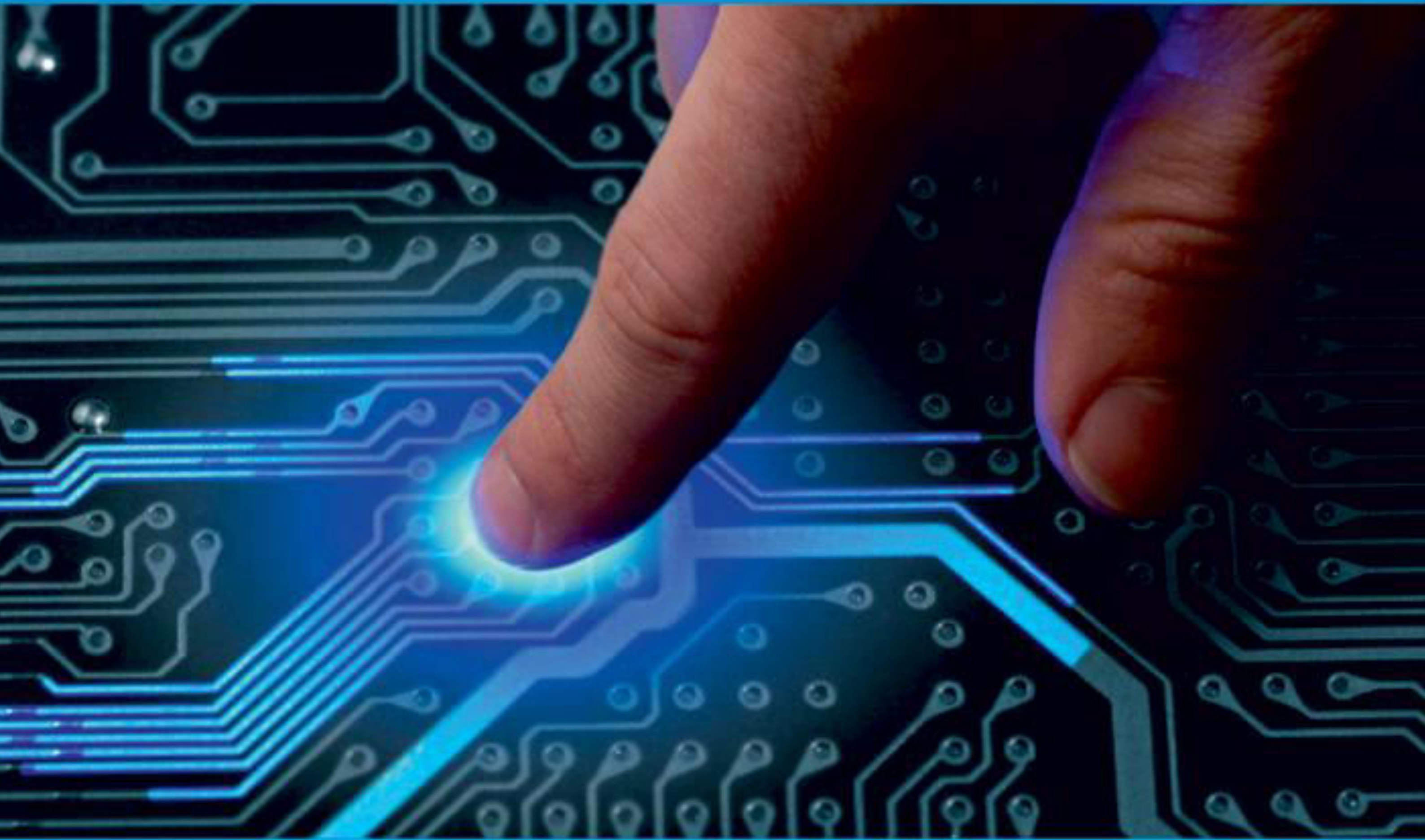
IJIRCCE

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



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 9, Issue 6, June 2021

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.542

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542 |

|| Volume 9, Issue 6, June 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0906284 |

Object Identification Using Convolution Neural Networks

Ms M Varalakshmi[1], G Naga Bala Mounika[2], G Pravallika[3], K Prameela Bhargavi [4]

Assistant Professor, Dept of Computer Science Engineering, Vasireddy Venkatadri Institute of Technology, Nambur, Guntur, Andhra Pradesh, India. [1]

UG Students, Dept of Computer Science Engineering, Vasireddy Venkatadri Institute of Technology, Nambur, Guntur, Andhra Pradesh, India. [2,3,4]

ABSTRACT: Every Object in this world is identified by its State, Behavior and its Identity. And objects are classified based on various factors like living and non-living things. A computer can also identify an object by using advanced technologies like Deep Learning. Image classification involves assigning a class label to an image, whereas object localization involves drawing a bounding box around one or more objects in an image. Object detection is more challenging and combines these two tasks and draws a bounding box around each object of interest in the image and assigns them a class label. Together, all of these problems are referred to as object recognition.

The objective of our proposed project is an extension of one of the many base papers which we referred in which identifies the object using Deep Learning. We extend this project to identify objects along with its name in audio format using Conventional Neural Networks (CNN) algorithm. Computer systems with this capability have basic and applied research areas, including security, psychiatry, education, robotics, etc.

I. INTRODUCTION

Convolutional neural networks (CNNs) have been widely used in visual recognition from due to its high capability in correctly classifying images. And CNNs become the foremost preferable choice for solving image classification challenges. Besides image classification, researchers have extend the application of CNNs to several other tasks in visual recognition such as localization, segmentation, generating sentences from image as well as object detection.In our project, we mainly specialize in the task of object detection which has tremendous application in our lifestyle. The goal of object detection is recognize objects during a single image, not only to return the arrogance of the category for every object, but also predict the corresponding bounding boxes. Among most of the works in object detection, region CNNs (CNN) is the most remarkable one that combines selective search , CNNs, support vector machines (SVM) and bounding box regression together to supply a high performance in object detection. In this paper, we'll provide an alternate approach of object detection by reducing the complexity of the CNN. First, we adopt edge box, a recent published algorithm to get region proposals, rather than selective search utilized in CNN albeit the mean average precision between edge boxes and selective search are almost an equivalent, edge boxes runs much faster than selective search. Secondly, we remove the entire category specific SVMs, and directly use the output of softmax within the last layer of CNN as our score.

II. LITERATURE REVIEW

Writing review is the main advance in programming improvement measure. Previously fostering the instrument it is important to decide the time factor, economy. When these things are fulfilled, then, at that point subsequent stages are to figure out which working framework and language can be utilized for fostering the instrument. When the developers start building the apparatus the software engineers need part of outer help. This help can be gotten from senior developers, from book or from sites. Prior to building the framework the above thought are considered for fostering the proposed framework.

W. Ouyang, X. Wang, X. Zeng, et al, “Deepid-net: Deformable deep convolutional neural networks for object detection,” 2015 IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 2403-2412.

The ImageNet picture arrangement and limitation dataset with 1,000 classes is picked to pre train the profound model. Its item recognition dataset has 200 article classes. In the exploratory area, the methodology is additionally applied to the PASCAL VOC. The pre training information keeps something similar, while the identification dataset just has 20

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542 |

|| Volume 9, Issue 6, June 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0906284 |

article classes. The pre training information keeps something very similar, while the identification dataset just has 20 item classes. The ImageNet Large Scale Visual Recognition Challenge (ILSVRC) 2014 contains two diverse datasets: 1) the grouping and confinement (Cls-Loc) dataset and 2) the identification (Det) dataset. The preparation information of Cls-Loc contains 1.2 million pictures with names of 1,000 classifications. It is utilized to pre train profound models. A similar split of train and approval information from the Cls-Loc is utilized for picture level comment and item level comment pre training. The Det contains 200 item classes and is parted into three subsets, train, approval (val), and test information.

III. IMPLEMENTATION

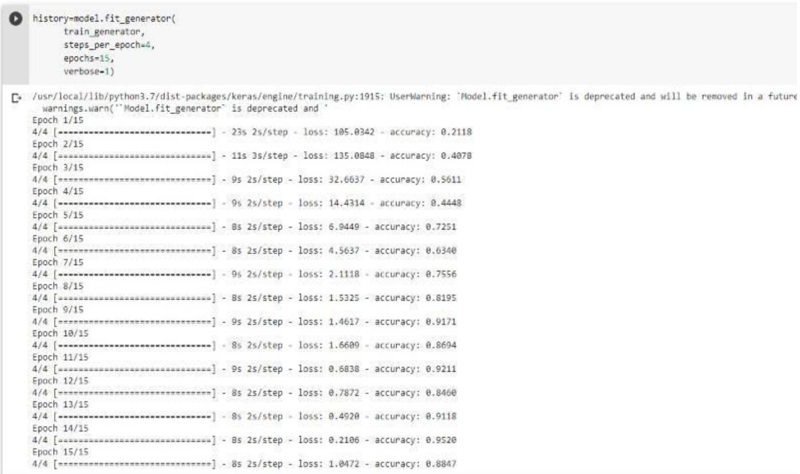
The obvious end goal of creating a Convolutional Neural Network is to identify the images that are given as input to the model. With the goal in mind, knowledge about the steps in dataset preparation for training the selected algorithm to perform desired detections is obtained.

Input layer is two-dimensional matrix. It's composed of image pixels. The model presented during this paper used four convolution layers. The grey scale image with 300 x 300 pixel matrix is that the input to the present layer. Every feature map is connected to its previous map. In every layer there are several feature maps. The convolution layer C1 uses 32 convolutional nuclei. The dimensions of convolution nuclei are 5x5. Layer C2 and C3 uses 128 nuclei.

Softmax layer contains 10 neurons. The feature of the output layer is assessed among the seven emotions. For visual perception several layers are involved within the process. The layers are Convolutional, Pooling, Fully Connected and Softmax Layers.

The number of epochs may be a hyper parameter that defines the amount times that the training algorithm will run through the whole training dataset. One epoch means each sample within the training dataset has had a chance to update the interior model parameters. An epoch is comprised of 1 or more batches. For instance, as above, an epoch that has one batch is named the batch gradient descent learning algorithm.

```
history=model.fit_generator(train_generator,
                            steps_per_epoch=7,
                            epochs=15,
                            verbose=1)
```



```
history=model.fit_generator(
      train_generator,
      steps_per_epoch=4,
      epochs=15,
      verbose=1)

/usr/local/lib/python3.7/dist-packages/keras/engine/training.py:1915: UserWarning: `Model.fit_generator` is deprecated and will be removed in a future
  warnings.warn('`Model.fit_generator` is deprecated and '
Epoch 1/15
4/4 [==============================] - 23s 2s/step - loss: 105.0342 - accuracy: 0.2118
Epoch 2/15
4/4 [==============================] - 11s 3s/step - loss: 135.0848 - accuracy: 0.4078
Epoch 3/15
4/4 [==============================] - 9s 2s/step - loss: 32.6637 - accuracy: 0.5611
Epoch 4/15
4/4 [==============================] - 9s 2s/step - loss: 14.4314 - accuracy: 0.4448
Epoch 5/15
4/4 [==============================] - 8s 2s/step - loss: 6.9449 - accuracy: 0.7251
Epoch 6/15
4/4 [==============================] - 8s 2s/step - loss: 4.5637 - accuracy: 0.6340
Epoch 7/15
4/4 [==============================] - 9s 2s/step - loss: 2.1118 - accuracy: 0.7556
Epoch 8/15
4/4 [==============================] - 8s 2s/step - loss: 1.5325 - accuracy: 0.8195
Epoch 9/15
4/4 [==============================] - 9s 2s/step - loss: 1.4617 - accuracy: 0.9171
Epoch 10/15
4/4 [==============================] - 8s 2s/step - loss: 1.6609 - accuracy: 0.8694
Epoch 11/15
4/4 [==============================] - 9s 2s/step - loss: 0.6838 - accuracy: 0.9211
Epoch 12/15
4/4 [==============================] - 8s 2s/step - loss: 0.7872 - accuracy: 0.8460
Epoch 13/15
4/4 [==============================] - 8s 2s/step - loss: 0.4920 - accuracy: 0.9118
Epoch 14/15
4/4 [==============================] - 8s 2s/step - loss: 0.2106 - accuracy: 0.9520
Epoch 15/15
4/4 [==============================] - 8s 2s/step - loss: 1.0472 - accuracy: 0.8847
```

Fig 1: *Number of steps for accuracy and loss*

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542 |

|| Volume 9, Issue 6, June 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0906284 |

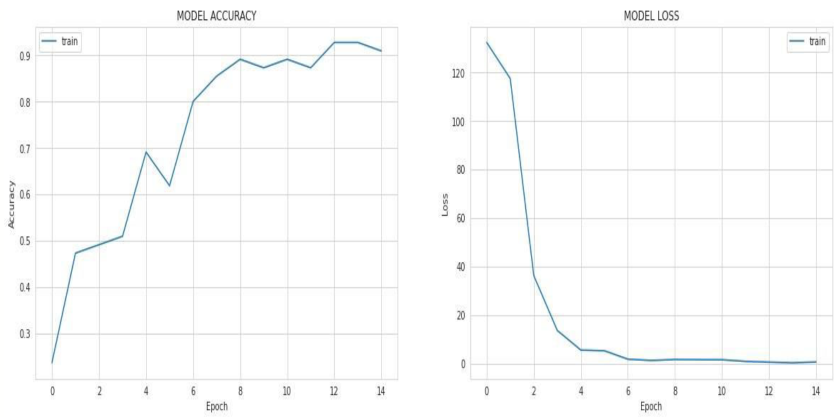


Fig 2: *Model accuracy and Model loss*

Images are present in Image_class folder. It contains train and test folders containing which successively contains corresponding images.
Then image is shipped as an input to the predict function. The max prediction keys returned. The worth of the returned key value is retrieved and is displayed. Like this for each cropped object within the image this process is completed and eventually the first image with the highlighted and object type is displayed, and also represented within the speech format.

IV. RESULT

The raw data as single images are collected as input from computer and internet. The model specifies the object's name. Upon tuning the parameters, the accuracy of 80 percentages is achieved. The model developed takes the image as input and detects the object in the given image and produces output image with its name and characteristics.
The design of input focuses on controlling the amount of input required, controlling the errors, avoiding delay, avoiding extra steps and keeping the process simple.Firstly, when the model successfully runs the input image is chosen from the computer.And then the model presents the predicted value.Thus by the predicting value of the model, the object class is chosen and the predicted value is displayed in the output along with the object's name and characteristics.After identifying the object in the picture which is represented in text format is converted in to the audio format is as shown in the below picture.

Fig 3: *Object identified as Chair*

Now, the input for the model is given with another new image as shown below, the image given as input is a shoe which is a foot wear, thus the model identifies the image as foot wear.

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542 |

|| Volume 9, Issue 6, June 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0906284 |

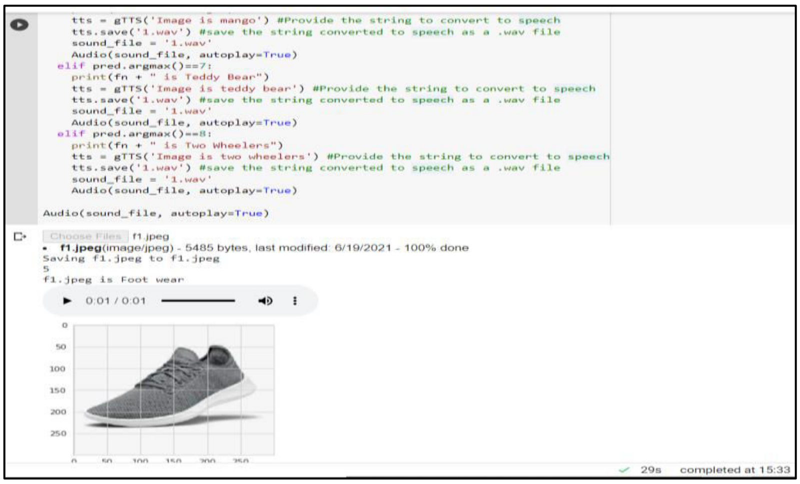


Fig 4: Object identified as Foot Wear

Here we inserted the image mango, and the model successfully identified it as mango also generates an audio that the object is a mango.

Fig 5: Object identified as Mango

V. CONCLUSION AND FUTURE SCOPE

Object Identification may be a key ability for many computer and robot vision system. Although great progress has been observed within the last years, and a few existing techniques are now a part of many consumer electronics or are integrated in assistant driving technologies, we are still away from achieving human-level performance, especially in terms of open-world learning. It should be noted that object Identification has not been used much in many areas where it might be of great help. The system proposed during this paper identifies the thing within the given images using Convolution Neural Networks i.e. Deep learning. The thing identification is combination of object localization and object classification CNN detects the thing within the model as we trained the model with great deal of knowledge.

The proposed system identifies several objects like dog, cat, chair, Beam bag, Bat, Bag. The info with images of even size collected is fed to model. Compared to past literature this model is in a position to acknowledge objects accurately and this model is strong to noise. We achieved a performance of identifying objects with 90 percent and above accuracy. The input images are present within the Trained Database. After the thing s are recognized from a picture the trail is about to trained data sets that are wont to find the object during this system.

Future scope of object identification is one of the main technologies that skyrockets the development of self-driving cars. Object identification powered by innovative deep learning and machine learning has already been embedded in a

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542 |

|| Volume 9, Issue 6, June 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0906284 |

number of fields with impressive success. It is used for automated image organization of huge databases and visual websites. Object recognition makes image classification for stock websites easier and even fuels marketers' creativity by enabling them to craft interactive brand campaigns.

REFERENCES

1. https://www.cv-foundation.org/openaccess/content_cvpr_2015/papers/Ouyang_DeepID-Net_Deformable_Deep_2015_CVPR_paper.pdf
2. https://www.cv-foundation.org/openaccess/content_cvpr_2014/papers/Erhan_Scalable_Object_Detection_2014_CVPR_paper.pdf
3. https://www.cs.toronto.edu/~kriz/imagenet_classification_with_deep_convolutional.pdf
4. https://www.semanticscholar.org/paper/Fast-kNN-classification-algorithm-based-on-partial-Hwang-Wen/d7def5e57a4d2e556e1283c92c7857a8ca164969#related-papers

INNO SPACE
SJIF Scientific Journal Impact Factor
Impact Factor: 7.542

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING