AlQalam
Alq J Med App Sci

https://journal.utripoli.edu.ly/index.php/Alqalam/index eISSN 2707-7179

Type of the Paper (Article, Review, Communication, etc.)

The Title of Your Article (Write Each Word in Capital Letter)

Firstname Surname[1]*, Firstname Surname[2]

[1]*Department of XXX, Faculty of XXX, University of XXX, City, Country*
[2]*Department of XXX, Faculty of XXX, University of XXX, City, Country*

ARTICLE INFO

Corresponding Email. *email address of the corresponding author*

Received: *00-00-2024*
Accepted: *00-00-2024*
Published: *00-00-2024*

Keywords. *Counting People, Face Detection, People Detection, Viola Jones LBP, Viola Jones CART.*

ABSTRACT

This paper focuses on the challenging task of counting the number of people in digital still images, which has important applications in many fields such as security, management, education, and commerce. This paper proposes a system that is based on the Viola-Jones face detection methods. This system consists of two parts: a) face detection and b) counting the detected faces. In the face detection part, the Viola-Jones (LBP and CART feature extraction) algorithm is applied to the input image. In the counting part, the detected faces are counted to estimate the number of people in the given image. The Viola-Jones algorithm is applied using 133 images from the People Image Groups dataset, and the best precision achieved is 96.9%. Overall, this paper presents a promising system for accurately counting the number of people in digital static images using a simple and cost-effective approach.

Cite this article. *Lastname Initial, Lastname Initial. Article Title. Alq J Med App Sci. 2024;7(1):816-821.*
https://doi.org/10.0000/ajmas.2020.002

INTRODUCTION

The introduction should briefly place the study in a broad context and highlight why it is important. ItPeople counting in digital images is one of the most active research areas in the field of computer vision, which can be applied in many areas of our daily lives. Estimating the number of people in an image can be used in many practical domains, such as, security, commerce, and management, where the output of counting people systems can be passed to other management systems like in public transport and airports. For instance, people counting systems provide an estimation of the total number of people in a building; such information can be used for security purposes to control the number of visitors. Also, knowing the number of people in retail stores can help to measure customers' flow, which in turn helps managers to optimize their staffing levels and improve customer services. Moreover, people counting can be used for pedestrian traffic management [1-4] and for fire management. People counting techniques are categorized into two main categories: vision-based counting systems and non-vision-based counting systems. The vision-based techniques are also categorized into tracking-based and non-tracking-based counting systems. The non-vision-based people counting systems include all systems that use data from devices such as a heat sensor, light from an infrared beam, or pressure sensors, in other words, not data from a video or an

image captured by a camera. Each class has strengths and weaknesses that are discussed in detail in [5]. The main problem discussed in this study is "how to count the number of people in different still images?". Face detection identifies human faces in an input digital image, hence it is the first step of automatic face recognition, where people can be counted based on the number of faces detected in a digital image. However, there canbe many challenges in the images that hinder the process of face detection and people counting. For instance, occlusion between people in the image, the dimensions and quality of the image, and varying environmental conditions are some of the challenges of people counting systems. The aim of this study is to overcome these challenges and to achieve this, the introduced system works on Viola-Jones face detection using LBP (Local Binary Patterns) feature extraction and Viola-Jones face detection using CART (Classification and Regression Trees) feature extraction algorithms [6-8]. This is to determine the number of people in a static image. The People counting systems based on Viola-Jones face detection using LBP and CART feature extraction provide more information, such as the location and appearance of people, that is not detected by visionless traditional counting systems like the infrared beam systems, which are limited to people counting data only. In addition, the performance of people counting systems based on an infrared counter is less accurate than digital image and video systems. This paper has been divided into five main sections: Section 1 provides a general background on people counting systems in the field of computer vision. Section 2 discusses related work on the field of people counting and its methods. Section 3 provides an overview of the proposed system including: the datasets used in this study, the pre-processing of the used images from this dataset, Viola-Jones face detection algorithms and how people are counted in still images using these algorithms. Then, in section 4, a discussion of the obtained results is introduced. Finally, in section 5, conclusions are presented.

Related Work

Face detection and counting people have been active research areas in the field of computer vision, and several studies have been conducted in this domain. Below is a summary of some of the important works related to face detection and counting people. Face detection is the process of locating and identifying human faces in digital images or videos. One of the most popular face detection algorithms is the Viola-Jones algorithm, which was proposed by Paul Viola and Michael Jones [9, 10] in 2001. It uses a cascade of classifiers based on Haar-like features and AdaBoost to detect faces in images. The algorithm has been widely used in various applications, including security, surveillance, and human-computer interaction. While traditional methods such as Viola-Jones algorithms have been successful in detecting faces under controlled conditions, they may struggle with complex scenarios such as occlusion and varying lighting conditions. Recent advancements in deep learning have led to improvements in face detection performance, particularly with the development of convolutional neural networks (CNNs). A notable method is the Single Shot MultiBox Detector (SSD) proposed by Liu et al. in 2016. The SSD is a convolutional neural network (CNN) that can detect faces in 3 real-time with high accuracy. Other deep learning-based methods, such as Faster RCNN and YOLO, have also been proposed for face detection [10,11]. The RetinaFace algorithm uses a single-stage CNN model to detect faces while achieving significant accuracy [12]. Similarly, the CenterFace algorithm uses a lightweight backbone network to detect faces in images and video streams [13]. Counting people is another challenging task in computer vision, especially in crowded or complex scenes. One approach is to use object detection algorithms, such as the Viola-Jones algorithm, to detect faces in images and then count them. Ittahir et al. proposed a system that presents a basic yet practical approach for people counting in still images using skin color face detection. The

proposed method provides reasonable accuracy given its simplicity, demonstrating its potential as an initial solution for applications where approximate people counts are sufficient [5]. Recently, deep learning-based methods have shown superior performance in crowd density estimation, where CNNs are used to estimate the density of people from input images. One notable study in this area is the work by Chen et al., where they proposed a method for counting people in crowded scenes using a combination of scaleinvariant feature transform (SIFT) and support vector regression (SVR) [15]. The SIFT algorithm is used to extract local features from the image, and the SVR is used to estimate the number of people based on the extracted features. The method was tested on several datasets and showed promising results. Zhang et al. proposed a method for counting people based on deep learning [16]. They used a convolutional neural network (CNN) to extract features from the image and then used a fully connected layer to estimate the number of people. Zhang et al. proposed a method for crowd counting using a deep neural network that is trained to estimate the density of people in an image [17]. The method uses a multi-column CNN architecture that can handle different scales of people in the image. The method was tested on several datasets and showed superior performance compared to state-of-the-art methods. A recent study by Chen et al. proposed a method for counting people in videos using a spatiotemporal attention mechanism. The method uses CNN to extract spatial features and a recurrent neural network (RNN) to capture temporal dependencies in the video sequence [18]. The spatiotemporal attention mechanism is used to focus on the most relevant regions of the image and video frames, leading to improved accuracy. The main difference between using Viola-Jones face detection methods and deep learning is the type of algorithms used. Viola-Jones face detection methods use supervised learning algorithms such as CART and LBP, while deep learning uses unsupervised learning algorithms such as convolutional neural networks. Viola-Jones face detection methods are more accurate for face detection and classification, while deep learning is more accurate for object recognition and image segmentation [19]. In conclusion, face detection and counting people are important research areas in computer vision with several applications. While deep learning-based methods have shown promising results in recent years, the Viola-Jones algorithms are popular face detection algorithms with high accuracy. Therefore, for this study, Viola-Jones algorithms are used for face detection and people counting in still images.

METHODS

In this paper, Viola-Jones methods were applied to detect multiple faces in a digital still image and then those faces were counted to estimate the number of people in an input image. This section introduces: the dataset used, the pre-processing required for the selected images from this dataset, and an overview of the Viola-Jones algorithms for the face detection and people counting processes (see Fig. 1 for the main steps). code.

RESULTS

This section should provide a concise and precise description of the experimental results, their interpretation, as well as the experimental conclusions that can be drawn. All figures and tables should be cited in the main text as Figure 1, Table 1, etc.

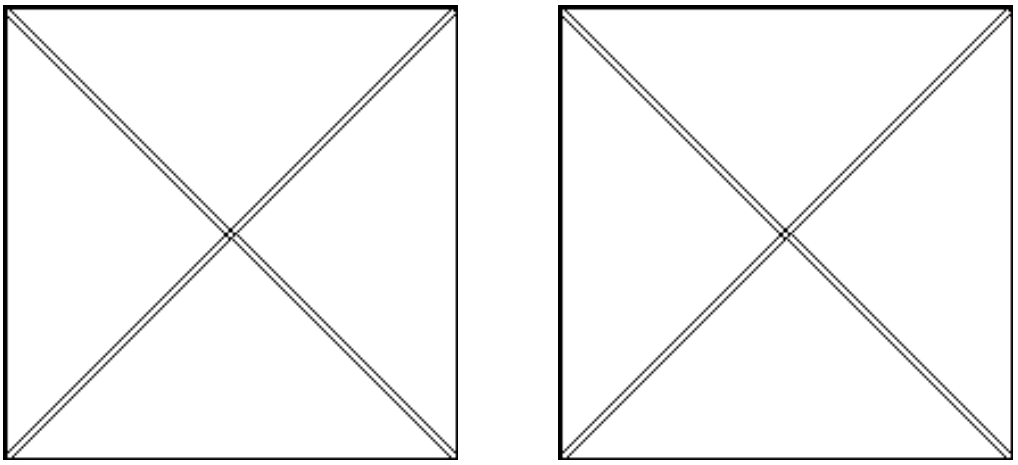
Table 1. This is a table. Tables should be placed in the main text near to the first time they are cited.

Title 1	Title 2	Title 3
entry 1	data	data
entry 2	data	data [1]

[1] *Tables may have a footer.*

The text continues here (Figure 2 and Table 2).

Figure 1. This is a figure. Schemes follow the same formatting.



(**a**) (**b**)

Figure 2. This is a figure. Schemes follow another format. If there are multiple panels, they should be listed as: (a) Description of what is contained in the first panel; (b) Description of what is contained in the second panel. Figures should be placed in the main text near to the first time they are cited.

DISCUSSION

Authors should discuss the results and how they can be interpreted from the perspective of previous studies and of the working hypotheses. The findings and their implications should be discussed in the broadest context possible. Future research directions may also be highlighted.

CONCLUSION

This section is not mandatory but can be added to the manuscript if the discussion is unusually long or complex.

Acknowledgments

In this section, you can acknowledge any support given which is not covered by the author contribution or funding sections. This may include administrative and technical support, or donations in kind (e.g., materials used for experiments).

Conflicts of Interest

Declare conflicts of interest or state “The authors declare no conflicts of interest.” Authors must identify and declare any personal circumstances or interest that may be perceived as inappropriately influencing the representation or interpretation of reported research results.

REFERENCES

References must be numbered in order of appearance in the text (including citations in tables and legends) and listed individually at the end of the manuscript. We recommend preparing the references with a bibliography software package, such as EndNote, ReferenceManager or Zotero to avoid typing mistakes and duplicated references.

In the text, reference numbers should be placed in square brackets [ ] and placed before the punctuation; for example [1], [1–3] or [1,3]. For embedded citations in the text with pagination, use both parentheses and brackets to indicate the reference number and page numbers; for example [5] (p. 10), or [6] (pp. 101–105).

1. Author 1 A, Author 2 C (up-to-six authors). Title of the article. Abbreviated Journal Name. Year;volume(issue):page range.
2. Author 1 A, Author 2 B. Book Title, 3rd ed.; Publisher: Publisher Location, Country, 2008; pp. 154–196.
3. Author 1, A.B.; Author 2, C.D.; Author 3, E.F. Title of Presentation. In Proceedings of the Name of the Conference, Location of Conference, Country, Date of Conference (Day Month Year).
4. Author 1, A.B. Title of Thesis. Level of Thesis, Degree-Granting University, Location of University, Date of Completion.
5. Title of Site. Available online: URL (accessed on Day Month Year).