**Introduction**

Melanoma is one of the most aggressive and life-threatening forms of skin cancer, making early and accurate detection crucial for effective treatment and improved survival rates. Traditional diagnostic methods rely on expert dermatologists, but they are time-consuming, subjective, and prone to variability. With advancements in artificial intelligence and deep learning, automated classification of skin lesions from dermoscopic images has emerged as a promising solution. By leveraging state-of-the-art computer vision techniques, such as convolutional neural networks (CNNs), vision transformers (ViTs), and U-Net variants, this project aims to enhance the accuracy and efficiency of melanoma detection.

**Problem statement**

1. **Challenges in Early Detection:** Melanoma, the deadliest form of skin cancer, can be difficult to distinguish from benign lesions in its early stages, leading to missed or delayed diagnoses.
2. **Dependence on Specialized Expertise:** Accurate diagnosis requires highly trained dermatologists, but access to specialized medical professionals is limited, especially in remote and underserved areas.
3. **Variability in Diagnosis:** Manual assessment of dermoscopic images can be subjective, leading to inconsistencies in diagnosis and potential misclassification of malignant and benign lesions.
4. **Need for Scalable and Reliable Solutions:** An AI-driven system can provide an objective, scalable, and automated approach to skin lesion classification, reducing human error and improving early melanoma detection rates.

**Objectives**

1. **Develop an AI-based Classification Model:** Design and implement a deep learning model to accurately classify skin lesions from dermoscopic images, distinguishing between benign and malignant cases.
2. **Enhance Diagnostic Accuracy and Consistency:** Improve melanoma detection by leveraging advanced computer vision techniques, reducing diagnostic subjectivity and human error.
3. **Optimize Preprocessing and Feature Extraction:** Apply image enhancement, augmentation, and segmentation techniques to improve model robustness and performance.
4. **Support Early Detection and Clinical Decision-Making:** Provide a reliable, scalable, and automated tool to assist dermatologists in early melanoma detection, enabling faster and more effective treatment planning.

**Literature review**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **No.** | **Title** | **Authors** | **Year** | **Methodology** | **Conclusion** | **Drawbacks** |
| **1.** | Hybrid Deep Learning Framework for Melanoma Diagnosis | Al-antari et al. | 2024 | Combines U-Net for segmentation, Inception-ResNet-v2 for feature extraction, and ViT for classification. | Achieved 98.65% accuracy, 99.20% sensitivity, and 98.03% specificity on the ISIC2020 dataset. | Limited to a single dataset, which may affect generalizability. |
| **2.** | Enhancing Melanoma Diagnosis with Advanced Deep Learning Models | 2Serra Aksoy,Ismail Bogrekci | 2024 | Uses ConvNeXt, ViT Base-16, and Swin Transformer V2 Small for classifying melanoma. | The study utilized ConvNeXt, Vision Transformer (ViT) Base-16, and Swin Transformer V2 Small (Swin V2 S), demonstrating the efficacy of state-of-the-art techniques in enhancing diagnostic accuracy. ConvNeXt Base stood out as the most robust model, | Issues such as image acquisition conditions, resolution, and the explainability and uncertainty of the models were not covered in this study. |
| **3.** | Segmentation and Classification of Dermoscopic Skin Images Using U-Net and Handcrafted Features | Abouche et al. | 2024 | Uses U-Net for segmentation, Dull Razor algorithm for hair removal, and handcrafted features for classification. | Competitive performance compared to state-of-the-art methods. | Handcrafted feature extraction may reduce generalization across lesion types. |

**Methodologies**

### **1: Vision Transformers (ViT) for Skin Lesion Segmentation**

**Why Use ViT?**

* Utilizes attention mechanisms for superior image segmentation performance.
* Captures long-range dependencies, improving feature extraction.

**Popular ViT Models:**

* **Swin Transformer (Swin-Tiny, Swin-Large)** – Hierarchical transformers with localized attention.
* **ViT-Google** – Outperforms other ViT models in skin lesion segmentation.
* **ViT-MAE** – Effective for segmentation but may not surpass ViT-Google.
* **ViT-ResNet50** – Combines transformer and CNN strengths for improved results.

**Implementation:**

* **Pre-trained Weights**: ImageNet-21k
* **Frameworks**: Hugging Face Transformers, PyTorch, TensorFlow

### **2**. **U-Net Variants for Skin Lesion Segmentation**

**Why Use U-Net?**

* Flexible and adaptable for different datasets and image resolutions.
* More computationally efficient than ViTs, making them ideal for real-time applications.

**Key U-Net Variants:**

* **Attention U-Net** – Integrates attention mechanisms for enhanced feature selection.
* **Dilated U-Net** – Uses dilated convolutions for a larger receptive field.
* **Residual U-Net** – Employs residual connections to ease training and enhance learning.

**Implementation:**

* **Frameworks**: PyTorch, TensorFlow/Keras
* **Libraries**: MONAI for medical image processing

**Conclusion**

1. **Enhanced Early Detection:** AI-driven classification improves melanoma detection accuracy and enables timely treatment.
2. **Reduced Diagnostic Variability:** Deep learning models provide consistent, objective, and reliable analysis.
3. **Increased Accessibility:** Automated systems can assist in remote areas with limited dermatological expertise.
4. **Scope for Improvement:** Future work includes optimizing models, integrating more data, and improving real-time performance.