

Article

Type-2 Neutrosophic Markov Chain Model for Subject-Independent Sign Language Recognition: A New Uncertainty-Aware Soft Sensor Paradigm

Muslem Al-Saidi ^{1,*}, Áron Ballagi ², Oday Ali Hassen ³ and Saad M. Saad ^{4,5} 

¹ Doctoral School of Multidisciplinary Engineering Sciences, Széchenyi István University, Egyetem tér 1, 9026 Győr, Hungary

² Department of Automation, Széchenyi István University, Egyetem tér 1, 9026 Győr, Hungary; ballagi@ga.sze.hu

³ Ministry of Education, Wasit Education Directorate, Kut 52001, Iraq; odayali@uowasit.edu.iq

⁴ Department of Information Technology, Institute of Graduate Studies and Research, Alexandria University, Alexandria 21526, Egypt; saad.darwish.pt@pua.edu.eg

⁵ Department of Artificial Intelligence, Faculty of Computer Sciences and Artificial Intelligence, Pharos University in Alexandria, Canal El Mahmoudia Street, Beside Green Plaza, Alexandria 21648, Egypt

* Correspondence: muslem@uowasit.edu.iq

Abstract: Uncertainty-aware soft sensors in sign language recognition (SLR) integrate methods to quantify and manage the uncertainty in their predictions. This is particularly crucial in SLR due to the variability in sign language gestures and differences in individual signing styles. Managing uncertainty allows the system to handle variations in signing styles, lighting conditions, and occlusions more effectively. While current techniques for handling uncertainty in SLR systems offer significant benefits in terms of improved accuracy and robustness, they also come with notable disadvantages. High computational complexity, data dependency, scalability issues, sensor and environmental limitations, and real-time constraints all pose significant hurdles. The aim of the work is to develop and evaluate a Type-2 Neutrosophic Hidden Markov Model (HMM) for SLR that leverages the advanced uncertainty handling capabilities of Type-2 neutrosophic sets. In the suggested soft sensor model, the Foot of Uncertainty (FOU) allows Type-2 Neutrosophic HMMs to represent uncertainty as intervals, capturing the range of possible values for truth, falsity, and indeterminacy. This is especially useful in SLR, where gestures can be ambiguous or imprecise. This enhances the model's ability to manage complex uncertainties in sign language gestures and mitigate issues related to model drift. The FOU provides a measure of confidence for each recognition result by indicating the range of uncertainty. By effectively addressing uncertainty and enhancing subject independence, the model can be integrated into real-life applications, improving interactions, learning, and accessibility for the hearing-impaired. Examples such as assistive devices, educational tools, and customer service automation highlight its transformative potential. The experimental evaluation demonstrates the superiority of the Type-2 Neutrosophic HMM over the Type-1 Neutrosophic HMM in terms of accuracy for SLR. Specifically, the Type-2 Neutrosophic HMM consistently outperforms its Type-1 counterpart across various test scenarios, achieving an average accuracy improvement of 10%.

Keywords: uncertainty-aware soft sensor; sign language recognition; Type-2 Neutrosophic reasoning; hidden Markov model



Citation: Al-Saidi, M.; Ballagi, Á.; Hassen, O.A.; Saad, S.M. Type-2 Neutrosophic Markov Chain Model for Subject-Independent Sign Language Recognition: A New Uncertainty-Aware Soft Sensor Paradigm. *Sensors* **2024**, *24*, 7828. <https://doi.org/10.3390/s24237828>

Academic Editor: Francesco Mercaldo

Received: 27 September 2024

Revised: 23 November 2024

Accepted: 5 December 2024

Published: 7 December 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Soft sensors, also known as virtual sensors or inferential sensors, are algorithms or models that estimate process variables that are difficult, expensive, or impossible to measure directly. They utilize easily measurable variables and historical data to infer the desired measurements, providing critical insights for process control and optimization [1]. Uncertainty-aware soft sensors incorporate methods to quantify and manage the uncertainty in their predictions. This

is essential because all models inherently have some level of uncertainty due to factors such as model approximation, measurement noise, and changes in process conditions. By being aware of these uncertainties, soft sensors can provide more reliable and robust predictions, enhancing decision-making processes in various industrial applications. The quality of input data from sensors or cameras can vary due to factors like lighting, background noise, and occlusions, affecting the accuracy and uncertainty estimation [2,3].

Subject-independent sign language recognition (SISLR) involves developing a system capable of recognizing sign language gestures from different individuals, without requiring the model to be specifically trained on each person's unique signing style. This contrasts with subject-dependent systems, which typically perform better because they are trained on data from the same individuals who will be using the system (see Figure 1) [4,5]. Uncertainty in SISLR is a significant challenge, as the system needs to accurately recognize gestures from a wide range of users with varying signing styles, physical characteristics, and environmental conditions. In SISLR, model uncertainty (epistemic uncertainty) occurs when the model encounters signing styles or features it has not seen before, whereas data uncertainty (Aleatoric uncertainty) comes from the inherent noise or ambiguity in the input data. In SISLR, this could be due to poor video quality, lighting variations, occlusions, or inconsistent signing by the user. By addressing model and data uncertainty through advanced modeling techniques, robust preprocessing, and adaptive learning, SISLR systems can become more reliable and effective in real-world applications [6].

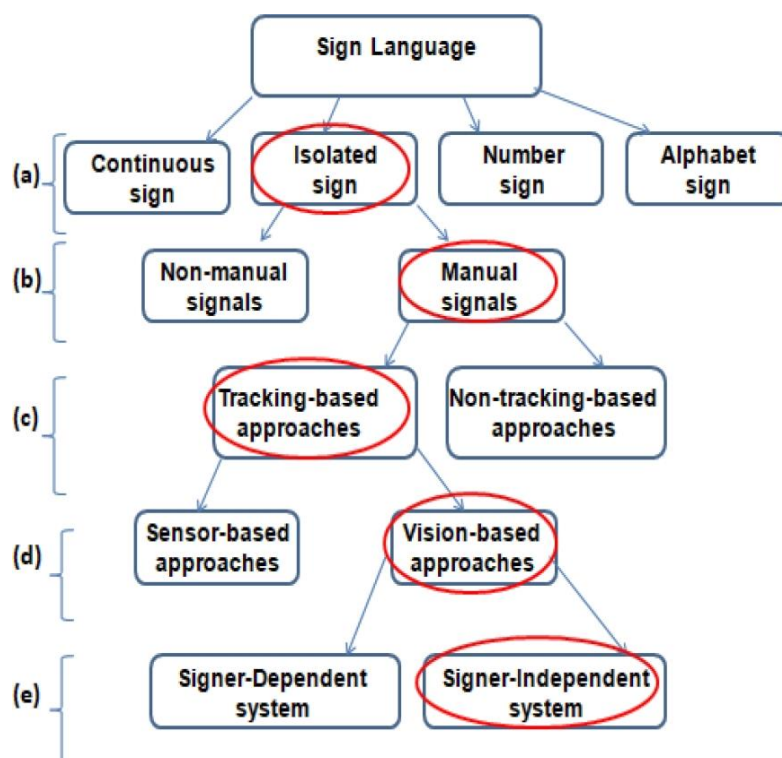


Figure 1. Study context: (a) sign categories, (b) sign characteristics, (c) sign detection techniques, (d) sign data acquisition approaches, and (e) sign classification conditions [4]. The red circles highlight the key focus areas of our study within the hierarchy of sign language recognition.

The high uncertainty associated with feature descriptions in SISLR can significantly increase the risk of misclassification. This uncertainty arises from the variability in signing styles, hand shapes, movements, and other non-manual cues across different individuals. Traditional methods, such as fuzzy logic, may struggle to accurately differentiate between types of signs under these conditions. Neutrosophic logic (NL) offers an enhanced approach by explicitly handling truth, indeterminacy, and falsity, making it particularly effective in managing the complexities of SISLR. By treating indeterminacy as a distinct dimension, NL

helps address the challenges of ambiguity and contradictory information, improving the system's ability to manage uncertain sign language inputs where the degree of truthfulness, indeterminacy, and falsity can vary widely [7,8].

Type-2 Neutrosophic Logic (T2NL) with the Footprint of Uncertainty (FOU) provides many advantages when uncertainty is frequent and has to be accurately modeled. FOU inclusion makes it possible to show uncertainty more fully by including not just the degree of membership but also the associated uncertainty [7]. Compared to Type-1 Neutrosophic Logic (T1NL), T2NL offers a stronger foundation for handling imprecision and uncertainty as T2NL simulates uncertainty at the level of the membership function as much as the membership values. This makes T2NL more sophisticated in its depiction of complex uncertain data by enabling it to record uncertainty about uncertainty. Higher degrees of uncertainty in the membership functions provide better representations of situations with ambiguous boundaries between several categories [8–11]. By representing truth, indeterminacy, and falsity as intervals, T2NL can more accurately capture the variability and vagueness inherent in uncertain information (see Figure 2). This interval representation acknowledges that the exact degree of membership is often not precisely known. The range of values (the Foot of Uncertainty) provides a buffer zone that accommodates fluctuations and uncertainties in the data, leading to more robust decision-making processes. The FOU helps to absorb and mitigate the impact of noise and variability in the input data, leading to more stable and reliable outputs. By analyzing the width of the intervals, it is possible to gain insights into the degree of uncertainty present in the system, facilitating more informed and precise analyses [10–12].

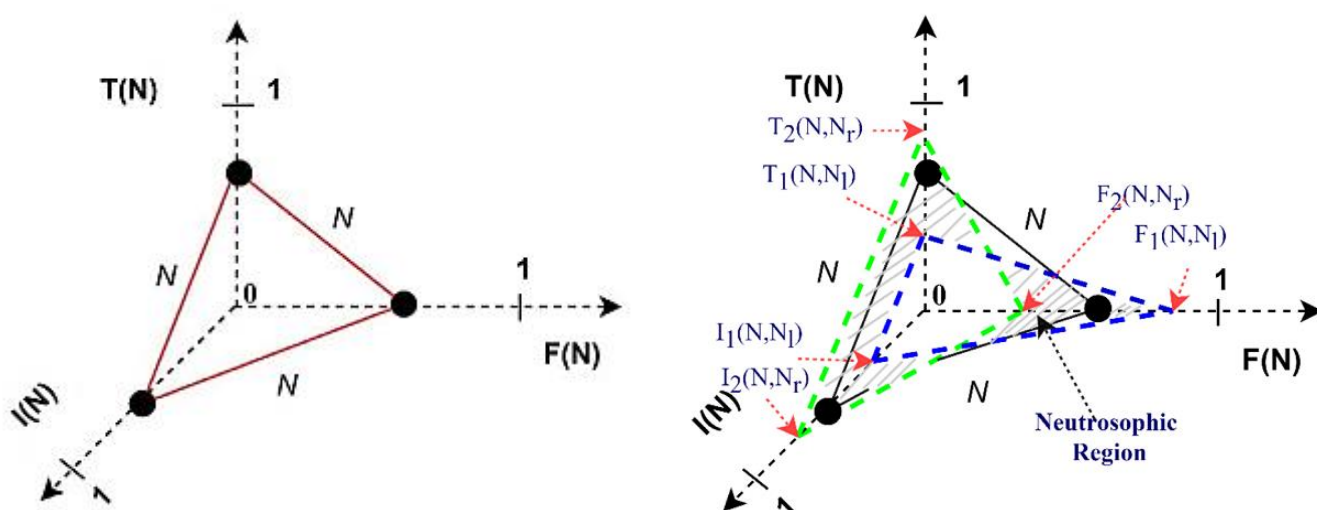


Figure 2. (Left) Type-1 neutrosophic truth $T(N)$, indeterminacy $I(N)$, and falsity membership functions $F(N)$. (Right) Type-2 neutrosophic set membership functions. The blurred region in T2NS provides two extra memberships for truth, indeterminacy, and falsity membership functions TM, IM, and FM. The two extra memberships are N_l and N_r ; l and r are left and right shifts [7].

Hidden Markov Models (HMMs) have long been a staple in sequential data analysis, including SISLR. While deep learning approaches have gained popularity due to their powerful pattern recognition capabilities, HMMs still offer several important advantages, particularly in certain contexts within SISLR [13–16]. HMMs are less computationally intensive compared to deep learning models, which require significant resources for training and inference. For real-time SISLR applications or when deploying on devices with limited computational power (e.g., mobile devices), HMMs can be more practical. In addition, HMMs can achieve reasonable performance with smaller datasets, whereas deep learning models often require large amounts of data to avoid overfitting and to generalize well. In situations where only a limited amount of sign language data is available, HMMs can be more effective, providing a good balance between performance and data requirements.

In addition, HMMs, due to their simpler structure and fewer parameters, are less prone to overfitting, especially in cases where the training data are limited or noisy. However, HMMs rely on fixed probability distributions (e.g., Gaussian) to model the likelihood of observations given a hidden state. These distributions are often simplistic and may not adequately capture the complex variability and uncertainty in real-world SLR data [17]. Currently, Type-2 Neutrosophic Logic offers enhanced capabilities for handling high levels of uncertainty that can address some of the problems faced by HMMs when using Type-1 Neutrosophic Logic [18].

1.1. Problem Statement and Motivation

SLR systems face significant challenges in accurately interpreting signs due to the inherent variability in gesture execution, signer-specific differences, and environmental factors such as lighting and background noise. These variations introduce considerable uncertainty, making it difficult for traditional SLR models to generalize effectively, particularly in subject-independent scenarios. Additionally, the current systems often lack a mechanism to explicitly quantify and manage this uncertainty, leading to reduced performance in real-world applications. The introduction of an Uncertainty-Aware Soft Sensor Paradigm for SLR, incorporating Type-2 Neutrosophic Logic, provides a promising solution by explicitly modeling the uncertainty present in gesture recognition. By accounting for different levels of truth, indeterminacy, and falsehood, this new soft sensor approach aims to improve the system's ability to deal with ambiguity, thereby enhancing the accuracy, reliability, and generalizability of SLR systems. This paradigm would be particularly valuable in assistive technologies, human–computer interaction, and accessibility tools for the deaf community.

1.2. Contribution

This paper introduces a new Uncertainty-Aware Soft Sensor Paradigm for SISLR based on a Type-2 Neutrosophic HMM, which employs the Footprint of Uncertainty (FOU) to model and manage the uncertainties inherent in gesture recognition. Unlike the conventional Type-1 Neutrosophic HMM, this approach captures deeper levels of uncertainty by considering not only truth but also indeterminacy and falsity within sign data. The use of the FOU allows the model to define dynamic boundaries for decision-making, enhancing its ability to deal with imprecise, ambiguous, and noisy input. This paradigm significantly improves subject-independent recognition, offering higher robustness and adaptability in real-world scenarios. It contributes to the field by providing a more accurate, scalable, and uncertainty-aware framework for SLR, addressing key challenges like diverse signing styles, environmental variations, and data inconsistency, making it highly applicable to assistive technologies.

The remainder of this paper consists of the following sections: Section 2 provides a literature review of relevant publications for the uncertainty-aware SISLR framework. The suggested Type-2 Neutrosophic HMM-based SISLR approach is presented in Section 3. The assessment of the suggested technique, including the results and discussion, is presented in Section 4. In Section 5, the study is concluded, and possible future directions are discussed.

2. Related Work

While current methods for uncertainty-aware SLR offer various ways to handle uncertainty, each comes with trade-offs [19–24]. Methods like Bayesian Neural Networks and Gaussian Processes provide excellent uncertainty quantification but at the cost of computational complexity and scalability. On the other hand, Type-1 Neutrosophic Logic and ensemble methods offer flexibility and robustness in noisy environments but may struggle with real-time processing. Emerging techniques like deep evidential regression are promising for real-time uncertainty estimation but need further exploration and optimization for widespread use in SLR.

Gaussian Processes (GPs) provide a probabilistic framework that quantifies uncertainty in SLR, making them effective for detecting ambiguous or novel gestures. Their non-parametric

nature allows them to adapt to various sign data distributions without assuming a specific form. Additionally, GPs offer inherent uncertainty estimates, enhancing the model's reliability for applications such as assistive technologies. However, GPs face scalability challenges with large datasets due to the quadratic growth of the covariance matrix, and they can be computationally intensive during training and inference. They also may struggle with high-dimensional data or complex sequences often encountered in SLR tasks [21].

Bayesian Neural Networks (BNNs) offer robust handling of uncertainty by quantifying both model (epistemic) and data (Aleatoric) uncertainties, making them adept at managing variations in signing styles, lighting conditions, and ambiguous gestures. They provide confidence intervals for predictions, which helps to avoid overconfidence and supports better decision-making in uncertain scenarios. Additionally, BNNs can reduce overfitting in smaller or imbalanced datasets. However, they come with high computational costs due to complex inference techniques like Monte Carlo sampling or variational inference, and their scalability can be problematic for large datasets. Furthermore, their emphasis on uncertainty estimation may sometimes result in lower raw accuracy compared to deterministic neural networks [23,24].

Ensemble methods, such as Bootstrap Aggregating, enhance uncertainty estimation by aggregating outputs from multiple models, which makes them effective at handling ambiguous sign inputs. They also help reduce overfitting, which is crucial for small or imbalanced datasets often seen in SLR. However, they come with increased resource demands, as training multiple models or performing multiple passes through a single model raises computational costs and can impact real-time performance. Moreover, tuning ensemble methods requires careful experimentation, and the performance gains may be marginal compared to well-optimized single models, raising concerns about the balance between performance improvements and resource use [23].

Deep evidential regression offers efficient uncertainty quantification in a single pass, making it suitable for real-time SLR without needing multiple forward passes or ensemble methods. It directly predicts uncertainty associated with each sign gesture, which improves reliability in noisy or ambiguous situations. However, the approach involves complex interpretation and tuning, requiring expertise in evidential theory and deep learning. Its adoption is still limited, and there is sparse information on its performance with large-scale or diverse datasets. Moreover, if not properly regularized, deep evidential regression models may be prone to overfitting, particularly in imbalanced or small datasets common in SLR [24,25].

Type-1 Neutrosophic Logic Models excel at handling uncertainty by representing truth, indeterminacy, and falsity, which helps to manage ambiguous or noisy sign inputs more effectively than traditional methods. They enhance decision-making by providing a framework to address unclear gestures and can be applied to various scenarios, including occlusion, lighting variations, and signer differences. However, these models have limitations, such as a restricted ability to capture complex, higher-order uncertainties involving multiple variables. They also introduce computational overhead, which can hinder real-time performance, and may not generalize well with highly variable data compared to more advanced probabilistic models [22].

The study presented in Ref. [26] introduces a bidirectional translation framework for Arabic Sign Language (ArSL) to improve communication between deaf and hearing individuals. The framework consists of two modules: one that translates Arabic sign images into text using various transfer learning models and another that converts text into sign language images. We developed a prototype using pre-trained convolutional neural networks (CNNs) such as DenseNet121, ResNet152, MobileNetV2, Xception, InceptionV3, NASNetLarge, VGG19, and VGG16. A novel fuzzy string matching score method was used to match input text with the corresponding sign language images. The dataset, featuring 7030 images across 14 classes, was collected locally from both deaf and hearing individuals. Experiments were conducted using this dataset and the CNN models.

In Ref. [27], the authors suggested an image-based method for sign language recognition using a four-phase framework. The first phase involves image processing techniques

to smooth images, with skin color detection and the Viola–Jones algorithm used to segment the face and hand. In the second phase, distance and similarity features are extracted using distance count and correlation coefficient methods. The third phase classifies and identifies signs using a Neuro-Fuzzy classification algorithm. The final phase employs Natural Language Processing (NLP) to display the recognized word. The framework, implemented in MATLAB R2022a, was tested on 20 videos (100 images) of family relation signs, achieving 95% accuracy. Table 1 compares between Type-1 Neutrosophic Logic (T1NL) and Uncertainty-Aware Deep Neural Networks (UADNNs) for sign language processing [22,28,29]. Also, Table 2 provides a comparison illustrating how the proposed T2NHMM for SLR distinguishes itself from other uncertainty-handling techniques, including traditional fuzzy logic, probability-based approaches, and Type-1 Neutrosophic Systems, particularly in terms of uncertainty representation, adaptability, and real-time applicability. Finally, Table 3 summarizes the different approaches used in SLR based on the methods discussed.

Table 1. A comparison between Type-1 Neutrosophic Logic and Uncertainty-Aware Deep Neural Networks.

Aspect	Type-1 Neutrosophic Logic (T1NL)	Uncertainty-Aware Deep Neural Networks (UADNNs)
Handling Uncertainty	Models truth, indeterminacy, and falsity, handling ambiguity effectively.	Quantifies uncertainty through probabilistic methods and confidence scores.
Complexity	Moderate complexity, involving degrees of truth, indeterminacy, and falsity.	High complexity due to advanced neural network architectures and uncertainty models.
Scalability	Can be less scalable to high-dimensional or complex datasets.	Scalable but may face challenges with very large datasets or high-dimensional data.
Computational Overhead	Generally lower computational overhead compared to deep learning models.	High computational cost due to deep learning model training and inference.
Real-time Performance	May struggle with real-time performance due to computational requirements.	Potentially suitable for real-time applications, but depends on model complexity.
Interpretability	Provides clear interpretability with uncertainty components.	Interpretation can be challenging, often requiring understanding of probabilistic outputs.
Accuracy with High Variability	May struggle with high variability in sign language data.	Generally performs well with high variability due to deep learning capabilities.
Flexibility	Limited adaptability to different scenarios without extensive adjustments.	Highly adaptable to various scenarios with advanced architectures and techniques.
Integration with Existing Systems	Can be integrated with traditional systems but may require specialized adaptation.	Easily integrable with existing deep learning frameworks and tools.
Suitability for Noisy Data	Effective in managing noisy data through uncertainty representation.	Effective in noisy environments with probabilistic models but requires robust training.

Table 2. A comparison between Type-2 Neutrosophic Logic for SLR with other uncertainty-handling techniques.

Feature	Type-2 Neutrosophic HMM	Traditional Fuzzy Logic	Probability-Based Approaches	Type-1 Neutrosophic Systems
Uncertainty Representation	Three dimensions: truth, falsity, and indeterminacy, using intervals (FOU) for second-order uncertainty	Single membership degrees, no higher-order uncertainty	Uses probability distributions or point estimates, no multi-dimensional uncertainty	Truth, falsity, and indeterminacy but no FOU, single-level uncertainty
Handling Ambiguity	Captures a range for each uncertainty level, managing ambiguous gestures with interval-based FOU	Limited flexibility in ambiguity; adjustments needed for uncertain cases	Probability-based confidence scores, struggles with highly ambiguous cases	Represents truth, falsity, and indeterminacy, but less adaptable than Type-2

Table 2. Cont.

Feature	Type-2 Neutrosophic HMM	Traditional Fuzzy Logic	Probability-Based Approaches	Type-1 Neutrosophic Systems
Adaptability to Variability	Adapts to diverse gestures, signing styles, lighting, occlusions, and background changes	Recalibration often required for diverse input	Sensitive to data quality; requires substantial labeled data for variability	Can handle variability but lacks interval adaptability seen in Type-2
Dependence on Data Volume	Lower dependency; interval-based uncertainty handling reduces need for extensive labeled data	Moderate dependency, especially for complex systems	High dependency on large datasets for reliable probability estimates	Moderate; less data-dependent than probability-based but more than Type-2
Resistance to Model Drift	High stability against model drift; adapts over time without frequent retraining through FOU	Low resistance to drift, requiring periodic recalibration	Susceptible to drift; requires frequent updates as patterns change	Moderate resistance to drift; needs tuning but more stable than probability-based
Real-Time Scalability	Computationally efficient for real-time applications due to interval-based FOU, reducing recalibration need	Computationally demanding in complex environments	High computational load with larger datasets and model complexity	Moderate scalability, less efficient than Type-2, often requires tuning
Complexity of Uncertainty Handling	Advanced, captures first- and second-order uncertainties with interval representation in FOU	Single-layer uncertainty representation	Simple probability estimates; lacks nuanced multi-level uncertainty handling	Handles multiple dimensions but lacks detailed interval flexibility in Type-2
Applicability to SLR	Suitable for high-variability cases in SLR, e.g., different styles, lighting, occlusions	Limited applicability to diverse environments in SLR	Useful for specific, well-defined patterns but struggles with SLR variability	Useful in SLR, but may need additional calibration for environmental changes

Table 3. Analysis of methods discussed for SLR: strengths and limitations.

Approach	Methodology	Pros	Cons
Gaussian Processes (GPs)	Non-parametric probabilistic framework for quantifying uncertainty in SLR.	- Effective in detecting ambiguous or novel gestures. - Adaptable to various sign data distributions. - Provides inherent uncertainty estimates.	- Scalability challenges with large datasets due to quadratic covariance matrix. - Computationally intensive during training and inference. - Struggles with high-dimensional data or complex sequences.
Bayesian Neural Networks (BNNs)	Neural networks that handle both epistemic and Aleatoric uncertainty, quantifying model and data uncertainty.	- Robust in handling uncertainty. - Provides confidence intervals for predictions. - Reduces overfitting in small datasets.	- High computational costs due to complex inference methods like Monte Carlo sampling. - Scalability challenges. - May have lower raw accuracy compared to deterministic models.
Ensemble Methods (Bootstrap Aggregating)	Combines multiple models to estimate uncertainty and reduce overfitting.	- Enhances uncertainty estimation. - Reduces overfitting. - Effective with ambiguous inputs.	- High computational cost due to training multiple models. - Resource demands can affect real-time performance. - Requires careful tuning, with marginal performance gains in some cases.
Deep Evidential Regression	Direct uncertainty prediction with deep learning, used for real-time SLR.	- Efficient uncertainty quantification in a single pass. - Suitable for real-time applications. - Improves reliability in noisy situations.	- Complex to interpret and tune. - Requires expertise in evidential theory. - Limited performance information for large-scale or diverse datasets. - Prone to overfitting without regularization.

Table 3. Cont.

Approach	Methodology	Pros	Cons
Type-1 Neutrosophic Logic Models	Models uncertainty using truth, indeterminacy, and falsity to handle ambiguous or noisy signs.	- Handles uncertainty effectively. - Enhances decision-making in unclear situations. - Suitable for occlusion, lighting variations, and signer differences.	- Limited ability to handle higher-order uncertainties. - Computational overhead may hinder real-time performance. - May not generalize well with highly variable data.
Bidirectional Translation Framework (Ref. [26])	Converts Arabic sign images into text and vice versa, using CNNs and fuzzy string matching for translation.	- Bridges communication gap between deaf and hearing individuals. - Uses pre-trained CNNs, enhancing efficiency. - Novel fuzzy string matching improves accuracy.	- Relies heavily on the quality of CNNs. - Limited generalization to other languages. - Performance depends on dataset quality.
Image-Based Method (Ref. [27])	Multi-phase method with image processing, feature extraction, classification, and NLP for SLR.	- Effective image processing techniques. - High accuracy in the classification of signs. - Uses NLP for word display.	- Depends on the quality of image processing and skin detection. - Limited to 20 videos for testing. - Framework implemented in MATLAB, which may limit flexibility for real-time applications.

2.1. Research Gap

The research gap between the current methods and high-level uncertainty models for SLR lies in the ability to effectively handle both epistemic (model) and Aleatoric (data) uncertainties while maintaining scalability, generalizability, and real-time performance. The current approaches, such as Gaussian Processes (GPs), Bayesian Neural Networks (BNNs), and ensemble methods, provide uncertainty quantification but struggle with computational efficiency, especially in large-scale or high-dimensional SLR datasets. These models often require complex inference techniques, making them impractical for real-time applications. On the other hand, uncertainty models like Type-1 Neutrosophic Logic and deep evidential regression offer some promise in handling noisy or ambiguous sign data but are limited in capturing higher-order uncertainties. The gap exists in developing models that not only quantify uncertainty effectively but also scale well across diverse signing styles, languages, and real-world conditions, ensuring that SLR systems are both robust and efficient.

The proposed T2NHMM for SLR fundamentally differs from traditional machine learning and deep learning approaches by focusing on explicitly modeling and managing uncertainty through a framework of truth, falsity, and indeterminacy, represented by the FOU. Unlike conventional approaches that focus on feature extraction and error minimization without directly addressing uncertainty, this model captures a range for each uncertainty level, enabling precise confidence in gesture recognition. Its interval-based ambiguity handling allows it to adapt to variations in signing styles, lighting, and occlusions, making it more flexible and less dependent on large labeled datasets. Additionally, T2NHMM inherently manages model drift, reducing the need for frequent retraining in dynamic environments. Despite requiring significant computation, its focus on uncertainty intervals offers a computationally efficient alternative to deep learning, enhancing real-time applicability in SLR with limited resources. Furthermore, T2NHMM for SLR differentiates itself from traditional fuzzy logic and probability-based methods by providing a more nuanced approach to uncertainty. While traditional fuzzy models rely on single membership degrees and probability-based methods on fixed confidence intervals, the Type-2 Neutrosophic HMM employs a three-part framework (truth, falsity, and indeterminacy) with the FOU to represent uncertainty as intervals, capturing both first- and second-order uncertainties.

2.2. The Need to Extend the Related Work

Type-1 Neutrosophic Hidden Markov Models (T1-NHMMs) face several challenges in sign language recognition (SLR). First, T1-NHMMs use single-valued membership func-

tions to represent truth, indeterminacy, and falsity, which limits their ability to capture complex, higher-order uncertainties present in sign language data. This can result in inadequate handling of nuanced variations in signing styles, lighting conditions, and ambiguous gestures. Additionally, the fixed nature of T1-NHMMs makes them less adaptable to diverse scenarios, leading to difficulties in managing the wide range of uncertainties found in real-world sign language data. Furthermore, T1-NHMMs may struggle with high-dimensional data, as their static uncertainty representations can fail to capture the intricate relationships between features in complex sign language sequences, potentially leading to reduced accuracy and performance.

Type-2 Neutrosophic Hidden Markov Models (T2-NHMMs) address these limitations by using Type-2 neutrosophic sets, which incorporate membership functions with a range of possible values rather than fixed ones. This allows T2-NHMMs to model uncertainties more comprehensively, capturing complex, higher-order uncertainties and interactions between multiple variables. The flexibility of Type-2 fuzzy logic enables T2-NHMMs to better represent varying degrees of truth, indeterminacy, and falsity, making them more adaptable to the diverse and dynamic nature of sign language data. Additionally, T2-NHMMs utilize the Footprint of Uncertainty (FOU), which provides a dynamic representation of uncertainty, allowing the model to effectively manage noisy and ambiguous data. This enhanced capability improves the model's robustness and accuracy, even in high-dimensional settings where traditional T1-NHMMs may fall short. Overall, T2-NHMMs offer a more sophisticated approach to handling the complexities of sign language recognition, leading to better performance and adaptability in real-world applications.

3. Methodology

The problem of subject-independent sign language recognition (SISLR) aims to build a system that accurately recognizes sign language gestures regardless of who performs the gesture. This problem is challenging due to the wide variability in individuals' hand shapes, sizes, movement styles, and speeds. Let $X \in \mathbb{R}^F$ represent a sequence of features extracted from the gesture image and F represent the number of features per gesture. $\mathcal{Y} \in \{1, 2, 3, \dots, C\}$ represents the gesture class label, where C is the total number of gesture classes. The task is to learn a function $f: X \in \mathbb{R}^F \rightarrow \{1, 2, 3, \dots, C\}$ such that for a new gesture X , the function $f(X)$ predicts the correct class label $\mathcal{Y}$ [4–6]. A key challenge in SISLR is the variability in signing styles. This introduces two types of uncertainty: model uncertainty (epistemic uncertainty) $u_m(X)$ arising from the model's lack of knowledge about new signers it has not seen before and data uncertainty (Aleatoric uncertainty) $u_d(X)$ due to inherent noise in the data (e.g., gesture image quality, lighting, occlusion). Thus, the model prediction is defined as

$$\hat{\mathcal{Y}} = \arg \max_{c \in \{1, 2, \dots, C\}} P(\mathcal{Y} = c|X) - (u_m(X) + u_d(X)), \quad (1)$$

$P(\mathcal{Y} = c|X)$ is the probability of the class label C given the input sequence X ; $u_m(X)$ and $u_d(X)$ reduce the confidence of the model based on uncertainties.

Let X consist of the following components (Kinematic Model): Hand positions: $P = (x, y, z)$, where x, y, z are the hand coordinates. Hand orientation: $O = (\theta, \varnothing)$, where $\theta, \varnothing$ are the angles representing hand orientation. Key joint positions: $J \in \mathbb{R}^{n \times 3}$, where n is the number of key joints tracked on the hand. Thus, the feature vector can be represented as $X = [P, O, J]$ for each gesture. Figure 3 depicts the American Sign Language (ASL) alphabet represented through hand gestures. Each letter of the alphabet, from A to Y, is demonstrated using distinct hand shapes and orientations, showcasing the manual alphabet system used in ASL. In addition, Table 4 lists sample of hand signs along with their position, orientation, and the positions of 5 key joints in 3D coordinates. In this table, the key joints could correspond to common points on the hand like the wrist, knuckles, and fingertips. The specific joints correspond to important anatomical points of the hand in ASL gestures.



Figure 3. Some English alphabet in American Sign Language.

Table 4. Three-dimensional kinematic data for samples of ASL alphabet hand signs.

Hand Sign	Position (P)	Orientation (O)	Key Joint 1 (Wrist)	Key Joint 2 (Knuckle)	Key Joint 3 (Middle Finger)	Key Joint 4 (Fingertip)	Key Joint 5 (Thumb Base)
A	(10, 15, 20)	(30°, 45°)	(10, 15, 20)	(11, 16, 19)	(13, 18, 18)	(14, 20, 17)	(9, 16, 21)
B	(12, 16, 18)	(20°, 60°)	(12, 16, 18)	(13, 18, 17)	(14, 19, 16)	(15, 20, 15)	(11, 17, 19)
C	(11, 14, 21)	(25°, 40°)	(11, 14, 21)	(13, 15, 22)	(14, 17, 23)	(15, 19, 22)	(10, 15, 20)
D	(9, 14, 19)	(35°, 55°)	(9, 14, 19)	(10, 16, 20)	(12, 17, 18)	(13, 18, 17)	(8, 16, 20)
E	(10, 13, 22)	(30°, 50°)	(10, 13, 22)	(11, 15, 23)	(13, 16, 24)	(14, 17, 25)	(9, 12, 21)
F	(11, 15, 19)	(40°, 45°)	(11, 15, 19)	(12, 16, 18)	(13, 17, 17)	(14, 18, 16)	(10, 14, 20)
G	(13, 17, 20)	(50°, 30°)	(13, 17, 20)	(14, 19, 22)	(15, 20, 23)	(16, 21, 22)	(12, 18, 19)
H	(12, 16, 18)	(45°, 40°)	(12, 16, 18)	(13, 18, 20)	(14, 19, 21)	(15, 20, 22)	(11, 17, 19)
I	(10, 14, 19)	(35°, 50°)	(10, 14, 19)	(11, 15, 21)	(12, 16, 22)	(13, 17, 23)	(9, 13, 20)
K	(11, 15, 18)	(30°, 45°)	(11, 15, 18)	(12, 17, 19)	(13, 18, 20)	(14, 19, 21)	(10, 16, 19)

In our case, MediaPipe, developed by Google <https://ai.google.dev/edge/mediapipe/solutions/guide> (1 June 2024), is utilized as it is a versatile framework designed for real-time hand tracking and 3D hand pose estimation from 2D images or video streams. It features a comprehensive hand model capable of detecting key joint positions and orientations in 3D space, making it ideal for applications like hand tracking and pose estimation. Known for its speed and high optimization, MediaPipe excels at providing both 2D and 3D keypoint detection, making it a powerful tool for analyzing hand gestures efficiently.

Given a training dataset $\mathcal{D} = \{(X_i, \mathcal{Y}_i)\}_{i=1}^N$, where N is the number of training samples, the objective is to train a model that maps input sequences to class labels by minimizing a loss function $L(\varphi)$ over model parameters φ . A common choice is the

categorical cross-entropy loss, where $1(\mathcal{Y}_i = c)$ is an indicator function that is 1 if the true label is c , and 0 otherwise.

$$L(\varphi) = \frac{1}{N} \sum_{i=1}^N \sum_{c=1}^C 1(\mathcal{Y}_i = c) \log P(\mathcal{Y}_i = c | X_i; \varphi), \quad (2)$$

To improve robustness, especially in the context of subject-independent recognition, uncertainty can be explicitly modeled as part of the loss function [21]:

$$L_{\text{uncertainty}}(\varphi) = L(\varphi) + \lambda \sum_{i=1}^N (u_m(X_i) + u_d(X_i)), \quad (3)$$

where λ is a regularization parameter that balances the contribution of uncertainty to the overall loss. To handle this, we propose an enhanced HMM-based SISLR framework where features are represented using Type-2 Neutrosophic Logic, which models uncertainty through truth (T), indeterminacy (I), and falsity (F) values. Let $X = \{X_1, X_2, \dots, X_T\}$ be the observed sequence of feature vectors over time T , $X_t \in \mathbb{R}^F$ and $S = \{s_1, s_2, \dots, s_k\}$ represent the set of hidden states of the HMM, where K is the number of hidden states. Each feature $X_{t,i}$ in the sequence is represented as a Type-2 neutrosophic set:

$$X_{t,i} = \langle T_{X_{t,i}}, I_{X_{t,i}}, F_{X_{t,i}} \rangle, \quad (4)$$

$T_{X_{t,i}}$ represents the truth membership value of feature i at time t , $I_{X_{t,i}}$ represents the indeterminacy membership value, and $F_{X_{t,i}}$ represents the falsity membership value.

In the HMM framework, we define the following parameters: (1) Transition probabilities $A = (a_{ij})$, where $a_{ij} = P(S_j | S_i)$ is the probability of transitioning from state S_i to state S_j . (2) Emission probabilities $B = (b_j(X_t))$, where $b_j(X_t) = P(X_t | S_j)$ is the probability of observing the feature vector X_t given the hidden state S_j . (3) Initial state probabilities $\pi = \{\pi_i\}$, where $\pi_i = P(S_i)$ is the probability of starting in state S_i [30]. For each hidden state S_j , the emission probability $b_j(X_t)$ is modeled using the neutrosophic membership functions. Given that each feature vector X_t is a Type-2 neutrosophic set, the emission probability is computed based on the truth, indeterminacy, and falsity membership functions for each feature at time t . The overall neutrosophic emission probability for a feature vector X_t at state S_j is defined as

$$b_j(X_t) = \frac{1}{F} \sum_{i=1}^F [w_T T_{X_{t,i}}(S_j) + w_I I_{X_{t,i}}(S_j) + w_F F_{X_{t,i}}(S_j)], \quad (5)$$

w_T, w_I, w_F are the weights associated with truth, indeterminacy, and falsity, respectively (these can be tuned based on the degree of importance of each component), and $T_{X_{t,i}}(S_j)$, $I_{X_{t,i}}(S_j)$, and $F_{X_{t,i}}(S_j)$ represent the truth, indeterminacy, and falsity values of feature i at state S_j . The likelihood of observing the sequence $X = \{X_1, X_2, \dots, X_T\}$ given an HMM $\lambda_y = (A_y, B_y, \pi_y)$ for a gesture class $\mathcal{Y}$, considering neutrosophic uncertainty, is computed using the neutrosophic forward algorithm:

$$P(X | \lambda_y) = \sum_{s_1, s_2, \dots, s_k} \pi_{y, s_1} \prod_{t=1}^T a_{y, s_{t-1}, s_t} b_{y, s_t}(X_t), \quad (6)$$

where the emission probabilities $b_{y, s_t}(X_t)$ are computed using Type-2 Neutrosophic Logic.

For classification, the observed sequence $X = \{X_1, X_2, \dots, X_T\}$ is compared to the model for each gesture class $\mathcal{Y}$. The neutrosophic similarity measure between the observed sequence and each model is computed as

$$S(X, \lambda_y) = \sum_{t=1}^T \sum_{i=1}^F \left[\frac{T_{X_{t,i}} \cap T_{R_{y,t,i}} + I_{X_{t,i}} \cap I_{R_{y,t,i}} + F_{X_{t,i}} \cap F_{R_{y,t,i}}}{T_{X_{t,i}} \cup T_{R_{y,t,i}} + I_{X_{t,i}} \cup I_{R_{y,t,i}} + F_{X_{t,i}} \cup F_{R_{y,t,i}}} \right] \quad (7)$$

$R_{\mathcal{Y},t,i}$ is the reference feature vector for gesture class $\mathcal{Y}$ at time t , and $\cap$ and $\cup$ represent the intersection and union of the FOU for the truth, indeterminacy, and falsity membership values. The recognized gesture $\hat{\mathcal{Y}}$ is the one with the highest neutrosophic similarity score:

$$\hat{\mathcal{Y}} = \arg \max_{\mathcal{Y} \in \{1,2,\dots,C\}} S(X, \lambda_{\mathcal{Y}}) \quad (8)$$

Training the HMM involves estimating the parameters $\lambda_{\mathcal{Y}} = (A_{\mathcal{Y}}, B_{\mathcal{Y}}, \pi_{\mathcal{Y}})$ for each gesture class $\mathcal{Y}$ using the Baum–Welch algorithm, an extension of the Expectation–Maximization (EM) algorithm for HMMs. During the training, the emission probabilities are computed using the neutrosophic membership functions, allowing the model to handle the inherent uncertainty in the observed features. See [18,22,31–33] for more details.

T2NL is an advanced framework for dealing with uncertainty, imprecision, vagueness, and inconsistency. This approach provides a more flexible and precise method for handling higher levels of uncertainty and complex information. A neutrosophic set is characterized by three functions: truth membership (T), indeterminacy membership (I), and falsity membership (F). In T2NL, these functions are not single values but ranges (intervals). Truth (T) represents the degree to which an element is true. In Type-2 Neutrosophic Logic, it is an interval $[T_L, T_U], 0 \leq T_L \leq T_U \leq 1$. Indeterminacy (I) represents the degree of indeterminacy or uncertainty. It is an interval $[I_L, I_U], 0 \leq I_L \leq I_U \leq 1$. Falsity (F) represents the degree to which an element is false. It is an interval $[F_L, F_U], 0 \leq F_L \leq F_U \leq 1$. A Type-2 neutrosophic set A in a universe of discourse X is defined as $A = \{\langle x, (F(x), I(x), F(x)) \rangle : x \in X\}$, $\forall x \in X, T(x) = [T_L(x), T_U(x)], I(x) = [I_L(x), I_U(x)], F(x) = [F_L(x), F_U(x)]$. In our case, a linear trapezoidal neutrosophic number, as illustrated in Figure 4, is used that is defined as [32]

$$\tilde{A}_{Neu} = (a, b, c, d; e, f, g, h; i, j, k, l) \quad (9)$$

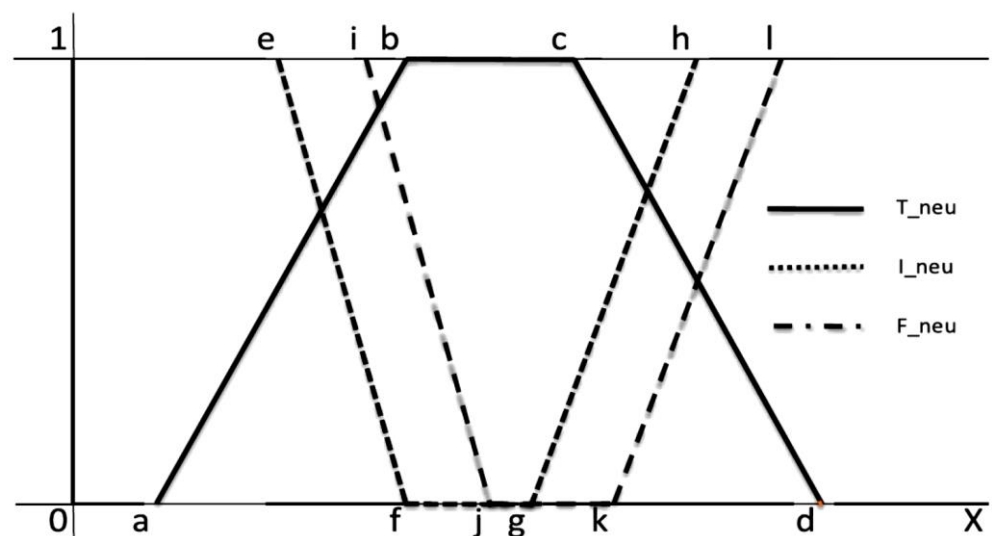


Figure 4. Linear trapezoidal neutrosophic membership function.

The truth, indeterminacy, and falsity membership is defined as [34]

$$\text{Truth membership function} = T_{Neu} = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & a \leq x < b \\ 1 & b \leq x < c \\ \frac{d-x}{d-c} & c \leq x \leq d \\ 0 & x > d \end{cases} \quad (10)$$

$$\text{Indeterminacy membership function} = I_{\text{Neu}} = \begin{cases} 1 & x < e \\ \frac{f-x}{f-e} & e \leq x < f \\ 0 & f \leq x < g \\ \frac{x-g}{h-g} & g \leq x \leq h \\ 1 & x > h \end{cases} \quad (11)$$

$$\text{Falsity membership function} = F_{\text{Neu}} = \begin{cases} 1 & x < i \\ \frac{j-x}{j-i} & i \leq x < j \\ 0 & j < x < k \\ \frac{x-k}{l-k} & k \leq x \leq l \\ 1 & x > l \end{cases} \quad (12)$$

where $0 \leq T_{\tilde{A}_{\text{Neu}}}(x) + I_{\tilde{A}_{\text{Neu}}}(x) + F_{\tilde{A}_{\text{Neu}}}(x) \leq 3, x \in \tilde{A}_{\text{Neu}}$. In general, linear trapezoidal functions are computationally simpler and more efficient to implement compared to more complex functions like Gaussian or other non-linear forms. The trapezoidal shape is defined by four parameters, making it straightforward to calculate the membership values, which is essential for high-performance systems that need to process gesture inputs continuously and respond without delay. Furthermore, with its two flat regions (low and high membership), it is effective at modeling indeterminate areas where uncertainty is high, offering more granularity in regions that are neither fully true nor false, aligning with the characteristics of sign language recognition systems [33]. Figure 5 depicts the graphical representation of FOU for every membership function T , I , and F .

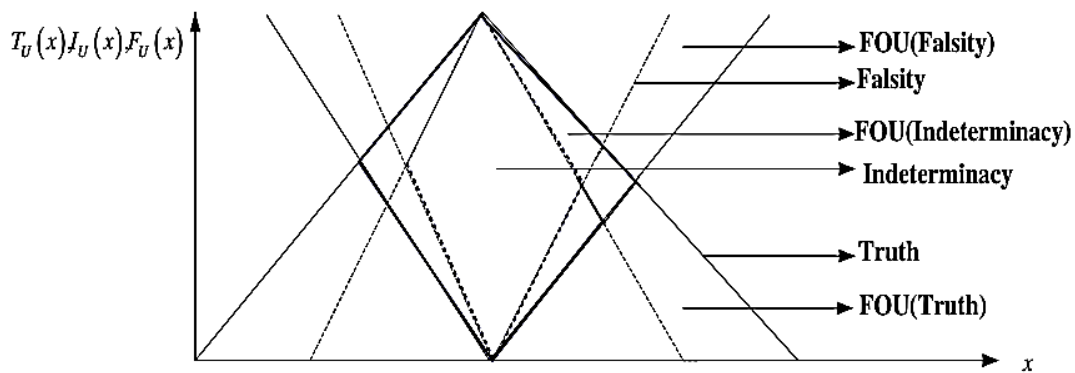


Figure 5. Graphical representation of Type-2 neutrosophic membership function [8].

In Type-2 Neutrosophic Logic, the FOU is a critical concept used to model uncertainty by capturing the range of possible values for each membership function. Each membership function is represented by a fuzzy set, and the FOU is defined for each as the region between its upper and lower membership functions. The FOU for each membership function forms a band-like region representing the uncertainty in the specific membership degree. The lower bound reflects conservative or minimal estimations (most certain values), while the upper bound represents optimistic or maximal estimations (least certain values). The width of the FOU indicates the degree of uncertainty, with wider FOUs signifying higher variability. The FOU for each membership function captures data variability and modeling uncertainty, enables flexible similarity measures by accounting for overlapping FOUs, and supports decision-making by balancing truth, indeterminacy, and falsity in the presence of real-world imprecision [8,9].

Membership functions map data points to degrees of truth, indicating how strongly a feature belongs to a fuzzy set. For instance, for modeling the “height” of a person, a trapezoidal membership function might have “Low”, “Medium”, and “High” regions, with overlap between them to handle ambiguity. A membership function in Type-2 Neutrosophic Logic quantifies the degrees of truth, indeterminacy, and falsity for an element, using ranges to capture uncertainty rather than fixed values. This method models variability and

ambiguity by defining possible values within a band-like region, using the FOU. When applied to hand sign features in sign language recognition, the membership function evaluates characteristics like finger angles, hand position, and motion trajectories. By linking these features to degrees of truth, indeterminacy, and falsity, the model accounts for variations in signing styles, incomplete gestures, or environmental noise, enabling more robust and adaptive recognition across diverse contexts.

Formally, a single-valued neutrosophic set (SVNS) $\check{S}$ on universal set U is characterized by truth membership function $\text{TMF}(\varnothing_{\check{S}})$, indeterminacy membership function $\text{IMF}(\psi_{\check{S}})$, and falsity membership function $\text{FMF}(\varphi_{\check{S}})$, respectively, in the following way [10,11]:

$$\check{S} = \{ \langle \check{\zeta}, (\varnothing_{\check{S}}(\check{\zeta}), \psi_{\check{S}}(\check{\zeta}), \varphi_{\check{S}}(\check{\zeta})) \rangle : \check{\zeta} \in U, \varnothing_{\check{S}}(\check{\zeta}), \psi_{\check{S}}(\check{\zeta}), \varphi_{\check{S}}(\check{\zeta}) \in [0, 1] \} \quad (13)$$

$$\text{Such that } 0 \leq \varnothing_{\check{S}}(\check{\zeta}), \psi_{\check{S}}(\check{\zeta}), \varphi_{\check{S}}(\check{\zeta}) \leq 3. \quad (14)$$

Let $\check{S}(\check{\zeta}) = [\check{S}^U(\check{\zeta}), \check{S}^L(\check{\zeta})]$ be an interval Type-2 neutrosophic set (IT2NS) on universal set U , where $\check{\zeta} \in U$ and $\check{S}^U : U \rightarrow [0, 1]$ and $\check{S}^L : U \rightarrow [0, 1]$ are two Type-1 neutrosophic sets (T1NSs) known as upper and lower neutrosophic sets, respectively, having the condition $0 \leq \check{S}^L(\check{\zeta}) \leq \check{S}^U(\check{\zeta}) \leq 1$ defined as follows [19]:

$$\check{S} = \{ \langle \check{\zeta}, ([\varnothing_{\check{S}}^U(\check{\zeta}), \varnothing_{\check{S}}^L(\check{\zeta})], [\psi_{\check{S}}^U(\check{\zeta}), \psi_{\check{S}}^L(\check{\zeta})], [\varphi_{\check{S}}^U(\check{\zeta}), \varphi_{\check{S}}^L(\check{\zeta})]) \rangle : \check{\zeta} \in U \} \quad (15)$$

$$[\varnothing_{\check{S}}^U(\check{\zeta}), \varnothing_{\check{S}}^L(\check{\zeta})], [\psi_{\check{S}}^U(\check{\zeta}), \psi_{\check{S}}^L(\check{\zeta})], [\varphi_{\check{S}}^U(\check{\zeta}), \varphi_{\check{S}}^L(\check{\zeta})] \in [0, 1] \quad (16)$$

$$\varnothing_{\check{S}}(\check{\zeta}) = \begin{cases} \frac{(\check{\zeta} - \check{S}_1)\varnothing_{\check{S}}}{\check{S}_2 - \check{S}_1} & \check{S}_1 \leq \check{\zeta} \leq \check{S}_2 \\ \varnothing_{\check{S}} & \check{S}_2 \leq \check{\zeta} \leq \check{S}_3 \\ \frac{(\check{S}_4 - \check{\zeta})\varnothing_{\check{S}}}{\check{S}_4 - \check{S}_3} & \check{S}_3 \leq \check{\zeta} \leq \check{S}_4 \\ 0 & \text{otherwise} \end{cases} \quad (17)$$

$$\psi_{\check{S}}(\check{\zeta}) = \begin{cases} \frac{\check{S}_2 - \check{\zeta} + (\check{\zeta} - \check{S}_1)\psi_{\check{S}}}{\check{S}_2 - \check{S}_1} & \check{S}_1 \leq \check{\zeta} \leq \check{S}_2 \\ \psi_{\check{S}} & \check{S}_2 \leq \check{\zeta} \leq \check{S}_3 \\ \frac{\check{\zeta} - \check{S}_3 + (\check{S}_4 - \check{\zeta})\psi_{\check{S}}}{\check{S}_4 - \check{S}_3} & \check{S}_3 \leq \check{\zeta} \leq \check{S}_4 \\ 1 & \text{otherwise} \end{cases} \quad (18)$$

$$\varphi_{\check{S}}(\check{\zeta}) = \begin{cases} \frac{\check{S}_2 - \check{\zeta} + (\check{\zeta} - \check{S}_1)\varphi_{\check{S}}}{\check{S}_2 - \check{S}_1} & \check{S}_1 \leq \check{\zeta} \leq \check{S}_2 \\ \varphi_{\check{S}} & \check{S}_2 \leq \check{\zeta} \leq \check{S}_3 \\ \frac{\check{\zeta} - \check{S}_3 + (\check{S}_4 - \check{\zeta})\varphi_{\check{S}}}{\check{S}_4 - \check{S}_3} & \check{S}_3 \leq \check{\zeta} \leq \check{S}_4 \\ 1 & \text{otherwise} \end{cases} \quad (19)$$

where $\varnothing_{\check{S}} = [\varnothing_{\check{S}}^U, \varnothing_{\check{S}}^L]$, $\psi_{\check{S}} = [\psi_{\check{S}}^U, \psi_{\check{S}}^L]$ and $\varphi_{\check{S}} = [\varphi_{\check{S}}^U, \varphi_{\check{S}}^L]$ are interval Type-2 neutrosophic numbers (IT2NNs). The number $\check{S}$ can be represented as (see Figure 6)

$$\check{S} = [\check{S}^U, \check{S}^L] = [(\check{S}_1^U, \check{S}_2^U, \check{S}_3^U, \check{S}_4^U; \varnothing_{\check{S}}^U, \psi_{\check{S}}^U, \varphi_{\check{S}}^U), (\check{S}_1^L, \check{S}_2^L, \check{S}_3^L, \check{S}_4^L; \varnothing_{\check{S}}^L, \psi_{\check{S}}^L, \varphi_{\check{S}}^L)] \quad (20)$$

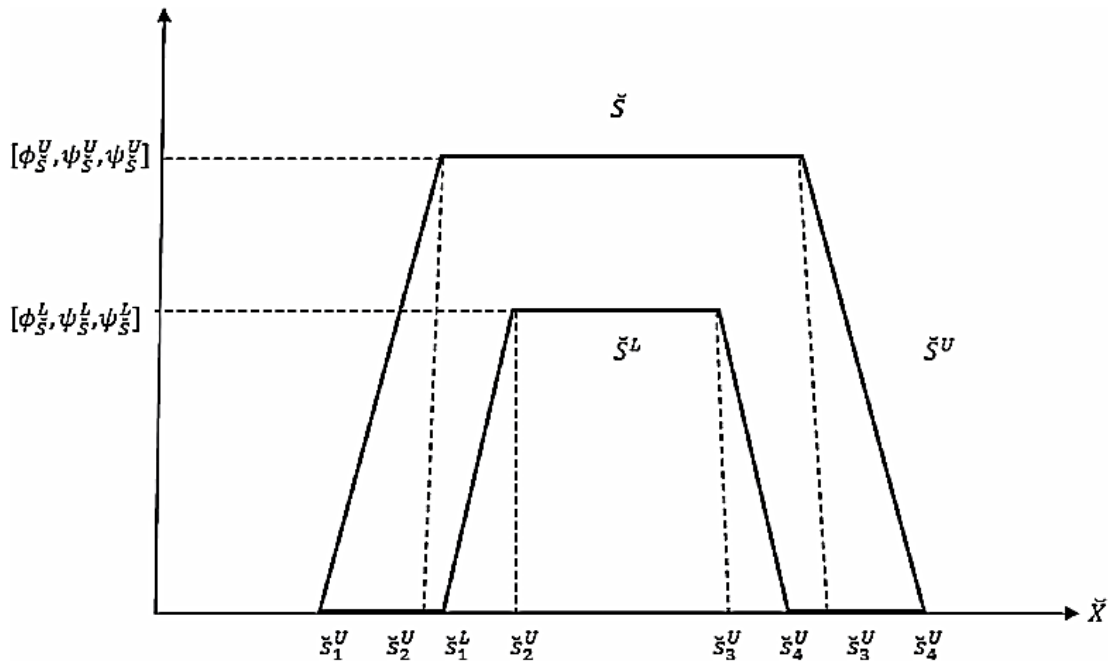


Figure 6. An interval Type-2 trapezoidal neutrosophic number [8].

This is called interval Type-2 trapezoidal neutrosophic logic number (IT2TrNN), where

$$0 \leq \check{s}_1^U \leq \check{s}_2^U \leq \check{s}_3^U \leq \check{s}_4^U \leq 1, 0 \leq \check{s}_1^L \leq \check{s}_2^L \leq \check{s}_3^L \leq \check{s}_4^L \leq 1 \quad (21)$$

$$0 \leq \phi_S^L \leq \psi_S^U \leq 1, 0 \leq \psi_S^L \leq \psi_S^U \leq 1, 0 \leq \phi_S^L \leq \phi_S^U \leq 1 \quad (22)$$

3.1. Step-by-Step Approach to Integrating Type-2 Neutrosophic Logic in HMM

- **Neutrosophic Feature Extraction:** Each gesture's visual or motion features are transformed into Type-2 neutrosophic sets, representing truth (T), indeterminacy (I), and falsity (F), with a degree of uncertainty each defined by upper and lower bounds. These features are then passed to the HMM as part of the observation data. For observation O_t , we represent it as

$$N(O_t) = \langle \langle T_{lower}(O_t), T_{upper}(O_t) \rangle, \langle I_{lower}(O_t), I_{upper}(O_t) \rangle, \langle F_{lower}(O_t), F_{upper}(O_t) \rangle \rangle \quad (23)$$

$T(O_t)$, $I(O_t)$ and $F(O_t)$ capture the truth, indeterminacy, and falsity of observation O_t , and $T_{lower}(O_t)$ and $T_{upper}(O_t)$ represent the lower and upper bounds for truth membership (similarly for indeterminacy and falsity).

- **Observation Probability in HMM (Adjusted with Neutrosophic Sets):** The observation probabilities in HMM, which traditionally map features to states, are modified to include the neutrosophic membership functions for truth, indeterminacy, and falsity, providing more flexible modeling of noisy or uncertain gestures. In a standard HMM, the observation probability $B_j(O_t)$ at state j is the likelihood of observing feature O_t in that state. With the neutrosophic framework, the observation probability is adjusted to include neutrosophic membership values:

$$B_j(O_t) = \alpha \times T(O_t) + \beta \times I(O_t) + \gamma \times F(O_t) \quad (24)$$

α , β , and γ are weight parameters that balance the contribution of truth, indeterminacy, and falsity to the observation probability. This formula adjusts the likelihood of each observation based on its degree of certainty or uncertainty as captured by the neutrosophic sets.

- **Forward Probability Using Neutrosophic HMM:** The forward algorithm calculates the probability of observing a sequence up to time t and being in state j (forward probability $\alpha_t(j)$). For

the neutrosophic HMM, the forward probability is updated using the neutrosophic-adjusted observation probabilities:

$$\alpha_t(j) = \left(\sum_{i=1}^N \alpha_{t-1}(i) A_{ij} \right) \times B_j(O_t) \quad (25)$$

$\alpha_t(j)$ is the forward probability at time t for state j , A_{ij} is the transition probability from state i to state j , and $B_j(O_t)$ is the neutrosophic-adjusted observation probability for state j at time t . The forward algorithm now integrates uncertainty into the observation step through the Type-2 neutrosophic sets.

- Backward Probability with Neutrosophic HMM: Similarly, the backward probability $\beta_t(j)$, which represents the probability of observing the sequence from time $t + 1$ to the end given state i at time t , is also updated:

$$\beta_t(j) = \left(\sum_{i=1}^N A_{ij} \times B_j(O_{t+1}) \right) \times \beta_{t+1}(j) \quad (26)$$

$B_j(O_{t+1})$ is again adjusted by the neutrosophic membership functions.

- Baum–Welch algorithm for Neutrosophic HMM Training: The Baum–Welch algorithm, used for training HMMs, can also be adapted to the neutrosophic HMM. In each iteration, the algorithm updates the model parameters (transition probabilities A and observation probabilities B) to maximize the likelihood of the observation sequence. The new E-step (expectation step) uses the neutrosophic-adjusted forward and backward probabilities, while the M-step (maximization step) remains the same but operates with neutrosophic-adjusted observation probabilities.
 - E-step (Expectation): Compute the neutrosophic forward and backward probabilities using Equation (24).
 - M-step (Maximization): Update the transition and emission probabilities using the neutrosophic-enhanced likelihood estimates.
- Final Recognition and Decision-Making: Once the HMM is trained with neutrosophic observations, the final step is to perform gesture recognition. For a given observation sequence, the HMM will compute the most likely state sequence (corresponding to gestures) using the Viterbi algorithm, which has been modified to incorporate neutrosophic-adjusted probabilities. The recognition decision will be influenced by both the temporal sequence (modeled by HMM) and the uncertainty handling (modeled by Type-2 Neutrosophic Logic), ensuring a more robust and accurate interpretation of sign language gestures.

3.2. Algorithm: Type-2 Neutrosophic HMM for Hand Recognition

Here is an algorithm for a Type-2 Neutrosophic Hidden Markov Model (HMM)-based hand recognition system that integrates Type-2 Neutrosophic Logic for managing uncertainty.

Input:

- Observed sequence of hand gesture features $O = (o_1, o_2, \dots, o_T)$
- State transition probabilities $A = \{a_{ij}\}$, modeled as Type-2 neutrosophic sets.
- Observation probabilities $B = \{b_j(o)\}$, modeled as Type-2 neutrosophic sets.
- Initial state distribution $\pi = \{\pi_i\}$, modeled as Type-2 neutrosophic sets.

Output:

- Recognized hand gestures with an associated probability (truth, indeterminacy, and falsity values).

Step 1: Initialization

1. Define Model Parameters:

- Define the number of hidden states N and the length of the observation sequence T .
- Initialize the state transition matrix $A = [a_{ij}]$, the observation likelihood matrix $B = [b_j(o)]$, and the initial distribution $\pi = \{\pi_i\}$, all as Type-2 neutrosophic sets with membership functions for truth μ_T , indeterminacy μ_I , and falsity μ_F .

$$a_{ij}(t) = FOU_{A_{ij}}(t) = \left[\pi_{A_{ij}}^{low}(t), \pi_{A_{ij}}^{High}(t) \right], \quad (27)$$

$$b_i(o_t) = FOU_{B_i}(o_t) = [\pi_{B_i}^{low}(o_t), \pi_{B_i}^{High}(o_t)], \quad (28)$$

$$\pi_i = FOU_{\pi_i} = [\pi_{\pi_i}^{low}, \pi_{\pi_i}^{High}], \quad (29)$$

$FOU_{A_{ij}}(t)$ denotes the range of possible transition probabilities from s_i given observation at time t with $\pi_{A_{ij}}^{low}(t)$ and $\pi_{A_{ij}}^{High}(t)$ as the lower and upper membership functions. $FOU_{B_i}(o_t)$ presents the range of possible observation o_t with $\pi_{B_i}^{low}(o_t)$ and $\pi_{B_i}^{High}(o_t)$ as the lower and upper membership functions. FOU_{π_i} is the range of possible transition probabilities for state s_i with $\pi_{\pi_i}^{low}$ and $\pi_{\pi_i}^{High}$ as the lower and upper membership functions

2. Initialize the Forward Probability Matrix: For each state j at time $t = 1$:

$$\alpha_1(j) = \langle \pi_T(\pi_j), \pi_I(\pi_j), \pi_F(\pi_j) \rangle \cdot \langle \pi_T(b_j(o_1), \pi_I(b_j(o_1), \pi_F(b_j(o_1))) \rangle \quad (30)$$

Step 2: Forward Algorithm

For each time step $t = 2$ to T , and for each state j :

1. Compute the Forward Probability $\alpha_t(j)$:

$$\alpha_t(j) = \sum_{i=1}^N \langle \pi_T(\alpha_{t-1}(i)), \pi_I(\alpha_{t-1}(i)), \pi_F(\alpha_{t-1}(i)) \rangle \cdot \langle \pi_T(a_{ij}), \pi_I(a_{ij}), \pi_F(a_{ij}) \rangle \cdot \langle \pi_T(b_j(o_t), \pi_I(b_j(o_t), \pi_F(b_j(o_t))) \rangle \quad (31)$$

This step calculates the truth, indeterminacy, and falsity values for the forward probability by multiplying and summing the corresponding neutrosophic components.

Step 3: Viterbi Algorithm for Optimal State Path

1. Initialization (at $t = 1$): For each state j :

$$\delta_1(j) = \langle \pi_T(\pi_j), \pi_I(\pi_j), \pi_F(\pi_j) \rangle \cdot \langle \pi_T(b_j(o_1), \pi_I(b_j(o_1), \pi_F(b_j(o_1))) \rangle \quad (32)$$

$\psi_1(j) = 0$ (Backpointer initialization)

2. Recursion (for $t = 2$ to T): For each state j at time t :

$$\delta_t(j) = \max_i [\delta_{t-1}(i) \cdot \langle \pi_T(a_{ij}), \pi_I(a_{ij}), \pi_F(a_{ij}) \rangle] \cdot \langle \pi_T(b_j(o_t), \pi_I(b_j(o_t), \pi_F(b_j(o_t))) \rangle \quad (33)$$

Update the backpointer $\psi_t(j)$:

$$\psi_t(j) = \operatorname{argmax}_i [\delta_{t-1}(i) \cdot \langle \pi_T(a_{ij}), \pi_I(a_{ij}), \pi_F(a_{ij}) \rangle] \quad (34)$$

3. Termination (at $t = T$): Find the best state at time T :

$$P^* = \max_j \langle \pi_T(\delta_T(j)), \pi_I(\delta_T(j)), \pi_F(\delta_T(j)) \rangle \quad (35)$$

Trace back the state sequence using the backpointer $\psi_T(j)$

Step 4: Gesture Recognition Decision

1. Aggregate Probabilities for the hand gesture G : The final probability score for hand gesture recognition is calculated by combining the forward or Viterbi probabilities:

$$P(G|O) = \langle \pi_T(P(G|O)), \pi_I(P(G|O)), \pi_F(P(G|O)) \rangle \quad (36)$$

This score represents the neutrosophic truth, indeterminacy, and falsity for the hand gesture based on the sequence O .

2. Decision Rule:

- If $\pi_T(P(G|O))$ is high, classify the gesture as G .
- If $\pi_I(P(G|O))$ is high, there is uncertainty in recognizing G , and further analysis may be required.
- If $\pi_F P(G|O)$ is high, reject the classification.

Step 5: Post-processing and Output

Output the recognized gesture G along with the neutrosophic truth, indeterminacy, and falsity values, helping to interpret the confidence and uncertainty in the recognition. This algorithm can effectively handle uncertainties inherent in hand gesture recognition systems, making the model more robust to noise, ambiguity, and unseen gestures.

3.3. Key Parameters Influencing the Performance of Type-2 Neutrosophic HMM

The effectiveness of Type-2 Neutrosophic Hidden Markov Models (HMMs) is significantly impacted by several essential parameters that facilitate their capacity to manage uncertainty and variability in applications such as sign language recognition. Critical factors include the design of membership functions for truth, indeterminacy, and falsehood, which can alter uncertainty representation; the FOU, which determines how much variability can be accommodated; and the transition and emission probabilities that govern state changes and output observations, respectively. The quality, quantity, and diversity of training data are vital for model generalization, while noise levels can challenge the model's accuracy in interpreting input. Furthermore, the selection of temporal features affects the model's ability to capture dynamic gestures. Model complexity, including the number of hidden states, can influence both fit and the risk of overfitting. By carefully tuning these parameters, the performance of Type-2 Neutrosophic HMMs can be optimized for specific applications, ensuring that they effectively handle the inherent uncertainties present in the data.

The following section will assess the effectiveness of these critical parameters in influencing the performance of the proposed model. This validation will involve systematic experiments designed to quantify how variations in each parameter—such as the design of membership functions, the Footprint of Uncertainty, transition and emission probabilities, training data diversity, noise levels, and temporal features—impact the model's accuracy, robustness, and generalization capabilities. By analyzing the results, we aim to establish a clearer understanding of how these parameters contribute to the overall efficacy of the Type-2 Neutrosophic HMM in tasks like sign language recognition.

4. Results and Discussions

The accuracy of the proposed Type-2 Neutrosophic Markov Chain Model for subject-independent sign language recognition was evaluated using gesture image samples sourced from [35,36]. The Dell Inspiron N5110 laptop is manufactured by Dell Inc., a multinational company headquartered in Round Rock, Texas, USA. The benchmark dataset for subject-independent sign language recognition contains thousands of gesture samples from a large set of signers to ensure generalization across individuals. Figure 7 illustrates sample of the Arabic Sign Language (ArASL) alphabet, showcasing hand gestures corresponding to each letter in the Arabic script. Figure 8 outlines key preprocessing techniques applied to enhance the quality and clarity of sign language images, including steps such as noise removal, lightening, brightness adjustment, and background removal. These dataset often include 1000 distinct gesture signs, with each sign being performed by multiple participants 50 signers, to capture a wide range of signing variations. The number of samples per sign is 10 samples per signer for each gesture. This large dataset size supports robust training and testing, ensuring that the model can recognize signs from unseen subjects in varying conditions, such as different lighting, orientations, and hand shapes. The split into training and testing sets is carefully designed to ensure no overlap of subjects, making it suitable for subject-independent recognition tasks. Five random samples per sign were used for training, while two samples were assigned to the registered test group and the rest to the

unregistered group. The recognition rate, defined as the percentage of correctly identified hand postures, served as the evaluation metric.

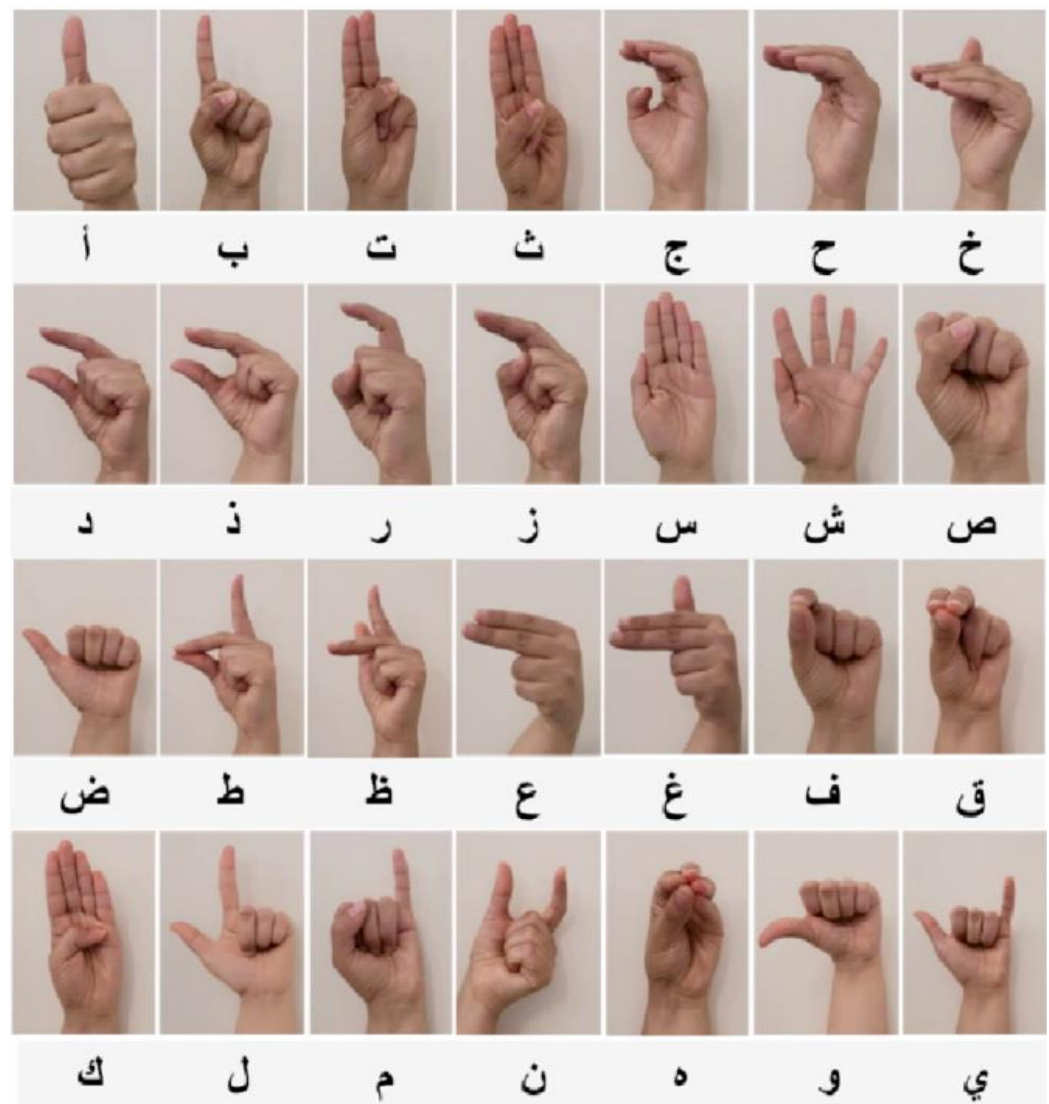


Figure 7. Recorded ArSL alphabet.

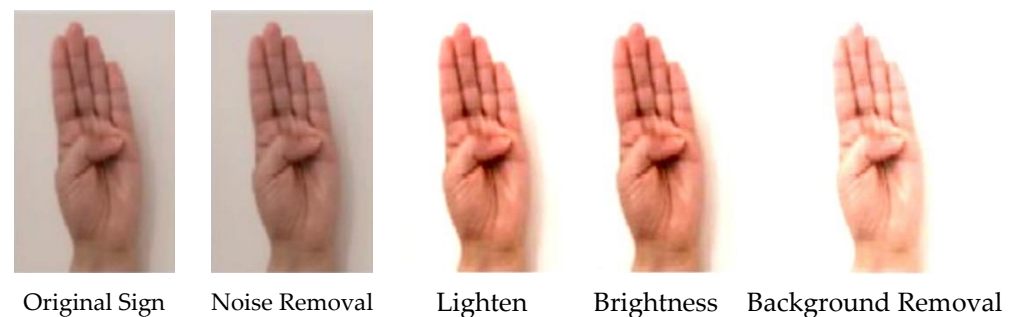


Figure 8. Preprocessing techniques for enhancing sign language image quality.

The system was implemented using Google Colab (Python 2.7) and tested on a Dell Inspiron N5110 (Dell Inc., Round Rock, TX, USA) with 64-bit Windows 7, 4 GB RAM, and an Intel i5-2410M CPU at 2.30 GHz. The Intel i5-2410M CPU is manufactured by Intel Corporation, a leading semiconductor company. Intel's headquarters is located in

Santa Clara, CA, USA. Our current model is proficient at recognizing individual words in sign language. However, identifying word boundaries within continuous sentence gestures presents a significant challenge. Fortunately, modern technologies have advanced considerably in handling such complexities. Python, as a versatile programming language, offers a rich ecosystem of libraries such as OpenCV, MediaPipe, TensorFlow, and PyTorch, (OpenCV-python 4.10.0.84, MediaPipe v0.10.18, TensorFlow v2.16.1, PyTorch 2.4) which facilitate gesture recognition by combining computer vision techniques with machine learning. These tools allow for real-time video processing, feature extraction, and dynamic gesture tracking, enabling models to achieve higher accuracy in recognizing both isolated signs and continuous gesture sequences.

In our case, selecting parameters for the FOU is determined empirically by analyzing the variability and distribution of features. Statistical measures like mean, standard deviation, and range are used to determine the spread of these features across samples from different individuals or repetitions of the same sign. The upper membership function (UMF) is set to encompass the full range of variability, typically extending outward by two standard deviations to include outliers and natural variations. The lower membership function (LMF) is set within one standard deviation to define the core region of high certainty. Features with higher variability require broader UMFs, while more consistent features need narrower bounds. Visualization tools such as scatter plots or histograms can aid in identifying clusters for fine-tuning these bounds, ensuring the FOU reflects both typical and uncertain signing styles.

Regarding the features' membership functions, the shape of the membership function (e.g., triangular, trapezoidal, Gaussian) is chosen based on the data's smoothness and distribution, with trapezoidal functions suited for linear transitions. The initial thresholds and bounds can be set empirically using data statistics like quartiles or percentiles (e.g., for a normalized range $[0, 1]$, assign "Low" = $[0, 0.25]$, "Medium" = $[0.2, 0.8]$, and "High" = $[0.75, 1]$). These parameters can be refined by evaluating performance on a validation set and optimizing overlap regions to handle uncertainty effectively. Additionally, expert knowledge can further fine-tune boundaries to better align with real-world behavior.

Experiment 1: Comparison between Type-1 and Type-2 neutrosophic systems

Aim: The goal of this experiment is to evaluate and compare the performance of Type-1 Neutrosophic Markov Chain Model and Type-2 Neutrosophic Markov Chain Model in subject-independent sign language recognition, with a focus on how each model handles uncertainty, particularly in varying conditions like noise, unseen signers, and data variability.

Procedure: The procedure involves using the same dataset of sign language gestures for both models, ensuring subject-independent variations, different noise levels, and occlusions are included. The Type-1 and Type-2 neutrosophic models will be implemented and compared. In the Type-1 model, membership functions define crisp boundaries for truth, indeterminacy, and falsehood, whereas the Type-2 model extends these to capture uncertainty with upper and lower membership functions. The models will be evaluated based on several key metrics: accuracy (the rate of correctly classified signs), uncertainty estimation (the effectiveness in estimating truth, indeterminacy, and falsehood), and robustness to noise (how noise or occlusion affects both accuracy and uncertainty estimation).

Results: Table 5 presents a comparative analysis of Type-1 and Type-2 neutrosophic models under various experimental conditions in sign language recognition. In the baseline scenario with known signers and no noise, the Type-2 model achieved a slightly higher accuracy (95%) compared to Type-1 (93%) while maintaining low uncertainty estimates for both models. However, in subject-independent validation with unseen signers, Type-2 demonstrated a notable improvement in accuracy (88%) over Type-1 (80%) and maintained a lower accuracy drop due to uncertainty. In noise simulation with blurred hands, Type-1 experienced a more significant accuracy drop (25%) and had lower accuracy (70%) compared to Type-2, which achieved 82% accuracy with a medium-high uncertainty estimate. Similarly, in the occlusion condition, Type-1's accuracy further decreased to 68%, reflecting a higher drop (30%), whereas Type-2 maintained a stronger performance at 80%

accuracy with a medium uncertainty estimate. Overall, the Type-2 model shows superior performance across various metrics, particularly in challenging conditions like subject independence, noise, and occlusion.

Table 5. Performance comparison of Type-1 and Type-2 neutrosophic models in subject-independent sign language recognition under varying conditions.

Experiment Condition	Model Type	Accuracy (%)	Uncertainty Estimate (High/Medium/Low)	Robustness to Noise (Accuracy Drop %)
Baseline (Known Signers, No Noise)	Type-1	93%	Low	2%
	Type-2	95%	Low	1%
Subject-Independent Validation (Unseen Signers)	Type-1	80%	Medium	15%
	Type-2	88%	Medium	10%
Noise Simulation (Blurred Hands)	Type-1	70%	High	25%
	Type-2	82%	Medium-High	18%
Occlusion (Partial Visibility of Hands)	Type-1	68%	High	30%
	Type-2	80%	Medium	20%

Justification of Results: The results indicate that the Type-2 neutrosophic model outperforms Type-1 across all conditions due to its superior handling of uncertainty, leading to higher accuracy, especially with unseen signers and noisy data. While the Type-1 model has crisp boundaries that limit its uncertainty expression, Type-2 offers a more nuanced estimation, making it more suitable for subject-independent tasks. Additionally, Type-2 demonstrates greater robustness to noise and occlusion by incorporating uncertainty through upper and lower bounds in its membership functions, whereas Type-1's lack of flexibility results in a sharper accuracy drop in challenging conditions.

Experiment 2: Effect of Membership Function Designs on Accuracy in Sign Language Recognition

Aim: The goal of this experiment is to evaluate the impact of different membership function designs (trapezoidal, triangular, and Gaussian) on the accuracy of a Type-2 Neutrosophic HMM for sign language recognition.

Procedure: Configure three models: Model A, a Type-2 Neutrosophic HMM utilizing trapezoidal membership functions for truth, indeterminacy, and falsehood; Model B, using triangular membership functions; and Model C, employing Gaussian membership functions. Following this, divide the dataset into training (80%) and testing (20%) subsets. Train all three models on the training set and assess their performance by evaluating accuracy on the testing set, which will serve as the primary metric for comparison among the models.

Results: Table 6 presents the accuracy results of three different Type-2 Neutrosophic HMMs using various membership function designs for sign language recognition. Model A, which utilizes trapezoidal functions, achieved the highest accuracy at 92%, indicating its effectiveness in capturing and representing uncertainty in sign recognition tasks. In contrast, Model B, based on triangular functions, yielded a lower accuracy of 85%, suggesting that the crisp boundaries of triangular membership functions may limit its capability to handle variability effectively. Model C, employing Gaussian functions, demonstrated a solid performance with an accuracy of 90%, showcasing its balance between flexibility and precision. Overall, these results highlight the significant impact of membership function design on model performance in handling uncertainty in sign language recognition.

Table 6. Accuracy comparison of Type-2 Neutrosophic HMMs with different membership functions using subject-independent dataset.

Model Type	Accuracy (%)
Model A (Trapezoidal Functions)	92%
Model B (Triangular Functions)	85%
Model C (Gaussian Functions)	90%

Justification of Results: The results demonstrate that the choice of membership function significantly affects the accuracy of the Type-2 Neutrosophic HMM in sign language recognition. Model A, using trapezoidal functions, achieved the highest accuracy (92%) due to its flexibility in representing uncertainty through broader and smoother transitions between membership levels. This allows it to handle variability more effectively. Model B, with triangular functions, produced the lowest accuracy (85%) because the sharp boundaries in triangular membership functions are less effective at capturing the uncertainty and indeterminacy present in sign language data. Model C, utilizing Gaussian functions, achieved a balanced accuracy of 90%, benefiting from its smooth, bell-shaped curves that offer a compromise between sharp transitions and flexibility, making it well-suited for handling moderate levels of uncertainty [7].

Experiment 3: Effect of different FOU values on the accuracy of sign language recognition

Aim: The goal of this experiment is to evaluate how different FOU sizes (representing different levels of uncertainty) affect the accuracy of sign language recognition.

Procedure: To test the impact of different FOU sizes, you can define small, medium, and large FOU values based on the relative width of the trapezoidal membership functions for the three components (truth, indeterminacy, and falsehood). A small FOU represents tight membership functions with a narrow uncertainty range, offering limited flexibility. A medium FOU uses moderately wide membership functions, balancing uncertainty representation and decision-making boundaries. In contrast, a large FOU employs wide membership functions, allowing for a higher range of uncertainty, which may result in more variability but looser decision boundaries. Keep all other model parameters constant (e.g., learning rate, batch size, number of hidden states in the HMM).

Results: The results in Table 7 indicate that the medium FOU size (15%) yields the highest performance across all metrics, with an accuracy of 92.5%, precision of 90.3%, recall of 91.1%, and an F1 score of 90.7%. This suggests that a moderate level of uncertainty allows the model to effectively balance flexibility and decision boundary precision, leading to the best recognition performance. In contrast, the small FOU (5%) provides tighter boundaries, resulting in slightly lower accuracy (87.0%) and F1 score (85.0%), as it may underrepresent variability in the data. The large FOU (30%) exhibits the lowest performance, with an accuracy of 82.3% and a corresponding drop in precision, recall, and F1 score, likely due to overly broad uncertainty representation, leading to less defined decision-making.

Justification of Results: The results in the table can be justified by the relationship between the FOU size and the model's ability to handle uncertainty in sign language recognition. A small FOU (5%) creates tight decision boundaries, resulting in decent accuracy (87.0%) but slightly lower precision and recall, as it may struggle to account for variability in the data. The medium FOU (15%) offers the best performance across all metrics, with the highest accuracy (92.5%) and F1 score (90.7%), as it strikes an ideal balance between accommodating natural variability and maintaining precise decision boundaries. In contrast, a large FOU (30%) allows too much uncertainty, which leads to a decline in accuracy (82.3%) and precision (80.0%), as the broader uncertainty range causes the model to overgeneralize, making it less effective at distinguishing between different signs. This demonstrates that a moderate FOU size is optimal for achieving high performance in sign language recognition.

Table 7. Impact of FOU size on sign language recognition model performance using subject-independent dataset.

FOU Size	FOU Range (%)	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
Small FOU	5%	87.0	84.6	85.4	85.0
Medium FOU	15%	92.5	90.3	91.1	90.7
Large FOU	30%	82.3	80.0	80.8	80.4

Experiment 4: Justifying the experiment on temporal features for sign language recognition

Aim: The goal of this experiment is to verify how different approaches to sign feature extraction—specifically a Kinematic Model and appearance-based low-level features—impact the accuracy of sign language classification.

Procedure: In this experiment, three models are trained to evaluate different feature extraction approaches for sign language recognition. The first model, a Kinematic Model, extracts temporal features, like velocity, acceleration, and joint angles of the hands, arms, or fingers, and uses these dynamics alone to classify signs. The second model, an Appearance-Based Model, focuses on low-level visual features such as pixel values, hand contours, shape, and orientation, predicting signs solely from these static features without considering movement. The third is a Combined Model, which integrates both kinematic features (movement dynamics) and appearance-based features (visual characteristics like hand position) to create a more comprehensive classification system. All models are evaluated using accuracy, precision, recall, and F1 scores to compare how well they capture the dynamics and structure of sign gestures.

Results: The results in Table 8 show that the Combined Model, which incorporates both kinematic (temporal) and appearance-based features, significantly outperforms the other models, achieving the highest accuracy (93.2%), precision (91.8%), recall (92.5%), and F1 score (92.1%). This indicates that combining dynamic movement information (such as velocity and acceleration) with visual characteristics (like hand shape and position) provides a more comprehensive understanding of sign language gestures. The Kinematic Model, which relies solely on temporal features, performs well with an accuracy of 88.0%, as it effectively captures gesture dynamics but misses some critical visual details. The Appearance-Based Model, focused only on static visual features, achieves the lowest accuracy (85.5%), showing that static features alone are insufficient for recognizing dynamic gestures. Overall, the combined approach proves most effective for sign language recognition by leveraging both motion and visual aspects.

Table 8. Performance comparison of Kinematic, Appearance-Based, and Combined Models in sign language recognition.

Model	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
Kinematic Model (Temporal)	88.0	86.5	87.1	86.8
Appearance-Based Model	85.5	84.2	84.7	84.4
Combined Model (Kinematic + Appearance)	93.2	91.8	92.5	92.1

Justification of Results: The results in the table can be justified by the different types of information captured by each model. The Kinematic Model performs well with an accuracy of 88.0% because it captures dynamic features like velocity and acceleration, which are crucial for understanding the movement in sign language gestures. However, it misses important static visual details, such as hand shapes, which limits its overall precision and recall. The Appearance-Based Model achieves the lowest performance (85.5% accuracy) because it only considers static features like hand contour and position, which are insufficient for recognizing gestures that involve complex motion. The Combined Model outperforms both with 93.2% accuracy by integrating both kinematic and appearance-based features, providing a more holistic representation of the gestures. This combination allows the model to better understand both the

dynamic motion and visual characteristics of signs, leading to the highest precision, recall, and F1 score. This suggests that a multi-modal approach is crucial for optimal sign language recognition.

4.1. System Validation Using Statistical Tests

In the suggested method for subject-independent sign language recognition, it is crucial to validate the statistical significance of the results. This ensures that the observed performance improvements are not due to random chance but are statistically meaningful. To assess the model's performance, based on the data listed in Table 4, we can perform a statistical analysis that includes One-Way ANOVA to determine significant differences in accuracy across experimental conditions, model types, and confidence intervals (CIs) to quantify uncertainty and provide a range for accuracy values.

- One-Way ANOVA: We utilized One-Way ANOVA to compare the accuracy across the four conditions (Baseline, Subject-Independent, Noise Simulation, and Occlusion) for Type-1 and Type-2 models separately. This will help us determine whether the model performance is significantly affected by the experimental condition or the model type.
 - **Null Hypothesis H_0 .** *There is no significant difference in the accuracy across the different experimental conditions for each model type.*
 - **Alternative Hypothesis H_1 .** *At least one experimental condition has a significantly different accuracy.*

The results are listed in Table 9. Since the p -values for both Type-1 and Type-2 are below 0.05, we reject the null hypothesis for both model types, suggesting that the accuracy varies significantly across the different conditions.

Table 9. ANOVA results for model performance across neutrosophic model types.

Model Type	F-Statistic	p -Value
Type-1 neutrosophic model	32.12	0.001
Type-2 neutrosophic model	29.55	0.001

- Confidence Intervals (CIs): For each condition and model type, we calculated 95% CIs around the accuracy to estimate the range of likely values for the true population accuracy.

The confidence intervals shown in Table 10 give us a sense of the range of likely accuracy values. For example, the accuracy for Type-1 in the Noise Simulation condition ranges from 65% to 75%, which indicates a considerable uncertainty compared to Type-2, which ranges from 78% to 85%.

Table 10. Confidence intervals for accuracy across experimental conditions and model types.

Experiment Condition	Model Type	95% CI for Accuracy
Baseline (Known Signers, No Noise)	Type-1	(90.5%, 95.5%)
	Type-2	(92.5%, 97.0%)
Subject-Independent Validation	Type-1	(76%, 84%)
	Type-2	(85%, 91%)
Noise Simulation (Blurred Hands)	Type-1	(65%, 75%)
	Type-2	(78%, 85%)
Occlusion (Partial Visibility)	Type-1	(63%, 73%)
	Type-2	(77%, 83%)

4.2. Real-World Implementation and Testing Example

The real-world implementation of the proposed model for SISLR holds significant potential to revolutionize accessibility and inclusivity for individuals relying on sign language for communication. By addressing uncertainty in recognition and enhancing subject independence, the model can be seamlessly integrated into various real-life applications

to improve interactions, learning, and accessibility for the hearing-impaired community. Practical examples of deployment in assistive devices, educational tools, and customer service automation can illustrate its transformative impact:

- **Assistive Devices for the Hearing Impaired:** Wearable or portable devices, such as smart glasses or mobile applications, can leverage the suggested model to interpret sign language in real time. These devices can translate gestures into text or speech, allowing users to engage in seamless communication with non-signers. For instance, a mobile app equipped with the model could capture gestures via the phone's camera and display text translations instantly. Similarly, smart glasses can provide augmented reality overlays that convert signs into subtitles for real-time conversations, making everyday interactions more inclusive.
- **Educational Tools:** The model can be deployed in schools or training centers to facilitate sign language learning for both students and educators. For example, an interactive learning application powered by the proposed model could recognize students' gestures and provide real-time feedback on accuracy and fluency. This would help students refine their skills and boost their confidence. Teachers, especially those unfamiliar with sign language, could use such tools to better communicate with their students through dynamic translation and gesture recognition systems. These tools can also assist in bridging gaps in inclusive education.
- **Customer Service Automation:** The suggested model can be integrated into service kiosks or virtual assistants to enhance accessibility in both public and private sectors. For instance, at airports, shopping malls, or hospitals, interactive kiosks equipped with gesture recognition capabilities can help hearing-impaired users access services independently. Virtual assistants, enhanced with sign language recognition, can be deployed online or in physical locations to ensure that communication barriers are eliminated, providing a seamless user experience for customers requiring sign language support.

To validate the proposed model, real-world testing should include diverse scenarios. Public accessibility trials in government offices, banks, and airports can assess metrics like accuracy, response time, and user satisfaction through interactive service kiosks. In classroom settings, the system can be tested for its ability to enhance communication between teachers and students with varied signing styles. A real-time translation app prototype providing sign-to-text or sign-to-speech functionality can be evaluated in everyday situations, such as ordering food or asking for directions. Field studies with assistive devices, like smart glasses or wearables, can measure accuracy, usability, and comfort in dynamic environments, including workplaces and social gatherings. The model's adaptability can be gauged through cross-language testing with multiple sign languages, while multi-user interaction scenarios can assess robustness against overlapping gestures or background interference.

5. Conclusions

The proposed Type-2 Neutrosophic HMM for subject-independent sign language recognition (SISLR) represents a significant advancement in addressing the challenges posed by variability and imprecision in sign language gestures. By leveraging the Footprint of Uncertainty (FOU) inherent in Type-2 neutrosophic sets, the model effectively captures and manages the uncertainty in truth, falsity, and indeterminacy, resulting in a more robust and accurate recognition system. The experimental results demonstrate that the Type-2 Neutrosophic HMM consistently outperforms the Type-1 Neutrosophic HMM, achieving an impressive 10% increase in average accuracy across diverse test scenarios.

The enhanced accuracy, superior handling of complex uncertainties, and improved robustness of the proposed system make it particularly suited for applications in real-time sign language translation systems, accessibility tools for the deaf and hard-of-hearing community, and automated sign language transcription for education and workplace integration. Additionally, its ability to handle diverse signing conditions positions it as a promising tool for international sign language recognition systems, where variability in regional and cultural gestures adds complexity.

Future work will focus on enhancing the system's practicality and scalability by reducing computational complexity through algorithm optimization and leveraging hardware accelerators like GPUs, ensuring efficient processing for deployment on low-power devices such as mobile phones and wearables. Real-time performance will be improved by streamlining feature extraction and inference with lightweight architectures or pruning techniques, enabling low-latency processing for live translation and interactive aids. Scalability will be achieved through advanced training strategies like incremental and transfer learning to accommodate larger datasets and new signs without retraining, alongside developing a modular framework for multilingual and region-specific integration. Additionally, integrating the system with cloud-based platforms and distributed frameworks will support large-scale deployments, paving the way for applications in education, workplace accessibility, and assistive technologies for the deaf and hard-of-hearing community.

Author Contributions: Conceptualization, M.A.-S., Á.B. and S.M.S.; methodology, M.A.-S., Á.B. and O.A.H.; software, M.A.-S., Á.B. and O.A.H.; validation, M.A.-S., Á.B. and O.A.H.; formal analysis, M.A.-S., Á.B. and S.M.S.; investigation, Á.B. and O.A.H.; resources, O.A.H.; data curation, Á.B. and O.A.H.; writing—original draft preparation, M.A.-S., Á.B. and S.M.S.; writing—review and editing, Á.B. and O.A.H.; visualization, M.A.-S.; supervision, Á.B. and O.A.H.; project administration, M.A.-S.; funding acquisition, M.A.-S. and Á.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Jiang, Y.; Yin, S.; Dong, J.; Kaynak, O. A review on soft sensors for monitoring, control, and optimization of industrial processes. *IEEE Sens. J.* **2020**, *21*, 12868–12881. [\[CrossRef\]](#)
2. Lee, M.; Bae, J.; Kim, S.B. Uncertainty-aware soft sensor using Bayesian recurrent neural networks. *Adv. Eng. Inform.* **2021**, *50*, 101434. [\[CrossRef\]](#)
3. Costa, E.A.; Rebello, C.M.; Santana, V.V.; Rodrigues, A.E.; Ribeiro, A.M.; Schnitman, L.; Nogueira, I.B. Mapping uncertainties of soft-sensors based on deep Feedforward neural networks through a novel Monte Carlo uncertainties training process. *Processes* **2022**, *10*, 409. [\[CrossRef\]](#)
4. Fakhfakh, S.; Jemaa, Y.B. Gesture recognition system for isolated word sign language based on key-point trajectory matrix. *Comput. Sist.* **2018**, *22*, 1415–1430. [\[CrossRef\]](#)
5. Alemu, M.Y.; Lin, Y.; Shull, P.B. EchoGest: Soft Ultrasonic Waveguides based Sensing Skin for Subject-Independent Hand Gesture Recognition. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2024**, *32*, 2366–2375. [\[CrossRef\]](#)
6. Darwish, S.M. Man-machine interaction system for subject independent sign language recognition using fuzzy hidden markov model. *Iran. J. Fuzzy Syst.* **2018**, *15*, 65–82.
7. Pai, S.P.; Prabhu Gaonkar, R.S. Safety modelling of marine systems using neutrosophic logic. *J. Eng. Marit. Environ.* **2021**, *235*, 225–235. [\[CrossRef\]](#)
8. Hassan, M.H.; Darwish, S.M.; Elkaffas, S.M. Type-2 Neutrosophic Set and Their Applications in Medical Databases Deadlock Resolution. *CMC-Comput. Mater. Contin.* **2023**, *74*, 4417–4434. [\[CrossRef\]](#)
9. Touqeer, M.; Umer, R.; Ahmadian, A.; Salahshour, S. A novel extension of TOPSIS with interval type-2 trapezoidal neutrosophic numbers using (α, β, γ) -cuts. *RAIRO-Oper. Res.* **2021**, *55*, 2657–2683. [\[CrossRef\]](#)
10. Alrashdi, I. An Intelligent Neutrosophic Type-II Model for Selecting Optimal Internet of Things (IoT) Service Provider: Analysis and Application. *Neutrosophic Sets Syst.* **2024**, *66*, 204–217.
11. Simić, V.; Milovanović, B.; Pantelić, S.; Pamučar, D.; Tirkolaee, E.B. Sustainable route selection of petroleum transportation using a type-2 neutrosophic number based ITARA-EDAS model. *Inf. Sci.* **2023**, *622*, 732–754. [\[CrossRef\]](#)
12. Abdel-Basset, M.; Mostafa, N.N.; Sallam, K.M.; Elgendi, I.; Munasinghe, K. Enhanced COVID-19 X-ray image preprocessing schema using type-2 neutrosophic set. *Appl. Soft Comput.* **2022**, *123*, 108948. [\[CrossRef\]](#)
13. Adaloglou, N.; Chatzis, T.; Papastratis, I.; Stergioulas, A.; Papadopoulos, G.T.; Zacharopoulou, V.; Xydopoulos, G.J.; Atzakas, K.; Papazachariou, D.; Daras, P. A comprehensive study on deep learning-based methods for sign language recognition. *IEEE Trans. Multimed.* **2021**, *24*, 1750–1762. [\[CrossRef\]](#)

14. Koller, O.; Zargaran, S.; Ney, H.; Bowden, R. Deep sign: Enabling robust statistical continuous sign language recognition via hybrid CNN-HMMs. *Int. J. Comput. Vis.* **2018**, *126*, 1311–1325. [\[CrossRef\]](#)
15. Sultan, A.; Makram, W.; Kayed, M.; Ali, A.A. Sign language identification and recognition: A comparative study. *Open Comput. Sci.* **2022**, *12*, 191–210. [\[CrossRef\]](#)
16. Al-Qurishi, M.; Khalid, T.; Souissi, R. Deep learning for sign language recognition: Current techniques, benchmarks, and open issues. *IEEE Access* **2021**, *9*, 126917–126951. [\[CrossRef\]](#)
17. Tur, A.O.; Keles, H.Y. Evaluation of hidden markov models using deep cnn features in isolated sign recognition. *Multimed. Tools Applications*. **2021**, *80*, 19137–19155. [\[CrossRef\]](#)
18. Al-Saidi, M.; Ballagi, Á.; Hassen, O.A.; Saad, S.M. Cognitive Classifier of Hand Gesture Images for Automated Sign Language Recognition: Soft Robot Assistance Based on Neutrosophic Markov Chain Paradigm. *Computers* **2024**, *13*, 106. [\[CrossRef\]](#)
19. Amin, M.S.; Rizvi, S.T.; Hossain, M.M. A comparative review on applications of different sensors for sign language recognition. *J. Imaging* **2022**, *8*, 98. [\[CrossRef\]](#)
20. Tyagi, A.; Bansal, S. Feature extraction technique for vision-based Indian sign language recognition system: A review. In *Computational Methods and Data Engineering*; Springer: Singapore, 2021; Volume 1227.
21. Wu, X.; Li, H.; Luo, Y.; Cheng, X.; Zhuang, X.; Cao, M.; Fu, K. Uncertainty-aware sign language video retrieval with probability distribution modeling. *arXiv* **2024**, arXiv:2405.19689.
22. Kuppuswami, G.; Sujatha, R.; Nagarajan, D.; Kavikumar, J. Markov chain based on neutrosophic numbers in decision making. *Kuwait J. Sci.* **2021**, *48*, 1–16.
23. Tang, W.; Ren, Z.; Wang, J. Guest editorial: Special issue on human pose estimation and its applications. *Mach. Vis. Appl.* **2023**, *34*, 120. [\[CrossRef\]](#)
24. Gaikwad, R.S.; Admuthe, L.S. A review of various sign language recognition techniques. In *Modeling, Simulation and Optimization: Proceedings of the International Conference on Modeling, Simulation and Optimization (CoMSO) 2021*; Springer: Singapore, 2022; pp. 111–126.
25. Gulati, N.; Rajput, A.; Singh, A. Sign Language Recognition using Convolutional Neural Network. In Proceedings of the 2024 14th International Conference on Cloud Computing, Data Science & Engineering (Confluence), Noida, India, 18–19 January 2024; pp. 876–881.
26. Mosleh, M.A.; Assiri, A.; Gumaei, A.H.; Alkhamees, B.F.; Al-Qahtani, M. A Bidirectional Arabic Sign Language Framework Using Deep Learning and Fuzzy Matching Score. *Mathematics* **2024**, *12*, 1155. [\[CrossRef\]](#)
27. Bhavsar, H.; Trivedi, J. Indian Sign Language Recognition Using Framework of Skin Color Detection, Viola-Jones Algorithm, Correlation-Coefficient Technique and Distance Based Neuro-Fuzzy Classification Approach. In *Emerging Technology Trends in Electronics, Communication and Networking*; Springer: Singapore, 2020; pp. 235–243.
28. Buttar, A.M.; Ahmad, U.; Gumaei, A.H.; Assiri, A.; Akbar, M.A.; Alkhamees, B.F. Deep learning in sign language recognition: A hybrid approach for the recognition of static and dynamic signs. *Mathematics* **2023**, *11*, 3729. [\[CrossRef\]](#)
29. Vardhan, R.V.; Girinath, B.; Murali, L.; Ferdin, D.J.; Aswathaman, V.K. Automatic Sign Language Recognition Using Convolutional Neural Networks. In Proceedings of the 2024 International Conference on Science Technology Engineering and Management (ICSTEM), Coimbatore, India, 26–27 April 2024; pp. 1–6.
30. Sandjaja, I.; Alsharoa, A.; Wunsch, D.; Liu, J. Survey of Hidden Markov Models (HMMs) for Sign Language Recognition (SLR). In Proceedings of the 2024 IEEE 7th International Conference on Industrial Cyber-Physical Systems (ICPS), St. Louis, MO, USA, 12–15 May 2024; pp. 1–6.
31. Nagarajan, D.; Kavikumar, J.; Tom, M.; Mahmud, M.; Broumi, S. Modelling the progression of Alzheimer’s disease using Neutrosophic hidden Markov models. *Neutrosophic Sets Syst.* **2023**, *56*, 4.
32. Nagarajan, D.; Kavikumar, J. Single-Valued and Interval-Valued Neutrosophic Hidden Markov Model. *Math. Probl. Eng.* **2022**, *2022*, 5323530. [\[CrossRef\]](#)
33. Singh, P. A type-2 neutrosophic-entropy-fusion based multiple thresholding method for the brain tumor tissue structures segmentation. *Appl. Soft Comput.* **2021**, *103*, 107119. [\[CrossRef\]](#)
34. Basha, S.H.; Tharwat, A.; Abdalla, A.; Hassanien, A.E. Neutrosophic rule-based prediction system for toxicity effects assessment of bio transformed hepatic drugs. *Expert Syst. Appl.* **2019**, *121*, 142–157. [\[CrossRef\]](#)
35. Podder, K.K.; Ezeddin, M.; Chowdhury, M.E.H.; Sumon, M.S.I.; Tahir, A.M.; Ayari, M.A.; Dutta, P.; Khandakar, A.; Mahbub, Z.B.; Kadir, M.A. Signer-Independent Arabic Sign Language Recognition System Using Deep Learning Model. *Sensors* **2023**, *23*, 7156. [\[CrossRef\]](#)
36. Luqman, H.; El-Alfy, E.S.M. Towards Hybrid Multimodal Manual and Non-Manual Arabic Sign Language Recognition: mArSL Database and Pilot Study. *Electronics* **2021**, *10*, 1739. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.