



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

IJRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 12, Issue 10, October 2023

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.423

9940 572 462

6381 907 438

ijirset@gmail.com

www.ijirset.com

Character Recognition on Forged Entities using AI

Alisha Inamdar, Ishwari Shendkar

#Computer Department, Jayawantrao Sawant Polytechnic University, Pune, Maharashtra, India

Jayawantrao Sawant Polytechnic Hadapsar, Pune, Maharashtra, India

ABSTRACT: The objective of a general Optical character recognition is a technology which is used to convert handwritten text ,text on images into editable , machine readable and electronic text . The aim of this paper is to scan and detect labels which are engraved on the different parts of machine or metals. Additionally these recognized labels are stored in the database in excel format with date,time and Label accuracy .The methodology which is used is YOLOV4 algorithm. This algorithm is used for recognition and detection. Darknet is used as the backbone in this methodology.

KEYWORDS: Darknet, Optical Character Recognition (OCR), You Only Look Once (YOLO).

I. INTRODUCTION

The character recognition works on basis of allowing computer to understand the printed or handwritten text. The implementation is carried out in three step process which will be discussed later in this paper.The previously implemented algorithms are Easy OCR and Keras OCR but these OCR techniques are inefficient in detecting the labels and the performance is quite poor. Because of this You Only Look Once(YOLOV4) was taken into consideration as this is used for objection detection tasks and reduces this problem.Darknet is considered as backbone of YOLOV4 model ,it is a type of technique used for feature extraction of convolution neural network. YOLOV4 has two types of pipeline which are one stage detector and two stage detector. The pipeline used is one stage detector where the objects are classified and detected at same time.

II. LITERATURE SURVEY

In today's era all the important paperwork processes are automated. For example the user accounts in the banks are created digitally by using identification documents such as PAN card ,Aadhar card and other important identification documents. Once the account is created to verify the user the OCR technology is implemented. These documents are passed to OCR system and required details such as name,birth date ,ID number are extracted and these help in verification of user. The OCR technology is also used in various sectors such as passport verification, in banking sector for verifying sign on the cheque , in recognition of number plates of vehicle .

In this system the accuracy level is 85.3%. Prior to implementing the YOLOV4 model the implementations which were taken into consideration and were implemented are Keras OCR and Easy OCR. But while working with them we faced major difficulties. These implementations had a high level accuracy only while working with documents and the dataset which we wanted to be recognized was on forged entities that is we wanted OCR to recognize the text which is engraved or carved.When implemented these traditional OCR techniques resulted in accuracy level less than 15%.Therefore after a survey the methodology we decided to implement was YOLOV4. This technique had a average precision of 69.5% then the other prior implemented techniques.

III. OPTICAL CHARACTER RECOGNITION

A OCR in general is used in converting the text related content. It detects document which can be in image format and convert it into electronic text. Furthermore it is also used to eliminate human interference and human error.The OCR detection is used for verifying multiple languages.

A. Architecture of OCR

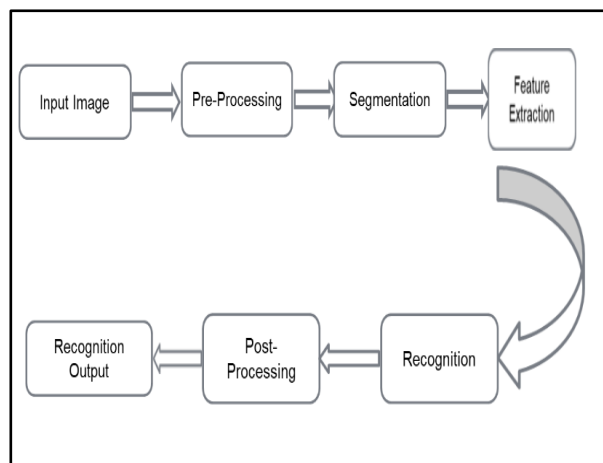
In this first the input image is taken from user which can be a document in png,jpg,txt format . Next the image is preprocessed i.e to make as easy as possible the the OCR model as to easily recognize the text only distinguishing from the background. Preprocessing is a preliminary or initial step which easily transforms data so that it could be effectively and efficiently processed further.

Next step is segmentation which includes dividing the dividing the image ,document into smaller partitions or segments . The main aim of segmentation is to simplify and transform the document or image into computer readable meaningful and thus easier to analyze.

Further step is feature extraction which involves converting the different documents into editable and searchable format. It involves extracting only relevant information from the provided in input image and recognize and interpret the text.

After feature extraction, character recognition is done which which is the key component of OCR.It includes classifying and identification of each character from the previously executed process which involves feature extraction.

Then later image post processing is done which improves character recognition accuracy.It checks grammatical errors. The architecture of OCR can be modified based goals and requirement of application.



IV. YOLOv4

A. Architecture of YOLOv4

In the One Stage Detector the four components are present:-

- Input
- Backbone
- Neck

The YOLOv4 is the successor of the previous version that is starting from YOLOv1. There were some issues in the YOLOv1 network which would take the same error for bigger and smaller objects which makes the model's prediction for the neighboring objects unsatisfactory. So, loss function improvement was done. Inception structure was being improved so that multi scale features can be more effectively extracted and the hidden information can be used more efficiently [1]. Spatial Pyramid Pooling (SPP) layer was being improved, as input images of any size can be given to the model and output images will be of fixed size and it can extract features at various scales [1]. These were the changes made to the YOLOv1 network and were modified and now YOLOv4 is the model that being used for the object detection tasks. Fig. 1 shows the four components used for the implementation [2-3]. Input images are given to the backbone and there are multiple backbones to choose from, namely VGG16, RESNET50, SpinNET, EfficientNet-B3 ,CSPResNext50, CSPDarkNet53. Model Backbone does the work of extracting important features

from the given input image using Cross Stage Partial Network (CSPNet). The Neck is used for generating feature pyramids that helps to identify the same object with different sizes and scales using Path Aggregation Network (PANet).

The Neck consists of Spatial Pyramid Pooling (SPP), Atrous Spatial Pyramid Pooling (ASPP), Receptive Field Block (RFB), Spatial Attention Module (SAM). The Path Aggregation Block has Feature Pyramid Network (FPN), Path Aggregation Network (PAN), Neural Architecture Search (NAS)-FPN, Fully Connected FPN, BiFPN, Adaptively Spatial Feature Fusion (ASFF) and Scale-wise Feature Aggregation Module (SFAM). The head block has the Dense Prediction Layer (DPL) and the Sparse Prediction Layer (SPL). Head is used for applying anchor boxes on features and generates final output vectors with class probabilities, objectiveness scores, and bounding boxes (similar to yolov3). In the DPL there are various models such as Region Proposal Network (RPN), Single Shot Detector (SSD), YOLO, RetinaNet, CornerNet, CenterNet, MatrixNet, FCOS whereas for the SPL there are Faster Region Based Convolutional Neural Network (RCNN), Region based Fully Convolutional Neural Network (RFCN), Mask RCNN. Leaky Rectified Linear Unit (ReLU) activation function is used in middle/hidden layers and sigmoid activation

function is used in the final detection layer. The PANet uses features from all layers and lets the network decide which parameters are useful and performs Region Of Interest align operation on each feature map to extract features for the object. In the Two Stage Detector the additional components after DPL is the SPL, which in total becomes five components. **The backbone** network of an object detector is mostly pre-trained on ImageNet classification. The main and the most used three backbones for YOLOv4 object detector is as follows:-

- CSPDarkNet53
- CSPResNext50
- EfficientNet-B3

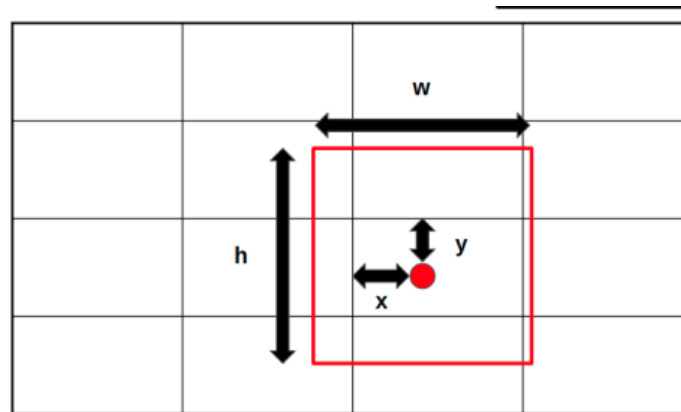
The CSPDarkNet53 and CSPResNext50 are based on DenseNet was developed to link it with the CNN layers as it reduces the vanishing gradient problem, supports feature propagation, encourages network to reuse features and reduces the number of network parameters. The idea of CSPDarkNet53 and CSPResNext50 are to remove the computational bottlenecks in the DenseNet and improve the learning by using an unedited version of feature map. **The 'backbone'** is responsible for feature formation, **'neck'** for feature aggregation means concatenating all the features extracted from the backbone and the last is the **'head'** where detection happens i.e. the DPL. There are some more techniques that are embedded into the yolov4 algorithm are used to improve accuracy during training and afterward. Bag of freebies and Bag of specials are added in to the YOLO Algorithm [3]. These bag of freebies help in training the model without interfering into the inference time. The bag of freebies refers to two techniques, one is the Bag of freebies for the backbone, which refers the cut mix and mosaic for data augmentation techniques and drop block for regularization, and second is bag of freebies for detection, which adds to the backbone, such as self-adversarial training, random training shapes and the rest. The bag of specials makes changes into the architecture and rises the inference time. The bag of specials also refers to two techniques, the first is the bag of specials for the backbone which has the mish activation function, cross stage partial connections. The second is the bag of specials for detection, which uses the SPP-block, the SAM block, and others. The SPP block is important as input of any resolution can be given but the output from the model will be of fixed size.

B. Components of Loss Function

Classification Loss refers to finding the object or class in the image for example Football, Face and Dog Localization Error refers to finding the exact location of the bounding box where the object is present with the parameters x, y, w, h.

C. Detection is YOLOV4

4x4 grid on an image Vectors for each
grid = Pc x y w h C1 C2

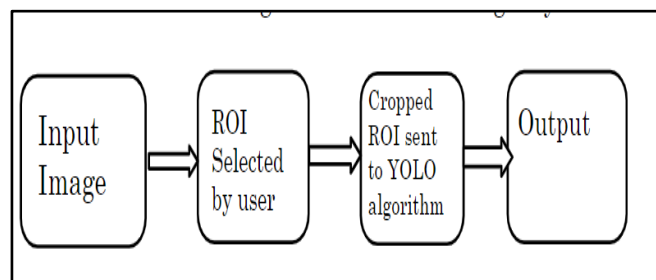


V. PROPOSED METHODOLOGY

In this application the important algorithm used is YOLOv4 algorithm (you only look once version 4). This algorithm is used to predict the labels. The methodology of system is that it works in stages. In first stage the input image is taken through dataset, then the area where the label is to be selected i.e the ROI (Region of interest), once it is selected the image is cropped and then using YOLO algorithm labels are predicted and stored in excel format with time and accuracy of recognized label.

The YOLOv4 algorithm works very well in terms of video feed. This implementation has the YOLOv4 algorithm as the backend for predicting the labels on the metal parts. Here the weights are being passed on the images so as to predict the labels on the material or the metal parts. Input of images, Region Of Interest Selector, Once Region Of Interest (ROI) is selected, Image is cropped and passed through YOLO, Output predicted labels are stored onto csv and excel format. The image input can be taken on an User Interface (UI) using pysimple GUI. This UI has some basic parameters to be set for the YOLO algorithm to perform the operations. Firstly the input images on which all these operations will take place. Second, config file where the layers and activation function is defined and also we have the weights and classes file to be given. All these things can be kept default or can be changed based on the trained dataset.

The application works in following way:-





a. Input image is taken from user



b. ROI is selected and image is cropped and predictions are done

Date	Label Detected	Detection Confidence	Order Time
2023-10-27 10:30:00	Person	0.95	10:30:00
2023-10-27 10:30:01	Person	0.95	10:30:01
2023-10-27 10:30:02	Person	0.95	10:30:02
2023-10-27 10:30:03	Person	0.95	10:30:03
2023-10-27 10:30:04	Person	0.95	10:30:04
2023-10-27 10:30:05	Person	0.95	10:30:05
2023-10-27 10:30:06	Person	0.95	10:30:06
2023-10-27 10:30:07	Person	0.95	10:30:07
2023-10-27 10:30:08	Person	0.95	10:30:08
2023-10-27 10:30:09	Person	0.95	10:30:09
2023-10-27 10:30:10	Person	0.95	10:30:10
2023-10-27 10:30:11	Person	0.95	10:30:11
2023-10-27 10:30:12	Person	0.95	10:30:12
2023-10-27 10:30:13	Person	0.95	10:30:13
2023-10-27 10:30:14	Person	0.95	10:30:14
2023-10-27 10:30:15	Person	0.95	10:30:15
2023-10-27 10:30:16	Person	0.95	10:30:16
2023-10-27 10:30:17	Person	0.95	10:30:17
2023-10-27 10:30:18	Person	0.95	10:30:18
2023-10-27 10:30:19	Person	0.95	10:30:19
2023-10-27 10:30:20	Person	0.95	10:30:20
2023-10-27 10:30:21	Person	0.95	10:30:21
2023-10-27 10:30:22	Person	0.95	10:30:22
2023-10-27 10:30:23	Person	0.95	10:30:23
2023-10-27 10:30:24	Person	0.95	10:30:24
2023-10-27 10:30:25	Person	0.95	10:30:25
2023-10-27 10:30:26	Person	0.95	10:30:26
2023-10-27 10:30:27	Person	0.95	10:30:27
2023-10-27 10:30:28	Person	0.95	10:30:28
2023-10-27 10:30:29	Person	0.95	10:30:29
2023-10-27 10:30:30	Person	0.95	10:30:30

c. The predicted labels are stored text format in database in excel format with date ,time,accuracy.



VI. CONCLUSION

In final stage, the pipeline designed is in such a way that images are given in the input and this input image is shown to the end user and the user selects the Region Of Interest (ROI). Once ROI is selected it is passed through the YOLOv4 weights and the detections are shown on the same page and the detected data is stored in excel and csv format. The UI has been developed using pysimple GUI wherein the UI has the option to browse folder from the personal computer and rest of the things are pipe-lined as written above. This UI can be converted on to

an application on android where all these processes can be run on a smartphone. The scope of work under this project is to make it self adversarial where in new images with new characters can be self- learnt by the model and the model can increase its performance by itself.

VII. FUTURE WORK

Developing an application for the end user Some of the images have circular labels, so in the existing dataset they have not been annotated so detecting and recognizing circular printed labels is not being done that needs to be included. Updating the Neural Network by itself by adding images to the training data and training the model in the backend automatically using Self-Adversarial Network. Need to develop a Android application wherein all the process gets streamlined and updates the data automatically to the server. If the label predicted is not as the actual label there must be an option for the end user to correct the wrongly predicted labels on the UI side.

VIII. ACKNOWLEDGMENT

The authors would like to acknowledge Jaywant Shikshan Prasarak Mandal (JSPM) at Handewadi, Hadapsar, Department of Computer Engineering for development of this application of Character recognition on Forged entities.

REFERENCES

- [1]C. Kumar B., R. Punitha and Mohana, "YOLOv3 and YOLOv4: Multiple Object Detection for Surveillance Applications," 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT), 2020, pp. 1316-1321, doi:10.1109/ICSSIT48917.2020.9214094.
- [2] Ahmad, Tanvir ma, Ying long Yahya, Muhammad Ahmad, Belal Nazir, Shah Haq, Amin Ali, Rahman. (2020), "Object Detection through Modified YOLO Neural Network".
- [3] Basu S, Das N, Sarkar R, Kundu M, Nasipuri M, Basu DK, " An MLP based Approach for Recognition of Handwritten Bangla Numerals " arXiv preprint arXiv:1203.0876. 2012 Mar 5.
- [4]H. Nisa, J. A. Thom, V. Ciesielski and R. Tennakoon, "A deep learning approach to handwritten text recognition in the presence of struck-out text", 2019 International Conference on Image and Vision Computing



INNO SPACE
SJIF Scientific Journal Impact Factor
Impact Factor: 8.423



ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

9940 572 462 **6381 907 438** **ijirset@gmail.com**



www.ijirset.com

Scan to save the contact details