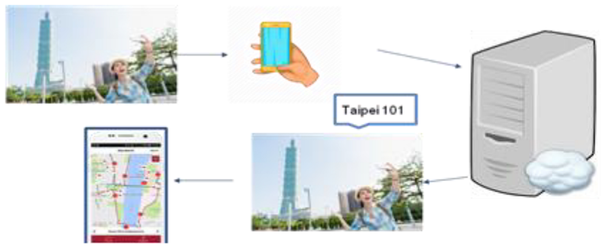
AN ATTENTION BASED-CONVOLUTIONAL NEURAL NETWORK FOR LANDMARK RECOGNITION IN ASIAN REGION

S. Perera and A. Ramanan
Department of Computer Science, Faculty of Science, University of Jaffna
shehanperera.office@gmail.com, a.ramanan@univ.jfn.ac.lk

Introduction

Landmark recognition is very helpful in many ways. Tourist can find out new attractive locations from social media and they can plan for a visit. Landmark recognition may also increase the interests of tourists to visit certain places through which it can contribute to the economy of that country.Also landmark recognition greatly helps people to better understand and organise their photo collections.



Objective

To build an intelligent system for recognising landmarks in the Asian region to improve tourism and business near to landmarks.

Dataset

- The Asian landmark dataset chosen from the **Google Landmark Recognition Challenge** dataset [4]. We choose 30 different landmarks from 30 different Asian countries.
- We used 35 images per landmark, for training and 15 images for testing the model.

Methodology

- At the initial stage of this study, we have tested different classifiers on landmark classification and found CNN to outperform SVM, k-NN and Random Forests (See Table 1).
- We fine-tuned the pre-trained VGG-11 model to obtain better results for the Google Landmark Recognition dataset. In fine-tuning process we note that using Optimizer as Adamax, Learning rate as 0.001, Loss function as Cross Entropy Loss give better result. Since we have small number of data we used 100 epochs.
- We also add an attention branch to the framework that combines the predicted *conv5* features and fine grained features to gate or magnify the *conv5* features to improve the precision of landmark classification (See Figure 1).
- We utilize the 14×14×512 predicted feature map of *conv5* to max pool the features in to 1×1×512. After that we used one convolution layer to obtain different number of features. From several trials we found that 1×1×1024 is the best (See Table 2). Thereafter the features are reconstructed through upsampling to yield 14×14×1024, thus producing dense feature map 14×14×(512+1024). The proposed method shows 94% of classification accuracy by contributing the CNN to yield attentive image features (See Table 2).

Methodology

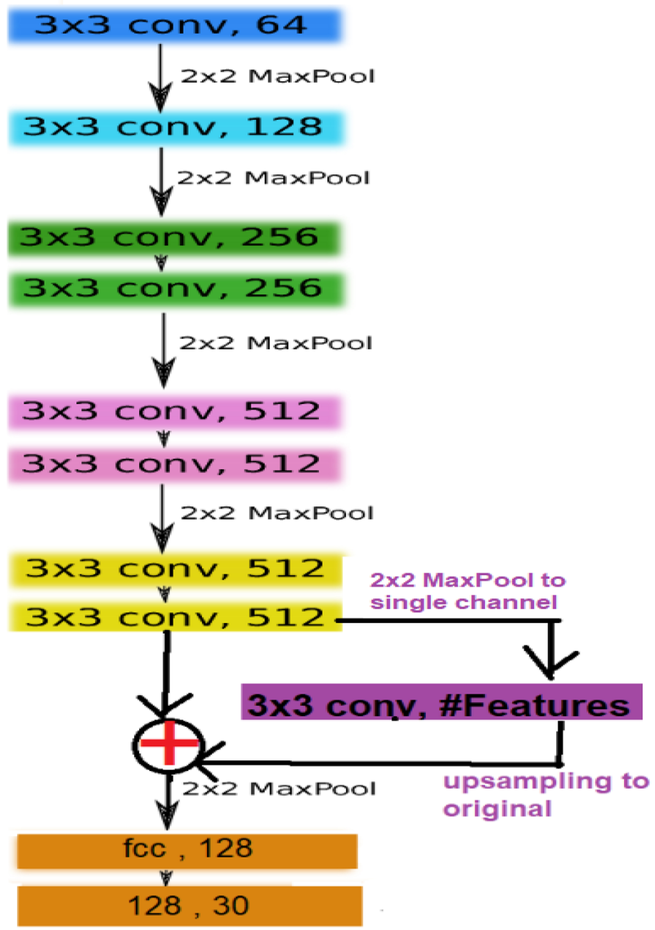


Figure 1: Modified VGG-11 Model

Conclusion

In this study we showed a modification to VGG-11that can recognize Asian landmarks with 94% classification accuracy. Table 1 shows that basically deep learning techniques like CNN works better on image classification rather than shallow learning methods. Even from several CNN models selected in this study VGG-11 model performs better in this landmark dataset. Adding attention branch for original network extracts more important features using max pooling. These important features combined with original features works to improve the knowledge of the network.

Test Results

Table 1: Performance comparison of different classifiers on Landmark classification

Classifier	Classification rate
Nearest Neighbor	35%
Random Forest	42%
SVM	70%
CNN	78%
Fine tune CNN	90%
Proposed Model	94%

Table 2: Performance comparison of different number of features selected at the attention branch on Landmark classification

#Features	Classification rate
16	89%
32	88%
64	90%
128	84%
256	91%
512	87%
1024	94%

Reference

[1]. O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang,A.Karpathy, A. Khosla, M. Bernstein, A.C. Berg and L. Fei-Fei., ImageNet Large Scale Visual Recognition Challenge. International Journal of Computer Vision (IJCV), pp:25-32,2015.

[2]. A. Karpathy, G. Toderici, S. Shetty, T. Leung,R. Sukthankar, and L. Fei-Fei. Large-scale video classification with convolutional neural networks. In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1725–1732, 2014.

[3]. H. Noh, A. Araujo, J. Sim, T. Weyand, and B. Han, "Large-Scale Image Retrieval with Attentive Deep Local Features", In Proceedings of IEEE Conference on Computer Vision (ICCV), pp:789-793,2017.

[4]. Google Landmark Recognition Challage Dataset : www.kaggle.com/c/landmark-recognition-challenge