



Natural Scene Images Classification

Group 2: Sen Fu, Pengkun Li, Yifan Luo, Yu Shi

Abstract

Our project aims to build and apply several image classification models in the computer vision field. Model inputs are the images in six different scenes and output is a classification category of the input images. Our group proposed four different models, from which CNN model proves to be the best one when comparing the errors in both training and testing set with other three. The best CNN model could be applied to unmanned aircraft, autonomous vehicles or other cases to identify the surrounding environment and then change the driving or controlling models.

Backgrounds

- Dataset is collected from Intel® Software on AnalyticsVidhya website, it contains 14034 images in training set and 3000 images in testing set. Each figure is recognized as one of the six categories: mountain, sea, glacier, building, street and forest.
- Dimension: Training set: [14034, 150, 150, 3], Test Set:[3000, 150, 150, 3]
- Features: color, texture, location. Moreover, images are in RGB color and normalized to a float matrix with value between [0,1].

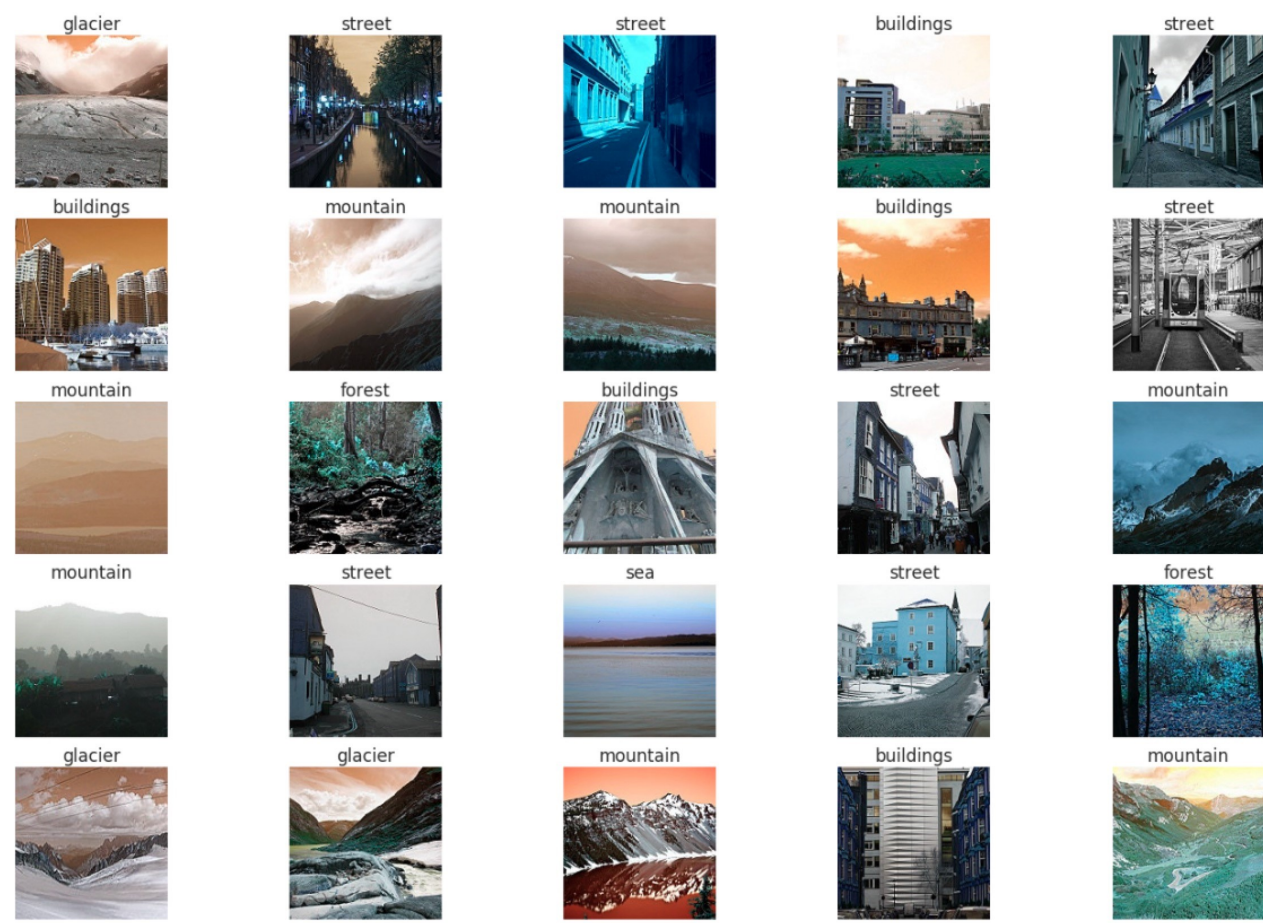


Fig.1 Examples of scene images dataset.

Models

1. Artificial Neural Network

- Reshape and standardize the images before feeding to network
- 20 layers: Input \rightarrow 128 \rightarrow 64 $\rightarrow \dots$ (64) $\dots \rightarrow$ 64 \rightarrow Output, activation function is ReLU for hidden layers, Softmax for output layer.

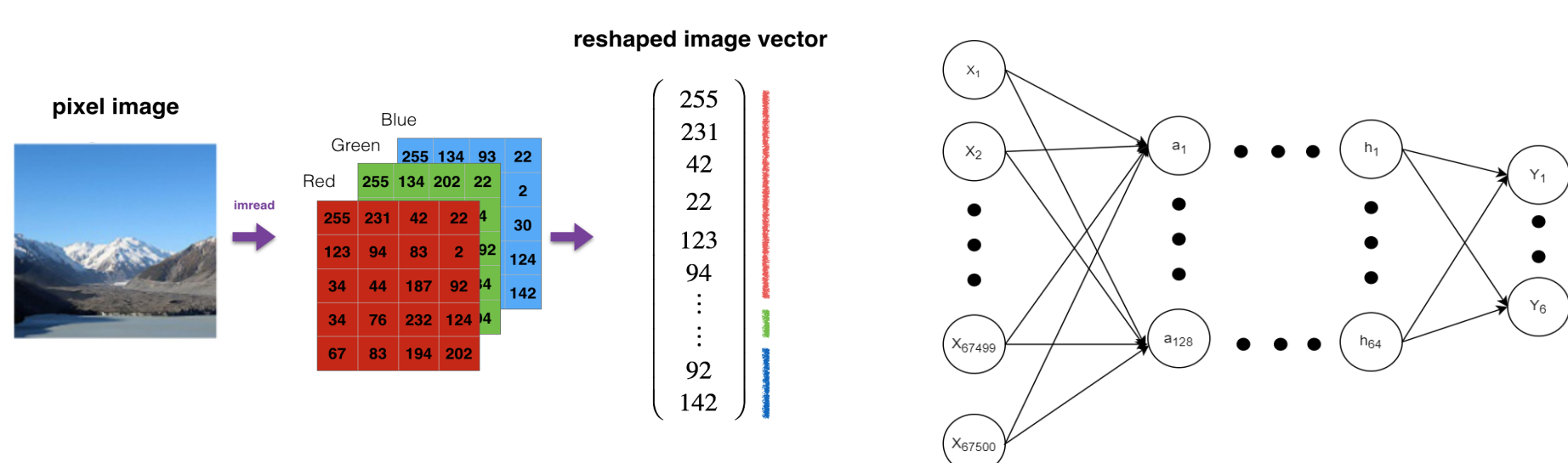


Fig.2 Data preprocessing and sample model structure

2. Random Forest

3. Support Vector Machines

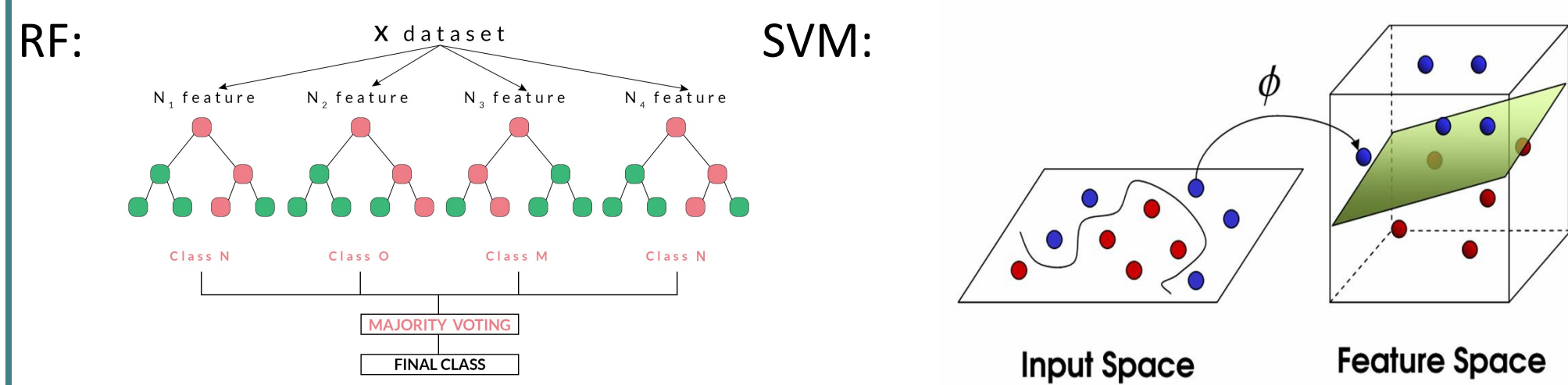


Fig.3 Sample Model structure of Random Forest and Support Vector Machines

4. Convolutional Neural Network

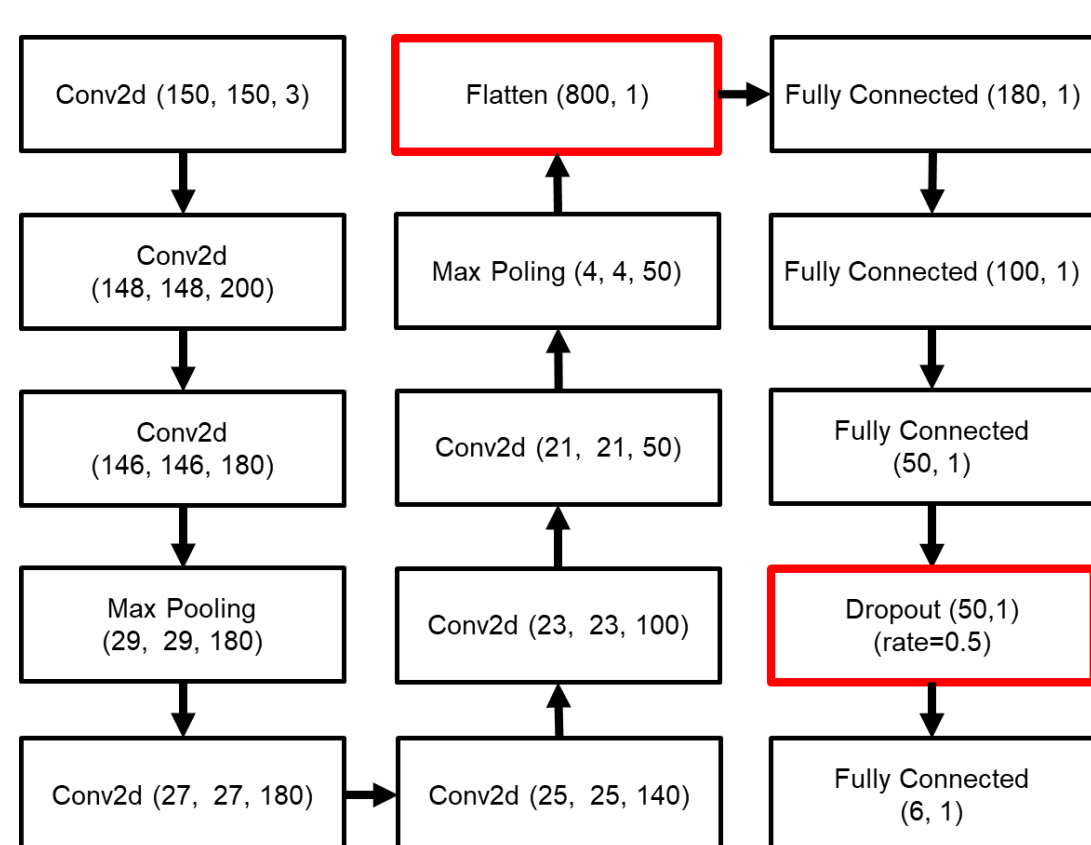


Fig.4 Model structure of Convolutional Neural Network

Results

1. Data Distribution

- The sample numbers and proportion of each categories in both training and testing set are shown below.

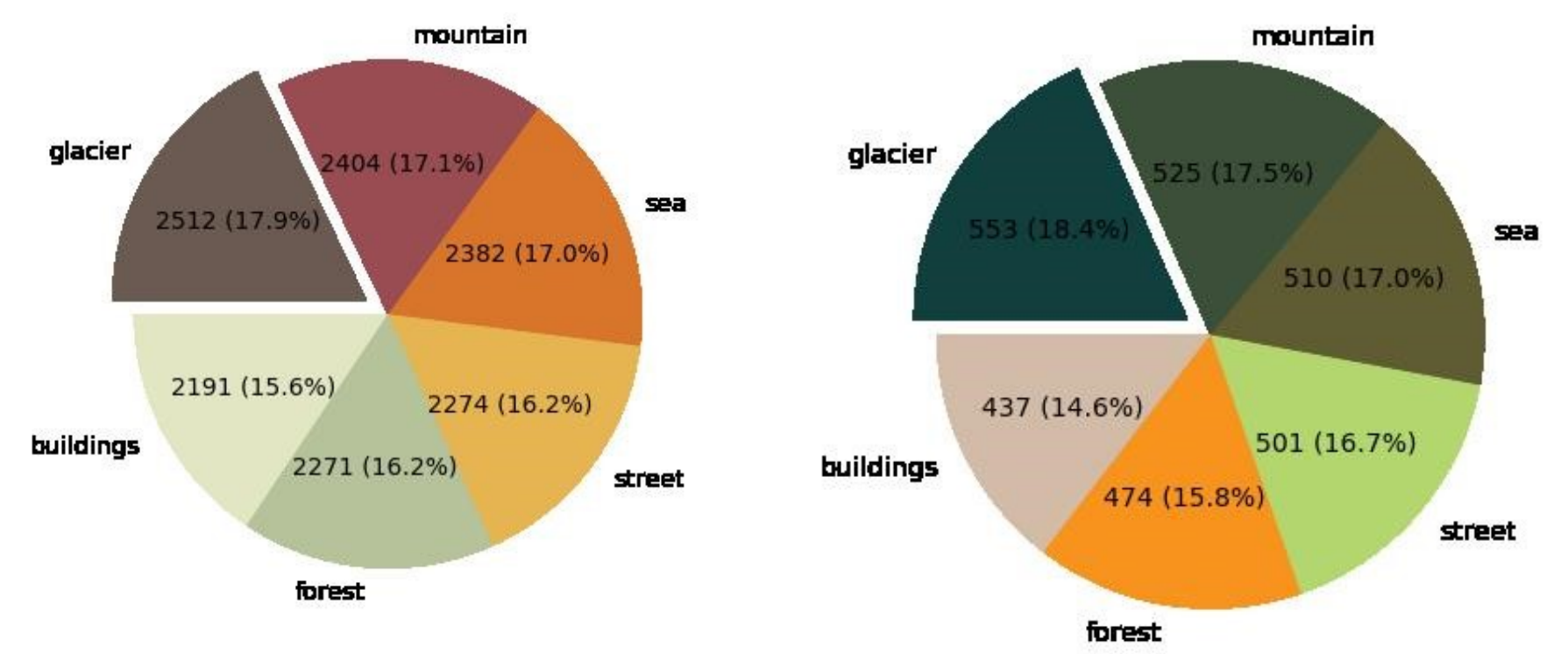


Fig.5 Sample number and proportion in training and testing set

2. Comparison with Mean Square Error among models

Table 1. Train and test error among 4 models

Index	Model Name	Train Error	Test Error
1	Artificial Neural network	29.97%	43.17%
2	Random Forest	0.072%	40.03%
3	Support Vector Machine	35.8%	50.4%
4	Convolutional Neural Network	7%	14.54%

3. Result metrics figures of CNN

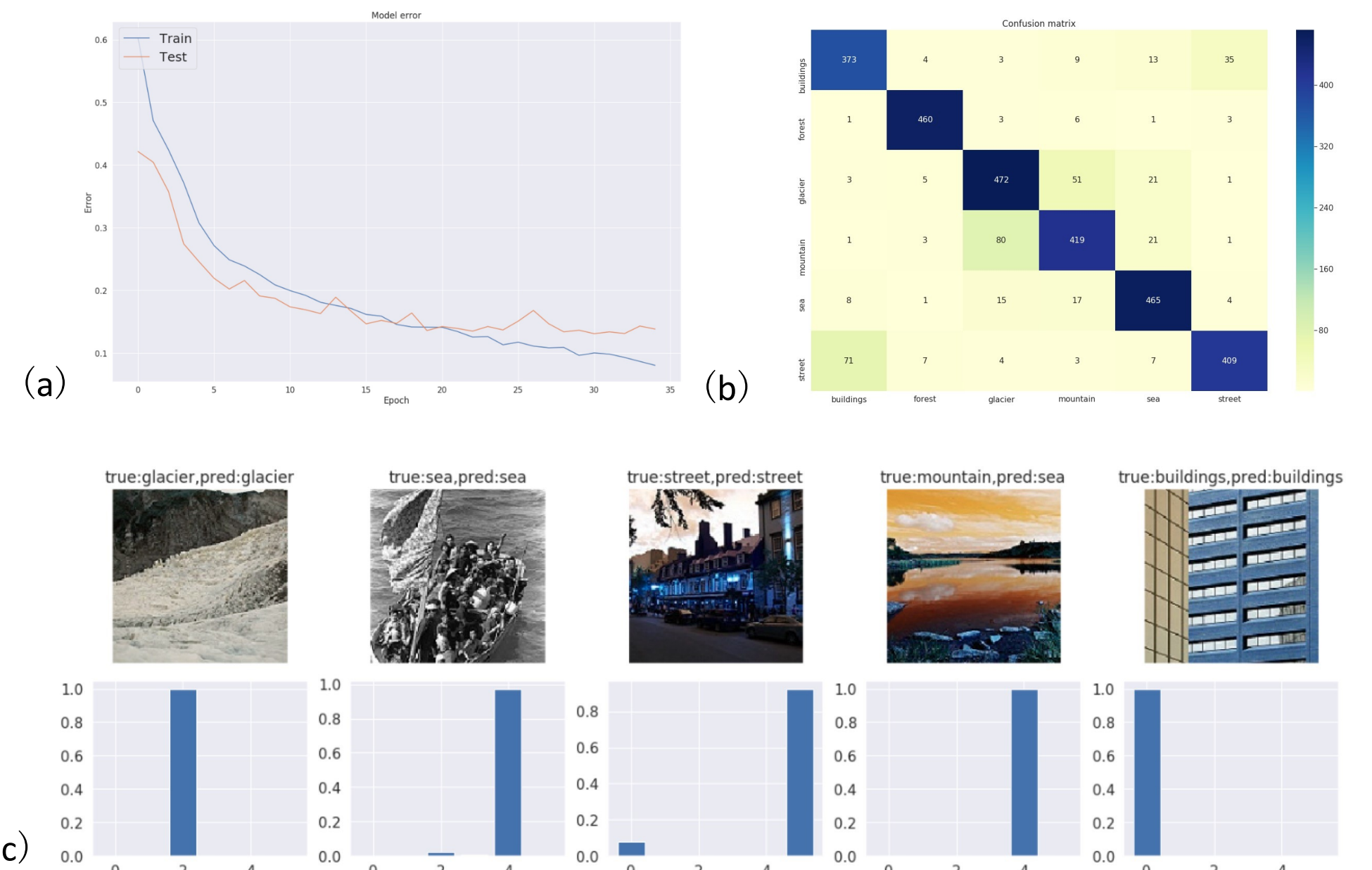


Fig.6 (a) Model error vs. epochs; (b) confusion matrix; (c) prediction examples

Discussion

- As shown above, four trained models can be evaluated and selected based on the accuracy of image classification. Supported by the excellent performance of feature extraction of CNN, we tried to adjust the parameters of CNN to achieve a lower error in both training and testing set.
- In our expectation, the convolutional neural network should have the best accuracy or lowest error. However, in our first attempt, the performance of CNN is worse than Random Forest. After the parameters adjustment, we overcome the problem of overfitting which causes the higher error of CNN.
- The final result with the selected CNN model error in testing set achieves 14%, which is lower than our goal with 10%. It is caused by the similarity among different categories, such as sea and glacier, or street and building. The error could be decreased by increasing the data size and construct a more advanced CNN structure.

Future Works and References

- Feature extraction engineering will be introduced, in our works above, features are extracted by our model but not selected manually, which will improve the performance of model. Also, other models like KNN will be introduced for comparison.
- References:
 - [1].Leo, Marco, et al. "Computer vision for assistive technologies." *Computer Vision and Image Understanding* 154 (2017): 1-15.
 - [2]. Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." *Advances in neural information processing systems*. 2012.