IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)
An International Open Access, Peer-reviewed, Refereed Journal

An Advanced Approach To Crowd Counting Using Deep Learning

[1]Mr.Bharath B, [2]Mr.Ashish L
[1]MCA Scholar,[2]Assistant Professor
[1] Department of MCA,
[1] Nehru College of Engineering and Research Centre, Pampady, India

***Abstract**:* Counting crowds is an important task in many regards, given the improvement of public safety, urban planning, and event management. Recent deep convolutional neural networks advancements greatly improve occlusion and view-point variation. This paper surveys the progression from the state-of-the-art methodologies, hybrid approaches focusing specifically on regression types for real-time applications, and indicates scaling frameworks with diverse datasets for future development.

***Index Terms* – Crowd counting, Deep learning, CNNs, Density estimation, Hybrid approaches.**

I. INTRODUCTION

Such instances as the conduct of a concert, rally by political camps, or a quick evacuation can hardly be accurately counted. Current conventional methods would easily break when facing a density of people for this is its downfall wherein it also exhibits problems involving occlusion and distortion in a viewpoint. For improvements in the use of CNN, however have opened avenues, enabling one to gain quite feasible density approximation. The paper discusses these changes, and it focuses on hybrid techniques to improve public safety and operational efficiency and the real-time applications. [1][2].

II. LITERATURE SURVEY

2.1 General Overview for Crowd Counting Methodologies and Algorithms in IoT

Mingliang Gao et al. (2023) discussed the applicability of machine learning, deep learning, and meta-heuristic algorithms for crowd estimation in IoT systems. It noted that significant issues include scale variability and cluttered environments and identified the scope for AI to solve dynamic crowd scenes. It further emphasized that there is no scalable solution to real-time deployment from the solutions in the current scenario [1].

2.2 Deep Learning in Crowd Counting: A Survey

Lijia Deng et al. (2023) even introduced new measures such as the Average Pixel Occupied (APO) index, also they suggested taxonomy for datasets evaluation to be considered for crowds counting models. They thought that dataset integration was challenging besides this, need for real time execution also came into place. It needed new approaches with excellent performance for it to scale out in applied applications [2].

2.3 Crowd Management through Simulation using Algorithm Deep Learning

Ibtehal Talal Nafea, in the year 2023, put forward a model based on CNN for image classification of the crowd into different densities. With accuracy rates of almost 98%, in real time, the article proved that such models could be efficiently applied in hazardous situations like that of the Hajj pilgrimage and that timely interventions based upon these simulations helped mitigate congestion-based risks effectively [3].

2.4 Crowd Dynamics Simulation using Deep Learning and Enhanced Social Force Model

Dapeng Yan et al. (2023) integrated deep learning with an advanced social force model to predict crowd dynamics. The approach considerably improved the simulation accuracy of crowd behavior, which is useful for designing public spaces and managing large-scale events [4].

2.5 Crowd Behaviour Analysis using Convolutional Neural Networks: A Survey

Gaurav Tripathi et al. (2023) presented a review of CNN-based approaches in crowd behavior analysis, such as density estimation and anomaly detection. Accuracy was improved with considerable magnitude; however, it pointed out problems in terms of generalizing the performance of crowds to the deployment phase. It highlighted the necessity of developing scalable architectures [5].

2.6 Congestion-aware Bayesian Loss for Crowd Counting

Jeong, J. et al., 2022 suggested congestion-aware Bayesian loss function that improved the localization of dense crowds. The approach decreased the errors for sparse and dense crowds, evaluated on the widely used datasets such as UCF-QNRF and ShanghaiTech [6].

2.7 Multi-Level Feature Fusion for Crowd Counting

Sindagi, V. A., & Patel, V. M. (2022) proposed a feature fusion model that comprises low- and high-level features from the networks to boost precision. The feature fusion model demonstrates excellent performance based on different crowd density scales [7].

III. RESEARCH METHODOLOGY

3.1 Existent Methodologies

Traditional crowd counting methods form the basis of this domain and are broadly categorized into:

- **Detection-Based Methods**: These involve identifying individual objects, such as heads or bodies, using sliding windows or part-based models. Though effective in sparse crowds, they fail with occlusions in denser settings [2].
- **Regression-Based Approaches**: This model directly estimates the total number of crowds from edges and textures from an image. They are relatively less computationally intensive but mostly inaccurate in their representation of spatial crowd variations [2].

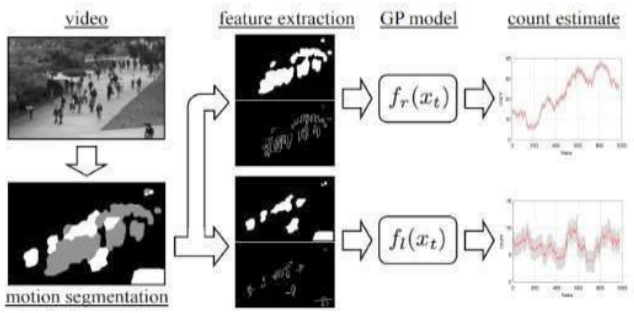

Fig. 1. Bayesian regression framework

- **Density Estimation Methods**: These generate spatial density maps to represent crowd distribution. Despite providing more reliable estimates, these methods rely heavily on handcrafted features, limiting their adaptability to complex scenarios [3].

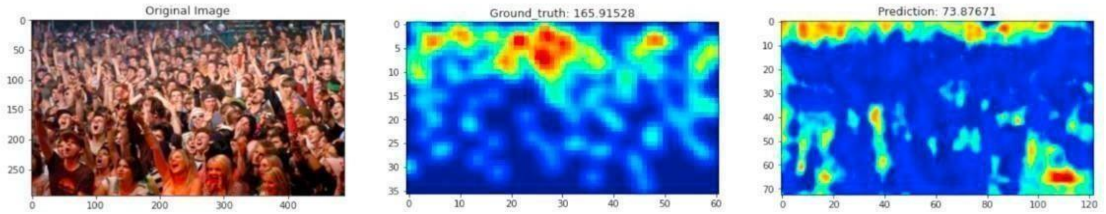

Fig. 2. Density map estimate for a particular picture is illustrated with GT and prediction

Density Estimation Techniques: These techniques output spatial density maps to represent crowd density. Since it presents an estimation more accurate than earlier techniques, but they rely significantly on handcrafted features and its usage in complex scenes is very limited [3].

3.2 Proposed Strategies

Deep learning has made the crowd counting task change from one which could get rid of all the inherent drawbacks of earlier approaches. Among these are

- **CNN-Based Regression Models**: This uses convolutional layers to increase precision and robustness across scales and viewpoints [4].
- **Density Estimation Using CNNs**: Deep learning facilitates the creation of high-resolution spatial maps, thereby providing accurate estimates even in dense conditions [5].
- **Hybrid Approach**: It is the fusion of a detection-based model with a density estimation model, thereby increasing flexibility and performance in real time [6].

- **Deep Learning Models**:
    1. Extended Bayesian Loss Method: It increases localization precision by employing scene geometry [6].
    2. Bayesian Loss Function: The labeling work decreases and it saves time since the probabilities can be assigned directly to the pixels [6].
    3. U-ASD Net: This is a hybrid model that was created specifically for sparse and dense crowds with diversity validation on multiple datasets [4].
    4. SRNet: It uses scale- and pixel-aware modules to utilize in variable-density scenes [5].
    5. FSCNet: The approach involves crowd counting and density mapping that yields robust results [4].
    6. ADCrowdNet: This employs attention maps for accurate detection and high-resolution density estimation [5].
    7. GTA5 Crowd Counting (GCC) is employed, where Spatial FCN architectures have simulated dynamic crowds scenario [6].
    8. VGG-16 framework combining both the feature extraction, as well as the spatial reconstruction, improve the accuracy to even more level of betterment
    9. Concept of Hierarchical Priors Cross Stage Refinement Network
    10. In the Scale Feature Extraction Network scale variation in the above case proves them an efficient model based on researches here
    11. Multi-Level MBTTBF Model: Integrates lowlevel and highlevel features to overcome problems of scale 7.

All these approaches seem to be towards deep learning-based techniques for crowd counting. These are mainly state-of-the-art architectures, which overcome scale variation, dynamic environments, and dense crowds, making use of fully convolutional networks, multi-level frameworks, and decoder-encoder designs. Innovation in datasets, modules, and loss functions is also current, thus putting these models at the borderline in terms of accuracy while adapting well to real-world considerations.

Datasets

Several key datasets have significantly advanced crowd counting research. The **ShanghaiTech** [21] dataset, with 1,198 images and 330,165 annotations, is divided into high and low-density sections. The **UCSD** [22] dataset, one of the first of its kind, contains 2,000 frames for training and testing. **UCF-CC-50** [23] includes 50 images that vary in crowd density and perspective. The **UCF-QNRF** [24] dataset features 1,535 images and 1.25 million annotations, capturing diverse environments and lighting conditions. The **MALL** [25] dataset, taken from mall CCTV footage, contains 2,000 frames and 62,325 pedestrian annotations. Lastly, the **WorldExpo'10** [26] dataset offers 3,920 frames and 199,923 annotations from the 2010 World Expo. In our research, we apply various deep learning-based crowd counting models to these datasets and evaluate their performance to better understand the strengths and weaknesses of each model in real-world scenarios.

Table 1. Datasets

Dataset	Year	Attribute	No of images	Resolution	Min	Max	Average	Train/Test
ShanghaiTe ch PartA [21]	2016	Real world Congested	482	Varied	33	3139	501	300/182
ShanghaiTe ch PartB [21]	2016	Real world	716	760x1024	9	123	578	400/316
UCSD [22]	2008	Real World	2000	320x240	13	-	46	800/1200
UCF-CC-50[23]	2013	Real World	50	Varied	94	1279	4543	-
UCF-QNRF [24]	2018	Real World	1535	2013x2902	49	815	12865	1201/334
MALL [25]	2012	Real World	2000	320x240	13	53	31	800/1200
WorldExpo’10[26]	2015	Real World	3980	576x720	1	50.2	253	-

IV. RESULTS AND DISCUSSION

Metrics like Mean Absolute Error (MAE) and Mean Squared Error (MSE) validate the effectiveness of CNN-based approaches. Hybrid models surpass traditional techniques by overcoming challenges like occlusions and scale variations. Among the models studied, the **Scale Feature Extraction Network (SFEN)** emerged as the most efficient, demonstrating superior reliability and precision in high-density environments [5].

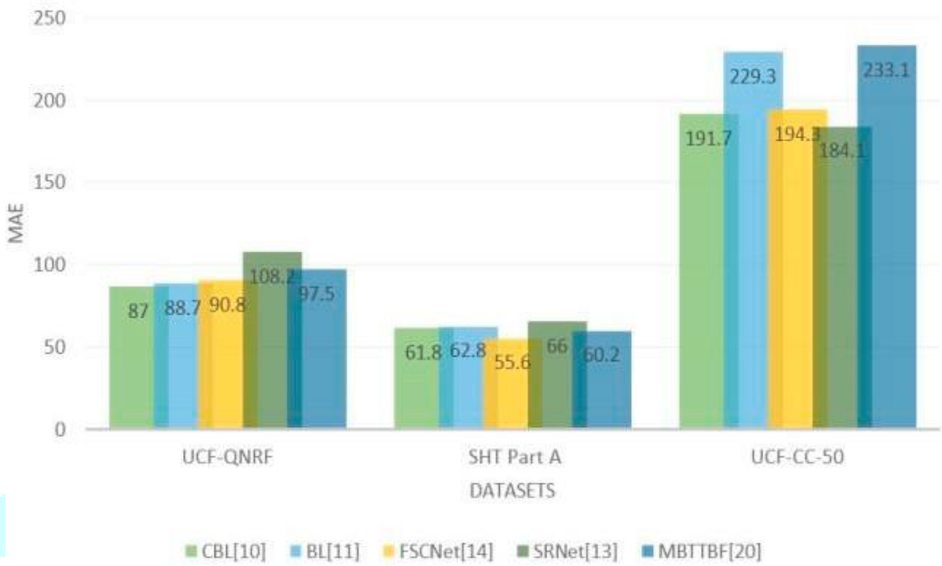

Fig 3. Graphical representation of the MAE of deep learning model on the most popular datasets. (Lower MAE corresponds to higher accuracy)

Despite these advancements, challenges persist in dataset diversity and adapting models for real-time applications.

V. CHALLENGES AND FUTUREWORK

Challenges:

- Variations in Crowd Scenes
- Need for Labelled Data
- Real-time Processing
- Model Adaptability

Future Work:

- Enhancing Model Flexibility
- Efficient Learning Methods
- Privacy Protection
- Seamless Integration

VI. CONCLUSION

This research thus highlights the great improvements that have been seen in crowd counting through the use of deep learning, especially with CNN-based and hybrid models. The improvements have alleviated many traditional challenges, including limited accuracy and scalability, since they can be able to carry out more reliable and efficient real-time processing of data. However, as we move ahead, it is important to refine these models so that they may be deployed at scale in real-world applications. Equally relevant are the privacy concerns that gather around crowd monitoring. The key to acceptance and practical use of such technology will lie in devising solutions that must balance individual privacy requirements against accuracy.

VII. REFERENCES

[1] Mingliang Gao, et al. (2023). *A Comprehensive Analysis for Crowd Counting Methodologies and Algorithms in IoT.*

[2] Lijia Deng, et al. (2023). *Deep Learning in Crowd Counting: A Survey*
.

[3] Ibtehal Talal Nafea (2023). *Simulation of Crowd Management Using Deep Learning Algorithm.*

[4] Dapeng Yan, et al. (2023). *Enhanced Crowd Dynamics Simulation with Deep Learning and Improved Social Force Model.*

[5] Gaurav Tripathi, et al. (2023). *Convolutional Neural Networks for Crowd Behaviour Analysis: A Survey.*

[6] Jeong, J., et al. (2022). *Congestion-aware Bayesian Loss for Crowd Counting.* IEEE Access.

[7] Sindagi, V. A., & Patel, V. M. (2022). *Multi-Level Feature Fusion for Crowd Counting.* IEEE International Conference on Computer Vision.

[8] ShanghaiTech Dataset: Includes 1,198 images and 330,165 annotations, divided into high-density and low-density parts.

[9] UCSD Dataset: Contains 2,000 frames, with 800 for training and 1,200 for testing, making it one of the first crowd counting datasets.

[10] UCF-CC-50 Dataset: Comprises 50 images with varying crowd densities and perspectives, subjected to 5-fold cross-validation.

[11] UCF-QNRF Dataset: Contains 1,535 images with 1.25 million annotations, covering diverse crowd densities and lighting conditions.

[12] MALL Dataset: Gathered from mall CCTV cameras, it includes 2,000 frames with 62,325 pedestrian annotations.

[13] WorldExpo'10 Dataset: Features 3,920 frames with 199,923 annotated people from the 2010 World Expo.

[14] Huang, L., Zhu, L., Shen, S., Zhang, Q., & Zhang, J. (2021). *SRNet: Scale-Aware Representation Learning Network for Dense Crowd Counting.* IEEE Access, 9, 136032-136044.

[15] Liu, N., Long, Y., Zou, C., Niu, Q., Pan, L., & Wu, H. (2019). *Adcrowdnet: An Attention-Injective Deformable Convolutional Network for Crowd Understanding.* In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 3225-3234).

[16] Wang, Q., Gao, J., Lin, W., & Yuan, Y. (2019). *Learning from Synthetic Data for Crowd Counting in the Wild.* In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 8198-8207).

[17] Xie, Y., Lu, Y., & Wang, S. (2020, October). *RSANet: Deep Recurrent Scale-Aware Network for Crowd Counting.* In *2020 IEEE International Conference on Image Processing (ICIP)* (pp. 1531-1535). IEEE.