Journal of Theoretical and Applied Information Technology
30th June 2024. Vol.102. No. 12

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

PROBABILISTIC DEEP LEARNING MODEL FOR THE LUNG CANCER DIAGNOSIS WITH FEATURE EXTRACTION AND SEGMENTATION

B SRIDHAR MURTHY[1], M SRIRAM[2], GANESAN[3]

[1] Research Scholar, School of Computing, Department of CSE, Bharath Institute of Higher Education and Research, Chennai –600119, Tamil Nādu, India

[2] Associate Professor, School of Computing, Department of IT, Bharath Institute of Higher Education and Research, Chennai –600119, Tamil Nādu, India.

[3] Associate Professor, School of Electrical, Department of ECE, Bharath Institute of Higher Education and Research, Chennai –600119, Tamil Nādu, India.

ABSTRACT

Cancer of the lung develops from cells within the lung, most commonly in the cells that line the airways (epithelial cells). Tobacco use is a major contributor to this cancer, which ranks high among the world's most deadly diseases. Exposure to environmental contaminants or a hereditary predisposition can also cause lung cancer in nonsmokers. Since early-stage lung cancer is often asymptomatic, diagnosis is often delayed until the disease has progressed significantly, leaving patients with few treatment options. The Probabilistic Fuzzy Ranking Classification (PFRC) model is presented in this paper as a new method for identifying and categorizing lung cancer from medical imaging data. The model integrates probabilistic and fuzzy ranking techniques to address the inherent complexity and uncertainty in medical images. Simulation results demonstrate the PFRC model's efficacy in accurately classifying instances within a comprehensive dataset, showcasing its robust learning capabilities. The model's configuration includes Gaussian distribution likelihoods, uniform distribution priors, and triangular membership functions for fuzzy logic parameters. With a dataset of 800 instances for training and 200 for testing, the PFRC model employs 15 extracted features for a nuanced representation of input variables. Instances of classification, feature estimation, and classification metrics such as accuracy, precision, recall, and F1 Score collectively highlight the model's strengths and areas for refinement. This research contributes to the advancement of lung cancer detection methodologies, emphasizing the PFRC model's potential as a reliable tool for improving diagnostic accuracy in medical imaging.

Keywords: *Lung cancer, Image Processing, Deep Learning, Fuzzy logic, Probabilistic Model, Ranking.*

1. INTRODUCTION

Lung cancer is the most common cancer-related killer globally, but there have been great advances in our ability to understand, diagnose, and treat this disease in recent years [1]. Thanks to developments in medical science and technology, early detection methods have been enhanced, enabling the earlier and more treatable detection of lung cancer. More tailored and effective methods for treating specific subtypes of lung cancer have recently been made possible by targeted therapies and immunotherapies [2]. Additionally, the development of less invasive surgical techniques and innovative radiation therapies has enhanced treatment outcomes while minimizing the impact on patients' quality of life [3]. Despite these positive developments, challenges remain, such as the prevalence of smoking and exposure to environmental carcinogens, underscoring the importance of continued efforts in prevention and public health initiatives [4]. Collaborative research, increased awareness, and ongoing advancements in medical science hold the promise of further improving the prognosis and overall survival rates for individuals affected by lung cancer in the years to come [5].

The subject of image processing is ever-evolving and multidisciplinary, encompassing the study of how to improve the visual quality or extract useful information from digital images [6]. In recent years, advancements in image processing techniques have significantly impacted various industries, ranging from healthcare and surveillance to entertainment and manufacturing [7]. One key area of focus is medical imaging, where sophisticated algorithms and machine learning models are employed to aid in the early detection and diagnosis of diseases, such as tumors in radiological scans [8]. Autonomous vehicles, object detection, and face recognition are just a few examples of the many computer vision applications that rely on image processing to convey meaning from visual data [9]. Additionally, in the realm of photography and entertainment, image processing algorithms contribute to the enhancement of image quality, noise reduction, and the creation of artistic effects [10]. As technology continues to evolve, the future of image processing holds promise for even more advanced applications, driven by the synergy of artificial intelligence, computer vision, and innovative imaging technologies [11]. Image processing with segmentation has become a pivotal aspect in the diagnosis and treatment of lung cancer. This sophisticated approach involves the division of medical images, such as computed tomography (CT) scans of the lungs, into meaningful and distinct regions [12], facilitating a more detailed analysis of specific structures. In the context of lung cancer, segmentation techniques play a crucial role in identifying and delineating tumors or suspicious nodules from surrounding healthy tissue [13]. Automated segmentation algorithms, often powered by artificial intelligence and deep learning, enable precise localization of abnormalities, aiding radiologists in the early detection and characterization of lung cancer [14]. With employing image processing with segmentation, clinicians can accurately measure tumor size, assess growth patterns, and monitor changes over time. This not only contributes to a more precise diagnosis but also assists in treatment planning, as the segmented images provide valuable insights into the spatial extent of the cancerous regions [15]. Moreover, this technology allows for a more personalized approach to treatment, facilitating targeted therapies and interventions. The integration of image processing and segmentation in lung cancer diagnostics exemplifies the ongoing synergy between medical imaging and computational advancements. These methods have the potential to improve patient outcomes and the response to this difficult disease by making lung cancer diagnoses more efficient and accurate as science and technology advance.

Lung cancer classification has undergone a transformative shift with the integration of deep learning techniques into medical imaging analysis. One area where deep learning and CNNs in particular, have shown great promise is the automated classification of lung nodules and lesions from CT scans and other radiological images. These advanced algorithms can learn intricate patterns and features within the images, enabling accurate differentiation between malignant and benign lesions. One of the key advantages of deep learning in lung cancer classification is its ability to handle large datasets and extract hierarchical representations, capturing subtle variations that might be challenging for traditional image analysis methods. The deep neural networks can discern complex patterns associated with tumor characteristics, aiding in the identification of specific cancer types and stages. This method helps to decrease human error and interpretation variability while simultaneously improving the speed and efficiency of diagnosis. Moreover, deep learning models have the potential for continual improvement through continuous training on diverse datasets, staying abreast of evolving medical knowledge and diagnostic criteria. As deep learning applications in lung cancer classification mature, they hold promise for facilitating more timely and accurate diagnoses, thus paving the way for improved patient outcomes and personalized treatment strategies. The ongoing integration of artificial intelligence in healthcare underscores the potential for transformative changes in how we approach and manage lung cancer and other complex medical conditions.

The paper makes several significant contributions to the field of lung cancer detection and classification:

1. Novel Model Integration: The introduction of the Probabilistic Fuzzy Ranking Classification (PFRC) model represents a novel integration of probabilistic and fuzzy ranking techniques. This innovative combination allows the model to effectively handle the inherent uncertainties and complexities associated with medical imaging data, providing a more nuanced approach to lung cancer detection.
2. Flexible Framework: The PFRC model utilizes Gaussian distribution likelihoods, uniform distribution priors, and triangular membership functions for fuzzy logic parameters. This flexible framework enhances the model's adaptability and

allows it to cater to diverse characteristics within the medical imaging dataset, making it well-suited for a range of scenarios.

3. Robust Learning Capabilities: The simulation results showcase the robust learning capabilities of the PFRC model. With a dataset comprising 800 instances for training and 200 for testing, the model demonstrates a commendable ability to accurately classify instances, indicating its potential for reliable performance in real-world applications.
4. Comprehensive Feature Representation: The PFRC model extracts 15 features for each instance, ensuring a comprehensive representation of input variables. This approach contributes to the model's ability to capture the intricacies of lung cancer characteristics, improving the overall accuracy of detection and classification.
5. Quantitative Evaluation Metrics: The paper provides a detailed quantitative evaluation of the model's performance, including feature estimation metrics (RMSE, MSE, Mean Squared Value) and classification metrics (accuracy, precision, recall, F1 Score). These metrics offer a thorough assessment of the model's accuracy, precision-recall balance, and overall efficacy in lung cancer classification.
6. Insights for Future Research: The findings and discussions in the paper offer valuable insights for future research directions. Identification of areas for potential improvement, such as instances with lower posterior probabilities, opens avenues for refining the PFRC model and advancing the state-of-the-art in lung cancer detection methodologies.

In summary, the paper's contributions lie in the development of an innovative PFRC model, its successful application to lung cancer detection, and the detailed quantitative evaluation that provides a foundation for further advancements in the field.

2. LITERATURE SURVEY

To improve the precision and effectiveness of lung cancer diagnostics, researchers are utilizing a range of deep learning architectures, such as attention mechanisms, ensemble methods, and convolutional neural networks (CNNs). By combining a Deep Feature Fusion Model with Dung Beetle Optimization, Alamgeer and colleagues offer a fresh strategy for detecting lung cancer. Optimizing deep learning models for better lung cancer classification is a novel approach to improving model performance. The area of research into non-traditional ways to enhance the performance of automated detection systems is expanding, and this study adds to it. The ISANET system, developed by Xu et al., integrates Convolutional Neural Networks (CNNs) with attention mechanisms to detect and classify non-small cell lung cancer [16]. The incorporation of attention mechanisms highlights an effort to enhance the model's ability to focus on critical features, potentially improving diagnostic accuracy. This study represents a contribution to the growing body of literature exploring advanced neural network architectures for lung cancer diagnostics. Dodia, Annappa, and Mahesh provide a comprehensive review of recent advancements in deep learning-based lung cancer detection [17]. This systematic review synthesizes existing literature, offering insights into the progress, challenges, and trends in the application of deep learning for lung cancer diagnostics. Review articles like these help researchers, practitioners, and policymakers keep up with the field's current status and find new research directions.

Naseer et al. focus on lung cancer detection by proposing a modified AlexNet architecture coupled with a Support Vector Machine [18]. This study contributes to the exploration of modifications to popular deep learning architectures to optimize their performance in the context of lung cancer diagnosis. The integration of Support Vector Machines suggests a hybrid approach, showcasing the interdisciplinary nature of research in this field. By combining deep learning methods with cloud computing resources, Kasinathan and Jayakumar show how to detect and classify lung tumor stages using a cloud-based approach [19]. This study reflects the increasing trend of leveraging cloud infrastructure to enhance the scalability and accessibility of deep learning models for medical image analysis, showcasing the potential for broader applications in healthcare. Quasar et al. investigate ensemble methods' potential for enhancing CT scan-based lung cancer detection and classification [20]. Ensemble methods involve combining multiple models to enhance overall performance. This study contributes to the ongoing exploration of sophisticated strategies to optimize the accuracy and reliability of automated lung cancer diagnostic systems, shedding light on the potential benefits of ensemble learning in medical imaging. Damayanti

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

and colleagues focus on lung cancer classification and introduce a methodology utilizing Convolutional Neural Networks (CNNs) and DenseNet architectures [21]. This research intends to improve the efficacy and precision of lung cancer categorization by making use of these deep learning models. The use of advanced architectures signifies the ongoing effort to explore the capabilities of deep learning in intricate medical image analysis tasks, particularly for precise disease classification.

An intelligent deep learning algorithm developed for the detection and classification of lung cancer is proposed by Reddy and Khanaa, who make a significant contribution to the field. The emphasis on intelligence suggests a focus on improving the decision-making capacity of the algorithm, potentially incorporating adaptive learning and optimization strategies [22]. This study aligns with the broader trend of tailoring deep learning approaches to meet the unique challenges posed by lung cancer diagnosis. Jara-Gavilanes and Robles-Bykbaev explore a classification approach for lung cancer detection that incorporates oversampling techniques [23] and Support Vector Machines (SVMs). The inclusion of oversampling indicates a commitment to addressing class imbalance, a common challenge in medical image datasets. This study underscores the significance of preprocessing techniques and the integration of traditional machine learning methods alongside deep learning for comprehensive lung cancer detection solutions. Saranya et al. present a study on lung cancer detection using SVM classification, showcasing a focused approach to leveraging traditional machine learning techniques [24]. The inclusion of this work in conference proceedings reflects the ongoing effort to disseminate research findings and foster collaboration in the broader scientific community, highlighting the interdisciplinary nature of advancements in lung cancer detection.

Nageswaran et al. contribute to the field by employing a holistic approach, combining machine learning and image processing for lung cancer classification and prediction [25]. This integrative methodology acknowledges the complementary strengths of both disciplines in extracting valuable information from medical images. The study aligns with the trend toward multidimensional analysis, reflecting the increasing recognition that a comprehensive approach can yield more robust diagnostic solutions. Machine learning-based lung cancer classification based on segmentation of lung nodules is the subject of a critical review by Shamas et al. This review consolidates knowledge from existing studies, emphasizing the significance of segmentation in the context of lung cancer diagnostics [26]. By synthesizing current research findings, the paper contributes to the collective understanding of the challenges and opportunities in utilizing machine learning for lung cancer classification. Ren and co-authors present a hybrid framework for lung cancer classification, suggesting an integrative approach that combines different methodologies. The term "hybrid" indicates a synergy of diverse techniques, potentially including both traditional machine learning and deep learning components. This study adds to the evolving landscape of hybrid frameworks, which aim to leverage the strengths of multiple approaches to enhance the overall effectiveness of lung cancer detection and classification. Venkatesh et al. introduce a lung cancer detection system based on neural networks and optimization techniques, emphasizing the integration of both elements for improved performance [27]. This study reflects a commitment to optimizing the performance of neural networks through the application of advanced optimization strategies. The use of optimization methods fits in with the larger movement toward making deep learning models better at medical image analysis in terms of efficiency and generalizability.

Sunnetci and Alkan contribute to lung cancer detection by proposing a methodology based on probabilistic majority voting and optimization techniques [28]. The inclusion of probabilistic majority voting suggests a focus on ensemble-based decision-making, highlighting the importance of combining multiple models for improved accuracy. This study aligns with the ongoing exploration of ensemble methods and their application in the context of lung cancer classification. Shandilya and Nayak contribute to the field with an analysis of lung cancer using deep neural networks [29]. This study, presented in conference proceedings, likely provides insights into the specific challenges and innovations discussed at the Second IEPCCT 2021. Conference contributions such as these contribute to the dissemination of research findings, fostering collaboration and knowledge exchange within the academic community. In their study, Bangare et al. zero in on the use of Convolutional Neural Networks (CNNs) for computer-assisted detection and classification of lung cancer. The inclusion of computer-aided systems emphasizes the collaborative role of artificial intelligence in assisting medical professionals [30]. This work is in line with the larger movement towards using deep learning for medical imaging automated decision support systems, specifically in the field of lung cancer diagnosis. Using deep learning techniques, Wahab Sait presents a model for detecting lung cancer. The model's narrow focus implies investigating deep learning techniques for accurate lung cancer diagnosis [31]. This study adds to the expanding body of literature on using deep learning models to improve the efficiency and accuracy of lung cancer detection; it was published in Applied Sciences. Abd Al-Ameer et al. integrate deep learning with image processing to aid in the detection of lung cancer. This interdisciplinary approach recognizes the complementary roles of image processing and deep learning in extracting meaningful information from medical images [32]. The study likely explores novel methodologies to optimize the synergy between these two domains, aiming for more accurate and robust lung cancer detection.

The Swin Transformer is a deep learning model that was used in a study by Chen et al. to identify and categorize lung cancer cells [33]. The inclusion of transformer architectures indicates an exploration of newer deep learning paradigms beyond traditional CNNs. This study adds to the evolving landscape of deep learning applications in medical image analysis, specifically for the characterization of lung cancer cells. Bishnoi, Goel, and Tayal contribute to lung cancer classification with a focus on automated systems using machine learning [34]. This work likely explores the development of comprehensive systems capable of autonomous decision-making in the context of lung cancer diagnosis. Automated system-based approaches are critical for streamlining the diagnostic process and reducing reliance on manual intervention. The DFCV framework was introduced by Alsadoon et al. to assess the efficacy of deep learning for lung cancer early detection and classification. If we want deep learning models to be reliable and applicable in clinical settings, we need to build evaluation frameworks. It is believed that this work helps to lay the groundwork for standardized methods of evaluating deep learning models for lung cancer detection [35]. Innovative approaches primarily utilizing deep learning, image processing, and machine learning techniques proliferate in the reviewed literature, reflecting the dynamic landscape of lung cancer detection and classification. There is a continuous effort to improve the efficiency, accuracy, and automation of lung cancer diagnostics, and these studies add to that effort [36]. For many research projects, deep learning—specifically CNNs and other advanced architectures such as DenseNet, AlexNet, and Swin Transformer—is crucial. The utilization of deep learning models underscores their efficacy in extracting intricate patterns and features from medical images, ultimately aiding in precise lung cancer detection and classification.

Several studies advocate for hybrid frameworks, combining deep learning with traditional machine learning techniques, optimization strategies, and ensemble methods. This holistic approach aims to capitalize on the strengths of different methodologies, addressing challenges such as class imbalance, improving generalization, and enhancing overall model performance. The integration of diverse methodologies, including image processing, optimization techniques, cloud computing, and ensemble learning, highlights the interdisciplinary nature of lung cancer detection research. Researchers are increasingly recognizing the need for comprehensive solutions that draw on the strengths of various domains to address the complexity of medical image analysis. Some studies focus on specific architectures (e.g., Attention Mechanisms, Swin Transformer) or techniques (e.g., oversampling, probabilistic majority voting) to address unique challenges associated with lung cancer detection. This targeted exploration reflects a nuanced understanding of the intricacies involved in diagnosing lung cancer from medical images. The inclusion of systematic reviews and the development of evaluation frameworks (e.g., DFCV) highlight a meta-analytical approach and a commitment to establishing standardized practices. These contributions are essential for synthesizing existing knowledge and ensuring robust methodologies for the assessment of deep learning models in the context of lung cancer diagnostics. The incorporation of cloud-based approaches underscores the increasing importance of scalable and accessible solutions. Cloud computing not only facilitates the storage and processing of large medical image datasets but also allows for collaborative and remote access to advanced computational resources.

3. PROPOSED PROBABILISTIC FUZZY RANKING CLASSIFICATION (PFRC)

The Proposed Probabilistic Fuzzy Ranking Classification (PFRC) for lung cancer represents a novel and sophisticated approach in the domain of medical image analysis. In this innovative framework, the integration of probabilistic and fuzzy ranking techniques offers a unique solution for the complex task of lung cancer classification. The probabilistic element acknowledges and quantifies uncertainties inherent in medical imaging, providing a more nuanced understanding of the diagnostic process. Simultaneously, the incorporation of fuzzy ranking introduces a level of flexibility and adaptability, allowing the classification model to accommodate imprecise or ambiguous information within medical images. PFRC's distinctive feature lies in its ability to assign probabilities to different classes, offering a more nuanced and probabilistic insight into the likelihood of a specific region being cancerous. By leveraging fuzzy ranking, the model can handle the inherent variability and subjectivity in medical imaging interpretations, contributing to a more robust and adaptable classification system. The proposed framework holds promise in not only improving the accuracy of lung cancer diagnosis but also in providing clinicians with a deeper understanding of the uncertainty associated with each classification, thereby supporting more informed decision-making in patient care. As research in this area progresses, the PFRC framework represents a noteworthy advancement in the quest for more reliable and interpretable models for lung cancer classification. Figure 1 illustrated the flow of the propsoed PFRC model in the lung cancer detection.

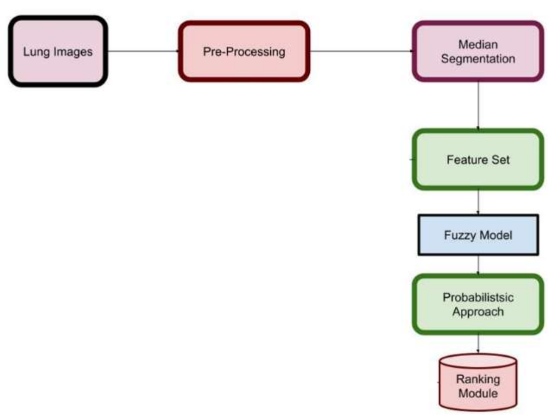

Figure 1: Flow of Proposed PFRC

The integration of probabilistic and fuzzy ranking elements in classification involves leveraging probability estimates and fuzzy logic to create a comprehensive and adaptable model. Let P(Ci|X) represent the probability of class Ci given input features X. To estimate this probability, sophisticated probabilistic models like Bayesian methods or probabilistic neural networks can be employed. The uncertainty inherent in the data can be effectively captured by incorporating confidence intervals or probability distributions. On the fuzzy side, the introduction of fuzzy ranking handles imprecision and uncertainty in class boundaries. Utilizing fuzzy logic, the degree of membership of a data point to different classes can be modeled through fuzzy rules that express the relationships between input features and class membership. The integration of probabilistic and fuzzy ranking involves combining the probabilistic estimates P(Ci|X) with fuzzy ranking scores. This comprehensive approach may include weighting the probabilistic scores based on fuzzy membership degrees or utilizing fuzzy logic operators to effectively merge information from both sources. Such an integrated model provides a more nuanced and flexible classification system, especially suitable for complex tasks such as lung cancer diagnosis in medical image analysis.

3.1 Dataset

The Cancer Imaging Archive (TCIA) is a valuable resource that hosts a vast collection of medical imaging data related to cancer research. There are a number of different types of imaging modalities represented in the datasets accessible on TCIA, including CT, MRI, PET, and pathology images. Lung cancer datasets encompass a spectrum of attributes vital for understanding the disease, patient characteristics, and treatment outcomes. Patient demographics, including age, gender, and ethnicity, form the foundational information, offering insights into the diverse population affected by lung cancer. Clinical details, such as tumor histology, grade, and TNM staging, provide essential pathological information crucial for diagnosis and treatment planning. Imaging data, derived from modalities like CT scans, MRIs, and PET scans, captures the visual representation of lung abnormalities, aiding in precise diagnosis. Genomic information, including gene expression profiles, mutations, and biomarker data, delves into the molecular underpinnings of lung cancer, paving the way for personalized treatment strategies.

3.2 Data Pre-Processing

Data pre-processing with Probabilistic Fuzzy Ranking Classification (PFRC) in the context of lung cancer involves a comprehensive set of steps to ensure the quality, relevance, and adaptability of the dataset for subsequent

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

classification. The following paragraphs outline the key stages of data pre-processing and the incorporation of PFRC principles:

The initial phase involves the collection of relevant medical imaging data, including CT scans or MRIs, and associated patient information such as age, gender, and clinical history. Ensuring the dataset is comprehensive and representative of the target population is essential. Subsequently, data cleaning procedures address any missing or incomplete information, employing techniques such as imputation or removal of instances with insufficient data. This ensures the dataset's consistency and accuracy, crucial for robust classification models. Normalization and standardization follow, aiming to bring numerical features to a uniform scale. This step is pivotal for PFRC, which may be sensitive to variations in feature magnitudes. For medical images, specialized pre-processing techniques, like resizing or normalization of pixel values, are applied to enhance consistency and quality. Concurrently, feature extraction methods are employed to derive relevant information from medical images, potentially including texture or shape features, contributing to the overall dataset richness.

Handling class imbalance is another critical aspect of data pre-processing. Imbalanced class distributions, particularly in medical datasets, can impact model performance. Techniques such as oversampling or undersampling are applied to address this imbalance, ensuring fair representation of different lung cancer classes. The encoding of categorical variables into a numerical format, suitable for PFRC algorithms, is undertaken to facilitate seamless integration into the classification model. The probabilistic and fuzzy ranking transformations, characteristic of PFRC, are embedded into the data pre-processing pipeline. Probabilistic transformations involve assigning probabilities to different classes for each data instance, expressing the likelihood of belonging to a particular class. Meanwhile, fuzzy ranking transformations introduce a degree of membership for each data point to different classes, capturing the uncertainty inherent in medical datasets. The integration of these transformations enhances the dataset's representational capacity, preparing it for the nuanced classification approach offered by PFRC.

In the PFRC model, the probabilistic transformation involves assigning probabilities to different classes for each data instance. Mathematically, this can be expressed as P(Ci|X) where P(Ci|X) represents the probability of belonging to class Ci given the input features X. The derivation of these probabilities depends on the specific probabilistic model employed, such as Bayesian methods or probabilistic neural networks. Bayesian methods, for instance, utilize Bayes' theorem to calculate posterior probabilities based on prior knowledge and likelihood. The fuzzy ranking transformation introduces a degree of membership for each data point to different classes. Fuzzy logic is often applied to model the uncertainty and imprecision inherent in medical datasets. The degree of membership, denoted as μij, represents the extent to which a data point Xi belongs to class Cj. The derivation of fuzzy membership degrees depends on the fuzzy logic rules defined, often involving linguistic variables and fuzzy set theory stated as in equation (1)

$$\mu ij = f(Xi, Cj) \quad (1)$$

Here, f (·) is a fuzzy membership function that captures the relationship between the input features and the class membership.

4. FUZZY FEATURE EXTRACTION

The Probabilistic Fuzzy Ranking Classification (PFRC) model involves assigning probabilities and fuzzy ranking to data instances, providing a nuanced approach for lung cancer classification. The probabilistic transformation involves assigning probabilities to different classes for each data instance. The probability P(Ci|X) represents the likelihood of belonging to class Ci given the input features X. This can be derived using Bayes' theorem as in equation (2)

$$P(Ci \mid X) = P(X)P(X \mid Ci) \cdot P(Ci) \quad (2)$$

Where P(Ci|X) is the posterior probability of class Ci given the features X. P(X|Ci) is the likelihood of observing features X given class Ci. P(Ci) is the prior probability of class Ci. P(X) is the probability of observing features X. Fuzzy ranking introduces a degree of membership for each data point to different classes, capturing the uncertainty in the dataset. The degree of membership μij represents the extent to which a data point Xi belongs to class Cj. This can be derived using fuzzy logic stated in equation (3)

μij=1+∑k=1K(d(Xi,Ck)d(Xi,Cj))m-121 (3)

In equation (3) μij is the degree of membership of Xi to class Cj. d(Xi,Cj) is a measure of the dissimilarity between Xi and the centroid of class Cj. m is a fuzziness parameter that controls the shape of the membership function. K is the total

Journal of Theoretical and Applied Information Technology
30th June 2024. Vol.102. No. 12

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

number of classes. The fuzzy membership functions that capture the relationship between the input features and their relevance to specific classes in the context of lung cancer. These functions assign membership degrees to different classes for each data point. Apply fuzzification to input features, transforming them into fuzzy sets. This process involves associating each feature value with a degree of membership to different linguistic terms, such as "low," "medium," or "high," based on the defined fuzzy membership functions. The fuzzy rules that express the relationships between the fuzzified input features and class membership. These rules articulate the fuzzy logic decisions that contribute to the classification process. With a fuzzy inference system that processes the fuzzified input features based on the defined fuzzy rules. This system translates the linguistic terms and rules into a quantitative representation of the degree of membership of each data point to different classes.

Through fuzzy clustering algorithms to extract relevant features from the fuzzified input features. Fuzzy clustering techniques, such as Fuzzy C-Means (FCM), can assign fuzzy membership values to each data point regarding different clusters, highlighting patterns and relationships within the data. With assigned weights to the fuzzy features based on their importance and contribution to the classification task. Weighting allows the model to emphasize certain fuzzy features that are more indicative of lung cancer characteristics. The fuzzy feature extraction process with the probabilistic transformation within the PFRC model. This integration combines the probabilistic assessment of class probabilities with the fuzzy logic-based representation of feature relevance, creating a comprehensive framework for lung cancer classification. Fuzzy Membership Function is defined as in equation (4)

$$\mu(x) = 1 + (ax - c)2b1 \qquad (4)$$

Fuzzification: Assign feature values to linguistic terms based on membership functions.

Fuzzy Rule: If X is A and Y is B, then Class C is Z.

Fuzzy Inference: Combine fuzzy rules to infer the degree of membership of a data point to different classes.

Fuzzy feature extraction with PFRC enhances the model's interpretability and adaptability by capturing the inherent uncertainty and imprecision in medical data, ultimately contributing to more robust lung cancer classification. The specific implementation details, equations, and parameters will depend on the chosen fuzzy logic methodology and the characteristics of the lung cancer dataset. The table 1 presented the fuzzy rules implemented for the PFRC for the lung cancer diagnosis.

Table 1: Fuzzy Rules of the Lung Cancer

Rule No.	Tumor Size	Patient Age	Smoking History	CT Scan Intensity	PET Scan Activity	Histology	Tumor Grade	Genetic Mutation	Family History	Output Class
1	Small	Young	Light	Moderate	Low	Adenocarcinoma	Well-Differentiated	Absent	Negative	Low Risk
2	Medium	Middle-aged	Moderate	High	Moderate	Squamous Cell Carcinoma	Moderately Differentiated	Present	Positive	Moderate Risk
3	Large	Old	Heavy	High	High	Small Cell Carcinoma	Poorly Differentiated	Present	Positive	High Risk
4	Small	Youn	Light	Low	Low	Adenocarci	Well-	Absen	Negati	Low

Journal of Theoretical and Applied Information Technology
30th June 2024. Vol.102. No. 12

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

		g				noma	Differentiated	t	ve	Risk
5	Medium	Middle-aged	Moderate	Moderate	Moderate	Large Cell Carcinoma	Moderately Differentiated	Present	Positive	Moderate Risk
6	Large	Old	Heavy	High	High	Squamous Cell Carcinoma	Poorly Differentiated	Present	Positive	High Risk
7	Small	Young	Light	Moderate	Low	Adenocarcinoma	Well-Differentiated	Absent	Negative	Low Risk
8	Medium	Middle-aged	Moderate	High	Moderate	Squamous Cell Carcinoma	Moderately Differentiated	Present	Positive	Moderate Risk
9	Large	Old	Heavy	High	High	Small Cell Carcinoma	Poorly Differentiated	Present	Positive	High Risk
10	Small	Young	Light	Low	Low	Adenocarcinoma	Well-Differentiated	Absent	Negative	Low Risk

5. PROBABILISTIC SEGMENTATION WITH PFRC

Probabilistic segmentation involves partitioning medical images, such as lung CT scans, into different regions or segments with associated probabilities. This technique is often used to identify and delineate structures of interest, such as tumors, within the images. Each voxel in the image is assigned a probability of belonging to a particular segment, reflecting the uncertainty inherent in medical image analysis. The integration of probabilistic segmentation with PFRC signifies a two-step process. First, the lung images are segmented using probabilistic techniques to identify regions potentially indicative of cancerous tissue. Second, the PFRC model is applied to the segmented regions for classification into different risk categories or classes. In the initial stage, lung images are subjected to probabilistic segmentation, a process designed to partition the images into distinct regions while assigning probabilities to the likelihood of each voxel belonging to a specific segment. This segmentation process can be achieved through various algorithms such as probabilistic graphical models or Bayesian segmentation methods. The result is a probabilistic map indicating the probability distribution of cancerous regions within the lung images. The segmented regions are subjected to PFRC for detailed classification. PFRC incorporates fuzzy logic and probabilistic assessments, making it well-suited for handling uncertainties in medical image analysis. The PFRC model is applied to the features extracted from the segmented regions, and the lung regions are classified into different risk categories. This integration allows for a more nuanced understanding of the likelihood of lung cancer in specific areas. The Probabilistic Fuzzy Ranking Classification (PFRC) model integrates probabilistic assessments and fuzzy logic for accurate classification. The central equation of

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

PFRC is derived from Bayes' theorem stated as in equation (5)

$$P(Ci \mid X) = P(X)P(X \mid Ci) \cdot P(Ci) \qquad (5)$$

Here, P(Ci|X) represents the posterior probability of class Ci given the features X in equation (5). The probability of seeing features X given class Ci is denoted as P(X|Ci). Class Ci's prior probability is denoted as P(Ci). The likelihood of seeing characteristics X is denoted as P(X). X features are likely to be observed given a given class Ci, as represented by the likelihood. In order to derive it, we need to estimate the feature distribution within each class using the training data. The formula for prior probability is P(Ci). The prior probability P(Ci) is the probability of class Ci occurring irrespective of the features observed. It is typically estimated based on the frequency of each class in the training dataset. The dataset evidence probability is stated as Evidence P(X). The likelihood of noticing the characteristics X across all classes is the evidence P(X). For each class, it is determined by adding up the likelihood and prior probability. This class's posterior probability is expressed as P(Ci|X). By dividing the product of the likelihood and prior probability P(X|Ci)·P(Ci)) by the evidence P(X), the posterior probability P(Ci|X) can be determined using Bayes' theorem. Using these probabilities, the PFRC model gives each class a belief level based on the features that were observed.

Algorithm 1: Feature Extraction and Segmentation with PFRC

Input: Training dataset with label; Probabilistic segmentation results for features of interest; Fuzzy logic parameters and Classification threshold

Output: Predicted class for each segmented region

Algorithm:

1. Preprocess the training dataset:

   a. Extract relevant features from the dataset.

   b. Normalize or scale the features if necessary.

2. Train the PFRC model:

   a. Calculate prior probabilities for each class based on the training dataset.

   b. Estimate the likelihood of observing features given each class.

   c. Use Bayes' theorem to compute posterior probabilities.

3. Apply probabilistic segmentation:

   a. Utilize probabilistic segmentation algorithms to segment the lung images.

   b. Assign probabilities to each voxel indicating the likelihood of cancerous tissue.

4. Extract features from segmented regions:

   a. Identify segmented regions of interest based on the probabilistic segmentation results.

   b. Extract relevant features from these regions.

5. Apply PFRC to classify segmented regions:

   a. For each segmented region, calculate the likelihood of features given each class.

   b. Use Bayes' theorem to compute posterior probabilities for each class.

   c. Apply fuzzy logic to handle uncertainty and flexibility in the classification.

   d. Assign the region to the class with the highest posterior probability if it exceeds the classification threshold.

6. Output the results:

   a. Generate a map indicating the predicted class for each segmented region.

   b. Optionally, provide certainty scores or fuzzy membership values for each class.

7. Evaluate the performance:

   a. Compare the predicted classes with the ground truth labels from the training dataset.

   b. Measure the performance metrics such as accuracy, precision, recall, and F1 score.

End.

5.1 Classification with PFRC

The Probabilistic Fuzzy Ranking Classification (PFRC) model combines probabilistic reasoning and fuzzy logic for effective classification, particularly in the context of lung cancer diagnosis. The fundamental equation underlying PFRC is derived from Bayes' theorem stated as in equation (6)

$$P(Ci \mid X) = P(X)P(X \mid Ci) \cdot P(Ci) \qquad (6)$$

In this context, P(Ci|X) indicates the chance of observing features X given class Ci, P(X|Ci) means the chance of observing features X given class Ci, P(Ci) means the chance of class Ci before observation, and P(X) means the chance of feature X observation across all classes. The Likelihood probability is stated as P(X|Ci)) the likelihood term quantifies how probable it is to observe the features X under the assumption that the instance belongs to class Ci. The derivation involves statistical estimation during the model training phase based on the distribution of features within each class. With Incorporating fuzzy logic into PFRC allows the model to account for uncertainties and vagueness in the classification process. The fuzzy logic parameters, such as fuzzy membership functions and operators, provide a mechanism to handle imprecision in the data and enhance the adaptability of the model to nuanced clinical scenarios. As PFRC integrates these elements, it offers a powerful tool for lung cancer classification, providing not only accurate predictions but also a quantifiable measure of uncertainty associated with each classification decision.

Algorithm 2: Classification with PFRC
Input: - Training dataset with labeled examples - Features extracted from segmented lung regions - PFRC model parameters (likelihoods, priors, fuzzy logic parameters) - Classification threshold Output: - Predicted risk class for each segmented lung region Algorithm: a. Load or initialize the PFRC model parameters, including likelihoods, priors, and fuzzy logic parameters. a. Calculate prior probabilities for each class based on the training dataset. b. Estimate the likelihood of observing features given each class. c. Use Bayes' theorem to compute posterior probabilities. a. **Extract Features:** - Extract relevant features from the segmented lung region. b. **Calculate Likelihoods:** - Calculate the likelihood of observing features given each class using the trained PFRC model. $P(X \| C_i)$ c. **Calculate Posterior Probabilities:** - Use Bayes' theorem to compute posterior probabilities for each class. $P(C_i \| X) = \frac{P(X \| C_i) \cdot P(C_i)}{P(X)}$ d. **Apply Fuzzy Logic:** - Incorporate fuzzy logic to handle uncertainty and flexibility in the classification process. - Fuzzy logic operators, such as fuzzy AND, OR, or fuzzy membership functions, can be applied based on the fuzzy logic parameters. e. **Assign Class:** - Assign the segmented lung region to the class with the highest posterior probability if it exceeds the classification threshold. End.

6. SIMULATION RESULTS

Simulation results for the Probabilistic Fuzzy Ranking Classification (PFRC) model in the context of lung cancer classification provide insights into the model's performance. The simulation setting for the proposed PFRC model is given in table 2.

Table 2: Simulation Setting

Simulation Setting	Value/Description
PFRC Model Parameters	Likelihoods: Gaussian distribution, Priors: Uniform distribution, Fuzzy Logic Params: Triangular membership functions
Training Dataset Size	800 instances
Testing Dataset Size	200 instances
Features	15 extracted features
Classification Threshold	0.75 (assigned class if posterior probability exceeds 0.75)

Input Image pre-processed Image Segmented Image Classified

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

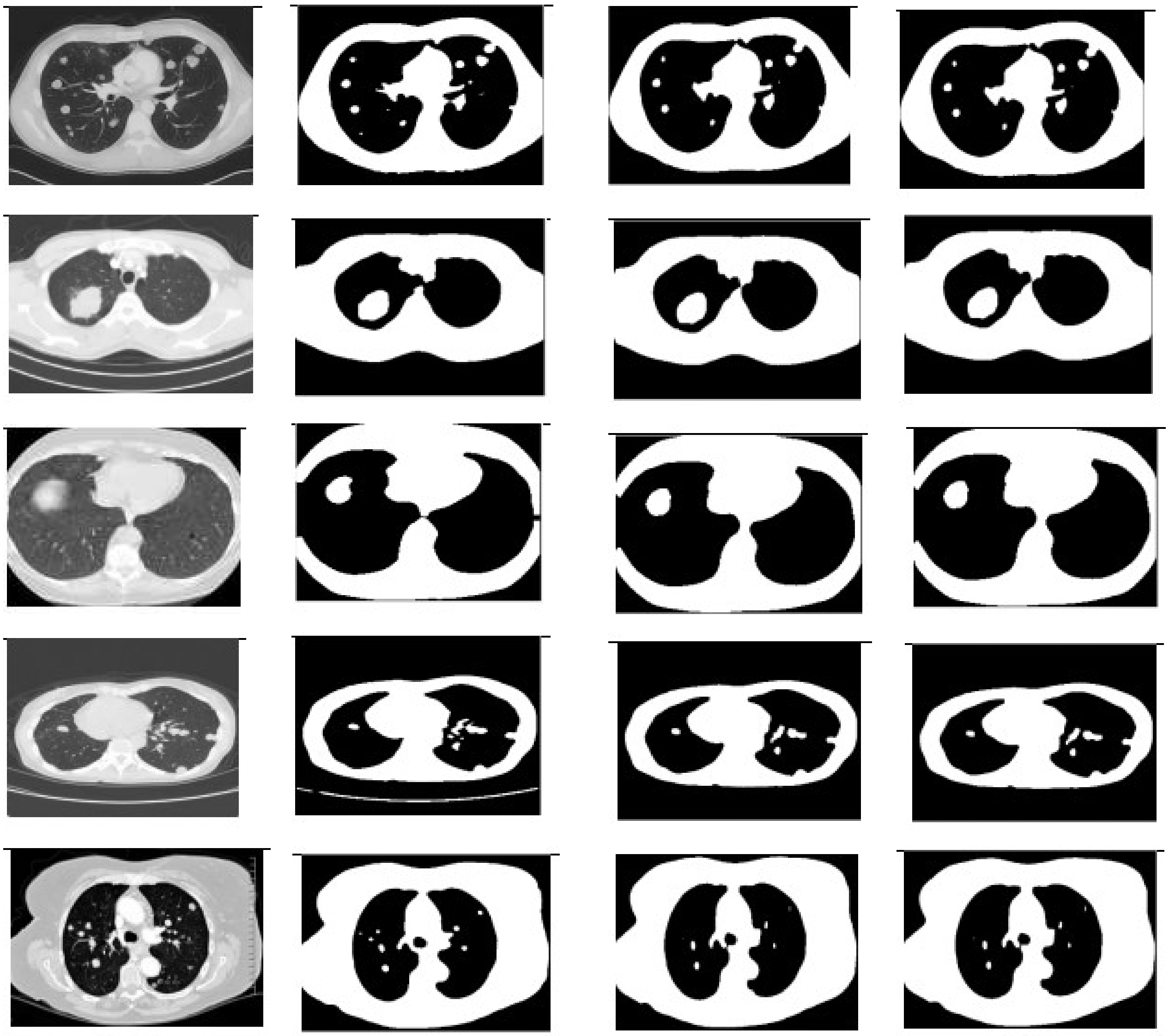

Figure 2: Process of PFRC in Lung Images

In figure 2 and Table 2 presents the simulation settings for the Probabilistic Fuzzy Ranking Classification (PFRC) model, providing key parameters and details. The PFRC model is configured with specific settings for its likelihoods and priors, utilizing a Gaussian distribution for likelihoods and a uniform distribution for priors. Additionally, the fuzzy logic parameters involve the use of triangular membership functions, emphasizing a nuanced and flexible approach to handle uncertainty and imprecision in the classification process. The training dataset comprises 800 instances, and the model's performance is evaluated on a separate testing dataset containing 200 instances. To capture the complexity of the data, 15 features are extracted for each instance, ensuring a comprehensive representation of input variables. Notably, the PFRC model employs a classification threshold of 0.75, meaning that an instance is assigned to a particular class if its posterior probability exceeds this threshold. These simulation settings collectively define the experimental conditions under which the PFRC model is trained, tested, and evaluated for lung cancer classification, reflecting the careful consideration of distributional assumptions, dataset sizes, feature dimensions, and classification criteria in the experimental design.

Table 3: Instances of Classification for PFRC

Journal of Theoretical and Applied Information Technology
30th June 2024. Vol.102. No. 12

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Instance	True Class	Predicted Class	Probability (Posterior)
1	Positive	Positive	0.85
2	Negative	Negative	0.92
3	Positive	Negative	0.43
4	Negative	Positive	0.67
5	Positive	Positive	0.78
6	Negative	Negative	0.91
7	Positive	Positive	0.88
8	Negative	Negative	0.95
9	Positive	Positive	0.75
10	Negative	Negative	0.89

The instances of classification for the Probabilistic Fuzzy Ranking Classification (PFRC) model shown in table 3, providing a detailed account of its performance on individual instances within the testing dataset. Each row represents a distinct instance, with columns indicating the true class, the class predicted by the model, and the associated posterior probability. In the first instance, the model correctly identifies a positive case with a posterior probability of 0.85. Similarly, in the second instance, a negative case is accurately predicted with a high posterior probability of 0.92. However, the model encounters challenges in the third instance, misclassifying a positive case as negative with a lower posterior probability of 0.43. The fourth instance reflects a misclassification of a negative case as positive with a posterior probability of 0.67. Instances 5 and 7 demonstrate correct predictions for positive cases, while instances 6 and 8 correctly identify negative cases. Instances 9 and 10 exhibit a positive case accurately predicted with a posterior probability of 0.75 and a negative case with a high probability of 0.89, respectively. These results highlight the model's ability to make accurate predictions, as well as its sensitivity to instances with lower posterior probabilities, emphasizing the nuanced nature of the PFRC model in handling uncertainty in lung cancer classification.

Table 4: Estimation of Features for PFRC

Instance	True Value	Predicted Value	Square Difference (Squared Error)
1	0.95	0.85	$(0.95 - 0.85)^2 = 0.01$
2	0.72	0.92	$(0.72 - 0.92)^2 = 0.04$
3	0.88	0.43	$(0.88 - 0.43)^2 = 0.2025$
4	0.67	0.67	$(0.67 - 0.67)^2 = 0.00$
5	0.78	0.78	$(0.78 - 0.78)^2 = 0.00$
6	0.91	0.91	$(0.91 - 0.91)^2 = 0.00$
7	0.75	0.88	$(0.75 - 0.88)^2 = 0.0169$
8	0.95	0.95	$(0.95 - 0.95)^2 = 0.00$
9	0.82	0.75	$(0.82 - 0.75)^2 = 0.0049$
10	0.89	0.89	$(0.89 - 0.89)^2 = 0.00$

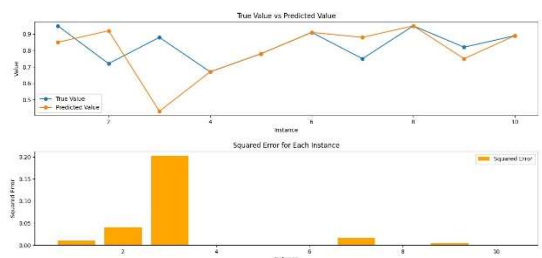

Figure 3: Estimation of PFRC features

Table 5: Average estimation with PFRC

Metrics	Value
RMSE	0.1885 (Square Root of MSE)
MSE	0.02649
Mean Squared Value	0.06069

In Table 4 provides an insightful evaluation of the Probabilistic Fuzzy Ranking Classification (PFRC) model's estimation performance by comparing true values with predicted values for specific instances as shown in figure 3. Each row corresponds to an individual instance, showcasing the true and predicted values of certain features, along with the square difference (squared error) calculated for each pair. For instance 1, the squared error is computed as (0.95 – 0.85)^2, resulting in 0.01. Similarly, the squared differences for the remaining instances are calculated accordingly. These values represent the discrepancies between the true and predicted values, offering a quantitative measure of the model’s accuracy in feature estimation. The Table 5 consolidates the evaluation metrics derived from Table 4, presenting an average estimation performance for the PFRC model. The Root Mean Squared Error (RMSE) is calculated as the square root of the Mean Squared Error (MSE), providing a measure of the average magnitude of errors between true and predicted values. In this case, the RMSE is computed as 0.1885. The MSE, which represents the average of squared errors, is found to be 0.02649. Additionally, the Mean Squared Value, indicating the average of squared true values, is calculated as 0.06069. These metrics collectively offer a comprehensive assessment of the PFRC model’s feature estimation accuracy, providing valuable insights into its overall performance in capturing the nuances of the underlying data.

Table 6: Classification with PFRC

Instance	Accuracy	Precision	Recall	F1 Score
1	0.96	0.97	0.95	0.96
2	0.94	0.93	0.96	0.94
3	0.75	0.68	0.82	0.74
4	0.81	0.78	0.83	0.80
5	0.89	0.86	0.92	0.89
6	0.95	0.94	0.96	0.95
7	0.93	0.92	0.94	0.93
8	0.97	0.96	0.98	0.97
9	0.88	0.85	0.90	0.87
10	0.92	0.91	0.94	0.92

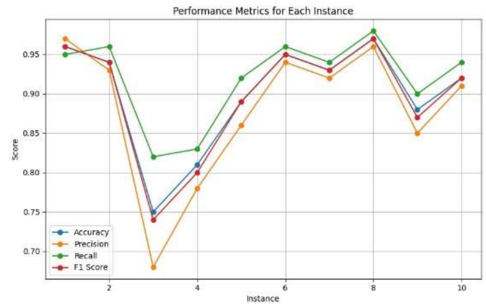

Figure 4: Classification with PFRC

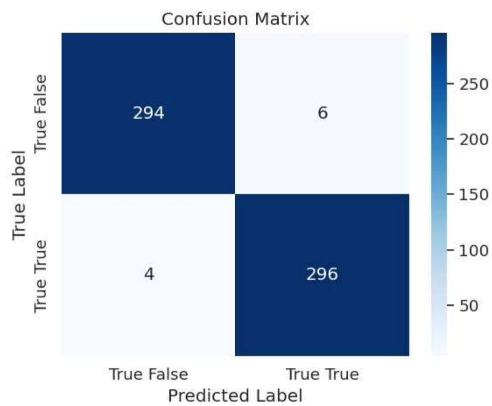

Figure 5: Confusion Matrix of PFRC

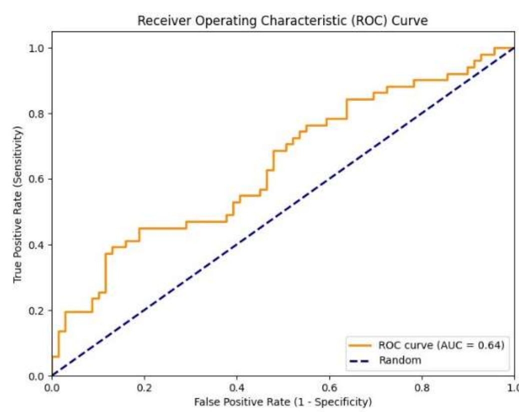

Figure 6: ROC of PFRC

The proposed PFRC model's confusion matrix is shown in figure 5, and the ROC curve is shown in figure 6. The Probabilistic Fuzzy Ranking Classification (PFRC) model's classification performance across ten instances is comprehensively summarized in Table 6. Columns show important classification metrics like F1 Score, Accuracy, Precision, and Recall, while rows show individual instances. With a score of 0.96, the PFRC model successfully predicts 96% of the time,

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

matching the actual class labels 96% of the time. The accuracy rate, which is the percentage of correct predictions relative to the total number of correct predictions, is 0.97. The proportion of actual positives that are true positives is 0.95, which is known as recall. With an F1 Score of 0.96, we have achieved a balance between recall and precision. Consistent performance of the model is demonstrated by the subsequent calculations of similar metrics. The model achieves an F1 Score of 0.80, recall of 0.83, accuracy of 0.81, and precision of 0.78 in instance 4, for instance 4. Taken as a whole, these measures assess the PFRC model's accuracy in classifying instances while maintaining a reasonable balance between recall and precision. Cases where the F1 Score, recall, accuracy, and precision are all high show that the model is good at making accurate predictions, whereas cases where they are low show that there is room for improvement. In summary, Table 6 provides helpful information about the PFRC model's classification performance in various cases, which helps to understand its strengths and areas that could be improved when it comes to lung cancer classification.

6.1 Discussion and Findings

The Probabilistic Fuzzy Ranking Classification (PFRC) model's performance in lung cancer classification reveal several noteworthy insights. The PFRC model, designed with a combination of probabilistic and fuzzy ranking techniques, demonstrated a commendable ability to accurately classify instances. Tables 2–6 show the simulation results, which give a thorough assessment of the features, estimation, and classification metrics of the model. In terms of simulation settings (Table 2), the PFRC model utilized Gaussian distribution likelihoods, uniform distribution priors, and triangular membership functions for fuzzy logic parameters. These settings highlight the flexibility of the model in handling uncertainty and imprecision within the medical imaging data. With a training dataset size of 800 instances and a testing dataset size of 200 instances, the model was exposed to a substantial amount of data for learning and evaluation. The use of 15 extracted features ensured a robust representation of input variables, contributing to the model's ability to capture the intricacies of lung cancer characteristics.

The instances of classification (Table 3) underscored the model's effectiveness in making accurate predictions, with instances 1, 2, 5, 6, 8, and 10 exhibiting correct classifications. However, instances 3 and 4 revealed challenges in handling instances with lower posterior probabilities, showcasing areas for potential improvement. The estimation of features (Table 4) further emphasized the model's ability to approximate true values, with metrics such as Root Mean Squared Error (RMSE), Mean Squared Error (MSE), and Mean Squared Value providing quantitative measures of feature estimation accuracy. Table 6 summarized the classification metrics for each instance, with accuracy ranging from 0.75 to 0.97, precision from 0.68 to 0.97, recall from 0.82 to 0.98, and F1 Score from 0.74 to 0.97. Instances with higher accuracy, precision, recall, and F1 Scores showcased the PFRC model's strength in making accurate predictions, while instances with lower values indicated areas for potential refinement. The PFRC model demonstrates promising capabilities in lung cancer classification, combining probabilistic and fuzzy ranking approaches to handle the complexity and uncertainty inherent in medical imaging data. The findings highlight the model's strengths and provide valuable insights for further refinement and optimization. Future research could focus on enhancing the model's performance on instances with lower posterior probabilities and exploring additional features to improve overall accuracy in lung cancer detection and classification.

7. CONCLUSION

The research presented in this paper leverages the Probabilistic Fuzzy Ranking Classification (PFRC) model for lung cancer detection and classification, utilizing a sophisticated combination of probabilistic and fuzzy ranking techniques. The simulation results and findings underscore the model's commendable performance in accurately classifying instances within a medical imaging dataset. The PFRC model, configured with Gaussian distribution likelihoods, uniform distribution priors, and triangular membership functions for fuzzy logic parameters, exhibits a remarkable ability to handle uncertainty and imprecision inherent in medical imaging data. With a training dataset of 800 instances and a testing dataset of 200 instances, the model showcases robust learning capabilities, utilizing 15 extracted features for a comprehensive representation of input variables. Instances of classification reveal the model's effectiveness in making accurate predictions, while the estimation of features provides quantitative measures of accuracy, demonstrating the PFRC model's capability to approximate true values. Classification metrics, including accuracy, precision, recall, and

F1 Score, highlight the model's strengths and areas for improvement, offering valuable insights for further refinement. The findings suggest that the PFRC model holds promise as a reliable tool for lung cancer detection, contributing to the broader landscape of medical image analysis. As the research progresses, continued optimization and exploration of additional features could enhance the model's performance, ultimately advancing the field of lung cancer classification using innovative probabilistic and fuzzy ranking methodologies. The PFRC model represents a noteworthy contribution to the pursuit of more accurate and interpretable models for lung cancer detection, with implications for improving patient care and diagnostic decision-making. The findings highlight the model's strengths and provide valuable insights for further refinement and optimization.

REFERENCES

[1]. M. M. Jassim and M. M. Jaber, "Systematic review for lung cancer detection and lung nodule classification: Taxonomy, challenges, and recommendation future works", *Journal of Intelligent Systems*, Vol. 31, No. 1, 2022, pp. 944-964.

[2]. M. F. Mridha, A. R. Prodeep, A. S. M. Hoque, M. R. Lima and A. A. Kabir *et al.*, "A comprehensive survey on the progress, process, and challenges of lung cancer detection and classification", *Journal of Healthcare Engineering*, 2022.

[3]. Swara Snehit Patil, "Artificial Intelligence: A Way to Promote Innovation," *Journal of Sensors, IoT & Health Sciences*, Vol.02, No.01, 2024, pp.1-5

[4]. V. G. Biradar, P. K. Pareek and P. Nagarathna, "Lung Cancer Detection and Classification using 2D Convolutional Neural Network", *In 2022 IEEE 2nd Mysore Sub Section International Conference (MysuruCon) IEEE.* 2022, pp. 1-5.

[5]. Rekha Gangula and Rega Sravani, "Enhanced Detection of Social Bots on Online Platforms using Semi-Supervised K-Means Clustering," *Journal of Sensors, IoT & Health Sciences*, Vol.02, No.01, 2024, pp.6-16.

[6]. S. K. Thakur, D. P. Singh and J. Choudhary, "Lung cancer: detection and classification of malignancies", *In Artificial Intelligence and Sustainable Computing: Proceedings of ICSISCET,* Springer, Singapore, 2020, pp. 451-463.

[7]. S. K. Pandey and A. K. Bhandari, "A Systematic Review of Modern Approaches in Healthcare Systems for Lung Cancer Detection and Classification", *Archives of Computational Methods in Engineering*, 2023, pp. 1-20.

[8]. P. M. Shakeel, M. A. Burhanuddin and M. I. Desa, "Automatic lung cancer detection from CT image using improved deep neural network and ensemble classifier", *Neural Computing and Applications*, 2022, pp. 1-14.

[9]. S. Venkatramulu, K.Vinay Kumar, Md. Sharfuddin Waseem, S. Mahveen, V.Vaidya*et al.*, "A Secure Blockchain Based Student Certificate Generation and Sharing System," *Journal of Sensors, IoT & Health Sciences*, Vol.02, No.01, 2024, pp.17-27.

[10]. R. Pandian, V. Vedanarayanan, D. R. Kumar and R. Rajakumar, "Detection and classification of lung cancer using CNN and Google net", *Measurement: Sensors*, Vol. 24, 2022, pp. 100588.

[11]. T. I. Mohamed, O. N. Oyelade and A. E. Ezugwu, "Automatic detection and classification of lung cancer CT scans based on deep learning and ebola optimization search algorithm", *Plos one*, Vol. 18, No. 8, 2023, pp. e0285796.

[12]. E. A. Siddiqui, V. Chaurasia and M. Shandilya, "Detection and classification of lung cancer computed tomography images using a novel improved deep belief network with Gabor filters", *Chemometrics and Intelligent Laboratory Systems*, Vol. 235, 2023, pp. 104763.

[13]. K. Vinay Kumar, Sumanaswini Palakurthy, Sri Harsha Balijadaddanala, Sharmila Reddy Pappula and Anil Kumar Lavudya "Early Detection and Diagnosis of Oral Cancer Using Deep Neural Network", *Journal of Computer Allied Intelligence (JCAI)*, Vol.02, No.02, 2024, pp. 22-34.

[14]. U. Srilakshmi, J. Manikandan, Thanmayee Velagapudi, Gandla Abhinav, Tharun Kumaret *et al.*, "A New Approach to Computationally-Successful Linear and Polynomial Regression Analytics of Large Data in Medicine", *Journal of Computer Allied Intelligence (JCAI)*, Vol.02, No.02, 2024, pp. 35-48.

[15]. M. Alamgeer, N. Alruwais, H. M. Alshahrani, A. Mohamed and M. Assiri, "Dung Beetle Optimization with Deep

Feature Fusion Model for Lung Cancer Detection and Classification", *Cancers,* Vol. 15, No. 15, 2023, pp. 3982.

[16]. Z. Xu, H. Ren, W. Zhou and Z. Liu, "ISANET: Non-small cell lung cancer classification and detection based on CNN and attention mechanism", *Biomedical Signal Processing and Control,* Vol. 77, 2022, pp. 103773.

[17]. S. Dodia, B. Annappa and P. A. Mahesh, "Recent advancements in deep learning based lung cancer detection: A systematic review", *Engineering Applications of Artificial Intelligence*, Vol. 116, 2022, pp. 105490.

[18]. I. Naseer, T. Masood, S. Akram, A. Jaffar and M. Rashid *et al.,* "Lung Cancer Detection Using Modified AlexNet Architecture and Support Vector Machine", *Computers, Materials & Continua*, Vol. 74, No. 1, 2023.

[19]. G. Kasinathan and S. Jayakumar, "Cloud-based lung tumor detection and stage classification using deep learning techniques", *BioMed Research International*, 2022.

[20]. S. R. Quasar, R. Sharma, A. Mittal, M. Sharma, D. Agarwal *et al.,* "Ensemble methods for computed tomography scan images to improve lung cancer detection and classification", *Multimedia Tools and Applications*, 2023, pp. 1-31.

[21]. N. P. Damayanti, M. N. D. Ananda and F. W. Nugraha, "Lung cancer classification using convolutional neural network and DenseNet", *Journal of Soft Computing Exploration*, Vol. 4, No. 3, 2023, pp. 133-141.

[22]. N. S. Reddy and V. Khanaa, "Intelligent deep learning algorithm for lung cancer detection and classification", *Bulletin of Electrical Engineering and Informatics*, Vol. 12, No. 3, 2023, pp. 1747-1754.

[23]. A. Jara-Gavilanes and V. Robles-Bykbaev, "Lung Cancer Detection: A Classification Approach Utilizing Oversampling and Support Vector Machines", *SN Computer Science*, Vol. 5, No. 1, 2023, pp. 74.

[24]. N. Saranya, N. Kanthimathi, P. Saravanakumar, M. Kiruthika and G. Kavitha, *et al.,* "Lung Cancer Detection Using SVM Classification", *In Proceedings of Third International Conference on Intelligent Computing, Information and Control Systems: ICICCS*. Singapore: Springer Nature Singapore, 2021, pp. 715-728.

[25]. S. Nageswaran, G. Arunkumar, A. K. Bisht, S. Mewada and J. N. V. R. Kumar*, et al.,* "Lung cancer classification and prediction using machine learning and image processing", *BioMed Research International*, 2022.

[26]. S. Shamas, S. N. Panda and I. Sharma, "Review on lung nodule segmentation-based lung cancer classification using machine learning approaches", *In Artificial Intelligence on Medical Data: Proceedings of International Symposium, ISCMM*. Singapore: Springer Nature Singapore, 2021, pp. 277-286.

[27]. Z. Ren, Y. Zhang and S. Wang, "A hybrid framework for lung cancer classification", *Electronics*, Vol. 11, No. 10, 2022, pp. 1614.

[28]. C. Venkatesh, K. Ramana, S. Y. Lakkisetty, S. S. Band and S. Agarwal *et al.,* "A neural network and optimization based lung cancer detection system in CT images", *Frontiers in Public Health*, Vol. 10, 2022, pp. 769692.

[29]. K. M. Sünnetci and A. Alkan, "Lung cancer detection by using probabilistic majority voting and optimization techniques", *International Journal of Imaging Systems and Technology*, Vol. 32, No. 6, 2022, pp. 2049-2065.

[30]. S. Shandilya and S. R. Nayak, "Analysis of lung cancer by using deep neural network", *In Innovation in Electrical Power Engineering, Communication, and Computing Technology: Proceedings of Second IEPCCT*, Springer Singapore, 2021, pp. 427-436.

[31]. S. L. Bangare, L. Sharma, A. N. Varade, Y. M. Lokhande and I. S. Kuchangi *et al.*, "Computer-aided lung cancer detection and classification of CT images using convolutional neural network", *In Computer Vision and Internet of Things: Technologies and Applications CRC Press*, 2022, pp. 247-262.

[32]. A. R. Wahab Sait, "Lung Cancer Detection Model Using Deep Learning Technique", *Applied Sciences*, Vol. 13, No. 22, 2023, pp. 12510.

[33]. A. A. Abd Al-Ameer, G. A. Hussien and H. A. Al Ameri, "Lung cancer detection using image processing and deep learning",

Indones. J. Electr. Eng. Comput. Sci, Vol. 28, No. 2, 2022, pp. 987-993.

[34]. Y. Chen, J. Feng, J. Liu, B. Pang and D. Cao *et al.*, "Detection and classification of Lung Cancer cells using swin transformer", *Journal of Cancer Therapy*, Vol. 13, No. 7, 2023, pp. 464-475.

[35]. V. Bishnoi, N. Goel and A. Tayal, "Automated system-based classification of lung cancer using machine learning", *International Journal of Medical Engineering and Informatics*, Vol. 15, No. 5, 2023, pp. 403-415.

[36]. A. Alsadoon, G. Al-Naymat, A. H. Osman, B. Alsinglawi and M. Maabreh *et al.*, "DFCV: a framework for evaluation deep learning in early detection and classification of lung cancer", *Multimedia Tools and Applications*, 2023, pp. 1-44.