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

Object Detection with Audio Feedback

Mr. S. Ranganath[1], Dr. K. Santhi[2], Dr. S. Jasmine Sabeena[3]
1 Dept. of CSE, SVCE, Tirupathi
2 Professor, Dept. of CSE, SVCE, Tirupathi
3 Associate Professor, Dept. of CSE, SVCE, Tirupathi

Abstract – Object recognition is one of the challenging applications of computer vision, which has been widely applied in many areas for e.g., autonomous cars, Robotics, Security tracking, Guiding Visually Impaired Peoples etc. With the rapid development of deep learning many algorithms were improving the relationship between video analysis and image understanding. All these algorithms work differently with their network architecture but with the same aim of detecting multiple objects within complex image. Absence of vision impairment restraint the movement of the person in an unfamiliar place and hence it is very essential to take help from our technologies and trained them to guide blind peoples whenever they need.

Index Terms: Tensor flow, Yolo_v3, Web Speech API, Deep Learning.

INTRODUCTION

Humans almost by birth are trained by their parents to categorize between various objects as children self is one object. Human Visual System is very accurate and precise that can handle multi-tasks even with less conscious mind. When there is large data then we need more accurate system to correctly recognize and localize multiple objects simultaneously. Here machines come into existence, we can train our computers with the help of better algorithms to detect multiple objects within the image with high accuracy and preciseness. Object Detection is the most challenging application of computer vision as it requires complete understanding of images. In other words, object tracker tries to find the presence of object within multiple frames and assigns labels to each object. There might be many problems faced by the tracker in terms of complex image, Loss of information and transformation of 3D world into 2 D image.

System Analysis & Feasibility Study

Existing Method:

In recent years many algorithms are developed by many researchers. Both machine learning and deep learning approaches work in this application of computer vision. This section outlines the journey of the different techniques used by the researchers in their study since 2012. Histograms of Gradient Descent (HOG) use SVM algorithm for detecting objects in real time. It is a feature detector used to

extract meaningful information from the image ignoring the background image. This algorithm works very effectively in detecting human and textual data.

Disadvantages:

- High complexity.
- Time consuming

Proposed System:

We propose a system that will detect every possible day to day multiple objects on the other hand prompt a voice to alert person about the near as well as farthest objects around them. To get audio we will use web speech API to produce speech.

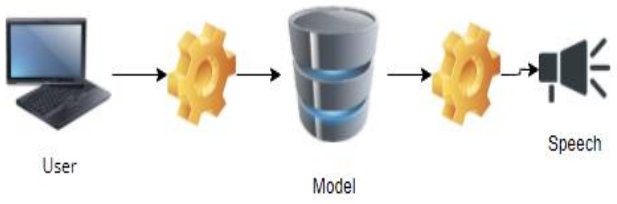

Fig: Architecture

Advantages:

- It increases the accuracy.
- It reduces the time complexity.

METHODOLOGY AND ALGORITHMS:

CNN:

In deep learning, a convolutional neural network (CNN, or ConvNet) is a class of deep neural networks, most commonly applied to analyzing visual imagery. They are also known as shift invariant or space invariant artificial neural networks (SIANN), based on the shared-weight architecture of the convolution kernels that shift over input features and provide translation equivariant responses.

Applications:

- Image recognition
- Video analysis
- Natural language processing
- Anomaly Detection
- Drug discovery
- Health risk assessment and biomarkers of aging discovery.

YOLO:

Yolo is a part of object detection, Object detection is a computer technology related to computer vision and image processing that deals with detecting instances of semantic objects of a certain class (such as humans, buildings, or cars) in digital images and videos. Well-researched domains of object detection include face detection and pedestrian detection. Object detection has applications in many areas of computer vision, including image retrieval and video surveillance.

Every object class has its own special features that helps in classifying the class – for example all circles are round. Object class detection uses these special features. For example, when looking for circles, objects that are at a particular distance from a point (i.e., the center) are sought. Similarly, when looking for squares, objects that are perpendicular at corners and have equal side lengths are needed. A similar approach is used for face identification where eyes, nose, and lips can be found and features like skin color and distance between eyes can be found.

Software Development Life Cycle – SDLC:

In our project we use waterfall model as our software development cycle because of its step-by-step procedure while implementing.

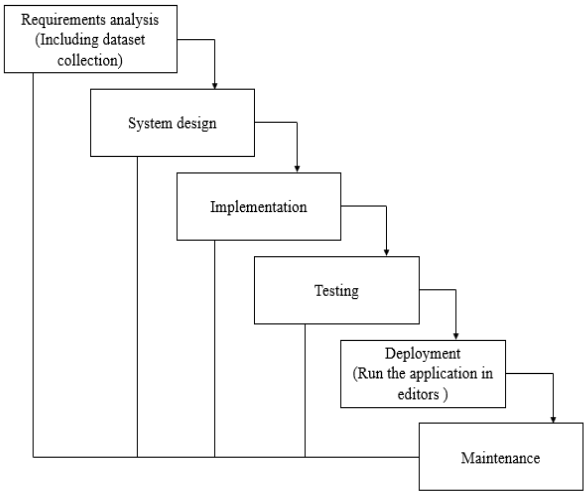

Fig: Waterfall Model

- **Requirement Gathering and analysis** − All possible requirements of the system to be developed are captured in this phase and documented in a requirement specification document.
- **System Design** − The requirement specifications from first phase are studied in this phase and the system design is prepared. This system design helps in specifying hardware and system requirements and helps in defining the overall system architecture.
  - **Implementation** − With inputs from the system design, the system is first developed in small programs called units, which are integrated in the next phase. Each unit is developed and tested for its functionality, which is referred to as Unit Testing.
  - **Integration and Testing** − All the units developed in the implementation phase are integrated into a system after testing of each unit. Post integration the entire system is tested for any faults and failures.
  - **Deployment of system** − Once the functional and non-functional testing is done; the product is deployed in the customer environment or released into the market.
- **Maintenance** − There are some issues which come up in the client environment. To fix those issues, patches are released. Also, to enhance the product some better versions are released. Maintenance is done to deliver these changes in the customer environment.

  They basically deal with issues like:

  • Portability
  • Security
  • Maintainability
  • Reliability
  • Scalability
  • Performance
  • Reusability
  • Flexibility

Examples of non-functional requirements:

1) Emails should be sent with a latency of no greater than 12 hours from such an activity.
2) The processing of each request should be done within 10 seconds
3) The site should load in 3 seconds whenever of simultaneous users are > 10000.

Objectives for Input Design:

The objectives of input design are −

- To design data entry and input procedures
- To reduce input volume
- To design source documents for data capture or devise other data capture methods
- To design input data records, data entry screens, user interface screens, etc.
- To use validation checks and develop effective input controls.

Output Design:

The design of output is the most important task of any system. During output design, developers identify the type of outputs needed, and consider the necessary output controls and prototype report layouts.

CONCLUSION:

In this work, we were able to detect objects using yolov3 and use web speech api to generate speech.

REFERENCES:

1. S. Cherian, & C. Singh, "Real Time Implementation of Object Tracking Through webcam," International Journal of Research in Engineering and Technology, 128-132, (2014) [2] Z. Zhao, Q. Zheng, P.Xu, S. T, & X. Wu, "Object detection with deep learning: A review," IEEE transactions on neural networks and learning systems.
2. N. Dalal, & B. Triggs, "Histograms of oriented gradients for human detection," In 2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05) (Vol. 1, pp. 886-893). IEEE, (2005, June).
3. R. Girshick., J. Donahue, T. Darrell, & J. Malik, "Region-based convolutional networks for accurate object detection and segmentation," IEEE transactions on pattern analysis and machine intelligence, 38(1), 142-158, (2015).
4. X. Wang, A. Shrivastava, & A. Gupta, "A-fast-rcnn: Hard positive generation via adversary for object detection," In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 2606- 2615), (2017).
5. S. Ren, K. H, R. Girshick, & J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," In Advances in neural information processing systems (pp. 91-99), (2015).
6. J. Redmon, S. Divvala, R. Girshick, & A. Farhadi, "You only look once: Unified, real-time object detection," In Proceedings of the EEE conference on computer vision and pattern recognition (pp. 779-788), (2016).