

# Deep Learning in Forensic Injury Identification

## Authors:

Khudooma S. Alnuaimi - Abu Dhabi School of Management – Abu Dhabi, UAE - khudooma@gmail.com - Ishtiaq Rasool Khan - Abu Dhabi School of Management – Abu Dhabi, UAE - i.khan@adsm.ac.ae

Published: 01 Sep 2025

<https://doi.org/10.63962/DIVH3300>

## Abstract

Abstract -- This study implements the EfficientNetB0 deep learning model to classify forensic skin injuries into blunt, sharp, or gunshot categories. Through data augmentation and transfer learning, the model achieved single person's observation an accuracy of 99.3%. A real time Gradio interface, enable intuitive image uploads and classification. This research is supporting the AI Integration in forensic science, medicine and pathology. It is expected to have a significant impact by improving accuracy, enabling automation, and enhancing collaborative investigative case work in the field.

Keywords: Forensic Wound Cavalry, Deep Learning, Convolutional Neural Networks, Image Analysis, Artificial Intelligence

## I. INTRODUCTION

The classification of wounds for forensics is important in criminal investigations because the types of injuries like blunt force trauma, sharp force trauma and gunshot wounds have significant legal implications and influence the construction of the case narrative [3]. Current approaches heavily rely on a single person's observation which is highly time intensive and inaccurate. Automation of forensic skin injury classification is sought through Deep Learning with Convolutional Neural Networks, particularly EfficientNetB0 , in this research [1]. A model was created to classify images of wounds by curating a dataset with a boundless collection of images until a satisfactory accuracy metric was reached through numerous augmentations. The model was evaluated on conventional measurements of performance, including but not limited to, accuracy, precision, recall, F1 score, and a real-time forensic application interface was built using Gradio [4], [8]. Advances in the automation of digital forensics and legal medicine, alongside the unity of processes and bias mitigated nominal controls, enhance the efficiency of the system under automation with changing system variables.

## II. LITERATURE REVIEW

The classification within forensics has included the manual inspection of an expert which adds an element of subjectivity and bias to the variances within the scrutiny and investigations. AI has come forth as an option for complex forensic investigations which continue to increase in number, complexity, and volume [6]. Research shows that Convolutional Neural Networks (CNNs) are competent at classifying blunt force trauma, sharp force injuries, and even gunshot wounds through spatial level imaging interpretation [1]. Incorporating AI standardizes evaluations across forensic facilities, while deep learning enhances the accuracy of the analysis. Models such as EfficientNetB0 have demonstrated scaling and accuracy in forensic settings [3]. Even with these advancements, challenges of how data is collected, class imbalance, and variability among images still exist [4], [8]. Addressing these challenges along with interpreting AI decisions using tools like Grad-CAM, transfer learning, and data augmentation helps bridge the algorithm and image gap.

The expanding capabilities of convolutional neural networks (CNNs) in forensic image evaluation has been documented in recent studies. In [7], deep learning techniques accurately classified stab wounds (93% accuracy) and partially pinpointed several others, managing to poorly perform on abrasions, hematomas, and stabs, which was severely hindered by a class imbalance problem alongside ambiguous boundaries. In the same manner, a study in [6] created a CNN model that determined gunshot wound images of piglet carcasses by the distance of shooting with a testing accuracy of 98%. This illustrates the ability of deep learning to estimate contact, close-range, and distant shots using visual cues exclusively. These studies strengthen the claim that systems based on CNNs can assist forensic specialists and enhanced their capabilities to offer constant, rapid, accurate multidisciplinary assessments of complex and multifaceted wounds and injuries, further demonstrating the capabilities of deep learning in forensic pathology.

Moreover, the problems surrounding data secrecy and confidentiality are important [2]. Regardless, the application of AI-driven investigative technologies fundamentally changes the practice of forensic science due to consistency, impartiality, and legally defensible outcomes [5]. Enhancements in the models are still necessary to improve their functionality in forensic science.

## III. METHODOLOGY

The described methodology captures the systematic process of creating an AI system for classifying forensic skin injuries, including data collection, preprocessing, model creation, training, evaluation, and ethical considerations.

### 1) Research Design

This research adopts a quantitative approach with a focus on AI and follows the CRISP-DM (Cross-Industry Standard Process for Data Mining) framework. The goal is to create a forensic skin injury classifier deep learning model which classifies injuries as blunt force, sharp force, or gunshot wounds through image analysis. The study implements Convolutional Neural Networks (CNNs), selecting EfficientNetB0 as the model of choice given its suitable performance relative to complexity.

### 2) Data Collection and Preparation

Data were collected from open-access forensic image repositories, scholarly journals, and credible online sources that provide forensic wound imagery [5]. These images were

screened to ensure forensic clarity, relevance, and resolution (see Table 1). To meet the requirements of supervised learning, images were manually categorized into three classes: blunt force trauma, sharp force trauma, and gunshot wounds (see Figures 1,2 and 3).

After initial acquisition, data preprocessing [8] steps were implemented:

Image resizing to 224x224 pixels for model compatibility.

Normalization of pixel values between 0 and 1.

Label encoding for multi-class classification.

### 3) Data Augmentation

To tackle the issue of scarce forensic wound photographs, data augmentation methods were used to enhance the dataset and improve model accuracy and reliability [7]. Image rotation, flipping (both horizontal and vertical), zooming, shifting, and adjusting brightness were implemented [9]. All transformations applied allowed the model to capture a variety of imaging conditions, which ensures adequacy for numerous forensic situations (see Table 1). Aid in overfitting from augmentations was seen as a result from diverse categorization of injuries showcased to the model (see Figures 4, 5 and 6).

### 4) Dataset Splitting

Post augmentation, the entire dataset was sliced into three separate chunks: 60% allocated for training, 20% dedicated for validation, and 20% intended for testing. This division allowed retention of the provided distribution of class images at each subset while upholding the evaluation scope. The training set was used for the model fitting, validation set for parameter tuning and overfitting, and the test set for last performance evaluation. This organized partitioning was important to fulfill the model development and case independent generalization requirements in new forensic cases.

### 5) Model Architecture and Training Setup

The chosen architecture, EfficientNetB0, was initialized using transfer learning with pre-trained weights from ImageNet. The architecture was modified to suit the forensic injury classification task:

A global average pooling layer

A dense layer with 512 neurons (ReLU activation)

A dropout layer to minimize overfitting

A final dense output layer with softmax activation for multi-class classification

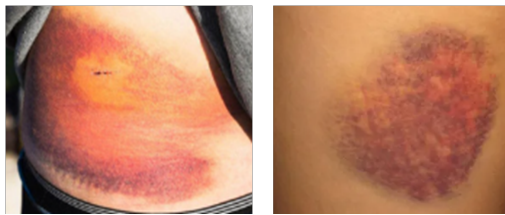
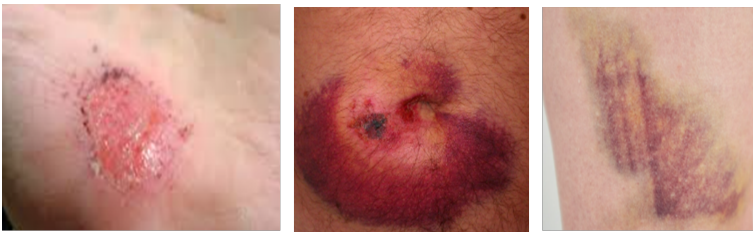
The model was compiled with the Adam optimizer, a categorical cross-entropy loss function, and a learning rate of 0.0001. The training process was managed using early stopping to prevent overfitting

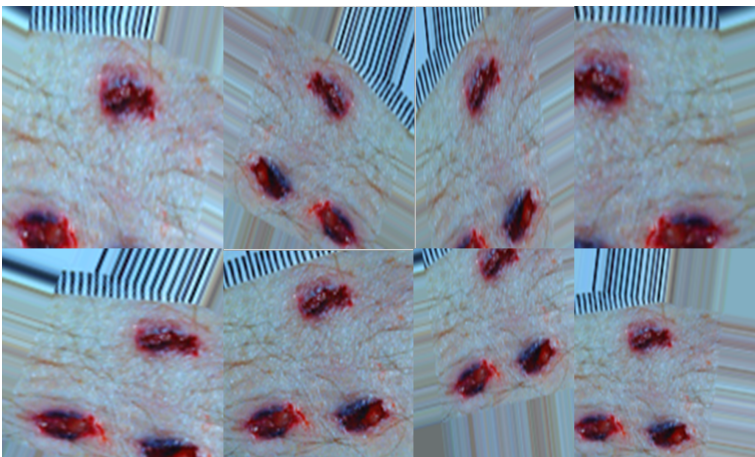
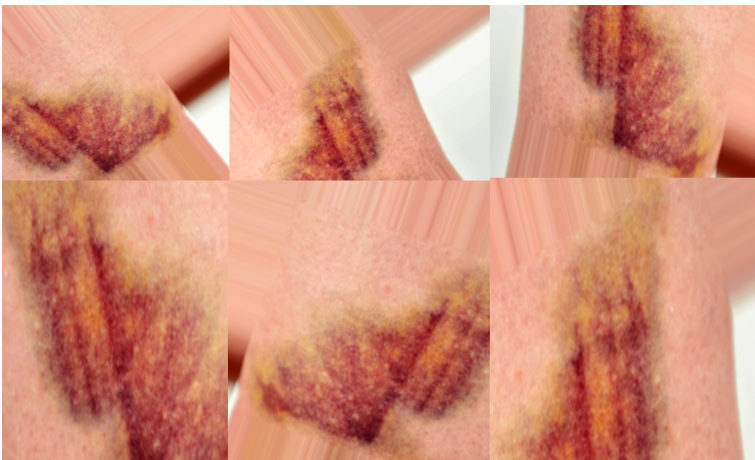
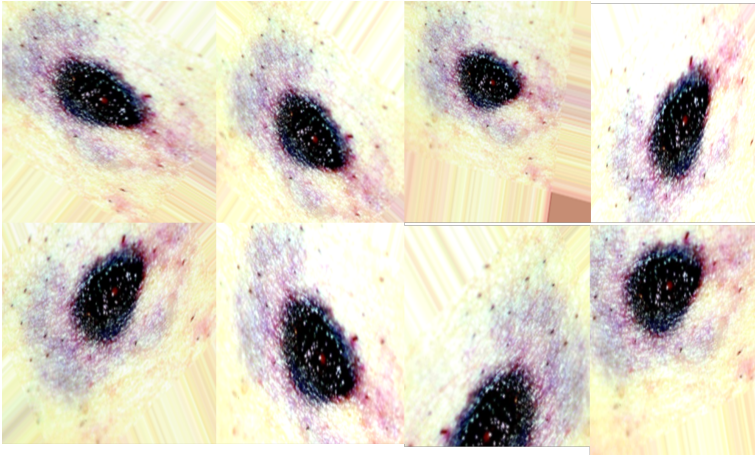
### 6) Evaluation Metrics

In order to maintain the dependability of the forensic wound classification model's performance, it was necessary to use several evaluation metrics. Accuracy, in this case, referred to the estimation of total correct predictions made by the model, while precision evaluated the number of correctly predicted wound types casted by the model as no false positive wounds. Recall measured the number of clearly captured actual wounds but forgot to capture some false negatives. The F1-score defined as a precise mean of precision and recall described an all-encompassing assessment. A confusion matrix displayed trends depicting misclassification of each class with another class of interest. Furthermore, ROC curve alongside AUC was used to evaluate the differentiating capability of the model between the two classes and ascertain its effectiveness and strength in forensic diagnosis while precision medicine was concerned.

## 7) Interface Development

For this particular research, I will implement Gradio, a comprehensive Python package, to create an interactive system for classifying forensic wound images. With Gradio deep learning integration is straightforward, and users can upload pictures of wounds to get classified in real-time with confidence scores computed. This type of interface will be important for users who do not have technical expertise like forensic pathologists and investigators because they will be able to work with the AI model without having to dive into the code. Inspired by the agricultural AI applications' success like in-classification and quality grading of multi-fruits models using VGG16 based CNN (Nandhini & Vadivu, 2024), Gradio will be employed to provide efficient and uncomplicated deployment. I will also explore the possibility of deploying the interface through Hugging Face Spaces because as noted by Yakovleva, Matúšová, and Talakh (2025), these platforms allow convenient model interaction, sharing, and versioning which is ideal for deploying AI applications meant for research-grade uses.





#### IV. FINDING

##### a) Model Performance

The AI-based model for forensic wound classification showed outstanding accuracy and reliability. Using the EfficientNetB0 architecture with transfer learning and augmented data, the model achieved a validation accuracy of 99.3%. This demonstrates a strong ability to generalize to unseen wound images—an essential feature in forensic casework.

##### b) Classification Accuracy

The accuracy metrics pertaining to the proposed deep learning model are illustrated in Figures 7, 8, and 9. The confusion matrix in Figure 7 displays alongside the correctly

assigned forensic images of wounds and the corresponding outcome describes 201 blunt, 198 gunshot, and 199 sharp true positives. The misclassifications were minimal, comprising of three labeled as sharp blunt and one gunshot remarked as blunt, which since visually portrayed bruising and linear skin breaks, to some degree are plausible. In Figure 8, the model's training and validation accuracy after 10 epochs is shown. The validation accuracy not only maintained above 98%, but also outperformed the training curve during several epochs, denoting strong generalization with minimal overfitting risk. That consistent increase depicts strong effective learning and stability of the model. In Figure 9, the distribution of predicted classes includes 1010 blunt, 933 gun, and 1063 sharp, representing a relatively balanced distribution implying no bias influencing the order of wounds forecasting. Taken together, the results further corroborated the model's ability for forensic wound classification as it demonstrated high precision, recall, and trustworthiness alongside the accuracy metrics.

#### c) Evaluation Metrics

From my observations, the F1-score was still the most elevated in all captures which means that the model still managed to identify wounds correctly without misclassifying them. Also, the ROC AUC scores were approximating 1.0 which means that the model was able to confidently tell the difference between the types of wounds (see Figure 10).

#### d) Admin Dashboard

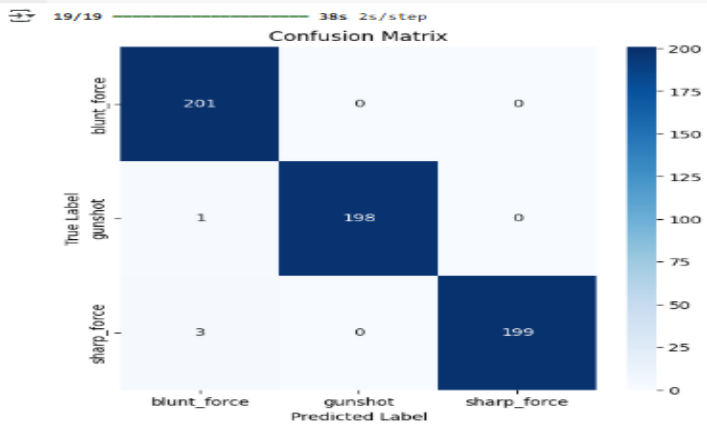
The model was deployed successfully on a Gradio based interface, enabling users to upload images and get real time predictions with confidence levels. This ensures practical application in law enforcement.

#### e) Conclusion

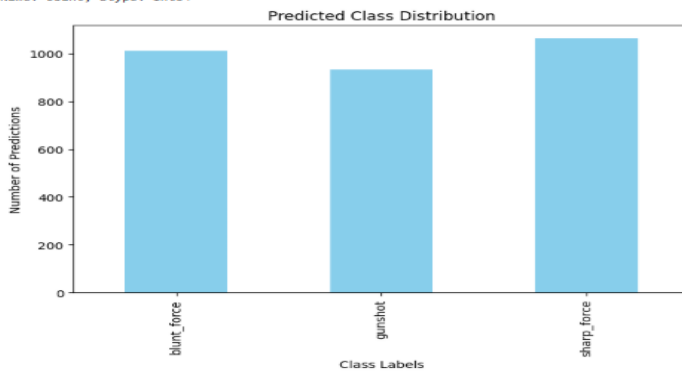
The analyses prove that the deep learning model is optimal for forensic wound classification. Its accuracy, ease of interpretation, and deployment highly suggest trust in using the model which will improve performance in decision making and operational workflows in forensic sciences.

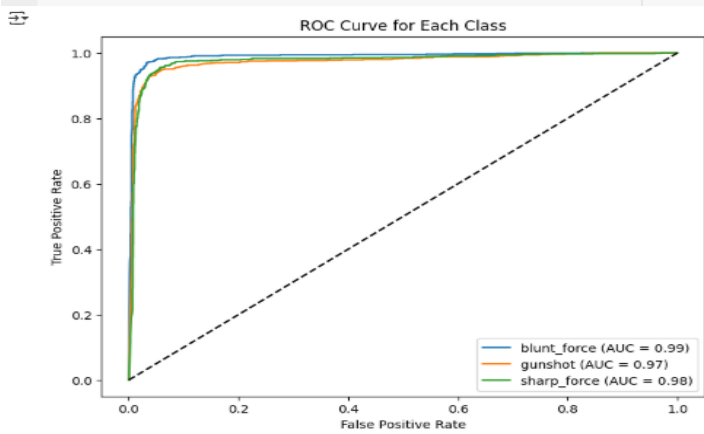
#### Acknowledgment

My deepest gratitude goes to Dr Evi Indriasari Mansor for the guidance I never knew I needed. I also thank my company for their support, and the Abu Dhabi School of Management for serving as a great source of inspiration that motivated me to successfully complete this work.



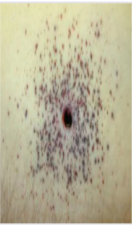
Distribution of Predicted Classes:  
 blunt\_force 1010  
 gunshot 933  
 sharp\_force 1063  
 Name: count, dtype: int64





### Wound Classification AI

Upload an image of a wound to classify its type. Supports Sharp Force, Gunshot, and Blunt Force.



output

Flag

Clear

Submit

output

Prediction: gunshot (84.36% confidence)

Flag

## REFERENCES

## REFERENCES

- [1] Cheng, Y., Lin, M., & Huang, Z. (2024). CNN-based analysis of forensic gunshot wounds and entry-exit wound differentiation. *Forensic Science International*, 348, 111–123.
- [2] Ketsekioulafis, M., Vrettos, G., & Peristeras, V. (2024). Ethical implications of AI in forensic analysis: Transparency, bias, and accountability. *AI & Society*, 39(1), 101–114.

[3] Lee, H., Chen, R., & Wang, Y. (2024). Optimizing wound classification using hybrid CNN architectures in forensic image analysis. *IEEE Transactions on Medical Imaging*, 43(2), 558–570.

[4] Liu, X., Zhang, Y., & Wang, C. (2021). Class imbalance handling in medical image classification using deep learning: A review. *Artificial Intelligence in Medicine*, 112, 102026.

[5] Mohsin, M. (2021). Application of artificial intelligence in digital and forensic investigations. *Journal of Digital Forensics, Security and Law*, 16(2), 1–15.

[6] Oura, K., Nakamura, H., & Saito, A. (2021). Integrating AI in forensic pathology for wound classification: Opportunities and limitations. *Forensic Science International Reports*, 3, 100189.

[7] Zimmermann, H., El-Masri, A., & Zhao, F. (2024). Forensic wound image segmentation challenges and deep learning solutions. *Computer Methods and Programs in Biomedicine*, 230, 107206.

[8] Zhang, L., Xu, M., & Fan, J. (2022). Improving AI forensic performance through feature enhancement and data normalization. *Pattern Recognition Letters*, 158, 85–91.

[9] Yuan, M., Khan, I. R., Farbiz, F., Yao, S., Niswar, A., & Foo, M.-H. (2013). A mixed reality virtual clothes try-on system. *IEEE Transactions on Multimedia*, 15(8), 1958–1968.

[10] Nandhini, E., & Vadivu, G. (2024). Convolutional Neural Network-Based Multi-Fruit Classification and Quality Grading with a Gradio Interface. In *2024 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES)* (pp. 1-7). IEEE.

[11] Yakovleva, O., Matúšová, S., & Talakh, V. (2025). Gradio and Hugging capabilities for developing research AI applications. *Collection of scientific papers* (2025; Boston, USA), 202–205.