CROWD COUNTING IS TO ESTIMATE THE NUMBER OF PEOPLE IN THE IMAGE WITH THE DEVELOPMENT OF CONVOLUTIONAL NEURAL NETWORKS

Mr. SK.MAHABOOB BASHA[1],
MUKALA NITISH REDDY[2], VAJRALA PAVANI[3], RAYALA PRAVALLIKA[4]

[1]Associate Professor , Dept. of INFORMATION TECHNOLOGY, NRI INSTITUTE OF TECHNOLOGY , A.P., India.
[2,3,4]Student, B.Tech (IT), NRI INSTITUTE OF TECHNOLOGY , A.P., India.

Abstract — In the study of crowd counting, this paper first proposed the method of adding deconvolution to multi-column network for the problem that the population size in the image does not affect the accuracy of counting, namely multi-column deconvolution neural network. Crowd counting is an important research topic in the field of computer vision. The multi-column convolution neural network (MCNN) has been used in this field and achieved competitive performance. However, when the crowd distribution is uneven, the accuracy of crowd counting based on the MCNN still needs to be improved. In order to adapt to uneven crowd distributions, crowd global density feature is taken into account in this paper. The global density features are extracted and added to the MCNN through the cascaded learning method. Because some detailed features during the down-sampling process will be lost in the MCNN and it will affect the accuracy of the density map, an improved MCNN structure is proposed. In this paper, the max pooling is replaced by max-ave pooling to keep more detailed features and the deconvolutional layers are added to restore the lost details in the down-sampling process.

INTRODUCTION

However, the largest pooling method used in pooling will result in loss of feature information. Authors of the paper [9] proposed to use a combination of shallow and deep networks to perform extensive data enhancement by sampling patches from multi-scale image representations. Although the problem of scale changes is solved, the number of parameters in the network is increased, resulting in network Performance has dropped dramatically. The paper [10] designed a U-net structured generation network, generated a density map from the input patch, and directly used the antagonistic loss to shrink the solution to a real subspace, thus weakening the fuzzy problem of density map estimation. A clearer density map is obtained, but the applicability of the perspective effect is poor. The paper [11] proposed to aggregate multi-scale convolution features extracted from the whole image into a compact single vector representation, which can be efficiently and accurately counted by the vector of local aggregation descriptors (VLAD), but with increased training. The difficulty of the training leads to inefficient training. In the paper [12], a crowd counting method for deep convolutional neural networks is proposed. The method uses the crowd count and density map to alternately return to two related learning objectives. The two learning objectives can be converted to each other, thereby improving the accuracy of crowd counting. Rate, but the application of this method is limited because it requires a perspective view of the image, which is not readily available during actual training and testing. Crowd counting is used to calculate the total number of people in images or video frames. The crowd counting methods can be divided into three categories: the direct counting method based on target detection, the indirect method based on feature regression and crowd counting based on deep learning. In the relevant researches based on target detection [1]–[5], Lin et al. [1] proposed to use Haar wavelet transform to extract the feature area of the head-like contour and build the SVM classifier to classify the feature area. Gardzinski et al. [2] proposed to use shape contour of body to achieve crowd detection and crowd density estimation. All of these methods are suitable for the scenes with low density crowd, but the detection accuracy will decrease in the case of high density crowd.

LITERATURE SURVEY

1.Zhang et al. proposed a cross-scene crowd counting model. It was trained alternately through two learning objectives, density map and global number. This algorithm is implemented based on single-column CNN. However, it is not suitable for the change in the scale of crowd. Zhang et al. proposed to use the MCNN with three branch networks for crowd counting. Different receptive fields were used in each branch network, and this improved MCNN could adapt to the change in the scale of the crowd. They also introduced a new dataset ShanghaiTech for crowd counting.

2.Boominathan et al. proposed to combine the features of shallow and deep convolutional neural networks to improve spatial resolution.

3.Sindagi and Patel proposed a multi-task network which combined the high-level prior with the density estimation.

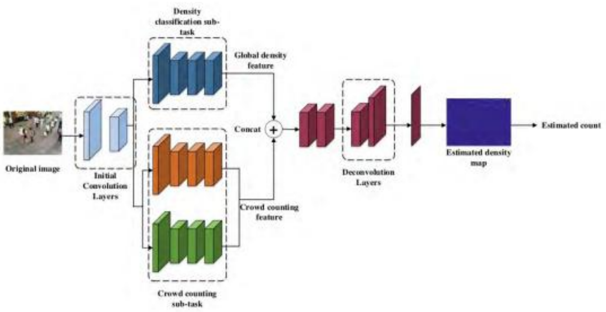
4.Sam et al. proposed Switch-CNN for crowd counting. In this network, a classifier was trained and an appropriate regressor was selected for input patches.

PROPOSED SYSTEM

Proposed a cross-scene crowd counting model. It was trained alternately through two learning objectives, density map and global number. This algorithm is implemented based on single-column CNN. However, it is not suitable for the change in the scale of crowd. The proposed to use the MCNN with three branch networks for crowd counting. Different receptive fields were used in each branch network, and this improved MCNN could adapt to the change in the scale of the crowd. They also introduced a new dataset ShanghaiTech for crowd counting,,

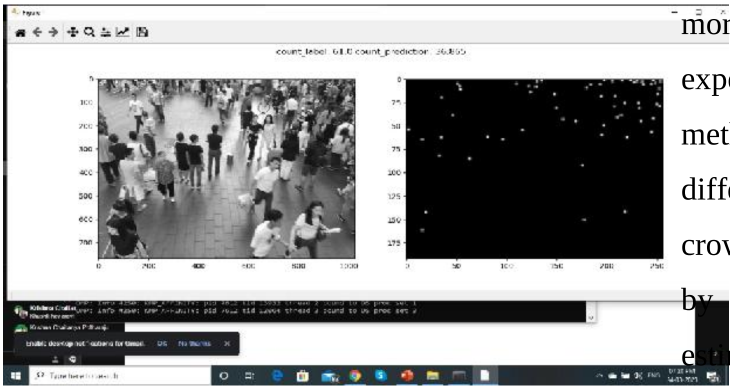
RELATED WORK

In this network, a classifier was trained and an appropriate regressor was selected for input patches. Shi et al. [20] proposed to aggregate multiscale features into



A compact single vector and used deep supervised strategy to provide additional supervision signal. Fu et al. [21] proposed to use the LSTM structure to extract the contextual information of crowd region. Liu et al. [22] proposed to add an attention module to adaptively select the counting mode used for different positions on the image. Yang et al. [23] proposed to use the MMCNN for robust crowd counting. In this work, the location, detailed information and scale variation were taken into account to generate density map in order to improve the robustness of crowd counting method. Generally, these algorithms have good performances in the crowd counting, but the performances of these methods were not effective when the crowd distribution is uneven [24], [25]. In order to solve the problem of inaccurate counting caused by uneven crowd distribution, the global density feature is extracted and used in this paper. A convolutional neural network with global density feature by using multi-task network cascades (MNCs) [18], [26] is proposed. In order to generate a more comprehensive density map, the max-ave pooling layers are used to keep more features of the image. Meantime, the deconvolutional layers are added to the convolutional neural network in order to restore the lost details in down-sampling process. It will help to improve the accuracy of density map and further improve the accuracy of crowd counting

SAMPLE OUTPUT SCREENSHOTS



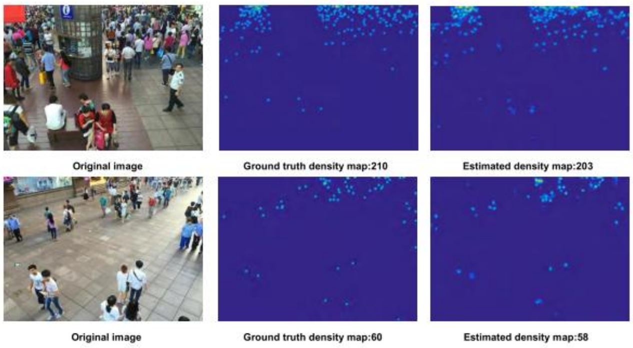
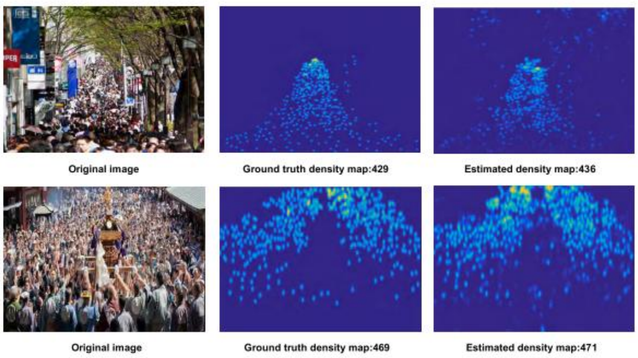


FIGURE 7. The ground truth and estimated density map of testing images in ShanghaiTech_PartB dataset.



CONCLUSION

In this paper, We proposed, an improved convolutional neural network combined with global density feature is proposed. It is different from existing crowd counting methods. The proposed method focuses on uneven crowd distribution. Moreover, the maxave pooling and deconvolutional layers are used to generate a more comprehensive density map. The experimental results show that the proposed method achieves competitive performance on different crowd datasets. Due to the high density crowd, some backgrounds will be taken as people by mistakes. It will bring about noise in the estimated density map and influence the counting results.

FUTURE SCOPE:

For the future work, we will focus on reducing the noise in the estimated density map and improving the accuracy of counting

References

[1] S.-F. Lin, J.-Y. Chen, and H.-X. Chao, ‘‘Estimation of number of people in crowded scenes using perspective transformation,’’ IEEE Trans. Syst., Man, Cybern. A, Syst., Humans, vol. 31, no. 6, pp. 645–654, Nov. 2001.

[2] P. Gardzinski, K. Kowalak, L. Kaminski, and S. Mackowiak, ‘‘Crowd density estimation based on voxel model in multi-view surveillance systems,’’ in Proc. Int. Conf. Syst., Signals Image Process. (IWSSIP), Sep. 2015, pp. 216–219.

[3] M. Li, Z. Zhang, K. Huang, and T. Tan, ‘‘Estimating the number of people in crowded scenes by MID based foreground segmentation and headshoulder detection,’’ in Proc. 19th Int. Conf. Pattern Recognit., Dec. 2008, pp. 1–4.

[4] T. Zhao, R. Nevatia, and B. Wu, ''Segmentation and tracking of multiple humans in crowded environments,'' IEEE Trans. Pattern Anal. Mach. Intell., vol. 30, no. 7, pp. 1198–1211, Jul. 2008.

[5] X. Kong, X. Song, F. Xia, H. Guo, J. Wang, and A. Tolba, ''LoTAD: Longterm traffic anomaly detection based on crowdsourced bus trajectory data,'' World Wide Web, vol. 21, no. 3, pp. 825–847, May 2018.

[6] Y.Zhang, D. Zhou, S. Chen, S. Gao, and Y. Ma. Single image crowd counting via multi-column convolutional neural network.[J]. In Proceedings of the IEEE CVPR, 2016, 23(3):589-597

[7] DB. Sam, S. Surya, and R. V. Babu.. Switching convolutional neural network for crowd counting.[J]. CVPR, 2017., 2017(1):1-7

[8] SHI Z L, YE Y D, WU Y P, et al.Crowd Counting Using Rank-based Spatial Pyramid Pooling Network[J].Acta Automatica Sinica, 2016, 42(6):866-874

[9] Xialei Liu, JWeijier.. Leveraging Unlabeled Data for Crowd Counting by Learning to Rank[J].CVPR, 2018, 24(4):7661-7669

[10] SINDAGI V A, PATEL V M.Cnn-based cascaded multi-task learning of high-level prior and density estimation for crowd counting[J].Advanced Video and Signal Based Surveillance (AVSS), 2017, :1-6

[11] ZShi, L. Zhang, Y.Sun, and Y.Ye.. Multiscale Multitask Deep NetVLAD for Crowd Counting [J].In IEEE,, 2018., 12(31):4953-4962

[12] C.Zhang, H. Li, X. Wang, and X. Yang.. Cross-scene crowd counting via deep convolutional neural networks. [J].IEEE Conference on Computer Vision and Pattern Recognition(CVPR), 2015 , 2(12):833-841

[13] HIdrees, I. Saleemi, C. Seibert, and M. Shah..Multi-source multi-scale counting in extremely dense crowd images.[J]. IEEE CVPR, 2013., 3(43):2547-2554

[14] L. Boominathan, S.S. Kruthiventi, and R.V. Babu. Crowdnet: A deep convolutional network for dense crowd counting. [J].In Proceedings of the 2016 ACM on Multimedia Conference, 2016, 2016(1):640-644

[15] CZhang, H. Li, X. Wang, and X. Yang. Cross-scene crowd counting via deep convolutional neural networks. [J]. ICCV, 2015., 1(2):1-7

[16] Jun Sang, Weiqun Wu, Hongling Luo et al. A crowd counting method based on single-column convolutional neural network[J]. Acta Automatica Sinica Sinica, 2018, 20(3):65-71

[17] L.Zhang, M. Shi, and Q. Chen. Crowd counting via scale-adaptive convolutional neural network.[J]. In 2018 IEEE Win-ter Conference on Applications of Computer Vision (WACV), 2018, 1(22):1113-1121