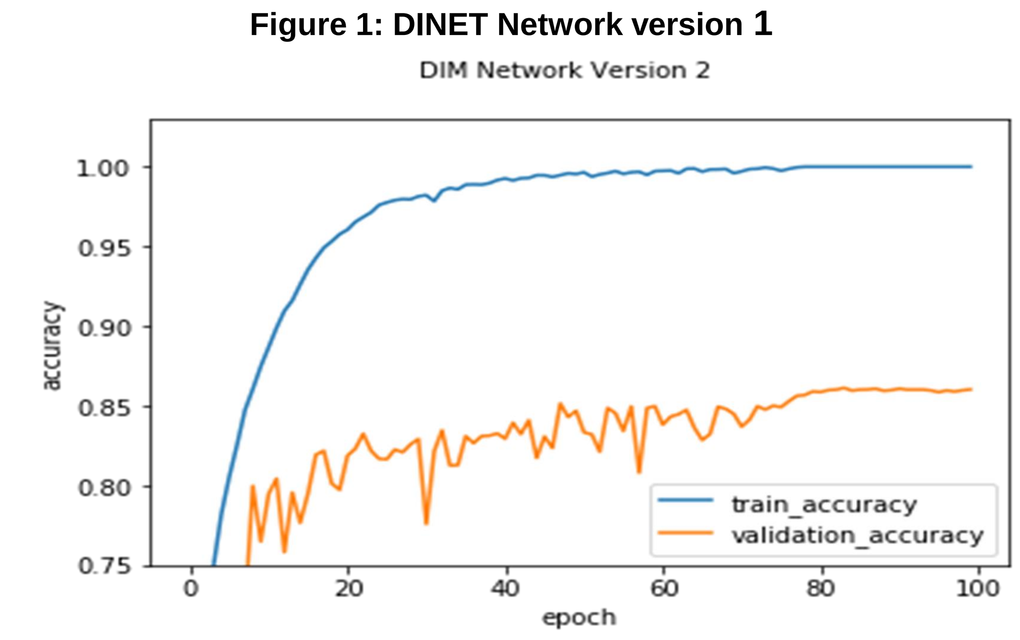


Figure 1: DINET Network version 2

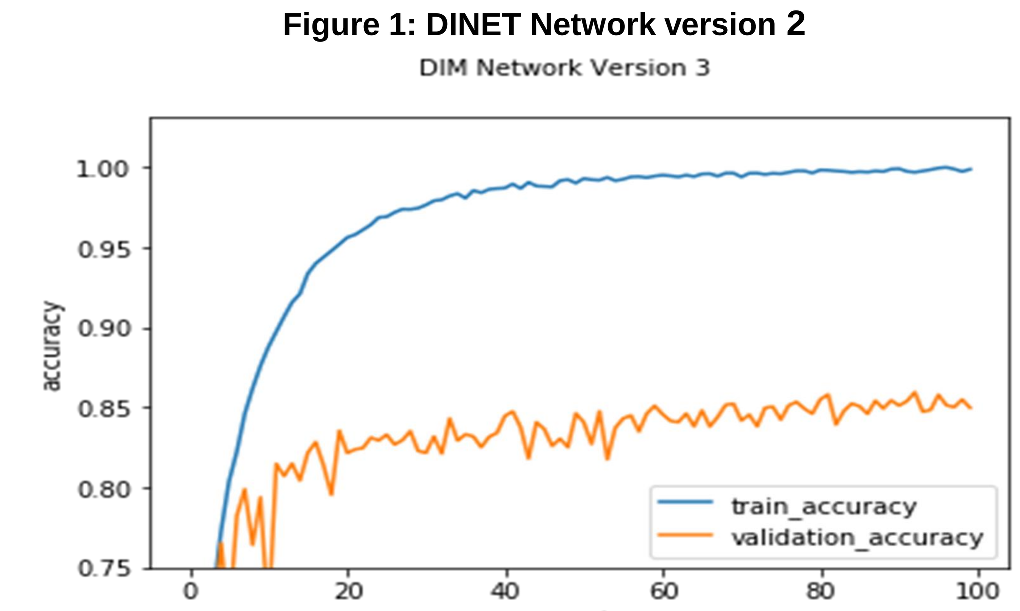


Figure 1: DINET Network version 3

Conclusion and Future Work

- We demonstrated the evaluation performance of our novel network on a general image dataset (CIFAR-10).
- Investigated the balance between model efficiency and parameter efficiency whilst mitigating the effect of vanishing gradient.
- Experiments show modest performance of using DINET modules over the traditional Inception network.
- NEXT Phase: Use Transfer learning techniques (Feature extraction and fine-tuning) and other methods such as data augmentation to investigate the effectiveness of using DINET on medical images, for example, the ChestX-ray 14 dataset (112,120 frontal-view chest X-ray images of 30,805 unique patients).
- In future: Use hyperparameter optimization, adding non-image patient data, finding optimal cut off layers, high resolution medical images using GANs. etc.
- Limitation: Small dataset; Inception V1 model.

References

1. Krizhevsky, A., I. Sutskever, and G.E. Hinton. Imagenet classification with deep convolutional neural networks. in Advances in neural information processing systems. 2012.
2. Simonyan, K. and A. Zisserman, Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556, 2014.
3. He, K., et al. Deep residual learning for image recognition. in Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.
4. Szegedy, C., et al. Going deeper with convolutions. in Proceedings of the IEEE conference on computer vision and pattern recognition. 2015.
5. Huang, G., et al. Densely Connected Convolutional Networks. in CVPR. 2017.
6. Wang, Q., G. Guo, and M.I. Nouyed, Learning Channel Inter-dependencies at Multiple Scales on Dense Networks for Face Recognition. arXiv preprint arXiv:1711.10103, 2017.
7. Greenspan, H., B. Van Ginneken, and R.M. Summers, Guest editorial deep learning in medical imaging: Overview and future promise of an exciting new technique. IEEE Transactions on Medical Imaging, 2016. 35(5): p. 1153-1159.
8. Wang, X., et al. Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. in Computer Vision and Pattern Recognition (CVPR), 2017 IEEE Conference on. 2017. IEEE.
9. Krizhevsky, A. and G. Hinton, Learning multiple layers of features from tiny images. 2009, Citeseer.

Acknowledgments

We would like to extend our gratitude to the Midwest Big Data Hub Organizers, and Dakota State University for the facilitation and support of this research work.