

# A Convolutional Neural Network Approach for Road Anomalies Detection in Bangladesh with Image Thresholding

Taoseef Ishtiak, Sajid Ahmed, Mehreen Hossain Anila, Tanjila Farah  
Department of Electrical and Computer Engineering, North South University, Bangladesh  
{taoseef.ishtiak, sajid.ahmed1, meehreen.anila, tanjila.farah}@northsouth.edu

**Abstract**— Machine Learning is becoming an important tool for road condition detection. It is used in Road Information System (RiS), Road Weather Information System (RWiS), Intelligent Transportation System (ITS). All these systems are based on image classification and detection. Road images vary country wise. This paper aims to propose a road condition detection system for less developed countries such as Bangladesh. The research work presents a novel approach of using the Fully Convolutional Neural Network (FCNN) algorithm to detect the types of roads from street images of Bangladesh, detect the anomalies of the roads and classify the roads accordingly with black and white image thresholding. Based on the dataset of the roads of Bangladesh, a model has been trained to detect five classes of roads of Bangladesh. The model has been tested and validated by roads of other countries and showed an average accuracy of 87% for all classes. The goal of the research is to detect the condition of streets or roads of Bangladesh, based on structural and surface condition, material characteristics and classify them accordingly, from perfect roads to severely damaged road.

**Keywords**— Machine learning, Convolutional Neural Network (CNN), Fully Convolutional Neural Network (FCNN), Anomaly detection, Road Weather Information System (RWiS), Road Information System (RiS), Intelligent Transportation System (ITS).

## I. INTRODUCTION

According to World Bank the traffic congestion in Dhaka is wasting around 3.2 million working hours daily and costing the economy billions of dollars. Increased population, under construction roads, and civilian's disobedience to traffic rules are the major reasons for traffic congestion. While choosing which road to take to reach destination, google maps helps find the fastest path now a days. But google maps is unable to detect under construction roads. The main motivation of this work is to provide a system that can also consider the road condition while suggesting the best route to the users. The goal of the research project is to detect roads condition from images of roads and placing the images into certain classification where the images fit the best to determine how well or bad the condition of the road is. The classes have been termed as: 1. Perfect Roads, 2. Good Roads, 3. Bad Roads, 4. Severely Bad Roads 5. Water on Surface Roads. To achieve this, we have applied the convolutional neural network approach, which is a very widely used machine learning algorithm to classify images [1], noise detection [2] and in many other machine learning projects as well. Our model thus uses the Fully Convolutional Neural Network to classify the roads and black and white image thresholding to detect the anomalies of roads such

as: cracks, building material of the roads from road-color, width, surface smoothness and presence of other objects on roads. The purpose of the project stands in a humanitarian ground where the project can be run on a web platform to let other users know about the condition of the roads of a certain place where the user might never visited before, also to let the authority of any city or town to know about certain area's bad road condition, the project is believed to bring great benefit to make the whole system automated, by classifying images based on our model so that necessary steps can be taken upon the emergencies. The model is believed to be implemented for autonomous cars to avoid anomalies on roads and control the speed, based on road condition. We are also confident to use the model for any country's purpose of making Road Weather Information System, Road Information System and Intelligent Transportation System. The project is primarily built on the datasets obtained from the road conditions of 53 stations of Bangladesh, which is believed to cover all types of roads including: highways, hilly tracts, local roads, bad roads and of other criteria as well. We have also tried to test the model with as many types of roads possible and determine the class of the roads.

The paper is organized as following: the literature review is discussing in section II. The working methodology of this system is discussed in section III. We discussed the preparation of training data set in section IV. The dataset categories are introduced in section V. The score function for road segmentation and classification is explained in section VI. The results obtained from training and testing dataset is analyzed in section VII. We conclude in section VIII.

## II. LITERATURE REVIEW

A number of researches have been performed to detect road, road-lane, road area extraction [3], rural roads detection [5], to detect number of objects on streets, vehicle detection and passer by detection, street signs detection. The works are most commonly based on classification and beneficial for autonomous driving. To build the methodology of our project we have gone through number of similar works on image processing [14], CNN based works and image classification research done before. For semantic reasoning and computational power efficiency [3][5] is a very good approach for real-time joint semantic reasoning for street classification [16], vehicle detection and road segmentation. For surface crack detection, in civil engineering projects, deep convolutional neural network approaches like; image classification and bounding box approaches is proposed in vision-based models. A crack

detection method likewise; has been proposed in [6] based on semantic segmentation on concrete crack images which is also a fully convolutional network (FCN). For detecting road surface condition [7] has used vibration and GPS sensor data to achieve the condition of road surface and then trained the data with a simple neural network-based model. Another robust work has been proposed to determine rural roads along with its boundary based on improved FCM [11] (Fuzzy C-Means) and Hough Transformation, where the RGB image was first transformed into HIS (Hue, Intensity, Saturation analysis) space. The image saturation histogram peak in HIS color space the road boundary has been detected with Hough Transformation. For road surface condition detection on various environmental effects like: dry, wet, snow, ice and water [15] has proposed a combined approach with SVM (Support vector machine) algorithm connected with a grid search algorithm PSO (Particle Swarm Optimization) to optimize the kernel function factor and penalty factor of SVM. GRU (Grated Recurrent Units) has been implemented with CNN for road segmentation [12]. Also CNN has been found very efficiently used in noise detection [4], local filters [6] and nonlocal mean filters [6] while transforming domain filters extend over curvelets [5], wavelets [5], principal components analysis with PCA-net which is also very much equivalent to CNN [6] based filters and block matching and 3-D filtering method [4]. Sparse de-noising model was proposed in [4] meanwhile applied for adaptive learning on image de-noising.

### III. METHODOLOGY

We have applied a two-step methodology to train our whole model to fit for anomalies detection in roads. In first step we feed our images into a 3 step 32 by 32 CNN layers to train the model with our dataset of 5,815 training set images. The raw data set needed preprocessing with turning digital images into analog via image gray-scaling and then thresholding, for detecting the cracks and other anomalies present in the images and the severity of the anomalies. Our training set data has been classified based on a number of parameters mentioned in Table I.

TABLE I. TABLE PARAMETERS CONSIDERED TO DETECT THE ROADS CONDITION

Sl No.	Name of the Parameter
1.	Cracks and their types
2.	Width of the Road
3.	Color of the Road in bright daylight
4.	Average weather condition of the place
5.	Height and surface condition
6.	Other anomalies standing in the road, e.g. drainage system and others.

The threshold images were synced accordingly with the training set RGB images for image segmentation [13][14] with respect to change of light intensity. The image thresholding works in a way that it takes a threshold value after image gray-scaling, [4] if a pixel value crosses a certain cut off, it turns the pixel into black or keeps it white otherwise [4]. We have preprocessed our training image to detect the cracks and other anomalies of roads observing the RGB image and using openCV libraries, we trained our model to get that cut-off value when the model will detect a certain change of pixel as crack or anomalies. Our model

has been trained and the preprocessing has been done to mark the road area as white threshold and mark the cracks and other anomalies as black threshold, as shown in Figure 1.



Figure 1. Image obtained from camera applied for thresholding.

After image transformation, 5,815 images have been analyzed for types of anomaly detection on roads, and our observed anomalies have been mentioned in Table I. We used a score function mentioned in Section VII which took the featured anomalies as input with respective defined weights for the training image matrix and generated a numerical value to classify the training set for us. The range of those values, helped us classify the primary training images of 5 classes from perfectly built roads, such as highway-RCC roads, Rigid Roads towards, roads which have few cracks, gradually towards imperfect roads and lastly roads that contain water in surface. The imperfect roads in our dataset are considered unfit for transportation. Following the anomaly features of Table I we have analyzed the cracks, roads' width, surface smoothness, building material, weather effects on roads like: presence of water, sand on road surface and train the model on this basis.

For building the model we have followed a two-step approach to classify the images. Step one was feeding the images to a deep FCNN (Fully Convolutional Neural Network) and training the model. We experimented with various range of epochs and lastly, we got our best result with respect to our computational resource, with 100 epochs with an average accuracy of 87% for all classes. Step-2 was training the model with threshold images, to learn the cutoff values to detect anomalies of roads with respect to different light intensity of the image. The 5,815 images that were primarily analyzed, we did thresholding for each of the images and recorded the cutoff values to all those images similar to the image shown in Figure 1. This method works by selecting a threshold value in the grayscale range which acts as a cutoff for determining which pixels should be converted to white or black in the resulting image as shown in Figure 1. The thresholding cutoff parameter value has been recorded for each image. We then built a simple model to predict the thresholding value for new images and applied thresholding for the validation set images. From our observation the thresholding has given satisfactorily good result to detect the presence of anomalies and roads in images of the four classes. We excluded thresholding for the roads that have surface water as our purpose was not to detect the cracks of this class but to train our model to identify roads that have water on the surface. For thresholding the images, the model was trained converting the good part of the road with transforming pixels to white and the anomaly part to black.

After thresholding the change of pixel intensity from black to white within the image has been detected and observed in the graphical presentation like Figure 2. A median filter modifies an image pixel such that the new pixel value is selected as the median of the neighboring values within a range and plots the output in a graph showing the transition of pixel density.

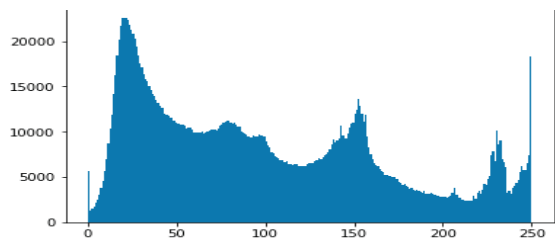


Figure 2. Image obtained from camera applied for thresholding.

From Figure 2 we could also observe a similar curve of pixel changing but the value segment of 0-50, 50-100, 100-150, 150-200, 200-255; present on the graph indicates which part of the images have more change in pixel. Figure 3 shows the work flow to build our methodology and training procedure.

Figure 3, represents the research methodology behind building the proposed architecture. After acquiring the training set images, the images were fed into the proposed FCNN model for validation. If the validation result works correctly, the following steps mentioned from Figure 3 were performed, otherwise the images were prepared for preprocessing and training again. Some of the acquired images, due to low quality issues were excluded from dataset in this step. If the image gets rightly classified in the validation step, the steps to mark the anomaly part in the images were to be performed. For acquiring this objective, the images were again preprocessed for training, by turning the RGB images into greyscale and then black and white thresholding. The thresholding has been done keeping the good part of the road as white threshold while the cracks and anomaly parts into black threshold with OpenCV library [13]. The cutoff value of the threshold image was recorded for further training. Then the thresholded images was trained as object detection model using the same architecture shown in Figure 4. The anomaly part was marked using Region of Interest (ROI) pooling [13]. So, after applying convolution steps, the ROI pooling was done and this went through the model's fully connected layers to train this part of the model with bounding box regression and then detect the anomaly part with bounding box.

Figure 4 illustrates the proposed FCNN model architecture. The RGB image was directly fed into a 3 x 3 convolutional layer using 32 filters. To reduce the dimension of parameters and computation in the network, 2 x 2 max pooling was applied. 20% dropout was applied to avoid overfitting in training model. Afterwards, the output of convolutional layer, gets flattened, to be used as single long feature vector to be used by the dense layer for final classification.

FCNN network that we have built consist of 32 by 32 convolutional layers with 5 output layers. We used 32 size filter, 2D MaxPooling with 2 by 2 pool-size, 512 dense

layers with activation function Relu with SoftMax classifier. In the dense layers, two step fully connected layers of size 512 with activation layer 'Relu' and have been kept after going through several trial and error steps to develop the network architecture.

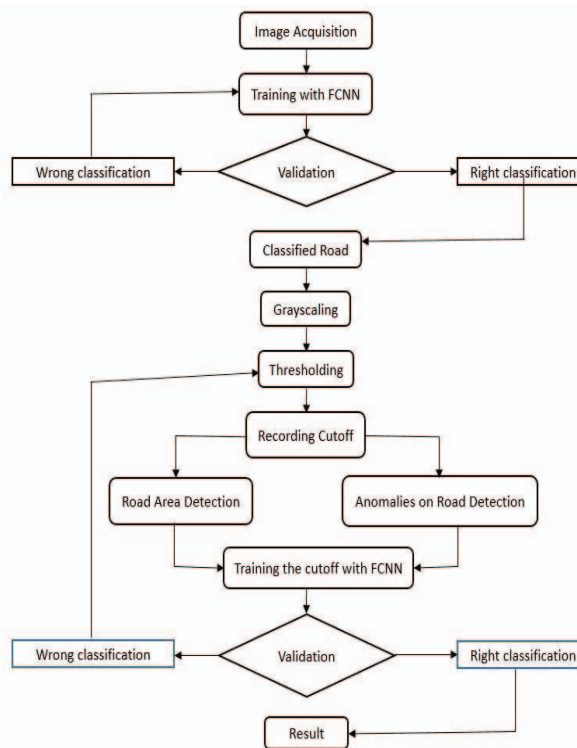


Figure 3. Methodology to build the proposed model.

The network at the final step uses 'Softmax' classifier to bring the output and classify the images correctly.

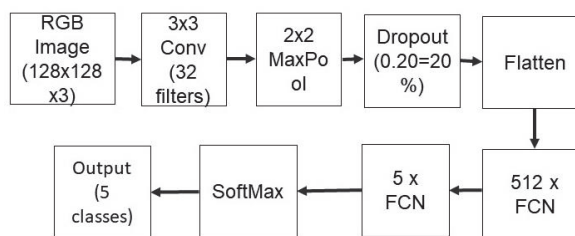


Figure 4. FCN Architecture.

The images' size was converted to 128 by 128 pixels. The model has been trained by Intel core i5 8th gen 6 core @4.00 GHz 32 GB DDR4 RAM PC NVidia GTX 1070 8 GB GDDR5 Memory: Dedicated 8GB + 8 GB Shared=16 GB GPU Memory in total. The training time with 100 epochs with 5,815 images took 130.43 minutes to train the cutoffs for thresholded images.

#### IV. IMAGE ANALYSIS

For image analysis, we considered the parameters present in Table I and classify the images based on our score function explained from Figure 7 in section V. The training set data has been fed into our convolutional neural

network. The network again is fed with images generating matrices of pixels from the black and white thresholded images shown in Figure 1. The network learns how the anomaly features of Table I present on a particular image and score the image on its own with forward and cross validation with backward propagation. With back propagation the network analyzes the cost or loss calculations of the scores that the network learned on itself and the score we have given to the analyzed training set. From image segmentation, the network will later classify the images according to the architecture shown in Figure 5.

## V. DATASET

### A. Brief Description:

Our training, validation and test dataset model consist of 8,370 images. An abstract view of our acquired dataset and how each class image looks like is presented with Figure 6. Among the acquired images, 5,815 images have been allocated for training data and have split into 5 classes, 625 of the rest images have been allocated for validation set and another rest 1,930 image has been allocated for test set as mentioned in Table II.

anomalies existing. For the classification of training set we considered all of the factors. Figure 6 shows a brief idea what types of images the dataset holds. The statistics of train, validation and test image set is compiled on Table II.

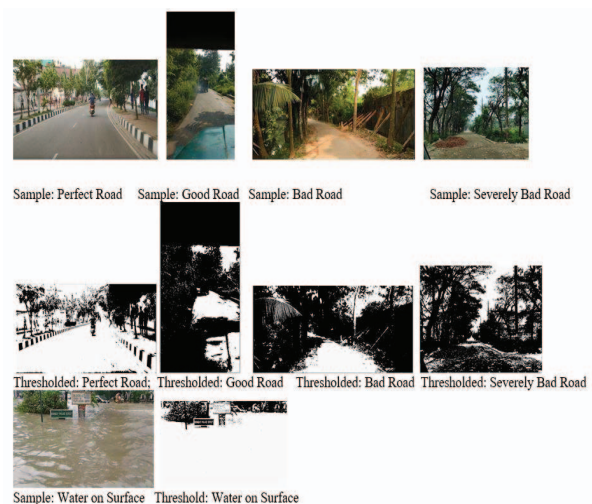


Figure 6. Dataset abstract view: Sample 5 classes of Roads

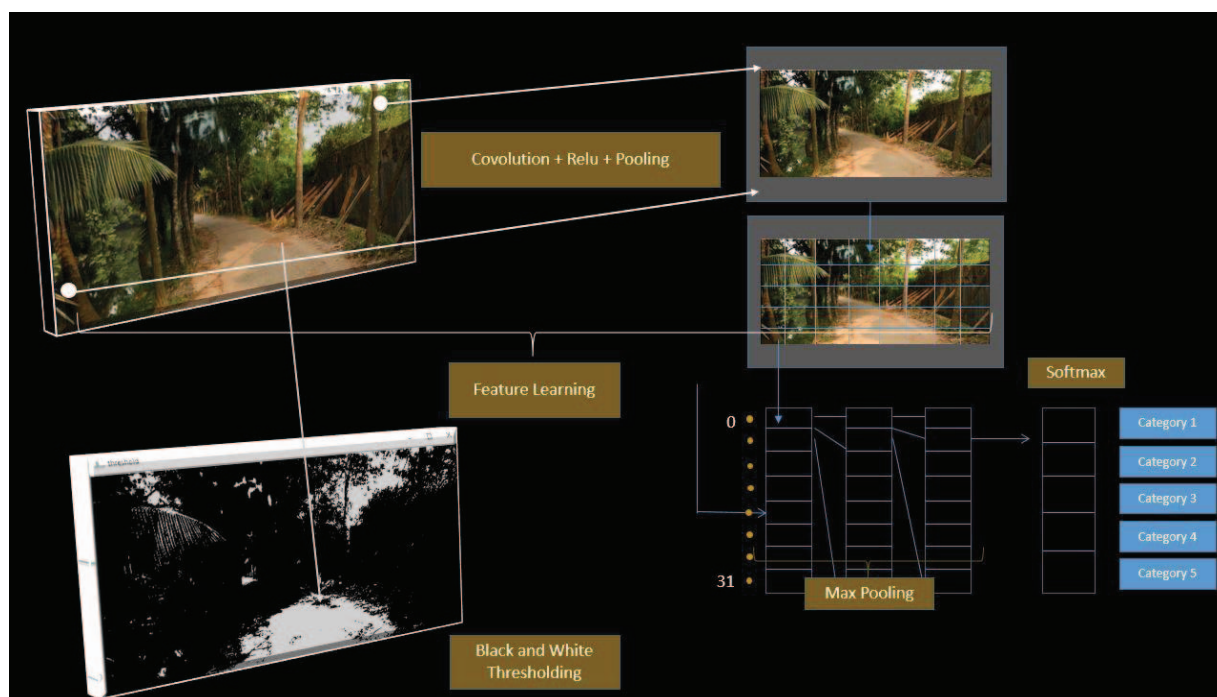


Figure 5. FCNN feature learning and classification steps.

6,785 images of the dataset have been collected by our research team over a time period of 4 months, another 1,285 images have been acquired from Local Government Engineering Department (LGED) of Government of People’s Republic of Bangladesh [8] and 300 images have been acquired from google images [9]. Our collected dataset contains images of 53 different stations inside Bangladesh of various kinds of roads and streets where there are both perfect roads and roads with many kinds of

TABLE II. DATASET SPLIT

Classes	Training	Validation	Test
Perfect Road	1,245	150	450
Good Road	930	150	450

Bad Road	1,360	125	450
Severely Bad Road	1,700	130	450
Water on Surface	580	70	130

**B. Description of Training Classes:**

1. Perfect Roads: The best category RCC and Rigid roads of Bangladesh most likely; the highway roads have been kept under this category. The road surface is very smooth, road width and surface material are best for transportation.
2. Good Roads: Under this class, the good local area roads have been allocated. 930 pictures have been trained for this class. The roads are slightly narrow compared to highway perfect roads, but made of RCC roads. The good roads of hilly tracts of Bangladesh have been allocated under this class as well.
3. Bad Roads: Roads with slight cracks, sand on surface, made with bricks have been allocated under this class. 1,360 images have been trained for this class.
4. Severely Bad Road: Roads which can be considered unfit for transportation have been allocated for severely bad roads. 1700 images have been identified for training for this class.
5. Water on Surface Road: We have experimented our model to learn to identify roads containing water taking motivation from [15]. 580 images have been collected for this category.

**C. Image processing on Dataset**

Based on the dataset of over 4000 images of roads of Dhaka city, a statistical analysis has been run on the training set images to understand the change of image preprocessing effects: gray scaling and thresholding over the change of surface, types of cracks, sunlight, and effect of shadow on roads. The images then have been scored based on the graphical analysis mentioned in Figure 1 and Figure 2 for categorization. Based on this categorical segmentation, the train data set is being processed to be fed into our convolutional layer. To analyze the image categorization, we have considered 8 parameters to analyze a road and score it accordingly based on the result generated from thresholding and gray thresholding. The parameters are listed below in Table. III.

TABLE III. PARAMETERS FOR IMAGE STATISTICAL ANALYSIS

SI No.	Name of the Parameter
1.	Roads Based Theoretical Score: before processing the images
2.	Road Color: 7 different color identified for roads in Bangladesh
3.	Road Material: 5 different materials identified for roads in Bangladesh
4.	Fourier Based Score: 5 categories based on the acquired results

5.	Gray Threshold Score: segmented into 5 categories based on the acquired results
6.	Road Width: 3 types of width analyzed
7.	Picture Quality: split into 3 qualities of images
8.	Relevancy of the image (based on the visibility of road in the image): 3 segmentation

Based on Table III, primarily 1,670 images have been categorized to understand the change of result for different types of images of roads. The model is being trained to filter out irrelevant images where the roads are not properly visible or false images being fed as input.

The steps mentioned in Figure 7, were followed for our score function to understand the type of road before feeding the images into training set. The score function was used in the project to classify any road for training set where human intervention was not very much needed. The steps in Figure 7 illustrate that, after picture acquisition, a theoretical score was given, based from ground the observation of roads which is weighted with 'x' term in the function. The road color was taken into consideration afterwards, which was termed as 'y'. Using OpenCV libraries, the Fourier Transformation based score and Threshold value, detecting the surface condition were applied in the function. These values have been termed as "z" and "w" respectively. Later on, the road width from observation, and picture quality termed as 'm' and 'n' respectively, were considered. Finally, to detect fault images, the proposed function has a separate parameter named as 'picture relevancy' which has been termed as 'p'. These terms were used and weighted following the score function mentioned in Section VI.

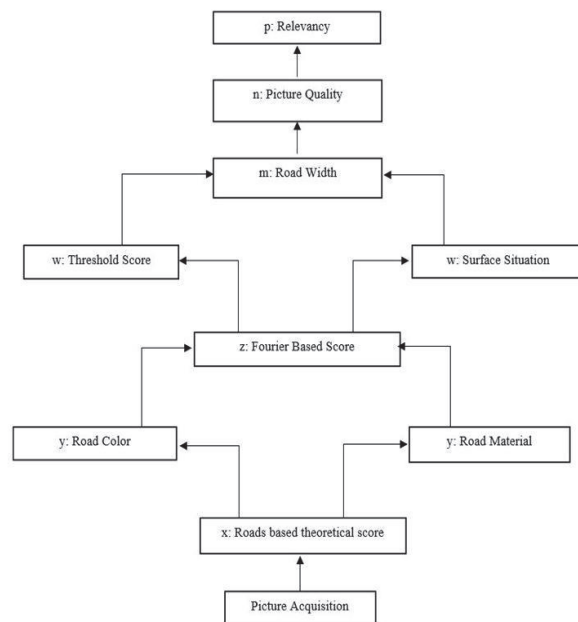


Figure 7. Flowchart of Statistical analysis.

## VI. SCORE FUNCTION

For road segmentation and classification, we have come up with a score function:

$$h(\Theta) = I [\Theta_1x^2 + \Theta_2y^3 + \Theta_3z + \Theta_4w + \Theta_5m + \Theta_6n + \Theta_7p]$$

Here,  $I$  represent the image matrix,  $\Theta$  represents road features and the variables  $x, y, z, w, m, n, p$  represents weights of the features. The score function helped us to understand and classify our train image set more perfectly. The score function has been generated following the research [10] where linear discriminant function has been used to detect the surface roughness, presence of water based on works done by [15], [16], [17]. The higher degrees in the polynomial function represents priority of the particular parameters. Likewise; on the basis of our observation of roads in Bangladesh, the road color in bright daylight is more effective to understand the condition of the road than other parameters. That is why the road color has higher priority than other parameters. The road based theoretical score is a weight given by us, experiencing the on-field road condition of ourselves. The polynomial function sets it parameter values on weight basis based on the parameters mentioned in Table III.

## VII. RESULT ANALYSIS

The dataset was split into training, validation and test sets, as mentioned in Section V. The proposed machine learning model's training and test set result has been compiled in Table IV and shown in Figure 8.

TABLE IV. RESULT SUMMARY FOR TRAINING AND TESTING DATA MODEL FOR ALL THE CLASSES.

Classes	Train Accuracy	Test Accuracy
Perfect Road	97%	89%
Good Road	93%	82%
Bad Road	94%	87%
Severely Bad Road	96%	82%
Water on Surface	91%	77%

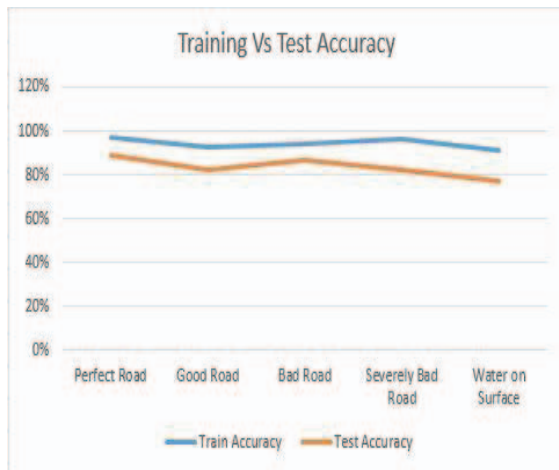


Figure 8. Train Vs Test Accuracies for Five Classes.

The model has been experimented from 0 to 100 epochs and the accuracy sharply raised from 5th epoch and flattened till 100 epochs. The whole dataset has been fed into our neural network model in each epoch. First, the low accuracy and high loss indicates our model was being underfit till 5<sup>th</sup> epoch. But with correcting the respective learning weights the network reached its optimal stage in 100<sup>th</sup> epoch. So, we kept our training till 100 epochs to avoid overfitting. The results mentioned in Table IV, are based on these 100 epochs. The performance of cost or loss function in the graph shown in Figure 9 describes gradual decrease in loss from 5th epoch. The figure also shows the performance of epoch and loss function of our CNN model.

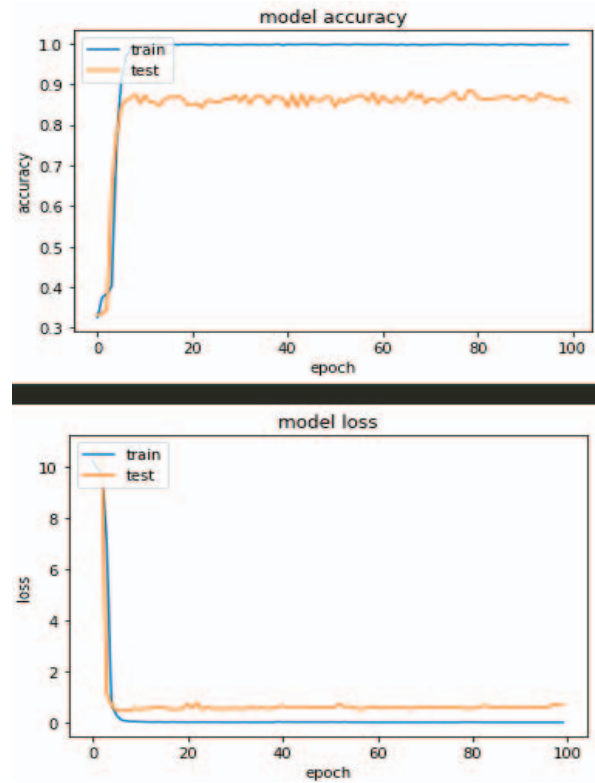


Figure 9. Model accuracy performance with respect to epoch and loss function.

The trained model has been built to detect the anomalies of roads through black and white thresholding. Figure 10 shows the detected anomalies on thresholded images for the five classes of our images.





Figure 10. Test Set results of classified road and anomalies detected with image black and white thresholding

### VIII. CONCLUSION

The test accuracy of our model is an average of 87% for all the classes which is believed to be satisfactory. We have developed a web interface to open the model for public use, so that new image data can be shared by users. The data will at first work as test set and later after validation, the data will be trained for increasing accuracy. One limitation of our model has come that, we have not been successful to avoid the shadow on roads while thresholding. That is why some thresholded results came up detecting object shadow as cracks. We also could not use thresholding for roads that have water on surface. Our future goal contains to resolve these limitations.

### REFERENCES

- [1] A. Goswami and H. Liu, "Deep Dish: Deep Learning for Classifying Food Dishes," (2017) [Online]. Available: <https://www.semanticscholar.org/paper/Deep-Dish-%203A-Deep-Learning-for-Classifying-Food-Goswami-Microsoft/2e596abd9d4616f75c9da45d88787cb2fd5bee81>. [Accessed April 29, 2019].
- [2] J. G. Haran, J. Dillenburger, and P. Nelson, "Realtime Image Processing Algorithms for the Detection of Road and Environmental Conditions," *Ninth International Conference on Applications of Advanced Technology in Transportation*, AATT 2006, Chicago, Illinois, August 13-15, 2006, pp. 55-60.

- [3] Z. Zhang, Q. Liu, Y. Wang, "Road extraction by deep residual U-net", *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 5, pp. 749-753, May 2018.
- [4] J. H. Chuah, H. Y. Khaw, F. C. Soon and C. Chow, "Detection of Gaussian noise and its level using deep convolutional neural network," in *Proc. 2017 IEEE Region 10 Conference, TENCON 2017*, Penang, Malaysia, Nov. 5-8, 2017, pp. 2447-2450.
- [5] Z. Wan-zhi and W. Zeng-cai, "Rural Road Detection of Color Image in Complicated Environment," *International Journal of Signal Processing, Image Processing and Pattern Recognition*, Vol.6, No.6, pp.161-168, 2013.
- [6] M. Teichmann, M. Weber, M. Zöllner, R. Cipolla and R. Urtaun, "MultiNet: Real-time Joint Semantic Reasoning for Autonomous Driving," *IEEE Intelligent Vehicles Symposium, IV 2018, Changshu, China, June 26-30, 2018*, pp. 1013-1020.
- [7] C. V. Dung and L. D. Anh, "Autonomous concrete crack detection using deep fully convolutional neural network," *Automation in Construction*, Volume 99, Pages 52-58, March 2019.
- [8] Local Government Engineering Department (LGED). <http://www.lged.gov.bd/> [Accessed November 11, 2018]
- [9] Google Images. <http://www.images.google.com>
- [10] J. Eriksson, L. Girod, B. Hull, R. Newton, S. Madden, and H. Balakrishnan, "The pothole patrol: using a mobile sensor network for road surface monitoring," *In Proceedings of the 6th international conference on Mobile systems, applications, and services (MobiSys '08)*, New York, USA, June 17 - 20, pp. 29-39.
- [11] J. Zhao, H. Wu, and L. Chen, "Road Surface State Recognition Based on SVM Optimization and Image Segmentation Processing," *Journal of Advanced Transportation*, Vol 2017, pp. 1-21, 2017.
- [12] Yecheng L. and Xinming H. "Road Segmentation Using CNN with GRU" *22nd International Conference on Information Visualization, 2018*.
- [13] K. Kang, X. Wang "Fully Convolutional Neural Networks for Crowd Segmentation," *Computing Research Repository (CoRR)*, arXiv, Cornell University 1411.4464, 2014.
- [14] K. Radhika, P. Vishalini "DIGITAL IMAGE PROCESSING – Sequence, Components and Pros" K. Radhika et al, / *(IJCSIT) International Journal of Computer Science and Information Technologies*, Vol. 7 (2), 2016, 922-924.
- [15] M. Yamada, K. Ueda, I. Horiba, S. Yamamoto and S. Tsugawa "Detection of Wet-Road Conditions from Images Captured by a Vehicle-Mounted Camera" *Journal of Robotics and Mechatronics* Vol.17 No.3 pp. 269-276 doi: 10.20965/jrm.2005.p0269
- [16] S. Yoda, H. Okabe, J. Takagi, and T. Yamashita, "Road surface recognition sensor using an optical spatial filter," in *Proceedings of the IEEE Conference on Intelligent Transportation System*, pp. 253–257, 1997.
- [17] K. Ueda, I. Horiba, and K. Iedaga, "A detecting method of wet condition on road using image processing," *Information Processing Society of Japan*, vol. 35, no. 6, pp. 1072–1080, 1994.