It's a Mole! It's a Mark! IT'S CANCER... or is it? Understanding Skin Cancer Detection using Data Science with Machine Learning

Rustam Guliyev
School of Computer Science
University of Nottingham
Nottingham, United Kingdom

Elakia V.M.
School of Computer Science
University of Nottingham
Nottingham, United Kingdom

Andrzej Kolodziej
School of Computer Science
University of Nottingham
Nottingham, United Kingdom

Akshita Bhatia
School of Computer Science
University of Nottingham
Nottingham, United Kingdom

***Abstract*—Melanoma, the deadliest form of skin cancer, poses significant diagnostic challenges, particularly across varying skin tones. Leveraging deep learning, this work investigates the performance of five pre-trained convolutional neural network (CNN) architectures—VGG16, ResNet50, Xception, DenseNet201, and MobileNetV2—on dermoscopic image classification. A comprehensive pipeline was implemented, including contrast enhancement using CLAHE, stratified data preprocessing, and real-time data augmentation to simulate clinical imaging conditions. Model performance was assessed not only by overall accuracy but also across FitzPatrick skin types, revealing disparities in detection rates. Xception and DenseNet201 demonstrated superior accuracy, especially for medium to darker skin tones, while performance was comparatively lower for lighter tones. These results highlight the significance of balanced datasets in medical picture analysis and the necessity of skin-tone-aware AI algorithms. The suggested method provides a starting point for creating equitable and useful diagnostic instruments for a range of populations.**

***Index Terms*—Skin cancer, deep learning, CNNs, transfer learning, skin tone analysis.**

I. Introduction

Prevention is better than cure, a principle articulated by Renaissance scholar Desiderius Erasmus, underscores the importance of proactive health management. This concept is particularly salient in the context of cancer, which often begins asymptomatically, leading to detection at later stages when treatment options are limited [1], [2]. The global burden of cancer is significant, with approximately 20 million new cases diagnosed annually [3]. In the United Kingdom alone, more than 385,000 new cancer cases are diagnosed each year, which equates to over 1,000 cases daily [4].

Unlike many internal cancers, skin cancer is visible and often accessible for early detection. The skin, being the largest organ of the body, reflects overall health and allows direct visual observation [5], [6]. However, visual accessibility does not equate to ease in diagnosis. Melanoma, the deadliest form of skin cancer, is the fifth most common cancer in the UK and remains a diagnostic challenge. Differentiating melanoma from benign lesions such as atypical nevi is complex, with even expert pathologists showing inter-observer variability [7]–[10].

Advancements in artificial intelligence (AI) and machine learning (ML) have enhanced the ability to detect skin cancer through automated analysis of dermoscopic images, thereby improving early diagnosis and patient outcomes [11]. However, diagnostic disparities remain, especially among patients with varying skin tones. Subtle clinical signs of melanoma may manifest differently in melanin-rich skin compared to lighter skin tones, potentially leading to misdiagnosis or delayed treatment [12].

This research aims to bridge this gap by analyzing how different skin types influence the performance of ML models in diagnosing melanoma. The central research question is:

RQ: How do different skin types affect skin cancer detection with different ML models?

A. Scope and Dataset Integration

This study addresses the research question by focusing on two main demographic factors:

- **Skin Complexion**
- **Race/Ethnicity**

Three datasets are utilized to explore these factors:

- ISIC Melanoma Dataset

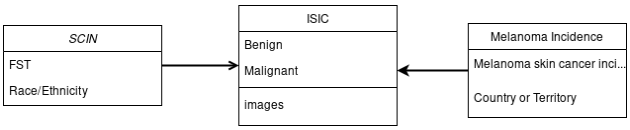

Fig. 1: Data relationship

- SCIN (Skin Condition Image Network)
- Melanoma Incidence from The Cancer Atlas

B. Skin Complexion

The primary dataset, the ISIC Melanoma Dataset [13], contains approximately 1,800 benign and 1,500 malignant biopsy images, predominantly of lighter skin tones. Two techniques are employed to assess complexion:

a) ***RGB Color Analysis****:* A Python-based model using OpenCV [14], NumPy, and Matplotlib clusters the images based on dominant RGB color values. K-means clustering identifies groups with similar color profiles.

b) ***Fitzpatrick Skin Type Classification****:* Developed by Dr. Thomas Fitzpatrick [15], this system categorizes skin based on UV response into six types (FST I–VI), ranging from very fair to very dark skin. This study correlates RGB clusters with Fitzpatrick types to better understand the dataset's complexion distribution.

To enhance this analysis, the SCIN dataset [16], developed by Google Research and Stanford Medicine, is incorporated. It contains over 5,000 labeled images of skin conditions, including attributes such as:

- Related category [e.g., `GROWTH_OR_MOLE`]
- Race and ethnicity [e.g., `ASIAN`]
- Fitzpatrick skin type [e.g., `FST1` to `FST6`]

C. Race and Ethnicity

To examine racial disparities, this study utilizes the Melanoma Incidence dataset from The Cancer Atlas [17], which provides melanoma case data for 213 countries. Attributes include:

- Country or Territory
- Melanoma Incidence (cases per 100,000 individuals)

By integrating SCIN race data and melanoma incidence rates, the study explores how race and ethnicity influence melanoma detection and outcomes across global populations.

II. Literature Review

Deep learning methods, especially Convolutional Neural Networks (CNNs), which are excellent at image recognition tasks, have been the main force behind the adoption of ML in dermatology. Using dermoscopic pictures, CNN-based models like ResNet-50 [18], InceptionV3, and DenseNet [19] have demonstrated exceptional accuracy in classifying skin lesions. Research has shown that these models can occasionally do better than dermatologists in differentiating between benign and malignant cancers, especially melanoma. Studies using huge datasets like HAM10000 and ISIC have demonstrated the effectiveness of AI-driven diagnostics by achieving classification accuracies of above 95%. [20], [21]

Despite improvements, a number of obstacles prevent broad adoption. Data bias is a serious problem since many AI models are trained mostly on lighter skin tones, which results in differences in the diagnosis of skin cancer in skin with more melanin. Research suggests that unusual lesions in darker skin may be difficult for AI systems to identify, raising the possibility of a misdiagnosis. To overcome this obstacle and increase diagnostic accuracy across various populations, more representative and varied training datasets are needed. [20], [21]

III. Methodology

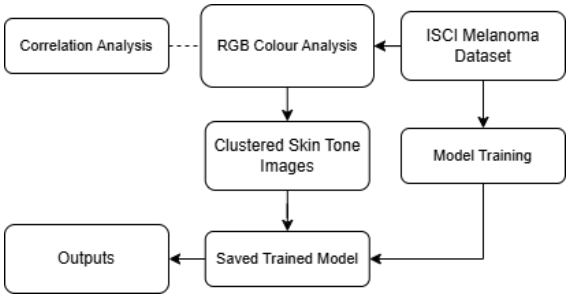

Fig. 2: Main pipeline

A. Exploratory Data Analysis

Two analysis were done to explore the data:

1) ***Correlation Analysis****:* To investigate the relationship between Fitzpatrick skin types and race/ethnicity, a chi-square test for independence was conducted on the SCIN dataset. Specifically, a contingency table was created cross-tabulating the frequencies of Fitzpatrick skin types against various self-reported race/ethnicity categories. The chi-square analysis returned a significant association ($\chi^2 = 685.93$, $p < 0.001$).

Chi-Square Test:

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} \tag{1}$$

where:

- χ^2 - statistical hypothesis test value
- O_i - observed frequencies in each category,
- E_i - expected frequencies under the null hypothesis,

Chi-Square test indicating a strong dependency between skin type classifications and reported race/ethnicity.

Visualization methods included heatmaps and bar charts. A heatmap visualization effectively depicted the frequency distributions across Fitzpatrick skin types and race/ethnicity, highlighting areas of significant overlap or disparity. Additionally, bar plots were utilized to visualize the frequency of mole/growth cases categorized by Fitzpatrick skin types within the SCIN dataset, clearly illustrating trends and distributions across skin tones.

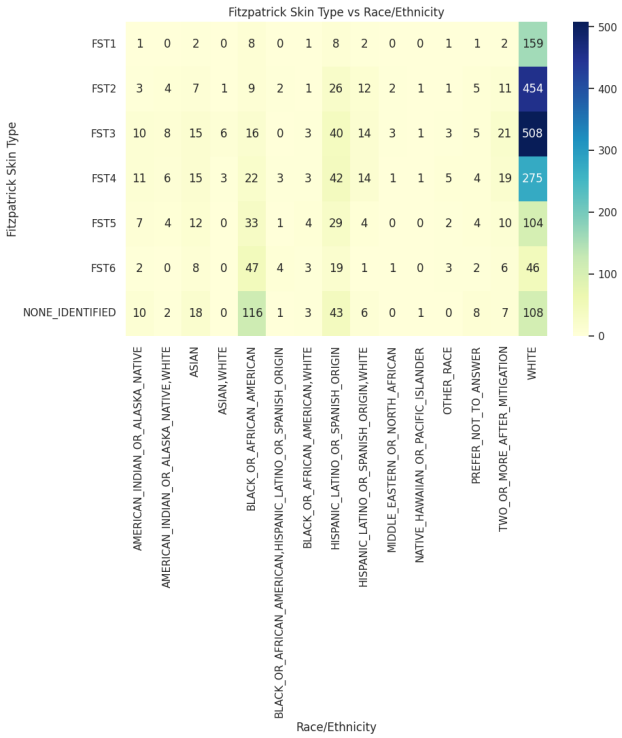

Fig. 3: Correlation analysis of Fitzpatrick Skin Type vs Race/Ethnicity

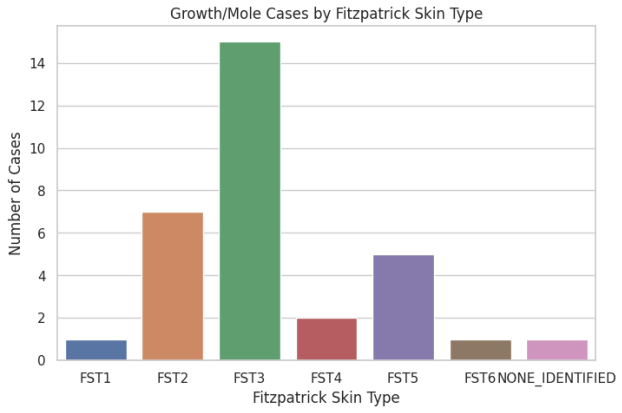

Fig. 4: Fitzpatrick Skin Types

For global mapping and comparative visualization, a choropleth digital map was created integrating ATLAS melanoma incidence data and SCIN-derived dominant race/ethnicity proportions per continent. Each country was color-coded according to the dominant race identified through SCIN data proportions per continent, with color intensity reflecting melanoma incidence severity, providing intuitive visual insights into potential racial/ethnic disparities in melanoma incidence globally.

*2) **RGB Colour Analysis**:* RGB Analysis will be fed into the machine learning model, which will produce results that can determine which skin type is easier to detect for skin cancer.Two approaches were used:

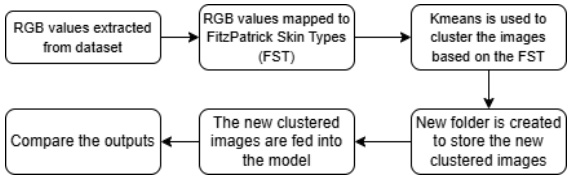

Fig. 5: Pipeline 1

Pipeline 1 used Python's OpenCV library to extract RGB values from the image dataset. The first step in the extraction process is to convert the image from BGR to RGB format.

A Gaussian blur is then applied to eliminate noise and small features. After converting the image to grayscale, a binary mask is created that uses a threshold value to identify brighter areas that are probably skin patches.

The important skin areas in the blurred RGB image are isolated using this mask, and the pixel values of those areas are extracted. If no skin pixels are found, [0, 0, 0] is returned. The average skin color is then given as a NumPy array after calculating the mean RGB values of the recognized pixels.These mean RGB values' average brightness were calculated,and assigned accordingly to the FitzPatrick Scale (Fig. 4), since it is predicated on the way the skin responds to sunlight, which is closely related to the amount of melanin, thus influencing skin brightness. [15] For the practical purpose of calculating melanin from photos, RGB brightness is used. Every class has a corresponding range of brightness values.

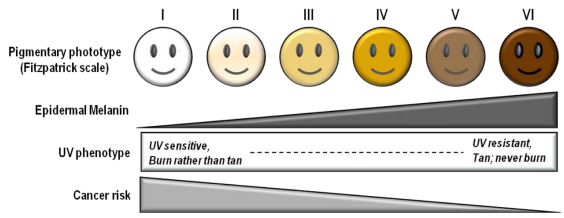

Fig. 6: FitzPatrick Scale

Unsupervised clustering was carried out using the K-Means technique to group the photos corresponding to their skin type. The six Fitzpatrick Skin Types

(FST1–FST6) were represented by the number of clusters, which was set at six. Each image was linked to one of the six clusters that was created from all average skin RGB values. The brightness-based logic previously explained was then used to translate each cluster center, which represented the average RGB of its group, to a Fitzpatrick type. Hence, each cluster index was successfully converted into a Fitzpatrick label (e.g., "Type II: Fair skin"), allowing skin tones in the dataset to be automatically categorised into a folder, which will be fed into the machine learning model(s).

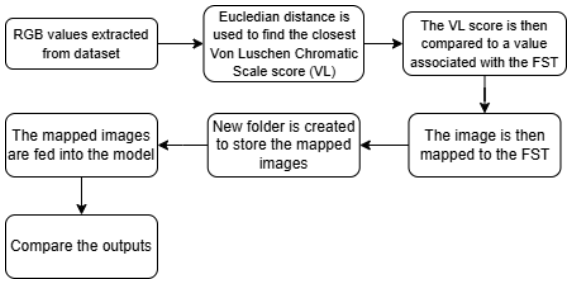

Fig. 7: Pipeline 2

Similar to pipeline 1 (Fig.3), OpenCV library was used for RGB value extraction. However, instead of using brightness mapping and K-means clustering, a different colour scale is used for bridging the image values to the FitzPatrick Skin Types (FST). This colour scale is the 'Von Luschan's Chromatic Scale' (VLS), developed by Felix von Luschan in 1897 by using coloured tablets or tiles of different colours and hues with which the colours of unexposed skin were matched. [22]

This led to the categorisation of skin tones into 36 distinct shades. While the FitzPatrick Scale simplifies classifications into six broader categories for clinical purposes, the VLS offers finer gradients, making it ideal for detailed skin tone analysis and studies of human diversity. [23], [24]

In pipeline 2, based on the Fig. 6, VLS RGB values will be linked to the FST Labels.[25] In order to achieve that, the values from Fig. 6, is used to calculate the Euclidean Distance between the RGB values of the dataset (R1, G1, B1), and the reference shades (R2, G2, B2).

$$D = \sqrt{(R_1 - R_2)^2 + (G_1 - G_2)^2 + (B_1 - B_2)^2} \quad (2)$$

When the closest match is identified, a Von Luschan Score (Fig. 6) is assigned, which is then mapped to one of six FitzPatrick Skin Types, ranging from Type 1 for pale, sensitive skin to Type 6 for deeply pigmented skin. These mapped images will be stored in a folder according to their FST Labels and will be fed into the machine learning model. Out of the 2 pipelines, pipeline

Von Luschan Score to Fitzpatrick Table with Color Reference

Von Luschan Score	RGB	Color Sample	Fitzpatrick Label
0 – 6	R244 G231 B44		FST I
7 – 13	R235 G202 B46		FST II
14 – 20	R201 G184 B22		FST IV
28 – 34	R162 G92 B92		FST V
35 +	R49 G33 B3		FST VI

Fig. 8: Mapping of Von Luschan skin tone scale to Flitzpatrick skin type classifications

1 (Fig.3) seemed to give better results than pipeline 2 (Fig. 5), hence pipeline 1 was used in integration.

B. Model Architecture and Justification

1) ***Data Preprocessing****:* To ensure reproducibility across training sessions, a fixed random seed ($seed = 42$) was set for both NumPy and TensorFlow libraries. This prevents fluctuations in results due to stochastic operations such as data shuffling, weight initialization, and augmentation, thereby enabling a fair comparison between models.

All input dermoscopic images were resized to 224×224 pixels to meet the input dimension requirements of the selected pre-trained architectures. Contrast enhancement was applied using CLAHE (Contrast Limited Adaptive Histogram Equalization) on the L-channel in the LAB color space. This technique improves lesion visibility by enhancing local contrast while preventing noise amplification, particularly in low-contrast skin regions.

The categorical labels ($'benign'$, $'malignant'$) were encoded numerically (0 and 1, respectively) using label encoding. Subsequently, the image data were normalized per channel using the training set's mean and standard deviation to center the input distribution. This normalization ensures stable convergence and prevents dominance of certain pixel ranges during training.

The dataset was then split into training and validation sets using a stratified 90/10 split to preserve class balance across subsets. TensorFlow's `tf.data` pipeline was used to efficiently load, preprocess, shuffle, and batch the data for training. A 90/10 stratified train-validation split was adopted to maximize training data availability while preserving class balance. This configuration is justified by the use of a separate skin tone–clustered test set, ensuring that the validation data remains unbiased and distinct from final model evaluation, thereby preventing

data leakage and enabling fair performance comparison across skin types.

*2) **Data Augmentation**:* To enhance model generalization and mitigate the risk of overfitting, real-time data augmentation techniques were employed during training using the `tf.image` module within TensorFlow's `tf.data` pipeline. Augmentations were applied stochastically on-the-fly, ensuring that each training epoch encountered a unique representation of the input data.

The augmentation strategy included the following transformations:

- **Random horizontal and vertical flipping**, simulating natural variations in lesion orientation.
- **Brightness modulation** with a maximum delta of $\pm 30\%$, accounting for differences in lighting conditions during image capture.
- **Contrast adjustments** within the range of 0.8 to 1.6, improving robustness to global illumination changes.
- **Saturation and hue shifts**, introducing color perturbations that mimic variations in skin pigmentation and imaging equipment.

These augmentation techniques serve to artificially expand the diversity of the dataset without requiring additional data collection. By simulating real-world imaging conditions, the model becomes more invariant to color, lighting, and spatial distortions, which is critical in clinical environments where such variability is common.

All augmentations were integrated directly into the training pipeline, ensuring efficient GPU utilization and preventing memory overhead.

*3) **Model Training**:* To address the task of melanoma classification from dermoscopic images, this study adopts a transfer learning-based approach using five well-established convolutional neural networks (CNNs): VGG16, ResNet50, Xception, DenseNet201, and MobileNetV2.

These models were selected due to their proven efficacy in image classification, particularly in medical domains, and their architectural diversity, which allows comprehensive evaluation across complexity, performance, and computational efficiency.

Each model was pretrained on the ImageNet dataset and repurposed for binary classification (benign vs. malignant). This transfer learning strategy enables leveraging low-level visual features learned from large-scale natural image data, addressing the limitations of small annotated medical image datasets

*4) **Workflow and Transfer Learning Strategy**:* The transfer learning strategy involved modifying each pretrained base model by removing its original classification layers and attaching a custom classification head. This process included two main components:

- **Feature Extractor:** The pretrained model was used as a fixed feature extractor by freezing its initial layers to preserve the learned representations.
- **Custom Classifier Head:** A new classification head was appended, consisting of the following layers:
  - `GlobalAveragePooling2D` - Condenses each feature map into a single value, which helps reduce model size and overfitting.
  - `Dense Layer` (128 units, ReLU activation) - Learns high-level, non-linear representations of the input features.
  - `Dropout Layer` (rate = 0.5) - Randomly disables 50% of neurons during training to help prevent overfitting.
  - `Dense Output Layer` (2 units, Softmax activation) - Produces class probabilities for distinguishing between benign and malignant cases.

These components create a minimal yet effective classification head adaptable to different pretrained backbones.

*5) **Key Hyperparameters and Configuration**:* Selected hyperparameters and their corresponding rationales are as follows:

TABLE I: Hyperparameter Settings

Hyperparameter	Value	Rationale
Optimizer	Adam (1×10^{-4})	Adaptive learning rate, efficient with sparse gradients.
Loss Function	Categorical Crossentropy	Suitable for multi-class classification using softmax output.
Batch Size	32	Balances training speed and memory usage.
Epochs	10	Reduces risk of overfitting during initial experimentation.
Input Size	$224 \times 224 \times 3$	Standard size for ImageNet pretrained models.
Dropout Rate	0.5	Mitigates overfitting in dense layers.
Activation	ReLU, Softmax	ReLU enables non-linearity; softmax supports class probability outputs.

*6) **Model Specific Justification**:*

I. **VGG16:** A deep, sequential architecture composed of stacked 3×3 convolutional layers followed by max-pooling. It offers a strong baseline due to its simplicity and historical success in vision benchmarks, especially in medical imaging tasks.

[26] **Key Layers:** Conv2D $\rightarrow$ MaxPooling2D ($\times$5 blocks) $\rightarrow$ Flatten $\rightarrow$ Dense
Limitation: High parameter count ($\sim$138M); lacks depthwise optimization.

II. **ResNet50:** Introduces residual connections to address the vanishing gradient problem. These skip

connections support identity mappings, enabling deeper architectures to preserve both low- and high-level features.

[18] **Key Layers:** Conv2D + BatchNorm + ReLU + Identity/Projection Skip Connections
Strength: Facilitates deep learning without degradation in performance.

III. **Xception:** Leverages depthwise separable convolutions for efficient feature extraction. This architecture is especially effective in dermoscopic image analysis due to its ability to capture fine spatial details.

[27] **Key Layers:** DepthwiseConv2D $\rightarrow$ PointwiseConv2D $\rightarrow$ ReLU + Residual
Strength: High representational power with lower computational cost.

IV. **DenseNet201:** Employs dense connectivity where each layer receives input from all previous layers. This improves feature propagation, reuse, and gradient flow, aiding in capturing subtle lesion textures.

[19] **Key Layers:** DenseBlock $\rightarrow$ Transition Layer (1×1 Conv + AvgPooling)
Strength: Promotes compact, efficient models with strong gradient behavior.

V. **MobileNetV2:** Designed for resource-constrained environments, this model uses inverted residuals and linear bottlenecks. It is ideal for lightweight deployments like mobile diagnostics.

[27] **Key Layers:** 1×1 Conv $\rightarrow$ DepthwiseConv $\rightarrow$ 1×1 Linear + Residual
Strength: Efficient and accurate with minimal computational load.

IV. Result

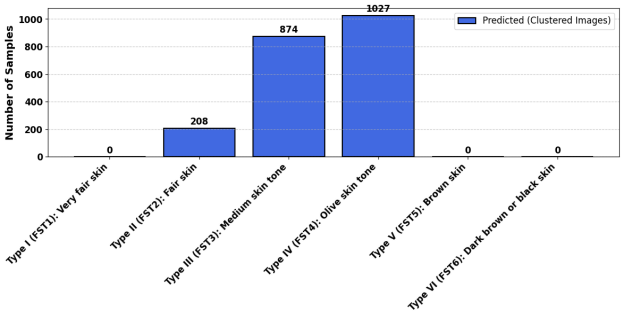

Fig. 9: Cluster image of skin tones

Based on expected RGB attributes, the distribution of clustered skin photos among Fitzpatrick skin types is displayed in the bar chart. Type IV (olive skin tone) accounts for 1,027 of the samples, followed by Type III (medium skin tone) with 874 and Type II (fair skin) with 208. A dataset imbalance or lack of diverse skin tones could be the reason why no photos were categorized as Type I (Very fair), Type V (Brown), or Type VI (Dark brown or black). Overall, the findings show that the dataset has a good representation of medium to olive skin tones.

a) Model Performance: Five CNN models were evaluated for binary classification of skin lesions. Xception achieved the highest validation accuracy (83.88%), followed by DenseNet201 (82.93%). MobileNetV2 and VGG16 also performed well with over 79% accuracy, while ResNet50 trailed slightly behind. These results confirm the effectiveness of transfer learning with pre-trained models for medical image analysis.

TABLE II: Validation accuracy comparison of different models

Model	Best Validation Accuracy (%)
Xception	83.88
DenseNet201	82.93
MobileNetV2	80.09
VGG16	79.62
ResNet50	77.73

b) Skin Tone Group Analysis: Models were tested on three Fitzpatrick skin types - FST II (Fair), FST III (Medium), and FST IV (Olive). All models showed higher accuracy on FST III and IV than on FST II. For example, VGG16 improved from 63.94% on FST II to 80.21% on FST III. This indicates reduced performance on lighter skin tones and suggests a need for more balanced data across all skin types.

c) Key Insight: The best-performing models (Xception, DenseNet201) generalize well across most skin tones but exhibit variability on fair skin. Addressing this requires more representative training data and targeted augmentation strategies

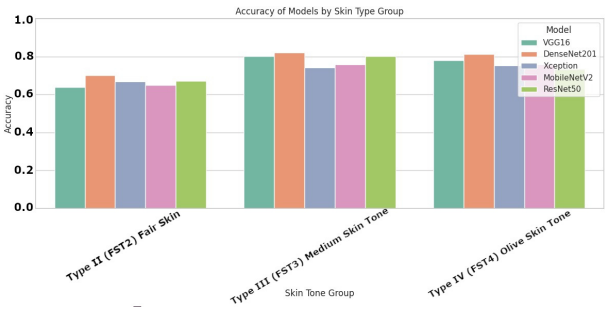

Fig. 10: Bar chart showing model performance across different skin types.

V. Discussion

The RGB analysis reveals differences in the distribution of skin tones, with dataset imbalances causing

olive tones to predominate. The absence of various complexions in the SCIN dataset has an impact on model accuracy, highlighting the need for more inclusive data. Depthwise separable convolutions provide Xception its strong performance, although fair skin recognition is hampered by low lesion contrast and underrepresentation. Performance may have been hampered by the model training being restricted to 10 epochs due to GPU limitations. Subsequent investigations should focus on optimizing models, integrating explainability techniques, broadening assessment to encompass all skin types, and utilizing a variety of clinically validated datasets. [28], [29]

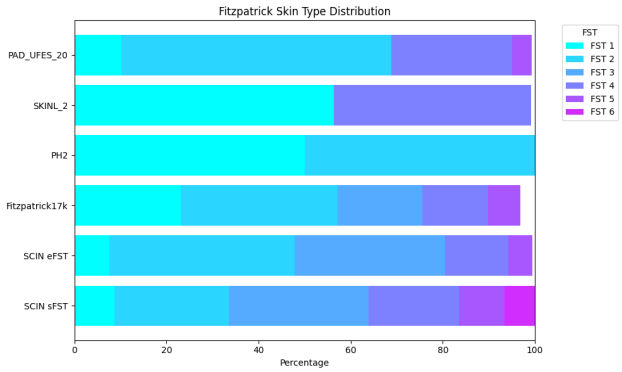

Fig. 11: Fitzpatrick Skin Type Distribution Imbalance

VI. Conclusion

Amongst the 5 machine learning models, Xception and DenseNet201 had the best performance especially with medium and olive skin tones, compared to the rest. This proves that skin types, along with the quantity of those skin types used in machine learning, play a vital role in cancer detection.Thus the collaboration between skin cancer detection and the categorisation of skin tones gives an approach that will be a benchmark for diversifying the datasets in order to have a seamless, smart, and accessible detection.

References

[1] S. Writer, "The Cancer Miracle Isn't a Cure. It's Prevention," Harvard T.H. Chan School of Public Health, Nov. 22, 2024. [Online]. Available: https://hsph.harvard.edu/news/the-cancer-miracle-isnt-a-cure-its-prevention-2/

[2] C. R. Uk, "A change of perspective for prevention?," Cancer Research UK - Cancer News, Mar. 13, 2025. [Online]. Available: https://news.cancerresearchuk.org/2025/03/12/a-change-of-perspective-for-prevention/

[3] "Cancer: Facts about the diseases that cause out-of-control cell growth." [Online]. Available: https://www.msn.com/en-us/health/other/cancer-facts-about-the-diseases-that-cause-out-of-control-cell-growth/ar-AA1DtZis

[4] "Cancer statistics for the UK," Cancer Research UK, Mar. 28, 2025. [Online]. Available: https://www.cancerresearchuk.org/health-professional/cancer-statistics-for-the-uk

[5] The Skin Cancer Foundation, "Early Detection - the Skin Cancer Foundation," Mar. 14, 2025. [Online]. Available: https://www.skincancer.org/early-detection/

[6] "Skin - anatomy, structure, diagram, function, significance." [Online]. Available: https://anatomy.co.uk/skin

[7] "MELANOMA FACTS & STATS," Melanoma UK. [Online]. Available: https://www.melanomauk.org.uk/melanoma-facts-stats

[8] SkinCancer.net, "The big three for diagnostic error," Sep. 17, 2019. [Online]. Available: https://skincancer.net/clinical/misdiagnosis

[9] J. Ellison, "Many pathologists agree overdiagnosis of skin cancer happens, but don't change diagnosis behavior," UW News, May 03, 2022. [Online]. Available: https://www.washington.edu/news/2022/05/03/many-pathologists-agree-overdiagnosis-of-skin-cancer-happens-but-dont-change-diagnosis-behavior/

[10] S. Courtaux, "What you need to know about claiming for misdiagnosed skin cancer," RWK Goodman, May 09, 2024. [Online]. Available: https://www.rwkgoodman.com/info-hub/what-you-need-to-know-about-claiming-for-misdiagnosed-skin-cancer/

[11] NICE, "Artificial intelligence (AI) technologies for assessing and triaging skin lesions referred to the urgent suspected skin cancer pathway: early value assessment," May 01, 2025. [Online]. Available: https://www.nice.org.uk/guidance/HTE24/chapter/3-committee-discussion

[12] A. Aquil et al., "Early detection of skin diseases across diverse skin tones using hybrid machine learning and deep learning models," Information, vol. 16, no. 2, p. 152, Feb. 2025. [Online]. Available: https://www.mdpi.com/2078-2489/16/2/152

[13] "ISIC Melanoma Dataset — IEEE DataPort." Available: https://ieee-dataport.org/documents/isic-melanoma-dataset

[14] GeeksforGeeks, "RGB color model in Python," Apr. 14, 2025. [Online]. Available: https://www.geeksforgeeks.org/rgb-color-model-in-python/

[15] J. D'Orazio et al., "UV radiation and the skin," Int. J. Mol. Sci., vol. 14, no. 6, pp. 12222–12248, Jun. 2013. [Online]. Available: https://www.mdpi.com/1422-0067/14/6/12222

[16] C. T. Chang et al., "DDI-2: A diverse Skin Condition image Dataset representing Self-Identified Asian Patients," J. Investig. Dermatol., Oct. 2024. [Online]. Available: https://www.jidonline.org/article/S0022-202X(24)02868-9/fulltext

[17] "The Cancer Atlas." [Online]. Available: https://canceratlas.cancer.org/data/map/

[18] N. A. W. Murdiyanto et al., "Performance comparison of CNN and ResNet50 for skin cancer classification using U-Net segmented images," Indones. J. Data Sci., vol. 5, no. 3, pp. 198–205, Dec. 2024. [Online]. Available: https://www.jurnal.yoctobrain.org/index.php/ijodas/article/view/200

[19] "Skin lesion classification using Pre-Trained DenseNET201 Deep neural Network," IEEE Conf. Publ. [Online]. Available: https://ieeexplore.ieee.org/document/9451818

[20] B. C. R. S. Furriel et al., "Artificial intelligence for skin cancer detection and classification for clinical environment: a systematic review," Front. Med., vol. 10, Jan. 2024. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fmed.2023.1305954

[21] F. Grignaffini et al., "Machine Learning Approaches for Skin Cancer Classification from Dermoscopic Images: A Systematic Review," Algorithms, vol. 15, no. 11, p. 438, Nov. 2022. [Online]. Available: https://www.mdpi.com/1999-4893/15/11/438

[22] N. G. Jablonski, "The evolution of human skin and skin color," Annual Review of Anthropology, vol. 33, pp. 585–623, 2004, Available: https://www.jstor.org/stable/25064866

[23] "Is there a spec for RGB values of the emoji skin tone modifiers?," Stack Overflow. [Online]. Available: https://stackoverflow.com/questions/75489304/is-there-a-spec-for-rgb-values-of-the-emoji-skin-tone-modifiers

[24] "Von Luschan's chromatic scale," DBpedia, 2025. https://dbpedia.org/page/Von Luschan

[25] V. R. Weir et al., "A survey of skin tone assessment in prospective research," npj Digit. Med.,

vol. 7, no. 1, Jul. 2024. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC11252344/
[26] V. Anand et al., "An enhanced transfer learning based classification for diagnosis of skin cancer," Diagnostics, vol. 12, no. 7, p. 1628, Jul. 2022. [Online]. Available: https://www.mdpi.com/2075-4418/12/7/1628
[27] H. F. Hassan and S. T. Ozer, "A comparative study of deep learning models for skin cancer detection: Leveraging Transfer Learning," Math. Model. Eng. Probl., vol. 12, no. 1, pp. 166–180, Jan. 2025. [Online]. Available: https://iieta.org/journals/mmep/paper/10.18280/mmep.120119
[28] "Datasets used to train AI to detect skin cancer lack information on darker skin and often incomplete," Haiku, Nov. 10, 2021. [Online]. Available: https://www.cancer.ox.ac.uk/news/lack-of-data-available-to-detect-skin-cancer-in-darker-skinsite-content
[29] "Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization," IEEE Conf. Publ. [Online]. Available: https://ieeexplore.ieee.org/document/8237336

Credit Statement

Akshita Bhatia: Conceptualisation, Methodology, Investigation, Writing, Visualisation, Supervision Resource and Project Administration **Elakia V M:**Methodology, Software, Validation, Formal analysis, Writing [Methodologies,Results,Discussions] and Visualization **Rustam Guliyev:**Methodology, Software, Validation, Writing [Methodologies, Report Editing in Latex], and Visualizations **Andrzej Kolodziej:**Methodology, Software, Validation and Visualisation.