IFERP
Explore Your Research Journey...

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Vol 12, Issue 04, April 2025

Smart AI- Powered Traffic Surveillance System for Identifying Using Optimized Deep Feature Engineering with Fuzzy Capsuled Convolution Neural Network

[1] Dhivya Thirisha. R, [2] Pavan kumar. K, [3] Akilesh Praveen. B

[1] [3] B.Tech - Artificial intelligence and data science, Paavai College of Engineering, Nammakkal, India
[2] BE (CSE) Artificial intelligence and machine learning, Paavai College of Engineering, Nammakkal, India
Emails ID: [1] dhivyathrisha0113@gmail.com, [2] pavanpriya8825@gmail.com, [3] akilesh2555@gmail.com

***Abstract*— Once a futuristic concept confined to science fiction, traffic surveillance systems are now indispensable to modern urban planning and traffic management strategies. These systems, leveraging a complex interplay of sensors, communication networks, and sophisticated software, provide real-time data and analytical insights that enable authorities to monitor, understand, and optimize traffic flow. Traditional methods for traffic surveillance often struggle with accurately identifying traffic feature entities and objects. In most forums, Preliminary methods failed to identify the traffic feature entities' object presence based on traffic, leading to patterns of poor performance accuracy due to higher image degradation and false positives. So, the complexity increases to finding the traffic surveillance irregularities is difficult. To resolve this problem, we propose an advanced traffic surveillance system based on deep feature engineering with a Fuzzy Capsuled Convolution Neural Network (FC-CNN). The traffic videos are initially collected from a real-time traffic monitoring system to convert videos into frames. Then, preprocessing is carried out with an Adaptive Mean Wavelet Filter (AMWF) to normalize the image frames. Next, segmentation is carried out by Entity Pattern Watershed Segmentation (EPWS) to identify the object entities of traffic patterns objects entities based on feature recognitions. Initially, a fuzzy capsuled Convolution neural network is applied to find the irregularities of the helmet missing, triples flow, wrong route, and other anomalies. The proposed system produces a high traffic surveillance detection rate to perform best at a higher precision rate and improved detection accuracy in true favorable rates with redundant false rates. The proposed system archives improved accuracy compared to the other preliminary methods supporting traffic pattern rules. This results in superior performance compared to existing preliminary methods and supports the enforcement of traffic pattern rules with improved accuracy. This novel approach offers a robust and reliable solution for enhancing traffic surveillance and promoting safer road conditions.**

Index Terms*— Traffic Surveillance, Deep Feature Engineering, Fuzzy Capsule Convolutional Neural Network, Anomaly Detection, Image Preprocessing, Watershed Segmentation, Wavelet Filter, Traffic Irregularities.

I. INTRODUCTION

The escalating demands of modern urban life have placed unprecedented strain on transportation infrastructure, leading to increasingly complex and congested traffic patterns. Effective traffic surveillance is paramount for managing this complexity, enabling proactive intervention and informed decision-making that enhances safety, efficiency, and sustainability. Multiple Deep Learning (DL) techniques in traffic surveillance research enhance object detection and real-time monitoring while optimizing computational performance [1]. The researched traffic object detection through DL, highlighting Convolutional Neural Networks (CNNs) and their processing efficiency, but they also pointed out high costs.

The suggestion proposes a Lightweight Model with an Optimized Feature Extraction (LM-OFE) design to reduce processing overhead. The author faced challenges with accident detection through DL systems because of restricted video resolution and an obstructed field of view [2]. The Adaptive Occlusion Handling Model (AOHM) represents an adaptive model that enhances detection capability during congested traffic situations.

Preliminary methods, often reliant on simple feature extraction and pattern recognition techniques, frequently suffer from high false positive rates and poor accuracy, hindering their ability to reliably detect critical events such as helmet violations, illegal vehicle occupancy, or deviations from designated routes. The intricate nature of traffic patterns, coupled with the challenges of noisy and degraded image data, significantly increases the complexity of identifying traffic irregularities and necessitates the development of more sophisticated and robust surveillance solutions.

This paper introduces a novel and intelligent traffic surveillance system designed to overcome the limitations of existing approaches. Our proposed system leverages the power of deep feature engineering coupled with a Fuzzy Capsuled Convolutional Neural Network (FC-CNN). Furthermore, integrating fuzzy logic into the capsule network architecture enhances the system's robustness to uncertainty and noise, enabling it to handle degraded images and ambiguous situations more confidently.

Our approach begins with acquiring real-time traffic video streams from existing monitoring infrastructure. These videos are then meticulously processed to extract individual frames, forming the foundation for subsequent analysis. A crucial pre-processing step involves the application of an Adaptive Mean Wavelet Filter (AMWF). This sophisticated filtering technique effectively normalizes the image frames by suppressing noise and enhancing relevant image details, thereby mitigating the impact of image degradation and improving the quality of the data fed into the subsequent stages of the system. This normalization process ensures consistent and reliable feature extraction, even under challenging environmental conditions.

Following pre-processing, a critical segmentation step is performed using Entity Pattern Watershed Segmentation. This advanced segmentation algorithm is designed to accurately identify and delineate distinct object entities within the traffic scene, such as vehicles, pedestrians, and motorcycles. The algorithm effectively distinguishes individual objects and defines their boundaries by leveraging feature recognition techniques, providing a structured scene representation that facilitates subsequent analysis. The CNN layers effectively extract hierarchical features from the segmented image data, capturing intricate details and patterns within the traffic scene. These features are then fed into the capsule network, which learns to represent objects as vectors, or "capsules," that encode not only the presence of an object but also its pose, deformation, and other relevant attributes. Incorporating fuzzy logic into the capsule network allows the system to handle uncertainty. This is crucial for detecting anomalies such as missing helmets, triple riding on motorcycles, vehicles traveling in the wrong direction, and other unconventional traffic behaviors.

The proposed FC-CNN-based system's performance is rigorously evaluated against traditional methods, demonstrating its superior capabilities in detecting traffic irregularities. Our results show that the proposed system achieves significantly higher traffic surveillance detection rates, boasting superior precision. The ability to accurately identify anomalies while minimizing false alarms is crucial for ensuring the effectiveness and reliability of the traffic surveillance system. Furthermore, the proposed system demonstrates improved accuracy compared to preliminary methods, thereby supporting the enforcement of traffic pattern rules and contributing to a safer and more efficient transportation environment.

The objective and contribution present a comprehensive and innovative approach to intelligent traffic surveillance. By leveraging deep feature engineering with a Fuzzy Capsuled Convolutional Neural Network, our system offers a robust and accurate solution for identifying traffic anomalies, overcoming the limitations of traditional methods. The AMWF pre-processing step and Entity Pattern Watershed Segmentation further enhance the system's performance by improving image quality and accurately delineating traffic entities. Future work will focus on further optimizing the FC-CNN architecture, exploring the use of temporal information to improve anomaly detection accuracy, and integrating the system with existing traffic management infrastructure for real-world deployment and evaluation.

II. LITERATURE SURVEY

Traditional traffic surveillance systems often fall short in accurately identifying and classifying traffic entities and anomalies, primarily due to limitations in handling image degradation, variations in lighting conditions, and the sheer complexity of real-world traffic scenes. The author's application of traffic surveillance imagery at low altitudes [3] faced issues because adverse weather conditions degraded image clarity. Image augmentation techniques form part of the Weather-Adaptive Drone Imaging Model (WADIM), which counteracts adverse weather effects on drone imaging. The [4] demonstrated a DL hybrid model for real-time detection, yet its complexity led to processing time delays. The Real-time Efficient Hybrid Model (REHM) achieves optimum accuracy and computational speed.

The [5] applied the Gaussian Mixture Model with an ensemble of DL for vehicle detection, yet their approach failed due to background noise interference. The proposed Noise-Resistant Ensemble Model (NREM) improves background removal precision in the system. Acoustic-based vehicle identification through the [6] faced challenges because of environmental noise disturbances. The proposed Acoustic Noise Filtering Model (ANFM) enhances signal processing ability in environments where detection becomes challenging.

The author's Multi-Object Detection and Tracking (MODT) model showed tracking deviations as sudden object movements occurred [7]. A Dynamic Object Tracking Model (DOTM) utilizes a predictive tracking system that enhances system stability. The authors applied machine learning techniques for vehicle detection according to [8]. The big data sets caused scalability issues during the execution. The proposed Scalable Vehicle Recognition System (SVRS) implements hierarchical feature learning to enhance dataset adaptability.

The author in [9] encountered difficulties tracking similar objects in real-time operations. The Adaptive Similarity Differentiation Model (ASDM) builds superior representation systems that improve tracking precision. The system proposed in [10] detected abnormal items for smart cities yet encountered false positive problems in highly trafficked areas. The False Positive Reaction Model (FPRM) suggests improving anomaly detection systems to minimize errors.

The traffic surveillance framework built by the author [11] faced difficulties handling different traffic conditions. The Generalized Traffic Monitoring Model (GTMM) demonstrates modified capabilities for shifting road situations to achieve dependable results. The [12] managed numerous wrong detections of potential weapons while conducting weapon detection on CCTV footage because of indistinct shapes. Contextual scene analysis functions within the proposed Context-Aware Weapon Recognition Model (CAWRM) to reduce false positive detections.

According to the author, the developed vehicle detection and tracking system for highways experienced issues tracking multiple objects in high-traffic situations [13]. The proposed Multi-Target Tracking Enhancement Model (MTTEM) employs a trajectory prediction algorithm for better tracking. An image enhancement model for accident detection was presented in [14], yet it faced challenges because of slow real-time processing. The Real-Time Accident Detection Model (RTADM) uses edge computing technologies to derive faster response times.

The research established a traffic incident detection system built on computer vision but experienced difficulties operating under different lighting environments [15]. Different lighting scenarios benefit from the Adaptive Illumination Compensation Model (AICM) because it improves system performance. These enhancements improve deep learning-based traffic monitoring systems by solving their weaknesses.

III. PROPOSED METHODOLOGY

Traffic management and surveillance are critical components of modern urban infrastructure, impacting safety, efficiency, and overall quality of life. Traditional traffic surveillance systems often rely on manual monitoring or basic automated processes that struggle to adapt to the complexities and nuances of real-world traffic patterns. These preliminary methods frequently suffer from poor performance accuracy due to image degradation, false positives, and an inability to effectively identify and classify traffic feature entities. Consequently, detecting traffic irregularities and anomalies becomes a challenging task. The development and design creates an advanced, intelligent traffic surveillance system leveraging deep feature engineering and a Fuzzy Capsuled Convolutional Neural Network (FC-CNN) to address these limitations and enhance the accuracy and reliability of traffic monitoring.

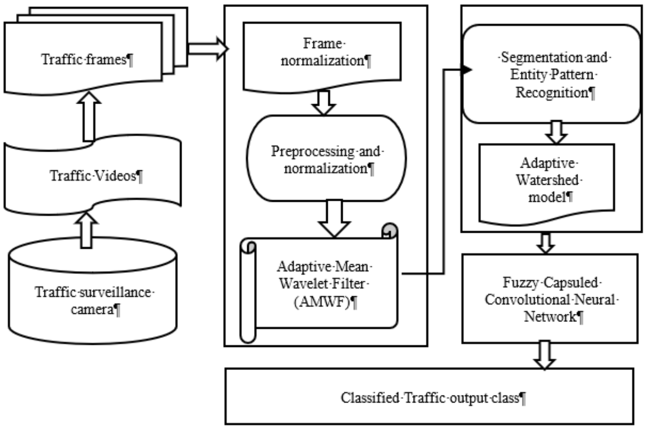

Fig. 1. Proposed architecture AWS-FCCNN

The core objective of this proposed system is to provide a robust and adaptive solution for identifying and responding to various traffic irregularities, including helmetless riders, triple-riding violations, wrong-way driving, and other anomalous events. Figure 1 explains the proposed architecture AWS-FCCNN. This is achieved through a multi-stage process encompassing data acquisition, preprocessing, segmentation, feature extraction, anomaly detection, and performance evaluation.

A. Data Acquisition and Pre-processing

The system begins with acquiring real-time traffic video feeds from strategically positioned surveillance cameras throughout the monitored area. These video feeds serve as the primary data source for the entire system. The videos are initially converted into individual frames to prepare the raw data for subsequent processing. This frame-by-frame representation allows for detailed analysis and the application of image processing techniques. This stage aims to enhance the quality of the image frames and mitigate the effects of noise, lighting variations, and other distortions that can negatively impact the performance of subsequent algorithms. Traditional image processing techniques often struggle to adapt to real-world traffic environments' diverse and dynamic conditions. Therefore, the proposed system employs an Adaptive Mean Wavelet Filter (AMWF). In video surveillance recognition, the Wavelet Filter Method is a vital preprocessing technique for images to enhance frames, improving object detection, face recognition, and license plate identification. Video frames decompose through Wavelet Transforms (WT) into several wavelet frequency components, thus reaching effective noise removal and edge sharpening alongside contrast optimization. The analysis of wavelet scales maintains vital image characteristics, including facial features, vehicle edges, and text, while filtering out disturbance from low-light conditions, image blur, and environmental elements. The preprocessing operation delivers better analytics accuracy through security systems, allowing improved detection of anomalies and more

IFERP
Explore Your Research Journey...
ISSN (Online) 2394-2320
International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 12, Issue 04, April 2025

accurate tracking and recognition.

$$W(a,b) = \int_{-\infty}^{\infty} f(t)\,\psi_{a,b}^{*}(t)dt \quad (1)$$

where,

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right) \quad (2)$$

Equations 1 and 2 divide video frames into multiple scales and positions to detect fine details and larger scene structures. The video frame pixel intensity function $f(t)$ operates with $\psi_{a \text{ and } b}$, representing the wavelet function that has undergone scaling and translation. Parameters a and b determine the level of detail alongside spatial positioning and object tracking in surveillance.

$$X(j,k) = \sum_m \sum_n I(m,n)\psi_{j,k}^{*}(m,n) \quad (3)$$

The efficiency of extracting structural features from digital video frames depends on the application of Equation 3. Through the wavelet function $\psi_{j,k}$, the input frame $I(m,n)$ allocates its components across different frequency bands. A surveillance system obtains better motion segmentation, face features, and vehicle tracking capabilities through the discrete wavelet transform (DWT), which leads to improved performance.

$$\hat{X}_{j,k} = \begin{cases} 0, & |X_{j,k}| \le T \\ X_{j,k} - T, & X_{j,k} > T \\ X_{j,k} + T, & X_{j,k} < -T \end{cases} \quad (4)$$

The noise reduction process described in Equation 4 maintains essential video elements, including faces, objects, and movement patterns, and it eliminates unnecessary noise. The threshold value T distinguishes between important visual details consisting of faces and text and unnecessary image artifacts such as low-light video noise. The accuracy levels of AI-based surveillance analytics increase due to this technique, which strengthens facial recognition capabilities, vehicle detection systems, and abnormal behavior monitoring. Through the Wavelet Filter Method, video surveillance recognition gains substantial improvement because it performs robust noise reduction, feature extraction, and contrast enhancement on frame quality. The Wavelet Filter Method is crucial when deployed in smart surveillance systems, law enforcement regions, and automated security monitoring environments.

AMWF is a sophisticated filtering technique that leverages wavelet transforms and adaptive mean filtering strengths. Adaptive mean filtering adjusts the filtering parameters based on the local characteristics of the image, ensuring that the filtering process does not blur or distort important details. Combining these two techniques, the AMWF effectively reduces noise while preserving critical image features, resulting in normalized image frames suitable for further analysis.

B. Segmentation and Entity Pattern Recognition

Accurate segmentation is essential for isolating individual objects and analyzing their characteristics. Traditional segmentation methods can be prone to errors, especially in complex traffic scenarios with overlapping objects, occlusions, and varying lighting conditions. To overcome these limitations, the proposed system utilizes Entity Pattern Watershed Segmentation. This technique combines the principles of watershed segmentation with entity pattern recognition. For video surveillance purposes, the Watershed method is a segmentation approach that detects objects and analyzes motion while separating backgrounds. The technique treats images as three-dimensional landscapes that use intensities to represent elevation heights. The algorithm floods the image from local minima until it reaches different regions where boundaries appear as watershed lines. Through watershed processing, overlapping objects get separated, edge detection improves, and feature extraction of surveillance data achieves better results. The accurate separation of background from foreground elements enables the Watershed method to serve essential roles in facial recognition, vehicle tracking, and suspicious activity detection.

$$G(x,y) = \sqrt{\left(\frac{\partial I}{\partial x}\right)^2 + \left(\frac{\partial I}{\partial y}\right)^2} \quad (5)$$

The $I(x,y)$ defines pixel intensity for brightness evaluation in the video frame. However, $\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y}$ calculates intensity differences to detect edges, while $G(x,y)$ calculates gradient magnitude for defining object boundaries during segmentation. The gradient magnitude of an image is calculated by Equation 5 through the assessment of intensity differences that occur in horizontal (x) and vertical (y) directions. The method assists in detecting major transitional areas between video objects and their background elements in surveillance footage analysis. When the gradient magnitude value is high, it indicates firm edges for segmentation boundary definition.

$$M(x,y) = \begin{cases} 1, & if\ (x,y)\ F \\ 0, & otherwise \end{cases} \quad (6)$$

The $M(x,y)$ discriminates between foreground object pixels (valued at 1) and background pixels (marked 0), while F denotes the objects of interest, including mobile humans and license plates. The segmentation process starts with Equation 7, which detects separate regions in video frame images. This procedure divides foreground elements (including humans, vehicles, and faces) from background areas into separate categories. The markers are essential guides for the flooding algorithm to carry out precise segmentation.

$$\frac{\partial R}{\partial t} = D\nabla^2 R - \mathrm{F}|\nabla R| \quad (7)$$

$R(x,y,t)$ refers to evolving regions that improve boundary definitions. In contrast, operator $D\nabla^2 R$ assists in eliminating excessive boundary fragmentation, and $\mathrm{F}|\nabla R||$ determines the speed at which segmentation expands to maintain precise object limits. The image region process of gradual filling is represented through Equation 7. The evolving area of segmentation appears as $R(x,y,t)$ through time. The speed of water propagation through $F|\nabla R|$ regulates how segmentation extends between objects, while $D\nabla^2 R$ supports the boundary-smoothing aspect.

$$W(x,y) = \begin{cases} 1, & if(x,y) \in L \\ 0, & otherwise \end{cases} \quad (8)$$

$W(x,y)$ determines watershed boundary pixels (1 value) and L to distinguish separate objects, preventing labeling errors between objects. $W(x,y)$ defines watershed lines through this mathematical expression that separates different identified regions. The $W(x,y)$ indication distinguishes watershed boundaries from other elements in video surveillance data, while L creates clear divisions between different objects within video surveillance frames. Video surveillance recognition depends heavily on the watershed method to perform facial recognition segmentation, motion tracking, and vehicle detection tasks.

Watershed segmentation is a region-growing algorithm that treats the image as a topographic surface, where pixel intensities represent elevation. The algorithm identifies "watersheds" or regional boundaries based on local intensity gradients. However, watershed segmentation can often lead to over-segmentation, resulting in numerous small regions. To address this issue, Entity Pattern Watershed Segmentation incorporates entity pattern recognition. This involves training the system to recognize specific patterns associated with different traffic entities. These patterns can be based on shape, size, texture, or other distinctive features. By integrating entity pattern recognition into the watershed segmentation process, the system can effectively suppress over-segmentation and identify the object entities of traffic patterns based on feature recognition with greater accuracy.

C. Fuzzy Capsuled Convolutional Neural Network (FC-CNN)

The core of the proposed system lies in its innovative application of deep feature engineering and a Fuzzy Capsuled Convolutional Neural Network (FC-CNN). This component is responsible for extracting relevant features from the segmented image regions and using these features to detect traffic irregularities. The proposed system further enhances the capabilities of capsule networks by incorporating fuzzy logic. Fuzzy logic allows the system to handle uncertainty and imprecision in the input data. The system can better represent the inherent ambiguity in real-world traffic scenarios by assigning fuzzy membership values to different features.

This procedure converges to a local minimum or saddle point of K_p.

Step 1: START -Established $W = [w_{xy}]$ Lattice, $W^{(0)}$

Step 2: At q -phase: computed objects $D^{(q)} = [d_y]\ through\ W^{(q)}$

$$D_y = \frac{\sum_{x=1}^{V} w_{xy}^{p} A_x}{\sum_{x=1}^{V} w_{xy}^{p}} \quad (8)$$

Step 3: Update $W^{(q)}, W^{(q+1)}$

$$w_{xy} = \frac{1}{\sum_{q=1}^{d}\left[\frac{\|a_x - d_j\|}{a_x - d_k}\right]^{\frac{2}{p-1}}} \quad (9)$$

If

$||\, W^{(q+1)} - W^{(q)}|| < \varepsilon$ then

End if

Otherwise

Return to step 2.

Step 4: Stop

FCM method has proven to be a powerful and efficient tool for image segmentation of Depends on Entity traffic pattern Recognition. On the other hand, Capsule Networks offer a more robust and expressive representation of visual information. Instead of simply predicting the probability of an object's presence, capsule networks predict a vector representing the object's pose, deformation, and other attributes. This allows capsule networks to capture the hierarchical relationships between object parts and to be more robust to variations in viewpoint and illumination. The deployed method is widely used for problems that demand accurate image understanding, including helmets, person patterns, autonomous vehicles, and object recognition tasks at scale. In equation 10, we perform the convolution operation,

$$F = \sum_{x=1}^{a} \sum_{y=1}^{b} X(x,y) \, . \, Q(x,y) + B \quad (10)$$

Let us assume F is the feature map, (x,y) is the position, X is the input image, Q is the kernel (filter), a and b are the kernel dimensions, and B is the bias term. Then we perform Residual Learning or ResNet Block through equation 11,

$$j = F(i, \{U_x\}) + i \quad (11)$$

Let us assume j is the residual block output, i is the input, U_x is the weights of the convolution layers in the residual block, and $F(i\{\ is\ the\ U_x\})$ as the learned mapping. This equation allows the network to learn residual functions instead of direct mappings. After performing the residual block, we perform the inception module through equation 12,

$$M = [f_1(i), f_3(i), f_5(i), f_p(i)] \quad (12)$$

Let us assume M is concatenated output from all branches, $f_p(i)$ is the max pooling operation, and $f_1(i), f_3(i), \square_{\square}(\square)$ and f5i have different scales. After we perform batch normalization D in equation 4 to stabilize training by normalizing activations,

$$D(i) = \frac{i-\mu}{\sqrt{\sigma^2 + \epsilon}} . \gamma + \beta \quad (13)$$

Let us assume μ as a mean of the batch and σ^2 as the variance for scaling and shifting. By following, we perform the activation (ReLU) function through equation 14,

$$f(i) = \max(0, i) \quad (14)$$

Here, we assume i for the input to the activation function; this equation introduces non-linearity. By following, we perform pooling layers; this layer contains two types of pooling layers; they are max and average pooling. In equation 6, we perform max pooling layer,

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Vol 12, Issue 04, April 2025

$$P_{max} = \max(i_x) \quad (15)$$

In equation 7, we perform the average pooling,

$$P_{avg} = \frac{1}{N}\sum_{x=1}^{N} i_x \quad (16)$$

These equations are used to reduce spatial dimensions. Then, we use Softmax for accurate classification in equation 17,

$$P(j_x) = \frac{e^{l_x}}{\sum_{y=1}^{K} e^{l_x}} \quad (17)$$

Let us assume x is class, l is logit, which is the output of the final layer, K is the number of classes, and $P(j_x)$ is the probability of the class. In this equation, we convert logits into probabilities. Then we perform loss function $\mathcal{L}$ through equation 18,

$$\mathcal{L} = -\sum_{x=1}^{K} j_x \log(\hat{j}_x) \quad (18)$$

Let us assume that the j_x true label in this 1 is assigned for the correct class, 0 is assigned for otherwise, and $\hat{j}_x$ is the predicted probability for the class. This equation determines the variances between the predicted and actual labels.

The Fuzzy Capsuled Convolutional Neural Network (FC-CNN) combines the strengths of CNNs, capsule networks, and fuzzy logic. It leverages CNNs for initial feature extraction and then passes these features to a capsule network, which learns to represent objects as vectors of attributes. Finally, fuzzy logic is used to refine the capsule outputs, considering the uncertainty and imprecision in the input data. The FC-CNN is trained on a large traffic video data dataset, labeled with various traffic irregularities, such as helmetless riders, triple-riding violations, wrong-way driving, and other anomalous features, into different categories of irregularities. Based on the training data, the fuzzy logic component is also trained to refine the capsule outputs. The system generates alerts and reports once the FC-CNN has identified potential traffic irregularities. These alerts can be transmitted to traffic management.

IV. PERFORMANCE EVALUATION AND OPTIMIZATION

The video surveillance images include detailed information about the detected irregularities present in video frames, such as the type of violation, the location of the incident, and the time of occurrence. The system can also be configured to trigger automated responses carried in a Python simulator, such as activating traffic signals, deploying law enforcement personnel, or issuing warning messages to drivers. This allows for a proactive and efficient response to traffic irregularities, improving safety and reducing congestion. The FC-CNN can be retrained on new data to improve its performance. The fuzzy logic component can also be adapted to account for changes in traffic patterns or environmental conditions.

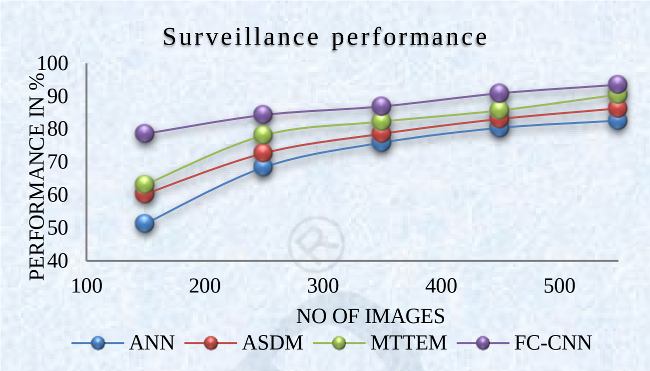

Fig. 2. Traffic surveillance detection

The proposed methods of Fc-CNN rendered an overall accuracy of 93.5% respectively. Figure 2 explains Traffic surveillance detection, and the chart also reveals that the proposed models' performance is significantly higher in the proposed models compared to the existing models for video traffic identification and classification. Regarding precision, recall, and F1-Score, the proposed models showed higher credibility and performance than the existing algorithmic methods. FC-CNN yielded 92.5%, 93.14% in sensitivity, 92.9%, 93.21% in specificity, and 93.1%, 94.35% for F1-Score, cumulating to the highest performance juxtaposed with the other models.

Table I: Analysis of Surveillance Detection Accuracy

Surveillance Detection accuracy in %				
Methods/ images	**ANN**	**ASDM**	**MTTEM**	**FC-CNN**
150	43.1	47.2	53.3	58.1
250	46.7	53.1	57.5	63.1
350	58.4	62.1	67.1	74.3
450	65.2	69.4	75.3	86.1
550	72.3	79.2	83.4	94.3

This section discusses the performance analysis of classification performance for traffic surveillance detection using different classification methods shown in Table 1. The overall performance of classification is validated using various parameters, such as precision, recall, and false classification performance.

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Vol 12, Issue 04, April 2025

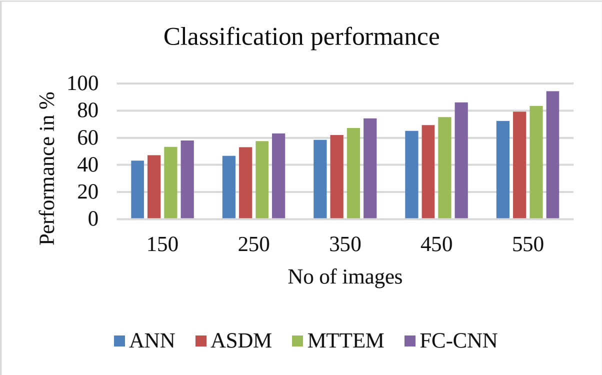

Fig. 3. Analyses of Accuracy performance

The FC-CNN's accuracy is higher than that of earlier techniques projected in Figure 3. As a result, our technique can swiftly identify flood photographs and has access to a vast data set. The FC-CNN is designed to achieve higher detection accuracy in true positive rates, reduce the number of false negatives, and ensure that more traffic irregularities are detected.

Table II: Comparative rate in precision

Analysis precision rate %				
Methods/ images	**ANN**	**ASDM**	**MTTEM**	**FC-CNN**
150	58.2	65.3	71.4	79.5
250	60.7	67.7	74.1	81.4
350	77.4	82.6	85.7	89.6
450	80.2	83.2	89.3	92.3
550	83.3	90.4	93.1	95.7

The FC-CNN system is expected to achieve higher precision rates in Table 2, meaning that a more significant proportion of the detected irregularities are true violations. The system is designed to adapt to changing traffic conditions and robust to variations in lighting, weather, and other environmental factors.

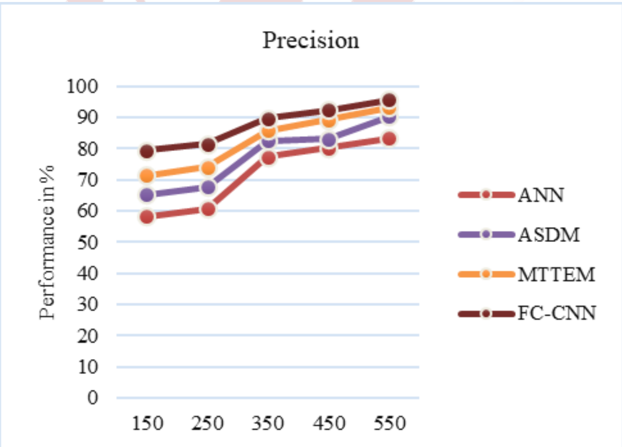

Fig. 4. Analysis of Precision Performance

In Figure 4, the FC-CNN system enables proactive traffic management with a higher precision rate by automatically detecting irregularities and triggering appropriate responses. By detecting and responding to traffic violations, the system can improve traffic safety and reduce accidents.

Table III: Analysis of Recall Performance

Performance in %				
Methods/ images	**ANN**	**ASDM**	**MTTEM**	**FC-CNN**
150	58.2	65.3	71.4	79.5
250	62.3	67.7	74.1	81.4
350	77.4	82.9	85.7	89.6
450	80.2	83.2	89.3	92.3
550	81.3	87.4	92.4	96.9

Table 3 shows the recall per romance of the proposed system, which is designed to adapt to changing traffic conditions and be robust to variations in traffic environmental factors. The system enables proactive traffic management by automatically detecting irregularities and triggering appropriate responses.

Fig. 5. Performance of Recall

Furthermore, Figure 5 shows that the proposed FC-CNN methods can be used to analyze a video surveillance dataset and assess the improvement in accuracy. Utilizing the offered FC-CNN technique for predictive precision analysis resulted in a significant accuracy enhancement of 93.46%.

Table IV: Comparative F1 performance

Performance in %				
Methods/ images	**ANN**	**ASDM**	**MTTEM**	**FC-CNN**
150	57.2	65.3	71.4	76.5
250	64.7	71.7	76.1	81.4
350	73.4	78.6	82.7	87.6
450	74.2	80.2	86.3	90.3
550	81.3	87.4	92.1	95.1

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Vol 12, Issue 04, April 2025

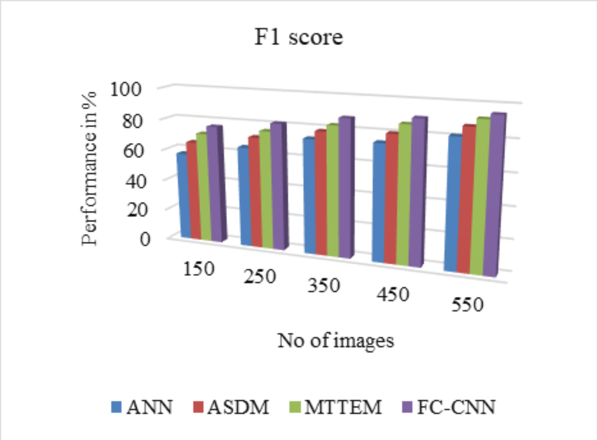

Fig. 6. Analysis of F1 Score Performance

As described in Figure 6, testing the features on the traffic video surveillance frames involves evaluating the false scores on subsets of their selected features to provide accurate prediction. Compared to other methods like ANN, ASDM, and MTEM from existing literature, the proposed method shows lower accuracy scores of 41.23%, 31%, and 28%, respectively. Similarly, employing the proposed FC-CNN method for accuracy prediction in error score analysis results in a 22.18% lower accuracy in emissions classification. The system aims to minimize the number of false positives, preventing unnecessary alerts and reducing the workload of traffic management personnel. The proposed system archives improved accuracy compared to the other preliminary methods supporting traffic pattern rules. The proposed method achieved less false classification performance than other methods.

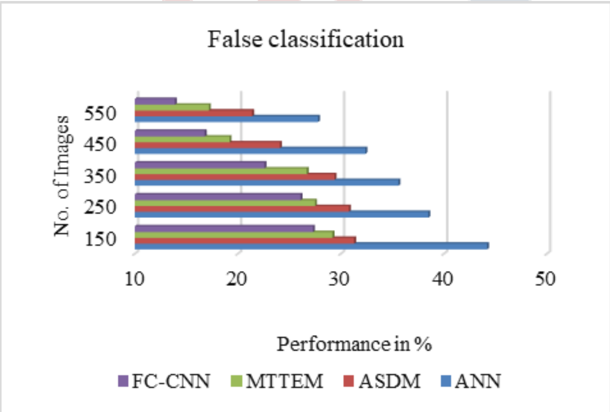

Fig. 7. Analysis of False classification

Figure 7 demonstrates that utilizing features from a traffic video frame dataset can offer accurate analysis for video surveillance utilizing selected feature subsets. Additionally, the suggested FC-CNN approach can select impotent traffic feature patterns from video surveillance features and enhance accuracy to 95.4%. Furthermore, compared to previous methods, the accuracy improves to ANN-91.39%, MTTEM 93.17%, and –ASDM 93.17%. The proposed smart AI-powered traffic surveillance system with deep feature engineering and a Fuzzy Capsuled Convolutional Neural Network is expected to offer significant advantages over traditional traffic monitoring methods.

V. CONCLUSION

The proposed smart AI-powered traffic surveillance system represents a significant advancement in traffic monitoring. By leveraging deep feature engineering and a Fuzzy-Capsuled Convolutional Neural Network, the system can accurately detect traffic irregularities, reduce false positives, and enable proactive traffic management. This advanced FC-CNN system has the potential to significantly improve traffic safety, reduce congestion, and enhance the overall quality of life in urban environments. Further research and development in this area will undoubtedly lead to even more sophisticated and effective traffic surveillance solutions in the future.

REFERENCES

[1] Azzedine Boukerche and Zhijun Hou. (2021). Object Detection Using Deep Learning Methods in Traffic Scenarios. ACM Comput. Surv. 54, 2, Article 30 (March 2022), 35 pages. https://doi.org/10.1145/3434398

[2] Pawar, K., & Attar, V. (2022). Deep learning-based detection and localization of road accidents from traffic surveillance videos. ICT Express, 8(3), 379-387. https://doi.org/10.1016/j.icte.2021.11.004

[3] Gupta, H., Verma, O.P. Monitoring and surveillance of urban road traffic using low altitude drone images: a deep learning approach. Multimed Tools Appl 81, 19683–19703 (2022). https://doi.org/10.1007/s11042-021-11146-x

[4] Li, L., Lin, Y., Du, B., Yang, F., & Ran, B. (2020). Real-time traffic incident detection based on a hybrid deep learning model. Transportmetrica A: Transport Science, 18(1), 78–98. https://doi.org/10.1080/23249935.2020.1813214

[5] Jagannathan, P., Rajkumar, S., Fonda, J., Divakarachari, P. B., & Subramani, P. (2020). Moving Vehicle Detection and Classification Using Gaussian Mixture Model and Ensemble Deep Learning Technique. Wireless Communications and Mobile Computing, 2021(1), 5590894. https://doi.org/10.1155/2021/5590894

[6] A. Shabbir, A. N. Cheema, I. Ullah, I. M. Almanjahie and F. Alshahrani, "Smart City Traffic Management: Acoustic-Based Vehicle Detection Using

Stacking-Based Ensemble Deep Learning Approach," in IEEE Access, vol. 12, pp. 35947–35956, 2024, doi: 10.1109/ACCESS.2024.3370867.

[7] Elhoseny, M. Multi-object Detection and Tracking (MODT) Machine Learning Model for Real-Time Video Surveillance Systems. Circuits Syst Signal Process 39, 611–630 (2020). https://doi.org/10.1007/s00034-019-01234-7

[8] M. Hasanvand, M. Nooshyar, E. Moharamkhani, and A. Selyari, "Machine Learning Methodology for Identifying Vehicles Using Image Processing", AIA, vol. 1, no. 3, pp. 154–162, Apr. 2023, doi: 10.47852/bonviewAIA3202833

[9] Jha, S., Seo, C., Yang, E. et al. Real time object detection and trackingsystem for video surveillance system. Multimed Tools Appl 80, 3981–3996 (2021). https://doi.org/10.1007/s11042-020-09749-x

[10] Ingle, P. Y., & Kim, Y. (2021). Real-Time Abnormal Object Detection for Video Surveillance in Smart Cities. Sensors, 22(10), 3862. https://doi.org/10.3390/s22103862

[11] Akhtar, M. J., Mahum, R., Butt, F. S., Amin, R., M., A., Lee, S. M., & Shaikh, S. (2021). A Robust Framework for Object Detection in a Traffic Surveillance System. Electronics, 11(21), 3425. https://doi.org/10.3390/electronics11213425

[12] M. T. Bhatti, M. G. Khan, M. Aslam and M. J. Fiaz, "Weapon Detection in Real-Time CCTV Videos Using Deep Learning," in IEEE Access, vol. 9, pp. 34366-34382, 2021, doi: 10.1109/ACCESS.2021.3059170.

[13] Azimjonov, J., & Özmen, A. (2021). A real-time vehicle detection and a novel vehicle tracking systems for estimating and monitoring traffic flow on highways. Advanced Engineering Informatics, 50, 101393. https://doi.org/10.1016/j.aei.2021.101393

[14] Pillai, M.S., Chaudhary, G., Khari, M. et al. Real-time image enhancement for an automatic automobile accident detection through CCTV using deep learning. Soft Comput 25, 11929–11940 (2021). https://doi.org/10.1007/s00500-021-05576-w

[15] Basheer Ahmed, M. I., Zaghdoud, R., Ahmed, M. S., Sendi, R., Alsharif, S., Alabdulkarim, J., Albin Saad, B. A., Alsabt, R., Rahman, A., & Krishnasamy, G. (2023). A Real-Time Computer Vision Based Approach to Detection and Classification of Traffic Incidents. Big Data and Cognitive Computing, 7(1), 22. https://doi.org/10.3390/bdcc701002