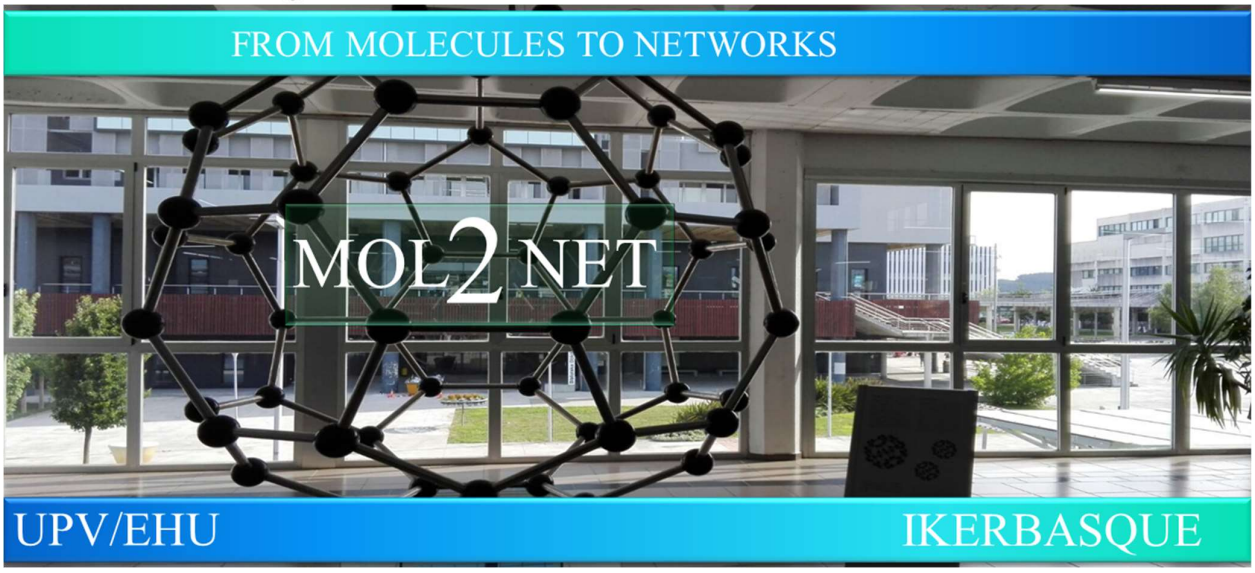
MDPI sciforum

MOL2NET'22, Conference on Molecular, Biomedical & Computational Sciences and Engineering, 8th ed.




# Computer Vision Based Skin Cancer Classification by Using Texture Features

*Samreen Naeem* [*, a], *Aqib Ali* [a] *and Sania Anam* [b]

[a] *College of Automation, Southeast University, Nanjing, China.*
[b] *Department of Computer Sci, Govt Associate College for Women Ahmadpur East, Bahawalpur, Pakistan.*

. * *Corresponding author: samreencsit@gmail.com*

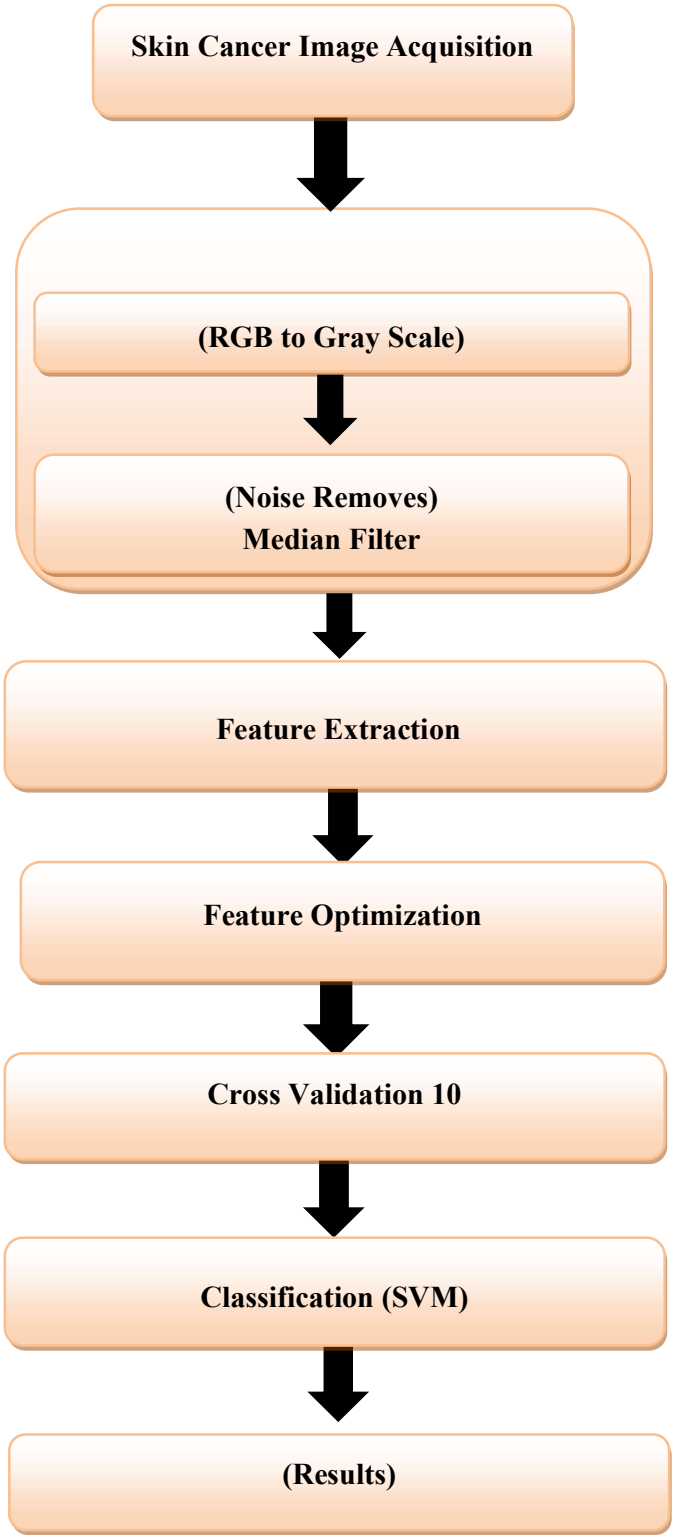
**Abstract.**

Cancer of the skin is now one of the most prevalent forms of the disease among people. As a result, accurate diagnosis of malignant lesions is of utmost significance in treating skin cancer. Dermatoscopic images can be used with computer-aided diagnostic tools, which may include machine learning models, to assist medical professionals in diagnosing skin cancer. In this particular research project, skin lesions were classified using image processing and machine learning strategies. Several distinct mathematical techniques have been implemented in the field of image processing in order to improve image quality. Image segmentation utilizing the watershed approach was conducted after an image preparation step, which included filtering the undesired pixels in the pictures. Following that, the lesioned regions were separated, and texture feature extraction was carried out. In the end, the classification was completed using the SVM algorithm, which stands for support vector machines. When the results acquired from the classifiers were compared, it was seen

that the SVM classifier had an accuracy of 94.33%.

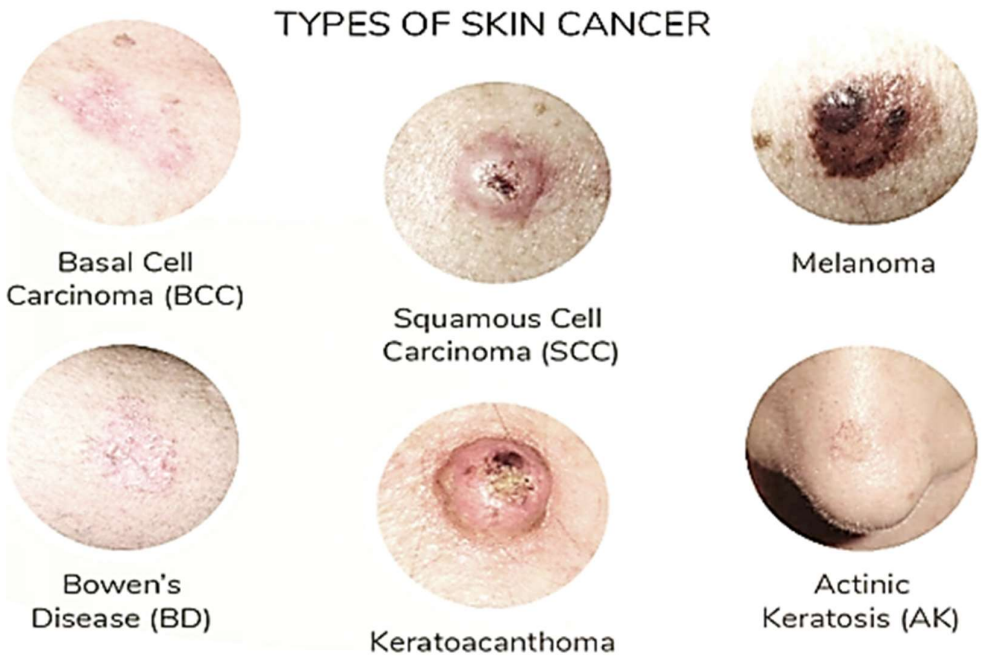
**Keyword:** Dermatoscopic Images, Skin Cancer, SVM, Machine Learning.

**Graphical Abstract**



## Introduction

The growth and division of healthy cells result in the formation of new cells. These new cells will eventually take the place of injured or elderly cells. *Cancer* is the name of the disease that occurs when the body's normal cells multiply uncontrollably. Cancer cells, rather than killing off, continue to proliferate and give rise to more aberrant cells. A tumor can develop when there is aberrant cell proliferation in the epidermal layer, which may destroy other healthy tissues. This process, known as metastasis, describes how cancer cells move from one location in the body to another through the circulatory and lymphatic systems. Cancer comes in a wide variety of forms [1].




The most frequent types of cancer include thyroid cancer, lung cancer, breast cancer, bladder cancer, kidney cancer, colon and rectum cancer, leukemia, pancreatic cancer, prostate cancer, and skin cancer. Other less common types of cancer include leukemia and pancreatic cancer. The surface of our skin is the most significant organ in our bodies [2]. This organ's primary responsibilities include acting as a protective barrier against potentially hazardous substances that may enter the body from the environment, regulating body temperature via the hair, sweat glands, and adipose tissue that it contains, ensuring proper fluid and electrolyte balance, and contributing to the overall health of our bodies. The layers of our skin may be broken down into three broad categories. The epidermis, the topmost layer of our skin, is the first layer we will discuss [3].

In most cases, this layer stops the loss of fluid and serves as protection for the tissues that lie underneath it. The dermis is the second layer of the skin. In addition to hair follicles and sebaceous glands, the dermis is home to blood vessels and nerve fibers. It is in charge of regulating the

temperature and maintaining the fluid-electrolyte balance [4]. The hypodermis is the third and last layer of our skin. This layer is composed of adipose tissue. It not only guards against the damaging effects of impact on structures like bone and muscle but also helps maintain a steady body temperature. These three layers of our skin are fertile ground for the growth of tumors. The specific sort of cancer that an individual has depends on which tissues and layers of their skin the altered cells originate in. As a result, the procedure of treating the kind of cancer that was detected in the patient may be challenging or straightforward [5].

Skin cancer is the most prevalent form of the disease in every region of the globe surveyed by the Skin Cancer Foundation (SCF). According to the statistics from the Skin Cancer Foundation (SCF), more than two persons lose their lives to skin cancer every hour in the United States [6].

## Materials and Methods

.

The ISIC 2020 training set, HAM10000, identified seven skin disorders in this investigation [7]. The dataset contains 300 150x2 RGB skin lesion photos. The significant number of lesions per category separates this dataset from others. These photos were classed as abnormal and usual. The dataset contains 300 malignant and 300 benign skin lesions. Preprocessing Ham10000 photos improved picture quality. Images received contrast-limited adaptive histogram equalization, morphological occlusion, and median filter. Image segmentation follows preprocessing. Image segmentation enhances meaning and analysis. Segmentation typically determines analysis success. Medical picture segmentation requires ROI extraction. Area-based segmentation uses the watershed transform to define the region of interest (ROI) and choose the ROI closest to the skin lesion.

Feature extraction is commonly used to classify segmented pictures in the training and test sets. This research extracted features using texture analysis [8]. Asymmetry is a crucial indicator of skin lesion malignancy. It measures how comparable the lesion form is along the primary axis. Vertical and horizontal asymmetry are computed independently. The compact index (CI) measures border irregularity as the ratio of the lesion circumference square to its area. Color is crucial to skin disease diagnosis. The normalized standard deviation of lesion red, green, and blue components measures color variance. The texture rule defines *diameter* as the most significant distance between any two sites of the lesion boundary. The lesion diameter is the diameter of the circle with the lesion area. Image texture determines pattern and color constancy. Haralick texture characterizes texture-based images by computing the gray-level co-occurrence matrix (GLCM). GLCM is used to extract texture features since it can calculate many characteristics and is easy to use. The co-occurrence matrix extracts 14 characteristic texture properties from the probability matrix. This research selected

Haralick's contrast, correlation, energy, and homogeneity [9]. GLCM is based on picture-pixel neighborhoods. It records the complete picture by searching for adjacent pairs of pixel values.

Correlation measures the combined probability of pixel pairings in each row and column. Energy, the square root of the total of square pixels, is the recurrence of pairs of pixels in the picture. Contrast distinguishes items by hue or color. Homogeneity; measures how near the GLCM component distribution is to its diagonal [10].

**Results and Discussion**

- Support vector machine (SVM)
- Time taken to build the model: 0.38 seconds
- Test mode: 10 fold

**Table 1:** SVM Classifier Summary

| | | |
|---|---|---|
| **Total Number of Instances** | 300 | |
| **Correctly Classified Instances** | 283 | **94.3333 %** |
| **Incorrectly Classified Instances** | 17 | **5.6667 %** |
| **Kappa statistic** | 0.8867 | |
| **Mean absolute error** | 0.0777 | |
| **Root mean squared error** | 0.2294 | |
| **Relative absolute error** | 15.5438 % | |
| **Root relative squared error** | 45.8745 % | |

**Table 2:** SVM Classifier Detailed Accuracy

| TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | Class |
|---|---|---|---|---|---|---|---|
| 0.913 | 0.027 | 0.972 | 0.913 | 0.942 | 0.888 | 0.953 | **Normal** |
| 0.973 | 0.087 | 0.918 | 0.973 | 0.945 | 0.888 | 0.953 | **Abnormal** |
| **0.943** | **0.057** | **0.945** | **0.943** | **0.943** | **0.888** | **0.953** | **Weighted Avg.** |

**Table 3:** Confusion Matrix result using SVM Classifier

| Classified as | A | B |
|---|---|---|
| A = Normal | 137 | 13 |
| B = Abnormal | 4 | 146 |

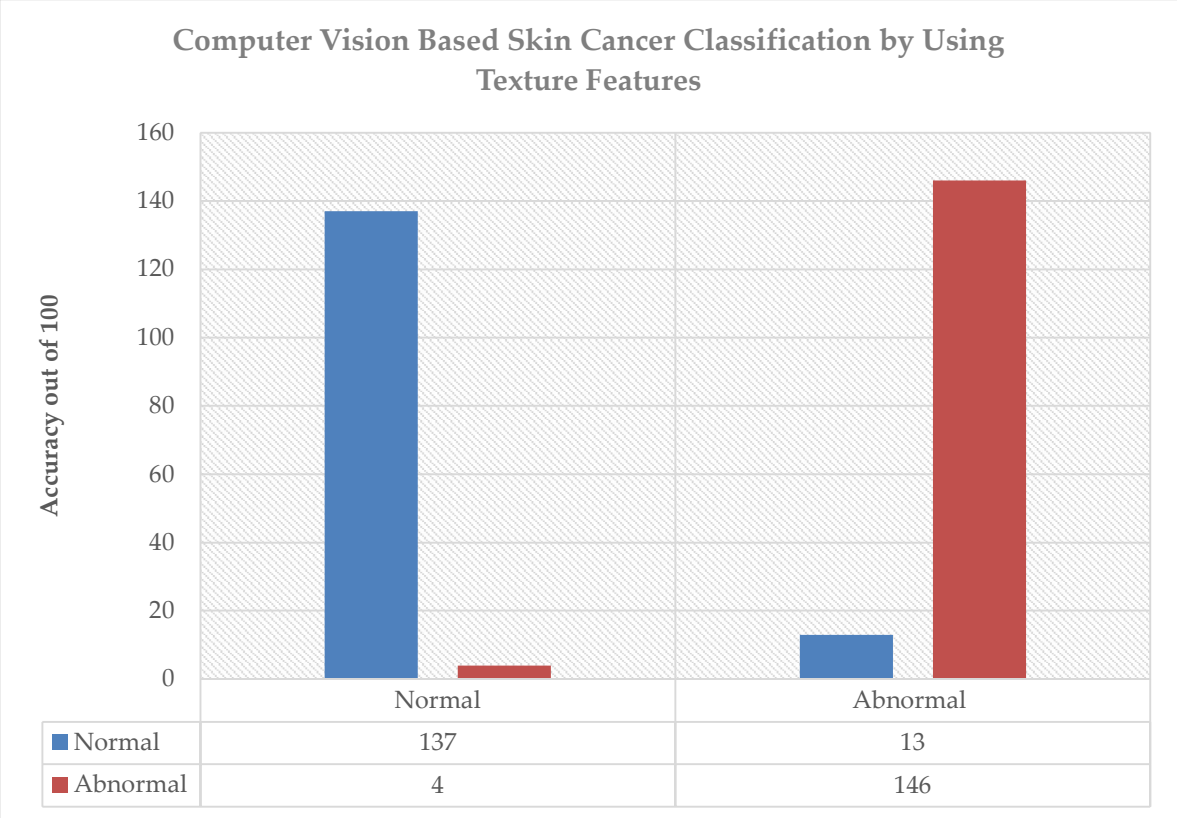



**Figure 1: Accuracy of Dataset using SVM Classifier**

## Conclusions

This work uses machine learning to classify skin lesions in HAM10000 dermatoscopic pictures. SVM were used to classify the dataset's pictures after texture analysis segmented and extracted lesion characteristics. The research found that the DVM approach had greater accuracy, precision, and f1-score. Combining an optimized classifier, a current segmentation technique, and a well-tested feature extraction strategy shows SVM classifier performance. This study's findings may be improved by increasing training samples and balancing labeled lesion groups. Using sophisticated classification models or deep learning techniques and training the models on the complete data set may also improve outcomes.

## References

[1]. Leiter, U., Keim, U., & Garbe, C. (2020). Epidemiology of skin cancer: update 2019. Sunlight, Vitamin D and Skin Cancer, 123-139.

[2]. Neagu, M., Constantin, C., Cretoiu, S. M., & Zurac, S. (2020). miRNAs in the Diagnosis and Prognosis of Skin Cancer. Frontiers in cell and developmental biology, 8, 71.

[3]. Cullen, J. K., Simmons, J. L., Parsons, P. G., & Boyle, G. M. (2020). Topical treatments for skin cancer. Advanced drug delivery reviews, 153, 54-64.

[4]. Tschandl, P., Rinner, C., Apalla, Z., Argenziano, G., Codella, N., Halpern, A., ... & Kittler, H. (2020). Human–computer collaboration for skin cancer recognition. Nature Medicine, 26(8), 1229-1234.

[5]. Zhang, N., Cai, Y. X., Wang, Y. Y., Tian, Y. T., Wang, X. L., & Badami, B. (2020). Skin cancer diagnosis based on optimized convolutional neural network. Artificial intelligence in medicine, 102, 101756.

[6]. Ilkhomovna, K. D. (2021). Modern Look of Facial Skin Cancer. BARQARORLIK VA YETAKCHI TADQIQOTLAR ONLAYN ILMIY JURNALI, 1(1), 85-89.

[7]. Razmjooy, N., Ashourian, M., Karimifard, M., Estrela, V. V., Loschi, H. J., Do Nascimento, D., ... & Vishnevski, M. (2020). Computer-aided diagnosis of skin cancer: a review. Current Medical Imaging, 16(7), 781-793.

[8]. Ali, A., Mashwani, W. K., Naeem, S., Uddin, M. I., Kumam, W., Kumam, P., ... & Chesneau, C. (2021). COVID-19 infected lung computed tomography segmentation and supervised classification approach.

[9]. Naeem, S., Ali, A., Qadri, S., Khan Mashwani, W., Tairan, N., Shah, H., ... & Anam, S. (2020). Machine-learning based hybrid-feature analysis for liver cancer classification using fused (MR and CT) images. Applied Sciences, 10(9), 3134.

[10]. Ali, A., Qadri, S., Khan Mashwani, W., Kumam, W., Kumam, P., Naeem, S., ... & Sulaiman, M. (2020). Machine learning based automated segmentation and hybrid feature analysis for diabetic retinopathy classification using fundus image. Entropy, 22(5), 567.