BILATERAL FILTER BASED DERMIC TUMOR CLASSIFICATION USING SVM AND CNN

KARTHIK B U[1], MUTHUPANDI G[2]

[1]Research scholar, Department of ECE, Presidency University, Bengaluru, India
[2]Associate Professor, Department of ECE, Presidency University, Bengaluru, India
E-mail: [1]karthikbu1994@gmail.com, [2]muthupandi@presidencyuniversity.in

ABSTRACT

A non-iterative edge-preserving image enhancement method is proposed for classifying human dermic tumors. Dermic tumor images typically feature more background residuals, which provide corrupted data for classification. In this instance, we separate the image data into two layers: Base layer and detailed layer. The artefacts of detailed layers were removed by applying bilinear interpolation. Additional amplification is used to produce high-pitched, detailed layer edges. The images are augmented with the enhancement ratio 1.5 and 2. To sharpen the image, additional Gaussian filtering is performed on the components of the detailed layer. The modified detailed layer's and the base layer's components are merged to get the enhanced image. For image denoise and edge preservation, the enhanced image is fed to the bilateral filter. Quantitative and qualitative techniques are employed to evaluate the image enhancement. For classification we use Support Vector Machine and Convolution Neural Networks. Accuracy of 97.29 and 97.75 is obtained for the classifiers SVM and CNN respectively.

Key words:*Gaussian filtering, Bilateral filter,Image decomposition, Image Enhancement, Image Classification.*

1.INTRODUCTION

Cancer is caused by uncontrolled cell growth that can invade or spread to other regions of the body. Among the most destructive and fatal kinds of cancer is skin cancer. The skin, biggest organ in the body, is vulnerable to harm from a range of factors, such as ultraviolet (UV) radiation from the sun, tanning, lifestyle choices, smoking, drinking, physical activity, infections, and the workplace environment [1,2]. These elements compromise its integrity and have a significant, disastrous influence on its health. Skin problems, which affect 1.5 million individuals, are the fourth most common cause of all illnesses in people [3]. Skin problems can have adverse effects such as harm to internal organs, relationship problems, limitations on everyday activities (like melanoma), and even death. They also pose a significant risk of psychological distress, which can result in isolation, grief, and even suicide [4]. Skin problems must be detected as soon as possible to save costs, lives, and morbidity. Skin cancer is treatable if detected early. Skin cancer research is an exciting field of study in medical imaging and machine learning [5]. Malignant and benign tumors are two major classes of skin cancer, with Malignant being the most dangerous and deadly type of cancer [6]. Benign tumor refers to any mole growth that develops on the skin, which is not cancerous. On the other hand, a malignant tumor is caused by a cancerous mole growth on the skin. The benign proliferates slowly and can be cured if detected early. Melanoma, on the other hand, is an extremely aggressive and fatal form of cancer that has no noticeable symptoms and is responsible for approximately 75% of all skin cancer deaths. Thus, if diagnosis is delayed, it spreads deeper, grows faster which is difficult to treat, and is not guaranteed to survive [7]. The medical community has spent an enormous amount of money on campaigns to increase people's knowledge of skin cancer, yet this does not assure their safety. So, it is vital to invest in the development of technologies that may be employed in early diagnosis procedures [8]. Recently, several experts in the field of computer vision (CV) have developed a computational model for the automatic classification and diagnosis of skin tumors [9-11].

2. LITERATURE REVIEW

Several research has been conducted over the years for the early diagnosis of ski tumor. **Hiam Alquran** et al. [12] suggested an OTSU threshold-based skin cancer diagnostic technique. Using the grey level co-occurrence matrix, features such as ABCD (atypical, border, color, and diameter), the total dermoscopic score (TDS), circulation, and textural characteristics after image pre-processing and segmentation (GLCM) are extracted. They achieved a classification accuracy of 92.1% by combining the RBF kernel and SVM classifier.

UzmaBano et al. [13] created a method for skin cancer detection. The result was pre-processed and then segmented for lesions using a maximum entropy threshold. By using GLCM, they retrieved several texture characteristics. With the help of the SVM classifier, they were able to attain a 95% classification accuracy.

Dalila et al. [14] devised a method for classifying benign and malignant skin lesions using an ant colony-based segmentation algorithm. Shape, texture, and color features are used. They employed KNN and ANN classifiers for classification, and their respective classification accuracy rates were 85.22% and 93.60%.

Hasan et al. [15] presented a CNN-based technique for identifying skin cancer. In the testing phase, they had a detection accuracy of 89.5%. The detection accuracy need to be improved, though, as it was insufficient. The research had a problem in that there was overfitting between the testing and training stages. It is difficult to develop an automated classification of skin cancer because of the variability and diversity of skin disease images.

Khan et al. [16] introduced a deep learning-based methodology in which skin lesions are preprocessed using the decorrelation formulation method and then segmented using MASK-RCNN. Dense Net is used to extract features from the segmented image. Finally, the best features are determined by using the SVM approach based on least squares with entropy control. The ISBI2016, ISBI2017, and HAM10000 datasets were used to conduct the experiment. The accuracy obtained on these datasets is 96.3%, 94.8%, and 88.5%, respectively.

Mahbod et al. [17] investigated the impact of picture scaling on a CNN network that has beentrained to classify skin lesions. The scientistsscaled the photos at six different sizes and evaluated the classification performance of three CNN architectures: EfficientNetB0, EfficientNetB1, and SeReNeXt-50. They designed and tested a multiscale multi-CNN fusion (MSM CNN) ensemble-based technique. This technique trains on a variety of cropped picture sizes and employs three distinct CNN models. Using the ISIC2018 dataset, the MSM CNN technique produced an accuracy of 86.2%. They concluded that cropping rather than scaling images yielded superior outcomes.

Dataset used: The International Skin Imaging Collaboration (ISIC) [18] is a partnership between academics and industry that intends to make it simpler to utilize Skin images to significantly reduce the death rate. The skin tumor images are categorized into benign and malignant tumor. The dataset is further subdivided into several age categories and genders within each kind.

3. PROPOSED WORK

A novel approach for the classification of dermal tumour images is proposed. The images are divided into base and detailed layers during the pre-processing step using the convolution and deconvolution methods. The bilinear edge interpolation approach is used on both the base and detailed layers. After being amplified, the detailed layer is passed through a Gaussian filter. The base layer is combined with the updated detailed layer and sent through the bilateral filter. Following enhancement, the classifier is fed features extracted from the enhanced images and dataset images. The classifier separates the images into benign and malignant tumor images.

3.1 IMAGE ENHANCEMENT

The main motive behind image enhancement is to increase the image's contrast and sharpness so that it may be processed or further examined. Image Enhancement's [19-20] purpose is to make digital images more appropriate for display or further image analysis by modifying them. The elimination of noise, the sharpening, or the brightening of an image might make it easier to categorize substantial elements. The proposed model is designed to perform better image enhancement and to preserve the edges for feature extraction and image classification.

3.1.1 IMAGE DECOMPOSITION

Although image decomposition [21] has gained significant attention in recent decades, it has remained a difficult problem to properly deconstruct the image efficiently. Image decomposition is a fundamental image processing technique used for a variety of tasks, including image smoothing, tone mapping, image abstraction, detail enhancement, and high-dynamic-range compression. Image decomposition's primary objective is to successfully isolate structure from a given image by keeping edge-like structural components and omitting fine-scale details. As a result, we intend to use the convolution and deconvolution methods to split the image into two separate layers: detailed layer and base layer. An image's base layer is made up of smoothly changing areas that transmit enormous structural data underlying the image. A detailed layer encapsulates limited scope disparities and encloses particulars of the image's appearance. The image decomposition method is explained below.

An image I can be represented as

$$I\ (i, j) = B\ (i, j) + D\ (i, j) \qquad (1)$$

where B is a Base layer and D is detailed layer, i and j are pixel coordinates.

The base layer comprises the smoothly changing regions and sharply defined areas. The base layer depicts the underlying data of a picture, like locales with different textures and shading. The detailed layer D holds minimal oscillating outlines and local intensity dissimilarities. The detailed layer D specifies the deviancies from smoothly fluctuating intensities in local texture regions. The image I in (1) is convolved with a Gaussian filter f, yields a blurred image x as follows:

$$x = I * f \qquad (2)$$

Where * is the convolution operator.

For restoring the blurred structural information, the blurred image is subjected to the deconvolution technique which yields the base layer. Deducting the base layer from the image results in the detailed layer in (3).

$$D = I - B \qquad (3)$$

3.1.2 EDGE PRESERVING INTERPOLATION

After image decomposition, the edge-preserving interpolation [22] is applied to both the layers for smoothing the data and not introducing spurious artefacts and preserving the details, i.e., edges. The Bilinear Interpolation technique is based on four adjacent pixels. If we want to determine the value of the unknown function I at the point X = (i, j), it is anticipated that we know the value of I at the four points Y11 = (i1, j1), Y12 = (i1, j2), Y21 = (i2, j1), and Y22 = (i2, j2). Fig. 2 shows the pictorial representation of these points and the interpolated point X (i, j).

We first do linear interpolation in the x-direction. This yield $f(Z_1)$ and $f(Z_2)$ as represented in (4) and (5)

$$f(Z_1) \approx \frac{i_2-i}{i_2-i1} f(Y_{11}) + \frac{i-i_1}{i_2-i_1} f(Y_{21}) \ (4)$$

$$f(Z_2) \approx \frac{i_2-i}{i_2-i1} f(Y_{12}) + \frac{i-i_1}{i_2-i_1} f(Y_{22}) (5)$$

where Z1= (i, j1), Z2= (i, j2)

We continue by interpolating in the y-direction. This gives us the desired estimate of I (i, j) as in (6).

$$f(X) \approx \frac{j_2-j}{j_2-j_1} f(Z_1) + \frac{j-j_1}{j_{2-j1}} f(Z_2) (6)$$

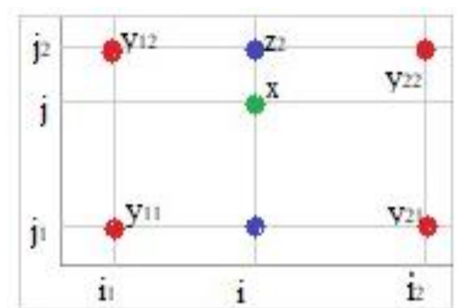


Figure 1: *Bilinear interpolation*

3.1.3 **BILATERAL FILTER**

After applying the edge-preserving interpolation to both the layers, the detailed layer is enhanced with enhancement ratio (ER) of 1.5 and 2 and then applied to the Gaussian filter to achieve sharper edges. The base layer and the augmented detailed layer are combined. During the process of amplification and sharpening, edge information might be lost and noise may be introduced, so the bilateral filter is applied to the merged image to denoise it while maintaining the edges.

Digital images usually suffer from noise due to the presence of unwanted signals such as specular reflections, camera lens flare, halogen lamp flicker, shadows, etc., before getting passed through a light detector. Such kinds of noisy images appear more often while capturing images outside indoor environments. There may be problems faced during the acquisition of those images. In these situations, advanced pre-processing techniques cannot provide high-quality imaging after filtering these noises separately. Image denoising filterssuch as bilateral normalisation, median blur (Median Blur), Gaussian blurring, MIP (Maximum Intensity Projection), mean/variance-based filters, and edge-preserving methods have played a crucial role in enhancing the image quality in this regard. The bilateral filter [23] is a non-iterative and nonlinear filter proposed by Tomasi and Manduchi for retaining the edges while lowering noise. It considers the geometric closeness of neighbouring pixels as well as their grey-level likenesses. BF calculates a local neighbourhood's weighted sum of pixels. Each pixel is replaced with the weighted average of pixels from its neighbours. The bilateral filter is described as a weighted average of adjacent pixels. The bilateral filter considers the value difference between neighbours to retain edges when smoothing. The bilateral filter's core idea is that two pixels cannot just influence one another by being in the same location; they also need to have the same values to do so. Weights can be computed based on the intensity and spatial distance of a pixel. The spatial distance in a pixel's neighbourhood is determined by the domain (spatial) filter, but the range (intensity) filter weights are determined by the pixel's radiometric distance. The bilateral filter is applied to the merged image.

The bilateral filter, represented by the symbol BF [I], is described by

$$BF[I]_p = \frac{1}{W_i}\sum_{j\in s} G_{\sigma_s}(\|i-j\|)G_{\sigma_r}(|Ii - Ij|)Ij \qquad (6)$$

Where W_i is the normalization factor confirms pixel weights sum to 1.

$$Wi = \sum_{j\in s} G_{\sigma_s}(\|i-j\|)G_{\sigma_r}(|Ii - Ij|) \qquad (7)$$

The level of filtering of the image I depends on thevalues of parameters σ_s and σ_r.
Where G_{σ_s} is spatial Gaussian that reduces the impact of distant pixels,G_{σ_r} is range Gaussian thatlessens the influence of pixels q when their intensity values diverge from Ii. First weights W_i of each pixel are calculated and normalised using (7) and using these weights and spatial Gaussian and range Gaussian, the bilateral filter denoises the image as described in (6). The spatial Gaussian and Range Gaussian values are tuned based on the images.

The skin lesions are pre-processed using bilateral filter with ER of 1.5 and 2. Figure 2 depicts each individual histogram and detailed layer enhancement. The bright areas of the image get brighter and the darker areas become less dark as the base layer is augmented. The entropy of the improved images is lower, indicating that original information may be lost. In this instance, we choose to improve via a detailed layer. The study is based on the ISIC dataset, which contains images of benign and malignant tumours.

Mean Square Error: The Mean Square Error is the average pixel-by-pixel difference between the original and modified images. It gives a measurement of the cover image error brought on by the data embedding process. The image quality improves as the MSE value decreases.

Peak Signal-to-Noise Ratio: PSNR is the ratio of a signal's greatest possible value to the strength of distortion noise. It is expressed in Db's. The better the image quality, the higher the PSNR value.

Structural Similarity Index: SSIM is an evaluation indicator used to see how similar the two images are. The perceived difference between the two images is measured. The SSIM Quality Index measure is based on the calculation of three elements: luminance, contrast, and structure.

Entropy: Entropy is the measure of the image's information content. The image quality increases if the Entropy value is high.

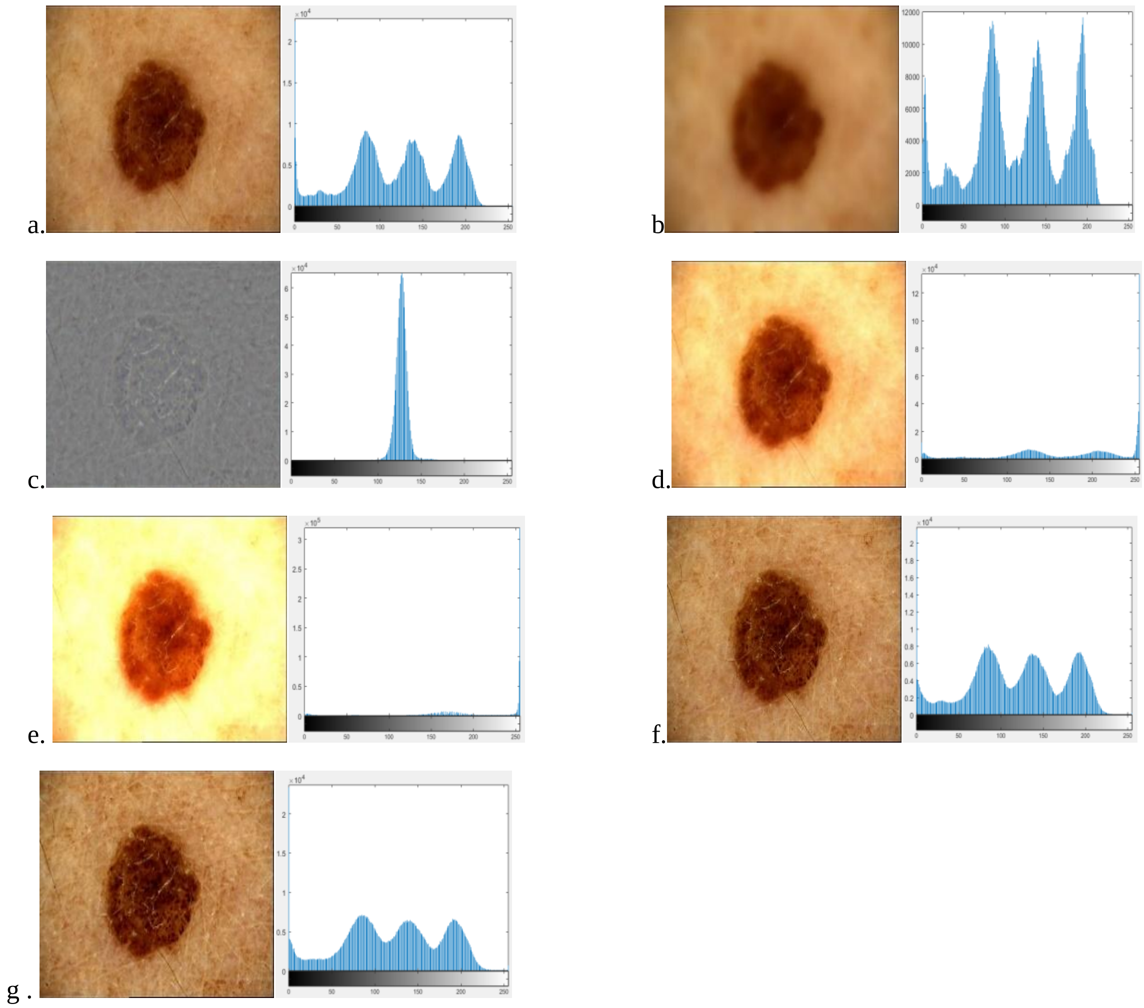


Figure 2: a. input image b. Base layer c. Detailed layer d. Base layer enhancement with ER 1.5 e. Base layer enhancement with ER 2 f. Detailed layer enhancement with ER 1.5 g. Detailed layer enhancement with enhancement ratio 2

			BENIGN				MALIGNANT			
Image ID	Age Group	Gender	MSE	PSNR	SSIM	Entropy	MSE	PSNR	SSIM	Entropy
ISIC3	10-20	Male	22.447	34.619	0.961	7.5411	5.512	40.717	0.990	5.8789
ISIC2	20-30	Female	30.731	33.255	0.972	7.2931	36.11	32.553	0.987	7.5300
ISIC4	30-40	Male	21.740	34.758	0.995	7.5695	20.00	35.118	0.984	7.3300
ISIC2	40-50	Female	14.108	36.636	0.993	6.9418	11.26	37.615	0.992	7.0264
ISIC1	50-60	Male	29.071	33.496	0.980	7.4158	11.03	37.703	0.994	7.0976
ISIC5	60-70	Female	18.888	35.368	0.994	7.5914	10.87	37.766	0.993	7.3085

Table 1:Comparison table for enhancement ratio 1.5

			BENIGN				MALIGNANT			
Image ID	Age Group	Gender	MSE	PSNR	SSIM	Entropy	MSE	PSNR	SSIM	Entropy
ISIC3	10-20	Male	82.116	28.986	0.899	7.6217	18.564	35.444	0.980	5.9791
ISIC2	20-30	Female	99.483	28.153	0.940	7.3873	129.47	27.009	0.959	7.5762
ISIC4	30-40	Male	80.175	29.090	0.989	7.2928	72.900	29.503	0.953	7.4291
ISIC2	40-50	Female	37.050	32.442	0.985	7.049	25.082	34.137	0.986	7.1225
ISIC1	50-60	Male	90.375	28.570	0.954	7.4752	27.625	33.717	0.989	7.2079
ISIC5	60-70	Female	45.598	31.541	0.989	7.6182	25.639	34.041	0.986	7.3903

Table 1:Comparison table for enhancement ratio 2

3.2 IMAGE SEGMENTATION

Image segmentation [24] is the method of breaking down the picture into many segments or areas in order to facilitate the analysis and understanding of its contents. This technique is widely used in computer vision, medical imaging, and other fields where it is necessary to retrieve meaningful data from images. Many methods, such as thresholding, region-based segmentation, and edge-based segmentation, can be used to segment images. Thresholding involves setting a threshold value that separates pixels into two classes based on their intensity values. Region-based segmentation groups pixels according to their colour or texture, whereas edge-based segmentation marks the borders between several areas in an image.

K-means clustering is a commonly used unsupervised machine learning algorithm that is used for partitioning data into groups or clusters based on their similarity. In this research we used K-means clustering in image segmentation.

Images are segmented using K-means clustering, which is described in the following steps

Step 1: It involves dividing the pixels in an image into K clusters, where K is a pre-defined number of clusters.

Step 2: Each cluster represents a segment or region of the image. The procedure begins by randomly picking K locations as cluster centroids.

Step 3: According to their colour, intensity, or texture characteristics, the image's pixels are then assigned to the closest centroid.

Step 4: The mean of the pixel values in each cluster is then determined, and the centroids are updated.

Step 5: This method is performed until the centroids reach a constant or until a predetermined number of iterations is reached.

K-means clustering in image segmentation has several advantages. It is computationally efficient and can be applied to large datasets. K-means clustering is also effective in segmenting images with clear boundaries between the regions, and it can be used for both colour and grayscale images. K-means clustering can also be combined with other segmentation techniques to improve the segmentation results.

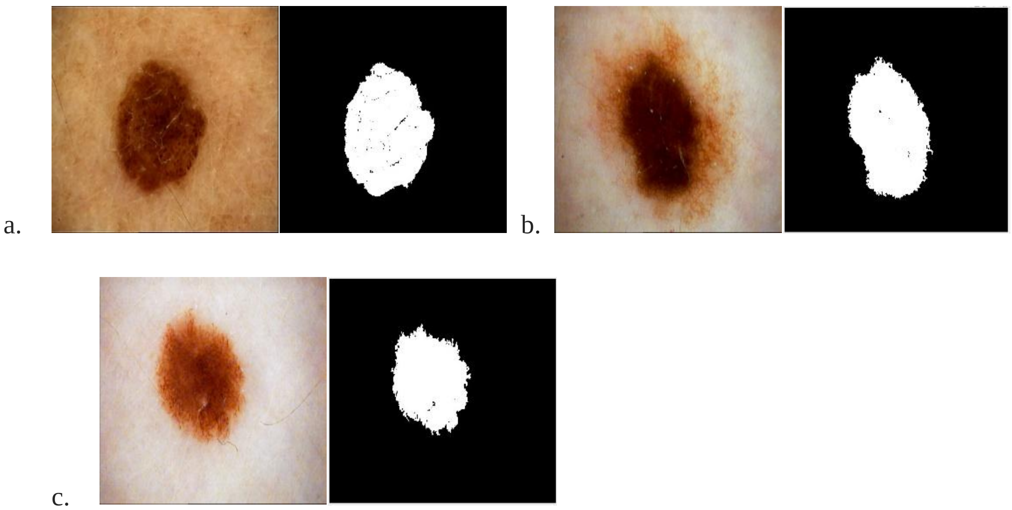


Figure 3: Image segmentation results of sample images

3.3 IMAGE CLASSIFICATION

A technique called image classification is used for the classification of images based on the visual information. We live in a data-driven age. As the Internet of Things (IoT) and artificial intelligence (AI) become more widespread, massive volumes of data are being generated. Image classification [25] applications include machine vision, traffic control systems, brake light detection, medical imaging, object recognition in satellite images, and many more. Training necessitates the use of algorithms and classifiers that ingest enormous volumes of data, analyse them, and extract relevant characteristics. The classification approach is meant to classify each feature in an image into separate classes. Identifying and classifying an image's attributes in terms of the subject actually reflects those aspects on the ground as a different grey level (or colour). The classification of medical images is a two-step procedure. The first step is to extract visual properties from image data using feature extraction techniques, and then use machine intelligence algorithms to group or categorise images based on these features. There are numerous effective classification methods, including artificial neural networks, decision trees, Bayesian networks, Adaboost, and SVM. Although K-nearest-neighbour approaches have the benefit of being simple to use, they are typically highly sluggish if the supplied data collection is sufficiently huge. However, these are extremely susceptible to the existence of irrelevant parameters. In classification problems, decision trees have also been extensively employed. During the training phase, they are often quicker than neural networks, but they lack versatility in parameter modelling. As a universal approach, neural networks are extensively used in a wide range of applications. However, many factors must be addressed while constructing a neural network to handle a given issue, such as the learning algorithm, architecture, number of neurons per layer, number of hidden layers, data representation, and much more.

3.3.1 FEATURE EXTRACTION

Feature extraction is a method that preserves the content of the original data set while converting raw information into numerical characteristics [26]. It is possible to manually extract features or do it automatically. Engineers and scientists have been working on feature extraction algorithms for text, signals, and pictures for decades. Automated feature extraction employs specialized algorithms to extract features from signals or images without the need for human interaction. This approach is quite helpful when you need to quickly switch from gathering raw data to developing algorithms. After enhancement, features are extracted to acquire the data contained in an image. A variety of techniques are employed to extract the characteristics depending on the image type.

Feature extraction plays a crucial role in recognising and understanding the content of images. One such method of feature extraction is the Grey-Level Co-occurrence Matrix (GLCM) method, which involves analysing the spatial distribution of grey-level intensities in an image. This report provides an overview of the GLCM method and its application in feature extraction.

The Grey-Level Co-occurrence Matrix (GLCM): The Grey-Level Co-occurrence Matrix (GLCM) is a statistical method used to describe the spatial relationship between pixels in an image. The GLCM is created by calculating the frequency of occurrence of pairs of pixels at different spatial relationships and with different grey-level intensity values. The resultant matrix describes the distribution of grey-level values in an image. The GLCM can be calculated by analyzing the relationship between pixels at different distances and angles.

The distance between the pixels is called the offset, while the angle between them is called the orientation. The resulting GLCM is a square matrix where the rows and columns represent the grey-level intensity values and the values in the matrix represent the frequency of occurrence of pairs of pixels with specific grey-level values at a given offset and orientation.

Feature Extraction Using GLCM: GLCM can be used for texture analysis, which involves extracting features that represent the textural properties of an image. The GLCM method is capable of detecting patterns and textures that are difficult to discern with the human eye.

The GLCM-based texture features can be extracted by analysing the statistical properties of the GLCM. Some of the commonly used features include:

1. Contrast: It calculates the variation in the grey-level intensity values of adjacent pixels. High contrast indicates a sharp transition between the grey-level values, while low contrast indicates a smooth transition.
2. Energy: It evaluates how uniformly the intensity values at different grey levels are distributed throughout a picture.
3. Homogeneity: It measures the similarity of the neighbouring pixels in terms of their grey-level intensity values.
4. Entropy: It determines how randomly or uncertainly the intensity values at different grey levels are distributed within a picture.
5. Correlation: It calculates the proportion between the intensity values of adjacent pixels' grey levels.
6. Skewness: Skewness describes how much the pixel intensity values in an image are skewed to the left or right of the mean. A distribution that is positively skewed implies that the majority of the pixel intensity values are lower than the mean, whereas a distribution that is negatively skewed suggests that the majority of the pixel intensity values are higher than the mean.
7. Kurtosis: Kurtosis describes how closely the pixel intensity values in a picture are clustered around the mean. A flatter distribution is indicated by a low kurtosis value, whereas a sharp peak is indicated by a high kurtosis value.

The extracted attributes are applied to the SVM classifier, which classifies the images into the appropriate classes.

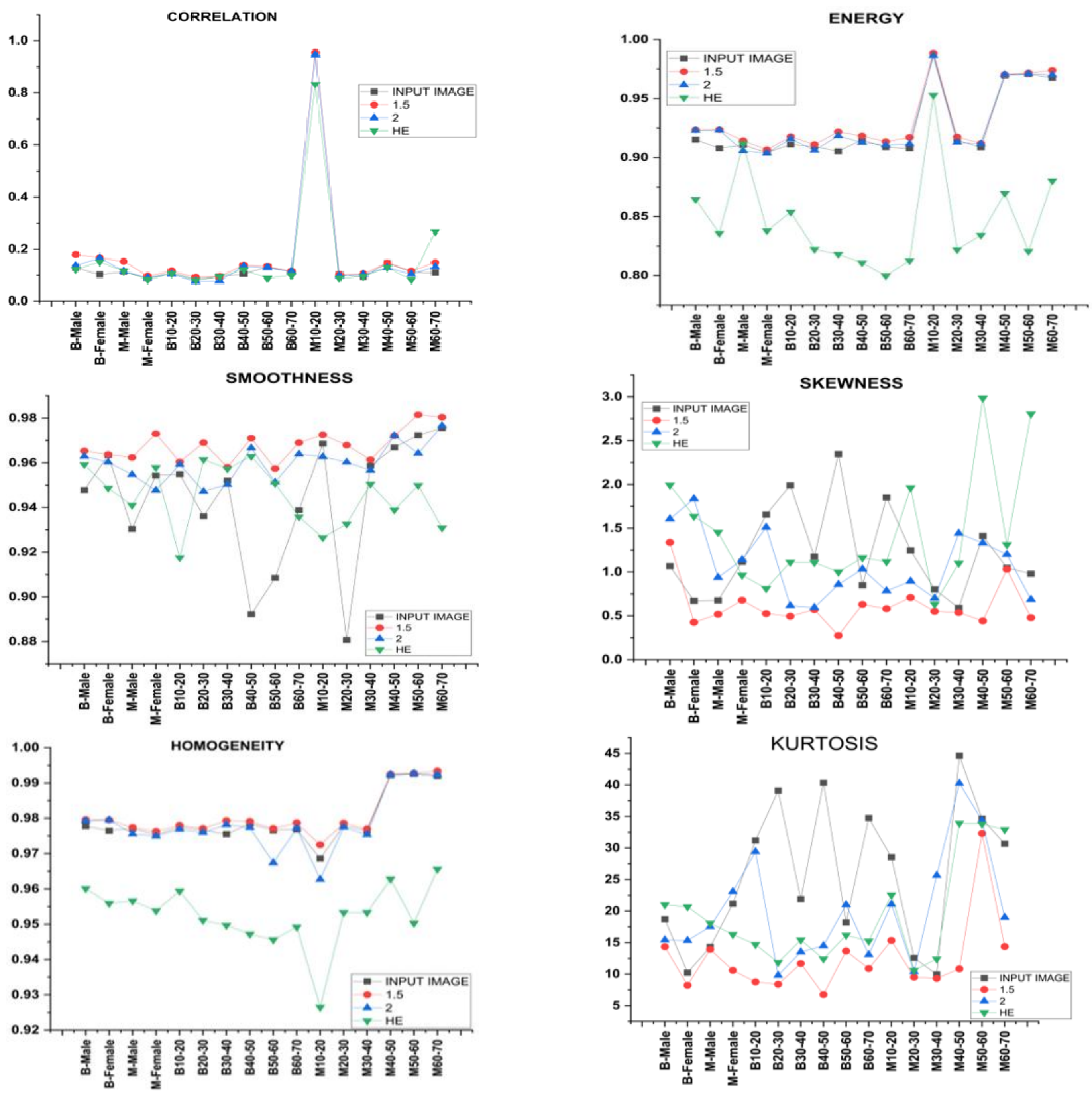


Figure 4: feature extraction results

3.3.2 SUPPORT VECTOR MACHINE

A number of recent studies have demonstrated that support vector machines, or SVMs, outperform other data classification methods in terms of classification accuracy. Support vector machines were used initially to solve the class classification problem, but as computer, network, and database technology advanced quickly to handle the classification and management of massive amounts of data, the class classification problem was unable to keep up with the demands of society. We'll extend this approach to problems involving multiple classes of classification. Support Vector Machines (SVM) is frequently used in classification problems, but they also solve regression problems.

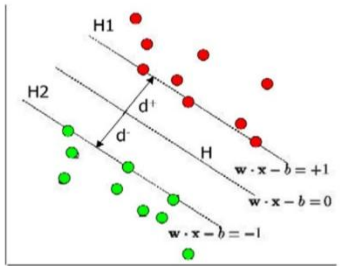


Figure 5: SVM

SVM's primary objective is to successfully separate the given dataset. The objective is to locate a hyperplane in the dataset with the biggest probable margin between support vectors. Margin is defined as the distance between the two nearest positions. The following steps are used by SVM to find the maximum marginal hyperplane:

1. Create hyperplanes that split the classes efficiently. Three possible hyperplanes are shown in the figure: black, blue, and orange. The black class distinguishes correctly between the two, but the blue and orange classes make more classification errors.

2. Choose the hyperplane that has the maximum segregation from the closest data points, as illustrated in fig 4.

Based on the attributes of the images, the SVM classifier classifies the images. In situations where the linear hyperplane cannot solve the problem, SVM employs a kernel method to raise the dimension of the input space, as seen on the right. Using linear separation, we can now easily separate these points.

SVM Kernels: In order to build the SVM, a kernel is used. A kernel is used to transform an input data space into the desired shape. In this instance, a low-dimensional input space is transformed to a higher-dimensional space using the kernel. In other words, it adds new dimensions to non-separable issues to make them separable. Non-linear separation issues are where it performs best. This kernel aids in creating classifiers withgreater accuracy. We have used Radial Basis Function (RBF) kernel in our implementation.

3.3.3 Convolution Neural Networks (CNN)

Yann LeCun first developed the concept of convolutional neural networks (CNN) in 1988. CNNs are one of the neural network subtypes that are widely applied in the area of computer vision. It gets its name from the kind of hidden layers that make up its arrangement. A Convolutional layer, pooling layers, fully linked layers, and normalizing layers constitute CNN's hidden layers. CNNs are very good at extracting features. CNNs, also known as deep artificial neural networks, are commonly used to classify images, organize them based on similarities, and identify objects in scenes. For instance, Facebook utilizes CNN for its automated tagging algorithms, Amazon for its product suggestion engine, and flattening layer and fully connected network. From images, it instantly learns. The classification, object recognition, segmentation, and image processing functions of image analysis can be performed using a CNN. Google for its user image search. A CNN is made up of one or more convolutional layers, followed by Max pooling layer,

The fundamental principles of all CNN designs include continually applying convolutional layers to the input, down sampling the input (via Max pooling), and increasing the number of feature maps. Furthermore, there are fully connected layers, activation procedures, and loss functions (such as cross-entropy or softmax). The most important CNN operations are convolutional layers, pooling layers, and fully connected layers.

The proposed convolutional neural network model is depicted in Fig. 6.

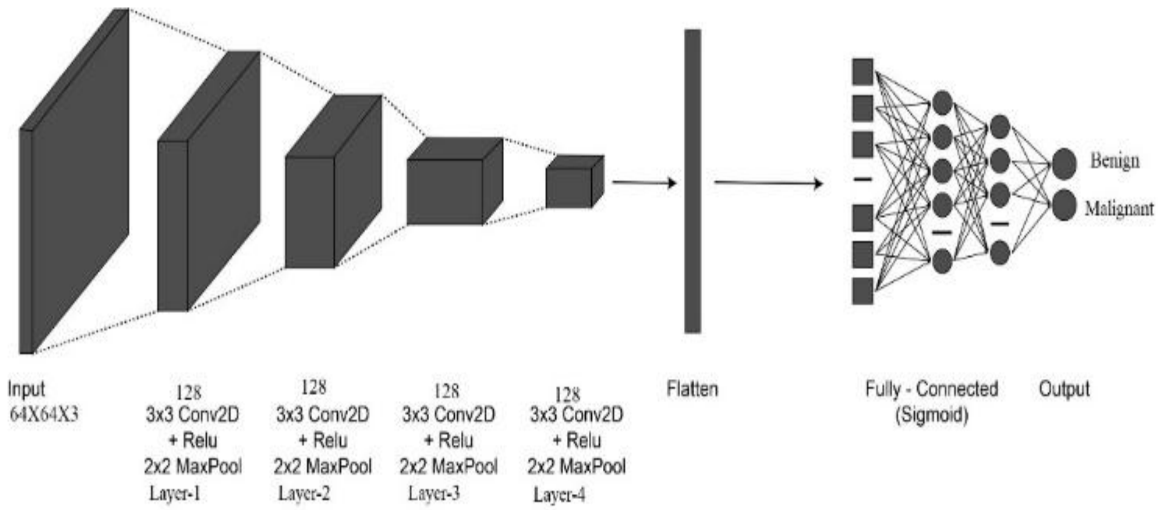


Figure 6: proposed hybrid CNN classifier

The images are resized to 64X64 and applied to convolution layer. The convolution layers use filter of size 3X3 with Rectified Linear Unit (ReLU) as an activation function and 2X2 Max Pool of stride 2 as pooling layer for feature extraction. The flatten layer maps the extracted features into a single column like vector which is 1 dimensional fed to fully connected layer.

Our model comprises of four Convolution layers and four max pooling layers and Flattening layer and fully connected layer.

- Convolution layer: A convolutional layer is the basic building block of a CNN. It consists of a number of filters (or kernels), the properties of which must be learned during training. The filters' sizes are frequently smaller than the input images sizes. After convolution with the picture, each filter generates an activation map.
- Max Pooling layer: The term "max pooling" refers to a pooling technique that selects the largest element from the feature map region that the filter has covered. The output of the max-pooling layer would then be a feature map that contains the most crucial features from the earlier feature map.
- Flattening Layer: This layer converts the multidimensional tensor output from the dropout layer into a one-dimensional tensor.
- Fully Connected Layer: It is the last layer of this classification algorithm, which classifies the images based on its features into benign or malignant.

Image classification results are evaluated by the metrics: accuracy, sensitivity and specificity.
The proportion of correct forecasts that the model generated over the total number of predictions made is referred to as accuracy.

$$Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)}$$

Sensitivity, sometimes referred to as recall or the true positive rate, quantifies the percentage of real positive occurrences that the model accurately recognized.

$$\text{Sensitivity} = \frac{TP}{(TP + FN)}$$

Specificity, on the other hand, measures the proportion of actual negative cases that are correctly identified by the model.

$$\text{Specificity} = \frac{TN}{(TN + FP)}$$

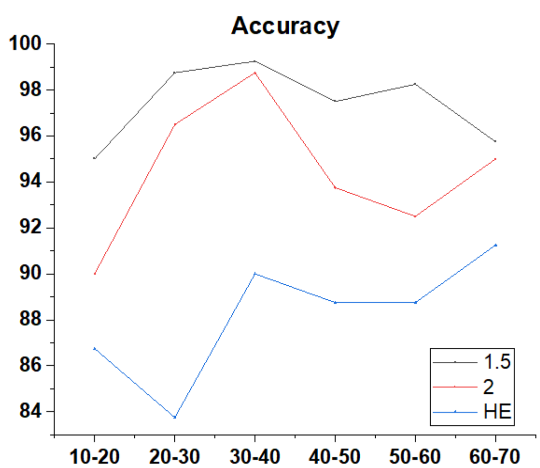
	Accuracy	Sensitivity	Specificity
BF+SVM	97.29	98.29	97.13
BF+CNN	97.75	98.28	97.13

Table 2: evaluation metrics for ER 1.5

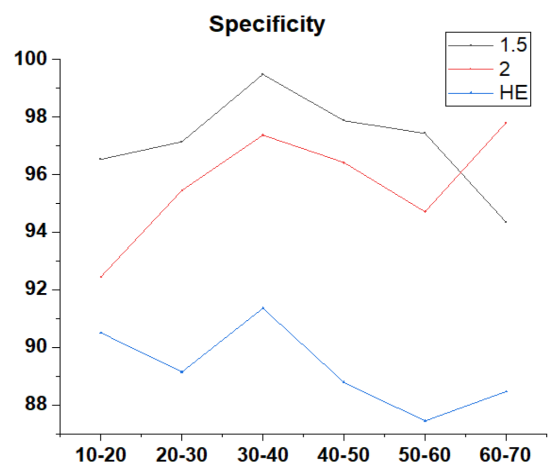
	Accuracy	Sensitivity	Specificity
BF+SVM	94.5	95.17	95.7
BF+CNN	95.25	94.08	95.35

Table 3 : evaluation metrics for ER 2

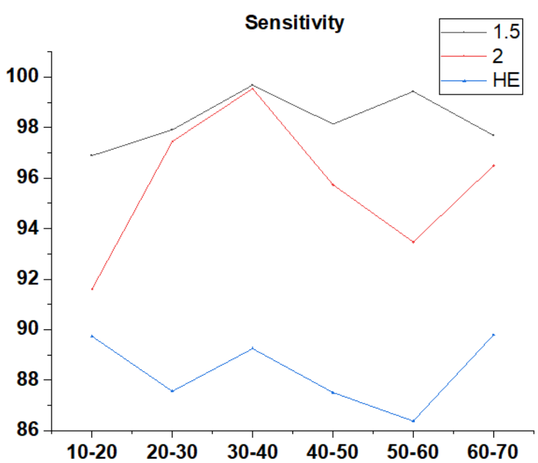
a.



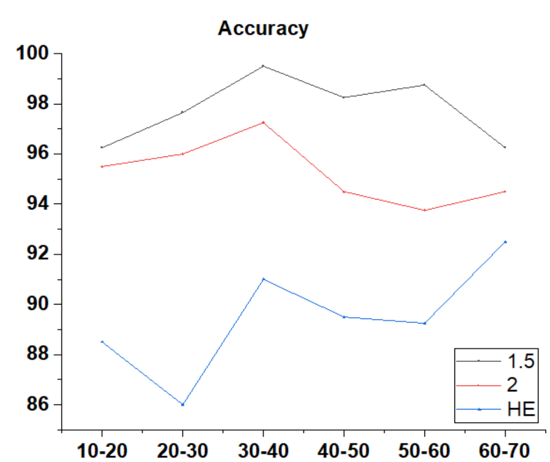
c.



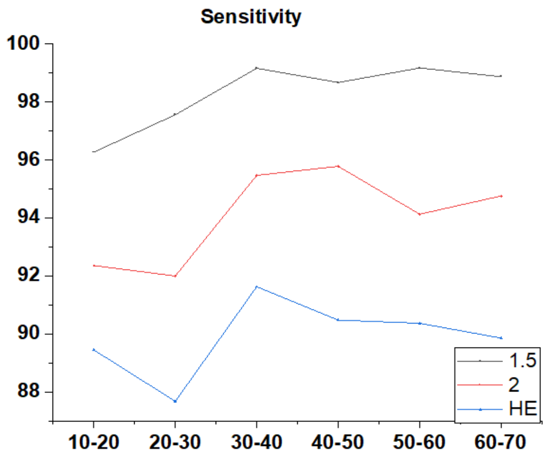
b.



d.



e.



f.

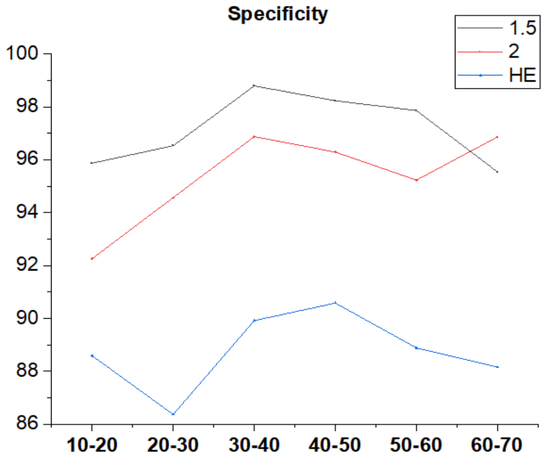


Figure 7: a, b and c represents the accuracy, sensitivity and specificity of Bilateral filter based enhancement and SVM classifier and d, e and f represents the accuracy, sensitivity and specificity of Bilateral filter based enhancement and CNN classifier.

4. **CONCLUSION:** In this paper, the enhancement of skin tumor images using bilateral filter and classification by using SVM and CNN is proposed. In skin tumor analysis, the bilateral filter-based image enhancement preserves the strong edges after enhancement and reduces the artefacts when compared with histogram equalization. It leads to good image enhancement practice for the correct diagnosis of skin-based tumor. After enhancement, the features are extracted and classified using the SVM and CNN classifier. The results are assessed by the qualitative and quantitative metrics, which give superior results when related to Histogram Equalization. The bilateral filter with enhancement ratio 1.5 has high PSNR, Low MSE and high SSIM value compared to HE and enhancement ratio 2. CNN image classification outperforms SVM in terms of classification ratio.

5. REFERENCES

1. Joanna Jaworek-Korjakowska, Pawel Kleczek, “Eskin: Study On The Smartphone Application For Early Detection Of Malignant Melanoma”, Wireless Communications And Mobile Computing, Vol 2018, Article Id 5767360, 11 Pages, 2018.
2. Salem C, Azar D, Tokajian S. “An Image Processing and Genetic Algorithm-based Approach for the Detection of Melanoma in Patients”, Methods Inf Med. 2018 Feb;57(1):74-80. doi: 10.3412/ME17-01-0061. Epub 2018 Apr 5. PMID: 29621833,2018.
3. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F. “Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries”, CA Cancer J Clin. 2021 May;71(3):209-249. doi: 10.3322/caac.21660. Epub 2021 Feb 4. PMID: 33538338.
4. Picardi, A., Lega, I., & Tarolla, E, “Suicide risk in skin disorders. Clinics in dermatology”, 31(1), 47–56,2013.
5. Baig, R., Bibi, M., Hamid, A., Kausar, S., & Khalid, S, “Deep Learning Approaches Towards Skin Lesion Segmentation and Classification from Dermoscopic Images - A Review”, Current medical imaging, 16(5), 513–533,2020.
6. Darrell S. Rigel; Julie Russak; Robert Friedman, “The Evolution of Melanoma Diagnosis: 25 Years Beyond the ABCDs”, 60(5), 301–316. doi:10.3322/caac.20074,2010.
7. Robertson, Fabienne Marie-Louise; Fitzgerald, Lisa, “Skin cancer in the youth population of the United Kingdom”, Journal of Cancer Policy, 12(), 67–71, 2017. doi: 10.1016/j.jcpo.2017.03.003.
8. Barata, C., Celebi, M. E., & Marques, J. S. (2019). “A Survey of Feature Extraction in Dermoscopy Image Analysis of Skin Cancer”, IEEE journal of biomedical and health informatics, 23(3), 1096–1109.
9. Kumar, V., & Choudhury, T, “Real-time recognition of malignant skin lesions using ensemble modelling”, Journal of Scientific

& Industrial Research Vol. 78, pp. 148-153, 2019.

10. Nida N, Irtaza A, Javed A, Yousaf MH, Mahmood MT, "Melanoma lesion detection and segmentation using deep region based convolutional neural network and fuzzy C-means clustering" Int J Med Inform. 2019 Apr; 124:37-48..
11. Saba T, Al-Zaharani S, Rehman A, "Expert System for Offline Clinical Guidelines and Treatment", Life Sci J 2012;9(4):2639-2658, http://www.lifesciencesite.com 393
12. Alquran, Hiam & Abuqasmieh, Isam & Alqudah, Ali & Al-Hammouri, Sajidah & Alawneh, Israa & Abughazaleh, Ammar & Hasayen, Firas, "The Melanoma Skin Cancer Detection and Classification Using Support Vector Machine". 10.1109/AEECT.2017.8257738, 2017.
13. U. B. Ansari, Tanuja Sarode, "Skin Cancer Detection Using Image Processing", Int. Res. J. Eng. Technol., vol. 4, no. 4, pp. 2395–56, 2017, [Online]. Available:
14. Fekrache Dalila, Ameur Zohra, Kasmi Reda, Cherifi Hocine "Segmentation and classification of melanoma and benign skin lesions", Optik - International Journal for Light and Electron Optics, 2017, 140, pp.749 - 761.
15. Mahamudul hasan, surajit das barman, samia islam, and ahmed wasif reza, "skin cancer detection using convolutional neural network", in proceedings of the 2019 5th international conference on computing and artificial intelligence (iccai '19). Association for computing machinery, new york, ny, usa, 254–258.
16. Khan M.A., Akram T., Zhang Y.D., Sharif M. Attributes based skin lesion detection and recognition: A mask RCNN and transfer learning-based deep learning framework Pattern Recognition Letters, 143 (2021), pp. 58-66.
17. Mahbod A., Schaefer G., Wang C., Dorffn er G., Ecker R., Ellinger I. Transfer learning using a multi-scale and multi-network ensemble for skin lesion classificationComputer Methods and Programs in Biomedicine, 193 (2020), Article 105475
18. ROTEMBERG, V., KURTANSKY, N., BETZ-STABLEIN, B., CAFFERY, L., CHOUSAKOS, E., CODELLA, N., COMBALIA, M., DUSZA, S., GUITERA, P., GUTMAN, D., HALPERN, A., HELBA, B., KITTLER, H., KOSE, K., LANGER, S., LIOPRYS, K., MALVEHY, J., MUSTHAQ, S., NANDA, J., REITER, O., SHIH, G., STRATIGOS, A., TSCHANDL, P., WEBER, J. & SOYER, P. A PATIENT-CENTRIC DATASET OF IMAGES AND METADATA FOR IDENTIFYING MELANOMAS USING CLINICAL CONTEXT. SCI DATA 8, 34 (2021).
19. Gupta, S., Gupta, R. & Singla, C. Analysis of image enhancement techniques for astrocytoma MRI images. Int. j. inf. tecnol. 9, 311–319 (2017). https://doi.org/10.1007/s41870-017-0033-8
20. Kumar, M., Priyanka Various image enhancement and matching techniques used for fingerprint recognition system. Int. j. inf. tecnol. 11, 767–772 (2019). https://doi.org/10.1007/s41870-017-0061-4.
21. Mun, H., Yoon, GJ., Song, J. et al., "Scalable image decomposition". Neural Comput & Applic 33, 9137–9151,2021.
22. Revaud, Jerome, Philippe Weinzaepfel, Zaid Harchaoui, and Cordelia Schmid. "Epic Flow: Edge-Preserving Interpolation of Correspondences for Optical Flow." In 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 1164– 72. Boston, MA, USA: IEEE, 2015.
23. Tomasi, C., and R. Manduchi. "Bilateral Filtering for Gray and Color Images". In Sixth International Conference on Computer Vision (IEEE Cat. No.98CH36271), 839–46. Bombay, India: Narosa Publishing House, (1998).
24. Shervin Minaee and Yuri Boykov and Fatih Porikli and Antonio Plaza and Nasser Kehtarnavaz and Demetri Terzopoulos, Image Segmentation Using Deep Learning: A Survey, 2020, 2001.05566.
25. Rajni Bala, Dr. Dharmender Kumar, "Classification Using ANN: A Review", International Journal of Computational Intelligence Research ISSN 0973-1873 Volume 13, Number 7, pp. 1811-1820,2017.
26. Khalid S, T. Khalil, and S. Nasreen, "A survey of feature selection and feature extraction techniques in machine learning," in 2014 Science and Information Conference. IEEE, 2014, pp. 372–378 (2014).