Nanotechnology Perceptions
ISSN 1660-6795
www.nano-ntp.com

Vision Bridge: Empowering Blind Individuals Using Machine Learning

Maguluri V. Lavanya, B. Venkateswarlu, Palvancha Shambhavi, P. Rahul Das

Geethanjali College of Engineering & Technology, CSE, Hyderabad, Telangana, India
Email: lavanyamaguluri@gmail.com

Visual impairment and blindness are the usual outcomes of eye disorders. The World Health Organization estimates that 250 million people worldwide suffer from some form of visual impairment and 40 million individuals are blind. In their daily lives, they encounter a lot of difficulties, particularly when they are traveling alone. To cover their basic needs, they usually depend on help from other sources. Thus, it's a challenging task to establish a technology solution to support them. To assist the blind, numerous devices have been developed or visually impaired. One such endeavor is to create on the other hand, teaches blind sufferers to recognize and categorize every conceivable everyday object a voice warning the listener of both nearby and distant objects in their environment. Two distinct algorithms are used in the development of the system: Yolo is used for object identification, while GTTS (Google Text to Speech) is used for audio warnings. Both techniques are tested using the MS-COCO Dataset, which has about 200 K pictures. Webcams are used to assess both algorithms in a range of scenarios to be able to gauge each algorithm's accuracy under all conditions.
Keywords: CNN, YOLO, Machine learning, Disease Prediction.

1. Introduction

The Internet and the use of breakthroughs in daily life are examples of how quickly data and structured technology are progressing. Innovation in object acknowledgment, following that Object detection, which is among the technologies to considering. This describes the ability to recognize the size, form, and position of various objects while recording their position using the device's camera. Identifying occurrences of actual objects in still images or Identifying objects in videos, like cars, bikes, TVs, flowers, and people, is called object detection. It improves our overall understanding the image by enabling us to identify, locate, and detect a wide range of objects inside an image. It has long been a major difficulty in

computer vision to develop precise machinery learning models that can locate and identify many items in a single image. However, it's now simpler than ever to create Object Detection apps because of recent developments in Deep Learning. Developed on top of Tensor Flow, the Object identification API of Tensor Flow is an open-source platform that simplifies the creation, training, and implementation of object identification models. Numerous techniques can be used to detect objects. The World Health Organization reports that more than 40 million people worldwide are blind, with another 250 million who have some vision impairment. Particularly when they are by independently, they encounter numerous difficulties and ongoing difficulties in Navigation. They frequently need to rely on someone else to handle their basic daily requirements. Making a mechanical arrangement for them is therefore an exceptionally challenging undertaking, and this is crucial. Developing an integrated machinery learning framework is one of the objectives of our project. This will make it possible for people who are blind or visually impaired to recognize and categorize common objects with voice assistance. Which decide the distance between the object and the person and then sounds an alarm. The same framework can be used to construct identifying obstacles devices.

2. Existing System

Technology is increasing day by day, advancement in object recognition technology incorporates autonomous vehicles etc. is huge. There is the current system on object detection in fact already there but that are not reaching the visually impaired who need it most. Those systems that is already existed they may be object recognition or voice guidance and those are not tremendously useful for sighted. In our framework we have both object detection and after that information is converted in to voice that will be useful for blind people.

3. Proposed System

The System proposal is to examine the surroundings for object detection and a voice guidance to assist the blind people by using various machine learning and deep learning algorithms. By using these we can provides an effective object-detection system that enables blind people to locate objects in a given area on their own, without the assistance of others. This system accurately displays its item detection and labeling capabilities. In addition, the model measures the separation between the object and the camera and provides audio cues as the camera operator gets closer to the object.

4. Algorithms

Convolutional Neural Network

Artificial neural network models known as convolutional neural networks are modeled after the visual cortex that is a fully linked, layered network with individual layers including cell sections that are perceptually aware of distinct fields of vision. There are multiple layers in a neural network, including the input layer, one or more hidden layers, and an output layer.

Their greatest application is in object detection, where they can identify different patterns such designs, colors, textures, and edges (horizontal and vertical). In this kind of neural network, which functions as a filter, the convolutional layers are known as the hidden layers. After receiving input, it is relevant a particular pattern or feature to change it before forwarding it to the subsequent layer. When there are additional convolutional layers, the inputs are altered in different ways every time they are sent to the following convolutional layer (2826). The first convolutional layer, for instance, might be able to distinguish 12 shapes or colors in a region; the second might be able to determine what kind of object it is (for example, a car or a table); and the third and final convolutional layer might be able to classify the object as the actual predicted outcome. In other words, the more layers the data passes through, the more complex patterns subsequent ones will be able to identify. Convolutional layers, fully-connected layers, an output layer, and an input layer are the standard components of a convolutional neural network design. A few modifications to the LeNet Architecture were made when designing CNN. When input and output are excluded, it has six levels. The following figure illustrates the Convolution Neural Network architecture utilized in the research.

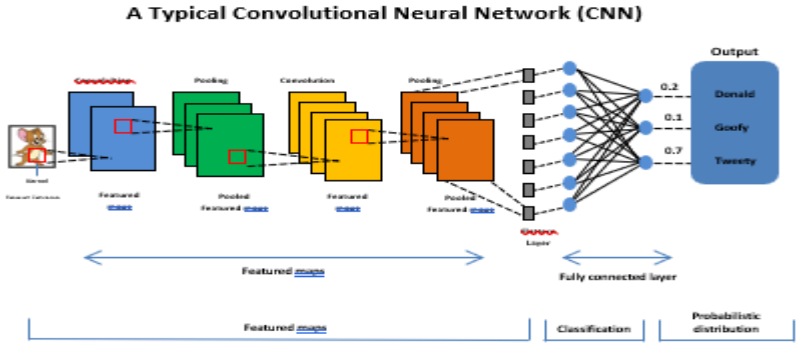

Figure 1.1 Convolutional Neural Networks

YOLO Algorithm

A well-liked object detection model called YOLO (You Only Look Once) is renowned for its quickness and precision. Since Joseph Redmon et al. first presented it in 2016; it has undergone multiple versions, the most recent of which is YOLO v7. Based on regression, the YOLO method predicts all of the classes and bounding boxes for an image in a single algorithm run, as opposed to focusing on the interesting portions of the image. The way Yolo operates is by dividing the input image into S×S cells, each of which is in charge of five bounding boxes.

Predictions(x,y,w and h) that describe the rectangle around the object and their confidence scores. Here, (x,y) represents the coordinates of the picture, and (w,h) represents width and height respectively. The confidence scores are predicted to be zero if no object is found. Else, it is predicted by Pr (Object) ∗ IOU()Intersection Over Union). Finally, the confidence prediction which represents the IOU between the predicted box and any ground truth box is found. Each grid cell can also forecasts C conditional class probabilities by using the formula Pr(Classi |Object). These probabilities are conditioned to be on the grid cell

containing an

object, where we only predict one set of class probabilities per grid cell, regardless of the number of boxes by using the formula Pr(Classi|Object) Pr(Object)_IOU = Pr(Classi) IOU, which gives us class-specific confidence scores for each box.

Once the bounding boxes are predicted.

The domain of object detection has seen tremendous change in recent years due to the introduction of cutting-edge architectures such as YOLOv4, which are supported by the sophisticated technologies discussed before. By segmenting a picture into smaller images and dividing them into a m×m square grid, YOLO divides an image into smaller images and uses a single neural network to conduct both prediction and classification for recognized items.

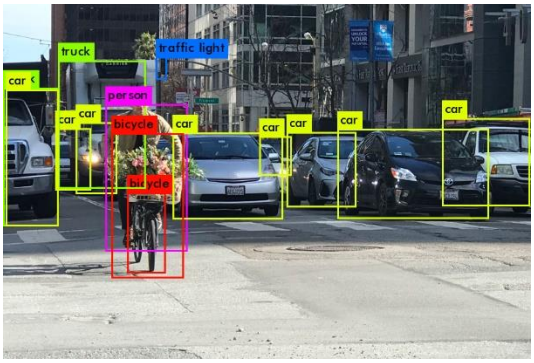

Figure 1.2 YOLO Algorithm

5. Summary & Outcomes

Output Screens

Figure 2.1 Output Screens with Warning Alerts

Figure 2.2 Output Screen when the objects are at safer distance

Figure 2.3

These in-depth examinations yielded definitive and very positive results. Remarkably, the method proved to be totally effective accomplishing all operational projects flawlessly and with remarkable 100% efficacy. Additionally, the product the detecting component performed fairly well, verifying its capacity to accurately and consistently identify objects under various conditions.

A warning signal will be sent to the user and bounding boxes will be displayed when the object is recognized by YOLO automatically. These favorable results show the system's durability and usefulness in realistic situations.

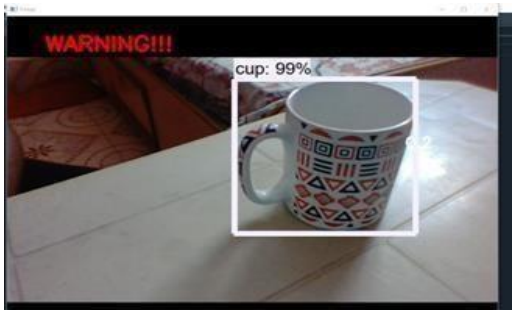

Figure 2.4 Output Screens with Warning Alerts

6. Conclusion

We successfully developed a model for the Subject Detection using YOLO (YOU ONLY LOOK ONCE) Algorithm which detects the objects more accurately and converts detected information into Voice Alerts with GTTS (Google Text to Speech). Many technologies have been developed to help those who are blind or visually impaired. One such endeavor would be to create an Integrated Machine Learning System that would enable the eye sightless victims to recognize and categorize objects in actual time while providing distance and speech feedback. Which, regardless of how close or how far they are from the object, also generates alerts? Sound guidance is made available by this technology for people who are blind. This method was developed especially to help those who are blind. However, there is room for improvement in terms of precision. In addition, the existing system may be modified to function with any convenient device because it is built on the Android operating system. It works better than other existing techniques for visually challenged people who

need it the most in their everyday existence. This technology gave us the capacity to develop those projects that will be of real benefit to the specific in need. Our system will help the visually impaired to navigate in the outside environments independently by detecting the object and the same location of it with necessary alert systems during navigation additionally during emergencies. Since the objects in the outdoor environment could be related to any subject, the dataset must cover sufficient ground so that the model is trained appropriately. It can provide a low-cost system for many in need.

References

[1] Ali JasimRamadhan, "Wearable Smart System For Visually Impaired People", Department of Computer Techniques Engineering Al-Kafeel University College Kufa 31003 Province of Najaf Republic of Iraq., pp. 843-856, 2018. Google Scholar

[2] Rupali A Tanpure, Pragathi D Borkar and Madhuri R Falke, "Advanced Voice Based Blind Stick with Voice Announcement of Obstacle Distance", International Journal for Innovative Research in Science & Technology, vol. 4, no. 8, pp. 85- 87, January 2018. Google Scholar

[3] R. Khlaikhayai, C. Pavaganun, B. Mangalabruks and P. Yupapin, "An intelligent walking stick for elderly and blind safety protection", Proceedings of the 2nd International Science Social Science Engineering and Energy Conference 2010: Engineering Science and Management Procedia Engineering, vol. 8, pp. 313-316, 2011.

[4] Rajeshvaree Ravindra Karmarkar (2021). Object Detection System for the Blind with Voice Guidance, International Journal of Engineering Applied Sciences and Technology,6(2): 67-70.

[5] CNN based Auto-Assistance System as a Boon for Directing Visually Impaired," International Conference on Trends in Electronics and Informatics C.-

[6] W. Lee, P. Chondro, S.-J. Ruan, O. Christen and E. Naroska, "Improving mobility for the visually impaired: a wearable indoor positioning system based on visual marker", IEEE Consumer Electronics Magazine, vol. 7, no. 3, pp. 12-20, 2018. Google Scholar.

[7] Vision Loss Expert Group of the Global Burden of Disease Study. Trends in prevalence of blindness and distance and near vision impairment over 30 years: an analysis for the Global Burden of Disease Study. Lancet Global Health 2020. doi: 10.1016/S2214-109X(20)30425-3.

[8] S. R. A. W. Alwi and M. N. Ahmad, "Survey on outdoor navigation system needs for blind people," 2013 IEEE Student Conference on Research and Development, Putrajaya, Malaysia, 2013, pp. 144-148, doi: 10.1109/SCOReD.2013.7002560.

[9] R. Ani, E. Maria, J. J. Joyce, V. Sakkaravarthy and M. A. Raja, "Smart Specs: Voice assisted text reading system for visually impaired persons using TTS method," 2017 International Conference on Innovations in Green Energy and Healthcare Technologies (IGEHT), Coimbatore, India, 2017, pp. 1-6, doi: 10.1109/IGEHT.2017.8094103.

[10] Mande Shen and Hansheng Lei, "Improving OCR performance with background image elimination," 2015 12th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD), Zhangjiajie, China, 2015, pp. 1566-1570, doi: 10.1109/FSKD.2015.7382178.

[11] K. Dhivya, G. Premalatha, and S. Monica, "Wearable Navigation Device for Virtual Blind Guidance," 2019 International Conference on Intelligent Computing and Control Systems (ICCS), Madurai, India, 2019,