Proceedings of the Third International Conference on Smart Systems and Inventive Technology (ICSSIT 2020)
IEEE Xplore Part Number: CFP20P17-ART; ISBN: 978-1-7281-5821-1

Multiclass Support Vector Machine based Plant Leaf Diseases Identification from Color, Texture and Shape Features

D. Anil Kumar, P. Sudheer Chakravarthi, K. Suresh Babu
Department of ECE, PACE Institute of Technology and Sciences, Ongole India
danilmurali@gmail.com, sudheerece2010@gmail.com, sureshbabu_k@pace.ac.in

Abstract:

In general, Indian economy highly depends on agricultural productivity. In the agricultural field, the identification and classification of leaf diseases play an important role. In developing countries, physical observation of plant leaf diseases can be prohibitively expensive due to the naked eye observation. The proposed research work has developed a framework to identify and classify different plant leaf diseases using K-means segmentation with a multiclass support vector machine (SVM) based classification. The proposed framework is implemented in four steps, step I performs the RGB to HSI colour transformation. In step-II, image segmentation using K-means clustering is performed. Next, colour, texture and shape features are extracted in step III. Finally, in step-IV, multiclass SVM is used for the extracted feature classification. Experimental results indicate that the proposed approach results in an improved detection and classification compared to other existing methods. Efficiency of the proposed algorithm recognizes the accuracy of leaf diseases at about 95.7%.

Keywords: Plant leaf diseases, Hue saturation intensity, K-means segmentation, colour feature, texture feature, shape feature, multiclass support vector machine.

1. Introduction

The economy of India was mainly depending on agricultural, throughout this area, plant leaf disease identification and classification/recognition that plays an important role. The quantity, quality or production of the respective agricultural field mainly depends on the proper care, without any care where it can have a large effect on plants. Over the last few decades, experts have witnessed the detection and classification of plant leaf diseases with naked eyes, this process is very slow and difficult in continuous monitoring. It is of a very high cost, when monitoring large farms and very less accuracy. Around the same time, few countries have no sufficient facilities and exports for detecting diseases. By these regions to solve these problems, an algorithm is proposed to remain helpful for continuously monitoring the large fields of crops. The proposed automatic system is cheaper and very easy to identify diseases by observing the symptoms of plant leaves. This system mainly supports computer vision machines to give automatic image-based process, observation and guidance [1, 2].

The detection of leaf diseases by visual monitoring requires more human power with less accuracy, it can be done in a very limited field. To solve these problems, a program for automatically identifying and classifying leaf diseases must be built. The proposed system used to reduce the expenses, time, efforts and give more accurate than manual monitoring. Nowadays, we observed that some general diseases like yellow and brown sports, colour variations in early and late growth as well as some other diseases are infectious, fungal and bacterial as seen in Figure 1. Recently, some of the researchers doing image processing-based disease identification and compare the colour difference for the observed region [1, 3].

Such automated systems with image processing techniques are used to solve problems in visual analysis, but the robustness and accuracy outcomes faced the major challenges in these systems. Within this work we develop an automated framework using the technique of image processing as well as learning-based methods for recognizing and classifying various plant leaf diseases.We propose this work a novel system for the detection and classification of diseases using image processing methods. First, we convert the RGB images into space converting HSI color. Furthermore, the images are segmented using the clustering process K-means. Thirdly, the color, form and texture features are extracted from the segmented image and finally the extracted features are evaluated using SVM multi-classto classify the defected plant leaves. The output of the proposed method tested with five specific leaf diseases such as Alternaria Alternate, Anthracnose, Bacterial Blight, Cercospora Leaf Spot, Mosaic and we tested stable leaves as well. Our proposed system received high level of acceptance over other state of the art approaches.

Figure.1. Example of plant leaf disease images.

The reaming paper is planned as follows; the brief related work of other existing algorithms is discussed in section 2. The detail discerption of the four-level methodology is given in section 3 in details: Transforming RGB to HSI, segmenting, removing features and eventually the classification system for diseases. Section 4 presents the findings and comments, as well as comparison of other state-of-the-art algorithms with our proposed algorithm and conclusions in section 5.

2. Literature review

The main purpose of this section is to demonstrate the potential of some commonly used image processing techniques that can manage the relevant goals efficiently. A comparative study of different image processing methods that are used for automatic plant leaf disease detection systems is also reviewed. For classification of leaf diseases, we discussed some learning techniques. In [4], the authors present a review of the different plant leaf disease classification by using different classification algorithms. In this paper, K-NN algorithm can detect the plant leaf diseases with good recognition rates when compared to other techniques. The authors in [1] detect the leaf diseases by using four image processing technique like colour transformation, threshold-based green pixels masking and segmentation, colour co-occurrence-based feature extraction and finally analyze the statistical features to detect the diseases. Babu et al. [5] designed and developed a backpropagation neural network to recognize the diseases of leaves. The gradient descent value was calculated with a neural network and back propagation method is used to minimize the error function.

The K-means clustering algorithm is one of the simplest unsupervised methods and is used for clustering analysis in recent image segmentation survey [6]. The main purpose of this method is to minimize the data given into K-subsets. The authors in [7] suggested an automated leaf disease detection method by using a neural networked K-means algorithm.This method detects and classifies diseases with about 93% precision. Mohammad Reza Larijani et al. suggested K-means clustering with the classifier K-NN for the identification of leaf diseases [8].

SVM is a popular machine learning and classification algorithm is claimed by [9]. It has many advantages like good recognition rates, robustness and fast speed etc. the authors in [10, 11] used SVM algorithm for solving detecting and recognition methods. Likewise, the authors in [12] proposed a combination of a simple linear iterative cluster with SVM based plant leaf identification is proposed. The overall recognition rate of the model is around 98.5%. T. Rumpf et al. [13] built an early detection and classification of sugar leaf diseases using SVM with approximately 97% accuracy. The authors in [14] proposed the identification and quantification of verticillium wilt in the olive method by using discriminant classifiers for SVM and Linear analysis.The number of structures proposed is respectively 79.2% and 75%. The combination of K-means clustering and SVM is a novel method for detecting and classifying different plant leaves, based on the above research.

In this work, we proposed a novel plant leaf disease identification and classification system is designed by using K-means and SVM algorithms. In this work, we test our proposed algorithms compared with K-nearest neighbour Classifier (KNN) [4], dynamic time wrapping (DTW) [15], decision tree [16], minimum distance criterion (MDC) [17] classification methods. In the next section, we discuss the different steps used in the structure of our methodology.

3. Methodology

The proposed automatic plant leaf detection system is illustrated as shown in figure 2. In this section, the conversion of RGB to HSI colour space, K-means clustering, features extraction and classification of different diseases are given in sections 3.1, 3.2, 3.3 and 3.4 respectively.

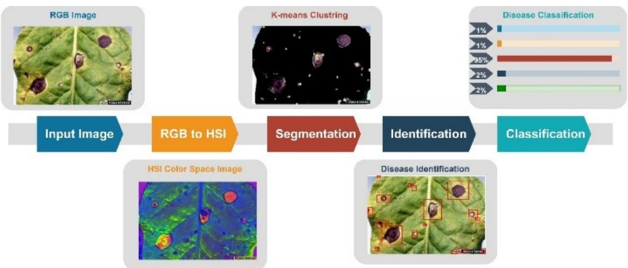


Figure.2. Flowchart of the proposed plant leaf detection procedure

3.1.Colour space transformation

Color is one of the essential features of image processing techniques for use in the image. -- pixel in the image will give in the RGB color space with a combination of the three color components Red (R), Green (G), and Blue (B). Figure 3 displays the effects of plant leaf disease thresholding using RGB images in colour.

(a) Original (RGB) Image (b) Detected plant leaf disease

Figure.3. Plant Leaf Disease Detection using RGB Colour Space.

From figure 3, it is observed that by using colour RGB images, the detection of plant leaf diseases is very difficult and high misclassification rates. So, the process is switched to HSI (hue, saturation, intensity) colour model in RGB. In HSI colour space in each pixel of the image and is divided into the components of Hue (H), Saturation (S) and Intensity (I). Colour conversion RGB-HIS [18] is calculated as

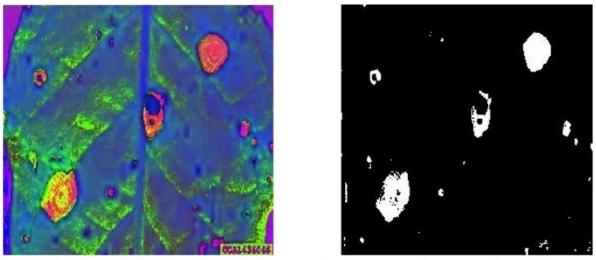
$$H=\begin{cases}\theta & \text{if } \mathrm{b}\le g\\ 360-\theta & \text{if } \mathrm{b}>g\end{cases}$$

$$\theta=\cos^{-1}\left\{\frac{\frac{1}{2}[(r-g)+(r-b)]}{\sqrt[2]{(r-g)^2+(r-b)(g-b)}}\right\} \quad (1)$$

$$S=1-\frac{3}{(r+g+b)}\times\left(\min(r,g,b)\right) \quad (2)$$

$$I=\frac{(r+g+b)}{3} \quad (3)$$

Here, θ is the angle of the Hue (H). From these equations, the saturation and intensity values will change by the effect of shifting and scaling operations, but the hue is kept constant through a hue preservation principle at the same time. Figure 4 shows the HSI binary images that obtain plant leaf diseases region.



(a) HSI color space Image (b) Detected plant leaf disease

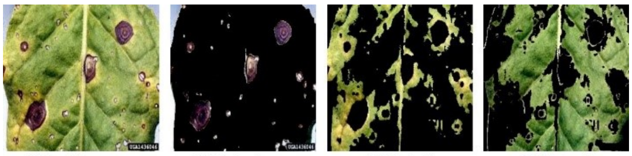
Figure.4. Plant Leaf Diseases Detection using HSI Colour Space.

In figure 3 and 4, the HSI colour space effetely detects the pant leaf diseases compared to RGB.

3.2. Segmentation

In this work, K-means clustering [19, 20] techniques are used for segmentation, it is a classical separation-based clustering method. In general, K-means is a technique of clustering that aims to detect the object nearest to them by separating the points in the space. Further, the proposed K-means algorithm was used to detect the plant leaf diseases. In this work, Euclidean distance is used for finding the similarity values of cluster centroids, such values shall be modified one by one before the best results of the clustering. For finding the clustering measure, the sum of square error criterion function was used. Although, the proposed method is efficient for segmenting the plant leaf diseases.

Let $I(x,y)$ is the input image with the resolution $x \times y$, to be cluster into k number of clusters and c_k be the centre cluster. The process of proposed algorithm as follows, the first number of cluster k and centreis initializing. Second, we calculate the Euclidean distance between each pixel of an image and the cluster center. Third, to find the nearest center pixels by using $d = \left\| I(x,y) - c_k \right\|$. Fourth, recalculate new centroid location using $c_k = \frac{1}{k} \sum_{y \in c_k} \sum_{x \in c_k} I(x,y)$. Finally, repeat this process to satisfy the sum of the square error function and reshape the cluster pixels into images. The K-means clustering experimental findings for a plant leaf disease, as shown in Figure 5.



(a) Input Image (b) Cluster-1 (c) Cluster-2 (d) Cluster-3

Figure.5. Output results of K-means Clustering for Plant Leaf Disease.

3.3. Extraction of features

In this section, we briefly explain the process of feature extraction techniques. In this work, we extract total 9 features including texture and colour features like energy, contrast, correlation, entropy, homogeneity, shape, prominence, mean and variance. Those mathematical equations are written as

$$Energy = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \left(P_{i,j} \right)^2 \quad (4)$$

$$Contract = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} P_{i,j} \left(i - j \right)^2 \quad (5)$$

$$Correlation = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} P_{i,j} \frac{(i-\mu)(j-\mu)}{\sigma^2} \quad (6)$$

$$Entropy = -\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} P_{i,j} \log\left(P_{i,j} \right) \quad (7)$$

$$Homogeneity = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \frac{P_{i,j}}{1 + \left(i - j \right)^2} \quad (8)$$

$$Shape = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \left(i - I_x + j - I_y \right)^3 P_{i,j} \quad (9)$$

$$\text{Prominence} = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \left(i - I_x + j - I_y \right)^4 P_{i,j} \quad (10)$$

$$mean\ (\mu) = \sum_{i=0}^{N-1}\sum_{j=0}^{N-1} iP_{i,j} \tag{11}$$

$$\text{var}\,ience\ (\sigma^2) = \sum_{i=0}^{N-1}\sum_{j=0}^{N-1} (i-\mu)^2 P_{i,j} \tag{13}$$

Here, $P_{i,j}$ is the co-occurrence matrix, P is the position operator of elements i, j. Where i represent the (x, y) location of the grey level and j represents the Euclidean distance of the grey level pixels. N represents the dimension of the co-occurrence matrix.

3.4.Diseases classification

SVM [21] is a supervised learning methodology evaluating data collection, used for recognition and classification analysis. In this work, we used multi SVM for classifying the 5 plant leaf diseases. Many researchers have used one vs one (OVO) [22], one vs all (OVA) [22], error-correcting performance codes (ECOC)[23], and directed acyclic graph (DAG)[24] throughout the last years.. OVO was used the pairwise SVMs $0.5N(N-1)$ and opting in pairs to help new samples overcome multi-class problems. For N number of categories, OVA creates N binary SVMs, where N is the number of classes. By using in this method, the computation time is very less but the efficiency was compromised. ECOC construct a code matrix by using output binary codes. Where every row contains a code, word linked to every class, and columns defines a binary pattern of classes. This method gives good recognition rates. DAG algorithm trained and tested from OVO and binary acyclic graph model. By the analysis of these models, the ECOC algorithm gives good classification rates and speed when compared to other algorithms. For implementing multi SVM in MATLAB, we observed ECOC gives better recognition rates and optimum speeds.

The finding plant leaf diseases using supervised learning and computer vision technique is proposed. The proposed ECOC multiclass SVM detect the plant leaf diseases better compared to other methods. The percentage of success of disease recognition is measured by the following equation as

$$\%\ of\ \text{Recognition} = \frac{number\ of\ correct\ classification}{Total\ no\ of\ test\ images} \times 100 \tag{14}$$

4. Results and discussion

In this study , we tested our proposed algorithm on five different plant leaf diseases including Alternaria Alternate, Anthracnose, Bacterial Blight, Cercospora Leaf Spot and images of mosaic disease are captured in Andhra Pradesh forin this work. In this work, to validate the proposed algorithm that uses three different parameters like precision, recall and % of recognition rate [25].

Figure 6 shows the detection results of the plant deaf disease by using our proposed method. Figure 6(a) represents the input image, Figure 6(b) shows the outcome of the clustering and Figure 6(c) shows the identification of the disease of the leaf.

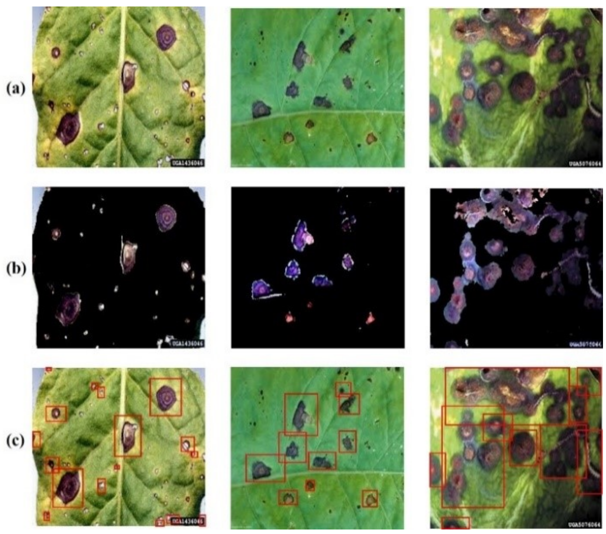


Figure.6. Detected Diseases Identification of Different Plant Leaves.

After detecting the diseases of plant leaves, we extract the library of 9 different colors and texture features. Plant leaf disease classification using multiclass SVM system. Therefore, to check the efficiency of the algorithm proposed compared to other classifierslike KNN [4], DTW [15], decision tree [16], MDC [17]. The comparison of various classifiers with the proposed classifier on plant leaf disease detection as shown in figure 7. From this precision-recall curve shows that the proposed SVM gives good detections compared to other classifiers.

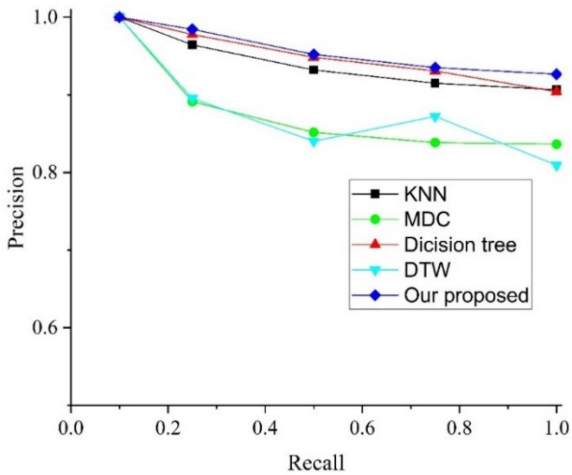


Figure.7. Comparison of Different Classifiers for Leaf Disease Classification

The classified five plant leaf diseases by using our proposed algorithm with SVM as shown in figure 8. As shown in this confusion matrix, can be seen that very few samples are misclassified because of early detection. Figure 9 demonstrates our proposed algorithm 's efficiency as opposed to other methods.Comparing our proposed algorithm 's performance results in terms of precision-recall and recognition rate, provides good results.The efficiency of our proposed algorithm gives approximately 95.7% accuracy.The proposed algorithm with SVM classifier provides good performance on the identification and classification of plant leaf disease, as shown in these results.

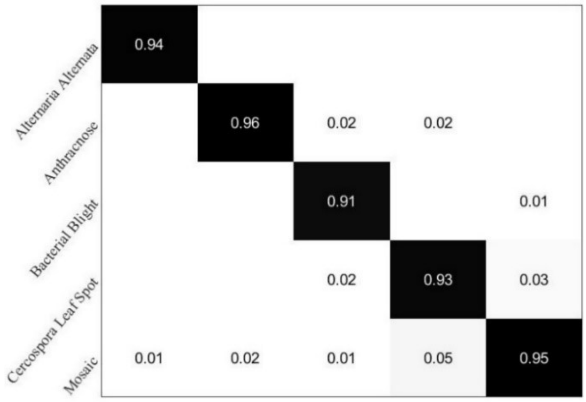


Figure.8. Confusion Matrix for Leaf Disease Classification using Proposed Algorithm

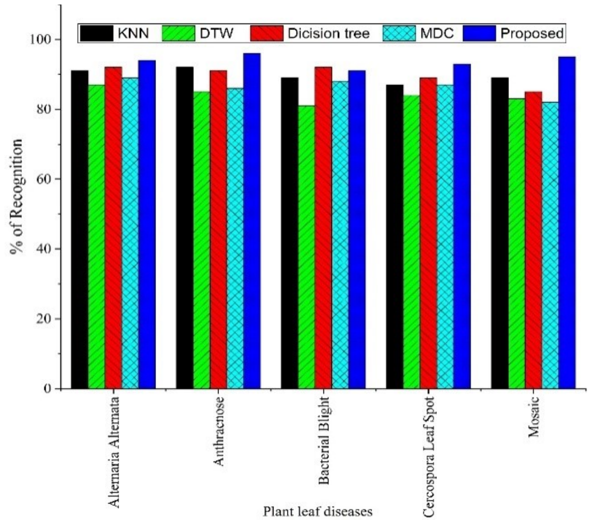


Figure.9. Comparison Results of Different Methods with our Proposed Method.

5. Conclusion

In this article, using image processing techniques, a novel automated identification and classification method is used for plant leaf diseases. A four-step based disease identification system is proposed to handle the different problems. The proposed system, we first convert RGB to HSI colour space, K-means based disease clustering, next extract colour, texture and shape features of segmented output, finally, leaf diseases are classified by using multiclass SVM method. To test the algorithm proposed as opposed to other classifiers. The average recognition of our multiclass SVM is around 95.7%, which is better than the other algorithms used to identify and recognize the plant-leaf diseases.

References

1. Dhaygude, S. B.,&Kumbhar, N. P. (2013). Agricultural plant leaf disease detection using image processing. *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering*, *2*(1), 599-602.
2. Arivazhagan, S., Shebiah, R. N., Ananthi, S., &Varthini, S. V. (2013). Detection of unhealthy region of plant leaves and classification of plant leaf diseases using texture features. *Agricultural Engineering International: CIGR Journal*, *15*(1), 211-217.
3. Bashir, S., & Sharma, N. (2012). Remote area plant disease detection using image processing. *IOSR Journal of Electronics and Communication Engineering*, *2*(6), 31-34.
4. Ghaiw at, S. N., & Arora, P. (2014). Detection and classification of plant leaf diseases using image processing techniques: a review . *International Journal of Recent Advances in Engineering & Technology*, *2*(3), 1-7.
5. Babu, M. P., & Rao, B. S. (2007). Leaves recognition using back propagation neural netw ork-advice for pest and disease control on crops. *IndiaKisan. Net: Expert Advisory System*.
6. Khatoon, R., & Das, A. (2019, January). A Survey on Image Segmentation Techniques. In *Proceedings of conference on Advancement in Computation, Communication and Electronics Paradigm (ACCEP-2019)* (p. 2).
7. Al Bashish, D., Braik, M., & Bani-Ahmad, S. (2011). Detection and classification of leaf diseases using K-means-based segmentation and. *Information Technology Journal*, *10*(2), 267-275.
8. Larijani, M. R., Asli-Ardeh, E. A., Kozegar, E., & Loni, R. (2019). Evaluation of image processing technique in identifying rice blast disease in field conditions based on KNN algorithm improvement by K-means. *Food Science & Nutrition*.
9. Kamavisdar, P., Saluja, S., & Agraw al, S. (2013). A survey on image classification approaches and techniques. *International Journal of Advanced Research in Computer and Communication Engineering*, *2*(1), 1005-1009.
10. Kumar, K. V. V., & Kishore, P. V. V. (2017). Indian Classical Dance Mudra Classification Using HOG Features and SVM Classifier. *International Journal of Electrical & Computer Engineering (2088-8708)*, *7*(5).
11. Al Farid, F., Hashim, N., & Abdullah, J. (2019, March). Vision-based hand gesture recognition from RGB video data using SVM. In *International Workshop on Advanced Image Technology (IWAIT) 2019* (Vol. 11049, p. 110491E). International Society for Optics and Photonics.
12. Sun, Y., Jiang, Z., Zhang, L., Dong, W., & Rao, Y. (2019). SLIC_SVM based leaf diseases saliency map extraction of tea plant. *Computers and electronics in agriculture*, *157*, 102-109.
13. Rumpf, T., Mahlein, A. K., Steiner, U., Oerke, E. C., Dehne, H. W., &Plümer, L. (2010). Early detection and classification of plant diseases w ith support vector machines based on hyperspectral reflectance. *Computers and electronics in agriculture*, *74*(1), 91-99.
14. Calderón, R., Navas-Cortés, J., &Zarco-Tejada, P. (2015). Early detection and quantification of Verticillium w ilt in olive using hyperspectral and thermal imagery over large areas. *Remote Sensing*, *7*(5), 5584-5610.
15. Romani, L. A., Goncalves, R. R. V., Zullo, J., Traina, C., &Traina, A. J. (2010, July). New DTW-based method to similarity search in sugar cane regions represented by climate and remote sensing time series. In *2010 IEEE International Geoscience and Remote Sensing Symposium* (pp. 355-358). IEEE.
16. Fan, L., Wang, Z., & Wang, H. (2013, December). Human activity recognition model based on decision tree. In *2013 International Conference on Advanced Cloud and Big Data* (pp. 64-68). IEEE.
17. Singh, V., &Misra, A. K. (2017). Detection of plant leaf diseases using image segmentation and soft computing techniques. Information processing in Agriculture, 4(1), 41-49.
18. Ford, A., & Roberts, A. (1998). Colour space conversions. *Westminster University, London*, *1998*, 1-31.
19. Zheng, X., Lei, Q., Yao, R., Gong, Y., & Yin, Q. (2018). Image segmentation based on adaptive K-means algorithm. *EURASIP Journal on Image and Video Processing*, *2018*(1), 68.
20. Dhanachandra, N., Manglem, K., &Chanu, Y. J. (2015). Image segmentation using K-means clustering algorithm and subtractive clustering algorithm. *Procedia Computer Science*, *54*, 764-771.
21. Sastry, A. S. C. S., Kishore, P. V. V., Kumar, D. A., & Kumar, E. K. (2018). Sign Language Conversion Tool (SLCTooL) Betw een 30 World Sign Languages. In *Smart Computing and Informatics* (pp. 701-711). Springer, Singapore.
22. Galar, M., Fernández, A., Barrenechea, E., Bustince, H., & Herrera, F. (2011). An overview of ensemble methods for binary classifiers in multi-class problems: Experimental study on one-vs-one and one-vs-all schemes. *Pattern Recognition*, *44*(8), 1761-1776.
23. Bai, X., Niw as, S. I., Lin, W., Ju, B. F., Kw oh, C. K., Wang, L., ... & Chew , P. T. (2016). Learning ECOC code matrix for multiclass classification w ith application to glaucoma diagnosis. *Journal of medical systems*, *40*(4), 78.
24. Zhang, X., Ding, S., & Sun, T. (2016). Multi-class LSTMSVM based on optimal directed acyclic graph and shuffled frog leaping algorithm. *International Journal of Machine Learning and Cybernetics*, *7*(2), 241-251.
25. Kumar, D. A., Sastry, A. S. C. S., Kishore, P. V. V., & Kumar, E. K. (2018). 3D sign language recognition using spatio temporal graph kernels. *Journal of King Saud University-Computer and Information Sciences*.