



A Neutrosophic Approach for Breast Mass Detection in digital Mammogram Images

Sofia Jennifer J¹, and Kalaivani C²

¹Sri Sivasubramaniya Nadar College of Engineering, Kalavakkam, Chennai-603 110, India.

E-mail: sofiajenniferj@ssn.edu.in

²Sri Sivasubramaniya Nadar College of Engineering, Kalavakkam, Chennai-603 110, India.

E-mail: kalaivaniC@ssn.edu.in

Abstract: Breast cancer is the most common type of cancer that affect women and yet there is no substantive cure. On early diagnosis of breast lesions with this size using mammogram would help the affected tumor patients have a good survival rate. Detecting cancer cells is a challenging task and hence we propose a single valued neutrosophic approach for segmenting breast tumor lesions in mammogram images. In this neutrosophic approach, the intensity values are represented as truth, indeterminacy and falsity membership in where the indeterminacy are noise, and the breast regions are the true and falsity values. The computations of these memberships are then pre-processed using α -mean and β -enhancement to minimize the indeterminacy. Finally, gamma clustering techniques are used to localize and segment the tumor from the background breast tissue region. Experiments are conducted on publicly available datasets and the proposed method is evaluated and achieves a better segmentation performance.

Keywords: Breast Mass detection; Neutrosophic domain; Alpha-mean; Beta enhancement; Gamma Clustering

1. Introduction

Cancer is an abnormal cell growth that invades all parts of the body bringing health issues. Early detection of tumour or benign cells helps us to save life when given proper treatment. Automatic diagnosis of tissue mass has been a great area of research for a few decades [1]. Breast cancer is developed in breast tissues are very common in women and the risk factors are obesity, alcohol consumption, genetics, early or late menopause etc [2]. In many medical cases, breast tumours are usually located in milk ducts tubes or in the milk secreting gland lobules. Studies have shown that on early detection it can be cured effectively.

Mammography is a low dose X-ray image capture technique used to screen breast for tumour detection. Patients with lumps or mass in breast are recommended for this screening. These lumps can be either benign or malignant. Hence the proposed work motivates to concentrate on localizing these breast lesions or mass. To detects these, the approaches have been categorized into model-based and non-model-based. In a model-based approach, techniques such as shape-based template matching, de-formable model, and deep learning model are involved. Al-Antari et al. [3] uses YOLO (You-Only-Look-Once) to detected the breast mass and FrCN (full resolution convolutional network) to segment the mass. Ahmed et al. [4] uses a hybrid method based on Deeplab-V3 and

Mask-RCNN for breast mass segmentation. Dhungel et al. [5] proposed a SSVM (structured support vector machine) model for locating breast lesion.

In non-model based, Kuo et al. [6] propose the enhancement of the suspected regions of breast mass in the mammographic image, then identified and located the breast mass by PSO. Stojić et al. and Amutha et al. [7] enhanced the edge details on the mammographic image using the mathematical morphology method. Liu et al. [8] uses the image template matching with a bright circular image template. By this method, region of interest is highlighted from the background to detect the mass effectively.

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, in image domain the intensity values classify it is an object or background in an image [9]. The uncertainty parameters such as noise or image acquisition negligence error are avoided, thus degrading the performance on using algorithms. To overcome this issue, a Neutrosophic set (NS) an extended version of fuzzy logic is introduced that deals these uncertainty or neutralities. The spatial images are converted into a set with degree to truth, falsity, and indeterminacy membership. Medical domain is the field where there is indeterminacy, unknown, hidden parameters, imprecision, high conflict between sources of information, non-exhaustive or non-exclusive elements of the frame of discernment so neutrosophy could be applied [10]. In the proposed application, indeterminacy is the medical information such as imprecise, partial, and vague or imperfect lab results or degradation of the X-ray equipment.

Our proposed work employs the positives of neutrosophic sets along with gamma clustering to localize the breast masses. This above approach leads to less computation time and effort claiming with best results. The rest of the paper is organized as follows: The proposed methodology is explained in Section 2. Section 3 explains the results of the experiment carried out using the proposed method and the conclusion of this work is discussed in Section 4.

2. Proposed Work

The proposed architecture consists of the following steps: (i) Neutrosophic domain (ii) Alpha Mean operation (iii) Beta enhancement operation (iv) Gamma Clustering. The system focuses on localization of the lesions and the boundary information related to it. Figure 1 is a sample input mammogram X-ray image.

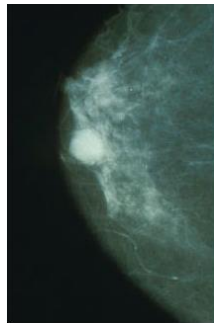


Fig 1: Sample Input X-ray image

2.1 Neutrosophic Domain

2.1.1 Neutrosophic Set (NS)

Neutrosophy derived on the concept of knowledge of neutral thought. It is defined as the study of origin, nature, and scope of neutralities, as well as their interactions with different ideational spectra [9]. It reveals and explore the world is full of indeterminacy.

Definition: Neutrosophic set (NS): Consider Z to be a universe of discourse and in where Neutrosophic set (S) is a part of it. In mathematical terms, an element z in set A is written as $z(t, i, f)$ and represented in NS logic using:

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

(1)

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:

$$n_{sup} = t_{sup} + i_{sup} + f_{sup}$$

(2)

$$\text{where } \sup_T = t_{sup}, \sup_I = i_{sup}, \sup_F = f_{sup}$$

$$n_{inf} = t_{inf} + i_{inf} + f_{inf}$$

(3)

$$\text{where } \inf_T = t_{inf}, \inf_I = i_{inf}, \inf_F = f_{inf},$$

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

In Equations (1) - (3), T , I and F are defined as the degree of the true, indeterminate and false membership function of set A respectively [11]. 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 [12].

2.1.2 Images in NS domain

For an image processing application, initially the X-ray image from spatial domain is represented as a neutrosophic image as represented as follows.

Definition: Neutrosophic image (NI): Let's consider Z to be a universe of the discourse and size of the image window is represented as $W = w_1 * w_2$ where w_1, w_2 indicates the rows and columns in spatial domain respectively. Thus, W holds the values of image intensity pixels range from 0 to 255 where $W \subseteq Z$ [13]. As per the Equation 4, the neutrosophic image is characterized by T , I and F membership sets [13]. For every input X-ray image with size $M*N$, each pixel $Image_x(m, n)$ is transformed to $Image_{NS}(m, n)$ in neutrosophic domain. In $Image_{NS}(m, n)$ the bright pixels represent the true domain $T_x(m, n)$, indeterminate pixels as $I_x(m, n)$ and dark pixels as $F_x(m, n)$. This representation of NS image is depicted as Fig 2.

$$Image_{NS}(m, n) = \{T_x(m, n), I_x(m, n), F_x(m, n)\} \tag{4}$$

$$T_x(m, n) = \frac{\overline{g(m, n)} - \bar{g}_{min}}{\bar{g}_{max} - \bar{g}_{min}} \tag{5}$$

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

where the $g(x,y)$ in Equation (5) is the input X-ray image and window size $w=3,5$ or 7 . On execution $w=3$ shows good performance results when compared to other values. In the equation, \bar{g}_{min} and \bar{g}_{max} are the minimum and maximum local mean intensity value of $\overline{g(m, n)}$.

$$I_x(m, n) = \frac{\delta(i,j) - \delta_{min}}{\delta_{max} - \delta_{min}} \tag{6}$$

where $\delta(m, n) = \text{abs}(g(m, n) - \overline{g(m, n)})$. The minimum of $\delta(m, n)$ is computed as δ_{min} and the maximum as δ_{max} .

$$F_x(m, n) = 1 - T_x(m, n) \tag{7}$$

where, $g(m, n)$ – Input X-ray image’s local mean value.
 $\delta(m, n)$ – the absolute difference between pixel intensity $g(m, n)$ and mean value $g(m, n)$

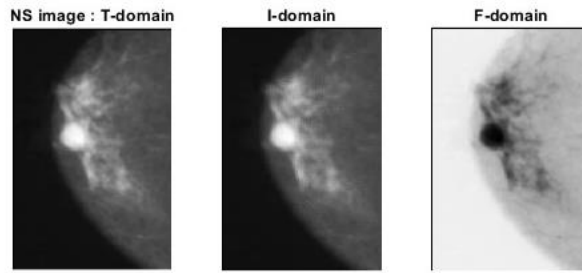


Fig 2: Images in NS domain

2.2 Alpha-mean operation

Various factor leads to uncertainty during the image acquisition errors stages such as intrinsically imperfect lab observations and the quantitative errors in measures. To reduce the indeterminacy degree $Image_{NS}(m, n)$ leads to better segmentation process. To maximize the entropy of true and false domain α -mean is used.

Definition: The α -mean operation for the input $Image_{NS} = Image_x(T_x, I_x, F_x)$ and the transformation is defined using Equation (8) – (11) [14].

$$Image_{NS}(\alpha) = Image_x(T_x(\alpha), I_x(\alpha), F_x(\alpha)) \tag{8}$$

The True α -mean set

$$\overline{T_x}(\alpha) = \begin{cases} T_x & \text{if } I_x < \alpha \\ \overline{T_x} & \text{if } I_x \geq \alpha \end{cases} \tag{9}$$

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

The False α -mean set

$$\overline{F_x}(\alpha) = \begin{cases} F_x & \text{if } I_x < \alpha \\ \overline{F_x} & \text{if } I_x \geq \alpha \end{cases} \tag{10}$$

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

The α value ranges within [0 1] and on experimenting many trial and error cases promising outcomes are shown when $\alpha=0.9$.

The Indeterminate α -mean set

$$\overline{I_x}(\alpha) = \overline{I_x}(m, n) = \frac{\overline{\delta}(x, y) - \overline{\delta}_{\min}}{\overline{\delta}_{\max} - \overline{\delta}_{\min}} \tag{11}$$

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

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

At the end of α -mean operation, the indeterminate pixel intensity is reduced thus enhancing the true and false domain as depicted in Figure 3.

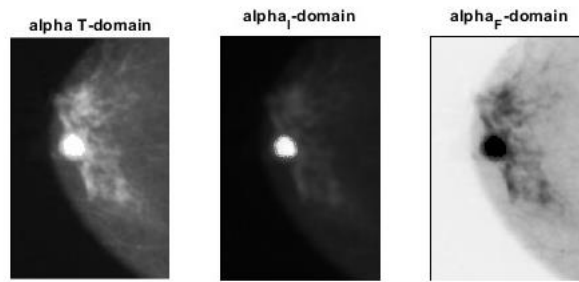


Fig 3: Alpha mean operation of NS image

2.3 Beta-enhancement operation

Definition: The β -enhancement operation for the input $Image_{NS}(\alpha) = Image_x(T_x(\alpha), I_x(\alpha), F_x(\alpha))$ is defined as Equation (12) – (22) [14].

$$Image_{NS}(\beta) = Image_x(T_x(\beta), I_x(\beta), F_x(\beta)) \tag{12}$$

The True beta mean set is calculated as follows:

$$\hat{T}_x(\beta) = \begin{cases} T_x & \text{if } I_x < \beta \\ \hat{T}_x & \text{if } I_x \geq \beta \end{cases} \tag{13}$$

where,

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

The False beta mean set is calculated as follows:

$$F_x(\beta) = \begin{cases} F_x & \text{if } I_x < \beta \\ \hat{F}_x & \text{if } I_x \geq \beta \end{cases} \tag{14}$$

where,

$$\hat{F}_\beta(m, n) = \begin{cases} 2F_x^2(m, n) & \text{if } F_x(m, n) \leq 0.5 \\ 1-2(1-F_x(m, n))^2 & \text{if } F_x(m, n) > 0.5 \end{cases}$$

On experimentation, β value also ranges within the interval 0 to 1 and hence multiple values are tested and an optimized results are yielded for X-ray images when $\beta = 0.85$.

The Indeterminate mean set is calculated as follows:

$$\hat{I}_x(\beta) = \hat{I}_x(m, n) = \frac{\delta(m, n) - \delta_{\min}}{\delta_{\max} - \delta_{\min}} \tag{15}$$

$$\text{Where, } \delta(m, n) = \text{abs}(\hat{T}_x(m, n) - \hat{I}_x(m, n))$$

$$\hat{T}_x(m, n) = \frac{1}{w \times w} \sum_{x=m-w/2}^{m+w/2} \sum_{y=n-w/2}^{n+w/2} T_x(x, y)$$

At the end of the β -enhancement operation the noises are reduced, thus retaining more edge information for segmentation as shown in Figure 4.

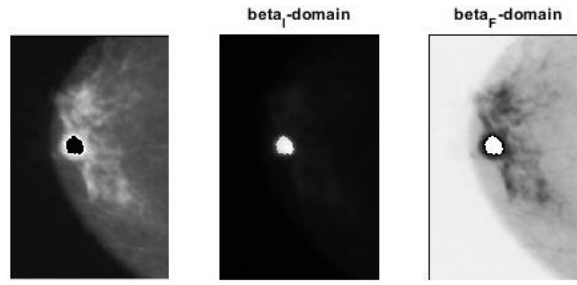


Fig 4: Beta enhancement operation of NS image

2.4 Gamma Clustering Algorithm

The final step involves segmentation of the breast masses using K-means clustering technique. As the name signifies, it groups the pixels of same intensity values as objects in the enhanced beta true images into K groups. The grouping is done using Equation 16 [24]:

$$Z = \sum_{j=1}^k \sum_{i=1}^{M_j} \|S_i - P_j\| \tag{16}$$

where P_j and M_j are the center and the total count of pixels of the j^{th} cluster and k represents total cluster numbers [14].

This classic clustering technique is further enhanced by data distribution with Gamma factor as it deals with skewed data. Minimize the value of Z using the following Equation 17

$$P_j = \frac{1}{M_j} \sum_{S_i \in C_j} S_i \tag{17}$$

where $S = \{s_i, i=1, 2, 3, \dots, t\}$, s_i represents the sample in the d -dimensional space and $C = \{C_1, C_2, C_3, \dots, C_n\}$ represents the segmented clusters that satisfies the Equation 18.

$$D = \bigcup_{i=1}^n C_i \tag{18}$$

The clustered data are represented in different color regions for differentiation purpose as depicted in Figure 5

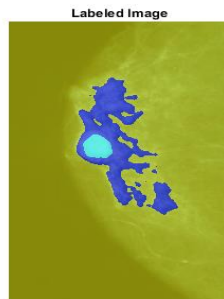


Fig. 5: Gamma-clustering operation (Clustered output)

<i>Algorithm 1: Breast Tumor Detection</i>
<i>Input: Input X-ray data from open-source dataset</i>
<i>Output: Localize the mass regions in breast.</i>

- i. Input the image from the open-source dataset.
- ii. Transform the input mammography image ($Image_x$) to Neutrosophic domain ($Image_{NS}$) using Equation (4) where T_x , I_x and F_x are calculated using Equation (5), (6) and (7) respectively.
- iii. Reduce the indeterminacy using alpha mean operation on each neutrosophic set images from (T_x , I_x , F_x) to ($T_x(\alpha)$, $I_x(\alpha)$, $F_x(\alpha)$) using Equation (8)-(11).
- iv. Enhance the true image using beta enhancement operation on each neutrosophic alpha set images from ($T_x(\alpha)$, $I_x(\alpha)$, $F_x(\alpha)$) to ($T_x(\beta)$, $I_x(\beta)$, $F_x(\beta)$) using Equation (12)-(15).
- v. Segment the resultant image using Gamma K-means Clustering
- vi. Draw the boundary box on the all the blobs of the segmented image to determine the masses.

3. Evaluation

3.1 Datasets

There are many open datasets available in the market such as Breast Imaging Reporting and Data System (BI-RADS), Emory BrEaSt imaging Dataset (EMBED), Mammographic Image Analysis Society (MIAS), INbreast database etc. Our proposed method is worked on samples from MIAS and INbreast database. In MIAS, each image had the dimensions 1024×1024 , and 90 images out of 322 contained tumor masses. In INbreast database holds about 410 images captured from 115 patients, out of which 90 cases with malignant masses and others are benign, calcifications or other types of masses.

3.2 Performance metrics

3.2.1 Entropy

Entropy calculates the intensity level distribution thereby quantifies the image information content. For better outcomes, a good clarity image with higher entropy value is needed at the end of enhancement stage. For a neutrosophic image entropy is defined as the addition of entropies of T, I and F domain using Equation 19 [14].

$$En_{NS} = En_T + En_I + En_F \quad (19)$$

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

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

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

where En_T , En_I and En_F - entropies of T, I and F domain

$p_T(i)$, $p_I(i)$ and $p_F(i)$ - probabilities of elements in T, I and F

3.2.2 Dice similarity co-efficient (DICE)

DICE is a statistical metric to validate the segmentation approach. It is a simple approach and is computed by the extent of overlapping pixel values between images that takes the segmentation area and the background with respect to the ground truth segmentation [15]. The Dice coefficients are defined using Equation 20:

$$DICE = \frac{2 \times P_{TP}}{2 \times P_{TP} + P_{FP} + P_{FN}} \quad (20)$$

where P_{TP} is the number of true positive, N_{FP} and N_{FN} is the is the number of false positive and false negative respectively. The positive and negative term refers to the pixels belonging to the mass tissue area and background area in comparison with ground truth [16].

3.2.3 Boundary localization error (BLE)

To detect the accuracy of the localization of breast mass region, the difference between the predicted and the closest actual ground truth (top, side, height, and width of boundary boxes) are calculated. The learning objective is formulated using Equation 21 [17]:

$$L_i^{box} = ||\hat{z}_i^{box} - z_i^{box}||^2 \quad (21)$$

where z_i^{box} is the original or manual annotated coordinate position of the mass regions and \hat{z}_i^{box} is the outcome of the proposed work coordinates including top, side, height and width.

3.2.4 Accuracy

The accuracy of the segmentation approach on mammography X-ray images using single valued neutrosophic approach is calculated using the Equation (22):

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (22)$$

In the above equation, TP defines the number of true positive images with all detected masses and TN is the number of true negative i.e. predicting the non-mass region by comparing them manually. FP is the number of false positives i.e. non-mass region as mass. FN is the number of false negatives i.e. falsely neglecting a mass region.

3.3 Experimental Results

To refine the input images for better robustness and effectiveness, the neutrosophic images, T, I, F are enhanced using alpha and beta operations using Equation 4 – 15. At each stage, the entropy is calculated using Equation 19 to ensure how much detailed information is extracted. The values are tabulated in Table 1.

Table 1: Entropy values of Images in each pre-processing stage

Methods	DICE	BLE	ACCURACY
Image Template [1]	0.914	82.4%	87.27
PSO algorithm [6]	0.931	83.2%	86.63
Deep structured learning [5]	0.955	87.5%	89.01
Adaptive threshold [3]	0.85	71.7%	72.2
Proposed Neutrosophic	0.972	92.8%	92.38

The tabulated values derive the idea that after beta enhancement the images shows well defined edges and corners for better segmentation. For segmentation, metrics such as DICE, BLE and accuracy are used to compare with the other existing approaches available in market. For comparison purposes, BLE with the relative error measure ($L_i^{box} \leq 0.05$) is only considered in the Table 2.

Table 2: Comparison of Breast detection methods

Techniques	Entropy value
Input X-ray Images	3.5521
Neutrosophic Image	3.7521
Alpha mean operation	4.8575
Beta enhancement operation	5.3424

As per the tabulated results, the proposed approach shows better results in terms of DICE, BLE and accuracy. As the indeterminate pixel value decreases using neutrosophic domain, better the outcomes in mass detection.

Conclusions

Detection of mass regions both benign or tumor on early stage would be a great help for good survival rate and thus proposed system focusses on breast lesion detection. To refine a medical X-ray input image for mass detection, this work focuses on neutrosophic image sets where the indeterminacy is minimized and maximizing the truth value. It is done using alpha-mean and beta enhancement methods to promote a well-defined edge for better segmentation results. Then, the gamma clustering technique is used segmentation. The above results demonstrate that the proposed neutrosophic method consistently outperform the state-of-the-art methods.

In future work, the detected mass will be classified as benign or tumor based on the mass intensity. It also focuses on developing a graphical user interface automatic tool that allows doctors for diagnosis usage.

References

1. Sun, L.; Sun, H.; Wang, J.; Wu, S.; Zhao, Y.; Xu, Y. Breast Mass Detection in Mammography Based on Image Template Matching and CNN. *Sensors* 2021, *21*, 2855. <https://doi.org/10.3390/s21082855>
2. Siegel, R.L.; Miller, K.D.; Fuchs, H.E.; Jemal, A. Cancer Statistics, 2021. *CA Cancer J. Clin.* 2021, *71*, 7–30.
3. Al-Antari M.A.; Al-Masni M.A.; Choi M.-T.; Han S.-M.; Kim T.-S. A fully integrated computer-aided diagnosis system for digital X-ray mammograms via deep learning detection, segmentation, and classification *Int J Med Inform*, 117 (2018), pp. 44-54.
4. Ahmed L.; Iqbal M.M.; Aldabbas H.; Khalid S.; Saleem Y.; Saeed S. Images data practices for semantic segmentation of breast cancer using deep neural network, *J Ambient Intell Humaniz Comput* (2020), pp. 1-17
5. Dhungel N.; Carneiro G.; Bradley A.P., Deep structured learning for mass segmentation from mammograms, 2015 *IEEE international conference on image processing (ICIP)*, IEEE (2015), pp. 2950-2954
6. Kuo, Y.C.; Lin, W.C.; Hsu, S.C.; Cheng, A.C. Mass detection in digital mammograms system based on PSO algorithm. In Proceedings of the 2014 *International Symposium on Computer, Consumer and Control*, Taichung, Taiwan, 10–12 June 2014; pp. 662–668.
7. Amutha, S.; Babu, D.R.; Shankar, M.R.; Kumar, N.H. Mammographic image enhancement using modified mathematical morphology and Bi-orthogonal wavelet. In Proceedings of

- the 2011 *IEEE International Symposium on IT in Medicine and Education*, Guangzhou, China, 9–11 December 2011; Volume 1, pp. 548–553.
8. Liu, F.; Zhang, F.; Gong, Z.; Chen, Y.; Chai, W. A fully automated scheme for mass detection and segmentation in mammograms. In *Proceedings of the 2012 5th International Conference on BioMedical Engineering and Informatics*, Chongqing, China, 16–18 October 2012; pp. 140–144.
 9. F. Smarandache, A unifying field in logics: Neutrosophic logic, neutrosophy, neutrosophic set, neutrosophic probability (Fourth edition), *University of New Mexico*, 332, pp. 5–21, 2002.
 10. Ansari, Abdul Quaiyum & Biswas, Ranjit & Aggarwal, Swati. (2011). Proposal for Applicability of Neutrosophic Set Theory in Medical AI. *International Journal of Computer Applications* (0975 – 8887). 27. 5-11. 10.5120/3299-4505.
 11. F. Smarandache, A unifying field in logics: Neutrosophic logic, Multiple-Valued Logic, vol.8, 489–503. 1999.
 12. F. Smarandache, A unifying field in logics: Neutrosophic logic, neutrosophy, neutrosophic set, neutrosophic probability (Fourth edition), *University of New Mexico*, vol.332, 5–21, 2002.
 13. Kuei-Hu Chang, A novel risk ranking method based on the single valued neutrosophic set. *Journal of Industrial & Management Optimization*, vol. 18(3), pp. 2237-2253, 2022.
 14. J, Sofia Jennifer, and Sree Sharmila T. "Neutrosophic Approach of Segmentation on Thermal Images--Case Study: Drowsy Driving Application." *NeuroQuantology: An Interdisciplinary Journal of Neuroscience and Quantum Physics*, vol. 20, no. 8, 15 July 2022, pp. 2217-2237
 15. Sofia Jennifer J and Sree Sharmila T: Real time blink recognition from various head pose using single eye, *Multimedia Tools and Applications*, vol.2, pp 1-15, 2018.
 16. Sofia Jennifer J, Sree Sharmila T, A Neutrosophic Set Approach on Chest X-rays for Automatic Lung Infection Detection, *Journal of Information Technology and Control*, Vol. 52 No. 1(2023).
 17. Sofia Jennifer J, Sree Sharmila T, Neutrosophic Approach for Glaucoma Detection in Retinal Images, *Proceedings of The Romanian Academy, Series A*, Vol.4, (2022).

Received: June 20, 2024. Accepted: August 13, 2024