Classifying types of inflammation on human skin

Nazgul Mamasheva[1]

[1]Sapienza University of Rome, Piazzale Aldo Moro 5, 00185 Rome, Italy

Abstract
A vast variety of human diseases and the damages causing the human body parts go along with inflammation. Many factors can cause inflammation from a simple sunburn to a potentially serious infection. Thus, inflammation detection is a key health test. A physician first inspects the inflammation on patient's skins, makes check-ups and then prescribes the corresponding treatments. The inflammation inspection step can be automated with a program, that can help speed up the provision of medical care to patient. In this project, we propose some possible approaches to develop such automation process.

Keywords
Feature extraction, Random Forest, VGG-19

1. Highlights

- skin detection
- skin segmentation
- inflammation detection
- inflammation classification

2. Introduction

Inflammation, which is part of the body's natural healing system, helps fight injury and infection. There are two types of inflammations, acute and chronic. Acute inflammation is the early (short term) response of the body to adverse stimuli. The chronic inflammation is an inflammatory reaction that lasts for months or years. [1] Clinically, acute inflammation is characterized by 5 cardinal signs: rubor (redness), calor (increased heat), tumor (swelling), dolor (pain), and functio laesa (loss of function). The chronic inflammation occurs without cardinal signs.

The classic description of inflammation refers to the visual changes on human skin. Since there are no cardinal signs for chronic inflammation, we will be dealing with acute or short-term inflammation. Among all five cardinal signs of acute inflammation we can visually inspect the redness and swelling as visible changes on human skin. Acute inflammation can be caused by allergic reaction, chemical irritants, infection, trauma injury, burns, laceration, cuts, wounds, frostbite, bacterial pathogen, ingrown toenail, bee sting, etc. Some allergic reactions, chemical irritants and bacterial pathogen can cause same visual changes such as rashes and hives. First degree of burns and frostbites can cause swelling and redness on human skin. Infection and insect bite can cause dermatitis. In terms of visual changes, the triggers that cause inflammation can be grouped into some categories, for instance, dermatitis, hives and rashes as shown in Figure 1. So, the dermatitis, hives, and rashes can be classified as different types of inflammatory skin conditions.

Figure 1: Samples of types of inflammation on human skin. On the left - dermatitis, in center - hives, on the right - rashes.

Dermatitis is a condition that is characterized by inflammation and tissue damage of the skin as a result of chemical, biological and external effects on it. Accompanied by itching, redness, peeling of the skin. The most common cause of the disease is an adverse external effect. The irritant can be a mechanical action, excessive ultraviolet radiation, or a specific allergen.

Hives, also known as urticaria, are red or skin-colored bumps or welts that appear on the skin in varying shapes and sizes. Triggers that cause hives can include: eating certain foods or spicy food, contact with certain plants, animals, chemicals and latex, cold water or wind, sweaty skin from exercise, emotional stress or a reaction to a

mamasheva.1814656@studenti.uniroma1 (N. Mamasheva)

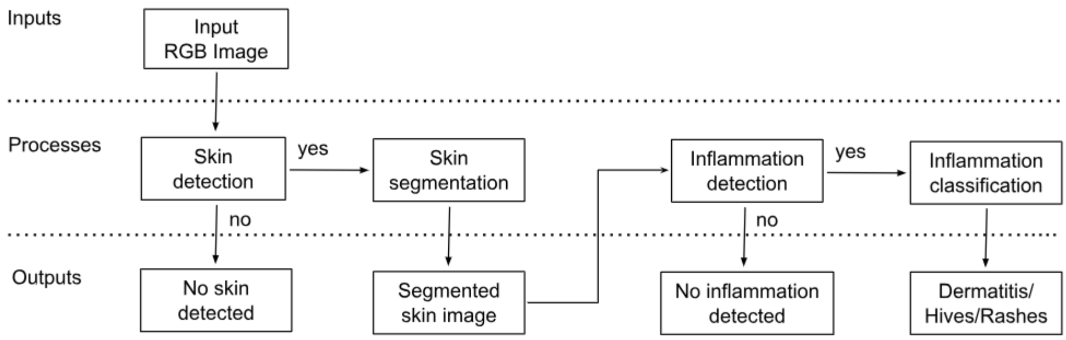

Figure 2: An automation process pipeline.

medicine, insect bite or sting, wearing itchy or tight clothing, an infection, a problem with an immune system, water or sunlight, etc.

Rashes are small red specks or pimples appearing on the body with certain triggers. The main cause is dermatitis, which is when your skin reacts to allergens or irritants. Bacteria, viruses, allergens and conditions including eczema, hives, and psoriasis can be the source of skin rashes.

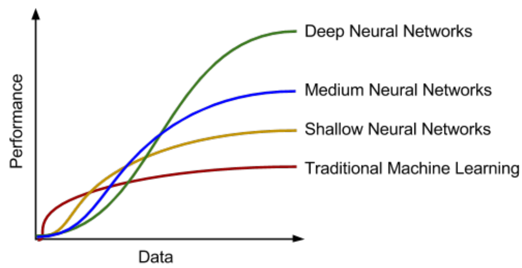

Figure 3: The choice of machine learning methods according to amount of data and performance rate.

3. Methods

The project's objective is to classify different types of skin inflammation on human skin and provide automation process of it. To accomplish this goal the automation program has to perform 4 steps: skin detection, skin segmentation, inflammation detection and inflammation classification. In automation process pipeline, that is illustrated in Figure 2, an RGB image will be given as an input by a user. The program must first define whether there is a human skin in the provided image. So, an input image will be given to skin detection task, if it detects skin in the image, then it will segment skin regions in the image and pass a segmented skin image to inflammation detection step. If an inflammation is present on the skin, then the inflammation classification step will take place and as a result it outputs the corresponding text with a name of detected type of inflammation such as "Dermatitis", "Hives" or "Rashes".

For skin detection task we implement features extraction process and utilize one of the traditional machine learning methods, Random Forest. Figure 3 illustrates how traditional machine learning method can be efficient by achieving higher performance rate than neural networks when there are a fewer data, as in our case.

The skin segmentation task is an extension of skin detection task. After training the model on skin dataset, we get a pre-trained skin detection model. We use this pre-trained model to perform a prediction on a given input image, it outputs an array that consists of 0 and/or 255 values. The value 255 in array indicates that the pixel belongs to the skin, while the value 0 indicates that the pixel is not a part of a skin. Thus, we produce a black and white image, i.e. skin mask of the given input image. We perform skin segmentation task by merging RGB input image and it's mask image.

The inflammation detection and inflammation classification tasks use same approaches with different datasets. These two tasks were implemented by using VGG-19 Convolutional Neural Network (CNN), which is trained on more than a million images from the ImageNet database. VGG is a commonly used neural network and it is one of the most prominent architectures used for image classification. To increase the size of our input dataset we perform a data augmentation technique.

For inflammation detection task we prepare a dataset with two classes of images: normal skin without inflammation and skin with inflammation. After training on this dataset we obtain a pre-trained inflammation detection model. We then employ the obtained pre-trained model to identify whether the segmented skin image contains inflammation or not.

We create a dataset for inflammation classification task with three classes: dermatitis, hives, and rashes. Similar to the previous task, we train the model on a dataset to get a pre-trained model, which we then use to identify if the inflammation on skin is dermatitis, hives, or rashes.

4. Related Works

The project is partly based on paper "Automated Detection, Extraction and Counting of Acne Lesions for Automatic Evaluation and Tracking of Acne Severity" [2]. In this work the authors discusses the development of an automated system for detecting, extracting, and counting acne lesions in images taken with mobile devices. The system could potentially provide a more objective and automated way to assess and monitor acne severity. As a part of their approach the authors performed a skin segmentation task in their work, which served to me as a guidance to implement a custom skin segmentation module. For skin detection task the authors implemented feature extraction techniques with various features like color, texture, spatial, shape, and unsupervised descriptors to achieve accurate segmentation. Based on these details I could develop a custom skin segmentation module from scratch, since the paper's code sources were not publicly available.

The work of E. Almansour and M. A. Jaffar "Classification of Dermoscopic Skin Cancer Images Using Color and Hybrid Texture Features" [3] guided me for understanding and utilizing the Gray Level Co-Occurrence Matrix (GLCM) and Local Binary Pattern for extracting texture features.

5. Implementation

5.1. Dataset

For the project work we prepared three datasets.

The first dataset is for skin detection task that consists of 144 training (72 - normal skin, 72 - inflammed skin) and 36 testing (18 - normal skin, 18 - inflammed skin) data.

The second dataset is for inflammation detection task, we prepared a dataset that consists of two classes, normal and inflammed skins. In training data there are 63 - normal skin data and 63 - inflammed skin data, for validation by 9 data for each class were proviede and for testing data 18 - skin data and 18 - inflammed data.

And the third dataset is for inflammation classification task, we provided only inflammed skin data with three classes, dermatitis, hives and rashes. For training data each of these classes were provided by 21 data, for validation - 3 and for testing - 6.

Figure 4: Samples of diverse human skin (without inflammation) in custom skin datasets.

We used ready datasets for normal skin data (without inflammation), which were provided with image-mask pairs. The images for inflamed skin were collected manually, since there were no proper and ready datasets for it. The collected data images were manually annotated by using Labelme software tool and their corresponding mask pairs were generated by a written script. All three datasets are balanced with equal amounts of data for each class. The datasets are also diverse, different types of skin tones are included as shown in Figure 1 and Figure 4.

5.2. Feature Extraction

In the domain of machine learning, features play a fundamental role as they serve as the foundational elements upon which models make their predictions. However, it's important to recognize that not all features are equally valuable for achieving desirable outcomes. In fact, for successful machine learning tasks, it's essential to focus only on features that truly impact the result of predictions. A feature, in this context, can be understood as a measurable attribute that characterizes a specific aspect of an entity, for example, a red-channel in RGB image. When data is arranged in a structured format, such as a table, these features are typically organized as columns, with each row representing a distinct instance. In table context, we consider the columns as are the predefined features and rows as pixel values on given image.

The process of feature extraction takes place after data pre-processing. During this phase, the goal is to transform raw data into numerical representations that are understandable and useful for machine learning models.

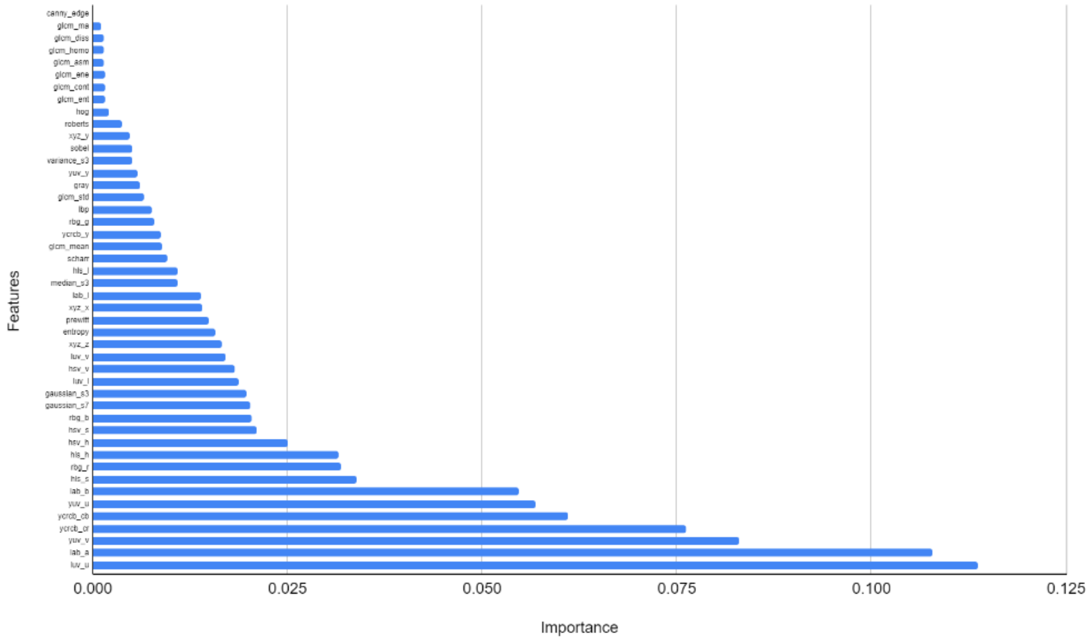

Figure 5: Features importance.

This involves identifying relevant patterns, relationships, and traits within the data that contribute to the model's ability to make accurate predictions.

5.2.1. Feature selection

Feature selection is a process in machine learning where relevant and informative features are chosen from the original set of features in a dataset. It aims to improve model efficiency, reduce overfitting, reduces computational complexity, simplifies data collection and improves model robustness. We used a feature selection technique in feature extraction. After training the model on provided features, we can get a feature list that shows us the features' importance. Then we can decide which features we should leave and which ones to eliminate. In Figure 5 we can see a chart that shows us features' importance which we defined for our feature extraction. According to this chart we can select the first 15-20 features, that show higher level of importance among all.

5.2.2. Color features

Color features in computer vision refer to the quantitative representations of colors within an image. They capture aspects like hue, saturation, and intensity, providing valuable information for tasks such as image analysis, object detection, and classification. The three main colors, such as red, green, and blue, can be studied in order to comprehend the concept of color space. In this project, the color feature technique aims to detect and characterize colors within segmented skin regions. This is achieved by extracting distinct channels from a variety of color spaces, including RGB, HSV, CIE L*u*v, CIE XYZ, YCbCr, CIE L*a*b, HSL and YUV.

5.2.3. Texture feature

Texture features in computer vision refer to patterns and variations in the visual appearance of surfaces. These patterns, like roughness or smoothness, can be quantified to help distinguish different textures in images. Texture features are important features for our feature extraction, since in our custom datasets we have not only smooth skins, but also skins with inflammation, that may have pimples, bumps and sores. For texture feature we utilized the Local Binary Pattern (LBP) and the Gray Level Co-Occurrence Matrix (GLCM) along with entropy and other features.

LBP is such type of a feature that transforms the image into an array. Hence, we have applied LBP operator on every pixel of the image in order to obtain the coded LBP image. The idea of LBP is to compare each pixel on the image with its neighbors. The procedure is as follows.

Figure 6: Samples of image augmentation.

Each pixel is compared with its 3 x 3 neighborhood that is comprised of eight other pixels. In that process the center pixel value is subtracted by all the neighbors. The resulting negative values are labeled as 0, and all the others with 1. Afterwards, for each pixel, the binary values, starting from the one of its top-left neighbor, are concatenated in a clockwise direction, creating a new binary number. The obtained decimal value is then used for labeling the given pixel and is referred to as LBP codes. [3]

GLCM is a tabulation of how often different combinations of pixel luminance values (grey levels) occur in a specific pixel pairing of an image. Using a two-dimensional gray-level co-occurrence matrix is commonly and widely used in the field of texture analysis In order to find the locative dependence of brightness (gray-level) values, which helps to find valuable information about the neighboring pixels in an image. We use five of the classical statistical texture measures: entropy, energy, contrast, correlation and homogeneity, which are derived from a grey level cooccurrence matrix (GLCM). [3]

5.3. Classification

The classification tasks are accomplished using the VGG-19 from Keras. VGG-19 is a variation on the VGG (Visual Geometry Group) model, and it has 19 layers (16 convolutional, 3 fully connected, 5 maxpool, and 1 softmax).

To enhance performance we enlarge our dataset with a data augmentation technique. Data augmentation is a machine learning technique that involves creating new training data by applying various transformations. In the case of image augmentation, it applies geometric and color space transformations (flipping, resizing, cropping, brightness, contrast) to existing images. It helps improve model performance by increasing dataset diversity and making models more robust to real-world variations. The sample results of image augmentation are shown in Figure 6.

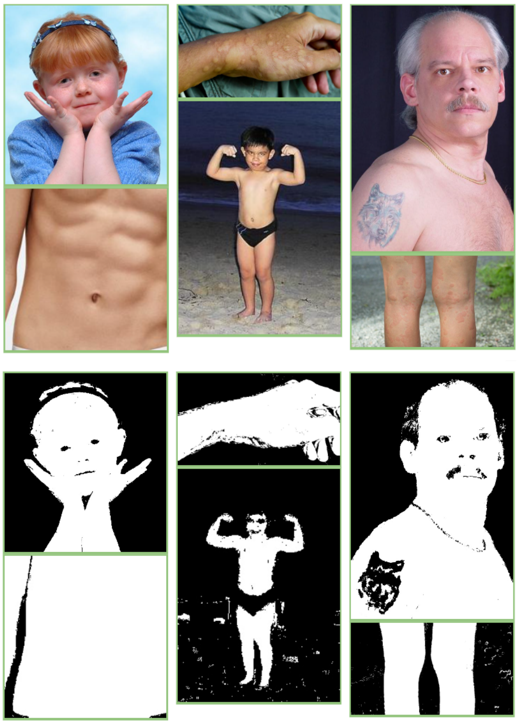

Figure 7: Skin detection results.

We build our training model by getting VGG19 as base model, then we add a Global Spatial Average Pooling layer and two dense layers with RELU activation function, and as an output layer we add another dense layer with Softmax activation function. For optimization we use Stochastic gradient descent with learning rate of 0.001. Categorical cross-entropy is used as a loss function for multi-class classification model, in which we classify three classes.

6. Results

The result on skin detection task we can see in Figure 7. We can say that the skin detection model performed fairly good with small amount of training data.

For inflammation detection and inflammation classification task we provided two confusion matrices in Figure 8 and Figure 9 that show us how the models performed on given test datasets. In Figure 8 we can see that the model predicted correctly inflamed and normal skins

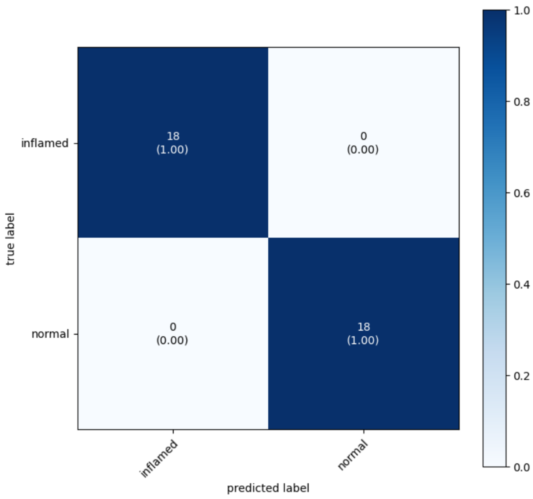

Figure 8: Confusion matrix of inflammation detection task.

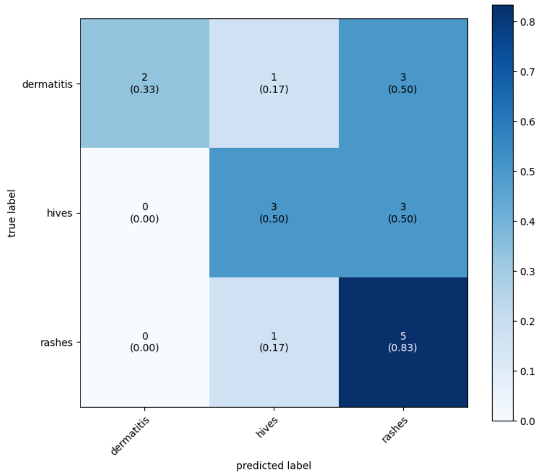

Figure 9: Confusion matrix of inflammation classification task.

with 100% accuracy. However, the inflammation classification model didn't perform as well as the inflammation detection model: it was only 33% correct in predicting dermatitis, 50% correct in predicting hives, and 83% correct in predicting rashes. Dermatitis and hives were predominantly mispredicted as rashes by up to 50%. In addition, dermatitis and rashes were mispredicted as hives by 17%. So, we can say that our inflammation classification model is good at recognizing and predicting the general type of inflammation, such as rashes, and not dermatitis and hives, which can vary quite within their classes. And our inflammation detection model performed excellently.

7. References

[1] Chapter 3. The Acute Inflammatory Response. In: Chandrasoma P, Taylor CR. eds. Concise Pathology, 3e. McGraw Hill; 1998.

[2] G. Maroni, M. Ermidoro, F. Previdi and G. Bigini, "Automated detection, extraction and counting of acne lesions for automatic evaluation and tracking of acne severity," 2017 IEEE Symposium Series on Computational Intelligence (SSCI), Honolulu, HI, USA, 2017, pp. 1-6, doi: 10.1109/SSCI.2017.8280925.

[3] E. Almansour and M. A. Jaffar, "Classification of dermoscopic skin cancer images using color and hybrid texture features," IJCSNS International Journal of Computer Science and Network Security, vol. 16, no. 4, pp. 135–139, 2016

8. Online Resources

The sources and materials are available via

- GLCM - GitHub,
- Color, Shape and Texture: Feature Extraction using OpenCV,
- Local Binary Patterns with Python & OpenCV.