Neutrosophic Sets and Systems, Vol. 100, 2026
University of New Mexico

Neutrosophic Approaches in Deep Learning for Ambiguous Medical Image Analysis: A Systematic Review Toward Wound Assessment

Zakaria Ahmed [1,*] and Reda Ahmed[2]

[1,2] Department of Computer science, Faculty of Computer and Informatics, Zagazig University, 44519, Egypt; 20912019100575@fci.zu.edu.eg.
***Correspondence:** 20912019100575@fci.zu.edu.eg.

Abstract

Deep Learning (DL) has contributed to the automated image analysis of wounds and skin lesions, but common architectures have not integrated an explicit representation of boundary ambiguity, tissue heterogeneity, and acquisition noise in images. Neutrosophic theory is relevant to address this problem by representing every pixel/region with an independent triplet (truth T, indeterminacy I, and falsity F memberships). This systematic literature review (SLR) conducted according to the PRISMA 2020 statement aims at mapping and critical analysis of the scientific literature hybridizing neutrosophic sets/logic and DL for wound/skin lesion/related soft tissues image analysis. A structured and multisource literature search over the time interval from the inception of the theoretical neutrosophic set theory until now revealed a small but coherent methodologically speaking literature; after two stages of selection according to clear criteria, 15 papers were selected for analysis. A four-axis taxonomy (level of integration, clinical application task, neutrosophic theory variant, and dataset imaging domain) is developed and populated using the selected literature. Among others, a significant result is that there is not a single study satisfying the inclusion criteria that addresses wound imaging; the body of evidence includes dermoscopy images and images of skin lesions, MRI brain tumor images, thyroid ultrasound images, radiographic bone fractures, and chest radiography, all of which have an identical ambiguity boundary-based structure compared to wound assessment. Neutrosophic fusion is shown to occur almost exclusively at the preprocessing and clustering phases of the DL workflow; loss-function-based and end-to-end approaches to neutrosophic-DL have not been explored yet. The indeterminacy (I) channel is revealed to show empirically well-justified diagnostic capacity better than the truth channel alone in some studies — thus being positioned as a potential intrinsic interpretability modality beyond post hoc explanation approaches. The review concludes by summarizing the results obtained into a critically reviewed comparative synthesis, highlighting translational challenges (high computational costs, lack of standardization, unexplored explainability synergy potential, and lack of deployment on the edge evidence) and providing a research agenda with prioritized needs (especially creation of a publicly available wound image benchmark including neutrosophic ground truth).

Keywords: Neutrosophic Set; Deep Learning; Wound Image Analysis; Skin Lesion Segmentation; Medical Image Uncertainty; Indeterminacy.

1. Introduction

Chronic and acute wounds like diabetic foot ulcers, pressure ulcers, venous ulcers, surgical wounds, and burns have to be visually assessed regularly in order to monitor their tissue composition and healing dynamics and cause a significant burden on global healthcare systems. Visual inspection by clinicians is still the gold standard and is subjective and inconsistent; hence, there has been a lot of wound image analysis research utilizing deep learning methods, including segmentation approaches like Detect-and-Segment [1] and tissue segmentation for diabetic foot ulcers [2, 3]. The problem with dermoscopic skin lesion images is the same segmentation of pathological lesions from normal tissues in various lighting conditions and under artifacts and occlusions and is considered a similar evidentiary field in this review (Section 4).

However, despite their excellent benchmark performance, several issues of structural nature can be identified in using DL classifiers and segmenters on wound and skin lesion images. These are: first, the lack of clear distinction between boundaries of wound bed, peri-wound skin, slough, and granulation tissue is gradual and non-linear [4]; second, softmax/sigmoid activation functions do not give a means of measuring epistemic uncertainty; in other words, there is no way to distinguish between ambiguous regions and the regions that are confidently classified in wrong ways. Finally, the data set is relatively small, inconsistently labeled, and unbalanced for skin types and wounds [1, 3]. The solution to this particular problem that of representing ambiguity as opposed to assigning a clear classification has been developed over forty years in a line of fuzzy sets, intuitionistic fuzzy sets, and finally, neutrosophic sets, as shown in Section 2.

1.1 Positioning Against Prior Surveys

The Table 1 below highlights the relation between this review and the most similar surveys. The hybridization survey [5] focuses on DL-neutrosophic hybrids in the context of medical imaging in general but does not focus on the application of neutrosophic theory and DL in the context of wound and skin injury imaging and does not utilize the PRISMA guidelines. General surveys of wound image analysis focus solely on DL while omitting any mention of neutrosophic or any other uncertainty formalism. Classical surveys of neutrosophic image processing predate the era of deep learning and hence overlook its possible use.

Table 1. Positioning of this review against the closest related prior surveys.

Prior Work	Covers DL	Covers Neutrosophy	Covers Wound/Skin Domain	PRISMA / Taxonomy
Mostafa et al. [5], 2021 (medical imaging, general)	Yes	Yes	Partial (not focused)	No / No
General wound-image-analysis DL surveys (e.g., underlying [1, 3])	Yes	No	Yes	Varies / No

Classical neutrosophic image-processing surveys (pre-DL era, e.g., foundations in [6, 7])	No	Yes	No	No / No
This review	Yes	Yes	Yes (with disclosed domain-broadening, §3.2)	Yes / Yes

1.2 Contributions

C1 : The first PRISMA 2020-compliant SLR scoped specifically to neutrosophic-DL hybridization for wound and skin-injury image analysis.

C2 : A historically grounded synthesis tracing how neutrosophic theory evolved from, and mathematically improves upon, fuzzy and intuitionistic fuzzy uncertainty modeling, and how this evolution was integrated into medical deep learning.

C3 : A four-axis taxonomy (integration level, clinical task, neutrosophic variant, dataset/imaging domain) for classifying the identified literature.

C4 : A consolidated, quality-appraised comparative synthesis of all included primary studies.

C5 : A stage-by-stage technical analysis of how indeterminacy (I) is computed, propagated, and exploited across the DL pipeline.

C6 : An evidence-based discussion of translation barriers: computational overhead, absent standardization, untested explainability synergy, and unproven edge-deployment readiness.

C7 : A prioritized future-research roadmap, foremost calling for a dedicated wound-specific neutrosophic-DL benchmark.

1.3 Structure of the paper is the following

The rest of the paper is structured as follows. In Section 2, the theoretical underpinnings and evolution of neutrosophy are reviewed, the neutrosophic ⟨T, I, F⟩ model is described, and its inclusion in medical image analysis and deep learning is explained. The PRISMA 2020 protocol is described in Section 3. The taxonomy is presented in Section 4. In Section 5, the technical synthesis and master comparison are given. Clinical feasibility and challenges are discussed in Section 6. Limitations are discussed in Section 7. Conclusion and future work follow in Section 8.

2. Background

This part lays out the background required to explain why the particular mathematical theory called neutrosophic, among all others, became the theoretical framework used in this literature for representing uncertainty in medical images. It provides an account of how each element of the theory became incorporated into medical image analysis using deep learning through time, in five stages: (2.1) the forty years of development of uncertainty theories preceding neutrosophic theory; (2.2) the philosophical basis and the definition of the neutrosophic set; (2.3) the computational tools within the neutrosophic theory; (2.4) the particular image transformation technique implementing this

framework to pixel values; and (2.5) the history of adoption of each of these elements into medical deep learning until this review's point of focus on wound imaging.

2.1 From Classical to Fuzzy to Intuitionistic Fuzzy Sets

Classical ("crisp") set theory admits only binary membership: an element either belongs to a set or it does not. Zadeh's fuzzy set theory [8] relaxed this by allowing a graded membership function:

$$A = \{ (x, \mu A(x)) \mid x \in X \}, \quad \mu A(x) \in [0, 1] \tag{1}$$

where μ_A(x) is the degree to which x belongs to A. Applied to a wound image, μ_A(x) might express "how much" a pixel resembles wound tissue already an improvement over a crisp mask, but it conflates two distinct situations: a pixel that is genuinely borderline between wound and healthy tissue, and a pixel about which the model is simply uncertain. Atanassov's intuitionistic fuzzy set (IFS) [9] partially addressed this by adding an explicit non-membership degree ν_A(x), independent of μ_A(x) up to a joint constraint:

$$A = \{ (x, \mu A(x), \nu A(x)) \mid x \in X \}, \quad \mu A(x) + \nu A(x) \leq 1 \tag{2}$$

with the residual hesitancy degree π_A(x) = 1 − μ_A(x) − ν_A(x) capturing whatever indeterminacy remains. This was a genuine conceptual advance membership and non-membership are no longer forced to be complementary but π_A(x) is not a free, independently specifiable quantity: it is whatever is left over once μ_A(x) and ν_A(x) are fixed. For a wound-boundary pixel, this means indeterminacy can only ever be reported as a residual, never as a primary, directly measured signal in its own right. This residual-only treatment of indeterminacy is precisely the gap that motivated Smarandache's neutrosophic set.

2.2 The Neutrosophic Set

Smarandache introduced neutrosophy as a philosophical framework for reasoning about a proposition together with its opposite and its neutral, indeterminate middle ground, and formalized the neutrosophic set as a mathematical instantiation of this idea [10, 11]. The defining move, relative to Section 2.1, is to make indeterminacy a first-class, independently valued component rather than a derived residual. In its original, philosophically general form, a neutrosophic set A on universe X assigns each element three membership functions T_A, I_A, F_A drawn from the non-standard unit interval, subject only to a loose sum bound:

$$0^- \leq TA(x) + IA(x) + FA(x) \leq 3^+ \tag{3}$$

with T_A, I_A, F_A permitted to vary completely independently of one another. This is the crucial theoretical departure from Section 2.1:

I_A(x) is no longer forced to be 1 − μ_A(x) − ν_A(x); it can be measured and reported on its own terms. For engineering and image-processing use where the non-standard analysis of Eq. (3) is unnecessary machinery the field almost universally adopts the Single-Valued Neutrosophic Set (SVNS) simplification, restricting each component to the standard real unit interval:

$$TA(x), IA(x), FA(x) \in [0, 1], \quad 0 \leq TA(x) + IA(x) + FA(x) \leq 3 \tag{4}$$

Table 2 summarizes the representational capacity this progression achieves each successive theory adds exactly the degree of freedom the previous one lacked, with the neutrosophic set the only one of the three in which indeterminacy is both explicit and fully independent.

Table 2. Representational capacity of fuzzy, intuitionistic fuzzy, and neutrosophic sets.

Property	Fuzzy Set	Intuitionistic Fuzzy Set	Neutrosophic Set
Membership (T/μ)	Explicit	Explicit	Explicit
Non-membership (F/ν)	Implicit (1−μ)	Explicit	Explicit
Indeterminacy (I/π)	Not represented	Residual only (π = 1−μ−ν)	Explicit and independent
Sum constraint	—	μ + ν ≤ 1	0 ≤ T+I+F ≤ 3 (no coupling)

2.3 Formal Neutrosophic Measures Used in the Reviewed Literature

Beyond the base definition, three derived measures recur throughout the neutrosophic-DL literature reviewed in Section 5 and are worth formalizing here. First, to rank or threshold a neutrosophic-valued pixel or decision back into a single comparable quantity, a score function is used, most commonly:

$$s(A) = (2 + TA - IA - FA) / 3 \quad (5)$$

alongside an accuracy function a(A) = T_A − F_A that ranks elements once scores tie. Second, when two neutrosophic representations must be compared — for instance, a candidate pixel against a class prototype a neutrosophic Euclidean distance is used:

$$d(A, B) = \sqrt{[\,(1/3)\,((TA - TB)^2 + (IA - IB)^2 + (FA - FB)^2)\,]} \quad (6)$$

with related cosine-similarity and correlation-coefficient measures formalized for SVNS decision-making by Ye [12], and subsequently re-used, in adapted form, as the Neutrosophic Similarity Score training-reweighting mechanism discussed in Section 5.2. Third, converting a full ⟨T,I,F⟩ triplet back to a single crisp output termed de-neutrosophication, by analogy with fuzzy defuzzification is most simply achieved via a weighted combination such as T_A − F_A · (1 − I_A), which discounts the truth−falsity contrast in proportion to how indeterminate the pixel is; Section 5.2 discusses how several reviewed studies instead keep the three channels separate throughout, precisely to avoid collapsing this information prematurely.

2.4 The Neutrosophic Image Transform: From Pixel to (T, I, F)

Guo and Cheng operationalized the SVNS formalism of Eq. (4) specifically for image data [6], mapping a pixel's local intensity, local homogeneity, and local non-homogeneity respectively onto the T, I, and F sub-images:

$$T(x, y) = [\, \bar{g}(x, y) - \bar{g}\min \,] / [\, \bar{g}\max - \bar{g}\min \,] \quad (7)$$

$$I(x, y) = [\, \delta(x, y) - \delta\min \,] / [\, \delta\max - \delta\min \,], \quad \delta(x, y) = |\, g(x, y) - \bar{g}(x, y) \,| \quad (8)$$

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

where g(x, y) is the raw pixel intensity, $\bar{g}(x, y)$ a local-window mean, and δ(x, y) the local homogeneity/inhomogeneity measure driving indeterminacy. The three resulting sub-images are most commonly stacked as a three-channel pseudo-RGB tensor, allowing unmodified CNN backbones to ingest the neutrosophic representation directly (Section 5.2, Fig. 1). Denoising [13] and fuzzy-c-means-based segmentation refinements [14] extended this same T/I/F decomposition before neutrosophic c-means (NCM) clustering [15] generalized it into a standalone unsupervised segmentation algorithm — the specific mechanism most re-used by the DL-hybrid studies in Section 5. Because I(x, y) is computed analytically and deterministically from image structure via Eq. (8), rather than inferred from model stochasticity as in Monte-Carlo dropout or deep ensembles, it is structurally distinct from post-hoc DL uncertainty-estimation techniques — a distinction central to the interpretability discussion in Section 6.

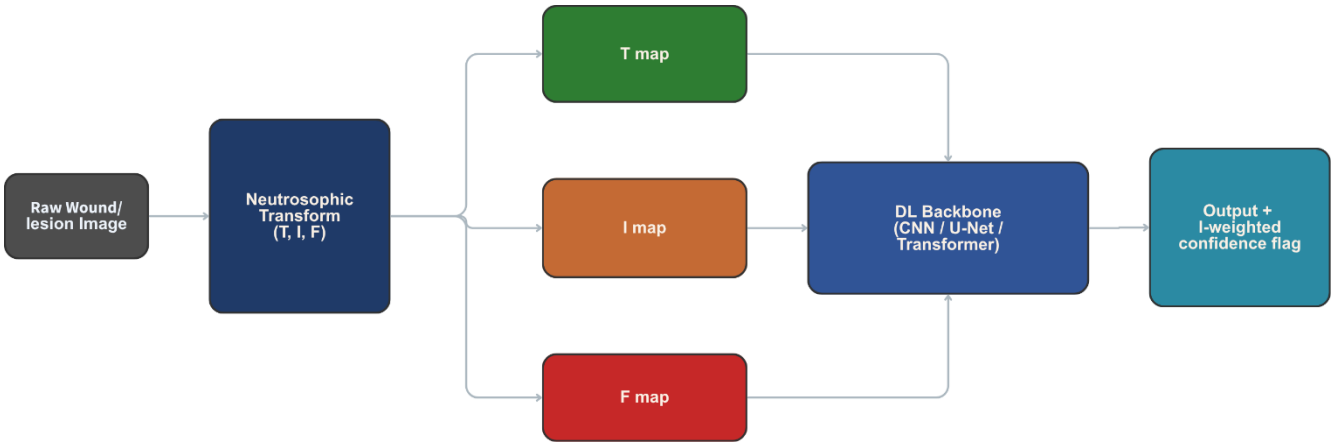

Figure 1. Raw image is decomposed into T/I/F maps and integrated into the DL backbone, with I serving as a confidence signal.

Illustrative worked example

Consider a 3×3 pixel neighborhood centered on a peri-wound boundary, with raw intensities g(x,y) ranging from 60 (clearly healthy skin) to 200 (clearly necrotic tissue) and a local mean $\bar{g}$ = 130. Applying Eq. (7) with observed $\bar{g}\min$ = 40 and $\bar{g}\max$ = 220 gives T ≈ (130−40)/(220−40) = 0.50 a mid-range truth value reflecting genuine ambiguity. If the local homogeneity measure δ for this neighborhood is high (the neighborhood mixes both tissue types), Eq. (8) yields a correspondingly high I, e.g., I ≈ 0.70, while F = 1 − T = 0.50 from Eq. (9). A conventional crisp classifier forced to threshold T alone at 0.5 would arbitrarily assign this pixel to either class; the neutrosophic

representation instead retains ⟨0.50, 0.70, 0.50⟩ and, via the score function of Eq. (5), s(A) = (2+0.50−0.70−0.50)/3 ≈ 0.43 a below-midpoint score that can be used to flag the pixel for clinician review rather than silently resolved. This is the concrete mechanism, worked at the single-pixel level, that Sections 4 , 6 examine at the scale of full studies and datasets.

2.5 A Complete Worked Scenario: From Pixel to Clinical Triage Decision

To make the end-to-end reasoning fully concrete, consider three representative regions cropped from a single hypothetical wound photograph, each pushed through Eqs. (7), (9) and then through the score function of Eq. (5), yielding a three-way triage decision rather than a binary label. Region R1 sits well inside clearly healthy peri-wound skin; R2 sits well inside the necrotic wound bed; R3 straddles the boundary between slough and granulation tissue precisely the kind of region a clinician would pause over during manual assessment. Table 3a reports the resulting T, I, F, and score values and the triage action a simple decision rule (accept if $s \geq 0.66$; refer if $0.33 \leq s < 0.66$; reject if $s < 0.33$) would take for each.

Table 3a. Worked scenario applying Eqs. (7) , (9) and the score function of Eq. (5) to three representative wound-image regions.

Region	T	I	F	s(A)	Triage Decision (Eq. 5)
R1 healthy skin	0.95	0.08	0.05	0.94	Accept as healthy (high confidence)
R2 necrotic wound bed	0.90	0.12	0.10	0.89	Accept as wound tissue (high confidence)
R3 slough/granulation boundary	0.50	0.70	0.50	0.43	Refer for clinician review (high indeterminacy)

The difference between R1/R2 and R3 encapsulates the whole practical argument for this review in a nutshell. A typical DL-based segmentation algorithm with a simple T-type confidence measure could theoretically also designate R3 as a "low confidence" instance depending on whether the softmax output was close to 0.5 – but there would be no way of knowing that this does not reflect class overlap as in R3 but instead out-of-distribution artifact or motion blur or sensor noise, since both would be reduced to the same single number. The neutrosophic model retains the ability to differentiate these cases by having a high I-value (0.70 in R3) that indicates structural ambiguity of the statistics of the image itself (see Eq. 8), and not merely low classifier confidence, which is the very diagnostic information needed by the wound-care practitioner. The discussion in Section 5.2 is focused on the fraction of the twelve primary studies that retain this ability to differentiate.

2.6 Summary of Notation

Table 3b consolidates the symbols introduced in Sections 2.1 , 2.4 for reference throughout the remainder of the paper.

Table 3b. Summary of notation used throughout Sections 2, 5, and 6.

Symbol	Meaning	Defined In
μ_A(x)	Fuzzy membership degree of x in set A	Eq. (1)
μ_A(x), ν_A(x), π_A(x)	Intuitionistic fuzzy membership, non-membership, and residual hesitancy	Eq. (2)
T_A(x), I_A(x), F_A(x)	Neutrosophic truth, indeterminacy, and falsity membership (independent)	Eqs. (3),(4)
s(A), a(A)	Neutrosophic score function and accuracy function	Eq. (5)
d(A,B)	Neutrosophic Euclidean distance between two neutrosophic values	Eq. (6)
g(x,y), $\bar{g}$(x,y), δ(x,y)	Raw pixel intensity, local-window mean, and local homogeneity measure	Eqs. (7),(8)
T(x,y), I(x,y), F(x,y)	Pixel-level neutrosophic image sub-domains	Eqs. (7),(9)

2.7 Evolutionary Timeline

Figure. 2 charts the resulting integration history from beginning to end. While Section 2.1's (1965–1986) theoretical heritage provided the vocabulary, it was Smarandache's (1998- 2003) formalization which provided the independent-indeterminate formalism [11] itself; the Guo and Cheng (2009) image transform (Eqs. 7, 9) made that formalism compute-able in pixel space; the years 2011-2015 completed the neutrosophic clustering into NCM as a segmented-primitive; Ye (2013)'s SVNS measures provided the scoring, distance, and similarity formalism of Section 2.3; and, finally, these elements were brought together with convolutional neural network backbones in the first neutrosophic-DCNN structures as early as 2019 [7, 16] which were all brought together into the field's first edited volume the same year. 2021 brought a dedicated hybridization survey [5], and 2022–2025 brought the accelerated application quantified in Section 5.1 below to skin lesions, brain tumors, bones, thyroids, and chest imagery. However, at no point during this thirty-year integration process did that trajectory ever reach wound imagery specifically: the gap filled by this SLR in Sections 3–6 is not a failure of this survey's search, but rather the very endpoint of its timeline.

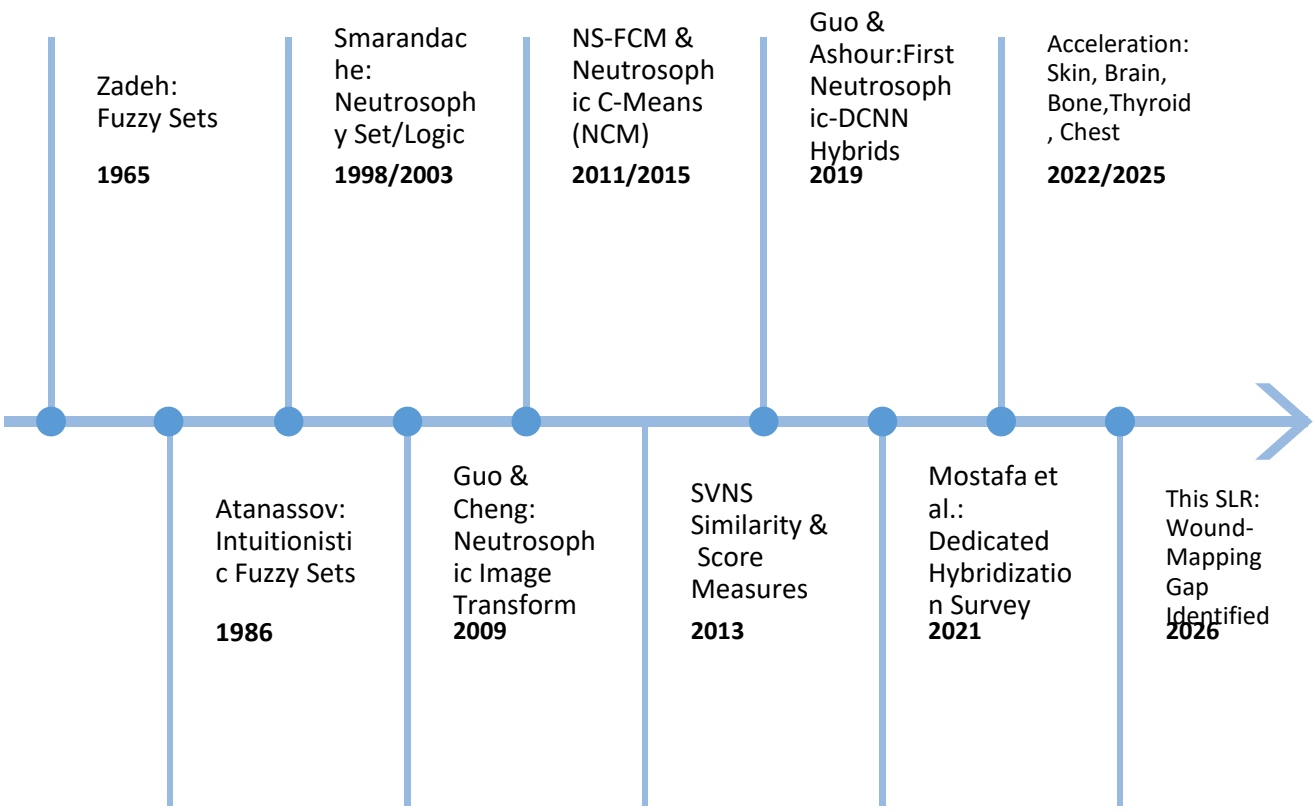

Figure. 2. Historical evolution and integration of neutrosophic theory into medical deep learning, from its fuzzy-set antecedents to the wound-imaging gap identified by this review.

2.8 Why Neutrosophic Theory Fits Wound and Skin-Injury Imaging Specifically

There are three aspects of wound and skin lesion images that make them an especially good fit to the theory described above, and which ensure that all four studies from Section 5, while not being applied specifically to wounds, can be used directly in any case. First, there is a multi-value and non-complementary nature of the progression of the condition (healthy skin → erythema near wound → slough/fibrin → granulation → necrotic tissue). Second, wound photographs are acquired in an uncontrolled environment with different lighting conditions, camera devices, dressing obfuscation, etc., thus increasing the local inhomogeneity term $\delta(x,y)$ of Eq. (8), which means that the indeterminacy channel is an actual, computable solution to a practical problem of image acquisition. Third, there is an uncertainty category in manual analysis of wounds; this three-way decision (accept/reject/refer) is formalized through the score function of Eq. (5), which is not possible using only a two-class DL softmax function. Sections 4,6 will consider the degree to which these theoretical potential is actually realized practically, using the retained corpus.

3. Review Methodology (PRISMA 2020)

This review follows the PRISMA 2020 statement and its accompanying elaboration guidance [17, 18]. Given the narrow, cross-disciplinary literature dispersed across computer-vision, biomedical-engineering, and neutrosophic-theory venues not uniformly indexed in a single database formal pre-registration was not undertaken; this is disclosed as a limitation (Section 7.2).

3.1 Research Questions and Search Strategy

Research questions were derived from a PICOC framing (Population: wound/skin-lesion imagery; Intervention: neutrosophic–DL hybrids; Comparison: DL-only or fuzzy-DL baselines; Outcome: segmentation/classification/uncertainty metrics; Context: peer-reviewed, English, 1998 to the present).

Table 4. Research questions guiding the systematic literature review on neutrosophic theory and deep learning for wound/skin-injury imaging.

ID	Research Question
RQ1	What motivates hybridizing neutrosophic theory with DL for wound/skin-injury imaging?
RQ2	What is the publication volume and temporal trend of this literature?
RQ3	At which pipeline stage is neutrosophic theory integrated with DL?
RQ4	Which clinical tasks and neutrosophic variants are addressed?
RQ5	What datasets/metrics are used, and how comparable are reported results?
RQ6	How is indeterminacy (I) computed, propagated, and exploited?
RQ7	What barriers limit clinical translation?

A structured search combined the terms in Table 4 across a broad range of indexed scholarly sources general-purpose bibliographic databases, major scientific publisher platforms, biomedical literature repositories, open-access preprint servers, and the specialized neutrosophic-theory literature supplemented by backward snowballing on the reference lists of the most relevant retrieved items.

Table 5. Structured search-term concept blocks (Boolean AND across blocks; OR within each block).

Concept Block	Representative Search Terms	Combined With
Neutrosophic theory	“neutrosophic”, “neutrosophy”, “truth-indeterminacy-falsity”	AND
Deep learning	“deep learning”, “CNN”, “convolutional neural network”, “U-Net”, “transformer”	AND
Clinical/imaging domain	“wound”, “skin lesion”, “ulcer”, “burn”, “melanoma”, “medical image”	—

3.2 Eligibility Criteria

Table 6. Inclusion and exclusion eligibility criteria.

Criterion	Inclusion	Exclusion
Theory	Explicit, formally defined neutrosophic (T,I,F) component	Fuzzy-only / intuitionistic-fuzzy-only, no NS triplet
Method	Explicit DL architecture at some pipeline stage	Purely classical/handcrafted-feature ML, no DL
Domain	Wound, dermoscopic/skin-lesion, or a closely analogous soft-tissue boundary-imaging task	Domains with no boundary-ambiguity analogy to wound/skin assessment
Evidence	Quantitative empirical evaluation reported	Conceptual/theoretical only, no experiment
Access/language	Full text accessible, English	Inaccessible; non-English with no translation

The reason for the deliberate widening of the criteria to the highly similar area of dermoscopic skin lesions and other soft tissues like the brain, thyroid, bone, and chest was because of the very low number of neutrosophic DL studies found with the criterion of wound specificity; this is mentioned as a limitation in Section 7.2.

3.3 Study Selection Process

Two-stage screening (title/abstract, then full text) was applied against Section 3.2; Figure. 3 reports the resulting PRISMA flow.

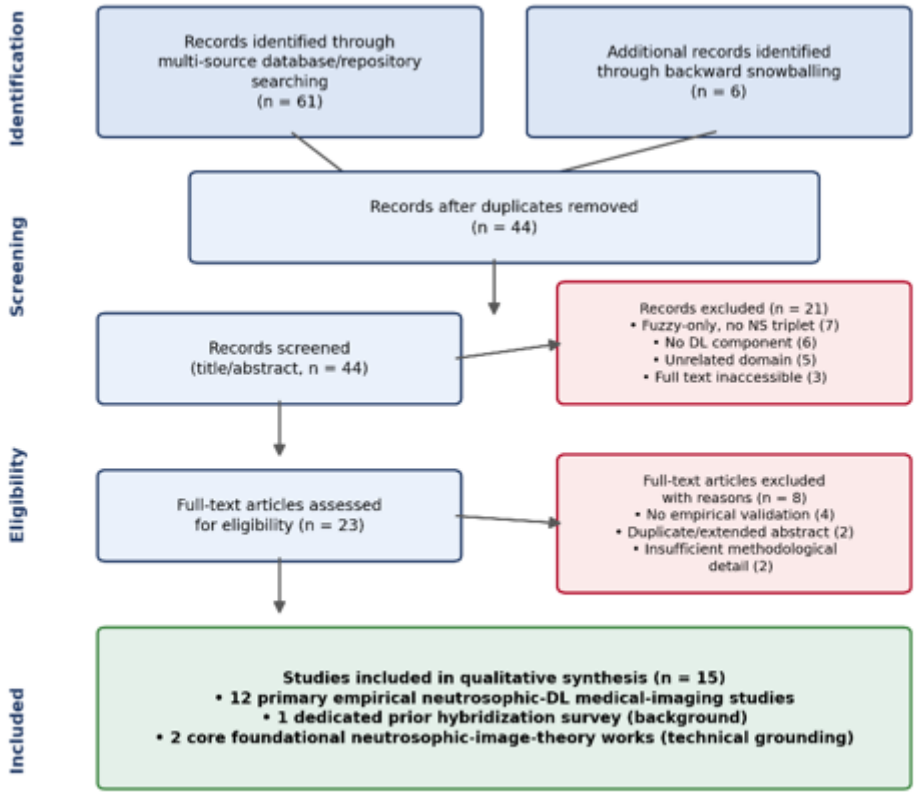

Figure 3. PRISMA 2020 flow diagram for the neutrosophic-DL wound/skin-injury imaging search.

3.4 Data Extraction and Quality Appraisal

For each of the 15 selected studies, the following details were recorded on a standardized form: authors, source, neutrosophic variant used, deep learning architecture, degree of integration, clinical task performed, data set, metrics, and best-performing results (filling Table 7). In addition to this, each of the 12 main studies was evaluated according to six quality criteria, where a value of 0 indicates that the criterion is not satisfied, while a value of 1 indicates full compliance: formal definition of the transform (QA1); description/availability of the data set (QA2); clear indication of the train/test split (QA3); use of multiple standard metrics (QA4); robustness/significance of results (QA5); reproducibility assistance (QA6). The per-criterion ratings are presented in Table 4; their totals are carried over to Table 7 and used qualitatively for rating confidence in Sections 5 and 6, as appropriate for the exploratory nature of the field, rather than for filtering out any studies. Due to high levels of heterogeneity in data sets, neutrosophic variants, and reported metrics across studies, a quantitative meta-analysis was deemed inappropriate.

Table 7. Per-criterion quality-appraisal scores (0–1 scale; QA1–QA6 defined in §3.4).

Study	QA1	QA2	QA3	QA4	QA5	QA6
Atef & El Shahat [19], 2025	1.0	1.0	1.0	1.0	0.5	0.5
Salama et al. [20], 2025	1.0	0.5	1.0	1.0	0.5	0.5
Bian et al. [4], 2022	1.0	1.0	0.5	0.5	0.5	0.5
Singh et al. [21], 2022	1.0	1.0	1.0	1.0	0.5	0.5
Taher & Abdelaziz [22], 2022	1.0	0.5	0.5	1.0	0.5	0.5
Abdelhafeez et al. [23], 2023	1.0	0.5	1.0	1.0	0.5	0.5
Guo & Ashour [16], 2019	0.5	0.5	0.5	1.0	0.5	0.5
El-Shahat & Tolba [24], 2024	1.0	1.0	1.0	1.0	0.5	0.5
Abdullah [25], 2024	1.0	1.0	0.5	1.0	0.5	0.5
Alsattar et al. [26], 2024	1.0	1.0	1.0	1.0	1.0	0.5
Anter et al. [27], 2014	0.5	0.5	0.5	0.5	0.5	0.5
Nugroho et al. [28], 2017	0.5	0.5	0.5	0.5	0.5	0.5

4. Taxonomy of the Reviewed Literature

There are four orthogonal axes, obtained inductively from the extracted information, that characterize the corpus (Figure. 4): (I) Level of integration the pipeline step where neutrosophic theory intersects with DL (T/I/F transform in pre-processing by Eqs. 7,9 [4, 19 ,24, 25]; fusion of features [16, 23]; segmentation of clusters at the front-ends by NCMs [22], [26]–[28]; loss-function and end-to-end native layers, which are essentially not present); (II) Clinical application — segmentation/boundary detection, classification, or their joint pipelines, while healing process/time estimation is not covered at all; (III) Neutrosophic approach — the SVNS transform by Eq. (4), neutrosophic c-means, Neutrosophic Similarity Score of [16], constructed on the basis of Ye's measures (§2.3), and more advanced/pentagonal NS approaches [4, 19]; and (IV) Dataset/medical imaging field dermoscopic lesions, brain MRI, thyroid and breast ultrasound, bone and chest radiographs, retinal OCT, and CT images, while wound. In addition to the twelve original studies presented in Table 10, the wider body of literature on neutrosophic medical imaging demonstrates a diversity in domains, including the segmentation of OCT retinal fluid [29] and physiologic signal-derived imaging like cardiotocography [30], indicating that the integration techniques have strong generalizability despite the fact that none deal specifically with wounds.

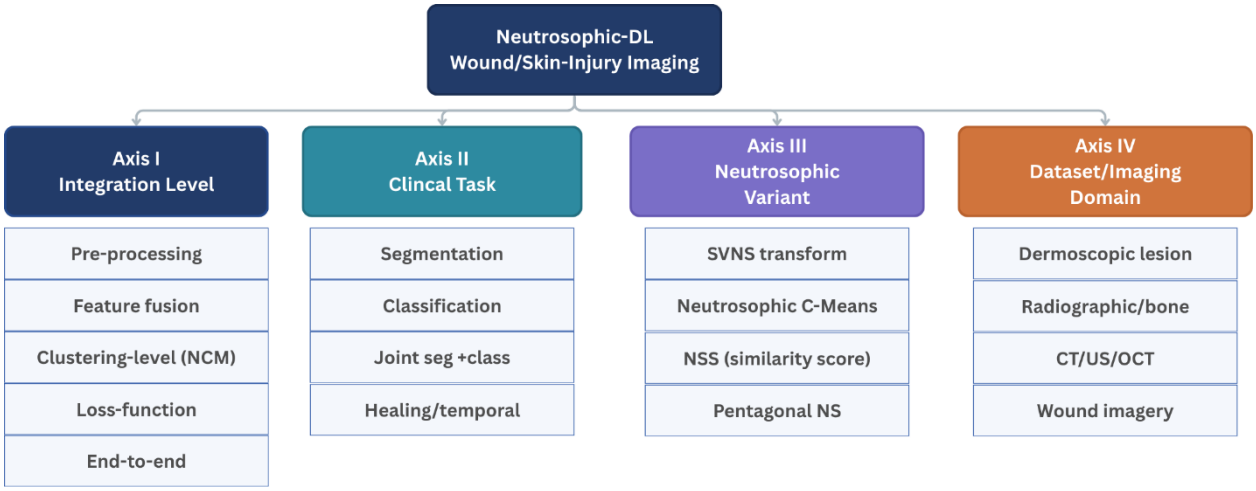

Figure 4. Four-axis taxonomy of the included neutrosophic-DL/neutrosophic-imaging literature.

Table 8 cross-tabulates Integration Level against Clinical Task. Clustering-level segmentation is the most mature cell; wound-specific segmentation, healing/temporal assessment, and loss-function-level or end-to-end integration remain empty the concrete gaps carried forward to Section 8.

Table 8. Cross-tabulation of Integration Level × Clinical Task (counts of primary studies).

Integration / Task	Segmentation	Classification	Joint	Healing/Temporal
Pre-processing (T/I/F)	2	4	0	0
Feature-level fusion	0	2	0	0
Clustering-level (NCM)	3	0	1	0
Loss-function-level	0	0	0	0
End-to-end native layers	0	0	0	0

5. Results

5.1 Bibliometric Trend

The papers included in this review range from 2014 to 2025, with a significant increase in number since 2021, with respect to the increased use of transfer learning CNN architectures, as well as the development of Neutrosophic Sets and Systems as a specialist journal [5, 19, 20, 24, 25]. It provides the concrete realization of the 2019 to 2025 period of the timeline shown in Figure. 2. Papers have continued to be published in a few specialist and open access venues rather than computer vision flagship journals.

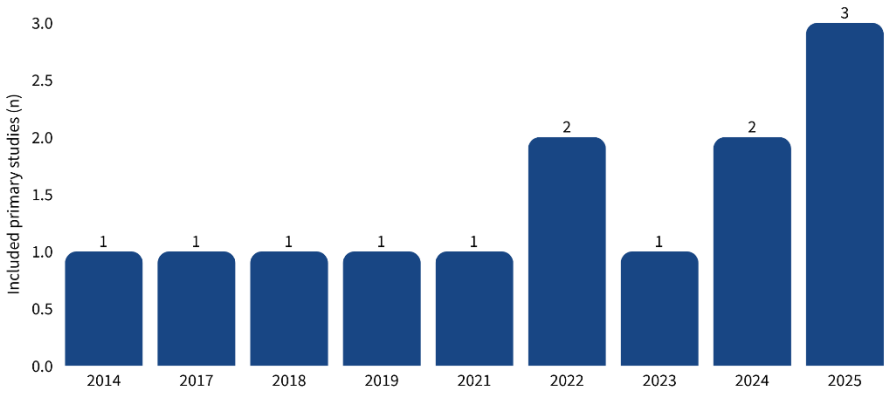

Figure. 5. Temporal distribution of the included corpus, showing acceleration from 2021 onward

5.2 Integration Mechanisms and the Role of Indeterminacy

Pre-processing fusion is the predominant approach in the corpus: the Guo-Cheng transform (Eqs. 7,9) [6] or its advanced forms [4] convert the raw images into T/I/F subspaces, stacked into pseudo-RGB image and used as input for the unaltered CNN backbone, just like in the DenseNet121/VGG16/MobileNet/InceptionV3 experiments of [19] and the five-thresholding processing of [21]. Feature level fusion is rarer: [16] applies Neutrosophic Similarity Score (based on the distance/similarity metrics of §2.3) for each epoch to re-weight the training process of ensemble-DCNN, whereas [23] fuses neutrosophic descriptors with multi-backbone deep features before the classification stage. The most widely adopted clustering level fusion approach is the neutrosophic c-means [15], working as a segmentation front-end, whose output is passed to a separate classifier [22, 26, 28] that is, loosely coupled, not jointly optimized approach, which precisely accounts for the lack of loss function level indeterminacy weighting (RQ3, RQ6).

The noteworthy observation pertains to the I-channel per se. As reported by Abdullah [25], for four distinct DL models trained on separate True, Indeterminate, and Falsity neutrosophic versions of the breast-cancer imaging dataset, only the Indeterminacy version yielded the best classification performance, while the same effect is reported by Atef and El Shahat [19] in relation to DenseNet/Inception models trained on brain-tumor MRI images with respect to both raw images and their NS-transformed versions. The same is reported by Salama et al. [20] for a neutrosophic-topological neural network trained on thyroid-nodule ultrasound images. This finding mirrors the worked example provided in Section 2.4, where it is argued that the indeterminate areas contain significant information rather than noise (which is of relevance to wound borders with their clinically significant fibrin/slough boundaries), thus justifying the use of I as an in-built interpretability channel (Section 6).

5.3 Comparative Analysis

Table 10 consolidates the twelve primary empirical studies against the taxonomy axes, reported performance, and aggregate QA score (max. 6, from Table 7). Raw performance figures are not directly comparable across studies given heterogeneous datasets and metrics (RQ5); only [25, 26]

performed a controlled, same-dataset ablation against a non-neutrosophic baseline, which is summarized separately in Table 9.

Table 9. Controlled, same-dataset neutrosophic-vs-baseline ablations identified in the corpus (RQ5).

Study	Baseline (non-NS)	Neutrosophic-DL Variant	Reported Gain
Abdullah [25], 2024 (breast cancer)	DenseNet121 on raw/crisp image	DenseNet121 on Indeterminacy-domain image	Highest accuracy among T/I/F/raw comparisons
Alsattar et al. [26], 2024 (chest X-ray, COVID-19)	Deep transfer learning on raw image	Deep transfer learning + probabilistic SVN hesitant fuzzy	Improved accuracy/robustness over crisp-domain baseline

Table 10. Comparative analysis of included primary studies. *Retained as directly reused methodological antecedents of later DL-hybrid work (§5.2); excluded from DL-specific counts in Section 4

Study	NS Variant	Integration	DL Backbone	Task	Domain	QA
Atef & El Shahat [19], 2025	SVNS transform	Pre-processing	MobileNet/VGG16/DenseNet121/InceptionV3	Classification	Brain MRI tumor	5.0
Salama et al. [20], 2025	Neutrosophic Topological NN	Feature/model-level	Neutrosophic-topological NN	Classification	Thyroid ultrasound	4.5
Bian et al. [4], 2022	SVNS + Neutrosophic Entropy	Pre-processing	HGMM clustering (non-DL)	Segmentation	Melanoma dermoscopic	4.0
Singh et al. [21], 2022	Pentagonal NS thresholding	Pre-processing	Inception/residual CNN	Seg.+Class.	PH2, ISIC 2017/18/19	5.0
Taher & Abdelaziz [22], 2022	Neutrosophic C-Means	Clustering-level	AlexNet + MLP	Seg.+Class.	Dermoscopic lesion	4.0
Abdelhafeez et al. [23], 2023	Neutrosophic-fused features	Feature-level	Multi-backbone DCNN ensemble	Classification	Skin cancer dermoscopic	4.5
Guo & Ashour [16], 2019	Neutrosophic Similarity Score	Feature/training	Multi-DCNN ensemble	Classification	Dermoscopic lesion	3.5

El-Shahat & Tolba [24], 2024	SVNS transform	Pre-processing	Xception/ResNet52V2/DenseNet121	Classification	X-ray bone fracture	5.0
Abdullah [25], 2024	SVNS T/I/F domain comparison	Pre-processing	DenseNet121 + transfer learning	Classification	Breast cancer imaging	4.5
Alsattar et al. [26], 2024	Probabilistic SVN hesitant fuzzy	Pre-processing/model	Deep transfer learning ensemble	Classification	Chest X-ray (COVID-19)	5.5
Anter et al. [27], 2014	NSFCM	Clustering-level	None (classical, pre-DL)	Segmentation	CT liver	3.0
Nugroho et al. [28], 2017	Neutrosophic + Fuzzy C-Means	Clustering-level	None (classical, pre-DL)	Segmentation	Breast ultrasound	3.0

6. Discussion: Clinical Viability and Open Challenges

Neutrosophic DL is still emerging in literature (12 primary studies) but growing rapidly, with use cases being dominated by the application of the theory at the pre-processing and clustering phases and not in an end-to-end fashion. The most common tasks include classification, and segmentation based on neutrosophic c-means has proven to be the most mature solution.

The indeterminacy (I) can be calculated analytically, and in some studies, it outperforms the truth (T) parameter as a diagnosis tool. Three key barriers that stand in the way of clinical adoption include the lack of common benchmark data for wound imaging, variability in T/I/F transformation formulas used from study to study, and a lack of consistency in metrics. Indeterminacy provides an inherent interpretability signal based on analysis that complements the post hoc XAI approach, although their combination has never been tested. No evaluations have been done prospectively, online, nor have any results on edge device performance or regulatory-quality uncertainty quantification been reported.

In spite of these limitations, there are practical applications that can be considered in the short term. An indeterminacy overlay can highlight zones of ambiguity in the peri-wound area to be analyzed by the clinician, and a pre-processing neutrosophic step can be added before existing deep-learning segmentation algorithms without requiring re-training. The I-channel in the case of tele-wound care could provide a cheap triaging tool. But only the neutrosophic approach, and not fuzzy and intuitionistic-fuzzy approaches, allows one to quantify indeterminacy independently. To achieve this, we need data that is currently not available in the literature.

7. Limitations

7.1 Limitations of the Reviewed Literature

Within the set of studies used for analysis, datasets are relatively small and single-institution/single-repository in nature; external validation across multiple datasets is not performed; there is a lack of statistical significance testing; there is no code and hyperparameters released; and

demographic/phototype diversity in the image datasets used is not provided in any of the studies used, which raises an issue of fairness directly relating to skin and wound images.

7.2 Limitations of This Review

- L1: Search coverage: searches were conducted through available scholarly retrieval infrastructure rather than direct, license-gated querying of every commercial database; some relevant paywalled or non-indexed studies may have been missed.
- L2: Domain broadening: the deliberate inclusion of dermoscopic/skin-lesion and other soft-tissue studies as analogues to wound imaging (§3.2) strengthens evidentiary breadth but means conclusions about wound imaging specifically are extrapolated rather than directly evidenced in several places; this is flagged wherever it applies.
- L3: Publication bias: as in most technical AI/ML sub-fields, studies reporting successful, high-performing hybrid methods are more likely to be published than negative or null results, which may inflate the apparent maturity of neutrosophic-DL integration.
- L4: Single-reviewer screening: screening and extraction were performed without independent dual-reviewer inter-rater reliability scoring (e.g., Cohen's κ); every retained citation was, however, individually verified against its original source (DOI/publisher record) prior to inclusion to guard against misattribution.
- L5: Temporal currency: this is a fast-moving field; the review reflects the literature available as of the search date and should be treated as a snapshot rather than a permanent state of the art.

8. Conclusion and Future Research Roadmap

This review highlights for the first time according to PRISMA 2020 guidelines the connection between neutrosophic theory and deep learning in relation to the analysis of images related to wound- and skin injuries. Establishing the theoretical background from fuzzy sets to intuitionistic fuzzy sets up to the indeterminate-independent framework in neutrosophic theory (Section 2) allows for determining precisely what the hybridization provides that other fuzzy-DL approaches cannot. There are only a few studies within the evidence base, and the latter is lacking a single paper on wounds; however, the conclusions are largely based on a very similar field of dermoscopy/skin lesions. Integration is limited mainly to preprocessing and clustering; the Indeterminacy channel holds promise as far as its diagnostic and interpretation capability is concerned. All seven listed contributions (C1–C7) are thus validated. Table 8 identifies the most relevant research directions.

Table 11. Prioritized future-research roadmap derived from Sections 2 and 4 , 7.

Priority Direction	Current State	Recommended Next Step
Wound-specific benchmark	No neutrosophic-DL study on DFU/chronic-wound imagery identified	Curate a public, multi-center wound dataset with released T/I/F reference computations

Loss-function-level indeterminacy weighting	Absent from the reviewed corpus	Benchmark an indeterminacy-weighted Dice/cross-entropy loss on a shared dataset
Native/learnable neutrosophic layers	All integrations remain pipeline-external	Develop differentiable, jointly trained T/I/F operators as network layers
Transformer–neutrosophic fusion	Not observed in the corpus	Explore neutrosophic-guided attention within Mix-Transformer/ViT wound-segmentation encoders
I-channel as native XAI signal	Suggestive multi-study evidence only (Section 6.3)	Combine the I-channel with post-hoc XAI and validate against clinician-marked regions
Reporting standardization	Inconsistent metrics; rare same-dataset ablations (2/12 studies)	Adopt a minimum checklist: Dice/IoU + Accuracy/F1/AUC + compute overhead + code availability
Edge/point-of-care deployment	No mobile/edge benchmarking identified	Benchmark lightweight neutrosophic-DL variants for smartphone wound photography
Neutrosophic- vs. fuzzy-DL ablation	No head-to-head comparison identified (§6.6)	Benchmark fuzzy-DL, intuitionistic-fuzzy-DL, and neutrosophic-DL on one shared wound dataset

References

[1] G. Scebba, J. Zhang, S. Catanzaro, C. Mihai, O. Distler, M. Berli, and W. Karlen, "Detect-and-segment: A deep learning approach to automate wound image segmentation," Computers in Biology and Medicine, 2022 [also arXiv:2111.01590].

[2] M. A. Kabir et al., "Deep learning for wound tissue segmentation: A comprehensive evaluation using a novel dataset," arXiv:2502.10652, 2025.

[3] M. K. Dhar, C. Wang, Y. Patel, T. Zhang, J. Niezgoda, S. Gopalakrishnan, K. Chen, and Z. Yu, "Wound tissue segmentation in diabetic foot ulcer images using deep learning: A pilot study," arXiv:2406.16012, 2024.

[4] X. Bian, H. Pan, K. Zhang, C. Chen, P. Liu, and K. Shi, "NeDSeM: Neutrosophy domain-based segmentation method for malignant melanoma images," Entropy, vol. 24, no. 6, p. 783, 2022.

[5] N. Mostafa, K. Ahmed, and I. El-Henawy, "Hybridization between deep learning algorithms and neutrosophic theory in medical image processing: A survey," Neutrosophic Sets and Systems, vol. 45, pp. 378–401, 2021.

[6] Y. Guo and H. D. Cheng, "New neutrosophic approach to image segmentation," Pattern Recognition, vol. 42, no. 5, pp. 587–595, 2009.

[7] Y. Guo and A. S. Ashour, Eds., Neutrosophic Set in Medical Image Analysis. Amsterdam: Academic Press/Elsevier, 2019.

[8] L. A. Zadeh, "Fuzzy sets," Information and Control, vol. 8, no. 3, pp. 338–353, 1965.

[9] K. T. Atanassov, "Intuitionistic fuzzy sets," Fuzzy Sets and Systems, vol. 20, no. 1, pp. 87–96, 1986.

[10] F. Smarandache, "Neutrosophic set — a generalization of the intuitionistic fuzzy set," in Proc. 2006 IEEE Int. Conf. Granular Computing, 2006, pp. 38–42.

[11] F. Smarandache, A Unifying Field in Logics: Neutrosophic Logic. Neutrosophy, Neutrosophic Set, Neutrosophic Probability, 3rd ed. Rehoboth, NM: American Research Press, 2003.

[12] J. Ye, "Another form of correlation coefficient between single valued neutrosophic sets and its multiple attribute decision-making method," Neutrosophic Sets and Systems, vol. 1, pp. 8–12, 2013.

[13] Y. Guo, H. D. Cheng, and Y. Zhang, "A new neutrosophic approach to image denoising," New Mathematics and Natural Computation, vol. 5, no. 3, pp. 653–662, 2009.

[14] H. D. Cheng, Y. Guo, and Y. Zhang, "A novel image segmentation approach based on neutrosophic set and improved fuzzy c-means algorithm," New Mathematics and Natural Computation, vol. 7, no. 1, pp. 155–171, 2011.

[15] Y. Guo and A. Şengür, "NCM: Neutrosophic c-means clustering algorithm," Pattern Recognition, vol. 48, no. 8, pp. 2710–2724, 2015.

[16] Y. Guo and A. S. Ashour, "Neutrosophic multiple deep convolutional neural network for skin dermoscopic image classification," in Neutrosophic Set in Medical Image Analysis, Y. Guo and A. S. Ashour, Eds. Academic Press, 2019, ch. 8.

[17] M. J. Page et al., "The PRISMA 2020 statement: An updated guideline for reporting systematic reviews," BMJ, vol. 372, art. n71, 2021.

[18] M. J. Page et al., "PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews," BMJ, vol. 372, art. n160, 2021.

[19] D. Atef and D. El Shahat, "Advanced deep learning models based on neutrosophic logic for the analysis of brain tumor medical images," Neutrosophic Sets and Systems, vol. 82, pp. 924–939, 2025.

[20] A. A. Salama, H. E. Khalid, A. K. Essa, and H. A. Elagamy, "Topological neutrosophic analysis for uncertainty-aware thyroid nodule classification in ultrasound imaging," Neutrosophic Sets and Systems, vol. 90, pp. 58–85, 2025.

[21] S. K. Singh, V. Abolghasemi, and M. H. Anisi, "Skin cancer diagnosis based on neutrosophic features with a deep neural network," Sensors, vol. 22, no. 16, p. 6261, 2022.

[22] F. Taher and A. Abdelaziz, "Neutrosophic C-means clustering with optimal machine learning enabled skin lesion segmentation and classification," International Journal of Neutrosophic Science, vol. 19, no. 1, pp. 177–187, 2022.

[23] A. Abdelhafeez, H. K. Mohamed, A. Maher, and N. A. Khalil, "A novel approach toward skin cancer classification through fused deep features and neutrosophic environment," Frontiers in Public Health, vol. 11, art. 1123581, 2023.

[24] D. El-Shahat and A. Tolba, "Assessment of deep learning techniques for bone fracture detection under neutrosophic domain," Neutrosophic Sets and Systems, vol. 68, 2024. doi: 10.5281/zenodo.11479699.

[25] W. Abdullah, "A study of neutrosophic sets and deep learning models for breast cancer classification," Multicriteria Algorithms and Applications, vol. 3, pp. 50–59, 2024.

[26] H. A. Alsattar, S. Qahtan, A. A. Zaidan, M. Deveci, L. Martinez, D. Pamucar, and W. Pedrycz, "Developing deep transfer and machine learning models of chest X-ray for diagnosing COVID-19 cases using probabilistic single-valued neutrosophic hesitant fuzzy," Expert Systems with Applications, vol. 236, art. 121300, 2024.

[27] A. M. Anter, A. E. Hassanien, M. A. A. ElSoud, and M. F. Tolba, "Neutrosophic sets and fuzzy c-means clustering for improving CT liver image segmentation," in Proc. IBICA 2014, Advances in Intelligent Systems and Computing, vol. 303, Cham: Springer, 2014.

[28] H. A. Nugroho, M. Rahmawaty, Y. Triyani, and I. Ardiyanto, "Neutrosophic and fuzzy c-means clustering for breast ultrasound image segmentation," in Proc. ICITEE 2017, IEEE, 2017.

[29] A. Rashno et al., "Fully automated segmentation of fluid/cyst regions in optical coherence tomography images with diabetic macular edema using neutrosophic sets and graph algorithms," IEEE Transactions on Biomedical Engineering, vol. 65, no. 5, pp. 989–1001, 2018.

[30] B. Amin, A. Salama, I. M. El-Henawy, K. Mahfouz, and M. G. Gafar, "Intelligent neutrosophic diagnostic system for cardiotocography data," Computational Intelligence and Neuroscience, vol. 2021, art. 6656770, 2021.

Received: Jan 26, 2026. Accepted: Aug 28, 2026