Estimating Human Mass Gathering on a Particular Time and Space Estimation by using Machine Learning

Vijay Kumar Sinha, and Shruti Aggarwal

Department of CSE, Chandigarh University, Punjab, India

Email: prof.vksinha@gmail.com ; drshruti.cse@gmail.com

Abstract

With the expanding populace, evaluating swarm thickness is a typical issue for swarm observation in Computer Vision. This issue stays a difficult assignment because of various varieties in scale issues created by various blocked uproars, changing shapes, and point of view variety. To handles these difficulties and to give a decent condition of precision we, in this way, center to gather a tremendous measure of datasets with shifting thickness levels and manufacture an Allied-CNN model. The assortment of the datasets is done from different sources like YouTube and some genuine recordings. The Allied-CNN model is worked in python and prepared on a named dataset of thousand item pictures from different points of view, for deciding thickness levels (as low thickness, medium thickness, and high thickness). Preparing results for thickness estimation show the preparation set precision arrives at 94.8%, the greatest approval exactness of just 88% is accomplished. Along these lines, this model aids in ordering a picture as low thickness, medium thickness, and high thickness. Estimations on this group datasets show that the proposed Allied-CNN performs modest outcomes contrasted with the cutting-edge strategies.

Keywords: Community, Modelling, Neural Network, Machine Learning, convolution neural network, perceptron

1 Introduction

Security has become a significant worry in concerning network wellbeing particularly in managing the mass social event in Malls, open fairs, or political meetings or discussing any open get-together occasions. These occasions viewed as exceptionally basic as they can in general structure a huge catastrophe if legitimate and severe group the board handled [1]. We have seen different examples where such casualty being accounted for in news because of groups gathering. As a pace of development of populace expanding it's turning into a test to comprehend and handle the group conduct adequately particularly for the security organizations.

DOI:
https://doi.org/10.54216/AJBOR.050101
Received: March 22, 2021 Accepted: August 29, 2021

A modest computerized machine-driven framework needs to fabricate with the goal that any observations can distinguish irregular conduct changes in the thickness levels startlingly. These strange changes may then be accounted for to the security group with the goal that legitimate move can be made. Checking and enhancing a group having shifting densities may embroil the location of individual people in the group. The primary testing assignment of checking such variety in the thickness possibly due to impediment, scale and viewpoint variety, and directions. Current following calculations show the issues in view of displaying foundation situations, between object-impediment that treat the test as a major issue to body shape models.

Different Researchers proposed various strategies have given techniques dependent on surface highlights, glob-channels, and so on. There have been exercises for head acknowledgment dependent on foundation demonstrating, omega models for head-shoulder-location, etc.

The Conventional strategies are not proper for a reconnaissance perspective. Numerous mishaps and rushes being accounted for in the loss of life over the world. A portion of the occurrences announced as of late during the mass occasion are shown in Figure 1 [2]. Demonstrating of the group is an extremely intense and complex undertaking as the group has the two elements and mental attributes. Henceforth there is a requirement for evaluating swarm thickness and characterize the conduct of the group in most reconnaissance recordings. Additionally, that can be reached out for future related occasion arranging and legitimate strategies can be applied for space structure for untamed life, traffic-related issues. The primary commitments of this examination are:

- Propound new datasets of the group with three thickness levels with various situations from this present reality.

- Propound another half and half traditional neural system with a completely associated layer to create the thickness of a scene utilizing the highlights removed from the mixture Allied-CNN model.

- Compare and deduce the best ideal classifier model that gives the best outcomes when we prepared with profound portrayals.

- Compare the productivity of the proposed structure with other cutting-edge strategies from the writing.

The remainder of the paper is sorted out as follows. Writing surveys of the related assignment of group thickness estimations are introduced in Sec. 2. In Sec. 3 the essential ideas of the Convolutional neural system, layers of CNN, how to limit the capacity, or upgrade the model. Sec. 4 gives the blueprint of the Datasets. Sec. 5 examined the proposed engineering of Allied-CNN that we worked in python programming. Sec. 6 depicts the reproduction arrangement and results with an appropriate chart. Ultimately, the paper is closed in Sec. 7

2. Literature Review

Numerous bits of exploration have been tended to these Crowd thickness estimations (CDE) issues. A portion of these surveys are examined as follows. Vishwanath [3] suggested a system dependent on CNN that includes learning by acclimatizing an elevated level-earlier (HLP) into the system. Utilizing Class mark HLP classifier ready to foresee the quantity of individuals in the information picture subsequently adapts exceptionally discriminative worldwide highlights. Muhammed Anees [4] utilize the profound learning system that fused an Inception Model, a module concocted by GoogleNet, for the thickness estimation of Crowded Videos containing two sections to be specific, highlight extraction and order assignments. LSTMN, a gathering of layers of an intermittent neural system, is utilized as the subsequent model to group the info recordings. Jianing Qiu [5] send profound learning for thickness estimation and including of individuals in pictures by utilizing Two-Column CNN gets from VGG-16 and Alexnet. Shiliang Pu [6] to a particular thickness of groups as "low, low, medium, high and high", utilizes GoogleNet, that balances between

DOI:
https://doi.org/10.54216/AJBOR.050101
Received: March 22, 2021 Accepted: August 29, 2021

the development of stature of the system and at the same time, the width of the system. These Models named "Beginning" which shapes the premise in PC vision. Xiaohang Xu [7] proposed a group learning method that includes tallying head width and incorporates the Support Vector relapse for assessing swarm thickness estimation. D. Kim [8] utilizes an Adaboast Algorithm as an apparatus for extricating close an element extraction to improve the getting to time length and time span. A.K. Pai [9] proposes a surface highlights approach for CDE that are Local-Binary Pattern, utilized in determining the highlights, and Gabor highlights are separated after playing out the numerical activity of the information picture with heaps of LOG-Gabor channels are determined at different extents and directions. [10],[12],[13] suggested a model of CNN based system for figuring the quantity of individuals and assessing thickness. They attempted to dispose of scale variety brought about by unpredictability and stature, a numerous goal misfortune that will lessen slope issue, thickness variety by applying different errand securing way to deal with train a system. N. Ilyas [11] recommended an engineering CASA that means to give inside and out highlights by various survey scales and touchy highlights by including an extended convolution with various channel measures that will extricate relevant data. Anuruddha L. [14] presents another novel calculation "various person on foot following" (MPT), a module-based programming framework for video reconnaissance that dispenses with observation frameworks lacking outsider extensibility, not have the option to have occasion discovery and have no type of irregularity location in exceptionally thick jam-packed scenes. Li Baiping [15] attempts to determine the issues progressively CDE including manual computation, managing insights techniques dependent on pixels of the frontal area picture. They utilize key-outline extraction innovation to assess Real-time CDE by choosing the CNN engineering of GoogleNet by looking at their exhibitions of AlexNet, VGGNet, GoogleNet, and ResNet. Mr Pankaj Badatia [16] proposed in Multi-segment DM based on GAN to take care of the issue of regular impediment and mess issue yet with the compromise of spatial data.

3. Convolution Neural Network

The historical backdrop of CNNs started in the year 1958 with the perceptron calculation. The best type of portraying CNN as models. These models will in general learn Visual channels to get the profound improved highlights of the picture. This section gives the nuts-and-bolts usefulness of Neural Network that empowers us to comprehend in an initial step how CNN functions. The accompanying area contains the perceptron, multilayer-layer perceptron (MLP), layers of CNN, and a few rudiments of CNNs. We have additionally characterized the fundamental ideas utilizing these CNN approaches.

3.1 Deep Learning

Profound learning (DL) is a subcategory of AI and one of a specific type of counterfeit neural systems, where procedures invigorated by the human mind. It gains from a lot of information. DL causes any mind-boggling issue to get settled in any event, when utilizing an informational collection that is exceptionally different, unstructured, and between associated.

As of late, DL has demonstrated extraordinary accomplishment in the PC vision field. DL gives preferred learning capacity over AI and along these lines helps in getting great utilization of datasets to remove spatial and worldly highlights from the content, pictures, and recordings.

3.2 The Perceptron

The Perceptron designed by Rosenblatt was the first Neural network in 1958. This neural network was made in order to simulate human learning. The concept was that the brain functioning carried out using huge Neurons and synapses. So to build a large neural network, perceptron was considered a single neuron.

DOI:
https://doi.org/10.54216/AJBOR.050101
Received: March 22, 2021 Accepted: August 29, 2021

Figure1 Crowd disaster in India

Consider some dataset arrange in logistic regression. Let input set S= R^d and Output set of T={ 0, 1}

Then the perceptron equation can be formed as:

$$H = \{d(s_1, s_{2,} \dots, s_d)$$

$$= f(\, l_0 + s1 * l_1 + s2 * l_1 \dots + sd * l_1) \qquad \dots (1)$$

where f is given by

$$f(s) = \begin{cases} 1 & s > 0 \\ 0 & s \le 0 \end{cases}$$

Here neuron's dendrites are modelled as weights (l_p). These are multiplied with input values (s_p). Summation of these multiplied values with bias (l_0) passed through the activation function to produce output. Thus our theory functions $f(s)$predict an input to be 1 if their linear combination with the given weights is greater than 1, and 0 otherwise. Conversely, while classifying logistic regression and the perceptron, this means that we can only correctly classify an entire data set if there exists a linear boundary between the two classes! This will motivate the generalization of the perceptron, the multilayer perceptron.

3.3 Neural Network

Since a single perceptron was unable to handle the numerous representational capabilities, a multi-layer perceptron comes into play. In the neural network (NN), various neurons are arranged in layers, each with its tunable weights and

DOI:
https://doi.org/10.54216/AJBOR.050101
Received: March 22, 2021 Accepted: August 29, 2021

biased. Each layer of the neural network receives some input, perform dot product, and passes it to the hidden or output layer [16]. The result obtained from the output layer is applied to the activations function that yields the input-output relationship. Combinations of inputs, output layers, and various hidden layers make the network complex and heavy in size. Fig illustrates this layered architecture.

In general, the layers are fully connected, meaning that every neuron is connected to each neuron of the previous and next layer. However, neurons of the same layer are not connected. The number of units in each layer and the number of layers must be chosen and result in networks that approximate functions of varying complexity. It can be difficult to find a good architecture, as using too many units and layers may result in a complex function that overfits training data and choosing too few units can produce a biased function that is too simple to represent the data distribution well.

Convolutional Neural Network (CNN) is a dedicated part of Neural Network (NN) which deals with Image Classification. They are highly in demand since they are resilient to any form of disturbance in the image and produces excellent results for the complex image classification related work. Neural Network only holds a finite number of hidden layers sandwiched between the input and output layers, but CNN has a very large number of hidden layers. This increases the size of the network and time consuming but the accuracy that obtained is very high in magnitude. Thus, CNN allows the train of the model using datasets and categorizes the datasets into different classes. CNN can also make use of different activation functions such as Rectified Linear Units (ReLu) that allows removing non-linearity in the datasets.

The structure of CNN is determined by the number of layers used in the network and the activation functions used in the layers.

$$S^{(x)} = \begin{cases} \emptyset_{(x)}(S^{(x-1)}, & if\ x \geq 1 \\ t_p & x = 0 \end{cases} \quad \ldots (2)$$

where, $S^{(x)}$: the output vector representing the outputs of nodes in layer (x) ;t_p : pth input data ;

$\emptyset_{(x)}$: function determine $S^{(x)}$given $S^{(x-1)}$

The first layer, i.e. Convolution is the process of selecting important features from the given image using some mathematical operation. Convolution is performed so that selected features will hold a relation among the pixels by choosing only limited squares of the input image.In conventional CNN, the Pooling Layer(PL) helps in diminishing the parameters when the size of the images is large enough in a NN. Pooling helps in reducing the dimensionality of each image by keeping essential. The common types of Pooling are Max Pooling (MP) layer, Average Pooling (AP) layer & Sum Pooling (SP) layer.Next, comesFully Connected layers (FCL) in which the features matrix is transformed into a feature vector that is carried out & put it into an FCL. The FCL merges these feature vector to produce a Classifier Model on which output is classified using some activation function.

$$S^{(x)} = \emptyset_{(x)}^{(FC)} * \left(S^{(x-1)}\right) = \left(S^{(x-1)}\right)^F w^{(x)} + (b^o)^{(x)} \quad \ldots(3)$$

DOI:
https://doi.org/10.54216/AJBOR.050101

Received: March 22, 2021 Accepted: August 29, 2021

where,$S^{(x-1)}$: output vector representing the output of nodes in layer x and $\emptyset_{(x)}^{(FC)}$: function to determine $S^{(x)}$when $S^{(x-1)}$ is given.

Grouping of FCLs can lead to an increase in the depth of a network. Simultaneously, we are increasing the approximation quality of a network. As defined in equation (2) is a linear function. Thus, by stacking multiple linear FCL, we finally achieve a linear model. Therefore, Non-linearity can be achieve using activation function to the outputs of f(x). Common activation functions are tanh() , sigmoid or ReLU.

The result of every layer, starting from the input is calculated as

$$O = \{\emptyset_d(o^{d-1}|d \in [1,2, \dots\dots, D]\} \qquad \dots(4)$$

The quality of an approximation can be shown by the term name 'Loss' or a 'Cost function'. For example, if we take a retailer who had purchased a product for Rs 50,000 and buts sells it for Rs 40000 then there would be a loss margin of Rs 10,000. And if the customers buy the same product at Rs 60,000 then also the customer will not go to purchase it because there will be a loss of Rs 10,000.

Figure 2 : Sample Images from the datasets

So, the retailer will try to have minimum loss would occur with maximum profit. Based on this, we can assume two properties of the cost function

- Non-Negative Value
- Lesser loss is better if two approximations need to compare.

To have good quality approximations, our aim will be better utilization of parameters in the CNN. All their various other optimizers present that work differently; we discuss the Adam Optimizers. Stochastic gradient descent (SGD) is an optimizer that maintains only one learning rate for all weights updates. During the training of the neural network, the learning rate does not change.

A learning rate defined for each network weight and separately adapted as learning unfolds. In this research perspective, we used Adam optimizer since this is the combination of extended versions of SGD which are the Adaptive Gradient algorithm (AdaGrad) and Root mean square propagation (RMSProp).

DOI:
https://doi.org/10.54216/AJBOR.050101

Received: March 22, 2021 Accepted: August 29, 2021

- AdaGrad sustains a per-parameter learning rate that upgrades performance on issues with limited gradients (e.g., computer vision problems).
- RMSProp also upholds per-parameter learning rates that are amended by looking at the mean obtained by the magnitudes of the weight of gradients.

Adam utilizes the merits of both the algorithm. Precisely, Adam optimizer can be expressed as:

$$w_k = w_{k-1} - \alpha \frac{p_k}{\sqrt{q_k} + \epsilon} \quad \ldots (5)$$

where w represents weights of the model at kth iteration, α denotes the step size of the iteration, p_k *and* q_k are the bias estimators for the gradients. [17]

4. Different Density Image Dataset

The fundamental strategy was to basically fabricate a decent arrangement of the database that will hold pictures from different foundation. Since the publically accessible dataset isn't appropriate for our errand of Density Estimation. In this way, our objectives of making the dataset gathered from different sources. At first, with the assistance of python content pictures downloaded from the Google Search Engine utilizing different settings like meetings rallies, gatherings, outdoors, and some more. Picking the setting of different structures empowers us to have datasets from different foundations. At that point we scan for recordings from different locales like recordings from YouTube, Videezy [18], and some continuous recordings are gathered. As our assignment, we pick an administered approach for CDE. Naming and Categorization of Images dependent on tallying boundaries are finished. We marked '1' as Low-Density Level, '2' as Medium Density Level, and '3' as High-Density Level. Our Datasets contain around 10000 Images.

Table I: Density level annotation scheme used during the construction of different density-level crowd

Density Level	Minimum Number of People Count	Maximum Number of People Count	Category
1	0	10	Low Density
2	11	100	Medium Density
3	101	>101	High Density

The following steps were performed before processing the images for Model Evaluation and Classification.

- Downloaded videos are sliced out using python script.
- These downloaded videos aid in producing about 204 frames using ffmpeg utility command.
- *Classification of Images*: All the frame images are classified individually as Low Density, Medium Density, and Highly Density.
- Then these classified images are again classified into training and testing datasets in the ratio 7:3

5. Proposed Architecture: Allied-Cnn

DOI:
https://doi.org/10.54216/AJBOR.050101
Received: March 22, 2021 Accepted: August 29, 2021

This exploration paper centers around the "Start to finish Deep Network" which came out from the possibility of Inception, ResNet18, RseNet50, and VGG16 model. Roused by the above named pre-prepared models, we stack different convolution layers in a profound CNN for this CDE issue. Basically it includes the Convolution layer (CL), PL, FCL, and Dense Layer(DL).

Since the convolution activity separates exceptionally significant relevant data of the info. The information picture of M*N*D size is convolved through the 3 layers of CL with down-examining layers every which target separating the essential highlights utilizing various channels. The M*N*D picture is convolved with K (M*M*D) portions bringing about K highlights. From the upper left of the info picture to upper right, every portion is moved in a left-to-right path each in turn. This rehashes for the following column also.

Since CNN is viewed as tough to clamor, in this manner these highlights drop by utilizing PL. In PL, this layer assumes a vital job is dropping the determination of the highlights against the commotion inclination. For this angle, we have utilized max-pooling that picks the most extreme incentive from the subsampling. CNN essentially dependent on the non-Linear work that will demonstrate diverse distinguishing proof of highlights in the shrouded layer. At every CL it is consolidated with the Non-Linear-Activation Function named ReLU "Amended Linear Unit". This is characterized as the capacity as,

$$y = \max(x, 0) \quad \text{…. (6)}$$

ReLU function helps in increasing the non-linear properties of the overall Network. Finally, an FCL will perform the sum of previous layers weights. All the features' elements of the previous one get used in this FCL to give Output.

5.1 Augmentation of data

Data Augmentation is the procedure that is used in Deep Neural Network so the training datasets can be modified with various forms of Images that will belong to the same class of the Original Datasets Image. This makes the transformation of Image datasets for a model of deep network with tons of variations that can simultaneously mend the ability to fit inside the model to generalize what new had been learned. The main mode of Augmentation can be like Shifting in the image, zooming, rotating, scaling up, flipping horizontal or vertical, increasing brightness of the image. Building CNN with data augmentation can help in changing the invariant approach and technique to train the image. This technique is not limited to training datasets but also for testing sets as well.

5.2 Architecture Details

The process of creating a deep model is chosen by sparsely connected network architecture with different filters. The image of size 224 x224 is given input through a heap of convolutional (conv.) layers of the Allied-CNN model. Allied-CNN model is a combination of 1*1 CL, 3*3 CL, 5*5 CL, and 7*7 CL and the activation function ReLU is used at each convolution layer. The order of kernels used in Allied-CNN can be seen in figure 2. Then the features are convolved over the Low-level Feature Extraction (LLFE).

LLFE is the parallel combination of 1x1 CL, 3x3 CL, and max pooling is used for concatenation. Further, these merged features are passed for Batch-Normalization and allied combinations of kernels are used. Then again, these features are passed over Medium Level Feature Extraction (MLFE). Like LLFE, these are the same but different kernels of 3x3, and 5x5 are used to get the medium level spatial feature of the image. After these, these medium

DOI:
https://doi.org/10.54216/AJBOR.050101
Received: March 22, 2021 Accepted: August 29, 2021

levels are again convolved over the third feature extraction part termed as High-Level Feature Extraction (HLFE). These are also similar to LLFE, MLFE but with the different kernels are used ie 1 x1 and 5x5. These are again operated on the Batch-Normalization process. Next, we passed these features set to Feature Fusion Level (FFL). FFL includes the kernel size 1 x1 with filters used in decreasing order. Then we go for FCL with the activation function Softmax is used with the parameters illustrated in table II where the crowd density classification is processed.

5.3 Density-Based Classification: Since we have three density levels described in Table I are mutually exclusive a more conventional classification approach is taken for the classification of density in the crowded images using the Softmax activation applied to the Density output and a categorical cross-entropy loss is minimized.T'_{pq} denotes to the predicted probability of category of class q on sample image p while T_{pq} denotes the predicted probability of the category of class q on sample image p for the ground truth.

$$D = -\frac{1}{d}\sum_{p=1}^{d}\sum_{q=1}^{3}(T_{pq})\log(T'_{pq})(6)$$

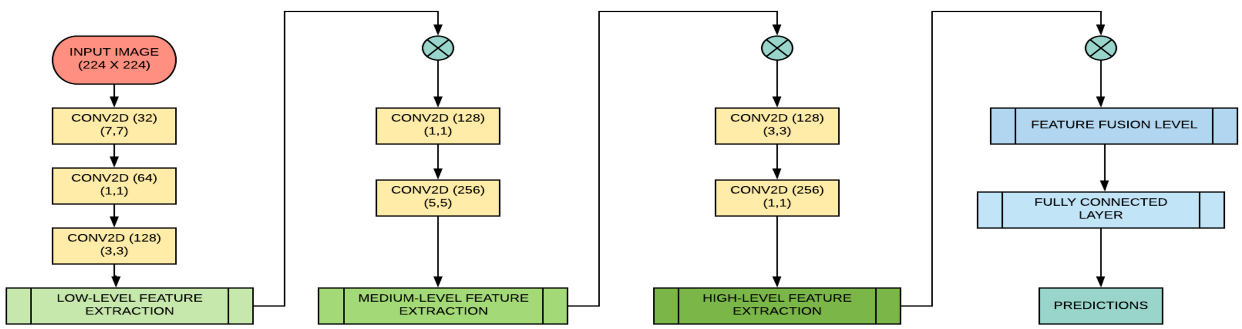

Figure 3. Architecture of Allied-CNN

6. Experimental Results

6.1 Implementation Details

Our model is operated on Intel(R) Core (TM) i5-8250U CPU @ 1.60GHz-1.80GHz with 8.00GBRAM, 64-bit Operating System, x64-based processor. The experiment is carried out using a Python 3.7.3 framework. Table 4.2 lists the parameters used during the training of the model. The batch size is set to be 32 during training. This architecture Allied-CNN is carried out and trained to 40 epochs. The activation functions are rectified linear Unit and for performing classification Softmax function is used.

Table 2: Parameters used in the Training Model

Parameters	**Value**
Learning Rate	0.001
Epochs	40

DOI: https://doi.org/10.54216/AJBOR.050101
Received: March 22, 2021 Accepted: August 29, 2021

Learning Policy	kept fixed

CDE provides a relationship between the image of a crowd and the density. The standardized form of this relationship helps to estimate density from the new image or videos. For the research aspect, we prepare the datasets described in Sec III, find the feature by using Allied-CNN and then analyze the feature map of the input image.

The experiment uses prepared datasets. In totals there are 12,589 images are chosen as the datasets that include different view angles and perspectives. These images are split into a training set with 9604 images and a testing set with 2985 images. In both, training and validation set, all data is categorized into 3 categories, labelled as 1, 2, and 3 the classification standard is shown in Table 1. The platform is Keras and Tensorflow is chosen.

6.2 Results

Train Network by using training data converges after the training accuracy reaches to 94.22% and the testing accuracy reaches to 88.14% with 40 iterations. The parameters used in training Allied-CNN is given in Table II. After training the performance of the Allied –CNN model is illustrated using Figure 5. Figure 7 shows the graph of loss and iteration. The accuracy of Allied-CNN can be illustrated in Figure 6. The accuracy and loss of Allied –CNN can be seen in Figure 4. Even the training and validating the Training dataset it takes is 17.5s per image.

The Allied-CNN model is being collated with exiting works available in the literature, especially that have performed density-based classification. Since we have created a contemporary database. The information being collated is analyzed. Shiliang Pu [19] acquired 91.73% accuracy for the 31 crowd Subway-carriage scene dataset to estimate the crowd density. [20] uses CNN and got 93.6%, 94.8%, and 76.2% accuracy for three different datasets PETS 2009, a Subway video, and a video of Chunxi Road in Chengdu respectively. Thusour Allied-CNN method performs better than these listed models in the literature.

DOI:
https://doi.org/10.54216/AJBOR.050101
Received: March 22, 2021 Accepted: August 29, 2021

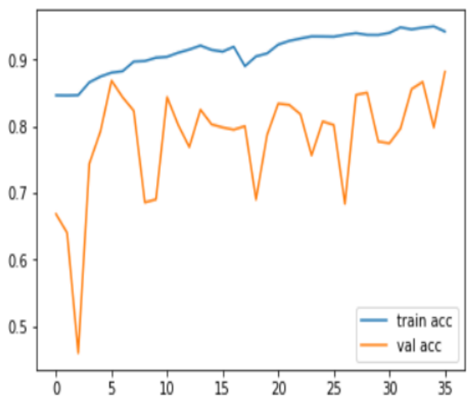

Figure 4 Plot of Train vs Validation Accuracy of Allied-CNN

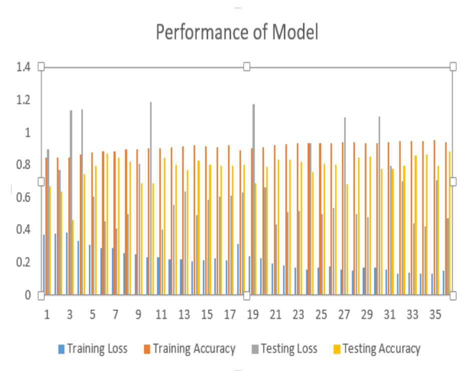

Figure 5 Performance of Allied-CNN

7. Conclusion

In this research paper, we propose a calculation for the human gathering estimation by using machine learning. The proposed strategies depend on the strategy LLFE, MLFE, HLFE, and FFN that helps in evaluating the thickness. Additionally, we attempted to fabricate the group datasets that could accomplish the objective of CDE. We assemble the Allied-CNN and train the engineering for around 40 ages, and we had the option to accomplish 94.8% preparing exactness. Assessments on this group datasets show that the proposed Allied-CNN performs serious outcomes contrasted with the cutting-edge techniques. Further work would embrace this system for swarm conduct investigation by making a few changes in accordance with this system.

References

[1] W. M. Shalash, A. A. AlZahrani and S. H. Al-Nufaii, "Crowd Detection Management System," 2019 2nd International Conference on Computer Applications & Information Security (ICCAIS), Riyadh, Saudi Arabia, 2019, pp. 1-8, doi: 10.1109/CAIS.2019.8769566.

[2] "Crowd Disaster in India" [Online] Available: https://blog.forumias.com/crowd-disasters-and-management-in-india/

[3] V. A. Sindagi and V. M. Patel, "CNN-Based cascaded multi-task learning of high-level prior and density estimation for crowd counting," 2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), Lecce, 2017, pp. 1-6, doi: 10.1109/AVSS.2017.8078491.

[4] M. V. Anees and S. G. Kumar, "Deep Learning Framework for Density Estimation of Crowd Videos," *2018 8th International Symposium on Embedded Computing and System Design (ISED)*, Cochin, India, 2018, pp. 16-20.

DOI:
https://doi.org/10.54216/AJBOR.050101
Received: March 22, 2021 Accepted: August 29, 2021

[5] J. Qiu, W. Wan, H. Yao and K. Han, "Crowd counting and density estimation via a two-column convolutional neural network," *4th International Conference on Smart and Sustainable City (ICSSC 2017)*, Shanghai, 2017, pp. 1-5.

[6] Shiliang Pu, Tao Song, Yuan Zhang, and Di Xie, "Estimation of Crowd Density In Surveillance Scenes Based On Deep Convolutional Neural Network". 8th International Conference on Advances in Information Technology, IAIT 2016, 19-22 December 2016, Macau, China.

[7] Xiaohang Xu, Dongming Zhang, and Hong Zheng, "Crowd Density Estimation Of Scenic Spots Based On Multi-Feature Ensemble Learning", Journal Of Electrical And Computer Engineering, Vol. 2017, Article Id 2580860, 12 Pages, 2017.

[8] D. Kim, Y. Lee, B. Ku and H. Ko, "Crowd Density Estimation Using Multi-class Adaboost," *2012 IEEE Ninth International Conference on Advanced Video and Signal-Based Surveillance*, Beijing, 2012, pp. 447-451.

[9] A.K. Pai, A. K. Karunakar and U. Raghavendra, "A Novel Crowd Density Estimation Technique using Local Binary Pattern and Gabor Features," *2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, Lecce, 2017, pp. 1-6.

[10] V. Huynh, V. Tran and C. Huang, "IUML: Inception U-Net Based Multi-Task Learning For Density Level Classification And Crowd Density Estimation," *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*, Bari, Italy, 2019, pp. 3019-3024.

[11] N. Ilyas, A. Ahmad, and K. Kim, "CASA-Crowd: A Context-Aware Scale Aggregation CNN-Based Crowd Counting Technique," in *IEEE Access*, vol. 7, pp. 182050-182059, 2019.

[12] Zeng, Xin & Wu, Yunpeng & Hu, Shizhe & Wang, Ruobin & Ye, Yangdong. (2019). DSPNet: Deep Scale Purifier Network for Dense Crowd Counting. Expert Systems with Applications. 141. 112977. 10.1016/j.eswa.2019.112977.

[13] Y. Tian, Y. Lei, J. Zhang, and J. Z. Wang, "PaDNet: Pan-Density Crowd Counting," in IEEE Transactions on Image Processing, vol. 29, pp. 2714-2727, 2020.

[14] A. L. Hettiarachchi, H. O. Thathsarani, P. U. Wickramasinghe, D. S. Wickramasuriya and R. Rodrigo, "Extensible video surveillance software with simultaneous event detection for low and high density crowd analysis," *7th International Conference on Information and Automation for Sustainability*, Colombo, 2014, pp. 1-6.

[15] B. Li, X. Han and D. Wu, "Real-Time Crowd Density Estimation Based on Convolutional Neural Networks," *2018 International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS)*, Xiamen, 2018, pp. 690-694.

[16] P. Badatia and P. P. Tasgaonkar, "Crowd counting and Density Estimation using Multicolumn Discriminator in GAN," *2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, Bangalore, 2018, pp. 1179-1183.

[17] Pavel Golik, Patrick Doetsch, and Hermann Ney. "Cross-entropy vs. squared error training: a theoretical and experimental comparison". In Interspeech, volume 13, pages 1756–1760, 2013.

[18] "Crowded People "[Online] Available: https://www.videezy.com/free-video/

[19] M. Fu, P. Xu, X. Li, Q. Liu, M. Ye, and C. Zhu, "Fast crowd density estimation with convolutional neural networks," Engineering Applications of Artificial Intelligence, vol. 43, pp. 81–88, 2015.

[20] P. V. V. Kishore, R. Rahul, K. Sravya, and A. S. C. S. Sastry, "Crowd Density Analysis and tracking," in Proceedings of the International Conference on Advances in Computing, Communications and Informatics (ICACCI '15), pp. 1209–1213, India, August 2015.

[21] M. Kumar, P. Mukherjee, K. Verma, S. Verma and D. B. Rawat, "Improved Deep Convolutional Neural Network based Malicious Node Detection and Energy-Efficient Data Transmission in Wireless Sensor Networks," in IEEE Transactions on Network Science and Engineering, doi: 10.1109/TNSE.2021.3098011.

[22] Tanvi Sharma, et al.. (2017) Intelligent Heart Disease Prediction System Using Machine Learning: A Review, International Journal of Recent Research Aspects, ISSN: 2349-7688, Vol. 4, Issue 2, pp. 94- 97.

[23] Akshat Srivastava et al, Analysis of Quality of Service in VANET, 2020 IOP Conf. Ser.: Mater. Sci. Eng. 993 012061

DOI:
https://doi.org/10.54216/AJBOR.050101
Received: March 22, 2021 Accepted: August 29, 2021

[24] Loveleen Gaur, Gurmeet Singh, Arun Solanki, Noor Zaman Jhanjhi, Ujwal Bhatia, Shavneet Sharma, Sahil Verma, Kavita, Nataša Petrović, Muhammad Fazal Ijaz, and Wonjoon Kim, Disposition of Youth in Predicting Sustainable Development Goals Using the Neuro-fuzzy and Random Forest Algorithms, Article number: 11:24 (2021)

[25] Monica Sood, et.al."Optimal Path Planning using Swarm Intelligence based Hybrid Techniques" Journal of computational and theoretical nanoscience (JCTN), ASPBS publisher. Vol. 16 No. 9, 2019, pp. 3717–3727, DOI:10.1166/jctn.2019.8240.

[26] Kaur Manjit; et al. "Flying Ad-Hoc Network (FANET): Challenges and Routing Protocols" Journal of Computational and Theoretical Nanoscience, Volume 17, Number 6, June 2020, pp. 2575-2581(7), https://doi.org/10.1166/jctn.2020.8932

[27] Ghosh, Gopal; et al. 'Internet of Things based video surveillance systems for security applications' Journal of Computational and Theoretical Nanoscience, Volume 17, Number 6, June 2020, pp. 2582-2588(7) https://doi.org/10.1166/jctn.2020.8933

[28] Gopal Ghosh, et al. 'A Systematic Review on Image Encryption Techniques' Turkish Journal of Computer and Mathematics Education, Vol.12 No.10 (2021), 3055-3059 M. Balazinska et al., "Data management in the worldwide sensor web," IEEE Pervasive Comput., vol. 6, no. 2, pp. 30–40, 2007, doi: 10.1109/MPRV.2007.27.

[29] A. Hussain et al., "A Resource Efficient hybrid Proxy Mobile IPv6 extension for Next Generation IoT Networks," in IEEE Internet of Things Journal, doi: 10.1109/JIOT.2021.3058982.

[30] Gandam A, Sidhu JS, Verma S, Jhanjhi NZ, Nayyar A, Abouhawwash M, et al. (2021) An efficient post-processing adaptive filtering technique to rectifying the flickering effects. PLoS ONE 16(5): e0250959. https://doi.org/10.1371/journal.pone.0250959

DOI:
https://doi.org/10.54216/AJBOR.050101
Received: March 22, 2021 Accepted: August 29, 2021