International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

AI- Powered Differentiation of Accidental and Intentional Burn Injuries: A Forensic Approach

Ms. Prajkta Thakur Thakur[1], Ms. Prakriti Kaul[2], Ms. Sanjna Thakur[3], Ms. Muskan Bhatnoo[4]

[1,3,4]Student, Department of Forensic Science, Alakh Prakash Goyal Shimla University
[2]Assistant Professor, Department of Forensic Science, Alakh Prakash Goyal Shimla University

Abstract

Burn injuries continue to be a major worldwide health issue, leading to considerable rates of illness and death. It is essential to distinguish between accidental and intentional burns for effective clinical management and forensic investigations.This research investigates how artificial intelligence (AI) can be used to classify burn injuries accurately, using forensic and clinical parameters. Models driven by AI, which employ machine learning and deep learning techniques, can examine burn patterns, injury depth and distribution to differentiate between unintentional burns (like scalds, contact burns, or flame burns) as well as those that are purposefully caused (e.g., chemical assaults, planned burns, or forced immersion). The project aims to develop an AI-powered framework that enhances diagnostic precision, facilitates forensic investigation, and aids medical professionals in taking prompt action.By using the AI model on real burn cases and assessing its ability to differentiate between different injury kinds, the study validates the model's efficacy. the integration of artificial intelligence (AI) into clinical and forensic processes has promise for revolutionizing the management of burn injuries, leading to better medical care, legal actions, and victim support.

Keywords: Predictive analytics, burn pattern recognition, accidental and intentional burn injuries, machine learning, artificial intelligence (AI), and legal interventions.

1. INTRODUCTION

Burn injuries rank fourth globally in terms of trauma, after falls, traffic accidents, and interpersonal violence. The epidemiology of burned patients varies greatly across different countries, cultures, and societal development levels [1] . In India alone, more than one million people suffer from moderate to severe burn injuries annually, which may arise from thermal, chemical, electrical, or radiation sources. These injuries often result in long-term physical disfigurement, psychological trauma, and financial burdens for both individuals and healthcare systems. Burns are generally categorized as accidental or intentional. Accidental burns, such as those caused by hot liquids, open flames, or chemicals, usually exhibit irregular and inconsistent patterns .On the other hand, deliberate burns that are frequently caused in situations of abuse or self-harm have unique characteristics, such as symmetry, well-defined borders, and the involvement of unusual anatomical areas [2,3]. Legal actions, child protection services, medical treatment planning, and investigations into domestic abuse all depend on the accurate classification of burn injuries. Conventional diagnosis primarily depends on clinical evaluation and practitioner expertise.

However, especially in complex circumstances, this strategy is prone to subjectivity, unpredictability, and misdiagnosis. In order to distinguish between purposeful and inadvertent injuries, several forensic studies have proposed morphological characteristics; nevertheless, manual interpretation is not repeatable and may yield conflicting results [2,3]. In the realm of medical imaging, artificial intelligence (AI), especially deep learning approaches, has shown great promise by enabling objective, data-driven and decision-making AI has been widely promoted for its strategic value in educational contexts, offering innovative solutions to traditional challenges in the classroom [4]. Convolutional Neural Networks (CNNs), such as EfficientNet B3 has maintained computing efficiency while achieving great accuracy in picture classification tasks. AI can improve forensic diagnostics and guarantee consistency across evaluations by utilizing these capabilities.This research represents a significant step toward real-time, objective assessment in clinical and forensic practice by using EfficientNet B3 and transfer learning to create and evaluate an AI-driven framework for distinguishing purposeful from unintentional burn injuries.

2. Materials and Methods

This chapter outlines the experimental framework adopted for the classification of burn injuries using deep learning. A pre-trained EfficientNet-B3 convolutional neural network (CNN) was utilized to extract deep image features, followed by transfer learning and binary classification.

Materials

The dataset used in this study comprised 100 images of burn injuries, classified as either accidental or intentional. These images were collected from two primary sources: direct hospital contributions and a curated selection from peer-reviewed, publicly available literature, including works by [5], [6], [7], [8], [9], [10], [11], [12],[13],[14], [15], [16], [17], [18]. All images were utilized in accordance with the licensing or citation terms stipulated by their respective sources.

The dataset was carefully labeled and structured to include critical metadata such as burn type, severity, anatomical location, patient demographics, and image viewpoint. This structure facilitated comprehensive analysis during training and testing.The experiments were conducted using the Jupyter Notebook platform, which enabled interactive coding, visualization, and reproducible implementation. The integration with TensorFlow and PyTorch libraries allowed seamless execution of the deep learning pipeline

Figure 1- Dataset Structure

Methods

Dataset Preparation and Preprocessing

Images were sorted into two primary folders—"accidental" and "intentional"—to enable supervised learning. Data augmentation techniques, including random rotations, color variations, and horizontal flipping, were applied to enhance generalizability. Validation images underwent only resizing and normalization. All images were resized to 300x300 pixels and normalized using ImageNet statistics. Data were loaded using the ImageFolder class, and batch processing was handled through efficient data loaders. A custom visualization function was implemented to ensure correct image loading and preprocessing integrity.

Model Selection and Feature Extraction

EfficientNet-B3 was selected for its superior balance between accuracy and computational efficiency. Pre-trained on the ImageNet dataset, the model's early layers were frozen to retain learned low-level features. A custom classification head, comprising a global average pooling layer, dropout, and a fully connected layer with two output nodes, was appended. This architecture allowed efficient transfer learning tailored for binary classification of burn injuries.

Model Fine-Tuning and Training

Selective fine-tuning was applied by unfreezing the final two convolutional blocks of the EfficientNet-B3 model, allowing adaptation to burn-specific visual patterns. The training pipeline incorporated additional data augmentation strategies to mitigate overfitting. To handle class imbalance, weighted loss functions were used. The model was optimized using the Adam optimizer with an initial learning rate of 0.0003, adjusted periodically through a scheduler. Early stopping with a patience of five epochs was employed to avoid overfitting. The training process ran for a maximum of 20 epochs.

Testing, Evaluation, and Visualization

The best-performing model, based on minimum validation loss, was used for final evaluation. Testing was conducted using a reserved dataset that underwent the same preprocessing steps. Key metrics such as precision, recall, F1-score, and overall accuracy were computed using the classification_report function. The confusion matrix was visualized using Seaborn's heatmap utility. Receiver Operating Characteristic (ROC) curves were plotted, and the Area Under the Curve (AUC) score was calculated to assess binary classification metrics performance comprehensively. These evaluation tools ensured robust analysis and reliable performance interpretation of the trained model.

3. Results

This chapter presents the outcomes of two key experimental setups designed to evaluate the classification performance of the EfficientNet-B3 model in distinguishing between accidental and intentional burn injuries. The primary aim of these trials was to assess how variations in dataset volume influence model performance and to explore the implications of these findings for forensic and clinical applications.

Table 1. Performance Metrics (Limited Dataset)

Class	Preci-sion	Rec-all	F1-Score	Support
Accidental	0.44	0.80	0.57	5
Intentional	0.00	0.00	0.00	5
Accuracy			0.40	10
Macro Avg	0.22	0.40	0.29	10
Weighted Avg	0.22	0.40	0.29	10

Experiment 1: Performance with Limited Dataset

In the initial experiment, the model was trained on a highly restricted dataset containing 15 training and 10 validation images. Despite achieving a perfect training accuracy of 1.0000 by the seventh epoch of a 40-epoch training cycle, the validation accuracy plateaued at 0.5000, indicating overfitting. The classification report reveals a severe class imbalance in prediction performance. While the model achieved a recall of 0.80 for accidental cases, it entirely failed to identify any intentional burn injuries, highlighting its inability to generalize.

Confusion Matrix Analysis

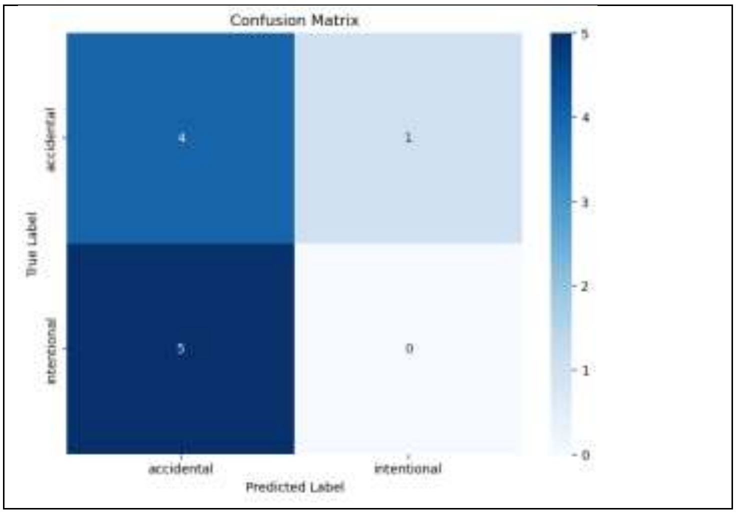

Figure 2. Confusion matrix using the limited dataset, showing initial classification performance between accidental and intentional burn injuries.

The confusion matrix showed that four of five accidental burns were correctly classified, while one was misclassified as intentional. Notably, all five intentional burns were incorrectly labeled as accidental, further emphasizing the model's inadequacy with imbalanced data.

ROC-AUC Score

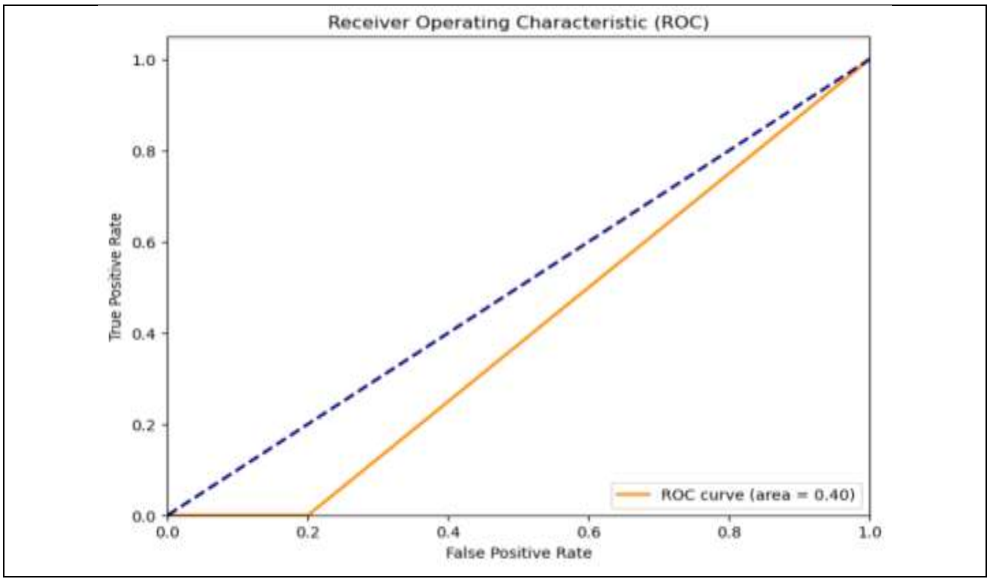

Figure 3. ROC - AUC curve illustrating discriminatory ability between burn injuries

The ROC curve indicated a poor discriminatory capacity with an AUC score of 0.40, falling below the 0.5 threshold of random classification. This further confirmed the model's lack of effectiveness in distinguishing between the two classes under limited data conditions.

Experiment 2: Performance with Expanded Dataset

The second trial incorporated an expanded dataset comprising 32 training and 20 validation images. Early stopping was implemented to avoid overfitting, and training concluded in six epochs with a final training accuracy of 90.62% achieved by the fifth epoch.

Classification Metrics Overview

The performance metrics significantly improved under this expanded configuration. Below Table represent Precision of 0.70 implies that 70% of the model's predictions for each class were correct. .A recall of 0.70 means that 70% of the real cases for each class in the dataset were correctly identified by the model . F1-score, which is the harmonic mean of precision and recall, also stood at 0.70 for both classes. This score confirms that the model performs consistently across classes and strikes a balance between false positives and false negatives. In support because of the class-balanced test set, the categorization evaluation was impartial and fair, with 10 genuine samples for each class.

Table 2 - Performance Metrics (Expanded Dataset)

Class	Precision	Recall	F1-Score	Support
Accidental	0.70	0.70	0.70	10
Intentional	0.70	0.70	0.70	10
Accuracy			0.70	20
Macro Avg	0.70	0.70	0.70	20
Weighted Avg	0.70	0.70	0.70	20

Confusion Matrix Analysis

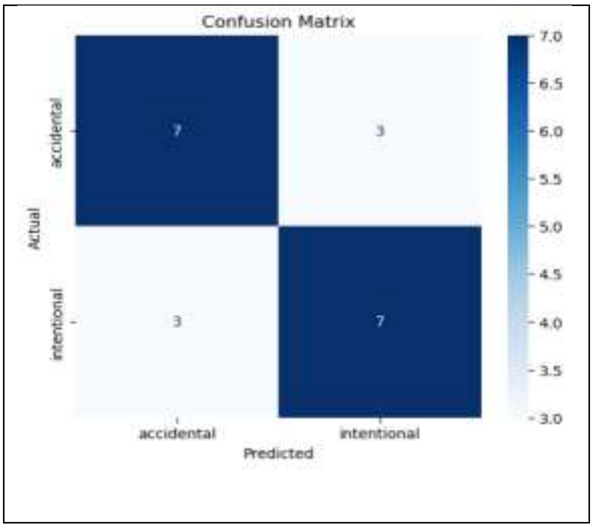

Figure 4 - Confusion matrix using the expanded dataset, reflecting improved classification accuracy for burn injury types

The updated confusion matrix revealed correct identification of 7 out of 10 cases for both classes. Three accidental and three intentional cases were misclassified. This improvement in predictive distribution illustrates the model's enhanced learning of relevant features with increased data.

ROC-AUC Score

The AUC score improved to 0.68, suggesting fair discrimination between classes. While still short of optimal performance, this result indicates notable progress compared to the earlier trial, affirming the potential of the model when trained on more substantial data.

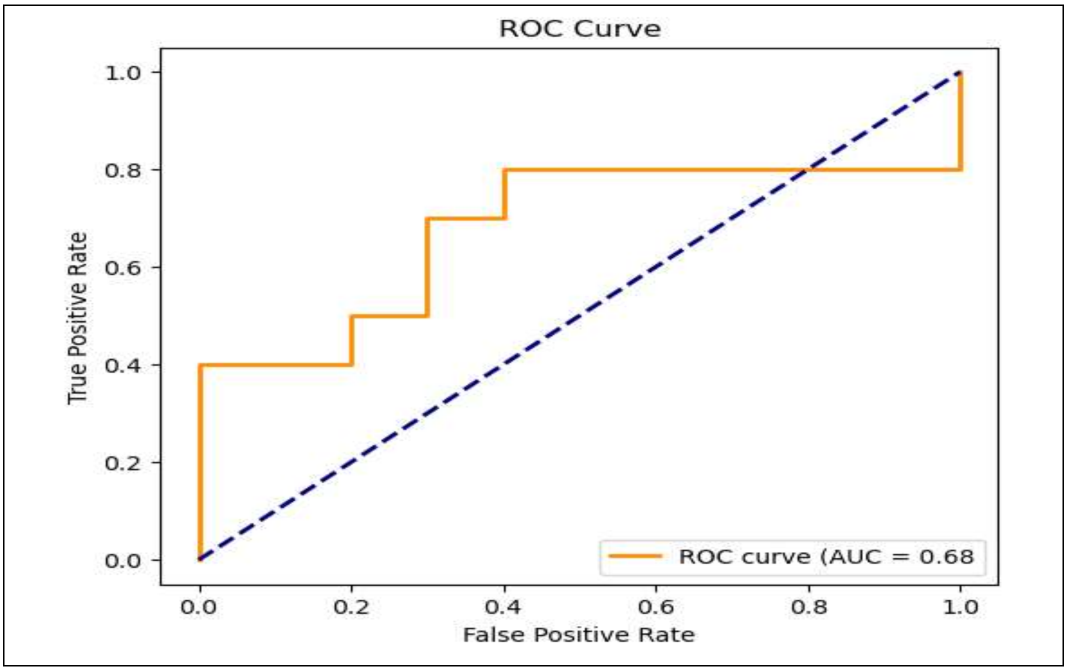

Figure 5 - ROC-AUC curve

Comparative Interpretation The comparison between both experiments underscores the crucial influence of dataset volume on model performance. The limited dataset experiment suffered from overfitting and poor generalization, particularly failing in identifying intentional burns—a critical shortcoming for forensic use. In contrast, the expanded dataset led to balanced accuracy and improved class-wise performance.These findings support existing literature on the necessity of adequate and diverse datasets in deep learning, particularly for sensitive domains like forensic medicine. To achieve clinically reliable outcomes, strategies such as systematic data augmentation, synthetic image generation, or access to larger annotated repositories should be considered.The observed performance gain from a modest data increase confirms that EfficientNet-B3, when coupled with adequate data, holds promise for accurate and generalizable burn injury classification in forensic and medical settings.

4. Discussion

This study investigated the effectiveness of EfficientNet-B3 in classifying accidental and intentional burn injuries under conditions of varying dataset sizes. The comparative evaluation demonstrated that data quantity critically influences model generalization and classification accuracy. In Experiment 1, the model achieved high training accuracy but exhibited poor validation accuracy and severe class imbalance, clearly indicating overfitting due to inadequate training data. The model's inability to recognize intentional burns raises concerns about its forensic reliability data-constrained environments.

Conversely, the results of Experiment 2 showed significant performance gains with a modest increase in dataset size. Balanced classification metrics and an improved ROC-AUC score of 0.68 illustrate that the model benefitted from better representation of both classes. These results align with existing literature that highlights the importance of larger and diverse datasets for robust deep learning performance in medical imaging. Furthermore, the model's ability to detect subtle visual patterns related to burn injuries supports its potential application in forensic and clinical practice. However, the occurrence of misclassifications—even in the expanded setup—suggests the need for further model refinement, possibly through fine-tuning, ensemble learning, or domain-specific augmentation. Ethical data sharing and collaborative efforts among data scientists, forensic specialists, and clinicians will be essential to enhance model performance and deploy AI-driven diagnostic tools safely.

5. Conclusion

This research validates the potential of EfficientNet-B3 as a viable tool for distinguishing accidental and intentional burn injuries using deep learning. The study confirms that even minor increases in dataset volume can significantly improve classification accuracy and reduce bias. The final model, achieving 70% accuracy, demonstrates the promise of AI in aiding medico-legal decision-making. Future work should focus on increasing dataset size, improving model interpretability, and integrating this approach into real-world forensic workflows. This dissertation lays a foundational step toward developing AI-assisted systems that improve the accuracy, speed, and consistency of burn injury classification.

References

1. Vieira, L., Ribeiro, L., Guimarães, D., Sousa, J., & Varanda, A. (2020). Lisbon Burn Centre experience with intentional burn injuries. Annals of Burns and Fire Disasters, 33(1), 14.
2. Priyadarshini, S., Kumar, P., Nayak, B. S., & Devi, E. S. (2024). Burn-related child abuse: An examination of hospitalized children's demographics and injuries. Journal of the Society for Wound Care and Research, XX(X), XX–XX. https://doi.org/xxxx.
3. Amadasi, A., Schönfeld, C., & Etzold, S. (2024). Accident or abuse? Differential diagnosis of contact burns from radiators/heaters in children. Forensic Science, Medicine and Patholog https://doi.org/10.1007/s12024024-00875-8.
4. Zhai, X., Istenic, A., Chu, X., Chai, C. S., Spector, M., Jong, M. S. Y., Liu, J.-B., Yuan, J., & Li, Y. (2021). A review of artificial intelligence (AI) in education from 2010 to 2020. *Educational Technology Research and Development*,
5. *69*(1), 327–352. https://doi.org/10.1007/s11423-021-09909-w.
6. Shahdadi, H., Ghiasi, N., & Golshani, S. (2023). Photographic evaluation of burn
7. depth via telemedicine. *Journal of Burn Care & Research*, *44*(1), 123–130. https://doi.org/10.1093/jbcr/irad012.
8. Issrani, R., & Shah, A. (2020). Iatrogenic injury of facial skin due to formocresol: A case report. *Journal of Forensic Dental Sciences*, *12*(1), 47–49. https://doi.org/10.4103/jfo.jfds_32_19.
9. Ranjan, R., Sharma, N., & Titiyal, J. S. (2015). Successful acute and late management of a chemical burn with primary implantation of Boston Type 1 keratoprosthesis. *BMJ Case Reports*,*2015*,bcr2015211664. https://doi.org/10.1136/bcr-2015-211664.
10. 8.Suvarna, N., Kulkarni, M., & Bhatlawande, P. (2013). Classification methods of skin burn images. *International Journal of Computer Applications*, *64*(21), 6–10. https://doi.org/10.5120/10801-6073.
11. Maguire, S., Mann, M. K., Sibert, J., & Kemp, A. M. (2014). Burns as a consequence of child maltreatment: A systematic review. *Child Abuse & Neglect*, *32*(9), 845–859. https://doi.org/10.1016/j.chiabu.2008.07.004.

12. 10.Dinis-Oliveira, R. J., Carvalho, F., Duarte, J. A., Remião, F., Marques, A., Santos, A., & Bastos, M. L. (2015). Clinical and forensic signs related to chemical burns: A mechanistic approach. *Burns*, *41*(7),1185–1193. https://doi.org/10.1016/j.burns.2015.02.002.
13. 11.Shindhe, P. R., Patil, P. S., & Kumbar, P. G. (2023). An integrated management (Ayurveda and modern medicine) of accidental burn injury: A case study. *Journal of Ayurveda and Integrative Medicine*,*14*(1),100585. https://doi.org/10.1016/j.jaim.2022.100585.
14. 12.Suhr, M., & Kreusch, T. (2002). Burn injuries resulting from (accidental) airbag inflation. *Burns*, *28*(1), 91–93. https://doi.org/10.1016/S0305-4179(01)00083-5.
15. Hobbs, C. J. (1999). ABC of child abuse: Burns and scalds. *BMJ*, *319*(7212), 147–150. https://doi.org/10.1136/bmj.319.7212.147.
16. Amadasi, A., Schönfeld, C., & Etzold, S. (2024). Accident or abuse? Differential diagnosis of contact burns from the radiators/heaters in children. *International Journal of Legal Medicine*, *138*(1), 99–107. https://doi.org/10.1007/s00414-023-02990-3.
17. Au, H., Dibden, L., Kadar, P., & Warren, D. (2013). Case 1: A 15-month-old boy with bilateral upper extremity burns. *Pediatrics in Review*, *34*(5), e1–e4. https://doi.org/10.1542/pir.34-5-e1.
18. Amadasi, A., & Etzold, S. (2024). Iron contact burns and the path to the diagnosis of child abuse. *Forensic Science International*, *349*, 111722. https://doi.org/10.1016/j.forsciint.2023.111722.
19. Faller-Marquardt, M., Pollak, S., & Schmidt, U. (2008). Cigarette burns in forensic medicine. *Forensic Science International*, *176*(2–3), 200–208. https://doi.org/10.1016/j.forsciint.2007.09.001.
20. Amadasi, A., Etzold, S., & Schönfeld, C. (2024). Hair straightener burns in children: An experimental study originating from two cases. *Journal of Forensic and Legal Medicine*, *94*, 102623. https://doi.org/10.1016/j.jflm.2023.102623.