International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 05 Issue: 08 | Aug - 2021 ISSN: 2582-3930

Object Identification and Detection Using Deep Learning and Machine Learning

Ibtesham Abid Sheikh[1], Prof. Hirendra Hajare [HOD][2]

[1 2]Department of Computer Science and Engineering

Ballarapur Institute of Technology ,Ballarapur ,Maharashtra,India

Abstract - Computer Vision is the branch of the science of computers and software systems that can recognize as well as understand images and scenes. Detecting and recognizing objects in unstructured environments is one of the most challenging tasks in computer vision research. Computer Vision is consists of various aspects such as image recognition, object detection, image generation, image super-resolution, and many more. Object detection is probably the most profound aspect of computer vision due to the number of practical use cases. Object detection is widely used for face detection, vehicle detection, pedestrian counting, web images, security systems, and self-driving cars. There are different algorithms and methods of object detection such as R-CNN, Fast-RCNN, RetinaNet, and fast yet highly accurate ones like SSD and YOLO. Using these methods and algorithms, based on deep learning which is also based on machine learning requires lots of mathematical and deep learning frameworks understanding by using dependencies such as TensorFlow, OpenCV, imageai, Keras, etc, we can detect every object in the image by the area object in highlighted rectangular boxes and identify each object and assigns its tag to the object. This also includes the accuracy of each method for identifying objects. Object detection refers to the capability of computer and software systems to locate objects in an image/scene and identify each object. Getting to use modern object detection methods in applications and systems, as well as building new applications based on the straightforward task. Early implementation of object detection involved the use of classical algorithms, like the ones supported in OpenCV, the popular computer vision library. However, these classical algorithms could not achieve enough performance to work under different conditions.

We present a new dataset to advance the state-of-the-art in object recognition by placing the question of object identification. Our primary goal is to identify the object that will be upload into our web application. As by seeing the image or any object it is difficult to identify the image our system will be helping to recognize the object within the database of our image which contains thousands of images of living and nonliving objects. So it can easily identify the object.

***Key Words*:-**Anaconda, python3.6, Keras, TensorFlow, Django, Computer Vision, YOLO, SSD, Fast-RNN, Machine Learning, Deep Learning

1. INTRODUCTION-

A few years ago, the creation of the software and hardware image processing system was mainly limited to the development of the user interface, in which most of the programmers of each firm were engaged. The situation has been significantly changed with the advent of the Windows operating system when the majority of the developers switched to solving the problems of image processing itself. However, this has not yet led to the cardinal progress in solving the typical task of recognition face, car numbers, toad signs, analyzing remote and medical images, etc

Each of these "eternal" problems is solved by trial and error by the efforts of numerous groups of engineers and scientists. As modern technical solutions are turn out to be excessively expensive, the task of automating the creation of the software tools for solving intellectual problems is formulated and intensively solved abroad. In the field of image processing, the required tool kit should be supporting the analysis and recognition of images of previously unknown content and ensure the effective development of applications by ordinary programmers. Just as the Windows toolkit supports the creation of interfaces for solving various applied problems.

Object recognition is to describe a collection of related computer vision tasks that involve activities like identifying objects in digital photographs. Image classification involves activities such as predicting the class of one object in an image. Object localization refers to identifying the location of one or more objects in an image and drawing an abounding box around their extent. Object detection does work of combines these two tasks and localizes and classifies one or more objects in an image. When a user or practitioner refers to the term "object recognition", they often mean "object detection". It may be challenging for beginners to distinguish between different related computer vision tasks.

The object identification system is a web application. Object understanding involves numerous tasks including recognizing what objects are present, determining the objects'. Image classification is straightforward, but the differences between object localization and object detection can be confusing, especially when all three tasks may be just as equally referred to as object recognition.

Object recognition refers to a collection of related tasks for identifying objects in digital photographs.Hence, we have developed a web application to identify the image of the object and predicting their name or which they are known. *Image classification* involves predicting the class of one object in an image. *Object localization* refers to identifying the location of one or more objects in an image and drawing an abounding box around their extent. *Object detection* combines these two tasks and localizes and classifies one or more objects in an image.

1. 2. Motivation of the Project:

The main motivation of this project is that as our works are dependable on the machine and technology. And emerging technology is based on artificial intelligence, deep learning, and machine learning. So it is a must for the machine to identify the object or the images or the pictures on the field in which the machine is working.

On behalf of such technologies and life moving fastly towards online platforms because the pandemic the situation we are going through. The world is moving forward to going contactless and through an online platform and world has being to towards everything digit.So we are focusing on images to be identified on the online platform.

e.g. For a car to decide what to do next: accelerate, apply brakes, or turn, it needs to know where all the objects are around the car and what those objects are. Since it requires an object to be identified.

2. What is Object Detection?

Object recognition refers to a collection of related tasks for identifying objects in digital photographs.Image classification is straightforward, but the differences between object localization and object detection can be confusing, especially when all three tasks may be just as equally referred to as object recognition.

a. Image classification:

Image Classification is a fundamental task that attempts to comprehend an entire image as a whole. The goal is to classify the image by assigning it to a specific label. Typically, Image Classification refers to images in which only one object appears and is analyzed. In contrast, object detection involves both classification and localization tasks and is used to analyze more realistic cases in which multiple objects may exist in an image.

b. **Object detection** is a computer vision technique that allows us to identify and locate objects in an image or video. With this kind of identification and localization, object detection can be used to count objects in a scene and determine and track their precise locations, all while accurately labeling them.

c. Object localization:

Object localization refers to identifying the location of one or more objects in an image and drawing an abounding box around their extent. Object detection combines these two tasks and localizes and classifies one or more objects in an image.

3. Related work

Deep Networks for Object Detection. The R-CNN method trains CNNs end-to-end to classify the proposal regions into object categories or background. R-CNN mainly plays as a classifier, and it does not predict object bounds (except for refining by bounding box regression). Its accuracy depends on the performance of the regional proposal module . Several papers have proposed ways of using deep networks for predicting object bounding boxes. In the OverFeat method, a Fully-connected layer is trained to predict the box coordinates for the localization task that assumes a single object. The fully-connected layer is the turned into a convolutional layer for detecting multiple class specific objects. The MultiBox Methods generate region proposals from a network whose last fully connected layer simultaneously predicts multiple class-agnostic boxes, generalizing the "single -box" fashion of OverFeat. These class-agnostic boxes are used proposal for R-CNN. The MultiBox Proposal network is applied on a single image crop or multiple large image crops, in contrast to our fully convolutional Scheme. MultiBox does not share features between the proposal and detection network. We discuss OverFeat and MultiBox in more depteh later in context with our work, the DeepMask method is devekoped for learning segmentation proposals.

4. Methodology

YOLO (You Only Look Once)
You only look once (YOLO) at an image to predict what objects are present and where they are. YOLO is refreshingly simple: see Figure 1. A single convolutional network simultaneously predicts multiple bounding boxes and class probabilities for those boxes. YOLO trains on full images and directly optimizes detection performance. This unified model has several benefits over traditional methods of object detection.

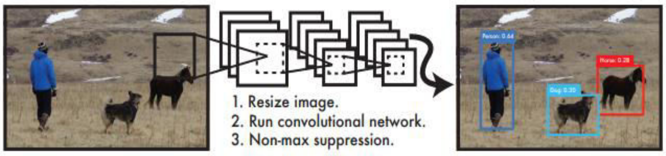


Fig -1: The YOLO Detection System.

YOLO is extremely fast. Since we frame detection as a regression problem we don't need a complex pipeline. We simply run our neural network on a new image at test time to predict detections. Our base network runs at 45 frames per second with no batch processing on a Titan X GPU and a fast version runs at more than 150 fps. This means we can process streaming video in real-time with less than 25 milliseconds of latency. Furthermore, YOLO achieves more than twice the mean average precision of other real-time systems.

YOLO reasons globally about the image when making predictions. Unlike sliding window and region proposal-based techniques, YOLO sees the entire image during training and test time so it implicitly encodes contextual information about classes as well as their appearance. Fast R-CNN, a top detection method , mistakes background patches in an image for objects because it can't see the larger context. YOLO makes less than half the number of background errors compared to Fast R-CNN.

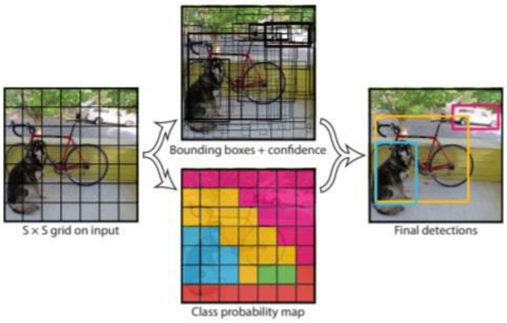
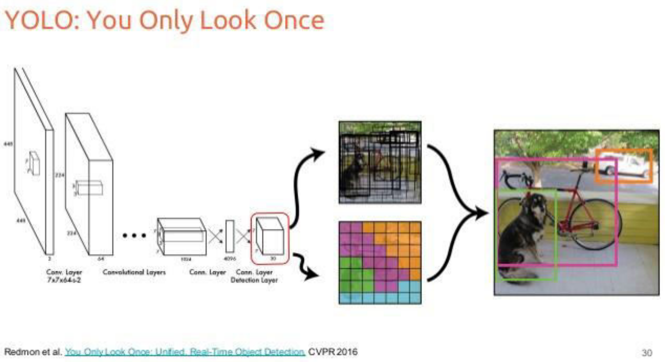


Fig -2: The Model

YOLO learns generalizable representations of objects. When trained on natural images and tested on artwork, YOLO outperforms top detection methods like DPM and R-CNN by a wide margin. Since YOLO is highly generalizable it is less likely to break down when applied to new domains or unexpected inputs. YOLO still lags behind state-of-the-art detection systems in accuracy. While it can quickly identify objects in images it struggles to precisely localize some objects, especially small ones. We examine these tradeoffs further in our experiments.

5. Architecture


YOLO: You Only Look Once

Redmon et al. You Only Look Once: Unified, Real-Time Object Detection. CVPR 2016 30

Fig -3: **YOLO Architecture**. Our System models detect as a regresion problem. It divides the image into an S*S grid and for grid cell predicts B bounding Boxes, confidence for thes box and C class Probabilities. These prediction are encoded S*S*(B*5+C) tensor

Our network uses features from the entire image to predict each bounding box. It also predicts all bounding boxes across all classes for an image simultaneously. This means our network reasons globally about the full image and all the objects in the image. The YOLO design enables end-to-end training and realtime speeds while maintaining high average precision.

Our system divides the input image into an S × S grid. If the center of an object falls into a grid cell, that grid cell is responsible for detecting that object.

Each grid cell predicts B bounding boxes and confidence scores for those boxes. These confidence scores reflect how confident the model is that the box contains an object and also how accurate it thinks the box is that it predicts. Formally we define confidence as Pr(Object) ∗ IOU . If no object exists in that cell, the confidence scores should be zero. Otherwise we want the confidence score to equal the intersection over union (IOU) between the predicted box and the ground truth

Each bounding box consists of 5 predictions: x, y, w, h, and confidence. The (x, y) coordinates represent the center of the box relative to the bounds of the grid cell. The width and height are predicted relative to the whole image. Finally the confidence prediction represents the IOU between the predicted box and any ground truth box. Each grid cell also predicts C conditional class probabilities, Pr(Classi |Object). These probabilities are conditioned on the grid cell containing an object. We only predict one set of class probabilities per grid cell, regardless of the number of boxes B.

At test time we multiply the conditional class probabilities and the individual box confidence predictions,which gives us class-specific confidence scores for each box. These scores encode both the probability of that class appearing in the box and how well the predicted box fits the object

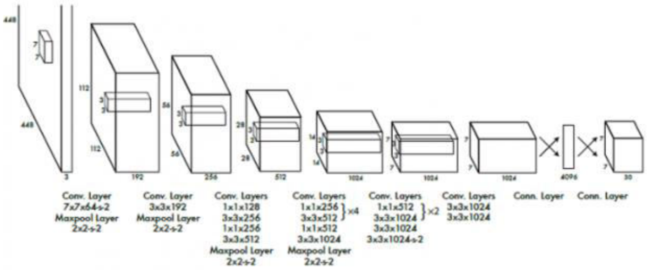


Fig -4: **The Architecture.** Our detection network has 24 convolution layers followed by 2 fully connected layers. Alternating 1*1 convolutional layer reduce the features space from preceding layers. We pretrain the convultional layers on the ImageNet Classification task at half the resolution(224*224 input image) and the double the resolution for detection.

A. Working

- ◆ Visit our Websites on any browser
- ◆ After visiting, Upload an image .jpg in the input field for image uploading
- ◆ Next, Submit the image by clicking on the Submit Button
- ◆ After Submission YOLO will work on the image as YOLO Algorithm is used for object detection and object predection.
- ◆ After this, The Desired result will be regerented

B. Implemention

a. Dependencies

1. Tensorflow (GPU version preferred for Deep Learning)
2. NumPy (for Numeric Computation)
3. Pillow/PIL (for Image Processing)
4. IPython (for displaying images in Jupyter Notebook)

5. Glob (for finding pathname of all the files)
6. Django(Framework of Webapplication)

b. Batch Normalization

It is a preprocessing step of features extracted from previous layers, before feeding it to the next layers of the network. We normalize the input layer by adjusting and scaling the activations. For example, when we have one feature in range 0 to 1 and other from 1 to 1000, we should normalize them to speed up learning. So, the Neural network does not assume the feature with a range from 1 to 1000 as a high priority in the features dependencies. This allows each layer of a network to learn by itself a little bit more independently of other layers. Almost every convolutional layer in Yolo has batch normalization after it. It helps the model train faster and reduces variance between units (and total variance as well).

```
Input: Values of $x$ over a mini-batch: $\mathcal{B} = \{x_{1...m}\}$;
       Parameters to be learned: $\gamma, \beta$
Output: $\{y_i = \mathrm{BN}_{\gamma,\beta}(x_i)\}$

    $\mu_{\mathcal{B}} \leftarrow \frac{1}{m}\sum_{i=1}^{m} x_i$                          // mini-batch mean
    $\sigma_{\mathcal{B}}^2 \leftarrow \frac{1}{m}\sum_{i=1}^{m}(x_i - \mu_{\mathcal{B}})^2$    // mini-batch variance
    $\widehat{x}_i \leftarrow \frac{x_i - \mu_{\mathcal{B}}}{\sqrt{\sigma_{\mathcal{B}}^2 + \epsilon}}$                  // normalize
    $y_i \leftarrow \gamma\widehat{x}_i + \beta \equiv \mathrm{BN}_{\gamma,\beta}(x_i)$          // scale and shift
```

Algorithm 1: Batch Normalizing Transform, applied to activation x over a mini-batch.

C. Data Flow Chart

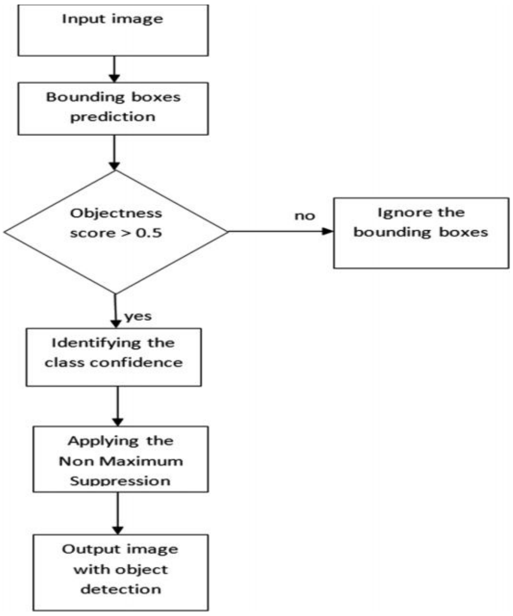
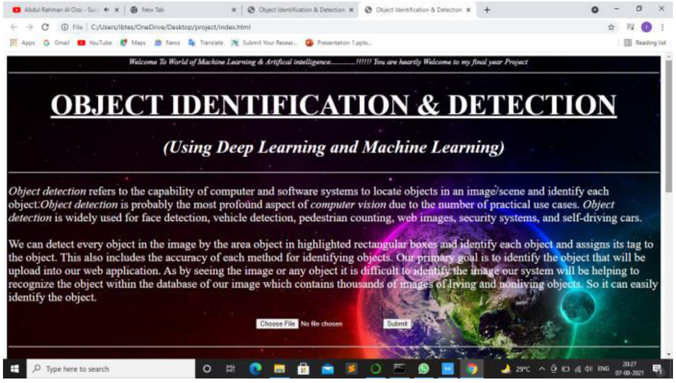

Input image
Bounding boxes prediction
Objectness score > 0.5
no
Ignore the bounding boxes
yes
Identifying the class confidence
Applying the Non Maximum Suppression
Output image with object detection

Fig -5. : Data flow

6. Result


Abdul Rahman Al Ossi - Su
New Tab
Object Identification & Detection
Object Identification & Detection
File | C:/Users/ibtes/OneDrive/Desktop/project/index.html
Apps
Gmail
YouTube
Maps
News
Translate
Submit Your Resear...
Presentation 1.pptx...
Reading list
Type here to search
29°C
ENG
20:27
07-08-2021

Welcome To World of Machine Learning & Artifical intelligence............!!!!!! You are heartly Welcome to my final year Project

OBJECT IDENTIFICATION & DETECTION

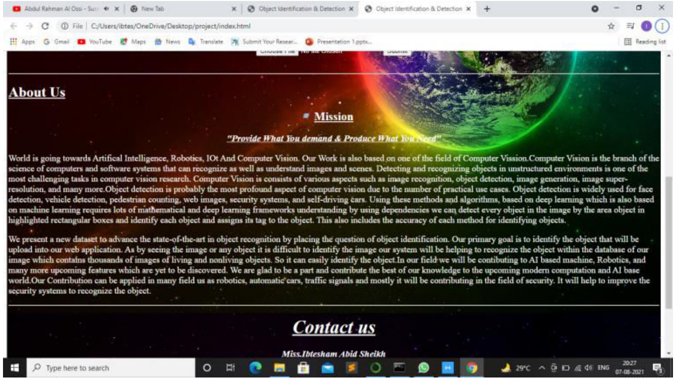
(Using Deep Learning and Machine Learning)

Object detection refers to the capability of computer and software systems to locate objects in an image/scene and identify each object.*Object detection* is probably the most profound aspect of *computer vision* due to the number of practical use cases. *Object detection* is widely used for face detection, vehicle detection, pedestrian counting, web images, security systems, and self-driving cars.

We can detect every object in the image by the area object in highlighted rectangular boxes and identify each object and assigns its tag to the object. This also includes the accuracy of each method for identifying objects. Our primary goal is to identify the object that will be upload into our web application. As by seeing the image or any object it is difficult to identify the image our system will be helping to recognize the object within the database of our image which contains thousands of images of living and nonliving objects. So it can easily identify the object.

Choose File No file chosen Submit

Fig -6.a. : Web Application Screenshot


Abdul Rahman Al Ossi - Su
New Tab
Object Identification & Detection
Object Identification & Detection
File | C:/Users/ibtes/OneDrive/Desktop/project/index.html
Apps
Gmail
YouTube
Maps
News
Translate
Submit Your Resear...
Presentation 1.pptx...
Reading list
Type here to search
29°C
ENG
20:27
07-08-2021

About Us

- Mission

"Provide What You demand & Produce What You Need"

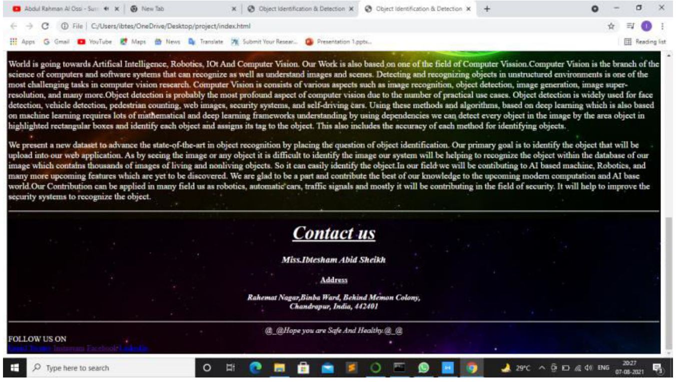
World is going towards Artifical Intelligence, Robotics, IOt And Computer Vision. Our Work is also based on one of the field of Computer Vission.Computer Vision is the branch of the science of computers and software systems that can recognize as well as understand images and scenes. Detecting and recognizing objects in unstructured environments is one of the most challenging tasks in computer vision research. Computer Vision is consists of various aspects such as image recognition, object detection, image generation, image super-resolution, and many more.Object detection is probably the most profound aspect of computer vision due to the number of practical use cases. Object detection is widely used for face detection, vehicle detection, pedestrian counting, web images, security systems, and self-driving cars. Using these methods and algorithms, based on deep learning which is also based on machine learning requires lots of mathematical and deep learning frameworks understanding by using dependencies we can detect every object in the image by the area object in highlighted rectangular boxes and identify each object and assigns its tag to the object. This also includes the accuracy of each method for identifying objects.

We present a new dataset to advance the state-of-the-art in object recognition by placing the question of object identification. Our primary goal is to identify the object that will be upload into our web application. As by seeing the image or any object it is difficult to identify the image our system will be helping to recognize the object within the database of our image which contains thousands of images of living and nonliving objects. So it can easily identify the object.In our field we will be contibuting to AI based machine, Robotics, and many more upcoming features which are yet to be discovered. We are glad to be a part and contribute the best of our knowledge to the upcoming modern computation and AI base world.Our Contribution can be applied in many field us as robotics, automatic cars, traffic signals and mostly it will be contributing in the field of security. It will help to improve the security systems to recognize the object.

Contact us

Miss.Ibtesham Abid Sheikh

Fig -6.b. : Web Application Screenshot


Abdul Rahman Al Ossi - Su
New Tab
Object Identification & Detection
Object Identification & Detection
File | C:/Users/ibtes/OneDrive/Desktop/project/index.html
Apps
Gmail
YouTube
Maps
News
Translate
Submit Your Resear...
Presentation 1.pptx...
Reading list
Type here to search
29°C
ENG
20:27
07-08-2021

World is going towards Artifical Intelligence, Robotics, IOt And Computer Vision. Our Work is also based on one of the field of Computer Vission.Computer Vision is the branch of the science of computers and software systems that can recognize as well as understand images and scenes. Detecting and recognizing objects in unstructured environments is one of the most challenging tasks in computer vision research. Computer Vision is consists of various aspects such as image recognition, object detection, image generation, image super-resolution, and many more.Object detection is probably the most profound aspect of computer vision due to the number of practical use cases. Object detection is widely used for face detection, vehicle detection, pedestrian counting, web images, security systems, and self-driving cars. Using these methods and algorithms, based on deep learning which is also based on machine learning requires lots of mathematical and deep learning frameworks understanding by using dependencies we can detect every object in the image by the area object in highlighted rectangular boxes and identify each object and assigns its tag to the object. This also includes the accuracy of each method for identifying objects.

We present a new dataset to advance the state-of-the-art in object recognition by placing the question of object identification. Our primary goal is to identify the object that will be upload into our web application. As by seeing the image or any object it is difficult to identify the image our system will be helping to recognize the object within the database of our image which contains thousands of images of living and nonliving objects. So it can easily identify the object.In our field we will be contibuting to AI based machine, Robotics, and many more upcoming features which are yet to be discovered. We are glad to be a part and contribute the best of our knowledge to the upcoming modern computation and AI base world.Our Contribution can be applied in many field us as robotics, automatic cars, traffic signals and mostly it will be contributing in the field of security. It will help to improve the security systems to recognize the object.

Contact us

Miss.Ibtesham Abid Sheikh

Address

Rahemat Nagar,Binba Ward, Behind Memon Colony,
Chandrapur, India, 442401

@_@Hope you are Safe And Healthy.@_@

FOLLOW US ON
Email Twitter Instagram Facebook LinkedIn

Fig -6.c. : Web Application Screenshot

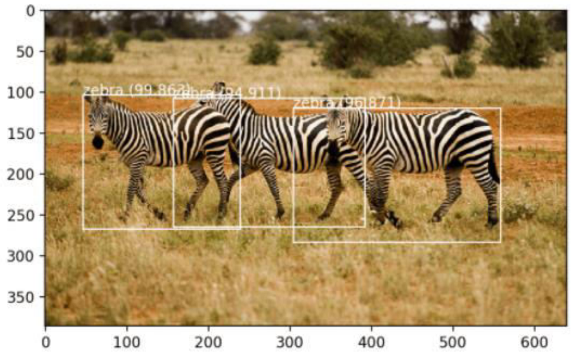


Fig -6.d. : Output

7. Conclusion

We introduce a Web Application for Object Identification and Detection using Deep learning and machine learning, YOLO model for object detection. Our Web is simple and easy to use as its construction is simple and can detect the multiple objects in an image and detection performance is fast than other model.

YOLO is the fastest general purpose object detector in the literature and YOLO pushes the state of the art in multiple object detection. YOLO also generalizes well for the Web Application that rely on fast, robust object detection.

Acknowledgement

I felt great pleasure in submitting this paper on "Object Identification and Detection using deep learning and machine learning" First and foremost, We express our deep dense of gratitude, sincere thanks to associate prof. Hirendra R. Hajare for the best support opinion, view , comments and thoughts have been really helped me.

REFERENCES

1. *P. F. Felzenszwalb, R. B. Girshick, D. Mcallester, and D. Ramanan, "Object detection with discriminatively trained part-based models," IEEE Trans. Pattern Anal. Mach. Intell., vol. 32, no. 9, p. 1627, 2010.*

2. *.Agarwal, S., Awan, A., and Roth, D. (2004). Learning to detect objects in images via a sparse, part-based representation. IEEE Trans. Pattern Anal. Mach. Intell. 26,1475–1490. doi:10.1109/TPAMI.2004.108*

3. *Alexe, B., Deselaers, T., and Ferrari, V. (2010). "What is an object?," in ComputerVision and Pattern Recognition (CVPR), 2010 IEEE Conference on (San Francisco,CA: IEEE), 73–80. doi:10.1109/CVPR.2010.5540226*

4. *Aloimonos, J., Weiss, I., and Bandyopadhyay, A. (1988). Active vision. Int. J.Comput. Vis. 1, 333–356. doi:10.1007/BF00133571*

5. *Andreopoulos, A., and Tsotsos, J. K. (2013). 50 years of object recognition: direc-tions forward. Comput. Vis. Image Underst. 117, 827–891. doi:10.1016/j.cviu.2013.04.005*

6. D. Erhan, C. Szegedy, A. Toshev, and D. Anguelov. Scalable object detection using deep neural networks. In Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on, pages 2155–2162. IEEE, 2014.O.

7. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei. ImageNet Large Scale Visual Recognition Challenge. International Journal of Computer Vision (IJCV), 2015. 3

8. P. Sermanet, D. Eigen, X. Zhang, M. Mathieu, R. Fergus, and Y. LeCun. Overfeat: Integrated recognition, localization and detection using convolutional networks. CoRR, abs/1312.6229, 2013

9. J. R. Uijlings, K. E. van de Sande, T. Gevers, and A. W. Smeulders. Selective search for object recognition. International journal of computer vision,

10. *Bengio, Y. (2012). "Deep learning of representations for unsupervised and transferlearning," in ICML Unsupervised and Transfer Learning, Volume 27 of JMLRProceedings, eds I. Guyon, G. Dror, V. Lemaire, G. W. Taylor, and D. L. Silver(Bellevue: JMLR.Org), 17–36.*

11. *Cadena, C., Dick, A., and Reid, I. (2015). "A fast, modular scene understanding sys-tem using context-aware object detection," in Robotics and Automation (ICRA),2015 IEEE International Conference on (Seattle, WA).*

12. *Erhan, D., Szegedy, C., Toshev, A., and Anguelov, D. (2014). "Scalable object detec-tion using deep neural networks," in Computer Vision and Pattern Recognition Frontiers in Robotics and AI www.frontiersin.org November 2015*

13. *Dalal, N., and Triggs, B. (2005). "Histograms of oriented gradients for humandetection," in Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEEComputer Society Conference on, Vol. 1 (San Diego, CA: IEEE), 886–893. doi:10.1109/CVPR.2005.177*