Pulmonary Fibrosis Prognosis Prediction using Quantum Machine Learning

Dr. Drakshaveni G
MTech(CST) , Ph.D.
Information Science
BMS Institute of Technology and Management
Bangalore, INDIA
drakshavenig@bmsit.in

Rathna Rohan T R
Information Science
BMS Institute of Technology and Management
Bangalore, INDIA
1by20is179@bmsit.in

Sanath Sunil Ingalgi
Information Science
BMS Institute of Technology and Management
Bangalore, INDIA
1by20is143@bmsit.in

Shweta Jayakumar
Information Science
BMS Institute of Technology and Management
Bangalore, INDIA
1by20is159@bmsit.in

Samarth R Upadhya
Information Science
BMS Institute of Technology and Management
Bangalore, INDIA
1by20is141@bmsit.in

***Abstract*—Lung fibrosis, characterized by the formation of scar tissue in the lungs, results in the deterioration of pulmonary function, impeding the oxygenation of blood. Despite the availability of CT scans, contemporary treatment methods exhibit limited efficacy. The accurate prognosis of pulmonary fibrosis is crucial for facilitating improved clinical trials and enhancing treatment strategies. This paper introduces a novel approach to predict the progression of pulmonary fibrosis, leveraging tabular patient data to estimate the volume of inhaled and exhaled air. The proposed method employs an ensemble of four machine learning algorithms to enhance prognostic accuracy for individual patients. Additionally, this solution offers forecast accuracy, a valuable metric in medical applications for evaluating the model's confidence in its predictions. The modeling of the proposed method demonstrates superior results compared to other forecasting techniques analyzed in the article.**

Keywords—Pulmonary fibrosis progression prognosis, Quantum Machine learning, Computer-aided diagnostics.

I. INTRODUCTION

Pulmonary fibrosis, is a progressive lung ailment resulting from lung tissue or cavity damage and scarring. In some instances, it is referred to as idiopathic PF (IPF) when the cause is unknown (Devaraj, 2014). The fibrous transformation of the pulmonary tissue due to scarring impedes the exchange of carbon dioxideand oxygen gases in the alveoli, the tiny air sacs at the end of airflow branches (U.S. Department of Health and Human Services, 2011). Consequently, the body is deprived of the oxygen necessary for blood oxygenation, leading to reduced lung volume (American Lung Association, 2022).

The substantial loss of lung capacity results in a significant decrease in the amount of air the patient can respire, leading to persistent dyspnea or shortness of breath (Pulmonary Fibrosis Foundation, 2022). Current medical practices acknowledge that the scarring of lung tissue is not completely reversible, necessitating symptom management through therapy and clinical drug trials (Mayo Foundation for Medical Education and Research, 2021). Pulmonary fibrosis has gained attention as one of the most prevalent and life-threatening forms of idiopathic interstitial lung disease, with a median survival rate of merely three years (Das and Chakraborty, 2015).

According to the Pulmonary Fibrosis Foundation, approximately one in every two hundred adults over 70 years of age may be affected by IPF in the United States alone, with over 250,000 cases diagnosed and 50,000 reported annually (Pulmonary Fibrosis Foundation, 2022). Alarmingly, about 40,000 individuals succumb to PF/IPF in the United States alone (Schwartz, 2018).

II. LITERATURE SURVEY

A. CoSANet using Convolutional Self Attention Network:

The Fibro-CoSANet model claims superior performance over CNNs and LSTMs in detecting pulmonary fibrosis. However, recent research by Furukawa et al. (2022) suggests that CNNs and transfer learning can yield effective results in pulmonary fibrosis detection. To assess methodology superiority, a thorough analysis of strengths and weaknesses is crucial. The use of the LIDC-IDRI dataset in Fibro-CoSANet training warrants scrutiny, as it consists of lung nodule CT images and may not be representative of pulmonary fibrosis cases, as emphasized by Zhou et al. (2022), who highlight the importance of a diverse and balanced dataset in medical image analysis.

B. Prediction Analysis of IPF Progression from OSIC Dataset:

Mandal et al. (2020) utilized the OSIC dataset to predict IPF progression with machine learning but omitted convolutional neural networks (CNNs) and attention processes. Furukawa et al. (2022) highlight the success of CNNs in detecting lung lesions and predicting disease progression in pulmonary fibrosis. Zhang et al. (2021) demonstrated high accuracy in classifying medical images using an ensemble of deep CNNs, showcasing the potential of CNNs in various medical image classification tasks, including lung disease detection, assupported by a meta-analysis conducted by Liu et al. (2020) comparing the performance of deep learning algorithms to healthcare professionals.

C. Fibrosis-Net: Deep Convolutional Neural Network:

Wong et al.'s (2021) Fibrosis-Net is a specialized deep convolutional neural network designed for predicting pulmonary fibrosis development from chest CT data. In contrast to Guan et al. (2019), who used typical machine learning, Furukawa et al. (2022) emphasized the effectiveness of deep learning, particularly 3D CNNs like those employed in Fibrosis-Net, for pulmonary fibrosis detection. Another study by G Liu et al. (2019a) and Li et al. (2020) further supported the superiority of 3D CNNs, with the former achieving better results in predicting IPF progression and the latter demonstrating high accuracy using a transfer learning approach for pulmonary fibrosis detection from CT images.

D. Quantum Particle Swarm Optimization:

Rathore et al. (2020) utilized Quantum Particle Swarm Optimization (QPSO), a metaheuristic algorithm, in predicting idiopathic pulmonary fibrosis progression from baseline CT images. Kaur et al. (2021) adopted an alternative approach, optimizing gradient boosting model hyperparameters with an Artificial Bee Colony (ABC) algorithm for accurate prediction of IPF development. Yang et al. (2020) employed a deep learning-based strategy using a 3D CNN architecture, achieving good accuracy and specificity in predicting IPF progression from CT images. In contrast, Flaherty et al. (2020) focused on radiometric characteristics from CT images, suggesting their potential in predicting IPF development and identifying candidates for early intervention.

E. Quantum Convolutional Neural Network:

Quanvolutional Neural Networks (QCNNs) leverage quantum mechanics concepts to enhance performance in machine learning applications, particularly in image classification tasks and quantum chemical applications (Houssein et al., 2022). Despite their potential advantages, a major challenge lies in the scarcity of quantum hardware, as current quantum computers have limited qubits, hindering the training and deployment of QCNNs in real-world scenarios (Romero, Olson, and Aspuru-Guzik, 2017). Overcoming this obstacle requires future research to focus on scalable quantum hardware development, improving interpretability, and comparing QCNN performance across diverse applications against classical neural networks.

III. OVERVIEW OF THE TECHNOLOGY

A. Machine Learning

Machine Learning, within the context of pulmonary fibrosis, constitutes a subset of Artificial Intelligence dedicated to crafting algorithms and statistical models that empower computer systems to learn and enhance their capabilities through experience, all without explicit programming.

In the realm of pulmonary fibrosis, Machine Learning finds application in:

Disease Progression Prediction: Utilizing patient data to forecast the progression of pulmonary fibrosis and assess its severity over time.

Treatment Response Evaluation: Analyzing medical interventions and patient responses to tailor effective treatment plans.

Prognosis Assessment: Predicting the future outcomes of pulmonary fibrosis cases based on individual patient characteristics and historical data.

B. Quantum Machine Learning

With the advance in quantum computing, Applying machine learning techniques to solve complex problems by developing native algorithms to optimize the use of quantum computers leverages the quantum computational capabilities such as quantum entanglement and superposition to achieve faster training speeds and inference in machine learning models by simply processing vast amounts of data at the same time

Application of quantum machine learning:

Complex Data Processing: Quantum machine learning can handle complex datasets more efficiently than classical machine learning algorithms.

Optimization of Treatment Plans: By considering a multitude of variables simultaneously, these algorithms could help tailor personalized treatment strategies that maximize effectiveness while minimizing side effects.

Prognosis Prediction: By leveraging quantum algorithms, it could better model the complex interactions of various factors influencing disease progression, leading to more precise and reliable prognostic assessments.

C. Quantum Computing

With the rapid emergence in our understanding in quantum physics and mechanics, scientists actively thrive to harness the physical capabilities and aspects of quantum mechanics to enhance and further develop our boundaries of communication and information processing

Applications of Quantum Computing:

Genomic Analysis: Quantum computing can accelerate the analysis of vast genomic datasets associated with pulmonary fibrosis, identifying genetic factors and mutations contributing to the disease.

Simulating Biological Systems: simulate biological systems with greater accuracy than classical computers.

IV. RESEARCH METHODOLOGY

In shaping the research philosophy, a pragmatic stance has been adopted, chosen for its versatility in integrating both quantitative and qualitative data to explore the intricacies of reality. The Deductive research approach guides this study, prioritizing the testing and substantiation of hypotheses withina quantitative framework, thereby opting for a deductive strategy over the inductive alternative.

The research strategy pivots around the primary use of surveysfor quantitative data collection, supplemented by interviews with medical professionals to infuse a qualitative dimension. The Mixed method stands as the preferred research choice, harmonizing surveys and interviews to comprehensively validate and fortify the research hypotheses. A Longitudinal time horizon governs the research, ensuring a temporal continuum in data collection before and after the system's development, affording a holistic understanding of the project's enduring impact.

This cohesive methodology, stemming from a pragmatic philosophy, employs a Deductive approach, strategically combining surveys and interviews within a Mixed method framework, all unfolding across a Longitudinal time horizon, thus ensuring a comprehensive exploration and substantiation of the research hypotheses.

V. ALGORITHM SELECTION

Quantum Particle Swarm Optimization:

Quantum Particle Swarm Optimization (QPSO) is an optimization technique that blends quantum mechanics with swarm intelligence principles. Initially proposed in 2006, QPSO as an extension of classical Particle Swarm Optimization (PSO) for application in quantum computing (Li, Zhan and Zhang, 2022). To find the best answer in a high-dimensional search space, QPSO combines the ideas of quantum mechanics with swarm intelligence. One of QPSO's key benefits is its capacity for global search, which enables it to look for a problem's global optimum rather than becoming bogged down in local optima(Qi et al., 2020).

Quanvolutional Neural Networks:

Convolutional neural networks (CNNs) and quantum mechanics are two different types of neural networks that work together to create Quanvolutional neural networks (QCNNs). They use the characteristics of quantum mechanics to improve CNN performance in tasks like image identification and classification. They are designed to function on quantum data, such as quantum pictures or quantum states (Acampora and Schiattarella, 2021). A high-level description of how QCNNs work is that they execute a number of quantum operations on the input quantum state or picture, then use classical convolutional operations to extract features from the output quantum state. The three primary parts of the QCNN architecture are the quantum layer, the classical layer, and the pooling layer. The Hadamard and phase gates, which enable the encoding of quantum information into the state, are examples of quantum operations the quantum layer conducts on theinput quantum state or picture. The resultant quantum state is then processed by the classical layer using traditional convolutional techniques to extract features. In order to decrease the dimensionality of the feature maps, the pooling layer down samples the data. The efficiency with which QCNNs can analyze and categorize quantum data is one of its main features.

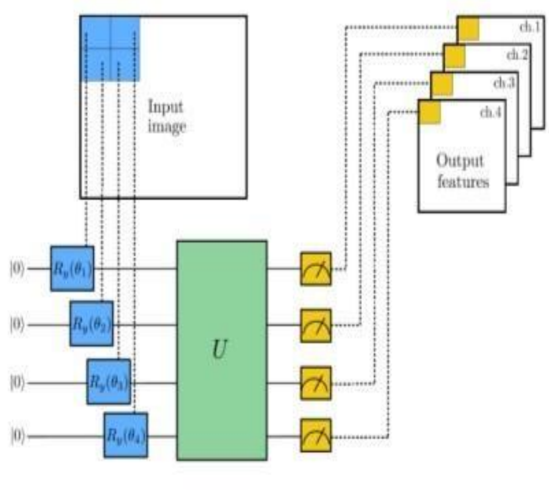

Figure 2.7: Typical Architecture of a Quanvolutional Neural Network

Quantum Convolutional Neural Networks :

quantum neural networks, a particular sort of neural network, utilize quantum computing concepts. QNNs process information using the principles of quantum physics, in contrast to traditional neural networks, which depend on classical computing concepts. One of the main distinctions between quantum neural networks and quantum neural networks is that the former is more general-purpose,whilst the latter are especially made to handle quantum pictures or quantum. In conclusion, quantum neural networks are a promising field of study that apply the ideas of quantum mechanics to machine learning challenges. QNNs have the potential to dramatically enhance the performance of machine learning algorithms in a variety of applications, even though research is still in its early phases.

Quantum Generative Adversarial Networks :

QGANs are a sort of quantum machine learning technique that combines the ideas of generative adversarial networks with quantum computing principles. QGANs are intended to create new quantum data that is comparable to a training set of quantum data and may be used for a range of tasks including quantum data reduction and synthesis. QGANs are made up of two neural networks: a generator network and a discriminator network.

Deep Convolutional Neural Networks:

Deep Quanvolutional Neural Networks are a sort of quantum machine learning technique that combines deep neural network principles with Quanvolutional neural network principles. DQNNs are intended to categorize quantum data and extract quantum features. DQNNs are made up of many layers of Quanvolutional filters that extract quantum characteristics from incoming quantum data. These characteristics are subsequently processed by numerous layers of fully connected neural networks that perform classification tasks. DQNNs, like QNNs, use quantum physics concepts such as superposition and entanglement to conduct certain computations more effectively than classical DNNs.

Hough Algorithm:

The Hough transform is a computer vision algorithm that detects fundamental forms in digital pictures such as lines, circles, and ellipses. The method operates by mapping the picture space into a parameter space and recognizing points or areas in that space that match to the forms being sought. The Hough algorithm has been intensively investigated in the computer vision field, with several modifications and extensions offered over the years. Gisin et al., (2002) for example, presented the Randomized Hough Transform (RHT) as a more efficient variant of the Hough transform that randomly picks points from the image space to minimize computing complexity.

VI. IMPLEMENTATION

The implementation phase of the project involves translating the devised methodology into actionable steps, beginning with the establishment of a robust computing infrastructure compatible with both classical and quantum computations. The diverse datasets are meticulously prepared, undergoing cleaning and preprocessing to ensure data integrity. The variational quantum circuit is designed and integrated into the hybrid- classical quantile regression model, with parameters fine-tuned to optimize its performance in predicting pulmonary fibrosis prognosis. Model training incorporates classical machine learning algorithms alongside the quantum circuit to capture disease nuances. The development of an automated system facilitates seamless prognosis prediction, adaptable to new data for continuous learning.

Data Set : At the heart of Fibro-QuanNet lies a data science initiative, necessitating a dataset encompassing pulmonary fibrosis prognosis predictions for the training and testing of models. Given the medical focus of this project, it is imperative that the data originates from reliable sources to ensure accuracy.

The requisite data includes :
-> Pulmonary cavity HRCT imagery data
-> Patient metadata, incorporating variables like smoking status, age, and FVC.

The necessary data was obtained through a dataset provided by an open-source imaging consortium. This dataset encompasses a pulmonology prognosis prediction dataset, featuring essential information such as HRCT imagery at baseline. Additionally, it includes crucial patient metadata, comprising details like smoking status, age, FVC, and lung deterioration percentage.

Programming Language

Given that Fibro-QuanNet is a machine learning application, Python has been chosen as the primary programming languagefor model development. Python, being a versatile and general- purpose language, extends its utility beyond machine learning to areas such as web development, scientific computing, and automation. Its widespread adoption in machine learning projects is attributed to its user-friendly nature and the availability of robust libraries like NumPy, Pandas, Matplotlib, and TensorFlow. These libraries facilitate seamless datapreprocessing, analysis, model building, and result visualization. Python's frameworks, versatility, and support for parallel and distributed computing further solidify its standing as a preferred language for machine learning endeavors.

Libraries

Penny Lane

PennyLane stands out as a versatile library with the capability to accommodate a diverse array of quantum computation engines. It empowers users to develop and train quantum machine learning models seamlessly. Furthermore, PennyLane facilitates the effortless execution of these models on various quantum platforms, enabling a straightforward comparison of results across different quantum systems.

Cirq

Cirq stands as a resilient Quantum Machine Learning (QML) library that distinguishes itself by enabling users to leverage and experiment with built-in quantum machines for creating and executing quantum models directly on Google's hardware. The notable advantage of utilizing Cirq lies in its user-friendly design, allowing for the straightforward definition of quantum circuits and algorithms through a simple and intuitive syntax.

Qiskit

Qiskit offers an array of tools and libraries designed for various quantum computing tasks, encompassing assistance for quantum algorithms, quantum circuit design, and quantum machine learning.

MONAI

MONAI, standing for Medical Open Network for AI, serves as a deep learning framework specifically tailored for medical imaging and associated tasks. It facilitates the development, training, and evaluation of deep learning models across a spectrum of applications, including image segmentation, classification, and generation. Offering a user-friendly and comprehensive framework, MONAI proves instrumental in the creation and training of deep learning models for diverse medical imaging tasks.

RAPIDS

RAPIDS, which stands for Acceleration of Diagnostics Deployment, is a software library engineered to deliver GPU acceleration for data science tasks. Tailored for high-performance computing, RAPIDS facilitates efficient and rapid data processing and analysis. By harnessing the capabilities of GPUs, this library significantly accelerates data processing, empowering data scientists to enhance their efficiency and focus on deriving meaningful insights from their datasets.

VII. EXISTING SYSTEMS

The Fibro-CoSANet model says it's better at finding lung issues than older methods like CNNs and LSTMs. However, some researchers, like Furukawa et al. (2022), still think CNNs work well. There's a concern that the dataset Fibro-CoSANet used may not be the best for studying lung problems like fibrosis, as pointed out by Zhou et al. (2022).

Mandal et al. (2020) tried to predict lung disease progression using a dataset but didn't use advanced methods like CNNs. In contrast, Furukawa et al. (2022) argue that CNNs and attention processes are essential for studying lung diseases. Zhang et al. (2021) showed that using multiple CNNs together is effective in telling if lungs are healthy or not. A study by Liu et al. (2020) found that computer programs using deep learning are as good as doctors at analyzing medical images and finding diseases.

VIII. CONCLUSION

In conclusion, the research successfully attained its objective through the creation, implementation, and assessment of a quantum-hybrid Quanvolutional neural network. This innovative model demonstrated the ability to precisely forecast the prognosis of pulmonary fibrosis by incorporating HRCT imagery input alongside additional patient meta-data. The incorporation of quantum machine learning techniques further enhanced the accuracy and efficacy of the prognosis predictions.

IX. REFERENCES

[1] Acampora, G. and Schiattarella, R. “Deep neural networks for quantum circuit mapping. Neural Computing and Applications”,2023.

[2] Alamri, A.K. ,“ Pulmonary Hypertension Association’s International Conference Scientific Sessions Overview” 2022.

[3] Al-jumaili S. “Recent Advances on Convolutional Architectures in Medical Applications: Classical or Quantum” , 2022.

[4] Amin, J, “Quantum Machine Learning Architecture for COVID-19 Classification Based on Synthetic Data Generation Using Conditional Adversarial Neural Network. “, December 2022.

[5] Azevedo, V., Silva, C, “Quantum transfer learning for breast cancer detection” ,2022.

[6] Beer, K, “Training deep quantum neural networks. Nature Communications, (11)”, 2021.

[7] Biamonte, J “Exploring Object Relation in Mean Teacher for Cross-Domain Detection.”, 2021.

[8] Das, S. and Chakraborty, K. “Review paper of Idiopathic Pulmonary Fibrosis. International Journal of Advanced Research, 3”, 2020