Journal of Theoretical and Applied Information Technology
15th December 2018. Vol.96. No 23

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

A CONVOLUTIONAL NEURAL NETWORK APPROACH TO ROAD CLASSIFICATION FROM SATELLITE IMAGES

[1]JOSE HORMESE, [2]CHANDRAN SARAVANAN

[1]Research Scholar, Dept of Computer Science and Engg, NIT, Durgapur, 713209, India
[2]Associate Professor, Dept of Computer Science and Engg, NIT, Durgapur, 713209, India

E-mail: [1]josehormese@gmail.com, [2]dr.cs1973@gmail.com

ABSTRACT

The significance of extracting roads from a satellite image of high resolution can help in road navigation, update geographic information systems, emergency rescue system that requires instantaneous maps. The difficulty lies in identifying and differentiating a road from its background. To overcome this difficulty more and more research is being conducted for devising efficient methods that can extract a road from a satellite image. The large presence of buildings and manmade structures along the roads or the presence of trees or the acquisition properties of the sensor can cause difficulties in identifying a road from a satellite image. The conventional way of identifying a road was using semi-automated approaches which is not feasible as well as consumed large amounts of time. A high accuracy technique for road extraction is needed. Also when compared with extraction from aerial images the extraction from satellite images is a challenging domain. The reason for that is satellite images have a resolution that is low and the presence of noise contents in any given image is quite high. This work deals with extracting a road network from high resolution satellite images. This work deals with estimating if a pixel in a satellite image is part of a road or not using Convolutional Neural Network. The advent of Tensor flow library has made this work feasible. The work proposes a new approach for making data sets for this complex problem and has concluded with a feasible solution for the problem.

Keywords: *Satellite Image, Geographic Information Systems, Road Network, Convolutional Neural Network, Tensor Flow*

1. INTRODUCTION

Taking photographs of Earth and other planets is becoming common as there is rapid development of the sensor technology employed in satellites. These satellite images have lot of uses in the areas of regional planning, agriculture, conservation of biodiversity, forestry, education, geology, intelligence applications and so on. The field of Geographic Information Systems have benefitted in a big way with the help of satellite images. Lot of studies and investigations were conducted over the years to reliably extract features like roads. A road network can be considered as a long narrow region with a particular length, width and orientations. It is seen that the width of the roads in any high resolution satellite image is several pixels and the length is longer than that which corresponds to buildings and other man-made structures.

It is important to identify a road network from satellite images that are having relatively high resolutions and be able to distinguish it with objects like buildings and rivers. The color information that is available in usually three or more spectral bands is taken as a significant attribute. Among the assortment of methods, the choice of the particular approach depends on the balance between speed, accuracy, complexity of the algorithm being used. It has to be considered that the expected accuracy is correlated to the good quality of the image, data that is available and also related to the resolution of the image available. Noise and occlusion can occur which makes the detection of roads very difficult and they can cause the creation of non-homogeneous regions that can lead to the inaccurate classification of road segments. There is a possibility that complex background and contextual structures like vehicles, shadow of trees other structures blocking the view can appear in a high resolution image. There is also a possibility that some road like segments having similar or even identical spectral and spatial properties like parking lots, railway lines or rivers can be misclassified as roads. The satellite images can be represented as raster images and digital raster images can be represented as portrayals of scenes, with imperfect renditions of objects. Imperfections in an image can result from the given imaging system, signal noise, atmospheric scatter and shadows as mentioned earlier. The task of identifying and extracting the required features from a raster image is based on a criteria developed to determine a particular feature (based on its characteristics within any raster image) and ignoring the presence of other features and imperfections in the image which are not useful.

2. LITERATURE REVIEW

There are two approaches for road extraction namely semi automated methods of extraction and automated methods of extraction. Both the methods are helpful for reducing the labour and time to a certain extent in the updation of the road database. It should be noted that semi automatic methods for the extraction of roads require some road seeds which has to be given as the initial points, that are normally given by the user and the particular road section can be seen to evolve under a particular model. Also, these methods use the black-and-white photographs of aerial imagery and also the panchromatic band of satellite images with good resolutions and hence the characteristics of the roads that are geometric in nature plays a crucial role. There are various algorithms for the classification of roads that have been discussed and proposed in early decades. Active contour models that are also known by the term snakes, have been used in the semi automatic extraction of roads [1]. The paper identifies an object based on a manual approach; Seed points have to be given which is used to understand the shape as well as position. Whenever there are any distortion that happens, the method gives a robust performance. Mena [2] and Das et al. had presented overviews of methods for detection of roads in this area. A novel method of classifying roads and tracking them based on analysis of morphological parameters, applying dynamic programming approaches is discussed. An Integrated Method that can be used for the Centerline Extraction of main road which is based on the classification of spatial properties to segment the images which are based on road as well as non-road groups [3]. The paper discusses on the classification based on spectral and spatial properties for segmenting the respective image either as a road or a non road type. Tensor voting techniques are used for the generation of the network of roads. Pixel based methods is used to classify road detection where Edge Detection techniques are applied to extract the appropriate road points [4]. Labeling of a road pixel is subjected to errors and the time taken is quite high, an automated system is proposed. A multistage framework was proposed and designed for the extraction of road networks that are based on the salient features of probabilistic SVMs [5]. The idea of the method is to use two road features namely linear trajectory and spectral contrast approaches. The statistical inference method was designed such that the linear features were modelled as a Markov point process as well as a geometric-stochastic model that can be applied on the width of the road, the direction and background intensity with a maximum a posteriori probability is used which helps in estimating the road network [6]. The approach uses a probabilistic – geometric model and also an estimation model for the generation of a road from a given image. Long and Zhao [7] address the automatic detection of roads in an optical image by using multiscale pretreatment steps. The urban environment always presents more difficulties compared to rural areas, as there is a high variability of gray shades on certain sections of the road. Shi et al. [8] present a novel way for the main-road centreline extraction from satellite images that are optical in nature which can integrate both the spectral as well as spatial information for classification purposes. This method can detect both curved and also rectilinear structures in a simultaneous way. Christophe and Inglada [9] have proposed an interesting road extraction technique, which employs spectral information to distinguish roads from other land uses. Roads can be considered as a segment that is elongated with the linear properties in the image that are local in nature. The Segments that correspond to a road form have a distinct direction, but at the same time the directions of these elongated regions are hard to estimate and the road segments are lost at the same time to some degree as a result of segmentation as only a small set of directions are taken into account. Hence it is difficult to classify a whole complete road segment [10].

Quackenbush [11] had given a review of the linear feature extraction from imagery that can be utilised for the extraction of roads. The importance of having an automated approach is to save the time as well as cost factors and provides an improvement in accuracy as well as detail. Poullis and You [12] classified the various road detection methods based on three categories namely pixel-

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

based or region-based and knowledge-based as well. The pixel-based methods mainly depend on the information that are computed and obtained from the pixels. Line [13], [14], and ridge [15], [16] detectors are primarily used for the extraction of potential road points and then the road points are interconnected that produces road segments which can be used as an input to a next level processing phase.

Road extraction methods have been proposed by several authors from different approaches and perspectives. Based on the homogeneous polygonal areas that are around each pixel, Hu et al. [17] has defined a pixel footprint which are useful for the extraction of road areas. Zhang et al. [18] had applied this detector for the extraction of roads in urban areas for getting better results. Movaghati et al. [19] had applied particle filtering approach as well as the Kalman filtering techniques for the extraction of road networks from aerial or satellite images. The Road intersection extraction from satellite images was identified and studied [20]–[22]. The studies show that the road intersection extraction alone can never denote a complete road network, but it is helpful for the understanding of road network topologies that can be used for higher processing.

3. PROPOSED METHOD

The main objective of the work is to formulate a unique method of road extraction from satellite images. All the works in road extraction so far are based on morphological operations on clusters of neighborhood pixels. The limitations of these methods are lack of generalization, limited success rate with change in dataset, terrain and spectral range. We live in a world where machine learning has progressed to a very great extent. Computer vision is one field where machine learning has been extensively used with commendable success ratio. We human beings are capable of estimating roads upon seeing an aerial photograph. There are certain visual patterns, which distinguishes a pixel to be classified as part of a road. As of now, there exist no mathematical classifiers to define the afore said patterns. In fact these patterns are difficult to define, but feel like intuitive to a human being. It is this thought, that intuitive patterns, which proposed Artificial Intelligence, Machine Learning and in turn, Artificial Neural Networks as a prospective solution for approaching the road extraction problem. The initial approach was to make a dataset of satellite views and corresponding road views. Then use these data sets to train an ANN. Upon starting working on it, two issues were identified.

1. It's not available and not easy to make a sufficiently large data set.

2. An ANN with input and output as a complete image is too complex and memory consuming, that it will be difficult to represent and train neural network with available computers.

Based on the work done above, a better lean approach was formalized. The problem has been refined as, estimate if a pixel is part of the road, based on the color profile of neighborhood of the pixel of interest. This makes the problem computationally approachable in two ways.

1. The ANN has been greatly thinned downs as input is a small block of pixels, approximately an array of 1000 numbers.

2. Output of ANN with new approach has become a scalar.

3. A single 1000 x 1000 image produces training cases of about a million with this approach.

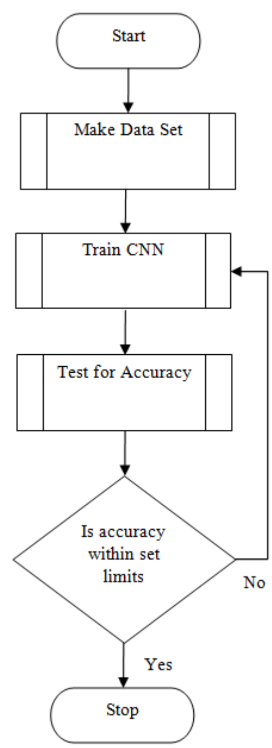

***Figure 1:** Flow Chart for Training CNN*

Figure 1 represents the Flow Chart for Training the Convolutional Neural Network. The first step is to make a Data Set. There are difficulties in making Data Sets that are comparatively large. The next step is to train the Convolutional Neural Network. This stage is a complex process and the amount of memory needed for training is too high. It is highly desirable to test for accuracy which happens in the next stage. If the accuracy is not within the set limits, the Convolutional Neural Network is again trained for getting better results.

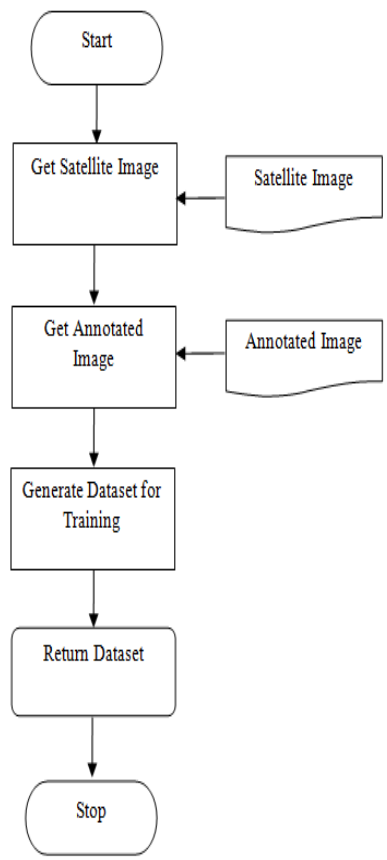

Figure 2: *Flow Chart for Making Dataset*

Figure 2 represents the flow chart for making the Dataset which is extremely important for any classification purposes like detection of roads. The first step is to get the Satellite Image from the repository. The next step is to get the Annotated Image which helps in generating the Dataset for training purposes. Annotation of images is also known by the term Image Tagging, where this technique is used for systems that need retrieval of images to identify and classify images from a database which is of interest. This is also a classification technique. From these sets of images the Dataset is generated.

3.1 Dataset

There are no datasets openly available for this problem. Datasets were created with satellite maps available from maps.google.com. Satellite view is used for satellite image and map view is used to create the output for the training and testing datasets.

3.2 Tools Used

Python is the global language of choice for machine learning experiments and projects. OpenCV is used for image manipulation and Numpy is used for general mathematical operations and data representations. In the initial days of the work, Scikit learn was used to attempt the machine learning approach for the problem. The limitation of this python library is speed and multiprocessing. Though library is easy to work with, it cannot make use multiple cores available in a PC these days for increased speed of network training. Upon facing limitations with Scikitlearn, yet another review of the trends across globe revealed that Tensorflow has matured enough to attempt deep learning problems with in manageable time lines with efficient utilization of multi core architectures and Graphic Processing Units (GPU). This library seems to be the best to work on this problem. Tensorflow spans computational process across multiple cores of a PC thereby, an increase in speed that is training time is found to reduce to 1/3 of time taken by Scikit Learn. Here Keras helps to integrate Tensorflow in Python. This heats up computer very much and any model with input vector bigger than 100 still takes more than 24 hours. Tensorflow with Keras is used to create, train and test the Deep Convolutional network discussed in this work.

Computational infrastructure requirement is a major bottleneck with Convolutional Neural Network solutions, particularly for training the network. Amazon Web Services (AWS) offers very sophisticated infrastructure with upto 16 GPUS for fast training of Convolutional Neural Networks. In additional large size RAM is available to work with big data sets. The hardware requirement of the infrastructure is met with time leased resources from AWS.

3.3 Solution Overview

With the new approach, a single satellite image can be used for generating a training data set of a

Journal of Theoretical and Applied Information Technology
15th December 2018. Vol.96. No 23
© 2005 – ongoing JATIT & LLS
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

million cases. Satellite Images comes in various forms starting from very high resolution images used for military applications to lower resolution images depending on the application where it has to be used. There are images of satellites which can represent portions of the land surface to less than 50 cm distance. A block of image with 255 x 255 pixels is one training input to the neural network. This block is used to predict, whether the pixel [127, 127] is represents a road or not. The initials input and output images are given in Figures 3 and 4. Figure 3 represents a sample satellite image which is having a quite high resolution. On close observation it can be noted that the roads can be identified. The aim of the experiment is to identify the roads segments of different dimensions

Figure 3: *A sample satellite image*

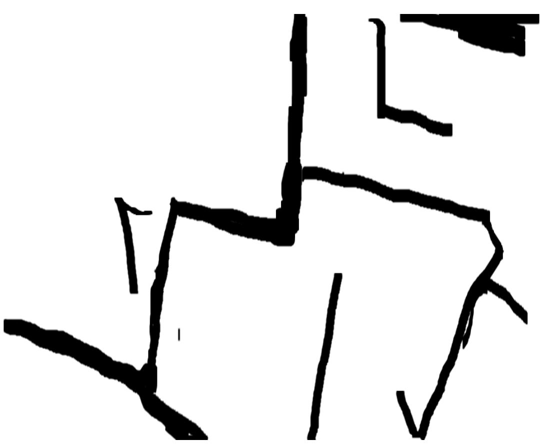

Figure 4: *Road extracted from the satellite image*

Figure 5 shows a single training input from the satellite image. The corresponding output for the block is positive indicating that central pixel is a road. For example, the above image gives a training set of 170,200 cases. For a CNN to get training effectively, the number of positive and negative cases should be equal. This particular image gives about 17,472 positive cases. Hence out of the negative cases, an equal number of random CSS are selected for the CNN training.

Figure 5: *A single case input to CNN for estimating the class of the centre pixel*

3.4 Convolution Neural Network Configuration

Choosing a optimum configuration for the network is trade off problem. A large network takes a long time to get trained and often ends up in over training. Overtraining is a problem where, instead of making a generalized learning, the network gets trained for the given dataset. It should be noted that small networks often gets trained fast, but fail to classify the complex non-linear cases.

3.4.1 Segmentation with Deep Convolutional Neural Network

Deep Convolutional Neural Network (DCNN) provides an effective and satisfactory result in the classification of roads from a satellite image. It is also an effective way to identify the road region from the satellite images. It should be noted that it is possible to identify the roads that are visible in the satellite images. Roads in satellite images are identified at the pixel level. Convolutional neural networks can handle images of any size whether it represents that of a small area or a very large area. Depending on the type of the satellite sensor, the type of the image that is obtained varies considerably. Convolutional neural networks have a flexible property that they have adaptive layers that can connect the main components that are present in the network. It can analyze and identify the satellite images and can output a group of predictions. The raw satellite image set will be down sampled with the help of pooled layers in relation to the set and hence become more and more meaningful. The last output in the pooled layer output is classified with the help of fixed sized layers that are connected fully. This can cause bottleneck in the network that may impose a limit to the size of the input. It is possible to take away the limit of the layers that are fully connected with the help of convolutional layers.

A Visual Geometric Group (VGG-19) Convolutional network acts as the backbone of a convolutional neural the network. The advantage and main purpose and use of this VGG-19 network is that it can down sample a given satellite image to upto a factor of 32 and hence the output obtained interpreted in a visual manner. The predictions made by the network are then processed and handled by the next component of the convolutional neural network which is nothing but a general up-sampling type of network. This up-sampling network can restore the properties of the predictions that are spatial in nature using convolutional layers that are backward in nature. This happens till the predictions made do have the same size as that of the given satellite image. A Fully Convolutional Network (FCN8) architecture is employed which can instill the results obtained from intermediate layers of a VGG-19 network to the up-sampling network. These given layers have a prediction resolution that is higher and helps in improving the accuracy of segmentation than other versions of fully convolutional networks. It is possible for a FCN32 that can up-sample a VGG-19 network output to 32 times which results in a more coarse segmentation. While at the same time a FCN16 can fuse only a single layer. It can be seen that by adding an additional layer, FCN8 has a much finer accuracy level.

Two approaches are used to categorize the output of a FCN8. One is a binary segmentation approach and the other is a regression approach. In the first approach a spatial tolerance of zero is associated with the binary segmentation. The scheme of this approach is to identify a pixel in the image either as a road or as a background. In the second approach namely the regression approach is associated with a spatial tolerance that is adjustable. The idea is to give a even target distribution that is centered along the road labels, having a maximum value of one that is assigned to the road labels which uniformly decreases to zero for the background other than the roads. The necessity for a tolerance is that if the network is penalized for predicting and identifying a road network that could be a few pixels away can have negative effects on the effectiveness of training. Hence the raw predictions that are made can be categorized by a softmax function while performing the binary segmentation approach or a sigmoid function wherever the regression approach is applied. Either of the two approaches produces a mapping of each and every data point that has a confidence score which varies between the values 0 and 1. The confidence score value suggests that a value that is closer and nearer to 1, indicates that the corresponding pixel do indeed belong to a road and in the other extreme a confidence score value closer to 0, indicates that the corresponding pixel represents the background respectively. The above mentioned softmax function indeed gives two different confidence score value for each and every pixel that can be represented in the form:

$$P_{j,\,road} + P_{j,\,background} = 1 \qquad (1)$$

where $P_{j,\,k}$ represents the probability of a class k that is predicted for the j-th pixel. The next stage in the segmentation part is that the confidence map undergoes thresholding so that a binary segmentation map is obtained. As far as the softmax output is concerned the largest confidence score value for all the pixels is selected with the help of argmax function. Another approach is taken for the sigmoid output where a threshold value of 0.5 is used.

3.4.2 Reduction in the effect of class imbalance

It has to be noted that a road network in a satellite image do appear as thin objects and there is a possibility that it could be easily outweighed when compared with the background class. This is a live problem which has to be overcome. Hence appropriate steps have to be taken that limits the imbalance of classes during the training phase. The implication is that the labels have to cover exactly the outline of roads and its embankments due to their visibility in the satellite images. So an effective way in the reduction of imbalance is to increase the spatial tolerance of the corresponding regression model. An alternative approach to handle this issue is to reweigh each class based on the loss calculation. The idea is to multiply the

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

pixel prediction loss with a coefficient value that is inversely related to the true class in the ground truth. It should be noted that these coefficient values is computed using the median class frequency values. As there are only two classes that exist here, the weighting coefficient of the background is set to 1 and experiments are being carried out with weighting coefficients of the road that is taken in the range W=[1,1/g_{road}] where g_{road} represents the ratio of pixels of road and the total pixels present in the ground truth.

Two different losses are being computed that depends on a type of model used. The Cross-Entropy loss namely (CE) is represented as:

$$Loss_{CE}(Y_{bin}, \hat{Y}_{bin}) = -\frac{1}{n}\sum_{i=1}^{n}\left(w_i \sum_{k} y_{i,k}\log(\hat{y}_{i,k})\right) \quad (2)$$

In the above equation $y_{i,k}$ and also $\hat{y}_{i,k}$ represents labels in binary ground truth namely Y_{bin} and the corresponding value of softmax in the predictions of the binary segmentation $\hat{Y}_{bin}$ for a given class k of the corresponding pixel i.

Also the computation of the Mean Squared Error loss namely (MSE) can be represented as:

$$Loss_{MSE}(Y_{tol}, \hat{Y}_{reg}) = \frac{1}{n}\sum_{i=1}^{n} w_i(y_i - \hat{y}_i)^2 \quad (3)$$

In Equation No 3, the attributes y_i and $\hat{y}_i$ represents the labels in the ground truth tolerant Y_{tol} and the corresponding sigmoid value that is present in the regression predictions namely $\hat{Y}_{reg}$ for a respective pixel i.

The total number of pixels in the satellite image is given by n and the corresponding loss weighting coefficient w_i for a given pixel i can be defined for both the losses as:

$$w_i = \begin{cases} \lambda & \text{if pixel } i \text{ is labeled as a road in } Y_{bin} \\ 1 & \text{if pixel } i \text{ is labeled as background in } Y_{bin} \end{cases} \quad (4)$$

λ Represents a fixed value taken from interval W

As given in the proposed solution the input layer of the convolutional neural network is a 255 x 255 x 3 matrix which is a block from the satellite image, which is expected contain relevant information to estimate the central pixel of the block is part of a road or not. The rest of the layers of the CNN is expected to have sufficient complexity to model the classification problem. Prior to passing the convolutional neural network, the mean of each RGB pixel is subtracted to remove redundant bias in the input data. The input matrix is then passed through a cascade of convolutional weight layers as shown in Table 1.

Table 1: *DCNN Training History*

SI.NO	Epoch	Error
1	500	2.7
2	1000	2.2
3	1500	0.8
4	2000	0.46
5	2500	0.2
6	3000	0.29
7	3500	0.27
8	4000	0.21
9	4500	0.13
10	5000	0.12

ReLU [23] is applied at multiple layers to ensure non-linearity decision making capabilities and simplify decision making by the end of the network layers. The final layer is a soft-max layer for final decision making. The network model is adapted from Karen Simonyan & Andrew Zisserman [24].

3.5 Training

There are no open public dataset for this problem. The lack of publicly available open datasets also shows some light on the novelty of the problem. Yet another reason for not having public open dataset is that, the computational complexity

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

of the problem is so high that, a single computer cannot meet the processing capability wise requirements of the deep neural convolutional neural network. It is the recent time, infact in the last couple of years, that cloud based computers have become available, that it has become feasible to have a access to very high computation power over internet. Another major contributor in making this problem approachable is the Tensorflow library from Google, which has dramatically parallelised the backpropagation gradient descent training .

The lack of publicly available open dataset has contributed to using publicly available online maps as the source of data for the work. The input is a satellite image from Google Maps, without any textual or graphical annotations on the satellite image. An attempt was made to use the vector maps of the area as the ground truth for training the deep convolutional neural network. But proper alignment with perspective distortion is a major hindrance in going with standard road data from the maps as source of ground truth. This resulted in manually creating a layer of roads on top of the satellite image with image editing tools. It is done this way to ensure proper and aligned overlap between input and the truth that is the road alignment. Later the satellite image and road image are separated as two independent files, the satellite image as input of the deep convolutional neural network and the roads file as result of the network, for the purpose of training.

No further processing is done on the input image, as it is hypothesized that a well designed deep neural network ought to have ability to identify patterns beyond absolute values and noisy environments. The roads used in the work are from Trivandrum, India and an actual survey of the location was made by walking on the roads and thus ensuring that there exists no disparity between the ground truth made and reality.

An Amazon Web Services Cloud machine with 60 GB RAM and 16 GPUs was used for training the network. The procedure used is Krizhevsky [23]. The back-propagation algorithm attempts to optimize a multinomial logistic regression objective of cross-entropy. Mini-batch gradient descent is used to minimize memory requirement of the training machine. The batch size used is 1024 with a momentum of 0.8. The initial weights are assigned with uniform random numbers. The number of epochs was set at 5000, which is a pretty large number, but a CNN of this complexity requires large number of epochs as there was no works of similar nature to start with from a trained CNN weights.

4. RESULTS

The evaluation metrics used for accuracy is the ratio of:

$$(TP + FN) / (TP + FN + TN + FP) \qquad (5)$$

Where TP is False Positives, FN is False Negatives, TN is True Negatives and FP is False Positives. It is the balancing of positive and negative cases equally in the validation dataset helping to use a simple ratio for accuracy studies.

The program was run using an Amazon cloud with 60 GB RAM and 16 GPUs. Figure 6 shows the learning curve for 5000 iterations which took about 58 hours to run. It gave an accuracy of 87 % on an average. The results obtained suggest that Convolutional Neural Networks is a prospective technology for road segmentation from satellite images. The results are expected to improve as further optimizations and refinements are carried out over the current network.

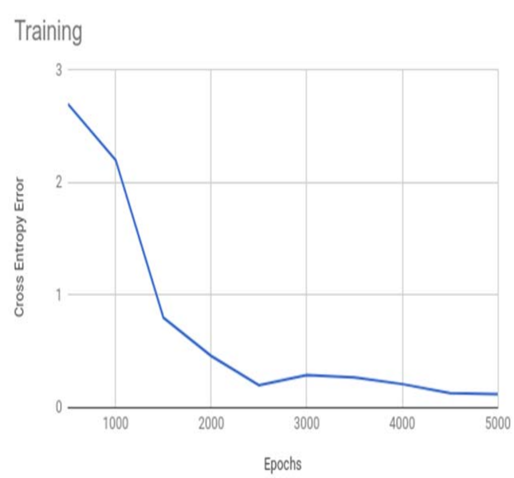

Figure 6: *Error over epochs during training*

It can be noted that getting results for simple images is quite easy as shown. Figure 7 shows an Input Image of a Rural Area. This input corresponds to a high resolution satellite image where there are no occlusions. But there is a presence of a river which looks similar in appearance. The challenge here is to clearly identify the road network.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Figure 7: *Input image of a rural area*

The corresponding output image obtained is as follows given by Figure 8. Using the automatic method the output obtained is as follows which clearly plots the road segments from the given input image. Roads are being plotted where the line segments corresponding to roads are being shown.

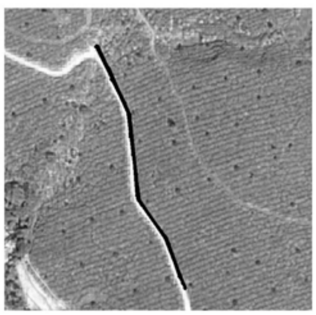

Figure 8: *Output image of a rural area*

Using the automatic method the output obtained is as follows which clearly plots the road segments from the given input image. Roads are being plotted where the line segments corresponding to roads are being shown.

4.1 Discussion

As established by Wang, W., et al. [25] road extraction is a difficult problem due to complexity of roads such as discontinuities, occlusion, parallel boundaries and sharp bends. Use of one kind of road feature can never solve the problem of road extraction from satellite images with sufficient accuracy to meet business requirements. Wang, W, et. al has left this as an open ended problem. Binary Tree Partition has been proposed by M. Li et al [26] for the road classification problem. The proposed approach is commendable for its simplicity in approach but fails to produce quality results. Hierarchical graph-based segmentation approach proposed and worked out by RashaAlshehii [27] has built extensively over Binary partition tree. Though the results are commendable, the approach demands optimum parameter values to be estimated for different terrains, which are more of heuristic in nature and there are no straight forward strategy for estimating parameter values. The complexity of the approach further adds restriction on parameter value estimation. Ruyi Liu [28] approach which relies on fundamental segmentation approaches but letting probabilistic patterns of roads as the critical classification criteria is yet another unique approach to the problem, but the results are not commendable and the approach is very restrictive in nature.

With refinements and optimizations better results are anticipated. The Scope for future work is to:

1. Optimize the CNN architecture for speed and accuracy.

2. Try out different network architectures and compare learning rate, speed and accuracy.

3. Manage dataset batch wise so that memory requirement of the machine can be minimized.

5. CONCLUSION

The Deep Convolutional Neural Network approach proposed in this work is able contain multiple features for road classification. Even though the fundamental mechanism behind the inferencing will never be inferred from the network, the results are satisfactory from an engineering perspective. The computational cost in terms of resource requirement and time are the major aspects demands further attention as continuation of this work, which can be further refined and improved with an extensive survey of deep learning networks and architectures. The Paper discusses about an unique approach in classifying roads from a satellite images using a convolutional neural network approach. Not much works have been undergone in this area of road classification. This area of road classification has a lot of potential and prospects in the near future.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

REFERENCES

[1] Grün, A., Li, H., Semi-automatic linear feature extraction by dynamic programming and LSB-snakes, Photogrammetric Engineering & Remote Sensing, Vol. 63, No. 8, pp. 985-995, (1997).

[2] J. B. Mena,State of the art on automatic road extraction for GIS update: A novel classification,Pattern Recognit. Lett., vol. 24, no. 16, pp. 3037– 3058, Dec. (2003).

[3] Wenzhong Shi, Zelang Miao, and Johan Debayle, An Integrated Method for Urban Main Road Centerline Extraction from Optical Remote Sensed Imagery, IEEE Transactions on Geo-science and Remote Sensing, Vol. 52, No. 6, June (2014).

[4] C. Unsalan and B. Sirmacek, Road network detection using probabilistic and graph theoretical methods, IEEE Trans. Geosci. Remote Sens., vol. 50, no. 11, pp. 4441–4453, Nov. (2012).

[5] S. Das, T. T. Mirnalinee, and K. Varghese, Use of salient features for the design of a multistage framework to extract roads from high resolution multispectral satellite images,IEEE Trans. Geosci. Remote Sens., vol. 49, no. 10, pp. 3906–3931, Oct. (2011).

[6] Barzohar, Mein, Cooper, David B. Automatic finding of main roads in aerial images by using geometric - stochastic models and estimation, IEEE Computer Vision and Pattern Recognition, pp. 459-464, (1993).

[7] H. Long and Z. Zhao, Urban road extraction from high-resolution optical satellite images, Int. J. Remote Sens., vol. 26, no. 22, pp. 4907–4921, (2005).

[8] W. Shi, Z. Miao, and J. Debayle, An integrated method for urban main-road centerline extraction from optical remotely sensed imagery, IEEE Trans. Geosci. Remote Sens., vol. 52, no. 6, pp. 3359–3372, Jun. (2014).

[9] E. Christophe and J. Inglada, Robust road extraction for high resolution satellite images,in Proc. IEEE Int. Conf. Image Process, vol. 5, pp. 437–440, (2007).

[10] D. Chaudhuri, N. Kushwaha, A. Samal, Semi-automated road detection from high resolution satellite images by directional morphological enhancement and segmentation techniques, IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 5 , 1538–1544,(2012).

[11] L. J. Quackenbush, A review of techniques for extracting linear features from imagery, Photogramm. Eng. Remote Sens., vol. 70, no. 12, pp. 1383–1392, Dec. (2004).

[12] C. Poullis and S.You, Delineation and geometric modeling of road networks, ISPRS J. Photogramm. Remote Sens., vol. 65, no. 2, pp. 165– 181, Mar. (2010).

[13] A. Gruen and H. Li, Semi-automatic linear feature extraction by dynamic programming and LSB-snakes,Photogramm. Eng. Remote Sens., vol. 63, no. 8, pp. 985–995, Aug. (1997).

[14] F. Dell'Acqua and P. Gamba, Detection of urban structures in SAR images by robust fuzzy clustering algorithms: The example of street tracking, IEEE Trans. Geosci. Remote Sens., vol. 39, no. 10, pp. 2287–2297, Oct. (2001).

[15] R. Nevatia and K. Babu,,Linear feature extraction and description, Comput. Graph. Image Process., vol. 13, no. 3, pp. 257–269, Jul. (1980).

[16] K. Treash and K. Amaratunga, Automatic road detection in gray scale aerial images,ASCE J. Comput.Civil Eng., vol. 14, no. 1, pp. 60–69, (2000).

[17] J. Hu, A. Razdan, J. C. Femiani, M. Cui, and P. Wonka, Road network extraction and intersection detection from aerial images by tracking road footprints, IEEE Trans. Geosci. Remote Sens., vol. 45, no. 12, pp. 4144–4157, Dec.(2007).

[18] J. Zhang, X. Lin, Z. Liu, and J. Shen, Semi-automatic road tracking by template matching and distance transformation in urban areas, Int. J. Remote Sens., vol. 32, no. 23, pp. 8331–8347, Dec. (2011).

[19] S. Movaghati, A. Moghaddamjoo, and A. Tavakoli, Road extraction from satellite images using particle filtering and extended Kalman filtering,IEEE Trans. Geosci. Remote Sens., vol. 48, no. 7, pp. 2807–2817, Jul. (2010).

[20] A. Barsi and C. Heipke, Artificial neural networks for the detection of road junctions in aerial images,ISPRS Arch., vol. 34, pt. 3/W8, pp. 113–118, Sep. (2003).

[21] M. Ravanbakhsh, C. Heipke, and K. Pakzad, Road junction extraction from high resolution aerial imagery,Photogramm. Rec., vol. 23, no. 124, pp. 405–423, Dec. (2008).

[22] J. J. Ruiz, T. J. Rubio, and M. A. Urena, Automatic extraction of road intersections from images based on texture characterisation, Survey Rev., vol. 43, no. 321, pp. 212–225, Jul. (2011).

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[23] Krizhevsky, A., Sutskever, I., and Hinton, G. E. ImageNet classification with deep convolutional neural networks. In NIPS, pp. 1106–1114, (2012).

[24] Karen Simonyan& Andrew Zisserman, Very deep convolutional networks for large scale image recognition, ICLR (2015).

[25] Weixing Wang , Nan Yang, Yi Zhang, Fengping Wang, Ting Cao, PatrikEklund, A review of road extraction from remote sensing images, Journal of traffic and transportation engineering, (2016).

[26] Mengmeng Li, Alfred Stein, WietskeBijker and Qingming Zhan, Region-based urban road extraction from VHR satellite images using Binary Partition Tree, International journal of Applied Earth Observation and Geoinformation, 44, (2016).

[27] RashaAlshehhi, Prashanth Reddy Marpu, HIerarchical graph-based segmentation for extracting road networks from high-resolution satellite images, ISPRS journal of Photogrammetry and Remote Sensing, 126, (2017).

[28] Ruyi Liu, Jianfeng Song, Qiguang Miao, PengfeiXu and Qing Xue, Road centerlines extraction from high resolution images based on an improved directional segmentation and road probability, Neurocomputing,(2016).