Integration of object detection and face recognition in real time

دمج كشف الأشياء والتعرف على الوجوه في الوقت الحقيقي

Mohammed Hashem Almourish[1], *
Ahmed Yousef Saeed[2]
Borhan Mohammed Radman[3]
Hazem Abdulah Mohammed[4]
Shakeep Fare`a Alban`a[5]
Abdulfattah E Ba Alawi[6]

[1] Department of communication & Computer Engineering - Taiz University - Yemen.

[2,3,4,5,6] Department of Software Engineering - Taiz University - Yemen.

Mohmedalmourish@gmail.com,ahmed12.12f@gmail.com,borhanpfaloubidy91@gmail.com,
haze770747@gmail.com,sh147444@gmail.com,baalawi.abdulfattah@gmail.com

ISSN: 2410-7727

2003 ١٤٢٤

جامعة الأندلس
للعلوم والتقنية
Alandalus University For Science & Technology

ISSN: 2410-7727

Integration of object detection and face recognition in real time

ABSTRACT

This paper presents a new system for integrating real-time object detection and face recognition technique. Firstly, object detection technology was used by model YOLOv3, which is trained on the darknet53 network that uses COCO dataset. The YOLOv3 model can detect 80 different classes of objects. Based on that, we will be able to detect people, buses, cats, dogs, cars, and more. while Haar cascade classifier and Local Binary Patterns Histogram (LBPH) algorithm were used to recognize faces and identify the names of people and objects on the camera in real time. Experiments showed the effectiveness of the proposed system for real-time object detection using YOLOv3-416 achieving a mAP-50 of 55.3, with an accuracy in the range of 81-92%. While the accuracy of real time face recognition using Haar cascade classifier and LBPH algorithm was 83%.

Keywords- Haar cascade, face recognition, object detection, real-time.

دمج كشف الأشياء والتعرف على الوجوه في الوقت الحقيقي

محمد هاشم المريش، احمد يوسف سعيد، برهان محمد ردمان ، حازم عبدالله البيطار،
شكيب فارع البناء ، عبدالفتاح إسماعيل باعلوي

الملخص

تقدم هذه الورقة نظامًا جديدًا لدمج تقنية اكتشاف الأشياء في الوقت الفعلي والتعرف على الوجوه. أولاً، تم استخدام تقنية اكتشاف الكائنات بواسطة الطرازYOLOv3 ، الذي تم تدريبه على شبكة darknet53 التي تستخدم مجموعة بيانات . COCOيمكن لنموذج YOLOv3 اكتشاف 80 فئة مختلفة من الكائنات. وبناءً على ذلك، سنتمكن من اكتشاف الأشخاص والحافلات والقطط والكلاب والسيارات والمزيد. بينما تم استخدام خوارزمية مصنف Haar cascade وخوارزمية المدرج التكراري للأنماط الثنائية المحلية (LBPH) للتعرف على الوجوه وتحديد أسماء الأشخاص والأشياء على الكاميرا في الوقت الفعلي. أظهرت التجارب فعالية النظام المقترح لاكتشاف الأشياء في الوقت الفعلي باستخدام YOLOv3-416 محققاً MAP-50 من 55.3، بدقة في نطاق 81-92٪. بينما كانت دقة التعرف على الوجوه في الوقت الفعلي باستخدام مصنف Haar cascade وخوارزمية LBPH 83٪.

ISSN: 2410-7727

I. Introduction

In recent years, object detection has witnessed a great improvement in reliability and accuracy, as well as face recognition techniques. There are many object detection systems and also face recognition systems that work independently, however the process of combining them faces many problems related to the long time to recognize objects and faces and poor accuracy. The first problem is the problem of detecting objects in real time and the problem of recognizing faces in real time independently. Object detection is the process of searching for objects of the same type and shape that have similar features. For example, detecting for a car without regard to its type or name, as the origin of all cars has the same exterior features. While recognizing human faces in real time requires speed in recognition as well as high accuracy. The second problem, which is the main problem, is the combination of object detection and face recognition, with high accuracy and small implementation time. To solve these problems, many techniques were used, but there were drawbacks in the performance or accuracy of these techniques in solving these problems, and therefore modern techniques and algorithms were proposed and gave high results.

In this paper the YOLOV3 model for object detection is proposed while a Haar cascade classifier with LBPH algorithm for real-time face recognition is proposed. The results of the proposed system showed high accuracy and small execution time.

The content of this paper is organized as follows:

The second section gives a brief overview of the related work. A methodology is described in section 3. Our results are shown in the fourth section. Conclusions are given in section 5.

2. Related work

This section reviews the literature related to work presented here. Many researchers have introduced many techniques to solve problems related to object and object recognition in real time. Several technologies have been proposed such as Deep Neural Networks (DNN), SSD [1], R-CNN [2], Fast R-CNN [3], Faster R-CNN [4], Mask R- CNN [5] and the YOLO Series. Most of the new techniques are versions of old techniques, as newest releases are developments of deep learning techniques. The R-CNN version is a real-time object detection technique that uses many layers to extract features and the SVM machine learning classifier, but its execution speed is slow. Therefore, it has been developed with new versions such as Fast R-CNN so that its performance is faster and the SVM classifier has been replaced by softmax function. However, the speed is still insufficient to identify

ISSN: 2410-7727

objects in real time as it takes two seconds. Because Fast R-CNN is still not fast enough, a third version of R-CNN was released, so this RPN version has been used in many applications to speed up real-time object detection [6]. Reference [7] presented a new model that called You-Look Only-Once (YOLO) and its first release was YOLOv1, and the size of the image input increased to 224 * 224 pixels, and this technique also detected objects in real time at speeds of 45/155 frames per second (fps). It has been trained on ImageNet-1000 dataset based on Darknet. One of the drawbacks of YOLOv1 technique is that it cannot detect small objects in the image, so it was developed into a new version of YOLOv2. The YOLOv2 version was developed by increasing the input image size to 448 * 448 pixels and also increasing its speed to 67 fps [8]. Then this version was developed into a new version called YOLOv3, which was used in this research to identify objects in real time with high accuracy, but its speed is less than YOLOv2, reaching 30 fps, due to the large number of layers in YOLOv3, which reaches 106 layer [9].

YOLOv3 technique was proposed to solve the problem of object detection in real time, which is considered this technique an evolution of deep learning technology that uses the Darknet53 network as its backbone. It is the latest version of the YOLO approach. The YOLOv3 Network is training on a COCO dataset that can detect 80 different classes of objects. Based on that, we will be able to detect people, buses, cats, dogs, cars, and more [10-11].

As for the problem of face detection and recognition in real time, many studies have been proposed to solve this problem such as Scale Invariant Feature Transform (SIFT) [12], Speeded-Up Robust Features (SURF) [13], LBPH [14-15], etc.

Reference [16] proposed three classifiers including Haar cascade to detect face, eyes and mouth, and the position predictor value was approximately 78.18%.

While the reference [17] proposed a system for face detection by combining a Haar cascade with an edge-direction matching algorithm. This system provided better matching, but the face detection speed was lower.

In this paper, a Haar cascade classifier and LBPH algorithm are proposed. The Haar cascade classifier is a classifier of machine learning technique to detect objects in real time. It identifies the faces in the images and extracts their features and then classifies them. Then they pass to the LBPH algorithm that compares the histogram of the tested images with the images in the database. Finally, the proposed algorithms for detecting objects in real time are combined with the proposed algorithms for recognizing faces in real time to solve the problem of identifying objects and the face in real time with high efficiency.

ISSN: 2410-7727

3. Methodology

In this section, we will present a methodology for real-time object detection and face recognition as shown in figure 1. The methodology steps of the proposed system can be summarized as follows: data collection, object detection using YOLOv3 model, and finally, objects and faces are recognized in real time using Haar cascade classifier with LBPH algorithm, and their names are identified in the camera.

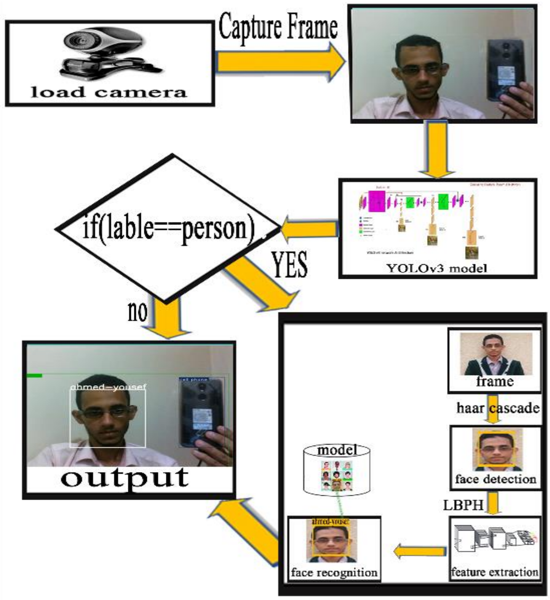

Figure 1. Methodology for object detection and face recognition in real time.

1. Data collection

Object and face datasets were collected from cell phone and laptop cameras. As for the face image, it was taken to the front face or facing the camera.

ISSN: 2410-7727

2.Methods

In this stage, we will show real time object detection method by YOLOv3 model and also Real time face recognition method by Haar cascade classifier and LBPH algorithm.

A. Real time Object detection

The YOLOv3 model is proposed for real-time object detection. It is the third version of YOLOv that is used to detect objects in real time, with a speed of 30 fps. Figure 2 shows the architecture of the YOLOv3 model.

Real-time object detection is done by following steps:

1- Training phase: The YOLOv3 model is trained on the darknet53 network, which is trained on a 106-layer neural network based on the MSCOCO dataset that contains many color images of objects, where the size of the images is 416 × 416 pixels.

2- The test phase: The YOLOv3 model is tested on new images of objects that are resized to 416 * 416 pixels to fit the model's input, and the images are predicted with high accuracy.

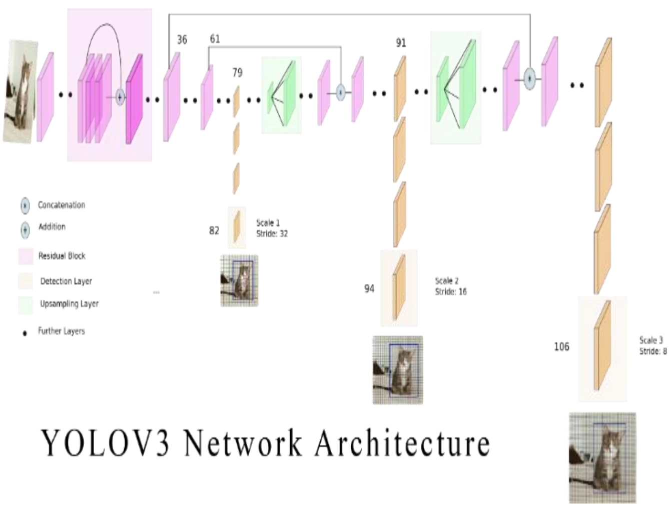

Figure 2. YOLOv3 architecture.

ISSN: 2410-7727

B. Real time face recognition.

In this part, faces will be recognized in real time as in the following steps: -

1- Face Detection: The Haar Cascade classifier was proposed based on machine learning, which detects faces in images. If it finds a face in the image, it subtracts the face from the image and then extracts the features for the face. Figure 3 shows the architecture of the Haar Cascade classifier.

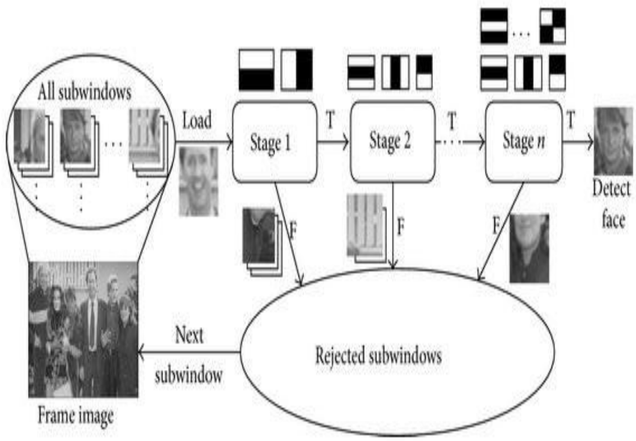

Figure 3. The architecture of the Haar Cascade classifier.

2- Face recognition: In this step, the LBPH algorithm was proposed, which recognizes the face after it is detected by the Haar Cascade by comparing the histogram of the images to be recognized and the histogram of the images stored in the database. If the histogram matches, the images will be recognized. Figure 4 shows the architecture of LBPH algorithm.

3- Finally, the proposed YOLOv3 real-time object detection model is integrated with the Haar Cascade classifier and LBPH real-time face recognition algorithm to solve the real-time object and face identification problem with high efficiency.

ISSN: 2410-7727

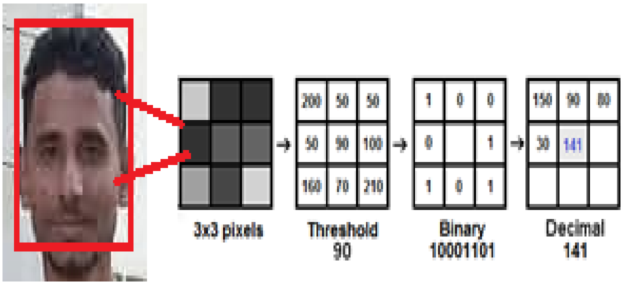

Figure 4. LBPH operations

4. Result

The proposed system was applied to objects and faces with different intensities using frontal and profile faces. Faces and objects are captured by a laptop camera and passed to the YOLOV3 model, which detects objects and faces and places a frame around them. After that, the faces in the images are recognized by a Haar cascade with the LBPH algorithm. The Figure 5 shows the practical detection of objects and faces with identifying the names of objects and also the names of people on the camera in real time. The cell phone and the bottle were detected as an object using the Yolov3 model in real time. Meanwhile, the faces were detected by a Haar cascade classifier and recognized by LBPH. The accuracy of recognizing objects and faces in real time is 83%. System tested on Core i5-2450M and Windows 8.1.

Figure 5. Detected object and face with label in real time.

ISSN: 2410-7727

5. Conclusion

In this paper, a system was proposed that combines face recognition and object detection, where we used the YOLOv3 model to detect different types of objects found in nature, such as humans, dogs, cats, cars, etc. In addition, a Haar cascade classifier has been proposed for face detection and LBPH for face recognition. This paper shows that Yolov3, Haar cascade classifier, and LBPH are suitable techniques for real-time object and face recognition. Experiments demonstrated the effectiveness of the proposed object detection system using YOLOv3-416 achieving a mAP-50 of 55.3, which was trained on dark network 53 using a COCO dataset where the accuracy of the system was in the range of 81-92%. While the accuracy of real time face recognition using Haar cascade classifier and LBPH algorithm was 83%.

References

[1] Wei, L., Dragomir, A., Dumitru, E., Christian, S., Scott, R., Cheng-Yang, F., Alexander, C.B.: SSD: Single Shot MultiBox Detector. Springer International Publishing, pp. 21–37 (2016).

[2] Girshick, R., Donahue, J., Darrell, T., Malik, J.: Rich feature hierarchies for accurate object detection and semantic segmentation. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 580–587 (2014).

[3] Girshick, R.: Fast r-cnn. In Proceedings of the IEEE international conference on computer vision, pp. 1440-1448 (2015).

[4] Ren, S., He, K., Girshick, R., Sun, J.: Faster r-cnn: Towards real-time object detection with region proposal networks. In Advances in neural information processing systems, pp. 91-99 (2015).

[5] He, K., Gkioxari, G., Dollár, P. Girshick, R.: Mask R-CNN. In Proceedings of the IEEE international conference on computer vision, pp. 2961-2969 (2017).

[6] PULKIT, S. Step-by-Step Introduction to the Basic Object Detection Algorithms (Part1). https://www.analyticsvidhya.com/blog/2018/10/a-step-by-step-introduction-to-the-basic-object-detection-algorithms-part-1/.[Accessed: 26- Dec- 2019].

[7] Redmon, J., Divvala, S., Girshick, R., Farhadi, A.: You only look once: Unified, real-time object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 779-788 (2016).

ISSN: 2410-7727

[8] Redmon, J., Farhadi, A.: YOLO9000: better, faster, stronger. In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 7263-7271(2017).

[9] Redmon, J., Farhadi, A.: Yolov3: An incremental improvement. arXiv preprint arXiv:1804.02767 (2018).

[10] Medium Terms of Service. https://towardsdatascience.com/yolo-v3-object detection-53fb7d3bfe6b. [Accessed: 31- Dec- 2019].

[11] Asif, S.: Human detection and distance estimation with monocular camera using YOLOV3 neural network. UNIVERSITY OF TARTU,faculty of science and technology (2019).

[12] FENG, Yuping, AN, Xuemei, et LIU, Ximei. The application of scale invariant feature transform fused with shape model in the human face recognition. In: Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC), 2016 IEEE. IEEE,. p. 1716-1720 (2016).

[13] VERMA, Reetika et KAUR, Rupinder. Enhanced Character Recognition Using Surf Feature and Neural Network Technique. IJCSIT) International Journal of Computer Science and Information Technologies, vol. 5, no 4, p. 5565-5570 (2014).

[14] Ahmed, A., Guo, J., Ali, F., Deeba, F., & Ahmed, A.: LBPH based improved face recognition at low resolution. In 2018 International Conference on Artificial Intelligence and Big Data (ICAIBD) (pp. 144-147). IEEE (2018).

[15] Abuzneid, M. A., Mahmood, A.: Enhanced Human Face Recognition Using LBPH Descriptor, Multi-KNN, and Back-Propagation Neural Network. IEEE Access, 6, 20641-20651(2018).

[16] Li Cuimei, Qi Zhiliang. “Human face detection algorithm via Haar cascade classifier with three additional classifiers”, 13th IEEE International Conference on Electronic Measurement & Instruments, pp. 01-03, (2017).

[17] Souhail Guennouni, Anass Mansouri.“Face Detection:Comparing Haar-like combined with Cascade Classifiers and Edge Orientation Matching”, International Conference on Wireless Technologies, Embedded and Intelligent Systems (WITS), pp. 02-04, (2017).