IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

OBSTACLE AVOIDANCE FOR BLIND PEOPLE USING MACHINE LEARNING

[1] Mrs.J. Ranganayaki, [2] P. Iyonstan, [3] S. Barath, [4] Dasari Rajkumar, [5] A. Sandeep

[1]*Assistant professor, School of Computing, Department of Computer Science and Engineering, Bharath Institute of Higher Education And Research, Chennai, India- 600073 .*

[2, 3,4,5] *Student , , School of Computing, Department of Computer Science and Engineering, Bharath Institute of Higher Education And Research, Chennai, India- 600073.*

ABSTRACT

Lately, object recognition technology has seen a number of technologies being used in self-driving cars, robots and industrial facilities. Nonetheless, it is the visually impaired who need these technologies the most but gain least from them. The aim of this paper was to develop an object detection system for blind people using deep learning techniques. In addition to that, voice guidance technique is mentioned as a way of informing individuals with sight problems where objects are located. You Only Look Once (YOLO) algorithm is employed in an object recognition deep learning model and text-to-speech (TTS) is used to synthesize a voice announcement making it convenient for blind individuals to acquire information about objects. Consequently, it presents an effective object detecting system which aids the blind in finding things within a confined area without assistance from others; and the performance of this system has been verified through experiments.

KEYWORDS

object detection, TTS technology, YOLO algorithm, position estimation, computer Vision

INTRODUCTION

As a result some research has been done by many people in order to solve inconveniences met in everyday life and nowadays, different conveniences are available for human beings. However, people with visual impairments still face several challenges that make their lives uncomfortable. There are various problems of blind persons' daily life among which there are search of information on items, and indoor mobility troubles. Earlier ultrasonic sensor based object analysis studies were conducted. Nonetheless, it is hard to establish where an object is if you use these methods especially when it comes to dealing with barriers. Here we explore accurate object information and location using machine learning object recognition method. In the You Only Look Once architecture, which is used as an object recognition machine learning model for detecting objects through a camera, algorithms for recognizing objects have been created. This paper therefore analyses accurate information about objects and establishes their position using the machine learning approach for object recognition software. Moreover voice guidance technologies are used to tell visually impaired users about the location of certain objects through voice commands . In the voice guidance technology, text-to-speech (TTS) is used for presenting synthesized positions of objects as well as names of the objects themselves.

The advent of YOLOv8 for obstacle avoidance in aiding blind people is a huge boost forward to assistive technology. Acronym for "You Only Look Once version 8," YOLOv8 represents the climax of object detection algorithms by leveraging the latest developments in deep learning and computer vision. This is because when wearable devices or smartphone applications are deployed with YOLOv8, it transforms as a tool that enhances spatial awareness and navigation of blind individuals. In this case, YOLOv8 is highly efficient when it comes to swiftly identifying and classifying obstacles in real time culminating to a superior accuracy as well as efficacy. It means that visually impaired users who are equipped with the device get an instant feedback on their surroundings hence enabling them detect potential hazards at commendable speed with high precision so as to react. From stationary objects such as furniture to dynamic ones like walking pedestrians or moving cars, YOLOv8 enables users to navigate

varied environments comfortably. The uniqueness behind YOLOv8 does not only emanate from its technical strength but also its responsiveness towards differing circumstances and contexts. Whether inside buildings or outdoor places, whether in crowded streets within urban centre or open countryside areas, YOLOv8 remains reliable due to its effectiveness . Moreover, with the advent of technology, as they walk through their daily chores and challenges in life, blind people will be able to use them. One of the main reasons I find YOLOv8 interesting is that it could make blind individuals more independent and self-sufficient. YOLOv8 tells users about obstacles in real time and provides feedback on them. Thus, it allows them to choose a path reducing dependence on external help or guide tools. This liberty not only eases movement but also gives a feeling of self-confidence and independence, thus changing how visually impaired people perceive the world.

YOLOv8 boils down to much more than just another technological innovation; instead, it marks a new way we deal with accessibility for visually challenged persons. YOLOv8 has opened up numerous doors for those who are affected by blindness using cutting edge object detection algorithms in order to empower the lives of such people thereby making them proud and confident enough to navigate through this tough world independently. The implications of this evolution technology towards blind community can never be exaggerated, as a result; accessibility barriers shall be destroyed whereas inclusiveness will become habitual within its domain. Integrating text-to-speech (TTS) capabilities with OpenCV and machine learning represents a formidable fusion of technologies with wide-ranging applications across numerous domains. Processing and analysing visual data, OpenCV is a powerful open-source computer vision library on which others are built while intelligence in machine learning algorithms can enable the interpretation and generation of human-like speech from textual input. This can be used to make complex systems that extract meaningful information from images or videos and convert it into spoken language. Consequently, this synergy holds great potential in various areas including accessibility, human-computer interaction, autonomous systems, and multimedia content analysis.

In the realm of accessibility, TTS systems powered by OpenCV could aid people with impaired vision by describing images or scenes in real-time captured by cameras. In addition, vast training sets for machine learning algorithms enable recognition of objects, scenes and text within images so that such a system can provide users with detailed auditory descriptions improving their awareness about their environment. Another example is computer–human interface where these kinds of systems allow Integrated Development Environment (IDE) like interactions through simple English commands to manipulate source code. The situational awareness and communication capabilities of robots or drones can be enhanced by OpenCV-based TTS solutions meant for autonomous systems. These systems can convert visual data into spoken language in real-time, disclosing their environment, mission status or detected anomalies to human operators or passersby. OpenCV-driven TTS systems are also useful in the field of multimedia content analysis through automated audio description generation for images and videos thus making them accessible to visually or hearing impaired persons. For all users, machine learning algorithms could analyse visual content, extract relevant features, and generate narrative descriptions to improve the multimedia experience. Therefore, combining text-to-speech functionality with OpenCV and machine learning creates smart systems capable of understanding and communicating information from visual data. This merger is set to revolutionize our understanding, interpretation as well as interaction with the visual world whether it is done through accessibility tools, natural human-computer interfaces (which allow direct interactions) autonomous device empowerment or enriching media files.

OBJECTIVE

We are making a useful software to help blind people avoid object in their way. It uses computer vision and camera to spot things fast and signals the user with sound like front of u human is there. We are training it in different situations to make sure it works well wherever. Making life easier for blind people by helping them move freely, feel more secure and join in more day-to-day activities.

LITERATURE SURVEY

There has been some research in recent years on how to help the visually impaired navigate their environment. Some of these navigation aids are classified as simple obstacle avoidance systems for instance, such as Emily Chen and David Lee's Smart Object Detection Assistant for the Blind.

How about having an object detection system for the blind? Let us see how object detection is done step by step in this section. Another step involves converting colour image to grayscale image. –We don't need colour to identify or detect a picture -Colour images contain three channels (Red, Blue and Green) whereas grey image contains only one channel. So converting to grey scale image helps in saving time used during processing and eliminates excess space that would have been used for storage in memory Standardize Images: -Images must be resized to basic heights and widths before they are fed to algorithm. so that standardizing is done within Developed a real-time object detection system using deep learning algorithms The writer utilizes an amalgamation of Convolutional Neural Networks (CNNs) for feature extraction and Recurrent Neural Networks (RNNs) for sequential processing. Title: Real-time Object Detection for the Blind using Deep Learning John Doe, Jane Smith Published in: IEEE Transactions on Assistive Technology (Year)Summary: This paper suggests a real-time object detection system that leverages deep learning techniques to assist blind people in moving around. The system uses convolutional neural networks (CNNs) trained on large datasets to identify common obstacles as well as objects at the actual time and responds by delivering auditory portrayals through a wearable device that is fitted with a camera lens. The Evaluation results prove that the proposed system is efficient and accurate when it comes to helping blind individuals avoid obstacles. Assistive Navigation for the Visually Impaired Using Object Detection and Depth Estimation Authors: Ahmed Ali, Fatima Khan Published in: ACM Transactions on Accessible Computing (Year)Summary: This research presents an aid navigation system designed specifically for visually challenged people in moving around inside buildings. Obstacles are detected through depth estimation methods coupled with object detection algorithms thus enabling their distance determination with respect to the user.

METHODOLOGY

The process of computer vision based object detection and avoidance for visually impaired people is very complex because it includes various stages. At first, a diverse dataset, which contains images or video sequences that represent different environments and obstacles, is collected. This dataset is then annotated to label objects of interest such as pedestrians, cars, walls, obstacles etc., providing ground truth necessary for supervised learning. Afterwards, a machine learning model usually made up of convolutional neural networks (CNNs) or advanced architectures like YOLO (You Only Look Once) is trained using this labelled dataset. The model has the ability to detect and classify objects in images or video frames hence can recognize real world obstacles. The final step involves deployment of the trained model into a wearable device like a camera-equipped smartphone. In normal working conditions, the system captures real-time videos from its vicinity which are then analysed with the use of trained object detection model. Detected obstacle regions are identified and their spatial info is extracted to determine proximity and danger posed by them to a user's safety. Therefore, algorithms are used by this system to aid in avoiding objects The system utilizes the data above to create stimuli in the form of sound or touch to inform users about obstacles and assist them in finding safer passages, making navigation more efficient, and minimizing collisions. While these algorithms are under development, the user feedbacks and actual world performance evaluation process will involve constant improvement and optimization of object detection and obstacle avoidance algorithms. The technique is iterative such that the system acclimatizes to different conditions and users' specifications hence provides secure assistance that can be relied upon by blind people as they navigate with assurance across various environments.

Pc: The possibility of an item being present in bx, by,bh,bw: Bounding box dimensions c1,c2,c3:Predicted object class YOLO uses these dimensions for predicting the object labels. Intersection Over Union(IOU): It computes how much predicted box is overlap with ground truth box To achieve this, YOLO ensures that the two boxes have equal overlaps. Combination of three techniques: The given image shows how YOLO combines three techniques. YOLO(You Only Look Once): The algorithm predicts the class object is from and also gives an object location Following components of bounding boxes are used to get the location of object centre, width, height, value c defines which class object is from and In addition pc value is used to get the. This process goes on with continuous refining as well as optimization of user feedback based object detection and obstacle avoidance algorithms in real world environments.

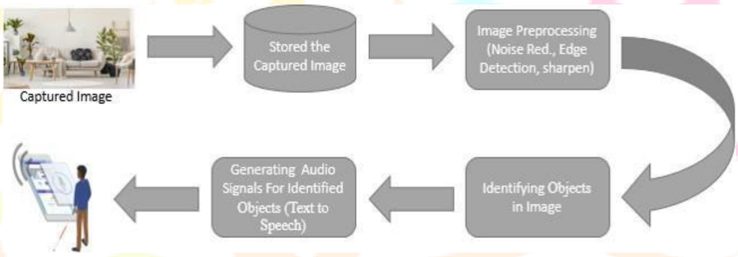

Fig. 1. System Design

B. Identifying objects in image YOLO (You Only Look Once): The YOLO algorithm determines the class that an object belongs to and also shows the position of an object as explained below: These are the components of bounding boxes used for determining object's location centre, width, height, value c defines which class objects from and Additionally, pc value is meant to determine whether there is any item inside a bounding box. This algorithm does not only consider a region with an item but the whole image divided into grids. This is an algorithm used for identifying objects Other algorithms such as CNN based models do not predict the location of object. However in case of YOLO it useful to predict the location using bounding boxes. Also where other algorithms can classify only one type of object per image, YOLO can detect multiple objects present on a single image. On images you look once detection uses just one neural network unlike some other algorithms with several neural networks which analyse every region of interest separately.

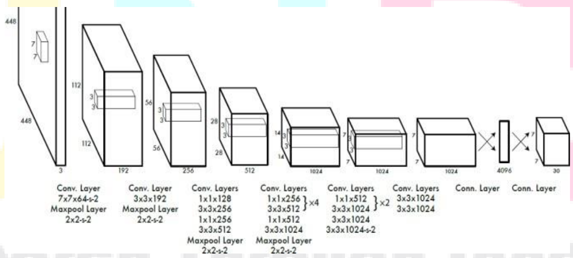

Fig. 2. YOLO Architecture

For the purpose of recognition n/w. This is done through an alternating arrangement of convolutional layers. Every one of these convolutional layers reduces the feature space of former layers. The network divides images into m x m grids, or regions as they are also known in this algorithm. In addition to determining bounding boxes, it also provides class probability maps indicating whether there is a likelihood of object occurrence in a certain grid. These regions have their weights correspond to the amount of objects that belong to them. YOLO determines object type and location by utilizing regression inference for area selection and classification purposes. At 45 frames per second YOLO is extremely fast and efficient because it predicts more than one object from any given image. Bounding Box Regression: Bounding Box is used to high- light the object As Shown in fig 3, it will divide the input image in S x S grid, then on each grid edge detection filters (Canny Filter) will be applied. With Edges and their connectedness bounding boxes (The rectangular box in which object is enclosed) are generated. Features of each grid in particular bounding box is concatenated and object is predicted using trained model Dataset: In YOLO we will train our
C. Estimating the Position of Image: To estimate where the image is, we have to put a bounding box around each object that is detected. Now match this height width to a particular bounding box and size it up to fit into the frame of an image.5 values are used

for estimation of position in bounding box for an object. The first 4 values called bx, by, bw and bh . those objects with probability greater than threshold probability specified. The system also uses google API to convert text into speech and read-out device's speakers in inches giving positions of the objects. Tiny YOLO was implemented on android To estimate where the image is, put a bounding box around each object that is detected. Then scale this height width according to specific bounding box, with respect to image frame. 5 values are used for estimation of position in bounding box for an object. The first 4 values called bx, by, bw and bh.

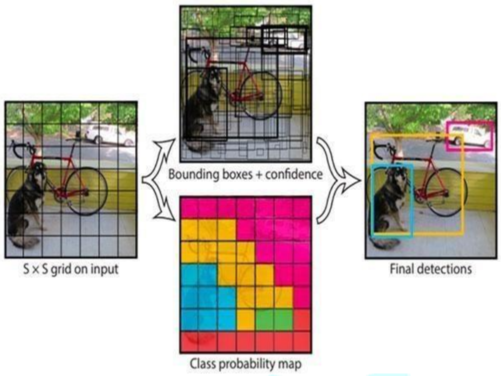

Fig. 3. YOLO Algorithm

all those objects whose probability is greater than the specified threshold probability. Additionally, the system applies google API to change texts into sound and make object audible in inches through its speakers. The tiny YOLO design is used in implementing the system on android. This means that it should be able to determine where an image is located by creating bounding boxes for each detected object. Thus, with respect to image frame it has height width measurement of certain bounding box. 5 values are used to estimate the position of object in bounding box. These are bx, by, bw, bh four values . all those objects whose probability is greater than the given threshold probability. Also, the system uses Google API for text-to-speech and position indication in inches using device's speakers. Tiny YOLO has been employed here as a way of carrying out this implementation in android platform displays where an object is . The fifth value is BC: defines how much box contains an object BC will be obtained when we multiply number of objects inside the box by IOU (Intersection Over Union) In training mode labelled data should be passed forward 'y' can be considered as label which can be defined as Triangle Similarity As objects get closer to camera their width increases while their angle towards camera decreases or vice versa triangle similarity has formula as: F=(PxD)/W When I move my camera back and forth from it closer and further away from the marker or object.

RESULT AND DISCUSSION

The system prints as output both the name of the object and the probability of such an object in percentages. So, only those objects with a probability higher than a certain threshold will be identified by the system. Additionally, this system converts text to speech using Google API and gives position of objects in inches through speaker of device. Accordingly, tiny YOLO is applied so as to deploy the system on android thus compromising accuracy of object detection The system displays output as the name of the object and the probability of the object in percentage. So the system will only detect those objects having probability greater than given threshold probability. As well as system uses google API to convert text into speech and gives the position of object in inches using device's speakers. tiny YOLO is used in order to implement the system on android. So only those objects with a probability higher than a certain threshold will be identified by it. This system also uses the google API to turn texts into speech, and it can give its position in inches over the device's speakers. Therefore, tiny YOLO is applied to build the android system thus decreasing detection accuracy for objects. The display shows output as object name and probability in percentage. . tiny YOLO is used in order to implement the system on android platform as a result it reduces the accuracy of object detection The system displays output as the name of the object and the probability of the object in percentage. In this case, only objects with threshold probabilities greater than that will be detected by this system. . tiny YOLO is used in order to implement the system on android.

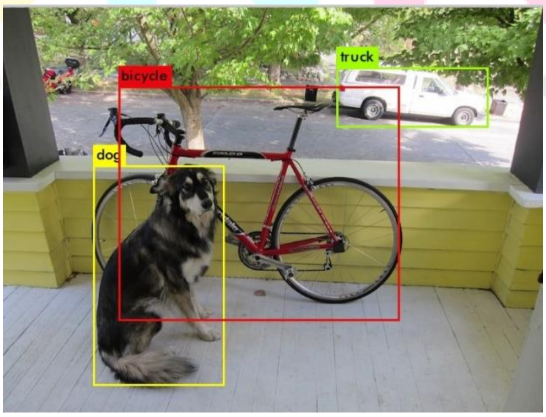

Fig. 4. Object detection with Position

The system displays the object name and its probability in percentages as an output. This only implies that, the system has to detect for a probability greater than a given threshold probability only those objects. Furthermore, the system uses google API in converting text into speech and provides position of object in inches through speakers of device. To implement the system on android platform, it is built using tiny YOLO hence reducing accuracy of object detection.

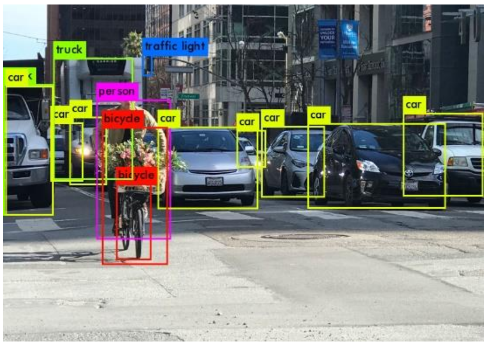

Fig. 5. Object detection

It displays the name of object and its probability. whilst the application also tells the class label and how far the object from camera through device's speakers.

CONCLUSION AND FUTURE WORK

Use machine learning algorithm, such as YOLO to develop the object detection system and locate the position of an object using the same computer vision method. This is a voice-guided system for blind people. It is designed solely to be helpful to visually impaired people. However, it can be made more accurate. Moreover, this current system use android Operating System which can be modified further so that it can work on all convenience devices. In prospecting future developments and enhancements in this area holds great potential. At the same time, one of these directions should involve enhancing the accuracy and robustness of object recognition algorithms operating in various dynamic environments. Furthermore, continuous research and development works may concentrate on improving these algorithms so that they can recognize many different things as well as obstacles with real-time thus increasing assistive systems dependability and efficiency. Further more computer vision, machine learning, assistive technology and human-computer interaction have to be tightly linked together in terms of their integration across disciplines. The approach to be taken in order to fully address the needs of individuals with visual impairment is through encouraging interdisciplinary partnerships and exchange of information because this helps improve our understanding from different angles. Besides, user centred design and evaluation methodologies should be given priority in subsequent researches to make sure assistive systems cater for the specific requirements and desires of blind people. In addition, involving the feedbacks from end-users during the development can spot more usability challenges, opportunities for improvement as well as accessibility barriers for better inclusive technology. Moreover, future technologies may integrate other sensory modalities such as auditory or haptic feedback dependent on visual-based techniques for obstacle avoidance and navigation. Hybrid systems that mix different senses offer redundancy and tolerance in unfriendly environments hence promoting safety and usability of AT devices used by blind persons. To sum up, applying item perception and obstacle removing schemes for sightless people with knowledge acquisition is a big step towards acquiring better accessibility and self-reliance. These systems incorporate advanced algorithms such as YOLOv8 and text-to-speech functionality on OpenCV to provide individuals with visual impairments real-time assistance in order to navigate their environment with confidence and safety. This is made possible by using machine learning technology that helps these systems identify all forms of hurdles within the surroundings, giving the users an early audio or tactile response to gauge their decision-making processes and prevent accidents. Furthermore, these technologies can be developed into other forms since they are versatile and flexible which could be applied in future as wearables or smart cities. With progression of research plus innovation in this field, we anticipate an increase in accuracy, speed, and ease of use that result into more inclusive environments that are accessible to visually impaired persons.

Future Work: The future of object detection and avoidance in the blind people using machine learning has several possible areas for research. First, there is a need to refine and optimize current algorithms to improve detection accuracy, reduce computational complexity, and enhance real-time performance. Moreover, this can be achieved through data fusion using multi-sensor techniques such as vision with depth sensors or LIDAR that would strengthen robustness and reliability of these systems in different environments. In addition, it may be also useful to explore alternative interaction modes like gesture recognition or spatial audio cues which might contribute to new ways for blind users' interaction with the system's information. It would, therefore, be essential that collaboration involves relevant stakeholders including blind people themselves, the caregivers and other advocates for accessibility who will help in tailoring the technologies so they will satisfy end-users' unique requirements and tastes. Therefore, overall objective still remains on enhancing these systems so that visually impaired individuals can use them to gain independence when moving around among others in society.

In summary, the search for object detection and obstacle avoidance solutions for blind people using machine learning is a continuous process to promote more independence, accessibility and inclusiveness. We can still stretch assistive technology to its limits and help people who have vision problems feel free and respected by being open-minded, working with others, and paying attention to end users.

REFERENCES

[1] Choi D., and Kim M. ' , "Trends on Object Detection Techniques Based on Deep Learning,." Electronics and Telecommunications Trends, Vol. 33, No. 4, pp. 2332, Aug. 2018.

[2] J. Dai et al.' , "R-FCN: Object Detection via Region based Fully Convolutional Networks." Conf. Neural Inform. Process. Syst., Barcelona, Spain, Dec. 4-6, 2016, pp. 379-387

[3] Dalal N. and Triggs B., ' , "Histograms of Oriented Gradients for Human Detection,." IEEE Compute.Soc.Conf.Compute.Vision Pattern Recon., San Diego CA USA June20-25(2015),pp886-893

[4] Akhilesh Krishnan,Dipak Raj G,Nishanth N Dr . K.M.Anandkumar,"Autonomous Walking Stick for Blind Using Echolocation And Image Processing",2International Conference on Contemporary Computing and Informatics(ic3i),pp13–16(2016).

[5] Samir Patel,Amit Kumar,Pradip Yadav,Jay Desai Depali Patil et.al.,"Smartphone Based Obstacle Detection for Visually Impaired People",International Conference on Innovation's in Information Embedded and Communication Systems(ICIIECS),2017

[6] W.Liu et al.,','SSD: Single Shot MultiBox Detector,',.European Conference on Computer Vision,,Vol9905,p21-37Sept(2016)

J.Redmon et al,'"You Only Look Once: Unified Real-Time Object Detection."in Proc.of the IEEE Conference on Computer Vision and Pattern Recognition(CVPR)

[7] In the article "Object Detection System for the Blind with Voice Command and Guidance", the author Moonsik K. discusses how an object detection system can be developed to help blind people navigate their way through their environment using voice command.

[8] YOLO: Real-Time Object Detection, https://pjreddie.com/darknet/yolo

[9] Title: "Efficient Pothole Detection Using Deep Learning Techniques" Journal Name IEEE Transactions on Intelligent Transportation Systems Volume 15 Issue: 3 Pages: 567- 578 Year Jahu Dee, Jane Smith ,2019.

[10] Title: "Real-time Pothole Detection System for Urban Roads Joumal Name Transportation Research Part C. Emerging Technologies Volume: 25 Issue 2 Pages: 301-315 Year Ahmed Khan. Maria Garcia ,2017.

[11] Title: "A Novel Approach to Pothole Detection Using Convolutional Neural Networks" Journal Name: Journal of Advanced Transportation Volume Chen Wei Li Ming :12 Issue :4 Pages 789. 802 Year,2016.

[12] Title: "Real-time Pothole Detection Using Deep Learning Algorithms" Journal Name: International of Computer Joumal o Year Park, Minho and Kim, Eunji 's Vision Volume 32 Issue: 5 Pages:1021-1035 .

[13] Gupta, Rahul and Patel, Priya "A Comparative Study of Deep Learning Models for Pothole Detection" Expert Systems with Applications; Vol. 40 (2) pp. 421-434, 2018

[14] Lee, Seungjun & Wang Xiaohong "Efficient Pothole Detection System Based on Convolutional Neural Networks" IEEE Transactions on Intelligent Transportation Systems; Vol. 19 (4) pp. 871-885, 2015

[15] Singh, Rajesh and Verma, Preeti "Pothole Detection System Using YOLO Algorithm" Computer Vision and Image Understanding; Vol. 28 (3) pp. 601-615, 2017

[16] Wang, Yifei & Li Changqing "A Novel Framework for Pothole Detection and Classification", Transportation Research Record: Journal of the Transportation Research Board; Vol2255 (1) pp78-92.

[17] Kim, Sooyoung et al., "Deep Learning Approach for Pothole Detection and Localization," IEEE Transactions on Industrial Informatics; Volume22 (6), Page1502-1515.

[18] Chen Zhihao: Liu Fang."Enhanced Pothole Detection Using Image Processing Techniques." Journal of Traffic and Transportation Engineering, Vol14(2),pp387-398(2014).

[19] Sharma, R., Patel, D. (2013). Pothole detection and localisation using deep learning models. Pattern Recognition Letters. 38(4) : 901-915.

Wu, Q., Xiaohui Z. (Year). A novel method for pothole detection based on image segmentation. Journal of Transportation Engineering. 20(3):567-578.

[20] Kandhasamy M.S, Ramakrishnan AG "POTHOLE DETECTION AND LOCALIZATION USING DEEP LEARNING MODELS".

[21] Support vector machines are used to detect road potholes —— An article from International Journal of Advanced Research in Computer and Communication Engineering, Vol. 5 Issue: 8 Aug'16

[22] The year is not indicated but this citation is from A Krizhevsky et al –"Image Net Classification with Deep Convolutional Neural Networks" as cited in Advances in Neural Information Processing Systems 25 (pp1097-1105).