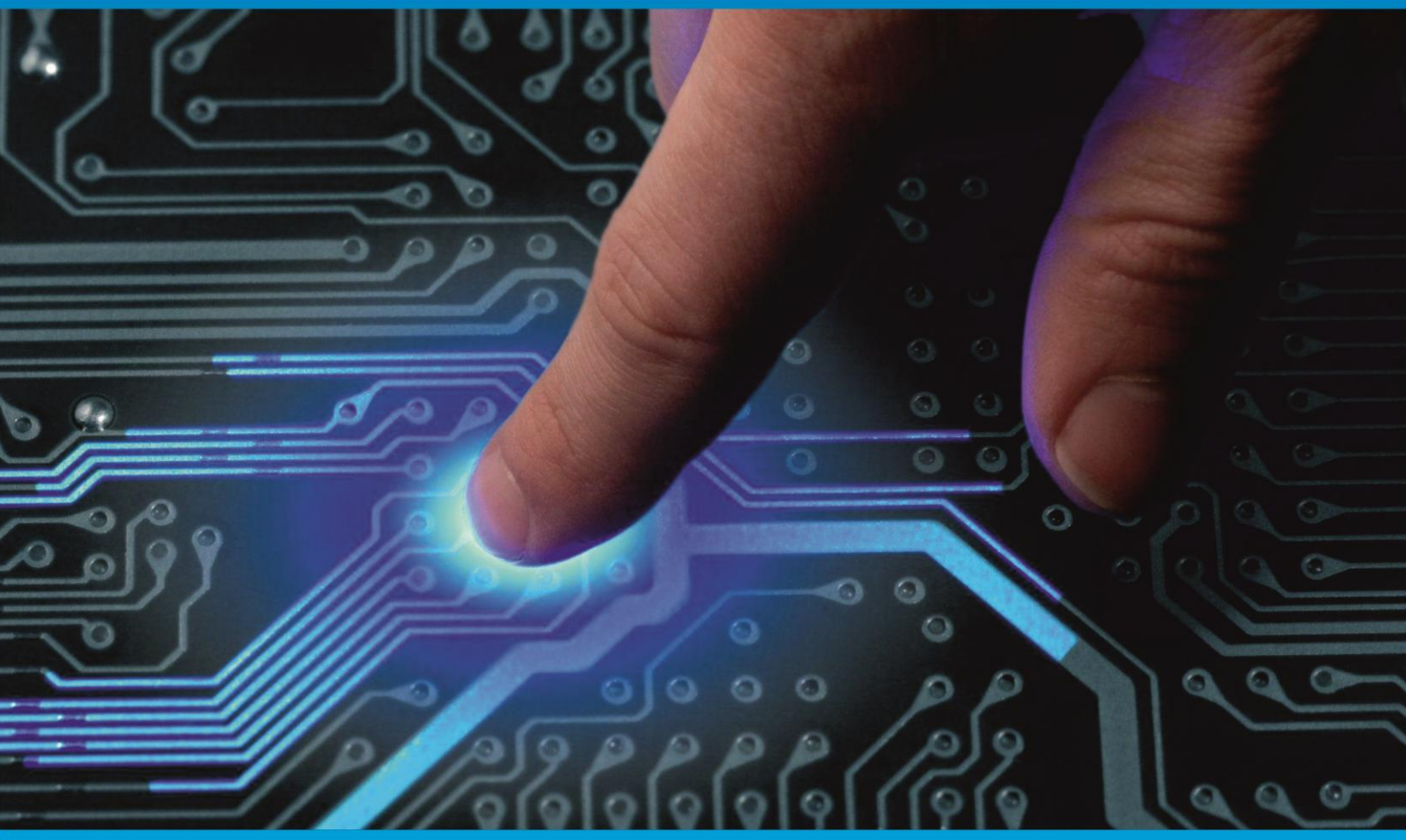
IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 9, Issue 5, May 2021

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.488

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 5, May 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0905130 |

YOLO Based Object Detection System

Prof. Sumeet Harale, Ankur Tripathi, Umesh walunjkar, Shubham mane,

Department of Computer, Indira College of Engineering and Management, Maharashtra, India

ABSTRACT: The Objective is to detect of objects using You Only Look Once (YOLO) approach. This method has several advantages as compared to other object detection algorithms. In other algorithms like Convolutional Neural Network, Fast Convolutional Neural Network the algorithm will not look at the image completely but in YOLO the algorithm looks the image completely by predicting the bounding boxes using convolutional network and the class probabilities for these boxes and detects the image faster as compared to other algorithms.

KEYWORDS: Convolutional Neural Network, Fast-Convolutional Neural Network, Bounding Boxes, YOLO.

I. INTRODUCTION

This project will be developed to make the life of blind people easy. This is a camera based system to scan the product behind the image and read the description of the product with the help of Id stored in the product. This is very beneficial in case of finding out the description of packaged goods to the blind people and thus helping them in deciding to purchase a product or not especially which are packaged. This is because it becomes very difficult for the blind people to distinguish between the packaged goods. In order to use this system, all the user needs to do is capture the image on the product in the device which then resolves the product which means it scans the image to find out the Id stored. Thus this application really benefits blind and visually impaired people and thus making their work of identifying products easy. This is very easy to use and affordable as it requires a scanner to scan the product and a camera phone to take the picture of the image containing the product. This is now easy to implement as most of the devices today have the required resolution in order to scan the product to identify the Id stored in it and read out the product description. This project can be implemented in any shopping mall, supermarket, Book stores, Medical stores etc.

II. LITERATURE SURVEY

You Only Look Once: Unified, Real-Time Object Detection, by Joseph Redmon. Their prior work is on detecting objects using a regression algorithm. To get high accuracy and good predictions they have proposed YOLO algorithm in this paper [1]. Understanding of Object Detection Based on CNN Family and YOLO, by Juan Du. In this paper, they generally explained about the object detection families like CNN, R-CNN and compared their efficiency and introduced YOLO algorithm to increase the efficiency [2]. Learning to Localize Objects with Structured Output Regression, by Matthew B. Blaschko. This paper is about Object Localization. In this, they used the Bounding box method for localization of the objects to overcome the drawbacks of the sliding window method [3]. System proposes a new approach to barcode decoding that bypasses binarization. This technique relies on deformable templates and exploits all of the gray-level information of each pixel. Due to this parameterization of these templates, system can efficiently perform maximum likelihood estimation independently on each digit and enforce spatial coherence in a subsequent step. System show by way of experiments on challenging UPC-A barcode images from five different databases that this approach outperforms competing algorithms. Implemented on a Nokia N95 phone, this algorithm can localize and decode a barcode on a VGA image (640 480, JPEG compressed) in an average time of 400-500 ms

III. BACKGROUND

RCNN
Region based convolutional neural networks (RCNN)
algorithm uses a group of boxes for the picture and then analyses in each box if either of the boxes holds a target. It employs the method of selective search to pick those sections from the picture. In an object, the four regions are used.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 5, May 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0905130 |

HOG: HOG is a feature descriptor that is extensively applied in various domains to distinguish objects by identifying their shapes and structures. Local object structure, pattern, aspect,and representation can usually be characterized by the arrangement of gradients of local intensity or the ways of edges.

RCNN:
Region based convolutional neural networks (RCNN)
algorithm uses a group of boxes for the picture and then analyses in each box if either of the boxes holds a target.

FASTER RCNN:
Faster RCNN is a transformed variant of fast RCNN.
The significant difference between both is that faster RCNN implements region proposal network (RPN)

IV. PROPOSED METHODOLOGY

1. Hardware set up:
Connect raspberry with camera and install drivers.
Connecting the hardware
Execute below command
raspistill -o image.jpg
Above command will start the camera and click the pic, and image will be save

2. Image Capture:
Capture real time image and send for processing.

3. Apply YOLO Algo:

◉ Applying the YOLO algo for Object Identification:

YOLO model divides an image into S×S grid. Each grid cell predicts B bounding boxes, and boxes' confidence scores for the prediction and detect if a class falls in the boxes. The confidence is defined as P r(object) × IOUtruth pred , which represents the confidence of a class in the box and accuracy of the box coordinates. Thus, each box has five parameters to predict: x, y, w, h and confidence. Each grid cell also predicts P r(Classi |Object). Thus the confidence for each box is P r(Classi |Object) × P r(object) × IOUtruth pred = P r(Classi) × IOUtruth pred . The overall variables to be predicted can be represented as a S × S × (B × 5 + C) tensor.

4. send voice output:
Process the image using image detection algorithm and send voice output.

The system send password to the user's mail after successful registration and also after in case user has forgotten the password.

Module 1: YOLO Algo:

This is done using the Yolo algorithm which accepts the images and converts it into gray scale and gives the results in the form of binary value.0 for black and 1 for white. It converts the large image into 50-50 cube and is then processed using Yolo algorithm. It divides the cube by 25 units each. It takes first image as input and then by pre-processing the final output is been drawn after respective expected output. The pre-processing has total 4*5=20 cubes for result analysis in which the first is input cube and the last is output cube.

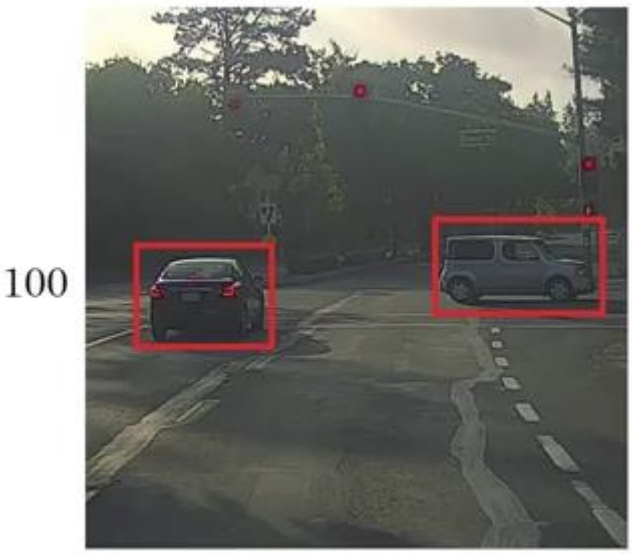
| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 5, May 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0905130 |

Yolo Algo

- Step 1: YOLO first takes an input image:



- Step 2: The framework then divides the input image into grids (say a 3 X 3 grid):
- Step 3: Image classification and localization are applied on each grid. YOLO then predicts the bounding boxes and their corresponding class probabilities for objects (if any are found, of course).

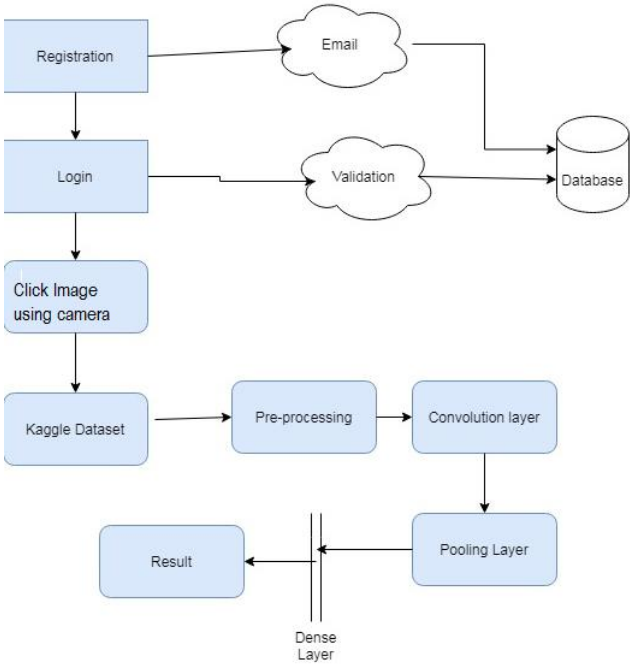
Module 2: Pooling Layer:

This allows in adjusting the plot accordingly. It has various parameters like top, bottom, etc. This is used to modify the clarity of the result and adjusting the focus in order to print the result of the analysed output.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

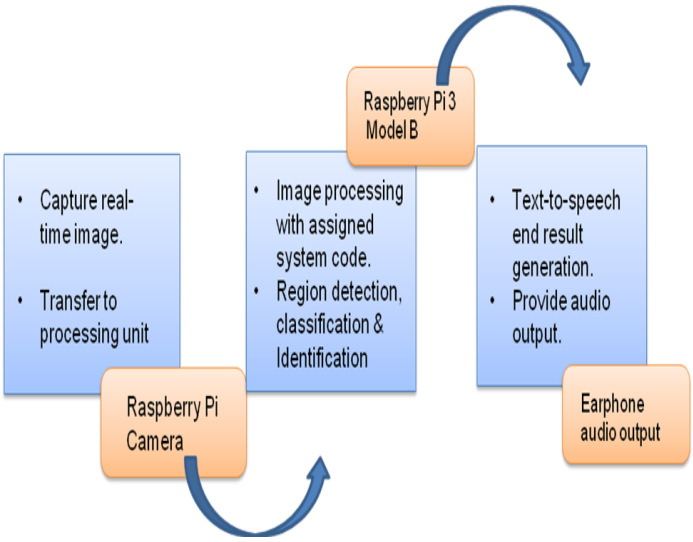
|| Volume 9, Issue 5, May 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0905130 |



V. SYSTEM ARCHITECTURE

The network has 24 convolutional layers with 2 fully connected layers. The ConvNet is to extract features from input images and the fully connected layers are to predict the probability of the boxes coordinates and confidence score. The accuracies of the predictions also depend on the architecture of the network. The loss function of the final output depends on the x, y, w, h, prediction of classes and overall probabilities. In our project, we use pretrained YOLO weight to detect Objects.



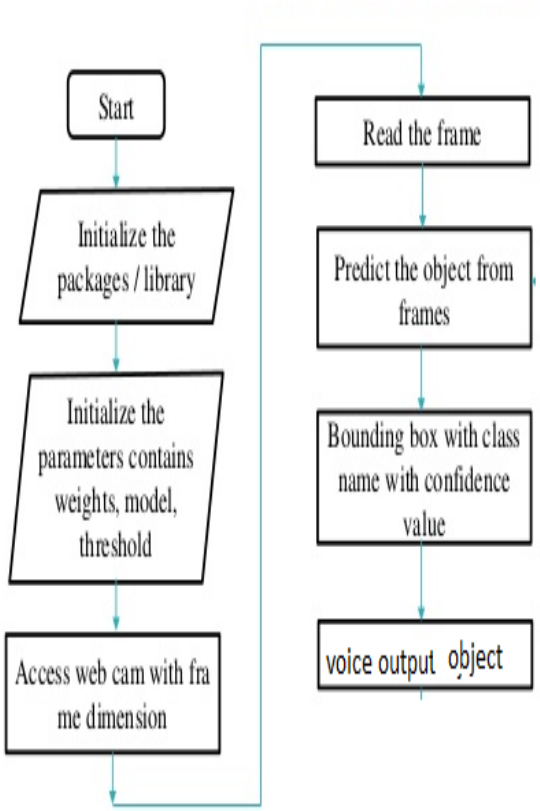
VI. FLOW DIAGRAM

- Visually Impaired person can identify the product
- Once product is identify there will be voice of that product.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 5, May 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0905130 |



VII. CONCLUSION

we will applied and proposed to use YOLO algorithm for object detection because of its advantages. This algorithm can be implemented in various fields to solve some real-life problems like security, monitoring traffic lanes or even assisting visually impaired people with help of audio feedback.

REFERENCES

[1] Joseph Redmon and Anelia Angelova, Real-Time Grasp Detection Using Convolutional Neural Networks (ICRA), 2015.

[2] A. Quattoni, and A.Torralba. Recognizing Indoor Scenes. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2009.

[3] Saurabh Gupta, Ross Girshick, Pablo Arbelaez and Jitendra Malik, Learning Rich Features from RGBD Images for Object Detection and Segmentation (ECCV), 2014.

[4] Tadas Naltrusaitis, Peter Robison, and Louis-Phileppe Morency, 3D Constrained Local Model for Rigid and Non-Rigid Facial Tracking (CVPR), 2012.

[5] Andrej Karpathy and Fei-Fei Li, Deep VisualSemantic Alignments for Generating Image Descriptions (CVPR), 2015.

[6] David Brown, Tom Macpherson, and Jamie Ward, Seeing with sound? exploring different characteristics of a visual-to-auditory sensory substitution device. Perception, 40(9):1120–1135, 2011.

[7] Liam Betsworth, Nitendra Rajput, Saurabh Srivastava, and Matt Jones. Audvert: Using spatial audio to gain a sense of place. In Human-Computer Interaction– INTERACT 2013, pages 455–462. Springer, 2013.

INNO SPACE

SJIF Scientific Journal Impact Factor

Impact Factor:
7.488

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details