WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

A Review on Face Recognition Systems Built using YOLO (You Only Look Once)

Manashri Gasti[1], Payal Das[2], Pranav Telangade[3] and Akshaya Prabhu[4]

[1]*Computer Engineering Department, Viva Institute of Technology, India*
[2]*Computer Engineering Department, Viva Institute of Technology, India*
[3]*Computer Engineering Department, Viva Institute of Technology, India*
[4]*Asst.Professor, Computer Engineering Department, Viva Institute of Technology, India*

ABSTRACT

Face detection and recognition are critical technologies used in a variety of applications, including security systems, access control, surveillance, and social media. YOLO (You Only Look Once) is a class of object identification models that employ deep learning methods to recognise things in real-time video and photos. Unlike standard object detection systems, which employ numerous steps to detect objects, YOLO uses a single neural network to forecast the bounding boxes and class probabilities for each identified object in the input picture or video frame. This makes it incredibly quick and efficient, allowing it to handle video streams at real-time speeds on a conventional GPU. In the course of extensive research conducted in various domains, researchers found that YOLO is not commonly utilised for face recognition tasks. Instead, YOLO can be used to discover faces in an image or video frame, and then a different face recognition algorithm can be used to recognise the found faces. This study presents the existing face detection and recognition systems built using YOLO.

Keywords: Deep Learning, Face detection, Face Recognition, Object Detection, YOLO

1. INTRODUCTION

CNN (Convolutional Neural Network) face recognition is a popular approach in deep learning-based face identification. CNNs are a form of neural network design used for image processing applications such as object recognition, segmentation, and classification. YOLO (You Only Look Once) is a collection of object identification models that recognise things in real-time video and photos using deep learning methods. It was created at the University of Washington by Joseph Redmon and his associates. YOLO, unlike standard object detection systems, uses a single neural network to forecast the bounding boxes and class probabilities for each identified object in the input picture or video frame. This allows it to handle video streams at real-time speeds on a conventional GPU, making it incredibly quick and efficient.

The YOLO model is created using a deep convolutional neural network (CNN) architecture and detects objects of various sizes and shapes utilising the notion of anchor boxes. To learn how to detect objects of interest, the model is trained on a large dataset of annotated photos. YOLO has achieved cutting-edge performance on various object detection benchmarks, including COCO and PASCAL VOC, and has grown in popularity in both academic and industrial areas due to its speed and accuracy. It has also been employed in a variety of applications like self-driving cars, robotics, surveillance, and others. The YOLO (You Only Look Once) algorithm has numerous applications, including image and video object detection, pedestrian detection, face detection and recognition, industrial inspection, medical imaging, and environmental monitoring. Overall, the YOLO algorithm's speed, precision, and versatility make it a valuable tool in a variety of businesses.

2. BACKGROUND

During extensive research conducted in various domains, researchers found that YOLO (You Only Look Once) is a popular object detection technique, but it is not commonly used for face recognition tasks. Instead, YOLO can be used to discover faces in an image or video frame, and then a different face recognition algorithm can be used to recognise the found faces. There are various face recognition algorithms available, including deep learning-based methods such as FaceNet and OpenFace, which can be used in conjunction with YOLO for face recognition. These algorithms function by storing facial traits into a high-dimensional vector space and comparing the vectors to see if two faces match.

To begin, YOLO can be trained on a face dataset to recognise faces in an image or video frame. The coordinates of the detected face bounding boxes can be output by the YOLO model. The discovered faces can then be cropped using the bounding box coordinates. Finally, to recognise the cropped faces, a separate face recognition algorithm, such as FaceNet or OpenFace, might be utilised. The face recognition system can encode the trimmed faces' facial traits and compare them to a database of known faces to see whether there is a match.

3. REVIEW OF LITERATURE

A survey of existing literature and methodologies was conducted to models built using YOLO (You Only Look Once). Considering that YOLO, a proficient object detection model, is proving its efficiency in face recognition as well, this study consists of the most relevant papers on face recognition using YOLO.

M P. Aneesa et al. [1] discuss the importance of CNN, different datasets used in face recognition systems, and different CNN models. Deep learning CNN can be used for face recognition to improve authentication security. This paper provides an overview of face recognition using CNN in various scenarios. CNN can detect the presence of a human face by taking real-time images from surveillance cameras. Face recognition can be accomplished using CNN models such as ResNet, Mobile Net, VGG, Alex Net, and Google Net. These architectures are applicable to both image and real-time video.

Bansal, Ankan, and colleagues [2] investigated the effects of dataset selection and training processes on deep convolutional neural networks for face recognition. This document presents some recommendations for making decisions. There is a wealth of video footage that has many more posture and expression changes than still photographs. This work offered a fresh dataset of 22,075 films and 3,735,476 annotated frames to guarantee that researchers may take advantage of this potential. The significance of reducing label noise from the dataset and choosing larger or deeper datasets cannot be overstated. Similarly, aligning faces with precise critical spots during training and assessment improves performance.

P. Jiang et al. [4] examined the evolution of YOLO variations throughout time. According to the study, the original YOLO design is made up of 24 convolution layers followed by two completely linked layers. YOLO predicts many bounding boxes per grid cell, but only those with the highest Intersection Over Union (IOU) with the ground truth are chosen, a process known as non-maxima suppression. Based on the investigation, the following observations were made: The YOLO versions are still in their early stages, so there is plenty of space for further study. Very useful for scenario implementations. There is still opportunity for advancement in the future. This article might concentrate on comparing implementations, such as scenario analysis. Also, the study on YOLO V1 in this paper is relatively restricted.

Ramachandra et al. [5] detailed the underlying CNN algorithm structure and offered an overview of the real-time object detection approach employed by YOLO. YOLO can train the whole model in parallel since it is a unified object detection model that is simple to create and train using its simple loss function. The study shows that the second major version of YOLO, known as YOLOv4, provides the optimum speed and accuracy trade-off for object detection. This article describes the underlying CNN algorithm structure in addition to offering an overview of YOLO's real-time object identification approach.

D. Garg et al. [6] use a deep learning model to improve the accuracy of identifying faces. Their research analyses YOLO (You only look once), a popular deep learning library, and compares the efficiency of recognising the face to the old technique. For recognising faces in movies, the suggested model employs the convolutional neural network as a deep learning strategy. Their model is trained and tested using the FDDB dataset. The proposed network accepts a 448 by 448 colour picture as input. The design is made up of 7 convolutional layers, followed by a max pooling layer with a size of 2 x 2. Next three completely linked layers are added, followed by the output layer.

Abhishek Rana et al. [7] described an architecture for a convolutional neural network to recognise the face using the YOLO framework. The training and testing of a model were both performed on two GPUs, which recognised faces at a quicker pace in real time. During the training and testing phases of YOLO (You Only Look Once), a single neural network was applied to the whole image. In their study, they predicted the bounding box using image characteristics, and they predicted parallel bounding boxes across a picture. The study concluded that YOLO provides for end-to-end training while keeping a real-time tempo, allowing for high average accuracy to be maintained.

Ali Gombe et al. [8] provide an experimental methodology for progressively reducing model size, with the goal of generating a small-scale model suited for deployment in a resource-limited context. This was accomplished by layer removal and filter resizing. Substantial research was conducted utilising the "You Only Look Once" approach (YOLOv3tiny). To examine the impact of the model's size reduction on a typical computer vision job such as face identification, the researchers employed two public datasets. The results clearly illustrate that a large reduction in model size has only a little influence on overall model performance.

Mardiana et al. [9] created a prototype of a library attendance system to aid library administration with face recognition of library visitors. The YOLOv5 technique is used in this work to find objects via image processing. The library attendance system incorporates three subsystems: an API service, YOLOv5 facial recognition, and a visitor identity system. The findings show that the library attendance system can work successfully, access the API service, and present information about face detection results. As a result, the system is compatible with the existing library automation system. The Library Attendance System may incorporate three subsystems, one of which is a facial recognition system based on YOLOv5.

Femi Emmanuel Ayo et al. [10] created a system for security and analytics that may be utilised to get analytics at a large event and a face mapping of someone involved in a crime scene for security purposes. This method used YOLO to recognise human faces and geometric analysis to map the human face rapidly and correctly. The research was carried out utilising NodeJS to run the Javascript application on the back end and to display web assets on the server. The study shows that utilising YOLO with Geometric analysis may reliably discern the gender of a face and forecast age. As a result, geometric analysis is a useful tool for facial identification and feature prediction.

Delong et al. [13] consider face detection to be a broad object detection job. They created YOLO5Face, a face detector based on the YOLOv5 object detector, by including a five-point landmark regression head and using the Wing loss function. They created detectors in a variety of model sizes, ranging

from a big model for maximum performance to a very tiny model for real-time detection on an embedded or mobile device. Experiment findings on the WiderFace dataset reveal that the face detectors could attain state-of-the-art performance in virtually all of the Easy, Medium, and Hard subsets, outperforming the more sophisticated designated face detectors. In their work, they proposed YOLO5Face, which is based on the YOLOv5 object detector.

4. ANALYSIS

Table 4.1 presents a detailed analysis of previous studies with a tabular view of the summary, benefits, and TechStack mentioned in each article.

Title	Summary	Advantages	TechStack
Face Recognition using CNN: A Systematic Review (2022) [1]	Importance of CNN models are analysed and different datasets used in face recognition	98.50% accuracy by diminishing overfitting.	Batch normalization CNN
CNNs for Face Detection and Recognition [3]	The proposed system follows two methods: First, Localization and next classification. Both are done in real time.	The results of our measurement are known to be False Positive = 0.0884% False Negative = 3.37% Accuracy = 91.0% Average IOU = 58.7% Average center deviation = 11.3 pixels Frames per Second = 11 on CPU, 89 on GPU	Two Stream CNN (3x3 Conv layer), Leaky ReLU, 2x2 Max Pooling, Cascade CNN
A review of YOLO Algorithm Developments (2022) [4]	This research paper gives a brief overview of the You Only Look Once (YOLO) algorithm and its subsequent advanced versions. This paper will compare the main differences among the five YOLO versions from both conceptual designs and implementations.	Main improvement measures of YOLO network from V1 to V5: The grid division is responsible for detection, confidence loss in version 1. Anchor with K-means added, two-stage training, full convolutional network in version 2. Multi-scale detection by using FPN in yolov3 SPP, MISH activation function, data enhancement Mosaic/Mixup, GIOU(Generalized Intersection over Union) loss function for yolov4 Flexible control of model size, application of Hardswish activation function, and data enhancement in yolo v5 [4]	YOLO V1 to V5
A Deep Learning Approach for Face Detection using YOLO (2018) [6]	The paper compares the accuracy of detecting the face in an efficient manner with respect to the traditional approach. A model is finetuned on various performance parameters and the best suitable values are taken into consideration. It is also compared the execution of training time and the performance of the model on two different GPUs.	The model was trained for 25 epochs with gradient descent optimizer algorithm. It was observed that accuracy remained nearly constant 92.2% after 20 epochs and the best value of learning rate is considered after trying different values and it is 0.0001	The proposed network accepts a 448 by 448 colour image as an input; Design is made up of 7 convolutional layers and a 2 x 2 max pooling layer.
Face Detection Using YOLO (2022) [7]	After fine-tuning all parameters and hyper-parameters of the proposed model, the accuracy of the proposed model was compared to other face identification techniques. After 20 epochs, it is discovered that the IoU accuracy obtained is the best, at 92.2 percent	The suggested model was shown to be more accurate than the haar cascade technique and the R-CNN based face detection model	YOLO, NMS (Non-Maximum Suppression)
Library Attendance System using YOLOv5 Faces Recognition (2021)	System integrates 3 sub- systems: API service, face recognition and visitor identification system.	mAP value is 0.4 Detection of multiple objects is 0.14267s.	YOLOv5 algorithm

[9]		Detection of multiple objects with the YOLOv5 algorithm on video is 76.92 fps.	
YOLO5Face: Why Reinventing a Face Detector (2022) [13]	YOLO5Face face detector with added five-point landmark regression head and the Wing loss function.	The system achieves an accuracy of 89.54%	CNN, YOLO, real time embedded device

5. CONCLUSION

The survey can conclude that YOLO, a popular object detection model can be used with other algorithms for face detection and recognition as well. YOLO is a single staged detector and processes the image in one go. The developments of YOLO versions are still ongoing and in future YOLO can prove to obtain speedy results in face recognition as there is no back propagation.

ACKNOWLEDGEMENT

We would like to express a deep sense of gratitude towards our major project guide Prof. Akshaya Prabhu, Computer Engineering Department, for her constant encouragement and valuable suggestions. The work that we could present is possible because of her timely guidance. We would like to thank our Department's Project Head Prof. Janhavi Sangoi for her support and coordination; Head of the Computer Engineering Department, Prof. Ashwini Save and In charge HOD Prof. Sunita Naik for their support and coordination.

REFERENCES

[1]. Aneesa M P, Saabina N, Meera K, 2022, Face Recognition Using Cnn: A Systematic Review, International Journal of Engineering Research & Technology (Ijert) Volume 11, Issue 06 (June 2022),

[2]. Bansal, Ankan & Castillo, Carlos & Ranjan, Rajeev & Chellappa, Rama. (2017). The Do's and Don'ts for CNN- based face Verification

[3]. Yicheng An, Jiafu Wu, Chang Yue "CNNs for Face Detection and Recognition", June 2017

[4]. Peiyuan Jiang, Daji Ergu, Fangyao Liu, Ying Cai, Bo Ma, A Review of Yolo Algorithm Developments, Procedia Computer Science,Volume 199,2022,Pages 1066-1073,ISSN 1877-0509,https://doi.org/10.1016/j.procs.2022.01.135.(https://www.sciencedirect.com/science/article/pii/S1877050922001363)

[5]. V. Viswanatha & R K, Chandana & A.C, Ramachandra. (2022). Real Time Object Detection System with YOLO and CNN Models: A Review. Xi'an Jianzhu Keji Daxue Xuebao/Journal of Xi'an University of Architecture & Technology. XIV. 10.37896/JXAT14.07/315415.

[6]. Garg, Dweepna & Goel, Parth & Pandya, Sharnil & Ganatra, Amit & Kotecha, Ketan. (2018). A Deep Learning Approach for Face Detection using YOLO. 1-4. 10.1109/PUNECON.2018.8745376.

[7]. Abhishek Rana. (2022). Face Detection Using YOLOhttps://doi.org/10.5281/zenodo.6567441

[8]. ali-gombe, Adamu & Elyan, Eyad & Moreno-García, Carlos & Zwiegelaar, Johan. (2021). Face Detection with YOLO on Edge. 10.1007/978-3-030-80568-5_24.

[9]. Mardiana, & Muhammad, Meizano & Mulyani, Yessi. (2021). Library Attendance System using YOLOv5 Faces Recognition. 68-72. 10.1109/ICCTEIE54047.2021.9650628.

[10]. Femi Emmanuel Ayo et al., "Geometric Analysis and YOLO Algorithm for Automatic Face Detection System in a Security Setting ", 2022 J. Phys.: Conf. Ser. 2199 012010

[11]. Rakshit Umesh Shetty, Prof. Ghanashyam Phadke, Shubham Niraj Kumar, Ashish Kantilal Tiwari, "Real Time Face Mask Detection using Yolov5 and OpenCv", International Journal for Research in Applied Science & Engineering Technology (IJRASET) ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538 Volume 10 Issue VI June 2022- Available at www.ijraset.com

[12]. Ultralytics, "Yolov5," 2021, February 2021, https://github.com/ultralytics/yolov5.

[13]. Qi, Delong & Tan, Weijun & Yao, Qi & Liu, Jingfeng. (2021). YOLO5Face: Why Reinventing a Face Detector.