

UNLEASHING THE POTENTIAL OF IMAGE FOR POTHOLE DETECTION

Submitted in partial fulfillment of the requirements for the award of
Bachelor of Engineering Degree in Electronics and Communication Engineering

by

PRABHAVATHY G (40130177)

SIVADHARSHINI S (40130221)



**DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING
SCHOOL OF ELECTRICAL AND ELECTRONICS**

SATHYABAMA

**INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)**

CATEGORY -1 UNIVERSITY BY UGC

Accredited with grade “A++” by NAAC | 12B Status by UGC | Approved by AICTE

JEPPIAAR NAGAR, RAJIV GANDHI SALAI, CHENNAI – 600119

APRIL- 2024



SATHYABAMA

**INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)**

Accredited "A++" Grade by NAAC | 12B Status by UGC | Approved by AICTE

www.sathyabama.ac.in

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

BONAFIDE CERTIFICATE

This is to certify that this Project Report is the bonafide work of PRABHAVATHY G (40130177) and SIVADHARSHINI S (40130221) who carried out the project entitled "UNLEASHING THE POTENTIAL OF IMAGES FOR POTHOLE DETECTION" under our supervision from November 2023 to April 2024.

Internal Guide

Dr.K.T. ILAYARAJAA, M.E., Ph.D.

Head of the Department

Dr. T. RAVI M.E., Ph.D.

Submitted for Viva voce Examination held on _____

Internal Examiner

External Examiner

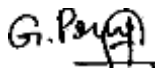
DECLARATION

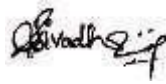
We, PRABHAVATHY G (40130177) and SIVADHARSHINI S (40130221) hereby declare that the Project Report entitled “Unleashing the potential of image for pothole detection” done by us under the guidance of Dr.K.T.ILAYARAJAA,M.E.,Ph.D., Assistant Professor, Dept.of Electronics and Communication Engineering, SATHYABAMA INSTITUTE OF SCIENCE AND TECHNOLOGY, CHENNAI, is submitted in partial fulfillment of the requirements for the award of Bachelor of Engineering degree in Electronics and Communication Engineering.

DATE:18.04.2024

PLACE: Chennai

SIGNATURE OF CANDIDATES

1. 

2. 

ACKNOWLEDGEMENT

We are pleased to acknowledge our sincere thanks to the **Board of Management of Sathyabama Institute of Science and Technology** for their kind encouragement in doing this project and for completing it successfully. We are grateful to them.

We convey our thanks to **Dr. N. M. NANDHITHA, M.E., Ph.D. Dean, School of Electrical and Electronics Engineering** and **Dr. T. RAVI, M.E., Ph.D., Head of the Department, Department of Electronics and Communication Engineering** for providing us necessary support and details at the right time during the progressive reviews.

We would like to express my sincere and deep sense of gratitude to our Project Guide **Dr. K.T.ILAYARAJAA,M.E.,Ph.D.**, Assistant Professor, Dept. of Electronics and Communication Engineering for his valuable guidance, suggestions and constant encouragement for the successful completion of our project work.

We wish to express my thanks to all Teaching and Non-teaching staff members of the Department of Electronics and Communication Engineering who were helpful in many ways for the completion of the project.

ABSTRACT

Potholes can generate damage such as flat tire and wheel damage, impact and damage of lower vehicle, vehicle collision, and major accidents. Thus, accurately and quickly detecting potholes is one of the important tasks for determining proper strategies in ITS (Intelligent Transportation System) service and road management system. Several efforts have been made for developing a technology which can automatically detect and recognize potholes. Roads are the main mode for transportation, now a days. As we use roads heavily and frequently it sometimes leads to potholes. Environmental factors also lead to potholes in several ways. These potholes are main reason for many accidents. A Scheduled and proper maintenance is required where we need to monitor each and every road. As this maintenance of all the roads at a time is not possible because it is not easy to monitor every single place or just because people ignore to check which causes formation of potholes that causes unnecessary traffic and many of accidents. To monitor all the roads this project for pothole detection using image processing techniques implemented. Techniques of Image processing which detects the potholes on roads and save the data of pothole for road maintenance department. This helps in keeping manual labor to the minimum number. Canny Edge Detecting technique and alone with the Box Detecting Technique are techniques of Image Processing we use in the proposed system. Hough transformation technique in the end gives effective output of potholes, very accurately.

TABLE OF CONTENTS

CHAPTER No.	TITLE	PAGE No.
	ABSTRACT	v
	LIST OF ABBREVIATIONS	ix
	LIST OF FIGURES	xi
	LIST OF TABLES	xii
1	INTRODUCTION	
	1.1 Pothole detection	1
	1.2 Objectives	2
	1.3 Motivation	2
2	LITERATURE SURVEY	4
	2.1 Automatic pothole detection	4
	2.2 Automatic detection and notification Humps on the road	6
	2.3 Smart sensing of pothole and obstacles	8
	2.4 Detection and notification on pothole and humps	9
	2.5 An approach for real time road deformation detection	11
	2.6 Pothole detection based on blob detection method	
3	EXISTING AND PROPOSED SYSTEM	15

3.1 Existing System	15
3.1.1 Disadvantages	15
3.2 Proposed System	16
3.2.1 Advantages	16
3.3 Flow chart	16
3.4 Comparison of Existing and Proposed System	17
3.5 Standards	19
3.6 Constraints	19
3.7 Trade offs	20
4 METHODOLOGY	21
4.1 Canny edge detection algorithm	21
4.1.1 Gaussian Filter	21
4.1.2 Intensity Gradient	22
4.1.3 Non-Maximum Suppression	23
4.1.4 Double threshold	24
4.1.5 Edge tracking by hysteresis	25
4.2 Image thresholding	25
4.2.1 Otsus's method	26
4.3 Hough transform	27
4.3.1 Theory	29
4.3.2 Example 1	30
4.3.3 Example 2	30
4.4 Feature extraction	31

4.5	General	32
4.6	Wavelets	32
4.7	Convolution Neural Network	33
4.8	What is CNN?	35
4.9	Why CNN?	36
4.10	System requirements for MATLAB	36
4.11	Block diagram	38
5	RESULTS AND DISCUSSION	39
5.1	Preprocessing Hsv Channel	39
5.2	Preprocessing RGB To Gray Conersion	40
5.3	Segmentation	42
5.4	Segmentation using Kmean and contour algorithm	43
5.5	Classification using CNN	44
6	CONCLUSION AND FUTURE ENHANCEMENT	
6.1	Conclusion	47
6.2	Future enhancement	47
	REFERENCES	

LIST OF ABBREVIATIONS

ABBREVIATION	EXPANSION
ITS	- Intelligent Transportation System
MLPNN	- Multi Layer Perceptron Neural Network
RFID	- Radio Frequency IDentification
FPGA	- Field Programmable Gate Arrays
GPS	- Global Positioning System
DSP	- Digital Signal Processing
MATLAB	- MATrix LABoratory
PCA	- Principal Component Analysis
NMR	- Nuclear Magnetic Resonance
FT	- Fourier Transform
CNN	- Convolutional Neural Networks
YOLO	- You Only Look Once
HSV	- Hue Saturation Value

LIST OF FIGURES

FIGURE No.	NAME OF THE FIGURE	PAGE No.
1.1	Pothole detection 1	3
1.2	Pothole detection 2	3
3.1	Flow chart	16
4.1	Gaussian smoothening	22
4.2	Hough transform	28
4.4	Example 1	29
4.5	Rendering of transform results	30
4.6	Convolution Neural Network	33
4.7	Training of Neural Network	34
4.8	Illustrate of biological neuron(left) and it's mathematical calculation(right)	34
4.9	Flow diagram	36
4.10	Histogram graph	37
4.11	Output display using YOLO	38
5.1	Preprocessing HSV channel	39
5.2	Preprocessing RGB to Gray conversion	40
5.3	Segmentation	42

FIGURE No.	NAME OF FIGURE	PAGE No.
5.4	Edge found in dilated image	43
5.5	CNN graph	44
5.6	Network estimation	45
5.7	Classification and loss	46
5.8	Confusion matrix	46
5.9	Accuracy of yolov7	46

LIST OF TABLES

TABLE No.	NAME OF TABLE	PAGE No.
3.1	Algorithm Comparison	18

CHAPTER 1

INTRODUCTION

1.1 POTHOLE DETECTION

Potholes are bowl-shaped openings on the road that can be up to 10 inches in depth and are caused by the wear-and-tear and weathering of the roads. They emerge when the top layer of the road, the asphalt, has worn away by lorry traffic and exposing the concrete base. Once a pothole is formed, its depth can grow to several inches, with rain water accelerating the process, making one of the top causes of car accidents. Potholes are not only main cause of car accidents, but also can be fatal to motorcycles. Potholes on roads are especially dangerous for drivers when cruising in high speed. At high speed, the driver can hardly see potholes on road surface. Moreover, if the car passes potholes at high speed, the impact may rupture car tires. Even though drivers may see the pothole before they pass it, it is usually too late for drivers to react to the pothole. Any sharp turn or suddenly brake, may cause car rollover or rear-end. Motivated from above reasons, we decided to investigate a system to detect potholes on roads while driving and the proposed system will produce the 3-dimensional information of potholes and determine the distance from pothole to car for informing the driver in advance. Also Currently, the main methods for detecting potholes still rely on public reporting through hotlines or websites, for example, the potholes reporting website in Ohio² . However, this reporting usually lacks accurate information of the dimensional and location of potholes. Moreover, this information is usually out of date as well. A method to detect potholes on road has been reported in a real-time 3D scanning system for pavement distortion inspection³ which uses high-speed 3D transverse scanning techniques. However, the high-speed 3D transverse scanning equipment is too expensive. Rajeshwari Madli et al. have proposed a cost-effective solution to identify the potholes on roads, and also to measure the depth and height of each pothole using ultrasonic sensors. All the pothole information is stored in database (cloud). Then alerts are provided in the form of a flash messages with an audio beep through android application. To detect the depth of pothole correctly, the ultrasonic sensor should be fixed under the car, which means the car should pass the pothole first vision- based solutions can detect potholes.

1.2 OBJECTIVE

To Identify the Pothole detection in real time. Lot of accidents are occurred by potholes, so this project objective is to detect the potholes and reduce accidents.

1.3 MOTIVATION

With the increase in world's population, there has been increasing load on the infrastructure. Roads have been flooded with the vehicular traffic. It has become increasingly difficult to manage this traffic. This is the prime motivation behind making a vehicle intelligent enough to aid driver in various aspects. One of the increasing problems the roads are facing is worsened road conditions. Because of many reasons like rains, oil spills, road accidents or inevitable wear and tear make the road difficult to drive upon. Unexpected hurdles on road may cause more accidents. Also because of the bad road conditions, fuel consumption of the vehicle increases; causing wastage of precious fuel. Because of these reasons it is very important to get the information of such bad road conditions, Collect this information and distribute it to other vehicles, which in turn can warns the driver. But there are various challenges involved in this. First of all there are various methods to get the information about the road conditions. Then this information must be collected and distributed to all the vehicles that might need this information. Lastly the information must be conveyed in the manner which can be understood and used by driver. Wein this project try to design and build such a system. In this system the access point collects the information about the potholes in the vicinity and store the information about the potholes for the future use. Ideally the vicinity is every rout till the next access point.

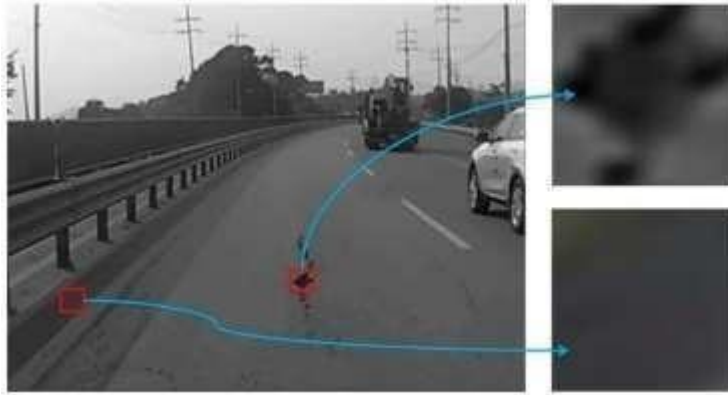


Fig : 1.1 Pothole detection 1

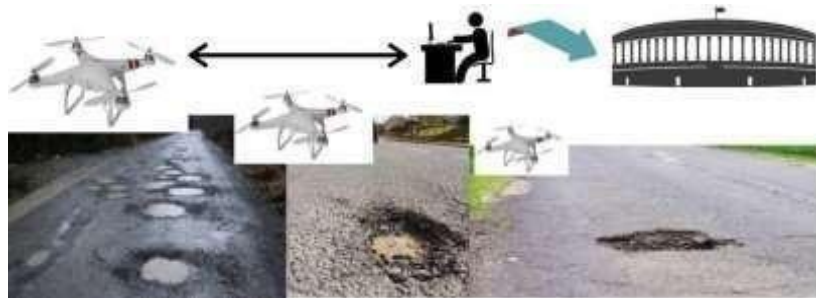


Fig :1.2 Pothole detection 2

CHAPTER 2

LITERATURE SURVEY

2.1 AUTOMATIC POTHOLE DETECTION

Most of the Indian rural and sub urban roads are not ideal for driving due to faded lanes, irregular potholes, improper and invisible road signs. This has led to many accidents causing loss of lives and severe damage to vehicles. Many techniques have been proposed in the past to detect these problems using image processing methods. But there has been little work specifically carried out for detecting such issues of Indian roads. Potholes can generate damage such as flat tire and wheel damage, impact and damage of lower vehicle, vehicle collision, and major accidents. Thus, accurately and quickly detecting potholes is one of the important tasks for determining proper strategies in ITS (Intelligent Transportation System) service and road management system. Several efforts have been made for developing a technology which can automatically detect and recognize potholes. In this project, a pothole two dimensional (2D) images based pothole detection method is used for improving the existing method. This system proposes a system Automatic detect the number of recently published papers dealing with crack detection and characterization of pavement surface distresses shows an increasing interest in this area. A recent publication a hierarchical method present in [1], which deals with detection of roads and slopes. In this paper, a novel framework is proposed for segmenting road images in a hierarchical manner that can separate the following objects: road and slopes with or without collapse, sky, road signs, cars, buildings and vegetation from the images. The experiments show that the approach in this paper can achieve a satisfied result on various road images. The roads are unstructured, which are more complex than the structured roads. In [2] multi scale approach based on Markov random field is proposed to segment fine structures (cracks) in road pavement surface images. Cracks are enhanced using a 1-D Gaussian smoothing filter and then processed by a 2-D matched filter to detect them. A total of 64 road pavement surface images representing several crack types are considered for experimentation, producing a qualitative evaluation. Details on image characteristics or the type of sensor used to capture them are not provided. Another paper [3] evidences the difficulty of cracks

detecting of less than 3 mm width when using edgedetectors. A non-sub sampled contourlet transform is the way that could be adopted in [4] to detect cracks, wherein a limited set of experimental results is presented. A complete methodology to automatically detect and characterize pavement defects is proposed in [5], using gray scale images captured by line scan cameras illuminated by lasers during road surveys performed using a high-speed image acquisition system. Crack detection uses a conditional texture anisotropy measurement to each image, and defect characterization uses a multilayer perceptron neural network with two hidden layers. The results presented are promising, but the experimental evaluation does not support the distinction of multiple cracks in the same image. In paper [6], Neural network method is used. The automated pavement defect detection can only identify crack type defects. To classify defect, a multi layer perceptron neural network (MLPNN) is used. Neural network is used to classify the images into four classes: defect-free, crack, joint and bridged. Experimental results are performed on real road images which are labelled by human operators. There are more additional filters required for this system. In paper [7], Vision-based approaches are used to address functionalities such as lane marking detection, traffic sign recognition, pedestrian detection, etc. This system is possible to detect the free road surface ahead of the ego-vehicle using an on board camera. Novelty Method is used for both Shadowed and un shadowed regions which provide highest performance. Road detection algorithm is devised by combining the illuminant invariant feature space and likelihood based classifier. The defect of this system is under saturation by improving image acquisition system. In paper [8], A neural network based technique for the classification of segments of road images into cracks and normal images. The features are passed to a neural network for the classification of images into images with and without cracks. Another approach extracts linear features (cracks) using two methodologies: one based on holistic thresholding and the second employing the Otsu algorithm.

Pros:

- **Informative Overview:** The content provides an informative overview of the challenges faced on Indian rural and suburban roads, specifically related to road conditions, potholes, and the need for technology-driven solutions.
- **Detailed Technical Insights:** It delves into various technical approaches and methods used for road condition detection and pothole detection, giving

readers a glimpse into the complexity and diversity of solutions.

- **Citation of Relevant Papers:** The content cites specific research papers and methods, which can be helpful for those looking to explore further research in this domain.

Cons:

- **Lack of Clarity:** The content lacks clarity in some areas and seems fragmented, making it difficult for readers to follow a coherent narrative. It jumps from one topic to another without a clear structure.
- **No Critical Evaluation:** While it presents various methods and approaches, it doesn't offer a critical evaluation or comparison of these methods, which could help readers understand their strengths and weaknesses.
- **Outdated Information:** The content does not mention any recent developments or advancements in the field beyond the knowledge cutoff date in September 2021. Road condition detection technology is a rapidly evolving field, and readers may miss out on the latest trends and breakthroughs.

2.2 AUTOMATIC DETECTION AND NOTIFICATION HUMPS ON

ROADS

One of the major problems in developing countries is maintenance of roads. Well maintained roads contribute a Major portion to the country's economy. Identification of pavement Distress such as potholes and humps not only helps drivers to avoid accidents or vehicle damages, but also helps authorities to maintain roads. This paper discusses previous pothole detection methods that have been developed and proposes a cost-effective solution to identify the potholes and humps on roads and provide timely alerts to drivers to avoid accidents or vehicle damages. Ultrasonic sensors are used to identify the potholes and humps and also to measure their depth and height, respectively. The proposed system captures the geographical location coordinates of the potholes and humps using a global positioning system receiver Taehyeong Kim and Seung- Ki [1] Ryu proposed a detection system which starts with noise removal, followed by adjustment of brightness and simplification of video by binarization. Then, noise removal is applied to the binarized image. After noise removal, the process of extraction of the outlines of the segmented objects is carried out. Extraction is followed by selection and square zoning for the objects. After all these processes, desired pothole area information is returned. Sudarshan Rode

[2] proposed a pothole detection system which is divided into three subsystems. First is sensing subsystem which senses the potholes encountered by it, by using accelerometer or by camera which scans the road. Both are mounted on the car. Then communication subsystem which transfers the information between Wi-Fi access point and mobile node. Access Point broadcasts the data about potholes in its area. Eriksson et al. [3] studied mobile sensing of roads to monitor and report any potholes. The system used accelerometer and GPS for detection and location respectively. Cars give detections which are fed to a central server.

Pros:

- **Highlighting a Critical Issue:** The content emphasizes the importance of road maintenance and the identification of pavement distress, which is a significant issue in developing countries. It underscores the economic impact of well-maintained roads.
- **Introduction of a Cost-Effective Solution:** It introduces a cost-effective solution for identifying potholes and humps on roads, which can potentially help reduce accidents and vehicle damages, making roads safer for drivers.
- **Mention of Existing Methods:** The content mentions existing pothole detection methods, providing context and showing that the proposed solution is built upon previous research.
- **Use of Sensor Technology:** The use of ultrasonic sensors and GPS receivers in the proposed system demonstrates the incorporation of modern technology for more accurate and reliable detection and location tracking.

Cons:

- **Lack of Technical Details:** The content briefly mentions the use of ultrasonic sensors and GPS receivers but does not delve into the technical specifics of how these sensors are employed or how the system works in detail. Readers may want more information on the technical implementation.
- **Limited Citations:** While it mentions previous work by other researchers, it could benefit from providing more comprehensive citations and referencing to support the claims made in the text.

2.3 SMART SENSING OF POTHOLES AND OBSTACLES

Byeoung-ho-Kang and Su-il-choi proposes the concept of sensing potholes by using 2d lidar method. It is a sensing method which uses light pulses to ascertain the surface of earth. The major drawback in this method is, it is highly affected by heavy rain, fog, etc. Also does not work well at huge reflections .The operating cost for this approach is comparatively high. The proposed the 3d laser method to detect potholes and obstacles. The 3D laser checking is one of the outstandingly flexible and productive advances for precisely catching extensive arrangements of 3D facilitates. This method uses laser pulses to detect the irregularities in road surfaces. It is applicable to 2d and 3d surfaces. The disadvantage of this approach is that it requires post-processing to produce a usable output i.e., the output requires manipulation. And high end hardware is required. AmilaAkagic, Emir Buza and Samir Omanavic used the RGB image processing technique, complex figuring devices that consolidate programming and equipment to process the pothole pictures are required in this technique. It is also consumes more time when compared to other methods. Also the accuracy rate is low in this approach. Moreover it needs prior knowledge of the images i.e., a large quantity of training sets have to be given. Wang didn't attainability concentrate to lead the comprehensive review of asphalt conditions by methods for of stereovision innovation. In this strategy, two advanced cameras are appended to the vehicle, which is utilized to cover an asphalt surface. The initial step is to dissect 2D pictures from every one of the two cameras to recognize and order any splits. To recuperate the 3D properties from given sets of 2D pictures on asimilar asphalt surface, the succession of steps, for example, camera alignment, twisting right, coordinating stereo focuses, 3D recreate, and profile report ought to be performed. It needs complex computation capabilities when compared to methods. Kana Azhary, FeerdMurtaza, Muhammad Herboon mohammed and Hafed Adman Habit stated an approach of finding and localizing the potholes based on computer vision in asphalt pavement images. Histograms from the input images are classified using naïve other bayes classifier using normalized graph cut segmentation scheme. This experimentation showed 90% accuracy on localizing the potholes from the pothole images.

Pros

- **Diverse Sensing Methods:** The content discusses various sensing methods for pothole detection, including 2D LiDAR, 3D laser, RGB image processing, and computer vision. This provides readers with a broad understanding of the available technologies.
- **Highlighting Advantages:** It mentions the advantages of each method, such as the flexibility and productivity of 3D laser, and the accuracy achieved in computer vision-based approaches

Cons

- **Lack of Technical Details:** The content lacks technical details about how each sensing method works. Readers may want more information on the underlying technology and algorithms used in these approaches.
- **No Comparison:** While it mentions the advantages and disadvantages of each method separately, it doesn't provide a direct comparison between them. A comparison could help readers understand which method might be more suitable for their specific needs.

2.4 DETECTION AND NOTIFICATION OF POTHOLE AND HUMPS

Vigneshwar. k et al. [1] has done Image pre-processing based on difference of Gaussian-Filtering and clustering based image segmentation strategies are actualized for better outcomes. The primary objective of this paper was to recognize a superior technique which was exceedingly productive and precise compared to the conventional techniques. Different image pre-processing and segmentation techniques for pothole identification were looked into utilizing execution measures. The Identification of various image processing strategies for pothole recognition was finished by comparing performance measures for various image segmentation strategies. Detriment of this paper, executing these image segmentation strategies utilizing hybrid classifiers like neural network and fuzzy rule base and to develop a standalone product for pothole detection. I Schiopu et al. [2] proposed a low complexity technique for identification and tracking of potholes in video sequences taken by a camera placed inside a moving car. The region of interest for the detection of the potholes is chosen as the picture territory where the street is seen with the highest resolution. The paper proposed an algorithm for pothole detection and tracking. The region of interest (ROI) was chosen off-line and candidate regions were

produced utilizing a threshold based algorithm. The upsides of this paper, great accuracy and a little runtime. Vinay Rishiwal et al. [3] shows a vibration based approach for programmed location of potholes and speed breakers alongside their coordinates. In this approach, a database is kept up for every street, which is made accessible to the general population with the assistance of worldwide database or through an entry. The proposed favourable circumstances are cost effective and extremely successful for street surface checking. discuss about the significance of street surface monitoring in terms of comfort and security required by the street travellers. This approach can be advantageously for secure travelling particularly in obscure street conditions. The demerits are more complex. Rajeshwari S. et a. [4] displays an intelligent traffic control framework to pass crisis vehicles easily. Each individual vehicle is equipped with special RFID tag (placed at a key area), which makes it difficult to remove or destroy. This paper utilizes RFID peruse, NSK EDK-125–TTL and PIC16F877A framework on-chip to peruse the RFID labels attached to the vehicle. As the whole framework is computerized, it requires less human intercession. With stolen vehicle exploration, the signal naturally swings to red, so that the cop can make fitting move, on the off chance that he/she is available at the intersection. From literature survey we can conclude that, Potholes can be detected by using image pre-processing, segmentation and ultrasonic sensor techniques for pothole identification. Location of potholes and speed breakers alongside their coordinates can be detected by using vibration based approach.

Pros

- **Focus on Performance:** It mentions the importance of performance measures and the comparison of various image segmentation strategies, highlighting the goal of achieving highly efficient and accurate methods for pothole detection.
- **Practical Applications:** The content discusses the practical applications of these methods, such as the use of a vibration-based approach for detecting potholes and speed breakers and the implementation of an intelligent traffic control system using RFID technology.

Cons

- **Lack of Technical Details:** While the content mentions various techniques and their advantages, it lacks technical details about how these methods

work, which may leave readers wanting more information about the underlying technology and algorithms.

- **Limited Discussion on Challenges:** It does not discuss the challenges or limitations of the proposed methods, which can be important for readers to understand the practical constraints and potential drawbacks.

2.5 AN APPROACH FOR REAL-TIME ROAD DEFORMATION

DETECTION

The study conducted for structuring this project focuses on the application of image processing for segmentation for differentiating of road defects from plain road and other surrounding objects. As mentioned earlier other two approaches are with the use of expensive stereo cameras and laser scanners for 3D reconstruction which have computing complexity and vibration based which has a few drawbacks like only the cracks and potholes will get logged from where vehicle wheels traverse through the deformities and not the ones which pass between the two side wheels. As described in paper by Emir Buza where Otsu thresholding method is used which segments images to obtain binary images from gray scale. Spectral clustering is then employed for determining shapes of irregular objects like road deformities from the selected images. In this shadows of surrounding objects and manhole covers can also get tagged as potholes and may not perform well in low light conditions[2]. Ajit Danti and Seungkiryu presented a novel approach by first using lane detection, then mapping a virtual polygon considering the lanes assides of the polygon and then defined it as Region of Interest thus extracting a smallpart of the incoming video to reduce the processing load on the system and then by inspecting various features, such as the length, area, variance, and trajectory potholeextraction has been carried out. The important part in this is that not all roads have lane lines are present and in such situation by fixing lane vanishing point the unknownlane line is plotted virtually. They have implemented this is a commercially available black box camera embedded system; due to which this implementation technically faces limitation of the processing capabilities of that embedded system and also future modifications may not be possible[2 & 3]. In paper by Christoph Mertz, another novel approach by using line laser light projected on the road surface

that helps in two ways: one for highly detailed 3D profiling of the road damages and other for plotting basic 2D graphical representation of the road deformations which can be well be used for filtering shadows and other objects that may get tagged as deformities which would be an improvement over the use of simple spectral clustering method [4]. But use of this system may not work effectively while in broad daylight but will generate one of the best results at night time. Thus this paper presents a combination of lane detection, laser and spectral clustering of the road for accurate detection of potholes and cracks well in advance before the vehicle passes over it.

Pros

- **Consideration of Limitations:** The content mentions the limitations and drawbacks of certain methods, such as the risk of tagging shadows and manhole covers as potholes in low-light conditions, which adds a realistic perspective to the discussion.
- **Integration of Technologies:** Some approaches, like the combination of lane detection, laser profiling, and spectral clustering, demonstrate the integration of multiple technologies for more accurate defect detection, which can be advantageous.

Cons

- **Limited Discussion of Challenges:** While some limitations are mentioned, a more in-depth discussion of the challenges and potential issues associated with each approach would be beneficial for readers to gain a complete understanding.
- **No Conclusion or Recommendations:** There is no conclusion or summary that synthesizes the key takeaways or provides recommendations for which method might be most suitable for specific scenarios or applications.

2.6 POTHOLES DETECTION BASED ON BLOB DETECTION METHOD

Some existing methods to detect pothole, such as; vibration-based methods, 3D reconstruction-based methods, and vision-based methods [5]. Varadharajadan et al. [6] introduces potholes inspection system using vision- based methods. Smart camera from a smart phone captures road images for segmentation. For

detecting cracks on road or other road damage, Super pixel SLIC variation algorithm is utilized. Road damages are confirmed using Multiple Instance Learning Algorithm. The result shows the system is able to detect road cracks, and has potential to detect potholes, patches, etc. Koch and Ioannis [5] classifying a road based on its defect and non-defect regions for potholes detection. The algorithm consists of three stages which is image segmentation, shape extraction, and texture extraction and comparison. Image segmentation accomplished by using histogram shape-based thresholding algorithm. There are two procedure to detect a pothole in elliptical shape based on the detected shades. First is morphological thinning to minimize the cracks effect from the potholes. The second one is elliptic regression to approximate an ellipse. A remote-controlled robot was used to collect the data and simulate a high speed vehicle. The data is classify by a high variety potholes such as by shapes and sizes, non-defect asphalt pavements and other defect such as cracking and patching, and shadows that caused by lighting condition. This method has 85.9% accuracy and 81.6% precision. Although it is inefficient for the computation due the redetected pothole from processing every single image of the road pavement videos [7]. Buza et al. [8] provides another option for pothole detection by utilize image analysis and spectral clustering. Data is collected by mounting the system on passenger vehicles. The defects region are detected from every frame, then being analyzed later. Spectral clustering is used on the shape extraction process. This method adds Otsu Image Thresholding to automatically decide threshold values every image with internal algorithms either. The estimation accuracy of potholes detection by this algorithm is around 81%. Pawade et al. [9] have presented a low-cost potholes detection method with Field Programmable Gate Arrays (FPGA). FPGA is a digital devices with simply design [10]. The main algorithms to detect potholes by using three edge detection technique which is Sobel, Prewitt, and Canny. The pre-processing using canny method as it is one of the best efficient noise removal technique. The three edge detection technique is processed in parallel threads by FPGA, which caused real-time detection is hard to be achieved. Nevertheless, this system has good result of the amount of potholes on the pavement road and can also alert the driver about location of the potholes via GPS technology. Simple image processing technique to detect some potholes in real world footage has been proposed by Nienaber et al. [1]. The road region is extracted based on the contour and colour of the road itself. The algorithm defines a pothole as a shape which has strong dark edge, so canny edge

detection method was capable to get potholes contour. Convex hull algorithm was applied to decrease the noise effect and a Gaussian filter is used to increase edge detection result. This pothole detection system method is tested by using a Go Pro

Pros

- **Mention of Accuracy:** Some methods mention accuracy rates, such as 85.9%, which gives readers an idea of the effectiveness of these approaches in detecting potholes.
- **Use of Image Processing:** Several methods utilize image processing techniques like edge detection, contour extraction, and Gaussian filtering, which are common and accessible tools for pothole detection.

Cons

- **No Comparison:** The content does not provide a direct comparison between the discussed methods in terms of their strengths and weaknesses. A comparison could help readers understand which method might be more suitable for specific scenarios.
- **No Evaluation Results for Some Methods:** While some methods mention accuracy rates, others do not provide information on the actual performance results or outcomes, making it challenging to assess their effectiveness.

CHAPTER 3

EXISTING AND PROPOSED SYSTEM

3.1 EXISTING SYSTEM

In an existing system a detection which starts with noise removal, followed by brightness adjustment and rendition of video by doing the process of binarization. Then, removal of noise is applied to the binarized image. After noise removal, the process of extracting the framework or outlines for the objects that are segmented is carried out. Extraction is followed by selection, square zoning for the objects.

It is categorised in to 3 subsystems. The encountered potholes are sensed by this first subsystem namely sensing subsystem, by using accelerometer or by camera which scans the road. Both are mounted on the car. Then there is another subsystem that is used to transfer the data between WiFi access points to the mobile nodes. The data about potholes is broadcasted by access point in its area in the existing system, the pothole is detected using the accelerometer sensor in smart phone. This system is not automotive in nature. The complaints if needed to be posted or to be informed to any governmental authority it will be done only with human intervention. This process may not provide the complete efficiency as many people may ignore the issue and will not post them. Even if people sends the complaint to an admin many pothole image may be repeated and thus it may cause a huge confusion. In this case, if prioritization has been done then it would be an optimized way to collect the frequent places that is being affected by potholes.

3.1.1. Disadvantages

- GSM technology has been used which causes delay in message delivery.
- It is a paid message service.
- The severity is less estimated.
- It is exceptionally costly in nature and it can't be connected to working class level vehicles.

3.2 PROPOSED SYSTEM

In this proposed system Image has been Browse using a MATLAB Tool. The image is firstly blurred using aver aging then it goes through Gaussian filet technique and lastly through median blur for removing unnecessary noise in the image. We modified the process with morphological operations just for achieving the most approximate detection of edge from the depth image. In general, these operations are a group of nonlinear operations that are carried out analogously on ordering of pixels without stirring their numerical values. Erosion, Dilation are morphological key operators. We have used erosion after blurring operations which is followed by two iterations of dilation. The pothole detection is utilizing canny edge detection technique. Detection techniques are multi - stage method to detect wide ranges of edges in images.

3.2.1 Advantages

- Easy to Replacement and maintenance.
- Fast in Identifying the potholes.
- Very cost efficient product.
- Simple method and Improves efficiency.

3.3 FLOW CHART

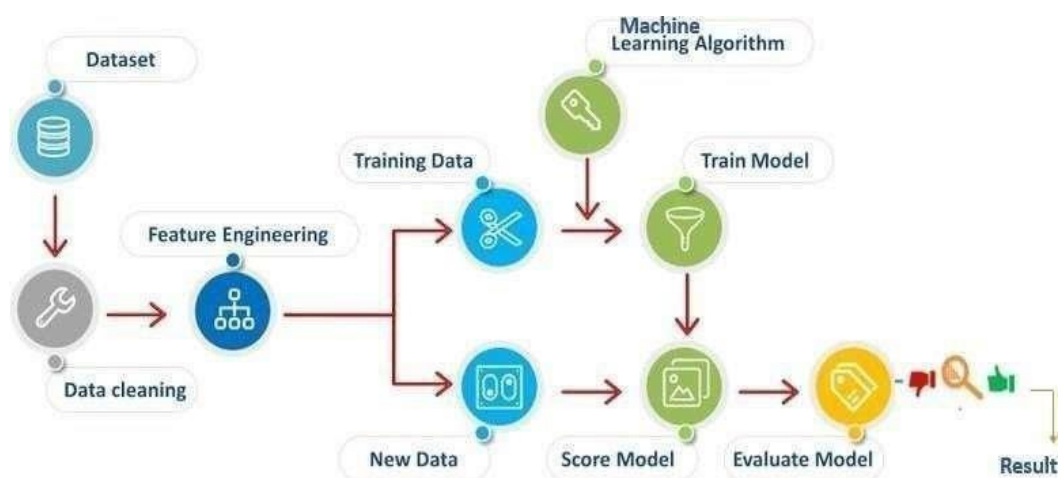


Fig: 3.1 Flow Chart

The most frequently mounting problems of the roads are facing is Damaged road conditions. For many reasons like earth quakes, accidents on roads, wear and tear of top layers of roads make the drivers to drive on those roads. Unexpected Obstacles on roads may cause number of accidents. Also because of worse road conditions, the vehicles consumes a lot of fuel unnecessarily. For these are the reasons it is very significant to get the data of such damaged roads. POTHOLEs are the crack-up's on roads that are formed on the roads for many reasons where the top road layers will get tattered by heavy traffic. Once after formation of pothole the size of hole will keep increasing till we ready that road or repair the pothole, which is main causes for Road mishap's. Potholes on roads are particularly affects for drivers rushing at high speed. Athigh speeds a driver almost inconceivably can notice the potholes on road surface. None the less the drivers may see the potholes before they cross it, it normally will be too late for a driver to respond to the pothole. A sudden turn or sudden braking will mostly causes car roll-over. Triggered to the above reasons, we introduced a system that detects potholes and the proposed system will produce the three dimensional data of potholes and save data information data and we can retrieve that data and get pothole image. In this paper we use different edge detection techniques to detect potholes on road. Canny edge detection technique, the most important algorithm we used here to detect potholes. It detects potholes very accurately and stores the data for future maintenance of roads. We use a web camera to record the roads and click the images because comparing to the costly high speed Three Dimensional traverse scanning equipment, web camera is cheap and expensive.

3.4 Comparison of Existing and Proposed system

Methodology:

Existing System: The existing system may use simple image processing techniques like edge detection or thresholding to identify potential potholes based on predefined criteria.

Proposed System: The proposed system integrates more advanced techniques such as semantic segmentation using K-means clustering and contour detection algorithms, allowing for a more precise and robust detection of potholes.

Accuracy:

Existing System: The accuracy of the existing system may vary depending on the complexity of the road conditions and the effectiveness of the chosen techniques.

Proposed System: The proposed system is likely to achieve higher accuracy due to its ability to segment the image into meaningful regions and identify potholes based on their visual characteristics.

Robustness:

Existing System: The robustness of the existing system may be limited by its reliance on handcrafted features or manual parameter tuning.

Proposed System: The proposed system is more robust as it leverages unsupervised learning techniques (K-means clustering) to automatically segment the image and contour detection algorithms to identify potential potholes, reducing the need for manual intervention.

Scalability:

Existing System: The scalability of the existing system may be limited by the complexity of the chosen techniques and the computational resources required.

Proposed System: The proposed system is scalable and can be applied to large datasets of road images, making it suitable for real-world applications such as monitoring road conditions over large areas

Tabel: 3.1 Algorithm comparison

Model	Precision	Recall	F1-score	Inference time(ms)
SSD-Mobilenetv2	0.42	0.56	0.479	7
YOLOv1	0.82	0.69	0.74	340
YOLOv2	0.81	0.76	0.78	33.7
YOLOv3	0.77	0.78	0.78	70.57
Tiny-YOLOv4	0.76	0.75	0.76	4.86
YOLOv4	0.81	0.83	0.82	52.51
YOLOv5	0.93	0.83	0.87	10

3.4 STANDARDS

The system must achieve high precision and recall rates in pathole detection to minimize false positives and negatives. This ensures that only genuine patholes are detected, enabling accurate notifications. Balancing computational resources to ensure efficient processing time without compromising accuracy is crucial. Optimal utilization of hardware resources enhances the scalability and responsiveness of the system. The algorithm should be robust against various environmental conditions such as lighting, weather, and road surface textures. Robustness ensures consistent performance across different scenarios, enhancing the reliability of pathole detection. The system should be capable of detecting patholes in diverse settings and environments, demonstrating its generalizability. Generalization ensures that the algorithm can adapt to new data and environments, making it applicable in real-world scenarios beyond the training dataset. The user interface, including voice notifications and messages, should be intuitive and user-friendly. Clear and concise alerts enhance usability, ensuring that users can easily understand and respond to notifications. The system should be compatible with various image formats and platforms to facilitate seamless integration into existing infrastructure. Compatibility ensures interoperability with different systems and technologies, enhancing the system's versatility and adoption potential. The system should be adaptable to changing requirements and conditions, allowing for easy updates and modifications. Adaptability ensures that the system can evolve over time to address emerging challenges and opportunities.

3.5 CONSTRAINTS

The system must accommodate various hardware configurations, ranging from high-performance servers to resource-constrained devices like smartphones. Ensuring the algorithm's compatibility with different hardware constraints is essential for widespread adoption and usability. Strict adherence to data privacy regulations and guidelines is paramount. The system must handle sensitive image data securely, ensuring that personal information is protected from unauthorized access or misuse. Implementing robust encryption and access controls safeguards user privacy and maintains trust in the system. The system must process images and generate notifications in real-time or near real-time to

provide timely alerts. Delays in processing may result in missed opportunities for intervention or repairs, compromising road safety and infrastructure integrity..The system should be scalable to handle a growing volume of images and notifications, particularly in densely populated urban areas or during peak usage times. Scalability ensures that the system can accommodate increasing demand without sacrificing performance or reliability.Compliance with relevant regulations governing image processing, data handling, and notification systems is essential. The system must adhere to legal requirements regarding privacy, data protection, and accessibility to ensure lawful operation and mitigate potential risks of non-compliance.The system must optimize resource utilization to minimize computational overhead and energy consumption. Efficient resource management ensures that the system remains cost-effective and environmentally sustainable, especially in deployments with limited resources or budget constraints.Meeting stringent accuracy requirements for pathole detection is critical. The system must minimize false positives and negatives to ensure reliable notifications and avoid unnecessary maintenance interventions. Continuous monitoring and performance evaluation are necessary to maintain high levels of accuracy over time.Integrating the YOLOv7 algorithm with voice and message notification systems may present technical challenges.

3.6 TRADE OFFS

Optimizing pathole detection accuracy with the YOLOv7 algorithm often requires complex neural network architectures and extensive training, potentially leading to longer processing times. However, achieving real-time detection is crucial for prompt notification delivery in high-traffic areas. Therefore, there exists a trade-off between accuracy and speed. Sacrificing some level of accuracy may be necessary to ensure timely alerts. Additionally, while YOLOv7 may offer superior accuracy compared to other YOLO algorithms, it could also demand more computational resources, impacting system scalability and cost-effectiveness. Balancing these factors is essential to create a pathole detection system that effectively balances accuracy, speed, resource efficiency, and cost.

CHAPTER 4

METHODOLOGY

4.1 CANNY EDGE DETECTION ALGORITHM

The Canny Edge Detector is an edge detection operator that uses a multi-stage algorithm to detect a wide range of edges in images. It was developed by John F Canny in 1986. Canny also produced a computational theory of edge detection explaining why the technique works. The Process of Canny edge detection algorithm can be broken down to 5 different steps :

1. Apply Gaussian Filter to smooth the image in order to remove the noise.
2. Find the intensity gradients of the image.
3. Apply non-maximum suppression to get rid of spurious response to edge detection.
4. Apply double threshold to determine potential edges.
5. Track edge by Hysteresis , Finalize the detection of edges by suppressing all the other edges that are weak and not connected to strong edges.

4.1.1 Gaussian Filter

Since all edge detection results are easily affected by the noise in the image, it is essential to filter out the noise to prevent false detection caused by it. To smooth the image, a Gaussian filter kernel is convolved with the image. This step will slightly smooth the image to reduce the effects of obvious noise on the edge detector. The equation for a Gaussian filter kernel of size $(2k+1) \times (2k+1)$ is given by

:

$$H_{ij} = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{(i - (k+1))^2 + (j - (k+1))^2}{2\sigma^2}\right); 1 \leq i, j \leq (2k+1)$$

Here is an example of a 5×5 Gaussian filter, used to create the adjacent image, with $\sigma = 1$. (The asterisk denotes a convolution operation.)

$$\mathbf{B} = \frac{1}{159} \begin{bmatrix} 2 & 4 & 5 & 4 & 2 \\ 4 & 9 & 12 & 9 & 4 \\ 5 & 12 & 15 & 12 & 5 \\ 4 & 9 & 12 & 9 & 4 \\ 2 & 4 & 5 & 4 & 2 \end{bmatrix} * \mathbf{A}.$$

It is important to understand that the selection of the size of the Gaussian kernel will affect the performance of the detector. The larger the size is, the lower the detector's sensitivity to noise. Additionally, the localization error to detect the edge will slightly increase with the increase of the Gaussian filter kernel size. A 5x5 is a good size for most cases, but this will also vary depending on specific situation.

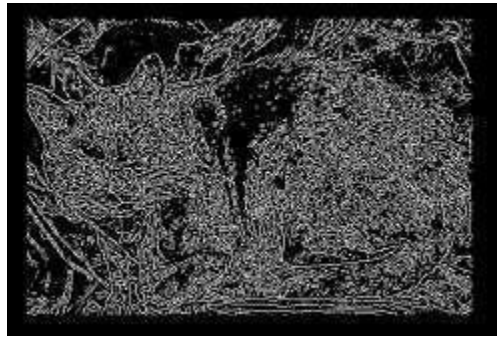


Fig: 4.1 Gaussian smoothing

4.1.2 Intensity Gradient

An edge in an image may point in a variety of directions, so the Canny algorithm uses four filters to detect horizontal, vertical and diagonal edges in the blurred image. The edge detection operator (such as Roberts, Prewitt, or Sobel) returns a value for the first derivative in the horizontal direction (G_x) and the vertical direction (G_y). From this the edge gradient and direction can be determined :

$$G = \sqrt{G_x^2 + G_y^2}$$

$$\Theta = \text{atan2}(G_y, G_x),$$

where G can be computed using the hypot function and a atan2 is the arctangent function with two arguments. The edge direction angle is rounded to one of four angles representing vertical, horizontal and the two diagonals (0° , 45° , 90° and 135°). An edge direction falling in each colour region will be set to a specific angle values, for instance θ in $[0^\circ, 22.5^\circ]$ or $[157.5^\circ, 180^\circ]$ maps to 0° .

4.1.3 Non-Maximum Suppression

Non-maximum suppression is an edge thinning technique. Non-maximum suppression is applied to find “the largest” edge. After applying gradient calculation, the edge extracted from the gradient value is still quite blurred. With respect to criterion 3, there should only be one accurate response to the edge. Thus non-maximum suppression can help to suppress all the gradient values (by setting them to 0) except the local maxima, which indicate locations with the sharpest change of intensity value. The algorithm for each pixel in the gradient image is:

1. Compare the edge strength of the current pixel with the edge strength of the pixel in the positive and negative gradient directions.
2. If the edge strength of the current pixel is the largest compared to the other pixels in the mask with the same direction (e.g., a pixel that is pointing in the y-direction will be compared to the pixel above and below it in the vertical axis), the value will be preserved. Otherwise, the value will be suppressed.

In some implementations, the algorithm categorizes the continuous gradient directions into a small set of discrete directions, and then moves a 3x3 filter over the output of the previous step (that is, the edge strength and gradient directions). At every pixel, it suppresses the edge strength of the center pixel (by setting its value to 0) if its magnitude is not greater than the magnitude of the two neighbours in the gradient direction. For example,

- If the rounded gradient angle is 0° (i.e. the edge is in the north-south direction) the point will be considered to be on the edge if its gradient magnitude is greater than the magnitudes at pixels in the **east and west** directions.
- If the rounded gradient angle is 90° (i.e. the edge is in the east-west direction) the point will be considered to be on the edge if its gradient magnitude is greater than the magnitudes at pixels in the **north and south** directions.
- If the rounded gradient angle is 135° (i.e. the edge is in the northeast-southwest direction) the point will be considered to be on the edge if its gradient magnitude is greater than the magnitudes at pixels in the **north west and south-east** directions.
- If the rounded gradient angle is 45° (i.e. the edge is in the north west–south east direction) the point will be considered to be on the edge if its gradient

magnitude is greater than the magnitudes at pixels in the **north east** and **south west** directions.

In more accurate implementations, linear interpolation is used between the two neighbouring pixels that straddle the gradient direction. For example, if the gradient angle is between 89° and 180° , interpolation between gradients at the **north** and **north east** pixels will give one interpolated value, and interpolation between the **south** and **south west** pixels will give the other (using the conventions of the last paragraph). The gradient magnitude at the central pixel must be greater than both of these for it to be marked as an edge.

Note that the sign of the direction is irrelevant, i.e. north–south is the same as south–north and so on.

4.1.4 Double Threshold

After application of non-maximum suppression, remaining edge pixels provide a more accurate representation of real edges in an image. However, some edge pixels remain that are caused by noise and colour variation. In order to account for these spurious responses, it is essential to filter out edge pixels with a weak gradient value and preserve edge pixels with a high gradient value. This is accomplished by selecting high and low threshold values. If an edge pixel's gradient value is higher than the high threshold value, it is marked as a strong edge pixel. If an edge pixel's gradient value is smaller than the high threshold value and larger than the low threshold value, it is marked as a weak edge pixel. If an edge pixel's value is smaller than the low threshold value, it will be suppressed. The two threshold values are empirically determined and their definition will depend on the content of a given input image.

4.1.5 Edge Tracking By Hysteresis

So far, the strong edge pixels should certainly be involved in the final edge image, as they are extracted from the true edges in the image. However, there will be some debate on the weak edge pixels, as these pixels can either be extracted from the true edge, or the noise/colour variations. To achieve an accurate result, the weak edges caused by the latter reasons should be removed. Usually a weak edge pixel caused from true edges will be connected to a strong edge pixel while noise responses are unconnected. To track the edge connection, blob analysis is

applied by looking at a weak edge pixel and its 8-connected neighbourhood pixels. As long as there is one strong edge pixel that is involved in the blob, that weak edge point can be identified as one that should be preserved.

The Canny algorithm is adaptable to various environments. Its parameters allow it to be tailored to recognition of edges of differing characteristics depending on the particular requirements of a given implementation. In Canny's original paper, the derivation of the optimal filter led to a Finite Impulse Response filter, which can be slow to compute in the spatial domain if the amount of smoothing required is important (the filter will have a large spatial support in that case). For this reason, it is often suggested to use Rachid Deriche's infinite impulse response form of Canny's filter (the Canny–Deriche detector), which is recursive, and which can be computed in a short, fixed amount of time for any desired amount of smoothing. The second form is suitable for real time implementations in FPGAs or DSPs, or very fast embedded PCs. In this context, however, the regular recursive implementation of the Canny operator does not give a good approximation of rotational symmetry and therefore gives a bias towards horizontal and vertical edges.

4.2 IMAGE THRESHOLDING

- We use Otsu's method for reduction of a grey level image to a binary image.
- The system then uses contour detection technique. For better accuracy, use binary images. So we have applied threshold and canny edge detection.
- While capturing video the hassle of adjusting speed is eliminated.

4.2.1 Otsu's Method

$$\sigma_w^2(t) = \omega_0(t)\sigma_0^2(t) + \omega_1(t)\sigma_1^2(t)$$

Weights ω_0 and ω_1 are the probabilities of the two classes separated by a threshold 't', and σ_0^2 and σ_1^2 are the variances of these two classes. The class probability

$\omega_{0,1}(t)$ is computed from the L bins of the histogram :

$$\omega_0(t) = \sum_{i=0}^{t-1} p(i)$$

$$\omega_1(t) = \sum_{i=t}^{L-1} p(i)$$

For 2 classes, minimizing the intra-class variance is equivalent to maximizing inter-class variance :

$$\begin{aligned}\sigma_b^2(t) &= \sigma^2 - \sigma_w^2(t) = \omega_0(\mu_0 - \mu_T)^2 + \omega_1(\mu_1 - \mu_T)^2 \\ &= \omega_0(t)\omega_1(t)[\mu_0(t) - \mu_1(t)]^2\end{aligned}$$

which is expressed in terms of class probabilities ω and class means μ , where the class means $\mu_0(t)$, $\mu_1(t)$ and μ_T are :

$$\mu_0(t) = \frac{\sum_{i=0}^{t-1} ip(i)}{\omega_0(t)}$$

$$\mu_1(t) = \frac{\sum_{i=t}^{L-1} ip(i)}{\omega_1(t)}$$

$$\mu_T = \sum_{i=0}^{L-1} ip(i)$$

The following relations can be easily verified :

$$\omega_0\mu_0 + \omega_1\mu_1 = \mu_T$$

$$\omega_0 + \omega_1 = 1$$

The class probabilities and class means can be computed iteratively. This idea yields an effective algorithm.

4.3 HOUGH TRANSFORM

The **Hough transform** is a feature extraction technique used in image analysis, computer vision, and digital image processing. The purpose of the technique is to find imperfect instances of objects within a certain class of shapes

by a voting procedure. This voting procedure is carried out in a parameter space, from which object candidates are obtained as local maxima in a so-called accumulator space that is explicitly constructed by the algorithm for computing the Hough transform. The classical Hough transform was concerned with the identification of lines in the image, but later the Hough transform has been extended to identifying positions of arbitrary shapes, most commonly circles or ellipses. The Hough transform as it is universally used today was invented by Richard Duda and Peter Hart in 1972, who called it a “generalized Hough transform” after the related 1962 patent of Paul Hough. The transform was popularized in the computer vision community by Dana H. Ballard through a 1981 journal article titled “Generalizing the Hough transform to detect arbitrary shapes”.

4.3.1 Theory

In automated analysis of digital images, a sub problem often arises of detecting simple shapes, such as straight lines, circles or ellipses. In many cases an edge detector can be used as a pre-processing stage to obtain image points or image pixels that are on the desired curve in the image space. Due to imperfections in either the image data or the edge detector, however, there may be missing points or pixels on the desired curves as well as spatial deviations between the ideal line/circle/ellipse and the noisy edge points as they are obtained from the edge detector. For these reasons, it is often non-trivial to group the extracted edge features to an appropriate set of lines, circles or ellipses. The purpose of the Hough transform is to address this problem by making it possible to perform groupings of edge points into object candidates by performing an explicit voting procedure over a set of parameterized image objects (Shapiro and Stockman, 304).

The simplest case of Hough transform is detecting straight lines. In general, the straight line $y = mx + b$ can be represented as a point (b, m) in the parameter space. However, vertical lines pose a problem. They would give rise to unbounded values of the slope parameter m . Thus, for computational reasons, Duda and Hart proposed the use of the Hesse normal form.

$$R = x \cos \theta + y \sin \theta$$

where 'r' is the distance from the origin to the closest point on the straight line, and θ (theta) is the angle between the x axis and the line connecting the origin with that closest point.

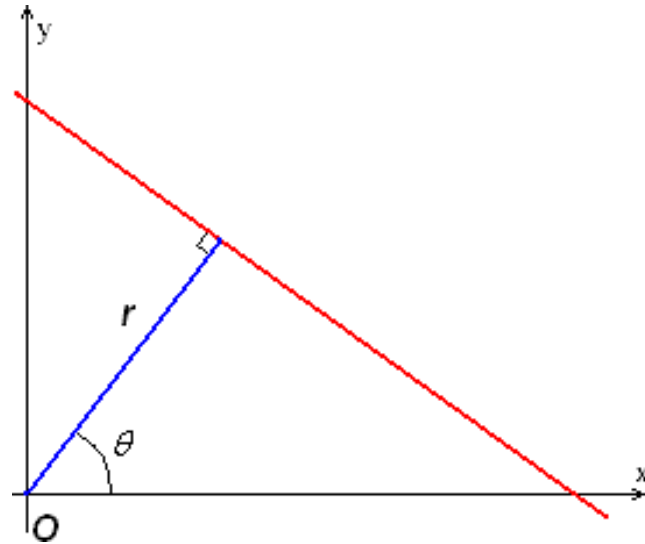


Fig: 4.2 Hough Transform

It is therefore possible to associate with each line of the image a pair (r, θ) . The (r, θ) plane is sometimes referred to as *Hough space* for the set of straight lines in two dimensions. This representation makes the Hough transform conceptually very close to the two-dimensional Radon transform. In fact, the Hough transform is mathematically equivalent to the Radon transform, but the two transformations have different computational interpretations traditionally associated with them.

Given a *single point* in the plane, then the set of *all* straight lines going through that point corresponds to a sinusoidal curve in the (r, θ) plane, which is unique to that point. A set of two or more points that form a straight line will produce sinusoids which cross at the (r, θ) for that line. Thus, the problem of detecting collinear points can be converted to the problem of finding concurrent curves.

4.3.2 Example 1

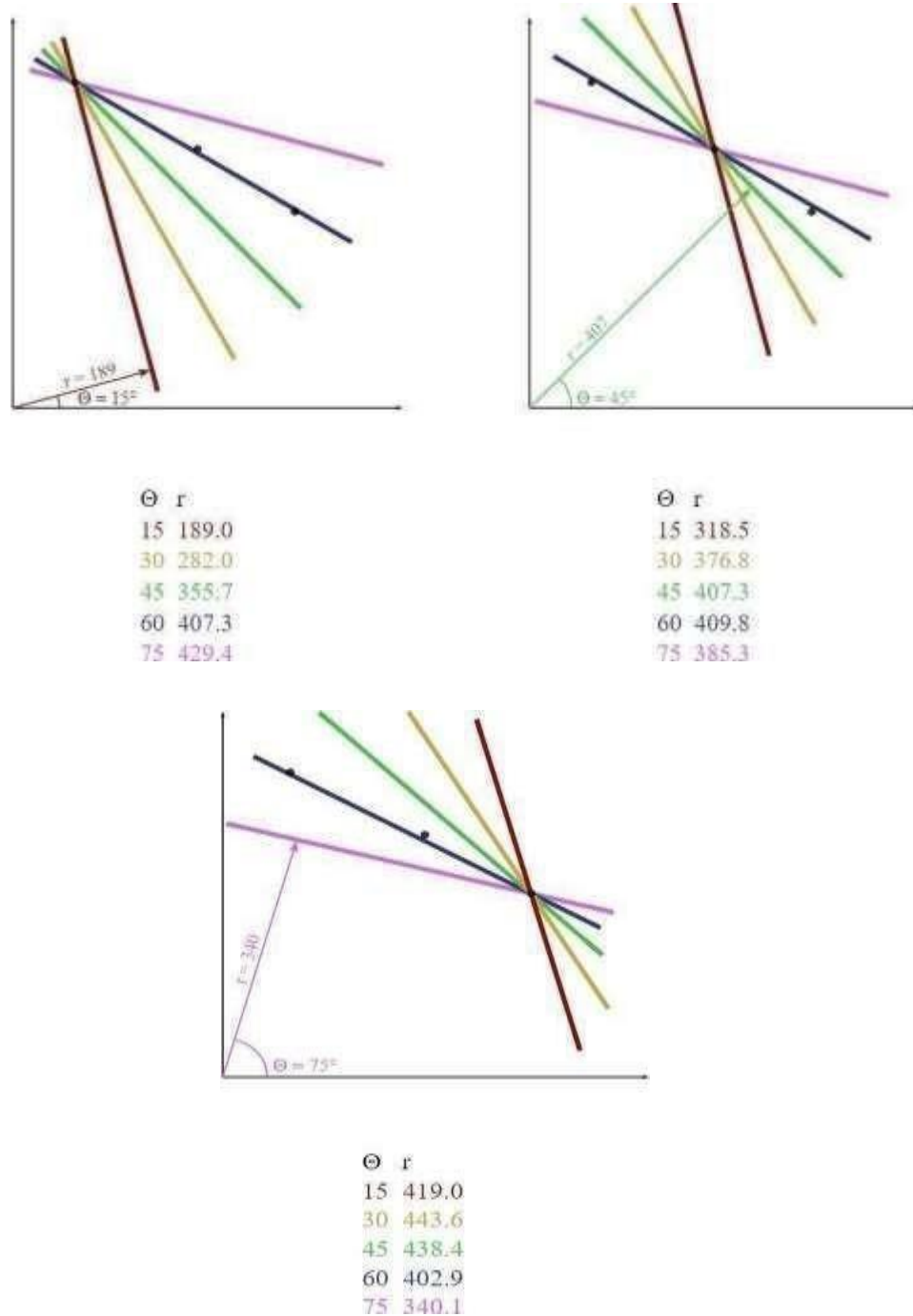


Fig : 4.3 Example 1

For each data point, a number of lines are plotted going through it, all at different angles. These are shown here in different colours. To each line, a support line exists which is perpendicular to it and which intersects the origin. In each case, one of these is shown as an arrow. The length (i.e. perpendicular distance to the origin) and angle of each support line is calculated. Lengths and angles are

tabulated below the diagrams. From the calculations, it can be seen that in either case the support line at 60° has a similar length. Hence, it is understood that the corresponding lines (the blue ones in the above picture) are very similar. One can thus assume that all points lie close to the blue line.

4.3.3 Example 2

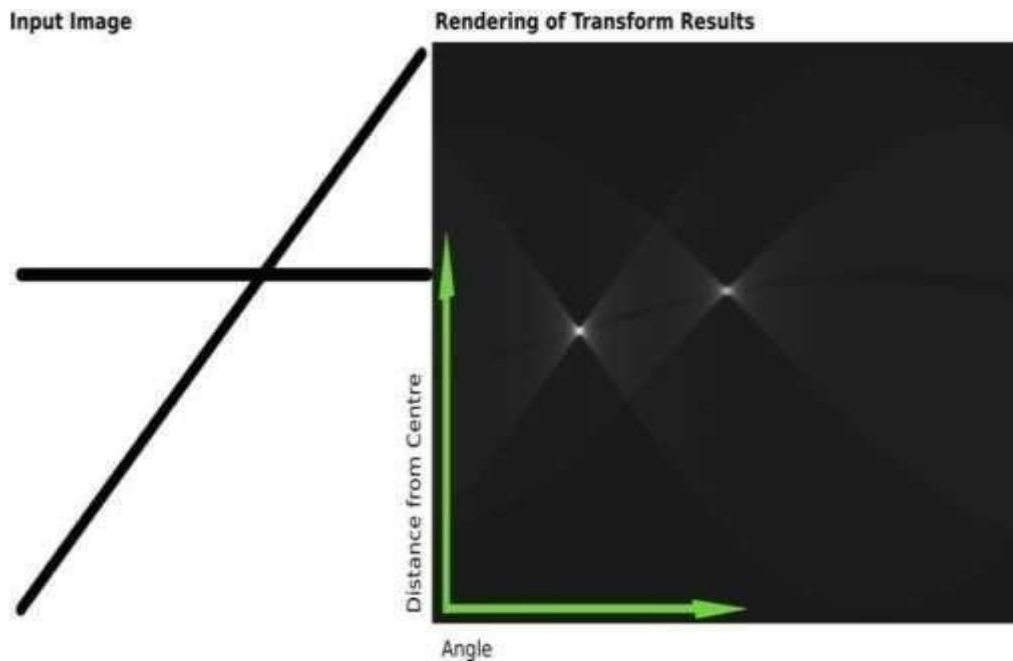


Fig:4.4 Rendering of Transform Results

The results of this transform were stored in a matrix. Cell value represents the number of curves through any point. Higher cell values are rendered brighter. The two distinctly bright spots are the Hough parameters of the two lines. From these spots' positions, angle and distance from image center of the two lines.

4.4 FEATURE EXTRACTION

In image processing, feature extraction starts from an initial set of measured data and builds derived values (features) intended to be informative and non-redundant, facilitating the subsequent learning and generalization steps, and in some cases leading to better human interpretations. Feature extraction is related to dimensionality reduction. When the input data to an algorithm is too large to be processed and it is suspected to be redundant (e.g. the same measurement in both feet and meters, or the repetitiveness of images presented as pixels), then it can be transformed into a

reduced set of features (also named a features vector). This process is called feature selection. The selected features are expected to contain the relevant information from the input data, so that the desired task can be performed by using this reduced representation instead of the complete initial data.

4.5 GENERAL

Feature extraction involves reducing the amount of resources required to describe a large set of data. When performing analysis of complex data one of the major problems stems from the number of variables involved. Analysis with a large number of variables generally requires a large amount of memory and computation power, also it may cause a classification algorithm to overfit to training samples and generalize poorly to new samples. Feature extraction is a general term for methods of constructing combinations of the variables to get around these problems while still describing the data with sufficient accuracy.

The best results are achieved when an expert constructs a set of application-dependent features, a process called feature engineering. Nevertheless, if no such expert knowledge is available, general dimensionality reduction techniques may help. These include:

- Independent component analysis
- Isomap
- Kernel PCA
- Latent semantic analysis
- Partial least squares
- Principal component analysis
- Multifactor dimensionality reduction
- Nonlinear dimensionality reduction
- Multilinear Principal Component Analysis
- Multilinear subspace learning
- Semidefinite embedding
- Autoencoder
- Deep feature synthesis

4.6 WAVELETS

Wavelets are a relatively new area of signal processing and offer a potentially very useful approach to image analysis. Because of the fact that one has local control over resolution, many cue features of different size can be extracted from an image. Wavelets also show promise as an aid in understanding how early vision can be emulated and further understood.

Wavelet Transforms

It is well known that in signal processing, the more compact your method of representing information the better. In 1822, Jean Joseph Fourier devised a very efficient way to represent the information content of a signal. His idea was to represent a signal as the sum of its frequencies. Transmitted power spectra, carrier frequencies, brain activity, NMR signals - all these global descriptions provide a lot of information in a compact manner. However, most of the power of this kind of representation vanishes when one tries to represent information that changes its nature during the course of signal recording. A good example of this kind of a signal is a musical score. A global analysis of a recording of a musical selection with a Fourier transform (FT) will indicate the specific notes played within the piece of music, but there is no way to recover the timing of the notes. The musical score, on the other hand, indicates each note that was played and the time it was played. Wavelets represent a signal in a way such that local frequency information is available for each position within the signal. Wavelets are able to analyze a signal based on the position-varying spectra. The multiresolution pyramidal decomposition that results is also well matched to fractals and shows great potential for the removal of background noise, such as static in recordings, and for pattern recognition and texture segmentation. In the following section we discuss the conceptual understanding of how wavelets can be used in signal analysis.

4.7 CONVOLUTIONAL NEURAL NETWORKS

Convolutional neural networks (CNNs) are widely used in pattern- and image-recognition problems as they have a number of advantages compared to other techniques.

4.8 WHAT IS A CNN?

A neural network is a system of interconnected artificial “neurons” that exchange messages between each other. The connections have numeric weights that are tuned during the training process, so that a properly trained network will respond correctly when presented with an image or pattern to recognize. The network consists of multiple layers of feature-detecting “neurons”. Each layer has many neurons that respond to different combinations of inputs from the previous layers. As shown in Figure 1, the layers are built up so that the first layer detects a set of primitive patterns in the input, the second layer detects patterns of patterns, the third layer detects patterns of those patterns, and so on. Typical CNNs use 5 to 25 distinct layers of pattern recognition

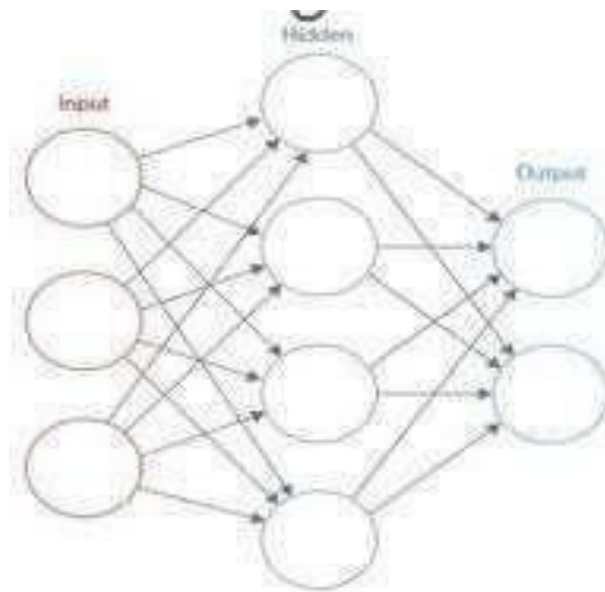


Fig :4.5 Convolution neural network

Training is performed using a “labeled” dataset of inputs in a wide assortment of representative input patterns that are tagged with their intended output response. Training uses general-purpose methods to iteratively determine the weights for intermediate and final feature neurons. Figure 2 demonstrates the training process at a block level.

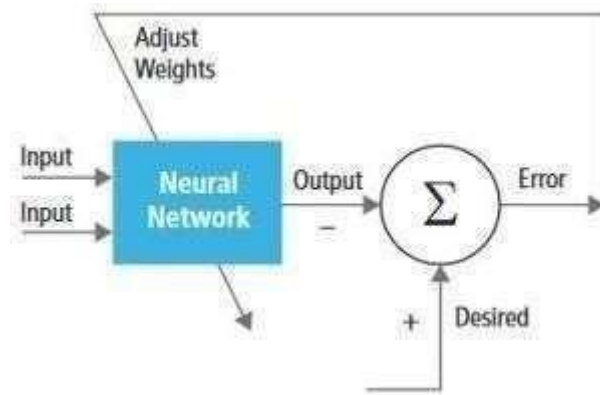


Fig :4.6 Training of Neural network

Neural networks are inspired by biological neural systems. The basic computational unit of the brain is a neuron and they are connected with synapses. Figure 3 compares a biological neuron with a basic mathematical model

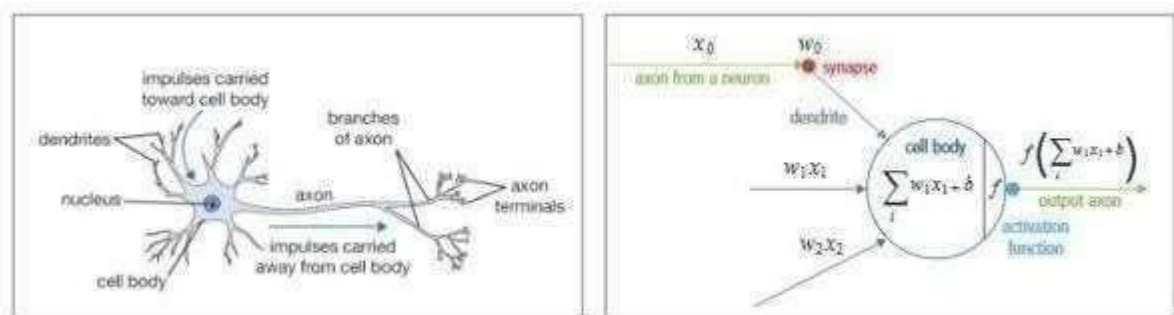


Fig :4.7 illustration of biological neuron(left) and its mathematical function (right)

In a real animal neural system, a neuron is perceived to be receiving input signals from its dendrites and producing output signals along its axon. The axon branches out and connects via synapses to dendrites of other neurons. When the combination of input signals reaches some threshold condition among its input dendrites, the neuron is triggered and its activation is communicated to successor neurons. In the neural network computational model, the signals that travel along the axons (e.g., x_0) interact multiplicatively (e.g., $w_0 x_0$) with the dendrites of the other neuron based on the synaptic strength at that synapse (e.g., w_0). Synaptic weights are learnable and control the influence of one neuron or another. The dendrites carry the signal to the cell body, where they all are summed. If the final sum is above a specified threshold, the neuron fires, sending a spike along its axon. In the computational model, it is assumed that the precise timings of the firing do not matter and only the

the firing communicates information. Based on the rate code interpretation, the firing rate of the neuron is modeled with an activation function f that represents the frequency of the spikes along the axon. A common choice of activation function is sigmoid. In summary, each neuron calculates the dot product of inputs and weights, adds the bias, and applies non-linearity as a trigger function.

4.9 WHY CNN?

While neural networks and other pattern detection methods have been around for the past 50 years, there has been significant development in the area of convolutional neural networks in the recent past. This section covers the advantages of using CNN for image recognition.

Ruggedness to shifts and distortion in the image: Detection using CNN is rugged to distortions such as change in shape due to camera lens, different lighting conditions, different poses, presence of partial occlusions, horizontal and vertical shifts, etc. However, CNNs are shift invariant since the same weight configuration is used across space. In theory, we also can achieve shift invariants using fully connected layers. But the outcome of training in this case is multiple units with identical weight patterns at different locations of the input. To learn these weight configurations, a large number of training instances would be required to cover the space of possible variations.

Fewer memory requirements: In this same hypothetical case where we use a fully connected layer to extract the features, the input image of size 32×32 and a hidden layer having 1000 features will require an order of 10000000 coefficients, a huge memory requirement. In the convolutional layer, the same coefficients are used across different locations in the space, so the memory requirement is drastically reduced. Using Convolutional Neural Networks for Image Recognition Easier and better training. Again using the standard neural network that would be equivalent to a CNN, because the number of parameters would be much higher, the training time would also increase proportionately. In a CNN, since the number of parameters is drastically reduced, training time is proportionately reduced. Also, assuming perfect training, we can design a standard neural network whose performance would be same as a CNN. But in practical training, a standard neural network equivalent to CNN would have more parameters, which would lead to more noise addition during the training process. Hence, the performance of a standard neural network equivalent to a CNN will always be poorer.

4.10 SYSTEM REQUIREMENTS FOR MATLAB

- Operating Systems
 - Windows 10
 - Windows 8.1
 - Windows 8
 - Windows 7 Service Pack 1
 - Windows Server 2016
 - Windows Server 2012
 - Windows Server 2008 R2 Service Pack 1
- Processors
 - Any Intel or AMD x86-64 processor
 - AVX2 instruction set support is recommended
 - With Polyspace, 4 cores is recommended
- Disk Space
 - GB for MATLAB only,
 - 4–6 GB for a typical installation.
- RAM
 - GB
 - With Simulink, 4 GB is required
 - With Polyspace, 4 GB per core is recommended
- Graphics
 - No specific graphics card is required.
 - Hardware accelerated graphics card supporting OpenGL 3.3 with 1GB GPU memory is recommended.

4.11 BLOCK DIAGRAM

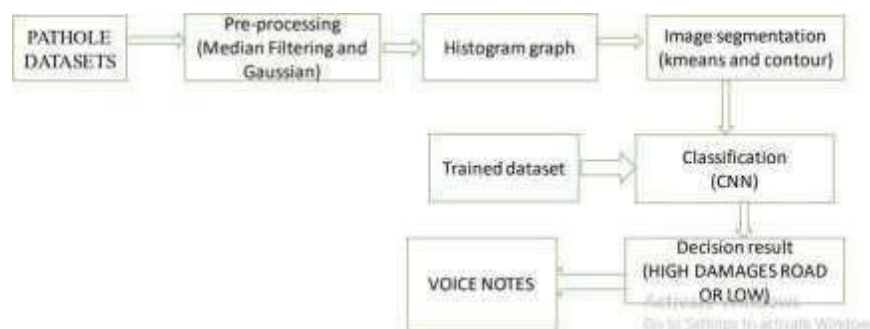


Fig:4.8 Flow Diagram

Trained Dataset:The trained dataset consists of a collection of road images annotated with pothole labels. These images are used to train the CNN for potholedetection. Each image is associated with a binary label indicating the presence or absence of a pothole.

Preprocessing (Median Filter and Gaussian Filter):Before feeding the images into the CNN, preprocessing steps are applied to enhance the quality of the imagesand remove noise. Median filtering and Gaussian filtering are common techniques used for noise reduction and smoothing in images. Median filtering helps in removingsalt-and-pepper noise, while Gaussian filtering blurs the image slightly, reducing high-frequency noise.

Histogram Graph:After preprocessing, a histogram graph may be generated to analyze the intensity distribution of the image pixels. Histogram analysis can provide insights into the distribution of potholes and background in the images, which can aid in setting appropriate thresholds for segmentation.

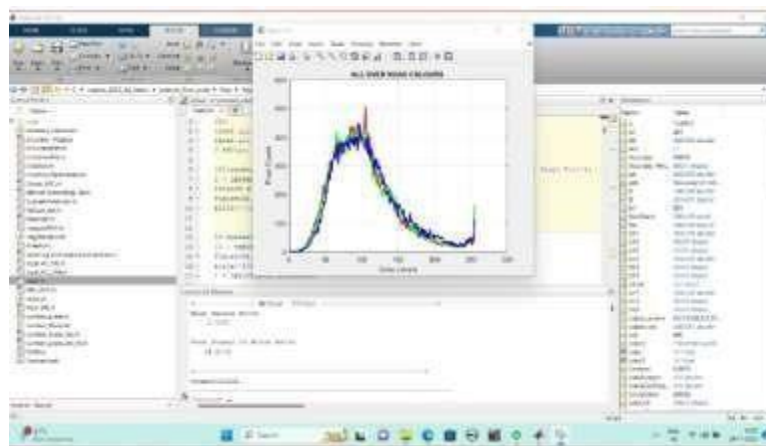


Fig:4.9 Histogram graph

Segmentation (K-means and Contour Algorithm):Segmentation is the process of partitioning an image into meaningful regions or segments. K-means clustering is a popular unsupervised algorithm used for image segmentation, where pixels are grouped into clusters based on their feature similarity. Contour detection algorithms,

such as the contour algorithm in OpenCV, can also be used to identify the boundaries of objects in an image. In the context of pothole detection, segmentation helps in isolating the regions of interest (i.e., potential potholes) from the background.

Classification using CNN: Once the regions of interest are identified through segmentation, they are fed into the CNN for classification. The CNN has been trained on the annotated dataset to recognize potholes from road images. It takes the segmented regions as input and outputs the probability of each region containing a pothole. This step utilizes the features learned by the CNN to make predictions based on the characteristics of potholes.

Output Display using YOLO: Finally, the output of the pothole detection system is displayed using the YOLO (You Only Look Once) algorithm. YOLO is an object detection algorithm that can detect and localize objects in images. In this context, YOLO is used to draw bounding boxes around the detected potholes and overlay them on the original road images. This provides a visual representation of the detected potholes, allowing users to easily identify their locations.

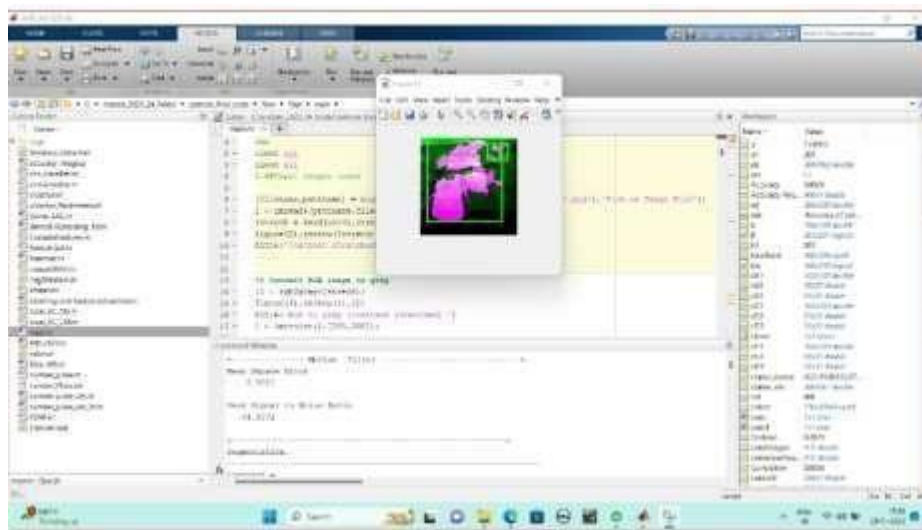


Fig:4.10 Output display using YOLO

The block diagram outlines the key steps involved in the pothole detection process, including data preprocessing, segmentation, classification, and output visualization, utilizing a combination of image processing techniques and deep learning algorithms.

CHAPTER 5

RESULTS AND DISCUSSION

5.1 PREPROCESSING HSV CHANNEL

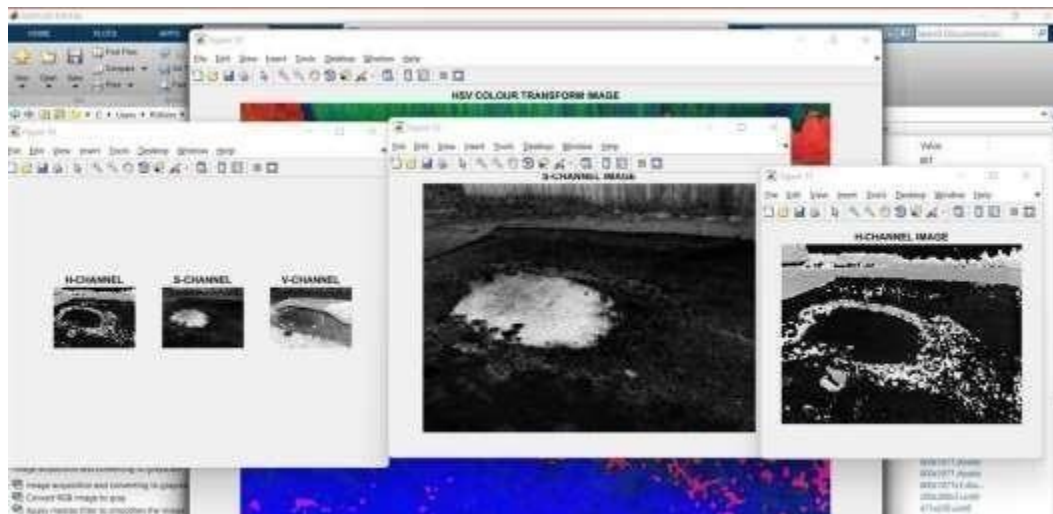


Fig:5.1 Preprocessing HSV Channel

- Preprocessing the HSV (Hue, Saturation, Value) channel output in image processing for pothole detection can be a crucial step to enhance the effectiveness of subsequent detection algorithms
- Color-Based Segmentation: HSV is a color space that separates color information. Preprocessing the Value (V) channel can be useful for color- based segmentation. You can threshold the V channel to extract regions that are likely to contain road surfaces or potholes.
- Noise Reduction: Apply noise reduction techniques such as Gaussian blurring or median filtering to the V channel output. This helps in reducing small-scale noise and irregularities in the image, making it easier to identify larger road defects like potholes.
- Contrast Enhancement: Adjust the contrast of the V channel to enhance the visibility of road defects. Histogram equalization or adaptive histogram equalization can be applied to achieve this.

- **Thresholding:** After noise reduction and contrast enhancement, perform thresholding on the V channel to create a binary image where potential potholes are highlighted. An appropriate threshold value should be chosen to distinguish between road defects and the background.
 - **Morphological Operations:** Use morphological operations like dilation and erosion to further refine the binary mask obtained from thresholding. This can help in connecting nearby pixels to form more coherent regions.
 - **Region of Interest (ROI) Selection:** Define a region of interest in the V channel output to focus the analysis on the road surface where potholes are likely to occur. This can reduce computational load and false positives.
 - **Adaptive Thresholding:** Consider using adaptive thresholding techniques that adapt the threshold value based on local image properties. This can be particularly useful when lighting conditions vary across the image.
 - **Edge Detection:** In addition to thresholding, you can perform edge detection on the V channel to identify the boundaries of potholes. Canny edge detection or Sobel operators can be applied for this purpose.
 - **Connected Component Analysis:** After preprocessing, use connected component analysis to identify and label individual potholes or road defects. This can help in quantifying the number and size of potholes.
- Validation and Post-Processing:** After detection, apply validation and post-processing steps to filter out false positives and refine the results. This may involve considering the shape, size, and context of detected regions.

5.2 PREPROCESSING RGB TO GRAY CONVERSION

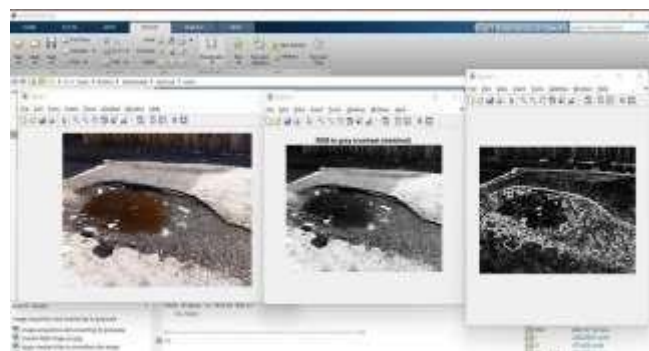


Fig:5.2 Preprocessing RGB to Gray Conversion

Preprocessing RGB to gray conversion is a common step in image processing for pothole detection, as it simplifies the image and reduces computational complexity.

- **Color Information Reduction:** RGB to gray conversion simplifies the image by eliminating color information. This can be advantageous when potholes are primarily detected based on intensity and texture rather than color.
- **Luminance Preservation:** The conversion typically retains the luminance information, which is crucial for detecting variations in intensity and shading, especially in grayscale images.
- **Reduced Data Size:** Grayscale images require less memory and computational resources compared to RGB images. This is important for real-time or resource-constrained applications.
- **Simplicity:** Grayscale images are simpler to process and analyze than RGB images. This simplification can lead to more straightforward algorithms and reduced complexity.
- **Consistent Representation:** Grayscale images provide a consistent representation, as each pixel encodes a single intensity value. This consistency is beneficial for feature extraction and classification.
- **Color Irrelevance:** In many cases, the color of the road and potholes is not relevant for detection. Converting to grayscale helps to focus on relevant intensity variations.
- **Noise Reduction:** Grayscale conversion can help reduce the impact of color noise present in the original RGB image, making it easier to identify potholes.
- **Contrast Enhancement:** Apply contrast enhancement techniques to the grayscale image to improve the visibility of road defects. Histogram equalization or adaptive histogram equalization can be effective for this purpose.
- **Thresholding:** After grayscale conversion and contrast enhancement, you can apply thresholding techniques to create a binary image where potential potholes are highlighted.
- **Edge Detection:** Grayscale images are suitable for edge detection algorithms, such as Canny edge detection or Sobel operators, which can help identify the boundaries of potholes.

- **Texture Analysis:** Grayscale images are conducive to texture analysis, allowing you to detect patterns and irregularities on the road surface.
- **Feature Extraction:** Extract relevant features, such as intensity gradients, from the grayscale image to characterize the road and potholes.
- **Validation and Post-Processing:** After detection, apply validation and post-processing steps to filter out false positives and refine the results. This may involve considering the shape, size, and context of detected regions.

5.3 SEGMENTATION

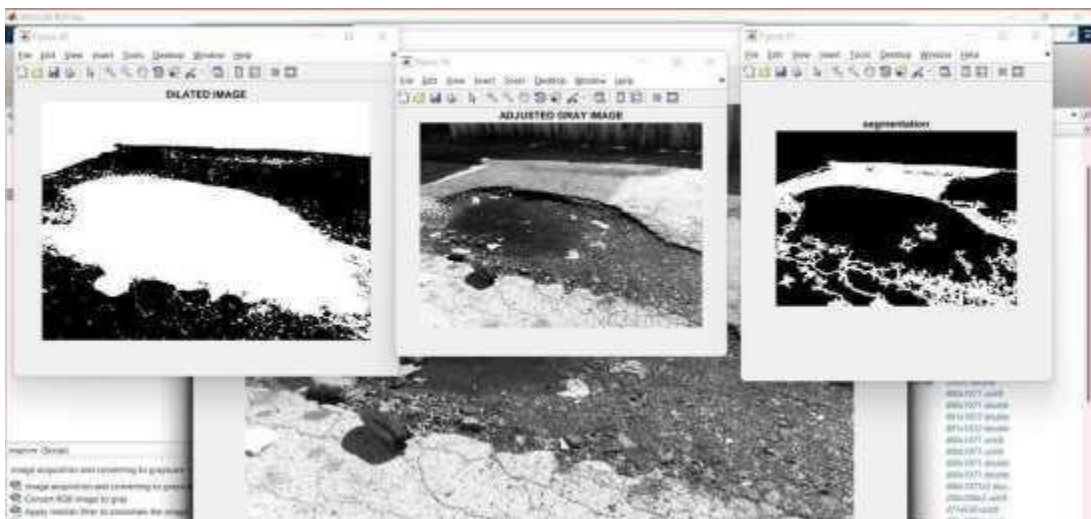


Fig:5.3 Segmentation

Machine Learning-Based Segmentation: Train machine learning models, such as semantic segmentation or instance segmentation models, to automatically segment potholes based on labeled training data.

- **Post-Processing:** After segmentation, apply post-processing techniques to remove noise, filter out small objects, and refine the segmented regions.
- **Region of Interest (ROI) Selection:** Define a region of interest where you expect potholes to be present. Segment within this ROI to reduce computational load and focus on relevant areas.
- **Adaptive Segmentation:** Implement adaptive segmentation techniques that adjust parameters based on local image properties. Adaptive methods can improve segmentation performance in varying conditions.
- **Validation and Quality Control:** Implement validation techniques to assess

the quality of segmentation results. These may involve comparing segmented regions to ground truth data. Parameter Tuning is a Fine-tune segmentation parameters and algorithms based on the specific characteristics of the images and potholes you are trying to detect.

5.4 SEGMENTAION USING KMEAN AND CONTOUR ALGORITHM

K-means Clustering: K-mean's clustering is an unsupervised learning algorithm used for clustering data into groups based on similarity. In the context of pothole detection, K-means clustering can be applied to group pixels in the image into clusters based on their color or intensity similarity. By specifying the number of clusters (K), the algorithm partitions the image into K distinct regions, each representing a different color or intensity range.

Contour Detection Algorithm: After applying K-means clustering, contour detection algorithms can be used to identify the boundaries of the segmented regions. Contours are curves joining continuous points along the boundary of an object in an image. In pothole detection, contour detection algorithms such as the one provided by OpenCV can help isolate the boundaries of potential potholes in the segmented regions.

Integration for Pothole Detection: Once the contours of the segmented regions are identified, additional criteria (e.g., size, shape, texture) can be used to filter out false positives and refine the detection of potholes. For example, potholes are typically circular or irregularly shaped depressions in the road surface, so contours with certain geometric properties can be selected as potential potholes.

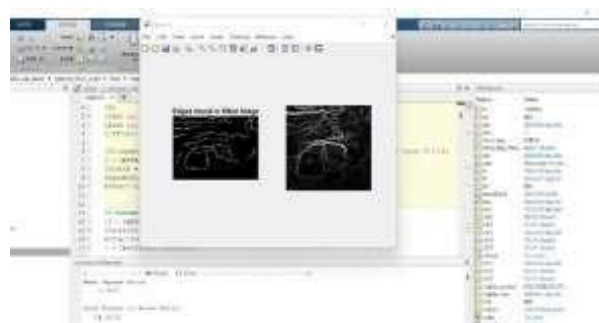


Fig:5.4 Edge found in dilated image

Post-processing and Classification: After semantic segmentation and contour detection, further post-processing steps may be applied to refine the results and improve accuracy. For instance, additional feature extraction techniques or machine learning classifiers (such as CNNs) can be used to classify the segmented regions as potholes or non-potholes based on their characteristics.

5.5 CLASSIFICATION USING CNN

Data Collection and Preprocessing: The first step in training a CNN for pothole detection is to collect a dataset of road images labeled with pothole annotations. These images can be captured using cameras mounted on vehicles or smartphones, or obtained from publicly available datasets. The dataset should include a diverse range of road conditions, lighting conditions, and types of potholes to ensure robustness of the trained model.

Before training the CNN, the dataset needs to be preprocessed to standardize the image sizes, adjust the contrast and brightness, and remove any irrelevant background noise. Preprocessing techniques such as resizing, normalization, and data augmentation can help improve the generalization and performance of the model.

Network Architecture: The architecture of the CNN plays a crucial role in its ability to learn and extract meaningful features from the input images. A typical CNN architecture for pothole detection may consist of multiple convolutional layers followed by pooling layers to extract hierarchical features from the input images. Additional layers such as batch normalization and dropout may be used to improve the stability and generalization of the model.

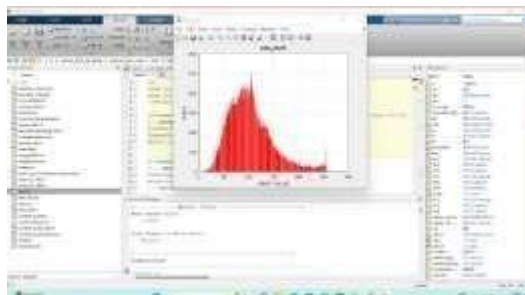


Fig:5.5 CNN graph

The final layers of the CNN architecture typically include fully connected layers followed by a softmax activation function to classify the input images into pothole and non-pothole classes. The output of the softmax layer represents the probability distribution over the two classes, and the class with the highest probability is chosen as the predicted class label.

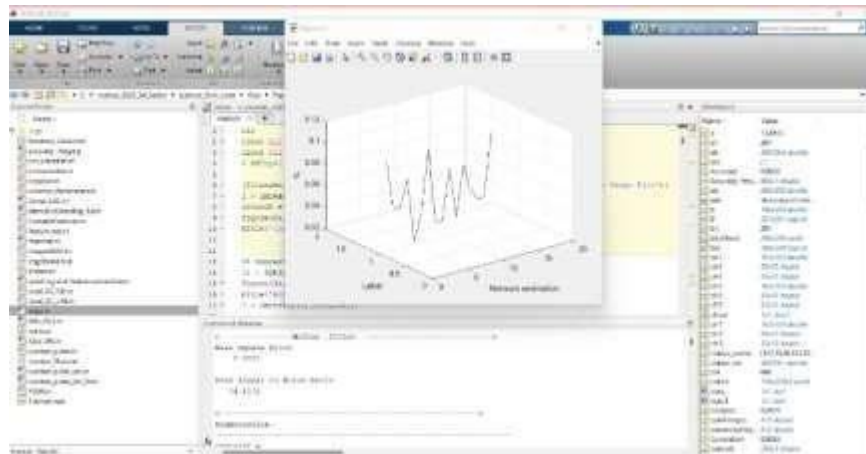


Fig:5.6 Network Estimation

Training Process: Once the CNN architecture is defined, the next step is to train the model using the labeled dataset. The training process involves iteratively optimizing the parameters of the CNN to minimize the classification error on the training data. This is typically done using gradient-based optimization algorithms such as stochastic gradient descent (SGD) or Adam.

During training, the CNN learns to identify discriminative features that are indicative of potholes from the input images. The loss function used for training is typically a categorical cross-entropy loss, which measures the difference between the predicted class probabilities and the ground truth labels.

Evaluation Metrics: After training, the performance of the CNN model is evaluated using various metrics such as accuracy, precision, recall, and F1-score. Accuracy measures the overall correctness of the predictions, while precision and recall quantify the model's ability to correctly identify potholes and avoid false positives, respectively. The F1-score provides a balance between precision and recall and is often used as a single metric to evaluate the overall performance of the model.

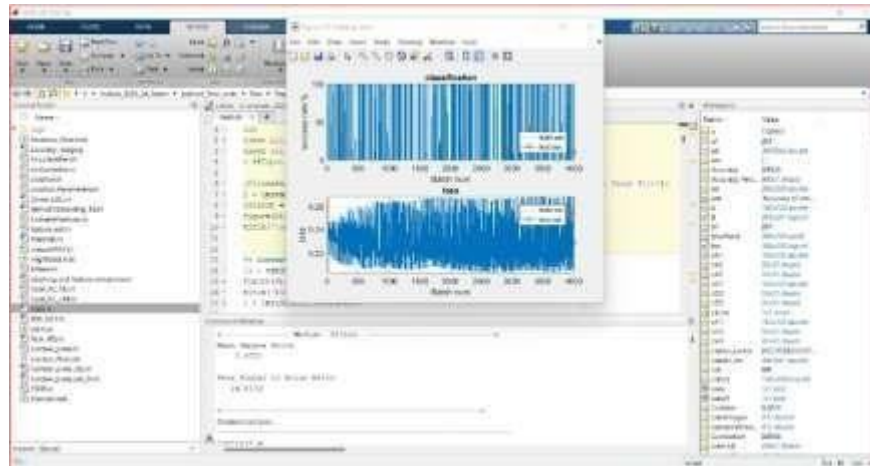


Fig:5.7 Classification and Loss

CNNs offer a powerful and effective approach for automated pothole detection from road imagery. By leveraging deep learning techniques, it is possible to train accurate and robust models that can detect potholes with high precision and recall. With further research and development, CNN-based pothole detection systems have the potential to significantly improve road maintenance and safety measures.

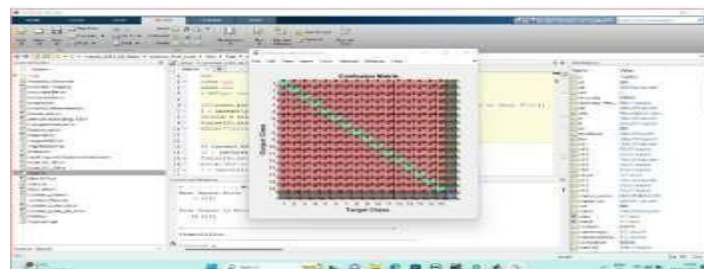


Fig:5.8 Confusion Matrix

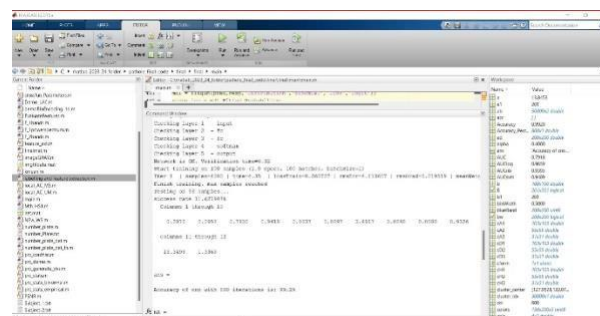


Fig 5.9 Accuracy of yolov7

CHAPTER 6

CONCLUSION AND FUTURE ENHANCEMENT

6.1 CONCLUSION

Various choices for implementing the System have been studied. These choices were also compared to each other on various criterion. We have specified the High level design choices of the Subsystems and justified the corresponding selections. We did experiments on the platform called MATLAB. We were also able to characterize road condition into categories like smooth, moderately uneven and highly uneven from the results of the experiments.

6.2 FUTURE ENHANCEMENT

In future we propose to do more experiments with variety of the scenario. Our next and immediate step is to use the Accelerometer on real vehicles and measure their response. Apply different scenarios like potholes on the slopes, turns and see how the accelerometer readings can characterize such condition. Along with that in remaining January and February we want to come up with and formalize the alternate solution for the Localization subsystem. In next months we perform simulations for given scenarios in Communication subsystem. And we plan to come up with values for the related parameters. For example Loss in throughput because of increase in number of vehicle, Loss in vehicle throughput because of the increase in speed. In the later months we perform the more experiments regarding the Localization system. If we are not able to come up with alternate solution; Then we will perform the experiments regarding integration of the rest of the system with GPS.

REFERENCES

- [1] Bradski, G. & Kaehler, A., 2008. In: Learning OpenCV. California: O'Reilly Media Inc., p.p154.
- [2] Buza, E., Omanovic, S. & Huseinovic, A., 2013. Pothole Detection with Image Processing and Spectral Clustering. Antalya, Turkey, 2nd International Conference on Information Technology and Computer Networks.
- [3] Danti, A., Kulkarni, J. & Hiremath, P., 2012. An Image Processing Approach to Detect Lanes, Pot Holes and Recognize Road Signs in Indian Roads. International Journal of Modeling and Optimization, 2(6), pp. 658-662.
- [4] Intelligent Robots and Computer Vision XIX: Algorithms, Techniques, and Active Vision. Proc. SPIE 4197, pp. 70-80.
- [5] Joubert, D., Tyatyantsi, A., Mphahlehle, J. & Manchidi, V., 2011. Pothole Tagging System. Pretoria, 4th Robotics and Mechatronics Conference of South Africa.
- [6] Kaehler, A. & Bradski, G., 2008. Learning OpenCV. 1st ed. California: O'Reilly.
- Karuppuswamy, J., Selvaraj, V., Ganesh, M. & Hall, E., 2000. Detection and Avoidance of Simulated Potholes in Autonomous Vehicle Navigation in an Unstructured Environment.
- [7] Koch, C. & Brilakis, I., 2011a. Pothole detection in asphalt pavement images. Advanced Engineering Informatics, August, 25(3), pp. 507-515.
- [8] Koch, C. & Brilakis, I., 2011b. Improving Pothole Recognition through Vision Tracking for Automated Pavement Assessment. Enschede, Netherlands, The18th EG-ICE Workshop on Intelligent Computing in Engineering, pp. 1-8.
- [9] OpenCV, 2014. The OpenCV Reference Manual Release 2.4.9.0. [Online] Available at: docs.opencv.org/opencv2refman.pdf [Accessed 18 July 2014].
- [10] Szeliski, R., 2011. Computer Vision - Algorithms and Applications. London: Springer.
- [11] Viola, P. & Jones, M., 2001. Rapid Object Detection using a Boosted Cascade of Simple Features. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 511-518.

