Integrative Model Fusion: A CNN-FCN Approach for Accurate Skin Cancer Diagnosis

Mehrad Hajati[1]

Abstract

Medical diagnostics using medical imaging is one of the important tasks in healthcare. Many machine learning algorithms were used to detect skin cancer only based on the images of skin. For example, Convolutional Neural networks are the most famous methods for image recognition. Adaptation of Trappenberg et al.'s paper, which covers a process to integrate a Convolutional Neural Network with a Fully Connected Network intended to increase the precision of diagnosing skin cancers provided in datasets by International Skin Imaging Collaboration. The paper uses two CNN ensembles, together with the Metadata, portraying an outstanding accuracy. Inspired by this paper, we constructed a project of a lighter dataset version. We analyze various CNNs for extracting relevant features from dermoscopic images while the FCN is processing-related metadata: age, location on the body, and sex of the patient. These two network outputs are then combined to construct a final model for detecting skin cancerous lesions. The proposed combined approach is intended to take advantage of both image-based and metadata-based information. The present paper will construct a model intended to overcome the challenges related to early melanoma detection.

Introduction

Skin cancer is one of the most common types of cancer worldwide, and early detection is quite vital for effective treatment. Recent advances in machine learning have shown great promise in medical image analysis, offering tools that can automate and improve diagnostic accuracy. Of these methods, Convolutional Neural Networks (CNNs) are particularly effective at image recognition tasks, while Fully Connected Networks (FCNs) can process supplementary metadata to enhance predictive performance. This work uses both methods in a unified framework to classify skin lesions by analyzing images and metadata from the International Skin Imaging Collaboration (ISIC). Recent studies have demonstrated the potential of deep learning in medical diagnostics. Trappenberg et al. (2020) showed that the VGG-19 architecture is effective for the task of skin cancer detection and outstanding in feature extraction from dermoscopic images. Despite progress, issues including data imbalance and limitations in computational resources persist as considerable obstacles to effective implementation (Tajbakhsh et al., 2016). Motivated by this literature, the objective of this project was to merge convolutional neural network (CNN) based image analysis with metadata-informed fully convolutional network (FCN) processing to develop a comprehensive diagnostic solution. The methodology here starts with the preprocessing of the ISIC dataset, which originally included more than 25,000 images. In consideration of computational efficiency, images were resized from 1024×1024 to 224×224 pixels, and the dataset was standardized to

[1] E-mail: M.hajati@dal.ca

balance the number of samples belonging to each type of lesion to 239, making a total of 1,912 images. The model contains two parts: a CNN to analyze the dermoscopic images and an FCN to process the metadata. The output of both networks is concatenated to predict which lesion type it is. This dual-input design allows the model to leverage both visual and contextual information for classification. With all the new approaches applied, the model showed a modest effectiveness, with an accuracy of about 30%. This suboptimal result could be explained by both the lower size of the dataset used and the degradation in quality due to the downscaling process. Underlining the necessity of both high-resolution images and bigger datasets for improving diagnostic outcomes in similar situations.

Literature review

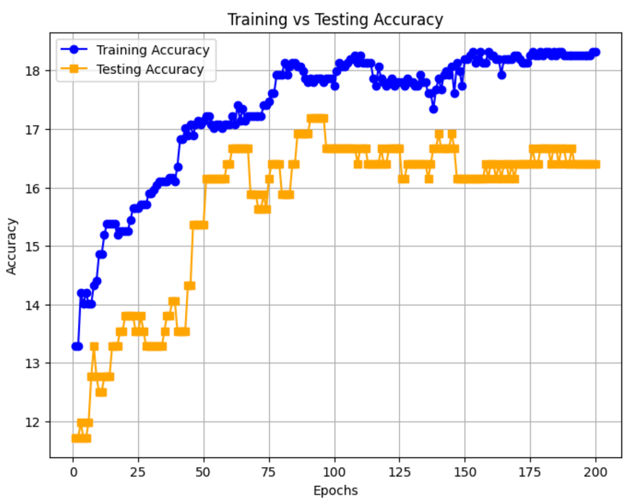
The referenced papers are directly related to our work since they emphasize the importance of integrating image data with metadata. In particular, Trappenberg et al. (2020) introduced the concept of combining CNN outputs with metadata through FCNs, which inspired us to explore a hybrid model approach. This work extends Trappenberg's framework by addressing issues like imbalanced datasets, resolution degradation, and lack of data augmentation. These, coupled with preprocessing improvements, provide substantial insights that help drive the future diagnosis of skin cancer. Other referenced works, on the whole, emphasize the use of deep learning models in analyzing medical images and especially skin cancer detection. Some of these references include Pacheco et al. (2019) and Codella et al. (2019), which underscores the usefulness of combining image data with metadata in improving diagnosis accuracy, an essential idea in our work. Litjens et al. (2017) provide a comprehensive review of deep learning methodologies applied to medical imaging, highlighting challenges such as data imbalance, preprocessing, and the importance of metadata—features that all informed our preprocessing protocols and model development. Tajbakhsh et al. (2016) discuss the benefits of transfer learning and fine-tuning in medical imaging, validating our choice of leveraging pre-trained models like VGG-19 for feature extraction.

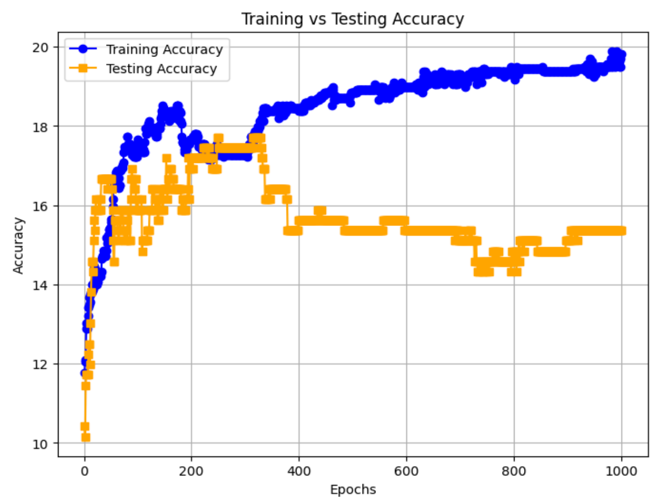
Data Processing

The initial dataset used in the ISIC challenge contained 25,331 images, all of which were 1024×1024 pixels in size and came with associated metadata. To make them easier to process and ensure computational feasibility, the images were later reduced in size to 224×224 pixels. Furthermore, to address the issue of data imbalance regarding the number of lesion types, the dataset was reduced by restricting each lesion type to the number of the least-represented class (DF class), resulting in 239 images per lesion category. This balancing process cut the dataset down to 1,912 images in total, evenly distributed among all lesion categories and making this project computationally feasible.

The raw metadata included fields for sex, lesion identification number, image identification number, age, and the approximate location of the lesion. The "sex" attribute, which was categorized as "male" and "female," is translated to the numerical values 0 and 1, respectively. The same treatment was done with the "approximate location" attribute, which contained eight different values and encoded into integer values ranging from 0 to 7. The "lesion ID" field was not informative to train on and was therefore removed. Rows with missing information in any field were excluded from the dataset to ensure data quality and consistency. Initially, the ground truth dataset included a relationship between image_id and a one-hot encoded array with eight columns, where each column represented a specific type of lesion. For easier interpretation of the ground truth, lesion types were systematically renumbered from 0 to 7, hence creating a new data frame linking each image_id to its corresponding lesion type in the form of a single integer. This approach allowed smooth integration into the machine-learning pipeline.

Results and analysis:





These two plots show the FCN model training accuracy vs testing accuracy. The X-axis are Epochs and the Y-axis shows the accuracy. One of them goes up to 100 Epoch but the other goes to 1000 Epoch. This is to show that our model cannot overfit the data using just the metadata. The codes for our programs can be found here at https://github.com/esysss/skin_cancer.

We cut the data into 8 equal classes to increase the running data and eliminate the unbalanced data. To create a baseline for comparison we ran a simple CNN on the images and ANN on the matadata to be able to compare our idea with the baseline and see the improvement. We divided the dataset into 80% for training and 20% for testing, using Cross Entropy and 40 Epochs.

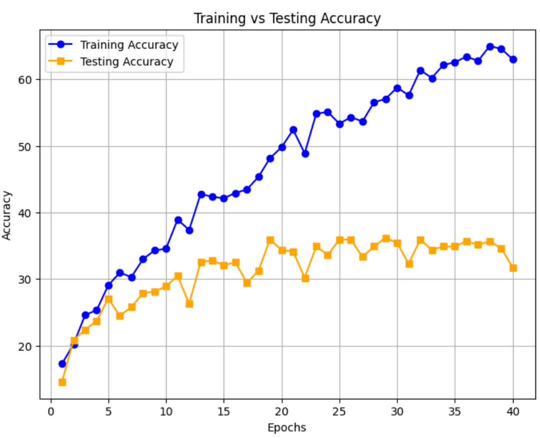
Final accuracy done on images and metadata:

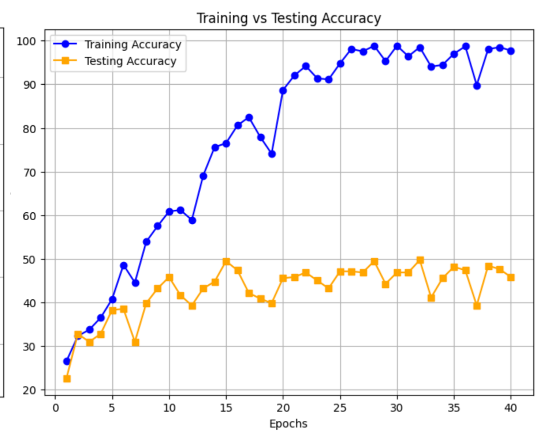
Approach (Images)	Accuracy	Approach (Metadata)	Accuracy
3-layer CNN	36.19%	FCN	17.70%
VGG16	48.95%	Random forest	10.94%

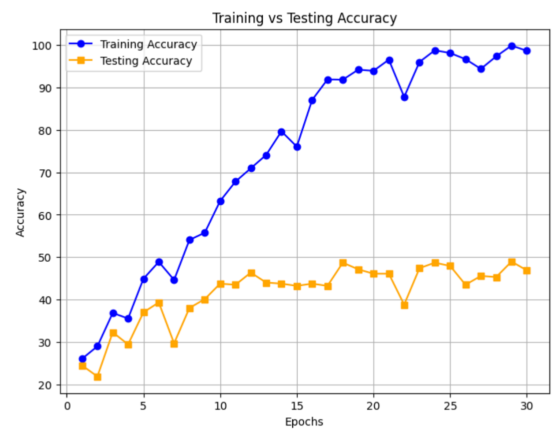
VGG19	49.73%	Logistic Regression	12.24%

Accuracy of the idea with a combination of images and metadata

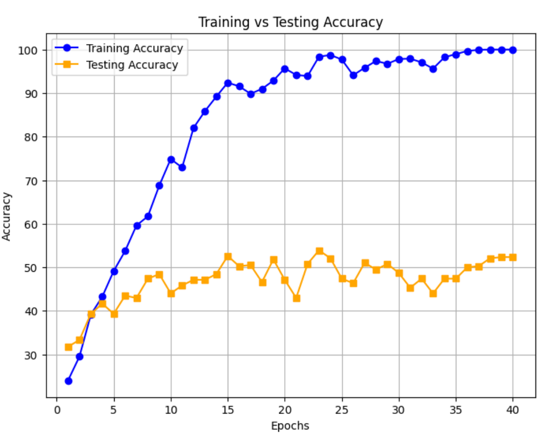
Approach	Accuracy
CNN + Metadata	40.62
VGG16 + Metadata	
VGG19 + Metadata	53.90

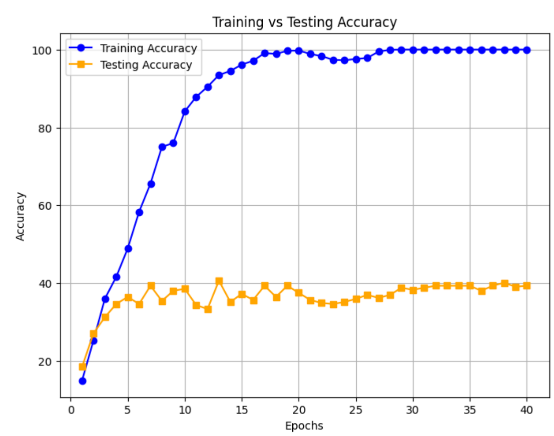






The first plot shows the training accuracy vs testing accuracy of simple 3-layer CNN. The second (upper right) shows the training accuracy vs testing accuracy for VGG-16. The third (lower right) shows the training accuracy vs testing accuracy for VGG-19. The X-axis shows the number of Epochs (30 were run) and the Y-axis shows the accuracy.





Future Works

This research has demonstrated a hybrid CNN-FCN method for skin cancer detection. However, several further research avenues can be explored to enhance model performance and effectiveness.

1. Clustering with Cosine Similarity: By creating a network of the images with cosine similarities, computing pairwise similarities in image features, thereby trying to improve the classification for lesions. Then clustering algorithms on the network to find out groups of lesion types, either unsupervised or semi-supervised to increase generalization across multiple datasets and expose some hidden patterns within the data.
2. Grayscale Image Analysis: While this study has used resized color images, another approach would be to convert the images to grayscale. This preprocessing step could dramatically reduce the data size and computational cost. Future experiments could investigate the impact of converting to grayscale on the accuracy of the model and the training time, which may yield useful insights into whether the color information is crucial for effective lesion classification.
3. In the main paper, extensive preprocessing was performed on the image data, including applying data augmentation using common image processing techniques: adjustments to brightness, contrast, saturation, and hue; horizontal and vertical rotations; translations; scaling; and shearing. In addition, before data augmentation was performed, shades of gray method (21) was used for all images. So, we would be concerned about how each preprocessing step would impact the model's performance and try to measure how good each of the different data augmentation techniques really is.

These directions, if explored, might lead to improved performance, reduced computational demands, and new methodologies in skin cancer diagnosis using machine learning.

References:

[1] Pacheco AG, Ali AR, Trappenberg T. Skin cancer detection based on deep learning and entropy to detect outlier samples. arXiv preprint arXiv:1909.04525. 2019 Sep 10.

[2] Litjens G, Kooi T, Bejnordi BE, Setio AA, Ciompi F, Ghafoorian M, Van Der Laak JA, Van Ginneken B, Sánchez CI. A survey on deep learning in medical image analysis. Medical image analysis. 2017 Dec 1;42:60-88.

[3] Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. CA: a cancer journal for clinicians. 2019 Jan;69(1):7-34.

[4] Tajbakhsh N, Shin JY, Gurudu SR, Hurst RT, Kendall CB, Gotway MB, Liang J. Convolutional neural networks for medical image analysis: Full training or fine tuning?. IEEE transactions on medical imaging. 2016 Mar 7;35(5):1299-312.

[5] Tschandl P, Rosendahl C, Kittler H. The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. Scientific data. 2018 Aug 14;5(1):1-9.

[6] Pacheco AG, Krohling RA. An attention-based mechanism to combine images and metadata in deep learning models applied to skin cancer classification. IEEE journal of biomedical and health informatics. 2021 Feb 26;25(9):3554-63.

[7] Varma PB, Paturu S, Mishra S, Rao BS, Kumar PM, Krishna NV. SLDCNet: Skin lesion detection and classification using full resolution convolutional network‐based deep learning CNN with transfer learning. Expert Systems. 2022 Nov;39(9):e12944.

[8] Pacheco AG, Sastry CS, Trappenberg T, Oore S, Krohling RA. On out-of-distribution detection algorithms with deep neural skin cancer classifiers. InProceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops 2020 (pp. 732-733).

[9] Villa-Pulgarin JP, Ruales-Torres AA, Arias-Garzon D, Bravo-Ortiz MA, Arteaga-Arteaga HB, Mora-Rubio A, Alzate-Grisales JA, Mercado-Ruiz E, Hassaballah M, Orozco-Arias S, Cardona-Morales O. Optimized convolutional neural network models for skin lesion classification. Computers, Materials & Continua. 2022 Feb 1;70(2).

[10] Ismail AR, Nisa SQ, Shaharuddin SA, Masni SI, Amin SA. The Utilising VGG-16 of Convolutional Neural Network for Medical Image Classification. International Journal on Perceptive and Cognitive Computing. 2024 Jan 28;10(1):113-8.

[11] Venugopal V, Raj NI, Nath MK, Stephen N. A deep neural network using modified EfficientNet for skin cancer detection in dermoscopic images. Decision Analytics Journal. 2023 Sep 1;8:100278.