Human counter using deep learning

Prof. Manisha Vaidya[1], Shravani Bante[2], kritika Badule[3], Samiksha Juwar[4], Arya katre[5]

[1] *Prof. Manisha Vaidya, Artificial Intelligence, Priyadarshini J.L College of Engineering, Maharashtra,*
[2] *Shravani Bante, Artificial Intelligence, Priyadarshini J.L College of Engineering, Nagpur, Maharashtra, India*
[3] *kritika Badule, Artificial Intelligence, Priyadarshini J.L College of Engineering, Maharashtra, India*
[4] *Samiksha Juwar, Artificial Intelligence, Priyadarshini J.L College of Engineering, Maharashtra, India*
[5] *Arya katre, Artificial Intelligence, Priyadarshini J.L College of Engineering, Maharashtra, India*

ABSTRACT

Counting people in visual inspection is a very difficult and complex problem. Automatic counting of people in public places is very important for security management. Many ideas and methods are planned in advance.
This model focuses on creating a physical counter using deep learning (DL) technology. The aim is to introduce a system that can count the number of people in a venue using image recognition algorithms.
Improve the efficiency and effectiveness of human search and calculation of specific medications for stores by creating a system with a graphical user interface and management functions.
This method does not provide accuracy and high performance in serious situations. To provide solutions that are more accurate to people counting and self-identification methods, a method based on expectation extraction and expectation maximization (EM) is now proposed. This study also means social distance measures along with covid-19 safety measures. The Single Shot Detection algorithm (SSD) uses the camera's live stream and the Convolutional Neural Network (CNN) will identify the person and provide an identification number for calculation. Keywords: real-time, human detection, deep-sort, yolov3.

Keywords— *Image Recognition, Cognitive Neural Network, People Counting and Tracking, IoT Devices, Machine Learning& Deep Learning Algorithms.*

1. INTRODUCTION

In Python project, we created a people detection and counting system from network camera photo, otherwise you can take your photo or photo. Generally, this intermediate, in-depth course will help you develop concepts in computer vision and provide some expertise in epistemology. Let us create an interesting project.

In recent years, the application of computer vision in products has made great progress. This growth is driven by the increasing demand for intelligent systems in robotics [1]-[2], surveillance [3], medical imaging [4], service [5], and automotive technology [6] involving humans or pedestrians. One of the intelligent machines that must find objects in images.

It can be used for many video analysis applications such as person recognition, people tracking [7], people counting [8], human computer interaction [9], crowdsourcing [10], and human activity search [11], but it is different from the original Compared to exploration, human exploration faces many problems due to complex environment, matter, and geometry and light changes. Counting and identifying people is an important task in many areas such as security, surveillance, crowd control and social distancing.

However, these tasks are difficult due to effects, design, lighting and camera angles. Traditional methods based on the characteristics of crafts and products often fail to solve these complex problems and achieve high performance.

Thus, deep sensing techniques have been proposed to use the power of neural networks to analyze changes in activity and enable humans to discover, track and sort.

In this article, we propose a new system for people using deep learning to perform calculations and get correct identities. Our system includes 3 additional features: Body Monitor, Body Tracker and Body Awareness. Human body detection is based on YOLOv3 version

and can instantly detect human body with high accuracy and performance.

Human tracker is based on Kalman inference and predicts and corrects the human body in the region throughout the frame. All human identities are based on the Pedestrian model, which assigns identity characters to people based on their appearance and activities. Our device can track multiple people in difficult situations and ensure everyone has the same identity during the video.

2. LITRETATUTRE SURVEY:

Anuj Mohan, Constantine Papageorgiou and Tomaso Poggio proposed a method based on a standard model for the detection of materials in still images with the help of equipment. This approach was realized with the creation of a device that could find people in complex situations. In particular, the system usually detects many parts of the human body in the image by using 4 good models, namely the head, left and right fingers and legs, rather than the whole bod as the detector. The device is then calibrated to ensure that the detected object is in the correct geometric configuration. Using laptop vision for human detection and computation now not only solves the problem of crowd control but also helps in monitoring human behavior in selected environments or situations. Velastin et al. (2020) used the same method to manage passengers to reduce traffic in public transport (Mokayed et al., 2022). Especially in rail systems, optimising self-adjusted passenger transportation time and traveling at maximum speeds is the most important factor affecting system efficiency and performance. For example, check and count how many passengers can enter and exit the service in each vehicle. This difficulty affects not only people's search, but also other searchable objects such as license number, engine, and other objects that provide transportation costs for smart drivers (Velastin et al., 2020). Therefore, an ingenious model was created to determine the difference between teaching and working on the platform in training and train and ground equipment, while ensuring passing access is efficient and fast, convenient, the length of stay, suitable body such as wide door.

In a comparative research paper (Raghavachari et al., 2015), the main challenge in the perspective of human knowledge and the use of mathematics is its accuracy in many cases and situations. Orienting different cameras will enable human analysis and analysis to estimate the cost of equipment needed for various applications. In addition, search accuracy may vary depending on the number of people, and the number of people may increase or decrease during the day. Acne occurs when people's conditions are so affected that it becomes difficult to distinguish the individual.

In addition, light or weather conditions, such as blurriness caused by camera shake, can also reduce the image sent to detection. Therefore, different algorithms or methods should overcome these four problems and be effective and suitable for different research situations. Another problem with real-time people analytics and statistics is the huge financial expense.

Jalled and Voronkov (2016) use image processing to detect people and cars, subtract the background, and use similar models to obtain masks to identify people and vehicle equipment. Preliminary results Artificial Intelligence and Applications Vol. 00 ISSS available. 00 202202. Share advice mostly with other things. Therefore, in order to identify the human product through classification, the face detection from the human product is sent to the classification necklace, and the product must pass all the distributed ones to get good results. As a test the neck classifier is only suitable for face recognition, when car recognition performance is inherently low, when tracking of faces or human objects is done with simple consistency of the face (such as the object does not change color turns or moves to complete).

3. PROPOSED WORK

The use of models for monitoring and verification is necessary and our models have been successfully applied. CNN is the main algorithm used and is trained using generated data. The file contains text appropriate for human and nonhuman images. In order to increase the diversity of the training data and prevent distortion, data improvement techniques such as rotation, translation and scaling are applied to the images. Stochastic Gradient Descent (SGD), Adam or RM Sprop optimisers can be used to reduce losses and adjust weights in the network. CNNs are trained iteratively on many images. During each iteration (period), the network adjusts its weights based on the calculated slope from failure. The training process continues until the model is connected and the loss is reduced to a minimum. Once the CNN is trained, it can be used to make predictions about new images. The input image is passed through CNN training, and the output of the output layer represents the probability that the input image contains humans. A threshold can be applied to the output probability to determine whether it contains humans. By using CNN architecture with appropriate techniques and training, we can capture people with good and accurate images.

4. WORKING OF MODEL

The YOLOv3 model is a realtime object detection algorithm that uses deep neural networks to identify objects in images. It splits the image into a grid of cells and estimates bounding boxes and confidence scores for each cell. The checkboxes are then filtered by threshold and not by maximum value to get the final result.The physical tracker is based on the Kalman filter, which uses realtime measurements, including noise and other errors, and produces estimates of unknown variables that are often difficult to achieve based on measurements alone. Predictions are more accurate.The Kalman filter uses state variable models and parameters to predict and adjust the visibility of human activity. The recogniser is based on the Pedestron model, a deep learning model designed to detect pedestrians in various situations.
It uses truth and the workings of many combinations to give each person a unique identity based on their characteristics.
YOLOv3 is known for its speed and accuracy, especially in realtime applications. It is designed to work in real time, making it one of the fastest tools for finding patterns. YOLOv3 is faster overall because it uses the same neural network for all images, predicting bounding boxes and class capabilities in a single measure. This is in contrast to RCNN and its variants (Fast RCNN, Faster RCNN), which use a twostep process of first recommending candidates in the region and then splitting them, which can take more time. and its variants are a set of models that focus on object detection in the field of computer vision.

This model has been highly successful at detecting objects and forms the basis of many modern computer vision systems. They are particularly famous for their accuracy and efficiency in identifying and locating various objects in images.

5. METHODOLOGY

5.1. Overall system architecture:

The planning process has three census elements: the number of people leaving the store, entering the store and going to the store, the number of people crossing the attack line to treat it as just one attack line a way to go up and down and vice versa to be pulled in. The second feature is the alert feature, which is also used to warn the user when the crowd exceeds a predetermined threshold. It includes special features like setting up auctions for stores, saving recurring videos for future searches and repeats, choosing traditional searches or quick discovery models to achieve speed, exposure to tradeoffs, and more. Finally, all these features are collected in a single image. The user interface (GUI) is located in the middle of the platform and has three builtin tabs such as the main application for setting the threshold, search and return for improving the threshold, and knowledge fair, as shown in figures 1, 2 and 3. The system has a cascading architecture human objects are tested first and then tracked in the framework. It
is a good idea to search for many products and calculations and in addition to study other people's products in the framework
and give more details to discover what is required. This needs to have two main components, including searchers for search and distribution, and trackers for surface analysis of detected objects; These will be explained in the following words.

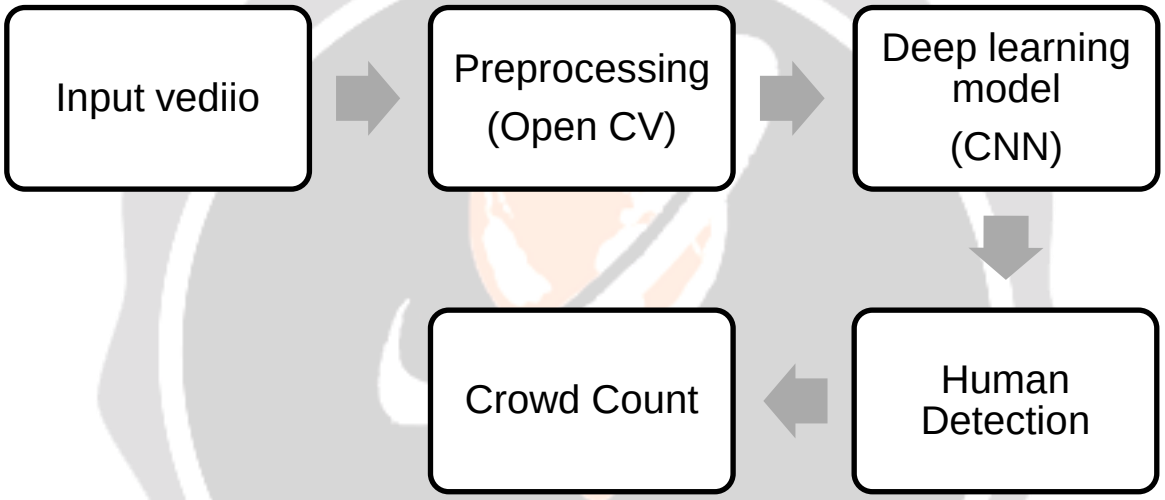
5.2. Human object detection and classification:

The first detail is to find the product by filtering the product classification and return it to the human body device to track it in the most convenient way. For immediate processing, YOLOv3, a deep learning neural network algorithm that can perform detection and classification simultaneously, was selected.
YOLOv3 model uses excels from the augmented and wellknown COCO dataset, with target weight and upper value set to 416. This configuration provides best accuracy of 55 in its class and performance study. Our average overall performance is 35. 2. Compared to

other highspeed searches using the official website model (FPS) system (Redmon, 2021). However, due to low hardware and the addition of deep SORT tracking algorithms, the test tool could not fulfils the FPS request. Therefore, other YOLOv3small models with the same target width and height will be available as search module options, allowing the user to choose between speed and accuracy. Actually, another highspeed version called Mobile Net SSD was also tested, but its speed and accuracy are lower than the YOLOv3-small version.

5.3. Human object tracking:

Another important aspect of tracking all human objects involves the use of countermeasures to determine the movement of objects from a particular direction. In general, terms, product recognition is a method that gives unique features to human products, and if it occurs throughout the frame, it continues until the product disappears. Additionally, when an object disappears from detection and reappears, there is no way to know whether it is the same object that appeared in the previous frame or a completely different object, and there is no way to compare the object's current motion with the previous frame. The previous motion of the object. . Therefore, tracking is necessary for computer vision, including the proposed process for real-time processing of video data. Deep SORT, an extension of the Simple Online Real-time Monitoring (SORT) algorithm, was chosen as the system monitor.



6. CONCLUSION

In summary, automation of the proposed system has been proven to be a more efficient, effective and accurate solution for congestion control compared to the manual system. model. It also achieves a higher accuracy of 91.07% compared to the ordinary YOLOv3 model. Deep learning YOLOv3 and Deep SORT algorithms overcome challenges in human identification and computation, such as different cameras, crowds, lighting, and human personal and computational hindrances. Ode's intuitive and innovative GUI for configuration, operation, search and playback. Additionally, using GPUs for computer vision tasks and first converting the YOLOv3 model to Tensor Flow format improved performance, showing significant results in terms of initial load and profile. Additionally, Nvidia GPU devices with Python applications currently only support the Google Tensor Flow format.We know

1.Be clear and concise:

A brief summary is helpful when the topic is general. Try to make everything as descriptive as possible to avoid losing important information.

2. Highlight innovation:

Highlight how your model differs from existing approaches. Will it lead to innovation or development?

3. Results and Implications:

Include a summary of the results or performance indicators of the application of the model. How do they compare to existing methods in terms of accuracy, speed, and other metrics? Also discuss the potential impact of your model on real-world applications.

In general, abstracts provide a framework that, with a little editing, can effectively communicate the importance and novelty of your work.

7. REFERENCE

1.[2] Arteta C, Lempitsky V, Noble J A and Zisserman A 2014 Interactive object counting Springer Cham pp 504-518.

2.[3] Brostow G and Cipolla R. 2006. Unsupervised Bayesian Detection of Independent Motion in Crowds University of Cambridge Vol 1.

1. [4] Chan A B, Liang Z S J and Vasconcelos N 2008 Privacy preserving crowd monitoring: Counting people without people models or tracking San Diego pp. 1-7.

2. [5] Change L C, Gong S, and Xiang T 2013. From semi- supervised to transfer counting of crowds. London, United Kingdom pp 2256-226.

1.[7] Faisal E, Sleit A, and Alsayyed R.2018. Crowd Counting Mapping to Make a Decision University of Jordan Amman, Jordan Vol 9.

2.[8] Gao G, Gao J and Liu Q 2020. CNN-based Density Estimation and Crowd Counting: A Survey.

3.[9] Good fellow I, Bengio, Y and Courville A .2016. Deep learning MIT Press.

4.[10] Idrees H, Saleemi I, Seibert C and Shah M. 2013. Multi- source multi-scale counting in extremely dense crowd images University of Central Florida pp 2547-2554.

5.[11] Kong D, Gray D and Tao H. 2006. A viewpoint invariant approach for crowd counting Santa Cruz, California vol 3 pp 1187-1190.

6.[12] Krausz B and Bauckhage C. 2011. Automatic detection of dangerous motion behaviour in human crowds, in Advanced Video and Signal-Based Surveillance (AVSS) pp 224-229.