International Journal of Scientific Research and Engineering Development-– Volume X Issue X, Year
Available at www.ijsred.com

RESEARCH ARTICLE OPEN ACCESS

Survey Paper on People Counting System

Aabshaar Shaikh*, Rohit Ramamoorthy**, Shweta Patil***, Maheshwari Biradar****
*(Computer Science and Engineering, D Y Patil International University, Akurdi, Pune
Email: 20200812016@dypiu.ac.in)
** (Computer Science and Engineering, D Y Patil International University, Akurdi, Pune
Email: 20190802071@dypiu.ac.in
***(Computer Science and Engineering, D Y Patil International University, Akurdi, Pune
Email: 20200812028@dypiu.ac.in)
****(Computer Science and Engineering, D Y Patil International University, Akurdi, Pune
Email: maheshwari.biradar@dypiu.ac.in)

--***********************--

Abstract-
The paper presents an investigation into the use of real-time people counting systems for gathering accurate and reliable data on human traffic in public spaces. Hardware and software components are used in the implementation of the proposed task. Cameras and sensors are examples of hardware elements. Different algorithms are used to execute the various approaches to data processing and analysis. The potential uses of real-time people-counting systems are also covered in this paper, including investigating the effects of redesigning public spaces and examining the effects of public health initiatives on the flow of pedestrian traffic. The paper illustrates the usefulness and efficiency of real-time people counting systems in gathering information on human movement in public spaces through an analysis of the literature and case studies. The paper emphasizes how real-time people counting systems have the potential to be an effective tool for researchers who need precise and trustworthy data on human traffic in public places. It ends by outlining potential directions for future study in this field, such as the creation of more sophisticated systems that can deliver even more comprehensive and nuanced information on human movement in real time.
Keywords - AI, FNNS, CNN, YOLO, BR, FBEM.

--***********************--

INTRODUCTION

A real-time people counting system automatically counts the population in a specific area in real-time. This kind of system is frequently used in studies that need precise and trustworthy information on the number of people moving through public areas like malls, railway stations, airports, and parks. Real-time people counting systems typically use a hardware and software combination to find and follow the human movement. Typically, the hardware consists of strategically placed cameras and sensors in the area under surveillance. The software component consists of computer algorithms that process the information gathered by the hardware and deliver real-time population counts. Real-time people counting methods come in a wide variety, each with unique advantages and limitations. While some systems use infrared sensors to detect body heat, others use video analytics to detect and monitor human movement. While some systems depend on machine learning algorithms to gain accuracy over time, others estimate the local population using straightforward statistical models. There are numerous possible uses for real-time population counting systems in research studies. These systems can be used by researchers, for instance, to examine the efficacy of public health initiatives like social exclusion and mask use. They can also be used to examine how changes in public area design affect the movement of pedestrians. All things considered, real-time people counting tools are an effective tool for researchers who require precise and trustworthy information on human traffic in public areas. We can anticipate seeing even more sophisticated systems that can provide in-the-moment data on human movement as technology continues to progress.

2. EXISTING MODELS

2.1 Feedforward Neural Network (FNNS)

The weights between the final hidden layer and the output layer are assessed using the linear least-squares approach in the suggested algorithm.A modified gradient-based training approach is then used to evaluate the weights between the input and hidden layers. Using a pure least-squares algorithm, the weights between the output layer and the final hidden layer are assessed. To address the stalling issue that the pure linear least-squares-based training method has, a hybrid training algorithm was created as per [1]. The stalling issue is resolved by this method, which does away with the challenge of selecting suitable transformation matrices in a pure linear least-squares-based algorithm.

2.2 EMPIRICAL METHOD

The job of face detection was carried out using a face detector based on a support vector machine (SVM). They then used the suggested filtering method to eliminate false hits picked up by the face detector mentioned earlier. In order to eliminate any candidates they noticed that appeared and vanished unexpectedly, theyadopted a temporal filter. Their percentage of falsepositive detection, however, is between 10 - 20%.Template matching algorithms have the ability to learn online appearance as per [2]. For online instruction in the suggested technique, a collection of templates for each viewer must be gathered. If a template's spatiotemporal characteristics resemble those of an earlier template, it might be considered to be sufficiently dependable. The online training starts counting as soon as the total number of gathered templates exceeds a predetermined threshold.

2.3 VISION-BASED PEOPLE COUNTING SYSTEM

This method includes a module for face detection, a module for face tracking, and a module for counting. Additionally, the synchronization results can show when new faces enter the video, when faces briefly vanish due to occlusion, and when faces leave the scene. They enhanced the original Kalman filter to track objects more precisely in this circumstance in [3]. They can keep following the obscured faces until they reappear within a few frames by using the Kalman filter's predicted face position. The face trajectories are sent to the counting module for additional processing after they have left.

2.4 VIOLA-JONES DETECTOR

Frontal human faces were identified in video sequences using the Viola-Jones detector. In the suggested approach [4], a new ID is assigned if the face is discovered for the first time in each linked component. If a face has already been detected, the presently detected face is given the previously detected face's ID. They established a threshold value, and if the number of detected faces occurring in the same connected component exceeds it, faces are counted.

2.5 FASTER R-CNN

The model counts the total number of humanheads while also identifying them using the R- CNN object reconstitution technique as per [5]. The best approach is to employ a graphical user interface (GUI). While using this, users can contribute both static and dynamic information. The faster R-CNN object reconstitution system to count the number of human heads and recognize them in the plural. In their model, the Resnet-50 network is split into two sections: layers from conv1 to conv4x, and layers from conv5x and above, and it also removes the final regression properties, classification recommended.

2.6 CASCADE CLASSIFIER

In order to identify people in a video frame, the Cascade Classifier looks for human faces with roughly constant facial aspect ratios.

An efficient object detection technique uses Haar feature-based cascade classifiers in [6]. Briefly said a cascade function is taught using a large number of positive (Pictures with faces in them) and negative images (Pictures without faces in them) in a machine-learning approach called a cascade classifier. The classifier in the cascade that is used to recognize objects in other photos will use information gathered from the outputfrom a specific classifier as extra information.

2.7 HOG TECHNIQUES

One of the best human detectors for static images is the HOG descriptor. Their technique simulates a detection window as a dense grid of HOG as per [6]. The HOG descriptor has been able to attain a real-time effect on human detection with the advancement of the hardware. An object in an image or video frame can be categorized as an individual based on these descriptors. The HOG approach is based on analyzing local histograms of image gradient orientations in a dense grid that has been well-normalized.

2.8 COUNTING ALGORITHM USING BLOB TRACKING METHOD

They used a hierarchical tracking framework, with the first level being an effective blob tracking technique. Two tracker containers are present in the system of [7]. The tracker container is the one that houses the trackers that keep tabs on those who are currently crossing the entry/exit line. The other container, known as the previously-counted tracker container, is for the trackers that maintain the color distribution of the already-counted individuals. A merge or split has occurred when a tracker is unable to locate its target, and only in this situation is the mean shift tracker triggered to resolve it.

2.9 HARRIS CORNER DETECTION

On the basis of video image processing and feature point extraction of optical flow, a method for counting the number of people is suggested. Each image's corners are found using the Harris corner detection algorithm, and the corners are then tracked using the LK optical flow method [8]. Setting a threshold allows for the person to count. The movement of people creates a continuous optical flow for the flow of people.

2.10 ALBIOL ET AL'S ALGORITHM

This model is based on utilizing SURF features. Compared to those provided by the Harris detector, the interest sites discovered by SURF are far more scale-independent and consequently distant from the camera. The steps in this method [9] include: Finding the interest point of the people, clustering the points, then implementing feature extraction and regression in it. Finally, they used a low-pass filter to compare the number of individuals across successive frames. In particular, the total number of people in the scene is determined by averaging the people counts on the last k frames.

2.11 MINKOWSKI FRACTAL DIMENSION

The input image is first subjected to edge detection, after which a binary image is produced. Then, n dilations between 1 and n are calculated. The image is finally divided into 5 classes of density, namely very low, low, moderate, high, and very high, using the fractal dimension as a characteristic. These classifications are regarded as Polus's suggestions. Polus has five different service flow levels: free, restricted, congested, very dense, and jammed. Each level is determined by the count of pedestrians present in the respective location as per [10].

2.12 CASCADE CLASSIFIER WITH DETECT MULTI SCALE USING OPENCV

The detectMultiScale technique in [11] is used to count the number of pupils in a class. The person in the image will be recognized using the detectMultiScale method; any items will be ignored.

2.13 YOLO ALGORITHM

YOLO stands for the phrase "You Only Look Once." Convolutional neural networks (CNN) are used in YOLO calculations to gradually distinguish between items. The following three methods are used by the YOLO algorithm [12] to function: Residual Blocks, Bounding Box Regression, and Intersection over Union (IOU). Their algorithm includes real-time video footage that is transformed into picture frames that are time-ordered in a sequence. An image is transformed into a N X N grid arrangement. filtering out people in each frame, then using bounding box regression to quantify them.

2.14 HEAD DETECTION AND TRACKING METHOD

Their model begins by gathering frame images from the input videos. The foreground extraction process is then carried out to extract areas of moving objects from frame pictures. A trained LBP feature-based Adaboost classifier is used to recognize human heads. Then a straight counting line is drawn, which could be either horizontal or vertical. The heads that pass the counting line are finally counted using the positions of the human heads.

2.15 BI-DIRECTIONAL PEOPLE COUNTER USING HSI HISTOGRAM

In the Proposed method, a single color video camera is mounted straight down on the gate's ceiling [14]. When a merge-split scenario is encountered, it can enable information for tracking and differentiating between multi-person and single-person patterns. Here, it is believed that the people-image pattern region can offer a reliable initial estimate of the population. A histogram can estimate the likelihood that a certain pixel value will be shown in an image. As a result, it is possible to distinguish one pedestrian from another using the histogram of color or intensity. Instead of using color coding (which could be problematic if an organization has a uniform), they have used bounding boxes to monitor each individual.

2.16 ASSUMPTION UNIVERSITY'S RASPBERRY PI CUSTOMER COUNTER (AU -PICC)

Computer vision is the foundation of the method used for people counting. To make a small, standalone device for counting people, they then implemented OpenCV in an embedded system of Raspberry Pi. This model was implemented to find the number of people interested in a product in a retail store. The main function of AU-PiCC [15] is to measure the number of individuals who are interested in a target product in a pre-defined area while using straightforward face identification to prevent duplication. This design could even be used as a straightforward, stand-alone people counter.

2.17 USING ANN WITH C++ AND OPENCV

This technique uses two ANNs to classify individuals, a single camera mounted over the access gate, and numerous image processing building blocks [16]. After the ROI is determined, an analysis is conducted to determine whether a person entered or exited the bus. The system makes use of the OpenCV and C++ tools. The result is provided by the second network, after which a zone of interest is created for counting. The algorithm then determines whether the number of passengers exceeds the vehicle's limit by counting how many people have entered and exited the vehicle. Instead of using barrier techniques like laser systems, which can easily cause injury or are very uncomfortable for people with disabilities, they opted to use artificial vision.

2.18 MIXTURE OF GAUSSIAN (MOG) AND BACKGROUND SUBTRACTION METHOD

The technique was developed using an OpenCV library in Python and a WiFi feed from an overhead IP (Internet Protocol) CCTV camera in [17]. For effective operation, the suggestedmethod also makes use of Background Subtraction, Blob Tracking, and Area contour.Also to remove noises and obtain appropriate blobs, morphological operations were performed on each frame of the live video feed as it was processed frame by frame. In order to count the blobs within a region of the frame, they then track each blob's coordinates using the contours technique. The tracking was completed using a Kalman filter, and the detection was carried out by correlating the heads of individuals in real-time from feed and preprocessed images in a data collection.

2.19 USING BAYESIAN REGRESSION

2.19.1 GAUSSIAN PROCESS REGRESSION (GPR)

A supervised learning technique called the Gaussian Process (GP) is used to resolve regression and probabilistic categorization issues. In the field of machine learning, GPR [18] is a nonparametric, Bayesian method to regression that is causing a stir. GPR has several advantages, working well on small datasets and having the ability to provide uncertainty measurements on the predictions.

2.19.2 BAYESIAN POISSON REGRESSION (BPR)

Poisson Regression models [18] work best when modeling situations where the results are tallies. Or, more specifically, count data, which are discrete data with non-negative integer values that measure something like the number of shoppers in a grocery line or the frequency with which an event happens over a specified period of time. The standard Poisson regression or NB regression is produced by modeling the outcome variable as Poisson or negative binomial (NB), with an arrival-rate parameter that is a function of theinput variables.

2.20 USING KALMAN FILTER

A filter subsequently uses the edge foreground image created during preprocessing to look for individuals using head detection. For more reliable identification, head detection uses data from the Kalman filter. They are then used for monitoring and counting. Here, they demonstrate the conventional method of counting using two virtual lines, one for entering and one for exiting. The in-out counts are updated as a person passes through the region.

2.21 USING PCA

The people-flow method for PCA [20] has been extensively used in feature extraction, classification, and dimensionality reduction of datasets. The eigenvectors of matrix A, which describe the central idea, are calculated. These eigenvectors can then be effectively used to fill the column space of matrix A. A collection of eigenvectors can be precisely calculated if they have a large number of composited face matrices. The initial linear arrangement variable is discovered by PCA.

2.22 USING A VISION SENSOR

The method is based on a low-power vision sensor. It dispatches the asserted pixels using an address-driven data representation, implements motion by comparing two succeeding frames, and extracts the contrast of the picture in binary form. Based on the virtual loop, a simple algorithm for counting individuals has been created. The vision sensor [21] is interfaced with the FPGA in the demonstrator, which is connected to a Computer with a GUI. Using the virtual loop algorithm they counted the number of people entering and leaving a location under observation.

2.23 FOREGROUND / BACKGROUND EDGE MODEL(FBEM)

The procedure builds a background model and then uses a difference operation between the current frame and the background model to retrieve the foreground picture. In order to extract edge information from each frame in video recordings, a clever algorithm is used. They employed two lines for entry and exit in order to accomplish bi-directional counting. They only used the minimal distance technique in their experiments because it works well. The people counting procedure doesn't start until the people center region moves into the space between these two lines as per [22]. The space between these two lines was known as the Region of Interest (ROI). As per the flow of the people, the count of entry and exit of the people is estimated.

2.24 USING SVM CLASSIFIER

To extract the foreground area in their method, adaptive background modeling is used. Through a series of frames, each individual is tracked using the Kalman filter and cost function. The foreground area is used to train the linear SVM classifier to look for the head-shoulder component in [23]. To count individuals arriving or leaving the scene, the resulting trajectories are examined.

2.25 BASED ON DENSE STEREOVISION

Stereo vision is a technique that involves the analysis of multiple images, typically two of the same object captured at various angles, along the camera's optical axis, or by tilting the acquisition system. A stereo-matching block that computes the disparity map for each set of images.Two binocular or three trinocular stereoscopic pictures are used in passive stereo vision. Then a segmentation block that identifies, in the height map, the heads of individuals by detecting round shapes with a constant height value. This module counts people when their head's trajectory approaches or exits the stereo field of view as per [24].

2.26 USING KLT - FEATURE TRACKER

Instead of using a background model, the system uses a KLT tracker to identify motion within the region of interest (ROI) and incorporates task-driven data for trajectory validation [25]. This method's stages are as follows: Temporal Features Based Trajectory Clustering, Feature Tracking and Candidate Creation, and Trajectory Validation. A novel appearance and disappearance time clustering of the feature point trajectory is used to accomplish the counting.

2.27 BASED ON VERTICAL KINECT SENSOR

They built a system that counts individuals using the vertical Kinect sensor [26], and they use depth information to eliminate the impact of appearance variation. Since the head is always nearer the Kinect sensor than other body parts, finding the appropriate local minimum regions is equivalent to finding the number of individuals. They suggested an unsupervised water-filling method that can identify these regions with the properties of robustness, locality, and scale-invariance in accordance with the specificity of the depth map.

2.28 USING A NETWORK OF SIMPLE SENSORS

A planar projection of the scene's visual hull is computed in their system by groupings of image sensors [27], aggregating the resulting silhouettes over a network, and segmenting foreground objects from the backdrop. Only a small amount of data is transmitted over the network, and the computational requirements scale well with the number of sensors and users. After removing phantom regions, they then introduce a geometric algorithm that determines bounds on the population in each projection region.

2.29 USING HOUGH TRANSFORM

The method makes use of image processing to estimate passenger crowd flow in real-time within a bus with a complex background. The fuzzy measure and the modified Hough Transform [28] are used to accurately obtain the contour feature. The position of each passenger is determined by utilizing the head's contour feature.

2.30 USING K-MEANS BASED SEGMENTATION

This method uses a zenithal camera approach that is mounted overhead. K-means clustering is used after the initial block-wise background subtraction to allow the segmentation of individual people in the scene as per [29]. The maximum number of clusters with a tolerable inter-cluster separation is calculated as the total number of individuals in the scene.

2.31 USING A STEREO CAMERA

The stereo camera is hung from the gate's ceiling in the proposed method [30], and its optical axis is positioned so that it can see people walking by directly overhead. Additionally, the stereo camera allows for precise segmentation of the road and human regions on the obtained images. The data of the moving people are not overlapped on the obtained images in this system setup if there is a large crowd at the entrance. The data area is thick along the timeline if the passer moves slowly, or thin if they move quickly. As a result, the passing speed, which is determined by template matching, is taken into account when counting the individuals who are passing.

3. ADVANTAGES AND LIMITATIONS OF MODELS

Advantages:

Compared to more conventional techniques, real-time people counting systems have a number of benefits for gathering information on pedestrian traffic in public areas. The methods described above have the following benefits:

According to the findings in [1], substantially fewer flops are needed for convergence in FNNS than are needed by other well-known algorithms.

As per Empirical Method [2] the majority of existing approaches must take care of the pose change issue before completing template matching, i.e., they must pre-define a set of thresholds to guarantee the feature set's proper functionality. Finding a feature set that can accommodate such modifications, though, is incredibly challenging.

It should be observed that in this vision-based people counting system [3], in comparison to the original, the inaccuracy of their Kalman filter falls more when face scales grow. In other words, when face scales increase, their Kalman filter performs more accurately.Their model overcame face occlusion and continued to track their face until it appeared.

As per the faster R-CNN model[5], while classification time has decreased, detection accuracy has increased.

The primary benefit of the Harris Corner Detector is that it is the best option for retrieving image data and stereo matching since it is invariant to factors like rotation, translation, and variations in light. In the presence of noise, it can reliably and precisely identify corners. Among other established corner identification methods, it is both the most repetitious and the most informative type of corner identification method [8].

The suggested strategy by Albiol et al [9] achieves good performance and has proven to be quite resilient with regard to parameter selection.

Coelho has examined the consequences of the restricted fractality of real things and put forth useful recommendations that could produce more accurate and insightful experimental findings than Minkowski's Fractional Dimension[10].

Using the Cascade Classifier with the detectMultiScale() method within the OpenCV module is mostly responsible for this improvement in accuracy. Expect the error rate of their counting approach to decline as more OpenCV updates are published [11].

The main advantages of YOLOv3-tiny are that the network is simple, the calculation is small, and it can run on the mobile terminal or the device side [12].

The approach used by the head detection and tracking method [13] effectively addresses the overlap issue brought on by head-shoulder identification of moving subjects and performs well in densely populated areas. Foreground extraction and the LBP feature-based Adaboost classifier helped reduce the false detection rate.

Additionally, with the exception of abrupt moving cases, patterns involving one or two individuals can also be counted with 100% accuracy. It should be noted that if there is no merge split situation to occur, 100% accuracy is unquestionably achieved [14].

The AU PiCC method's [15] limit will be the minimum distance of 30 cm, which will ensure an accuracy of 90%, but it is well suited for the job of counting customers who are genuinely interested in a product rather than passersby.

The benefit of using ANN is that it can tell whether someone gets on or off the bus even though the background subtraction stage produces a noisy image. Another advantage of this system [16] is that it is much more mobile-friendly than conventional methods, making it much more accessible to disabled users.

Each pixel has a distinct "threshold" chosen for it in the MOG and background subtraction methods [17]. These pixel-level "thresholds" change over time. Without erasing the current backdrop model, objects are permitted to blend into the background. Offers quick recovery.

When comparing the two Bayesian regression techniques [18], it was discovered that BPR was more accurate for denser crowds while GPR did better for less dense crowds (in which case, the regression mapping is more linear).

While using Kalman's filter the one crucial element is the feedback between the tracking and detection phases, which enables the development of a more reliable algorithm and addresses potential temporal errors and partial occlusions in real-image sequences like the test films used in [19].

PCA [20] has a few pros it reduces overfitting and also improves data visualization.

The model implemented using vision sensors [21] has an accuracy of 77% outdoors and 91% indoors. To substitute expensive physical sensors and reduce installation and maintenance costs, it is a cost-effective software option.

The FBEM [22] results show that the proposed method can reach a very high level of accuracy and be used for real-time tasks.

To quickly identify people, the system in [23] adheres to Dalal's HOG - SVM - Bootstrapping framework and makes use of the discriminative power of HOG features. Their experimental findings showed that the effective detection rate varied between 85% and 100%.

The comparison between physically counted values and the ones calculated with their algorithm of dense stereovision [24] leads to a counting accuracy that is around 99% for laboratory and 97% for bus data sets.

They believe that the KLT feature system [25] can manage more difficult conditions in a traffic bus in real-time than the single-CCD camera-based passenger counting system that is currently in use.

The people detector using a vertical Kinect sensor [26] is highly accurate and can be used for tracking algorithms for people counting that are reasonably easy to implement.

These characteristics allow their system [27] to operate in real time and to be set up as an untethered wireless sensor network. Implementing a decentralized communication architecture suitable for a much larger sensor network is the next step. As a result, network traffic would be reduced even more, and robustness and scalability would increase.

As it is largely unaffected by image noise, the Hough Transform [28] can be very helpful when trying to detect lines with brief noise-induced breaks in them or when objects are partially obscured.

The suggested K-means method [29] for counting people gets real-time performance while being straightforward and effective. The people segmentation algorithm performs very well despite being straightforward. The high precision and recall values of the counting process, which result in an F-score of 0.97, are greatly influenced by the correct segmentation rate of nearly 95% of all non-empty frames.

The accuracy of proposed model using a stereo camera [30] has an accuracy of above 90%. Stereo cameras see more information, which results in fewer occlusions. This in turn lowers the proposed model's error rate.

Overall, real-time people-counting systems provide researchers with a technique that is more precise, effective, scalable, and non-intrusive for collecting information on pedestrian activity in public areas. They serve as a potent instrument for guiding policies, design selections, and operational plans that advance security, effectiveness, and well-being in a variety of contexts. Given their thorough processing chain, they have a decent chance of maintaining an error rate of less than 2%.

Limitations:

While real-time people counting systems have several benefits over more conventional approaches for gathering information on pedestrian traffic in public areas, they also have some drawbacks. The following are a few drawbacks and restrictions of the methods listed above:

As per the Empirical Method [2] the people who were in the frame for short intervals were considered as noise and eliminated. This would not give an output of high accuracy.

Due to movement and head distortion, themajority of the errors in the viola - jones detector model [4] involved counting people twice.

In the case of [6] for both Cascade Classifier and HOG the accuracy is of moderate levels (around 56.87% and 59.04% respectively).

An ellipse can be used to accurately represent the overhead perspective of moving persons, and when compared to morphological processes, this approach allows for the successful filling of all holes and missing pieces, regardless of their location. This is a preprocessing phase, so it is not desirable to use a different, computationally expensive strategy for real-time performance. Moreover, BGS can make mistakes when a person is dressed in clothing that is close to the color of the floor and when a large FG blob is broken into smaller blobs as per [7].

Harris corner detection has one significant flaw in that we must set various threshold values for each image to find the most pronounced interest locations [8].

The Albiol et al approach only gives a reliable estimate when the people are at an average distance from the camera. When people are close to or distant from the camera, it tends to overestimate or underestimate the number of people. The fact that Albiol's method does not account for the density of the identified interest sites when estimating the population is a second flaw in it [9].

By using a better method to estimate the fractal dimension, the results might be improved due to the crowd photos' low fractality in comparison to Minkowski Fractional Dimension [10].

The drawbacks of Haar cascades include their propensity for false-positive detections, the need for parameter adjustment when used for inference or detection, and their general lack of accuracy in comparison to more "contemporary" techniques available today [11].

The YOLOv3-tiny performance is influenced by several variables. The first consideration is the video's image quality. It is evident that while digital color photos display respectable performance, black-and-white images do not. The camera's position is the following factor. The camera should be positioned so that it can see clearly without being obstructed by anything else. And the accuracy is relatively low [12].

The merge-split phenomenon frequently occurs as individuals pass through a gate in the majority of real-world scenarios, so the accuracy is not 100% in real-world scenarios. Some fast-moving circumstances will cause the counting accuracy to be slightly reduced as the number of people contained in a moving pattern rises above two, particularly for a mix of fast bi-directional merge-split cases.

In AU - PiCC [15] has cons such as lighting conditions, headgear, processing time, camera shutter speed, and the number of customers.

If the ANN is not properly trained, the system [16] may not be able to perform the count correctly or have a decent classification.

The MOG and background subtraction method [17] cannot handle abrupt or extreme lighting shifts. The Gaussians must be initialized (median filtering). There are a lot of factors, and you should pick them carefully. It also uses too much CPU power, which prevents it from being used for real-time work.

The requirement for training for each distinct perspective in Bayesian regression [18] is one of its limitations for crowd counting. This limitation is appropriate for systems that are always watching. The need for training, however, might make it more difficult to rapidly implement a crowd-counting system.

With more individuals present, Kalman's Filter [19] effectiveness starts to slightly decline. The test films' complete occlusions are primarily to blame for this. Also, poor pattern-head coincidence can lead to some false negatives.

PCA [20] has a few limitations such as independent variables becoming less interpretable, data standardization is a must before implementing PCA.

Errors occur in the FBEM [22] when the foreground edge is divided into smaller curves and when the individual is wearing clothing that is similar in hue to the background.

The severe mutual occlusion was to blame for the majority of miss detections [23].

The use of color images was avoided because they would require additional processing time and were avoided because they desired a real-time counting system from the start. However, the use of color would provide improvements in the choice of homologous pixels for the stereo-matching [24] procedure because we have more information for neighborhood comparison.

The algorithm's restriction, however, states that the water filling [26] cannot manage a scenario in which a moving object is closer to the vertical Kinect sensor than the head, such as when hands are raised above the head.

The system's [27] sensitivity to noise at the silhouette intersection is a flaw. An undercount may result from noisy silhouettes that grossly overestimate the size of the objects.

They continue to encounter some issues with hough transform [28]. The main issue they faced was dark hair with dark clothing or light hair with light backgrounds made it difficult to extract the head's contour; as a result, edge information is insufficient to build a passenger's head's contour.

The k-means model [29] is not optimized, and simultaneous tracking and segmentation, mean-shift tracking, and multidimensional dynamic cluster assignment could all be used to further enhance the algorithm's performance.

Real-time people counting systems have many benefits for collecting information on pedestrian behavior in public areas, but they also have drawbacks that researchers should take into account when choosing a data collection technique. In order to decide whether these systems are suitable for their study objectives and ethical considerations, researchers must weigh the advantages and limitations of using them.

4. APPLICATIONS

Real-time people counting systems have a wide range of potential applications in research studies. Here are a few examples.
Determining the success of initiatives in public health. The effect of public health interventions like social isolation, mask use, and hand hygiene on pedestrian traffic flow can be studied using real-time people-counting devices. Researchers can learn more about the efficacy of these interventions by comparing statistics onpedestrian behavior before and after their implementation. It might also be applied to studying the results of modifications to public areadesign. To assess how changes in the design of public spaces will affect the flow of pedestrian traffic, real-time people-counting devices can be used. For instance, researchers can learn moreabout the impacts of new benches, lighting, or signage on pedestrian flow and safety by comparing statistics on pedestrian behavior before and after their installation. The planning and improvement of public transportation networkscould also use it. At transportation hubs like railway stations and airports, statistics on the flowof pedestrian traffic can be collected in real-time using people counting systems. Researchers can use this data to analyze patterns in pedestrian activity and use this knowledge to improve the layout and functionality of public transportation systems. The suggested FNNS algorithm in [1] was used to implement a similar model for crowdestimation at underground stations, and promising results were seen. Enhancing security and safety for the people. By identifying and reacting to unusual patterns in pedestrian behavior, real-time people counting systems can improve public safety and security. Security personnel can beinformed to examine, for instance, if a sudden increase in pedestrian traffic is observed in a specific location. Post-secondary educational institutions want to include people counts at events, as described by the author in [4], but it could also be used to estimate stadium capacity and plan evacuation strategies for crises like fires and weather-related emergencies. Enhancing commercial and retail activities. To collect information on customer traffic movement in retail and commercial spaces, real-time people-counting devices can be used. Business owners can improve their store layout, product placement,

and other parts of their operations by analyzing this data to gain insights into customer behavior. In [2], the author discusses comparing the number of people who pass by a digital billboard to the number who actually read it, as this information would be used to identify the specifics of the people who are interested in the merchandise. In general, real-time people counting tools are effective tools for researchers who need precise and trustworthy information on human traffic in public areas. They may help to guide policies, design choices, and practical plans that advance security, effectiveness, and well-being in a variety of contexts.

5. CONCLUSION

Using real-time people counting systems in research projects offers a creative and practical method for obtaining precise and trustworthy information on human traffic in public areas. Real-time data on pedestrian movement is made available to researchers through the combination of hardware, such as cameras, sensors, and software algorithms. Real-time people counting systems are useful and effective in a variety of settings, such as malls, train stations, and parks. These systems give researchers the chance to understand pedestrian behavior, examine the results of public health initiatives, and contribute to the design of public spaces. However, there are drawbacks to using real-time people counting systems, including problems with accuracy and privacy. Future studies in this field should concentrate on overcoming these difficulties and creating more sophisticated systems that can deliver real-time data on human action that is even more precise and nuanced. In general, real-time population counting devices are a promising field for study and invention. They have the potential to greatly improve our comprehension of pedestrian behavior in public areas and contribute to the development of policies and design choices that advance safety, effectiveness, and well-being.

6. REFERENCES

[1] T. W. S. Chow, J. Y.-F. Yam, S.-Y. Cho, „Fast training algorithm for feedforward neural networks: application to crowd estimation at underground stations", Artificial Intelligence in Engineering, Vol. 13, pp. 301-307, 1999.

[2] Duan-Yu Chen and Kuan-Yi Lin, A NOVEL VIEWER COUNTER FOR DIGITAL BILLBOARDS, 2009 Fifth International Conference on Intelligent Information Hiding and Multimedia Signal Processing

[3] Xi Zhao, Emmanuel Dellandr´ea and Liming Chen, A People Counting System based on Face Detection and Tracking in a Video, 2009 Advanced Video and Signal Based Surveillance

[4] H.Hakan Cetinkaya, Muammer Akcay, People Counting at Campuses, 4th WORLD CONFERENCE ON EDUCATIONAL TECHNOLOGY RESEARCHES, WCETR- 2014

[5] Saravana Kumar, Kranthi Kumar, M Vishnu Menon, K. Anirudh Reddy, B, Mahender Yadav, B. M. N. Sai Pavan, People Counting in Crowd: Faster R-CNN, 2022

[6] Poo Kuan Hoong, Ian K.T. Tan, and Chai Kai Weng, A COMPARISON OF PEOPLE COUNTING TECHNIQUES VIA VIDEO SCENE ANALYSIS, ARPN Journal of Engineering and Applied Sciences VOL. 10, NO. 23, DECEMBER 2015 ISSN 1819-6608

[7] Senem Velipasalar, Ying-Li Tian, Arun Hampapur, AUTOMATIC COUNTING OF INTERACTING PEOPLE BY USING A SINGLE UNCALIBRATED CAMERA

[8] Xia Wu, Design of Person Flow Counting and Monitoring System Based on Feature Point Extraction of Optical Flow, 2014 Fifth International Conference on Intelligent Systems Design and Engineering Applications

[9] D. Conte, P. Foggia, G. Percannella and M. Vento, A Method Based on the Indirect Approach for Counting People in Crowded Scenes, 2010 Seventh IEEE International Conference on Advanced Video and Signal Based Surveillance

[10] A. N Marana, L. da F. Costa, R. A. Lotufo, S. A. Velastin, ESTIMATING CROWD DENSITY WITH MINKOWSKI FRACTAL DIMENSION,

IEEE 1999 IEEE International Conference on Acoustics, Speech, and Signal Processing

[11] Jahanvi Mehariya, Chaitra Gupta, Niranjan Pai, Sagar Koul, Prashant Gadakh, Counting students using OpenCV and Integration with Firebase for Classroom Allocation, Proceedings of the International Conference on Electronics and Sustainable Communication Systems (ICESC 2020)

[12] Abhijith A, Adarsh Balan V V, Adarsh K V, Advait Nikesh, Priya V V, Human Detection And Counting In Visual Surveillance. 2021 IJCRT | Volume 9, Issue 10 October 2021

[13] Bin Li, Jian Zhang, Zheng Zhang, Yong Xu, A People Counting Method Based on Head Detection and Tracking, 2014

[14] Thou-Ho (Chao-Ho) Chen, Tsong-Yi Chen and Zhi-Xian Chen, An Intelligent People-Flow Counting Method for Passing Through a Gate, 2006 IEEE

[15] Tussanai Parthornatt, Natchaphon Burapanonte, Wisarute Gunjarueg, People Identification and Counting System Using Raspberry Pi, 2018

[16] Patricia Chato, Darío José Mendoza Chipantasi, Nancy Velasco, Sylvia Rea, Vicente Hallo, Patricia Constante, Image processing and artificial neural network for counting people inside public transport, 2018

[17] Kartik Madhira, Aditya Shukla, Pedestrian Flow Counter using Image Processing, International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS-2017)

[18] Antoni B. Chan, Nuno Vasconcelos, Counting People With Low-Level Features and Bayesian Regression, IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 21, NO. 4, APRIL 2012

[19] Jorge García, Alfredo Gardel, Ignacio Bravo, José Luis Lázaro, Miguel Martínez, and David Rodríguez, Directional People Counter Based on Head Tracking. IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS, VOL. 60, NO. 9, SEPTEMBER 2013

[20] Jyun-Ciang Huang, Mao-Cheng Huang, Gwo-Jiun Horng, and Gwo-Jia Jong, The Application of Face and People-Flow Recognition for Kanban Broadcast system, 2007

[21] L. Rizzon, N. Massari, M. Gottardi, L. Gasparini, A low-power people counting system based on a vision sensor working on contrast, 2009 IEEE

[22] Shengsheng Yu, Xiaoping Chen, Weiping Sun, Deping Xie, A Robust Method for Detecting and Counting People, 2008 IEEE

[23] Huazhong Xu, Pei Lv, Lei Meng, A People Counting System based on Head-shoulder Detection and Tracking in Surveillance Video, 201O International Conference On Computer Design And Applications (ICCDA 2010)

[24] Tarek Yahiaoui, Louahdi Khoudour, Cyril Meurie, Real-time passenger counting in buses using dense stereovision, Journal of Electronic Imaging 19(3), 031202 (Jul–Sep 2010)

[25] Tao Yang, Yanning Zhang ,Dapei Shao, Ying Li, Clustering method for counting passengers getting in a bus with single camera, Optical Engineering 493, 037203 (March 2010)

[26] Xucong Zhang, Junjie Yan, Shikun Feng, Zhen Lei, Dong Yi, Stan Z. Li, Water Filling: Unsupervised People Counting via Vertical Kinect Sensor, 2012 IEEE Ninth International Conference on Advanced Video and Signal-Based Surveillance

[27] Danny B. Yang, H´ector H.Gonz´alez-Ba˜nos, Leonidas J. Guibas, Counting People in Crowds with a Real-Time Network of Simple Image Sensors, Proceedings of the Ninth IEEE International Conference on Computer Vision (ICCV'03), IEEE 2003

[28] Haibin Yu, Zhiwei He, Jilin Liu, A VISION-BASED METHOD TO ESTIMATE PASSENGER FLOW IN BUS, Proceedings of 2007International Symposium on Intelligent Signal Processing and Communication Systems Nov.28- Dec.1, 2007 Xiamen, China

[29] Borislav Antic´, Dragan Letic´, Dubravko C´ulibrk and Vladimir Crnojevic, K-MEANS

BASED SEGMENTATION FOR REAL-TIME ZENITHAL PEOPLE COUNTING, [IEEE 2009 16th IEEE International Conference on Image Processing ICIP 2009 - Cairo, Egypt]

[30] E;. Terada, D. Yoshida, S. Oe, J. Yamaguchi, A Method of Counting the Passing People by Using the Stereo Images, 1999 IEEE