Ministry of Higher Education and
Scientific Research
University of Mosul
College of Computer Science and
Mathematics
Department of Computer Sciences

Person Detection in Video File Based on Hybrid Intelligent Technique

A Thesis Submitted to the Council of the College of
Computer Science and Mathematics
University of Mosul
as a Partial Fulfillment of Requirements
for the Degree of Doctor of Philosophy in
Computer Sciences

BY

Lubna Thanoon Ehmeedy Al_Kahla

Supervised by

Asst. Prof. Dr. Jamal Salahaldeen Majeed Al_Neamy

2022 A.D. **1443 A.H.**

Abstract

Several security threats have spread recently, offset by the development of surveillance cameras in public and even private places, which played a role in limiting the spread of these breaches. Despite this, the human role in examining and analyzing videos recordings to identify a person's movements can not be neglected. Detecting people and tracking their movement accurately and effectively has become of great interest. With the rapid development of artificial intelligence in recent years, personal identification by relying on faces has gained more attention with its many advantages, the most important of which are non-contact and high synchronization, Ease of use, and a high potential for use in government, public utilities and security.

In this thesis, a person detection system for video files is designed to verify the presence of persons in the video file or not.

This system consists of several stages: The first stage is to convert the video files into a set of frames, then face areas are identified using the Viola-Jones method, The second stage is extracting the characteristics of these faces based on two techniques, the first technique is deep feature extraction, which is the pre-trained network VGG16-CNN. The second technique is shallow feature extraction, which is represented by three methods Principal Component Analysis, Local Binary Patterns And, Gabor Filter. The new addition in this thesis is represented by adding the self-organization map (SOM)algorithm into VGG16-CNN to cluster each person’s face images into separate groups for identifying the label for each person in the video file.

Additionally, This layer is compared by the K-Means algorithm, The last stage is classification, which is the stage of detecting persons based on the multilayer percepton neural network.

The system is tested on five video files of different times, and each video contains different numbers of people. The results obtained through the confusion matrix showed that the system gives a high classification accuracy and a good detection rate when using the VGG16-CNN algorithm for feature extraction and the SOM algorithm, clustring where the highest classification accuracy was obtained which is (93%) and the highest precision rate was (100%).Through the test results represented by measures of precision, recall, and accuracy, which are obtained from the confusion matrix, it is found that the system is strong and has a high detection rate.

وزارة التعليم العالي والبحث العلمي
جامعة الموصل
كلية علوم الحاسوب والرياضيات
قسم علوم الحاسوب

كشف الشخص فـــي ملف فيديو بالاعتماد على تقنية ذكائية هجينة

اطروحة مقدمة
الى مجلس كلية علوم الحاسوب والرياضيات في جامعة الموصل
كجزء من متطلبات نيل شهادة دكتوراه فلسفة في
علوم الحاسوب

من قبل

لبنى ذنـــون إحميـــدي الكحلـــة

بإشراف

أ.م.د. جمال صلاح الدين مجيد النعيمي

1443 هـ

2022 م

الخلاصة

انتشرت في الاونة الاخيرة العديد من التهديدات الامنية المتنوعة قابلها تطور وانتشار اصناف من كاميرات المراقبة في الاماكن العامة وحتئ الخاصة التي ساعدت من الحد في انتشار هذه الخروقات وعلئ الرغم من ذلك لم يستغنى عن الدور الذي يلعبه التدخل البشري في عمليات فحص وتحليل التسجيلات الفيديوية لمعرف تحركات شخص معين والاماكن التي ينتقل فيها ويتطلب ذلك وقت وجهد بشري لذلك اصبحت عملية تحديد الاشخاص وتتبع حركتهم بدقة وفعالية من الامور بالغة الاهتمام ومع التطور السريع للذكاء الاصطناعي في السنوات الاخيرة اكتسب التحديد الشخصي بالاعتماد علئ الوجوه المزيد من الاهتمام بما له من مزايا عديدة اهمها عدم الاتصال والتزامن العالي وسهولة الاستخدام وامكانيته العالية من الاستخدام في الحكومة والمرافق العامة والامن .

في هذه الاطروحة تم تصميم نظام بحث عن اشخاص داخل ملفات فيديوية لتحقق من وجود الاشخاص في الملف الفيديو من عدمه ,هذا النظام مكون من عدة مراحل المرحلة الاولى تحويل الملفات الفيديوية الى مجموعة من الاطر ثم يتم تحديد مناطق الوجه باستخدام طريقة Viola Jones ,المرحلة الثانية هي استخلاص الخواص الخاصة بتلك الوجوه بالاعتماد على تقنيتين الاولى هي deep feature extraction والتي يتم استخدام شبكات التعليم العميق وهي الشبكة المدربة مسبقا VGG16-CNN, والتقنية الثانية هي shallow feature extraction والمتمثلة بالطرق التقليدية لاستخلاص الصفات وهي Principal Component Analysis, Local Binary Patterns و Gabor Filter,المرحلة الثالثة هي عنقدة صور الوجه التابعة لكل شخص الى مجاميع منفصلة لغرض تحديد lable لكل شخص بجميع لقطاته بالملف الفيديوي وتتم هذه المرحلة باستخدام خوارزمية ال self organization map (SOM clustring) ومقارنة دقة وكفاءة هذه الخوارزمية مع خوارزمية ال K-Means ، المرحلة الاخيرة هي التصنيف وهي مرحلة تحديد

الاشخاص بالاعتماد على Artificial Neural multilayer percepton Network وتدريب هذا المصنف بالاعتماد على Backpropagation Artificial Neural Network .

تم اختبار النظام على خمس ملفات فيديوية مختلفة الاوقات وكل فيديو يحتوي على اعداد مختلفة من الأشخاص، اظهرت النتائج التي تم الحصول عليها من خلال الconfusion matrix ان النظام اعطى دقة تصنيف عالية ومعدل كشف جيد عند استخدام خوارزمية VGG16-CNN في استخلاص الخواص وخوارزمية SOM clustring, حيث تم الحصول على اعلى دقة تصنيف وهو (93%) واعلى معدل precision وهو(100%) .

من خلال نتائج الاختبار المتمثلة بالمقاييس precision , recall , accuracy والتي تم الحصول عليها من ال confusion matrix تبين ان النظام قوي ولديه معدل كشف عالي فقد اثبت النظام قدرته على تحديد اشخاص داخل الملفات الفيديوية من حيث وجودهم داخل هذه الملفات من عدمه مما يجعله مناسبا للاستخدام في تحديد الهوية والبحث عن المطلوبين ومراقبتهم وتتبع حركتهم.