# **Detection and Tracking of People in a Dense Crowd through Deep Learning Approach Figures and Tables**

Muhammad Firdaus Mohamed Badauraudine1, Dr. Megat Norulazmi Megat Mohamed Noor1, Assoc. Prof. Ts. Dr. Haidawati Mohamad Nasir1, Dr. Mohd Shahizan Othman2

Universiti Kuala Lumpur Malaysian Institute of Information Technology1, Universiti Teknologi Malaysia2

Kuala Lumpur, Malaysia1, Johor Bharu, Malaysia2

[mfirdaus.badauraudine@s.unikl.edu.my](mailto:mfirdaus.badauraudine@s.unikl.edu.my)

+601162694030

**Figures**

FIGURE 1. SLR’s Methodology

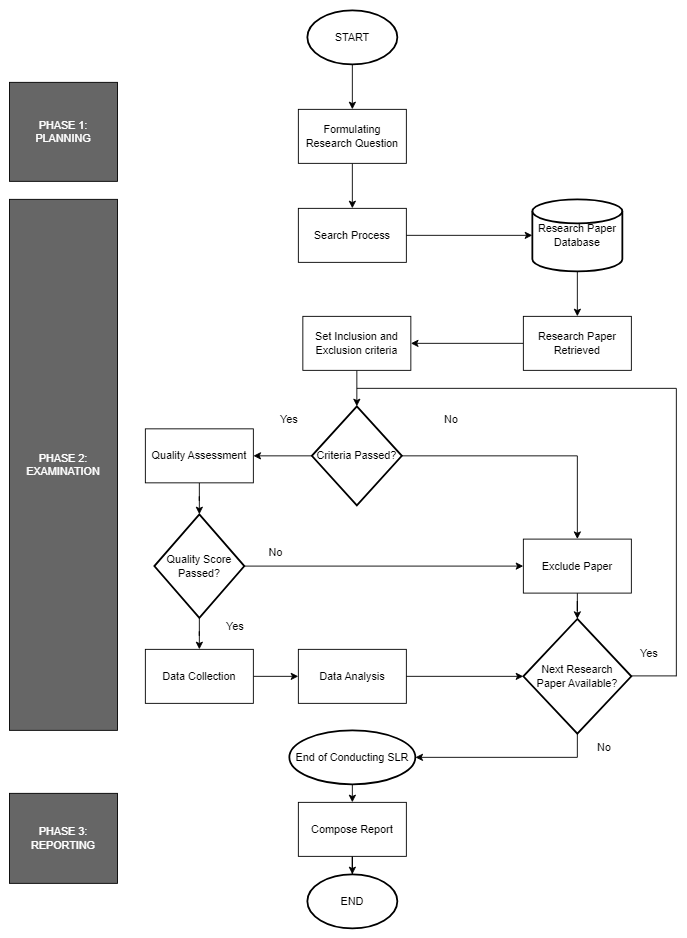


FIGURE 2. PRISMA Diagram for the current SLR

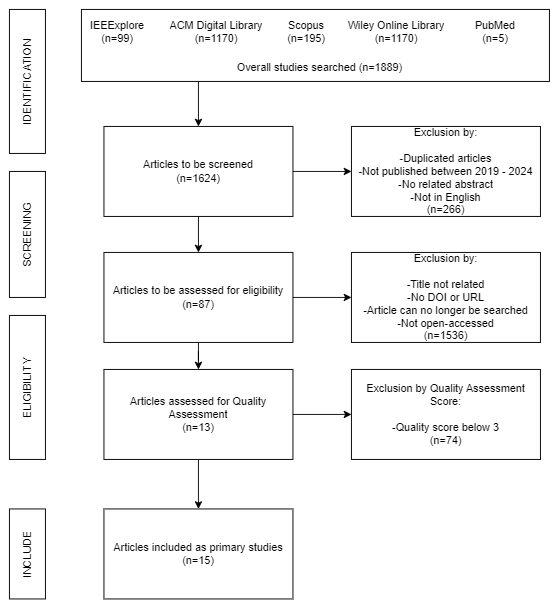


FIGURE 3. Key items of the primary studies visualised by VOSviewer

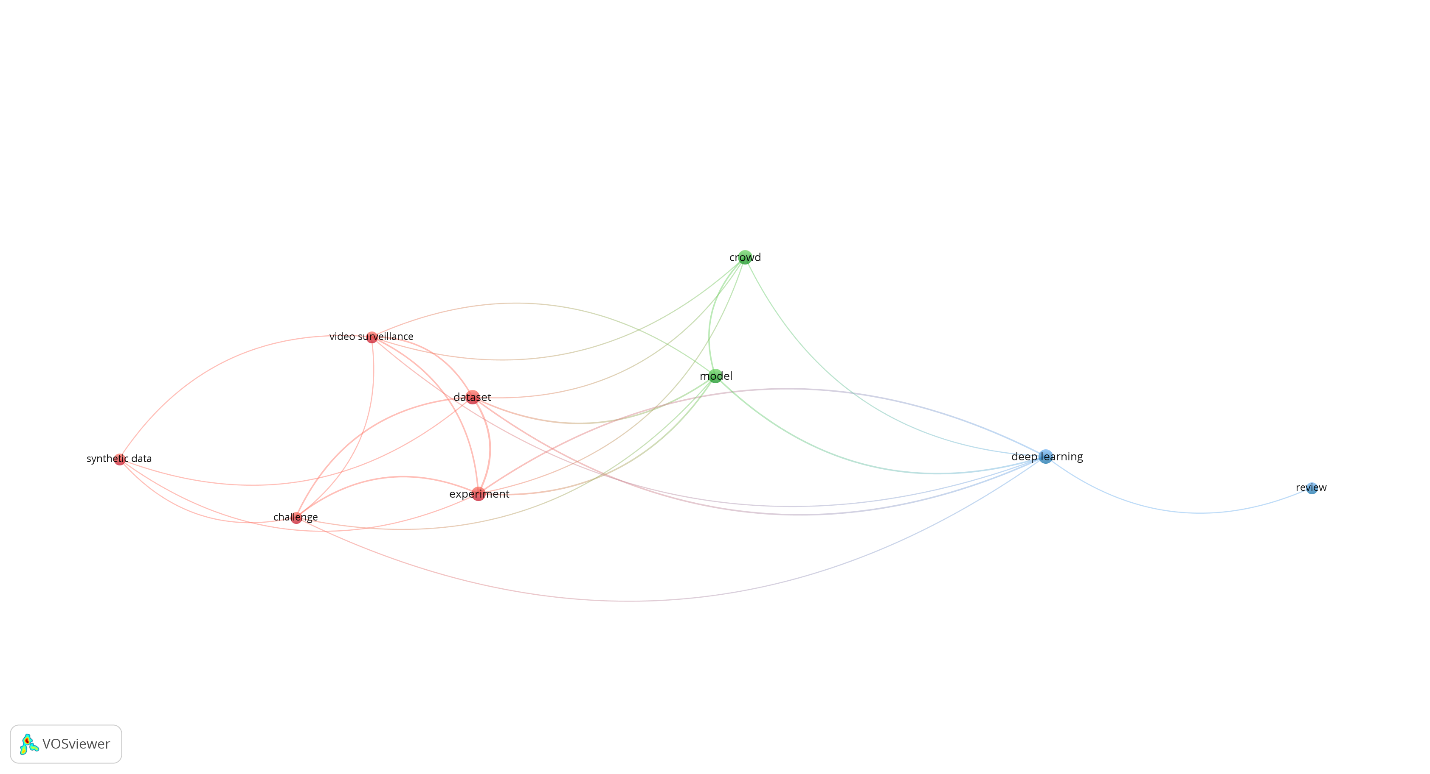


FIGURE 4. Terms of key items from primary studies that are visualised by VOSViewer

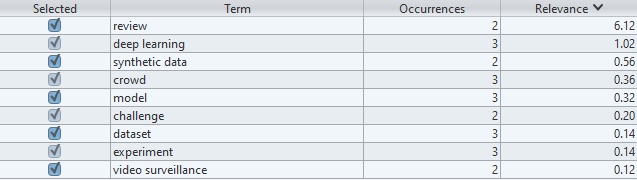


FIGURE 5: How the three DCNNs work together with the PF for tracking people in a dense crowd.

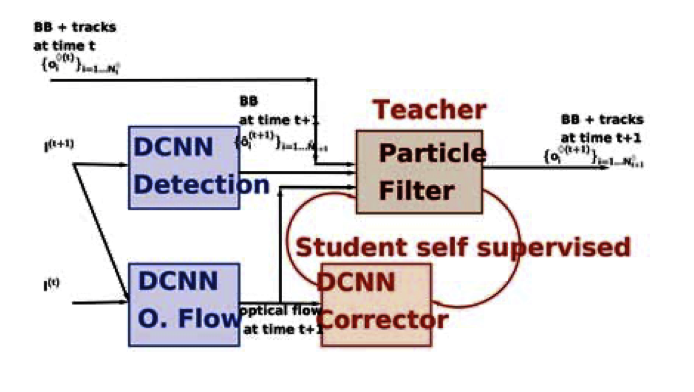
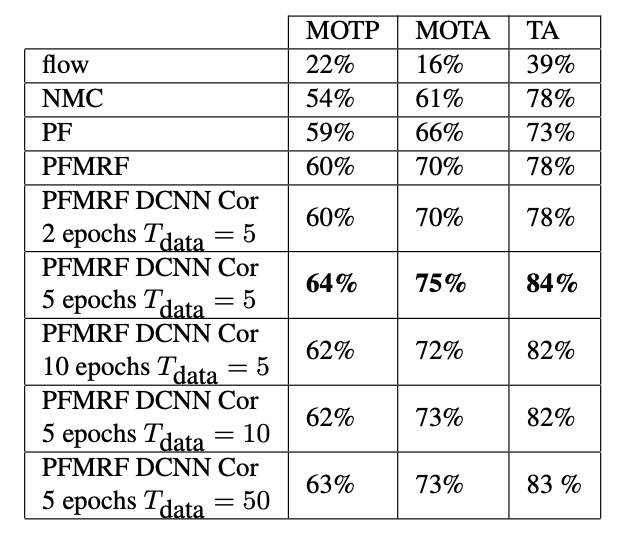


FIGURE 6: Results of PS2's experiments



**Tables**

TABLE 1. Search query used for each database and number of papers gotten from them.

|  |  |  |
| --- | --- | --- |
| Database | Search Query | Number of Papers |
| IEEExplore | *("CROWD") AND ("DEEP LEARNING") AND ("TRACK" OR "TRACKING")* | 99 |
| ACM Digital Library | *("CROWD") AND ("DEEP LEARNING") AND ("TRACK" OR "TRACKING")* | 1170 |
| Scopus | *("CROWD") AND ("DEEP LEARNING") AND ("TRACK" OR "TRACKING")* | 195 |
| Wiley Online Library | *("CROWD") AND ("DEEP LEARNING") AND ("TRACK" OR "TRACKING")* | 420 |
| PubMed | *("CROWD") AND ("DEEP LEARNING") AND ("TRACK" OR "TRACKING")* | 5 |

TABLE 2. Inclusion and Exclusion criteria and papers excluded.

|  |  |  |
| --- | --- | --- |
| Inclusion Criteria | Exclusion Criteria | Number of Papers Excluded |
| Articles must be unique | There are duplicated articles | 140 |
| Paper must be published within 5 years of when the SLR is written (2019 to 2024). | Paper is not published within 2019 to 2024 | 52 |
| Have an abstract related to the current research | No abstract or abstract not related to current research | 71 |
| Paper must be in English | Paper is in a language other than English | 2 |
| Articles’ titles must relate to the current research | Articles not related to current research. Focus is on crowds. If crowd is not present, the article is excluded | 1497 |
| Have a DOI or URL | Doesn’t have both URL and DOI | 11 |
| The same article could still be obtained | The same article cannot be found through the search engine used | 2 |
| The article can be fully accessed | The article is not open-accessed even when using tools such as Zotero and Sci-Hub | 26 |

TABLE 3. Quality Assessment Questions

|  |  |
| --- | --- |
| Quality Assessment | Respective RQ |
| QA 1: The research uses localisation as its deep learning approach | Q1 |
| QA 2: The crowd involved in the study is a dense crowd | Q1 |
| QA 3: The research uses auto annotation for tracking the individuals in a dense crowd | Q2 |
| QA 4: One trained dataset is used on different test datasets | Q2 |
| QA 5: Re-identification of the same person is implemented in the tracking | Q3 |

TABLE 4. Quality Assessment Scoring Criteria

|  |  |
| --- | --- |
| Quality Assessment Scoring | Score |
| The author(s) demonstrated a clear, detailed and explicit explanation on the answers for the specific RQ | HIGH = H = 1 |
| The author(s) provided a general, non-detailed, and non-explicit explanation on the answers for the specific RQ | MEDIUM = M = 0.5 |
| The author(s) either provided no or very few technical details on answering the specific RQ | LOW = L = 0 |

TABLE 5. Papers, the QAs they answer, and their score

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Paper Title | QA answered (Score) | QA Score | Total Score | Include/Exclude |
| Pedestrian Tracking Algorithm for Dense Crowd based on Deep Learning | 1  2  3  4  5 | 0  1  0  0  0.5 | 1.5 | Exclude |
| Simulation of Evacuating Crowd Based on Deep Learning and Social Force Model | 1  2  3  4  5 | 0  1  0  1  0.5 | 2.5 | Exclude |
| Deep Learning and One-class SVM based Anomalous Crowd Detection | 1  2  3  4  5 | 0  0  0  0  0.5 | 0.5 | Exclude |
| Pedestrian Detection and Feedback Application Based on YOLOv5s and DeepSORT | 1  2  3  4  5 | 0  1  0  0.5  1 | 2.5 | Exclude |
| Real-Time Pedestrian Detection using YOLO | 1  2  3  4  5 | 0  1  0  0  1 | 2 | Exclude |
| A Review on Unconstrained Real-Time Rotation-Invariant Face Detection | 1  2  3  4  5 | 0.5  0  0  0  0 | 0.5 | Exclude |
| Selective face de-identification scheme using multiple face recognition and classification techniques | 1  2  3  4  5 | 0.5  0  0  0  1 | 1.5 | Exclude |
| Learning how to analyse crowd behaviour using synthetic data | 1  2  3  4  5 | 0  1  1  1  0 | 3 | Include |
| An Overview of Crowd Counting on Traditional and CNN-based Approaches | 1  2  3  4  5 | 0  0.5  0  0.5  0.5 | 1.5 | Exclude |
| People Counting Based on Multi-scale Region Adaptive Segmentation and Depth Neural Network | 1  2  3  4  5 | 0  0  0  1  0 | 1 | Exclude |
| AcousticID: Gait-based Human Identification Using Acoustic Signal | 1  2  3  4  5 | 0  0  0.5  0  0.5 | 1 | Exclude |
| Person Tracking Using Ankle-Level LiDAR Based on Enhanced DBSCAN and OPTICS | 1  2  3  4  5 | 0.5  0.5  0  0  0.5 | 1.5 | Exclude |
| Deep Learning-Based Crowd Scene Analysis Survey | 1  2  3  4  5 | 0  0.5  0  0  0 | 0.5 | Exclude |
| Tracking Hundreds of People in Densely Crowded Scenes With Particle Filtering Supervising Deep Convolutional Neural Networks | 1  2  3  4  5 | 0  1  1  1  0 | 3 | Include |
| People Tracking System Using DeepSORT | 1  2  3  4  5 | 0  1  0  0.5  0.5 | 2 | Exclude |
| Performance Comparison and Analysis for Large-Scale Crowd Counting Based on Convolutional Neural Networks | 1  2  3  4  5 | 0  1  0  0  0.5 | 2 | Exclude |
| Application of Cognitive Computing for Smart Crowd Management | 1  2  3  4  5 | 0  1  0  0.5  0 | 1.5 | Exclude |
| REGROUP: A Robot-Centric Group Detection and Tracking System | 1  2  3  4  5 | 0  0  0.5  0  1 | 1.5 | Exclude |
| DenseCAvoid: Real-time Navigation in Dense Crowds using Anticipatory Behaviors | 1  2  3  4  5 | 0  1  0  0  0 | 1 | Exclude |
| The Indoor People Tracking and Counting System | 1  2  3  4  5 | 1  0  0  0.5  1 | 2.5 | Exclude |
| Informative scene decomposition for crowd analysis, comparison and simulation guidance | 1  2  3  4  5 | 0  1  0  0  0 | 1 | Exclude |
| Pedestrian motion in simulation applications using deep learning | 1  2  3  4  5 | 0  1  0.5  0  0 | 1.5 | Exclude |
| Crowd Abnormal Behavior Detection Combining Movement and Emotion Descriptors | 1  2  3  4  5 | 0  1  1  1  0 | 3 | Include |
| Informative scene decomposition for crowd analysis, comparison and simulation guidance | 1  2  3  4  5 | 0  1  1  0.5  0 | 2.5 | Exclude |
| 3DGCN: 3-Dimensional Dynamic Graph Convolutional Network for Citywide Crowd Flow Prediction | 1  2  3  4  5 | 0.5  0  0  1  0 | 1.5 | Exclude |
| A sound-based crowd activity recognition with neural network based regression models | 1  2  3  4  5 | 0  0  1  1  0 | 2 | Exclude |
| A Survey on Device-free Indoor Localization and Tracking in the Multi-resident Environment | 1  2  3  4  5 | 0.5  1  0  0  0 | 1.5 | Exclude |
| Crowd activity recognition in live video streaming via 3D-ResNet and region graph convolution network | 1  2  3  4  5 | 0.5  1  0  0.5  0.5 | 2.5 | Exclude |
| Enriched and discriminative convolutional neural network features for pedestrian re-identification and trajectory modeling | 1  2  3  4  5 | 0  0  0.5  1  1 | 2.5 | Exclude |
| Algorithms for Microscopic Crowd Simulation: Advancements in the 2010s | 1  2  3  4  5 | 0.5  0.5  0.5  1  0 | 2 | Exclude |
| Transfer learning-based online multiperson tracking with Gaussian process regression | 1  2  3  4  5 | 0  1  0  1  0 | 2 | Exclude |
| Fast intensive crowd counting model of Internet of Things based on multi-scale attention mechanism | 1  2  3  4  5 | 0  1  0.5  0.5  1 | 3 | Include |
| Passenger Flow Statistics Algorithm of Scenic Spots Based on Multi-Target Tracking | 1  2  3  4  5 | 0  1  0  0.5  1 | 2.5 | Exclude |
| UUCT - HyMP: Towards Tracking Dispersed Crowd Groups from UAVs | 1  2  3  4  5 | 0.5  0  1  1  1 | 3.5 | Include |
| Multi-Scale Occluded Pedestrian Detection Based on Deep Learning | 1  2  3  4  5 | 1  1  0.5  1  0.5 | 4 | Include |
| Crowd Evacuation Simulation Using Hierarchical Deep Reinforcement Learning | 1  2  3  4  5 | 0  1  0  0  0 | 1 | Exclude |
| Automated Pedestrian Tracking Based on Improved ByteTrack | 1  2  3  4  5 | 0  1  0  0  1 | 2 | Exclude |
| Group Tracking for Video Monitoring Systems: A Spatio-Temporal Query Processing Approach | 1  2  3  4  5 | 0  1  0  1  0 | 2 | Exclude |
| STP-TrellisNets: Spatial-Temporal Parallel TrellisNets for Metro Station Passenger Flow Prediction | 1  2  3  4  5 | 0  1  0  0.5  0 | 1.5 | Exclude |
| Scene invariant crowd counting using multi-scales head detection in video surveillance | 1  2  3  4  5 | 0  1  0  0.5  0 | 1.5 | Exclude |
| Practical model with strong interpretability and predictability: An explanatory model for individuals' destination prediction considering personal and crowd travel behavior | 1  2  3  4  5 | 0  0.5  0  0  0 | 0.5 | Exclude |
| Research of Pedestrian Object Tracking Algorithms | 1  2  3  4  5 | 0  0.5  0  0  1 | 1.5 | Exclude |
| People Tracking and Counting using Jetson TX2 Kit with Tracking Algorithm | 1  2  3  4  5 | 0  0.5  0.5  0  0 | 1 | Exclude |
| Locating people in Real-World for Assisting Crowd Behaviour Analysis Using SSD and Deep SORT Algorithm | 1  2  3  4  5 | 0  0.5  0.5  0.5  0 | 1.5 | Exclude |
| Dynamic Gallery for Real-Time Multi-Target Multi-Camera Tracking | 1  2  3  4  5 | 0  0.5  0  0.5  1 | 2 | Exclude |
| Covid-19 crowd detection and alert system using image processing | 1  2  3  4  5 | 0  0  0  1  0 | 1 | Exclude |
| Collision-Line Counting Method Using DeepSORT to Count Pedestrian Flow Density and Hungary Algorithm | 1  2  3  4  5 | 0  0  0  0  0 | 0 | Exclude |
| Deep Learning in Smart Video Surveillance for Crowd Management: A Systematic Literature Review | 1  2  3  4  5 | 1  0  0  0  0 | 1 | Exclude |
| Two-stream Adaptive Convolutional Neural Network for Crowd Counting | 1  2  3  4  5 | 1  0  0  0  0 | 1 | Exclude |
| Cooperative Hierarchical Framework for Group Activity Recognition: From Group Detection to Multi-activity Recognition | 1  2  3  4  5 | 0  1  0  0  0 | 1 | Exclude |
| Development and testing of a pedestrian traffic monitoring system on a convolutional neural network platform | 1  2  3  4  5 | 0  0  0  0  0 | 0 | Exclude |
| Pedestrian Trajectory Prediction in Crowd Scene Using Deep Neural Networks | 1  2  3  4  5 | 0  0.5  0  0.5  0 | 1 | Exclude |
| Deep Person Detection in Two-Dimensional Range Data | 1  2  3  4  5 | 0  0  1  1  0 | 2 | Exclude |
| Deep neural network-based Wi-Fi/pedestrian dead reckoning indoor positioning system using adaptive robust factor graph model | 1  2  3  4  5 | 0  0  0  0  0 | 0 | Exclude |
| Automated Physical Distance Estimation and Crowd Monitoring Through Surveillance Video | 1  2  3  4  5 | 0  1  0.5  0  0 | 1.5 | Exclude |
| Deep Learning Based Crowd Monitoring And Person Identification System | 1  2  3  4  5 | 0  0  0  0  1 | 1 | Exclude |
| Context Aware Crowd Tracking and Anomaly Detection via Deep Learning and Social Force Model | 1  2  3  4  5 | 0  1  0.5  0  0 | 1.5 | Exclude |
| People Counting in Public Spaces using Deep Learning-based Object Detection and Tracking Techniques | 1  2  3  4  5 | 0  1  0  0  0 | 1 | Exclude |
| Implementation of Realtime design of crowd Enumeration via tracking using AI system | 1  2  3  4  5 | 0  0  0  0  0 | 0 | Exclude |
| Realtime Crowd Monitoring—Estimating Count, Speed and Direction of People Using Hybridized YOLOv4 | 1  2  3  4  5 | 0  1  0  0  0 | 1 | Exclude |
| Anomaly Detection Approach for Human Detection in Crowd Based Locations | 1  2  3  4  5 | 0  1  0.5  0.5  0 | 2 | Exclude |
| Rapid Detection of Pilgrims Whereabouts During Hajj and Umrah by Wireless Communication Framework : An application AI and Deep Learning | 1  2  3  4  5 | 0  1  0  0  0 | 1 | Exclude |
| AerialMPTNet: Multi-Pedestrian Tracking in Aerial Imagery Using Temporal and Graphical Features | 1  2  3  4  5 | 0  1  0  0.5  0 | 1.5 | Exclude |
| DensePeds: Pedestrian Tracking in Dense Crowds Using Front-RVO and Sparse Features | 1  2  3  4  5 | 0  1  0  1  0 | 2 | Exclude |
| An Aerial Crowd-Flow Analyzing System for Drone Under YOLOv5 and StrongSort | 1  2  3  4  5 | 0  1  1  0  1 | 3 | Include |
| On the Fine-Grained Crowd Analysis Via Passive WiFi Sensing | 1  2  3  4  5 | 0  1  0  0.5  0 | 1.5 | Exclude |
| Density-based clustering with fully-convolutional networks for crowd flow detection from drones | 1  2  3  4  5 | 0  1  0.5  1  0 | 2.5 | Exclude |
| Fusion of CCTV Video and Spatial Information for Automated Crowd Congestion Monitoring in Public Urban Spaces | 1  2  3  4  5 | 0.5  1  0.5  0.5  1 | 3.5 | Include |
| Deep Reinforcement Learning with Pedestrian Trajectory Prediction Model for Service Robot Navigation in Crowded Environments | 1  2  3  4  5 | 0  0  0  0  0 | 0 | Exclude |
| Person Monitoring by Full Body Tracking in Uniform Crowd Environment | 1  2  3  4  5 | 0  0  0  0  0 | 0 | Exclude |
| Development of a Realistic Crowd Simulation Environment for Fine-Grained Validation of People Tracking Methods | 1  2  3  4  5 | 0  1  0  0  0 | 1 | Exclude |
| Realtime Crowd Monitoring - Estimating Count, Speed and Direction of People Using Hybridized YOLOv4 | 1  2  3  4  5 | 0  1  0  0  0 | 1 | Exclude |
| Multiple Pedestrian Tracking in Dense Crowds Combined with Head Tracking | 1  2  3  4  5 | 0  1  0  0  1 | 2 | Exclude |
| Deep Learning driven automated person detection and tracking model on surveillance videos | 1  2  3  4  5 | 0  0.5  0  1  1 | 2.5 | Exclude |
| Research on Pedestrian Detection and DeepSort Tracking in Front of Intelligent Vehicle Based on Deep Learning | 1  2  3  4  5 | 0  1  0  0.5  0 | 1.5 | Exclude |
| Intelligent Deep Learning Enabled Crowd Detection and Classificai Model in Real Time | 1  2  3  4  5 | 0  0.5  0  0.5  0 | 1 | Exclude |
| Deep Transfer Learning Enabled Intelligent Object Detection for Crowd Density Analysis on Video Surveillance Systems | 1  2  3  4  5 | 0  1  0  0  0 | 1 | Exclude |
| Multi-person multi-camera tracking for live stream videos based on improved motion model and matching cascade | 1  2  3  4  5 | 0  0  0  0  1 | 1 | Exclude |
| A Real-time improved pedestrian dead reckoning trajectory tracking algorithm | 1  2  3  4  5 | 0  1  0  0.5  0 | 1.5 | Exclude |
| Enhancing Real-Time Human Tracking using YOLONAS-DeepSort Fusion Models | 1  2  3  4  5 | 0  1  0  1  1 | 3 | Include |
| Less Is More: Learning from Synthetic Data with Fine-Grained Attributes for Person Re-Identification | 1  2  3  4  5 | 0  0.5  1  1  1 | 3.5 | Include |
| A Survey on Multi-Target Multi-Camera Tracking Methods | 1  2  3  4  5 | 0.5  1  0.5  0  1 | 3 | Include |
| A survey on deep learning-based real-time crowd anomaly detection for secure distributed video surveillance | 1  2  3  4  5 | 0  0.5  0  0.5  1 | 2 | Exclude |
| Handling Heavy Occlusion in Dense Crowd Tracking by Focusing on the Heads | 1  2  3  4  5 | 0  1  0.5  0.5  1 | 3 | Include |
| Multi-objective deep reinforcement learning for crowd-aware robot navigation with dynamic human preference | 1  2  3  4  5 | 1  1  0.5  0.5  0.5 | 3.5 | Include |
| Topology and channel affinity reinforced global attention for person re-identification | 1  2  3  4  5 | 1  1  0  0  1 | 3 | Include |
| A convolutional neural-network-based pedestrian counting model for various crowded scenes | 1  2  3  4  5 | 0  1  0.5  0.5  0 | 2 | Exclude |

TABLE 6. Primary Studies and its most respected RQ

|  |  |
| --- | --- |
| Primary Study | Respected RQ |
| Learning how to analyse crowd behaviour using synthetic data | RQ2 |
| Tracking Hundreds of People in Densely Crowded Scenes With Particle Filtering Supervising Deep Convolutional Neural Networks | RQ2 |
| Crowd Abnormal Behavior Detection Combining Movement and Emotion Descriptors | RQ2 |
| Fast intensive crowd counting model of Internet of Things based on multi-scale attention mechanism | RQ3 |
| UUCT - HyMP: Towards Tracking Dispersed Crowd Groups from UAVs | RQ2 |
| Multi-Scale Occluded Pedestrian Detection Based on Deep Learning | RQ1 |
| An Aerial Crowd-Flow Analyzing System for Drone Under YOLOv5 and StrongSort | RQ2 |
| Fusion of CCTV Video and Spatial Information for Automated Crowd Congestion Monitoring in Public Urban Spaces | RQ1 |
| Enhancing Real-Time Human Tracking using YOLONAS-DeepSort Fusion Models | RQ3 |
| Less Is More: Learning from Synthetic Data with Fine-Grained Attributes for Person Re-Identification | RQ2 |
| A Survey on Multi-Target Multi-Camera Tracking Methods | RQ3 |
| Handling Heavy Occlusion in Dense Crowd Tracking by Focusing on the Heads | RQ1 |
| Topology and channel affinity reinforced global attention for person re-identification | RQ1 |

TABLE 7. Summary of Primary Studies

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| ID | TITLE | REFERENCE | PUBLISHER | TYPE | RESPECTED RQ | YEAR |
| PS1 | Learning how to analyse crowd behaviour using synthetic data | (A.R. Khadka et al., 2019) | ACM Digital Library | Conference Paper | RQ2 | 2019 |
| PS2 | Tracking Hundreds of People in Densely Crowded Scenes With Particle Filtering Supervising Deep Convolutional Neural Networks | (Gianni Franchi et al., 2020) | IEEE | Conference Paper | RQ2 | 2020 |
| PS3 | Crowd Abnormal Behavior Detection Combining Movement and Emotion Descriptors | (Xiao Li et al., 2020) | ACM Digital Library | Conference Paper | RQ2 | 2020 |
| PS4 | Fast intensive crowd counting model of Internet of Things based on multi-scale attention mechanism | (Dong Liu et. al., 2022) | The Institute of Engineering Technology | Article | RQ3 | 2022 |
| PS5 | UUCT - HyMP: Towards Tracking Dispersed Crowd Groups from UAVs | (Tonmoay Deb et. al., 2021) | IEEE | Conference Paper | RQ2 | 2021 |
| PS6 | Multi-Scale Occluded Pedestrian Detection Based on Deep Learning | (Fang Li et. al., 2022) | IEEE | Journal Article | RQ1 | 2022 |
| PS7 | An Aerial Crowd-Flow Analyzing System for Drone Under YOLOv5 and StrongSort | (Kuan-Hao Yeh et. al., 2022) | IEEE | Conference Paper | RQ2 | 2022 |
| PS8 | Fusion of CCTV Video and Spatial Information for Automated Crowd Congestion Monitoring in Public Urban Spaces | (Vivian W. H. Wong et. al., 2023) | MDPI | Article | RQ1 | 2023 |
| PS9 | Enhancing Real-Time Human Tracking using YOLONAS-DeepSort Fusion Models | (Athilakshmi R. et. al., 2024) | ResearchGate | Conference Paper | RQ3 | 2024 |
| PS10 | Less Is More: Learning from Synthetic Data with Fine-Grained Attributes for Person Re-Identification | (Suncheng Xiang et. al., 2021) | arXiv | Article | RQ2 | 2021 |
| PS11 | A Survey on Multi-Target Multi-Camera Tracking Methods | (Temitope Ibrahim Amosa et. al., 2023) | ScienceDirect | Survey Paper | RQ3 | 2023 |
| PS12 | Handling Heavy Occlusion in Dense Crowd Tracking by Focusing on the Heads | (Yu Zhang et. al., 2023) | arXiv | Article | RQ1 | 2023 |
| PS13 | Topology and channel affinity reinforced global attention for person re-identification | (Xile Wang et. al., 2021) | Wiley Online Library | Research Article | RQ1 | 2021 |