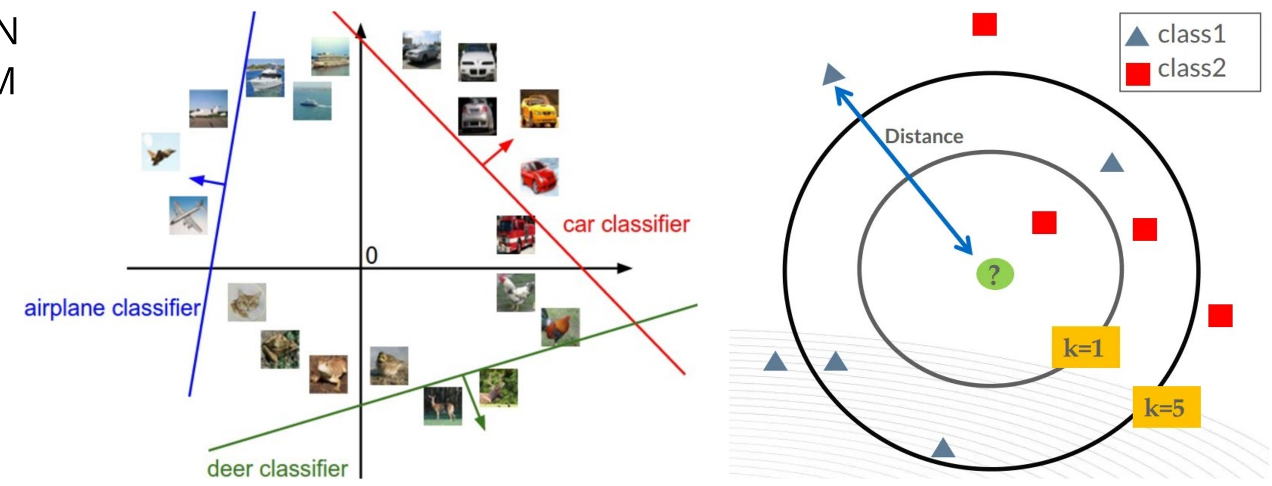
Medical Image Classification Model Based on CNN

Wei Deng, Kang Zhu
Beijing University of Posts and Telecommunications

Image Classification

Image classification is one of the basic research topics in the field of artificial intelligence. There are many classical image classification algorithms:
- KNN
- SVM

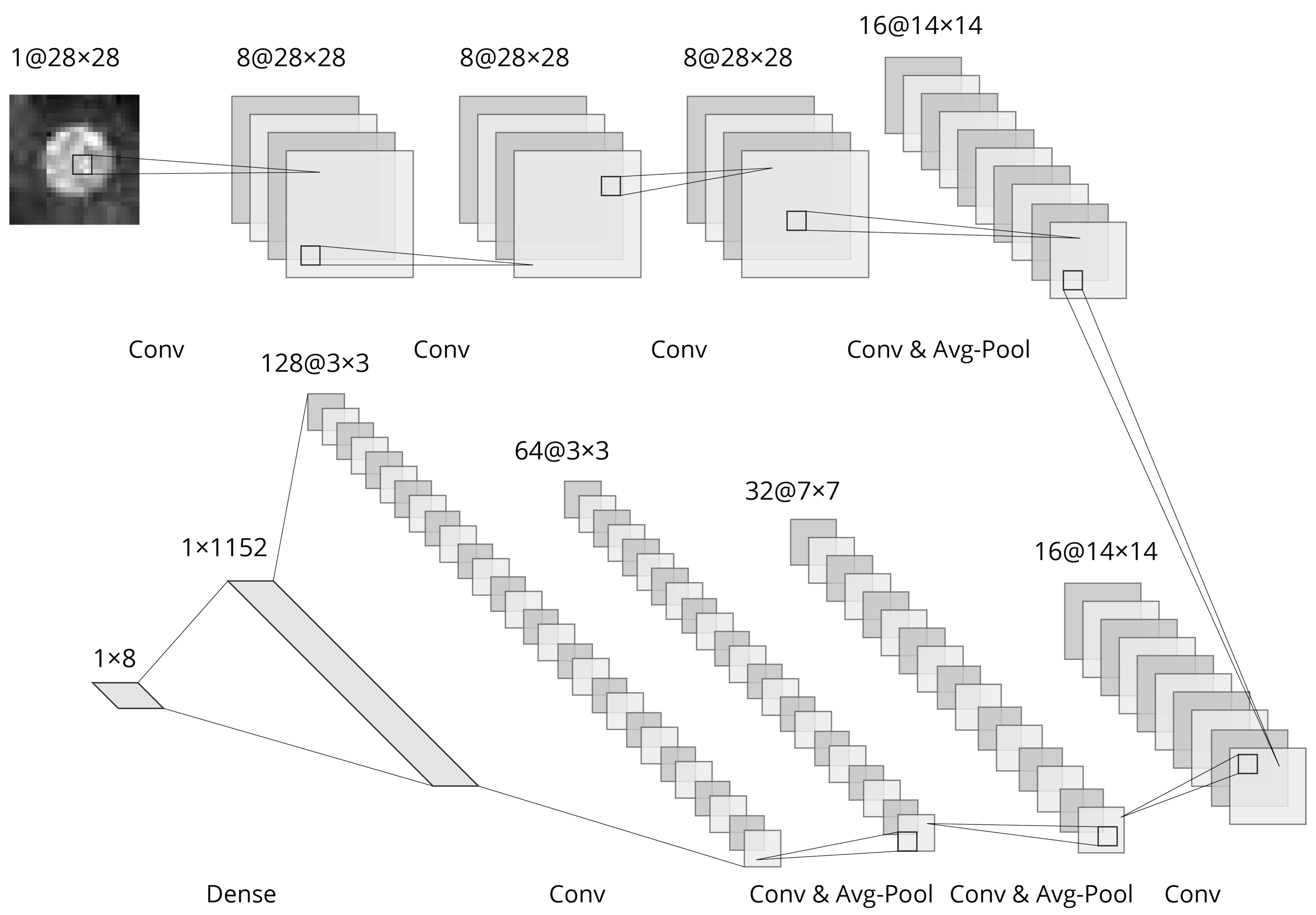


Limitation of KNN&SVM:
- KNN has a large amount of computation and memory overhead
- SVM performance mainly depends on the selection of kernel function, and doesn't suitable for large-scale training samples

Medical Image Classification Model Based on CNN

In image classification tasks, CNN is a benchmark method. Andre Woloshuk[1], based on CNN, proposes a convolutional neural network called NephNet2D, which performs well in medical image classification.

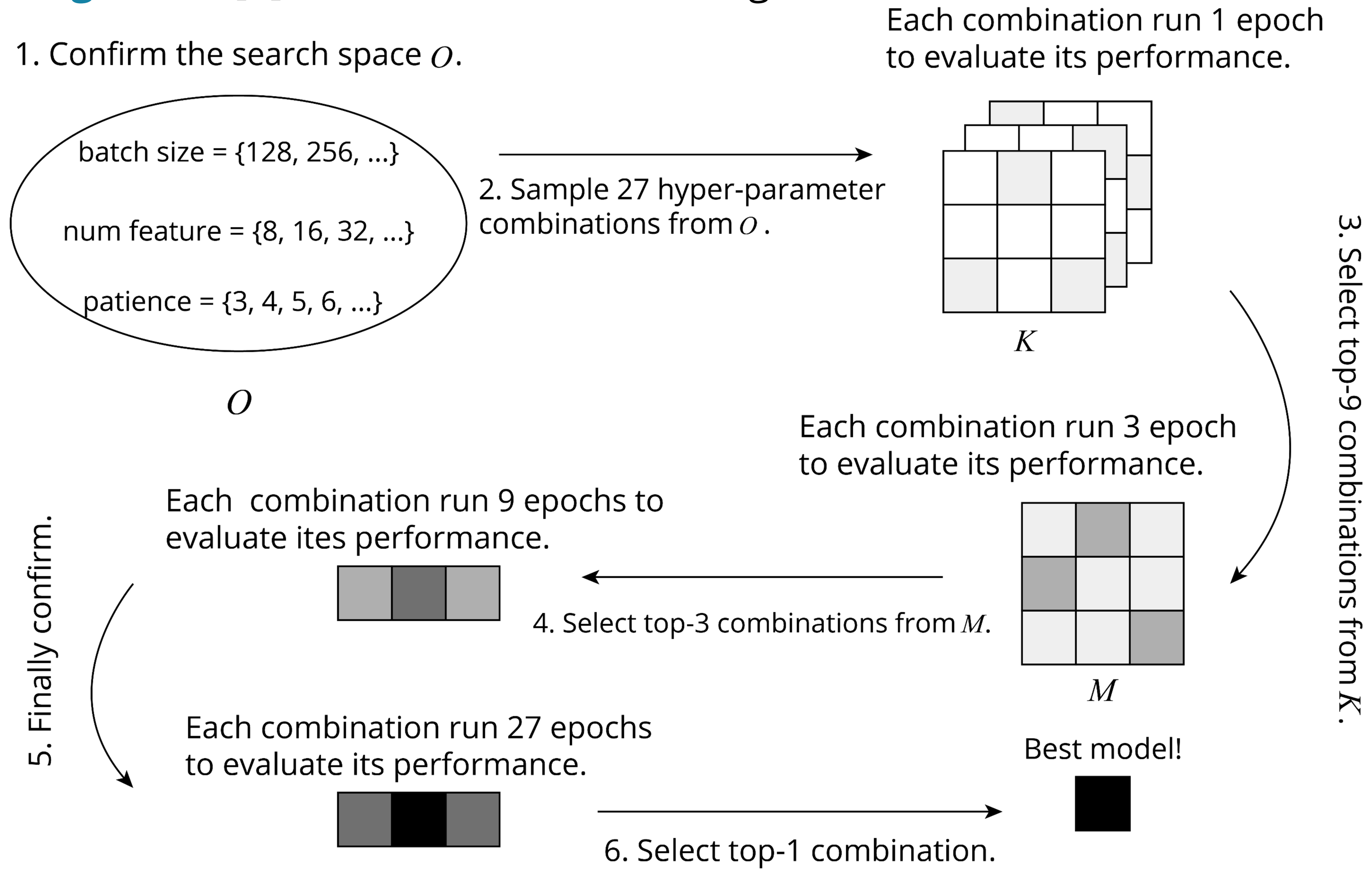
Based on **NephNet2D**, we proposed a medical image classification model.



- Filter: size=3×3, stride=1, padding=1
- Average pooling: size=2×2, stride=2
- 2D-Batch normalization is used
- Activation: ReLU
- Dropout is used

The Trick of Hyperparameter Optimization

In order to improve the efficiency of hyperparameters adjustment, we utilize an automatic hyperparameters regulator based on **Hyperband Algorithm**[2]. It works as the following flow:



With the automated hyperparameter regulator, we can:
- Make hyperparameter selection adjustments within the range we expect
- Under the condition of limitated number of iterations, find the satisfactory hyperparameters

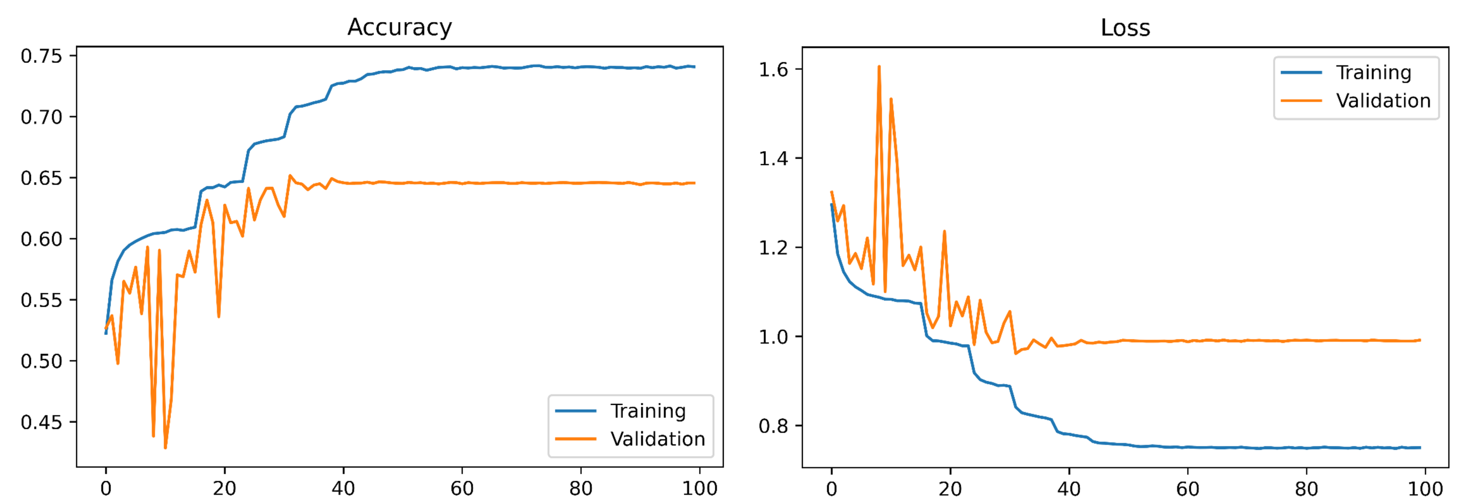
Experimental Environment and Setup

- PyTorch deep-learning framework
- Apex library for mix precision training to accelerate the training and reduce memory usage
- Using a subset of the whole dataset(10,000 samples) for adjusting the model roughly
- 80%(120,000) of dataset for training and 20%(30, 000) of dataset for validation. Test dataset contains 50,000 pieces of data.
- Scikit-learn library for metric computing

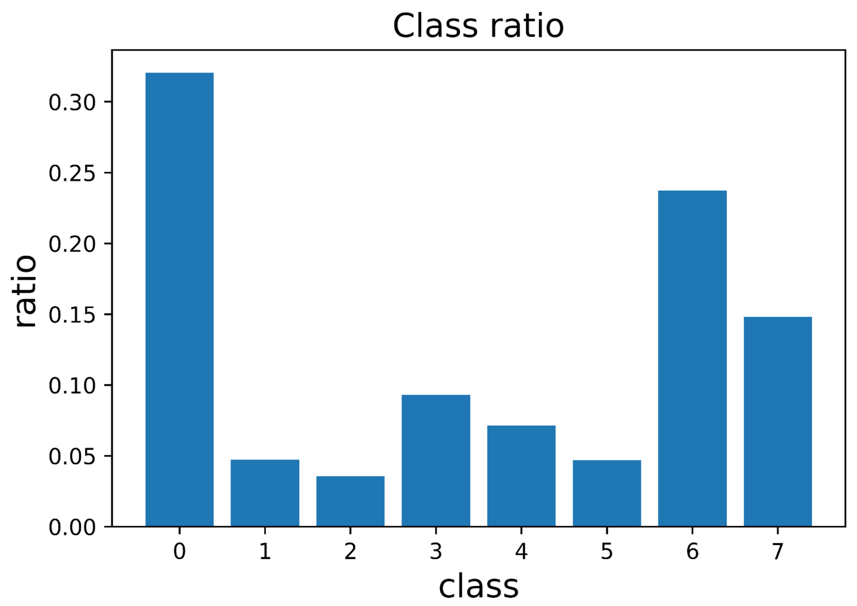
- Optimizer: SGD with momentum=0.8
- Batch size: 128
- Weight decay: 0.004
- Initial learning rate: 0.025
- Dropout rate: 0.5

- Learning rate scheduler: Reduce learning rate when validation metric stopped improving, with hyperparameter patience=5 and factor=0.29.

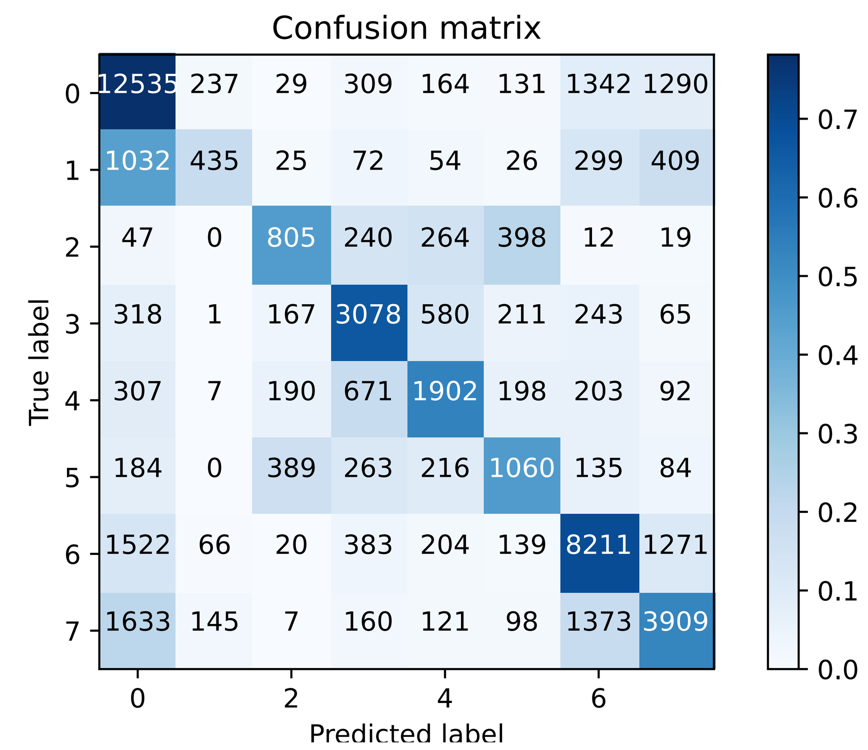
Training Results



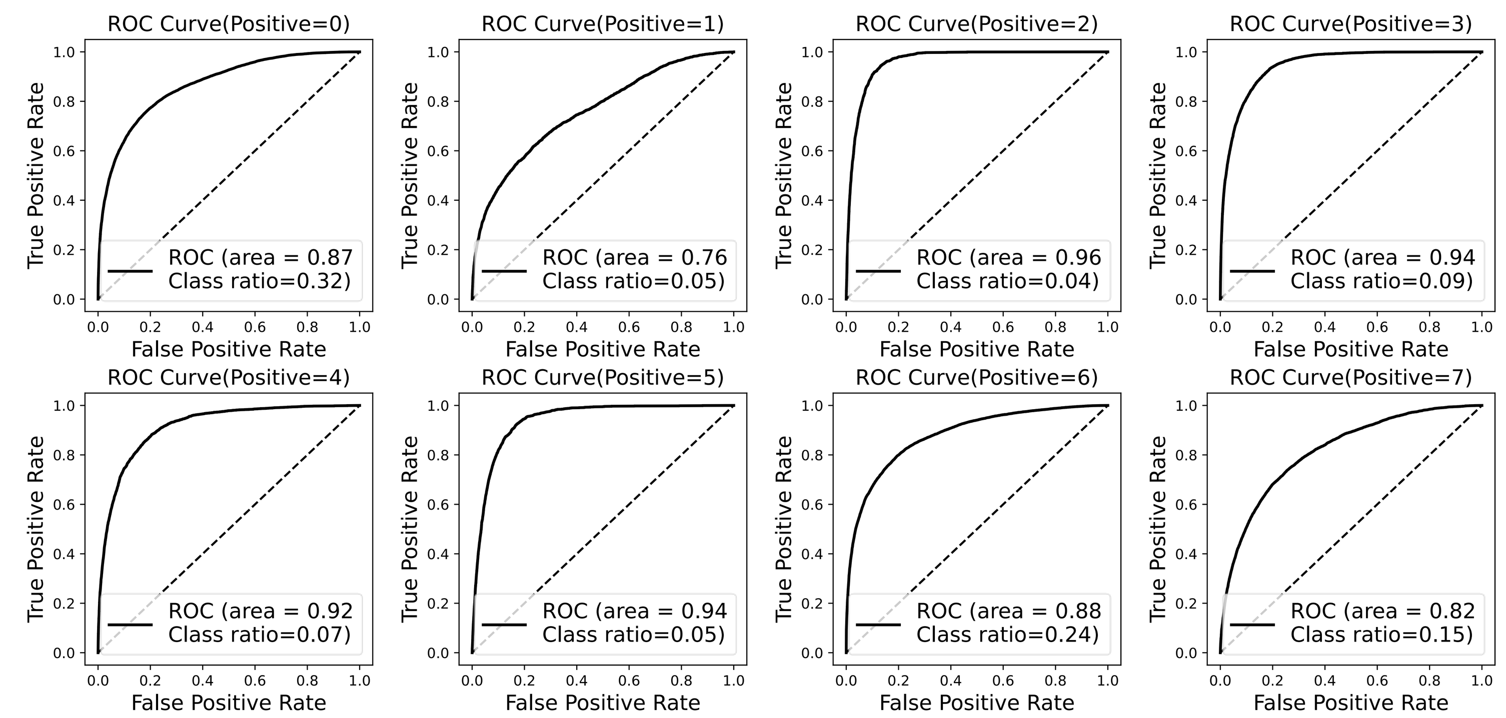
Testing Results



Metric	Result
Accuracy	0.6387
F1-Score	0.5426
Recall	0.5356
Precision	0.5678



- For each class, regard itself as positive and others negative
- Compute macro-average of F1-Score, recall and precision to evaluate the whole performance
- For each class, illustrate the ROC curve and compute AUC



Reference

[1] A. Woloshuk et al. In situ classification of cell types in human kidney tissue using 3dnuclear staining.Cytometry Part A, 99(7):707–721, 2021
[2] Lisha Li∗, Kevin Jamieson∗∗, Giulia DeSalvo†, et al. Hyperband: Bandit-Based Configuration Evaluation for Hyperparameter Optimization.