**Vision based dermatological disease(acne) detection**

Prof. Jyoti Madake, Tejal Salunke, Riddhi Selkar, Shivangi Shardul, Dr. Shripad Bhatlawande, Dr. Swati Shilaskar, Dr. Rajesh Jalnekar.

***Abstract*—*Now-a-days skin illness is becoming one of major issues all over the world. Frequent changes in the climatic conditions is the main cause of skin problems. Many techniques and technologies are used for classification of different types of skin diseases. Skin disease diagnosis by observations is a very challenging task for both patients as well as doctors. Also in developing countries it is very difficult to consult a doctor for treatment because the treatment is very expensive and all can not afford it. So, considering this as one of the objectives this acne detection system is implemented. Skin pigmentation types in people vary from person to person and also skin types are different like oily, dry and mixed. In this model, the detection of acne is done by using machine learning and computer vision techniques. SIFT is used for feature extraction and PCA technique is used for dimensionality reduction. Many different classification algorithms are used like Decision Tree, LGBM. Along with SIFT, GLCM is also used as a feature extraction algorithm. And it was found that GLCM gave maximum accuracy and was fast and more efficient.***

***Keywords - Acne, Dermatology, GLCM, LGBM, DT***

1. Introduction

Dermatology focuses on hair, skin, nails and also the diseases that impact them. In the widest sense, a dermatologist examines skin, scalp, hair, and nail problems, and some of the cosmetic issues. Acne is now the most frequently observed facial condition. Acne affects all the age groups, teenagers as well as adults. At first acne is converted into inflammatory acne then it transforms into cystic acne. This cystic acne is very painful and it goes deep under the skin. Acne is mainly caused by the hormonal changes that take place in our body and it also depends on other factors like stress , food habits , poor food quality, changes in weather and also the medications. Patients and clinicians, especially in developing countries, require an automated skin disease diagnosis method. Especially considering the fact that a number of research studies have been conducted to apply computer vision-based algorithms to identify dermatological skin diseases, almost all of them have concentrated on only two or three diseases. Skin examination is very critical to complete before starting any other treatment. Hand count and mark acne on the face of patients are mostly used skin analysis methods but these methods are not really safe and it also results in pimples occurring on the skin and it also takes lots of our time. Using Machine learning and Computer vision techniques we achieved very good accuracy.

1. Literature survey

Verma, et. al, used different techniques[1] of classification like PAC, BNB, LDA, RNC and NB for skin disease prediction and to improve the accuracy of algorithms Bagging , Gradient boosting and Adaboost like ensemble techniques were used.With 99.68% of high accuracy the prediction is done. The other approach[2] applied the two-stage method ,in the first stage the images of skin disorders undergo many types of pre-processing techniques like grayscaling, sharpening, smooth and median filter , RGB extraction and sobel operator before performing feature extraction. In the second stage[6] to detect histopathological characters of skin disease, machine learning algorithms were utilized. After training,testing the system shows accuracy of 95% for six different disorders.

Roy, K. et. al,[3] with the help of different image segmentation techniques has proposed a system for detection of different skin diseases. The techniques used were adaptive thresholding, k-means clustering, edge detection and morphological-based image segmentation and noise reduction was also implemented. Using OpenCV and python they have divided the picture on the basis of region or edge detection. Four types of skin disorders were considered and four segmentation techniques were applied. The resulting image was displayed on the basis of signal to noise ratio. [4]In one of the related studies the authors have presented texture and variation for healthy skin and few skin disorders using techniques like colour segmentation and imaging.

Alamdari, N. et. al. have proposed a very efficient system using HSV methods[5] and k-means but the texturing approach used in this is needed to be improved because of insufficient system. The clustering method is better than color-based method when the HSV and k-means were compared at two levels.For classifying the images they used the best segmentation method,k-means clustering and this two-level clustering improved the accuracy of the system.

Firstly the reduction of noise[7] is done by using enhancement techniques and after that two noise reductions were applied. The comparison of median and gaussian filters were used for increasing the efficiency in addition to that the four color spaces RGB, HSV, YUV ,YCbCr were employed for feature extraction.This is a mobile application with simple and easy interface. In the similar way the authors implemented the system [8] that consists of HSV conversion, RGB to gray conversion , binary threshold , ROI extraction and Non-ROI features elimination. About calculating efficiency, it observed that if the facial problem is similar to acne there is a high chance of the result to be shown as false positive. In the last accuracy, sensitivity and precision were measured.

G. Maroni,et. Al, [9] designed the system for improving skin segmentation model performance by using unsupervised features that consist of 10 Random Forest models and on the FSD dataset achieved the good accuracy in acceptable computation time. The CIELab model’s channel a\* has been proven to enhance acne lesion and Adaptive Thresholding on this channel produced satisfactory results in separating acne lesion from healthy skin.

In this work,[10] researchers started locating the face region of the input image. The ROI was found using facial characteristics and skin tone and to determine a reasonable threshold for removing possible faults, an approximate poisson distribution is employed. Then, to pick relevant features for classification, the SFFS is used and for classifying possible issues into normal,acne and spot , two SVM made decision tree classifiers are used. According to results, the suggested approach can identify facial skin defects and also recognize lesions.

Ajith, et. al, [11] have used image preprocessing techniques for skin disease detection. In this model, the patient will first have to input the image for his/her skin disorder. The model has a database which includes 6 types of skin disorders that can be used by the patient for detecting the acne and that will help in getting the information about the disorder before visiting the doctor. All the steps are applied on the input images and the output image is displayed with disease that has been diagnosed. Comparative study is done using transforms like DCT, DWT and SVD. [12] In this paper, they have proposed a system which is automated and examines the image of the skin illness and diagnoses the disorder. They have classified the dataset into 9 skin disorders. The accuracy obtained in the model is 90%. They have classified the dataset into 9 skin disorders.

Since 2016, Researchers reviewed 45 research studies that used deep learning [13] technologies to identify skin diseases. The disease kind, data set, data processing technology, data augmentation technology, model for skin disease image recognition, deep learning framework, evaluation indicators, and model performance are all examined in these works. In addition, they present a comparison of traditional and machine learning-based approaches for diagnosing and treating skin diseases. They also look at the current state of this field and predict four probable study directions for the future. Their findings reveal that dermatologists and other computer-aided therapy approaches in skin disease diagnosis are outperformed by deep learning-based skin disease picture recognition strategies, notably the multi deep learning model fusion method.

1. Methodology

This paper presents an acne detection system which will detect the acne and categorize it into its type. Firstly, the dataset of around 20k images was collected and categorized into 6 types. Preprocessing of the images was done which included resizing, gray scaling, rotation and flipping. Then, feature extraction and data augmentation was performed on the images. Feature extractors used for the process were SIFT and for texture feature extraction GLCM was used. And along with this feature selection and dimensionality reduction was done using K-means and PCA respectively. Classification of the acne type was done using algorithms like Decision tree and LGBM.

A) System block diagram

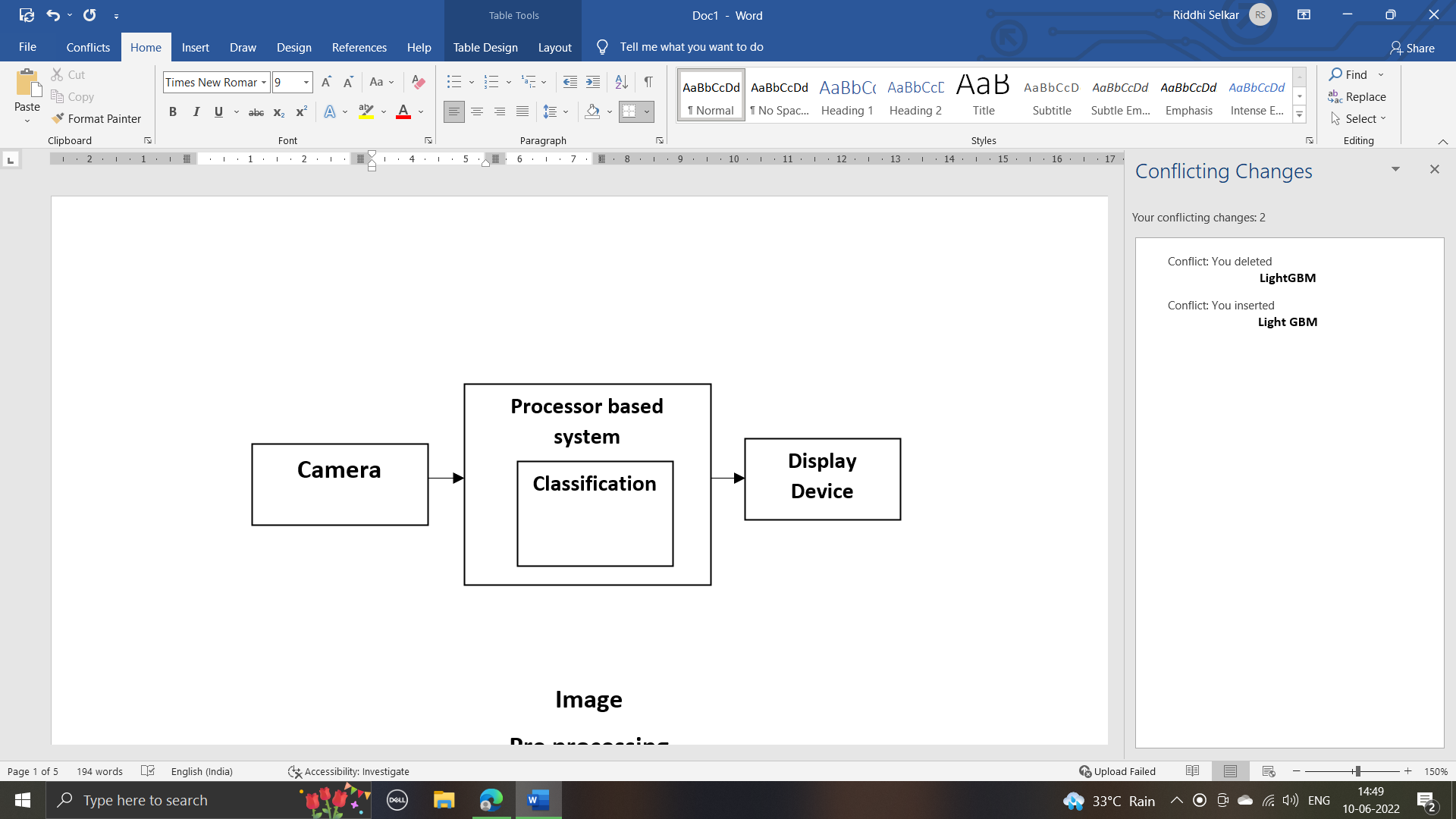


Fig1. Block diagram of the project

This project hardware consists of a camera (of laptop) for taking inputs from the user. Then the processor based system for classifying and detecting acne. At last the display device to output the results.

# B ) Details of the dataset preparation, preprocessing

Dataset was prepared of 20k images that were downloaded from the standard dataset named dermnet.in. The collected 20k images of acne disorders were combined and then categorized into 6 types which are acne vulgaris, scars, infantile, back, cyst, moles and also included healthy faces for classification. The distribution of the types of acne is shown in table no.1.

Table 1. Distribution of types of acne.

|  | **Type of acne** | **Number Of Images** |
| --- | --- | --- |
| 1 | Acne vulgaris | 8241 |
| 2 | Acne infantile | 8627 |
| 3 | Scars | 608 |
| 4 | Back | 385 |
| 5 | Moles | 77 |
| 6 | Healthy faces | 4511 |
| 7 | Total | 22,449 |





Fig 2. Few examples of acne

# c)Image Preprocessing and Data Augmentation

From the dataset around 9k images, which included two acne types were preprocessed to reduce the distortions in them. Preprocessing of the images included gray scaling, rotation and flipping of the images, brightening. Renaming and labeling was performed on the dataset and the images were resized to dimension of (280x430).

Data augmentation was used to improve the performance and results of our model. After augmentation, the dataset had around 14k images in the dataset.

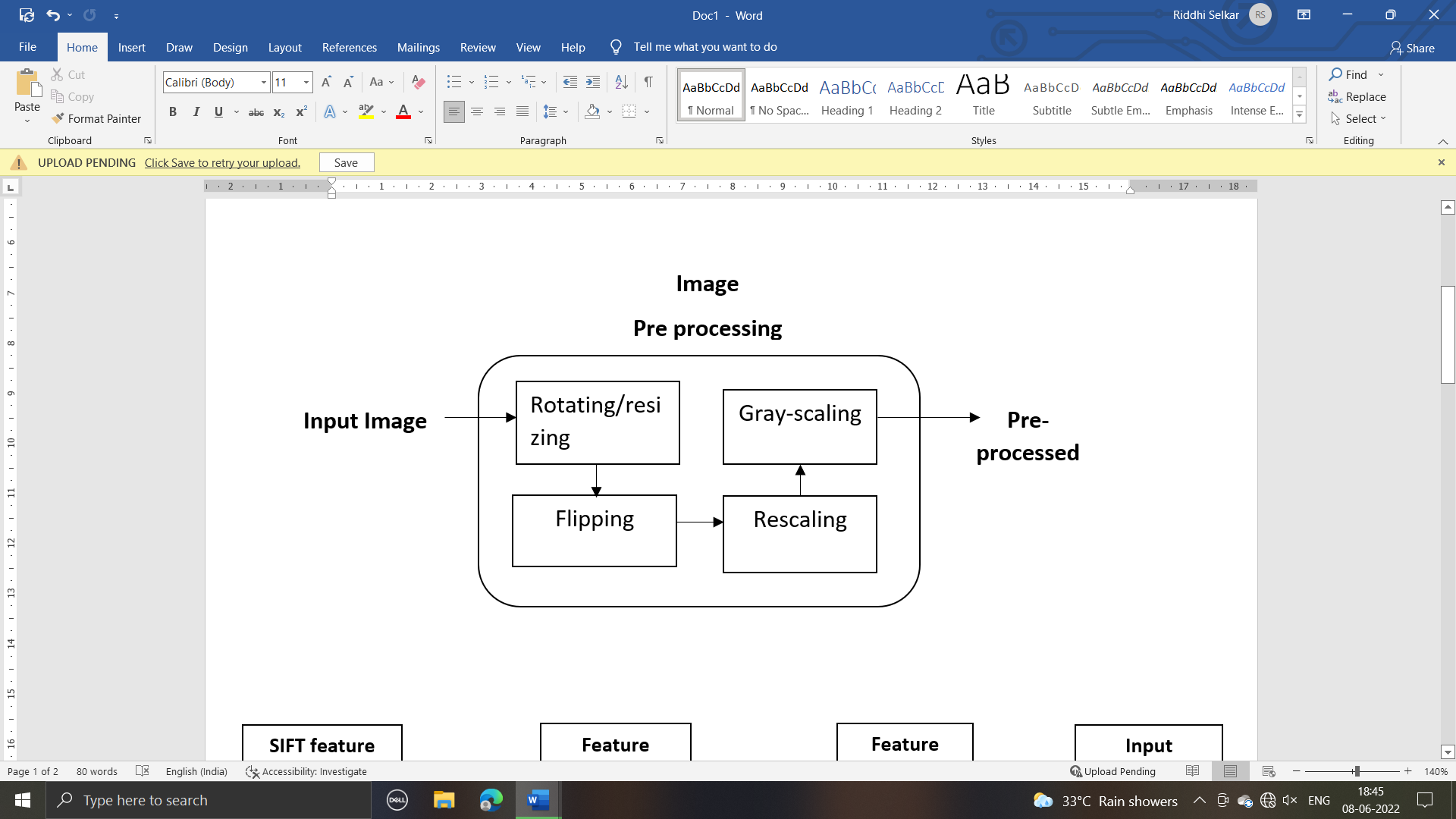


Fig 3. Data Augmentation

# d) Feature Extraction and Dimension Reduction

In the proposed system, feature extraction on the dataset is performed using two algorithms and dimensionality reduction is done using PCA technique. The two approaches used for feature extraction are as follows:

**A**] **Feature extraction using SIFT:** SIFT (scale invariant feature transform) is a feature extraction approach which is used widely. This feature is independent of viewpoints as well as image scale and orientation. The feature vector of (4286640,128) was obtained in by using scale invariant feature transform technique, after doing clustering using k-means with k=5 the feature vector of (5850,7) was obtained and at last after dimensionality reduction feature vector of (5850,4) using principal component analysis was obtained. These features are then given as input to Decision Tree classifier which further classifies the acne infantiles and acne vulgaris images.

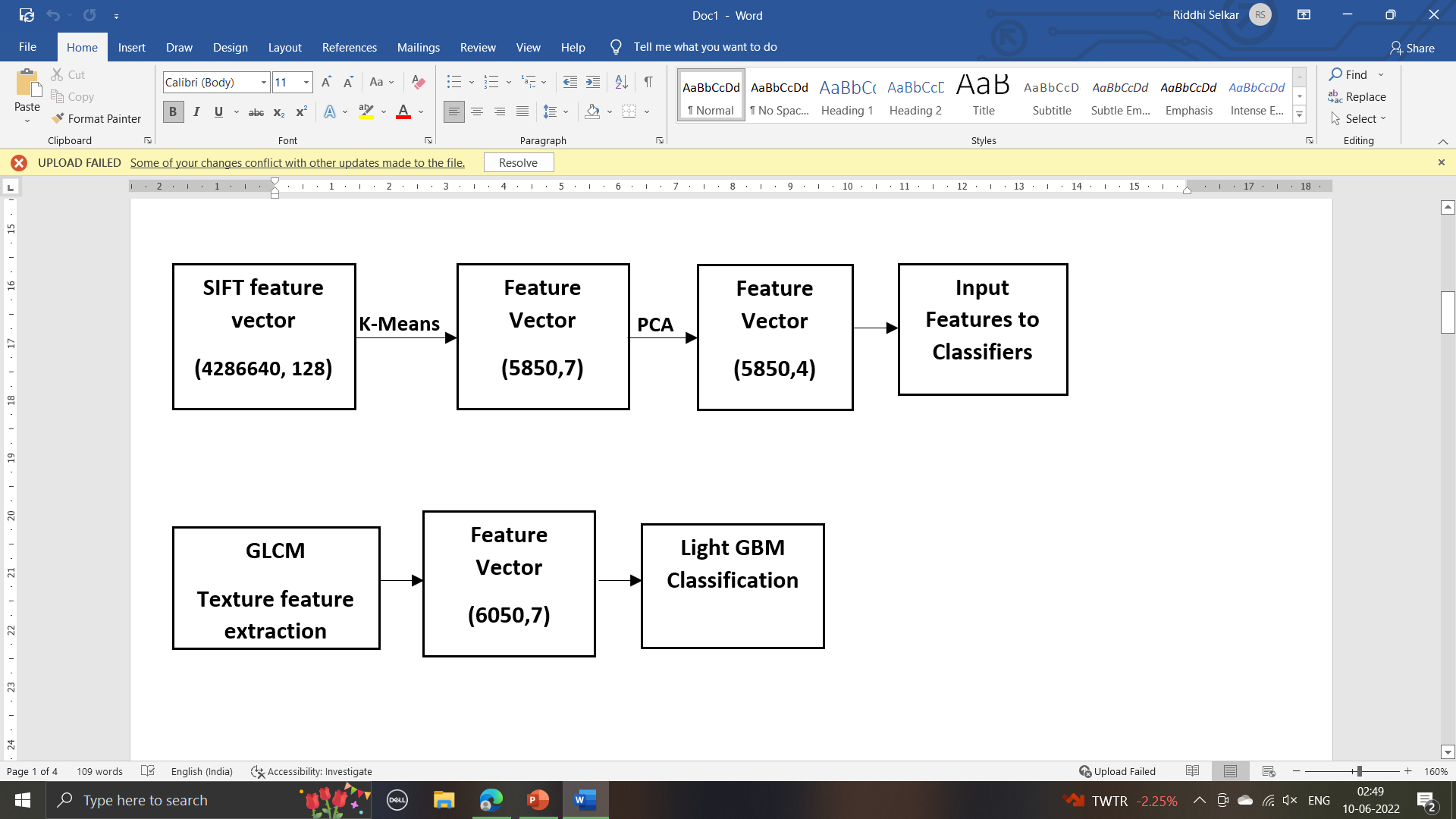


Fig. 4 Block diagram of SIFT

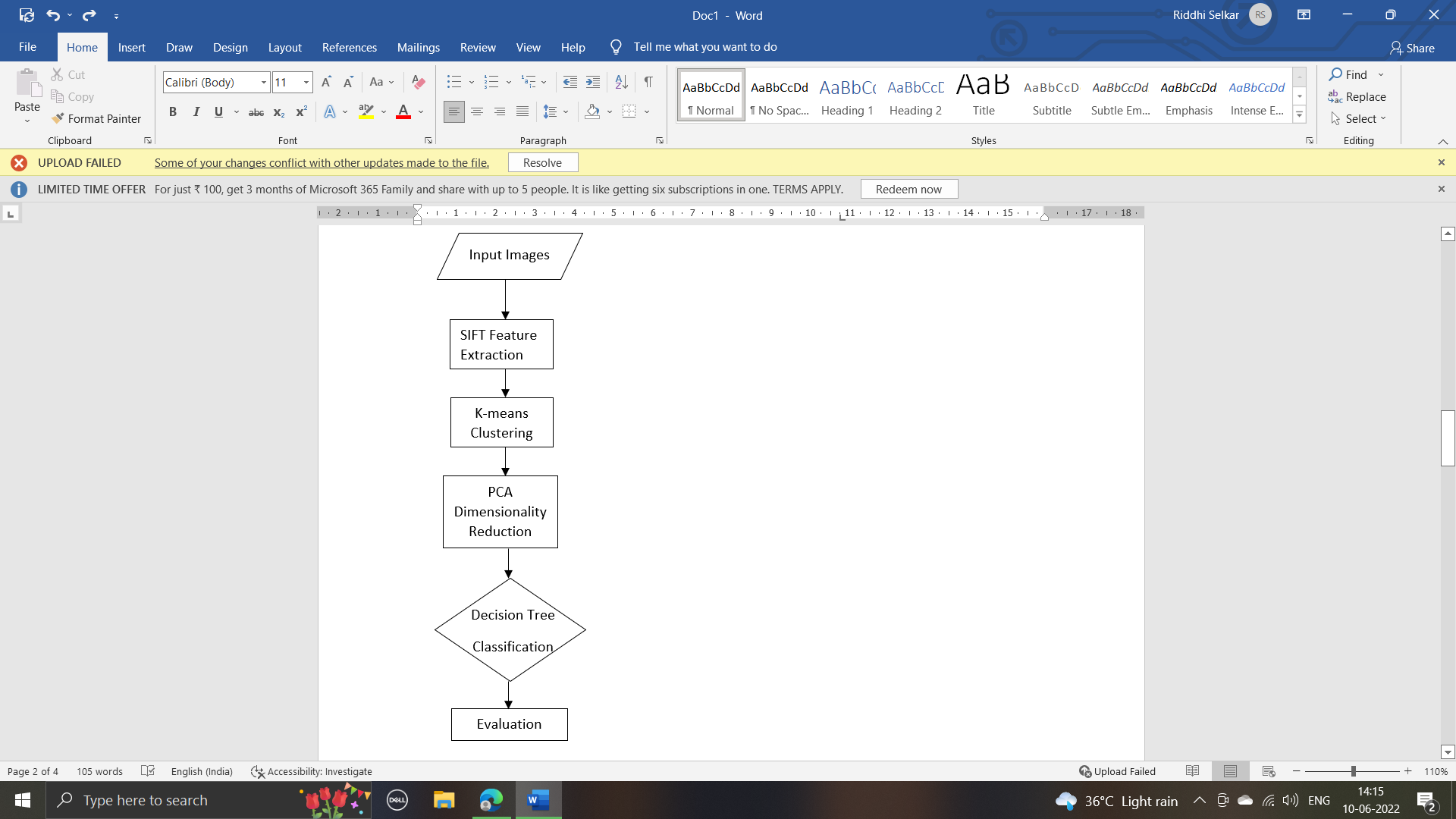
****

Fig. 5: Flowchart of feature extraction by SIFT

**B**] **Texture feature extraction using GLCM:** Gray-Level Co-occurrence Matrix is a feature extraction approach that performs feature analysis. This method is used to extract the properties like contrast, homogeneity from the pictures. After the texture feature extraction of all these properties the feature vector of (6050,7) was generated and further features are given as an input to Light GBM classifier which classifies the acne types and detects the actual and predicted types along with the images.

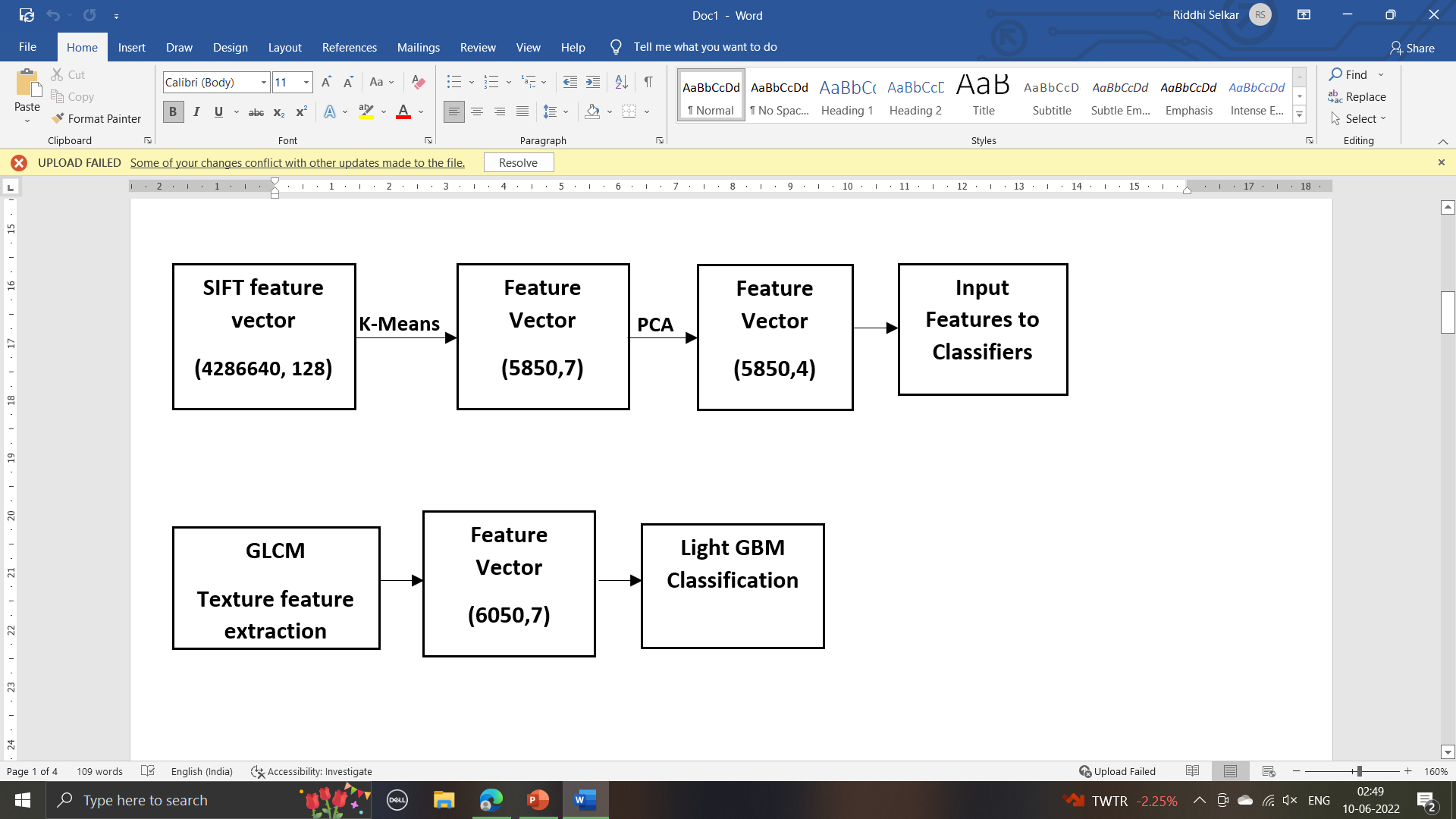


Fig 6 Block diagram of GLCM

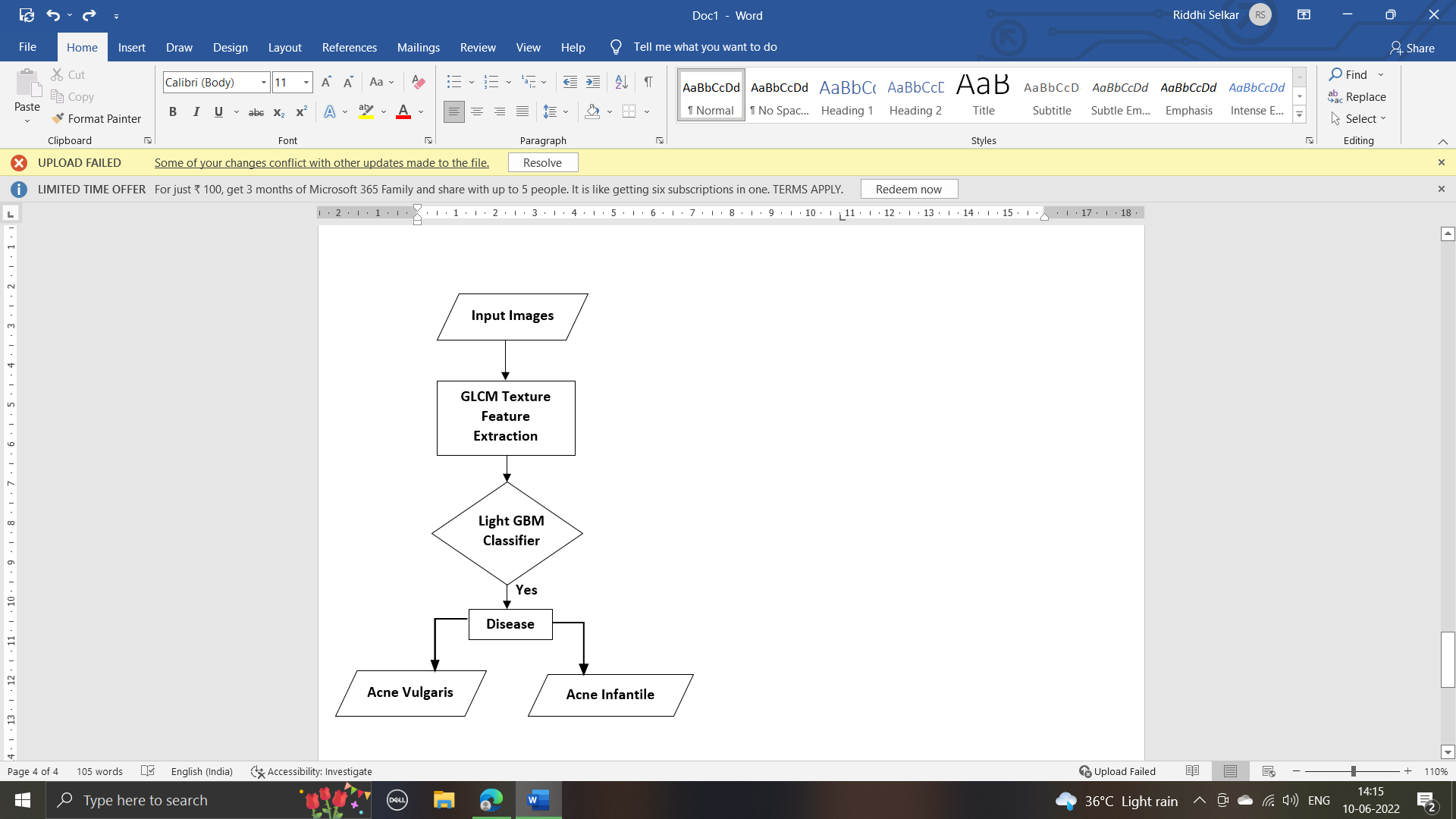


Fig. 7: Flowchart of texture based feature extraction using GLCM

# d) Algorithms

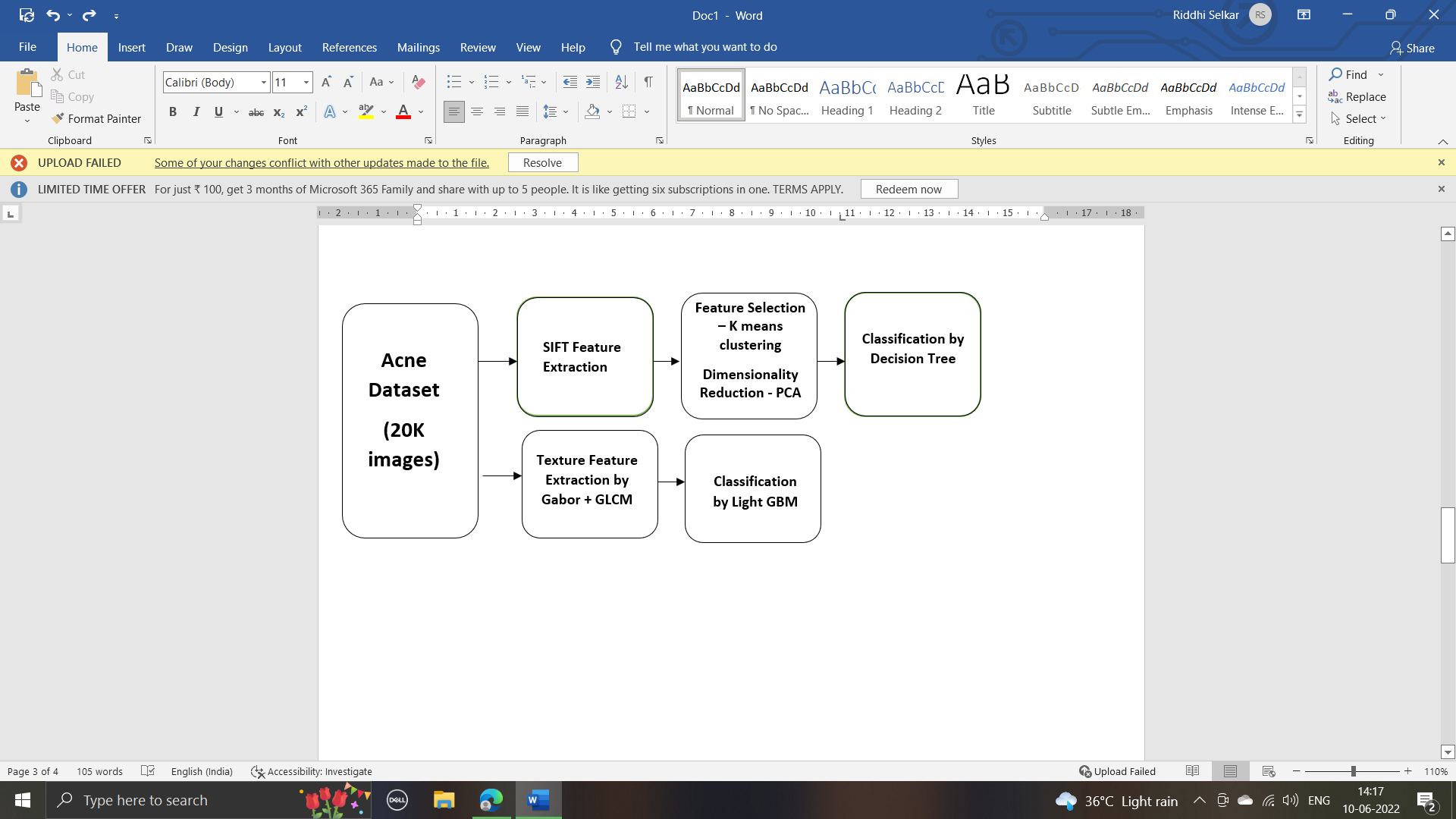


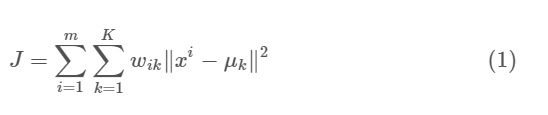
Fig 8. Algorithms used in project

The figure describes the actual flow of the project. The model is divided into two parts one is feature extraction using SIFT and classification using Decision Tree and the other is Texture based feature extraction using GLCM and classification and detection using Light GBM algorithm.

The detailing of each algorithm and approach is given below.

**Algorithms :**

**Feature selection using K-Means:** It is a type of unsupervised learning that clusters an unlabeled dataset into groups. In this clustering, each cluster has its own centroid so it is called a centroid- based approach. It helps in reducing the sum of distance between the data points and the clusters. The objective function of k-means is written as:



where wik=1 for data point xi if it belongs to cluster *k*; otherwise, wik=0. Also, μk is the centroid of xi’s cluster.

**Feature dimensionality reduction using PCA:** Principal Component Analysis helps in dimensionality reduction of features using orthogonal transformation in which observations of correlated features are converted into a set of linearly uncorrelated data. It is also an unsupervised approach.

**Algorithm 1**: Algorithm for dimensionality reduction

**Input:** Image

**Output:** Reduced feature vector

*Initialization:*

*LOOP Process*

1: **for** each image in directory **do**

2:Extract SIFT features

3: Transform using k-means (k=5)

4: Normalize using MinMaxScaler

5: Append to feature vector

6: **end for**

7: Fit PCA (n\_components=5)

8: Transform using PCA

9: **return** reduced feature vector

**Approach 1:**

**Feature classification using DT:** This algorithm is a type of supervised learning that helps in predicting or categorizing the outcomes on the basis of answers that are obtained from the previous set of questions. This aims at finding almost all possible results for an issue. The entropy equation is given as:

where pi is the probability of class i. Entropy helps in determining the gain in information in each node of the decision tree.

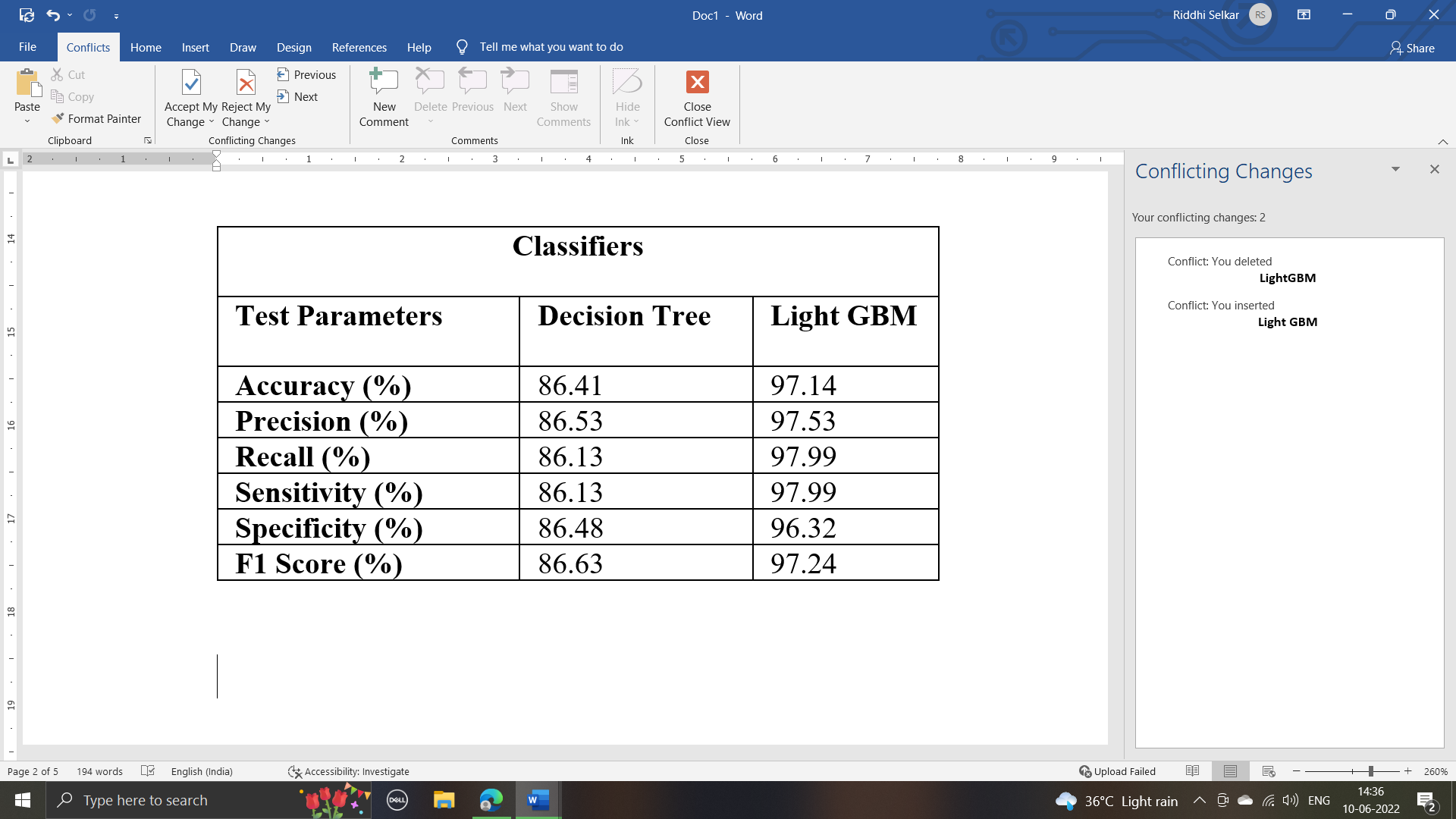
**Approach 2:**

**Feature classification using LGBM:** LightGBM is an algorithm that aims at reducing the efficiency of the model and also helps in reducing the memory usage. It is a decision tree algorithm which is a gradient boosting framework that is fast and high performance.

## IV.Result and Discussion

The model was trained with a dataset of 14k images of disorders of acne. The model was built on the Desktop - UVHT37464 bit OS with intel core i5-10210U CPU @ 1.60GHz 2.11 GHz frequency and RAM of 16.0 GB. Different algorithms were implemented in the model which included SIFT and GLCM for feature extraction respectively and Decision tree and LGBM as classification algorithms. The accuracy obtained from the above algorithms was 86% and 97% resp. The highest accuracy was obtained by the LGBM algorithm compared to that of DT. The accuracies obtained by implementing different results are shown below in table no.2

Table 2. Table describing accuracy with different models



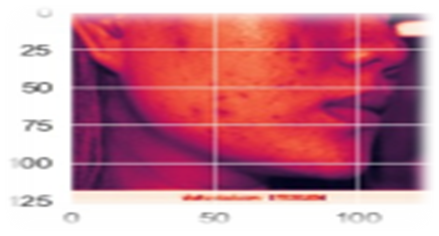


Fig 9. Resultant image 1

The prediction of this image is : Acne Vulgaris

The actual label for this image is : Acne Vulgaris

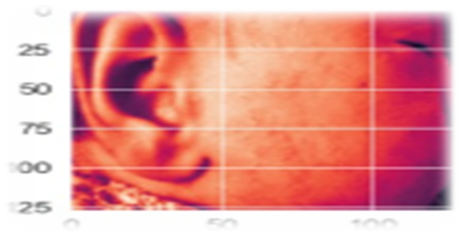


Fig 10. Resultant image 2

The prediction of this image is : Acne Infantile

The actual label for this image is : Acne Infantile

V. Conclusion

In this project, we have proposed a system which helps in detection of acne. Dataset of 20k images, categorized into 6 types of acne disorders has been trained and tested. Different techniques and algorithms have been implemented on the model and the accuracy obtained by using Decision tree classifier was found to be 86% and by using LGBM was found to be 97%. Feature extraction algorithms used were SIFT and GLCM respectively.

This is a model which can be used by medical practitioners, dermatologists, skin disease doctors and by any person having a little bit of knowledge about the dermatology field and wanting to diagnose the type of it.

##### VI. References

1. Verma, A.K., Pal, S. and Kumar, S., 2020. “ Prediction of skin disease using ensemble data mining techniques and feature selection method—a comparative study ”. Applied biochemistry and biotechnology, 190(2), pp.341-359. Kumar, V.B., Kumar, S.S. and Saboo, V., “ Dermatological disease detection using image processing and machine learning ”. In 2016 Third International Conference on Artificial Intelligence and Pattern Recognition (AIPR) (pp. 1-6). IEEE.
2. Roy, K., Chaudhuri, S.S., Ghosh, S., Dutta, S.K., Chakraborty, P. and Sarkar, R., 2019, March. “Skin Disease detection based on different Segmentation Techniques”. In 2019 International Conference on Opto-Electronics and Applied Optics (Optronix) (pp. 1-5). IEEE.
3. Alam, M.N., Munia, T.T.K., Tavakolian, K., Vasefi, F., MacKinnon, N. and Fazel-Rezai, R., 2016, August. “Automatic detection and severity measurement of eczema using image processing”. In 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (pp. 1365-1368). IEEE.
4. Alamdari, N., Tavakolian, K., Alhashim, M. and Fazel-Rezai, R., 2016, May. “Detection and classification of acne lesions in acne patients: A mobile application”. In 2016 IEEE international conference on electro information technology (EIT) (pp. 0739-0743). IEEE.
5. Kumar, V.B., Kumar, S.S. and Saboo, V., 2016, September. “Dermatological disease detection using image processing and machine learning”. In 2016 Third International Conference on Artificial Intelligence and Pattern Recognition (AIPR) (pp. 1-6). IEEE.
6. Haddad, A. and Hameed, S.A., 2018, September. “Image analysis model for skin disease detection: framework”. In 2018 7th International Conference on Computer and Communication Engineering (ICCCE) (pp. 1-4). IEEE.
7. Chanthara Phaichi, T., Uyyanonvara, B., Sinthanayothin, C. and Nishihara, A., 2015, March. “Automatic acne detection for medical treatment.” In 2015 6th International Conference of Information and Communication Technology for Embedded Systems (IC-ICTES) (pp. 1-6). IEEE.
8. Maroni, G., Ermidoro, M., Previdi, F. and Bigini, G., 2017, November. “Automated detection, extraction and counting of acne lesions for automatic evaluation and tracking of acne severity”. In 2017 IEEE symposium series on computational intelligence (SSCI) (pp. 1-6). IEEE.
9. Chang, C.Y. and Liao, H.Y., 2011, October. “Automatic facial skin defects detection and recognition system.” In 2011 Fifth International Conference on Genetic and Evolutionary Computing (pp. 260-263). IEEE.
10. Ajith, A., Goel, V., Vazirani, P. and Roja, M.M., 2017, June. “Digital dermatology: Skin disease detection model using image processing.” In 2017 International Conference on Intelligent Computing and Control Systems (ICICCS) (pp. 168-173). IEEE.
11. Yasir, R., Rahman, M.A. and Ahmed, N., 2014, December. “Dermatological disease detection using image processing and artificial neural network.” In 8th International Conference on Electrical and Computer Engineering (pp. 687-690). IEEE.
12. Li, L.F., Wang, X., Hu, W.J., Xiong, N.N., Du, Y.X. and Li, B.S., 2020. “Deep learning in skin disease image recognition: A review.” IEEE Access, 8, pp.208264-208280.
13. ALEnezi, N.S.A., 2019. “A method of skin disease detection using image processing and machine learning”. Procedia Computer Science, 163, pp.85-92.
14. Goceri, E., 2019, October. “Skin disease diagnosis from photographs using deep learning”. In ECCOMAS thematic conference on computational vision and medical image processing (pp. 239-246). Springer, Cham.
15. Zhang, B., Zhou, X., Luo, Y., Zhang, H., Yang, H., Ma, J. and Ma, L., 2021. “Opportunities and Challenges: Classification of Skin Disease Based on Deep Learning.” *Chinese Journal of Mechanical Engineering*, *34*(1), pp.1-14.