Image Classification Optimization

How to improve image classification model performance from data augmentation and model structure?

Presenters: Mengzhi Zhou, Guanzhi Wang

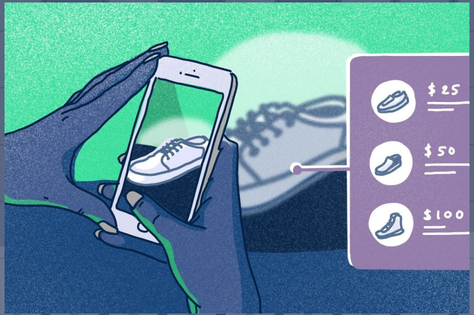
Introduction

- Definition
- Applications
- Methods
- Problem

- The task of extracting information classes from a multiband raster image.

- A process in computer vision that can classify an image according to its visual content.

- Face ID
- Instagram filters
- Visual search
- Etc...



- SVMs
- Decision trees
- Neural networks
- Bayesian, Markov chain Monte Carlo, etc…

- How to improve image classification model accuracy while keeping fast-paced data processing speed?

Related Work

- Anne Bonner gave a basic overview for image classification and CNN. She talked about the foundation of CNN and image classification, logics behind models, basic structure of CNN, etc.
- Le and James presented a performance comparison between image classification models with MNIST dataset.
- Wang and Luis illustrated an experiment to compare impact of different data augmentation techniques: traditional transformation, generative adversarial networks, and neural net augmentation.

- Fong and Andrea mentioned photographer bias, which can strongly affect neural network performance. They discussed about occlusion augmentation techniques: batch augmentation, joint training and dropout method, and set up an experiment to compare those occlusion augmentation effects.

- Vailaya, Figueiredo, Jain and Zhang modified the binary Bayesian classifiers that achieved classification accuracy that higher than 90% for all four classification problems with database of 6931 photos.

- McCann and Lowe created the Local Naive Bayes Nearest Neighbor.
- Zhou, Cui, Li, Liang and Huang discussed about the hierarchical Gaussian mixture model whose parameters are learned in a Bayesian framework that is useful in image classification .

Data

- From Kaggle
- 25,000 RGB images
- Two classes: cat and dog
- Labeled

- Irrelevant images dropped
- Random subset method applied
- Use GPU rather than CPU

Methods

Horizontal and vertical comparison among image classification models

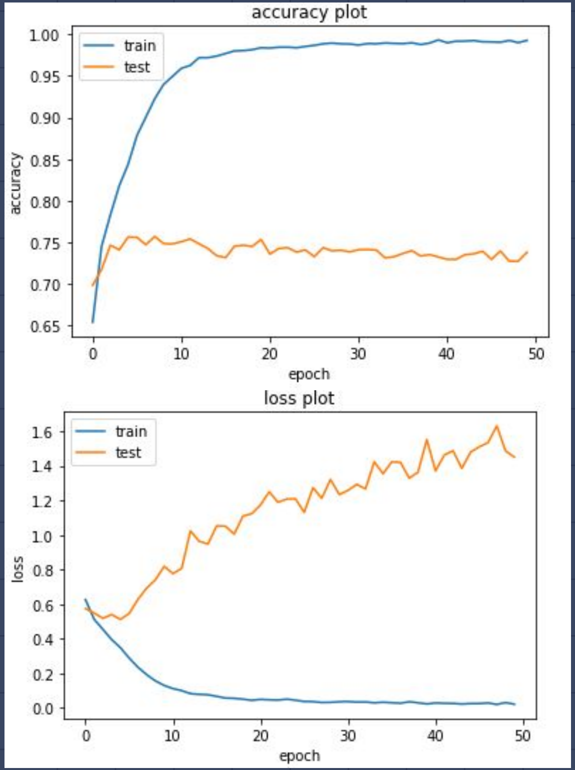
5 Different Image Classification Models

- Baseline model: 1 convolution layer + no data augmentation process
- DA model: 1 convolution layer + data augmentation process
- 3CNN model: 3 convolution layers + no data augmentation process
- 3CNN + DA model: 3 convolution layers + data augmentation process
- Pretrained model: VGG16 image classification model (ImageNet)

1. Baseline Model

- Model Structure:

  1 convolution layer with 32 filters, relu activation function, 0.3 dropout rate, adam as optimizer
- No data augmentation process



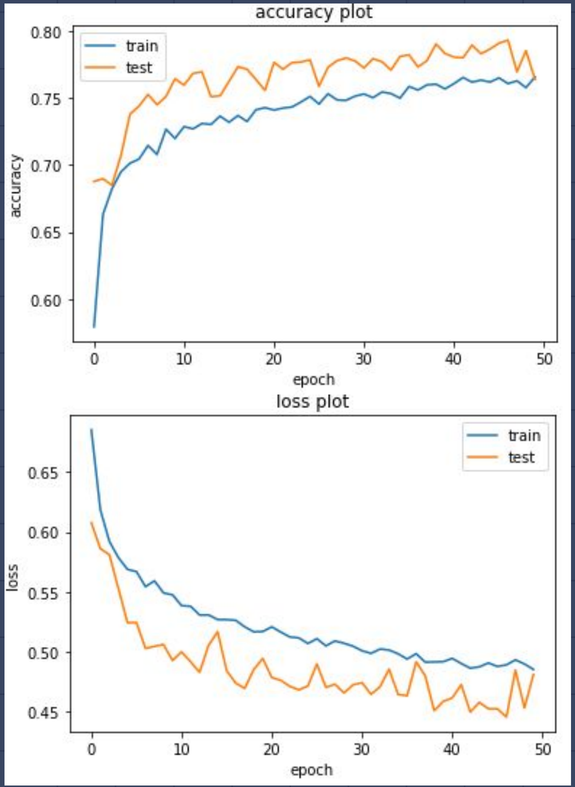
2. DA Model

- Model Structure:

  Same as baseline model

- Data augmentation process:

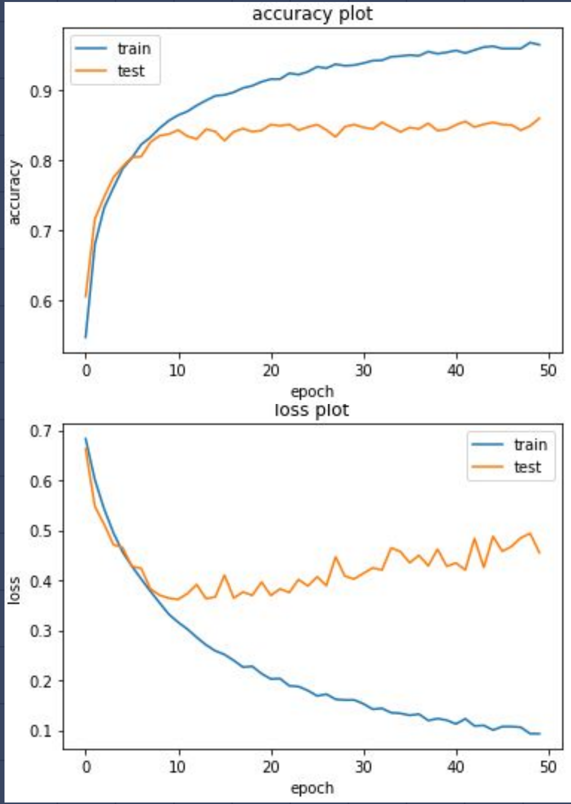
  Rotation, width and height shift, shearing, horizontal flip and zoom



3. 3CNN Model

- Model Structure:

  3 convolution layers with up to 128 filters, relu activation function, 0.3 dropout rate, adam as optimizer
- No data augmentation process

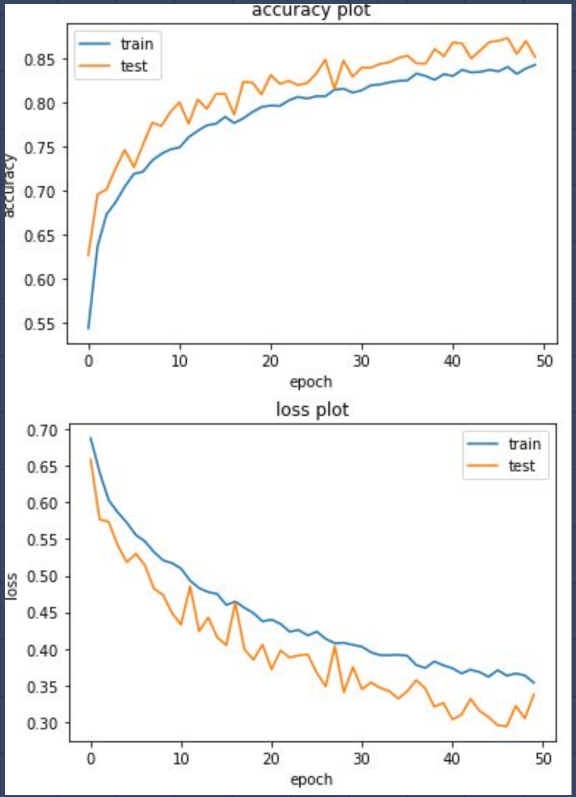


4. 3CNN + DA Model

- Model Structure:

  Same as 3CNN model
- Data augmentation process:

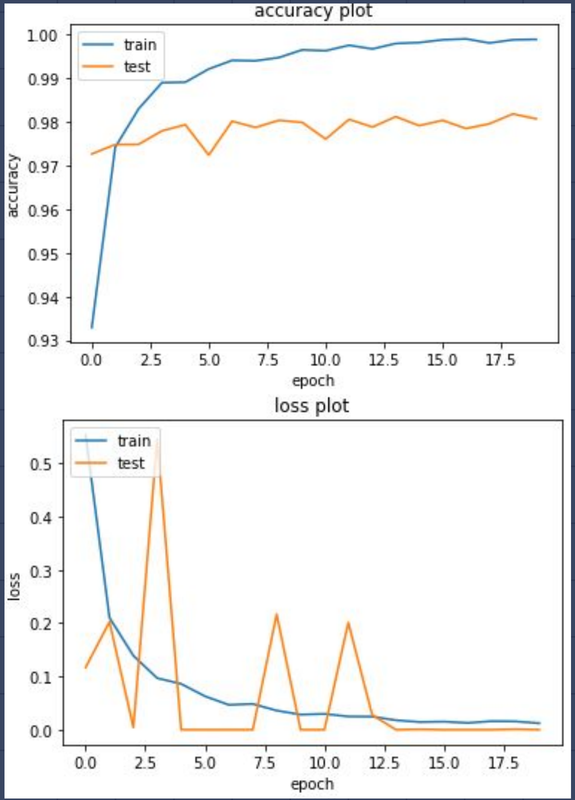
  Same as DA model



5. Pretrained Model

- VGG16 model (transfer learning)
- Based on ImageNet dataset:

  Include over 1000 categories with animal, plant, sports, material, food, etc.



Overall Performance Comparison

- Except pretrained model, 3CNN+DA model has the best overall performance
- Pretrained model has highest accuracy among 5 image classification models

TABLE V
PERFORMANCE EVALUATION

	Convolutional Neural Network				
	Baseline	***DA***	***3CNN***	***3CNN+DA***	***VGG16***
Accuracy(val)	0.7382	0.7637	0.8590	0.8522	0.9818
Loss(val)	1.4504	0.4806	0.4552	0.3373	0.2631
AUC score	0.7382	0.7637	0.8590	0.8522	-
Time(mins)	13.52	22.59	14.03	22.54	23.47

*DA:Data Augmentation

Conclusion

What is the effect of data augmentation and neural network structure on model performance?

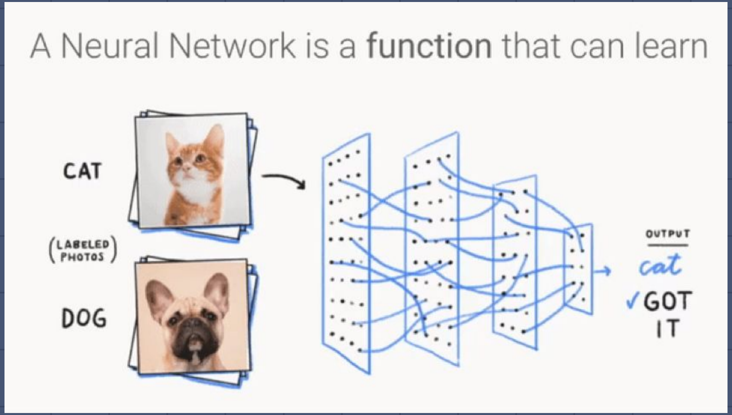
- Data augmentation process could avoid overfitting problem but substantially slow down processing speed of model

- As complexity of neural network structure increase, model gets stronger computation ability but also greater potential risk for overfitting problem.

The cooperation between data augmentation and neural network structure is significant for image classification model performance!

Future Research

- Parameter tuning: filter number, dropout rate, etc.
- Improve neural network structure
- Developing Bayesian framework
- Different optimizer comparison: adam, SGD, ADA, etc.
- Explore effect of each data augmentation components separately



Reference

1. Buyya, R., Calheiros, R. N., Dastjerdi, A. V. (Eds.). (2016). Big data: principles and paradigms. Morgan Kaufmann.
2. Vailaya, A., Figueiredo, M. A., Jain, A. K., Zhang, H. J. (2001). Image classification for content-based indexing. IEEE transactions on image processing, 10(1), 117-130.
3. Bonner, Anne. "The Complete Beginner's Guide to Deep Learning: Convolutional Neural Networks." Medium, Towards Data Science, 1 June 2019, towardsdatascience.com/wtf-is-image-classification8e78a8235acb.
4. Le, James. "The 4 Convolutional Neural Network Models That Can Classify Your Fashion Images." Medium, Towards Data Science, 7 Oct. 2018, towardsdatascience.com/the-4-convolutional-neural-networkmodels-that-can-classify-your-fashion-images-9fe7f3e5399d.
5. Wang, Jason, and Luis Perez. "The Effectiveness of Data Augmentation inImageClassificationUsingDeepLearning."TheEffectivenessofData Augmentation in Image Classification Using Deep Learning, 13 Dec. 2017.
6. Fong, Ruth, and Andrea Vedaldi. "Occlusions for Effective Data Augmentation in Image Classification." 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW), 23 Oct. 2019, doi:10.1109/iccvw.2019.00511.
7. McCann, S., Lowe, D. G. (2012, June). Local naive bayes nearest neighbor for image classification. In 2012 IEEE Conference on Computer Vision and Pattern Recognition (pp. 3650-3656). IEEE.
8. Zhou, X., Cui, N., Li, Z., Liang, F., Huang, T. S. (2009, September). Hierarchical gaussianization for image classification. In 2009 IEEE 12th International Conference on Computer Vision (pp. 1971-1977). IEEE.
9. "Dogs vs. Cats Kaggle", Kaggle.com, 2020. [Online]. Available: https://www.kaggle.com/c/dogs-vs-cats/overview.
10. Sarkar. (2018, November 17). A Comprehensive Hands-on Guide to Transfer Learning with Real-World Applications in Deep Learning. Retrieved from https://towardsdatascience.com/a-comprehensive-handson-guide-to-transfer-learning-with-real-world-applications-in-deeplearning-212bf3b2f27a

THANKS!

Any questions?