Journal of Information & Optimization Sciences
ISSN 0252-2667 (Print), ISSN 2169-0103 (Online)
Vol. 45 (2024), No. 4, pp. 1105–1115
DOI : 10.47974/JIOS-1695

A method to estimate crowd size and crowd count

Ravi Saharan *
Nishtha Kesswani†
Basant Aggarwal §
Department of Computer Science and Engineering
Central University of Rajasthan
Kishangarh
Ajmer 305817
Rajasthan
India

Abstract

Crowd is known when people in large number gather together. When crowd gather for any reasons, it may cause any serious problem for administration. So authority must know about actual situation of crowd. In recent years need for crowd size estimation, has arisen to manage the crowd well. It can help in managing crowd before happening of huge event, traffic management, providing security to crowd and in many other areas where to manage crowd, some estimation on crowd size is needed. After reviewing existing applications, it can be concluded that there are many approaches for estimating density of a crowd and to count the individual's, but most of the available approaches did not estimate both crowd density and crowd count together using single model. To overcome these restrictions, a single model is proposed to estimate crowd count and density in the crowded scenes.

Here two models are prepared in four stages that work individually as well as forms steps for combined approach. These models firstly pre-process the crowd images extract features and then classify into different classes. Crowd density model classifies images into several density classes. Firstly, dataset of crowd images is prepared according to the requirement using crowd dataset. These crowd images are then sorted and provided to next models. Combined Crowd density and Count Estimation model classifies images into various densities and count classes.

Subject Classification: *68T10, 68U10.*

Keywords: *Crowd, Crowd size, Crowd count, Feature extraction.*

* *E-mail:* **ravisaharan@curaj.ac.in (Corresponding Author)**
† *E-mail:* **nishtha@curaj.ac.in**
§ *E-mail:* **basant@curaj.ac.in**

1. Introduction

When people in big scale gather at one place is called crowd. Around the world on public places like metro stations, airport, beaches, concerts, conferences, rallies, sports, malls and on many political, social or professional events people generally gather. So, the need arises for crowd to manage, to secure and to better distribute the available resources among people and better planning at crowded places. A system was required that can estimate the density and count of the people present so that better facilities, security and management can be provided. The problem of crowd analysis can be separated into two parts: Estimating crowd density and estimating count of people in a crowd. The primary goal to predict the crowd density is to change an input crowdy image into a density-map that shows density of people present in the image, whereas the goal of prediction of crowd count is to know about the count of individuals in a crowded image.

Clutter, occlusions, non-uniform distribution of population in a picture, blurred illumination, intra as well as inter-image differences in presence, incorrect lighting, size, and angle of image capturing are just a few of the issues that crowd analysis faces[2].

Here Figure 1(a) shows image of sports, Figure 1(b) shows image of shopping malls and Figure 1(c) shows image of railway station.

Figure 1

(a) sports pic; (b) shopping malls; (c) railway station

1.1 *Justification of proposed work*

Mass gathers for different purposes at different places, therefore to provide security, to manage and to better facilitate the crowd, crowd analysis is needed. In absence of such a system situation may become worse. Some instances of human stampedes in India like; Sept 29, 2017: Mumbai Stampede, Oct 3,2014: Gandhi Maidan, Patna, Bihar, Oct,13,2013:

Ratangarh Temple, Datia, MP, Sepr,12,2012, Rajgir,Patna, Sept 30,2008, Chamunda Devi temple, Jodhpur, Rajasthan. So to know exact information which provides estimation of crowd density and crowd count, this method is proposed.

2. Literature Survey

After studying literature, it has been concluded that crowd estimation can be broadly classified as follows [2]:

In Table 1, various crowd estimation techniques and their comparative process and uses are shown.

Table 1

Comparative examination of crowd estimation techniques

Approaches for crowd estimation **Direct Approaches:**	**Process**	**Use**
Model-based analysis	Segments or detects each object in image using appearance	Suitable for sparse crowd.
Trajectory-based clustering	used for videos	well suited for sparse crowd
Indirect Approaches:		
Pixel-based analysis	based on foreground pixels counting	As density increases error in estimation increases. Occlusion in images also increases the error
Methods based on texture and gradient	Low-level feature such as texture, area, and edges are used	Suitable for dense crowded scene. But it is vulnerable to background change
Methods based on interest Points	Interest points are pixels of interest features that include corner points detected in images.	Well suited to the moving crowd

Based on above mentioned comparative examination, it can be concluded that existing techniques (direct approaches/ indirect approaches) either useful to count crowd density or crowd count only.

Literature survey is as follows:

1. Thanh-Sach Le et al. (2015) presented an alternative counting strategy, in which without detecting actual person, offered a solution to the human-crowd density estimation. Using this strategy, a method based on texture feature extraction having Gabor filters was proposed. [1].
2. Sami Abdulla Mohsen Saleh et al. (2015) conducted a survey that covered two approaches, first was object-based target detection and second was target detection using a network (texture based, pixel based, and contour based analysis. [2].
3. Vandit Chauhan et al. (2016) presented an epsilon support value regression strategy based on fusion to acquire data of overall crowd mages and hence to know about persons [3].
4. Yaocong Hu et al. (2016) examined how a deep learning method could be used to predict the possible number of people in a mild / dense crowd in a single photograph. To begin, features were extracted using a ConvNet structure. [4].
5. Ven Jyn Kok et al. (2016) published an overview that gave a review of crowd behavior analysis approaches from physical and biological standpoint [5].
6. Muhammad Saqib et al. (2016) proposed a texture-depend feature mining technique to estimate crowd density[6].
7. Yongsang Yoon et al. (2016) illustrated about conditional point process-based crowd count method that well suits for scenes that are sparsely and moderately crowded with an average accuracy of 81% [7].
8. Hiliang Pu et al. (2017) suggested a deep convolutional NN (neural network) - based technique for estimating crowd density. Deep networks were utilized to estimate crowd density, which first contributing, then subway-carriage scenes with density photos data set was utilized to assess crowd density prediction [8].
9. Xiaolong Jiang et al. (2019) suggested an encoding–decoding network for crowds' counting which was based on producing high density destination maps. The work is divided into several categories. Back-propagation process applied by distributing Ly and avoids the gradient vanishing problem by using combinatorial losses on intermediate output [9].
10. JiweiChen et al. (2020) Crowd Attention Convolutional Neural Network was offered as a model (CAT-CNN). Using automated

encoding to a confidence mapping. CAT-CNN adaptively assesses the human head at each pixel point. [10].

11. Qi Wang et al. (2021) created the NWPU-Crowd dataset, a large crowd count with localized dataset with 5109 pictures and 2,133,375 labeled heads with boxes and points. It has the widest density range and incorporates varied illumination situations when compared to other genuine datasets [11].
12. Zhang Q et al. (2022) proposed a deep neural network framework for multi-view crowd counting, which fuses information from multiple camera views to predict a scene-level density map on the ground-plane of the 3D world [12].
13. Xiong H et al. (2023) provided a mathematical analysis and a controlled experiments on synthetic data, to show the working of closed set [13]

3. Methodology used

The main purpose of this paper is to prepare a model to estimate crowd density with crowd count. Once we will get the exact information of crowd density and crowd count, we can prepare better crowd strategy and possibly incidents may be overcome. Here, a step by step method is presented in form of models of crowd density with crowd count. It has following steps:

Database preparation : Firstly, images taken from crowd dataset.

Image pre-processing : These images are then resized, rescaled and converted to gray images.

Feature extraction : These gray images are supplied to further to extract unique features. In **Classification**, these images are then classified to different crowd classes based on crowd count and density .

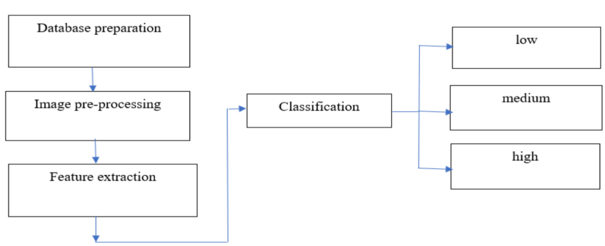

Figure 2

Model 1 Block Diagram

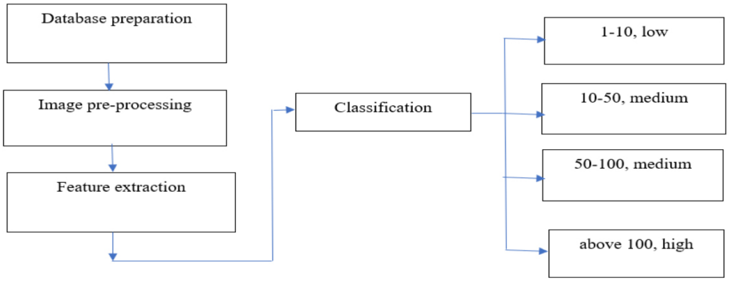

Figure 3

Model 2 Block Diagram

Crowd density estimation model 1: Here model takes inputs as images and classifies based on density under various parameter like 'low', 'medium' and 'high'. Figure 2, describes the crowd density estimation model 1, which classifies the images as 'low', 'medium' and 'high'.

Figure 3, describes the crowd density estimation model 2, which classifies the images based on crowd count and density.

Crowd density and count estimation model 2: Here model takes inputs as images and classifies based on crowd count under various parameter like '1-10 low', '10-50 medium', '50-100 medium', 'above 100, high'

3. Experiments and Result Analysis

Data Set Overview: Crowd images from NWPU-crowd dataset are taken from website http://www.kaggle.com. NWPU consists of 5,109 images and contains 2,133,375 annotated instances of images. Comparing with the previous datasets, the main advantages are, Fair Evaluation, High Resolution and Large Appearance Variation. Images are

Figure 4

(a) Class "low" (b) Class "medium"(c) Class "high"

Figure 5

(a) Class "1-10 low"(b) Class "50-100 medium" (c) Class "above 100 high"

edited and stored to prepare dataset. After each stage images of previous dataset are rearranged to prepare dataset for the next stage. The dataset contains images of different crowded places or events with different angles, illumination, crowd density, crowd count and physical distance among people present in the crowded scenes.

DATASET FOR MODEL 1 & 2:- We take several images which is used as a dataset for model 1 & model 2.

Here Figure 4(a) , 4(b) & 4(c) describe the input images for to be classified by model 1 as 'low', 'medium' & high'

Here Figure 5(a) , 4(b) & 4(c) describe the input images for to be classified by model 2 based on count and density.

Results are as follows:

Various results of training and testing with validation, accuracies and losses for Model 1are plotted here:

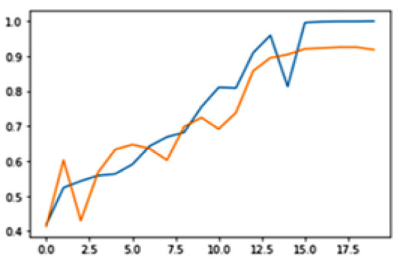

Figure 6

Plot of training(validation) (orange) and testing (blue) accuracies

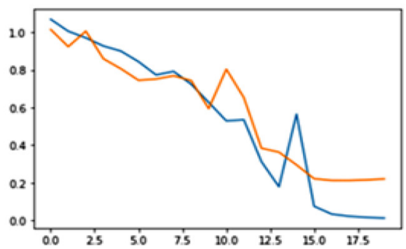

Figure 7

Plot of training (validation) (orange) and testing (blue) losses

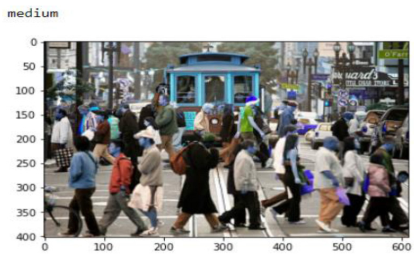

Figure 8

Result Output (prediction)

Figure 6 depicts the variation of the training and testing accuracy with respect to the number of epochs of proposed model, and Figure 7 describes the relation between training and testing loss of proposed model with number of epochs.

Figure 8 depicts sample outputs (predictions) from proposed model 1, in which Predicted crowd density is 'Medium'.

Confusion matrix for model:

In Figure 9, confusion matrix is shown with parameters 'true value' and 'predicted value'.

76	0	8
0	158	12
7	5	162

Figure 9

Confusion matrix for model

Classification matrix for crowd density estimation model

	precision	recall	f1-score	support
0	0.92	0.90	0.91	84
1	0.97	0.93	0.95	170
2	0.89	0.93	0.91	174
accuracy			0.93	428
macro avg	0.93	0.92	0.92	428
weighted avg	0.93	0.93	0.93	428

Figure 10

Classification matrix for model

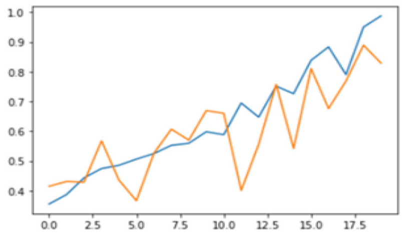

Figure 11

Plot of training(validation) (orange) and testing (blue) accuracies

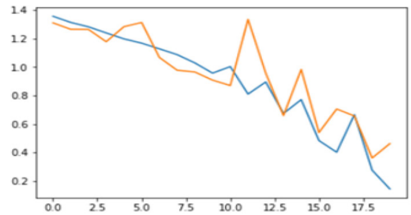

Figure 12

Plot of training(validation) (orange) and testing (blue) losses

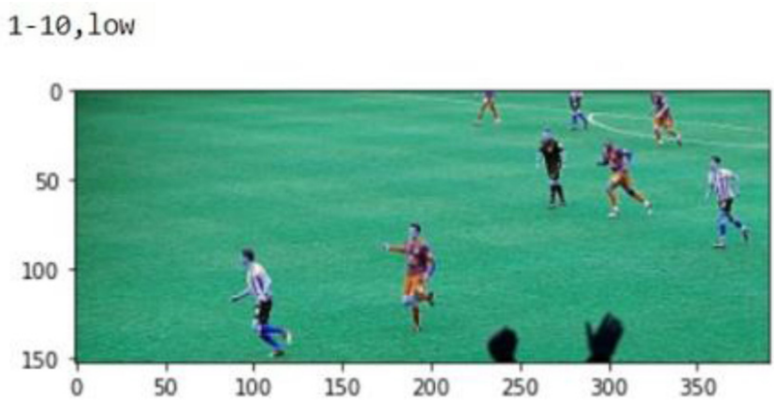

Figure 13

Result Output (prediction)

In Figure 10, classification matrix is shown with parameters like precision, recall, F1-score etc.

Various results of training and testing with validation, accuracies and losses for Model 2 are plotted here:

Figure 11 depicts the variation of the training and testing accuracy with respect to the number of epochs of proposed model, and Figure 12 describes the relation between training and testing loss of proposed model with number of epochs.

Figure 13 depicts sample outputs (predictions) from proposed model 2, in which Predicted Crowd count is between ‘1-10 ’, crowd density is ‘Low’.

Table 2 demonstrates the results for testing accuracy, validation accuracy, testing loss for proposed model 1 & model 2.

Table 2

Efficiency matrices

Models (Direct Approaches)	Model 1	Model 2
Testing Accuracy	.9924	.9873
Validation accuracy	.9182	.8287
Testing Loss	.0126	.1458

5. Conclusion and Future Scope

After studying the available literature, it has been concluded that no single model that provides crowd density estimation with crowd count estimation for an image at the same time. So, a model that provides crowd density estimation with crowd count estimation together for an image is developed. The work is completed in stage wise designing models. Each model serves individually as well as forms steps for next stage. The accuracy of model 1 developed in stage 1 is 0.9994, the accuracy of model 2 developed in stage 2 is 0.9873. If we compare proposed model with existing one (Yongsang Yoon et al. (2016) based on conditional point process-based crowd count method, it works for sparse and moderate crowd with average accuracy of 0.81. The proposed model is evaluated on the standard dataset. The experimental results show the effectiveness of the proposed model. The Proposed work can be extended for videos as the work done here, is for images only. More crowd classes can be added in the present work like more density, count and mask classes etc.

References

[1] Thanh-Sach Le, Chi-Kien Huynh, Human-crowd density estimation based on gabor filter and cell division, *Advanced Computing and Applications (ACOMP)*, IEEE, 157– 161, (2015).

[2] S. Abdulla M. Saleh, Shahrel Azmin Suandi, and Haidi Ibrahim, Recent survey on crowd density estimation and counting for visual surveillance, *Engineering Applications of Artificial Intelligence*, (41):103–114, (2015).

[3] V. Chauhan, S. Kumar, S.Kumar Singh, Human count estimation in high density crowd images and videos, *Parallel, Distributed and Grid Computing (PDGC)*, IEEE, 343–347, (2016).

[4] Yaocong Hu, Huan Chang, Fudong Nian, Yan Wang, and Teng Li, Dense crowd counting from still images with convolutional neural networks, *Journal of Visual Communication and Image Representation*, 38:530–539, (2016).

[5] Ven Jyn Kok, Mei Kuan Lim, Chee Seng Chan, Crowd behavior analysis: A review where physics meets biology, *Neurocomputing,* 177:342–362, (2016).

[6] Muhammad Saqib, Sultan Daud Khan, Michael Blumenstein, Texture-based feature mining for crowd density estimation: A study, *Image and Vision Computing New Zealand (IVCNZ)*, 1–6, (2016).

[7] Yongsang Yoon, Jeonghwan Gwak, Jong-In Song, Moongu Jeon, Conditional marked point process-based crowd counting in sparsely and moderately crowded scenes, *Control, Automation and Information Sciences*, IEEE 215–220, (2016).

[8] Shiliang Pu,Tao song ,Yuan Zhang,Di Xie, Estimation of crowd density in surveillance scenes based on deep convolutional neural network, *Procedia Computer Science* 111:154–159, (2017).

[9] Poonia, Ramesh C., and Linesh Raja. "Smart Technologies in Engineering." Recent Advances in Computer Science and Communications (Formerly: Recent Patents on Computer Science) 13.6 : 1172-1172 (2020).

[10] Patel, Hima, et al. "Neural network with deep learning architectures." *Journal of Information and Optimization Sciences* 39.1 : 31-38 (2018).

[11] Qi Wang, Junya Gao, Wei Lin,Xuelong Li., NWPU-Crowd: A Large-Scale Benchmark for Crowd Counting and Localization, *IEEE Transactions on Pattern Analysis and Machine Intelligence*,43(6): 2141 – 2149, (2021).

[12] Zhang Q, Chan AB, Wide-area crowd counting: Multi-view fusion networks for counting in large scenes, *International Journal of Computer Vision* 130(8) : 1938–1960, (2022).

[13] Xiong H, Lu H, Liu C, Liu L, Shen C, Cao Z, from open set to closed set: Supervised spatial divide-and conquer for object counting, *International Journal of Computer Vision* 131(7) : 1–19, (2023).