IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol. 8, Issue.4, Apr 2020 E-ISSN: 2347-2693

Image Classification based on Feature Extraction with AlexNet Architecture

Zarli Cho[1*], Khin Myo Kyi[2], Kyi Thar Oo[3]

[1]University of Computer Studies, Taungoo, Myanmar
[2]University of Computer Studies, Taungoo, Myanmar
[3]University of Computer Studies, Thaton, Myanmar

[*]*Corresponding Author: zarlicho85@gmail.com*

DOI: https://doi.org/10.26438/ijcse/v8i4.1418 | Available online at: www.ijcseonline.org

Received: 02/Apr/2020, Accepted: 19/Apr/2020, Published: 30/Apr/2020

Abstract —Deep learning has emerged as a new area in machine learning and is applied to a number of signal and image applications. Although the existing traditional image classification methods have been widely applied in practical problems, such as unsatisfactory effects and weak adaptive ability. The main purpose of the work presented in this paper, is to apply the concept of image feature extraction with AlexNet Convolutional Neural Networks (CNN) in Digital Elevation Map and Topological Map boundary classification of Yangon City in Myanmar. The automated derivation of topographic data from DEMs is faster, less subjective and provides more reproducible measurements than traditional manual techniques applied to topographic maps. Data are acquired from the United States Geological Survey (USGS) database. This study is supposed to handle of geospatial information and production of maps. Geospatial users have to understand the distortion characteristics of each maps. The analysis of this result is revealed that has a good classification accuracy for all the tested maps based on the proposed system.

***Keywords*—**AlexNet, CNN, Elevation Map, USGS

I. INTRODUCTION

Recent automation and increasing user-friendliness of geospatial systems has made the production of maps easier, faster and more accurate. Map boundary classification and recognition is important role in the research areas of Geospatial users which process are based on image processing techniques. Image processing involves some basic operations namely image restoration/ rectification, image enhancement, image classification, images fusion etc. Image classification forms an important part of image processing, the objective of image classification is the automatic allocation of image to thematic classes. Two types of classification are supervised classification and unsupervised classification. There are various machine learning approaches for solving this problem such as K nearest neighbor, support vector machine (SVM), decision tree and artificial neural network.

Deep learning or hierarchical learning, has emerged as a new area of machine learning research [2,3]. Deep learning is a class of machine learning techniques that exploit many layers of non-linear information processing for supervised or unsupervised feature extraction and transformation and for pattern analysis and classification [4]. This work aims at the application of convolutional neural network or CNN for image classification with AlexNet architecture. Deep Neural Networks (DNN) have shown significant improvements in several application domains including computer vision and speech recognition. In computer vision, a particular type of DNN, known as CNN, have demonstrated state-of-the-art results in object recognition and detection [5,8].

Convolutional neural networks show reliable results on object recognition and detection that are useful in real world applications. Training CNNs to perform this kind of automated feature extraction typically comes with the onus of requiring large volumes of labelled training data. The aim of the study is to apply the combination of image feature extraction technique with AlexNet architecture of CNN in digital elevation map boundary classification. The rest of the paper is classified as follows: section 2 describes the proposed system and briefly discuss the procedure of the system. Section 3 describes the related methodologies of this study and explains the analytical results in section 4.

II. PROPOSED SYSTEM OF IMAGE CLASSIFICATION

The aim of the study is to propose digital elevation maps classification based on AlexNet convolutional neural network which is combined with image features analysis. The process of the proposed framework which are shown in figure 1. In the data acquisition stage, the image data used for testing the algorithm includes Digital Elevation Map and Topological Map of Yangon, Myanmar from United States Geological Survey (USGS) database. Figure 2 describes the topographical map of Yangon City, Myanmar. The selected region are extracted by using ArcGIS software. In the pre-processing step, size of each image is 227x 227 and is comprised of 4 bands-Red, Green, Blue and Near Infrared. The execution is done on open source of MathWorks library. The major contribution of this work involves the following:

1. Co-occurrence Matrices: image feature extraction for proposed classification framework.
2. The extracted feature which are classified with AlexNet architecture of convolutional neural network.

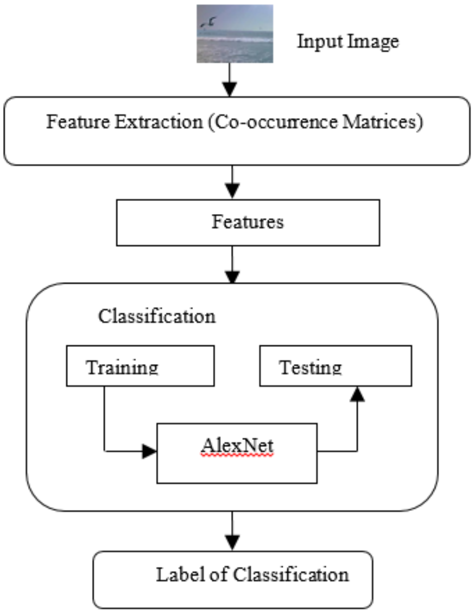

Figure 1. Proposed Framework

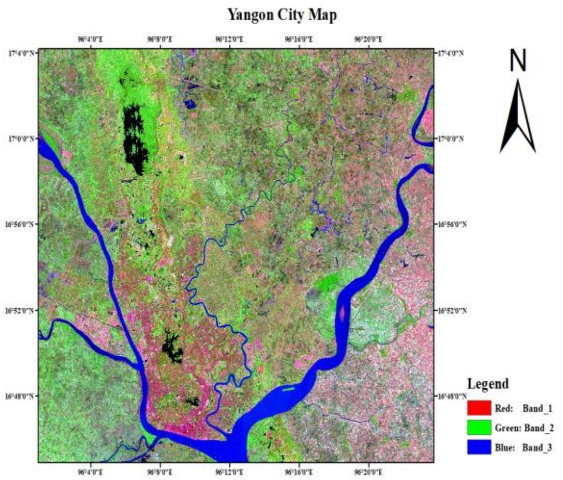

Figure 2. Topographical Map of Yangon City

III. METHODOLOGY

There are two fundamental technique for proposed system analysis. The input images are extracted with co-occurrence matrix and the extracted features are classified based on AlexNet architecture. The related methodologies of proposed system will describe in the following section.

A. Features Extraction

Among all statistical methods, the most popular one which is based on the estimation of the second order statistics of the spatial arrangement of the gray values is the gray level co-occurrence matrices. A co-occurrence matrix is a square matrix whose elements correspond to the relative frequency of occurrence of pairs of gray level of pixels separated by a certain distance in a given direction. Formally, the elements of a G*G gray level co-occurrence matrix P_d for a displacement vector

$d = (d_x, d_y)$ is defined as :

$$P_d(i,j) = \left| \{((r,s),(t,v)) : I(r,s)=i,\ I(t,v)=j\} \right| \quad (1)$$

where $I(\cdot,\cdot)$ denote an image of size NxN with G gray values, (r, s), (t, v)∈ NxN, (t,v)=(r + dx, s + dy) and | . | is the cardinality of a set. Haralick, Shanmugan and Dinstein [7] proposed 14 measures of textural features which are derived from the co-occurrence matrices, and each represents certain image properties as coarseness, contrast, homogeneity and texture complexity.

Those that are used, in this work, for extracting features in the defect detection of textured images are:

1) Entropy :

$$ENT = -\sum_i \sum_j p(i,j) \log\ p(i,j) \quad (2)$$

Entropy gives a measure of complexity of the image. Complex textures tend to have higher entropy.

2) Contrast :

$$CON = \sum_i \sum_j (i-j)^2\, p(i,j) \quad (3)$$

Contrast feature is a measure of the image contrast or the amount of local variations present in an image.

3) Angular Second Moment:

$$ASM = \sum_i \sum_j \{p(i,j)\}^2 \quad (4)$$

Angular second moment is a measure of the homogeneity of an image. Hence it is a suitable measure for detection of disorders in textures. For homogeneous textures value of angular second moment turns out to be small compared to non-homogeneous ones.

4) Inverse Difference Moment:

$$IDM = \sum_i \sum_j \frac{1}{1+(i-j)^2}\, p(i,j) \quad (5)$$

In Equation (2) - (5), $p(i,j)$ refers to the normalized entry of the co-occurrence matrices. That is $p(i,j) = P_d\,(i,j)/R$ where R is the total number of pixel pairs (i,j). For a displacement vector d = (dx,dy) and image of size NxM R is given by (N-dx)(M-dy).

B. Convolutional Neural Network

Convolutional Neural Network formed with the help of different layers to perform the image classification task [6]. The architecture of the CNN contains the different layers: Input layers, Convolution layer, ReLU (Rectified Linear Unit), Pooling, Fully Connected Layer and Softmax Layer.

1) AlexNet Architecture

There are many reputed CNN architectures. In this paper, the proposed image classification technique is based on AlexNet architecture. The first famous CNN architecture is AlexNet, which popularizes the convolutional neural network in Computer Vision, developed by [1]. In 2020 AlexNet was presented to the ImageNEt ILSVRC challenge and considerably performed better than the second runner-up. The network had the similar architecture to LeNEt; however, it was most profound, most significant architecture with all the convolution layers stacked together rather than the altering convolution and pooling layers as it was in LeNet. The process of AlexNet architecture which are shown in figure 3.

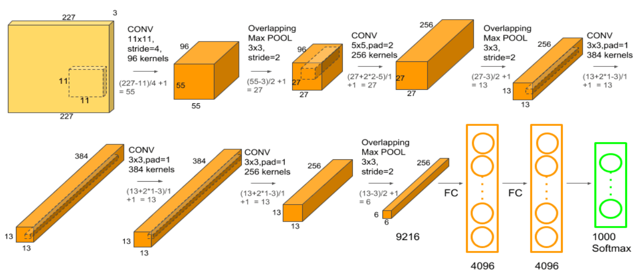

Figure 3. AlexNet Architecture

IV. RESULTS AND DISCUSSION

The performance of digital elevation map classification of the proposed system is evaluated on a publically available USGS database. Training the architecture has been carried out with 32 Topographical Map and 32 Digital Elevation Map. The results are summarized in term of precision, overall accuracy, class wise accuracy, number of trainable parameters.

$$Accuracy = \frac{TP+TN}{TP+FN+FP+TN} \quad (6)$$

$$Precision = \frac{TP}{TP+FP} \quad (7)$$

TP is the amount of effectively classified patches, TN is the amount of the patches that don't have a place with the specific lass and not assembled accurately. FN is the amount of patches that have a place with the specific class however were not gather precisely, FP is the amount of patches that don not have a place with the specific class and have been wrongly ordered.

A. Preprocessing stage of Weight Matrix Analysis

In this study, the tested image belong to the downtown area of Yangon city, Myanmar. Some of the topographical map can be classified by visual analysis but could appear similar when reduced in size and changed to digital elevation map. This section describes the size of weight matrix in each layer of the proposed framework. Computation of weight matrix size in each layer in the proposed system can be calculated based on the number of output and input channels in each layer and filter size. Figure 4 describes the training and testing data of proposed system.

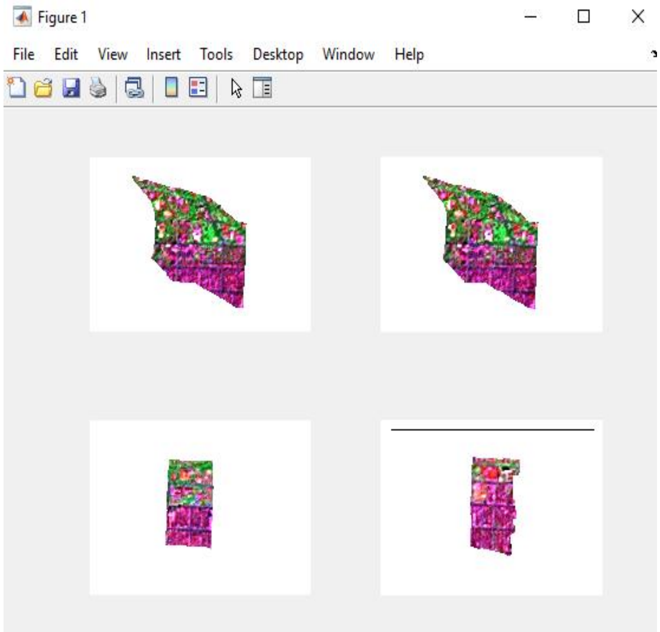

Figure 4. Different types of training and testing data

B. Analytical Results

In the feature extraction process, the co-occurrence matrix is a statistical method of examining texture of a grayscale image. Which is more suitable for analysis of digital elevation map of boundary classification. The textural features extracted from the images by co-occurrence matrix were helpful in identification of different regions in the tested data.

The number and size of channels influence the quantity of parameters in convolutional systems. A 25 layer convolutional neural network with ReLUs (solid line) reaches a 34% training error rate on CIFar 10 six times faster than an equivalent network with tanh neuron (dashed line). The learning rates of each network were chosen independently to make training as fast as possible. No regularization of any kind was employed. The magnitude of the effect demonstrated here varies with network architecture, but network with ReLUs consistently learn several times faster than equivalents with saturating neurons. The validation accuracy of training progress is above 66% which is shown in figure 5.

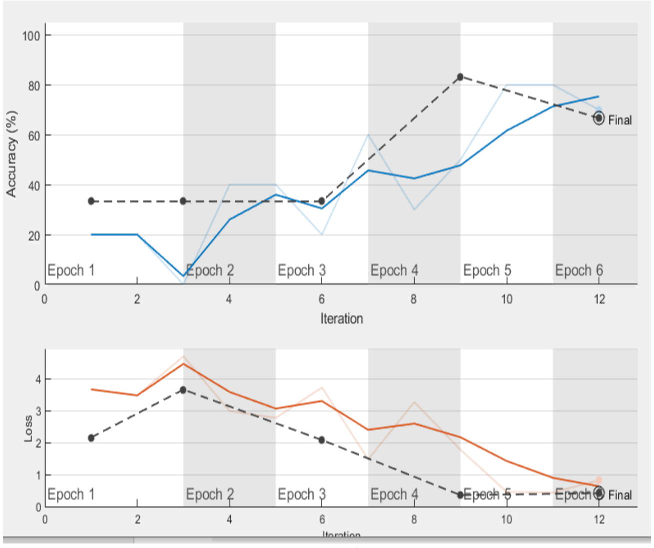

Figure 5. Training Progress of Validation Accuracy

Figure 6 describes the labelled of townships for tested data which are classified with AlexNet CNN. In this analysis, training and testing data of topographical map are gray scale image although digital elevation map are color image. In this figure, topographical map such as "Kyauktada, Pabetan and Latha" township and digital elevation map such as "Lanmadaw" township are correctly classified with the proposed system.

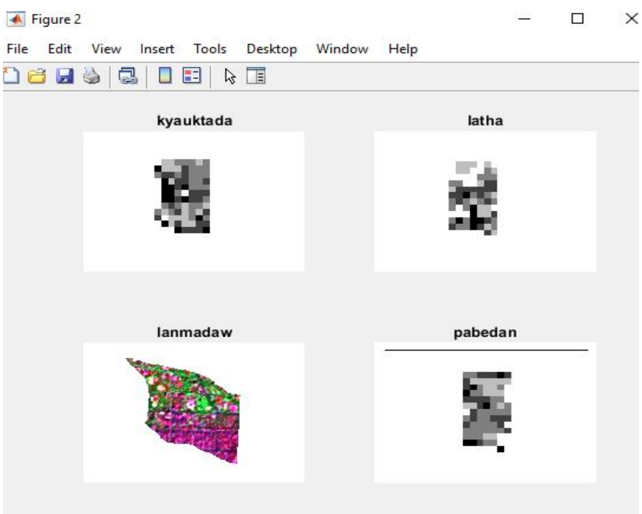

Figure 6. Image Labeling for Digital Elevation Map (Township)

The ROC curve shows the trade-off between sensitivity and specificity. Figure 7 shows that analysis of the four classes (township) which indicates a better performance because the classifier gives curves closer to the top-left corner.

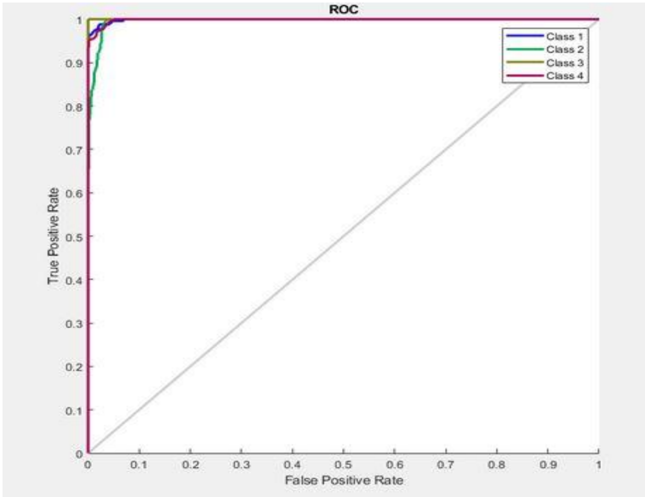

Figure 7. ROC curve for Map Classification

V. Conclusion and Future Scope

In this analysis, image features extraction based on AlexNet architecture which is applied in the digital elevation map and topographical map boundary classification. The traditional classification algorithm has the disadvantages of low classification accuracy and poor stability in maps boundary classification tasks. It showed that the combination of image feature extraction based AlexNet CNN is significantly better than the traditional classification algorithm. By applying co-occurrence matrix, each matrix prepared the data to emphasize primarily structure in given direction and a grain size that is at least as large as the selected distance. Combine to use with deep learning with feature extraction is able to detect complicated interactions from features, learn lower level features from nearly unprocessed original data. The experimental results show that the proposed method not only has a higher accuracy than other mainstream methods but also can be good adapted to various image databases. The analysis of this study revealed that gave a better accuracy for boundary image classification and automatically detected the important features without any human supervision by using AlexNet convolutional neural network.

Reference

[1] A. Krizhevsky, et al, "Imagenet classification with deep convolutional neural networks," in Advances in neural information processing systems, 2012, pp, 1097-1105.

[2] Bragilevsky L, Baji IV. (2017) "Deep learning for Amazon satellite image analysis." Communications, Computers and Signal Processing (PACRIM).:1–5

[3] Giacinto G, Roli F. "Design of effective neural network ensembles for image classification purposes". Image Vision Comput 2001;19(9–10):699–707.

[4] Lu, Dengsheng and Weng, Qihao. (2007) "A survey of image classification methods and techniques for improving

classification performance.” International journal of Remote sensing 28(5):823–870
[5] Meng T, Wu C, Jia T, Jiang Y and Jia Z, ‘Recombined convolutional neural networks for recognition of macular disorders in SD-OCT images’, In 2018 37th Chinese control conference (CCC), pp 9362–9367, IEEE.
[6]M. M. R. Khan, et al., "Study and Observation of the Variation of Accuracies of KNN, SVM, LMNN, ENN Algorithms on Eleven Different Datasets from UCI Machine Learning Repository," arXiv preprint arXiv:1809.06186, 2018
[7] Ojala, T., & Pietikäinen, M.; "Texture Classification, Machine Vision and Media Processing Unit", University of Oulu, Finland, Available at.
[8]Y. Kim, "Convolutional neural networks for sentence classification," arXiv preprint arXiv:1408.5882, 2014

Authors Profile

She received the M.Sc degree of physics in 2000. She got a Ph.D degree of computer hardware technology in 2005. Her research interests include image processing, data analysis, cryptography and network security.

She got a bachelor of computer science degree in 2005 from University of Computer Studies, Taungoo, Myanmar. In 2010, she received the MASc degree from University of Computer Studies, Thaton, Myanmar. Her research interests is image processing, digital signal processing and image security.

She received an M.Sc degree of physics in 2000. She got a doctor of philosophy in computer hardware technology in 2005. Her research interests is image processing and network security.