Uncertainty-infused Representation Learning Using Neutrosophic-based Transformer Network

Mohamed Abdel-Basset [iD]

Faculty of Computers and Informatics, Zagazig University, Zagazig, Sharqiyah, 44519, Egypt

Email: *mohamedbasset@ieee.org ; mohamedbasset@zu.edu.eg*

Abstract Representation learning and interactive modeling of visual content are critical aspects for advancing visual interpretation in different computer vision tasks. However, visual data's inherent uncertainty and ambiguity remain a critical challenge facing representation learning algorithms. In response to this challenge, this study presents a simple but effective model, namely, the neutrosophic-based transformer network (NTN), which integrate the theory of neutrosophic logic and transformer architecture to offer unprecedented challenge in managing uncertainties. The design of NTN includes three primary building blocks: neutrosophic encoding, multipath network, and fusion and decision modules. Motivated by the success of neutrosophic in interpreting indeterminacy involved in visual data, we introduce a neutrosophic encoding module that applies a convolving window to map image data into the neutrosophic domain (truth, indeterminacy, and falsehood). This helps the NTN mitigate spatial and intensity uncertainties present in image patches, thereby enhancing boundary and uniformity retention while minimizing discontinuities. Then, multipath networks are built with visual transformer encoding blocks (composed of multi-head self-attention, feed-forward network, and residual link) to take the responsibility of learning rich representations from the generated neutrosophic image. By the end of NTN, the multiplicative fusion module is presented to fuse diverse knowledge from different network paths to obtain insightful representation that can assist in making informed decisions about the input. A set of proof-of-concept experiments are conducted to evaluate the proposed NTN against cutting-edge approaches using two image recognition datasets (namely Fashion-MNIST and CIFAR-10) with different uncertainty settings, and the findings demonstrate the potential of NTN in maintaining high representation power through efficient modeling of uncertainty information within visual recognition tasks.

Keywords: Uncertainty; Neutrosophic set; Representation learning; Vision Transformer; Machine learning; Image Recognition.

1. Introduction

Representational learning has been recognized as a core concept in machine learning (ML) that usually refers to the ability of an algorithm/model to extract complex and meaningful features from the training or inference data. This motivates research communities to develop a wide variety of ML techniques that can have effective representational power for different tasks, including image identification, natural language processing, and audio analysis. With this representation power, the ML systems can decode intricate patterns and make informed decisions [1]. The quality of data being passed to the model contributes significantly to the ability of

machines to understand and engage with the real world more efficiently. Nevertheless, the real-world data is not usually clean but unpredictable and ambiguous limiting the representation power. This highlights the need for a new method to augment the representational learning capabilities with uncertainty handling capabilities to adequately handle the intricate and diverse nature of uncertainties seen in real data [2].

Vision data obtained from dynamic and unexpected surroundings inherently contain some level of uncertainty caused by a variety of factors, including occlusion, fluctuations in lighting conditions, changes in viewpoint, and deformations of the object [3]. In traditional ML methods, like probabilistic modeling, the range and likelihood of results are estimated to put measurable bounds on uncertainty in the data. However, the problem with these methods lies in their inability to handle situations encountering both ambiguities and contradictions simultaneously. In another way, the fuzzy logic was presented by Zadeh, to handle uncertainty by reconsidering system parameters as fuzzy numbers (instead of crisp values), a membership function characterized each. However, it may be insufficient to address more complex uncertainty categories such as epistemic or probabilistic uncertainty [4]. These constraints bring a critical need to explore other methodologies that can address the above limitation and thereby provide a more inclusive handling of uncertainties.

In 1995, Florentin Smarandache introduced the concept of Neutrosophic logic, in which each proposition is projected to have a level of truthfulness, a level of indeterminacy, and a level of falseness. With its triadic nature, the neutrosophic logic offers a great opportunity to deal with situations encountering uncertainties and contradictions, which substantially correlate with the complexities of visual recognition problems [5]. With the introduction of the idea of "neutrosophic membership", it could be easy to get a detailed measurement of the levels of truth, indeterminacy, and falsehood in a statement, offering a comprehensive depiction of different types of uncertainty [6]. The literature on computer vision has witnessed many breakthroughs in recent years due to the continuous evolution of deep image recognition models. Among these models, Vision transformers (ViTs) have been achieving remarkable success in capturing hierarchical patterns within images. ViTs were developed as customization of Transformer architecture, which was initially established for language modeling tasks, but with some key edits to suit image processing. One of the main distinctions about ViT lies in representing input images as a sequence of patches of equal size. These patches are passed to self-attention layers to enable the model to have long-range dependencies between them, as it grants the model to learn how the representations of distinctive fragments of an image contribute to a final decision. However, the ViT is not designed to deal with the inherent ambiguity and uncertainties in the input images, which limit their representation power and lead to poor classification performance [7]. This significant research gap highlighted the need for a novel solution that helps keep the

representational power of ViTs under serious and multidimensional uncertainty prevalent in visual data [8].

This paper presents a hybrid framework, called a neutrosophic-based transformer network (NTN), that integrates neutrosophic logic into the visual learning of ViTs, aiming to provide a consistent mechanism for modeling uncertainties inherent in visual recognition tasks. The proposed NTN is designed to introduce a neutrosophic encoding module to transform the noisy and ambiguous images into a neutrosophic domain before being fed into a representational layer within the attention layers of our model. Proof-of-concept experimentations are performed on two image recognition datasets (namely Fashion-MNIST and CIFAR-10) under different uncertainty settings, and the results demonstrated the effectiveness of our NTN in maintaining high representation power through effective modeling of diverse uncertainty patterns in visual recognition tasks.

2. Literature Review

In this section, we provide an overview of background literature and related works that emphasize the representation of learning and uncertainty modelling, which shape the foundations of our search.

In their book, Smarandache and Said [6] laid the groundwork for incorporating neutrosophic theories into different application domains (e.g., product acceptance determination). They also studied the concept of neutrosophic graphs by investigating their association with different machine-learning algorithms. They also discussed the use of neutrosophic representation to solve the systems of linear equations. Pamucar et al. [7] introduced an approach for evaluating and selecting suppliers based on fuzzy neutrosophic decision-making, in which a Dombi aggregator was used as a weight aggregator that applies pairwise comparison based on trapezoidal neutrosophic linguistic variables. Their approach used MABAC (multiattribute border approximation area comparison) tool to analyze the suppliers in a resilient supply chain management (RSCM) system with an uncertain environment and numerous factors; meanwhile, sensitivity analysis tests were applied to evaluate the model examined. This study proposes a novel fuzzy-neutrosophic-based approach for resilient supplier selection. Besides, Haq et al. [8] integrated entropy–MultiAtributive Ideal-Real Comparative Analysis (MAIRCA) Interval-Valued Neutrosophic Sets (IVNSs) in a unified framework to concurrently handle the subjective measures with vague or uncertain data and objective measures with crisp inputs. They used a wing-spar of a Human-Powered Aircraft (HPA) to prove the applicability of their framework, in which a committee of three subordinate experts makes decisions on the substitutes or criteria lingually via IVNSs. Then, the level of skill of an expert implicated in the decision-making process is evaluated utilizing a weighting methodology with verbal information. Finally, they used the MAIRCA method to assess materials for the HPA spar using the entropy weighting method.

Abdelhafeez et al. [9] studied the clustering of breast cancer through the application of a broadly adopted c-means algorithm and its improved versions: fuzzy c-means algorithm and neutrosophic c-means algorithm. They conducted an in-depth comparative study between these algorithms to analyze and interpret both the qualitative and quantitative efficiency metrics. In addition, Essameldin et al. [10] introduced a neutrosophic-based approach for opinion mining on Twitter, aiming to tackle perspectivism, its effects, and indeterminateness. For the first aim, they used Graphistry to conduct social network analysis (SNA), whereby a neural network was applied to impact weighting based on SNA-generated output and the public's responses to analyzed texts. Banerjee et al. [11] explored the diagnosis of melanoma lesions using a hybrid approach that integratess the triangular neutrosophic concept into the deep convolutional network, aiming to handle the uncertainty in dermoscopic and digital pictures and providing improved and reliable classification decision. Moreover, Karam and Ali [12] introduced a structural method for evaluating the performance of functioning onshore wind resources, respecting the three scopes of sustainability. Om their method, the energy manufacture, ecological validity, and practicability of onshore wind accommodations are all conditional on precise evaluations. They also discussed many criteria, including site accessibility, transmission, environmental impact, wind resources, permitting and regulatory requirements, turbine technology, and others. Lo et al. [14] developed an inclusive approach for the selection of strategic alliance partners, in which the neutrosophic ITARA technique was presented to create a group of criteria objective weights, and the neutrosophic TOPSIS method was used to regulate the performance and precedence of strategic alliance partners. They also used neutrosophic fuzzy logic to replicate the uncertainty in complicated problems inherent in realistic data from a multinational reactive factor manufacturing company. Furthermore, Singh [15] introduced a three-way n-valued neutrosophic concept lattice, illustrating the integration of neutrosophic frameworks into formal concept analysis. The author also introduced a mathematical mechanism to autonomously portray the n-valued neutrosophic according to its n-valued indeterminacy, n-valued falsity n-valued truth. His method presented a granular-centered computing model to find some n-valued neutrosophic theories, which offer many ways to transform the given n-valued neutrosophic context into binary context at the user's essential degree of granulation. Thanikachalam et al. [16] integrated the Interval Neutrosophic Set into a deep-learning model to assist in distinguishing the contaminated regions in the fundus image. Their model included three feature extractors: texture features, histogram features, and wavelet features, where the Optimal Deep Belief Network (ODBN) and Shuffled Shepherd Optimization (SSO) algorithm were used to make classification and optimize hyperparameters, respectively. From the above discussions, it can be noted that there is growing interest in the applicability of neutrosophic methods to handle uncertainty in different applications. Motivated by that, this work seeks to advance the understanding of uncertainty-infused representation learning through the novel integration of neutrosophic logic into ViTs.

3. Methodology

This section explains the methodology of designing the proposed NTN to give insight into the role of neutrosophic logic in empowering the representational learning capabilities with an effective ability to handle uncertainty during the image recognition process. This is achieved by taking advantage of ViTs, with the unique capacities of neutrosophic logic, to build a novel representation learning solution that explicitly accounts for uncertainties. In the following, the details of the structural design of the NTN are presented along with its main components, including the integrated neutrosophic logic, for learning informative representations from data.

3.1. Preliminaries

Derived from the concept of neutrosophic logic, the neutrosophic set was introduced as an extension of classical sets by enabling representation and dealing with indeterminacy exhibited in real-world data. In other words, the Neutrosophic Set provided a step toward representing incomplete and uncertain information, and it's composed of three essential components, namely truth membership $\omega_{\hat{A}_{\mathcal{N}}}(x)$, indeterminacy membership $\sigma_{\hat{A}_{\mathcal{N}}}(x)$, and falsity membership $\eta_{\hat{A}_{\mathcal{N}}}(x)$. mathematically speaking, given a neutrosophic set $\hat{A}_{\mathcal{N}} = \{(x; [\omega_{\hat{A}_{\mathcal{N}}}(x), \sigma_{\hat{A}_{\mathcal{N}}}(x), \eta_{\hat{A}_{\mathcal{N}}}(x)]) \vdots x \in X\}$, the corresponding three components $\omega_{\hat{A}_{\mathcal{N}}}(x): X \to]0^-, 1^+[$, $\sigma_{\hat{A}_{\mathcal{N}}}(x): X \to]0^-, 1^+[$, $\eta_{\hat{A}_{\mathcal{N}}}(x): X \to]0^-, 1^+[$ represents the degree to which an element x belongs to the set; the degree to which an element x is indeterminate in its membership status; and the degree to which an element x does not belong to the set, respectively [11-14]. In this context, the definition of superior sum of the above components is given as follows:

$$n_{sup} = sup\left(\omega_{\hat{A}_{\mathcal{N}}}(x)\right) + sup\left(\sigma_{\hat{A}_{\mathcal{N}}}(x)\right) + sup\left(\eta_{\hat{A}_{\mathcal{N}}}(x)\right) \in]^-0, 3^+[, \tag{1}$$

In the above expression, the term *Sup* symbolizes the supremum, or, in other words, the least upper bound. The definition of the superior sum of the Neutrosophic components is given as follows:

$$n_{inf} = inf\left(\omega_{\hat{A}_{\mathcal{N}}}(x)\right) + inf\left(\sigma_{\hat{A}_{\mathcal{N}}}(x)\right) + inf\left(\eta_{\hat{A}_{\mathcal{N}}}(x)\right) \in]^-0, 3^+[, \tag{2}$$

According to the above two expressions, the set is called an intuitionistic set in the case of $n_{sup} < 1$, and is referred to as a paraconsistent set in case of $n_{sup} > 1$.

The literature contains popular operations that can be applied to neutrosophic sets including union, intersection, and complement, which is defined using the following expressions for two neutrosophic sets $\hat{A}_{\mathcal{N}}$ and $\hat{B}_{\mathcal{N}}$.

$$\hat{A}_{\mathcal{N}} \cup \hat{B}_{\mathcal{N}} = \left\{x \mid \max\left(\omega_{\hat{A}_{\mathcal{N}}}(x), \omega_{\hat{B}_{\mathcal{N}}}(y)\right), \max\left(\sigma_{\hat{A}_{\mathcal{N}}}(x), \sigma_{\hat{B}_{\mathcal{N}}}(y)\right), \max\left(\eta_{\hat{A}_{\mathcal{N}}}(x), \eta_{\hat{B}_{\mathcal{N}}}(y)\right)\right\}. \tag{3}$$

$$\hat{A}_{\mathcal{N}} \cap \hat{B}_{\mathcal{N}} = \left\{x \mid \min\left(\omega_{\hat{A}_{\mathcal{N}}}(x), \omega_{\hat{B}_{\mathcal{N}}}(y)\right), \min\left(\sigma_{\hat{A}_{\mathcal{N}}}(x), \sigma_{\hat{B}_{\mathcal{N}}}(y)\right), \min\left(\eta_{\hat{A}_{\mathcal{N}}}(x), \eta_{\hat{B}_{\mathcal{N}}}(y)\right)\right\}. \tag{4}$$

$$\neg\hat{A}_{\mathcal{N}} = \left\{x \mid 1 - \omega_{\hat{A}_{\mathcal{N}}}(x), 1 - \sigma_{\hat{A}_{\mathcal{N}}}(x), 1 - \eta_{\hat{A}_{\mathcal{N}}}(x)\right\}. \tag{5}$$

Generally, neutrosophic sets can be leveraged in image processing tasks to represent composite and uncertain pixel information that classical set theory might not sufficiently tackle.

When the three components of neuromorphic sets satisfy the following $\omega_{\hat{A}_{\mathcal{N}}} \in [0,1]$, $\sigma_{\hat{A}_{\mathcal{N}}}(x) \in [0,1]$, $\eta_{\hat{A}_{\mathcal{N}}}(x) \in [0,1]$, then $\hat{A}_{\mathcal{N}}$ is known as single-valued neutrosophic (SVNS), which obeys the following constraints.

$$0 < \omega_{\hat{A}_{\mathcal{N}}}(x) + \sigma_{\hat{A}_{\mathcal{N}}}(x) + \eta_{\hat{A}_{\mathcal{N}}}(x) < 3, \tag{6}$$

Assume that there is SVNS denoted as $\hat{A}_{\mathcal{N}}$ in X, with triplet $\langle \omega_{\hat{A}_{\mathcal{N}}}(x), \sigma_{\hat{A}_{\mathcal{N}}}(x), \eta_{\hat{A}_{\mathcal{N}}}(x) \rangle$, then each element $x \in X$ if this set is called a singled-valued neutrosophic number (SVNN). To make things simpler, an SVNN c or e in SVNS can be written as $\langle \omega_c, \sigma_c, \eta_c \rangle$ or $\langle \omega_e, \sigma_e, \eta_e \rangle$, where each element has a unique value for its truth, indeterminacy, and falsity membership components [15-18]. The SVNNs are subject to many operations to manipulate their membership values, and these operations include Addition, Subtraction, Multiplication, Division, Power, aggregation, and variance, which are defined as follows:

$$c + e = \langle \omega_c + \omega_e, \sigma_c + \sigma_e, \eta_c + \eta_e \rangle, \tag{7}$$

$$c - e = \langle \omega_c - \omega_e, \sigma_c - \sigma_e, \eta_c - \eta_e \rangle. \tag{8}$$

$$c \cdot e = \langle \omega_c \cdot \omega_e, \sigma_c \cdot \sigma_e, \eta_c \cdot \eta_e \rangle. \tag{9}$$

$$c/e = \langle \omega_c / \omega_e, \sigma_c / \sigma_e, \eta_c / \eta_e \rangle \tag{10}$$

$$c^n = \langle \omega_c^n, \sigma_c^n, \eta_c^n \rangle \tag{11}$$

To aggregate SVNNs, we can define a weighted sum as follows:

$$X_{\text{aggregated}} = \sum_{i-1}^{n} w_i \cdot x_i = \left\langle \sum_{i-1}^{n} w_i \cdot \omega_{\hat{A}_{\mathcal{N}}}(x_i), \sum_{i-1}^{n} w_i \cdot \sigma_{\hat{A}_{\mathcal{N}}}(x_i), \sum_{i-1}^{n} w_i \cdot \eta_{\hat{A}_{\mathcal{N}}}(x_i) \right\rangle, \tag{12}$$

In addition, to quantify the distance between two SVNNs, we can use normalized hamming distance or Hausdorff distance, which can be expressed as follows:

$$d_{hamm}(c, e) = \frac{1}{3}\{|\omega_c - \omega_e| + |\sigma_c - \sigma_e| + |\eta_c - \eta_e|\}, \tag{13}$$

$$d_{haus}(c, e) = \max\{|\omega_c - \omega_e|, |\sigma_c - \sigma_e|, |\eta_c - \eta_e|\}, \tag{14}$$

3.2. Proposed Network

Herein, we debate and discuss the detailed architecture of the proposed NTN, which is composed of three main building modules: neutrosophic encoding, multipath network, and fusion and decision modules. These modules jointly enable the NTN to handle ambiguities and uncertainties throughout representation learning effectively. In the following subsection, we dive into details of the design of each module, its contribution to the representation learning process, and its relation to other modules in the NTN.
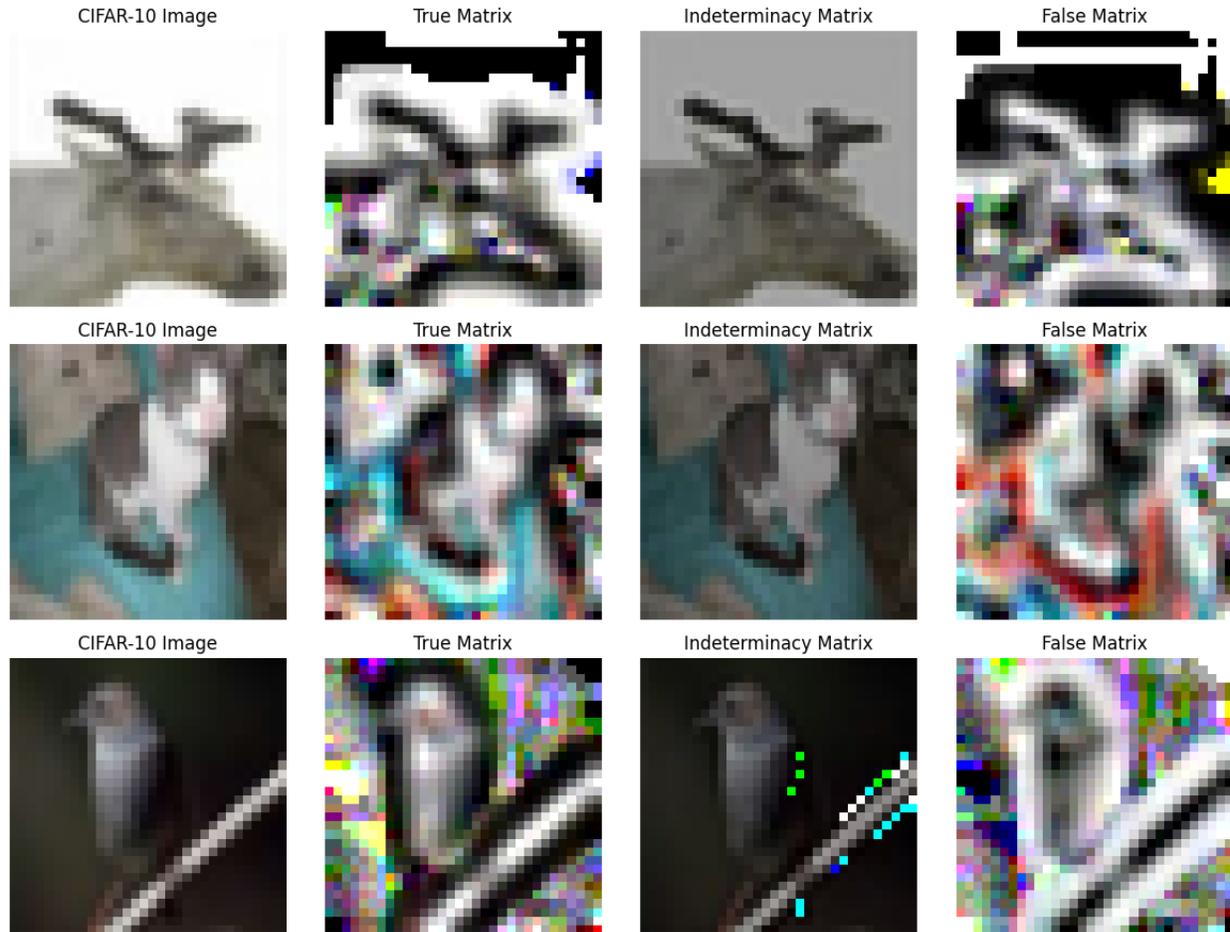
Figure 1. Visualization of the samples of CIFAR-10 data in the original and neutrosophic domains.

3.3. *Neutrosophic Encoding Module*

In the early phase of the proposed NTN framework, the input image is received and encoded be passed to subsequent feature extraction layers. Unlike transition ViT, in which the input images get patched and linearly projected before being passed to the network, the proposed encoding module takes the responsibility of transforming original pixel values into neutrosophic numbers. This encoding process is made before the inputs are projected to the feature extraction layer, leading to three types of maps corresponding to the components of neutrosophic. This mapping process, which is integral to encoding images in the neutrosophic domain, is tailored to the specific image processing application at hand. In mathematical terms, the pixel in location (i, j) of the input image can be denoted as $P(i, j) = \left(\omega_{\hat{A}_{\mathcal{N}}}(i, j), \sigma_{\hat{A}_{\mathcal{N}}}(i, j), \eta_{\hat{A}_{\mathcal{N}}}(i, j) \right)$, or for simplicity $P\left(\omega_{\hat{A}_{\mathcal{N}}}, \sigma_{\hat{A}_{\mathcal{N}}}, \eta_{\hat{A}_{\mathcal{N}}} \right)$. This representation conveys valuable information about the pixel's composition, specifying the percentage of its truth membership (white), indeterminacy membership (noise), and falsity membership (black). The calculation of these membership values $\omega_{\hat{A}_{\mathcal{N}}}$, $\sigma_{\hat{A}_{\mathcal{N}}}$, and $\eta_{\hat{A}_{\mathcal{N}}}$ are a serious characteristic of encoding images before they get patched. Motivated by the existing literature [4,5], the proposed encoding module computes the $\omega_{\hat{A}_{\mathcal{N}}}$ as

the likelihood of a pixel being white, $\sigma_{\hat{A}_N}$ as the degree of uncertainty or noise, and $\eta_{\hat{A}_N}$ as the likelihood of a pixel being black.

$$\omega_{\hat{A}_N}(i,j) = \frac{\overline{g}_{(i,j)} - \overline{g}_{min}}{\overline{g}_{max} - \overline{g}_{min}} \tag{15}$$

$$\sigma_{\hat{A}_N}(i,j) = \frac{\partial_{(i,j)} - \partial_{min}}{\partial_{max} - \partial_{min}} \tag{16}$$

$$\eta_{\hat{A}_N}(i,j) = 1 - \omega_{\hat{A}_N}(i,j) = \frac{\overline{g}_{max} - \overline{g}_{(i,j)}}{\overline{g}_{max} - \overline{g}_{min}} \tag{17}$$

In the above formula, $\overline{g}_{(i,j)}$ denotes the average concentration, calculated as follows:

$$\overline{g}_{(i,j)} = \frac{1}{p*p}\sum_{m=i-p/2}^{m=i+p/2} \sum_{n=j-p/2}^{n=j+p/2} g(m,n), \tag{18}$$

where p denotes the spatial dimension of the squared subwindow representing the influence of the adjacent points.

$$\overline{g}_{max} = Max(\overline{g}_{m(i,j)}), \overline{g}_{min} = Min(\overline{g}_{m(i,j)}) \tag{19}$$

$$\partial = |g_{(i,j)} - \overline{g}_{(i,j)}| \tag{20}$$

$$\partial_{max} = Max(\partial_{(i,j)}) \text{ and } \partial_{min} = Min(\partial_{(i,j)}) \tag{21}$$

These calculated values of $\omega_{\hat{A}_N}$, $\sigma_{\hat{A}_N}$, and $\eta_{\hat{A}_N}$ provide a nuanced description of each pixel's composition and offer insights into the extent of white, noise, and black contents within the pixel. Figure 1 illustrates the results of applying neutrosophic encoding to samples of CIFAR-10 data. This visualization provides an insightful view of how the encoding process translates the typical color images into the neutrosophic domain, allowing us to analyze representations in terms of truth, indeterminacy, and falsity. In addition to capturing the pixel values, the neutrosophic domain representation also helps model the inherent ambiguities and uncertainties in the images. This graphic depiction of the encoding output makes it easier to understand the encoding process and its potential advantages for managing uncertainties in visual data.

3.4. Multipath Network Module

Following the encoding process, the multipath network is presented as the next building module in the proposed NTN. This multipath structure is designated with transformer encoding blocks to take the responsibility of feature extraction and representation learning from the neutrosophic encoded inputs. To recap, the architecture of the transformer encoder is primarily composed of multihead self-attention (MHSA) mechanisms, feedforward neural networks (FNNs), and residual connections that jointly process the received patches of input to learn intricate data patterns, including those related to uncertainties. The MHSA was designed to learn global interdependencies between received sequences of patches [19]. Given neutrosophic encoded input sequence $X = (x_1, x_2, \dots, x_n)$, where n is the sequence length, the linear transformations are applied to encode sequence X into three matrices, namely the query matrix (Q), the key matrix (K), and the value matrix (V):

$$Q = X \cdot W_Q, K = X \cdot W_K, V = X \cdot W_V, \tag{22}$$

where W_Q, W_K, and W_V are learnable weight matrices. Then, a matrix of attention scores (A) is calculated based on the dot product of the query matrix with the transposed key matrix, followed by scaling, as shown as follows:

$$A = \text{softmax} \left(\frac{oK^T}{d_k} \right), \tag{23}$$

where d_k is the dimension of the key vectors [20]. The matrix A is later adopted to weight the value matrix, leading to self-attention output given as follows:

$$SA(Q, K, V) = A \cdot V. \tag{24}$$

The above computations are performed concurrently across different multiple heads to allow the TNT to focus on different parts of the neutrosophic encoded input sequence:

$$MHSA_{\text{multihead}} = \text{Concat}(SA_1, SA_2, \dots, SA_h) \cdot W_o, \tag{25}$$

where h is the number of attention heads. W_O is an additional learnable weight matrix for output transformation. Then, the output of different attention heads is passed to FNN with two linear transformations separated by a nonlinear activation function:

$$FFN_1 = MHSA_{\text{multihead}} \cdot W_1 + b_1. \tag{26}$$

Then, the activation function, typically rectified Linear Units ($ReLU$), is applied element-wise, as shown as follows:

$$FFN_2 = \text{ReLU}(FFN_1). \tag{27}$$

The result is further transformed with another set of weight matrices:

$$FFN_{\text{output}} = FFN_2 \cdot W_2 + b_2, \tag{28}$$

where W_1, W_2, b_1, and b_2 are learnable weight and bias matrices. The building of each network path in TNT is composed of a stack of three multiple transformer blocks, in which each block takes the output of the previous block as input, tolerating the network to capture progressively composite patterns and dependencies in the neutrosophic-encoded data.

3.5. *Fusion and Decision Modules*

By the end of NTN architecture, the fusion module is introduced as a pivotal building block for consolidating the various knowledge obtained from different paths in the previous network. We apply multiplicative fusion to combine the outcomes of the various representational learning paths, which aim to inform the model's final decision with uncertainty-infused representation. The choice of multiplicative fusion can be attributed to its ability to capture non-linear and complex interactions among features, empowering the model to emphasize relevant information by defeating inputs according to the learned weights. In mathematical terms, given the outputs $O_1, O_2, \dots, O_n$ from the respective paths, the multiplicative fusion can be described as follows:

$$F_{\text{multiplicative}} = \prod_{i=1}^{n} O_i, \tag{29}$$

where O_i represents the output of the ith path. Finally, the classification decision is made based on the cross-entropy given below:

$$L_{CCE} = -\frac{1}{N} \sum_{i=1}^{N} \sum_{c=1}^{C} \left(y_{actual}^{i,c} \right) \log \left(y_{model}^{i,c} \right). \tag{30}$$

In the above formula, N, C symbolize the number of samples and number of classes, respectively.

4. Experimental Setup

This section debates the empirical preparation and related setups for the proof-of-concept experiments conducted in this work. In other words, our discussion here covers data description, evaluation metrics, training hyperparameters, execution settings, and.

First, we used popular datasets (Fashion-MNIST, CIFAR-10) to train and evaluate NTN and competing methods. The former dataset encompasses 70,000 images ($size = 28 \times 28$) fitting to distinct classes, where 7,000 images exist per class. Among them, 60,000 images are used as a training set, while 10,000 images are used for testing. The second dataset contains 60,000 images ($size = 32 \times 32 \times 3$) distributed across 10 classes, where the training set is selected to contain only 50,000 images, and the test set contains 10,000 images. During the evaluation phase, the model performance is measured by displaying a confusion matrix, from which the following metrics are calculated:

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}}, \tag{31}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}}, \tag{32}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}}, \tag{33}$$

$$\text{F1} - \text{measure} = 2 * \frac{\text{Recall} \times \text{Precision}}{\text{Recall+Precision}}. \tag{34}$$

As an essential step for the reproducibility of our work, we chart out the implementation setup associated with our experiment. Also, the training hyper-parameters are presented to facilitate interpreting the different configurations taken into account throughout our experiments. Table 1 introduces a comprehensive detail of our implementations in terms of parameters and corresponding values.

Table 1. Summary of implementation setup for our experiments

Parameter	Value
Device	Dell Workstation
CPU	Intel(R) Core (TM) i5-3317U CPU @ 1.70–1.70 GHz
GPU	NVIDIA RTX 2060
RAM	32 GB
OS	Windows 10
Frameworks	TensorFlow 2.8.0, Sk-learn, Sci-Py, MatPlot
Batch Size	64
Learning Rate	0.001
Training Epochs	60

Given that we study uncertainty modeling, we apply Gaussian noise to the datasets in our experiments. This approach can be achieved in different scenarios according to the level of noise injected into the data. Table 2 summarizes the different scenarios adopted in our experiments, in

which the "Mean (μ)" column specifies the mean of the Gaussian noise, and the "Standard Deviation (σ)" column indicates the standard deviation of the noise.

Table 2: Noise setting summary

Scenario	Noise Level	Mean (μ)	Standard Deviation (σ)
Sc-0	Clean (No Noise)	0.00	0.00
Sc-1	Low Noise	0.10	0.05
Sc-2	Medium Noise	0.30	0.10
Sc-3	High Noise	0.50	0.15
Sc-4	Very High Noise	0.70	0.20

5. Results and Analysis

In this section, we provide a detailed explanation of the experimental results, offering an inclusive evaluation of the performance of the proposed model across various datasets. The results act as a lens by which we can obtain valuable insights about the interaction between the visual deep network and neutrosophic logic when modeling the representational patterns under high uncertainty settings.

To interpret the performance ability of the proposed model in modeling uncertainties in visual representational learning tasks, we conduct fair experimental comparisons to quantitively evaluate the performance of our model against the cutting-edge baselines. In these experiments, the data from scenarios Sc-3, Sc-4, and Sc-5 are used. Table 3 summarizes the numerical classification results on Fashion-MNIST data under different scenarios. Our model achieves remarkable improvements over competing models under different noise levels. Table 4 summarizes the numerical classification results on CIFAR-10 data under different scenarios. Remarkably, the classification performance of our model outperforms the competing models under different noise levels. As observed, the performance improvement becomes more significant in the case of high noise scenarios, further validating the advantage of our model for modeling uncertainties in noisy data.

Table 3. Comparison of quantitative results of different classifiers on Fashion-MNIST dataset.

	Sc-2			Sc-3			Sc-4		
Model	Accuracy	F1-score	AUC	Accuracy	F1-score	AUC	Accuracy	F1-score	AUC
SVM	94.19±2.49	94.15±2.42	97.11±0.05	91.28±5.52	92.88±5.82	95.46±0.1	88.47±5.29	87.34±2.64	93.1±4.77
LeNet	95.21±3.07	94.87±2.88	97.76±5.43	92.65±1.36	92.43±5.43	97.08±5.51	90.06±5.78	89.67±5.2	95.88±1.13
ResNet-18	96.88±3.44	96.74±3.24	98.99±2.54	93.20±1.17	95.29±3.00	98.03±1.38	90.78±3.71	90.12±1.8	94.69±0.36
ViT	97.02±3.54	96.64±5.13	98.26±1.23	93.61±2.83	93.56±3.25	96.97±3.79	90.93±5.29	90.54±1.05	94.84±1.9
CCT	97.31±2.1	96.87±2.45	98.74±3.37	95.13±1.48	94.89±3.14	97.53±1.34	91.87±1.56	91.45±2.37	96.92±2.13
Proposed	98.28±1.4	97.99±1.91	99.31±0.94	97.17±1.89	97.01±2.67	98.33±1.02	94.29±1.91	94.22±2.33	97.33±1.11

Table 4. Comparison of quantitative results of different classifiers on Fashion-MNIST dataset.

	Sc-2			Sc-3			Sc-4		
Model	Accuracy	F1-score	AUC	Accuracy	F1-score	AUC	Accuracy	F1-score	AUC
SVM	90.12±0.03	88.69±5.59	96.13±4.48	85.97±0.7	84.49±1.76	95.31±4.95	83.04±0.35	82.38±0.23	91.1±2.11
LeNet	92.01±5.16	91.55±5.51	97.08±0.94	87.02±5.52	85.49±4.08	94.18±4.15	82.32±5.81	81.25±1.66	90.04±3.97
ResNet-18	92.22±4.99	91.06±3.92	97.9±1.88	87.6±2.8	87.14±3.91	96.25±2.97	83.48±5.4	82.43±3.15	90.43±4.78
ViT	94.95±5.94	94.4±1.95	98.05±5.54	90.6±1.68	89.26±0.78	95.93±0.92	86.78±3.71	85.92±4.9	91.33±2.99
CCT	95.16±3.95	94.55±3.08	98.07±4.48	91.89±5.35	90.8±5.93	95.83±1.33	87.26±5.19	85.77±1.64	90.12±4.48
Proposed	96.41±1.22	96.33±2.22	98.44±1.87	94.64±2.27	94.21±1.69	97.66±2.16	95.41±2.42	95.39±2.21	97.18±1.02

Beyond the numerical improvements achieved by the proposed NTN in the above comparisons, the t-test is used to provide a comprehensive statistical significance analysis to ensure that the achieved improvements are not occurring by chance. This experimental analysis computes p-value statistics (refer to Table 5), which are compared with a suitable threshold (e.g., 0.05), that is typically determined according to a predefined confidence level (e.g., 95%). As notable in Table 5, the obtained p-values are below the significance threshold, which leads to drawing strong conclusions regarding the statistical significance of improvements achieved by the proposed NTN, demonstrating its competitive advantage.

Table 5: Statistical significance analysis results.

Model	Fashion-MNIST			CIFAR-10		
	Sc-2	Sc-3	Sc-4	Sc-2	Sc-3	Sc-4
Proposed vs SVM	8.86E-03	8.71E-03	1.25E-02	1.33E-02	5.27E-03	3.31E-04
Proposed vs LeNet	8.94E-03	1.34E-03	1.34E-04	1.57E-02	1.58E-03	1.33E-05
Proposed vs ResNet-18	1.68E-05	1.39E-04	8.17E-03	3.18E-03	1.37E-02	1.11E-06
Proposed vs ViT	7.01E-03	6.42E-06	6.66E-05	1.21E-04	7.17E-09	1.49E-08
Proposed vs CCT	1.12E-07	1.44E-08	1.03E-07	1.19E-08	1.50E-06	6.10E-04
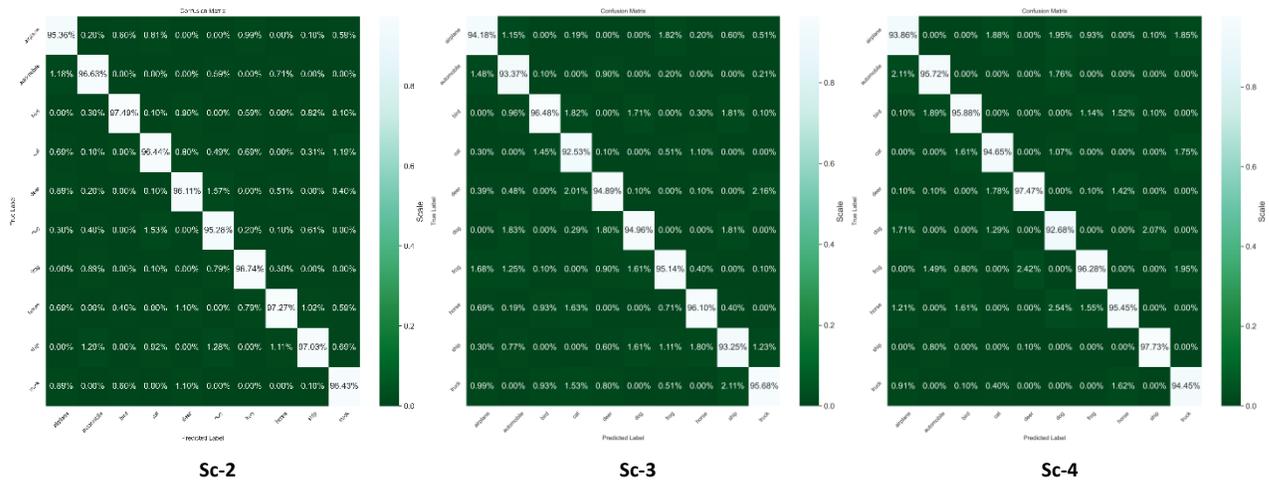
Figure 2. Visual illustration of confusion matrices for NTN under different uncertainty settings on Fashion-MNIST data.

In the previous experiments, the focus was given to the global view of model performance, but, to get a more detailed view of the class-level performance of the proposed NTN, we display its confusion matrix in both Figures 2-3. Figure 2 displays three confusion matrices corresponding to different uncertainty scenarios in the Fashion-MNIST dataset. It is worth noting that the average recognition accuracy of each class is almost similar with only 1%-2% variations, which reflect the representational power of NTN, demonstrating their consistent detection performance across various uncertainty levels. Moreover, in Figure 3, we display three confusion matrices corresponding to different uncertainty scenarios in the CIFAR-10 dataset. By observing the class-level precision, we can further drive critical insights on the consistent discrimination ability of NTN under RGB settings, which conform to our findings in grayscale scenario.
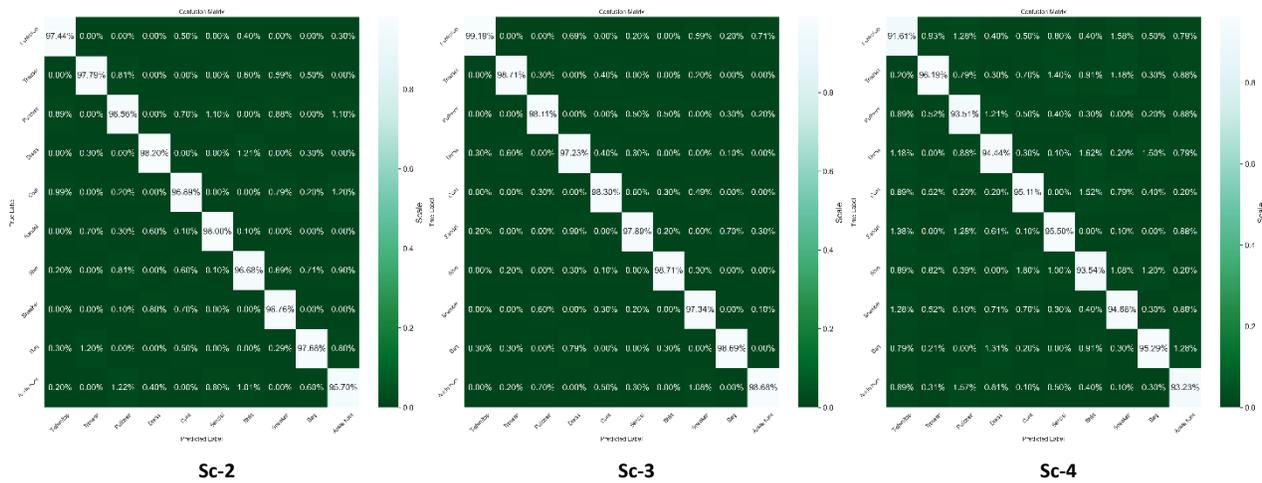
Figure 3. Visual illustration of confusion matrices for NTN under different uncertainty settings on CIFAR-10 data.
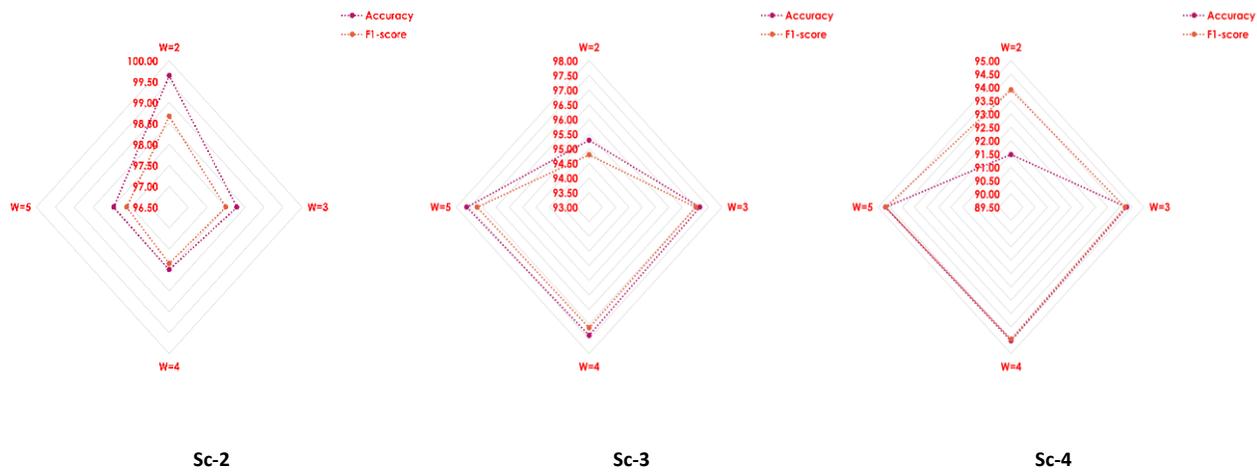
Figure 4. Ablation analysis for NTN under different uncertainty settings on Fashion-MNIST data.

Moreover, to capture the effectiveness of window size in the neutrosophic encoding module, a set of ablation experiments is performed to provide an in-depth evaluation of this hyperparameter. Different window sizes are independently used to assess their influences on the model's performance, where the corresponding numerical results on Fashion-MNIST are presented in Figure 4. As displayed, we can draw the conclusion that smaller window sizes work better on Sc-1, whereas larger window sizes are preferable in noisy scenarios (Sc-3 and Sc-4). This can be attributed to the fact that when the window size grows, more contextual information around each pixel is contemplated in the neutrosophic domain, making the image more tolerant to artifacts produced by noisy pixels in the near locality. Conversely, in Sc-2 data, we require only nearby details, and this effect may not exist, alleviating the need for such background locality knowledge.

6. Conclusion

This study presents an exploration of the impact of uncertainty on representational learning through developing a multi-path framework for modeling uncertainty in visual inputs during the process of feature extraction. A neutrosophic encoding is applied to map the image patches into triplet neutrosophic components, while visual transformer encoding is applied to extract insightful representation across different paths. The experimental comparisons on the Fashion-MNIST and CIFAR-10 datasets demonstrated the efficiency of the proposed NTN over the state-of-the-art methods even under scenarios with increased levels of uncertainty, showcasing its compliance and flexibility. The experimental findings demonstrate the promise of neutrosophic logic in revolutionizing the architectural design of the existing ML systems to be robust against ambiguity in different types of data, including time series, text, and graphs. Furthermore, the implications of this study can extend to support the explainability of ML decisions in complex data scenarios.

Conflict of interest
The authors declare that there is no conflict of interest in the research.
Funding
This research has no funding source.
Ethical approval
This article does not contain any studies with human participants or animals performed by any of the authors.

References

[1]. Ala, A., Simic, V., Pamucar, D., & Jana, C. (2023). A Novel Neutrosophic-Based Multi-Objective Grey Wolf Optimizer for Ensuring the Security and Resilience of Sustainable Energy: A Case Study of Belgium. *Sustainable Cities and Society*, 104709.

[2]. Ahmed, R., Nasiri, F., & Zayed, T. (2021). A novel neutrosophic-based machine learning approach for maintenance prioritization in healthcare facilities. *Journal of Building Engineering*, *42*, 102480.

[3]. Bhat, S., & Koundal, D. (2021). Multi-focus image fusion using neutrosophic based wavelet transform. *Applied Soft Computing*, *106*, 107307.

[4]. Koundal, D., & Sharma, B. (2019). Challenges and future directions in neutrosophic set-based medical image analysis. In *Neutrosophic Set in Medical Image Analysis* (pp. 313-343). Academic Press.

[5]. El-Ghareeb, H. A. (2020). Neutrosophic-based recommender for eLearning systems. In *Optimization Theory Based on Neutrosophic and Plithogenic Sets* (pp. 169-213). Academic Press.

[6]. Smarandache, F., & Said, B. (2020). Neutrosophic theories in communication, management and information technology.

[7]. Pamucar, D., Yazdani, M., Obradovic, R., Kumar, A., & Torres-Jiménez, M. (2020). A novel fuzzy hybrid neutrosophic decision-making approach for the resilient supplier selection problem. *International Journal of Intelligent Systems*, *35*(12), 1934-1986.

[8]. Haq, R. S. U., Saeed, M., Mateen, N., Siddiqui, F., & Ahmed, S. (2023). An interval-valued neutrosophic based MAIRCA method for sustainable material selection. *Engineering Applications of Artificial Intelligence*, *123*, 106177.

[9]. Abdelhafeez, A., Mohamed, H. K., Maher, A., & Abdelmonem, A. (2023). A Neutrosophic based C-Means Approach for Improving Breast Cancer Clustering Performance. *Neutrosophic Sets and Systems*, *53*(1), 19.

[10]. Essameldin, R., Ismail, A. A., & Darwish, S. M. (2022). Quantifying Opinion Strength: A Neutrosophic Inference System for Smart Sentiment Analysis of Social Media Network. *Applied Sciences*, *12*(15), 7697.

[11]. Banerjee, S., Singh, S. K., Chakraborty, A., Basu, S., Das, A., & Bag, R. (2021). Diagnosis of Melanoma Lesion Using Neutrosophic and Deep Learning. Traitement du Signal, 38(5).

[12]. Sallam, K. M., & Mohamed, A. W. (2023). Neutrosophic MCDM Methodology for Evaluation Onshore Wind for Electricity Generation and Sustainability Ecological. Neutrosophic Systems with Applications (NSWA), Vol. 4, 2023, 53.

[13]. Alenizi , J. and Alrashdi, I. (2023) "SFMR-SH: Secure Framework for Mitigating Ransomware Attacks in Smart Healthcare Using Blockchain Technology", *Sustainable Machine Intelligence Journal*, 2. doi: 10.61185/SMIJ.2023.22104.

[14]. Lo, H. W., Chang, D. S., & Huang, L. T. (2022). Sustainable strategic alliance partner selection using a neutrosophic-based decision-making model: a case study in passive component manufacturing. *Complexity*, *2022*.

[15]. Singh, P. K. (2018). Three-way n-valued neutrosophic concept lattice at different granulation. *International Journal of Machine Learning and Cybernetics*, *9*, 1839-1855.

[16]. Thanikachalam, V., Kavitha, M. G., & Sivamurugan, V. (2023). Diabetic Retinopathy Diagnosis Using Interval Neutrosophic Segmentation with Deep Learning Model. *Computer Systems Science & Engineering*, *44*(3).

[17]. A. El-Douh, A. (2023) "Assessment the Health Sustainability using Neutrosophic MCDM Methodology: Case Study COVID-19", Sustainable Machine Intelligence Journal, 3. doi: 10.61185/SMIJ.2023.33101.

[18]. Colhon, M., Vlăduţescu, Ş., & Negrea, X. (2017). How objective a neutral word is? A neutrosophic approach for the objectivity degrees of neutral words. *Symmetry*, *9*(11), 280.

[19]. Khalaf, O. I., Natarajan, R., Mahadev, N., Christodoss, P. R., Nainan, T., Romero, C. A. T., & Abdulsahib, G. M. (2022). Blinder Oaxaca and Wilk Neutrosophic Fuzzy Set-based IoT Sensor Communication for Remote Healthcare Analysis. *IEEE Access*.

[20]. Sleem, A. (2022) "Empowering Smart Farming with Machine Intelligence: An Approach for Plant Leaf Disease Recognition", Sustainable Machine Intelligence Journal, 1(1). doi: 10.61185/SMIJ.2022.1013.