Skin cancer detection (Paper)

Presented by:

1- Ehab Nafea 20108922

2- Makarious Ghanamy 19108679

Presented to : Dr.Nashwa Elbendary

1-Introduction

-Background:-

Skin cancer, which includes melanoma, basal cell carcinoma (BCC), and squamous cell carcinoma (SCC), is a major global health concern. The traditional method for detecting skin lesions has been visual inspection by dermatologists, however it is vulnerable to subjectivity and diagnostic problems due to the small variations between benign and malignant growths.

Automated skin cancer detection systems powered by machine learning and computer vision provide a game-changing solution to these issues. These systems scan photos of skin lesions using advanced algorithms, particularly convolutional neural networks (CNNs), to extract complicated patterns and features that aid in correct classification.

Datasets such as the International Skin Imaging Collaboration (ISIC) have helped train these machine learning models. Despite its potential, issues such as data shortages must be addressed.

-purpose:-

The goal of skin cancer detection is to aid in the early detection, correct categorization, and timely intervention of skin abnormalities. Automated detection systems powered by machine learning and computer vision are critical in attaining this goal since they:

1- Early detection: The ability to recognize skin lesions and anomalies in their early stages, before they proceed to advanced and sometimes life-threatening disorders.

2- Accurate Classification: Using advanced algorithms to differentiate benign and malignant tumors, minimizing the possibility of misdiagnosis and enabling suitable treatment choices.

3- Enhanced Diagnostic Accuracy: Delivering objective and consistent assessments, supplementing dermatologists' expertise, and reducing errors associated with

4- subjective visual inspection.

2- Literature Review:

Summarize Previous Work:

Skin cancer classification has been investigated using known models such as VGG16. Transfer learning, particularly from pre-trained convolutional neural networks (CNNs) such as VGG16, has been widely used. When trained on huge datasets, these models may learn complicated patterns from photos. Important elements relevant to skin lesion classification are effectively captured by adapting pre-trained models without sacrificing previously learned information.

Methods of the Future:

Recent advances include the use of pre-trained CNN architectures (e.g., VGG16) that have been trained on various datasets such as ImageNet. This method enables models to benefit from information gained from large data representations. Techniques for data augmentation, such as ImageDataGenerator, play an important role in improving the training dataset. Furthermore, the claimed validation accuracy is an important indicator of a model's performance on unknown data.

3- Data set and Methodology:

1- Data set description:-

The dataset used for skin cancer categorization is from the ISIC (International Skin Imaging Collaboration). It is made up of photographs of various skin lesions. The training set has 2239 photos divided into 9 classes, whereas the test set contains 2357 images divided into 1 class. Image resizing to 224x224 pixels and normalization (scaling pixel values between 0 and 1) are two preprocessing procedures. During training, data augmentation techniques were used.

2- Model Architecture:

The VGG16 architecture, a pre-trained convolutional neural network (CNN), is used in the model's transfer learning. The top layers of the base VGG16 model were deleted, and bespoke fully connected layers were added. A Flatten layer, a Dense layer with 256 units with ReLU activation, and a final Dense layer with a sigmoid activation function for binary classification comprise the improved architecture. The VGG16 layers were frozen in order to retain their learned characteristics.

3- Training Methodology:

The Adam optimizer was employed, with a learning rate of 0.0001.

Batch Size: The batch size for producing batches during training, validation, and testing was set to 16.

Epochs: The model was trained across a total of ten epochs.

ImageDataGenerator was used to enhance training photos using modifications such as rescaling, shearing, zooming, and horizontal flipping.

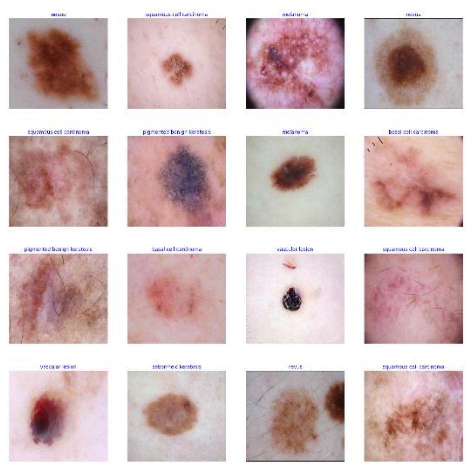
---

4 -Results and Discussion:-

1- Performance Metrics: Accuracy, Precision, Recall, and F1-score are used as evaluation metrics.

Confusion Matrix: This matrix compares the model's predictions to the actual labels for each class.
Precision, recall, and F1-score for each class, as well as overall metrics, are included in the Classification Report.

2- Experimental Outcomes:
Confusion Matrix: Displays the model's performance in predicting each class versus the true data.
Classification Report: Describes the precision, recall, and F1-score for each class, demonstrating how successfully the model classified various skin lesions.
Overall Metrics: Discuss the model's overall accuracy on the test dataset and highlight specific classes where it excelled or underperformed.



3- Comparison:

Comparison with current approaches: In skin cancer classification tasks, compare the model's performance measures (accuracy, precision, recall, etc.) with state-of-the-art approaches or current models.
Areas for Development: Identify certain classes or scenarios where the model excelled and others where it suffered.
Suggest prospective changes or tactics to increase the model's performance, such as fine-tuning, adopting newer architectures, or leveraging larger datasets.

5- Conclusion and Discussion:

1- Interpretation of Findings:
Implications for Clinical Practice: Interpret the findings in terms of their significance in skin cancer diagnosis, potential impact on early identification, and subsequent treatment outcomes.
Model Performance Importance: Discuss how the obtained accuracy and categorization metrics can benefit dermatology and skin cancer diagnosis.

2- Limitations:
Dataset Constraints: Address any dataset restrictions, such as dataset size, class imbalance, or fluctuations in image quality.

3 -Model Limitations:
Discuss any model architecture or training approach limits that may have influenced the results.
Future Prospects:
Enhancing Datasets: Emphasize the significance of larger and more diversified datasets for building robust skin cancer classification models.

Addressing Class Imbalance: Provide solutions for overcoming class imbalances in the dataset, as well as their possible impact on model performance.

6. References:

National Cancer Institute. (2021). Understanding Cancer. [Online]. Available: https://www.cancer.gov/about-cancer/understanding/what-is-cancer

ISIC - International Skin Imaging Collaboration. (2021). ISIC Archive. [Online]. Available: https://www.isic-archive.com/#!/topWithHeader/onlyHeaderTop/gallery

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR) (pp. 770-778).

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.

Chollet, F. et al. (2015). Keras. [Online]. Available: https://keras.io

Abadi, M. et al. (2015). TensorFlow: Large-scale machine learning on heterogeneous systems. [Online]. Available: https://www.tensorflow.org/