iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 Issue: IV Month of publication: April 2026

DOI: https://doi.org/10.22214/ijraset.2026.81557

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

Tongue Colour Classification using Image Processing and Deep Learning

Mrs. Md. Jareena Begum, Tandra Hima Sri, Kallagunta Sai Pallavi, Javisetty Maheswari
Department of Computer Science and Engineering Bapatla Women's Engineering College

Abstract: Problem: The technique of tongue diagnosis is an important aspect of Traditional Chinese Medicine, wherein the color and appearance of the tongue provide clues to the health and disease of the individual. However, the conventional methods of diagnosis are quite subjective and are based on the observation and judgment of the individual, which sometimes leads to inconsistent results. Therefore, the paper proposes an intelligent system for classifying the colors of the tongue, which is based on the concept of artificial intelligence. The system classifies the tongues into five categories, which include Dark Red, Light Red, Purple, Red, and White. For the segmentation of the tongue region from the image, the paper proposes the use of OpenCV, which is a computer vision library. For the extraction of the relevant features, the paper proposes the use of the Residual Network with the attention mechanism, which selects the relevant features from the image. The classification is done by the Random Forest algorithm. The proposed system has been tested with a data set of 1,068 images, and the accuracy of the system is above 96%, which is higher than the conventional methods of image processing. In addition to the above, the system has been implemented as an interactive web page, which facilitates the real-time analysis of the tongues and the diagnosis of the disease.
Keywords: Traditional Chinese Medicine, Tongue Diagnosis, Deep Learning, ResNet, Computer Vision, Medical AI.

I. INTRODUCTION

In traditional Chinese medicine, observation is considered one of the four key diagnostic techniques, together with auscultation and olfaction, inquiry, and palpation. Among these observation techniques, tongue diagnosis plays an especially important role in determining the inner condition of the body. In TCM, the tongue is considered an external reflection of inner organ functions and diseases, with different characteristics of the tongue, especially color, being an important indicator of the inner condition of the body.

For example, a pale or white color of the tongue is considered an indication of Yang deficiency, cold diseases, or anemia, while a deep red or purple color is considered an indication of inner heat, inflammation, or blood stasis, which is associated with cardiovascular diseases or liver problems. Early identification of these signs enables practitioners to provide timely prevention and treatment.

Despite its importance, traditional tongue diagnosis has many limitations, especially in that it is based on observation, perception, experience, and judgment of the practitioner, which may lead to different conclusions regarding the same patient. Other external factors, such as light, camera quality, and patient position, may also introduce more variability in results.

Recent advancements in Artificial Intelligence (AI) and computer vision provide solutions for such challenges. Machine learning techniques can automatically recognize useful features from images and classify images objectively and accurately. Moreover, combining deep learning techniques and image processing can create standard diagnostic tools, which are less dependent on human subjectivity and environmental factors.

This study proposes an automated framework using AI techniques for accurate tongue color classification, which is useful for TCM diagnosis. It combines computer vision and deep learning techniques for an efficient approach for tongue image analysis.

A. The contributions of this study are:

- An automated approach for accurate tongue image segmentation using HSV masking for accurate extraction of the tongue region from cluttered scenes.
- A hybrid approach for accurate tongue color classification using attention-augmented ResNet and Random Forest classifier.
- A web-based diagnostic tool for real-time tongue analysis and TCM diagnosis.

II. RELATED WORK

In recent years, the integration of artificial intelligence and computer vision for medical diagnosis has received considerable attention. Concerning Traditional Chinese Medicine (TCM), many researchers have explored the automated analysis of tongue images.

In the initial studies, researchers relied on traditional image processing and classical machine learning. To be more specific, Wang et al. [1] proposed a computerized tongue diagnosis system using SVMs along with color histogram features. The proposed approach obtained satisfactory results for tongue image classification. Nevertheless, the approach proved to be highly sensitive to changes in the environment, i.e., variations in the level of illumination and background noise.

Recently, the rise of deep learning has inspired researchers to utilize CNNs for the analysis of tongue images. Zhang et al. [2] proposed a deep CNN to extract high-level features from tongue images for medical diagnosis. The proposed approach obtained satisfactory results for the analysis of tongue images. Nevertheless, the CNN-based approach failed to consider the attention of the model towards the relevant areas of the tongue image.

Recently, researchers have proposed various machine learning-based models to improve the performance of the classifier. To be more specific, Li et al. [3] proposed a hybrid model using CNNs along with Random Forest classifiers for medical image analysis. The proposed approach obtained satisfactory results for the analysis of medical images. The proposed approach proved to be highly efficient for the analysis of medical images.

In the proposed approach, the attention of the model towards the relevant areas of the tongue image is considered. The proposed approach utilizes a combination of the attentionbasedResNet model along with the Random Forest classifier for the analysis of the tongue image.

III. PROBLEM STATEMENT

In general, traditional tongue diagnosis in TCM depends on visual observation by experienced practitioners. Even though this method has been proven over time, it is still subjective, based on practitioners' experience and their ability to observe visual clues. Therefore, it is possible that two different practitioners may come up with different conclusions when they examine the same patient.

Environmental factors also play an important role in confusing practitioners when they try to diagnose patients based on their tongues' images. These factors include different lighting conditions, quality of images, and position of patients when their tongues are photographed. All these factors make it more challenging for practitioners to diagnose patients manually. In other words, it is challenging for practitioners to accurately determine tongue color patterns.

Therefore, it is necessary to have an automated method that can diagnose patients' tongues and determine their colors objectively with minimal human intervention. A smart method that relies on computer vision can be very instrumental in improving consistency in initial diagnosis by practitioners.

IV. PROPOSED SYSTEM AND METHODOLOGY

Our proposed system will be an end-to-end AI system for the analysis of tongue images to identify the different patterns of tongue colors used in TCM. The system will be a combination of computer vision, deep learning for feature extraction, and machine learning for the final classification. The methodology will be a four-step process: data collection, image preprocessing, feature extraction using a deep neural network, and the final classification using ensemble learning.

A. Dataset Description

This study will be conducted using 1,068 tongue images collected from various public medical image datasets and TCM diagnostic image collections. The collected images will be classified into five different diagnostic categories based on the tongue color patterns used in TCM. The five categories will be Light Red, Dark Red, Purple, Red, and White, which correspond to different physiological states.

Prior to the commencement of the study, the images will be standardized to a fixed size. The images will be resized to a 32x32x3 matrix, where 32x32 corresponds to the spatial dimensions of the image and 3 corresponds to the three RGB values of each pixel. The pixel values will be normalized to the range [0,1].

For the purpose of fair evaluation, the data set was randomly shuffled and divided into a training data set and a testing data set. The data set for training consisted of 854 images (80

B. Image Preprocessing and Segmentation

Accurate extraction of the tongue region is vital for the determination of a reliable classification result. The images captured by cameras or uploaded by users may contain faces, backgrounds, or lighting issues, which could affect the result. Therefore, image preprocessing was applied.

Images are first converted from the BGR color space to the HSV color space. The HSV color space is more convenient for color-based segmentation.

A double-thresholding mask was applied for the extraction of the tongue region. The region of interest, the tongue, was extracted by maintaining the pixel values within a certain range of the HSV color space, corresponding to the color of the tongue. The unwanted regions are then removed. Morphological operations, including erosion and dilation, are applied for the purpose of eliminating noise from the extracted region. This binary mask is then used to filter the original image using bitwise operations. This results in a clean extraction of the tongue area.

C. Feature Extraction using Deep Learning

Finally, the segmented tongue area is passed through a deep convolutional neural network to extract useful visual features. Here, a Residual Network (ResNet) is employed to perform feature extraction. ResNet is a deep neural network used to overcome the problem of vanishing gradients. It uses residual connections to connect the convolutional layers.

Residual learning allows the network to learn the identity mapping using the shortcut connections. The residual learning equation is given by:

H(x) = F(x) + x

Here, H(x) is the output of the residual block, F(x) is the residual mapping, and x is the input feature map.

To improve feature discrimination, an attention module is used along with the ResNet. The attention module helps the network to focus on the informative parts of the tongue image. It helps the network to highlight the subtle changes in the color of the tongue.

The high-dimensional feature vector generated by the deep network represents the spatial and color characteristics of the image of the tongue.

D. Random Forest Classification

Once the feature extraction is completed, the feature vectors are passed through the Random Forest classifier for the final prediction.

A Random Forest classifier is an ensemble learning approach that uses multiple decision trees for the final prediction.

Each decision tree works independently and predicts the class label for the feature vectors. Finally, the prediction of the Random Forest classifier is made based on the majority voting results of all the trees. The prediction of the RF classifier can be defined as follows:

$$F(x) = \frac{1}{N}\sum_{i=1}^{N} T_i(x) \qquad (1)$$

where $F(x)$ is the final prediction of the Random Forest classifier and $T_i(x)$ represents the prediction produced by the i-th decision tree.

Random Forest is an effective classifier for medical image classification tasks, particularly for high-dimensional feature spaces, and reduces the problem of overfitting.

E. System Workflow

Here’s how the proposed system works from start to finish:

1) Image Capture/Upload: Tongue images are obtained either in real time using a camera or by uploading them through the web application interface.
2) Pre-processing: The acquired tongue images are resized and normalized to satisfy the input requirements of the neural network.
3) Segmentation: Tongue regions are separated from background elements using HSV color thresholding implemented with the OpenCV library.
4) Feature Extraction: The segmented tongue image is processed by an Attention-enhanced ResNet model to extract meaningful visual features.
5) Classification: The extracted features are classified into one of the predefined tongue color categories using a Random Forest classifier.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

6) Result Interpretation: The predicted class is mapped to a knowledge base that provides possible Traditional Chinese Medicine (TCM) health interpretations.

V. SYSTEM ARCHITECTURE

The system architecture of the proposed system is composed of three layers, each with a specific function to ensure modularity, scalability, and efficiency in processing tongue image data. These layers include the presentation layer, processing layer, and analytical layer. These layers are expected to work harmoniously, as shown in Figure 1, to provide a tongue color classification system.

A. Presentation Layer

This layer essentially focuses on how users can interact with the system. As such, this layer was implemented using standard web technologies, including HTML, CSS, and JavaScript. Users can either use their webcam to take pictures of their tongues in real-time or use the web application to upload existing tongue images. The frames captured are encoded in Base64 before being sent to the backend for further processing.

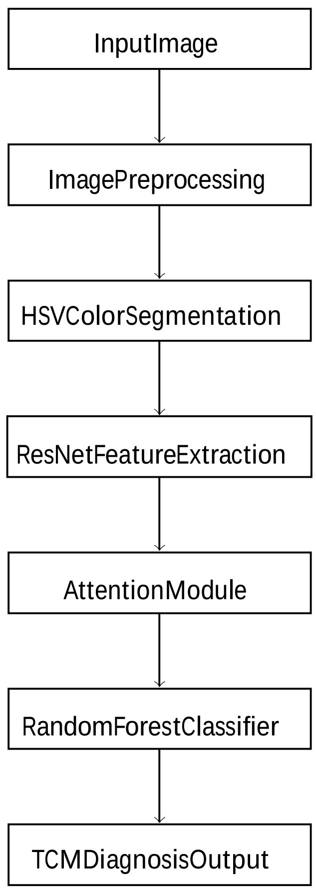


Fig. 1. Software architecture of the proposed AI-based tongue color classification system

B. Processing Layer

This layer essentially acts as a bridge between the user interface and the analytical components. As such, this layer was implemented using the Flask web application framework, which handles HTTP requests and image uploads. At this layer, necessary preprocessing of the tongue image data takes place before being sent to the analytical layer.

C. Analytical Layer

This layer essentially contains the main components of the system that perform necessary machine learning and image processing tasks. These include applying HSV thresholding using OpenCV to segment the tongue from the background. A deep learning model based on ResNet, with an attention mechanism, was implemented to extract features from the tongue.

These features are classified using a Random Forest classifier implemented using Scikit-learn. The system finally produces results based on tongue color classification, including Traditional Chinese Medicine interpretations.

VI. IMPLEMENTATION

We have implemented the tongue color classification system using popular data science and web development tools. We have used the Python 3.10 version of the language for this purpose, given the availability of rich resources for ML and computer vision.

For the web interface, we have used the Flask framework to handle the backend routes, HTTP requests, and interactions between the UI and the analysis modules. We have used Bootstrap for the frontend to create a user-friendly interface for the users. The users have the option to either click a photo of the tongue using the webcam or upload the image directly. For the image processing module, we have used OpenCV for the heavy lifting. We have used the HSV conversion of the images to segment the image using a binary mask. We have used the contour detection method to segment the tongue from the background.

For the feature extraction module, we have used TensorFlow along with Keras. We have used a ResNet model along with an attention module to extract the discriminative features from the segmented images of the tongues.

For the final prediction of the tongue color, we have used a Random Forest classifier. We have used the Scikit-learn library for this purpose. The models are serialized for easy deployment. We have used joblib for the Random Forest model, while the deep learning model is saved in the .h5 format.

VII. RESULTS AND DISCUSSION

In order to evaluate the effectiveness of the proposed tongue color classification system, we used another set of 214 images of tongues, which are completely different from the 854-image set used for the model. This design is intended to evaluate the proposed model against other models.

In order to evaluate the proposed model effectively, we have included two baseline models. The first model uses handcrafted color filtering, where the model uses a set of thresholds to filter the colors of the tongues. The second model uses a CNN-based deep learning model to learn the features of the tongues. The results of the proposed model compared to the baseline models are given in the table.

TABLE I
COMPARISON WITH EXISTING METHODS

Method	Accuracy	Processing Time (ms)	Resilience to Noise
Handcrafted Filters	71.3%	<50	Low
Pure CNN	89.4%	~120	Moderate
Proposed System	96.5%	~140	High

The handcrafted model achieved an accuracy of 71.3 The CNN-based model achieved an accuracy of 89.4
The proposed model achieved the highest accuracy of 96.5

As shown in Fig. 2, the proposed model achieves the highest accuracy compared to the other models. The bar chart shows the accuracy of the proposed model compared to the other models.

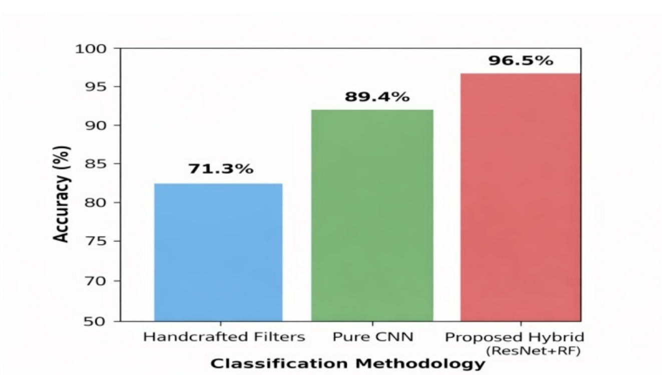


Fig. 2. Accuracy comparison of traditional handcrafted filters, pure CNN, and the proposed hybrid model.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

Besides accuracy, the proposed system was also tested for precision, recall, and F1 score, as shown in Fig. 3. The proposed system achieved 96.5

A high precision indicates that the model does not make many errors in the predicted tongue color, whereas a high recall indicates that the model correctly predicts relevant tongue color patterns. The F1 score also indicates the stability of the proposed framework.

Therefore, the experiments demonstrate that the proposed AI-based hybrid model significantly improves the accuracy of the automated diagnosis of tongue color in comparison to traditional methods and the CNN-based method.

VIII. ADVANTAGES OF THE PROPOSED SYSTEM

The proposed tongue color classification system has some clear advantages over the conventional methods of diagnosis.

Fig. 3. Performance metrics of the proposed tongue color classification system.

Firstly, the system improves the objectivity and consistency of the diagnosis of the tongue by utilizing machine learning to analyze the image of the tongue, which reduces the variability of the diagnosis depending on the judgment of the individual and the tiredness of the individual.

Secondly, the proposed system presents a cost-effective approach to initial medical screenings. The system does not require specialized equipment for imaging, as it utilizes the usual imaging equipment such as webcams and cameras, which are readily available and easily accessible, especially in remote locations.

Lastly, the framework combines the predictions of the algorithm with some diagnostic outputs to help improve the decisions made by the individual. The system improves the accuracy and reliability of the initial diagnosis of the color of the tongue in Traditional Chinese Medicine.

IX. CONCLUSION

In this study, the automated tongue color classification system using artificial intelligence in the form of a combination of computer vision and deep learning was proposed as a diagnostic tool in the practice of Traditional Chinese Medicine. The model combines the use of a deep learning model and a

Random Forest algorithm in the classification of tongue colors.

As shown in the experimental results, the proposed model achieved a classification accuracy of 96.5

X. FUTURE WORK

The current system demonstrates great potential in the classification of tongue colors. However, there is always room to grow and improve. A natural extension of the current work is the development of a mobile app that can carry out the analysis in real time using the smartphone camera.

Another direction that can be taken is the expansion of the diagnostic features beyond the colors alone. For example, the model can be trained to look at the texture of the tongue, such as the presence of teeth marks, fissures, and the coating on the tongue. This will provide a more complete diagnostic picture.

Another improvement that can be made is the integration of external health data to improve the accuracy of the predictions. For example, the analysis of the tongue images can be combined with geographic information to look at the correlations between the state of the body and the environment.

REFERENCES

[1] Z. Zhihua, J. Yan, C. Qinglin, X. Jialin, Z. Yue, and W. Jian, ”TongueColor Classification Based on Deep Convolutional Neural Network,” in Proc. IEEE Int. Conf. Bioinformatics and Biomedicine (BIBM), 2019, pp. 2289–2294.

[2] Y. Xu et al., ”A Novel Classification Method of Tongue Images Based on Deep Learning,” IEEE Access, vol. 9, pp. 49666–49675, 2021.

[3] X. Wang, C. Zhang, and L. Zhu, ”Computerized Tongue Diagnosis Based on SVM,” Journal of Medical Systems, vol. 41, no. 5, pp. 210– 215, 2017.

[4] B. Zhang, H. Wang, Z. Wang, and S. Hao, ”Computer-Aided Tongue Diagnosis System Based on Computer Vision,” in Proc. IEEE Int. Symp. Computer, Consumer and Control (IS3C), 2020, pp. 272–275.

[5] L. Hu, M. Chen, K. Zheng, and B. Huang, ”Hybrid CNN-RF Models for Medical Image Classification,” IEEE Trans. Medical Imaging, vol. 40, no. 2, pp. 433–444, 2021.

[6] J. Zhang, T. Huang, and Y. Hu, ”Deep Learning Applied to Tongue Diagnosis,” in Proc. IEEE Int. Conf. Bioinformatics and Biomedicine (BIBM), 2019, pp. 1102–1108.

[7] X. Cui et al., ”Application of ResNet in Tongue Image Classification,” in Proc. IEEE 4th Int. Conf. Information Technology, Networking, Electronic and Automation Control (ITNEC), 2020, pp. 1438–1442.

[8] Y. Zhao, W. Xu, and Y. Liu, ”Computer-Assisted Tongue Diagnosis System Based on an Integrated Deep Learning Architecture,” Artificial Intelligence in Medicine, vol. 114, p. 102035, 2021.

[9] W. Ma, Z. Li, C. Li, Y. Zhang, and Y. Shi, ”Automated Tongue Image Segmentation Based on Deep Convolutional Neural Networks,” IEEE Access, vol. 8, pp. 154247–154256, 2020.

[10] S. Huang, G. Jiang, and D. Li, ”An Automated Deep Learning Framework for Tongue Color and Texture Recognition in Traditional Chinese Medicine,” Computational and Mathematical Methods in Medicine, vol. 2020, pp. 1–12, 2020.

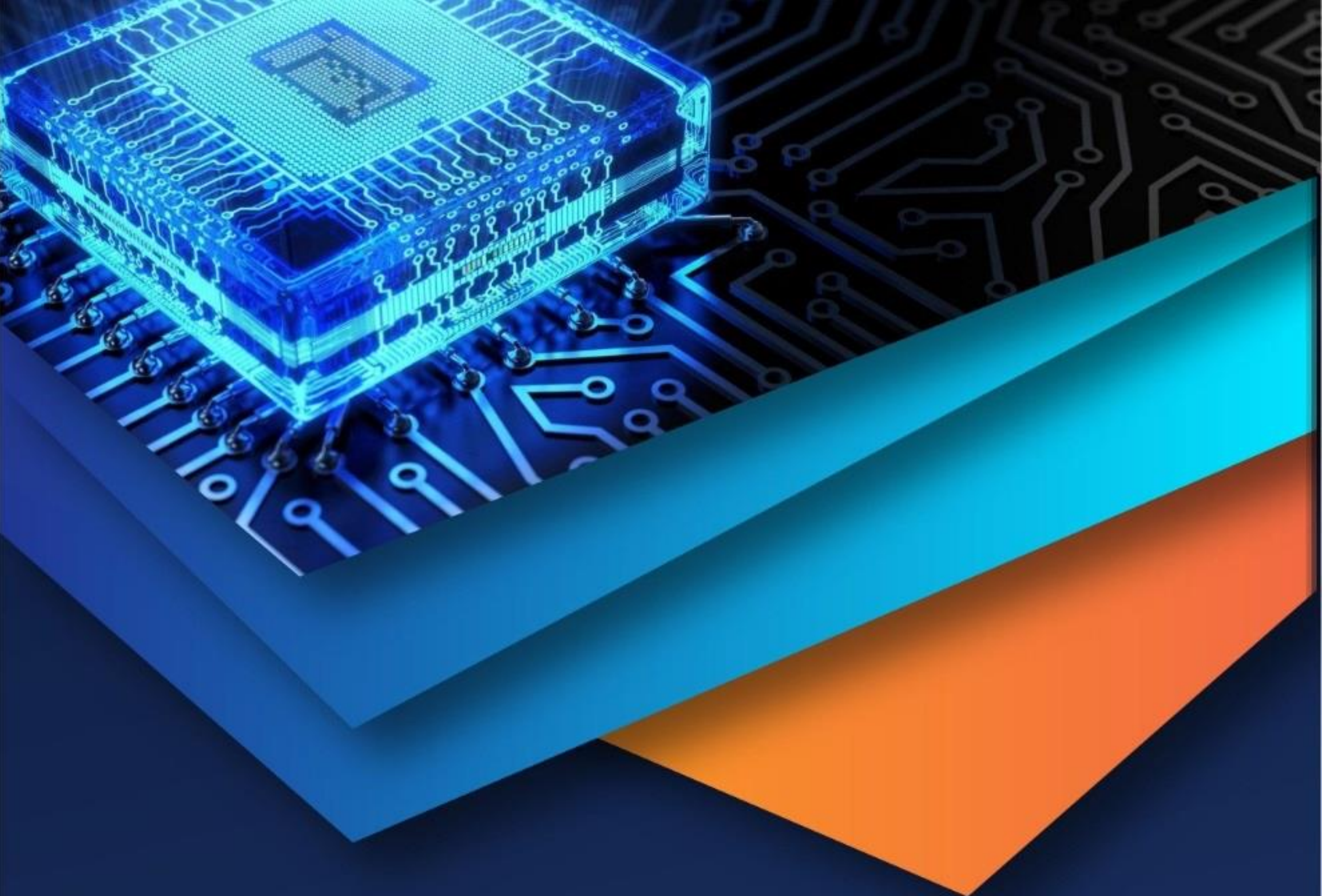
[11] K. He, X. Zhang, S. Ren, and J. Sun, ”Deep Residual Learning for Image Recognition,” in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770–778.

[12] L. Breiman, ”Random Forests,” Machine Learning, vol. 45, no. 1, pp. 5–32, 2001.

[13] J. Hu, L. Shen, and G. Sun, ”Squeeze-and-Excitation Networks,” in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2018, pp. 7132–7141.

[14] G. Bradski, ”The OpenCV Library,” Dr. Dobb’s Journal of Software Tools, 2000.

[15] T. Ojala, M. Pietikainen, and T. Maenpaa, ”MultiresolutionGrayScale and Rotation Invariant Texture Classification with Local Binary Patterns,” IEEE Trans. Pattern Analysis and Machine Intelligence, vol.24, no. 7, pp. 971–987, 2002.



10.22214/IJRASET

INDEX COPERNICUS

45.98

IMPACT FACTOR: 7.129

TOGETHER WE REACH THE GOAL

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)