Journal of Information Systems Engineering and Management
2025, 10(30s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Unusual Activity Detection for Fish-eye Camera : A Review

Dr. Abhijit Patil, Akanksha Pinto, Diptee Sable, Niharika Sinha
Department of Computer Engineering, K. J. Somaiya Institute of Technology, Mumbai, India

ARTICLE INFO

Received: 30 Dec 2024

Revised: 12 Feb 2025

Accepted: 26 Feb 2025

ABSTRACT

Fisheye cameras provide an expansive field of view that surpasses 180° and are frequently employed in high-security environments, including retail establishments and parking facilities. Nevertheless, these cameras pose significant obstacles for conventional surveillance frameworks, since the inherent distortion in their images complicates the identification of atypical behaviors such as shoplifting or vehicle theft, necessitating human oversight. Current systems frequently encounter difficulties in effectively recognizing such occurrences. To answer this question, we present a study in several algorithms that could help in the detection of anomalous behaviors of fisheye images. Our approach entails a new deep learning methodology wherein automatic and accurate detection will be made possible through the application of convolutional neural networks. We apply the YOLOv8 model with Xception as our base model to identify abnormal activities correctly and locate frames in which suspicious actions occur. This method enhances crime deterrence through instantaneous detection and, in addition, can act as a cost-effective alternative to current monitoring methods, thereby significantly reducing reliance on labor-intensive monitoring.

Keywords: Fish-eye camera, Unusual Activity, YOLOv8, Xception, Deep Learning, CNN

I. INTRODUCTION

The fisheye camera [6] is rather uncommon since it is a surveillance camera with the capability of taking views at more than 180°. This makes it rather useful in important security areas. The unique feature about the camera is that it makes radial distortion on images such that an object's edges appear squeezed. This distortion can blur things and make it hard to keep track of and identify things. But with a fisheye camera, one gets to enjoy the advantage of wide coverage using only one camera. Fisheye cameras are best used in high-security places like banks, highways, parking lots, and stores.

With the rise of crime, unusual activities [3] like car theft, burglaries, and shoplifting are increasingly perpetrated in public as well as private places. Such criminal activities stand tall financially and risk lives lost in extreme cases. For example, in the old surveillance systems, a security officer has to view video feeds visually and attempt to trace the activities of individuals and their vehicles. This is very inefficient and also a high-risk operation due to human imperfection. It is very time-consuming and error-prone as well. With such huge volumes of data from the new surveillance systems, humans alone can't be relied on to flag suspicious activities. Besides, fisheye camera footage is usually processed with distortion inherent in it and rarely gives the right results when applying traditional computer vision techniques.

Object tracking [7] is also referred to as the method to maintain object recognition stable across successive frames in a video. It is an inevitable process of surveillance applications since the behavior of objects' movement will be understood over time, and it helps to find abnormal activities, such as theft. Tracking in dynamic fisheye video is harder since the motion of objects and camera motion change the appearance of objects and complicate the good tracking ability expected from traditional methods.

With the arrival of AI and ML [1] in recent years, the surveillance field has changed by creating tools for new challenges. Moreover, deep learning can handle vast, complex data, allowing the CNNs to automatically identify patterns in images without the need for handcrafted features or feature extraction. We can thus design systems that find objects and odd behaviors for themselves and automatically follow them. Then we need not maintain constant human watching.

In laying out a deep survey of the prior work in applying fisheye cameras in surveillance systems regarding object detection and behavior, discussions regarding the inadequacies of traditional methods when applied for dealing with distortion in fisheye images are held. We also present a new approach using cutting-edge deep learning techniques to boost object and anomalous activity detection in video from fisheye cameras. For fast, single-shot object detection, we use the YOLOv8 [6] system. The YOLOv8 system uses Xception as the backbone model; Xception is a CNN [9]-based model and has exceptional feature extraction capability even with images distorted by fisheye distortion.

The proposed system is set to manage video from fisheye cameras automatically, and cases of shoplifting or car theft can be spotted precisely. The system will send the real-time alert and capture the moment when suspicious activity occurs. This system will work very efficiently without human intervention and will improve accuracy and speed significantly compared to the current methods in place.

This proposed system is supposed to work autonomously without any human intervention. It is going to significantly improve detection accuracy and response time with all the many constraints present in state-of-the-art methodologies. Integrate the YOLOv8 with Xception and create an effective system that is very efficient and will detect robberies in shopping malls and car parking so that all measures can be taken to prevent such crimes and also reduce the cost of operation.

The rest of this paper is structured as follows: Section II gives a short overview of work that might be related to object detection and tracking using fish-eye camera videos. Section III clarifies what our proposed method on object detection and tracking is. Section IV reviews the paper and presents ideas for further research.

II. LITERATURE REVIEW

Detecting suspicious human activity and unusual behaviors has grown in importance in recent years in domains including crowd monitoring, self-driving, and security surveillance. Machine learning has been widely used to improve the accuracy of detection and reliability in these difficult contexts, especially deep learning and ensemble approaches. Advanced models such as YOLO, SVM, and 3D Convolutional Neural Networks (3D-CNN) are employed in studies to identify unusual activity, detect humans, and recognize poses. Methods such as ensemble classifiers, which integrate models like XGBoost and Random Forest, have demonstrated efficacy in enhancing classification performance.

Ensemble machine learning techniques improve detection performance while detecting suspicious human activity and anomalous behaviors. A study by [1] used SVM to identify human poses along with YOLO and OpenCV for person detection. For precise detection, unusual activity frames were isolated, and pose classification performance was improved by an ensemble classifier that used Random Forest and XGBoost. Due to limitations of public datasets, the system requires a custom dataset, and the accuracy depends on the quantity of the dataset. Additionally, frames must be manually filtered. Another approach [2] uses transfer learning and 3D Convolutional Neural Networks (3D-CNN) to detect anomalous activities in video footage. It enhances detection capabilities using feature extraction. Dynamic backgrounds and occlusions reduce detection accuracy, and training effectiveness is impacted by data scarcity.

Phong et al. [3] used OpenCV and Motion Influence Maps to examine movement patterns across frames for crowd scenarios. K-mean clustering is used to identify unusual activities. The efficacy of this approach declines with the number of users and in complex backgrounds, making generalization for different scenarios hard. Deep learning models such as Single Shot Detector (SSD) are combined with transfer learning in security contexts. Bounding boxes are used for object detection, and frames are processed through OpenCV. It works well for predetermined tasks. Mushtaq et al. [4] found that it has trouble with real-time implementation and tasks beyond its predefined scope.

Li et al. [5] investigated the use of Markov Logic Networks (MLNs) for the classification of suspicious activities. It classifies actions by extracting sub-actions using fuzzy classification, training on labeled examples. Large datasets provide computational challenges for MLNs and are best suited for controlled environments where performance can be degraded by noisy data. A 2023 study [6] tested YOLOv8 and Mask-RCNN for people counting in fisheye camera applications and found that the performance of YOLOv8-N was better because it used fewer resources. The results were hindered by fisheye distortion and human labeling errors. Another study [7] developed a Fisheye Multi-Object Tracking (FEMOT) system using distorted Fisheye Image Augmentation (DFIA) and Hybrid Data Association (HDA) to support object detection and tracking. However, fisheye-specific augmentation continues to be difficult due to high computational cost.

Inspired by YOLOv3 and FPN architectures, CFPN was created for small object identification on fisheye cameras in traffic flow estimation [8]. The system faces difficulty in detecting tiny things and maximizes computational efficiency on embedded devices. Other approaches, such as the Artificial Fisheye Image Augmentation used OpenCV and COCO2017 distortion techniques [9], and object detection is done by YOLOv8. The method is not best to generalize due to reliance on fisheye distortion. Training requires a lot of computing power.

Researchers applied curved bounding box approaches to autonomous driving on the Woodscape dataset to increase detection precision [10] using YOLO with a ResNet18 and non-max suppression. Training time rises as curve and polygon fitting complexity grows. The Fisheye8k dataset [11] used a modified YOLOv3 with a ResNet18 encoder and a ranger optimizer for easier detection of objects on fisheye cameras for self-driving. Results were affected by distortions and uneven class distributions. Semi-supervised anomaly identification was the last method used for autonomous cars, using labeled and unlabeled data for better detection [12]; nonetheless, fisheye distortion was offered.

Table. 1. Literature Review

Ref. No	Authors	Methodology	Limitations
[1]	Aqil Shamnath, Meena Belwal	Human detection is done using YOLO and OpenCv. SVM model is used for human pose detection. XGBOOST/ Random forest classifier is used for pose classification. Ensemble voting classifier is used for superior prediction performance. Frames are extracted from the video for the unusual activity detection. The alert is sent to the admin and the authorities.	It requires custom data creation as it lacks availability of public dataset for suspicious activity. The system accuracy is affected by both dataset size and manual filtering of irrelevant frames.
[2]	Rajshri C. Mahajan, Namrata K. Pathare, Vibha Vyas	To identify and study unusual human behavior in video footage, 3D Convolution Neural Network along with transfer learning is used. 3D CNN layers used for feature extraction.	The accuracy of detection is damaged by environmental variations such as moving backgrounds and occlusions. The dataset faces limitations due to insufficient training data.
[3]	Arun Kumar Jhapate, Sunil Malviya, Monika Jhapate	Recognizes human activity in crowds using OpenCV and Motion Influence map to identify usual or unusual actions. The movement between frames is analyzed by Motion Influence. The detected activities then clustered using K-Means.	The accuracy is limited by multi user activities and complex backgrounds. The method may not generalize well to all scenarios.
[4]	Ajeet Sunil, Manav Hiren Sheth, Shreyas E	The study utilizes Single Shot Detector (SSD) along with transfer learning to detect usual and unusual activities. OpenCV is used for frames processing and object detection is done by bounding boxes.	Limited to a predefined set of activities and objects. Unusual activities outside this scope are not detected. Challenges in real-time implementation.

[5]	Aditi Kapoor, K.K.Biswas, M.Hanmandlu	Markov Logic Networks (MLNs) is used for classification of unusual activities. The actions and sub-actions from video data are extracted through fuzzy classification. MLNs are trained using labeled examples to classify new activity as either usual or unusual. The dataset used is the KTH dataset.	While dealing with large datasets MLNs introduce computational complexity. The system is designed to detect activities in a controlled environment. Inaccuracies can be established while handling noisy data.
[6]	Jevgenijs Telicko, Andris Jakovics	The dataset is generated from three fisheye cameras namely OV3660, OV2640, and IMX219. The labeling was done using the VIA tool. YOLOv8 and Mask-RCNN were compared for object detection and instance segmentation. YOLOv8-N achieved better performance due to lower resource demands.	Difficulty is obtained to adapt models trained on projective camera data to fisheye distortions. Human labeling errors and the need for fisheye correction impacted model performance.
[7]	Ping-Yang Chen, Jun-Wei Hsieh, Ming-Ching Chang, Munkhjargal Gochoo, Fang-Pang Lin, Yong-Sheng Chen	The study involves developing the Fisheye Multi-Object Tracking (FEMOT) system to track objects in fisheye camera views. Distorted Fisheye Image Augmentation (DFIA) is used to convert normal camera images to fisheye samples and Hybrid Data Association (HDA) for object detection. YOLOX is used for detection along with BoT-SORT-R for tracking, and OSNet for re-identification.	The system reliance on pre-existing dataset created for perspective cameras is challenging. High computational cost associated with fisheye-specific data augmentation.
[8]	Ping-Yang Chen, Jun-Wei Hsieh, Munkhjargal Gochoo, Chien-Yao Wang, Hong-Yuan Mark Liao	Concatenated Feature Pyramid Network (CFPN) is developed for small object detection using fish-eye cameras. Inspired by the architectures of YOLOv3 and FPN.	Struggles with the detection of extremely small objects. Computational efficiency is limited on some embedded devices due to network complexity.
[9]	Maris Broks, Jevgenijs Telicko, Andris Jakovics	The dataset used is COCO2017 and OpenCV fisheye distortion techniques are applied. Object detection is done by YOLOv8.	The generalizability is reduced due to reliance on fisheye distortions. The training process was constrained by computational resources.
[10]	Hazem Rashed, Eslam Mohamed, Ganesh Sistu, Varun Ravi Kumar, Ciaran	They have used the woodscape dataset. The disadvantages of general approach have been discussed. They propose a new curved bounding box technique for precise detection. YOLO with Resnet18 has been used along with	The training time may increase due to computational complexity of fitting curve and polygon representations.

	Eising, Ahmad El-Sallab ,Senthil Yogamani	non-max suppression.	
[11]	Munkhjargal Gochoo,Munkh-Erdene Otgonbold, Erkhembayar Ganbold, Jun-Wei Hsieh, et. al.	Valeo fisheye dataset has been used. YOLOVv3 has been modified using Resnet 18 encoder for detection of suspicious activities and ranger optimizer. Various bounding box techniques have been used.	Dataset has significant distortion from the fisheye lens. It suffers from an imbalanced class distribution.
[12]	Arjun S. Dileep, Nabilah S. S., Sreeju S., Farhana K., Surumy S.	The pose of the human while doing a suspicious activity has been analyzed using the points such as knees, elbows, etc. This is then fed to the CNN which will classify the frame with the label of what the unusual activity is.	The model is limited to activities detectable via pose analysis. The system may struggle with generalization. The reliance on 2D pose estimation can limit the model's accuracy.
[13]	Dimitris Tsiktsiris, Antonios Lalas, Minas Dasygenis, Konstantinos Votis	The research employs a semi-supervised anomaly detection approach using overhead fisheye cameras. It integrates deep learning techniques to analyze vehicle behavior and identify anomalies, leveraging both labeled and unlabelled data to enhance detection accuracy.	Fisheye distortion can make it harder to accurately track vehicle behavior at extreme angles.

III. PROPOSED METHODOLOGY

Convolutional neural networks have been consistently used in object detection models. This is because of their ability to understand, recognize, and detect complex patterns. CNNs use a variety of complex mathematical functions to process the training data and learn the patterns in it. It then proceeds to find the same patterns in the input given, followed by classifying that input. YOLOv8, also known as You Look Only Once version 8, is a very popular algorithm. It is used for object detection and owes its popularity to the fact that it is not only accurate but also fast. It has five variants, namely nano, small, medium, large, and extra large. Depending on the user's needs, a suitable model can be selected. YOLOv8's architecture can be divided into three parts, namely the head, neck, and backbone. The backbone forms the major part of the architecture. The backbone is responsible for recognizing and extracting the various features from the images in the training dataset. The backbone of YOLOv8 uses a custom CSPDarknet53, which has 53 convolutional layers. This backbone requires a lot of computations to train since it has over 50 layers. We propose a model that will use Xception as the backbone rather than using CSPDarknet53. Xception has been especially designed to extract features from images, which will in turn increase the overall accuracy of the model. We have specifically chosen Xception as it uses depthwise separable convolutions. This helps it to perform more precise spatial and channel-wise feature representation that too in a faster and more accurate manner. The neck of YOLOv8 combines the semantic features with the spatial information. This increases the chances of correctly classifying the image. The head of YOLOv8 is responsible for the final classification that is done by the model. In our case, it will classify the video frame as whether it depicts unusual activity—shoplifting and vandalism in parking lots or not.

I. Dataset creation and preprocessing

The initial step involves selecting a real-world dataset containing videos that have recorded the activities of shoplifting and vandalism in parking lots. These will be taken from various open-source platforms such as Roboflow, Kaggle, and GitHub. The datasets will be downloaded, and the videos will then be broken down into frames. This will be done using VLC media player, which offers the option to perform such an action satisfactorily. The frames will then undergo the process of being warped using Python libraries like OpenCV and NumPy. This ensures that the images will have the same distortions as those of the images that are captured by a real fish-eye camera.

This procedure has been followed to create a synthetic dataset. This will compensate for the lack of availability of datasets in the domain of our study. The created synthetic dataset will further be used to train the model. For validating the working of the model, the collection of real-world images will be used. These images are captured by fish-eye lenses originally and not synthetically prepared. Thus, the model's working in the real world will be validated.

II. Model Development

Model development involves training the YOLOv8 model on the new synthetic dataset. The model will then be trained on the customized model, which uses Xception as its backbone, further enhancing feature extraction.

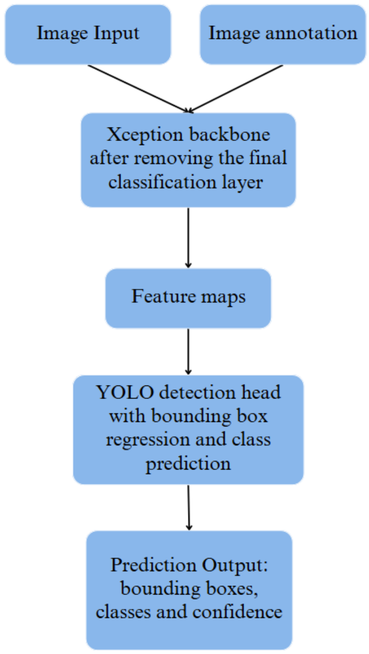

Fig. 1. Block Diagram for the proposed hybrid model

1. Understanding the architecture of YOLOv8 and Xception plays a crucial role in the development of a hybrid model. To integrate the two architectures, we need to understand the in-depth workings of both the models. Xception is mainly used for classification tasks, while YOLO is used for object detection. We plan to remove the top layers of Xception and then use the remaining in YOLO. This will act as an efficient feature extractor.
2. Xception is a convolutional neural network that has been inspired by the architecture of Google's Inception modules. Xception performs well on large-scale datasets as compared to other networks like the ResNet family.
3. Xceptions architecture has three main parts. The initial entry flow is for low-level extraction. This is followed by the middle flow to capture the high-level features. The exit-level flow consolidates the features and prepares them for classification. The middle flow comprises the depthwise separable convolutional layers for high-level feature extraction.
4. Yolo's detection head will receive the output of Xception. Using these high-level features, we will effectively be able to identify the features in the images and draw bounding boxes around the suspects.
5. GPU acceleration will be required to train the model. The dataset is huge, and the networks are intricately designed, thus requiring more computational power.
6. This new hybrid model will then be trained on the prepared dataset.

III. Metrics used for evaluation and validation

The evaluation and validation of the model play an important role in gauging how effective the developed model will be. In the paper [1], a confusion matrix along with evaluation metrics such as accuracy, F1-score, precision, specificity, false positive rate, and false negative rate has been used. The developed model will provide bounding boxes around people participating in the act of shoplifting. The mentioned parameters will be used to evaluate the correctness of the predicted bounding boxes. In [2], the various evaluation metrics, such as accuracy, precision, specificity, and sensitivity, have been calculated class-wise. The mean absolute error has been calculated as well, along with the previously mentioned metrics in [6]. The paper compares 9 pretrained models using those evaluation metrics.

From this we can conclude that the most effective metric for the evaluation and validation depends on the task at hand, the dataset used for training, and the output given by the model. We have decided to use a confusion matrix and calculate the accuracy, F1 score, false positive rate, false negative rate, and precision. This will be calculated for the two classes, namely suspicious activity and normal, predicted by our model. The confusion matrix will give us the overall view of how many of the boxes predicted by the model are correct for each of the classes. The accuracy and precision will let us know how well the model has captured the intricacies of predicting over a warped image. The number of false negatives will prove how well the model is suited for real-world deployment. Each false negative means an escaped perpetrator, and thus efforts will be made to keep this number low. These metrics are applied over the count of bounding boxes. For further scrutiny, metrics like Intersection over Union will also be evaluated in the validation dataset's bounding boxes. Such a metric will help us understand the development of the bounding box itself. This is crucial for understanding how well the bounding box itself was made, especially in the case of images taken by fish-eye lenses.

IV. CONCLUSION

In summary, fisheye cameras are crucial for surveillance, especially in high-security locations where broad perspectives are advantageous. However, conventional object detection and analysis of behavior are challenged by their intrinsic visual distortions. While dataset restrictions, environmental unpredictability, and fisheye distortions continue to be obstacles, models like YOLO, 3D-CNN, and ensemble classifiers (e.g., Random Forest, XGBoost) show better accuracy in detecting and classifying abnormal activities. Even though methods like transfer learning and specific data augmentation (like fisheye augmentation) have shown potential, they can be difficult to generalize in many circumstances and frequently come with large computing costs. In order to overcome these constraints and increase the accuracy, the research in recent development in deep learning, specifically YOLO8 with Xception. Our suggested method greatly lessens the need for manual monitoring by using convolutional neural networks to automatically identify anomalous activity in real time. By investigating improved data augmentation techniques and different network topologies designed specifically for fisheye cameras, future research could further optimize model performance while guaranteeing robustness and scalability.

REFERENCES

[1] A. Shamnath and M. Belwal, "Human Suspicious Activity Detection Using Ensemble Machine Learning Techniques," 2022 2nd International Conference on Intelligent Technologies (CONIT), Hubli, India, 2022, pp. 1-5, doi: 10.1109/CONIT55038.2022.9848183.

[2] R. C. Mahajan, N. K. Pathare and V. Vyas, "Video-based Anomalous Activity Detection Using 3D-CNN and Transfer Learning," 2022 IEEE 7th International conference for Convergence in Technology (I2CT), Mumbai, India, 2022, pp. 1-6, doi: 10.1109/I2CT54291.2022.9824607.

[3] A. K. Jhapate, S. Malviya and M. Jhapate, "Unusual Crowd Activity Detection using OpenCV and Motion Influence Map," 2nd International Conference on Data, Engineering and Applications (IDEA), Bhopal, India, 2020, pp. 1-6, doi: 10.1109/IDEA49133.2020.9170704.

[4] A. Sunil, M. H. Sheth, S. E and Mohana, "Usual and Unusual Human Activity Recognition in Video using Deep Learning and Artificial Intelligence for Security Applications," 2021 Fourth International Conference on Electrical, Computer and Communication Technologies (ICECCT), Erode, India, 2021, pp. 1-6, doi: 10.1109/ICECCT52121.2021.9616791.

[5] A. Kapoor, K. K. Biswas and M. Hanmandlu, "Unusual human activity detection using Markov Logic Networks," 2017 IEEE International Conference on Identity, Security and Behavior Analysis (ISBA), New Delhi, India, 2017, pp. 1-6, doi: 10.1109/ISBA.2017.7947700.

[6] J. Telicko and A. Jakovics, "Comparative Analysis of YOLOv8 and Mack-RCNN for People Counting on Fish-Eye Images," 2023 3rd International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME), Tenerife, Canary Islands, Spain, 2023, pp. 1-6, doi: 10.1109/ICECCME57830.2023.10252265.

[7] P. -Y. Chen, J. -W. Hsieh, M. -C. Chang, M. Gochoo, F. -P. Lin and Y. -S. Chen, "Fisheye Multiple Object Tracking by Learning Distortions Without Dewarping," 2023 IEEE International Conference on Image Processing (ICIP), Kuala Lumpur, Malaysia, 2023, pp. 1855-1859, doi: 10.1109/ICIP49359.2023.10222872.

[8] P. -Y. Chen, J. -W. Hsieh, M. Gochoo, C. -Y. Wang and H. -Y. M. Liao, "Smaller Object Detection for Real-Time Embedded Traffic Flow Estimation Using Fish-Eye Cameras," 2019 IEEE International Conference on Image Processing (ICIP), Taipei, Taiwan, 2019, pp. 2956-2960, doi: 10.1109/ICIP.2019.8803719.

[9] M. Broks, J. Teličko and A. Jakovičs, "Artificial Fish-Eye Image Augmentation Approach for CNN Based Object Detection," 2024 IEEE 11th Workshop on Advances in Information, Electronic and Electrical Engineering (AIEEE), Valmiera, Latvia, 2024, pp. 1-6, doi: 10.1109/AIEEE62837.2024.10586656.

[10] H. Rashed et al., "Generalized Object Detection on Fisheye Cameras for Autonomous Driving: Dataset, Representations and Baseline," 2021 IEEE Winter Conference on Applications of Computer Vision (WACV), Waikoloa, HI, USA, 2021, pp. 2271-2279, doi: 10.1109/WACV48630.2021.00232.

[11] LGochoo, Munkhjargal & Otgonbold, Munkh-Erdene & Ganbold, Erkhembayar & Hsieh, Jun-Wei & Chang, Ming-Ching & Chen, Ping-Yang & Dorj, Byambaa & Jassmi, Hamad & Batnasan, Ganzorig & Alnajjar, Fady & Abduljabbar, Mohammed & Lin, Fang-Pang. (2023). FishEye8K: A Benchmark and Dataset for Fisheye Camera Object Detection. 5305-5313. 10.1109/CVPRW59228.2023.00559.

[12] A. S. Dileep, N. S. S., S. S., F. K. and S. S., "Suspicious Human Activity Recognition using 2D Pose Estimation and Convolutional Neural Network," 2022 International Conference on Wireless Communications Signal Processing and Networking (WiSPNET), Chennai, India, 2022, pp. 19-23, doi: 10.1109/WiSPNET54241.2022.9767152.

[13] D. Tsiktsiris, A. Lalas, M. Dasygenis and K. Votis, "Enhancing the Safety of Autonomous Vehicles: Semi-Supervised Anomaly Detection With Overhead Fisheye Perspective," in IEEE Access, vol. 12, pp. 68905-68915, 2024, doi: 10.1109/ACCESS.2024.3399841.

[14] J. Kannala and S. S. Brandt, "A generic camera model and calibration method for conventional, wide-angle, and fish-eye lenses," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 28, no. 8, pp. 1335-1340, Aug. 2006, doi: 10.1109/TPAMI.2006.153.

[15] C. -C. Hsu, W. -H. Huang, W. -H. Tseng, M. -H. Wu, R. -J. Xu and C. -M. Lee, "OmniDet: Omnidirectional Object Detection via Fisheye Camera Adaptation," 2024 IEEE 7th International Conference on Multimedia Information Processing and Retrieval (MIPR), San Jose, CA, USA, 2024, pp. 335-341, doi: 10.1109/MIPR62202.2024.00060.

[16] L. H. Pham et al., "Improving Object Detection to Fisheye Cameras with Open-Vocabulary Pseudo-Label Approach," 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Seattle, WA, USA, 2024, pp. 7100-7107, doi: 10.1109/CVPRW63382.2024.00704.

[17] A. Oner, M. S. Albayrak, F. Guner and I. M. Atakli, "An Activity Recognition Application Based on Markov Decision Process Through Fish Eye Camera," 2019 IEEE Second International Conference on Artificial Intelligence and Knowledge Engineering (AIKE), Sardinia, Italy, 2019, pp. 251-258, doi: 10.1109/AIKE.2019.00050.

[18] A. Kasthuri, A. Balamurali, A. K. Srinivasan and P. Shanmuga Priya, "Jagriti via Mudra: A pose based surveillance anomaly detection system," 2024 International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI), Chennai, India, 2024, pp. 1-6, doi: 10.1109/ACCAI61061.2024.10602098.

[19] S. Phatangare, S. Kate, D. Khandelwal, A. Khandetod and A. Kharade, "Real Time Human Activity Detection using YOLOv7," 2023 7th International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), Kirtipur, Nepal, 2023, pp. 1069-1076, doi: 10.1109/I-SMAC58438.2023.10290168.

[20] N. Bordoloi, A. K. Talukdar and K. K. Sarma, "Suspicious Activity Detection from Videos using YOLOv3," 2020 IEEE 17th India Council International Conference (INDICON), New Delhi, India, 2020, pp. 1-5, doi: 10.1109/INDICON49873.2020.9342230.