A Secure Federated Hybrid Deep Learning Model with Blockchain-Assisted Aggregation for Skin Cancer Diagnosis: A Review

Naresh Alapati[1,2], BVVS Prasad[3]

[1]Lincoln University College, Petaling Jaya, Selangor, Malaysia.

[2]Department of Computer Science and Engineering, Vignan's Nirula Institute of Technology and Science for Women, Peda Palakaluru, Guntur, India.

[3]School of Engineering (CSE), Anurag University, Venaktapur, Ghatkesar, Hyderabad, Telangana, India

E-mail:[1]pdf.Naresh@lincoln.edu.my

Abstract

Skin cancer is one of the most common and fastest-growing cancers worldwide. When diagnosed late, can a serious health threat, whether it's melanoma or non-melanoma skin cancer. In the medical imaging domain, deep learning has revolutionized the way medical images are analysed, with the ability to automatically extract discriminative features from dermoscopic images with impressive accuracy. The existing centralised solutions, however, demand a large number of annotated datasets, which are generally hard to access due to strict data privacy laws (HIPAA, GDPR) and the distribution of Data across geographically widely spread healthcare institutions. The current models are Ill-Generalizable in the presence of Variations in imaging conditions, biased towards common lesion types due to class imbalance, and susceptible to adversarial and poisoning attacks in shared computing infrastructures. The complementary and synergistic solutions are recent advances in Federated learning (FL), Hybrid deep learning architectures and Blockchain Technology. FL provides privacy-preserving collaborative training without moving data; the hybrid architectures benefit from the strengths of various architectures for better and more robust feature representation; blockchain ensures tamper-resistance and auditable aggregation of distributed model updates. This review systematically summarises the current state of the art in each of these areas, critically appraises the strengths and weaknesses, identifies research gaps and proposes an integrated conceptual framework for an accurate, scalable, real-world, multi-institutional secure skin cancer diagnosis.

Keywords—Skin Cancer, Federated Learning, Hybrid Deep Learning, Blockchain, Medical Imaging, Privacy Preservation, Transfer Learning.

I. Introduction

Skin cancer is one of the most prevalent and fastest-growing cancers worldwide with more than 1.5 million cases of diagnosis being reported each year by the World Health Organisation. Melanoma is the rarest of the three common skin cancers but it is the most dangerous, and it is responsible for most deaths from skin cancer in its Later stages. The importance of early detection for timely diagnosis is consistently shown to have significantly increased 5-year survival rates, from about 20% to over 95%. Traditionally, the diagnosis was done by dermoscopy, visual inspection and histopathological biopsy, which are all subject to time-consuming and require dermatological expertise, thus with high inter-observer variability.

Recently, Deep learning (DL) and specifically Convolutional Neural Networks (CNNs) have revolutionized automated skin lesion analysis. Architectures like EfficientNet, DenseNet, ResNet, and InceptionNet have been shown to be more efficient and effective for hierarchical feature extraction and lesion classification, and several benchmarking evaluations have confirmed CNN accuracy at least when compared to board-certified dermatologists in controlled conditions. However, there are number of important challenges to deployment in the real world: First, large, diverse, centralised data is needed for effective model training, but this is difficult because of data privacy regulations. Second, class imbalance in the data sets available results in poor performance in rare but clinically important lesion categories. Third, shifts across amazing devices and clinical protocols negatively impact performance in the real world. Fourth, centralised data repositories are vulnerable to security issues such as data leakage and adversarial attacks.

The research community has been looking to three transformative paradigms to solve these interconnected challenges. Federated learning (FL) allows multiple institutions to train the same model without sharing the raw data, but instead sharing only the model parameters, thus maintaining patient privacy. The Hybrid deep learning (HDL) architecture is a new one that uses a combination of complementary models to capture a wide range of features representations and improve overall classification robustness, as compared to single-architecture approaches. Blockchain Technology offers an immutable and decentralized ledger to record updates to the models, which ensures audits and prevents any model poisoning or Byzantine attack by any of the participating clients. Although some recent developments have been made in all these areas, previous studies have not brought together all three into a unified skin cancer diagnosis pipeline. The aim of this review is Therefore to: (1) overview the latest developments in the field of AI and dermatology, Federated learning, and blockchain integrated healthcare;(2) critically compare existing works on several key points;(3) list the open research Questions and (4) suggest a framework that aims to be integrated secure and scalable.

II. Literature Survey

This section summarizes recent advances in the areas of skin cancer classification, federated learning in medical Imaging, hybrid deep learning architectures and AI systems integrated with blockchain technology in healthcare. Table 1 shows a tabular comparison of the works reviewed with respect to five important evaluation parameters.

Table 1. Comparison of related works and this work across key parameters

Reference	Privacy-Preserving FL	Hybrid DL Architecture	Blockchain Integration	Rare Class Handling
[1] Yang et al., 2025	No	No	No	No
[2] Aboulmira et al., 2024	No	Yes	No	No
[3] Farea et al., 2024	No	Yes	No	Partial
[4] Balaha et al., 2023	No	Yes	No	No
[5] Chen et al., 2022	Yes	No	No	No
[6] Pacheco et al., 2020	No	Yes	No	No
[7] Barata et al., 2019	No	No	No	No
[8] Haenssle et al., 2018	No	No	No	No
[9] Nguyen et al., 2022	Yes	No	No	No
[10] Kumar et al., 2024	Yes	No	Yes	No
This Work	**Yes**	**Yes**	**Yes**	**Yes**

The traditional machine learning methods such as Artificial Neural Networks (ANN), Support Vector Machines (SVM) and K-Nearest Neighbours (KNN) have been extensively reviewed for skin disease classification by Yang et al. [1]. Although these methods showed promise for automatic diagnosis, they were very sensitive to noise and imaging artifacts, had limited ability to deal with high dimensional dermoscopic patterns and performed poorly in multi-class cases where the class distribution was imbalanced.

Aboulmira et al. [2] introduced an ensemble of the EfficientNet modules (B1 to B5) trained using noisy student self-training for the design of SkinHealthMate: a web and mobile dermatology system. The framework performed at state-of-the-art accuracy on ISIC benchmarks, was robust to changes in image quality, but failed to achieve good performance on rare skin conditions and the centralized training paradigm posed high data privacy concerns in multi-institutional contexts.

Farea et al. [3] Proposed a hybrid deep learning systems which synergically integrated DenseNet, ResNet, and Inception architecture with artificial bee colony (ABC) metaheuristic optimisation for automatic hyperparameter tuning. This multi-architecture fusion enhanced feature extraction richness and increased classification accuracy on the ISIC dataset but had high computational overhead and long training time, which was not practical for deployment in resource-limited clinical settings.

Balaha et.al. [4] designed a transfer learning model with the addition of spatial self-attention(SpaSA) mechanisms to guide model attention towards clinically relevant dermoscopic features. This method boosted the accuracy of the classification and the interpretability of the results, but was still very data-and computation- intensive relaying heavily on pre-trained weights and demanding sophisticated computational resources for training and inference.

In order to address the challenge of highly personalizable data, Chen et al. [5] introduced a personalized Federated learning framework that combines progressive Fourier aggression (PFA) for Frequency-domain Model Aggregation, Deputy-Enhanced Transfer (DET) for cross-client knowledge transfer, and Conjoint Prototype Alignment (CPA) for feature space harmonization. The framework showed better scalability for the various data distributions but came with high hyperparameter tuning cost and aggregation complexity.

Pacheco et al. [6] investigated combinations of CNNs With structured clinical metadata with late-fusion, obtaining as much as 5 – 7% improvement in classification accuracy over image only baselines, but sensitive to missing clinical attributes and challenging to distinguish visually similar lesion classes. Hierarchical classification using lesion taxonomy was introduced by Barata et al. [7] to enhance interpretability and Haenssel et al. [8] demonstrated the viability of using CNNs for Melanoma detection at dermatologist level of accuracy, Albeit within a limited training set homogeneity and the lack of Federated training. Even with persistent communication overhead challenges, Nguyen et al. [9] have shown FL's ability to provide privacy-preserving medical image analysis. Kumar et al. [10] Introduced blockchain based Federated learning that guarantees audibility and accountability of healthcare data, although it faces challenges due to its limited transaction speed and complexity.

III. Research Gaps and Motivation

Although significant advances have been reported in the literature review, the following six key unanswered issues inspire the proposed integrated framework:

- No unified integrated frameworks: Current literature only deals with FL, hybrid DL, or blockchain separately. None of the studies to date brings all three together in a single skin cancer diagnosis pipeline, leaving the privacy-accuracy-security trade-offs suboptimal and the synergistic benefits missed.
- Lack of privacy-preserving mechanisms: most high-accuracy models are based on the centralised training paradigm, which is not suitable for HIPAA and GDPR. There are federated approaches, but they are not frequently used in combination with competitive diagnostic performance on complex thermoscopic datasets, and do not provide formal guarantees of differential privacy.

- Poor handling of rare disease classes: This is due to the clinical imbalance of rare classes of lesions in the dataset, which current model architectures and training strategies are not sufficiently able to handle.
- Advanced hybrid models and Federated aggregation protocols come with significant computational and communication requirements, which are not suitable for resource-constrained clinical settings where there is limited bandwidth and limited GPU infrastructure.
- Poisoning gradient inversion and Byzantine Attacks: There are weak defences in most FL implementations. While a principled solution is blockchain-based aggregation, it has not been systematically used in FL contexts for skin cancer.
- Limited real-world validation: Most systems are validated under idealised conditions, and there is limited evidence of performance and reliability in multi-institutional deployments configured with various imaging equipment, patient populations and acquisition protocols.

IV. Proposed Framework Overview

This study introduced a concept of secure federated hybrid deep learning systems with blockchain-assisted aggregation, based on the identified research gaps. The framework combines four main elements to tackle privacy, accuracy, security and rare-class performance all at the same time.

The Hybrid Deep Learning Module (HDLM) Is a fusion of two feature extractions, EfficientNet-B4 and DenseNet-121, with the two feature extractions being parallel in order to process the dermoscopic image inputs. EfficientNet B4 is based on compound scaling in its depth, width, and resolution Dimensions to capture fine-grained spatial features, while DenseNet-121 is based on dense connectivity patterns to promote feature reuse and gradient flow. The maps of both streams are fused together through a multi-scale attention gate based on channel-wise squeeze and excitation and spatial attention, which focuses on the diagnostically salient parts of the dermoscopic image, such as typical vascular patterns, asymmetric pigment patterns and irregular borders, by downweighting irrelevant background information. The fused representation goes through fully connected layers, of which the first two have batch normalization, and the last two have dropout regularization ($p=0.4$) to yield class probabilities for nine standard ISIC lesion categories.

The Federated Learning Protocol is a way to have N hospitals throughout the world train the global model without sharing patient data such as images or clinical records. Every client executes T local training epochs On its own local data set, The aid of differential privacy mechanisms such as Gaussian noise injection combined with adaptive gradient clipping that is calibrated to guarantee ε-differential privacy. The central aggregation server at the end of each communication round will only receive noise-protected weight gradients. The FedAvg Algorithm takes advantage of a weighted distribution based on the amount of data contributed by each client (institutional data size), allowing for balanced participation from large academic centres and smaller clinical Practises. The blockchain-assisted aggregation component saves the gradient submission made by each client and the aggregated global model parameters as cryptographically signed, Immutable transactions on Permission Hyper Ledger fabric blockchain. Smart contracts automatically verify the integrity of updates identify and punish statistical outliers that may be signs of poisoning attacks, require a minimum number of participants for each round and complete the aggregation once all the client updates are received. The training goal is formulated using a Focal Loss function with class weights inversely proportional to their frequency, and systematic oversampling of rare categories of lesions is applied using Conditional Generative Adversarial Networks (CGAN) prior to the local training rounds to tackle class imbalance.

V. Conclusion and Future Work

This review has critically surveyed the contributions on Federated learning, hybrid deep learning approaches, and blockchain adaptation in healthcare AI, systematically exploring the state of the art in the field of skin cancer diagnosis using deep learning. The analysis shows that even though progress in each of these directions is significant, with contemporary CNN-based systems reaching dermatologist-level accuracy, Federated learning enabled privacy-preserving collaborative learning, and blockchain enabling tamper-resistant auditability, no current work has the ability to combine all three paradigms in order to achieve the full objective of privacy-preserving, dermatologist-level, accurate, rare-class-sensitive, and system-secure collaborative Learning. The gap is huge for responsible deployment of AI support In dermatology in the real world on institutional scale.

The proposed integrated conceptual framework tackles the aforementioned challenges by seamlessly integrating the EfficientNet, DenseNet, hybrid feature extraction, federated learning, differential privacy, secure aggregation by blockchain and smart contract governance with the focal loss and CGAN augmentation for rare disease class balancing. The architecture generates a holistic and executable empirical research and system development agenda. Specific future research directions include: (1) testing the diverse multi-institutional Dermoscopic datasets (ISIC, HAM10000, and real clinical deployments) (2) Optimizing communication efficiency via gradient comparison and quantization, and asynchronous aggregation strategies (3) exploring alternative blockchain consumes mechanisms to enhance scalability and transaction throughput (4) extending to multimodal inputs that include structured clinical metadata and patient history and (5) developing post-hoc and inherent explainability methods for generating clinical interpretable diagnosis rationales that enable trustworthy human-AI collaboration in dermatological practice.

References

[1] G. Yang, S. Li, Q. Zhao, and W. Zhang, "A comprehensive review of ML and DL techniques for skin disease classification," IEEE Access, vol. 13, pp. 11234–11258, 2025.

[2] A. Aboulmira, R. Hachimi, and H. Lakhbab, "SkinHealthMate: Web and mobile AI for skin disease using EfficientNet ensemble," Expert Syst. Appl., vol. 241, p. 122683, 2024.

[3] E. Farea, A. Al-Qahtani, and S. Hameed, "Hybrid DL model with ABC optimization for skin disease classification," Biomed. Signal Process. Control, vol. 88, p. 105634, 2024.

[4] H. M. Balaha and A. E. Hassan, "Skin disease classification using transfer learning and spatial self-attention mechanisms," IEEE Access, 2023.

[5] Z. Chen, W. Liang, and K. Li, "Personalized FL with progressive Fourier aggregation and conjoint prototype alignment for skin lesion classification," IEEE Trans. Med. Imag., vol. 41, no. 12, pp. 3597–3610, 2022.

[6] A. G. C. Pacheco et al., "PAD-UFES-20: A skin lesion dataset composed of patient data and clinical images from smartphones," Data Brief, vol. 32, p. 106221, 2020.

[7] C. Barata, M. E. Celebi, and J. S. Marques, "Improving dermoscopy image classification using color constancy," IEEE J. Biomed. Health Inform., vol. 19, no. 3, pp. 1146–1152, 2015.

[8] H. A. Haenssle et al., "Man against machine: Deep learning CNN for dermoscopic melanoma recognition vs. 58 dermatologists," Ann. Oncol., vol. 29, no. 8, pp. 1836–1842, 2018.

[9] D. C. Nguyen, M. Ding, P. N. Pathirana, A. Seneviratne, J. Li, and H. V. Poor, "Federated learning for smart healthcare: A survey," ACM Comput. Surv., vol. 55, no. 3, pp. 1–37, 2022.

[10] S. Kumar, R. Patel, M. Singh, and A. Sharma, "Blockchain-enabled FL framework for secure healthcare applications," IEEE Access, vol. 12, pp. 45000–45015, 2024.

[11] P. Tschandl, C. Rosendahl, and H. Kittler, "The HAM10000 dataset of multi-source dermatoscopic images," Sci. Data, vol. 5, p. 180161, 2018.

[12] N. Codella et al., "Skin lesion analysis toward melanoma detection: ISIC 2018 challenge," arXiv:1902.03368, 2019.

[13] T. J. Brinker et al., "Skin cancer classification using CNNs: Systematic review," J. Med. Internet Res., vol. 20, no. 10, p. e11936, 2018.

[14] A. Mahbod et al., "Fusing fine-tuned deep features for skin lesion classification," Comput. Med. Imaging Graph., vol. 71, pp. 19–29, 2019.

[15] R. Sharma, S. Gupta, and A. Verma, "Hybrid DL with multi-scale feature extraction for skin disease," Expert Syst. Appl., vol. 215, p. 119433, 2023.