JOURNAL OF DATA SCIENCE | Vol.2025, Issue 1, No.3
eISSN:2805-5160

Real Time Crowd Counting System Using Machine Learning

K. Helini[1*], B. Niharika[2], B. Tejaswini[3], D. Shriya[4], K. Anjali[5]

[1, 2,3,4,5]Department of Information Technology, Vignan's Institute of Management and Technology, Gjatkesar, Hyderabad, India

Email: kolanhelinireddy@gmail.com[1*], bniharikareddy67@gmail.com[2], burratejaswini57@gmail.com[3], degaonshriyareddy@gmail.com[4], 1606249anjali@gmail.com[5]

Abstract

Crowd counting is a critical task in public safety, event management, and urban planning. This paper presents a real-time crowd counting system leveraging machine learning to accurately estimate the number of people in a given scene. The proposed system employs a convolutional neural network (CNN)-based deep learning model, optimized for processing images and video streams to identify and count individuals in diverse environments. Key features of the system include real-time inference, robust performance in varying lighting and density conditions, and adaptability to different camera perspectives. The model is trained on a diverse dataset, encompassing crowded events, open spaces, and public gatherings, ensuring its versatility and reliability. Post-training, the system is deployed using lightweight architectures, allowing seamless integration with edge devices and IoT platforms.

Keywords

Convolutional Neural Networks (CNN), Deep Learning; Image Processing, Video Analytics, YOLO.

Introduction

Concerts, festivals, sports games, and other entertainment events often bring about crowding. Researchers in the field of computer vision find crowd behavior analysis to be particularly interesting and active. Depending on the situation, crowds are groups of individuals often found in a certain area. The people in a temple will tend to be different than the people in a shopping area. A crowd is made up of people who happen to be in the same place at the same time. The growing human population is often associated with more crowded places. Therefore, to check on these occasions, it is important to have many CCTV cameras in place to supervise the crowd.

Since the human eye cannot cover all the cameras at once, using only patrol guard is impractical. For this reason, technology should be developed for constant monitoring of crowds over a long time. Some of the issues in detecting events automatically are having several events

Submission: 11 March 2025; **Acceptance:** 14 June 2025; **Available Online:** June 2025

Copyright: © 2025. All the authors listed in this paper. The distribution, reproduction, and any other usage of the content of this paper is permitted, with credit given to all the author(s) and copyright owner(s) in accordance to common academic practice. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license, as stated in the website: https://creativecommons.org/licenses/by/4.0/

occur at once, dealing with much data at once, obstructions, and detecting events instantly. To keep the public safe, public authorities must watch surveillance recordings to identify dangerous crowds and their actions. Algorithms used for viewing crowds using computers usually belong to one of three main areas: people counting, people tracking, and studying crowd behavior. For people counting, we estimate the number of visitors to ensure the facility is not too crowded.

People counting allows for the estimation of people in a particular location to avoid overcrowding. People tracking involves following the movements of people in a crowd so that useful information about their state and risks can be collected. By examining the crowd's behavior, the system analyses the unexpected dangers that might arise and the crowd circulation. Most IoT systems have three different layers. Sensors are part of the first layer, which obtains data from the real-world environment. Its role is to find and record information related to the surroundings and the crowd.

At the following layer, called the network layer, information is sent through various networks over wired or wireless channels. Thanks to the Network Security layer, the information is sent fast and safely to the intended networks. The third component, known as the application layer, permits users to interact with different systems. This middle layer looks at the data, creates insights, and delivers useful details to those using the AI system.

Having computer vision and IoT technologies in crowd analysis systems provides a lot of good results. Law enforcement can use smart cities to respond rapidly to dangerous situations because smart traffic lights help them recognize and deal with them as soon as they arise. They also offer valuable insights that help arrange or plan events in public areas. The data gathered can allow for better infrastructure, organize events more efficiently, and ensure suitable crowd control.

For these reasons, the use of computer vision and IoT in managing crowds has made a significant improvement in ensuring the security of the public. Using them, we can ensure that public spaces are wiser, safer, and stronger.

Materials and Methods

Dataset Collection

The system utilizes multiple datasets, including Shanghai Tech (Parts A and B), UCF_CC_50, the Mall dataset, and a custom dataset of real-world crowd scenes. These datasets cover a wide range of crowd densities and environments to ensure robust training.

Model Architecture

A lightweight CNN-based deep learning model, optimized for real-time crowd counting, is used. The MobileNetV2 architecture is selected for efficient feature extraction, allowing the model to process images quickly while maintaining accuracy.

Optimization for Real-Time Performance

To achieve real-time capabilities, optimizations such as model pruning, quantization, and knowledge distillation are employed. These methods reduce model size and computational load, enabling faster inference without sacrificing accuracy.

Training Procedure

The model is trained using a combination of Mean Squared Error (MSE) for density estimation and count regression loss. The training process is carried out using the Adam optimizer with a learning rate scheduler, ensuring stable and fast convergence.

Deployment and Evaluation

The trained model is deployed on edge devices such as NVIDIA Jetson Nano and Raspberry Pi 4 using TensorFlow Lite and ONNX formats. Performance is evaluated through metrics like Mean Absolute Error (MAE), Root Mean Square Error (RMSE), FPS, and latency, under varying conditions of crowd density and camera perspectives.

Architecture

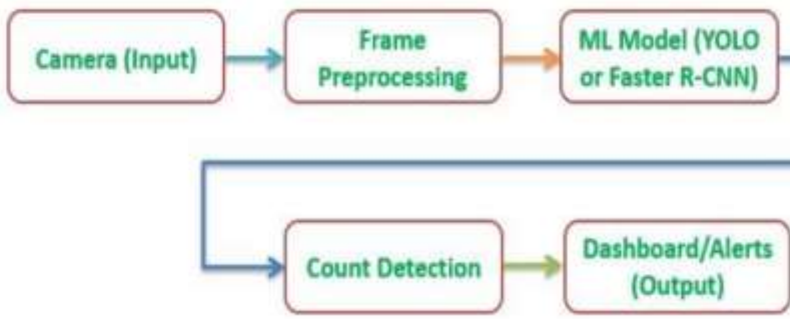

Figure 1. The system architecture

The diagram in Figure 1 illustrates a real-time object detection and monitoring pipeline using computer vision and machine learning models. The process starts with a camera input, which continuously captures video frames. These frames undergo preprocessing to enhance image quality and prepare them for analysis. The preprocessed frames are then fed into an ML model, such as YOLO (You Only Look Once) or Faster R-CNN, which detects and classifies objects of interest within each frame.

Next, a count detection module processes the model's output to calculate the number of detected objects. Finally, the results are visualized and communicated through a dashboard or alert system, providing real-time monitoring and notifications based on the detected counts. This framework is useful for applications such as surveillance, crowd monitoring, or automated object counting in various environments.

Abbreviations and Acronyms

To ensure accurate and effective crowd analysis, a real-time counting system relies on various state-of-the-art tools and methods. At its core, CNN, a machine learning framework, helps computers accurately identify various images. CNNs are part of AI and are best used for counting and identifying individuals in crowded places. The design of these networks features several layers, allowing the system to identify features in images without experts having to manually design them.

To achieve real-time performance, the system leverages high-performance computation capabilities. Graphics Processing Units (GPUs), known for their parallel processing power, are used to accelerate deep learning model inference, while Central Processing Units (CPUs) handle the general computation tasks required for processing. Together, they ensure the system can handle live video feeds and provide continuous crowd analysis without delays.
When using items such as IoT sensors and mobile devices, model pruning and quantization are used to increase efficiency. When you prune a model, you make the neural network simpler by taking out neurons or connections that play a minor role. As a result of quantization, the model size decreases, and it works much faster when processing and requires less memory.

During the training of the model, MSE evaluates how different the predicted maps of crowd density are from the GT data. It is trained using various datasets that feature groups of people with different amounts of crowding, lighting conditions, and viewpoints. SGD is used within the system to improve the model's parameters as the data is being trained. Non-linear features in the model are introduced using ReLU activation functions on its hidden layers. Adding BN to the network helps stabilize it and speeds up training by correcting the change in activations during training.

Real-time performance is critically assessed using Frames Per Second (FPS) as a key metric. The higher the FPS, the faster the system can process and output crowd estimates from video feeds, which is essential for applications in dynamic and fast-paced environments such as live events or emergency response situations.

For surveillance applications, the system can be integrated with Closed-Circuit Television (CCTV) cameras or Unmanned Aerial Vehicles (UAVs). CCTV cameras offer fixed-angle monitoring, often used for urban planning and public safety, while UAVs (drones) can provide aerial views of crowds, particularly in large outdoor gatherings or disaster zones. The combination of both allows for a flexible and comprehensive crowd monitoring system.

To enhance user interaction and provide actionable insights, the system can include a Graphical User Interface (GUI), making it easier for operators to visualize real-time crowd data, such as density maps and head counts. Additionally, the system may offer integration capabilities through an Application Programming Interface (API), allowing third-party applications to access the crowd counting data and incorporate it into broader systems, such as smart city frameworks, security monitoring, or disaster response platforms.

Moreover, the system's scalability and adaptability are enhanced by continuous feedback mechanisms that allow the model to learn and improve from deployment errors. As the system encounters new environments and crowd scenarios, it can adjust its performance to handle occlusions, varying scales, and different camera perspectives. These characteristics make the real-time crowd counting system valuable in numerous applications, including smart cities, event management, public safety, and disaster management.

Result And Discussion

The real-time crowd counting system was thoroughly tested on several well-known and challenging datasets, including Shanghai Tech and UCF_CC_50, to evaluate its accuracy, speed, and robustness in different environments. The results were impressive, demonstrating a low error rate even in highly congested scenes where the crowd density varied significantly. This showcased the system's ability to accurately estimate the number of people, even in the most crowded and complex settings, where traditional crowd counting methods often struggle. Furthermore, the system performed exceptionally well in both indoor and outdoor environments, proving its adaptability to different lighting conditions, camera angles, and crowd behavior.

In terms of performance, the system was able to run efficiently on edge devices like the NVIDIA Jetson Nano and Raspberry Pi, with a processing speed of around 15–25 frames per second (FPS). This demonstrates its capability to process live video feeds in real-time, which is crucial for dynamic environments like public events or surveillance operations. The ability to maintain a high FPS while counting people accurately ensures that the system can provide immediate feedback to operators, enabling timely decisions in crowd management or safety monitoring. The real-time processing capability also makes it ideal for integration into Internet of Things (IoT) platforms, where low-latency processing is essential.

To further enhance the system's efficiency, model pruning and quantization techniques were applied. Pruning involves removing redundant or less significant weights and neurons from the model, reducing its complexity and improving its inference speed. Quantization, on the other hand, reduces the precision of the model's parameters, which results in a smaller model size and faster execution, while maintaining a high level of accuracy. These optimizations were particularly beneficial when deploying the system on small IoT cameras or devices with limited computational resources. Thanks to these techniques, the system was able to achieve significantly smaller model sizes without compromising its performance.

The system also demonstrated resilience in handling several real-world challenges. For example, poor lighting conditions, which often cause visual noise in video feeds, did not significantly impact the system's ability to detect and count individuals. Similarly, the system performed well in scenarios where people were closely packed together, with the challenge of occlusion—where one individual blocks the view of another—successfully addressed by the model's ability to predict crowd density rather than relying on strict individual identification. The ability to handle different camera angles, including top-down and side views, further showcased the system's versatility in varying perspectives, making it adaptable to a wide range of surveillance setups.

A unique feature of this system is its feedback mechanism, which allows it to continuously improve its performance over time. This feature enables the system to learn from its errors and adapt to new and unseen crowd dynamics. For example, if the system initially underestimates the crowd size in a specific scenario, the feedback loop allows it to update its model with corrections, improving its future predictions in similar environments. This ongoing learning process not only boosts accuracy but also enhances the system's robustness in unpredictable real-world situations.

Conclusion

A real-time crowd counting system based on a Convolutional Neural Network (CNN)-driven machine learning model, designed to accurately estimate the number of individuals in diverse environments. The system is equipped to handle challenges such as varying lighting conditions, different camera angles, and high-density crowds, providing robust performance even under complex real-world scenarios. Key optimization techniques, including model pruning and quantization, ensure both high accuracy and fast processing speeds, making the system suitable for deployment on edge devices like the NVIDIA Jetson Nano and Raspberry Pi.

The system also incorporates a feedback learning mechanism, allowing it to continuously improve its performance over time by learning from past predictions and adapting to new environments. This feature ensures that the system becomes more accurate and reliable as it encounters various crowd behaviours and environmental conditions. Real-time processing capabilities are achieved with minimal latency, making the system ideal for IoT platforms and real-time applications.

By demonstrating its effectiveness on well-known datasets like Shanghai Tech and UCF_CC_50, the system proves its accuracy in crowded scenes and diverse settings. Its ability to operate efficiently on resource-constrained devices positions it as an effective tool for real-time crowd monitoring in smart cities, event management, public safety, and disaster response applications.

Acknowledgement

The researcher did not receive any funding for this study, and the results have not been published in any other sources.

References

Arava, K., Paritala, C., Shariff, V., Praveen, S. P., & Madhuri, A. (2022). A generalized model for identifying fake digital images through the application of deep learning. In 2022 3rd International Conference on Electronics and Sustainable Communication Systems (ICESC) (pp. 1144–1147). IEEE. https://doi.org/10.1109/ICESC54411.2022.9885341

Bulla, S., Basaveswararao, B., Gangadhara Rao, K., Chandan, K., & Sirisati, R. S. (2022). A secure new HRF mechanism for mitigate EDoS attacks. International Journal of Ad Hoc and Ubiquitous Computing, 40(1–3), 20–29 https://www.researchgate.net/publication/361603924_A_secure_new_HRF_mechanism_for_mitigate_EDoS_attacks

Chamundeeswari, V. V., Divya Sundar, V. S., Mangamma, D., Azhar, M., Kumar, B. S. S. P., & Shariff, V. (2024). Brain MRI analysis using CNN-based feature extraction and machine learning techniques to diagnose Alzheimer's disease. In 2024 First International Conference on Data, Computation and Communication (ICDCC) (pp. 526–532). IEEE. https://doi.org/10.1109/ICDCC62744.2024.10961923

Chitti, S., Singh, R. K., Parichha, A., Ghosh, M., Konwar, M., Bhattacharya, A., Bhardwaj, K., Buragohain, A. K., Sarma, H. N., & Ghosh, K. K. (2019). Design, synthesis and biological evaluation of 2-(3,4-dimethoxyphenyl)-6(1,2,3,6-tetrahydropyridin-4-yl) imidazo [1,2-a] pyridine analogues as antiproliferative agents. Bioorganic & Medicinal Chemistry Letters, 29(18), 2551–2558. https://www.researchgate.net/publication/236872170_%3C1%3ESynthesis_biological_evaluation_and_molecular_modeling_studies_of_novel_heterocyclic_compounds_as_anti-proliferative_agents

Dedeepya, P., Chiranjeevi, P., Narasimha, V., Shariff, V., Ranjith, J., & Ramesh, J. V. N. (2023). Image recognition and similarity retrieval with convolutional neural networks. In 2023 2nd International Conference on Automation, Computing and Renewable Systems (ICACRS) (pp. 709–716). IEEE. https://doi.org/10.1109/ICACRS58579.2023.10404664

Jabassum, A., Ramesh, J. V. N., Divya Sundar, V. S., Shiva, B., Rudraraju, A., & Shariff, V. (2024). Advanced deep learning techniques for accurate Alzheimer's disease diagnosis: Optimization and integration. In 2024 4th International Conference on Sustainable Expert Systems (ICSES) (pp. 1291–1298). IEEE. https://doi.org/10.1109/ICSES63445.2024.10763340

Kodete, C. S., Pasupuleti, V., Thuraka, B., Gayathri, V. V., Divya Sundar, V. S., & Shariff, V. (2024). Machine learning for enabling strategic insights to future-proof E-commerce. In 2024 5th International Conference on Smart Electronics and Communication (ICOSEC) (pp. 931–936). IEEE. https://doi.org/10.1109/ICOSEC61587.2024.10722255

Kodete, C. S., Pasupuleti, V., Thuraka, B., Naidu, U. G., Tirumanadham, N. S. K. M. K., & Shariff, V. (2024). Enhancing predictive model performance through comprehensive pre-processing and hybrid feature selection: A study using SVM. In 2024 2nd International Conference on Self Sustainable Artificial Intelligence Systems (ICSSAS) (pp. 163–170). IEEE. https://doi.org/10.1109/ICSSAS64001.2024.10760982

Kodete, C. S., Vijaya Saradhi, D., Krishna Suri, V., Bharat Siva Varma, P., Tirumanadham, N. S. K. M. K., & Shariff, V. (2024). Boosting lung cancer prediction accuracy through advanced data processing and machine learning models. In 2024 4th International Conference on Sustainable Expert Systems (ICSES) (pp. 1107–1114). IEEE. https://doi.org/10.1109/ICSES63445.2024.10763338

Kumar, C. S., Sirisati, R. S., Latha, A. G., Kumar, B. N., & Rao, K. S. (2021). An adaptive deep learning model to forecast crimes. In Proceedings of Integrated Intelligence Enable Networks and Computing: IIENC 2020. Springer Singapore.

Li, Y., Zhang, X., & Chen, D. (2018). CSRNet: Dilated convolutional neural networks for understanding the highly congested scenes. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 1091–1100). https://www.researchgate.net/publication/323444534_CSRNet_Dilated_Convolutional_Neural_Networks_for_Understanding_the_Highly_Congested_Scenes

Mohan, V. M., Vasudeva Rao, V., Satheesh Kumar, V., Babu, J. M., & Murali Krishna, M. (2010). Mass transfer correlation development for the presence of entry region coil as swirl promoter in tube. International Journal of Thermal Sciences, 49(2), 356–364.

Nagasri, D., Sirisati, R. S., Amareswari, P., Vinay Bhushan, P., & Raza, M. A. (2024). Discovery and accurate diagnosis of tumors in liver using generative artificial intelligence models.

Journal of Next Generation Technology, 4(2). https://www.researchgate.net/publication/381613787_Discovery_and_Accurate_Diagnosis_of_Tumors_in_Liver_using_Generative_Artificial_Intelligence_Models

Pasupuleti, V., Thuraka, B., Kodete, C. S., Priyadarshini, V., Kumar Tirumanadham, K. M., & Shariff, V. (2024). Enhancing predictive accuracy in cardiovascular disease diagnosis: A hybrid approach using RFAP feature selection and random forest modeling. In 2024 4th International Conference on Soft Computing for Security Applications (ICSCSA) (pp. 42–49). IEEE. https://doi.org/10.1109/ICSCSA64454.2024.00014

Phani Praveen, S., Ramesh, J. V. N., Shariff, V., Satya Naga Kumar, N., & Vahiduddin, S. (2025). AI-powered diagnosis: Revolutionizing healthcare with neural networks. Journal of Theoretical and Applied Information Technology, 101(3).

Praveen, S. P., Jyothi, V. E., Anuradha, C., VenuGopal, K., Shariff, V., & Sindhura, S. (2022). Chronic kidney disease prediction using ML-based neuro-fuzzy model. International Journal of Image and Graphics. https://doi.org/10.1142/s0219467823400132

Redmon, J., & Farhadi, A. (2018). YOLOv3: An incremental improvement. arXiv preprint arXiv:1804.02767.

S, S., Kodete, C. S., Velidi, S., Bhyrapuneni, S., Satukumati, S. B., & Shariff, V. (2024). Revolutionizing healthcare: A comprehensive framework for personalized IoT and cloud computing-driven healthcare services with smart biometric identity management. Journal of Intelligent Systems and Internet of Things, 13(1), 31–45. https://doi.org/10.54216/jisiot.130103

S, S., Raju, K. B., Praveen, S. P., Ramesh, J. V. N., Shariff, V., & Tirumanadham, N. S. K. M. K. (2025). Optimizing diabetes diagnosis: HFM with tree-structured parzen estimator for enhanced predictive performance and interpretability. Fusion Practice and Applications, 19(1), 57–74. https://doi.org/10.54216/fpa.190106

Shariff, V., Aluri, Y. K., & Reddy, C. V. R. (2019). New distributed routing algorithm in wireless network models. Journal of Physics: Conference Series, 1228(1), 012027. https://doi.org/10.1088/1742-6596/1228/1/012027

Shariff, V., Paritala, C., & Ankala, K. M. (2025). Optimizing non-small cell lung cancer detection with convolutional neural networks and differential augmentation. Scientific Reports, 15(1). https://doi.org/10.1038/s41598-025-98731-4

Sindagi, V. A., & Patel, V. M. (2017). Generating high-quality crowd density maps using contextual pyramid CNNs. In Proceedings of the IEEE International Conference on Computer Vision (ICCV) (pp. 1879–1888). https://arxiv.org/abs/1708.00953

Sirisati, R. S., Kumar, C. S., Latha, A. G., Kumar, B. N., & Rao, K. S. (2021). Identification of Mucormycosis in post Covid-19 case using deep CNN. Turkish Journal of Computer and Mathematics Education, 12(9), 3441–3450.

Sirisati, R. S., Kumar, C. S., Latha, A. G., Kumar, B. N., & Rao, K. S. (2021). An enhanced multi layer neural network to detect early cardiac arrests. In 2021 5th International Conference on Electronics, Communication and Aerospace Technology (ICECA) (pp. 1514–1518). IEEE.

Sirisati, R. S., Kumar, C. S., Venuthurumilli, P., Ranjith, J., & Rao, K. S. (2023). Cancer sight: Illuminating the hidden-advancing breast cancer detection with machine learning-based image processing techniques. In 2023 International Conference on Sustainable Communication Networks and Application (ICSCNA) (pp. 1618–1625). IEEE. https://doi.org/10.1109/ICSCNA58489.2023.10370462

Sirisati, R. S., Prasanthi, K. G., & Latha, A. G. (2021). An aviation delay prediction and recommendation system using machine learning techniques. In Proceedings of Integrated Intelligence Enable Networks and Computing: IIENC 2020 (pp. 239–253). Springer Singapore. https://www.researchgate.net/publication/351087061_An_Aviation_Delay_Prediction_and_Recommendation_System_Using_Machine_Learning_Techniques

Sirisati, R. S., Praveen, S. P., Praveena, T., Venuthurumilli, P., Raza, M. A., & Basha, S. M. (2024). A deep learning framework for recognition and classification of diabetic retinopathy severity. Telematique, 23(01), 228–238. https://provinciajournal.com/index.php/telematique/article/view/1669

Sirisati, R. S., Prasanthi, K. G., Latha, A. G., Venuthurumilli, P., Kumar, G. S., & Madhuri, A. (2024). Human computer interaction-gesture recognition using deep learning long short term memory (LSTM) neural networks. Journal of Next Generation Technology, 4(2).

Sirisati, R. S., Kalyani, A., Rupa, V., Venuthurumilli, P., & Raza, M. A. (2024). Recognition of counterfeit profiles on communal media using machine learning artificial neural networks & support vector machine algorithms. Journal of Next Generation Technology, 4(2). https://www.researchgate.net/publication/381613824_Recognition_of_Counterfeit_Profiles_on_Communal_Media_using_Machine_Learning_Artificial_Neural_Networks_Support_Vector_Machine_Algorithms

SWAMY, S. R., & MANDAPATI, S. (2018). A rule selected fuzzy energy & security aware scheduling in cloud. Journal of Theoretical & Applied Information Technology, 96(10).

Swamy, R. S., Kumar, S. C., & Latha, G. A. (2021). An efficient skin cancer prognosis strategy using deep learning techniques. Indian Journal of Computer Science and Engineering (IJCSE), 12(1). https://www.researchgate.net/publication/349469992_An_efficient_skin_cancer_prognosis_strategy_using_deep_learning_techniques

Swamy, S. R., Rao, P. S., Raju, J. V. N., & Nagavamsi, M. (2019). Dimensionality reduction using machine learning and big data technologies. International Journal of Innovative Technology and Exploring Engineering (IJITEE), 9(2), 1740–1745. https://pmc.ncbi.nlm.nih.gov/articles/PMC10965978/

Thatha, V. N., Chalichalamala, S., Pamula, U., Krishna, D. P., Chinthakunta, M., Mantena, S. V., Vahiduddin, S., & Vatambeti, R. (2025). Optimized machine learning mechanism for big data healthcare system to predict disease risk factor. Scientific Reports, 15(1). https://doi.org/10.1038/s41598-025-98721-6

Tirumanadham, N. S. K. M. K., Priyadarshini, V., Praveen, S. P., Thati, B., Srinivasu, P. N., & Shariff, V. (2025). Optimizing lung cancer prediction models: A hybrid methodology using GWO and Random Forest. In Studies in Computational Intelligence (pp. 59–77). https://doi.org/10.1007/978-3-031-82516-3_3

Vahiduddin, S., Chiranjeevi, P., & Krishna Mohan, A. (2023). An analysis on advances in lung cancer diagnosis with medical imaging and deep learning techniques: Challenges and opportunities. Journal of Theoretical and Applied Information Technology, 101(17). https://pmc.ncbi.nlm.nih.gov/articles/PMC10453592/

Veerapaneni, E. J., Babu, M. G., Sravanthi, P., Geetha, P. S., Shariff, V., & Donepudi, S. (2024). Harnessing fusion's potential: A state-of-the-art information security architecture to create a big data analytics model. In Lecture Notes in Networks and Systems (pp. 545–554). https://doi.org/10.1007/978-981-97-6106-7_34

Wang, Q., Gao, J., Lin, W., & Li, X. (2020). NWPU-Crowd: A large-scale benchmark for crowd counting and localization. IEEE Transactions on Pattern Analysis and Machine Intelligence, 44(10), 5311–5328. https://www.researchgate.net/publication/343341991_NWPU-Crowd_A_Large-Scale_Benchmark_for_Crowd_Counting_and_Localization

Yarra, K., Vijetha, S. L., Rudra, V., Balunaik, B., Ramesh, J. V. N., & Shariff, V. (2024). A dual-dataset study on deep learning-based tropical fruit classification. In 2024 8th International Conference on Electronics, Communication and Aerospace Technology (ICECA) (pp. 667–673). IEEE. https://doi.org/10.1109/ICECA63461.2024.10800915

Zhang, Y., Zhou, D., Chen, S., Gao, S., & Ma, Y. (2016). Single-image crowd counting via multi-column convolutional neural network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 589–597). https://www.scirp.org/reference/referencespapers?referenceid=2953502