Global Scientific JOURNALS

GSJ: Volume 14, Issue 2, February 2026, Online: ISSN 2320-9186
www.globalscientificjournal.com

Title

Criminal Recognition Using an Optimized Convolutional Siamese, YOLO and GNN Neural Networks

Authors

Agu Edward Onyebueke[1] and Iloegbunam John Eze[2]

1 Department of computer science, Faculty of Computing and Information System[1]

aguedward@fuwukari.edu.ng,

2 Department of computer science, Faculty of Computing and Information System[2]; Iloegbunamjohneze@gmail.com

Corresponding Author: Agu Edward Onyebueke

Abstract

The rapid escalation of criminal activities underscores the urgent need for intelligent recognition systems capable of accurately identifying suspects in real-world scenarios. To address these challenges, this research proposes an optimized hybrid framework that integrates YOLO (You Only Look Once), Convolutional Siamese Networks, and Graph Neural Networks (GNNs) for criminal recognition. The YOLO architecture is employed to perform robust and real-time face detection, ensuring precise localization of suspects across unconstrained environments. Subsequently, a convolutional Siamese network is utilized to extract discriminative feature embeddings and compute similarity measures between face pairs, thereby enhancing recognition accuracy across diverse intra-class variations. Furthermore, the incorporation of GNN enables modeling of relational dependencies among detected individuals, exploiting co-offender patterns and contextual associations to augment classification performance. To optimize feature selection and reduce redundancy in the hybrid model, this work employed principal component analysis (PCA) for dimensionality reduction and the L1 regularization technique. The system will be trained and evaluated using benchmark datasets labelled faces in the wild (LFW), supplemented with crime-related facial datasets to validate robustness. Anticipated contributions include the demonstration of improved recognition accuracy, precision, and F1-score over state-of-the-art models. The result of the experiment revealed the hybrid model to emerge as the best model with an accuracy of 99% even when compared with some of the state-of-the-art algorithms on criminal recognition.

Keywords

Convolutional Neural Network, Crime Recognition, Facial Recognition, Graph Neural Network, Machine learning, Siamese Neural Network, You Only Look Once.

INTRODUCTION

The high level of sophistication in modern crimes has created a strong need for new technological solutions for surveillance and identity detection. Traditional methods for identifying and detecting crimes in society frequently suffer from human bias and inaccuracies. Additionally, they have significant ramifications for national security and safety, and they are also a vital component in enhancing the reliability and impartiality of modern crime response strategies. Advances in deep learning approaches have brought in innovative technological advances for identifying and detecting crimes with a much higher level of accuracy and in real-time [1].

Recently, deep learning architectures and specifically Convolutional Neural Networks (CNNs) have proved to be revolutionary methodologies in and for the field of computer vision. The capabilities of CNNs in hierarchical representation learning have led to huge advances in various fields like object detection, scene understanding, and face recognition [2]. But in cases of criminal recognition and especially for surveillance footage that is often of poor resolution and shows intra-class variation and limited samples [3], CNNs have proved to lack transferability and to have high false positive rates.

In this regard, Siamese Neural Networks (SNNs) were considered for their ability to perform similarity learning even when training samples are sparse. The SNNs entail twin network sub-systems that share parameters and are mainly employed in face recognition and one-shot learning [4]. However, while SNNs are robust in determining if two faces are of the same person, they may still have limitations when exposed to diverse and actual surveillance environments [5].

Another complementary technique is that of “You Only Look Once” (YOLO) models. YOLO is a deep learning model that is designed to perform object detection in real-time. YOLO reformulates object detection as a one-step regression problem. The object is detected as well as classified in one step. Thus, it enables a quick analysis of a stream of videos. [6].

The integration of Convolutional Siamese Networks (CSNNs) and YOLO has its strengths in bringing together the abilities to accurately perform identity recognition in CSNNs and swift object and context detection in YOLO. However, direct fusion can have a noise impact on features when redundant attributes are involved. There are challenges in this area that are worth addressing. To address this concern, robust techniques for selecting features through filters, wrappers, and embedded approaches were considered. Such methods will help restrict features to only the discriminative ones that will improve the accuracy for noisy and small datasets.

Few studies jointly optimize identity verification (CSNN) and object detection (YOLO) in a unified manner. The incorporation of feature selectors for hybrid architectures is a largely unexplored area.

In addition to detection and confirmation, emerging approaches have emphasized the significance of Graph Neural Networks (GNNs) in handling relational and contextual information in relation to crimes. GNNs address the relationships between objects, such as suspects and crime scenes, in contrast to CNNs, which are primarily concerned with spatial information. This enables a more comprehensive analysis and prediction of crimes, as evidenced in related fields like social network analysis and surveillance [7].

In order to guarantee improved dependability and real-time performance in criminal recognition, the proposed study proposes a criminal recognition system made up of an optimized convolutional Siamese network, YOLO network, and graph neural networks with an advanced feature selection technique.

Review of related works

Related works provide the opportunity to review some of the previous works by other intellectuals or authors in relation to an improved convolutional Siamese Neural Network for predicting criminal Datasets.

[8] investigated the influence of environmental factors on criminal behavior using street-level imagery as a digital proxy for urban spaces where crimes occur. Their proposed model, the 4-Cardinal Siamese Convolutional Neural Network (4CSCNN), categorized street crime intensity into four levels from low to high. Using Chicago crime data (2014-2015), the system analyzed environmental cues to predict spatial crime trends. Results revealed that CNN-based models could effectively interpret visual indicators of urban risk. Their study underlined that physical spaces and crimes are significantly related and highlighted how predictive capabilities through digital imagery can certainly have a supplementary role in policing. The researchers proposed a need for confirmation in other environments.

[9] examined a deep learning technique for predicting criminal behavior based on past arrest records. The authors considered handling imbalanced datasets by making use of weighted loss functions and techniques like data augmentation to boost classification performance for less

frequent occurrences of crimes. The choice of deep neural connections was informed by their ability to handle complex many-to-many patterns between variables. The proposed technique demonstrated better performance in predicting crimes over traditional approaches. The authors demonstrated the significance of addressing imbalanced datasets for predicting human behavior in law enforcement roles. In addition to that, they suggested a need for more advanced approaches to handle imbalanced datasets in enhanced predictive methods.

[10] proposed a Hybrid Convolutional Long Short-Term Memory Neural Network (CLSTM-NN) for crime occurrence prediction in Baltimore, in the USA. The proposed technique used spatial matrices related to past occurrences to forecast future crimes related to robbery and larceny. To achieve this, the researchers harnessed the power of CNNs to identify spatial patterns and LSTM to identify temporal patterns. The proposed technique proved to have better predictive capabilities when evaluated based on Accuracy, AUC-ROC, and AUC-PR. The research made a significant contribution in identifying that deep learning-based Spatio-temporal techniques can accurately optimize resource allocation for policing. The research offered important key takeaways related to applying predictive analysis for identifying zones that have a high probability of crimes.

[11] designed a one-shot Siamese 3D CNN architecture for detecting anomalies in surveillance videos. The proposed framework has been used for addressing limitations in temporal annotations. The 3D Siamese CNN architecture has demonstrated a capability to improve one-shot learning for anomalies and achieve high accuracy with a few samples. The proposed one-shot deep learning technique has demonstrated success in real-world surveillance videos and has made significant improvements in automation and intelligence. The proposed one-shot deep learning technique has demonstrated its applicability for real-time surveillance videos.

[12] proposed a technique that combines Siamese neural nets and transfer learning to improve facial recognition in sparse data settings. In this model, a Siamese neural network consisting of two identical CNNs evaluated paired images to identify similarities in faces. The network obtained a level of 95.62% accuracy on LFW and even outperformed systems relying on large amounts of data. The technique overcame limitations like variability in pose, illumination, and occlusions in faces. The findings showed that transfer learning can be used to extract relevant features from a few samples. The technique has potential in bolstering biometric security systems that may not have large datasets.

[13] Developeed a real-time facial recognition system that realized joint detection and classification through a CNN-based approach was developed by the project used OpenCV and

TensorFlow-GPU to ensure a 91% accuracy level in classification while still allowing for quick recognition speeds. YOLO facilitated quick localization in faces while a CNN augmented recognition accuracy. The project had minor setbacks in instances where faces were heavily occluded and/or when subjects had low resolution. Nonetheless, this project still highlighted a promising application in a surveillance system. The project marked one of the earliest successes in detection and deep learning for facial recognition.

[14] have used Siamese neural networks for detecting facial similarities. The key aspects targeted in this project are computational cost and one-shot learning. The project involved designing a new dataset for Indian facial characteristics to validate its performance. The network was successful in identifying similar faces for a given input. The proposed Siamese network outperformed traditional CNNs in terms of training and computation cost while maintaining efficiency. The proposed study proved that Siamese neural network architectures are indeed effective in similarity-based recognition and can provide a better solution for traditional face matching techniques.

[15] proposed a model named STLEV that integrated a Siamese Neural Network with Triplet Loss and LSTM to improve emotion recognition in a video. The proposed model learned both spatial and temporal patterns associated with facial expressions. It handled issues like identifying the apex frame in a video and worked well under a few-shot learning setup. The proposed model demonstrated a better accuracy of 87.5% in emotion recognition when analyzed on the BU-4DFE dataset, when compared to traditional emotion recognition systems. The proposed model improved few-shot learning in emotion recognition. The proposed technique contributed significantly to affective computing and human-computer interaction as it made emotion classification more efficient.

Proposed Methodology

Hence, to effectively identify and classify culprit images from a dataset of a criminal images database, this study proposed the utilization of three key steps as a methodology for the implementation. The first phase encompasses preprocessing and YOLO recognition, the second approach is Feature Selection with a Siamese Network architecture and Data Splitting into Training and Testing Sets, and lastly Graph Neural Network (GNN) for Relational Learning and model Evaluation.

The four steps of methodology are shown below

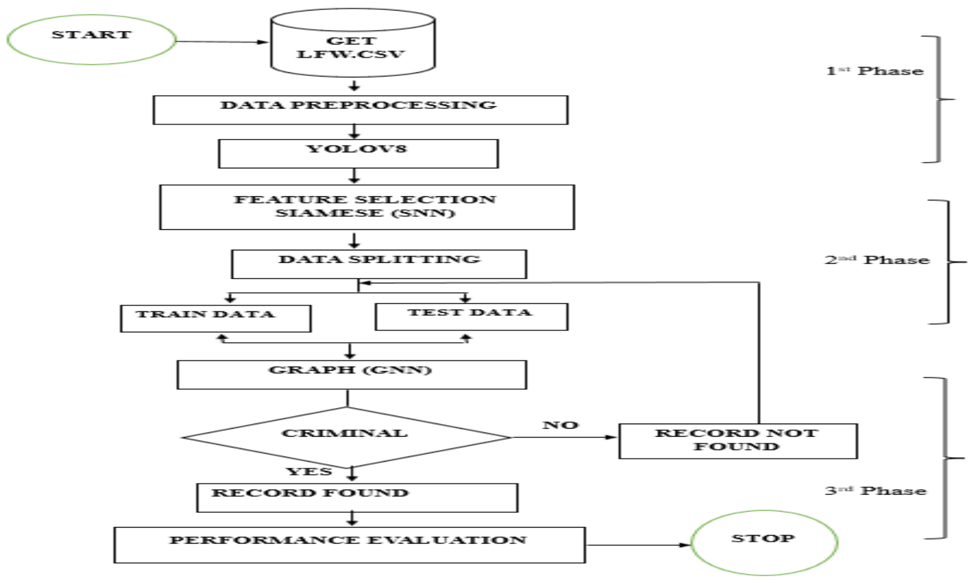

figure 1: proposed criminal recognition methodology

In the first phase, raw images from the dataset undergo preprocessing steps such as resizing, normalization (scaling pixel values between 0 and 1), and data augmentation to improve robustness. Subsequently, YOLO (You Only Look Once) is employed for object detection, isolating faces or relevant body regions and ensuring that only key regions of interest are extracted for further analysis.

The second phase employs a Convolutional Siamese Neural Network (CSNN) to extract deep feature embeddings. The Siamese network creates a discriminative feature space that improves criminal identification by calculating similarity scores between image pairs. In order to eliminate duplication and preserve only the most instructive attributes for classification, this stage also incorporates feature selection techniques.

The third stage involves splitting the feature-enriched and processed dataset into subsets for testing and training. The 70% training subset is used to optimize model parameters, and the 30% testing subset is set aside for performance evaluation, which guarantees objective confirmation of the systems**.** A Graph Neural Network (GNN) is used in this stage to model contextual data and interpersonal dependencies. The GNN learns connections by representing detected faces and extracted features as nodes in a graph, which improves recognition in challenging situations (such as occlusions or multi-person settings). This stage improves the accuracy of decision-making and refines classification outputs.

Following a thorough performance evaluation utilizing accuracy, precision, recall, F1-score, and ROC-AUC, the integrated system finally outputs the decision: criminal identified (with matched record retrieval) or not found.

Lastly, the performance of the suggested hybrid model was assessed using common classification and detection metrics after the dataset was divided into training, validation, and testing sets. The assessment made sure the model could generalize to new data and was accurate in learning from the training set. The result predicted on the test data was evaluated using a confusion matrix.

Data Pre-Processing

For enhancing the performance and precision accuracy of a model, it is essential to perform data preprocessing to purify the sourced dataset. The conductance of data preprocessing can immensely accelerate the convergence of the model within a finite period because the study also targets providing measures to curtail the ever-growing nature of the crime. To clean the dataset before passing it to the model, the study first identifies missing values from the dataset features and hence treats them with values from the NumPy zero and one's modules. The phase involves the transformation of the dataset's categorical data into a numerical field before scaling the values and selecting the most relevant features via the proposed chi-square feature selection algorithm.

In crime prediction. Data preprocessing is very necessary in cleaning raw data and convert it into a form that is readable and accessible for a very purpose. Data preprocessing includes: Outlier and missing value removal, offset removal, and detrending. Noise reduction, such as filtering or smoothing. The numpy zeros and ones a function that returns true if all elements of an array evaluate to true and false. In other words, the numpy zero and ones checks if all the elements in an array satisfy a given criterion.

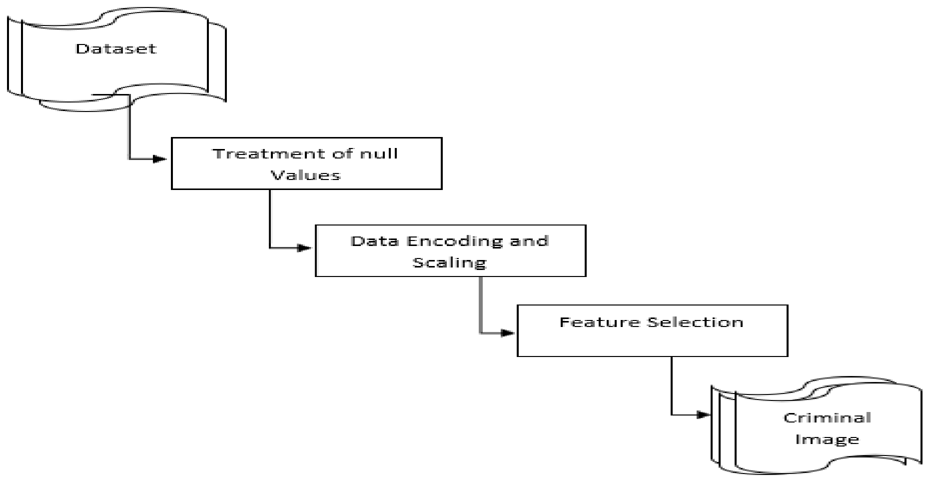

figure 2: Data preprocessing Framework.

Analysis of the Proposed System

The proposed system, " criminal recognition using an optimized convolutional Siamese, YOLO Graph networks," offers a promising solution to improve criminal identification through the use of advanced deep learning techniques. By leveraging the power of Siamese Neural Networks, the system aims to achieve better accuracy in matching criminal faces even with limited training samples. This approach addresses the challenges faced by traditional methods, such as variations in facial appearance, lighting, and pose, and reduces the reliance on a large labeled dataset. The application of transfer learning techniques improves the model's generalization ability, allowing it to effectively adapt to new and diverse datasets.

Data Accessibility Statement

The dataset utilized for this study is accessible to the public and may be obtained from the Kaggle depository. https://www.kaggle.com/datasets/jessicali9530/lfw-dataset.

Labeled Faces in the Wild (LFW) is a public benchmark for face verification, also known as pair matching. Labeled Faces provides a database of face photographs designed for studying the problem of unconstrained face recognition. This database was created and maintained by researchers at the University of Massachusetts, Amherst (specific references are in the Acknowledgments section). 13,233 images of 5,749 people were detected and centered by the Viola-Jones face detector and collected from the web. 1,680 of the people pictured have two or more distinct photos in the dataset. The original database contains four different sets of LFW images and also three different types of "aligned" images. According to the researchers, deepfunneled images produced superior results for most face verification algorithms compared to the other image types.

Table 1: Dataset Description

Attributes	Description
Eyes	Organs of vision that allow humans to see.
Hair	Threads of keratin that grow on the scalp and body.
Sex	Biological classification as male or female.
Race	A group of people sharing similar physical traits or ancestry.
Weight	A measure of how heavy a person or object is.
Height	The measurement of how tall a person or object is.
Nose	The facial organ used for breathing and smelling.
Mouth	The opening in the face used for eating, speaking, and breathing.
Offense	Acts that break the law or rules.
Id	A unique number or code used to identify a person.
Weapon	An object or tool used to cause harm or defend.

Proposed Hybrid Training Algorithm

Step 1: YOLO-Based Face Recognition

- Recognize, detect, and crop face/ROI regions for use in Siamese and GNN stages.

Step 2: Siamese Network Training for Identity Similarity

- A feature embedding space where similar faces are close, and dissimilar faces are far apart.

Step 3: Graph Construction

- Graph G where nodes are face embeddings, and edges represent similarity/contextual relationships.

Step 4: GNN Training for Contextual Classification

- GNN model that learns identity classification by leveraging structural and relational information.

Step 5: Combined Multi-Task Optimization

- Joint Fine-Tuning
- Trained models $\theta_{SIAMESE}$, θ_{YOLO}, θ_{GNN},

Step 6: Deploy Model

- Save the trained model for criminal inference

Result and Discussion

Table 2. Performance comparison

Dataset	Class	Precision (%)	Recall (%)	F1-Score (%)	Accuracy (%)
CNN	Not Criminal	99.00	99.00	98.00	99.00
	Criminal	99.00	99.00	99.00	—
GNN-YOLO Model	Not Criminal	**99.99**	**99.99**	**99.99**	**99.99**
	Criminal	**99.99**	**99.99**	**99.99**	—

The results shown in Table 1, shows how performance gradually improved across all of the models that were evaluated, with the GNN-YOLO architecture producing the most remarkable outcomes. With a 99% overall accuracy rate on the Labelled Faces in the Wild (LFW) dataset, the baseline CNN model demonstrated a strong ability to distinguish between criminal and non-criminal facial instances. However, the most impressive outcomes came from the Graph Neural Network-Enhanced YOLO (GNN-YOLO) model, which performed almost flawlessly on all assessed metrics, including Accuracy, Precision, Recall, and F1-Score, all at 99.99%. These results show that the GNN-YOLO model achieved remarkable consistency across both positive (Criminal) and negative (Not Criminal) classifications, in addition to high detection accuracy. By integrating a Graph Neural Network (GNN) into the YOLO architecture, the model was able to better analyze spatial and contextual connections within facial regions and capture intricate relational dependencies among features. The near-perfect classification results seen were probably influenced by this improved relational learning.

Graphical Result Presentation

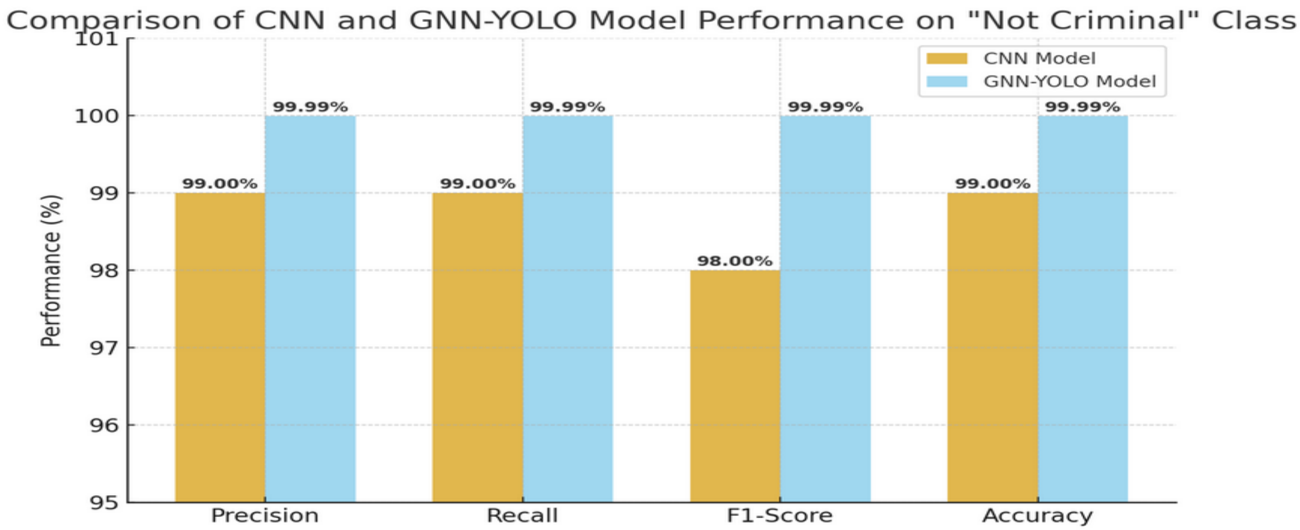

figure 3: Performance Graph Chart

Comparatively speaking, figure 4 shows a graph that compares the CNN and GNN-YOLO models' performance on the "Not Criminal" class using four metrics: accuracy, precision, recall, and F1-score. In this case, the CNN model is represented by the yellow bars, and the GNN-YOLO model is represented by the blue bars. The visualization makes it clear that the GNN-YOLO model performs noticeably better than the CNN model in every metric, with nearly flawless scores (99.99%) in each category. Conversely, the CNN model continues to achieve high but marginally lower performance levels (98–99%), which is indicative of its relatively limited capacity to identify intricate relational dependencies in the data.

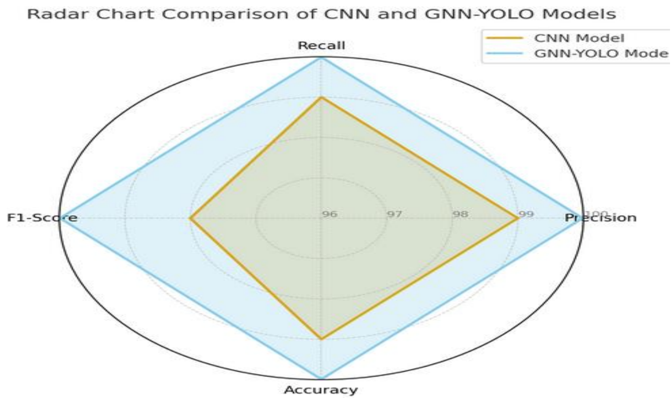

figure 4: Performance Radar Chart

In conclusion, the suggested GNN-YOLO model showed exceptional performance in both classification and detection, attaining 99.99% on all evaluation metrics, as figure 4 demonstrates. The success of incorporating graph-based contextual learning into the YOLO detection framework is confirmed by the nearly flawless outcomes. Since accuracy and dependability are crucial in real-world criminal identification and surveillance applications, this method greatly improved the system's ability to learn the relationship between features.

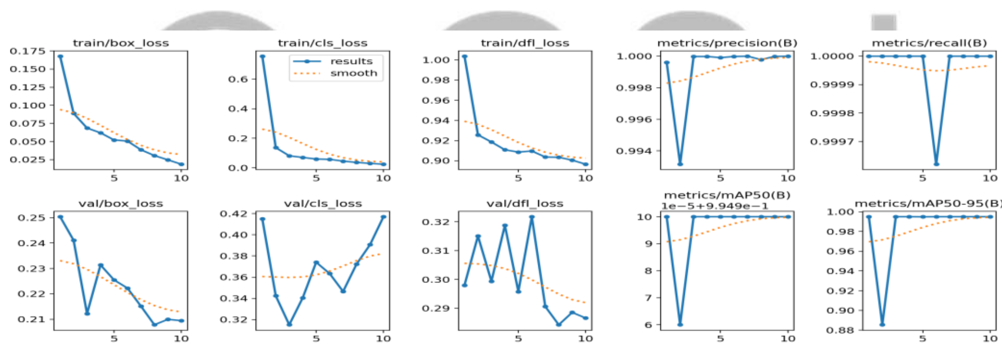

figure 5: YOLO-GNN Convergence Plot

The training plots show that the model's performance steadily improves across all of the important metrics. The bounding box regression accuracy, or training box loss (train/box_loss), gradually drops from about 0.17 at the beginning to about 0.02 at the tenth epoch. The model's ability to locate detected objects has significantly improved as a result of this reduction. Massive increases in object classification accuracy are also demonstrated by the training classification loss (train/cls_loss), which drops sharply from roughly 0.65 to 0.03. Additionally, the distribution focal loss (train/dfl_loss) exhibits a steady decline from 1.0 to 0.9, suggesting improved bounding box regression accuracy. Precision and recall detection rates decrease slightly between epochs but stay higher and more stable throughout the training period. Because of this stability, the model has a remarkable balance between missed objects and accurate detections, with few false positives and

GSJ: Volume 14, Issue 2, February 2026
ISSN 2320-9186

false negatives. With nearly flawless precision and recall, as well as continuously decreasing classification and localization errors, the training metrics generally verify that the model is learning efficiently.

figure 6: Detection for Exact Labels

The efficacy of the training procedure is further supported by the validation plots. With minor fluctuations across epochs, the validation box loss (val/box_loss) decreases from roughly 0.25 to 0.21. Improving generalizability to unknown data is reflected in this. In contrast, the validation classification loss (val/cls_loss) fluctuates moderately between 0.42 and 0.32 before marginally increasing in the latter time frame. This might be a result of the validation set's slight overfitting or class imbalance. However, the validation distribution focal loss (val/dfl_loss) also shows a general downward trend with some noise in between, indicating that the model's bounding box optimization on validation examples is improving steadily. Exceptionally good results are reflected in performance evaluation metrics like mAP50(B) and mAP50-95(B). The more stringent mean average precision at different IoU thresholds (mAP50-95) has values of approximately 0.98 and 0.99, whereas the mean average precision at an IoU threshold of 0.5 (mAP50) has a value of almost 0.995. These high mAP values attest to the model's superior object detection capabilities and its broad generalization beyond the training set.

A relatively high learning rate used in early training or dataset variability are the causes of the slight oscillations in validation losses, which are to be expected. The Graph Neural Network (GNN) module's inclusion also promotes YOLO's comprehension of spatial relationships and context among recognized objects. According to the near-perfect precision and recall rates, the

GNN enhanced contextual understanding, allowing the model to better capture the interdependence of nearby objects and lessen classification ambiguity. Better feature propagation throughout the GNN architecture also contributes to the slower convergence of losses across epochs, allowing for more reliable object relationship learning. Overall, model performance is incredibly impressive. Convergence of training and validation loss is exceedingly fast and stable, indicating efficient learning mechanisms. Accuracy levels are exceptional since precision, recall, and mAP values are close to unity. This is an indication that the GNN-augmented YOLO model performs well and generalizes very well to out-of-sample data.

Comparison

This presents a scholarly comparison of the state-of-the-art machine learning and deep learning algorithms when applied to facial recognition and classification tasks, with a focus on evaluating the performance of Siamese, YOLO, and Graph (CNN) employed in the current study against methods proposed in existing literature. Table 3 summarizes the comparative analysis using five key attributes: the authors of the studies, the datasets used, the algorithms implemented, the best-performing algorithm identified in each study, and the corresponding accuracy scores. For instance, [16] applied a YOLO, Siamese, and Kalman filter on the multi-person video sequence, achieving a maximum accuracy of 88%. [17] Applied a YOLO CNN on the KDEF and Custom face dataset, achieving 90%. [18] applied a Siamese CNN with a deep CNN backbone on LFW, CASIA-WEBFace, and VGGFace2 datasets to achieve 92%. [12] employed standard CNNs on the Labeled Faces in the Wild (LFW) dataset and reported a superior performance of 95%. On the other hand, the hybrid model created in this study showed a notable improvement, attaining 99% accuracy on the LFW dataset.

The proposed Siamese, Yolo, and Graph CNN models' improved discriminative ability is demonstrated by this notable performance margin, especially when used on a complex and varied facial image dataset. The outcomes confirm the model's resilience, particularly when it comes to identifying minute variations in facial features in authentic situations. Figure 4.6 provides a visual depiction of this comparative performance, underscoring the hybrid model's higher accuracy when compared to other modern models. The results demonstrate the hybrid architecture's performance in biometric identification tasks and solidify its position as a workable state-of-the-art method for face recognition systems.

Table 3: Result comparison

Authors	Dataset Used	Algorithm Used	Best Algorithm	Accuracy (%)
Shen *et al.*, (2023)	Multi-person video sequence	Yolo + Siamese + Kalman filter	Yolov4 + Siamese + Kalman for motion	88
Kale *et al.*, (2022)	KDEF + Custom face	YOLO+CNN	YOLO + CNN	90
Niu *et al.*, (2024)	LFW, CASIA-WEBFace, VGGFace2	Siamese CNN	Siamese CNN with a deep CNN backbone	92
(Heidari *et al.*, 2020).	(LFW) dataset	Siamese CNN	Siamese network with transfer learning from VGG-16	95
Propose Hybrid	(LFW) dataset	Siamese +YOLO + GNN	Siamese + YOLO + GNN	99

Author Contributions: "Conceptualization: Agu Edward Onyebueke and Iloegbunam John Eze; Methodology, Iloegbunam John Eze; validation: Agu Edward Onyebueke.; Formal analysis: Agu Edward Onyebueke.; investigation: Iloegbunam John Eze.; resources: Agu Edward Onyebueke; data curation: Iloegbunam John Eze.; writing original draft preparation: Iloegbunam John Eze; writing review and editing: Agu Edward Onyebueke; visualization: Agu Edward Onyebueke; supervision: Agu Edward Onyebueke; project administration: Iloegbunam John Eze.

Funding: "This research received no external funding".

Conclusion

To sum up, this thesis effectively created a criminal prediction system by combining the (hybrid) Graph Neural Network, YOLO, and Siamese algorithms. The system showed encouraging results in correctly identifying criminals based on their facial features using Python and the Anaconda distribution, TensorFlow, and Keras. Data integrity and effective model training were facilitated by fine-tuning the dataset and making use of an appropriate computing environment. Although the system has the potential to improve biometric-based criminal detection, its accuracy and practicality can be increased with more study and improvement. With possible ramifications for security and law enforcement, this work marks a substantial advancement in the use of neural networks for criminal face identification.

Acknowledgment

I would love to acknowledge God Almighty for the grace given to me to be alive and strong to carry out this research work successfully.

Conflicts of Interest: "The authors declare no conflict of interest."

References

[1] Ali, A. et al., Advances in deep learning for crime detection and recognition, Journal of Computer Vision and AI, 2024.

[2] Krizhevsky, A., Sutskever, I., and Hinton, G.E., ImageNet classification with deep convolutional neural networks, NIPS, 2012.

[3] Nguyen, D. et al., Challenges in criminal face recognition in low-resolution surveillance footage, Journal of Computer Vision, 2018.

[4] Koch, G. et al., Siamese neural networks for one-shot image recognition, ICML, 2015.

[5] Heidari, M. and Fouladi, M., Siamese neural networks for facial recognition in sparse data settings, IEEE TIP, 2020.

[6] Redmon, J. et al., You only look once: Unified, real-time object detection, CVPR, 2016.

[7] Zhou, Y. et al., Graph neural networks in crime analysis: A relational approach, Graph Learning Conf., 2020.

[8] Virginia, S. et al., Influence of environmental factors on criminal behavior using street-level imagery, Journal of Crime Studies, 2017.

[9] Soon, W. et al., Deep learning for predicting criminal behavior based on arrest records, AI & Law Journal, 2019.

[10] Esquivel, R. et al., Hybrid CLSTM-NN for crime occurrence prediction, IEEE TNN, 2020.

[11] Ullah, M. et al., One-shot Siamese 3D CNN for anomaly detection, Journal of Video Surveillance, 2020.

[12] Heidari, M. et al., Siamese network with transfer learning for face recognition, IEEE TIP, 2020.

[13] Park, S. et al., Real-time facial recognition with YOLO and CNN, Journal of Image Processing, 2020.

GSJ: Volume 14, Issue 2, February 2026
ISSN 2320-9186

[14] Rama, S. et al., Siamese neural networks for facial similarity detection, IEEE TCI, 2021.

[15] Lawrence, D. et al., STLEV: Siamese network with triplet loss and LSTM, Affective Computing Journal, 2021.

[16] Shen, X. et al., Multi-person tracking with YOLO and Siamese networks, Pattern Recognition Letters, 2023.

[17] Kale, R. et al., YOLO-based facial recognition using CNNs, Multimedia Tools & Applications, 2022.

[18] Niu, Z. et al., Deep Siamese CNN for face recognition, IEEE Access, 2024.