Skin Cancer Identification Using Deep Learning Technique

J.Anitha 21WH1A1267, K.Padma Sree 21WH1A1270 ,CH.Siri Meghana 21WH1A12A2 , T.Manaswini 21WH1A12A5

Under the esteemed guidance of

Dr. B. Srinivasulu

Assistant Professor

Bachelor of Technology

Department of Information Technology

BVRIT HYDERABAD college of engineering for Women

Overview

Introduction

- **Skin Cancer** is among the most rapidly increasing types of cancer globally, ranking as the sixth most common.
- It occurs due to the abnormal growth of skin cells, often triggered by factors such as UV ray exposure, a weakened immune system, and family history.
- Tumors can be classified as benign (non-harmful, like moles) or malignant (life-threatening and capable of damaging surrounding tissues).
- Benign tumors are non-cancerous, grow slowly, and do not spread, but may cause issues by pressing on nearby tissues.
- Malignant tumors are cancerous, grow aggressively, invade surrounding tissues, and can spread to other parts of the body.
- Early diagnosis and treatment are crucial for managing malignant tumors.

Benign & Melignant Images

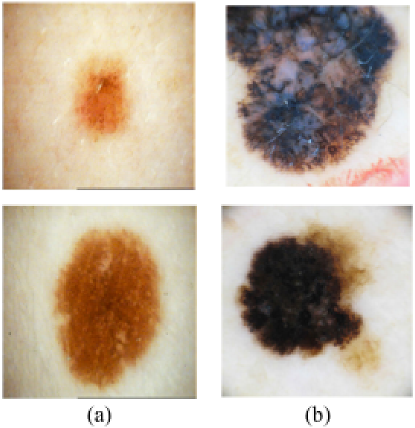


Figure: Image of (a)Benign Case (b)Malignant Case

Problem Statement

- Skin cancer diagnosis, particularly melanoma, is hindered by the subjectivity, cost, and limited availability of skilled professionals in traditional methods.
- While CNNs offer potential for automated detection, challenges such as limited annotated data, class imbalance, and lack of interpretability restrict their clinical adoption.
- A solution is needed to enhance accuracy, transparency, and accessibility for early and reliable diagnosis.

Literature Survey

Title	Author(s)	Year	Volume no.	Journal name	Methodology	Result
Skin Cancer Identification Using Deep Learning Technique	Gaurav Kumar Gautam, Sofia Singh, Archana Singh	2024	10	2024 International Conference on Communication, Computer Sciences and Engineering (IC3SE), Gautam Buddha Nagar, India.	• Utilized EfficientNet Convolutional Neural Network (CNN) for feature extraction and classification. • Implemented fairness evaluation algorithm to address biases in model predictions. • Applied federated learning approach for image classification with datasets like HAM10000.	• Achieved an accuracy of 82.14 percentage with EfficientNet for classifying seven skin lesion categories. • Strong performance metrics: AUC = 0.9545, F1-score = high values near unity.
Performance Enhancement of Skin Cancer Classification Using Computer Vision	Ahmed Magdy, Hamdeer Hussein, Rehab F. Abdel-Kader, and Khaled Abd El Salam	2023	11	IEEE Access	• AlexNet ResNet-18, ResNet-50, ResNet-101 VGG-16, VGG-19 DenseNet-201 EfficientNet-B0 Inception-v3 MobileNet-v2	• KNN-PDNN: Accuracy greater than 99 percentage with VGG-16 as the best feature extractor. • AlexGWO: Accuracy 99 percentage, outperforming standalone AlexNet and other tested models.

Literature Survey

Title	Author(s)	Year	Volume no.	Journal name	Methodology	Result
A Deep Learning-Based Convolutional Neural Network Model with VGG16 Feature Extractor for the Detection of Alzheimer Disease Using MRI Scans	Shagun Sharma, Kalpna Guleria, Sunita Tiwari, Sushil Kumar	2022	24	Measurement: Sensors	• Developed a Deep Learning (DL) model using Convolutional Neural Networks (CNN). • Employed VGG16 feature extractor for image-based Alzheimer detection. • Used two MRI datasets for training and testing with accuracy, precision, recall, AUC, and F1-score metrics to evaluate performance.	• Dataset 1: Achieved 90.4 percentage accuracy, precision of 0.905, recall of 0.904, AUC of 0.969, and F1-score of 0.904. • Dataset 2: Achieved 71.1 percentage accuracy, precision of 0.71, recall of 0.711, AUC of 0.85, and F1-score of 0.71. • The proposed model outperformed state-of-the-art methods in early diagnosis tasks

Literature Survey

Title	Author(s)	Year	Volume no.	Journal name	Methodology	Result
Early Detection of Skin Cancer Using Deep Learning Approach	Ibrahim Al-shourbaji, Ghassan Samara, Hussam Abu Munshar, Waleed A. Zogaan, Faheem A. Reegu, Shadab Alam, Muhammad Saidu Aliero	2021	20	Elementary Education Online	• Utilized Convolutional Neural Networks (CNNs) for classifying skin cancer images. • The dataset used was obtained from Kaggle's ISIC archive and contained 3297 labeled images. • Performed data augmentation (rotation, flipping, zooming) to enhance dataset diversity and prevent overfitting. • The CNN architecture included layers for convolution, pooling, and fully connected classification.	• Successfully classified skin cancer images into melanoma (malignant) and benign classes. • Demonstrated the effectiveness of CNNs in early detection, achieving high accuracy and reliable classification.

Literature Survey

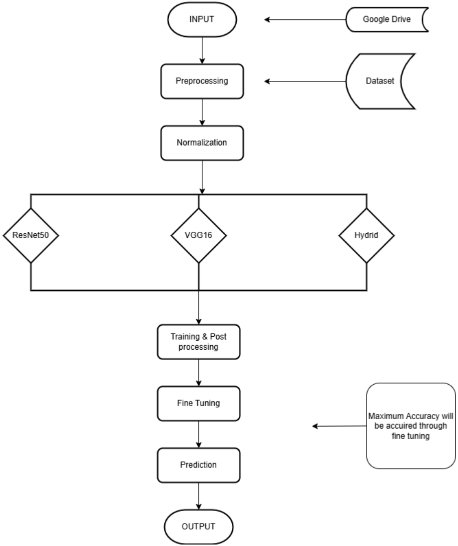
Title	Author(s)	Year	Volume no.	Journal name	Methodology	Result
Comparative Analysis of Skin Cancer (Benign vs. Malignant) Detection Using CNNs	Mohammed Rakeibul Hasan	2021	2021 Article ID: 5895156	Journal Of Healthcare Engineering	• Comparison of CNN models (VGG16, ResNet50, SVM, sequential models) on Kaggle dataset with 6594 images; accuracy comparison and fine-tuning.	• VGG16 achieved the highest accuracy of 93.18 percentage, followed by SVM (83.48 percentage) and ResNet50 (84.39 percentage).
Comparing the Performance of Linear Regression vs. Deep Learning on Detecting Melanoma Using Apple Core ML	Herry Sujaini	2021	10	Bulletin of Electrical Engineering and Informatics	• Developed models using linear regression and CNN on the ISIC dataset; metrics compared included accuracy, sensitivity, specificity, and false-positive/negative rates.	• CNN achieved 70 percentage accuracy with a 25 percentage false-negative rate, outperforming linear regression (68 percentage).

Proposed System

The following are the proposed models:

- **ResNet-50**: Used to address the vanishing gradient problem and extract deep features for accurate classification in skin cancer detection.
- **VGG-16**: Utilized for capturing detailed spatial hierarchies with its structured and efficient architecture.
- **Hybrid Model**: Combines the strengths of ResNet-50 and VGG-16 to improve overall classification accuracy and performance.

Architecture



Tools & Technologies

- **NumPy:** Efficient handling of large numerical arrays.
- **Pandas:** Data manipulation and analysis
- **Matplotlib:** Visualization of results
- **Sklearn:** Model training and evaluation
- **TensorFlow/Keras:** Deep learning model development
- **Seaborn:** Enhanced statistical visualizations

Feasibility Study

- Public datasets like **ISIC** on Kaggle provide annotated skin images, essential for training accurate skin cancer detection models.
- Tools like **TensorFlow** and **Keras** offer easy-to-use tools for building and refining deep learning skin cancer detection models.
- Models like **CNN**,**ResNet-50**,**VGG-16** and **Hybrid** are highly effective in detecting skin cancer and identifying melanoma from dermoscopic images.
- Communities on **GitHub** and **Google Scholar** provide research, open source code, and support for skin cancer detection projects. They keep developers up-to-date on the latest advances.
- Ongoing research ensures that skin cancer detection models continue to improve. Integrates new techniques for better accuracy and interpretability in clinical use.

Societal Impact

- The proposed system can significantly reduce deaths related to melanoma by enabling accurate and early diagnosis, particularly in areas limited by resources.
- It provides a cost-effective and scalable solution that alleviates the burden on healthcare professionals and improves patient outcomes, ultimately improving global healthcare equity and reducing the socioeconomic impact of skin cancer.
- In general, this approach leads to improved global public health, improved access to timely medical care, and a reduction in health disparities.
- Sustainable Development Goals : GOOD HEALTH AND WELL BEING (3).

Partial code implementation and results

```
# loading the training data
data_dir = '/content/drive/MyDrive/Major Project/train'

train_df = loading_the_data(data_dir)

train_df
```

	filepaths	labels
0	/content/drive/MyDrive/Major Project/train/ben...	benign
1	/content/drive/MyDrive/Major Project/train/ben...	benign
2	/content/drive/MyDrive/Major Project/train/ben...	benign
3	/content/drive/MyDrive/Major Project/train/ben...	benign
4	/content/drive/MyDrive/Major Project/train/ben...	benign
...	...	...
2632	/content/drive/MyDrive/Major Project/train/mal...	malignant
2633	/content/drive/MyDrive/Major Project/train/mal...	malignant
2634	/content/drive/MyDrive/Major Project/train/mal...	malignant
2635	/content/drive/MyDrive/Major Project/train/mal...	malignant
2636	/content/drive/MyDrive/Major Project/train/mal...	malignant

2637 rows × 2 columns

Importing the dataset

Partial code implementation and results

```
# crobed image size
batch_size = 8
img_size = (224, 224)

tr_gen = ImageDataGenerator(rescale=1. / 255)
ts_gen = ImageDataGenerator(rescale=1. / 255)

train_gen = tr_gen.flow_from_dataframe( train_df, x_col= 'filepaths', y_col= 'labels', target_size= img_size, class_mode= 'categorical',
                                        color_mode= 'rgb', shuffle= True, batch_size= batch_size)

valid_gen = ts_gen.flow_from_dataframe( valid_df, x_col= 'filepaths', y_col= 'labels', target_size= img_size, class_mode= 'categorical',
                                        color_mode= 'rgb', shuffle= True, batch_size= batch_size)

test_gen = ts_gen.flow_from_dataframe( test_df, x_col= 'filepaths', y_col= 'labels', target_size= img_size, class_mode= 'categorical',
                                        color_mode= 'rgb', shuffle= False, batch_size= batch_size)
```

```
Found 2637 validated image filenames belonging to 2 classes.
Found 334 validated image filenames belonging to 2 classes.
Found 334 validated image filenames belonging to 2 classes.
```

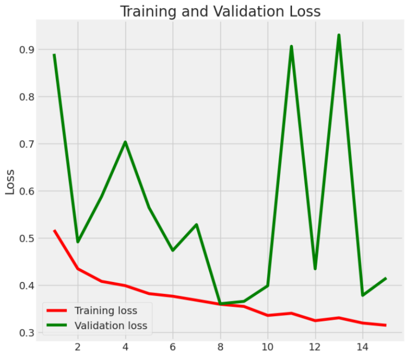
Pre-processing

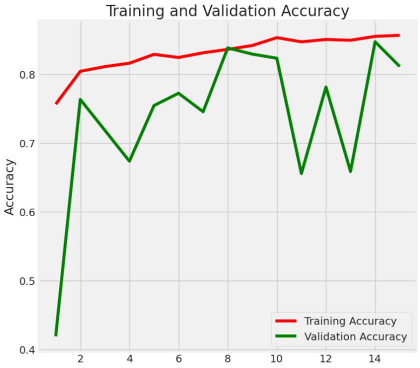
```
# get the pre-trained model (ResNet50)
base_model = ResNet50(weights='imagenet', include_top=False, input_shape = img_shape, pooling= None)

# freeze the layers in conv5_block3
for layer in base_model.layers:
    if 'conv5_block3' in layer.name:
        layer.trainable = True
    else:
        layer.trainable = False

# fine-tune ResNet50 (Adding some custom layers on top)
x = base_model.output
x = GlobalAveragePooling2D()(x)
x = BatchNormalization()(x)
x = Dense(128, activation = 'relu')(x)
x = Dropout(0.5)(x)   # Dropout layer to prevent overfitting
x = Dense(32, activation = 'relu')(x)
predictions = Dense(class_counts, activation = "sigmoid")(x)    # output layer with softmax activation

# the model
ResNet50_model = Model(inputs = base_model.input, outputs = predictions)
```

```
Downloading data from https://storage.googleapis.com/tensorflow/keras-applications/resnet/resnet50_weight
94765736/94765736 ━━━━━━━━━━━━━━━━━━━━ 3s 0us/step
```

Model Building

Partial code implementation and results





Partial code implementation and results

```
# Model evaluation
model_evaluation(ResNet50_model)
```

```
330/330 ━━━━━━━━━━━━━━━━━━━━ 418s 1s/step - accuracy: 0.8573 - loss: 0.3104
42/42 ━━━━━━━━━━━━━━━━━━━━ 54s 1s/step - accuracy: 0.8551 - loss: 0.3068
42/42 ━━━━━━━━━━━━━━━━━━━━ 92s 2s/step - accuracy: 0.7926 - loss: 0.4674
Train Loss:  0.2900655269622803
Train Accuracy:  0.8642396926879883
--------------------
Validation Loss:  0.41477176547050476
Validation Accuracy:  0.811377227306366
--------------------
Test Loss:  0.5126500725746155
Test Accuracy:  0.811377227306366
```

Model Evaluation & Accuracy

Project Timeline

Phase	Timeline	Description
Phase 1: Image Pre-processing and Dataset Preparation	October 2024 - November 2024	Collecting and pre-processing skin cancer image dataset from ISIC, applying augmentation to address class imbalance and improve diversity.
Phase 2: Model Development	November 2024 - December 2024	Develop deep learning models, including ResNet-50, VGG-16, and hybrid, for skin lesion classification.
Phase 3: Model Interpretability and Validation	January 2025 - February 2025	Ensure reliable performance with extensive cross-validation and robustness testing under diverse conditions, including image quality and lighting variations.

Project Timeline

Phase	Timeline	Description
Phase 4: Deployment and Integration	March 2025	Deploy the model as a web or mobile-based diagnostic tool, enabling clinicians and users to upload dermoscopic images for real-time analysis. Ensure the system is user-friendly and scalable.
Phase 5: Paper Publication	April 2025 - May 2025	Prepare detailed documentation of the project's methodology, results, and contributions. Submit the findings to a peer-reviewed journal or conference and share insights with the broader research and medical communities.

References I

[1] [2] [3] [4] [5] [6]

G. K. Gautam, S. Singh, and A. Singh.
Skin cancer identification using deep learning technique.
2024 International Conference on Communication, Computer Sciences and Engineering (IC3SE), Gautam Buddha Nagar, India, 10.:553–559, 2024.

Ahmed Magdy, Hadeer Hussein, Rehab F. Abdel-Kader, and Khaled Abd El Salam.
Performance enhancement of skin cancer classification using computer vision.
IEEE Access, 11:72120–72133, 2023.

References II

Sharma S., Guleria K., Tiwari S., and Kumar S.
A deep learning based convolutional neural network model with vgg16 feature extractor for the detection of alzheimer disease using mri scans.
Measurement: Sensors, 24:100506, Dec. 2022.

Alshourbaji Ibrahimarabia.
Early detection of skin cancer using deep learning approach.
Elementary Education Online, 20(5), 2021.

Mohammed Rakeibulfatemi Hasan, Mohammed Ishraafmonirujjaman Khan, Manjitzaguia Mohammadkaur, and Atef.
Comparative analysis of skin cancer (benign vs malignant) detection using convolutional neural networks.
Journal Of Healthcare Engineering, 2021 Article ID:5895156, 2021.

References III

Herryramadhan Sujaini, Enriko Yudhistiranovriando, and Haried.
Comparing the performance of linear regression versus deep learning on detecting melanoma skin cancer using apple core ml.
Bulletin of Electrical Engineering and Informatics, 10(6), 2021.

Thank You