International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

# CROWD COUNTING METHOD BASED ON CNN

**T Sarada [1], A Dharani Guptha[2], Achana Manasa[3], Arepally Navya Sri[4], K Yuvaraj[5]**

[2,3,4,5] UG Scholars, Department of CSE, **AVN Institute of Engineering and Technology,**Hyderabad, Telangana, India.
[1]Assistant Professor, Department of CSE, **AVN Institute of Engineering and Technology**, Hyderabad, Telangana, India.

**Abstract:**

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. The experimental results in the UCF_CC_50 dataset and the ShanghaiTech dataset show that the proposed method has higher accuracy and stability.

## Introduction

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 headlike contour and build the SVM classifier to classify the feature area. Kowalak 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. In the relevant researches based on feature regression [6]-[10], the regression relationships between image features and the number of people are established for crowd counting. Chan et al. [7] proposed to use low-level features and Bayesian regression to improve the robustness and adaptability of the regression model. Idrees et al. [8] proposed to use multiplesources information to estimate the number of people in a single image, and the UCF_CC_50 dataset was introduced in this work.

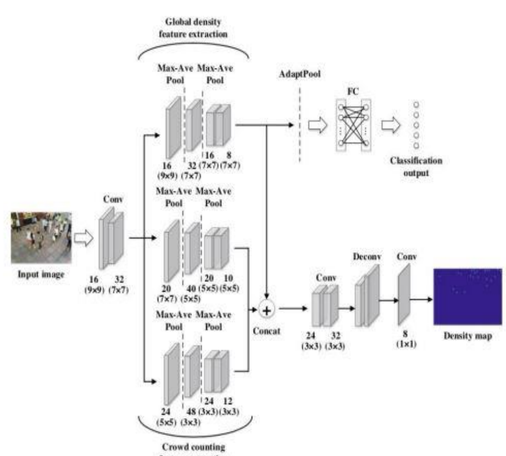
International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

Recently, with the rapid development of deep learning and big data [11]-[14], crowd counting methods based on deep learning are proposed gradually. Zhang et al. [15] 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 singlecolumn CNN. However, it is not suitable for the change in the scale of crowd. Zhang et al. [16] 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. Boominathan et al. [17] proposed to combine the features of shallow and deep convolutional neural networks to improve spatial resolution. Sindagi et al. [18] proposed a multi-task network which combined the high-level prior with the density estimation.

Sam et al. [19] proposed Switch-CNN for crowd counting. 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.

**System Architecture**



International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

**Existing System:**

The crowd counting method based on MCNN has achievedgood counting effects so far. However, uneven crowd distributions have not been taken into account in the existed crowdcounting methods. In order to solve the problem of unevencrowd distribution, the global density feature is taken intoaccount. In this paper, the network is constructed with the twoaspects: extracting feature maps with global density featureand generating a more comprehensive density map.

**Disadvantages:**

Crowd counting is used to calculate the total number ofpeople in images or video frames. The crowd counting methods can be divided into three categories: the direct countingmethod based on target detection, the indirect method basedon feature regression and crowd counting based on deeplearning. In the relevant researches based on target detection.

**Proposed System:**

with the rapid development of deep learning andbig data , crowd counting methods based on deeplearning are proposed gradually. Zhang et al. [15] proposed across-scene crowd counting model. It was trained alternatelythrough two learning objectives, density map and global number. This algorithm is implemented based on single-columnCNN. However, it is not suitable for the change in the scaleof crowd. Zhang et al. [16] proposed to use the MCNN withthree branch networks for crowd counting. Different receptivefields were used in each branch network, and this improvedMCNN could adapt to the change in the scale of the crowd.They also introduced a new dataset ShanghaiTech for crowdcounting. Boominathan et al. [17] proposed to combine thefeatures of shallow and deep convolutional neural networks toimprove spatial resolution. Sindagi and Patel [18] proposed amulti-task network which combined the high-level prior withthe density estimation. Sam et al. [19] proposed Switch-CNNfor crowd counting. In this network, a classifier was trainedand an appropriate regressor was selected for input patches.Shi et al. [20] proposed to aggregate multiscale features into

**Advantages:**

The max-ave pooling combines the advantages of theabove two pooling methods. The max-ave pooling is obtainedby superimposing the average pooling and the max poolingwith the same weight.

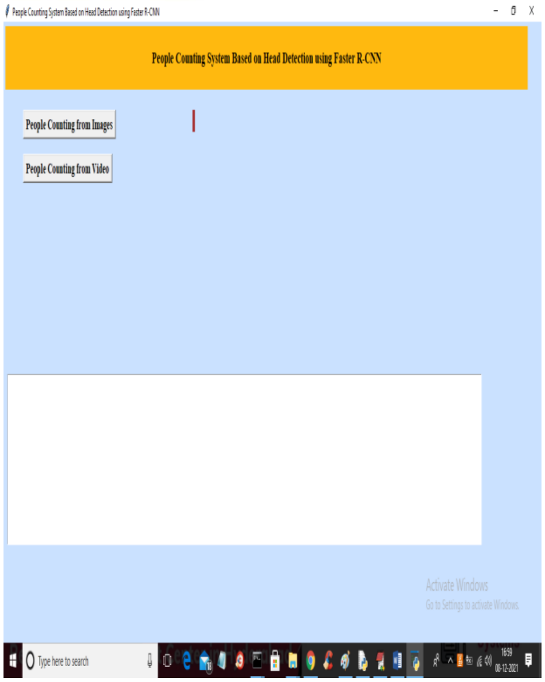
**7.SCREENSHOTS**

In this project we are using Faster RCNN model to count humans head from images and videos. In below screen you can see we are loading Faster RCNN model

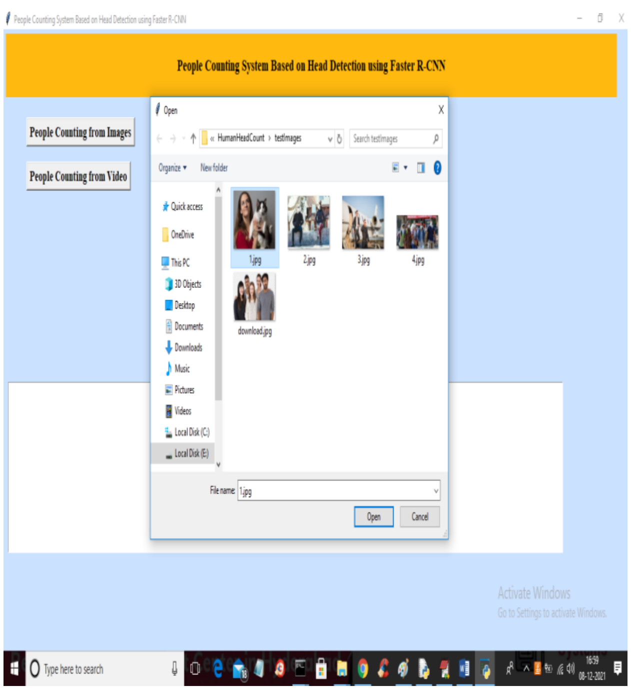
In above screen read red colour comment to know about RCNN model

To run project double click on ‘run.bat’ file to get below screen

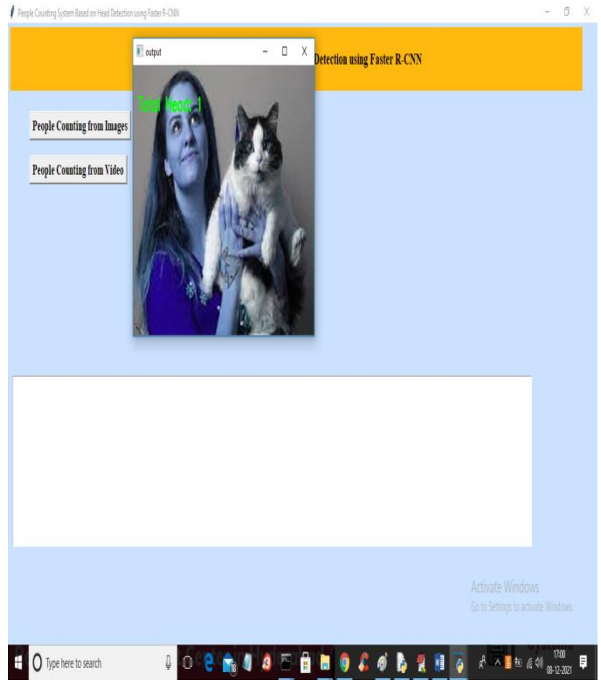
International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362
IJARST



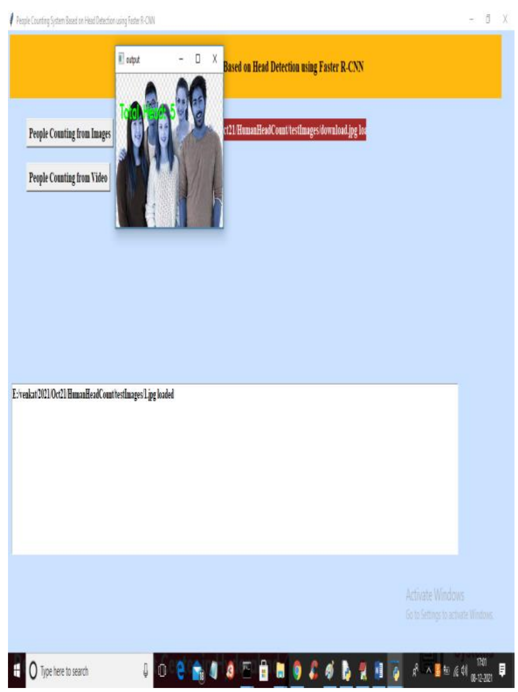

In above screen click on 'People Counting from Image' button to upload image like below screen




In above screen selecting and uploading '1.jpg' file and then click on 'Open' button to get below output

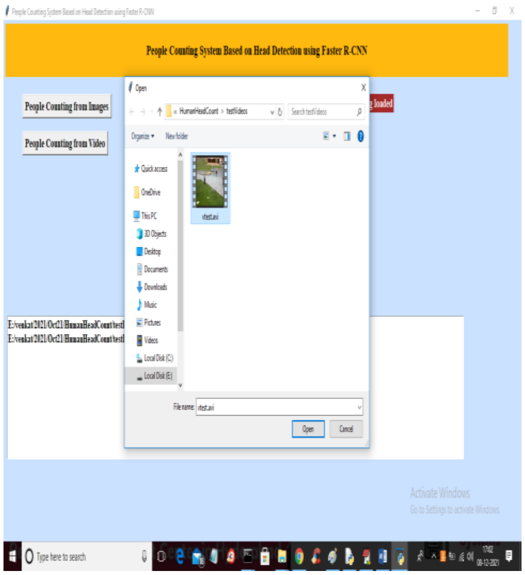



In above screen we got output as 'Total Head: 1" and now test other image

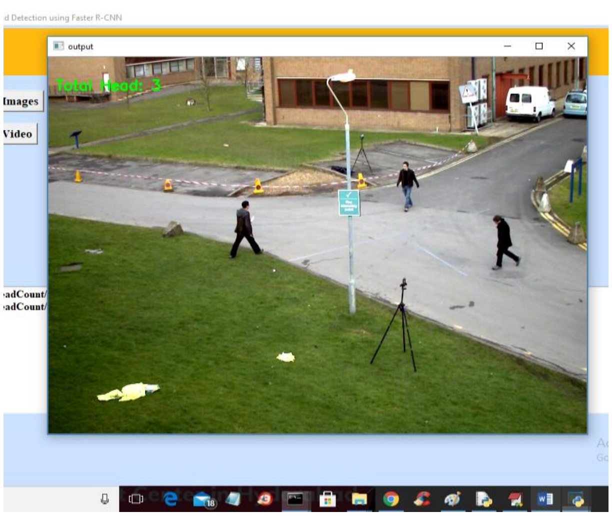


International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362
IJARST

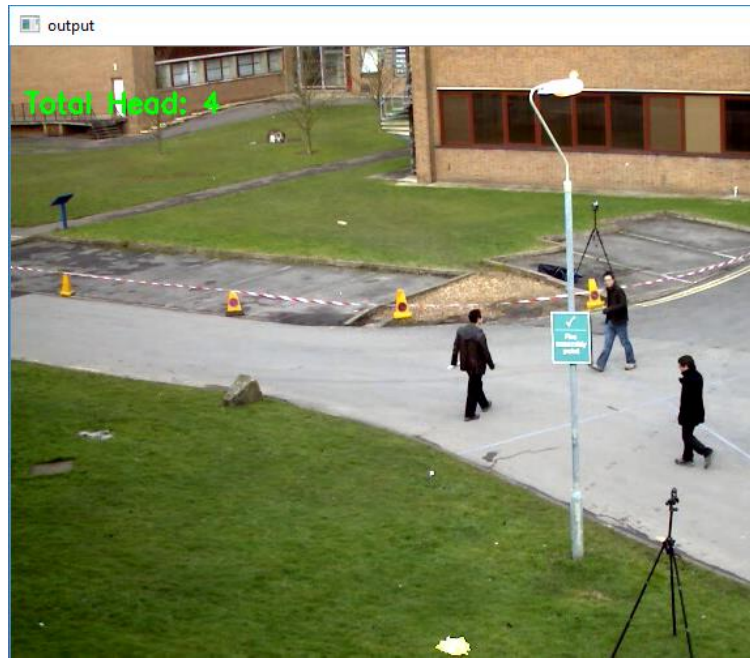
In above screen we got Total Head as 5. Now click on 'People Counting from Videos' button to get below screen




In above screen I am selecting and uploading 'vtest.avi' file and then click on 'Open' button to load video and start human head counting and based on you system speed video will be processed and if your system fast then video will be process faster else process slower and below is the output




In above frame we got 3 head as total humans are 3 and in below screen we got as 4




At any time press 'q' key on video to stop processing

**Conclusion:**

In this paper, 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 de convolutional 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 densitymap and influence the counting results. For the future work,we will focus on reducing the

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

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] G. Paweł, K. Kowalak, K. Łukasz, and S. Maćkowiak, “Crowd density estimation based on voxel model in multi-view surveillance systems,” in Proc. IEEE Int. Conf. Syst., Signals, Image Process., 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 head-shoulder detection,” in Proc. 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: long-term traffic anomaly detection based on crowdsourced bus trajectory data,” World Wide Web, vol. 21, no. 3, pp. 825–847, May. 2018.

[6] A. B. Chan, Z. S. J. Liang, and N. Vasconcelos, “Privacy preserving crowd monitoring: Counting people without people models or tracking,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2008, pp. 1-7.

[7] A. B. Chan, and N. Vasconcelos, “Counting people with low-level features and Bayesian regression,” IEEE Trans. Image Process., vol. 21, no. 4, pp. 2160-2177, Apr. 2012.

[8] H. Idrees, I. Saleem, C. Seibert, and S. Mubarak, “Multi-source multiscale counting in extremely dense crowd images,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2013, pp. 2547-2554.

[9] W. Ma, L. Huang, and C. Liu, “Crowd density analysis using cooccurrence texture features,” in Proc. IEEE Int. Conf. Comput. Sci., Convergence Informat. Techn., Nov. 2010, pp. 170-175.

[10] D. Kong, D. Gray, and H. Tao, “A viewpoint invariant approach for crowd counting,” in Proc. Int. Conf. Pattern Recognit., Aug. 2006, pp. 1187-1190.

[11] X. Kong et al., “Mobility dataset generation for vehicular social networks based on floating car data,” IEEE Trans. Vehicular Techn., vol. 67, no. 5, pp. 3874–3886, May. 2018.

[12] X. Kong, F. Xia, J. Wang, A. Rahim, and S. K. Das, “Time-locationrelationship combined service recommendation based on taxi trajectory data,” IEEE Trans. Ind. Informat., vol. 13, no. 3, pp. 1202–1212, Jun. 2017.

[13] J. Zhang, P. Liu, F. Zhang, and Q. Song, "CloudNet: ground-based cloud classification with deep convolutional neural network," Geophys. Res. Lett., vol. 45, no. 16, pp. 8665-8572, Aug. 2018.

[14] L. Ye, Z. Liu, L. Li, L. Shen, C. Bai, and Y. Wang, "Salient object segmentation via effective integration of saliency and objectness," IEEE Trans. Multimedia, vol. 19, no. 8, pp. 1742-1756, Aug. 2017.

[15] C. Zhang, H. Li, X. Wang, and X. Yang, "Cross-scene crowd counting via deep convolutional neural networks," in Proc IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2015, pp. 833-841.

[16] Y. Zhang, D. Zhou, S. Chen, S. Gao, and Y. Ma, "Single-image crowd counting via multi-column convolutional neural network," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2016, pp. 589-597.

[17] L. Boominathan, S. S. S. Kruthiventi, and R. V. Babu, "CrowdNet: A deep convolutional network for dense crowd counting," in Proc. 24th ACM Int. Conf. Multimedia, Oct. 2016, pp. 640-644.

[18] V. A. Sindagi, and V. M. Patel, "Cnn-Based cascaded multi-task learning of high-level prior and density estimation for crowd counting," in Proc. 14th IEEE Int. Conf. Adv. Video Signal Based Surveill., Aug. 2017, pp. 1-6.

[19] D. B. Sam, S. Surya, and R. V. Babu, "Switching convolutional neural network for crowd counting," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jul. 2017, pp. 4031-4039.

[20] 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.

[21] J. Fu, H. Yang, P. Liu, and Y. Hu, "A CNN-RNN neural network join long short-term memory for crowd counting and density estimation," in Proc. IEEE Int. Conf. Adv. Manufacturing, Nov. 2018, pp. 471-474.

[22] J. Liu, C. Gao, D. Meng, and A. Hauptmann, "DecideNet: Counting varying density crowds through attention guided detection and density estimation," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2018, pp. 5197-5206.

[23] 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 Comm., vol. 64, pp.118-129, Mar. 2018. [24] V. A. Sindagi, and V. M. Patel, "Generating high-quality crowd density maps using contextual pyramid CNNs," in Proc. IEEE Int. Conf. Comput. Vis., Oct. 2017, pp. 1879-1888.

[25] P. Chen, Y. Tang, L. Wang, and X. He, "Crowd density estimation based on multi-level feature fusion," Journal of Image, Graphics, vol. 23, no. 8, pp. 1181-1192, Mar. 2018.

[26] J. Dai, K. He, and J. Sun, "Instance-aware semantic segmentation via multi-task

network cascades," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2016, pp. 3150-3158.

[27] P. Wang, G. Tian, and H. Chen, "A saliency detection model combined local and global features," in Proc. Chin. Automation Congress, Oct. 2017, pp. 2863-2870.

[28] X. Tan, H. Zhu, Z. Shao, X. Hou, Y. Hao, and L. Ma, "Saliency detection by deep network with boundary refinement and global context," in Proc. IEEE Int. Conf. Multimedia, Expo, Jul. 2018, pp. 1-6.

[29] W. C. Hung et al., "Scene parsing with global context embedding," in Proc. IEEE Int. Conf. Comput. Vis., Oct. 2017, pp. 2650-2658.

[30] H. Fradi, X. Zhao, and J. L. Dugelay, "Crowd density analysis using subspace learning on local binary pattern," in Proc. IEEE Int. Conf. Multimedia, Expo, Jul. 2013, pp. 1-6.

[31] J. S. Bridle, "Probabilistic interpretation of feedforward classification network outputs, with relationships to statistical pattern recognition," Neurocomputing Algorithms Architectures, Appl., vol. 68, pp. 227-236, 1990.

[32] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: surpassing human-level performance on ImageNet classification," in Proc. IEEE Int. Conf. Comput. Vis., Dec. 2015, pp. 1026-1034.

[33] C. Bai, L. Huang, J. Chen, X. Pan, and S. Chen, "Optimization of deep convolutional neural network for large scale image classification," Journal of Software, vol. 29, no. 4, pp. 1029-1038, Dec. 2018.

[34] C. Bai, L. Huang, X. Pan, J. Zheng, and S. Chen, "Optimization of deep convolutional neural network for large scale image retrieval," Neurocomputing, vol. 303, pp. 60-67, May. 2018.

[35] M. D. Zeiler, and R. Fergus, "Visualizing and understanding convolutional networks," in Proc. Eur. Conf. Comput. Vis., Sep. 2014, pp. 818-833.

[36] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," IEEE Trans. Image Process., vol. 13, no. 4, pp. 600-612, Apr. 2004.

[37] V. Lempitsky, and A. Zisserman, "Learning to count objects in images," in Proc. Conf., Workshop on Neural Informat. Process. Syst., Dec. 2010, pp. 1324-1332.

[38] D. Onoro-Rubio, and R. J. López-Sastre, "Towards perspective-free object counting with deep learning," in Proc. Eur. Conf. Comput. Vis., Oct. 2016, pp. 615-629.

[39] L. Zeng, X. Xu, B. Cai, S. Qiu, and T. Zhang, "Multi-scale convolutional neural networks for crowd counting," in Proc. IEEE Int. Conf. Image Process., Sep. 2017, pp. 465-469.

[40] D. B. Sam, and R. V. Babu, "Top-down feedback for crowd counting convolutional neural network," in Proc. Thirty-Second AAAI Conf. Artificial Intell., Feb. 2018, pp. 7323-7330.

[41] M. Malacz, and G. Sarwas, “Crowd counting using complex convolutional neural network,” in Proc. Signal Process. Alogorithms, Archit., Arrang., Appl., Sep. 2018, pp. 88-92.

[42] L. Zhang, M. Shi, and Q. Chen, “Crowd Counting via Scale-Adaptive Convolutional Neural Network,” in Proc. IEEE Winter Conf. Appl.Comput. Vis., Mar. 2018, pp. 1113-1121.