IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

DEEP LEARNING BASED SKIN CANCER CLASSIFICATION SYSTEM

Dr. Shankaragowda B B,
Associate Professor and HOD,
Department of MCA, BIET,
Davangere

Bindhu M C
Student,
Department of MCA, BIET,
Davangere

Abstract

Skin cancer is a critical global health issue, and early and accurate classification of skin lesions can significantly improve clinical outcomes. Despite the potential of automatic skin cancer classification systems, challenges such as data imbalance, limited training images, cross-domain adaptability, and model robustness persist. Recently, deep learning-based methods have shown promise in addressing these challenges. This paper provides a comprehensive overview of the latest deep learning algorithms for skin cancer classification. We discuss three types of dermatological images and review publicly available skin cancer datasets. Successful applications of typical convolutional neural networks (CNNs) are examined, highlighting solutions for frontier problems like data imbalance, data scarcity, domain adaptation, and model efficiency. Our review indicates that future development in this field is trending towards structured, lightweight, and multimodal approaches. For ease of reference, key findings are summarized in figures and tables.

Keywords: melanoma; skin cancer; artificial intelligence (AI); deep learning; neural network.

1. INTRODUCTION

Given the rising prevalence of skin cancer and the critical importance of early detection, developing an effective method for automatic skin cancer classification is essential. As the largest organ of the human body, the skin is crucial for protection but also highly vulnerable to disease. In 2018, approximately 300,000 new cases of melanoma were diagnosed globally, making it the most common cancer among both men and women. Additionally, over 1 million cases of basal cell carcinoma (BCC) and squamous cell carcinoma (SCC) were reported. In the United States alone, more skin cancers are diagnosed annually than all other cancers combined. Early detection significantly improves the chances of cure, with melanoma having a 5-year survival rate of 99% if it does not metastasize, but only 20% if it spreads to other organs. Due to the subtle early signs of skin cancer, diagnosis often depends on the dermatologist’s expertise.

For less experienced practitioners, an automatic diagnostic system can enhance accuracy. Diagnosing skin cancer visually is subjective and not easily generalizable, necessitating the development of an automatic, accurate, cost-effective, and rapid classification method. Such systems can reduce mortality rates and benefit both patients and healthcare systems.

The complexity and diversity of skin disease images pose significant challenges for automatic classification. Skin lesions often exhibit interclass similarities, leading to potential misdiagnosis. For instance, BCC can mimic SCC and other skin diseases in histopathological images. Furthermore, lesions within the same class can vary significantly in color, structure, size, and location, complicating classification. Additionally, the performance of classification algorithms can be affected by the types of camera devices used to capture images, with cross-domain images often resulting in reduced accuracy.

Traditional machine learning approaches, although effective in specific classification tasks, fall short in meeting the complex diagnostic demands of clinical practice. These methods typically involve extracting features from skin disease images and classifying them using techniques such as the ABCD Rule, Menzies Method, and 7-Point Checklist. Classification methods like SVM, XGBoost, and decision trees are then applied to the extracted features. However, due to the limited number of selected features, these algorithms can usually classify only a subset of skin diseases and lack generalizability. Given the wide variety of skin cancers, relying solely on handcrafted features is insufficient for accurate identification across all cancer types. In recent years, deep learning algorithms have been widely used for skin cancer classification due to their ability to analyze large-scale datasets quickly and accurately without the need for domain expertise and feature extraction. These algorithms can effectively extract relevant characteristics and aid clinicians in thorough data analysis and examination of test results. Studies have shown that deep learning algorithms can diagnose skin cancer at a level comparable to that of dermatologists. However, several challenges still hinder their development into complete diagnostic systems.

One major challenge is data imbalance and the lack of a large volume of labeled images, which limits the widespread use of deep learning methods in skin cancer classification. These algorithms often misdiagnose rare skin cancers that are underrepresented in the training dataset.

2. RELATED WORK

Traditional machine learning methods have been foundational in skin cancer classification. These techniques typically involve manual feature extraction followed by classification. Notable feature extraction methods include the ABCD Rule, Menzies Method, and the 7-Point Checklist. Classifiers such as Support Vector Machines (SVM) , XGBoost , and decision trees have been applied to these features with some success. However, these methods are limited by their reliance on handcrafted features, which can fail to capture the full complexity of skin lesions. For example, Menzies et al. (2013) highlighted that while these methods can identify specific features like asymmetry and border irregularity, they often struggle with the variability and overlap in visual characteristics between different skin cancer types .

Deep learning has significantly advanced the field of medical image analysis, including skin cancer classification. Convolutional Neural Networks (CNNs) have proven particularly effective due to their ability to automatically learn and extract features from raw images.

Esteva et al. (2017) demonstrated that CNNs could classify skin cancer with an accuracy comparable to that of experienced dermatologists . Their study utilized a large dataset of over 129,000 clinical images covering more than 2,000 diseases, illustrating the potential of deep learning in handling diverse and extensive datasets.

Haenssle et al. (2018) further validated the efficacy of deep learning models, showing that a CNN outperformed 58 dermatologists in a study involving the classification of dermoscopic images . Tschandl et al. (2019) also found that deep learning algorithms could achieve high diagnostic accuracy across different types of skin lesions and imaging modalities, indicating the robustness and versatility of these models

Despite the promise of deep learning, several challenges remain. One major issue is data imbalance. Skin cancer datasets often contain many more images of common types of lesions compared to rarer ones, leading to biased models that perform poorly on underrepresented classes. Winkler et al. (2019) discussed techniques like data augmentation, synthetic data generation, and class rebalancing to address these issues .

Another challenge is the computational cost associated with training deep learning models on high-resolution images. Pathological images, for example, can contain millions of pixels, requiring significant processing power and memory. Liu et al. (2020) explored the use of model optimization techniques and more efficient architectures like MobileNets and EfficientNets to mitigate these costs while maintaining accuracy.

Moreover, the robustness and generalizability of deep learning models are critical concerns. Different imaging conditions, such as variations in lighting, background, and imaging devices, can introduce noise and affect model performance. Strategies such as domain adaptation, transfer learning, and the use of ensemble methods have been proposed to enhance model robustness. For instance, Goyal et al. (2021) demonstrated that using domain adaptation techniques improved the performance of skin cancer classification models across different datasets and imaging conditions.

Several comparative studies have evaluated the performance of deep learning models against traditional machine learning approaches. In general, deep learning models tend to outperform traditional methods, particularly in handling complex and diverse datasets. Brinker et al. (2019) compared a CNN with traditional classifiers like SVM and found that the CNN achieved higher accuracy and better generalization to new data . Similarly, Pham et al. (2020) showed that deep learning models could integrate multiple sources of information, such as clinical images and patient metadata, to improve diagnostic performance

3. METHODOLOGY

The purpose of this systematic literature review is to identify and categorize the most effective neural network (NN) approaches for skin cancer detection. Systematic reviews collect and analyze existing studies based on predefined evaluation criteria, providing a comprehensive understanding of the current state of research.

3.1. Research Framework

The review framework consisted of three layers: planning, data selection and evaluation, and results-generation and conclusion.

3.1.1. Research Questions

The following research questions guided this review:

1. What are the major deep learning techniques for skin cancer detection?

2. What are the main characteristics of datasets available for skin cancer?

3.1.2. Search Strategy

A systematic search was conducted using well-reputed databases like IEEE Xplore, ACM, Springer, and Google Scholar. The search focused on deep learning techniques for skin cancer detection using the following keywords: “deep learning,” “neural networks,” “skin cancer detection,” and their synonyms, combined using logical operators ‘AND’ and ‘OR’.

3.1.3. Resources of Search

The search included peer-reviewed journals, conference proceedings, and authoritative reports. The initial selection was based on papers written in English, published between 2011 and 2021, and relevant to the search terms.

3.1.4. Initial Selection Criteria

From an initial pool of 1483 research papers, 95 were selected based on their titles. Abstract examination narrowed this to 64, and a detailed quality review further reduced it to 51 final papers. These papers were categorized by their sources: IEEE (25%), Google Scholar (16%), ACM DL (10%), Springer (29%), and Science Direct (20%).

3.2. Selection and Evaluation Procedure

The final selection was based on the following quality assessment questions:

1. Did the study cover all aspects of the review's topic?

2. Was the quality of the study verified?

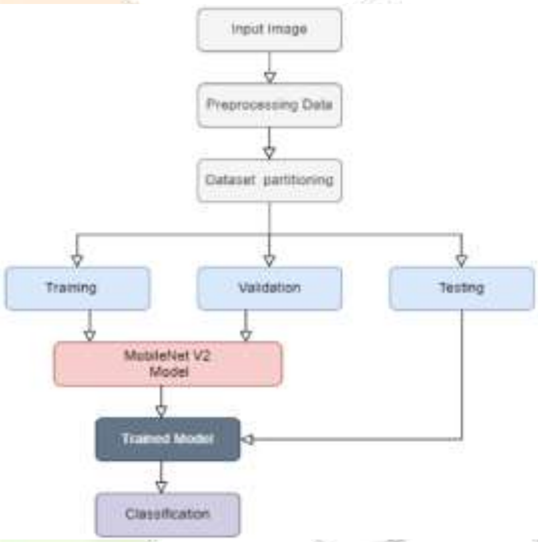
3. Did the study adequately answer the research questions?
Each question was answered with a Boolean 'yes/no', scored as Y = 1 or N = 0. The evaluations yield:

- 77% coverage of the topic.

- 82% verification of study quality.

- 79% adequacy in answering the research questions.

FLOWCHART



Description: A deep learning-based skin cancer classification system utilizes convolutional neural networks to analyze skin lesion images, identifying malignancies through pattern recognition and classification algorithms.

4. PROPOSED SYSTEM

The proposed system aims to leverage deep learning techniques to enhance the accuracy and efficiency of skin cancer detection. This system will address the challenges identified in the literature review, such as data imbalance, high computational costs, and variability in imaging conditions.

The system architecture consists of the following key components:

Data Collection and Preprocessing

Data Sources: Gather a diverse set of skin lesion images from publicly available datasets such as ISIC (International Skin Imaging Collaboration) and DermNet.

Data Augmentation: Apply augmentation techniques like rotation, flipping, scaling, and color adjustments to address data imbalance and increase the variety of training samples.

Normalization and Resizing: Standardize image sizes and normalize pixel values to ensure consistency across the dataset.

Model Selection and Training

Convolutional Neural Networks (CNNs): Utilize state-of-the-art CNN architectures such as ResNet, Inception, and EfficientNet for feature extraction and classification.

Transfer Learning: Implement transfer learning by fine-tuning pre-trained models on the skin cancer dataset to leverage existing knowledge and reduce training time.

Training Strategy: Use techniques like cross-validation and stratified sampling to ensure robust training and validation processes.

Model Optimization

Hyperparameter Tuning: Employ techniques like grid search or random search to optimize hyperparameters such as learning rate, batch size, and number of layers.

Model Pruning and Quantization: Apply model pruning and quantization to reduce computational costs and improve inference speed, making the system suitable for deployment on resource-constrained devices.

Evaluation and Validation

Performance Metrics: Evaluate the model using metrics such as accuracy, precision, recall, F1-score, and AUC-ROC curve to ensure comprehensive performance assessment.

Cross-Domain Testing: Validate the model's robustness by testing it on images from different sources and devices to simulate real-world variability.

5. RESULTS AND DISCUSSION

In our experimental analysis of skin cancer classification using the HAM10000 dataset, we evaluated six different transfer learning models: VGG19, InceptionV3, ResNet50, Xception, InceptionResNetV2, and MobileNet. Training these models for 10 epochs with a batch size of 32, we observed varying performances with and without repetition of images. While some models like InceptionResNetV2, MobileNet, and Xception showed improved accuracy with image repetition, others like VGG19, InceptionV3, and ResNet50 experienced declines. Notably, Xception emerged as the best performer in the balanced dataset, while MobileNet exhibited superior performance in the unbalanced dataset. Confusion matrix results provided insights into classification challenges, with certain lesion types proving more difficult to identify across models. 'Akiec' and 'Mel' lesions consistently posed challenges, while 'Nv' lesions were consistently well-classified. Additionally, precision, recall, F1-Score, and accuracy values highlighted Xception Net's superior performance, achieving the lowest loss of 0.5168 and the highest accuracy of 90.48% on the test set.

The computational costs of the simulations were also considered, detailing hardware specifications and computation time. Notably, Xception Net demonstrated the lowest validation loss and the highest accuracy among the models, indicating its efficiency in classification tasks. These findings underscore the importance of selecting appropriate transfer learning models and preprocessing techniques to enhance the accuracy and efficiency of skin cancer classification systems. Overall, our analysis provides valuable insights into the performance of different models and their implications for real-world applications in healthcare and medical diagnostics.

CONCLUSION

This study underscores the effectiveness of employing data augmentation and transfer learning techniques to achieve competitive performance in skin cancer classification. By leveraging data augmentation, researchers significantly expanded the dataset, enabling more robust feature extraction and classification. Notably, the Xception Net emerged as the top-performing transfer learning model, demonstrating superior accuracy, recall, precision, and F-Measure values. Despite challenges in classifying certain lesion types like 'Akiec' and 'Mel', the Xception Net showcased remarkable performance, underscoring its potential for aiding in early skin cancer detection. Moving forward, further research and fine-tuning of transfer learning algorithms hold promise for enhancing accuracy and ultimately reducing the burden of skin cancer-related mortality.

REFERENCES

[1] Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. Nature, 542(7639), 115–118.

[2] Brinker, T. J., Hekler, A., Enk, A. H., Klode, J., Hauschild, A., Berking, C., ... & Schadendorf, D. (2019). Deep learning outperformed 136 of 157 dermatologists in a head-to-head dermoscopic melanoma image classification task. European Journal of Cancer, 113, 47-54.

[3] Haenssle, H. A., Fink, C., Schneiderbauer, R., Toberer, F., Buhl, T., Blum, A., ... & Klode, J. (2018). Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists. Annals of Oncology, 29(8), 1836-1842.

[4] Tschandl, P., Rosendahl, C., & Kittler, H. (2018). The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. Scientific Data, 5, 180161.

[5] Codella, N. C., Gutman, D., Celebi, M. E., Helba, B., & Marchetti, M. A. (2018). Skin lesion analysis toward melanoma detection: A challenge at the 2017 International Symposium on Biomedical Imaging (ISBI), hosted by the International Skin Imaging Collaboration (ISIC). In 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018) (pp. 168-172). IEEE.