International Journal of Scientific Research in Computer Science, Engineering and Information Technology
ISSN : 2456-3307 (www.ijsrcseit.com)
doi : https://doi.org/10.32628/CSEIT23903104

Machine Learning-Based Crowd behavior Analysis and Forecasting

Sachin Bhardwaj*1, Apoorva Dwivedi2 , Ashutosh Pandey3, Dr. Yusuf Perwej4, Pervez Rauf Khan5

1*Assistant Professor, Department of Computer Science & Engineering, Ambalika Institute of Management & Technology, Lucknow, Uttar Pradesh, India

2Assistant Professor, Department of Computer Science & Engineering, Invertis University, Bareilly, Uttar Pradesh, India

3Assistant Professor, Department of Computer Science & Engineering, Pranveer Singh Institute of Technology (PSIT), Kanpur, Uttar Pradesh, India

4Professor, Department of Computer Science & Engineering, Ambalika Institute of Management & Technology, Lucknow, Uttar Pradesh, India

5Assistant Professor & HoD, Department of Computer Science & Engineering, Azad Institute of Engineering and Technology (AIET), Lucknow, Uttar Pradesh, India

ARTICLEINFO

Article History:

Accepted: 01 June 2023

Published: 10 June 2023

Publication Issue

Volume 9, Issue 3

May-June-2023

Page Number

418-429

ABSTRACT

In many places today, the world's overcrowding causes crowded conditions. Analysis of crowd activity is a developing field of study. It is common knowledge that mob activity can forecast what might happen during an event. Crowd management could be very effective if situations like riots, mass lynchings, traffic jams, accidents, stampedes, etc. could be predicted beforehand. In this paper, we propose a new multicolumn convolutional neural network (MCNN) based technique for predicting mob behavior. The features of the incoming image are first analyzed and extracted. The approximated number of the gathering is then established, and image cropping is completed. For each area of the image, low level characteristics are retrieved. The objects in the picture are then created as density images. Using our method, the gathered characteristics and their object density maps are then linearly mapped. At last, we forecast and quantify the population using the MCNN algorithm. For the ShanghaiTech dataset, we have evaluated our method using actual data.

Keywords: Abnormal Activity Detection, Image Pre-Processing, Crowd Analysis, Machine Learning, ShanghaiTech Dataset, Rectified Linear Unit (ReLU), Multicolumn Convolutional Neural Network (MCNN).

I. INTRODUCTION

As is common knowledge, the word "crowd" is used to describe a gathering or group of people. Since it has so many uses in video surveillance, crowd analysis and scene comprehension have received a lot of focus lately [1]. At some point, we have all seen or participated in a gathering. Therefore, it is safe to state that the crowd plays a significant role in our existence. Most of the locations we go, like markets, streets, parks, arenas, shops, etc., are constantly crowded with people [2]. This profound connection to the multitude leads us to the most crucial job, namely crowd analysis. Many current technologies of detection [3], tracking, and activity identification, which are only relevant to sparse scenes, do not perform well in congested scenes due to the large number of people gathered there with frequent and heavy occlusions [4]. As a result, numerous recent study projects have been conducted, many of which focus on crowd situations.

The process of understanding a crowd's general activity and using that knowledge to draw significant conclusions, such as estimating the crowd's size or composition, is known as crowd analysis. Numerous real-world occurrences, such as mob lynchings, traffic gridlock, uprisings, stampedes, violence, etc., can be predicted with the aid of crowd analysis [5]. We can alert the appropriate officials in advance to take preventive measures using these crucial forecasts. For instance, if we can foresee that there will be more people than usual near a crossroad, the traffic officials can make plans to avoid any traffic jams there and, as a result, any likely mishaps. An image [6] or film of the gathering is used as the input for the analysis. The collective behavior of every person in the throng makes up the mob's behavior. To comprehend the behavior of the throng, we must analyze this collective feeling. It's important to keep in mind that crowd behavior or feeling is not simply the accumulation of individual moods [7]. In a gathering, people are positioned differently and frequently move in various directions. Finding the typical behavior of a crowd and then contrasting it with the present behavior is very challenging.

Once the definition of a gathering is clear, the task of defining the area's boundaries arises. For the job of analyzing crowd behavior, there are numerous viable explanations. In general, the primary objective in this field is to be able to comprehend how a group of people act using data [8] primarily retrieved from video sources. However, there are many diverse elements of crowd activity that we can find interesting. For instance, knowing how many people are present at a particular place and how their numbers change over time can help avoid dangerous pedestrian stampedes [9]. Additionally, being aware of the major flows of a population in motion (such as when entering a sporting event) can help you spot individuals whose movements deviate from those flows and pinpoint the reasons behind them.

II. Related Work

Numerous studies on crowd analysis have been published over the past ten years. While some of them concentrate on one of its subtopics, others study the research line as a whole. The most significant evaluations we found and those that caught our attention are presented in this part. The experts in this article [10] have suggested a novel method for conducting Motion Information Images (MII) and Convolutional Neural Networks (CNN) [11] for anomalous crowd event identification. It begins by outlining the concept of feature engineering for each of crowd counting, crowd measuring, action tracking, and action identification. The study on crowd analysis by [12] provides a fascinating summary of the state of the literature as of the end of 2008. The study offers a broad overview of crowd analysis work, including crowd data and crowd behavior analysis. This study uses 2D Convolutional Neural Networks (ConvNets) to develop a unique way for identifying crowd emotions [13] and behaviors. The objective of this network is to classify [14] the typical crowd behaviors. The researchers created a collection of pictures that focused on six different emotional states, including neutral, neutral, excitement, happiness, and excitement. For crowd counting in actual scenes, Wang et al. [15] generate massive amounts of synthetic data and convert it into photorealistic photos. According to the approach suggested by [16], changes in consistency, entropy, and contrast three descriptors for the co-occurrence matrix are used to analyses crowd behavior.

Highlighting how crowded and occluded environments prevent manual approaches from accurately simulating crowd dynamics [17]. The characteristics that are employed for crowded scene analysis are mentioned in their evaluation, and they are divided into three categories: flow-based, local spatiotemporal features, and trajectory/tracklets. They divided the examined research into three categories: crowd tracking, crowd behavior analysis, and crowd counting. The objective of this study is to identify utilizing a hybrid strategy of compressed sensing and deep learning networks [18]. The researchers have developed a brand-new hybrid random matrix (HRM) to achieve this goal. The Restricted Isometry Property (RIP) is aided by this matrix. A motion heat map was employed by Mostafa et al. [19] to locate the area of interest. Different classifiers [20] are utilized after determining the motion structure, and the CNN classifier produces good results for detecting anomalies [21] in crowd scenes. the analysis of movies for abnormalities using deep learning. It is difficult to find anomalies in videos [22]. The clarity of the video and the range of changes that might occur inside the film, including human motions [23] and environmental variables, are what make this work challenging. The researchers in this study learn several data mining techniques and enhance the precision of computer identification of anomalous activity in crowd footage [24]. In this article, a brand-new IDS video crowd is introduced. It performs experiments utilizing shell command data from Unix users. Three sections make up the experiment. Additionally, Lamba et al. provide a taxonomy for crowd analysis and feature extraction, whereas Grant and Flynn [25] propose one for crowd analysis and related datasets. In order to detect anomalous events from surveillance footage, the authors suggested [26] the Deep One-Class (DOC) model, which combines the one-class Support Vector Machine (SVM) with Convolutional Neural Network (CNN). Using CNN, high-level characteristics are retrieved. Both classification and CNN model parameter optimization employ a single-class SVM [27] layer. In order to analyse crowds in crowded settings, this research offers [28] an end-to-end network that is lightweight and minimal in complexity.

Given that it only needs 0.86 M parameters, this network is referred to as being lightweight. Comparatively speaking to other neural networks, this is a pretty tiny number. A Bayesian model has been put out by Wu et al. [29] for the identification of crowd behavior. According on the density levels of picture patches, L2SM [30] attempts to forecast rescaling factors, which are subsequently applied to resample feature maps to provide the final prediction. Crowd analysis is studied by the author [31] and is broadly divided into two categories: crowd counting and crowd behavior analysis. It is possible to see a visualization [32] of the proposed taxonomy. The authors demonstrate that earlier research on the identification of human activities concentrated on specific settings. Interest in collective or crowd behavior emerged later. The review mentions research on group analysis, the identification of anomalous occurrences, and crowd motion for crowd behavior analysis.

A motion heat map was employed by Mostafa et al. [33] to locate the area of interest. Different classifiers are utilized after determining the motion structure, and the CNN [34] classifier produces favorable outcomes for detecting anomalies in crowd scenes.

III. Importance of Crowd Analysis

When planning crowd management tactics for public gatherings, public space design, visual surveillance, and virtual environments, people employ crowd analysis. Crowd behavior analysis is a technology that can correctly and quickly identify anomalies and provide an alarm. This implies that based on trustworthy signals, security officers may swiftly identify anomalous circumstances from a variety of footage. Extreme significance must be given to crowd control [35]. Uncontrolled crowds have always led to risky scenarios like stampedes and accidents. We require some practical measures and methods to efficiently and successfully bring the throng under control given its daily growth. These days, cameras that are positioned in public spaces provide us with a wealth of information about crowds in the form of pictures and movies. For instance, airports, stadiums, and train stations all have video surveillance equipment. By making the input so readily available, this provides crowd analysis an additional boost. We have data that is not only plentiful but also of diverse quality and has been gathered over a lengthy period of time. However, despite having input at our disposal, we have not been able to utilise the resources to their full potential. With improved methodologies, there is still a lot of room for improvement in the fields of anomalous crowd detection and crowd behaviour analysis.

IV. Human Conduct

Each individual's behavior in the actual world is distinct and is influenced by a plethora of variables that evolve through time, as well as the experiences that each person has. Additionally, physiological considerations characterize it, and capabilities that restrict some segments of the population from carrying out particular tasks. Each agent should act differently in the crowd simulations, exactly as they would in real life [36]. But the amount of system resources required for each virtual agent to display a unique behavior would make it impossible to simulate this process in real time. People's age, weight, and height were taken into consideration for determining the travel speed, step cadence, and stride length [37]. To accomplish this, the researcher studied men and women between the ages of 19 and 90 as they crossed a 12-meter footbridge at their typical speed. The suggests a few algorithms to determine how quickly people walk while engaging in daily tasks. The equations suggested by Samson were re-applied in the case study in order for the virtual agents to replicate these displacements with the characteristics that set men and women apart in the simulation [37]. The link between the displacement speeds of men and women is seen in Table 1.

Table 1. Samson's Equations for Human Displacement Velocity in Meters Per Second

Woman	Man
Speed = 1,420	Speed = 1460
Speed = -0.003 age (*) + 1.552	Speed = -0.002 age (*) + 1.582
Speed = -0.002 age (*) + 0.618 height (*) + 0.484	Speed = -0.002 age (*) + 0.442 height (*) + 0.750
Speed = -0.001 age (*) + 0.827 height (*) - 0.003 weight (*) + 0.316	Speed = -0.001 age (*) + 0.486 height (*) - 0.001 weight (*) + 0.720

V. Components of Crowd Analysis

The most active field of study is crowd analysis. Crowd, in the broadest sense, refers to a collection of people; its behavior has a collective feature, such as "angry crowd" or "peaceful crowd." The major crowd analysis [38] components, including crowd behavior comprehension, crowd tracking, crowd motion detection, and crowd density estimate, are summarized below as well as the state of the art for each.

5.1 Crowd Behavior Understanding

Recognizing crowd behavior is crucial in video monitoring for public spaces. Regular motion direction, slow (walking) pace, pausing, etc. are typical characteristics typical behaviors. Fighting, bifurcation, deviation, and opposing crowd movement are examples of unusual behaviors that exhibit the appropriate motion properties. Therefore, crowd motion data conveyed by the long-term and short-term motion vectors may be used to infer crowd behavior. The crowd vector's length and speed are exactly proportional [39]. Cameras used for video surveillance are typically fixed. Therefore, the vector length

corresponding to the typical population pace may be calculated if the camera is calibrated, the average walking speed is measured, and the video frame rate is known. The quantity of legitimate featured points is used to calculate crowd movement. The texture of a crowd image does not [40] vary much during a reasonably short time period, hence the number of features should remain constant. If the quantity of valid feature points suddenly varies, this property might be utilized to identify suspicious events.

5.2 Crowd Tracking

Many applications, including motion-based detection, automated surveillance, video indexing and rescue, human computer communication, and traffic monitor, among others, employ object tracking [41]. Object recognition, object segmentation, backdrop removal, object illustration, and characteristic selection for tracking are some of the phases that make up object tracking. One of the most important mechanisms in a wide range of computer visualization applications is object tracking. When people are being monitored, the ranking is calculated at the individual level. However, instead of focusing on one specific object the entire time, when we talk about crowd tracking, the focus is on the movement of a crowd as a collection of small elements whose organization changes continually [42]. Additionally, tracking is carried out by the human head, which is challenging when it comes to supervising and controlling big groups in public spaces like railway stations and airports [43]. With this method, it is hard to follow big groups of people in densely populated settings when most of the human bodies are either fully or partially occupied and the majority of the picture is in motion. We propose a method for monitoring pedestrians by detecting and tracking only their heads rather than their entire bodies [37], as we think that the human head is the only body component that can be consistently recognized and tracked under these circumstances.

5.3 Crowd Motion Detection

The Background Subtraction Method can be used to identify crowd movements. The Background Subtraction technique, as the name indicates, separates the foreground from the background in a series of video frames. A foreground item is a focus of attention that helps reduce the quantity of data that has to be processed and provides crucial information for the job at hand [44]. For purposes like surveillance, a group of methods known as background subtraction can be used to isolate items of interest from a scene. Calculating a reference image, deducting each new frame from it, and thresholding the outcome are all steps in the background removal process [45].

5.4 Crowd Density Assessment

In the investigation of aberrant crowd behavior, assessment of crowd density is an important subject. Currently, the population-to-area ratio is used to calculate crowd density. The safety of public events has long been a top issue because of the elevated danger of deterioration [46]. Because they are so effective at capturing information and need so little in the way of human resources, video analysis methods are becoming more and more common in the visual monitoring of public spaces. Two basic objectives of crowd density estimation are to estimate the approximate population size of the target scene. Second, provide a variety of crowd members, i.e., calculate the density in broad classifications [47].

5. Dataset

The Shanghaitech dataset is a sizable dataset for crowd counting. It includes 1198 crowd photos with annotations. The dataset is split into two sections, Part-A and Part-B, each of which contains 482 and 716 photos, respectively [48]. In contrast to most current datasets, as seen in figure 1, the crowd density greatly fluctuates across the two groups, making accurate assessment of the population more difficult. The train and test subsets of Part-A each include 300 and 182 pictures, respectively. The train and test subsets of Part-B are made up of 400 and 316 pictures, respectively. A location in the center of the head is labelled on each individual in a crowd shot. The collection has 330,165 annotated individuals in total.

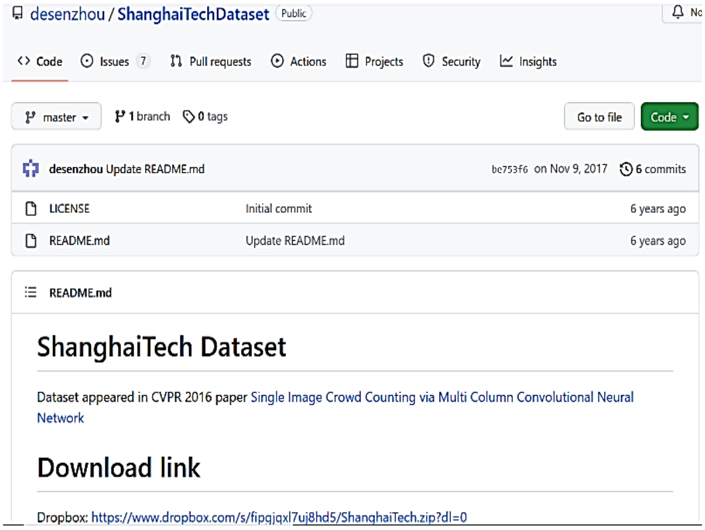

Figure 1 The ShanghaiTech Dataset

VI. Methodology

This part gives a thorough explanation of our suggested technique. The methodology's summary is shown in Figure 1. Using a Multicolumn Convolutional Neural Network (MCNN), we have introduced a unique method for forecasting crowd behaviour [49]. The first stage is the pre-processing of the input picture. For the aim of pre-processing the picture, we employ OpenCV. The extraction of low-level characteristics is the goal of this pre-processing. To do so, we first obtain the image spectrum, filter it, and then crop it into patches in accordance with the frame set. We extract low level characteristics for each of these patches. To get the ground truth density of the picture, we construct density maps of the objects from this and linearly map them with the collected features.

6.1 Multi-column Convolutional Neural Network (MCNN)

To map the picture to its population density map, we propose a straightforward but efficient Multi-column Convolutional Neural Network (MCNN) architecture. Special CNNs known as multi column CNNs employ multiple CNNs [50] inside the same design. Different resolution input photos are utilized for each CNN. The outputs of the several CNNs utilized are combined linearly to determine the final result.

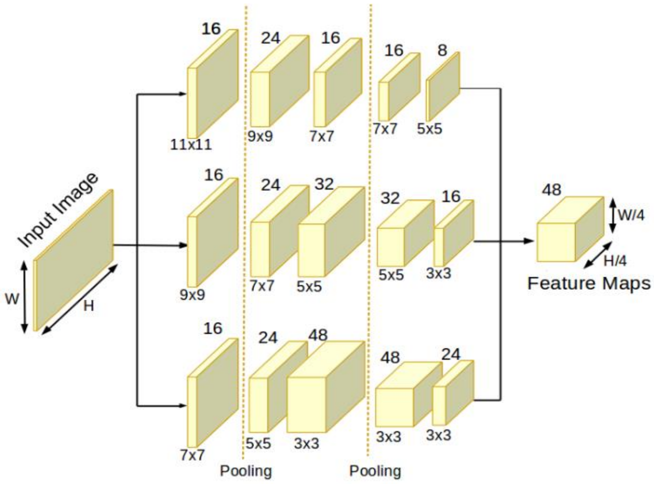

Figure 2 The Feature Extraction Module

The best aspect of this CNN is that it automatically extracts features, as demonstrated in figure 2. In essence, the convolution layer is a filter that is applied to an input picture to produce a filtered image. A mapping of different filtered pictures of the same input that shows the intensity of the visual characteristics in figure 3 is produced by repeatedly applying the same filter.

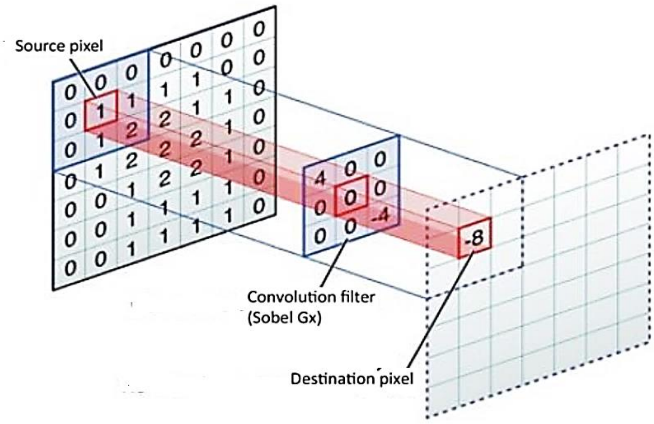

Figure 3 The Working of a Convolution Layer Functions

The maximum element is taken from each location on the feature map that is covered by the filter by the max pooling layer. As a result, at the conclusion of the max pooling layer, we are left with a reduced feature map that only shows the key components of the input picture. Because it down samples feature detection in feature maps, we apply max pooling in CNNs. It benefits in two ways. One is that we capture the image's key elements, and two is that we lessen the complexity of the processing. Max pooling [51] is frequently employed in CNNs for tasks involving object identification since it aids in identifying an item's most distinguishing characteristics, such as its edges and corners see figure 4. We have a filter that is 2*2 in size, and the stride of the filter is 2,2.

```
import numpy as np
from keras.models import Sequential
from keras.layers import MaxPooling2D
# define input image
image = np.array([[1, 2, 6, 2],
                  [8, 3, 9, 7],
                  [7, 6, 2, 4],
                  [4, 1, 6, 1]])
image = image.reshape(1, 4, 4, 1)
# define model containing just a single max pooling
layer
model = Sequential(
   [MaxPooling2D(pool_size = 2, strides = 2)])
# generate pooled output
output = model.predict(image)
# print output image
output = np.squeeze(output)
print(output)
```

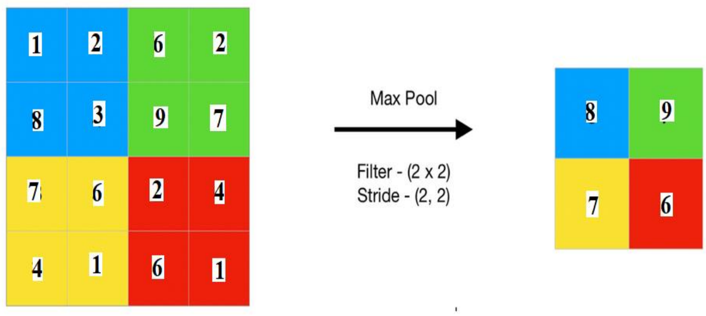

Figure 4 The Max pooling layer

Multicolumn Convolutional Neural Network (MCNN) is the approach we suggest. The data set is filtered before a training and testing data set is created. Then, in the following stage, we will use photos of the crowd to extract their characteristics, write down their dimensions and shapes, and use grey scaling to determine the density of the ground truth. Then, using the MCNN algorithm [52], we will input the photos to determine their estimated density. To determine the model's final weights, we compare the estimated account and density with the ground truth count and density. Once the model has been trained, we can use it to determine the density ratio of the photos and make predictions about the crowd's normality or abnormality. Currently, the MCNN algorithm is used to estimate crowd density, or anticipate crowd density. A crowd density map is produced by the MCNN. These density maps are combined to give us the overall number of people present. Our MCNN model's turn to be trained is now. For this, we compare the estimated density (ED) to the previously computed ground truth density (GTD). We can save the model by finalizing the weights if the difference between the estimated density (ED) and the ground truth density (GTD) [53] is less than the predetermined threshold. However, if the difference exceeds the threshold, a fresh set of weights are used to train the model. After our model has been trained, we assess the density ratio to separate crowds [54] that are normal from those that are aberrant.

This experiment transforms each picture in the dataset into grayscale images before to model training due to the ShanghaiTech dataset's combination of color photos and grayscale images and to minimize the influence of sample imbalance. There are two sections to the ShanghaiTech dataset: A and B. Part A has 300 photos, whereas Part B contains 400 photographs. Since the crowd density [55] varies among the dataset's images, we divide each image into 9 parts before training, and the related density map is divided into 9 pieces as well. Prior to the network's output density map, the image features should be sampled four times after convolution to guarantee that the features' dimensions matched throughout training.

Figure 5 The Crowd Density Original Image Map

The final output projected density map is a fourth of the original picture due to the usage of two up sampling with a scale factor of 2 and four times of maximum pooling across the whole model shown in figure 5 and 6. To guarantee that the parameter dimensions match, the original picture is scaled to an integer multiple of 4, and the actual density map is scaled to 1/4 of the original. The activation function we employed was the Rectified Linear Unit (ReLU) [56].

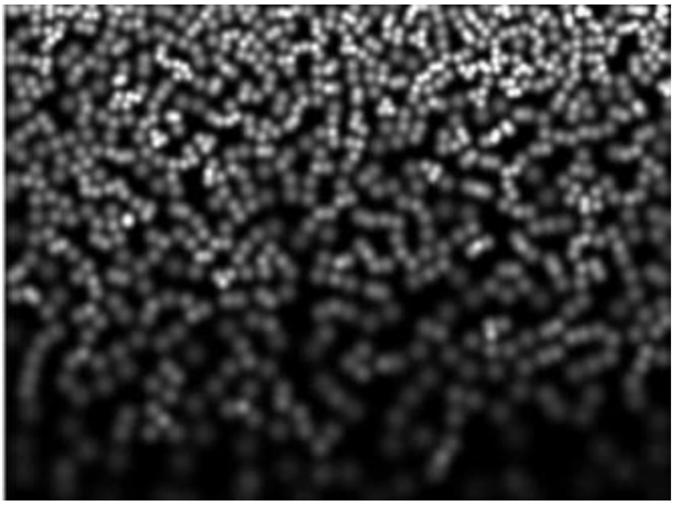

Figure 6 The Crowd Density Ground Truth Map

VII. Results Evaluation

There are two sections in this paragraph. In the first section, we'll go through the numerous measures, such MAE and MSE, that were employed as assessment indicators for our suggested technique. The outcomes of our tests evaluating the precision of our suggested technique in comparison to models like MCNN are shown in the second part.

7.1 Evaluation Index

We used two significant assessment criteria in our experiment to gauge how well our suggested approach (MCNN) performed. They are Mean Squared Error (MSE) and Mean Absolute Error (MAE). The following definitions apply to these markers.

7.1.1 Mean Absolute Error (MAE)

MAE, or Mean Absolute Error The mean of all mistakes we receive for any given set of forecasts is known as the mean absolute error, or MAE [57]. Here, the discrepancy between the actual observations and the expected values is referred to as the error. The Mean Absolute Error is the average of this discrepancy. The Ground Truth Density (GTD), which is the real observation in our experiment, was determined for it. The anticipated figure is the crowd's Estimated Density (ED). As a result, the absolute inaccuracy is determined by the difference between the ED and the GTD. We can determine its average to get the prediction's MAE.

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N}|x_i - x_i'|$$

7.1.2 Mean Squared Error (MSE)

The square of the mean of the square of the mistakes is the mean squared error, or MSE. The direction of the difference is taken into account by this metric. As a result, MSE [58] results are always favorable. We figure out the square of the difference between the actual observation and the projected values. The Mean Squared Error is calculated by taking the average of this difference.

$$\mathrm{MSE} = \left(\frac{1}{N}\sum_{i=1}^{N}|x_i - x_i'|^2\right)^{1/2}$$

N test photos are present. The *i-th* image's projected and actual population sizes are $x_{i'}$ and x_i, respectively.

We have shown that our technique outperforms others after using it on the Shanghaitech dataset. The graph in tables 1 and 2 compares the accuracy of various approaches to our suggested method (MCNN) [59]. This provides proof that crowd counting and crowd density analysis, which are illustrated in figure 7 and 8, are more accurately predicted by MCNN [60].

Table 1 ShanghaiTech Dataset Experimental Crowd Count Accuracy Results

Crowd No	Crowd 100	Crowd 200	Crowd 500	Crowd 1000
MCNN	97	93	89	85

Table 2 ShanghaiTech Dataset Experimental Crowd Density Accuracy Results

Crowd No	Crowd 100	Crowd 200	Crowd 500	Crowd 1000
MCNN	98	95	93	91

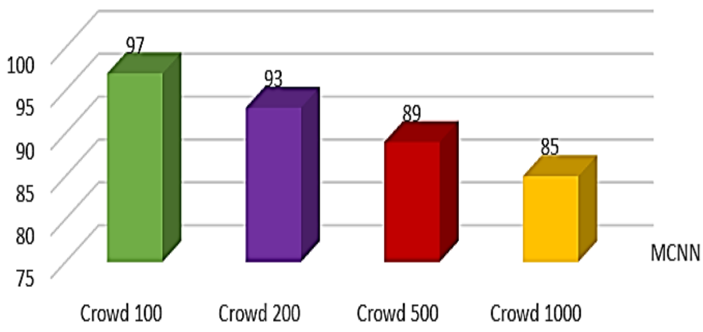

Figure 7 The Count Accuracy Results

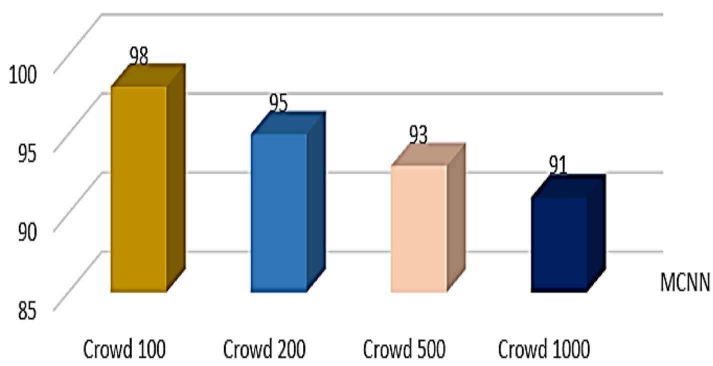

Figure 8 The Density Accuracy Results

VIII. Conclusion

There are several locations where large crowds or groups of people may be seen, including airports, sports venues, and numerous religious, educational, and entertaining activities. In such busy areas, video surveillance may be highly beneficial for crowd control. Recognizing the presence of a crowd and counting the number of individuals present is crucial. To prevent riots, this can be highly helpful in identifying unexpected troop build-up. This research investigates a multicolumn convolutional neural network (MCNN) model for accurate counting in densely populated areas. The Shanghaitech Dataset, which is further separated into Part A and Part B, is utilized here. In terms of the annotated heads for crowd counting, this is the biggest dataset to date. The suggested methodology beats modern crowd counting techniques across all assessment datasets. The network functions successfully, as demonstrated by studies on the ShanghaiTech dataset using the evaluation indicators average absolute error (MAE) and mean square error (MSE). In next work, we'll investigate how to manage various image sizes and resolutions while generating the true picture.

IX. REFERENCES

[1]. Camille Dupont, Luis Tobıas & Bertrand Luvison." Crowd-11: A dataset for fine-grained crowd behaviour analysis". . In: IEEE Xplore Computer Vision Foundation, 2011

[2]. Khan, A.; Shah, J.; Kadir, K.; Albattah, W.; Khan, F. Crowd Monitoring and Localization Using Deep Convolutional Neural Network: A Review. Appl. Sci., 10, 4781, 2020

[3]. Yusuf Perwej, Nikhat Akhtar, Firoj Parwej, "The Kingdom of Saudi Arabia Vehicle License Plate Recognition using Learning Vector Quantization Artificial Neural Network", International Journal of Computer Applications (IJCA), USA, Vol. 98, No.11, Pages 32 – 38, 2014, DOI: 10.5120/17230-7556

[4]. Motlagh, N.H.; Bagaa, M.; Taleb, T. UAV-Based IoT Platform: A Crowd Surveillance Use Case. IEEE Commun. Mag., 55, 128–134, 2017

[5]. Chan, A.B.; Vasconcelos, N. Bayesian Poisson regression for crowd counting. In Proceedings of the 2009 IEEE 12th International Conference on Computer Vision, Kyoto, Japan, pp. 545–551, 2009

[6]. Yusuf Perwej, "An Optimal Approach to Edge Detection Using Fuzzy Rule and Sobel Method", International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, Volume 4, Issue 11, Pages 9161-9179, 2015, DOI: 10.15662/IJAREEIE.2015.0411054

[7]. Rabaud, V.; Belongie, S. Counting crowded moving objects. In Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, New York, NY, USA, Volume 1, pp. 705–711, 2006

[8]. Yusuf Perwej, Md. Husamuddin, Fokrul Alom Mazarbhuiya,"An Extensive Investigate the MapReduce Technology", International Journal of Computer Sciences and Engineering (IJCSE), E-ISSN : 2347-2693, Volume-5, Issue-10, Page No. 218-225, 2017, DOI: 10.26438/ijcse/v5i10.218225

[9]. Zhang, C.; Li, H.; Wang, X.; Yang, X. Cross-scene crowd counting via deep convolutional neural networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12, 2015

[10]. C. Direkoglu, "Abnormal Crowd Behaviour Detection Using Motion Information Images and Convolutional Neural Networks," in IEEE Access, vol. 8, pp. 80408-80416, 2020, doi: 10.1109/ACCESS.2020.2990355

[11]. Shubham Mishra, Mrs Versha Verma, Nikhat Akhtar, Shivam Chaturvedi, Yusuf Perwej, "An Intelligent Motion Detection Using OpenCV" , International Journal of Scientific Research in Science, Engineering and Technology (IJSRSET), Print ISSN: 2395-1990 , Online ISSN : 2394-4099, Volume 9, Issue 2, Pages 51-63, 2022, DOI: 10.32628/IJSRSET22925

[12]. Zhan, B., Monekosso, D. N., Remagnino, P., Velastin, S. A., & Xu, L. Q.. ,"Crowd analysis: A survey. Machine Vision and Applications", 19(5–6), 345–357, 2008

[13]. G. Tripathi, K. Singh and D. K. Vishwakarma, "Crowd Emotion Analysis Using 2D ConvNets," 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT), Tirunelveli, India, 2020, pp. 969-974, doi: 10.1109/ICSSIT48917.2020.9214208

[14]. Yusuf Perwej, Shaikh Abdul Hannan, Nikhat Akhtar, "The State-of-the-Art Handwritten Recognition of Arabic Script Using Simplified Fuzzy ARTMAP and Hidden Markov Models", International Journal of Computer Science and Telecommunications, Volume, Issue 8, Pages 26 - 32, 2014

[15]. Qi Wang, Junyu Gao, Wei Lin, and Yuan Yuan. Learning from synthetic data for crowd counting in the wild. In CVPR, pages 8198–8207, 2019

[16]. XuguangZhang, Zhang, Q., ShuoHu, ChunshengGuo, and Yu, H.,"Energy level based abnormal crowd behavior detection", Sensors, MDPI, 2018

[17]. Lamba, S., & Nain, N.,"Crowd monitoring and classification: a survey". In Advances in computer and computational sciences (pp. 21–31). Springer, 2017

[18]. M. Gao et al., "Violent crowd behavior detection using deep learning and compressive sensing," 2019 Chinese Control And Decision Conference (CCDC), Nanchang, China, 2019, pp. 5329-5333, doi: 10.1109/CCDC.2019.8832598

[19]. Mostafa, T. A., Uddin, J., and Ali, M. H..,"Abnormal event detection in crowded Scenarios", In 3rd International Conference on Electrical Information and Communication Technology (EICT), Khulna, Bangladesh, 2017

[20]. Y. Perwej, "Unsupervised Feature Learning for Text Pattern Analysis with Emotional Data Collection: A Novel System for Big Data Analytics", IEEE International Conference on Advanced computing Technologies & Applications (ICACTA'22), SCOPUS, IEEE No: #54488 ISBN No Xplore: 978-1-6654-9515-8, Coimbatore, India, 2022, DOI: 10.1109/ICACTA54488.2022.9753501

[21]. Y. Perwej, Prof. (Dr.) Syed Qamar Abbas, Jai Pratap Dixit, Nikhat Akhtar, Anurag Kumar Jaiswal, "A Systematic Literature Review on the Cyber Security", International Journal of Scientific Research and Management (IJSRM), ISSN (e): 2321-3418, Volume 9, Issue 12, Pages 669 - 710, 2021, DOI: 10.18535/ijsrm/v9i12.ec04

[22]. Chong, Y., & Tay, Y.."Modeling representation of videos for anomaly detection using deep learning: A review", arXiv:1505.00523,2015

[23]. Yusuf Perwej, "A Literature Review of the Human Body as a Communication Medium using RedTacton", Communications on Applied Electronics (CAE), ISSN: 2394-4714, Foundation of Computer Science FCS, USA, Volume 9, No.4, Pages 7 – 17, 2016, DOI: 10.5120/cae2016652161

[24]. Xie, S., Zhang, X. & Cai, J. Video crowd detection and abnormal behavior model detection based on machine learning method. Neural Comput & Applic 31, 175–184, 2019

[25]. Grant, J. M., & Flynn, P. J.," Crowd scene understanding from video: a survey. ACM Transactions on Multimedia Computing", Communications, and Applications (TOMM), 13(2), 19, 2017

[26]. C, S. J. S. J. H.,"Abnormal event detection for video surveillance using deep one-class Learning", Multimedia Tools and Application 78, pp.36333647, 2019

[27]. Y. Perwej, "The Bidirectional Long-Short-Term Memory Neural Network based Word Retrieval for Arabic Documents", Transactions on Machine Learning and Artificial Intelligence (TMLAI), Society for Science and Education, United Kingdom (UK), ISSN 2054-7390, Volume 3, Issue 1, Pages 16 - 27, 2015, DOI: 10.14738/tmlai.31.863

[28]. X. Ma, S. Du and Y. Liu, "A Lightweight Neural Network For Crowd Analysis Of Images With Congested Scenes," 2019 IEEE International Conference on Image Processing (ICIP), Taipei, Taiwan, 2019, pp. 979-983, doi: 10.1109/ICIP.2019.8803062

[29]. Wu, S., Wong, H. S., and Yu, Z. ,"A bayesian model for crowd escape behavior detection", IEEE Transactions on Circuits and Systems for Video Technology Vol.24, No.1, pp.85–98, 2014

[30]. Chenfeng Xu, Kai Qiu, Jianlong Fu, Song Bai, Yongchao Xu, and Xiang Bai. Learn to scale: Generating multipolar normalized density maps for crowd counting. In ICCV, pages 8382–8390, 2019

[31]. Grant, J. M., & Flynn, P. J., "Crowd scene understanding from video: a survey. ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)", 13(2), 19, 2017

[32]. Akhtar, N., Tabassum, N., Perwej, A., Perwej, Y.: Data analytics and visualization using Tableau utilitarian for COVID-19 (Coronavirus). Glob. J. Eng. Technol. Adv. 3(2), 028–050 (2020), https://doi.org/10.30574/gjeta.2020.3.2.0029

[33]. Mostafa, T. A., Uddin, J., and Ali, M. H. 2017. Abnormal event detection in crowded scenarios. In 3rd International Conference on Electrical Information and Communication Technology (EICT), Khulna, Bangladesh.

[34]. Nikhat Akhtar, H. Pant, Apoorva Dwivedi, Vivek Jain, Y. Perwej, "A Breast Cancer Diagnosis Framework Based on Machine Learning", International Journal of Scientific Research in Science, Engineering and Technology (IJSRSET), Print ISSN: 2395-1990, Volume 10, Issue 3, Pages 118-132, May-June-2023, DOI: 10.32628/IJSRSET2310375

[35]. G. Brostow and R. Cipolla. Unsupervised bayesian detection of independent motion in crowds. Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on, 1:594-601, June 2006

[36]. Zou, Q., & Chen, S.,"Simulation of Crowd Evacuation under Toxic Gas Incident Considering Emotion Contagion and Information Transmission", Journal of Computing in Civil Engineering, 34(3), 04020007, 2020

[37]. Samson, M., Crowe, A., De Vreede, P., Dessens, J., Duursma, S., Verhaar, H.: Differences in gait parameters at a preferred walking speed in healthy subjects due to age, height and body weight. Aging Clinical and Experimental Research 13(1), 16–21, 2001

[38]. C. Zhang, H. Li, X. Wang and X. Yang, "Cross-scene crowd counting via deep convolutional neural networks", IEEE Conference on Computer Vision and Pattern Recognition, 2015

[39]. Miyazaki Shinji, Miyano Hiroyoshi, e al., " New Congestion Estimation System Based On the " Crowd Behavior Analysis Technology", NEC Technical Journal/Vol.9 No.1

[40]. Min Sun, Dongping Zhang, Leyi Qian and Ye Shen, Crowd Abnormal Behavior Detection on Label Distribution Learning, IEEE, 2015

[41]. B. Solmaz, B. Moore, and M. Shah. Identifying behaviors in crowd scenes using stability analysis for dynamical systems. IEEE Tran. on PAMI, 34(10):2064-2070, Oct. 2012

[42]. shad Ali and Matthew N. Dailey, " Multiple Human Tracking in High Density Crowds"

[43]. W. Ge, R. T. Collins, and R. B. Ruback. Vision-based analysis of small groups in pedestrian crowds. IEEE Trans. on PAMI, 34(5):1003-1016, May 2012

[44]. Li, H. Chang, M. Wang, B. Ni, R. Hong and S. Yan, "Crowded Scene Analysis: A Survey", IEEE Transactions on Circuits and Systems for Video Technology, vol. 25, no. 3, pp. 367-386, 2015

[45]. S. Heldens, N. Litvak and M. van Steen, "Scalable Detection of Crowd Motion Patterns," in IEEE Transactions on Knowledge and Data Engineering, vol. 32, no. 1, pp. 152-164, 2018, doi: 10.1109/TKDE.2018.2879079.

[46]. X. Jiang, Z. Xiao, B. Zhang, X. Zhen, X. Cao, D. S. Doermann, et al., "Crowd counting and density estimation by trellis encoder-decoder networks", 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 6126-6135, 2019

[47]. M. O. Osifeko, G. P. Hancke and A. M. Abu-Mahfouz, "Surveilnet: A lightweight anomaly detection system for cooperative iot surveillance networks", IEEE Sensors Journal, vol. 21, no. 22, pp. 25293-25306, 2021

[48]. https://github.com/desenzhou/ShanghaiTechDataset

[49]. Y Y Zhang, D S Zhou et al., "Single-image crowd counting via multi-column convolutional neural network", CVPR, pp. 589-597, 2016
[50]. Saurabh Sahu, Km Divya, Dr. Neeta Rastogi, Puneet Kumar Yadav, Dr. Yusuf Perwej, “Sentimental Analysis on Web Scraping Using Machine Learning Method” , Journal of Information and Computational Science (JOICS), ISSN: 1548-7741, Volume 12, Issue 8, Pages 24-29, August 2022, DOI: 10.12733/JICS.2022/V12I08.535569.67004
[51]. Y. Ling et al., "A RRAM based max-pooling scheme for convolutional neural network", Proc. 5th IEEE Electron Devices Technol. Manuf. Conf. (EDTM), pp. 1-3, Apr. 2021
[52]. S. Sudholt and G. A. Fink, "PHOCNet: A Deep Convolutional Neural Network for Word Spotting in Handwritten Documents", ICFHR, pp. 277-282, 2016
[53]. Li, Y., Zhang, X. & Chen, D. Csrnet: Dilated convolutional neural networks for understanding the highly congested scenes. In Proc. IEEE Conference on Computer Vision and Pattern Recognition, 1091–1100, 2018
[54]. Al-Mushayt O., Haq Kashiful, Yusuf Perwej, “Electronic-Government in Saudi Arabia; a Positive Revolution in the Peninsula”, International Transactions in Applied Sciences, India, ISSN-0974-7273, Volume 1, Number 1, Pages 87-98, July-December 2009
[55]. Liu, W., Salzmann, M. & Fua, P. Context-aware crowd counting. In Proc. IEEE Conference on Computer Vision and Pattern Recognition, 5099–5108, 2019
[56]. M. D. Zeiler, M. Ranzato et al., "On Rectified Linear Units For Speech Processing", Proceeding of ICASSP, pp. 3517-3521, 2013
[57]. W. Wang and Y. Lu, "Analysis of the Mean Absolute Error (MAE) and the Root Mean Square Error (RMSE) in Assessing Rounding Model", IOP Conference Series: Materials Science and Engineering, vol. 324, no. 1, 2018
[58]. J. Qi, J. Du, S. M. Siniscalchi, X. Ma and C.-H. Lee, "Analyzing upper bounds on mean absolute errors for deep neural network-based vector-to-vector regression", IEEE Trans. Signal Process., vol. 68, pp. 3411–3422, 2020
[59]. Z. Shi, L. Zhang, Y. Sun and Y. Ye, "Multiscale multitask deep NetVLAD for crowd counting", IEEE Trans. Ind. Informat., vol. 14, no. 11, pp. 4953-4962, Nov. 2018
[60]. B. Yang, J. Cao, N. Wang, Y. Zhang and L. Zou, "Counting challenging crowds robustly using a multi-column multi-task convolutional neural network", Signal Process. Image Commun., vol. 64, pp. 118-129, Mar. 2018

Cite this article as :

Sachin Bhardwaj, Apoorva Dwivedi, Ashutosh Pandey, Dr. Yusuf Perwej, Pervez Rauf Khan, "Machine Learning-Based Crowd behavior Analysis and Forecasting", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 9, Issue 3, pp.418-429, May-June-2023. Available at doi : https://doi.org/10.32628/CSEIT23903104
Journal URL : https://ijsrcseit.com/CSEIT23903104