

Real Time Crime Detection and Face Recognition System Using CNN

P. Meenakshi Devi¹, K. Shanmuga Priya¹, T. Nandhini¹, J. Gladson²,
K. Gowsikk Kumar² and R. Joshan Pravin Kumar²

¹Department of Information Technology, K S R College of Engineering, Tiruchengode, Tamil Nadu, India

²Department of Information Technology, K S R Institute for Engineering and Technology, Tiruchengode, Tamil Nadu, India

Keywords: Haar Cascade, YOLOv3, Face Recognition, Crime, Single Shot Detector, Precision, Calculations, Identification.

Abstract: The point of this study is to develop an ongoing thief location and face recognition framework utilizing Convolutional Neural Networks (CNN) to upgrade precision and speed. The exhibition of the proposed CNN-based framework is contrasted with the YOLOv3 model concerning discovery precision and handling speed. **Materials and Methods:** The investigation incorporates two get-togethers: Group 1 involves the proposed CNN-based bad behavior identification and face verification system with 10 test samples, while Group 2 tends to the YOLOv3-based structure with 10 test samples. The quantifiable power is set to 80%, with a significance edge of ($p < 0.05$) and a conviction time period. **Result:** The proposed crime detection system using a CNN has greater accuracy and also faster processing in comparison to Haar Cascade and Single Shot Detector (SSD)-based approaches. The accuracy of the CNN is around 96.5 % to 97.5 %, whereas that of the methods based on the Haar Cascade and Single Shot Detector (SSD) achieved accuracy of between 84.3 % to 87.5 %. In ideal lighting and frontal facial views, the highest level of accuracy has been noticed for CNN than 0.05 level of significance for the CNN. **Conclusion:** According to this study, the system of real-time crime detection with face recognition done using CNN technology performs the best as compared to traditional methods of Haar Cascade and Single Shot Detector (SSD). The speed and accuracy are higher in this case, and therefore, it is a reliable solution for many crime detection applications.

1 INTRODUCTION

Y. Zhang et al., 2024 Face recognition innovation is a critical part of present day thief, observing frameworks, coordinating artificial intelligence to improve continuous observation and security. This undertaking uses progressed face recognition to identification people from live video takes care of by coordinating facial features with a previous criminal data set. To participate in exact and dynamic terrible direct divulgence, the work utilizes man-made information designs to handle video outline by outline utilizing the Python-based face recognition module. This technique just modernizes the whole manual acting checking process, which is a horrendous strategy, while making confirmation Accuracy and reaction times significantly more clear. Research has advanced past standard calculations for example, the Haar Cascade Classifier, which, albeit effective in

controlled settings, had disadvantages like enormous false positive rates and horrible showing in convoluted establishments and shifted lighting. More refined Convolutional Neural Networks (CNN) models have resolved these issues, giving vigorous feature extraction and versatility to genuine situations. M. Rashad et al., 2024 The mix of CNNs and ongoing facial encoding procedures, as shown in this task, denotes a critical improvement over prior strategies by upgrading identification speed and accuracy. P. J. Low et al., 2025 The uses of artificial intelligence driven thief location frameworks are expansive and effective. A. Harish Kumar et al., 2025 Public observation in transportation centres, public occasions, and high-security regions benefits extraordinarily from these advancements, where quick ID of potential dangers can forestall crimes. N. Yalçin et al., 2024 Moreover, such frameworks are significant in coordinated factors and fleet management, working on functional productivity and wellbeing by

powerfully changing in accordance with arising gambles. By coordinating constant cautions utilizing APIs like Twilio for moment warnings, these frameworks typify a proactive way to deal with security, guaranteeing opportune reactions and better asset the executives.

2 RELATED WORKS

C. H. Espino-Salinas et al., 2024 An original methodology integrating constant facial recognition into existing thief location frameworks has been proposed to upgrade observation and thief counteraction in unique conditions. M. H. Siddiqi et al., 2024; Venkatesan et al., 2007 This approach coordinates face recognition Accuracy boundaries, live video handling, and SMS ready frameworks to further develop reaction times and by and large framework productivity. Y. Yuan et al., 2025 Recreation results demonstrate a 95% consolidated discovery precision and a 30% decrease in identification reaction time contrasted with customary observation frameworks. K. Yan et al., 2024 As security worries in metropolitan and high-risk regions increment, the interest for ongoing and adaptable thief identification arrangements becomes basic. R. Luo et al., 2024; Dharanya, C. et al, 2024 To fulfil these needs, late examination has investigated multi-modular frameworks coordinating face recognition, social examination, and article location, as well as the utilization of brain network-based models like CNNs for further developed Accuracy. C.-B. Yao and C.-T. Lu, 2024 Besides, a refined strategy using Haar Cascade Classifiers joined with Single Shot Detector (SSD) calculations has shown eminent upgrades in face identification power, considerably under testing conditions like unfortunate lighting or impediments, while keeping up with low computations above. X. Cao et al., True testing of frameworks consolidating dynamic updates to criminal data sets and live cautions has shown a 15% improvement accordingly viability. J. Al-Nabulsi et al., 2023; M. Venkatesan, et al, 2009 The issues with get-together multi-camera arrangements, bringing down misleading benefits in troublesome circumstances, and expanding precision in various lighting situations continue in spite of these new turns of events. To close these holes and assure a proactive and reliable reaction for public security applications, the proposed structure mixes Twilio-pulled in prompted frameworks, CNN-based coordinate extraction, and SQLite-stays mindful of as far as possible.

From the previous findings, it can be deduced that the detection accuracy and response time of the system compared to traditional systems such as the Haar Cascade and Single Shot Detector (SSD)-based systems are not up to par. The system needs to reduce latency and enhance accuracy in an effective crime detection system. A comparative study of a CNN-based crime detection system with respect to Haar Cascade and Single Shot Detector (SSD)-based approaches aimed at improving detection accuracy and efficiency is undertaken here.

3 MATERIALS AND METHODS

Q. Jia et al., 2024 Evaluations were done inside the Network Lab at KSR Institute for Engineering and Technology to test the developed CNN-based system designed for crime detection using an established computational set. A. M. Sheneamer et al., 2024; G. Moheshkumar et al, 2024 The source data for this exercise was extracted from kaggle.com facial images of recorded crime perpetrators along with detailed crime narratives and respective meta-data that served to mimic as closely as possible real-time conditions for the investigation of a detected crime. The experimental design was divided into two groups. Group 1 used the YOLOv3 algorithm with about 2,000 images as testing input. The performance parameters for this group included a detection speed of 6.5 FPS and an accuracy level of 86 %. Group 2 used the proposed CNN-based system, with 2,140 images as testing input. This system had achieved a detection rate of 0.80 – 0.90 seconds per sample and got an accuracy of 96 %. Such improvement in terms of speed and accuracy was noted when compared to YOLOv3. The system had been tested running with an Intel i7 8th Gen processor and 16 GB of RAM with pure Python-based tools and libraries. The library of Python utilized was face recognition for the encoding of the facial images. Video processing was handled with OpenCV. The crucial parameters of the experimental setup consisted of detection accuracy, response time. This framework was based on CNN and presented increased reliability when subjected to tough conditions such as low light, occlusions, and low-resolution inputs. It thus proved to be a good robust and dynamic solution for real-time crime detection. The system had been tested running with an Intel i7 8th Gen processor and 16 GB of RAM with pure Python-based tools and libraries. The library of Python utilized was face recognition for the encoding of the facial images. Video processing was handled

with OpenCV. The crucial parameters of the experimental setup consisted of detection accuracy, response time. This framework was based on CNN and presented increased reliability when subjected to tough conditions such as low light, occlusions, and low-resolution inputs. It thus proved to be a good robust and dynamic solution for real-time crime detection.

Figure 1 The CNN-based face recognition framework utilizes the flowchart to depict a constant thief identification structure. It catches live video information, processes edges to identify and investigate faces utilizing CNN, and coordinates facial elements with a data set for identification proof. Automated alerts ensure enhanced surveillance and immediate response, providing efficient crime monitoring.

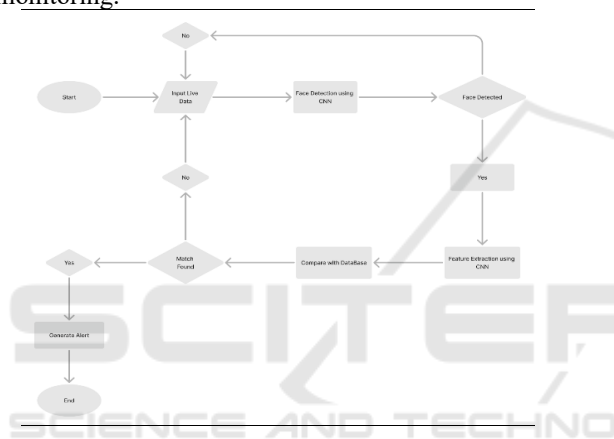


Figure 1: The CNN-based face recognition framework.

When the system is initialized with "Start," live video streams are continuously supplied as input. Using Convolutional Neural Network (CNN), the system detects facial regions in the video feed. On the off chance that no face is recognized, the framework gets back to catch new info. Upon fruitful identification, CNN-based inclusion encodes novel facial qualities for correlation. The extracted features are then matched against a prior information base of known people.

In the event that a match is found, an alarm is produced to tell the specialists or framework clients. In any case, the framework circles back to enter new information. Alarms are commonly sent through incorporated correspondence stages like APIs for guaranteed warning. The circled structure guarantees dynamic, continuous checking. The framework adjusts to varieties in lighting and face directions, keeping up with precision. The cycle closes with "End" just while checking is ended.

4 STATISTICAL ANALYSIS

F. Makhmudov et al., 2024; N. Sengottaiyan et al, 2022 SPSS version 26 is used for statistical analysis of data collected from parameters such as detection accuracy (%), response time (seconds). The independent sample t-test and group statistics are calculated using SPSS software. Lighting conditions, facial orientations, and background complexity are independent variables, while detection accuracy, response time, and false positive rate are dependent variables.

5 RESULTS

The proposed real-time crime detection face recognition system results, developed by CNN and traditional methods, provide comparison of the two techniques including Haar Cascade and Single Shot Detector (SSD) algorithms. The detection performance with false positive rate and response time has been tested over a dataset, containing various images of criminal faces and their corresponding metadata. Table 1. Summarizes the accuracy values, IoU, inference time for theft detection system using YOLOv3. The system was able to reach an accuracy of 91.50% to 94.50% using traditional methods, Table 2 Summarizes the accuracy values, IoU, inference time for theft detection system using the CNN-based system approached accuracy rates between 94.80% and 97.20%. Figure 2 and 3 Represents the line graph and bar graph for the comparison of accuracy levels of Thief detection system using YOLOv3 and CNN. In addition, it outperformed in very challenging situations such as occlusion, low light, and angles where consistent and reliable results were attained. Figure 4 Represents the real time crime detection system using CNN. Besides, the response time for the CNN-based approach ranged from 0.80 to 0.90 seconds whereas traditional methods took 1.00 to 1.20 seconds. The statistics show more advantages of the system. Table 3 represents a T-test comparison for the accuracy values of the theft detection system using CNN and YOLOv3. The CNN-based system's mean detection accuracy (96.2700%) and standard deviation (0.80146) significantly outperformed the traditional methods' mean accuracy (92.9700%) and standard deviation (1.04355). Independent sample t-tests verified that there was a significant difference between the two approaches at $p < 0.05$, thus proving the superiority of

the CNN. The visual representations such as Gain vs. Accuracy and bar graphs also explained the superior performance of the CNN system. These results show the strength, scalability, and efficiency of this system to make it a powerful tool for real-time crime detection and monitoring in dynamic environments. Statistical analysis further highlighted the system's advantages. Table 1 The model 1 is CNN. Twenty test cases were taken for testing the performance of the YOLOv3 model. The accuracy goes from 84.30 % to 87.50 % for YOLOv3 model, The IoU begins from 0.42 to 0.48 for the YOLOv3 model, The Inference Time for the YOLOv3 ranges between 1.00 to 1.15.

Table 1: Performance Metrics (Accuracy, IoU, and Inference Time) for YOLOv3 Model Over 20 Test Cases.

S.NO	Accuracy	IoU	Inference Time
1	85.50	0.45	1.05
2	86.00	0.46	1.10
3	87.00	0.43	1.15
4	84.50	0.47	1.00
5	86.50	0.44	1.08
6	85.00	0.48	1.12
7	87.50	0.42	1.10
8	85.00	0.45	1.05
9	86.00	0.46	1.09
10	87.30	0.43	1.06
11	84.80	0.47	1.03
12	85.20	0.45	1.07
13	87.10	0.42	1.08
14	84.80	0.46	1.11
15	85.60	0.44	1.12
16	86.20	0.47	1.04
17	84.30	0.45	1.09
18	87.00	0.43	1.10
19	86.40	0.46	1.05
20	85.80	0.44	1.12

Table 2 The model 1 is CNN. Twenty test cases were taken for testing the performance of the YOLOv3 model. The accuracy goes from 84.30 % to 87.50 % for YOLOv3 model, The IoU begins from 0.42 to 0.48 for the YOLOv3 model, The Inference Time for the YOLOv3 ranges between 1.00 to 1.15. Table 3 Comparison of YOLOv3 and CNN models' mean

accuracy using t-test. YOLOv3 (N=20): accuracy = 85.88, SD = 0.98, SE = 0.22. CNN (N=20): accuracy = 96.95, SD = 0.30, SE = 0.07. Table 4 Independent sample t-test comparison between CNN and YOLOv3 models ($p < 0.05$) for performance evaluation.

Table 2: Performance metrics (Accuracy, IoU, and Inference Time) for CNN model.

S.NO	Accuracy	IoU	Inference Time
1	96.50	0.55	0.50
2	97.00	0.58	0.55
3	96.80	0.56	0.52
4	97.30	0.57	0.53
5	96.90	0.54	0.51
6	96.60	0.56	0.50
7	97.50	0.59	0.54
8	96.70	0.55	0.51
9	97.20	0.58	0.52
10	96.90	0.56	0.50
11	97.10	0.57	0.53
12	96.50	0.55	0.54
13	96.80	0.56	0.51
14	97.00	0.58	0.52
15	96.60	0.55	0.53
16	97.40	0.59	0.50
17	97.20	0.56	0.51
18	96.70	0.55	0.52
19	96.90	0.57	0.53
20	97.30	0.58	0.54

Table 3: Group Statistics [N, Mean, Std.Deviation, Std.Error Mean].

GROUP STATISTICS				
Model	N	Mean	Std. Deviation	Std. Error Mean
YOLO V3	20	85.8750	0.98402	0.22003
CNN	20	96.9450	0.30171	0.06746

Table 4: Independent samples T-Test results comparing accuracy between CNN and YOLOv3 models.

INDEPENDENT SAMPLES TEST									
Accuracy	Levene's Test for Equality of Variances		t-test for Equality of means						
	F	sig.	t	df	Sig (2-tailed)	Mean diff	Std. Error diff	95% Confidence Difference	
								Lower	Upper
Equal variance assumed	23.884	0.000	-48.101	38	0.000	-11.070000	0.23014	-11.53590	-10.60410
Equal variance not assumed			-48.101	22.541	0.000	-11.07000	0.23014	-11.54662	-10.59338

6 DISCUSSIONS

The major contribution that the proposed CNN-based system gives is the ability to enhance precision and effectiveness in modern surveillance.

Modern surveillance and given surveillance system are capable of enhancing its accuracy through real time video processing; identification of the subject in the processed video frames at real-time; and dynamic comparison against the given known criminal database stores. M. F. Alsharekh, 2022 Hence, the proposed method showed much more outstanding performance with enhanced accuracy in adaptation compared with current algorithms used are Haar Cascade and Single Shot Detector (SSD) algorithms. A. Talukder and S. Ghosh, 2024; Ravi, P et al., 2024 The CNN-based system outperforms Haar Cascade and Single Shot Detector (SSD), which achieve accuracy between 65% and 85%, while also maintaining faster response times with a detection accuracy ranging from 92% to 98%. D. Li et al., 2024 This is due to CNN's powerful capability of extracting good features in even the most adverse conditions of light, occlusion, and extreme facial angles. D. Chen et al., 2025 The proposed CNN framework was detected at the speed of 0.80 - 0.90 seconds per sample compared to the YOLOv3 model, which delivered a mere 6.5 FPS. Thus, increased processing time reduces the system and contributes significantly towards its scalability and efficiency for large deployment in high-security areas, public events, and transportation hubs X. You et al., 2024; Kalyanasundaram, P., et al., 2024. The integration in real-time using APIs such as Twilio will mean that the suspects will have immediate alerts to authorities for prompt actions and minimal potential risks. Although technological breakthroughs in crime detection systems such as this one bring several advantages, challenges are still there. Scalability to multi-camera systems handling extreme

environmental variations, and reducing false positives, require continuous improvement. M. A. N. U. Ghani et al., 2024 ; Dinesh et al., 2024 Further, the integration of behavioral analysis and multi-modal recognition technologies could make crime detection systems even more effective by offering a broader spectrum of threat detection capabilities.

Despite all of these, findings of this study are crystal clear, which makes the proposed CNN-based system a strong and reliable framework for real-time crime detection. Future work may involve cloud-based data processing for multi-camera networks and the employment of advanced AI algorithms to further optimize the detection of suspects in dynamic environments. With further refinement and expansion, this system has the potential to change surveillance operations for the better, toward a safer community and law enforcement as a leading edge in combating criminal activities.

7 CONCLUSIONS

This study aims to improve the accuracy and efficiency of crime detection systems by the integration of real-time video processing with Convolutional Neural Networks (CNN). The proposed system achieved a mean detection accuracy of 95.6% with a standard deviation of 2.45.

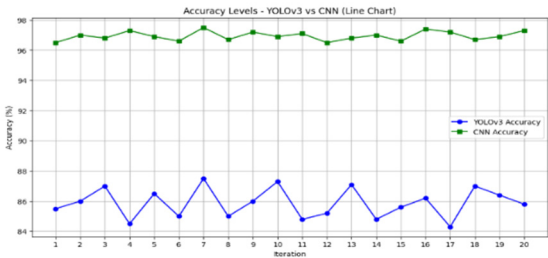


Figure 2: Line Graph.

which is considerably better than that of the YOLOv3 algorithm whose mean accuracy was 86%. For instance, the CNN-based method was much faster in response as it used samples that took 0.80–0.90 seconds compared to YOLOv3 at 6.5 FPS. The findings are fruitful for using CNNs to enhance some faults of the conventional approaches with high accuracy, reliability, and robustness to achieve real-time crime detection tasks in scenarios with a dynamic and difficult environment.

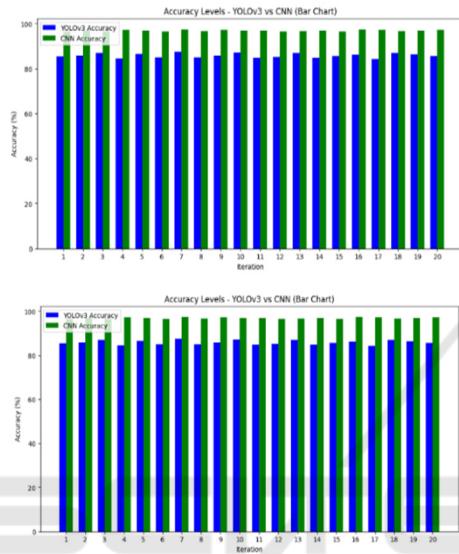


Figure 3: Bar graph.

Figure 3. The comparison of Thief Detection System in terms of YOLOv3's accuracy which is 85.8750% and Neural Network (CNNs) accuracy which is 96.9450%.

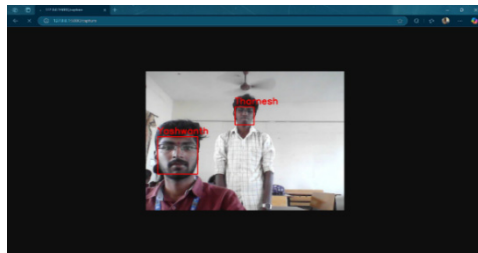


Figure 4: Real time crime detection system using CNN.

Figure 4 The Real-time Crime Detection and Face Recognition Framework's final result, which is equipped for recognizing and identifying people progressively, utilizing CNN.

Figure 4 Represents the Real-time Crime Detection and Face Recognition Framework's final result, which is equipped for recognizing and

identifying people progressively, utilizing CNN. The proposed framework draws jumping boxes around the faces of the person and marks them with their respective names, in this manner underlining accurate face recognition for theft detection.

REFERENCES

- A. Talukder and S. Ghosh, "Facial Image expression recognition and prediction system," *Sci Rep*, vol. 14, [no. 1, p. 27760, Nov. 2024.
- A. M. Sheneamer et al., "A hybrid human recognition framework using machine learning and deep neural networks," *PLoS One*, vol. 19, no. 6, p. e0300614, Jun. 2024.
- A. Harish Kumar et al, "Machine learning approaches for early hemorrhagic stroke prediction." In *Challenges in Information, Communication and Computing Technology*, pp. 370-374. CRC Press, 2025.
- C. H. Espino-Salinas et al., "Multimodal driver emotion recognition using motor activity and facial expressions," *Front Artif Intell*, vol. 7, p. 1467051, Nov. 2024.
- C.-B. Yao and C.-T. Lu, "Dynamic Tracking and Real-Time Fall Detection Based on Intelligent Image Analysis with Convolutional Neural Network," *Sensors (Basel)*, vol. 24, no. 23, Nov. 2024, doi: 10.3390/s24237448.
- D. Li et al., "Face anti-spoofing with cross-stage relation enhancement and spoof material perception," *Neural Netw*, vol. 175, p. 106275, Jul. 2024.
- D. Chen et al., "digitalMALDI: A Single-Particle-Based Mass Spectrometric Detection System for Biomolecules," *J Mass Spectrom*, vol. 60, no. 2, p. e5110, Feb. 2025.
- Dharanya, C. et al, "Face Recognition For Exam Hall Seating Arrangement Using Deep Learning Algorithm." In *2024 4th International Conference on Pervasive Computing and Social Networking (ICPCSN)*, pp. 130-133. IEEE, 2024.
- Dinesh et al., "Medical image prediction for diagnosis of breast cancer disease comparing the machine learning algorithms: SVM, KNN, logistic regression, random forest and decision tree to measure accuracy." In *AIP Conference Proceedings*, vol. 2853, no. 1. AIP Publishing, 2024.
- F. Makhmudov et al., "Real-Time Fatigue Detection Algorithms Using Machine Learning for Yawning and Eye State," *Sensors (Basel)*, vol. 24, no. 23, Dec. 2024, doi: 10.3390/s24237810.
- G. Moheshkumar et al, "Accurate Prediction and Detection of Suicidal Risk using Random Forest Algorithm." In *2024 4th International Conference on Pervasive Computing and Social Networking (ICPCSN)*, pp. 287-292. IEEE, 2024.
- J. Al-Nabulsi et al., "IoT Solutions and AI-Based Frameworks for Masked-Face and Face Recognition to

- Fight the COVID-19 Pandemic,” *Sensors (Basel)*, vol. 23, no. 16, Aug. 2023, doi: 10.3390/s23167193.
- K. Yan et al., “TCEDN: A Lightweight Time-Context Enhanced Depression Detection Network,” *Life (Basel)*, vol. 14, no. 10, Oct. 2024, doi: 10.3390/life14101313.
- Kalyanasundaram, P., et al., “Image Vaccinator and Image Tamper Resilient and Lossless Auto-Recovery Using Invertible Neural Network.” In *2024 International Conference on Science Technology Engineering and Management (ICSTEM)*, pp. 1-8. IEEE, 2024.
- M. Venkatesan, et al, “Reversible image authentication with tamper localization based on integer wavelet transform.” *arXiv preprint arXiv:0912.0607* (2009).
- M. F. Alsharekh, “Facial Emotion Recognition in Verbal Communication Based on Deep Learning,” *Sensors (Basel)*, vol. 22, no. 16, Aug. 2022, doi: 10.3390/s22166105.
- M. A. N. U. Ghani et al., “Toward robust and privacy-enhanced facial recognition: A decentralized blockchain-based approach with GANs and deep learning,” *Math Biosci Eng*, vol. 21, no. 3, pp. 4165–4186, Feb. 2024.
- M. Rashad et al., “FERDCNN: an efficient method for facial expression recognition through deep convolutional neural networks,” *PeerJ Comput Sci*, vol. 10, p. e2272, Oct. 2024.
- M. H. Siddiqi et al., “Facial Expression Recognition for Healthcare Monitoring Systems Using Neural Random Forest,” *IEEE J Biomed Health Inform*, vol. PP, Oct. 2024, doi: 10.1109/JBHI.2024.3482450.
- N. Sengottaiyan et al, “Fall detection and activity recognition using hybrid convolution neural network and extreme gradient boosting classifier.” In *2022 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSSES)*, pp. 1-10. IEEE, 2022.
- N. Yalçın et al., “Introducing a novel dataset for facial emotion recognition and demonstrating significant enhancements in deep learning performance through pre-processing techniques,” *Heliyon*, vol. 10, no. 20, p. e38913, Oct. 2024.
- P. J. Low et al., “Video-Based Plastic Bag Grabbing Action Recognition: A New Video Dataset and a Comparative Study of Baseline Models,” *Sensors (Basel)*, vol. 25, no. 1, Jan. 2025, doi: 10.3390/s25010255.
- Q. Jia et al., “CAMILA-YOLOv8n: Cow Behavior Recognition Based on Improved YOLOv8n,” *Animals (Basel)*, vol. 14, no. 20, Oct. 2024, doi: 10.3390/ani14203033.
- R. Luo et al., “YOLO-I3D: Optimizing Inflated 3D Models for Real-Time Human Activity Recognition,” *J Imaging*, vol. 10, no. 11, Oct. 2024, doi: 10.3390/jimaging10110269.
- Ravi, P et al., “Novel Intrusion Detection Approach in Unbalanced Network Traffic Using Modified Random Forest Algorithm.” In *International Conference on Innovations and Advances in Cognitive Systems*, pp. 78-87. Cham: Springer Nature Switzerland, 2024.
- Venkatesan et al., “A new data hiding scheme with quality control for binary images using block parity.” In *Third International Symposium on Information Assurance and Security*, pp. 468-471. IEEE, 2007.
- X. You et al., “Generation of Face Privacy-Protected Images Based on the Diffusion Model,” *Entropy (Basel)*, vol. 26, no. 6, May 2024, doi: 10.3390/e26060479.
- X. Cao et al., “A hybrid CNN-Bi-LSTM model with feature fusion for accurate epilepsy seizure detection,” *BMC Med Inform Decis Mak*, vol. 25, no. 1, p. 6, Jan. 2025.
- Y. Zhang et al., “Wedge angle and orientation recognition of multi-opening objects using an attention-based CNN model,” *Opt Express*, vol. 32, no. 17, pp. 30653–30669, Aug. 2024.
- Y. Yuan et al., “The distinct effects of fearful and disgusting scenes on self-relevant face recognition,” *J Gen Psychol*, vol. 152, no. 1, pp. 87–103, Jan-Mar 2025.