International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 07 July 2026
Page No: 5025-5031
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0661

Hybrid AI Framework for Bloodstain Detection, Classification and Forensic Scene Reconstruction

Divit P[1], A R Hemanth Reddy[2], M. Nithya[3]
[1,2,3]Artificial Intelligence and Data Science, St Joseph's College of Engineering, Chennai, India.
Email ID: *divitpandi5@gmail.com[1], hem.red1903@gmail.com[2], nithyasmail6@gmail.com[3]*

Abstract
Bloodstain Pattern Analysis is used in forensic investigations for reconstructing the sequence of events that occurred at a crime scene. To improve the forensic reasoning, we use Retrieval augmented Generation that is integrated in the pro- posed system. To classify the bloodstain into the given categories we use YOLO based detection model. The bloodstain is further analyzed by Morphological and spatial features and the results are sent to RAG engine. This RAG engine uses Forensic literature and do aim specific resources and historic case to reconstruct the crime scene. Feedback from the expert forensic officers is used to reinforce and increase the accuracy of the reasoning engine. The framework connects AI driven pattern recognition with expert forensic knowledge, resulting in explainable and reliable tool which assists investigators in decision making and evidence interpretation faster.

Keywords: *component, formatting, style, styling, insert.*

1. Introduction

In Forensic Science the Bloodstain Pattern Analysis plays a key part role that provides insights in direction and dynamics of event occurring during violent crimes. The Experts from the forensic can only predict information like weapon used, number of impact events and position and movement of individual by analyzing the bloodstain patterns that can be categorized into flows, pools, wipes, splatters and drops. By advancing the Artificial intelligence and computer vision and deep learning in Bloodstain Pattern analysis we increase and enhance the acccuracy and objective of bloodstain pattern analysis by using YOLO model we automate object detection that show strong performance in visual recognition tasks and are well suited for identifying and classifying bloodstain patterns from crime scene images. By automating the process, we make quick progress and less human error and not biased towards one thing unlike human's decision making. Even though deep learning models are effective in object detection they lack in domain based context reasoning. For this lackness we created a Retrieval augmented generation that is a Large language model that takes knowledge base from various document materials that are forensic science and bloodstain pattern based. these document are fed into the RAG and makes them a forensic reasoning expert. The retrieval augmented Generation model gives a full contexual reasoning explanation and also converting them into visual representation with interpretable forensic narratives. One of the significant thing in this proposed model is that we have added a human in the loop that is a feedback mechanism that experts can view and make changes to the resultant reasoning that is proposed by the system. This feedback gives the user an interactive and improved system. This feedback mechanism improves the reasoning accuracy and whenever the proposed makes a mistake the human in the loop that is the forensic expert gives a correct feedback that corrects the system from the error [1-10].

2. Related Work

2.1. Traditional BPA Techniques

The traditional bloodstain pattern analysis techniques use trigonometry, estimation of impact angle, and recreation of trajectory to provide information about the origin of the source of blood and the chronology of events. The further movement of blood droplets are determined by methods like stringing and area-of-origin analysis. These approaches can be viewed as effective, their accuracy depends on the competence of the analyst and the lack of generally accepted norms that may lead to different interpretations and possible mistakes.

2.2. Machine Learning in Forensic Image Analysis

In recent research, convolutional neural networks have been used to computerize bloodstain patterns. YOLO model have been used to identify stains in the

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 07 July 2026
Page No: 5025-5031
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0661

complexity images which is highly accurate and recalling. Other techniques add feature extraction (like gray level co-occurrence matrix texture analysis and shape descriptors) in order to enhance the precision of classification. Models usually stop at the visual categorization and do not provide any contextual forensics, thus limiting their practical usefulness [11-20].

2.3.AI Reasoning and Knowledge Retrieval in Forensics

Large Language Models (LLMs) and RAG architectures provide the capability to combine visual detection with domain knowledge. RAG systems retrieve relevant literature or case reports and generate context-aware outputs. In forensic applications, this enables the AI to not just detect stains, but infer likely weapon type, movement patterns, and event sequences, producing outputs that are interpretable and actionable.

3. Proposed Methodology

3.1.System Overview

The framework integrates detection, feature analysis, reasoning, and feedback (Figure 1). Input images are processed by YOLOv8, producing bounding boxes and class probabilities. Morphological and texture features are extracted from the detected stains. The RAG module retrieves relevant literature and generates a scene reconstruction narrative, which can be reviewed and corrected by forensic officers. Feedback is used to fine-tune the detection and reasoning models, creating an adaptive system.

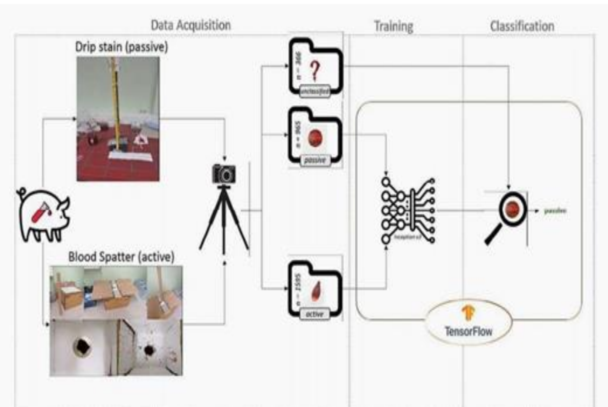

Figure 1 Proposed System Architecture. The pipeline: Input Image → YOLOv8 Detection → Feature Extraction → RAG Reasoning → Feedback Loop

3.2.Dataset Preparation

Usually The Existing Bloodstain pattern analysis technique use trigonometry, estimation of impact angle and recreation of trajectory to recreate and get information about the event. By using other methods like stringing and area of origin analysis we determine the further movement of the blood droplets. These approaches are effective and their accuracy depends on the competence of the analyst lack of generally accepted norms that may lead to different interpretations and possible mistakes. The pertained yolov8 model is a robust model that is used multiple object detection in a single image and this takes only few training data. When a new image that contains multiple bloodstain patterns of various bloodstain types this yolov8 model creates a bounding boxes representing each bloodstain pattern and with their class or label mentioned in the bounding box. This allows for simultaneous localization and classification of multiple stains within a single image, effectively automating what would otherwise be a time-consuming manual process. The outputs obtained from the detection module can be integrated into a Retrieval-Augmented Generation (RAG) or LLM based reasoning framework, where the identified stain classes are used for querying forensic knowledge bases and available reference literature Shown in Figure 2.

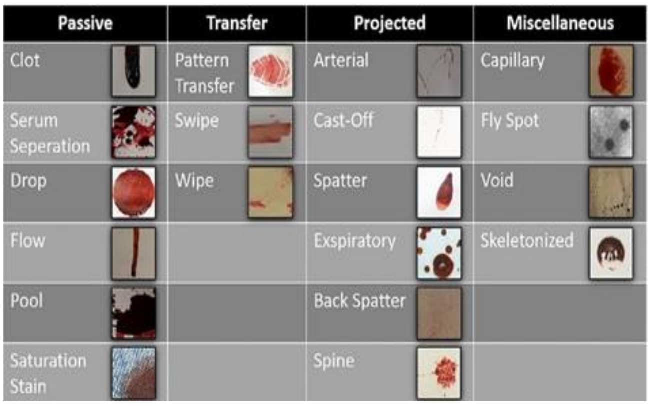

Figure 2 The bloodstain dataset image. The dataset contains multiple classes including drop, flow, high splatter, low splatter, pool, medium splatter, pattern, swipe, and wipe

Through this process, context- aware narratives related to the possible dynamics of the blood- shed event are generated by the system, which can support crime scene reconstruction. By combining accurate detection, stain classification and reasoning, a scalable and interpretable forensic analysis framework is provided. Human errors are reduced, investigation efficiency is improved, and evidence-based reconstruction of complex crime scenes is facilitated [21-30].

3.3.YOLO-Based Detection and Classification

The YOLOv8 model is employed to detect and classify bloodstains in forensic images. YOLOv8 is a single-stage object detector capable of real-time multi-class detection. It uses the CSPDarkNet backbone for feature extraction and PANet (Path Aggregation Network) for multi-scale feature aggregation, allowing it to detect stains of varying sizes and shapes. Each detected bloodstain is output with:

- Bounding box coordinates (x,y,w,h)(x, y, w, h)(x,y,w,h)
- Class label (e.g., drop, flow, pool, swipe, etc.)
- Confidence score representing detection probability

YOLOv8 Loss Function: The YOLOv8 training objective is a weighted sum of three loss components:

$$L = \lambda_{box} L_{box} + \lambda_{obj} L_{obj} + \lambda_{cls} L_{cls} \qquad (1)$$

3.4.Morphological and Texture Feature Extraction

Each identified stain is then evaluated for a wide range of geometric and morphological parameters, such as shape parameters (circularity, solidity, perimeter), area, orientation (angle with respect to a reference axis), and eccentricity, which describes the elongation of the stain.to extract advanced spatial and texture analysis, we not only employ a geometric feature, we employ a paradigm to adequately describe the patterns of bloodstain patterns. The most important statistics are extracted from texture analysis using Grey Level Co-occurrence Matrices. Contrast quantifies that checks the intensity differences between neighboring pixels, and we employ correlation to obtain linear relationships between grey levels, energy that describes textual uniformity. Homogeneity that measures the smoothness of intensity distribution altogether, these describe the stain texture that enhances the distinction of bloodstain patterns. To capture the global and local shape variations of blood- stains, enabling robust differentiation among stains such as drops, pools, flows and wipes we use moment in variations and Fourier descriptors. These shape descriptors are combined with forensic reasoning pipeline, here they serve as interpretable inputs for classification models and help on reconstruction of dynamic crime scene event.by making use of these geometric properties of blood, orientation information and texture based metrics the proposed solution builds a multi dimension representation of each bloodstain that enhances the probability of pattern recognition accurately and improves the crime scene reconstruction Shown in Figure 3.

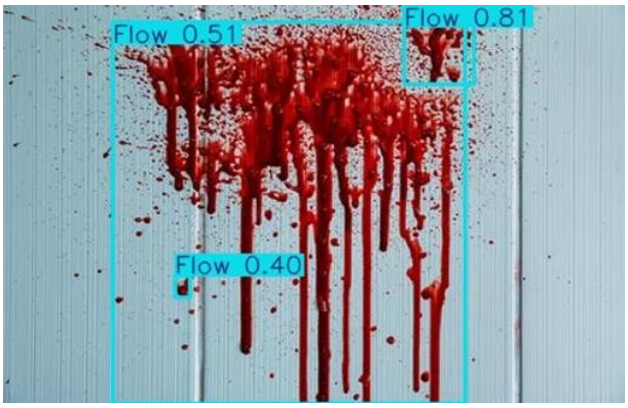

Figure 3 Detected bloodstains analysis showing geometric and texture features such as area, orientation, eccentricity, and GLCM-based properties

The Retrieval Augmented Generation based reasoning model gets knowledge base from various forensic literature, case reports and domain specific books and journals to generate hypothesis based on the sequence of events.by integrating retrieved domain knowledge and model's reasoning, the solution provides context based explanations and backed inter- predations. This enhances the forensic decision making. We also created a feedback based learning mechanism that combines corrections and validations from forensic experts, this helps refinement of model detection and that shows improvement in accuracy, reliability and long term

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 07 July 2026
Page No: 5025-5031
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0661

system performance.

3.5.Retrieval-Augmented Generation (RAG)-Based Reasoning

This RAG model is used here as a bridge between visual detection and forensic knowledge base. Once the Yolo model detects bloodstain patterns this result is passed as a query to the forensic LLM database that consist of data from previous case reports and many domain specific documents and literature that includes past case study. the system first looks up similar forensic knowledge and then explains what the detected bloodstain patterns most likely indicate, resulting in a clear and contextual narrative of the events that occurred.

3.6.Feedback Mechanism and Continuous Learning

To adapt the Proposed Solution to adapt in real world setting we incorporate a human in this loop approach. After the completion of the above pipeline process, the forensic experts analyze the output generated by the proposed solution and provide feedback that helps in the reinforcement learning process. When the experts provide corrections and explanations, the feedback is again incorporated into the system to improve the predictions Shown in Figure 4.

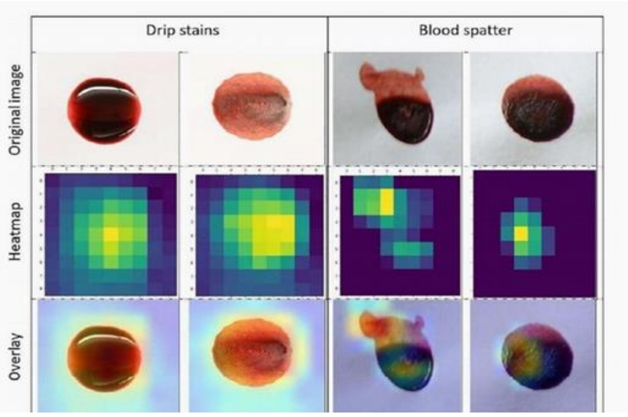

Figure 4 Gray-Level Co-occurrence Matrices (GLCM), deriving statistics such as contrast (measuring intensity variation)

With time, the system gradually learns from the expert feedback and improves its ability to recognize different stain patterns and explain their forensic values. This process helps in the reduction of errors and provides more reliable results. As the system is exposed to new stain patterns, difficult lighting conditions, and different crime scene settings, it learns from these experiences and thus becomes a scalable solution that develops along with the advancements in the forensic field.

4. Experimental Setup and Results

4.1.Implementation Environment

The Proposed Solution is developed using python in pytorch that is used for building the deep learning model and Open CV that is used for image Preprocessing and data augmentation tasks. Lang Chain is integrated to support RAG pipeline that enables contexual reasoning and knowledge based interpretation. This workstation is equipped with NVIDIA RTX series GPU and 16 gb of ram that speed up the process. The dataset was divided into 80 percent for training and 20 percent for testing, ensuring that all bloodstain categories were evenly represented in both sets.

4.2.Evaluation Metrics

To Determine the performance of the Model we use Metrics such as Precision, Recall, F1 Score and mean Average Precision. To evaluate the reasoning, we make forensic expert to check for any corrections and review it. And another assessment to review the reconstruction of the possible crime scene, reflecting its effectiveness in supporting forensic analysis.

4.3.Results

This Proposed hybrid consist of YOLO v8 Based Detection Module and a RAG driven reasoning component. The mean average prediction of the Detection module is 94.3 percentage accuracy in all bloodstain Category. The Reasoning RAG Module gave an 87 percentage of the evaluated cases showing its ability to reconstruct meaning full and context aware forensic explanation. The Reason Behind these inaccuracies where the scenes involving overlapping or distorted stains, this remains Challenging due to inherent complexity of real world crime scenes.

4.4.Comparative Analysis

The Yolo v8 Detection Module has a clear Advantage Over the Conventional CNN based classifier. This Connects the Geometric and texture based features so that this improves interpretability with the generated forensic explanation. The RAG based reconstruction creates a crime scene based on the input images. Moreover, the Human in this loop who is a forensic expert gives feedback that creates a reinforcement

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 07 July 2026
Page No: 5025-5031
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0661

learning for the proposed system. Comparatively this hybrid module that contains RAG Based Reasoning and Yolo v8 detection gives a more reliable and practically useful tool.

5. Discussion

This Proposed solution is a hybrid module that contains different models for each Pattern Detection, Reasoning and Feedback Mechanism. For The Blood Stain Pattern detection, we use a Yolo v8 model that directly connects geometric and texture analysis to improve the interpretability. This Yolo v8 model also helps detect more than one bloodstain pattern in a image by drawing a boundary box on each stain pattern. This is fast and accurate method to detect multiple bloodstain pattern in an image. Next is a Reasoning module that uses RAG based model that has a LLM that retrieved data from various domain specific documents by various Authors. So from the image we calculate various metrics, this creates a query that consist these metrics and the detected bloodstain from the Yolo v8 model. This RAG reasoning module enhances this Proposed solution as this reconstructs the crime scene drawing know- edge from forensic literature and other domain specific documents.so from these reconstructions the user can get a clear idea on the crime scene.so create a more accurate reconstruction we create a feedback mechanism on the crime scene with a forensic expert so that it is reliable, accurate and practically useful over time.

Conclusion

This work presents a comprehensive hybrid AI framework for bloodstain pattern analysis (BPA) and crime scene reconstruction, combining deep learning detection, morphological feature analysis, retrieval-augmented reasoning, and expert feedback. The YOLOv8 model enables fast and accurate detection of diverse bloodstain types, while the extraction of quantitative descriptors such as stain area, shape, and trajectory provides a robust basis for forensic interpretation. The reasoning module using the retrieval-augmented generation (RAG) approach enhances understanding by contextualizing observed patterns in a knowledge base of forensic literature, thus facilitating hypotheses based on evidence and improving the accuracy of scene reconstruction. By incorporating human feedback, the system allows for continuous learning and adaptation, mitigating potential weaknesses in automated analysis and ensuring consistency with expert opinion. The system has excellent performance, interpretability, and applicability in digital forensic analysis. Remaining challenges, such as overlapping stains, occlusions, and knowledge gaps, still exist but expected improvements, such as three-dimensional reconstruction, multi-camera analysis, and domain-specific fine-tuning of large language models, have the potential to further enhance accuracy and usability. In summary, this work presents a new system that combines state-of-the-art artificial intelligence approaches with forensic knowledge, thus promoting faster, more accurate, and evidence-based crime scene reconstruction and bloodstain pattern analysis.

References

[1]. S. Kannan, "A dataset of drip patterns for teaching and research in bloodstain pattern analysis," Data in Brief, vol. 45, p. 108348, 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2352340925000848
[2]. Y. Liu and X. Zhang, "Enhancing bloodstain analysis through AI- based segmentation: Leveraging Segment Anything Model for crime scene investigation," arXiv preprint arXiv:2308.13979, 2023. [Online]. Available: https://arxiv.org/abs/2308.13979
[3]. Y. Li and C. Zhang, "Machine learning for blood pattern classifi- cation," arXiv preprint arXiv:2501.02151, 2025. [Online]. Available: https://arxiv.org/abs/2501.02151
[4]. ACM Research, "ForensicSight: A Forensic Image Analysis Tool," 2023. [Online]. Available: https://github.com/ACM-Research/ForensicSight
[5]. R. M. Arthur, "An automated approach to the classification of impact spatter bloodstains," Forensic Science International, vol. 291, pp. 1–7, 2018. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/29933203/
[6]. H. Jung, "Bloodstain pattern classification," IoT Lab, 2024. [Online]. Available: https://iotlab.skku.edu/publications/international- journal/Bloodstain-Pattern-

Classification-FSI2024.pdf
[7]. S. Yildirim Yayilgan, "Enhancing criminal investigation analysis with KriRAG," Proceedings of COLING 2025, pp. 334–344, 2025. [Online]. Available: https://aclanthology.org/2025.coling-main.334.pdf
[8]. F. Y. Loumachi, "Advancing cyber incident timeline analysis through retrieval-augmented generation," Journal of Cybersecurity, vol. 14, no. 2, p. 67, 2025. [Online]. Available: https://www.mdpi.com/2073- 431X/14/2/67
[9]. "Retrieval-Augmented Generation (RAG) for LLMs in 2025," Rapid Innovation, 2024. [Online]. Available: https://www.rapidinnovation.io/post/retrieval-augmented-generation- using-your-data-with-llms?ref=chitika.com
[10]. "What Is Retrieval-Augmented Generation (RAG)? An Introduction," Palo Alto Networks, 2024. [Online]. Available: https://www.paloaltonetworks.com/cyberpedia/what-is-retrieval- augmented-generation
[11]. "Bloodstain Pattern Data Set," Mendeley Data, 2022. [Online]. Available: https://data.mendeley.com/datasets/234yvmkcys
[12]. "Blood Cell Detection with YOLOv8 from Scratch," Kaggle, 2025. [Online]. Available: https://www.kaggle.com/general/436883
[13]. "Bloodstain Pattern Data Set," CSAFE Forensic Science Dataset Portal, 2017. [Online]. Available:https://forensicstats.org/data/
[14]. "A Novel, Data-Driven Approach to the Classification of Bloodstain Pat- terns," Center for Statistics and Applications in Forensic Evidence, 2021. [Online]. Available: https://forensicstats.org/blog/portfolio/a-novel-data- driven-approach-to-the-classification-of-bloodstain-patterns/
[15]. "Machine Learning for Blood Pattern Classification," arXiv preprint arXiv:2501.02151, 2025. [Online]. Available: https://arxiv.org/abs/2501.02151
[16]. "Bloodstain Pattern Classification," IoT Lab, 2024. [Online]. Available: https://iotlab.skku.edu/publications/international-journal/Bloodstain-Pattern-Classification-FSI2024.pdf
[17]. "Retrieval-Augmented Generation (RAG) for LLMs in 2025," Rapid Innovation, 2024. [Online]. Available: https://www.rapidinnovation.io/post/retrieval-augmented-generation- using-your-data-with-llms?ref=chitika.com
[18]. "What Is Retrieval-Augmented Generation (RAG)? An Introduction," Palo Alto Networks, 2024. [Online]. Available: https://www.paloaltonetworks.com/cyberpedia/what-is-retrieval- augmented-generation
[19]. "Bloodstain Pattern Data Set," Mendeley Data, 2022. [Online]. Avail- able: https://data.mendeley.com/datasets/234yvmkcys
[20]. "Blood Cell Detection with YOLOv8 From Scratch," Kaggle, 2025. [Online]. Available: https://www.kaggle.com/general/436883
[21]. "Bloodstain Pattern Data Set," CSAFE Forensic Science Dataset Portal, 2017. [Online]. Available: https://forensicstats.org/data/
[22]. "A Novel, Data-Driven Approach to the Classification of Bloodstain Pat- terns," Center for Statistics and Applications in Forensic Evidence, 2021. [Online]. Available: https://forensicstats.org/blog/portfolio/a-novel-data- driven-approach-to-the-classification-of-bloodstain-patterns/
[23]. "Machine Learning for Blood Pattern Classification," arXiv preprint arXiv:2501.02151, 2025. [Online]. Available: https://arxiv.org/abs/2501.02151
[24]. "Bloodstain Pattern Classification," IoT Lab, 2024. [Online]. Available: https://iotlab.skku.edu/publications/international-journal/Bloodstain-Pattern-Classification-FSI2024.pdf
[25]. "Retrieval-Augmented Generation (RAG) for LLMs in 2025," Rapid Innovation, 2024. [Online]. Available:

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 07 July 2026
Page No: 5025-5031
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0661

https://www.rapidinnovation.io/post/retrieval-augmented-generation- using-your-data-with-llms?ref=chitika.com
[26]. "What Is Retrieval-Augmented Generation (RAG)? An Introduction," Palo Alto Networks, 2024. [Online]. Available: https://www.paloaltonetworks.com/cyberpedia/what-is-retrieval- augmented-generation
[27]. "Bloodstain Pattern Data Set," Mendeley Data, 2022. [Online]. Avail- able: https://data.mendeley.com/datasets/234yvmkcys
[28]. "Blood Cell Detection with YOLOv8 from Scratch," Kaggle, 2025. [Online]. Available: https://www.kaggle.com/general/436883
[29]. "Bloodstain Pattern Data Set," CSAFE Forensic Science Dataset Portal, 2017. [Online]. Available: https://forensicstats.org/data/
[30]. "A Novel, Data-Driven Approach to the Classification of Bloodstain Pat- terns," Center for Statistics and Applications in Forensic Evidence, 2021. [Online]. Available: https://forensicstats.org/blog/portfolio/a-novel-data- driven-approach-to-the-classification-of-bloodstain-patterns/