Neutrosophic Approach of Segmentation on Thermal Images - Case Study: Drowsy Driving Application

Sofia Jennifer J[1*], Sree Sharmila T[2]

Abstract

Image segmentation is a critical and vital component in many application domains. The proposed work focuses on performing segmentation in thermal images using neutrosophic idea of truth, indeterminacy, and falsity membership. The computations of these memberships are then pre-processed using α-mean and β-enhancement to minimize the indeterminacy. Then, as a part of the convergence criterion, the gamma clustering is used to define segmented parts. Thermal imaging is a visual representation of different colour regions based on thermo gram. These images are not affected by variation in illumination is commonly used because it uses the intrinsic emissivity of thermal radiation from the human body. Drowsy driving application uses this significant advantage as both hardware and software components need to adapt both high intensity of daylight and zero intensity of night light situations. Drowsiness factors such as yawn and eye-blink are both considered for efficient performances outcomes. On experimental analysis of the proposed methodology provides an unbiased result with the drowsy detection accuracy of 97.01%.

Key Words: Blink detection, Clustering, Drowsy driving application, Neutrosophic Sets, Thermal Images, Yawn detection

DOI Number: 10.14704/nq.2022.20.8.NQ44244 **NeuroQuantology 2022; 20(8):2217-2229**

Introduction

Image segmentation paves way for applications focusing on object recognition, pattern recognition, medical imaging and robot vision etc in a image-processing domain. In the image segmentation approach the foreground is individually separated from background or any other object background. On the onset of image segmentation improved fuzzy logic is used for differentiation. Many researches have concluded that the fuzzy theory on image segmentation holds more significant information than traditional methods [32].

Neutrosophy is the study of origin, nature, and scope of neutralities [33]. This is an extended branch of fuzzy logic or set and is generalized as neutrosophic set (NS) similarity [33]. The NS further enhances the independent uncertainty measures by generalizing into truth, indeterminacy, and falsity membership. This approach takes one more step forward towards segmentation approach by retaining more persistent information. Few notable researches are being focused on segmentation of visible images using neutrosophic set.

Segmentation on thermal images is a challenging task as the thermal camera captures the absolute temperature values of the objects and their geometric outlines and edge features are more blurred than visible images. On using the advantage of acquiring more persistent information using neutrosophic set, the thermal images are segmented from the background.

For study and analysis a drowsy driving application is taken into consideration. In this application, the driver's vigilance level is considered to be the major reasons of traffic accidents and thus posting a serious issue of safety for the society.

Corresponding author: Sofia Jennifer J
Address: [1,2]SSN College of Engineering, Kalavakkam, Chennai 603110, India
E-mail: [1]sofiajenniferj@ssn.edu.in, [2]sreesharmilat@ssn.edu.in

On proving this statement a statistical report shows the traffic accidents due to driver's diminished vigilance level is about 20% [1]. The suggested solution focuses on prevention of accidents, by observing the driver's vigilance level continuously and alert him with alarm when drowsy [2, 3]. To achieve this, segmentation of the objects such as the eyes and the mouth state is done to indicate the drowsiness factors.

Related Works

There is numerous research works based on image segmentation approaches. For gray-scale image segmentation, follows either of the two approaches, such as homogeneity or discontinuity. The concept of homogeneity is used in methods such as clustering, threshold, region growing and merging whereas the idea of discontinuity is used for object edges and corners.

Then segmentation approaches uses mathematical fuzzy theory as it retains more detailed information than many other existing approaches as it handles uncertainty values. The Fuzzy c-means (FCM), clusters into individual components based on the cluster size and intensity values for segmentation results.

In a traditional fuzzy set, the membership of the set A defined on universe x $\mu_A(x) \in [0, 1]$. If $\mu_A(x)$ can hold values from 0 to 1 and based on the values it decides it is object or background in an image [11]. There are cases such as uncertainty where defining them in a crisp value is a difficult process. To overcome this issue, a Neutrosophic set (NS) are introduced as a new branch that extends fuzzy logic and deals with neutralities. The spatial images are converted into a set with degree to truth, falsity, and indeterminacy membership. These retrieved indeterminacy membership information is used for segmentation approaches such as color and texture segmentation [4 -7], clustering and thresholding. Y. Guo [12] proposed a list of fuzzy C-means clustering algorithms, the interval and neutrosophic fuzzy clustering algorithm (NCM) for segmentation.

A very few researches [14-17] focus on segmentation of thermal images and mostly on tumor detection in breasts. To isolate the region of interest is a challenging task as the thermal images does not hold a specific shape. Wang [31] uses global variance using Otsu's threshold, R, M et al. uses K-means technique [17] and [16] uses projection profile. No new research is done on using segmentation of thermal images using neutrosophic set.

Our proposed work concentrates on using NS based segmentation of thermal images for a case study of Drowsy driving application. Many researches have been done on drowsy driving such as Khan et al. [15] proposes an automatic driver fatigue recognition where the facial expressions are taken into consideration, Wang et al. [8] proposes uses SVM classifiers etc. To adapt to the different illumination scenarios of both day and night time thermal images found to be the suitable solution. The captured thermal images comprises of only the absolute temperature and lacks the generic geometrical and appearance features making segmentation a difficult task. Researchers such as Strakowska et al. [9] proposes an idea of using eye corners to track thermal face and Marzec et al. [13] suggests an eye localization based on CNN technology in thermal images.

Thermal images do not have shape and specific limits

and therefore, ROI extraction from these types of images is challenging Thermal images do not have shape and specific limits and therefore, ROI extraction from these types of images is challenging Our proposed work uses the advantage of using the indeterminacy membership values of NS in a thermal image to segment the eyes and mouth state such as eye blink and yawning to determine the drowsiness in an efficient way.

For segmentation of breast thermal images S. Motta et al. [2] applied global variance by Otsu's threshold to subtract the background with the final aim of segmentation of both breasts as symmetrically as possible and were satisfied with the results of their study. With the aim of mass detection in thermal images, R., M. et al. [3] applied thresholding based methods like grey threshold, RGB technique, and K-means technique and pointed out specifically that the RGB technique suffered from the limitation of data loss. Dayakshini, D. et al. [4] proposed a novel method of breast segmentation called Projection Profile Analysis for identifying the boundaries in the breast thermograms. For edge detection of the breast, they successfully used the Sobel operator while adaptive thresholding was applied for changing the ROI's color to white and reported that the overall results of their approach were satisfactory. For processing breast images by automated segmentation, De Oliveira et al. [5] applied various image processing steps like threshold, clustering, identification and improvement of corners on 328 lateral thermal

images of breasts. A thresholding method was applied by them to separate the background and patient body in two modes. 96% and 97% of mean values of accuracy and sensitivity were obtained by them. Duarte A. et al. [1] developed an application using Matlab for characterizing thermal images that facilitates selection of any ROI without considering its geometry and provides options for its further optimizations. However, the thresholding algorithm implemented by themis not specified. Barcelos E. et al. [6] performed a study on 28 subjects who were athletes for the purpose of preventing muscle injury. They proposed automated recognition of flexible ROIs that have shape of body by optimizing the Otsu's method of segmentation and further performing progressive analyses of thermograms.

Case Study

Drowsy Driving Application

The proposed methodology system follows a step-by-step procedure and is described in block diagram as in Fig. 1: (i) Face detection using Projectiles (iii) Preprocess the image (iv) Transform to Neutrosophic set (v) Eye /Mouth Segmentation and (iv) Eye/Mouth state.

For segmentation of breast thermal images S. Motta et al. [2] applied global variance by Otsu's threshold to subtract the background with the final aim of segmentation of both breasts as symmetrically as possible and were satisfied with the results of their study. With the aim of mass detection in thermal images, R., M. et al. [3] applied thresholding based methods like grey threshold, RGB technique, and K-means technique and pointed out specifically that the RGB technique suffered from the limitation of data loss. Dayakshini, D. et al. [4] proposed a novel method of breast segmentation called Projection Profile Analysis for identifying the boundaries in the breast thermograms. For edge detection of the breast, they successfully used the Sobel operator while adaptive thresholding was applied for changing the ROI's color to white and reported that the overall results of their approach were satisfactory. For processing breast images by automated segmentation, De Oliveira et al. [5] applied various image processing steps like threshold, clustering, identification and improvement of corners on 328 lateral thermal images of breasts. A thresholding method was applied by them to separate the background and patient body in two modes. 96% and 97% of mean values of accuracy and sensitivity were obtained by them. Duarte A. et al. [1] developed an application using Matlab for characterizing thermal images that facilitates selection of any ROI without considering its geometry and provides options for its further optimizations. However, the thresholding algorithm implemented by them is not specified. Barcelos E. et al. [6] performed a study on 28 subjects who were athletes for the purpose of preventing muscle injury. They proposed automated recognition of flexible ROIs that have shape of body by optimizing the Otsu's method of segmentation and further performing progressive analyses of thermograms.

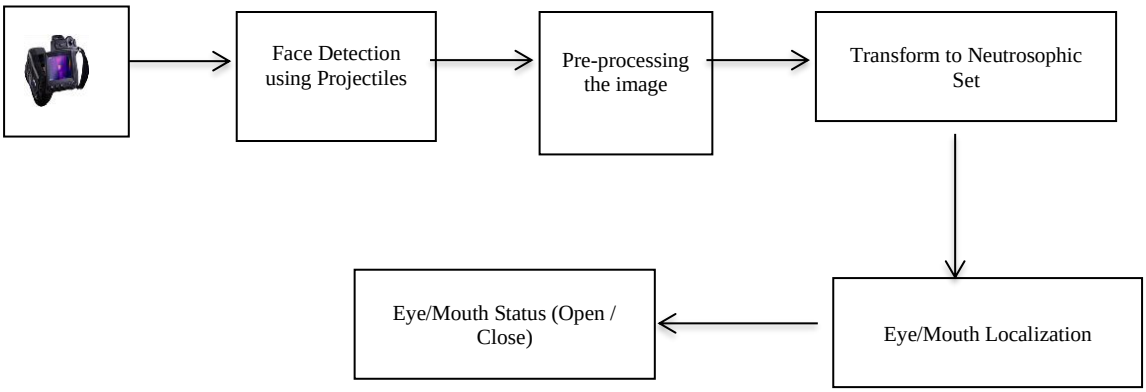


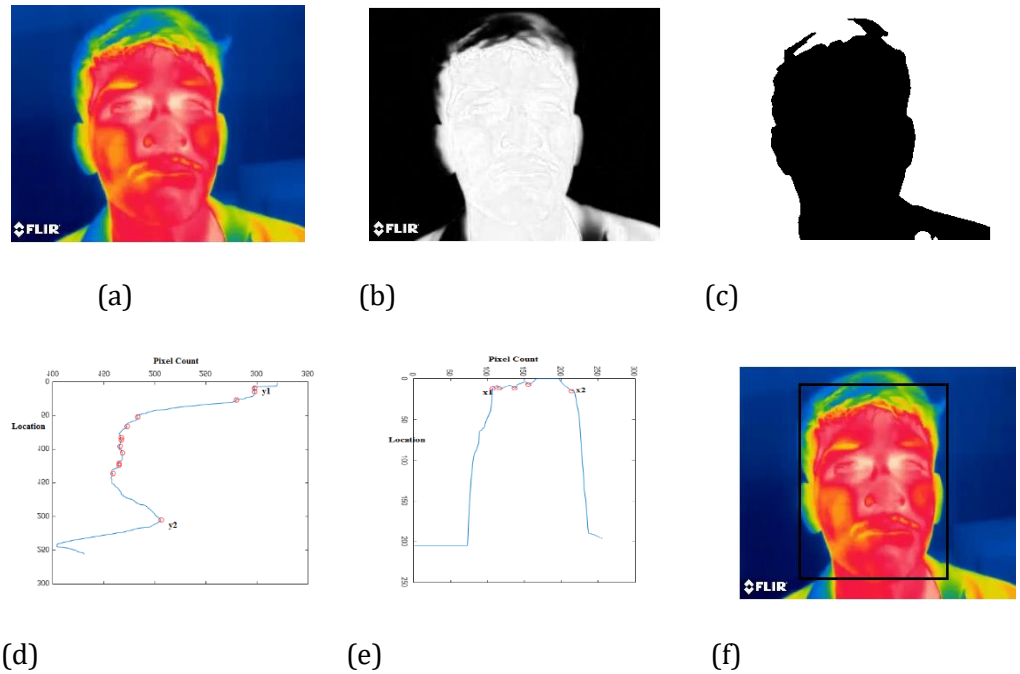
Fig. 1: Proposed Drowsy System Architecture

Face detection using Projectiles

The face is segmented using head curve geometry as the traditional cascaded algorithms such as fail in thermal images [19]. This step is done to focus only on the face by determine its boundary box area and eradicating the irrelevant information such as background to simplify the processing task. A step-by step procedure is described in algorithm 1 with pictorial representation in Figure 2 as follows:

Algorithm 1: Thermal Face detection

Input: Ithermal
Output: Boundary box area of the face [x1, y1, width, height]
Extract the red component Fig 2 (b) of the input image Ithermal Fig 2 (a)
Perform morphological operations such as dilation and closing to obtain the object boundary and inverse the resultant image. (Fig 2 (c)).
Compute the vertical (V) projections of the processed image and locate the local maxima points as plotted in Fig 2 (d).
The first peak point locates the y1 and following the maximum variation point indicates point y2. The distance between them indicates the height of the boundary box.
Compute the local maxima points with its location from horizontal (H) projections within the range of y1 and $\frac{y_2}{2}$ as plotted in Fig 2 (e).
Mark the first peak point as x1 and following the maximum variation point indicates point x2. The distance between them indicates the width of the boundary box.
Draw the boundary box with the above x1, y1, width and height on Ithermal as on Fig 2 (e).



(a) (b) (c)

(d) (e) (f)

Fig. 2 : (a) Input thermal image (b) Red component (c) Preprocessed image (d) Vertical projection graph (e) Horizontal projection graph (f) Face segmentation

Image Pre-processing

The images edges are highlighted using the Laplacian of Gaussian (LoG) filter. The Laplacian filter detects sudden intensity transitions in the images i.e it highlights the edges. The LoG (x, y) of the input thermal image (Fig. 3 (a)) with Gaussian standard deviation σ is written using Equation 1.

$$\mathrm{LoG}(x, y) = -\frac{1}{\pi\sigma^4}\left[1-\frac{x^2+y^2}{2\sigma^2}\right]e^{\frac{x^2+y^2}{2\sigma^2}} \quad (1)$$

The value returns a symmetric Laplacian of Gaussian filter by using a window filter size 7 with standard deviation $\sigma = 0.4$ as depicted in Fig 3 (b). To simplify the implementation, the processed image is then transformed to gray scale image as pictured in Fig. 3 (c). The processed image is then dilated by adding pixels to object boundaries that enhance the edges as portrayed in Fig. 3 (d).

(a) (b) (c) (d)

Fig. 3: Image Pre-processing of thermal face (a) Input thermal face Ithermal (b) LoG filtered image (c) Gray scale image (d) Dilated image

Neutrosophic Domain

Neutrosophic Set

In 1998, Smarandache introduced a new idea such as the neutrosophic fuzzy set (NFS) [19]. The NFS is represented as uncertain, incomplete, imprecise, and indeterminate information that prevail in many real world issues [19]. The definition, concepts and properties of neutrosophic set (NS) are mentioned in many research papers [20, 21].

Definition 3.3.1. Neutrosophic set (NS) : Consider Z to be a universe of discourse and a neutrosophic set S is a part of Z. An element z in set A is denoted in mathematical terms as $z(t, i, f)$ and represented in NS logic using Equation 2:

$$S = \{[z, (T_S(z), I_S(z), F_S(z))] z \in Z\} \quad (2)$$

where $T_S(z)$, $I_S(z)$ and $F_S(z)$ are the neutrosophic components and are real standard or non-standard sets of $]0^-, 1^+[$ and is defined using Equation 3 – 4 [21].

$$n_{sup} = t_{sup} + i_{sup} + f_{sup} \quad (3)$$

$$\text{where sup\_T} = t_{sup}, \quad \text{sup\_I} = i_{sup}, \quad \text{sup\_F} = f_{sup}$$

$$n_{inf} = t_{inf} + i_{inf} + f_{inf} \quad (4)$$

$$\text{where inf\_T} = t_{inf}, \quad \text{inf\_I} = i_{inf}, \text{inf\_F} = F_{inf},$$

$$\text{so}, 0^- \leq T_S(z) + I_S(z) + F_S(z) \leq 3^+$$

In the above equations T, I and F are defined as the degree of the true, indeterminate and false membership function of set A respectively [21]. An element x(t, i, f) belongs to set A and is represented in the following way: t% true, i% indeterminacy, and f% false. In this t varies in T, i varies in I, and f varies in F domain [20].

Neutrosophic Images

In neutrosophy, for any image-based application, the image from spatial domain is represented as a neutrosophic image as follows.

Definition 2.3.2 Neutrosophic image (NI): Consider Z to be a universe of the discourse as in definition 2.3.1 and the image window W = w * w i.e rows and columns in spatial domain. Thus W is a collection of image intensity pixels, where W ⊆ Z and it holds with bright pixels [22].

As per the Equation 5, the neutrosophic image is generally characterized by three membership sets T, I and F [21]. For the proposed thermal image with dimensions M * N, each pixel PT(m, n) is represented as PNS (m, n) in the neutrosophic image. PNS(m, n) interpret the memberships to bright, indeterminate and black intensity values through true TT(m, n), indeterminate I T(m, n) and false F T (m, n) as portrayed in Fig 4 (a), (b) and (c) respectively. It is represented using the following Equation 5 - 8. [23]:

$$PNS(m, n) = \{TT(m, n), I\ T(m, n), F\ T(m, n)\} \quad (5)$$

$$TT(m, n) = \frac{\overline{g(m, n)} - \bar{g}_{min}}{\bar{g}_{max} - \bar{g}_{min}} \quad (6)$$

$$\text{where } \overline{g(m, n)} = \frac{1}{w \times w} \sum_{x=m-w/2}^{m+w/2} \sum_{y=n-w/2}^{n+w/2} g(x, y)$$

$$IT(m, n) = \frac{\delta(i,j) - \delta_{min}}{\delta_{max} - \delta_{min}} \quad (7)$$

$$\text{where } \delta(m, n) = abs(g(m, n) - \overline{g(m, n)})$$

$$FT(m, n) = 1 - TT(m, n) \quad (8)$$

where,

g(m, n) – Thermal image's local mean value.

δ(m, n) – the absolute value of difference between intensity g(m, n) and its local mean value g(m, n)

(a) (b) (c)

Fig. 4: Neutrosophic Thermal Image: (a) TT domain (b) IT domain (c) FT domain

Alpha mean operation

As mentioned, the indeterminacy degree of PNFS holds significant information for segmentation. To increase the entropy of the indeterminacy domain, operations such as α-mean and β-enhancement is done.

Definition 2.3.3 The α-mean operation for PNFS = P (TT, IT, FT) and is defined using Equation (9) – (15).

PNFS(α) = P (TT (α), I T (α), F T (α))

(9)

The True α-mean set

$$\overline{TT}(\alpha) = \begin{cases} TT & if \quad IT < \alpha \\ \overline{TT} & if \quad IT \geq \alpha \end{cases}$$

(10)

where $\overline{TT}(m, n)=\frac{1}{w \times w}\sum_{x=m-w/2}^{m+w/2}\sum_{y=n-w/2}^{n+w/2} TT(x, y)$

(11)

The False α-mean set

$$\overline{FT}(\alpha) = \begin{cases} FT & if \quad IT < \alpha \\ \bar{F}T & if \quad IT \geq \alpha \end{cases}$$

(12)

where $\overline{FT}(m, n)=\frac{1}{w \times w}\sum_{x=m-w/2}^{m+w/2}\sum_{y=n-w/2}^{n+w/2} FT(x, y)$

(13)

The α value usually ranges within [0 1]. For experimentation various α values have been worked on, where α=0.9 holds promising results for thermal images.

The Indeterminate α-mean set

$$\overline{I_T}(\alpha) = \overline{I_T}(m, n)=\frac{\bar{\delta}(x, y)- \bar{\delta}_{min}}{\bar{\delta}_{max}- \bar{\delta}_{min}}$$

(14)

where $\delta(m, n)=abs(\overline{T_T}(m, n)- \overline{\overline{T_T}}(m, n))$

where $\overline{\overline{T_T}}(m, n)=\frac{1}{w \times w}\sum_{x=m-w/2}^{m+w/2}\sum_{y=n-w/2}^{n+w/2} \overline{T_T}(x, y)$

(15)

At the end of α-mean operation, among the NS images, the indeterminate image shows a uniform distribution of the pixel intensity values as shown in Fig. 5 (a) – (c).

(a) (b) (c)

Fig. 5: Alpha mean operation: (a) $\overline{T_T}(\alpha)$ domain (b) $\overline{I_T}(\alpha)$ domain (c) $\overline{F_T}(\alpha)$ domain

Beta-enhancement operation

Definition 3.3.4 [21]: The β-enhancement operation for PNFS = P (TT, IT, FT) is defined as Equation (16) – (22).

P`NFS(β) = P (TT`(β), IT`(β), FT`(β))

(16)

The True beta mean set is calculated as follows:

$$\grave{T}_T(\beta) = \begin{cases} TT & if \quad IT < \beta \\ \grave{T}_\beta & if \quad IT \geq \beta \end{cases}$$

(17)

where,

$$\grave{T}_\beta(\grave{m}, n)= \begin{cases} 2TT^2(m, n) & \text{if} \quad TT(m, n) \leq 0.5 \\ 1\text{-}2(1\text{-}TT(m, n))^2 & \text{if} \quad TT(m, n) > 0.5 \end{cases}$$

(18)

The False beta mean set is calculated as follows:

$$\grave{FT}(\beta) = \begin{cases} FT & \text{if} \quad IT < \beta \\ \grave{F}_\beta & \text{if} \quad IT \geq \beta \end{cases}$$

(19)

where,

$$\grave{F}_{\beta}(\grave{m}, n)= \begin{cases} 2FT^{2}(m, n) & \text{if} \quad FT(m, n) \le 0.5 \\ 1\text{-}2(1\text{-}FT(m, n))^{2} & \text{if} \quad FT(m, n) > 0.5 \end{cases}$$

(20)

The β value usually ranges within the interval [0 1]. For experimentation trials, various β values are computed, where α=0.85 yields good outcomes for thermal alpha images.

The Indeterminate mean set is calculated as follows:

$$\grave{I}_{T}(\beta) = \grave{I}_{T}(m, n)= \frac{\grave{\delta}(m,n)- \grave{\delta}_{min}}{\grave{\delta}_{max}- \grave{\delta}_{min}}$$

(21)

where $\grave{\delta}(m, n) = abs(\grave{T}_{T}(m, n) - \overline{\grave{T}_{T}}(m, n))$

$$\overline{\grave{T}_{T}}(m, n) = \frac{1}{w \times w}\sum_{x=m-w/2}^{m+w/2}\sum_{y=n-w/2}^{n+w/2}\grave{T}_{T}(x, y)$$

(22)

The β-enhancement operation removes the noise and retains detailed information as shown in Fig. 6 (a) – (c).

(a) (b) (c)

Fig. 6: Beta enhancement operation: (a) $\grave{T}_{T}(\beta)$ domain (b) $\grave{I}_{T}(\beta)$ domain (c) $\grave{F}_{T}(\beta)$ domain

Gamma-clustering operation

Finally, the images are segmented using K-means clustering technique. This technique groups the objects of same intensity values in the thermal images into K groups. Grouping is done using Equation 23 [24]:

$$M = \sum_{j=1}^{n}\sum_{i=1}^{o_j}\|D_i - Z_j\|$$

(23)

where Z_j and o_j are the center and the total count of pixels of the jth cluster and n represents total cluster numbers.

To achieve efficient clustering, value of M is minimized and made to satisfying the following Equation 24:

$$Z_j = \frac{1}{o_j}\sum_{D_i \in C_j} D_i$$

(24)

where $D=\{d_i, i=1, 2, 3.....t\}$, di represents a sample in the d-dimensional space. $C = \{C_1, C_2, C_3 \dots \dots . . C_n\}$ represents the partition segmentation that satisfies the Equation 25.

$$D = \bigcup_{i=1}^{n} C_i$$

(25)

The T and I subsets holds a new value by the changes in the indeterminacy after optimizing γ using Equation 26:

$$D(m, n) = \begin{cases} TT(m, n) & \text{if} \quad IT(m, n) < \gamma \\ \overline{T_{\gamma}}(m, n) & \text{if} \quad IT(m, n) \ge \gamma \end{cases}$$

(26)

Fig. 7 presents the outcome of the clustering operation for the optimized NS.

Fig. 7: Gamma-clustering operation

Eye and Mouth Localization

To localize the eyes and mouth region the resultant image is divided into halves where the upper half represents the eyes and the lower half represents the mouth region. To locate the region of interest (eyes, mouth) the following steps are considered.

eISSN 1303-5150

Determine the vertical projection on both the upper and lower half of face edge separately using Equation 27. The outcome of this step detects the right and the left eye/mouth region boundaries as depicted in Fig. 8 (a) and (b).

$V(n)= \sum_m grad(m, n)$ (27)

Similarly compute horizontal projection on the face edge using Equation 28. Here the upper and lower limits of the eye/mouth region boundaries is detected as shown in Fig. 7 (a) and (b).

$V(m)= \sum_n grad(m, n)$ (28)

(a) (b)

Fig. 8: (a) Localized Eyes (b) Localized Mouth

Eye and Mouth Status

After eye localization, the irrelevant information around the eyes is eliminated to classify the eye state as open or close. In image representation, the open state includes all eye parts but a closed state involves a slight line with eyelid alone [26]. Experiment analysis in done in paper [26, 30] about using the difference in upper and lower outer eye boundary by splitting into half. As mentioned, for an open state the upper eye frame pixel count is higher as it includes upper eyelids, iris etc. While in a closed eye state only lower eyelids [25, 26].

For mouth status, the upper, lower, left and right co-ordinates positions are noted using Equation 28. To detect yawning, a Circular Hough Transform (CHT) is used around the mouth area to extracts circles from edge images that indicate wide-open mouths [28].

Drowsy Detection

A driver's drowsiness is determined using both high yawning frequency and slow blinking frequency. Both the eye and mouth status is continuously monitored for consecutive frames. If the eye state is closed or the mouth is widely open for few consecutive frames based on a threshold value, it is considered as a blink or yawn.

A medical study [29] indicates drowsiness is defined when a person's mouth state remains open for four seconds or eyes remain closed for two or more seconds. Henceforth, the threshold value for blink is 60 (2 seconds for a 30 fps video) and for yawn is 120 (4 seconds for a 30 fps video) and when the proposed system exceeds this value it indicates a fatigue or drowsy and issues an alarm signal warning.

Algorithm 2: Proposed Drowsy Detection

Input: Thermal Face Image Ithermal

Output: Drowsy Detection

Compute and crop the thermal face along with this boundary box area using Algorithm 1.

For pre-processing, apply the LoG filter using Equation 1 as shown in Fig. 3 (b) and convert to gray scale image as Fig. 3 (c).

Dilate the edge boundary of the resultant gray scale image as depicted in Fig. 3(d).

Transform the pre-processed thermal face image ImgT into a neutrosophic fuzzy set PNS using Equation 6 – 8 as shown in Fig. 4 (a) – (c).

Apply alpha mean operation on PNS using Equation 7 - 13 for pixel distribution uniformity as shown in Fig.5 (a) – (c).

Perform beta enhancement operation on the NS image as portrayed in Fig.6 (a).

Segment clusters using gamma-clustering operation on the True set as depicted in Fig.7.

Divide the image into upper halve for eyes and lower half for mouth localization as shown in Fig. 8.

Determine the eye status using the method difference between upper and lower eye frame and mouth status using Circular Hough Transform.

Detect the drowsiness based on eye and mouth status for monitoring the consecutive frames.

Evaluation

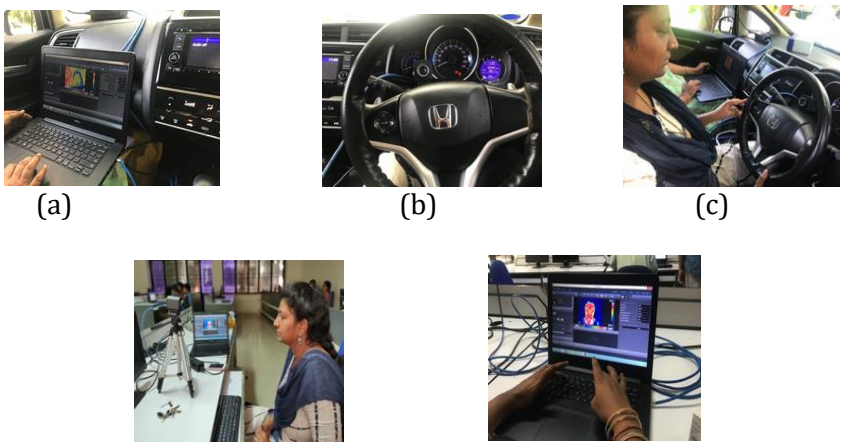
Experimental setup

For image acquisition, a thermal camera FLIR A35 is used. The camera used is 4.1 × 1.9 × 1.8 inches in physical size, and brings thermal imaging to the smallest space anywhere in a car setting at an affordable price. It produces a high quality with 320 × 240 pixel images and the captured images can track even slight changes as the temperature differences as small as 50 mK. The frame rate is about 30 images per second. On holding the advantage of small device, this experimental setup is

a non-invasive approach and it reduces the diversions while driving.

In both in-car and laboratory sets of experimental trials as photographed in Fig. 9 (a) - (e), the participants are seated at a distance of about 30 cm from the thermal camera and is tilted to visual angle of 0.10. In the laboratory simulation set-up, the thermal camera is placed in front of observers with manual adjustments set at eye level. The participants are instructed to behave like driving a car stimulating real road conditions such as prolonged eye closure and yawning. For in-car test inputs, the camera is placed on the dashboard just above and beyond the steering column of the car (2018 Honda Jazz). It is carefully noted that the camera is positioned left of the speedometer, so that the real-world events are not obstructed by the driver's view. Both tests inside laboratory and in-car recordings are conducted in steady ambient temperature of 66.2° F -71.6° F. The instructions are explained to the subjects and the scenario tests are done in a stationary vehicle.

For open source, thermal images database available in market [30, 34, 33, 37] is used. The images in open source database are captured using Phoenix Indigo IR which is a high-resolution thermal camera.



(a) (b) (c)

(d) (e)

Fig. 9 : (a)-(c) In-car setup (d)-(e) Laboratory setup

Performance measures

Bounding box regression

In face detection task, for each candidate window, the offset between the face and the nearest ground truth (top, side, height, and width of boundary boxes) is predicted. The Euclidean loss for each sample is employed using Equation 29 [31, 32]:

$$L_i^{box} = ||\hat{y}_i^{box} - y_i^{box}||^2 \quad (29)$$

where $\hat{y}_i^{box}$ is the ground-truth coordinate of the thermal face and y_i^{box} is the resultant four coordinates including top, side, height and width based on Algorithm 1.

Entropy

Thermal images generally have poor details and contrast values. Entropy quantifies the image information content i.e. it is used to calculate the intensity level distribution. The more detailed information indicates higher clarity of the image leading to higher entropy value. For neutrosophic image entropy is defined as the addition of entropies of T, I and F domain as shown in Equation 30 - 33 [35, 39].

$$En_{NS} = En_T + En_I + En_F \quad (30)$$

$$En_T = -\sum_{i=\min\{T_T\}}^{\max\{T_T\}} p_T(i) \ln p_T(i) \quad (31)$$

$$En_I = -\sum_{i=\min\{I_T\}}^{\max\{I_T\}} p_I(i) \ln p_I(i) \quad (32)$$

$$En_F = -\sum_{i=\min\{F_T\}}^{\max\{F_T\}} p_F(i) \ln p_F(i) \quad (33)$$

where EnT, EnI and EnF - entropies of T, I and F domain

pT(i), pI(i) and pF(i) – probabilities of elements in T, I and F

Sensitivity, Precision, Specificity, Accuracy and F1-Score

For eye/mouth state detection performance metrics such as sensitivity, precision, specificity, accuracy and F1- Score are used. Sensitivity or recall is defined as the average detection rate [10] and are formulated using Equation 34.

$$\text{Sensitivity} = \frac{TP}{TP+FN} \quad (34)$$

Precision is defined as true positive rate whereas specificity is true negative rate [10] and is formulated using Equation 35-37 as follows:

$$\text{Precision} = \frac{TP}{TP+FP} \quad (35)$$

$$\text{Specificity} = \frac{TN}{TN+FP} \quad (36)$$

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (37)$$

where, TP defines the number of true positive frames with an open state as open and TN is the number of true negative i.e. predicting the exact status as closed by comparing them manually. FP is the number of false positives i.e. unidentified eye/mouth image or an opened state is predicted as closed one. FN is the number of false negatives i.e. falsely identified face image or a closed eye/mouth is identified as opened one.

To resolve the issue of trade-off between precision and recall, the F1-score is used. It determines the detector's absolute performance, so that they can be easily compared to one another [12]. The F1-score is defined in Equation 38 as the harmonic mean of the precision and recall [18]:

$$\text{F1-score} = 2 * \frac{1}{\frac{1}{\text{recall}} + \frac{1}{\text{precision}}}$$

Experimental results

The error measure of the bounding box regression for the test frames using Equation 29 is tabulated in Table 1.

Table 1 : Relative error measure (L) for various face localization methods

Methods/Relative Error	$L_i^{box} \leq 0.05$	$L_i^{box} \leq 0.15$	$L_i^{box} \leq 0.25$
Projection profile	93.1%	95.5%	98.5%
Viola Jones algorithm	36.6%	56.8%	87.2%
SVM with Gabor Features	47.3%	75.4%	96.0%

Table 1 indicates Projection profile achieves a better performance with correct identification rate of 93-98%. The mean distance to the true position over all detections that are considered to be correct is 0.02.

Table 2: Entropy values of Images

Techniques	Entropy value
Proposed	6.1455
Neutrosophic Image	18.8102
Alpha mean operation	19.1575
Beta enhancement operation	20.234

To improve the robustness and effectiveness of the transformed neutrosophic thermal images, T, I, F are enhanced using alpha and beta operations using Equation 5 – 26. The entropy of these images calculated using Equation 30 – 33 signifies how much detailed information is extracted. Table 2 (a) portrays an increase in entropy values of the thermal images after alpha and beta operation.

The performance metrics sensitivity, specificity, precision, F1-score, and accuracy values are tabulated in Table 4 on both open and customized databases. It is observed that the proposed model achieves an average accuracy of 97.012% in determining the eye and mouth status correctly when compared with other existing segmentation methods as shown in Table 3. As shown by the test results, this eye and mouth state method is considered to be more robust and successful, with a very few false negatives.

Table 3: Accuracy of different Segmentation approaches

Techniques	Accuracy (%)
Superpixel [34]	72.358
LSACM [35]	88.491
FDMT [36]	92.436
Proposed Neutrosophic	97.012

Table 4: Sensitivity, specificity, precision, F1-score, and accuracy values

Database	Trials	Performance Metrics (%)				
		Sensitivity	Precision	Specificity	Accuracy	F1 Score
Customized	Trial-1	96.39	93.02	93.75	94.97	94.68
	Trial-2	98.13	99.06	99.33	98.83	98.59
	Trial-3	94.81	97.33	98.29	96.91	96.05
	Trial-4	97.41	98.26	98.36	97.9	97.83
	Trial-5	95.24	83.33	91.55	94.68	88.89
Open	Trial-1	96.74	93.68	95.38	95.95	95.19
	Trial-2	98.04	99.01	99.25	98.72	98.52
	Trial-3	95.6	97.75	98.29	97.12	96.66
	Trial-4	96.84	97.87	98.59	97.89	97.35
	Trial-5	97.39	93.33	93.94	95.55	95.32
Average		96.659	95.264	96.673	97.012	95.908

Results of eye and mouth state frame by frame are presented in Fig. 10 and Fig. 11 respectively. From analyzing a sequence of frames in Figure 9 the correspondence between eyelid closures is noted at frame 1230. As the eye state remains closed for above the threshold for 90 frames, an alarm sounds are triggered indicating the driver is in a drowsy state.

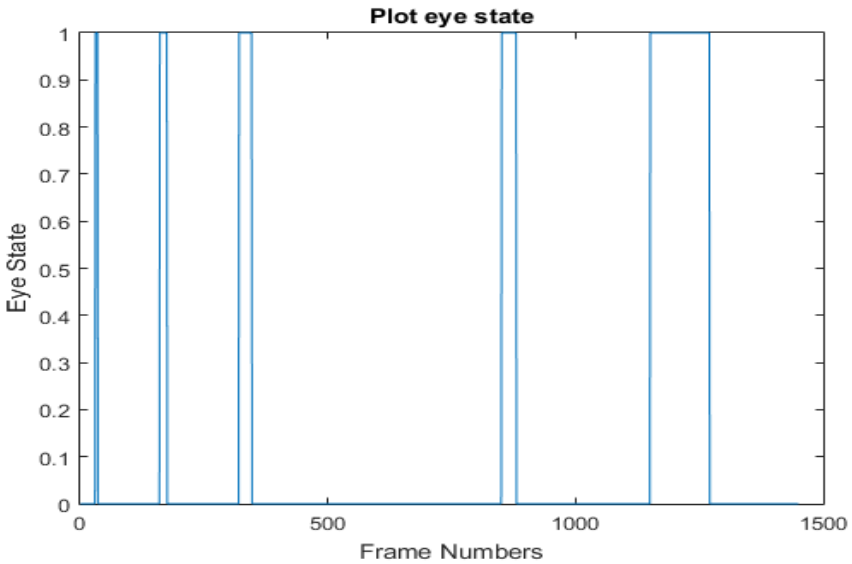


Fig. 10 Frame by frame eye state visualization

The yawning reflex leads to temperature deviations in thermal mouth area. The reason is due to inhalation of air makes the temperature decrease inside the mouth area when compared to close state. On analysis in parallel from the sequence of frames in Fig. 11 mouth openness is noted at frame 860. The proposed system indicates drowsiness, when a person's mouth state remains open for four seconds or eyes remain closed for two or more seconds.

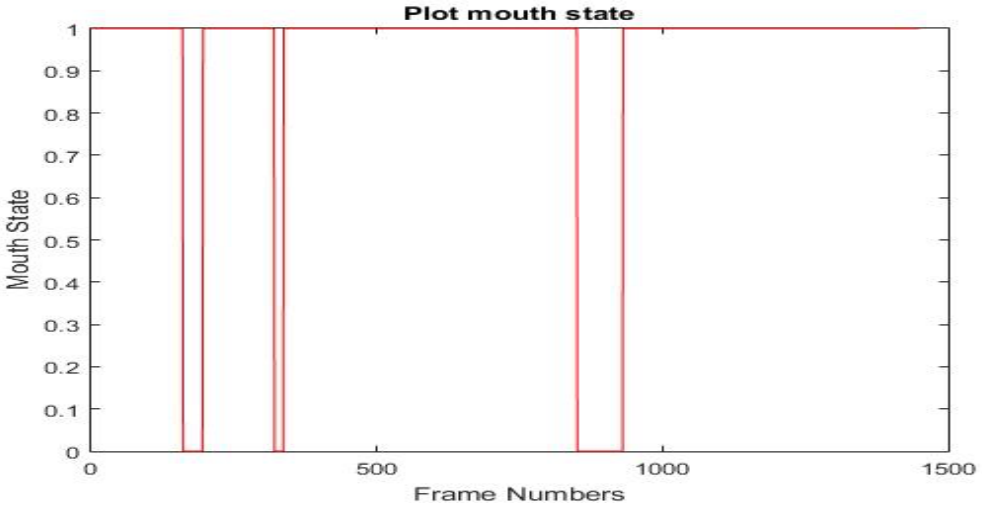


Fig. 11 Frame by frame mouth state visualization

The proposed framework is compared with other existing frameworks on both open and customized datasets and graphed in Fig.12.

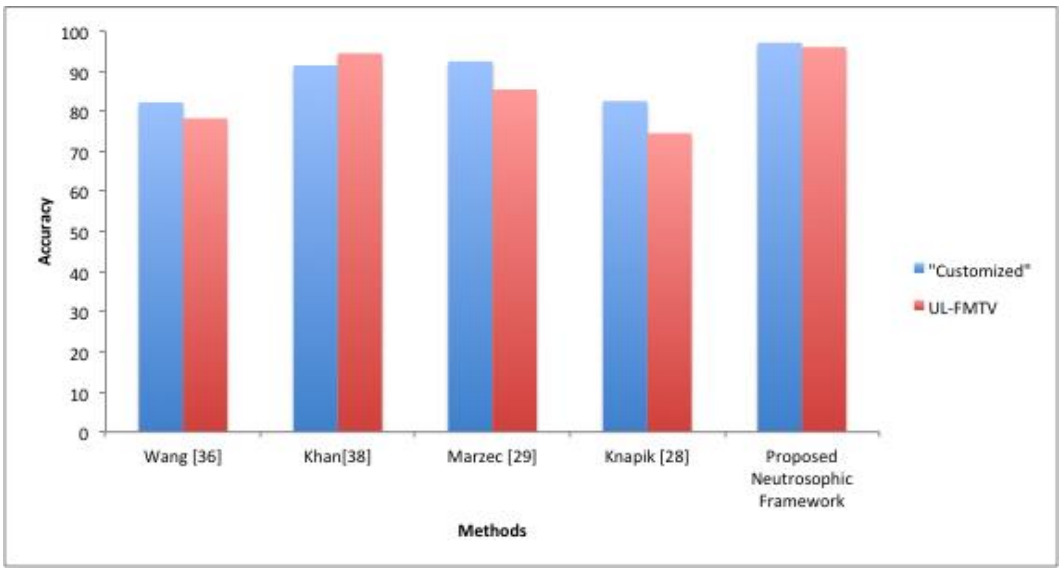


Fig.12 Accuracy obtained and compared with other frameworks

Conclusion

The proposed work deals with neutrosophic segmentation approach on thermal images for a study on drowsy detection. At first, the input images are pre-processed and forwarded to a neutrosophic framework. The neutrosophic images are enhanced using alpha mean and beta enhancement to work on neutrality. The outcome is then clustered with finite feature extraction of eyes and mouth. Then the eyes and mouth are localized and the cropped facial parts are binary classified as open or closed eye/mouth state for drowsy detection. The differentiation in states is based on the difference in pixels in sub frame and circular Hough transform for eyes and mouth respectively. Experimental results concludes that the segmentation using neutrosophic consistently outperform the state-of-the-art methods. Then the system also studies and analysis the drowsy driving application with other available methods.

In future work, more focus is needed on computational speed, as even on high-end hardware the developing cycle is long for real time applications.

References

Bergasa L, Nuevo J, Sotelo M, and Vazquez M.: Real-time system for monitoring driver vigilance, IEEE Transactions on Intelligent Transportation Systems, vol. 7, no. 1, pp. 63–77, 2006.

Fu H., Wei Y., Camastra F., Arico P., and Sheng H.: Advances in

Eye Tracking Technology: Theory, Algorithms, and Applications, Computational Intelligence and Neuroscience, vol.3, 2016.

Zhang L., Cao Y., Yang F., and Zhao Q.: Machine Learning and Visual Computing, Applied Computational Intelligence and Soft Computing, vol. 2, 2017.

N. M. Zaitoun and M. J. Aqel, "Survey on image segmentation techniques", Procedia Comput. Sci. 65, 797–806, 2015.

A. Comelli, A. Bruno, M. L. Di Vittorio, F. Ienzi, R. Lagalla, S. Vitabile and E. Ardizzone, "Automatic multi-seed detection for MR breast image segmentation," in ICIAP 2017: Image Analysis and Processing, Lecture Notes in Computer Science, Vol. 10484 (Springer, 2017), pp. 706–717

N. Shrivastava and J. Bharti, "A comparative analysis of medical image segmentation," in Proc. 2018 4th Int. Conf. Advanced Computing Networking and Informatics, Advances in Intelligent Systems and Computing, Vol. 870 (Springer, 2019), pp. 459–467.

V. P. Singh, S. Srivastava and R. Shrivastava, "Automated and effective content-based image retrieval for digital mammography," J. X-Ray Sci. Technol. 26, 29–49, 2018.

Wang S, Liu Z, Shen P, Ji Q: Eye localization from thermal infrared images, Pattern Recogn, Vol. 46(10): pp. 2613–2621, 2013.

Stra¸kowska M, Stra¸kowski R: Automatic eyecorners detection and tracking algorithm in sequence of thermal medical images, Measure Autom Monitor, Vol.61(6):199–202, 2015.

F. Smarandache, A unifying field in logics: Neutrosophic logic, Multiple-Valued Logic, 8, pp. 489–503, 1999.

F. Smarandache, A unifying field in logics: Neutrsophic logic, neutrosophy, neutrosophic set, neutrosophic probability (Fourth edition), University of New Mexico, 332, pp. 5–21, 2002.

Kuei-Hu Chang, A novel risk ranking method based on the single valued neutrosophic set. Journal of Industrial & Management
Optimization, 2021 doi: 10.3934/jimo.2021065

Marzec M, Lamz˙a A,Wro´bel Z,Dziech A, Fast eye localization from thermal images using neural networks. Multimedia Tools and Applications, 2016.

Cyganek B, Gruszczyn´ski S: Hybrid computer vision system for drivers' eye recognition and fatigue monitoring, Neurocomputing ,Vol.126, pp. 78–94, 2014.

Khan S A, Hussain S, Xiaoming S, Yang S, An effective framework for driver fatigue recognition based on intelligent facial expressions analysis. IEEE Access Vol.6, pp. 67459–67468, 2018.

Knapik M, Cyganek B: Driver's fatigue recognition based on yawn detection in thermal images. Neurocomputing, Vol.338, pp.274–292, 2019.

Farfade S. S., Saberian M. J., and Li L. J.: Multi-view face detection using deep convolutional neural networks, ACM on International Conference on Multimedia Retrieval, pp. 643-650, 2015.

Ghiass, Reza, Bendada, Hakim & Maldague, X.. Université Laval Face Motion and Time-Lapse Video Database (UL-FMTV). 10.21611/qirt.2018.051, 2018.

Zhang Y, Chu J, Leng L, Miao J :Mask-refined r-cnn: a network for refining object details in instance segmentation. Sensors, Vol. 20(4):1010, 2020.

F. Smarandache,: A Unifying Field in Logics Neutrosophic Logic. Neutrosophy, Neutrosophic Set, Neutroso- phic Probability, American Research Press, 2003.

Guo Y, Cheng HD, New neutrosophic approach to image segmentation. Pattern Recogn 4:587–95, 2009.

Guo Y, Cheng HD, A new neutrosophic approach to image denoising. New Math Nat Comput 5:653–662, 2009

M. Zhang, L. Zhang, H. D. Cheng, "A Neutrosophic approach to image segmentation based on Watershed method," Signal Processing, vol. 90, no. 5, pp.1510-1517, May 2010.

Y. Guo and Y. Akbulut et al., An efficient image segmentation algorithm using neutrosophic graph cut, Symmetry, 9, 185, 2017.

Sofia Jennifer J and Sree Sharmila T: Real time blink recognition from various head pose using single eye, Multimedia Tools and Applications, pp 1-15, 2018.

Sofia Jennifer J, Sree Sharmila T and Srinivasan R: "Facial feature extraction for head tilt images based on eye canthus, IEEE TENCON, 2016.

Sofia Jennifer J and Sree Sharmila T: Edge Based Eye-Blink Detection for Computer Vision Syndrome, IEEE ICCCSP Conference, 2017.

Knapik M, Cyganek B: Driver's fatigue recognition based on yawn detection in thermal images. Neurocomputing, Vol.338, pp.274–292, 2019.

Marzec M, Lamz˙a A,Wro´bel Z,Dziech A, Fast eye localization from thermal images using neural networks. Multimedia Tools and Applications, 2016.

Farfade S. S., Saberian M. J., and Li L. J.: Multi-view face detection using deep convolutional neural networks, ACM on International Conference on Multimedia Retrieval, pp. 643-650, 2015.

Wang, W.; Duan, L.; Wang, Y. Fast Image Segmentation Using Two-Dimensional Otsu Based on Estimation of Distribution Algorithm. J. Electr. Comput. Eng. 2017, 2017, 12.

Jing Zhao, Xiaoli Wang and Ming Li, A Novel Neutrosophic Image Segmentation Based on Improved Fuzzy C-Means Algorithm (NIS-IFCM), International Journal of Pattern Recognition and Artificial Intelligence, Vol. 34, No. 05, 2019

Jinyu Wen, Shibin Xuan, Yuqi Li, Qing Gao and Qihui Peng," Image-segmentation algorithm based on wavelet and data-driven neutrosophic fuzzy clustering", The Imaging Science Journal, Vol. 67, No. 2, 2018.

J. Cheng, J. Liu, Y. Xu et al., "Superpixel classification based optic disc and optic cup segmentation for glaucoma screening," IEEE Transactions on Medical Imaging, vol. 32, no. 6, pp. 1019–1032, 2013.

K. Zhang, L. Zhang, K.-M. Lam, and D. Zhang, "A level set approach to image segmentation with intensity in-homogeneity," IEEE Transactions on Cybernetics, vol. 46, no. 2, pp. 546–557, 2016.

R. Priyadharsini, A. Beulah and T. Sree Sharmila," Optic disc and cup segmentation in fundus retinal images using feature detection and morphological techniques",Current Science, vol. 115 (4), pp. 748 – 752, 2018.

*e*ISSN 1303-5150