Quest Journals
Journal of Electronics and Communication Engineering Research
Volume 12 ~ Issue 3 (May - June 2026) pp: 08-20
ISSN(Online): 2321-5941
www.questjournals.org

Research Paper

A Comprehensive Review of Explainable Fuzzy Inference Systems for Disease Diagnosis Based on Interpretable Rule-Based Artificial Intelligence

Dr Pankaj Mohindru
Department of Electronics & Communication Engineering, Punjabi University
(A State Government University), Patiala 147002, Punjab, India

Abstract
Artificial Intelligence (AI) has revolutionized healthcare, providing support in disease diagnosis, clinical decision-making, risk prediction, and personalized medicine. But there are many machine learning and deep learning models that are "black boxes," which lack transparency and therefore clinical trust. Explainable Artificial Intelligence (XAI): This overcomes this shortcoming by offering a more specific explanation of the reasoning behind AI forecasts. This review study gives a comprehensive overview of Explainable Fuzzy Inference Systems (EFISs) for disease diagnosis using interpretable rule-based artificial intelligence. The healthcare application is an area where fuzzy logic can handle the uncertainty, vagueness, and gradual clinical transitions well, using linguistic variables, membership functions, and IF–THEN rules. Discussed fuzzy inference architectures, rule-based knowledge representation, expert knowledge integration, and applications in cardiovascular disease, diabetes, cancer, neurological disorders, liver and kidney disease, respiratory disease, and infectious disease. Also, hybrid systems such as neuro-fuzzy systems, ANFIS, fuzzy–machine learning, fuzzy–deep learning, and novel fuzzy–LLM systems are analyzed. Evaluation of diagnostic performance and explainability is based on accuracy, sensitivity and specificity, F1-score, ROC-AUC, rule complexity, fidelity, stability, and clinical expert assessment. The problems are rule explosion, high dimensionality of data, lack of standardized explainability metrics, bias of the datasets, privacy, and lack of clinical validation. Multimodal, personalized, robust, and human-centered fuzzy AI systems for trustworthy clinical diagnosis are to be emphasized for future research.

Keywords: *Artificial Intelligence (AI); Explainable Artificial Intelligence (XAI); Explainable Fuzzy Inference Systems (EFISs); Fuzzy Logic*

I. Introduction

In the realm of healthcare, the application of Artificial Intelligence (AI) is transforming the way the industry operates, with several AI applications gaining relevance, including disease diagnosis, clinical decision support, medical image analysis, risk prediction, and personalized medicine (Fahim et al., 2025). In recent years, the vast amount of Electronic Health Records (EHRs), medical images, laboratory results, physiological data, genetic information, and patient-generated health data has opened the door to using computational techniques to support healthcare professionals (Maleki Varnosfaderani & Forouzanfar, 2024). In the field of healthcare, machine learning (ML) and deep learning (DL) tools have demonstrated their capability of providing valuable insights for diagnosis in large-scale medical data and their ability to identify intricate patterns within these data (Ogut, 2025). But many AI models make predictions without giving an explicit explanation of the logic behind the prediction, even when they are successful. However, when the predictive accuracy of the AI is applicable to healthcare, e.g., diagnosis, which could affect the treatment and the patient outcomes, there is a need for validation and understanding of the process that the AI is using to make the prediction (Pinto-Coelho, 2023). This has given rise to research in explainable and interpretable AI, as it has become an important research field in medical AI (Sadr et al., 2025).

Explainable Artificial Intelligence (XAI) refers to the techniques and models used to make AI (Ahsan et al., 2022) understandable. Explainability allows doctors in the clinical setting to understand on what basis a specific diagnosis or prediction of risk has been made, and whether this logic is clinically justified or not. Explanations may include symptoms that impact, markers, imaging characteristics, and patient risk factors, and help the clinician understand potential for error, bias, or inappropriate behavior of the model (Dwivedi et al., 2023). The greater the interpretability, the better it can be used to support human – AI collaboration, build trust in computational recommendations, foster accountability, and facilitate clinical validation (Kalasampath et al., 2025). However, finding a good balance between predictive performance and meaningful interpretability, especially for complex ML and DL architectures applied to high-dimensional and heterogeneous medical data (Chamola et al., 2023), is still a challenge.

The use of fuzzy logic provides a good basis for the design of explainable AI-driven diagnostic systems. In contrast, fuzzy logic allows for membership in the truth value range of 0 to 1, but classical binary logic has only two possible values, true or false (Kostopoulos et al., 2024). The situation is particularly critical in the medical field, where data may be uncertain, ambiguous, and may involve gradual transitions. The symptoms can be either mild, moderate, or severe; the laboratory measurements can be either at or between uncertain boundaries (Shukla et al., 2025). These can be linguistically represented with "low," "normal," "moderate," and "high," which is closer to clinical reasoning that can be done by fuzzy logic. Fuzzy Inference Systems (FISs) are systems that are founded on fuzzy sets, membership functions, and IF–THEN rules to translate the clinical observations into a diagnosis and/or risk assessment. A rule might be: If blood glucose level is elevated and body mass index is elevated, then the risk for diabetes is high. In addition, there is a benefit in such explicit rules can be evaluated and accredited by clinicians, unlike opaque computational models (Muhammad et al., 2021).

Fuzzy inference combined with interpretable rule-based artificial intelligence further expands the potential of fuzzy AI to interpret disease diagnosis. Fuzzy rules can be used to represent relationships between symptoms, biomarkers, patient characteristics, and disease outcomes, and they can be used to deal with uncertainty as well as with traditional thresholds (Ennab & Mcheick, 2025). There are several models to represent and process uncertain clinical information, such as Mamdani, Sugeno, and Takagi–Sugeno–Kang models, and type-1 and type-2 fuzzy systems (Mjahad & Rosado-Muñoz, 2025). The same methods have been studied in cardiovascular diseases, diabetes, cancer, neurological diseases, liver and kidney diseases, respiratory diseases, and infectious diseases. They can integrate expert medical knowledge, which is particularly beneficial in situations where datasets are restricted, noisy, unbalanced, or hard to understand (Srivastava et al., 2025).

Despite this, there are important challenges that still need to be addressed. As more clinical variables are added, more and more rules will be added to the rulebase, leading to the formation of what is known as rule explosion. Redundancy and linguistically complex structures are also possible in automatically generated rules (Leelavathy et al., 2025). Moreover, as the model gets more complicated, interpretability is reduced, but accuracy gets better (Navin & Krishnan, 2024). Comparisons between different existing studies are also difficult because of variation in datasets, the types of diseases under study, representations of features, disease evaluation, and definitions of explainability. Several proposed systems have also been tested by analyzing databases of past data, instead of actual clinical settings (George Kerry et al., 2021).

With this perspective, this review critically analyzes an explainable fuzzy inference system for disease diagnosis with interpretable rule-based AI (Singh et al., 2019). It analyzes its theoretical foundation and rule representation, its incorporation of expert knowledge, application across different disease realms, hybrid AI solutions, evaluation measures, challenges, and future research avenues (Al-Ali & Qidwai, 2025). Special focus is placed on the connection between diagnostic accuracy, uncertainty management, interpretability, rule complexity, clinical usability, and trustworthiness (Ye et al., 2025). This review aims to shed light on the potential of explainable fuzzy AI in making reliable, transparent, human-centered, and clinically meaningful disease diagnosis (Ambalavanan et al., 2025).

II. Literature Review

Artificial Intelligence (AI) has become a key technology in the field of healthcare, especially for disease diagnosis, clinical decision support, medical image analysis, and patient risk assessment. But with the rise of complex machine learning and Deep Learning models, there have been concerns about these models' interpretability and transparency. However, in the healthcare domain, while predictive power is crucial, so is the ability to explain why a specific diagnosis is being made by AI. Cao et al. (2024) emphasized the importance of interpretable AI models in fostering trustworthy AI in health care and the ability of fuzzy rules to relate to human-comprehensible reasoning as one of the best means to make complex computational models understandable by humans. Their review also found very few studies on fuzzy-rule-based medical diagnosis, especially for multimodal medical data like sequential signals, medical images, and tabular clinical data.

Research into AI clinical decision support systems (CDSSs) has also shown the value of interpretability in healthcare. The interpretability in healthcare has also been demonstrated with research on clinical decision support systems (CDSSs) based on AI. In their literature review, Xu et al. (2023) identified the following concepts that are associated with interpretability in AI-based CDSSs: Interpretability of the model structure, Interpretability of the relationship between the input variables and output variables, and interpretability of the explanation of the algorithmic decisions. Among the important models they found for developing interpretable models, they mentioned fuzzy logic, decision rules, logistic regression, and decision trees. For the explanation of black-box models, other post-hoc methods were identified, including feature importance, sensitivity analysis, visualization, or maximization of activation. The study also showed that the kinds of explanations required are dependent upon the nature of the data, the biomarkers, the clinician's needs, the patient's needs, and the interaction between the human and the AI. The results show that interpretability should be part of the design of the AI system itself, and not just an afterthought.

Many aspects of medical diagnosis call for fuzzy logic, since in medical care information is often ambiguous, imprecise, incomplete, and subjective. Shukla et al. (2025) surveyed fuzzy-set-based methods for medical diagnosis and showed their applications in building intelligent diagnostic and expert systems. Fuzzy logic is different from traditional binary logic in that clinical variables can be expressed in terms of linguistic variables (low, moderate, high), etc., and therefore fuzzy logic allows for partial transitions between disease states. This property makes fuzzy inference systems suitable for some clinical applications where there are not always well-defined boundaries between symptoms and laboratory measurements.

The addition of fuzzy logic to neural networks has enhanced the capabilities of interpretable AI even more. Singh (2025) discussed recent neuro-fuzzy architectures and reported that these architectures are a fusion of the adaptive learning property of neural networks and the explainable reasoning property of fuzzy logic. Recent architectures have shown potential applications in image recognition, streaming data analysis, and biomedical diagnostics. The study did uncover some key problems, however, such as the compromise between predictive power and explainability, scalability, and the absence of any common set of metrics for assessing explanations. Neuro-fuzzy systems are therefore a potential approach to build AI models with learnable complex relationships that retain understandable reasoning structures.

New studies in disease-specific areas also confirm the effectiveness of fuzzy and neuro-fuzzy approaches. Sharma et al. (2025) have analysed the Multilayer Fuzzy Inference System (MLFIS) and Adaptive Neuro-Fuzzy Inference System (ANFIS) for the early identification of chronic kidney disease. Their MLFIS performed diagnosis based on symptoms and risk factors to classify probable CKD and biochemical parameters to classify the stage of the disease, and ANFIS automatically learned the diagnosis rules from data. ANFIS achieved an accuracy of 99.5%, whereas MLFIS had an accuracy of 99.33% using 500 anonymized patient records. The authors did not consider whether the dataset was large enough or readily available for broader clinical validation, even though the reported performance was high.

Likewise, Navin and Krishnan (2024) proposed a fuzzy rule-based evidence-based clinical decision support system based on EHRs. The system implemented fuzzy rules and a knowledge base to obtain the clinical features that are relevant and assist in the diagnostic suggestions. The results reported showed acceptable predictive and discriminatory effectiveness, and they also revealed a degree of agreement with expert assessments. To validate the interpretability of fuzzy logic with post-hoc XAI techniques, Khanna et al. (2025) later combined them to detect oral cancer and make treatment recommendations, which also highlights the benefits of incorporating intrinsic fuzzy-rule interpretability along with post-hoc XAI techniques.

The literature surveyed suggests that the EFS could be a potential alternative to all black-box AI models for diagnosis of disease. Fuzzy rules can be used to capture the hazy clinical knowledge in comprehensible terms, and the use of neuro-fuzzy and fuzzy-deep-learning techniques can enhance adaptability and predictability. However, issues with the complexity of the rules, scalability, integrating multimodal data, standardizing interpretability metrics, and external clinical validation persist. In the future, more compact and clinically meaningful fuzzy rule bases with good diagnostic performance, uncertainty handling, transparency, and clinically useful explanations should be developed.

2. Fundamentals of Fuzzy Inference Systems and Explainable AI

Uncertain knowledge, imprecise knowledge, and knowledge expressed in natural language have been found to be important problems to handle in the field of intelligent decision-making systems, which have been addressed using Fuzzy Inference Systems (FISs) (Maria et al., 2021). Fuzzy inference is based on the concept of fuzzy logic, which is an extension of binary logic, where the variables in the logic can take on values that are not limited to being only true or false. This property is especially important in the healthcare sector, as medical data often lacks precision, ambiguity, and transitions (Singh, 2025). Clinical terms such as 'slightly elevated blood pressure' and 'moderate pain' do not always need to be solely interpreted as 'normal' or 'abnormal' with absolute thresholds, as with the term 'high glucose level' (Cao et al., 2024). Fuzzy logic enables these observations to be incorporated into computational models in a more clinically-sounding manner (Kowalski et al., 2025).

A key idea of fuzzy logic is the fuzzy set. Each element in a fuzzy set receives a membership function value between 0 and 1. Membership functions are used to establish the degree of membership of a numerical observation into one of the linguistic categories, e.g., low, medium, or high. The membership functions that are used most frequently include triangular, trapezoidal, Gaussian, and bell-shaped functions (D'Aniello, 2023). Membership functions can convert quantitative measurements like blood pressure, glucose, cholesterol, BMI, or lab markers into linguistic representations that assist in the diagnosis of diseases. For example, in disease diagnosis, membership functions can be applied to convert quantitative measurements such as blood pressure, glucose, cholesterol, body mass index, or lab markers into linguistic representations. Linguistic variables are therefore of great value as a link between medical numerical data and reasoning that is understandable by human beings. For instance, instead of a number, a fuzzy system can characterize glucose as “normal,” “moderately high,” or “very high,” and thus enable diagnostic reasoning in terms of clinically meaningful statements (Yerkin et al., 2025).

The main way in which knowledge is represented in a fuzzy inference system is by means of fuzzy rules (Shukla et al., 2025). These rules are typically written in an interpretable format, that is, an IF–THEN format, for example, “IF glucose is high AND body mass index is high THEN diabetes risk is high.” Rules can be created based on expert knowledge, historical data, machine-learning algorithms, or all of the above (Bui, 2025). Their explicit structure enables users to explore interdependencies between the input variables and diagnostic outcomes, which is why they are especially well suited for interpretable artificial intelligence (AI) (Fernández et al., 2024). In the typical FIS, the inference process is a sequence of operations such as: fuzzification of crisp input(s), processing the fuzzy rules and aggregating the results, and the defuzzification process if a crisp decision is needed. Multiple ambiguous clinical observations can be combined in a single diagnostic recommendation (Pabarja et al., 2024).

The Mamdani Fuzzy Inference Systems are one of the most popular fuzzy inference architectures that are linguistically interpretable. Mamdani models have fuzzy sets in both the rule antecedent and consequent, and are more suitable for applications in which the rules must be expressed in human-readable form (Kumar et al., 2025). Sugeno and Sugeno Takagi–Kang (STK) models, on the other hand, normally have mathematical functions or constants in their consequent parts (Koohsari et al., 2026). While the consequents of a TSK system are not as directly interpretable as the consequents of a linguistic Mamdani rule, a TSK system can be useful for computational efficiency and for optimization and hybrid learning methods. The choice, therefore, is between these architectures and will depend on the balance between the desired computational performance, the learning ability, and the interpretability (Hostetter, 2025).

Another significant difference lies with regard to the Type - 1 and Type - 2 fuzzy systems. In Type-1 fuzzy systems, regular membership functions are employed, and each input is assigned just one membership degree. The use of type-2 fuzzy systems further increases the degree of uncertainty by using the Membership degree as a fuzzy set (Kambalimath & Deka, 2020). This is useful in medical applications when patients' measurements are uncertain, but also when there is uncertainty in the definition of linguistic categories and uncertainty in expert knowledge. Type-2 fuzzy methods can provide increased flexibility in coping with incomplete, noisy, and subjective clinical data, but the computational complexity might limit applicability in real-time use and interpretation (Mogotsi et al., 2025).

The concept of Explainable Artificial Intelligence (XAI) (Chamola et al., 2023) is closely related to fuzzy inference principles. XAI includes approaches and models for explaining AI predictions to human users. Interpretability means the capacity to comprehend how a model works or makes decisions, and transparency is the degree of visibility and clarity of the model's mechanisms (Ortigossa et al., 2024). Explainability is more general and takes into account the need to provide understandable explanations for a specific prediction or decision. They are critical in healthcare applications, where the accuracy of the AI's recommendations is paramount, as doctors and nurses need to determine whether the recommendation is supported by relevant medical evidence (Kaya et al., 2026).

Explainable AI has an intrinsic relationship with fuzzy systems, as their linguistic variables, membership functions, and IF–THEN rules can be used to explicitly represent the decision logic (Sofianidis et al., 2021). An explainable fuzzy system can not only provide a diagnosis of the disease but also give the user an indication of which clinical parameters led to the diagnosis and the degree of influence on the inference process (Hussain et al., 2021). However, the number of rules, the complexity of membership functions, and the redundancy of conditions, along with automatically generated knowledge, can degrade the comprehensibility of the process (Islam et al., 2021). Therefore, the design of interpretable fuzzy diagnostic systems needs to be carefully addressed to ensure that the rule base is simple, consistent, and clinically meaningful while maintaining good prediction capabilities. These together provide a solid ground to create reliable and trustworthy AI systems that are designed for human health diagnosis (Hassija et al., 2024).

3. Explainable and Interpretable Rule-Based Artificial Intelligence

Explainable and interpretable rule-based Artificial Intelligence (AI) provides an interpretable approach to clinical decision-making support system which incorporates the knowledge of diagnosis in the form of explicit rules (Xu et al., 2023). Some rule-based systems can provide reasoning for the prediction, such as a Black-Box model, that can be read and checked by health care professionals (Domingues, 2025). A typical clinical rule is written in the form of an IF–THEN statement, with antecedents being patient characteristics/symptoms/laboratory data/imaging data, and the consequent being the diagnosis/risk category/clinical recommendation. Such rules enable systematic capture of medical knowledge and facilitate an easy understanding of the decisions (Navin & Krishnan, 2024).

This is key to rule-based AI and could use data from medical literature, medical guidelines, patient history, or medical domain experts. Physicians' experience can be used to guide the creation of computational diagnostic systems, particularly when the information available is limited or conflicting (Kim et al., 2024). Rules may be either hand-built and written by experts or learned from machine learning, evolutionary algorithms, decision trees, or fuzzy learning techniques. But auto-generated rule bases can have too many rules, rules that are repetitive, and rules that are not easily interpretable (Cao et al., 2024). In this case, the number of rules can be simplified and pruned to eliminate redundant conditions, reduce the number of rules, and increase their readability without significantly reducing the diagnostic capability (Turgunbaev, 2025).

Linguistic interpretability is also used in fuzzy diagnostic systems to make the rules easier to read, for clinical variables (e.g., low, moderate, high). Explanations may be global or local. Explanations at the global level describe the overall decision structure of the model, and explanations at the local level describe the reason for a particular diagnostic decision for a particular patient (Abbas et al., 2025). The clinical rules can then be interpreted by humans to increase clinician confidence and aid in the validation of AI-generated recommendations (Kostopoulos et al., 2024).

Rule bases should be consistent, complete, and minimally redundant for dependable diagnostic reasoning. Fuzzy rules can produce conflicting results, and the lack of rule coverage can also result in insufficient rules to cover some patients' conditions (Alnattah et al., 2025). A key difficulty is thus the need to balance accuracy and interpretability. The more complex and/or numerous the rules, the more accurate the predictions, but the less understandable the model. Thus, the successful explainable rule-based AI seeks to identify compact, clinically meaningful, clinically valid rule structures which enable reliable human–AI interaction when tackling the disease diagnosis task (Osman & Owolabi, 2023).

4. Explainable Fuzzy Inference Systems for Disease Diagnosis

Explainable Fuzzy Inference Systems (EFISs) are a formal approach to the problem of converting vague and inexplicable medical data into understandable medical diagnoses. The general procedure of a typical system is: Data preprocessing, Feature selection, Fuzzy representation, Rule construction, Inference, and Generation of explainable output (Cao et al., 2024). Medical data is pre-processed by cleaning, normalizing, and transforming it to address missing values, noise, and inconsistent measurements (Ennab, 2025). Clinical features can then be selected according to statistical approaches, expert knowledge, or machine learning (Nemade, 2025). Fuzzy linguistic variables are used to represent the symptoms, biomarkers, and physiological measurements, and this allows for the modelling of the gradual nature of clinical conditions, instead of depending on hard thresholds.

The diagnostic knowledge is then converted into IF–THEN rules that can be understood by humans. For instance, the rule Glucose is high and BMI is high is considered High diabetes risk (Alsadie, 2024). In inference, the observations are patient-specific and are fuzzified, and then relevant rules are selected and fired, and the output of the rules is aggregated to give a diagnostic classification or risk score (Awan, 2023). Useful for dealing with uncertainty due to incomplete measurements, ambiguous symptoms, overlapping symptoms of the diseases, and subjective clinical evaluation. The level of confidence or risk estimation can also convey the impact of a recommended diagnosis (Sirocchi et al., 2024).

A big benefit of the EFISs is that they can integrate clinical knowledge from experts with patient-specific data. The rule base can also be validated, modified, and extended by physicians, and data-driven learning can discover patterns from clinical databases that are not known (Kopitar et al., 2024). The final diagnostic output may, therefore, include not only the predicted disease, but also important clinical variables and activated rules that are used to predict the disease (Naseef et al., 2026). These supporting explanations allow the human-AI interaction, in which the clinician can examine and query the algorithmic suggestions and confirm them prior to making clinical choices (Xu et al., 2026).

Table 1. Key Components of Explainable Fuzzy Inference Systems for Disease Diagnosis

Component	Function in Diagnostic System	Explainability Contribution
Data preprocessing	Cleans and normalizes medical data	Improves reliability of inputs
Feature selection	Identifies relevant clinical variables	Highlights important diagnostic factors
Fuzzy representation	Converts numerical data into linguistic terms	Makes clinical conditions understandable
Rule construction	Represents diagnostic knowledge using IF–THEN rules	Provides transparent reasoning
Fuzzy inference	Combines activated rules to generate diagnosis	Shows how evidence contributes to decisions
Uncertainty handling	Manages imprecise and incomplete information	Reflects real-world clinical uncertainty
Risk estimation	Provides diagnostic confidence or risk level	Communicates decision strength
Expert integration	Incorporates physician knowledge	Supports clinically meaningful reasoning
Explainable output	Presents diagnosis and contributing factors	Enables human verification
Human–AI interaction	Allows clinicians to review AI decisions	Strengthens trust and clinical acceptance

5. Applications in Disease Diagnosis

Because they can combine heterogeneous diseases' clinical indicators and maintain an interpretable reasoning structure, Explainable Fuzzy Inference Systems (EFISs) have been utilized in a wide variety of diseases (Cao et al., 2024). Fuzzy systems have been applied to the fusion of various risk factors such as age, blood pressure, cholesterol level, heart rate, electrocardiographic indexes, etc., to estimate the probability of cardiovascular disease diagnosis. They have some principles for language use that apply to understanding cardiovascular risk associated with combinations of risk factors. The fuzzy logic techniques that are frequently used in diabetes mellitus include glucose level, body mass index, age, blood pressure, insulin-related parameters, and family history (Ennab, 2025). The ability to show a gradual transition between readings that are normal and abnormal is very helpful in identifying patients with varying degrees of diabetes risk.

In cancer detection and classification, fuzzy systems can combine clinical, pathological, imaging, and biomarker information. Interpretable rules can be used to link combinations of tumor attributes to benign or malignant outcomes and aid in the clinical evaluation of model decisions (Nemade, 2025). In the case of Parkinson's disease, fuzzy and neuro-fuzzy models have been studied based on motor symptoms, speech characteristics, gait parameters, tremor-related features, and other clinical measurements. These are also applicable to other neurodegenerative diseases, such as Huntington's disease, in which symptoms may differ significantly from patient to patient (Alsadie, 2024). To classify Alzheimer's disease and estimate disease progression, fuzzy approaches can integrate cognitive measures, demographic information, neuropsychological evaluations, and features derived from neuroimaging in a way that strengthens the classification and estimation (Abdullakutty et al., 2024).

Other applications where fuzzy diagnostic systems have shown potential include liver and kidney disease, where several laboratory biomarkers may have overlapping or unclear diagnostic ranges (Loey & Khalil, 2026). In the field of respiratory diseases, symptoms, pulmonary-function measurements, parameters related to oxygen, and imaging findings can be described by fuzzy variables to aid in identification and severity evaluation of the disease (Halawani et al., 2026). Another important use is in the diagnosis of infectious diseases, as symptoms and laboratory parameters can change over time, and they might not be diagnostic in the early stages. These uncertain indicators can be combined to estimate the risk of infection using fuzzy models (Zhao et al., 2026).

In addition to these applications, EFISs could also be used for chronic and complex diseases with compound risk factors (Mayeta-Revilla et al., 2025), such as metabolic diseases, arthritis, thyroid diseases, and obesity, among others. The broad applicability of fuzzy diagnostic systems arises from the features of combining numerical patient information with expert knowledge and linguistic reasoning (Srivastava & Srivastava, 2026). Due to the differences in datasets, features selected, rule generation techniques, validation approaches, and clinical settings, direct comparisons between existing studies are challenging (Abbas et al., 2024).

Table 2. Applications of Explainable Fuzzy Inference Systems in Disease Diagnosis

Disease/Domain	Typical Diagnostic Inputs	Potential Explainable Fuzzy Output
Cardiovascular disease	Blood pressure, cholesterol, age, heart rate, ECG features	Low/medium/high cardiovascular risk
Diabetes mellitus	Glucose, BMI, insulin, age, blood pressure	Diabetes risk or diagnostic category
Cancer	Tumor characteristics, biomarkers, imaging features	Benign/malignant classification or risk
Parkinson’s disease	Tremor, gait, speech, motor symptoms	Disease classification/severity
Huntington’s disease	Motor, cognitive, behavioural and imaging features	Disease status/progression risk
Alzheimer’s disease	Cognitive scores, imaging, neurological measures	Disease classification/progression
Liver disease	Liver enzymes, bilirubin, age, clinical indicators	Disease likelihood/severity
Kidney disease	Creatinine, urea, GFR, blood pressure	Kidney disease risk/stage
Respiratory disease	Symptoms, oxygen level, pulmonary measures, imaging	Disease probability/severity
Infectious diseases	Symptoms, biomarkers, laboratory findings	Infection risk/classification
Other chronic diseases	Demographic, clinical, laboratory variables	Personalized risk or diagnostic category

In summary, the applications presented here highlight the potential of XFS as a tool to connect data-driven medical AI with human clinical reasoning, especially in contexts where uncertainty, linguistic knowledge, and personalized diagnostic insights are significant.

6. Hybrid Explainable Fuzzy AI Approaches

Hybrid Explainable Fuzzy Artificial Intelligence (AI) approaches are designed to boost the diagnostic power of the intelligent healthcare system and retain an interpretable reasoning structure through the integration of fuzzy inference with machine learning, optimization, and advanced AI methods (Singh, 2025). A Neuro-Fuzzy Inference System combines artificial neural networks and fuzzy logic: it is possible to learn the membership functions and fuzzy rules from the data, but the rule-based representation of the diagnostic knowledge is preserved (Cao et al., 2024). The learning algorithms to optimize fuzzy parameters with medical datasets are especially crucial for adaptive neuro-fuzzy inference systems (ANFIS). ANFIS can thus be used to model nonlinear relationships between the clinical variables and produce relatively understandable diagnostic rules (Mohanraj & Sujatha, 2025).

Furthermore, fuzzy systems can be used together with decision trees, with the concepts of fuzzy memberships being uncertain borders between the clinical categories. These models can be used to mitigate overlapping patient characteristics while maintaining the hierarchical nature of decision-making without losing the decision-making process (Ennab, 2025). The Fuzzy – Genetic Algorithm (FG) approaches include evolutionary optimization techniques for selecting relevant features, optimizing the parameters of the membership function, and removing irrelevant and redundant rules (Sharma et al., 2025). Similarly, Fuzzy–Support Vector Machine (SVM) classifiers can combine fuzzy incomplete physician information with the high classification power of SVM, particularly when there are conflicting or noisy data sets (Zacarias-Morales et al., 2025).

Fuzzy inference could also be combined with ensemble learning to further enhance robustness, which means that multiple classifiers or rule systems can be combined together. Meanwhile, the fusion of fuzzy logic and deep learning offers a chance to leverage the high prediction power of neural architectures while adding fuzzy representations to the otherwise complex models (Abdullah et al., 2026). Explainable fuzzy–neural architectures can convert the learned features into linguistic concepts, and then diagnose using fuzzy rules. Recently, the integration of Fuzzy systems and Large Language Models (LLMs) has become a new field of research (Khanna et al., 2025). LLMs can help identify clinical narratives as structured language variables, create candidate rules, or simplify technical model results and output them as clinician-readable explanations. The challenge with such integration is to ensure that the explanations produced by the language model are aligned with the actual reasoning of the diagnostic model, which may differ from the explanations generated by the language model. (Nemade, 2025).

Table 3. Hybrid Explainable Fuzzy AI Approaches

Hybrid Approach	Primary Function	Explainability Potential	Major Limitation
Neuro-Fuzzy Systems	Combines neural learning with fuzzy reasoning	Interpretable rules and membership functions	Increasing model complexity
ANFIS	Learns fuzzy parameters from data	Moderate–high	Rule explosion with many inputs
Fuzzy–Decision Tree	Combines fuzzy uncertainty with hierarchical classification	High	Complex trees may reduce readability
Fuzzy–Genetic Algorithm	Optimizes rules, features, and membership functions	High if rules remain compact	Computationally expensive
Fuzzy–SVM	Handles uncertainty and nonlinear classification	Moderate	SVM decision structure can be difficult to explain
Fuzzy–Ensemble Models	Improves robustness through multiple models	Moderate	Multiple models increase complexity
Fuzzy–Deep Learning	Combines representation learning with fuzzy reasoning	Potentially high	Deep feature representations may remain opaque
Fuzzy–Neural Architectures	Integrates neural features with linguistic rules	High when rules are well designed	Difficult optimization
Fuzzy–LLM Systems	Integrates linguistic AI with fuzzy reasoning	Potentially very high	Hallucination and explanation-faithfulness concerns

The key benefit of hybrid fuzzy AI is its ability to achieve a balance between learning capability, predictive performance, uncertainty management, and interpretability. However, each extra computational unit can add complexity, extra computational costs, and problems of model validation. Thus, hybrid systems cannot be said to be intrinsically explainable just because fuzzy logic is part of the system. To make explainability effective, it is necessary to have clear rules, comprehensible linguistic variables, plausible explanations, and clinically valid results. In the future, hybrid architectures promising high diagnostic accuracy and being simultaneously easy to use and reliable for clinical decisions should be developed.

7. Evaluation of Explainable Fuzzy Diagnostic Systems

The assessment of Explainable Fuzzy Diagnostic Systems (EFDSs) must be performed on a multi-dimensional basis, considering both the diagnostic effectiveness and the quality of explanations. While the traditional measures of performance provide data on the system's ability to accurately diagnose a disease, the measures of explainability provide information about whether the system's reasoning can be understood and assessed clinically (Nemade, 2025). Thus, it is important to consider predictive performance, the properties of the rules, the quality of the explanations produced, and the clinical usefulness in combination within an effective evaluation framework (Bui, 2025).

The diagnostic performance is typically evaluated by accuracy, sensitivity, and specificity. In the present study, accuracy shows the percentage of correctly classified cases, and sensitivity and specificity show the ability to accurately classify patients with a disease and the ability to accurately classify patients without the disease (Tharwat, 2021). Sensitivity is critical in medical diagnosis that could have severe implications if the disease is not detected. Precision, recall, and F1 score give complementary information, especially in the case of imbalanced datasets. Precision refers to the percentage of correctly predicted positive patients, and recall is the same as sensitivity (Carrington et al., 2022). The F1 score gives a balance between precision and recall.

ROC-AUC is a measure that is often employed to assess the discriminatory power of diagnostic models over a range of classification levels. Also, Precision–Recall curves can be very informative for imbalanced datasets in the medical domain, as it highlights performance for positive cases. But if a fuzzy system is able to predict well, this does not prove that it is truly explainable (Hicks et al., 2022).

There are several ways to measure the explainability of rules: rule complexity, number of rules, and rule length (De Diego et al., 2022), linguistic complexity, fidelity, stability, and consistency. Usually, a model with fewer rules and shorter antecedents is easier to understand by clinicians, as long as important diagnostic relationships are not lost (Chicco & Jurman, 2023). Linguistic Interpretability is the ability to determine whether the membership functions and linguistic labels are meaningful clinical concepts. Fidelity evaluates if the explanation captures the model's decision-making process accurately, while stability evaluates if the same inputs lead to consistent explanations. Consistency is also relevant, as conflicting explanations might decrease clinical trust (Fraiwan & Faouri, 2022).

Table 4. Evaluation Criteria for Explainable Fuzzy Diagnostic Systems

Evaluation Dimension	Metric/Indicator	Purpose
Diagnostic performance	Accuracy	Overall classification correctness
Diagnostic performance	Sensitivity/Recall	Ability to detect diseased patients
Diagnostic performance	Specificity	Ability to identify non-diseased patients
Diagnostic performance	Precision	Reliability of positive predictions
Diagnostic performance	F1-score	Balance between precision and recall
Discrimination	ROC-AUC	Overall class-separation capability
Imbalanced-data performance	Precision–Recall AUC	Evaluation of positive-class performance
Rule interpretability	Number of rules	Measures rule-base complexity
Rule interpretability	Rule length	Assesses cognitive complexity
Linguistic interpretability	Membership-function clarity	Determines clinical meaningfulness
Explanation quality	Fidelity	Measures faithfulness to model reasoning
Explanation quality	Stability	Evaluates consistency across similar cases
Explanation quality	Consistency	Detects contradictory reasoning
Clinical evaluation	Expert assessment	Determines practical clinical usefulness
Overall assessment	Accuracy–interpretability trade-off	Balances predictive power and transparency

Clinical validity and expert assessment are key components, as mathematical interpretability does not necessarily translate into clinical utility. Physicians may determine if fuzzy rules reflect accepted medical knowledge and if explanations are clear enough, and if the system gives them information they can act on clinically. Trust, comprehension, decision self-confidence, and willingness to use are additional factors that could be analyzed in user studies.

One of the challenges is thus the balance between accuracy and interpretability. More complicated models can yield high predictive accuracy, but generate too many rules or explanations that are hard to understand. On the other hand, very basic rule sets can be very easily understood but not very accurate for use in clinical application. The goal should then be to find a balance that allows the model to have a high level of diagnostic accuracy with compact, clinically relevant, faithful, and understandable reasoning. Such a multi-dimensional assessment is crucial in designing reliable fuzzy AI systems that can be used in healthcare decision-making in real-world applications.

8. Challenges, Research Gaps, and Future Directions

Although Explainable Fuzzy Diagnostic Systems (EFDSs) have the potential to be widely used in clinical practice, there are still some challenges that have to be addressed. Medical data can be complex, containing structured clinical data, medical images, laboratory data, genomic data, physiological signals, and unstructured clinical narratives (Singh, 2025). The integration of these heterogeneous information sources into a single interpretable fuzzy framework. Medical information is also uncertain, vague, noisy, and has missing values (Ennab, 2025). Fuzzy logic is very suitable for dealing with gradual and uncertain data, but the data set may be incomplete, and the measurement values may not be consistent, thus causing the rule to be activated and affecting the reliability of the diagnosis.

Other challenges that are important are rule explosion. The combinations of fuzzy rules can be large as the number of input variables and linguistic categories grows, as described in Bui (2025). Too many rules make the algorithm complex and complicated to understand for clinicians. The problem gets worse when dealing with high-dimensional medical data sets that have hundreds or thousands of possible features. Some of these methods include feature selection, rule pruning, dimensionality reduction, and automated knowledge extraction to preserve interpretability without compromising significantly on the diagnostic performance (Talpur et al., 2023).

One research gap is that there are no consistent metrics for explainability. Existing studies often have achieved very good accuracy, sensitivity, specificity, and F1 score but failed to have a meaningful quantitative evaluation of the quality of explanation (Azad et al., 2026). Evaluation frameworks should be systematically developed to include measures like rule complexity, explanation fidelity, stability, consistency, and human comprehension. In the same way, it is difficult to compare existing fuzzy diagnostic systems directly due to differences in datasets, pre-processing techniques, validation methodologies, and clinical populations used. A further challenge is dataset bias, as models developed using the data from a specific hospital, geographic area, demographic group or disease population may not generalise to other clinical contexts (Kondaveeti & Simhadri, 2025).

Another significant research gap is the need for clinical validation and in vivo application. Prospective clinical studies are rarely used to evaluate many proposed fuzzy diagnostic systems, and instead retrospective or publicly available datasets are used. Physicians, patients, and healthcare institutions should continue to be involved in future research in both developing and validating models. The clinical utility, usability, robustness, and safety of the system must also be shown, along with the statistical performance. In addition to statistical performance, explainable systems should also demonstrate clinical utility, usability, robustness, and safety (Muhammad & Bendechache, 2025). Moreover, privacy, security, and ethical concerns are to be taken into account when patient-sensitive information is used for model training and inference. It is crucial to have the right data governance, anonymization, secure computation, and responsible data-sharing mechanisms in place.

Additionally, regulatory compliance and trust among users are also essential for the adoption of medical AI. Healthcare professionals must have trust in an AI system to ensure reliability, transparency, reproducibility, and proper validation. Explanations should be correct and not give a misleading interpretation. So, future evidence requirements for regulation will increasingly include evidence of model performance, explanation fidelity, robustness, fairness, and accountability (Wyatt et al., 2024).

Table 5. Challenges, Research Gaps, and Future Directions

Challenge/Research Gap	Current Limitation	Future Direction
Complex medical data	Heterogeneous and high-dimensional inputs	Develop multimodal fuzzy architectures
Uncertainty and missing data	Incomplete or imprecise observations	Advanced type-2 and uncertainty-aware fuzzy models
Rule explosion	Excessive combinations of rules	Rule pruning and automated rule optimization
High-dimensional data	Reduced interpretability	Feature selection and interpretable dimensionality reduction
Explainability metrics	Lack of standardized measures	Develop unified explanation evaluation frameworks
Dataset bias	Limited population generalizability	Diverse, multicenter, representative datasets
Clinical validation	Limited real-world testing	Prospective and multicenter clinical trials
Privacy and security	Sensitive patient information	Privacy-preserving and secure AI methods
Regulation and trust	Unclear requirements for explanations	Clinically validated and auditable AI frameworks
Multimodal data	Difficulty integrating different data types	Unified fuzzy representation of clinical modalities
Personalization	Generic diagnostic recommendations	Patient-specific adaptive fuzzy models
Human-centered AI	Limited clinician involvement	Interactive physician–AI decision-support systems

Multimodal and personalized explainable fuzzy AI should be further studied in the future. All these data, such as clinical records, laboratory data, medical images, physiological signals, genetic information, and patient narratives, could be integrated into a single system within multimodal systems. These systems could yield more complete diagnostic reasoning without obscuring intermediate representations that are interpretable. Personalized fuzzy models could also tune membership functions, rules, and even temporal health status changes according to the individual patient.

Lastly, the future of EFDSs should focus on human-centered clinical AI – where AI is used to augment and not replace clinicians. Physicians might be able to view active rules, adjust their assumptions, provide feedback, and assess other diagnostic scenarios in interactive systems. Incorporating validated clinical knowledge into models may be possible through the use of continuous learning mechanisms, while maintaining regulatory and safety controls. Fuzzy logic coupled with explainable machine learning, multimodal learning, neuro-symbolic AI, and other new language-based systems, therefore, might offer a way to develop diagnostic technologies that are accurate, transparent, robust, ethical, personalized, and clinically trustworthy.

9. Comparative Review and Critical Discussion

Based on the literature reviewed, conventional machine learning (ML), black-box AI, and fuzzy diagnostic models have various combinations of prediction power, transparency, and clinical usability. When structured medical information is available in large quantities, conventional ML techniques including decision trees, support vector machines, random forests, and ensemble classifiers could yield good diagnostic performance. But they have differing levels of interpretability. Unlike fuzzy diagnostic models, however, these are models that represent uncertainty in terms of linguistic variables and explicit IF–THEN rules, and which can be more easily examined by clinicians when reasoning. While deep learning models can excel in processing complex data like medical images and physiological signals, understanding their internal representations is often challenging, and further XAI methods are required.

In healthcare, the black-box vs. interpretable AI issue is critical. While black-box models are able to find very complicated nonlinear relationships, they may lack the ability to explain why a prediction was made. Interpretable fuzzy systems provide access to their decision mechanisms by means of membership functions and rules, thus enabling the clinician to analyze the link between the symptoms, biomarkers, and the diagnostic outcomes. However, the smaller the number of rules that are automatically generated, or the more complex the membership functions, the less interpretable the fuzzy models become. Therefore, fuzzy logic is not automatically explainable, but depends on the design and complexity of the whole system, and hence, on the quality of explanations.

There are also differences in the advantages of the various fuzzy inference methods. The advantage of the Mamdani systems is that they tend to be highly interpretable from a linguistic perspective since the antecedents and consequents can be expressed in linguistic terms. Although the Sugeno and TSK models are computationally efficient and can be optimized and adapted to learning, the mathematical consequents may not be as easily understood by clinical users. However, the flexibility of type-2 fuzzy systems comes at the cost of increased complexity in computation and concept, while adding greater capabilities in representing uncertainty in membership functions and expert knowledge.

Table 6. Comparative Analysis of AI Approaches for Disease Diagnosis

Approach	Diagnostic Capability	Interpretability	Uncertainty Handling	Expert Knowledge	Clinical Trust
Conventional ML	High	Low–moderate	Moderate	Limited	Moderate
Decision Tree	Moderate–high	High	Limited	Moderate	High
Deep Learning	Very high for complex data	Low	Limited–moderate	Limited	Moderate–low
Mamdani FIS	Moderate–high	Very high	High	High	High
Sugeno/TSK FIS	High	Moderate	High	Moderate–high	Moderate–high
Type-2 FIS	High	High	Very high	High	High
ANFIS	High	Moderate	High	Moderate	Moderate–high
Hybrid Fuzzy AI	High–very high	Moderate–high	High	High	Potentially high

One key problem that arises from the literature is the accuracy–interpretability paradigm. There is a possibility of having very complex models that can capture the subtle relationships and get very good predictive performance, but if the model is too complex, it's hard to be sure of their decisions. On the other hand, very simple fuzzy rule bases can be understood easily, but cannot capture the complexity of multifactorial diseases. The most promising systems are thus those that have a balanced architecture, with good predictive performance, a relatively small number of rules, an easy-to-understand set of linguistic variables, and clinically meaningful relationships.

There is also interdomain variation in performance. Fuzzy approaches are specifically suited for diseases where the symptoms, risk factors, and biomarkers are gradual or involve significant uncertainty, such as diabetes, cardiovascular disorders, chronic diseases, etc. For neurological disorders like Parkinson's disease, Huntington's disease, and Alzheimer's disease, the symptoms and the progression of the disease can be very different between patients, and thus fuzzy reasoning could be beneficial. Deep learning could offer more robust pattern recognition capabilities in cancer diagnosis and medical imaging, and fuzzy layers or interpretable rule systems could help to explain the prediction. Thus, there is no one most effective intervention for all clinical domains.

One of the key advantages of fuzzy diagnostic systems is the application of expert knowledge. Clinical experts can specify meaningful linguistic variables, set initial rules, automatically validate relationships, and extract clinically implausible outcomes. This provides a chance for "collaborative" AI, where computational learning and medical knowledge are used together. But relying too heavily on rules that have been set by experts can add subjectivity and restrict scalability. Data-driven rule learning and expert knowledge are thus more desirable than just relying on one of the two methods.

Transparent reasoning alone is not enough for clinical usability and trustworthiness. Clinicians need systems that are accurate, stable, understandable, robust and compatible with their current workflows. An explanation should contain meaningful diagnostic factors and an accurate representation of the models' reasoning. Techno-complex and inconsistent explanations that diverge from accepted medical information may diminish, not build, trust. Human evaluation should thus be used in addition to traditional performance measures.

Table 7. Critical Synthesis of Reviewed Approaches

Dimension	Major Finding	Critical Observation
Accuracy	Hybrid and learning-based models often achieve strong performance	Accuracy alone does not establish clinical usefulness
Interpretability	Fuzzy rule-based models provide strong transparency	Large rule bases can reduce interpretability
Uncertainty	Fuzzy systems effectively represent gradual and imprecise information	Missing and noisy data remain challenging

Expert knowledge	Improves clinical relevance and rule validation	Subjectivity may affect consistency
Generalization	Performance may decline across institutions and populations	Multicenter validation is required
Explainability	Linguistic rules offer intuitive explanations	Standardized explanation metrics remain limited
Clinical trust	Transparency can improve acceptance	Trust requires evidence of reliability and safety
Scalability	Hybrid systems can handle complex relationships	Increasing complexity may compromise interpretability

In general, the literature that has been reviewed shows that explainable FIS are a middle ground between purely symbolic clinical decision systems and highly complex data-driven systems of AI. The primary advantage of them is the ability to integrate uncertainty management, expert knowledge, linguistic reasoning, and interpretable decision rules. They also suffer from drawbacks such as rule explosion, scalability issues, reliance on the definition of the membership functions, substandard evaluation approaches, and inadequate prospective clinical validation.

The critical synthesis thus shows that further research should take the next step beyond the mere aim of obtaining maximum accuracy of classification. Among these, faithful explanations and compact rule structures, clinical validation, fairness and robustness, multimodal integration, and human–AI collaboration should receive greater emphasis. Explicit and clinically meaningful fuzzy reasoning with the representation-learning capability of advanced machine learning, especially in the form of hybrid fuzzy architectures, seems to hold especially promise. The usefulness of an explainable fuzzy diagnostic system should be measured not only by its prediction accuracy of the disease, but also by the ability of the clinician to understand, validate, and apply the logical justification in actual clinical situations.

III. Conclusion

Explainable Fuzzy Inference Systems (EFISs) are a promising way of creating transparent and clinically meaningful Artificial Intelligence (AI) systems for the diagnosis of diseases. With the help of fuzzy logic, this review has shown how to tackle the uncertainty, vagueness, imprecision, and gradual transition inherent in medical information without losing the interpretability of the linguistic variables and IF–THEN rules. By combining fuzzy inference with rule-based artificial intelligence, the diagnostic decision can be presented in a way that allows a healthcare provider to analyse, validate, and/or revise the decision. These strategies have been applied to a range of cardiovascular diseases, diabetes, cancer, neurological diseases, liver and kidney diseases, respiratory diseases, and infectious diseases, and have shown promise in many of these areas. The paper also emphasizes the importance of hybrid architectures such as neuro-fuzzy, ANFIS, fuzzy–machine learning, and fuzzy–deep learning systems to enhance adaptability and diagnostic accuracy. However, hurdles such as rule explosion, high-dimensional data, dataset bias, varying explainability measures, privacy issues, and limited clinical validation in the real world still exist. The authors recommend using modular and clinically relevant rule bases, multi-modal and individual diagnostic systems, standardized explainability assessment, robust external validation, and human-centred clinical AI in future research. In total, EFISs can provide valuable linkages between computational intelligence and reliable clinical reasoning.

Reference

[1]. Abbas, Q., Jeong, W., & Lee, S. W. (2025, August). Explainable AI in clinical decision support systems: a meta-analysis of methods, applications, and usability challenges. In *Healthcare* (Vol. 13, No. 17, p. 2154). MDPI.

[2]. Abbas, S., Asif, M., Rehman, A., Alharbi, M., Khan, M. A., & Elmitwally, N. (2024). RETRACTED: Emerging research trends in artificial intelligence for cancer diagnostic systems: A comprehensive review. *Heliyon*, *10*(17).

[3]. Abdullah, A. S., Mishra, U., Geetha, S., & Prasad, P. (2026). A Synergistic Deep Learning and Fuzzy Logic Framework for Medical Diagnostics. *Procedia Computer Science*, *283*, 5316-5328.

[4]. Abdullakutty, F., Akbari, Y., Al-Maadeed, S., Bouridane, A., & Hamoudi, R. (2024). Advancing histopathology-based breast cancer diagnosis: insights into multi-modality and explainability. *arXiv preprint arXiv:2406.12897*.

[5]. Ahsan, M. M., Luna, S. A., & Siddique, Z. (2022, March). Machine-learning-based disease diagnosis: A comprehensive review. In *Healthcare* (Vol. 10, No. 3, p. 541). MDPI.

[6]. Al-Ali, A., & Qidwai, U. (2025). Rule-based modeling of low-dimensional data with PCA and Binary Particle Swarm Optimization (Bpso) in Anfis. *arXiv preprint arXiv:2502.03895*.

[7]. Alnattah, A., Jajroudi, M., Fadafen, S. A. N., Manzari, M. N., & Eslami, S. (2025). Artificial Intelligence in Clinical Decision-Making: A Scoping Review of Rule-Based Systems and Their Applications in Medicine. *Cureus*, *17*(8), e91333-e91333.

[8]. Alsadie, D. (2024). Artificial intelligence techniques for securing fog computing environments: trends, challenges, and future directions. *IEEE Access*, *12*, 151598-151648.

[9]. Ambalavanan, R., Snead, R. S., Marczika, J., Towett, G., Malioukis, A., & Mbogori-Kairichi, M. (2025). Ontologies as the semantic bridge between artificial intelligence and healthcare. *Frontiers in Digital Health*, *7*, 1668385.

[10]. Awan, M. B. (2023). *Evaluation and improvement of energy flexibility and performance of building heating, ventilation, and air-conditioning systems* (Doctoral dissertation, University of Wollongong).

[11]. Azad, M., Fahad, N. M., Raiaan, M. A. K., Anik, T. R., Khan, M. F. K., Toyé, H. M. K., & Muhammad, G. (2026). A systematic review of diffusion models for medical image-based diagnosis: Methods, taxonomies, clinical integration, explainability, and future directions. *Diagnostics*, *16*(2), 211.

[12]. Barzegar, Y. (2026). Technical and economic impact of the huge internet of things and artificial intelligence in the creation of the intelligent water network.

[13]. Bui, K. (2025). Deep neuro fuzzy systems for explainable AI in healthcare: A literature review.

[14]. Cao, J., Zhou, T., Zhi, S., Lam, S., Ren, G., Zhang, Y., ... & Cai, J. (2024). Fuzzy inference system with interpretable fuzzy rules: Advancing explainable artificial intelligence for disease diagnosis—A comprehensive review. *Information Sciences*, *662*, 120212.

[15]. Carrington, A. M., Manuel, D. G., Fieguth, P. W., Ramsay, T., Osmani, V., Wernly, B., ... & Holzinger, A. (2022). Deep ROC analysis and AUC as balanced average accuracy, for improved classifier selection, audit and explanation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *45*(1), 329-341.

[16]. Chamola, V., Hassija, V., Sulthana, A. R., Ghosh, D., Dhingra, D., & Sikdar, B. (2023). A review of trustworthy and explainable artificial intelligence (XAI). *IEEE Access*, *11*, 78994-79015.

[17]. Chicco, D., & Jurman, G. (2023). The Matthews correlation coefficient (MCC) should replace the ROC AUC as the standard metric for assessing binary classification. *BioData mining*, *16*(1), 4.

[18]. D'Aniello, G. (2023). Fuzzy logic for situation awareness: a systematic review. G. D'Aniello. *Journal of Ambient Intelligence and Humanized Computing*, *14*(4), 4419-4438.

[19]. De Diego, I. M., Redondo, A. R., Fernández, R. R., Navarro, J., & Moguerza, J. M. (2022). General Performance Score for classification problems: IM De Diego et al. *Applied Intelligence*, *52*(10), 12049-12063.

[20]. Domingues, N. S. (2025). A hybrid decision support system using rule-based and AI methods: the OnCATs knowledge-based framework. *International Journal of Medical Informatics*, 106144.

[21]. Dwivedi, R., Dave, D., Naik, H., Singhal, S., Omer, R., Patel, P., ... & Ranjan, R. (2023). Explainable AI (XAI): Core ideas, techniques, and solutions. *ACM Computing Surveys*, *55*(9), 1-33.

[22]. Ennab, M. M. H. (2025). *A hybrid convolutional-fuzzy model for interpretable AI in healthcare: improving transparency and accuracy in chronic disease management* (Doctoral dissertation, Université du Québec à Chicoutimi).

[23]. Fahim, Y. A., Hasani, I. W., Kabba, S., & Ragab, W. M. (2025). Artificial intelligence in healthcare and medicine: clinical applications, therapeutic advances, and future perspectives. *European journal of medical research*, *30*(1), 848.

[24]. Fernández, I., Puente, J., Ponte, B., & Gómez, A. (2024). Integration of AHP and fuzzy inference systems for empowering transformative journeys in organizations: Assessing the implementation of Industry 4.0 in SMEs. *Applied Intelligence*, *54*(23), 12357-12377.

[25]. Fraiwan, M., & Faouri, E. (2022). On the automatic detection and classification of skin cancer using deep transfer learning. *Sensors*, *22*(13), 4963.

[26]. George Kerry, R., Ukhurebor, K. E., Kumari, S., Maurya, G. K., Patra, S., Panigrahi, B., ... & Patra, J. K. (2021). A comprehensive review on the applications of nano-biosensor-based approaches for non-communicable and communicable disease detection. *Biomaterials Science*, *9*(10), 3576-3602.

[27]. Halawani, H. T., Senan, E. M., Asem Othman, E. M., Khan, M. A., Asiri, Y., Alyousef, A., & Fati, S. M. (2026). Interpretable hybrid CNN–ViT framework with neuro-symbolic clinical decision support for early cervical cancer diagnosis. *Scientific Reports*.

[28]. Hassija, V., Chamola, V., Mahapatra, A., Singal, A., Goel, D., Huang, K., ... & Hussain, A. (2024). Interpreting black-box models: a review on explainable artificial intelligence. *Cognitive Computation*, *16*(1), 45-74.

[29]. Hicks, S. A., Strümke, I., Thambawita, V., Hammou, M., Riegler, M. A., Halvorsen, P., & Parasa, S. (2022). On evaluation metrics for medical applications of artificial intelligence. *Scientific reports*, *12*(1), 5979.

[30]. Hostetter, J. W. (2025). *Gradient-Based Neuroplastic Adaptation for Concurrent Optimization of Neuro-Fuzzy Networks*. North Carolina State University.

[31]. Hussain, F., Hussain, R., & Hossain, E. (2021). Explainable artificial intelligence (XAI): An engineering perspective. *arXiv preprint arXiv:2101.03613*.

[32]. Islam, S. R., Eberle, W., Ghafoor, S. K., & Ahmed, M. (2021). Explainable artificial intelligence approaches: A survey. *arXiv preprint arXiv:2101.09429*.

[33]. Kalasampath, K., Spoorthi, K. N., Sajeev, S., Kuppa, S. S., Ajay, K., & Maruthamuthu, A. (2025). A literature review on applications of explainable artificial intelligence (XAI). *IEEE Access*, *13*, 41111-41140.

[34]. Kambalimath, S., & Deka, P. C. (2020). A basic review of fuzzy logic applications in hydrology and water resources. *Applied Water Science*, *10*(8), 191.

[35]. Kaya, O., Shah, A. S., Karabulut, M. A., Karahan, S. N., Osmanca, M. S., & Acır, N. (2026). Explainable artificial intelligence (XAI): Concepts, applications, challenges, and future perspectives. *IEEE Access*.

[36]. Khanna, S., Khatri, S. K., & Sharma, N. K. (2025). Explainable AI integrated Fuzzy Rule-Based Machine Learning and Nonlinear Variation Inequalities for Oral Cancer Disease Detection and Treatment Methodology in Healthcare. *Advances in Nonlinear Variational Inequalities*, *28*, 362-363.

[37]. Kim, S. Y., Kim, D. H., Kim, M. J., Ko, H. J., & Jeong, O. R. (2024). XAI-based clinical decision support systems: a systematic review. *Applied Sciences*, *14*(15), 6638.

[38]. Kondaveeti, H. K., & Simhadri, C. G. (2025). Evaluation of deep learning models using explainable AI with qualitative and quantitative analysis for rice leaf disease detection. *Scientific Reports*, *15*(1), 31850.

[39]. Koohsari, H., Alielahi, H., Najafi, A., & Adampira, M. (2026). Evaluation and optimization of effective depth of dynamic compaction using Sugeno fuzzy inference system and PSO. *Geomechanics and Engineering*, *45*(4), 485-516.

[40]. Kopitar, L., Fister Jr, I., & Stiglic, G. (2024). Using generative AI to improve the performance and interpretability of rule-based diagnosis of type 2 diabetes mellitus. *Information*, *15*(3), 162.

[41]. Kostopoulos, G., Davrazos, G., & Kotsiantis, S. (2024). Explainable artificial intelligence-based decision support systems: A recent review. *Electronics*, *13*(14), 2842.

[42]. Kowalski, P. A., Casari, M., & Po, L. (2025). Explainable AI for rule reduction in fuzzy models for air pollution measurement adjustment. *Environmental Modelling & Software*, 106734.

[43]. Kumar, A., Singh, S. K., Samanta, B., & Bhattcherjee, A. (2025). Risk assessment in sociotechnical systems based on functional resonance analysis method and hierarchical fuzzy inference tree. *Scientific Reports*, *15*(1), 23827.

[44]. Leelavathy, S., Ramprasath, J., Mary, I. M., Nagendran, R., & Kumar, R. D. (2025). Harnessing fuzzy logic framework to quantify diagnostic uncertainty in medical decision support. *Fuzzy Information and Engineering*, *17*(4), 425-445.

[45]. Loey, M., & Khalil, H. M. (2026). Hierarchical Fuzzy System Integrated with Deep Learning for Robust and Interpretable Classification of Breast Malignancies Using Radiomics Features from Ultrasound Imaging. *Computers*, *15*(3), 147.

[46]. Maleki Varnosfaderani, S., & Forouzanfar, M. (2024). The role of AI in hospitals and clinics: transforming healthcare in the 21st century. *Bioengineering*, *11*(4), 337.

[47]. Maria, A. J., Castiello, C., Luis, M., & Mencar, C. (2021). Explainable fuzzy systems: Paving the way from interpretable fuzzy systems to explainable AI systems. *Studies in Computational Intelligence*, *970*, 1-253.

[48]. Mayeta-Revilla, L., Cavieres, E. P., Salinas, M., Mellado, D., Ponce, S., Torres Moyano, F., ... & Salas, R. (2025). Radiomics-driven neuro-fuzzy framework for rule generation to enhance explainability in MRI-based brain tumor segmentation. *Frontiers in Neuroinformatics*, *19*, 1550432.

[49]. Mjahad, A., & Rosado-Muñoz, A. (2025). Combining Deep Learning Architectures with Fuzzy Logic for Robust Pneumonia Detection in Chest X-Rays. *Applied Sciences*, *15*(19), 10321.

[50]. Mogotsi, L. L., Rasool, A., Matlotse, E., Ali, S., & Ali, A. (2025). Intelligent fault detection of MV/HV transformers using fuzzy logic based on DGA. *Eng*, *6*(9), 228.

[51]. Mohanraj, S., & Sujatha, R. (2025). A novel CNN–fuzzy–XAI approach for Alzheimer's disease severity classification using brain MRI scans. *Cogent Engineering*, *12*(1), 2575105.

[52]. Muhammad, D., & Bendechache, M. (2025). More than just a heatmap: elevating XAI with rigorous evaluation metrics. *Frontiers in Medical Technology*, *7*, 1674343.

[53]. Muhammad, K., Obaidat, M. S., Hussain, T., Ser, J. D., Kumar, N., Tanveer, M., & Doctor, F. (2021). Fuzzy logic in surveillance big video data analysis: Comprehensive review, challenges, and research directions. *ACM Computing Surveys (CSUR)*, *54*(3), 1-33.

[54]. Naseef, R. S., Mukhtar, A. M., Mahammed, R. A., Attallah, S. B., Abdulmohsin, A. S., Alsabah, M., ... & Nyangaresi, V. O. (2026, July). An Interpretable FP-Growth-Based Association Rule Mining Framework for Multi-Source Obesity Risk Analysis Using NHANES Data. In *2026 18th International Conference on Electronics, Computers and Artificial Intelligence (ECAI)* (pp. 1-10). IEEE.

[55]. Navin, K., & Krishnan, M. (2024). Fuzzy rule-based classifier model for evidence-based clinical decision support systems. *Intelligent Systems with Applications*, *22*, 200393.

[56]. Nemade, M. (2025). An Explainable Fuzzy Deep Learning Framework for Uncertainty-Based Medical Diagnosis. *International Journal of Fuzzy Mathematical Archive*, *24*(2), 47-58.

[57]. Ogut, E. (2025). Artificial intelligence in clinical medicine: challenges across diagnostic imaging, clinical decision support, surgery, pathology, and drug discovery. *Clinics and Practice*, *15*(9), 169.

[58]. Ortigossa, E. S., Gonçalves, T., & Nonato, L. G. (2024). Explainable artificial intelligence (XAI)—from theory to methods and applications. *IEEE Access*, *12*, 80799-80846.

[59]. Osman, A. B., & Owolabi, A. (2023). Explainable AI Models for Clinical Decision Support Using Large-Scale Healthcare Data.

[60]. Pabarja, R., Jamali, G., Salimifard, K., & Ghorbanpur, A. (2024). Analysis of the LARG of the hospital medical equipment supply chain using the fuzzy inference system. *International journal of research in industrial engineering*, *13*(2), 116-151.

[61]. Pinto-Coelho, L. (2023). How artificial intelligence is shaping medical imaging technology: a survey of innovations and applications. *Bioengineering*, *10*(12), 1435.

[62]. Sadr, H., Nazari, M., Khodaverdian, Z., Farzan, R., Yousefzadeh-Chabok, S., Ashoobi, M. T., ... & Yamaghani, M. R. (2025). Unveiling the potential of artificial intelligence in revolutionizing disease diagnosis and prediction: a comprehensive review of machine learning and deep learning approaches. *European journal of medical research*, *30*(1), 418.

[63]. Sharma, A., Singh, D., Singh, A., Kantipudi, M. P., Potharaju, S., & Babu, B. S. (2025). Comparative development and evaluation of multilayer fuzzy and adaptive neuro-fuzzy inference systems for early diagnosis of chronic kidney disease. *Informatics in Medicine Unlocked*, 101724.

[64]. Shukla, A. K., Mehra, P., & Muhuri, P. K. (2025). Fuzzy sets-based approaches for improved medical diagnosis: an analysis and overview of major research directions. *ACM Computing Surveys*.

[65]. Singh, A. P., Biswas, A., Shukla, A., & Maiti, P. (2019). Targeted therapy in chronic diseases using nanomaterial-based drug delivery vehicles. *Signal transduction and targeted therapy*, *4*(1), 33.

[66]. Singh, S. (2025). Neuro-fuzzy architectures for interpretable AI: A comprehensive survey and research outlook. *Journal of Machine Learning Research*, *1*(11).

[67]. Sirocchi, C., Bogliolo, A., & Montagna, S. (2024). Medical-informed machine learning: integrating prior knowledge into medical decision systems. *BMC medical informatics and decision making*, *24*(Suppl 4), 186.

[68]. Sofianidis, G., Rožanec, J. M., Mladenic, D., & Kyriazis, D. (2021). A review of explainable artificial intelligence in manufacturing. *Trusted artificial intelligence in manufacturing: a review of the emerging wave of ethical and human-centric AI technologies for smart production*, *24*, 93-113.

[69]. Srivastava, A., Bhanot, D., Jasim, L. H., Varshney, N., & Patil, V. (2025). Advancements in fuzzy logic applications for diagnostic decision support systems in healthcare. *Fuzzy Information and Engineering*, *17*(3), 284-297.

[70]. Srivastava, K., & Srivastava, R. (2026). Explainable deep learning approaches and clinical insights for cancer biomarker identification. *Frontiers in Oncology*, *16*, 1810793.

[71]. Talpur, N., Abdulkadir, S. J., Alhussian, H., Hasan, M. H., Aziz, N., & Bamhdi, A. (2023). Deep Neuro-Fuzzy System application trends, challenges, and future perspectives: A systematic survey. *Artificial Intelligence Review*, *56*(2), 865-913.

[72]. Tharwat, A. (2021). Classification assessment methods. *Applied computing and informatics*, *17*(1), 168-192.

[73]. Turgunbaev, R. (2025). Rule-based reasoning and its role in intelligent decision-making. *Technical Science Integrated Research*, *1*(2), 11-14.

[74]. Wyatt, L. S., van Karnenbeek, L. M., Wijkhuizen, M., Geldof, F., & Dashtbozorg, B. (2024). Explainable artificial intelligence (XAI) for oncological ultrasound image analysis: a systematic review. *Applied Sciences*, *14*(18), 8108.

[75]. Xu, Q., Xie, W., Liao, B., Hu, C., Qin, L., Yang, Z., ... & Luo, A. (2023). Interpretability of clinical decision support systems based on artificial intelligence from technological and medical perspectives: a systematic review. *Journal of Healthcare Engineering*, *2023*(1), 9919269.

[76]. Xu, Z., Zhang, Y., & Zhang, H. (2026). Identification of BMI-related high-risk feature combinations for diabetes among young adults with normal baseline fasting plasma glucose using interpretable machine learning: a health check-up cohort study. *Frontiers in Endocrinology*, *17*, 1850071.

[77]. Ye, Y., Pandey, A., Bawden, C., Sumsuzzman, D. M., Rajput, R., Shoukat, A., ... & Galvani, A. P. (2025). Integrating artificial intelligence with mechanistic epidemiological modeling: a scoping review of opportunities and challenges. *Nature Communications*, *16*(1), 581.

[78]. Yerkin, A., Igali, A., Kadyrgali, E., Shagyrov, M., Ziyada, M., Muratbekova, M., & Shamoi, P. (2025). Fuzzy theory in computer vision: A review. *arXiv preprint arXiv:2507.18660*.

[79]. Zacarias-Morales, N., Pancardo, P., Hernández-Nolasco, J. A., & García-Constantino, M. (2025). Artificial neural network, attention mechanism, and fuzzy logic-based approaches for medical diagnostic support: A systematic review. *AI*, *6*(11), 281.

[80]. Zhao, J., Liu, Y., Wang, G., Zhou, M., & Yuan, L. (2026). Compact Fuzzy-Rule Decision-Level Fusion for Ovarian Cancer Survival Prediction with Controlled Modality Extension. *IEEE Transactions on Fuzzy Systems*.