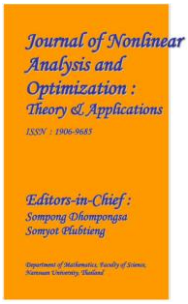


Journal of **N**onlinear **A**nalysis and **O**ptimization
Vol. 15, Issue. 1, No.10 : 2024
ISSN :**1906-9685**

A CNN CROWD COUNTING MODEL USING DEEP LEARNING

Dr. S.Om Prakash [1], P.CHATHURYA MAHATHI SRIJA [2], G.VISHNU PRIYA [3], V.SURYA KEERTHI [4], P.VAMSI REDDY [5]

[1] Associate Professor, Department Of Computer Science And Engineering,
RAGHU ENGINEERING COLLEGE (AUTONOMOUS)
Affiliated to JNTU GURAJADA, VIZIANAGARAM
Email: omprakash.samantray@raghuenggcollege.in

[2] B.Tech, Department Of Computer Science And Engineering,
RAGHU ENGINEERING COLLEGE (AUTONOMOUS)
Affiliated to JNTU GURAJADA, VIZIANAGARAM
Email: srijap2303@gmail.com

[3] B.Tech, Department Of Computer Science And Engineering,
RAGHU ENGINEERING COLLEGE (AUTONOMOUS)
Affiliated to JNTU GURAJADA, VIZIANAGARAM
Email: vsmur391@gmail.com

[4] B.Tech, Department Of Computer Science And Engineering,
RAGHU ENGINEERING COLLEGE (AUTONOMOUS)
Affiliated to JNTU GURAJADA, VIZIANAGARAM
Email: keerthivatsavayi@gmail.com

[5] B.Tech, Department Of Computer Science And Engineering,
RAGHU ENGINEERING COLLEGE (AUTONOMOUS)
Affiliated to JNTU GURAJADA, VIZIANAGARAM
Email: reddyvamsi32@gmail.com

ABSTRACT

In addressing the worldwide Covid-19 sub-variant pandemic situation, the process of flattening the curve for coronavirus cases will be difficult if the citizens do not take action to prevent the spread of the virus. One of the most important practices in these outbreaks is to identify crowd count in public places. This proposed system presents the detection of people and identifies the heavily crowded areas and helps as a precautionary measure in reducing physical contact between people. Using Convolution Neural Networks we will find count in crowd, if count exceeds given threshold, we can identify it as an alert zone with heavy crowd.With the proliferation usage of video surveillance for safety, traffic control, and privacy purposes and with the constant growth of population, it is important to keep monitoring using Closed-Circuit Television (CCTV). With new upcoming developed technologies, new systems and algorithms are introduced and implemented to the crowd counting system today retrieving live video surveillance from the CCTV. However, recent studies show that there are some challenges still faced regarding the crowd counting system which uses the density estimation.

The problems that occurred have resulted from the inaccuracy of the system that is caused by several factors. Factors such as the perspective distortion which is caused by the lack of data training and the method such as face detection is an ineffective method to determine the population density. Studies proposed have projected the idea of developing a more robust crowd counting methodology by implementing crowd counting by detection, clustering, and regression. Implementing these methods using the Convolutional Neural Network (CNN) will better the result of the detection since in CNN the image can be inputted and it will undergo several layers which will result in the system being able to differentiate one image from the other. With CNN the process of crowd counting will be able to be more advanced.

1. INTRODUCTION

In the field of effective computing, many techniques and tools are proposed to detect images for various purposes. In addressing the worldwide Covid-19 pandemic situation, the process of flattening the curve for coronavirus cases will be difficult if the citizens do not take action to prevent the spread of the virus. One of the most important practices in these outbreaks is to ensure a safe distance between people in public. This proposed system presents the detection of people with social distance monitoring as a precautionary measure in reducing physical contact between people. Using CNN we will find count in crowd, if count exceeds given threshold, we can identify whether social distance is maintained or not.

1.1 OBJECTIVE

- Analyze the incoming and outgoing of people at a public place by surveillance.
- Images will be collected at some instances and predicts number of people.

1.2 ABOUT THE PROJECT

Proposed system is to utilize machine learning-learning tools for counting people in a image taken at a public place.

The dataset is organized into 2 folders (train, test) and contains subfolders for each folder. There are hundreds of images(jpeg/jpg) of peoples at public places like malls etc. Using these images,we trained images to count the number of people.

1.3 PURPOSE

- Our purpose is to help government organizations with information of public data for more awareness in this pandemic situation.
- Government needs accurate information to take effective measures to maintain social distancing inorder to stop the spread of the virus.
- We try to find out from the count of an each image give and come to a conclusion whether the social distancing is maintained or not.

1.4 SCOPE

Pictures of pubic at a possible crowded place collected through surveillance systems will be sent to our machine learning model and then it will give the output as count of people in that image.

2. LITERATURE SURVEY

With the rising growth of video surveillance usage, it has revolutionized the industry. There is more research on video surveillance as well as exploring systems that can detect crowds. Brostow and Cipolla discover a system where they are able to detect specific people in crowds. However, there is a fault in the system where they encounter noises or other objects that exist such as stores and kiosks. As for Pathan et al. they work on the system where it counts by the erroneous movements in a public place. The system has encountered difficulty when it comes to the accuracy of the result due to the methods used which are the subtraction process in detecting people. Meanwhile, Krausz and Bauckage present the idea of a system that will automatically identify the critical situation during congestion by using an alarm system. However, they encounter an error when the system is implemented. Moreover, in the research of Zhao and Nevatia, they use the articulated ellipsoids to be able to further model the human form, because every individual appearance of a person is different. Implementing the augmented Gaussian distribution for the model of the background. As the moving head pixels are detected, the approach of MCMC is implemented to assist a higher chance in the detection probability of the crowds in a crowded environment. Counting by detection is a technique notably implemented the sliding window-based detection algorithms to count the number of instances within an image. By using these methods, it will inflict the background clutter and the presence of high-density crowd.To overcome this, researchers used another method to count the crowd, which is to count by global regression. Counting by global regression has been projected as a method to crowd-count the

pedestrians utilizing clustering or detection. As for crowd counting the method is restricted by several barriers between individuals. There are numerous prevailing ways to aid global regression prediction which is done by utilizing the regressors. The method is appropriate for the crowded places where it is trained in the low-level features. A familiar feature with semi-supervised regression as well as the method of transfer is proposed to scale back the quantity of the data however, there is the obstruction of labels that is lacking to supports its efficiency. There are reports on the estimation crowd counting from images however, there is the absence of data from surveillance video. Counting by global regression dismisses some data such as the data of pedestrian, however there is another method in which the counting is done by density estimation. A method that counts the crowd through pixel-level density map regression. Counting by density estimation also can estimate object count from images. Also, the interactive object counting system is introduced to assist the relevancy of the feedback.

3. EXISTING SYSTEM

- Existing Systems acquire the information as images through surveillance present at the crowded places.
- This Existing model will count the number of people which takes more time to give the output and also require manpower and error occurrence will be high.
- This model requires CCTV surveillance systems, offices and stations in order to acquire data.

DRAWBACKS OF EXISTING SYSTEM:

- More manpower
- Complexity of work
- Time consuming

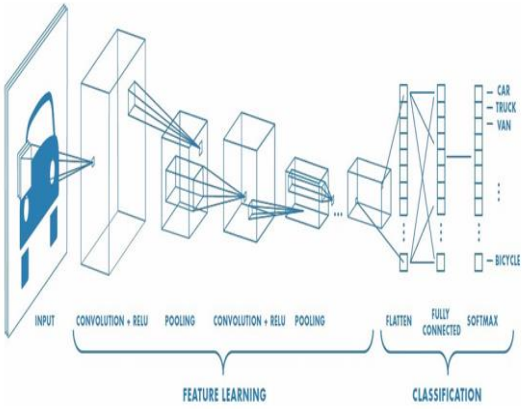
4. PROPOSED SYSTEM

To overcome the drawbacks of the existing system, the proposed system has been evolved.

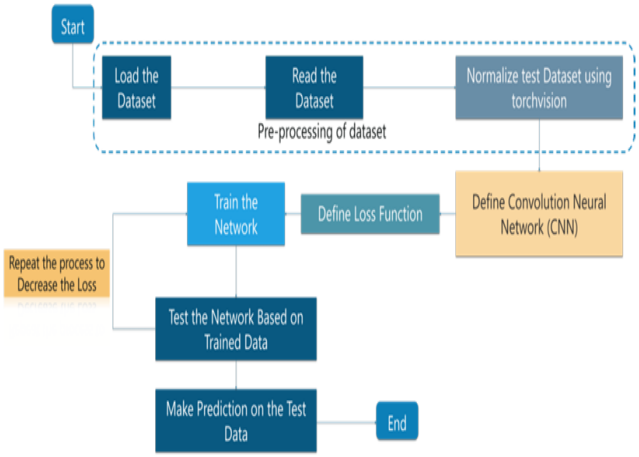
- Using machine-learning models for counting the number of people to differentiate the status of social distancing.
- Input images can be taken remotely from our mobile phones at the areas where CCTV cameras are not available.
- Input is taken as image and doesn't require the use of hardware devices to retrieve the output.

4.1 System Architecture

The main objective of the project is to count the number of people at expected crowded place.This model initially takes the images of people which is our dataset as the input. These inputs are fed into the deep learning classification algorithms to analyze the count of people and this analysis can be helpful to assess the current pandemic situation. This project replaces external manpower for social distancing maintenance. This model can be used by the government for better maintenance with less manpower in a small period of time.



Flowchart for Proposed model



4.2 Implementation

Step 1: Convolution Layer Convolution Layer performs the convolution operation. The three main elements of convolution operation are:

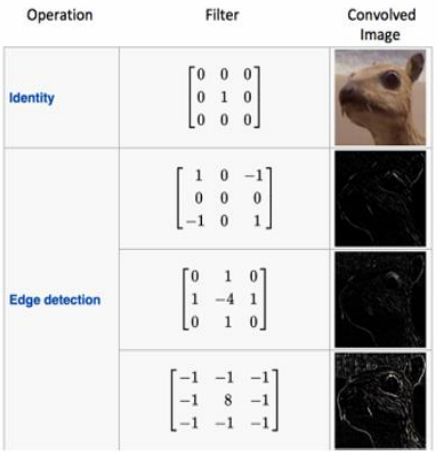
- Input image
- Feature Map
- Feature Detector

Working of Convolution Operation:

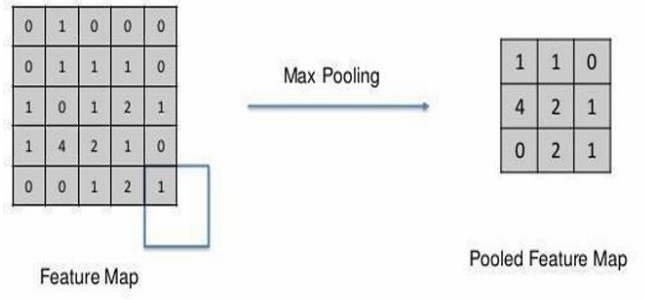
- Place the Feature detector on the top-left corner of the input image within the

borders.Count the number of cells in which the feature detector matches the input image.

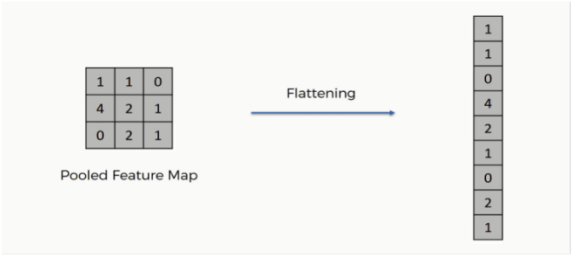
- The number of matching cells is then inserted in the top-left cell of the feature map.
- Move the feature detector one cell to the right and do the same operation. since we are moving the feature detector one cell at a time, that would be called a stride of one pixel.

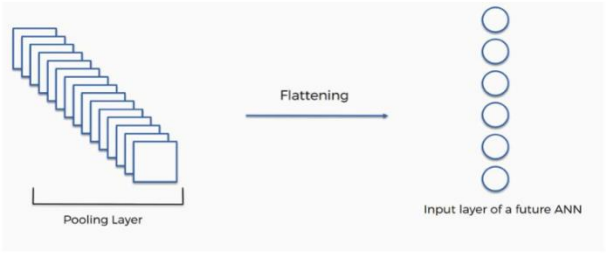


Step 2: Max Pooling The purpose of max pooling is enabling the convolutional neural network to detect the object in the image when presented with the image in any manner. Max pooling is a pooling operation that selects the maximum element from the region of the feature map covered by the filter. Thus, the output after max-pooling layer would be a feature map containing the most prominent features of the previous feature map.



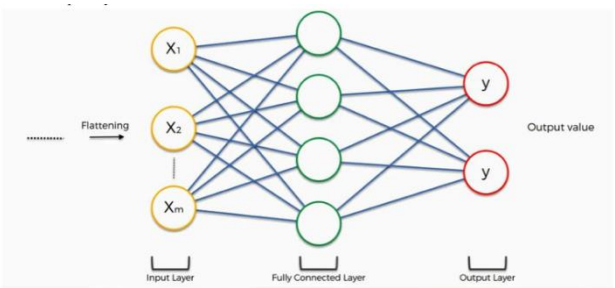
Step 3 : Flattening Flattening is converting the data into a 1-dimensional array for inputting it to the next layer. We flatten the output of the convolution layers to create a single long feature vector. And it is connected to the final classification model, which is called a fully-connected layer. In this Flattening layer, we are supposed to have a pooled feature map by now. we are going to flatten our pooled feature map into a column like in the image below.





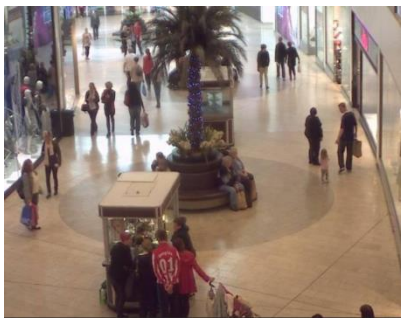
Step 4: Full Connection It's here where the process of creating a convolutional neural network begins to take a more complex and sophisticated turn. Fully Connected Layer is simply, feed forward neural networks. Fully Connected Layers form the last few layers in the network. The input to the fully connected layer is the output from the final Pooling or Convolutional Layer, which is flattened and then fed into the fully connected layer. There are three layers in the full connection step:

- Input layer
- Fully-connected layer
- Output layer



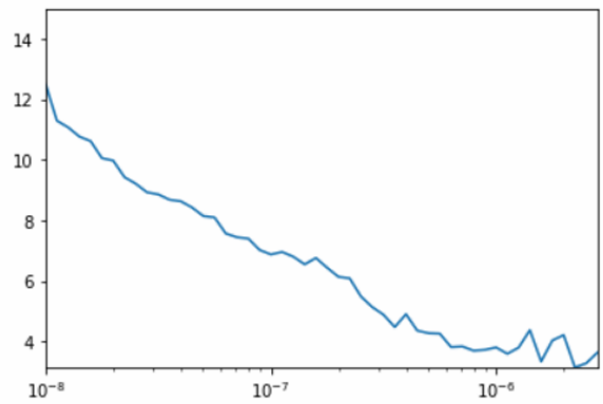
The role of this fully connected layer is to take this data and combine the features into a wider variety of attributes that make the convolutional network more capable of classifying images, which is the whole purpose of creating a convolution neural network.

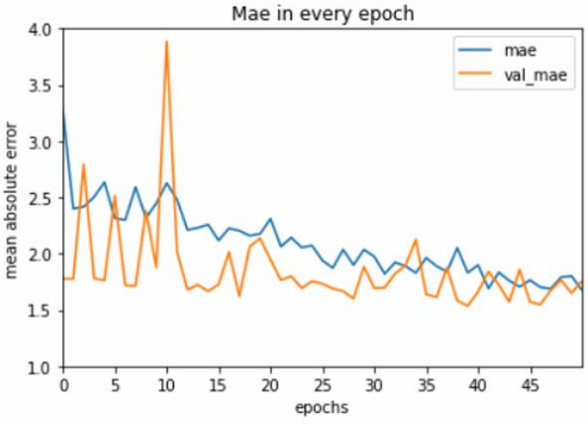
5. RESULTS

S.NO	Test Cases	Outputs	Test Case report
1.	**Case 1:** Picture of people at a shopping mall		Successfully predicts the count
2.	**Case 2:** Picture of people at a shopping mall with less crowd		Successfully predicts the count
3.	**Case 3:** Picture with single person		Successfully predicts the count
4.	**Case 4:** Picture of people at a public place		Successfully predicts the count

5.	**Case 5:** Picture without person		Fails to predict because the count was given even though there are no persons in the picture
6.	**Case 6:** Picture without person		Fails to predict because the count was given even though there are no persons in the picture

GRAPHS





5. CONCLUSION

A CNN CROWD COUNTING MODEL USING DEEP LEARNING is trained successfully using python with an accuracy of 90% against custom data set. This can further be applied to various real-time analyses by taking snapshots of the public from a video and can be implemented at possible different crowded places like parks, restaurants, traffic signals and stadiums. This analyzation and prediction of data will give more accurate predictions for government which will help them to face any other situation which demand social distancing type of precaution from getting in danger not only for pandemics but also helps monitoring and administering over the population whether to increase or decrease the necessary precautions and actions in order to face any kind of situation. We can increase the working capacity and compatibility of this model in respect to the work and result we need to acquire from this model.

REFERENCES

1] Andy S. 2018. Algorithm Spotlight: Crowd Counter https://algorithmia.com/blog/algorithm-spotlight-crowdcounter .

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

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

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

[6] Dailey K 2013. The rise of CCTV surveillance in the US https://www.bbc.com/news/magazine-22274770.

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

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

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

[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.

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

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

[13] Kurama V .2019. Dense and Sparse Crowd Counting Methods and Techniques: A Review https://nanonets.com/blog/crowd-counting-review/#what-is-crowd-counting.

[14] Lee W Y, Ko K E, Geem Z W and Sim K B 2017 Method that determining the Hyperparameter of CNN using HS algorithm vol 27 pp 22-28.

[15] Lempitsky V and Zisserman A 2010. Learning to count objects in images. In Advances in neural information processing systems pp 1324-1332.

[16] Lin S F and Lin C D 2006. Estimation of the pedestrians on a crosswalk. Savitribai Phule Pune University Pune, India pp 4931–4936.

[17] Loy C, Chen K, Gong S and Xiang T. 2013. Crowd Counting and Profiling: Methodology and Evaluation London, United Kingdom pp 347-382.

[18] Luo H, Sang J, Wu W, Xiang H, Xiang Z, Zhang Q, and Wu Z. 2018. Applied Chongqing, Chine Vol 8 pp 1–12.

[19] Nakatsuka M, Iwatani H, and Katto J. 2008. International Conference on Mobile Computing and Ubiquitous Networking Tokyo, Japan pp 1–6