Project Proposal for COMS 4995: Neural Networks and Deep Learning

Xiaofei Gao
Columbia University
New York, NY 10027
xg2429@columbia.edu

Hellen Zhao
Columbia University
New York, NY 10027
hz2843@columbia.edu

Thea Zhu
Columbia University
New York, NY 10027
wz2636@columbia.edu

Abstract

1 Introduction

2 Related Work

Convolutional Neural Networks in Image Classification Convolutional Neural Networks (CNNs) are proved to be very effective in the image classification tasks. Reviewing the winners of the ImageNet Large Scale Visual Recognition Challenge (ILSVRC), we can see that almost all of them used CNNs. Currently, the populat CNNs with lower error rates are VGGNet, GoogLeNet/Inception V1-V4, ResNet, EfficientNet, etc [1]. Among them, Inception V4 reached a top-1 error rate of 16.4% on the ImageNet dataset [2], while EfficientNet V2 reached a top-1 error rate of 12.7% on the ImageNet dataset [3].

Transfer Learning Training a model from scratch is time-consuming and requires a lot of computing resources. By reusing a pre-trained model, we only need to fine tune the model on a smaller dataset to solve a new problem. There are various techniques we can apply during transfer learning, such as freezing certain layers in the pre-trained model, or fine tune the layers.

Transfer Learning in Image Classification In image classification, transfer learning is very helpful because there're a lot of available models that were pre-trained on ImageNet dataset. Kolesnikov et al. (2020) introduced Big Transfer (BiT) [4], a simple recipe that combined selected components and achieved strong performance on over 20 datasets. There are many works have been done applying transfer learning into different image classification tasks, such as food image classification [5] and dog breed classification [6]. In particular, it is widely used in the medical image classification tasks. For example, Davila et al. (2024) evaluated eight fine-tuning strategies based on three pre-trained models across different types of medical images [7], and found the pre-trained model and fine-tuning method need to be chosen based on the target dataset.

Preprint. Under review.

3 Methodology

References

[1] Chen, Leiyu & Li, Shaobo & Bai, Qiang & Jing, Yang & Jiang, Sanlong & Miao, Yanming. (2021). Review of Image Classification Algorithms Based on Convolutional Neural Networks. Remote Sensing. 13. 4712. 10.3390/rs13224712.

[2] Szegedy, Christian & Ioffe, Sergey & Vanhoucke, Vincent & Alemi, Alexander. (2016). Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning. AAAI Conference on Artificial Intelligence. 31. 10.1609/aaai.v31i1.11231. Review of Image Classification Algorithms Based on Convolutional Neural Networks. Remote Sensing. 13. 4712. 10.3390/rs13224712.

[3] Tan, M., & Le, Q. (2021, July). Efficientnetv2: Smaller models and faster training. In International conference on machine learning (pp. 10096-10106). PMLR.

[4] Kolesnikov, A., Beyer, L., Zhai, X., Puigcerver, J., Yung, J., Gelly, S., & Houlsby, N. (2020). Big transfer (bit): General visual representation learning. In Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part V 16 (pp. 491-507). Springer International Publishing.

[5] Burkpalli, Vishwanath & Patil, P.C.. (2020). An efficient food image classification by inception-V3 based cnns. International Journal of Scientific and Technology Research. 9. 6987-6992.

[6] Agarwal, Ambuj & Kiran, Vidhu & Jindal, Rupesh & Chaudhary, Deepak & Tiwari, Raj. (2022). Optimized Transfer Learning for Dog Breed Classification. 10.13140/RG.2.2.35522.04807.

[7] Davila, Ana & Colan, Jacinto & Hasegawa, Yasuhisa. (2024). Comparison of fine-tuning strategies for transfer learning in medical image classification. Image and Vision Computing. 146. 105012. 10.1016/j.imavis.2024.105012.