# DeiT-OEL: Data Efficient Image Transformer with Ontology-Based Evidence Linking for Intelligent Crime Scene Reconstruction

*Pramodhini.* R[1*], *Urvashi* Sharma Mishra[2], and *Ashok* Kumar Pamidi Venkata[3]

[1]Department of Electronics and Communication Engineering, Nitte Meenakshi Institute of Technology, Nitte (Deemed to be University), Bengaluru, India
[2]Department of Computer Science, Hans Raj Mahila Maha Vidyalaya, (Managed by DAVCMC), Jalandhar, India
[3]Megan Soft INC, Livonia, United States

**Abstract.** Intelligent crime scene reconstruction is critical in forensic informatics to provide an accurate analysis that supports criminal investigations and legal proceedings. This task becomes more challenging because of the complexity of visual crime evidence, including varying lighting conditions, occlusions, and scene clutter. However, traditional models struggle to maintain accuracy and effectively link evidence at complex crime scenes. To overcome this problem, this paper proposes a Data-efficient Image Transformer with Ontology-based Evidence Linking (DeiT- OEL) method for crime scene image classification. To improve evidence linking and reasoning, ontology-based evidence linking using a geographic query language for RDF data (GeoSPARQL) is integrated to connect the spatial and semantic relationships between crime scene elements. The proposed DeiT-OEL method is evaluated using the UCF Crime and Chicago Crime datasets, where preprocessing includes image normalization, resizing, and augmentation to improve the data quality. The experimental results show that the proposed DeiT-OEL method achieves a classification accuracy of 99.99%, significantly outperforming existing traditional models.

## 1 Introduction

Crime scene investigation aims to gather and preserve evidence from the scene of a crime to connect it to the perpetrator [1]. Forensic science is based on cases and depends on scientific knowledge, investigative methods, and logical reasoning [2]. In complex crime scene investigations, an interdisciplinary and multiagency approach is often required, with concise forensic strategies in the early stages to incorporate and define the operational parameters of various forensic disciplines in a clearly communicated plan [3]. 3D scanning and modeling in forensic applications have been widely documented, including methods like Capturing geometric data in crime scene reconstruction is performed using terrestrial laser scanning, structured blue-light scanning, and photogrammetry, RTC investigations, and shooting reconstructions [4]. The objective of this paper is to investigate how UAV technology contributes to evidence detection at outdoor crime scenes [5]. The forensic science field is under scrutiny from both legal and scientific circles, which question the dependability of certain forensic methods, particularly those relying mainly on visual assessments and human interpretation, including pattern evidence examinations like friction ridge analysis, firearms and toolmarks, footwear, tire tracks, and handwriting comparisons [6]. Crime scene investigations involve complex environments that demand a crime scene investigator's (CSI's) advanced technical skills and analytical thinking to uncover and interpret unclear evidence [7]. Recently, crime scene investigators have been increasingly trained across multiple disciplines by policing agencies [8]. Several investigations have pointed out a shortfall in EMT training concerning crime scene management alongside the provision of emergency medical care [9]. Blood samples collected from violent crime scenes and toxicological analyses of individuals can reveal details about the cause and manner of death by identifying and measuring consumed drugs [10]. Ahmed Sedik et.al [11] developed an enhanced AR system for holographic platforms such as Microsoft HoloLens. The system employed two main approaches: image classification using lightweight Convolutional Neural Networks and ConvLSTM and image segmentation based on the fuzzy active contour model (FACM). This methodology was implemented on a bloodstain pattern dataset to evaluate its effectiveness in AR applications. This approach introduced improved detection and visualization of bloodstain patterns, thereby supporting forensic investigations. The system offers advantages in immersive training and analysis; however, its limitations include potential hardware constraints, variability in real-world scenarios, and dependency on data quality for accurate performance. Sarfaraz Natha

* Corresponding author: pramodhini.r@nmit.ac.in

et.al [12] developed Composite Recurrent Bi-Attention (CRBA) model is used for anomaly detection in surveillance videos. It utilizes DenseNet201 to extract spatial features and BiLSTM networks to model temporal dependencies. A multi-attention mechanism helps focus on important spatiotemporal regions, enhancing the differentiation between normal and abnormal behaviors. The methodology introduced enhanced detection and classification capabilities evaluated on the UCF and Road Anomaly Dataset (RAD). This approach offers advantages in resource efficiency and rapid response to public safety, but its limitations include high computational requirements and challenges under adverse weather conditions, indicating areas for future improvement. Amjad Rehman et.al [13] developed automated human activity recognition systems, which have become a major research focus in recent years, particularly for video surveillance applications in indoor and outdoor environments. These systems employed sensors such as cameras, radars, and mobile devices to detect abnormal human behavior. Deep-learning architectures, including CNN and LSTM, were implemented to automatically extract features from video frames and classify activities. YOLOv4 and 3D CNN models have been introduced for real-time activity detection using centralized IoT frameworks. These methods offer scalable monitoring and rapid anomaly detection; however, their limitations include high computational demands and reliance on multiple camera feeds. Abdulrahman Alshalawi et.al [14] developed A hybrid model combining fire modules and time-distributed convolutional layers for spatial feature extraction, along with a swin transformer for self-attention to capture local details and long-range dependencies for training and testing, the model used UCF Crime and CUHK Avenue datasets, which include realistic anomalous events. This approach offers improved detection of anomalies with high accuracy in binary classification and multiclass tasks, thereby enhancing surveillance capabilities. However, the performance slightly decreased for multiclass classification, highlighting the need for further research to improve the real-time detection across diverse anomaly types. Liu Yang et.al [15] developed a model that integrates reinforcement learning (RL) with a scope loss function, where the RL mechanism prioritizes correct predictions for minority class instances. The scope loss function ensures a balance between exploiting known data and exploring new data, while maintaining accuracy and generalizability. The model employed convolutional neural networks (CNNs) for feature extraction from textual content and utilized artificial bee colony (ABC) optimization to refine hyperparameters. The datasets included Chicago Police Department crime reports and crime-related tweets. This approach offers enhanced predictive accuracy for crime-related activities. However, performance is influenced by evolving social media language and data sparsity.

The key contribution of the paper:

1. Proposes the DeiT-OEL Method for Crime Scene Image Classification, combining the Data-efficient Image Transformer (DeiT) with ontology-based evidence linking through GeoSPARQL to connect spatial and semantic relationships between crime scene elements.
2. Integrates Deep Learning with Ontology-based Reasoning, enabling enhanced evidence linking and semantic understanding of crime scene features, improving classification accuracy and supporting better interpretation of spatial patterns and crime context.
3. Achieves Superior Classification Performance on the UCF Crime and Chicago Crime datasets, significantly outperforming traditional CNN models in terms of both precision and efficiency.
4. Validates the Method through Rigorous Evaluation, showing that the DeiT-OEL method excels in crime scene image classification, leveraging deep learning and semantic reasoning for improved crime scene analysis and evidence detection.

### 1.1 Problem Statement

Crime scene image classification and evidence linking for intelligent reconstruction are challenging because of the complex visual evidence, varying lighting conditions, occlusions, and scene clutter. Traditional forensic methods and existing deep learning models struggle to effectively link spatial and semantic relationships between crime-scene elements, rely heavily on manual interpretation, and often require large datasets for training. These limitations reduce the accuracy, interpretability, and computational efficiency of automated forensic systems, thereby hindering their real-time applicability in criminal investigations.

### 1.2 Objective

To address these challenges, this paper proposes the DeiT-OEL method, which integrates a Data-efficient Image Transformer (DeiT) with Ontology-based Evidence Linking (GeoSPARQL). This method efficiently extracts global image features from crime-scene images using DeiT and enhances semantic and spatial reasoning by mapping evidence relationships through GeoSPARQL. This method reduces computational complexity and improves classification F1-score, recall, precision, and accuracy. The proposed method is evaluated on the UCF Crime and Chicago Crime datasets, achieving superior performance while enabling real-time, reliable, and explainable forensic decision-making.

This paper is structured as follows: Section II describes the proposed method for crime scene image classification. Section III presents the results and Section IV provides a detailed discussion of the experimental findings. Finally, Section V concludes the paper.

## 2 Proposed methodology

This paper proposes the DeiT-OEL method, which integrates a Data-efficient Image Transformer (DeiT) with Ontology-based Evidence Linking (GeoSPARQL) for crime-scene image classification. The architecture first preprocesses images from the UCF and Chicago crime datasets through normalization, resizing, and augmentation to enhance data quality. Visual features are then extracted using DeiT, where a Teacher-Student model strategy guides the student model through a combination of KLDivLoss and Cross-Entropy Loss (CELoss), improving the classification accuracy and training efficiency. To enhance evidence reasoning, GeoSPARQL is employed to map semantic and spatial relationships between crime-scene elements using a virtual knowledge graph linked to PostgreSQL + PostGIS storage. User SPARQL queries are translated into SQL for dynamic knowledge retrieval, thereby supporting spatial-semantic reasoning. Experimental evaluation demonstrated that the proposed DeiT-OEL method achieves 99.99% and 99.95% accuracy on the UCF Crime and Chicago Crime datasets, respectively, as shown in Figure 1, the model significantly outperforms existing models.

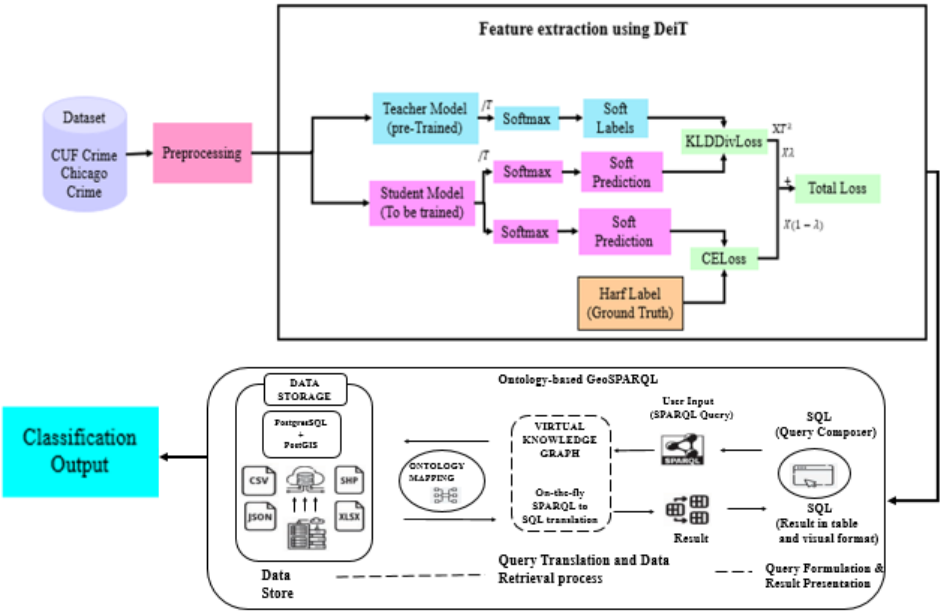


**Fig. 1.** Overall architecture of the DeiT-OEL method.

## 2.1 Dataset Description

The UCF-Crime dataset [16] is a large-scale video dataset for anomaly detection, containing 1,900 surveillance videos across 13 crime categories and one normal class, averaging 80–100 videos over 128 hours. Videos are captured at roughly 30 fps and usually have a resolution near 320×240 pixels. Sourced from real-world CCTV footage, the dataset includes challenges like lighting changes, crowded environments, and complex backgrounds, making it valuable for developing effective DL-based crime detection models. The Chicago Crime [17] dataset contains information about crimes reported in the City of Chicago from 2001 to the present, except for the last seven days. Data were obtained from the Chicago Police Department system (CLEAR). To protect privacy, exact addresses are not shown; only block-level locations are shown. The dataset includes reports that have not been fully verified, and the type of crime updated as investigations continue. There also be mistakes in humans or machines. Therefore, the Chicago Police Department does not guarantee that the data is completely accurate, up-to-date, or error-free, and used to compare crime trends over time. First, image normalization is applied to scale the pixel intensity values to the [0, 1] range, improving the model convergence during training. Next, the images are resized to a standardized dimension of 256 × 256 pixels (width × height) using computer vision libraries in Python to ensure a uniform input size and reduce computational complexity. Data augmentation techniques, like random horizontal flipping, rotation, and scaling, are employed to increase dataset diversity and reduce overfitting, thereby, enhancing the model's generalization capability. After preprocessing, the fragmented video frames are shuffled and split into training, validation, and testing subsets. The dataset is partitioned into 80% training and 20% testing. For a robust evaluation, both k-fold cross-validation (k = 10) and hold-out validation strategies are used. The k-fold cross-validation ensures stable performance estimation by averaging the results across folds, whereas hold-out validation calculates the average performance over multiple simulation runs between the train and test subsets.

## 2.2 Feature Extraction using DeiT

The particle size effect correction model mainly modifies, using particle size distribution data, the measured ash content in coal samples, with the goal of minimizing errors due to size variations and improving measurement accuracy. The initial step involves calculating to determine the ash error, the difference

between the actual and measured ash content. Through normalization and regularization, the error is subsequently adjusted to serve as a target for ash measurement correction. Three-channel color images with a resolution of 224 × 224 pixels are used as input for training in the model design. Mask images provide detailed physical characteristics of larger particles on the surface of coal samples, such as size, shape, color, and spatial position. Recent advancements in deep learning have significantly improved image analysis, resulting in the creation of many efficient models. Models like ViT have shown strong performance in image classification, as transformers excel in feature extraction and capturing global information. Self-attention is the primary mechanism behind transformer models, enabling them for recognizing global patterns and capturing spatial dependencies across the image. To enlarge the receptive field progressively, CNN models, in contrast, utilize local receptive fields and require several convolution and pooling layers. Since large datasets are required for training conventional transformer models, analysing coal composition becomes challenging because of the significant expense involved in obtaining large-scale coal image datasets. The data-efficient image transformer (DeiT) model is employed to overcome this challenge utilized with transfer learning. Applying a training technique known as "distillation," DeiT is an enhanced transformer model designed to improve performance on small datasets. In distillation, the student model (DeiT) learns more efficiently under the guidance of a well-trained teacher model, which is often a traditional, high-performance model, such as CNN. By using this method, even from small datasets, DeiT effectively extracts meaningful image features, which is crucial for coal image analysis when data availability is limited.

DeiT's multi-head self-attention mechanism enables it to capture interactions between global features, in contrast to conventional CNNs, which rely on convolutions focused only on local regions. By utilizing pre-trained models, DeiT achieves good performance on smaller labeled datasets due to its data efficiency. This approach tailors the regression head to meet specific task requirements, allowing for a direct transformation of coal particle images into ash content predictions, enhanced by analyzing complex relationships and multimodal feature fusion. DeiT automatically extracts complex image features, overcoming the challenges and subjectivity of manual feature design that are typical of traditional methods relying on handcrafted features. Capturing the overall structure of coal images is enhanced by the global receptive field, which is crucial for composition analysis that depends on understanding these global patterns. DeiT works alongside CNNs, offering for ash correction, a robust modeling approach that boosts generalization and prediction accuracy. Consequently, using DeiT for regression tasks in ash error correction proves highly effective.

### 2.3 Ontology-based GeoSPARQL

To create an ontology, the available data sources are analysed, followed by the use of a top-down or bottom-up method. The method is grounded in this conceptual framework, developed by identifying key concepts, properties, and spatial relationships within the domain. Rather than building an ontology from scratch, it is recommended to enhance existing ontologies by incorporating custom elements to address specific domain requirements. Utilizing established ontologies brings several benefits, including recognized standardization within the geospatial community, efficiency in terms of cost and time, seamless compatibility with other systems and datasets. To manage and incorporate geospatial and observational data, an extended ontology is developed for this project. Includes properties for recording specific measurements, identifiers and spatial coordinates. The ontology includes classes for features such as weather stations, lakes, farmland plots, buildings, and points of interest. GeoSPARQL is well suited due to its simplicity, encompassing fundamental geospatial aspects like data types, data properties for various geospatial features, and object properties for modeling relationships, especially since environmental geospatial data is the focus of the use cases. geospatial data. By employing this strategy, the ontology handles spatial and non-spatial data, which simplifies the integration, querying, and analysis of geospatial information from multiple sources for environmental management. Users create mappings using the Ontop Protégé plugin to link the database with the ontology, enabling seamless integration and efficient querying of the underlying data.

## 3 Result

For the experiments, a machine with Windows OS, an Intel Core i7 10th generation processor, 16 GB of RAM, and an 8 GB NVIDIA RTX 3090Ti GPU is used. CUDA nvJPEG 12.3.3.54 is employed to boost GPU processing. The software setup consisted of Python 3.9.1, TensorFlow 2.2, and Keras 2.3.1 libraries.

### 3.1 Evaluation Metrics

The performance is evaluated for each method of crime-scene reconstruction, and the definitions of these matrices are provided in Table 1.

**Table 1.** Evaluation metrics.

| Metric | Formula |
|---|---|
| Accuracy | $\frac{TP + TN}{TP + TN + FP + FN}$ |
| Precision | $\frac{TP}{TP + FP}$ |
| Recall | $\frac{TP}{TP + FN}$ |
| F1 Score | $2 \times \frac{Precision \times Recall}{Precision + Recall}$ |

### 3.2 Performance Evaluation

The performance of the proposed DeiT-OEL method on two benchmark crime datasets, namely UCF Crime and Chicago Crime, is evaluated and compared with popular transformer-based models, including BERT, RoBERTa,

DeBERTa, and the Swin Transformer. In terms of accuracy, the proposed DeiT-OEL achieved 99.99% for UCF Crime and 99.95% for Chicago Crime, significantly outperforming all the baseline models. Similarly, the proposed method demonstrates higher F1-score, precision, and recall than existing models, confirming its robustness and reliability for crime detection. Overall, these results indicate that the DeiT-OEL model is highly effective in identifying crime events from video and textual data, achieving a consistently high performance across multiple datasets. Table 2 presents a detailed performance comparison of the DeiT-OEL with the baseline transformer models.

**Table 2.** performance analysis of proposed DeiT-OEL method.

| Datasets | Models | Accuracy (%) | Precision (%) | Recall (%) | F1-score (%) |
|---|---|---|---|---|---|
| UCF Crime | BERT | 86.5 | 84.7 | 83.9 | 84.3 |
| | RoBERTa | 91.2 | 90.5 | 90.0 | 90.2 |
| | DeBERTa | 93.7 | 93.2 | 92.8 | 93.0 |
| | Swin Transformer | 92.5 | 94.0 | 95.8 | 95.9 |
| | Proposed (DeiT-OEL) | 99.99 | 99.92 | 99.87 | 99.79 |
| Chicago Crime | BERT | 83.8 | 83.0 | 82.5 | 82.7 |
| | RoBERTa | 89.5 | 88.7 | 88.2 | 88.4 |
| | DeBERTa | 92.0 | 91.5 | 91.0 | 91.2 |
| | Swin Transformer | 93.2 | 92.7 | 93.5 | 94.6 |
| | Proposed (DeiT-OEL) | 99.95 | 98.90 | 99.87 | 98.74 |

### 3.3 Computational Complexity

The proposed DeiT-OEL method is evaluated for training and inference times and compared with the baseline transformer models. The training time of 70 s is significantly lower than that of BERT (220 s). Similarly, the inference time of the proposed DeiT-OEL model is 80 s, compared with 232 s for BERT. The results demonstrate that DeiT-OEL reduces both the training and inference times, making it highly efficient. The proposed model is computationally lightweight and suitable for real-time crime detection applications. Figure 2 presents a comparison of the computational complexity of all baseline and proposed models.

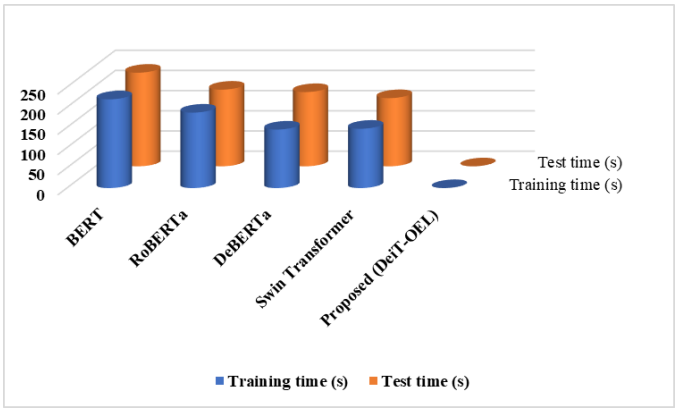


**Fig. 2.** Computational complexity of DeiT-OEL method.

### 3.4 Statistical Analysis

Compared with the traditional BERT, RoBERTa, DeBERTa, and Swin Transformer models, the proposed DeiT-OEL method demonstrated superior accuracy and adaptability across the UCF and Chicago crime datasets. The improvement in the classification accuracy was statistically significant ($p < 0.05$), as validated by a paired t-test conducted over 10 independent runs. These results confirm the robustness, reliability, and effectiveness of the proposed DeiT-OEL method for scalable and precise real-time crime-detection applications.

### 3.5 Ablation Study

The ablation study in this section demonstrates that each component of the proposed DeiT-OEL method enhances performance on the UCF Crime and Chicago Crime datasets. Starting with the base ViT model, the accuracy and overall metrics are comparatively lower; however, introducing the DeiT architecture improved data-efficient learning, resulting in better accuracy and F1-score. When GeoSPARQL integration is included, the model became more effective at incorporating spatial and semantic relationships, leading to further improvements in precision, recall, and F1-score. Finally, the proposed DeiT-OEL framework optimizes attention and feature learning, reduce redundancy, and maximizes discriminative power, which enables the method to achieve superior performance. Each added component contributed significantly, thereby enhancing the model to achieve high classification accuracy and reliability, as shown in Table 3.

**Table 3.** Ablation study.

| Dataset | Models | Accuracy (%) | Precision (%) | Recall (%) | F1-score (%) |
|---|---|---|---|---|---|
| UCF Crime | Vit | 88.5 | 87.8 | 86.9 | 87.3 |
| | DeiT | 92.3 | 91.9 | 91.2 | 91.5 |
| | DeiT+ GeoSPARQL | 93.8 | 92.5 | 92.2 | 93.3 |
| | Proposed (DeiT-OEL) | 99.99 | 99.92 | 99.87 | 99.79 |
| Chicago Crime | Vit | 85.2 | 84.6 | 84.0 | 84.3 |
| | DeiT | 90.5 | 90.0 | 89.7 | 89.8 |
| | DeiT+ GeoSPARQL | 94.3 | 93.8 | 94.0 | 93.7 |
| | Proposed (DeiT-OEL) | 99.95 | 98.90 | 99.87 | 98.74 |

### 3.6 Comparative Analysis

The proposed DeiT-OEL method is evaluated on the UCF Crime and Chicago Crime datasets and compared with existing crime scene classification methods. Using the UCF Crime dataset, the CRBA [12] method reported an accuracy of 86.20%, whereas YOLO-v4 + 3DCNN [13] achieved 94.21% accuracy. In contrast, the proposed DeiT-OEL method attained an accuracy of 99.99%, indicating a significant performance gain. Similarly, on the Chicago Crime dataset, CNN+ABC [15] achieved an accuracy of 96.41%. The proposed DeiT-OEL method achieved 99.90% accuracy, surpassing all baseline methods. These results confirm the effectiveness and superiority of the DeiT-OEL method for accurate and reliable crime-scene image classification, as shown in Table 4.

**Table 4.** Comparative analysis of existing models.

| Datasets | Models | Accuracy (%) |
|---|---|---|
| UCF crime | CRBA [12] | 86.20 |
| | YOLO-v4 + 3DCNN [13] | 94.21 |
| | Proposed method (DeiT-OEL) | 99.99 |
| Chicago crime | CNN+ABC [15] | 96.41 |
| | Proposed method (DeiT-OEL) | 99.90 |

## 4 Discussion

The experimental results indicate that the proposed DeiT-OEL method achieves superior performance compared with existing models on the UCF Crime and Chicago Crime datasets. The integration of DeiT with GeoSPARQL ensures the efficient extraction of global image features and effective mapping of semantic and spatial relationships between crime scene elements, strengthening the feature representation for accurate crime scene classification. Preprocessing techniques, including image normalization, resizing, and data augmentation, further enhance quality and diversity of input data and improve the model's generalization capability. The ablation study confirms that each component contributes positively to performance, with DeiT enabling data-efficient learning, GeoSPARQL enhancing spatial-semantic reasoning, and the full DeiT-OEL framework achieving optimal accuracy, precision, recall, and F1-score. Additionally, the method exhibits superior computational efficiency, with significantly reduced training and inference times compared to the BERT, RoBERTa, and Swin Transformer models, making it highly suitable for real-time crime scene reconstruction and forensic analysis applications.

## 5 Conclusion

This paper proposes the DeiT-OEL method, a Data-efficient Image Transformer with Ontology-based Evidence Linking, for intelligent crime scene reconstruction. DeiT-OEL combines efficient image feature extraction using DeiT with GeoSPARQL for semantic and spatial evidence reasoning, thereby improving feature representation and classification accuracy. Experimental results using the UCF Crime and Chicago Crime datasets demonstrate that DeiT-OEL significantly outperforms the existing models, achieving classification accuracies of 99.99% and 99.95%, respectively. The ablation study validated the contribution of each component, while computational analysis confirmed the reduced training and testing times, demonstrating its computational efficiency. Overall, DeiT-OEL provides a reliable, scalable, and high-performance solution for automated crime-scene analysis and forensic investigations. Future work will focus on expanding the ontology for broader forensic applications and integrating multimodal data sources to further improve the evidence linking and decision-making accuracy.

## Reference

1. M. Albeedan, H. Kolivanda, R. Hammady, Designing and evaluation of a mixed reality system for crime scene investigation training: a hybrid approach. Virtual Reality **28**, 127 (2024). https://doi.org/10.1007/s10055-024-01018-8
2. C. Roux, R. Bucht, F. Crispino, P. De Forest, C. Lennard, P. Margot, M.D. Miranda, N. NicDaeid, O. Ribaux, A. Ross, S. Willis, The Sydney declaration–Revisiting the essence of forensic science through its fundamental principles. Forensic Sci. Int. **332**, 111182 (2022). https://doi.org/10.1016/j.forsciint.2022.111182
3. S.H. Yu, G. Thomson, V. Rinaldi, C. Rowland, N. Nic Daeid, Development of a Dundee Ground Truth imaging protocol for recording indoor crime scenes to facilitate virtual reality reconstruction. Sci.

Justice **63**, 238–250 (2023). https://doi.org/10.1016/j.scijus.2023.01.001

4. P.H. Home, D.G. Norman, K. Wade, E. Spearing, M.A. Williams, 3D scanning a crime scene to enhance juror understanding of Bloodstain Pattern Analysis evidence. Sci. Justice **64**, 333–338 (2024). https://doi.org/10.1016/j.scijus.2024.04.007
5. A. Georgiou, P. Masters, S. Johnson, L. Feetham, UAV-assisted real-time evidence detection in outdoor crime scene investigations. J. Forensic Sci. **67**, 1221–1232 (2022). https://doi.org/10.1111/1556-4029.15009
6. H. Swofford, S. Lund, H. Iyer, J. Butler, J. Soons, R. Thompson, V. Desiderio, J.P. Jones, R. Ramotowski, Inconclusive decisions and error rates in forensic science. Forensic Sci. Int. Synergy **8**, 100472 (2024). https://doi.org/10.1016/j.fsisyn.2024.100472
7. A.S. Knes, M. de Gruijter, M.C. Zuidberg, C.J. de Poot, CSI-CSI: Comparing several investigative approaches toward crime scene improvement. Sci. Justice **64**, 63–72 (2024). https://doi.org/10.1016/j.scijus.2023.11.009
8. H.V. Wilkins, V. Spikmans, R. Ebeyan, B. Riley, Application of augmented reality for crime scene investigation training and education. Sci. Justice **64**, 289–296 (2024). https://doi.org/10.1016/j.scijus.2023.11.009
9. M.A. Daibes, M.R. Alsadi, M.M. Alruwaili, M.Y. Alzaatreh, M.T. Mrayyan, H.Y. Abunab, M.J. ALhemedi, Knowledge about crime scenes and evidence management among emergency medical team professionals. BMC Emerg. Med. **25**, 75 (2025). https://doi.org/10.1186/s12873-025-01230-y
10. R.J. Nathan, B. Okeleye, R. Abdullahi, W. Oyebode, Identification of ketamine and norketamine in dried bloodstains on crime-scene surfaces. Egypt. J. Forensic Sci. **15**, 1 (2025). https://doi.org/10.1186/s41935-024-00418-w
11. A. Sedik, H. Kolivand, M. Albeedan, An efficient image classification and segmentation method for crime investigation applications. Multimed. Tools Appl. **84**, 19399–19423 (2025). https://doi.org/10.1007/s11042-024-19773-w
12. S. Natha, F. Ahmed, M. Siraj, M. Lagari, M. Altamimi, A.A. Chandio, Deep BiLSTM attention model for spatial and temporal anomaly detection in video surveillance. Sensors **25**, 251 (2025). https://doi.org/10.3390/s25010251
13. A. Rehman, T. Saba, M.Z. Khan, R. Damaševičius, S.A. Bahaj, Internet-of-Things-Based Suspicious Activity Recognition Using Multimodalities of Computer Vision for Smart City Security. Secur. Commun. Netw. **2022**, 8383461 (2022). https://doi.org/10.1155/2022/8383461
14. A. Alshalawi, W. Abdul, G. Muhammad, Fire Swin: Video Anomaly Detection Using a hybrid model with convolutional layers, fire module, and Swin transformer. J. Eng. Res. (2025). https://doi.org/10.1016/j.jer.2025.08.016
15. L. Yang, J. Guofan, Z. Yixin, W. Qianze, Z. Jian, R. Alizadehsani, P. Pławiak, A reinforcement learning approach combined with scope loss function for crime prediction on Twitter (X). IEEE Access. **12**, 149502-149527 (2024). https://doi.org/10.1109/ACCESS.2024.3473296
16. UCF crime dataset Link: https://www.kaggle.com/datasets/odins0n/ucf-crime-dataset (Accessed on 07.09.2025)
17. Chicago crime dataset Link: https://www.kaggle.com/datasets/chicago/chicago-crime (Accessed on 07.09.2025)