



Enhancing Medical X-Ray Image Classification with Neutrosophic Set Theory and Advanced Deep Learning Models

Walid Abdullah^{1,*}

¹Faculty of Computers and Informatics, Zagazig University, Zagazig, Sharqia,
44519, Egypt, waleed@zu.edu.eg.

* Correspondence: waleed@zu.edu.eg.

Abstract

The classification of medical images presents significant challenges due to the presence of noise, uncertainty, and indeterminate information. Traditional deep learning models often struggle to manage this, leading to reduced diagnostic accuracy, especially when dealing with low-quality or ambiguous conditions. This paper proposes a hybrid approach that integrates Neutrosophic Set (NS) theory with deep learning models to enhance X-ray image classification under uncertain conditions. NS theory introduces three domains: True (T), Indeterminate (I), and False (F) to manage image uncertainty and noise, allowing deep learning models to better interpret complex, ambiguous visual information. To evaluate the approach, five state-of-the-art deep learning models—MobileNet, ResNet50, VGG16, DenseNet121, and InceptionV3 are utilized, and their performance was evaluated on two different medical image datasets: Cervical spine injuries detection and chest disease classification. The results indicate that models trained on NS-transformed data, particularly DenseNet and MobileNet, yield superior outcomes compared to those trained on the original data, achieving significantly higher accuracy, precision, and recall. This demonstrates that incorporating NS theory into deep learning models significantly enhances their ability to classify uncertain and noisy X-ray images, providing a robust solution for improving diagnostic accuracy in medical imaging.

Keywords: Neutrosophic Set (NS), Image Classification, X-ray Imaging, Deep Learning, Cervical spine, Chest Diseases.

1. Introduction

Medical imaging plays a pivotal role in diagnosing and managing diseases, especially in critical areas such as chest diseases and Cervical spinal injuries. X-rays are the most widely used due to their cost-effectiveness and rapid results, particularly in the detection of chest diseases and bone fractures [1]. However,

despite their ubiquity, the interpretation of X-ray images presents significant challenges for both clinicians and automated systems. Artificial Intelligence, particularly deep learning models have shown remarkable success in image classification tasks, offering state-of-the-art performance in a range of medical applications [2]. Despite the power of AI, medical images, especially X-rays, often exhibit ambiguity and incomplete information due to varying factors such as exposure settings, image noise, and patient positioning inconsistencies can hide critical details and introduce uncertainty [3]. Addressing these uncertainties is essential for building more robust and reliable classification models Which lead to more accurate diagnosis process [4].

Deep learning has revolutionized medical image analysis, with convolutional neural networks (CNNs) being the most prominent technique used [5]. CNNs such as MobileNet, ResNet, VGG16, DenseNet121, and Inception offer powerful feature extraction capabilities, enabling them to capture intricate patterns in medical images [6, 7]. These models have been widely adopted for tasks ranging from disease detection to segmentation. However, despite their impressive performance in image classification, most existing methods lack mechanisms to effectively manage the uncertainty and noise inherent in medical imaging data, which can lead to suboptimal diagnostic accuracy [8]. Deep learning models such as CNNs, while powerful in many contexts, struggle when dealing with noisy or ambiguous data. Additionally, the inherent complexity of human anatomy makes it difficult to distinguish between subtle abnormalities in cases such as spinal fractures or chest diseases [9]. These gaps underscore the need for models capable of managing uncertainty while maintaining high classification accuracy.

Neutrosophic Set (NS) theory offers a promising approach to handle uncertainty in image classification. NS theory was introduced by Florentin Smarandache in 1995 and unified 1999 [10]. It extends classical binary logic by introducing three domains: truth (T), falsity (F), and indeterminacy (I) [11]. These domains allow for the representation of incomplete, inconsistent, and uncertain information, which is often present in medical X-rays. By converting pixel values into these three domains, NS theory helps to emphasize the certainty of certain image features while minimizing the impact of uncertain or noisy areas. Neutrosophic Sets operate by assigning three membership values (T, I, F) to each element within a set. Each value lies in the interval $[0, 1]$, representing the degrees to which the element belongs to the truth, indeterminacy, and falsity domains [12]. This tripartite model is uniquely advantageous in scenarios where information is incomplete or conflicting, such as medical imaging [13, 14]. For instance, in an X-ray of a fracture, some regions may be clearly indicative of a fracture (high truth), some may not indicate a fracture (high falsity), while other regions, perhaps due to poor image quality or ambiguous features may fall into an indeterminate zone (high indeterminacy).

By explicitly modeling the areas of uncertainty, NS provides a more refined approach to interpreting medical images. This framework has been applied in various image processing tasks, such as segmentation and classification, particularly in the medical field [14, 15]. In the context of X-ray image classification, few studies have integrated uncertainty management techniques, such as NS theory, into the preprocessing or training pipeline of deep learning models, leaving a significant gap in leveraging uncertainty-aware methods for medical X-ray analysis. In our approach, the NS framework is applied to X-ray images, aiding in the classification task by enhancing images feature representation and improving robustness in uncertain scenarios. Thus, integrating uncertainty management methods, such as NS theory, into deep learning pipelines could potentially improve performance in challenging medical scenarios.

In this paper, we propose a hybrid approach that integrates NS theory with deep learning models to enhance the classification of medical X-ray images. Our methodology leverages NS theory during the preprocessing phase to address image uncertainty, such as noise and low contrast, and transform the images into three NS images domains (T, I, F), which may improve the interpretability and robustness of deep learning models. To demonstrate the efficacy of this approach, we conduct case studies on two distinct datasets cervical spine fracture and dislocation, and chest diseases, which show the framework's ability to generalize across diverse medical imaging tasks. To evaluate the approach, five state-of-the-art deep learning models were used including MobileNet, ResNet50, VGG16, DenseNet121, and InceptionV3. The Experimental results indicate that models trained on NS-transformed data, particularly DenseNet121 and MobileNet, demonstrate superior performance compared to those trained on the original data in terms of accuracy, precision, and recall. The key contributions of this work include:

- a. Introducing an NS-based pre-processing step to manage uncertainty in X-ray medical images, making it easier for deep learning models to extract meaningful patterns.
- b. Demonstrating the impact of NS theory in improving classification performance across five state-of-the-art models: MobileNet, ResNet50, VGG16, DenseNet121, and InceptionV3.
- c. Evaluating the effectiveness of our approach on two distinct medical datasets (spinal fracture X-rays and chest disease images)

This hybrid methodology addresses limitations in existing deep learning approaches by offering a more reliable solution for medical image classification in uncertain environments. Our results underscore the potential of NS theory as a transformative tool for handling uncertainty in x-rays images, to be more understood for DL models, which make this a interesting area for further investigation.

The remainder of this paper is structured as follows: Section 2 provides a detailed overview of the related work in the areas of Neutrosophic Sets and deep learning for medical image classification. Section 3 describes the methodology used in this study, including the NS-based preprocessing technique, and the architecture of the deep learning models (MobileNet, ResNet, VGG, DenseNet, Inception). In Section 4, we present the experimental setup, the datasets, and evaluation metrics. Section 5 includes the results, the comparisons of the models, and a detailed discussion of the results. Finally, Section 6 concludes the paper with a summary of findings and suggestions for future work.

2. Related Work

In this section, we will review the application of Neutrosophic Set Theory and deep learning techniques in the classification of medical images, particularly focusing on X-rays images. The review will cover various approaches used to enhance accuracy, highlight key methodologies and utilize models that have been employed in recent studies, and discuss their strengths, limitations, and the gaps that remain in achieving more accurate and reliable approaches.

Cei et al. [16] introduced a novel NS-based deep learning approach for analyzing digital mammograms, specialized X-ray images of the breast used to detect and diagnose breast cancer, specifically for the detection and classification of microcalcifications (MCs), which are critical early indicators of breast cancer. The technique utilized the membership sets of NS to map digital mammograms into three distinct domains: T, I, and F. These domains were then employed to train a convolutional neural network (CNN) model, facilitating tasks such as lesion detection and regional clustering. The proposed method achieved a sensitivity of 92.5% for detecting MC clusters, with an area under the curve (AUC) of 0.908 and 0.872 on the validation and test sets, respectively. These results underscore the effectiveness of combining NS and DL techniques in automating the detection and classification of MC clusters in digital mammography images.

Khalifa [17], produced a study to explore the impact of neutrosophic sets on deep learning models using a limited dataset of COVID-19 X-ray images. The images were transformed into the NS domain, which consists of three categories: true (T), indeterminacy (I), and false (F) images. These transformed images were then used to train various DL models includes AlexNet, GoogLeNet, and ResNet18. The performance of the models was evaluated across four domain-original images and the three NS domains—comparing results based on accuracy, precision, recall, and F1 score. The findings demonstrated that incorporating NS into DL models could significantly enhance testing accuracy, particularly in the context of limited COVID-19 datasets, highlighting its potential for improving diagnostic accuracy. The authors in [18] explored the use of chest CT scans for early detection of COVID-19, they proposed a hybrid method which integrates binary cross-entropy, transfer

learning, and deep convolutional neural networks techniques for enhanced classification accuracy. Pre-trained models like ResNet (50), VGG (19), VGG (16), and Inception V3 were applied to the DCNNs to improve the classification process. The results showed that the pre-trained models achieved accuracies of 99.07%, 98.70%, 98.55%, and 96.23%, respectively, using the Adam optimizer. Another study addressed the feature selection problem in COVID-19 detection using chest X-ray images by integrating metaheuristic algorithms with a deep learning model [19]. In this approach, the VGG19 deep network was employed for feature extraction, and a feature selection method was applied to identify the most significant features, enhancing the classification performance. The selected features were then fed into an optimized neural network, with optimization driven by a hybrid metaheuristic optimizer. The results demonstrated that the proposed method achieved an accuracy of 99.88%, showcasing its effectiveness in improving COVID-19 detection.

The study presented in [20] introduces the BoneNet-NS technique for classifying fractures in X-ray images. This approach integrates deep learning with neutrosophic set methodologies to effectively manage aleatoric uncertainty in medical imaging. The research proposes two frameworks for combining NS with DL models, named BoneNet-NS1 and BoneNet-NS2. Utilizing a dataset of 4,924 X-ray images, various DL models, including Xception, ResNet52V2, DenseNet121, and a customized CNN, were evaluated to distinguish between fractured and non-fractured classes. Statistical analyses demonstrated that BoneNet-NS2 achieved remarkable performance metrics across most DL models when working with NS image domain instead of the original images. Specifically, with the ResNet52V2 model, BoneNet-NS2 achieved an accuracy of 99.7%, a log loss of 0.006, F1-score of 99.7.

Jennifer et al. [21] employed a neutrosophic approach to differentiate between lung infection types. By classifying chest x-ray images into True (T), False (F), and Indeterminacy (I) set memberships, this approach effectively reduces fuzziness while retaining critical information for feature extraction of lung opacity. The preprocessing stage utilizes alpha-mean and beta-enhancement operations to decrease indeterminacy and enhance relevant image components. Subsequently, the enhanced neutrosophic images are classified using several deep learning models, including ResNet-50, VGG-16, and XGBoost. Experimental evaluations conducted on the ActualMed COVID-19 Chest X-ray and COVID-19 radiography datasets reveal that the enhanced neutrosophic images achieve a notable accuracy of 97.33%, surpassing the performance of other domain sets.

In [22], a deep neural network model utilizing neutrosophic features for skin cancer diagnosis was introduced. The model applied neutrosophic-based lesion segmentation to reduce noise and improve classification accuracy on PH2, ISIC 2017, ISIC 2018, and ISIC 2019 datasets. The network, built with Inception and residual blocks, achieved high accuracy rates of 99.50%, 99.33%, 98.56%, and 98.04% on the respective datasets, outperforming many existing classifiers. These results highlight

the effectiveness of combining neutrosophic techniques with deep learning for skin cancer detection.

Authors in [23] introduced a hybrid to address the need for rapid and accurate COVID-19 detection using chest X-ray images. The study utilized Neutrosophic techniques (NTs) combined with machine learning (ML) methods to create an automated tool for classifying X-rays into COVID-19 positive or negative cases. Morphological features (MFs) and principal component analysis (PCA) were employed to extract key features from the images. Compared to RT-PCR tests, which are costly and require expert personnel, chest X-rays provide a more accessible and efficient alternative for COVID-19 diagnosis. The model achieved high performance metrics, including 98.46% accuracy, 98.19% precision, 98.18% sensitivity, demonstrating its effectiveness and superiority over other diagnostic methods.

In [24], an encoder-decoder deep neural network incorporating neutrosophic set theory and indeterminacy fusion is proposed for segmenting White Blood Cells (WBCs). The model leverages indeterminacy within the NS domain to enhance the segmentation of WBCs into distinct nucleus and cytoplasm regions. This approach addresses the limitations of prior methods, which often overlook internal structures. The model surpasses three original encoder-decoder networks, achieving high precision rates and the greatest mean segmentation accuracy of 0.95301, demonstrating the effectiveness of integrating NS-based indeterminacy in improving medical image segmentation.

Another hybrid method combining neutrosophic and convolutional neural networks (NS-CNN) is proposed in [25] for the classification of brain tumors as benign or malignant. The method first segments MRI images using the neutrosophic set – expert maximum fuzzy-sure entropy (NS-EMFSE) approach. In the classification stage, features from the segmented brain images are extracted using a CNN and then classified using SVM and KNN classifiers. Based on a 5-fold cross-validation on 80 benign and 80 malignant tumor samples, the experimental results demonstrated that the CNN features performed best with the SVM classifier, achieving an average accuracy of 95.62%.

Despite notable advancements in the application of NS within DL methodologies for addressing uncertainty and noise in medical image analysis, significant gaps remain regarding robustness and accuracy. Many studies report limitations in classification performance, often relying on small datasets, which may yield unpredictable results over time. Furthermore, there is a pressing need for more comprehensive evaluations of NS combined with DL techniques in the medical field. In this context, testing the proposed approach across multiple datasets not only enhances the robustness and generalizability of the model but also provides a more holistic understanding of its performance across diverse medical imaging scenarios.

3. Methodology

The methodology is designed to combine the Neutrosophic Set (NS) theory and deep learning models for effective classification of medical X-ray images. It consists of three core components: first, the theoretical foundation and equations governing the NS domain and its application to image uncertainty handling are introduced. This is followed by an explanation of the deep learning models utilized, including MobileNet, ResNet50, VGG16, DenseNet121, and inceptionV3, highlighting their architecture and suitability for the task at hand. Finally, the proposed approach, which integrates the NS-based pre-processing with the deep learning models to address the challenges of image noise, ambiguity, and the need for accurate diagnosis through automated classification, is detailed.

3.1. Neutrosophic Set (NS) Definitions and Preliminaries

Neutrosophic Set (NS) theory extends traditional fuzzy set theory by introducing the concepts of indeterminacy, allowing for a more comprehensive representation of uncertainty in data. NS decomposes information into three distinct components: Truth (T), Indeterminacy (I), and Falsity (F). NS is particularly useful for handling uncertain, incomplete, and inconsistent data, which are common in medical imaging due to noise and low contrast in X-rays. The Key Definitions are:

- Truth (T): Represents the degree of truth or certainty in the image's pixel values.
- Indeterminacy (I): Captures the uncertainty or ambiguity present in the image.
- Falsity (F): Denotes the degree of falsehood or error in the pixel values.

• Neutrosophic Sets (NS) Domains

In NS, each element x is characterized by three independent degrees: truth, indeterminacy, and falsehood. Mathematically, this can be expressed as follows:

$$NS(x) = \{T(x), I(x), F(x)\} \quad (1)$$

$$0 \leq T(x), I(x), F(x) \leq 1 \quad (2)$$

$$0 \leq T(x) + I(x) + F(x) \leq 3 \quad (3)$$

where: $T(x)$ represents the degree of truth, $I(x)$ denotes the degree of indeterminacy, $F(x)$ indicates the degree of falsehood. In image transformation process to Neutrosophic Domains, each pixel of an image is preprocessed by Neutrosophic logic to calculate its T, F, I components to determine its domain. Given an X-ray image G with pixel intensities, Pixel in image is represent as $P(x, y)$ the image can be transformed into the domains using the following:

- First, to handle image noise and uncertainty, the local mean $g(x, y)$ of an image G is calculated using a convolution operation. Then Calculate absolute difference between the original image and its mean $O(x, y)$.

$$g(x, y) = \frac{1}{25} \sum_{i=-2}^2 \sum_{j=-2}^2 G(x + i, y + j) \quad (4)$$

$$O(x, y) = \text{abs}(G(x, y) - g(x, y)) \quad (5)$$

- Second the NS domains for the pixel $P(x, y)$ can be represented as follows:

$$P_{NS}(x, y) = \{T(x, y), I(x, y), F(x, y)\} \quad (6)$$

$$T(x, y) = \frac{g(x, y) - g_{min}}{g_{max} - g_{min}} \quad (7)$$

$$I(x, y) = \frac{O(x, y) - O_{min}}{O_{max} - O_{min}} \quad (8)$$

$$F(x, y) = 1 - T(x, y) \quad (9)$$

where x and y variables typically represent the pixel coordinates in a 2D image, $G(x, y)$ refers to the pixel value at the coordinate (x, y) in the image G , i and j are indices used to denote the neighboring pixels around a given pixel (x, y) . g_{max} is the highest average pixel value within a defined neighborhood. g_{min} is the lowest average pixel value within that same neighborhood. O_{max} is the highest absolute difference observed across the neighborhood. O_{min} is the lowest absolute difference observed [25, 26]. By applying these transformations to each pixel of the image, the X-ray is decomposed into its NS components. This allows for better feature extraction in images with uncertainty, aiding in medical diagnoses where details are often unclear.

• Entropy in Neutrosophic Sets

Entropy is typically used to quantify the uncertainty or randomness in the image data [27]. It can be used to assess the information content of the image. Higher entropy indicates more complexity or uncertainty in the pixel values, which could influence how to define T , I , and F . and to adjust the parameters in those equations, essentially influencing how you interpret the pixel values around each $G(x, y)$. The Entropy can be employed to evaluate the uncertainty associated with the image data. The entropy H of a neutrosophic set can be defined as:

$$\text{Entropy}(E) = - \sum_{k=1}^N P(K) \log(P(K)) \quad (10)$$

$$E_{NS} = E_T + E_I + E_F \quad (11)$$

This equation provides a measure of the uncertainty within the neutrosophic representation of the image, where $p(k)$ is the probability of occurrence of pixel value k in the image, and E_T , E_I , and E_F are entropies for True, Indeterminacy, and False, respectively.

To adjust the influence of the entropy of the image on the degree of truth and balance the response of the NS approach under different image conditions, additional parameters can be used such as *Alpha* (α) and *Beta* (β). *Alpha* (α) parameter controls the weight of entropy in the classification. A higher value indicates a greater reliance on the image's information content, enhancing the representation of areas of interest in the image. While *Beta* (β) acts as a complementary measure that reflects the degree of uncertainty in the classification process. It allows for a balance between certainty and uncertainty, making the model robust against noise and variations in the input data.

$$\alpha = \alpha_{min} + \frac{(\alpha_{max} - \alpha_{min})(E_n - E_{min})}{(E_{max} - E_{min})} \quad (12)$$

$$\beta = 1 - \alpha \quad (13)$$

where E_{max} is the maximum possible entropy for an image of size $m \times n$, E_{min} represents the minimum entropy, which occurs when the image is completely uniform (complete certainty), while α_{min} and α_{max} provide bounds for the influence of entropy on the NS framework, allowing you to control how the model reacts to uncertainty in different images.

3.2. Deep Learning Models

Deep learning has emerged as a powerful approach for solving complex problems in various fields, including medical image analysis. At the core of deep learning is the use of neural networks, particularly Convolutional Neural Networks (CNNs) [28], which have revolutionized image classification tasks by automatically learning hierarchical feature representations directly from data. CNNs are specifically designed to process grid-like data, such as images, by utilizing convolutional layers that apply filters to detect important features like edges, textures, and patterns. These features are learned through multiple layers of abstraction, allowing CNNs to capture both low-level and high-level details within the images. In the medical imaging, CNNs can automatically identify complex and subtle patterns that may not be easily visible to the human eye, making them invaluable for tasks such as disease diagnosis and fracture detection. However, training CNNs from scratch requires vast amounts of labeled data and significant computational resources. To overcome this limitation, we leverage transfer learning [29], a technique that allows pretrained models to be adapted to new domains. Transfer learning uses models that have already been trained on large datasets, such

as ImageNet, and fine-tunes them on the target medical datasets [30]. This approach not only reduces the need for large amounts of training data but also significantly accelerates the learning process.

In this study, deep learning models are trained to work within the **neutrosophic domain**, where medical images are transformed to capture true, false, and indeterminate information based on neutrosophic sets. By extracting these features, the models are equipped to handle the inherent uncertainty present in medical imaging data. CNNs are well-suited for this task as they can capture important structural features in medical images and can be adapted to work with the nuanced information provided by the neutrosophic transformation. This allows the models to focus on extracting the most relevant and informative features from noisy or ambiguous data, leading to more robust classification results.

To further enhance performance, we employed transfer learning techniques to adapt pretrained CNN models to the new medical image domains. This process allows the models to retain useful general features from their original training (on natural images) while adapting to the specific task of medical image classification. By fine-tuning the pretrained models on the medical X-ray datasets, we can extract more meaningful features specific to fractures and chest diseases. This hybrid approach of combining neutrosophic sets with deep learning models provides a powerful framework for handling complex medical images. In this study, we applied a set of well-established deep learning models with advantages in terms of computational efficiency and feature extraction capabilities to classify medical X-ray images:

- MobileNet: Is a streamlined CNN architecture designed for computational efficiency by utilizing depthwise separable convolutions [31].
- ResNet50: Introduces residual connections, allowing deeper networks by bypassing the vanishing gradient problem. These shortcut connections enable better gradient flow, making it easier to train very deep models [32].
- VGG16: Is a deep CNN model with a uniform layer structure, known for its simplicity and effectiveness in image classification [33].
- DenseNet121: provides enhanced feature propagation and reduced vanishing gradient issues, which makes it possible to retrieve complex image features more reliably [34].
- InceptionV3: This model incorporates techniques like factorized convolutions to enhance efficiency and reduce computational costs without sacrificing performance [35].

These models have been widely used for image classification tasks and have demonstrated strong performance in various medical imaging applications. They were selected for their proven ability to learn complex features from data and their potential to address the challenges of classification tasks with x-ray images. By

employing these models, we harnessed the complementary strengths of these models to classify medical X-ray images, each offering unique benefits in terms of feature extraction and computational efficiency.

3.3. The Proposed NS-DL Approach

The integration of NS with Deep Learning DL techniques presents a promising avenue for enhancing the classification of medical X-ray images. This approach aims to leverage the strengths of both methodologies, addressing uncertainties inherent in medical imaging while utilizing the powerful feature extraction capabilities of deep learning models. By combining NS with DL, it can achieve more robust classification outcomes that account for ambiguous, incomplete, or inconsistent information present in medical images. Neutrosophic Logic offers a framework for representing and managing the uncertainty associated with medical images, allowing for more comprehensive information representation. It enables deep learning to extract relevant features from images while NS addresses the vagueness and ambiguity in pixel information.

The proposed approach consists of a multi-step architecture that seamlessly integrates NS and DL for enhanced medical image classification:

1. Image Acquisition and Preprocessing: Obtain X-ray images and preprocess them to ensure uniformity in size and format.
2. NS Domains Conversion: Apply neutrosophic transformation to the images to convert them into a neutrosophic representation. This involves producing the NS domains: truth, indeterminacy, and falsity, providing a richer representation of the image data.
3. Feature Extraction Using Deep Learning: using pre-trained deep learning models (MobileNet, ResNet50, VGG16, DenseNet121, and inceptionV3) to extract essential features from the neutrosophic images.
4. Classification and Fine-Tuning: Adding a classifier that integrates the features obtained from the deep learning model with the neutrosophic representations, enhancing the model's ability to classify the images.
5. Evaluation and Comparison: Assess the performance of the integrated model against traditional deep learning models that do not utilize NS.

In our case. NS theory can be applied to represent the uncertainty and noise present in the X-rays image data by converting the image into the neutrosophic domain, where each pixel is assigned to three memberships: truth (T), indeterminacy (I), and falsity (F). following the formula:

$$T(x, y) = \mu T(x, y), I(x, y) = \mu I(x, y), F(x, y) = \mu F(x, y) \quad (14)$$

where $\mu T(x, y)$, $\mu I(x, y)$ and $\mu F(x, y)$ are the actual calculated membership functions that map the pixel's values into the truth, indeterminacy, and falsity components respectively. These values are typically normalized between 0 and 1. These three

components represent different aspects of the pixel's characteristics. (T) represents the degree to which a pixel belongs to a region of interest (e.g., a bone or fracture). (I) captures the uncertainty or vagueness in the pixel, which may be due to noise, low contrast, or blurring in the X-ray image. (F) represents the degree to which a pixel does not belong to the region of interest. Fig.1. shows an sample of x-ray images and their corresponding T, I, and F neutrosophic image domains.

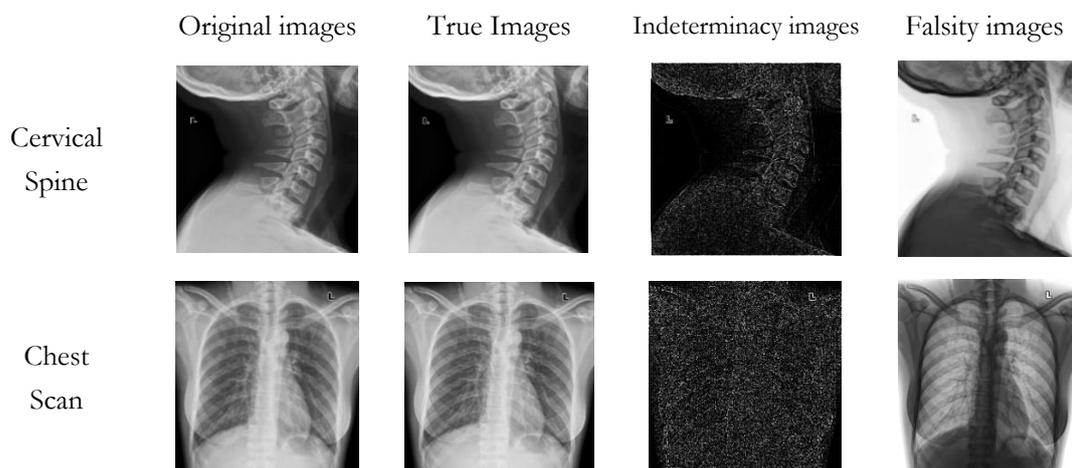


Figure 1: Sample X-ray Images with Corresponding True (T), Indeterminacy (I), and False (F) Neutrosophic Image Domains

Transfer learning technique [29] is used to optimize the performance of these models, where the pretrained versions MobileNet, ResNet50, VGG16, DenseNet121, and InceptionV3, originally trained on the ImageNet dataset, were fine-tuned for our specific medical X-ray datasets. Additionally custom dense layers were added on top of the pretrained models for classification into the target classes of the spine fracture and chest disease datasets. The models were trained using the Adam optimizer, with a categorical crossentropy loss function [36].

$$\text{Minimize: } \text{loss}(CCE) = - \sum_{i=1}^M y_i \cdot \log \check{y}_i \quad (15)$$

where y_i is true value \check{y}_i is shorthand for a vector that contains all the outputs that were predicted based on the training samples. Fine-tuning allowed the models to adapt to the specific features of the medical images while leveraging the powerful feature extraction capabilities they had already learned from ImageNet. Fig.2 show the architecture for the proposed NS-DL approach, While Algorithm.1 showing the main its main steps.

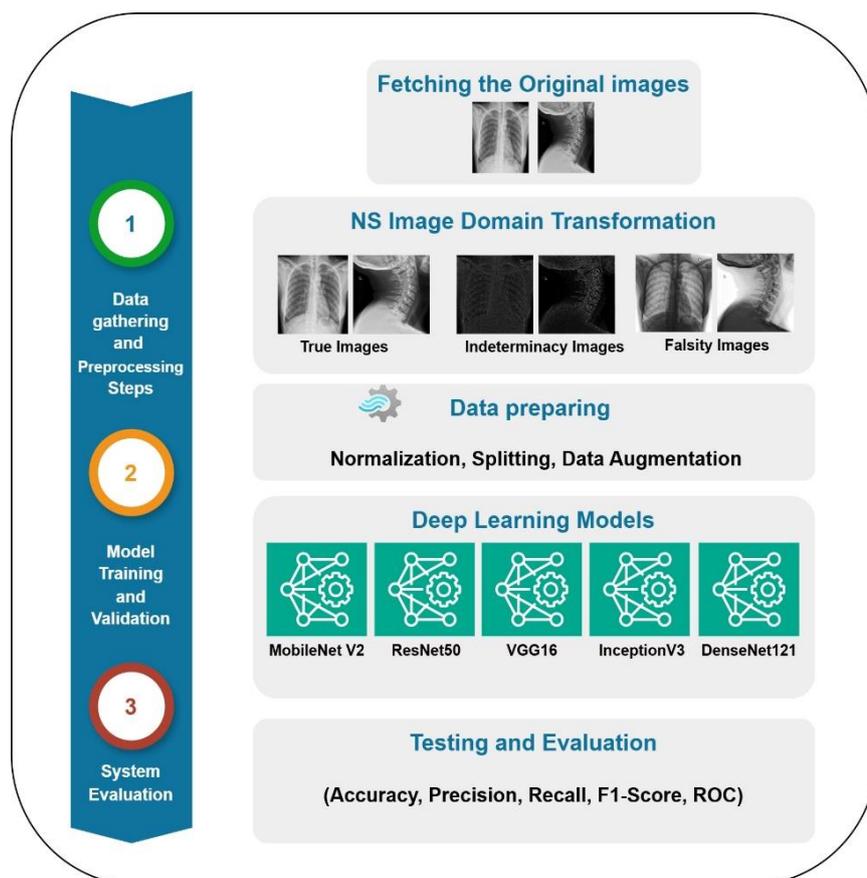


Figure 2: Architecture of the Proposed NS-DL approach

Algorithm 1: The Proposed NS-DL Approach

Input: Input_dataset(G), number of epochs (T).

Output: Model Score (S).

1. Preprocessing the Input Dataset:
 - Convert G to grayscale to simplify pixel intensity calculations and reduce computational complexity.
 2. Pixel-Level Calculations:
 - Calculate pixel intensity $G(x, y)$, which represents the value of each pixel at coordinates (x, y)
 - Calculate local mean O_{min} , which helps identify localized intensity variations in the image.
 3. Generate Neutrosophic Domains: Compute Truth, Indeterminacy, Falsity domains.
 - $T(x, y) = (G(x, y) - G_{min}) / (G_{max} - G_{min})$
 - $I(x, y) = \text{abs}(G(x, y) - O_{min}) / (O_{max} - O_{min})$
 - $F(x, y) = 1 - T(x, y)$
 - Normalize the resulting image output for consistent input to the DL models.
 5. Model Implementation:
 - Utilize five state-of-the-art deep learning models: MobileNet, ResNet, VGG, DenseNet, and Inception
-

-
- Fine-tune each model to adapt to the specific features of the NS images.
6. Train each model separately on the three NS domains images.
 - $t = 0$ // for the current epoch
 - While $t < T$ // number of epochs
 - Computing the score function
 - Check Early stopping conditions (if satisfied \rightarrow end training)
 - Updates the model's weights.
 - $t = t + 1$
7. After training, Calculate the final score S of the model and evaluate it.
-

4. Experimental Setup and Analysis

In this section, we describe the datasets, environment setup, and evaluation metrics used to assess the performance of the proposed approach. The experiments are designed to evaluate the effectiveness of integrating Neutrosophic Sets (NS) with deep learning models (MobileNet, ResNet, VGG16, DenseNet121, and Inception) for classifying X-ray images. We focus on the classification of medical X-rays with the goal of improving diagnostic accuracy through feature extraction and uncertainty management using NS.

4.1. Utilized Dataset

In this study, the proposed models were evaluated using two distinct X-ray image datasets. The first dataset focuses on Cervical Spine Fracture and Dislocation, while the second is utilized for Chest Disease classification [37]. Each dataset comprises images that have been meticulously labeled by medical experts to ensure accuracy. The Chest Disease X-ray Dataset contains images categorized into four labels: Hemothorax, Pneumothorax, Flail Chest, and Normal. In contrast, the Cervical Spine X-ray Dataset includes three labels: Fracture, Dislocation, and Normal. Both datasets consist of images in JPEG and JPG formats, each with a resolution of 256×256 pixels. Table 1 presents the distribution of images for each class in both datasets, while Fig.3 shows the proportional distribution of Image Classes, providing insight into the balance and representation of each category used for training and evaluation.

Table 1: Image Class Distribution for utilized X-ray Datasets

Chest Disease Dataset			Cervical Spine dataset		
Class ID	Class Name	Number of images	Class ID	Class Name	Number of images
0	Flail	525	0	Dislocation	530
1	Hemothorax	422	1	Fracture	772
2	Normal	481	2	Normal	707
3	Pneumothorax	522			

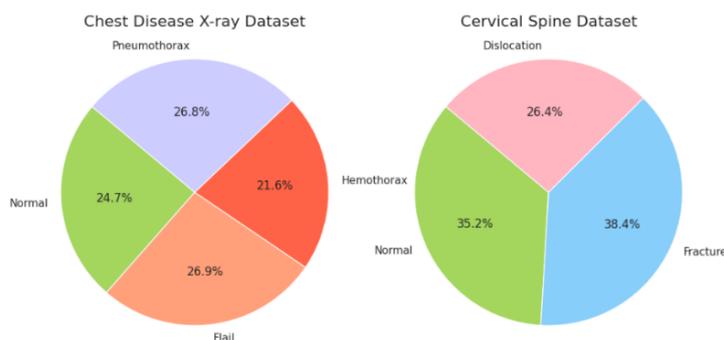


Fig 3: Percentage distributions of Image Classes in X-ray Datasets

In the image preprocessing steps. Each image is converted to grayscale to simplify the data and focus on essential features without color distractions. Following this, neutrosophic transformation is applied, generating three NS domains: True, Indeterminate, and False. This transformation aids in capturing the nuanced details of the X-ray images, facilitating better feature extraction. Pixel normalization is then performed on the transformed images. This process standardizes the pixel values, ensuring they fall within a specific range, it helps accelerate the training process and improves the model's convergence by preventing issues related to varying scales of pixel values. Furthermore, Data augmentation techniques are applied next to enhance the diversity of the training dataset and mitigate the risk of overfitting [38]. Common augmentation strategies include rotation, flipping, zooming, and shifting of images. These techniques allow the model to learn more robust features and improve its generalization to unseen data. Finally, the datasets are split into training, validation, and test sets with a ratio of 70%, 15%, and 15%, respectively. The validation set plays a critical role in monitoring the model's performance during training, allowing for early stopping and hyperparameter tuning without overfitting to the training data. This systematic approach ensures that the model is well-prepared to achieve optimal performance on unseen test data.

4.2. Environment Setup and Hyperparameters Tuning

All experiments are conducted in a computational environment with the following specifications:

- Hardware: NVIDIA Tesla P100 GPU with 16GB of memory, 30GB RAM, and an Intel Xeon processor.
- Software: The experiments are implemented using Python 3.10, TensorFlow 2.15, Karas API 3.3. The system runs on a Kaggle environment.

For deep learning models, the hyperparameters are carefully tuned to optimize performance, all utilized models (MobileNet, ResNet, VGG16, DenseNet121, and Inception) were Initialized with pretrained ImageNet weights, a learning rate of

0.001, batch size of 32, and an Adam optimizer [39]. The model is fine-tuned for 100 epochs with early stopping technique [40]. To fine-tune the pretrained models, we incorporated custom dense layers on top to classify the target categories in both the spine fracture and chest disease datasets. This fine-tuning process enabled the models to adjust to the unique characteristics of the medical images while still benefiting from the robust feature extraction capabilities previously acquired from ImageNet [28]. The specific configurations of the new layers and their hyperparameters for are presented in [Table2](#).

Tabel 2: The configurations of the new layers and hyperparameters settings

Layer	Parameter	Value
Dense layer_1	Number of units	128
	Activation function	Rule
Dense layer_2	Number of units	64
	Activation function	Rule
Dense layer_3	Number of units	N_classes
	Activation function	softmax
	Optimizer	Adam
	Learning rate	0.001
	Epochs	100
	Early stopping	Monitor=loss, Patience=10

4.3.Evaluation Metrics

To evaluate the performance of the proposed approach across the cervical spine fracture and chest disease classification tasks, a comprehensive set of evaluation metrics was employed: (Accuracy, Precision, Recall, F1-Score) These metrics provide insight into the model's effectiveness in identifying and categorizing X-ray images, The mathematical formulas for the utilization of evaluation metrics can be defined as follows:

$$\text{Accuracy} = (TP + TN) / \text{total_predictions} \quad (16)$$

$$\text{Precision} = TP / (TP + FP) \quad (17)$$

$$\text{Recall} = TP / (TP + FN) \quad (18)$$

$$\text{F1 - score} = 2 * (\text{precision} * \text{recall}) / (\text{precision} + \text{recall}) \quad (19)$$

where TP is the correctly classified positive cases, TN is the correctly classified negative cases, FP is the incorrectly classified positive cases, and FN is the incorrectly classified negative cases. These four metrics allow us to assess the performance of different models from various perspectives, *Accuracy* measures the overall correctness of the model by calculating the ratio of correctly classified images to the total number of images, *Precision* evaluates the model's ability to correctly

identify positive instances, providing a focus on the relevance of the true positive predictions. High precision indicates that the model makes fewer false positive errors, which is crucial in medical diagnosis to avoid misidentifying normal cases as diseased, *Recall*, also known as sensitivity or true positive rate, reflects the model's capability to retrieve all relevant instances. A higher recall demonstrates the model's ability to correctly detect diseased cases, making it an essential metric in the context of medical imaging, where missed diagnoses could have severe consequences, and *F1 score* combines precision and recall into a single metric, acting as their harmonic mean. This score is particularly useful when there is an imbalance in the dataset, as it provides a more nuanced evaluation by considering both false positives and false negatives.

Receiver Operating Characteristic (ROC) curves and the associated Area Under the Curve (AUC) are also used to offer a graphical representation of the model's ability to differentiate between the positive and negative classes [41]. The ROC curve plots the true positive rate against the false positive rate, providing insight into the trade-off between sensitivity and specificity at various classification thresholds. A higher AUC value suggests that the model has a strong ability to distinguish between classes, which is critical for clinical decision-making.

$$AUC = \int_0^1 TPR(FPR) d(FPR) \quad (20)$$

where *TPR* is the true positive Rate, and *FPR* is the false positive rate. The Area Under the ROC Curve (AUC-ROC) gives a single scalar value summarizing the performance of the classifier, where the perfect classifier has an AUC-ROC of 1. This collection of metrics provides a comprehensive evaluation of the model's performance, covering both overall accuracy and the balance between false positives and false negatives, and ensures that the evaluation process thoroughly captures the strengths and limitations of the models, it is particularly important in the medical domain, where misclassification can have significant consequences.

5. Results and Discussion

In this section, we present the comparative performance of deep learning models trained on both the original and NS transformed data. The models were evaluated on two medical imaging datasets: spinal injury detection and chest disease classification. The results comparison focuses on the impact of NS image transformation, the classification performance and the. The primary evaluation metrics used include accuracy, precision, recall, and F1-score. Results are broken down by the original dataset and the NS-transformed data, specifically the True (T), Indeterminacy (I), and Falsity (F) domains of the NS theory. We highlight the models' improvements in uncertainty handling, leading to better classification accuracy and overall performance.

5.1. Results in Cervical Spine dataset

The results for the spinal injury detection dataset demonstrate the effectiveness of applying NS transformations to the image data. As shown in Table 3, on the original dataset, MobileNet demonstrated the highest overall accuracy of 99.00%, with F1-score value of 98.96%. Upon applying NS transformations, significant enhancements were observed. In the True (T) domain, both InceptionV3 and DenseNet121 attained notable accuracies of 99.67%, accompanied by F1-scores of 99.64% and 99.69%, respectively. The confusion matrices and ROC curves for InceptionV3 and DenseNet121 are presented in Fig 4 and Fig 5, respectively. Furthermore, in the Indeterminacy (I) and Falsity (F) domains, both DenseNet121 and MobileNet exhibited remarkable improvements, each achieving an accuracy of 99.67% with balanced performance metrics. These results indicate that NS transformations significantly bolster the performance of deep learning models in spinal injury detection, improving their accuracy and robustness across various conditions.

Table 3: Performance Metrics of DL Models on Original and NS Domains C-Spinal Injury Detection Datasets

Data	Model	Accuracy	Precision	Recall	F1-Score
Original Data	MobileNet	99.00	98.81	99.15	98.96
	ResNet50	95.99	95.72	96.15	96.91
	VGG16	98.33	98.21	98.32	98.26
	InceptionV3	96.99	97.05	97.21	97.11
	DenseNet121	99.00	99.81	99.10	99.94
True (T) NS domain Data	MobileNet	99.33	99.18	99.31	99.23
	ResNet50	97.32	97.05	87.31	97.17
	VGG16	98.66	98.63	98.25	98.43
	InceptionV3	99.67	99.60	99.67	99.64
	DenseNet121	99.67	99.68	99.71	99.69
Indeterminacy (I) NS domain Data	MobileNet	99.33	99.17	99.38	99.26
	ResNet50	96.99	96.71	96.89	96.69
	VGG16	99.33	99.05	99.41	99.22
	InceptionV3	98.33	97.81	98.56	98.13
	DenseNet121	99.67	99.60	99.67	99.64
Falsity (F) NS domain data	MobileNet	99.67	99.72	99.57	99.64
	ResNet50	96.99	96.65	97.04	96.83
	VGG16	98.33	97.98	98.13	98.05
	InceptionV3	98.33	98.26	97.75	97.98
	DenseNet121	99.33	99.33	99.31	99.31

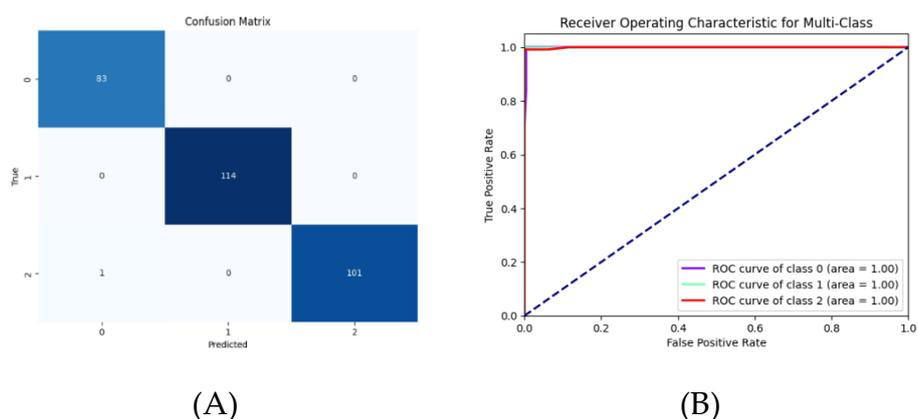


Figure 4: Performance Analysis of the Inception Model on True (T) Domain Images: (a) Confusion Matrix; (b) ROC Curve.

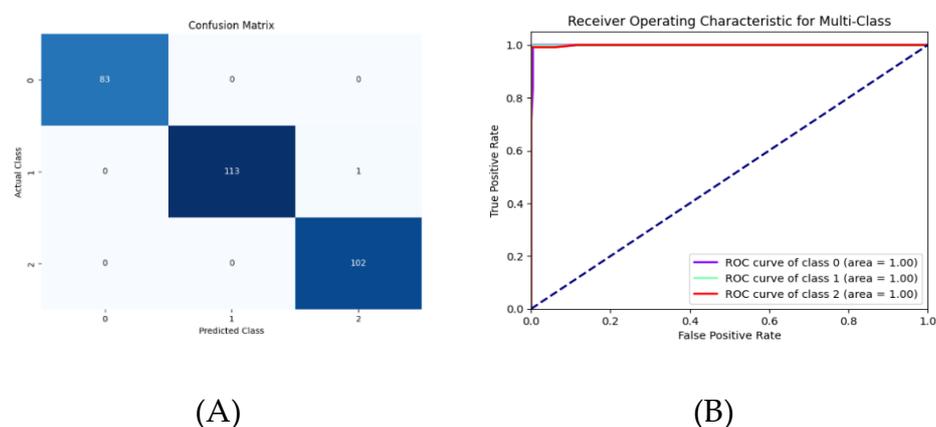


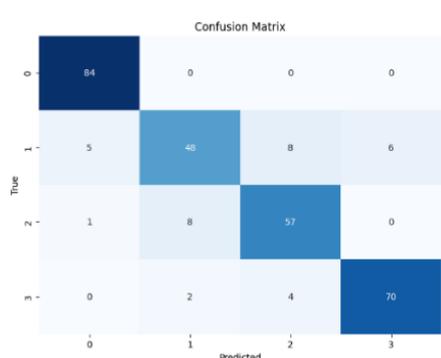
Figure 5: Performance Analysis of the DenseNet Model on True (T) Domain Images: (a) Confusion Matrix; (b) ROC Curve.

5.2. Results in Chest Disease Classification

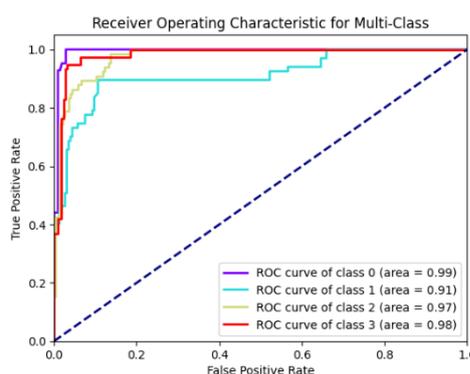
For the chest disease classification dataset, similar trends were observed, with NS-transformed data providing a notable boost in performance. As shown in Table 4, the DenseNet model achieved the highest performance on the original dataset, with an accuracy of 86.69%. In the NS-transformed images, DenseNet excelled in the True domain, achieving an accuracy of 88.40%, followed closely by the MobileNet model in the Indeterminacy domain, which significantly improved from its original performance, attaining an accuracy of 88.05%. The confusion matrices and ROC curves for these two best models are presented in Figures 6 and 7. These results underscore the effectiveness of NS transformations in enhancing the performance of deep learning models for chest disease classification, particularly highlighting DenseNet121's capabilities in accurately detecting conditions within this dataset.

Table 4: Performance Metrics of DL Models on Original and NS Domains Chest Diseases Classification Datasets

Data	Model	Accuracy	Precision	Recall	F1-Score
Original Data	MobileNet	85.32	86.27	85.25	85.20
	ResNet50	75.43	75.68	75.37	75.43
	VGG16	83.96	84.39	83.79	83.34
	InceptionV3	83.28	82.49	82.44	82.37
	DenseNet121	86.69	86.43	86.21	86.06
True (T) NS domain Data	MobileNet	87.37	86.97	86.73	86.64
	ResNet50	76.79	76.07	75.78	75.50
	VGG16	84.30	84.62	84.10	84.28
	InceptionV3	86.35	86.24	85.95	85.62
	DenseNet121	88.40	87.70	87.53	87.48
Indeterminacy (I) NS domain Data	MobileNet	88.05	87.46	87.19	87.19
	ResNet50	79.86	79.65	79.98	79.39
	VGG16	81.91	81.69	81.40	81.40
	InceptionV3	83.62	83.24	82.68	82.78
	DenseNet121	87.37	87.07	86.63	86.65
Falsity (F) NS domain data	MobileNet	86.69	86.20	86.58	86.32
	ResNet50	76.45	76.28	75.93	75.21
	VGG16	83.96	83.41	83.52	82.94
	InceptionV3	82.94	82.84	83.85	82.74
	DenseNet121	87.71	87.47	88.26	87.63



(A)



(B)

Figure 6: Performance Analysis of the DenseNet Model on True (T) Domain Images: (a) Confusion Matrix; (b) ROC Curve.

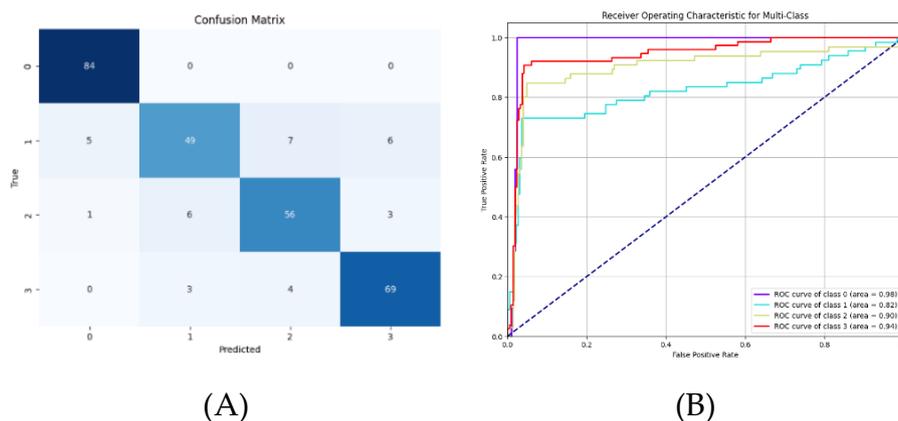


Figure 7: Performance Analysis of the MobileNet Model on Indeterminacy (I) Domain Images: (a) Confusion Matrix; (b) ROC Curve.

5.3. Summary of Experiments

The results of the experiments clearly demonstrate the positive impact of Neutrosophic Set (NS) transformations on the performance of deep learning models for medical image classification across both the cervical spine and chest disease classification datasets. For the C-spinal injury detection dataset, NS-transformed models exhibited significant improvements in performance metrics. DenseNet121 and Inception emerged as the top performer in the True (T) domain, achieving an accuracy of 99.67%, while the original dataset accuracy was 99.00%, 96.99%, respectively. Inception also showed notable enhancement, achieving an accuracy of 99.33% in the Indeterminacy domain compared to its original accuracy of 99.00%. These improvements underscore the effectiveness of NS transformations in reducing uncertainty and enhancing the models' ability to extract relevant features. In the chest disease classification dataset, DenseNet121 again demonstrated the highest accuracy among the models, achieving 88.40% on the True domain, a significant improvement from 86.69% on the original dataset. MobileNet performed well too, with an accuracy of 87.37% on the True domain, showing a marked increase from 85.32% on the original data. Other models, such as InceptionV3 and VGG16, also benefited from NS transformations, indicating that these methods enhance overall classification performance across various conditions. These results highlight the potential of integrating NS theory into medical image classification workflows, particularly in scenarios characterized by uncertainty, enabling more accurate capture of key features associated with spinal injuries and chest diseases, which could lead to improved diagnostic accuracy in real-world clinical applications.

6. Conclusion and future Work

This study presents a hybrid approach that integrates Neutrosophic Set (NS) theory with deep learning models to enhance the classification of medical X-ray images across two challenging datasets: cervical spine injuries and chest disease classification. Traditional deep learning models often struggle with the inherent

noise and uncertainty present in medical images, leading to reduced diagnostic accuracy. By incorporating NS theory, which categorizes information into True (T), Indeterminate (I), and False (F) domains, we effectively addressed these challenges. Our results demonstrate that models trained on NS-transformed data consistently outperformed those trained on original datasets. Specifically, DenseNet121 achieved the highest accuracy in both datasets, with 88.40% for chest disease classification and 99.97% for cervical spine injury detection. MobileNet and Inception modes also exhibited significant improvements in the different NS domain. These findings highlight the effectiveness of combining NS theory with state-of-the-art deep learning models, as the transformed data allowed for better feature extraction and more accurate predictions across various medical imaging scenarios. The incorporation of NS theory not only enhances the ability to manage uncertainty in medical images but also provides a robust solution for improving diagnostic accuracy. However, this study has certain limitations. One limitation is the reliance on only X-ray medical datasets (cervical spine and chest diseases). While these datasets are valuable, they may not fully represent the diversity and complexity of medical images encountered in real-world clinical environments. The generalizability of our approach to other medical imaging domains requires further investigation. Additionally, while NS theory provides an improved mechanism for handling uncertainty, it may introduce extra computational overhead, which could be a challenge for resource-constrained environments. Future work will focus on expanding the application of NS theory to additional medical imaging modalities, such as MRI and CT scans, to further validate its effectiveness. Additionally Future studies will also aim to optimize the computational efficiency of the NS-DL approach, particularly in terms of processing time and resource utilization. This will help ensure that our approach is not only accurate but also feasible for deployment in real-time clinical settings.

Funding: This research received no external funding.

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

References

1. Abhisheka, B., et al. (2024). Recent trend in medical imaging modalities and their applications in disease diagnosis: a review. *Multimedia Tools and Applications*, 83(14), 43035-43070.
2. Adams, S.J., et al. (2021). Artificial intelligence solutions for analysis of X-ray images. *Canadian Association of Radiologists Journal*, 72(1), 60-72.
3. Zou, K., et al. (2023). A review of uncertainty estimation and its application in medical imaging. *Meta-Radiology*, 100003.
4. Sambyal, A.S., N.C. Krishnan, and D.R. Bathula. (2022). Towards reducing aleatoric uncertainty for medical imaging tasks. Paper presented at the 2022 IEEE 19th International Symposium on Biomedical Imaging (ISBI).

5. Kayalibay, B., G. Jensen, and P. van der Smagt. (2017). CNN-based segmentation of medical imaging data. arXiv preprint arXiv:1701.03056.
6. Salehi, A.W., et al. (2023). A study of CNN and transfer learning in medical imaging: Advantages, challenges, future scope. *Sustainability*, 15(7), 5930.
7. Hassan, E., et al. (2023). COVID-19 diagnosis-based deep learning approaches for COVIDx dataset: A preliminary survey. *Artificial intelligence for disease diagnosis and prognosis in smart healthcare*, 107-122.
8. Rodríguez-Sánchez, Á., et al. (2020). Review of the influence of noise in X-ray computed tomography measurement uncertainty. *Precision Engineering*, 66, 382-391.
9. Gour, M. and S. Jain. (2022). Uncertainty-aware convolutional neural network for COVID-19 X-ray images classification. *Computers in biology and medicine*, 140, 105047.
10. Smarandache, F. (1999). A unifying field in Logics: Neutrosophic Logic Philosophy (pp. 1-141): American Research Press.
11. Zadeh, L.A. (1965). Fuzzy sets. *Information and control*, 8(3), 338-353.
12. Smarandache, F. (2015). Neutrosophic masses & indeterminate models. *Advances and Applications of DSmT for Information Fusion*, 133.
13. Guo, Y. and A.S. Ashour. (2019). Neutrosophic sets in dermoscopic medical image segmentation Neutrosophic set in medical image analysis (pp. 229-243): Elsevier.
14. Guo, Y. and A.S. Ashour. (2019). Neutrosophic set in medical image analysis: Academic Press.
15. Mostafa, N.N., A.K. Kumar, and Y. Ali. (2024). A comparative study on x-ray image enhancement based on neutrosophic set. *Sustainable Machine Intelligence Journal*, 7, (2): 1-10.
16. Cai, G., et al. (2019). Neutrosophic set-based deep learning in mammogram analysis Neutrosophic Set in Medical Image Analysis (pp. 287-310): Elsevier.
17. Khalifa, N.E.M., et al. (2021). A study of the neutrosophic set significance on deep transfer learning models: An experimental case on a limited covid-19 chest x-ray dataset. *Cognitive Computation*, 1-10.
18. Hassan, E., et al. (2024). Detecting COVID-19 in chest CT images based on several pre-trained models. *Multimedia Tools and Applications*, 1-21.
19. Samee, N.A., et al. (2022). Metaheuristic Optimization Through Deep Learning Classification of COVID-19 in Chest X-Ray Images. *Computers, Materials & Continua*, 73(2).
20. El-Shahat, D. and A. Tolba. (2024). Assessment of deep learning techniques for bone fracture detection under neutrosophic domain. *Neutrosophic Sets and Systems*, 68, 109-135.
21. Jennifer, J.S. and T.S. Sharmila. (2023). A neutrosophic set approach on chest X-rays for automatic lung infection detection. *Information Technology and Control*, 52(1), 37-52.
22. Singh, S.K., V. Abolghasemi, and M.H. Anisi. (2022). Skin cancer diagnosis based on neutrosophic features with a deep neural network. *Sensors*, 22(16), 6261.
23. Yasser, I., et al. (2022). A hybrid automated intelligent COVID-19 classification system based on neutrosophic logic and machine learning techniques using chest X-Ray images. *Advances in Data Science and Intelligent Data Communication Technologies for COVID-19: Innovative Solutions Against COVID-19*, 119-137.

24. Guo, Y., A.I. Shahin, and H. Garg. (2024). An indeterminacy fusion of encoder-decoder network based on neutrosophic set for white blood cells segmentation. *Expert Systems with Applications*, 246, 123156.
25. Özyurt, F., et al. (2019). Brain tumor detection based on Convolutional Neural Network with neutrosophic expert maximum fuzzy sure entropy. *Measurement*, 147, 106830.
26. Guo, Y. and H.-D. Cheng. (2009). New neutrosophic approach to image segmentation. *Pattern Recognition*, 42(5), 587-595.
27. Wu, Y., et al. (2013). Local Shannon entropy measure with statistical tests for image randomness. *Information Sciences*, 222, 323-342.
28. Krizhevsky, A., I. Sutskever, and G.E. Hinton. (2012). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25.
29. Torrey, L. and J. Shavlik. (2010). Transfer learning Handbook of research on machine learning applications and trends: algorithms, methods, and techniques (pp. 242-264): IGI global.
30. Kornblith, S., J. Shlens, and Q.V. Le. (2019). Do better imagenet models transfer better? Paper presented at the Proceedings of the IEEE/CVF conference on computer vision and pattern recognition.
31. Howard, A.G. (2017). Mobilenets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:1704.04861.
32. He, K., et al. (2016). Deep residual learning for image recognition. Paper presented at the Proceedings of the IEEE conference on computer vision and pattern recognition.
33. Simonyan, K. and A. Zisserman. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.
34. Huang, G., et al. (2017). Densely connected convolutional networks. Paper presented at the Proceedings of the IEEE conference on computer vision and pattern recognition.
35. Szegedy, C., et al. (2017). Inception-v4, inception-resnet and the impact of residual connections on learning. Paper presented at the Proceedings of the AAAI conference on artificial intelligence.
36. Zhang, Z. and M. Sabuncu. (2018). Generalized cross entropy loss for training deep neural networks with noisy labels. *Advances in neural information processing systems*, 31.
37. . X-ray ChestPelvis CSpine Scans dataset. Retrieved Accessed on 5 June 2024 . Available online; from <https://www.kaggle.com/datasets/pardonndlovu/chestpelviscspinescans>
38. Mikołajczyk, A. and M. Grochowski. (2018). Data augmentation for improving deep learning in image classification problem. Paper presented at the 2018 international interdisciplinary PhD workshop (IIPhDW).
39. Jais, I.K.M., A.R. Ismail, and S.Q. Nisa. (2019). Adam optimization algorithm for wide and deep neural network. *Knowl. Eng. Data Sci.*, 2(1), 41-46.
40. Prechelt, L. (2002). Early stopping-but when? *Neural Networks: Tricks of the trade* (pp. 55-69): Springer.
41. Carrington, A.M., et al. (2021). Deep ROC analysis and AUC as balanced average accuracy to improve model selection, understanding and interpretation. arXiv preprint arXiv:2103.11357.

Received: Sep 3, 2024. Accepted: Feb 16, 2025