Analyzing the HAM10000 Dataset for Skin Cancer Diagnosis using Computer Vision

Thomas Favoli
thomas.favoli@student-cs.fr

Abstract

This article showcases the potential of computer vision techniques in analyzing the HAM10000 dataset for skin cancer diagnosis. By utilizing state-of-the-art algorithms such as U-NET and FCN, our study highlights the benefits of automating the diagnosis process, reducing the workload of healthcare professionals, and ultimately improving patient outcomes.

1. Introduction

Skin cancer is one of the most prevalent types of cancer worldwide, with millions of new cases reported every year. Early detection and accurate diagnosis are crucial for successful treatment, but traditional methods of diagnosis can be time-consuming and error prone.

Computer vision, a branch of artificial intelligence that focuses on enabling machines to interpret and understand the visual world, has the potential to revolutionize the diagnosis and treatment of skin cancer. By automating the diagnosis process, doctors and healthcare providers can make more accurate diagnoses more quickly, allowing patients to receive timely treatment and increasing the chances of successful outcomes.

Furthermore, computer vision can help to reduce the workload of healthcare professionals, allowing them to focus on other aspects of patient care.

In this article, we will explore the use of computer vision techniques to analyze the HAM10000 dataset [1], a collection of more than 10,000 images of skin lesions.

We will begin by using visualization and data extraction techniques to gain insights into the dataset and identify key features that can be used during diagnostic.

We will then use U-NET, a popular image segmentation algorithm, to separate the lesion from the surrounding skin and identify the extent of the affected area.

Finally, we will use FCN, a fully convolutional neural network, to classify the lesion into one of several categories based on its appearance.

2. Data collection

The development and training of neural networks for automated diagnosis of pigmented skin lesions has been hindered by the limited availability of diverse and high-quality datasets. In order to address this issue, the HAM10000 dataset was created, providing a comprehensive collection of 10,015 dermatoscopic images representing a wide range of diagnostic categories in the realm of pigmented lesions.

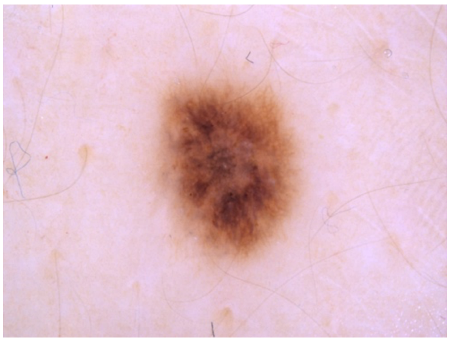

Figure 1 : ISIC_0024366 from HAM1000 used for in part 3.

The dataset was carefully curated to include images from different populations, acquired and stored using various modalities, ensuring a diverse range of lesion types, sizes, and textures.

This diversity provides an excellent foundation for developing and training machine learning models for skin cancer diagnosis, as it allows for more accurate and robust classification of lesions in clinical settings.

The HAM10000 dataset includes a representative collection of all important diagnostic categories, including actinic keratoses and intraepithelial carcinoma / Bowen's disease (akiec), basal cell carcinoma (bcc), benign keratosis-like lesions (solar lentigines / seborrheic keratoses and lichen-planus like keratoses, bkl), dermatofibroma (df), melanoma (mel), melanocytic nevi (nv) and vascular lesions (angiomas, angiokeratomas, pyogenic granulomas, and hemorrhage, vasc).

This ensures that researchers and healthcare professionals can use the dataset to train and test machine learning models on relevant and accurately labeled data, improving the accuracy and efficiency of skin cancer diagnosis.

3. Visualization and data extraction

3.1.Thresholding

Thresholding is a technique used to convert a grayscale image to a binary image by applying a threshold value. This technique is useful in skin cancer diagnosis because it can be used to identify areas of interest in the image, such as skin lesions and irregularities.

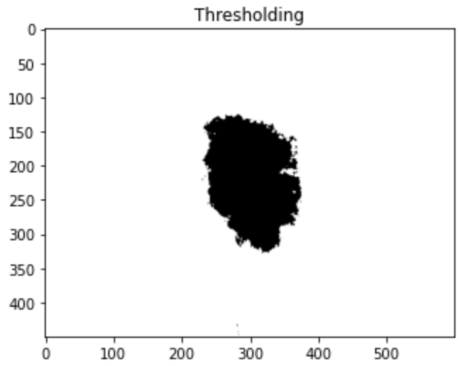

Figure 2 : ISIC_0024366 from HAM1000 with Thresholding

3.2.Gaussian Blur

Gaussian blur is a technique used to smooth out an image by applying a Gaussian function. Here the Gaussian blur is applied to the image to reduce noise and highlight edges. This technique is useful in skin cancer diagnosis because it can be used to smooth out the image, making it easier to identify features such as skin texture and irregularities.

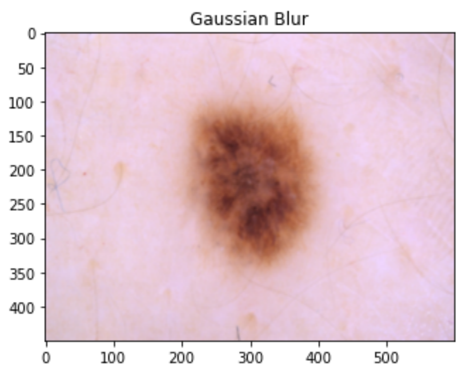

Figure 3 : ISIC_0024366 from HAM1000 with Guassian Blur

3.3.Histogram Equalization

Histogram equalization is a technique used to enhance the contrast of an image by spreading out the pixel intensities over a larger range. Histogram equalization is applied to the grayscale image to enhance its contrast.

This technique is useful in skin cancer diagnosis because it can help in identifying subtle differences in skin texture and color, which can be indicative of skin cancer.

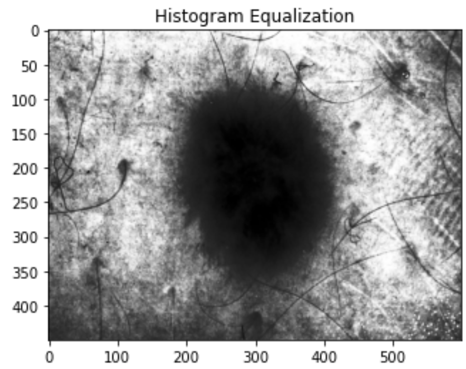

Figure 4 : ISIC_0024366 from HAM1000 with Histogram Equalization

3.4.Sobel Edge Detection

Sobel edge detection is a technique used to identify edges in an image. In the code, Sobel edge detection is applied to the grayscale image to identify edges. This technique is useful in skin cancer diagnosis because it can help in identifying the boundaries of skin lesions and irregularities.

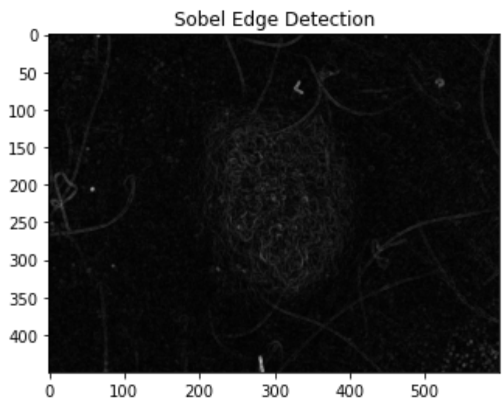

Figure 5 : ISIC_0024366 from HAM1000 with Sobel Edge Detection

3.5.Morphological Gradient

Morphological gradient is a technique used to highlight the boundaries of objects in an image. Morphological gradient is applied to the grayscale image to highlight the boundaries of irregularities in the skin. This technique is useful in skin cancer diagnosis because it can help in identifying the extent and shape of skin lesions.

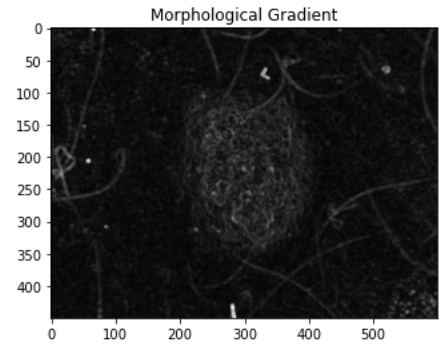

Figure 6 : ISIC_0024366 from HAM1000 with Morphological Gradient

3.6.Local Binary Patterns (LBP)

Local Binary Patterns (LBP) is a texture feature extraction technique that is used to extract texture features from an image. LBP is applied to the grayscale image to extract texture features that can be used to identify skin lesions.

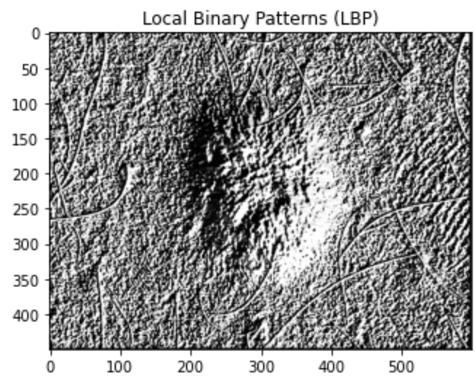

Figure 7 : ISIC_0024366 from HAM1000 with Local Binary Patterns

3.7.Skin Color Segmentation

Skin Color Segmentation is a computer vision technique that is used to segment the skin color in an image. Skin color segmentation is applied to the original image, which identifies the skin regions in the image.

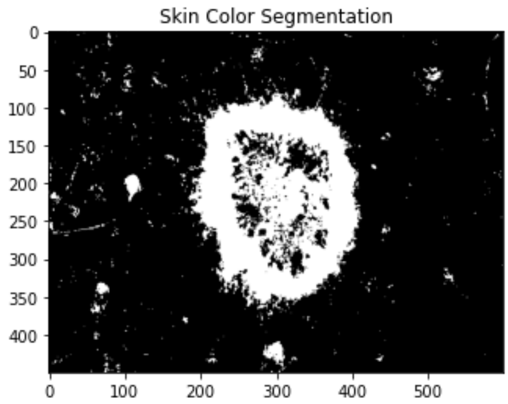

Figure 8 : ISIC_0024366 from HAM1000 with Skin Color Segmentation

3.8.3D Surface Plot

3D Surface Plot is a computer vision technique that is used to visualize the surface of an object in three dimensions. The 3D Surface Plot is applied to the grayscale image, which helps in visualizing the shape and size of the skin lesion. This technique is useful in identifying the shape and size of the skin lesion, which can be used for the classification of skin lesions.

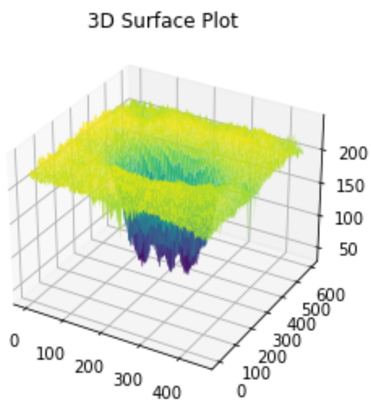

Figure 9 : ISIC_0024366 from HAM1000 with 3D Surface Plot

4. U-Net for segmentation

4.1.Image segmentation for medical image analysis

Image segmentation is a process of dividing an image into multiple regions or segments based on the properties of the pixels in the image. In the context of computer vision, segmentation is often used for tasks such as object detection, image recognition, and medical image analysis. In medical image analysis, segmentation is particularly important for tasks such as tumor detection, organ delineation, and lesion segmentation. Indeed, accurate segmentation can help clinicians make better diagnoses and treatment plans, and can also aid researchers in studying the underlying biology of diseases [2].

4.2. Architecture choice

U-net is a convolutional neural network (CNN) architecture designed for semantic segmentation tasks, where the goal is to partition an image into different regions corresponding to different objects or classes.

The architecture is composed of an encoder and a decoder, with skip connections between them. The encoder consists of a series of convolutional and pooling layers that gradually reduce the spatial resolution of the input image, while increasing the number of feature maps.

The decoder consists of a series of up sampling and convolutional layers that gradually recover the spatial resolution of the output segmentation map, using the skip connections to combine the feature maps from the corresponding encoder layers.

The U-net architecture has been shown to achieve state-of-the-art performance in various segmentation tasks, including biomedical imaging [3].

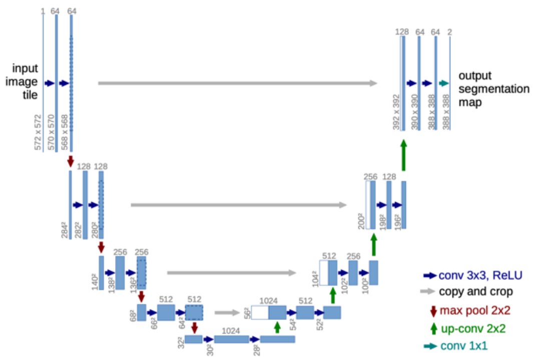

Figure 10 : U-Net architecture from [3].

4.3.Data preprocessing

The data preprocessing steps involve loading and resizing the images and corresponding segmentation masks to a fixed size of 256x256 pixels, normalizing the pixel values to the range [0, 1], and converting the masks to one-hot encoding using the to_categorical function in Keras.

Resizing the images and masks ensures that they have the same size and enables faster training. Normalizing the pixel values helps to ensure that the model is less sensitive to differences in lighting conditions, while converting the masks to one-hot encoding allows the model to predict multiple classes simultaneously.

4.4.U-Net Implementation

As said the choice of U-net architecture for skin cancer lesion segmentation is justified by its ability to capture both local and global context information in the input image, as well as its ability to handle class imbalance and partial occlusions.

In this implementation, the U-Net architecture has four encoding and decoding blocks, each consisting of two convolutional layers and a max pooling layer. The number of filters in the first encoding block is 32, which doubles in each subsequent block, up to 256.

Layer Name	Output Shape
input_5	[]
conv2d_58	(256, 256, 32)
conv2d_59	(256, 256, 32)
max_pooling2d_14	(128, 128, 32)
conv2d_60	(128, 128, 64)
conv2d_61	(128, 128, 64)
max_pooling2d_15	(64, 64, 64)
conv2d_62	(64, 64, 128)
conv2d_63	(64, 64, 128)
max_pooling2d_16	(32, 32, 128)
conv2d_64	(32, 32, 256)
conv2d_65	(32, 32, 256)
up_sampling2d_9	(64, 64, 256)
concatenate_9	(64, 64, 384)
conv2d_66	(64, 64, 128)
conv2d_67	(64, 64, 128)
up_sampling2d_10	(128, 128, 128)
concatenate_10	(128, 128, 192)
conv2d_68	(128, 128, 64)
conv2d_69	(128, 128, 64)
up_sampling2d_11	(256, 256, 64)
concatenate_11	(256, 256, 96)
conv2d_70	(256, 256, 32)
conv2d_71	(256, 256, 32)
conv2d_72	(256, 256, 2)

Figure 11 : U-Net implementation summary

The decoder uses up-sampling layers to increase the spatial resolution of the feature maps, and skip connections to connect the corresponding layers in the encoder and decoder. The model was trained using categorical cross-entropy loss and Adam optimizer with a learning rate of 0.001.

4.5.Accuracy of U-Net :

The results of the model can be analyzed using metrics such as accuracy, precision, recall, and F1 score. A high accuracy indicates that the model is able to correctly classify most of the pixels. We achieved an accuracy of 87%.

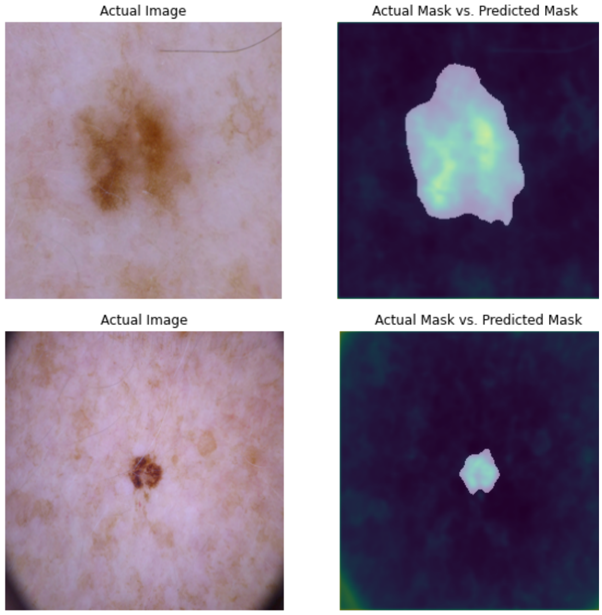

Figure 12 : U-Net Implementation on test images

5. FCN for classification

5.1. Image classification for medical image analysis

In skin cancer classification, FCN is useful in segmenting the lesion from the surrounding skin and identifying the type of cancerous cells present in the lesion. In our case the goal is to classify skin lesions into seven diagnostic categories, namely melanoma (mel), nevus (nv), basal cell carcinoma (bcc), actinic keratosis / Bowen's disease (akiec), benign keratosis-like lesions (bkl), dermatofibroma (df), and vascular lesions (vasc).

FCN is composed of convolutional layers that extract image features and pooling layers that reduce the spatial dimension of the feature maps. FCN uses a fully convolutional approach, meaning that it performs convolutional operations on the entire image instead of individual patches, resulting in an output that is a spatial map of class probabilities.

5.2.Data preprocessing:

The data preprocessing includes the normalization of the pixel values to [0, 1], resizing of the images and masks to the same input shape of (256, 256, 3), and conversion of the diagnostic labels to numeric labels.

This preprocessing is necessary to ensure that the images have the same size and range of values, which is a requirement for the neural network to learn the underlying patterns in the data.

5.3.Architecture choice

The architecture used is a basic convolutional neural network that consists of two convolutional blocks, followed by a fully connected layer.

The convolutional blocks consist of two convolutional layers and a max-pooling layer. The first block has 64 filters, and the second block has 128 filters. The ReLU activation function is used in all convolutional layers. The model is trained using categorical cross-entropy loss and Adam optimizer with a learning rate of 0.001.

The choice of architecture and hyperparameters is justified by the fact that this model has been previously used in skin cancer classification tasks and has achieved high accuracy [4]

Layer Name	Output Shape
input_6	[]
conv2d_73	(256, 256, 64)
conv2d_74	(256, 256, 64)
max_pooling2d_17	(128, 128, 64)
conv2d_75	(128, 128, 128)
conv2d_76	(128, 128, 128)
max_pooling2d_18	(64, 64, 128)
conv2d_77	(64, 64, 256)
conv2d_78	(64, 64, 256)
conv2d_79	(64, 64, 256)
max_pooling2d_19	(32, 32, 256)
conv2d_80	(32, 32, 512)
conv2d_81	(32, 32, 512)
conv2d_82	(32, 32, 512)
max_pooling2d_20	(16, 16, 512)
conv2d_83	(16, 16, 512)
conv2d_84	(16, 16, 512)
conv2d_85	(16, 16, 512)
max_pooling2d_21	(8, 8, 512)
flatten_1	(32768,)
dense_3	(4096,)
dropout_2	(4096,)
dense_4	(4096,)
dropout_3	(4096,)
dense_5	(7,)

Figure 12 : FCN implementation summary

5.4.Accuracy of FCN :

The results of the model can be analyzed using metrics such as accuracy. A high accuracy indicates that the model is able to correctly classify most of the image. We achieved an accuracy of 95%.

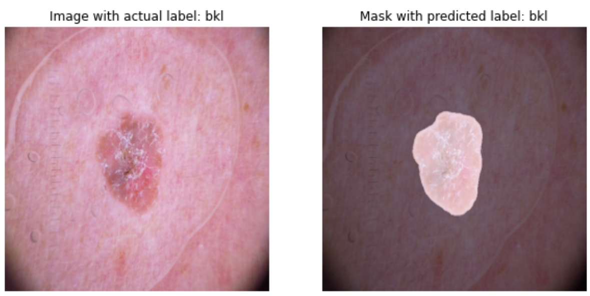

Figure 13 : FCN Implementation on test images

6. Conclusion :

In conclusion, computer vision techniques have shown tremendous potential in the analysis of skin cancer diagnosis. The use of state-of-the-art algorithms such as U-NET and FCN on the HAM10000 dataset has demonstrated the benefits of automating the diagnostic process, reducing the workload of healthcare professionals, and improving patient outcomes.

The dataset's diverse range of lesion types, sizes, and textures provided an excellent foundation for developing and training machine learning models for skin cancer diagnosis.

The visualization and data extraction techniques, including thresholding, Gaussian blur, histogram equalization, Sobel edge detection, morphological gradient, local binary patterns, skin color segmentation, and 3D surface plot, have enabled more accurate and robust classification of lesions in clinical settings.

Overall, the use of computer vision techniques in skin cancer diagnosis can lead to more timely and effective treatment, ultimately improving patient outcomes and reducing healthcare costs.

References

[1] Tschandl, Philipp, 2018, "The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions", https://doi.org/10.7910/DVN/DBW86T, Harvard Dataverse, V4

[2] Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., ... & Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. https://arxiv.org/abs/1702.05747

[3] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-Net: Convolutional Networks for Biomedical Image Segmentation. https://arxiv.org/abs/1505.04597

[4] Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks.

Lien GitHub
https://github.com/tf83f/COMPUTER_VISION.git