International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 7, Issue 11, pp 79-86, November-2023

https://doi.org/10.47001/IRJIET/2023.711011

Assistive System to Identify and Manage Lung Cancer

[1]D.A.H.K. Rathnayaka, [2]E.M.V.Y. Ekanayake, [3]K.K. Paranavithana, [4]K.B.A.S.M. Dissanayake, [5]Wishalya Tissera [6]Dasuni Nawinna

[1,2,3,4,5]Department of Information Technology Sri Lanka Institute of Information Technology Malabe, Sri Lanka

[6]Department of Computer Systems Engineering Sri Lanka Institute of Information Technology Malabe, Sri Lanka

Authors E-mail: [1]kavindyahashini60@gmail.com, [2]vishmiekanayake@gmail.com, [3]kalanikavitha09@gmail.com, [4]savinudissanayake112@gmail.com, [5]wishalya.t@sliit.lk, [6]dasuni.n@sliit.lk

***Abstract* - Lung cancer, a leading cause of cancer-related deaths worldwide, necessitates early detection and effective management tools. This research introduces an assistive system leveraging machine learning to identify risk factors and prioritize treatments based on severity. Critical socio-demographic data, like age, gender, and smoking history, enable accurate risk assessment and personalized profiles. The system adapts to lifestyle changes post-diagnosis, offering tailored healthcare solutions. Integrating Mask R-CNN, a deep learning algorithm for medical imaging, enhances lung cancer diagnosis precision and treatment strategies. The user-friendly interface allows healthcare providers, caregivers, and patients easy access via mobile or computer applications. The implementation includes natural language processing for questionnaires and data preprocessing. The system's design consists of three components: front-end interface, back-end server, and Mask R-CNN for tumor identification. Correlation analysis for patient prioritization to generate a list. The proposed system aims to revolutionize lung cancer care, delivering accurate risk assessments, personalized treatment plans, and continuous monitoring, ultimately improving patient outcomes and saving lives.**

Keywords: Lung cancer, CT Scan images, Machine Learning, Random Forest Classifier, Natural Language Processing, tokenization, Mask R CNN, Image processing, Logistic Regression.

I. INTRODUCTION

Cancer is one of the most prevalent and destructive diseases affecting a significant percentage of the world's population. Among the various types of cancer, lung cancer is a leading cause of cancer-related deaths worldwide. Unfortunately, the early stages of this disease often do not display any symptoms, making it difficult to detect at an early stage. A lung cancer diagnosis can also create anxiety and uncertainty, significantly impacting a patient's quality of life. Due to these challenges, it is vital to provide tools and systems that aid in the early detection, management, and treatment of lung cancer. An assistive system that can identify and manage lung cancer risk is crucial in effectively managing this disease.

With the integration of machine learning algorithms, an assistive system can be designed to identify cancer risk factors and make appropriate decisions based on the patient's unique cancer risk status. The system can also generate a treatment priority list for patients based on the severity of their condition. The main benefit of such a system is its ease of use, as it can be accessed by healthcare professionals, caregivers, or even patients themselves through an application on a mobile phone or work on a computer. Additionally, the system can integrate with other medical devices to provide healthcare providers with a comprehensive view of the patient's health status. To accurately assess cancer risk, an assistive system must consider the various factors that contribute to the development of lung cancer. These factors include critical socio-demographic information such as the patient's age, gender, and smoking history, all of which contribute significantly to the risk assessment and staging of the disease. By incorporating this data into the model, the system can deliver accurate and tailored results that match the patient's unique profile. Another important aspect of an assistive system is its ability to display changes in a patient's lifestyle after a cancer diagnosis. An assistive system that takes these changes into account can help to maintain a patient's quality of life even after a cancer diagnosis. By identifying changes in the patient's lifestyle, the system can provide tailored healthcare solutions that cater to the patient's individual needs. The Mask R-CNN (Region-based Convolutional Neural Network) architecture is utilized by the assistive system for the automated detection and localization of lung cancer in medical images. By integrating deep learning algorithms with medical imaging, we can significantly enhance the accuracy and efficiency of lung cancer diagnosis. With its ability to detect even small tumors with high precision, the Mask R-CNN architecture can revolutionize how we approach cancer treatment, allowing for more effective strategies to be implemented. One of the most significant challenges in treating lung cancer is determining the most effective treatment plan for each patient based on the various stages of the disease. By integrating machine learning algorithms into the system, the system can generate a treatment priority list that prioritizes treatment plans according to the patient's condition. This ensures that healthcare providers can deliver

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 7, Issue 11, pp 79-86, November-2023

https://doi.org/10.47001/IRJIET/2023.711011

the most effective treatment possible to the patient, leading to better patient outcomes.

An assistive system that can identify and manage lung cancer risk is essential for the early detection and effective management of the disease. Machine learning algorithms can be integrated into the system to provide accurate and tailored results that match the patient's unique profile. The system can also detect and identify the cancer stage and affected areas, generate a treatment priority list, and display changes in a patient's lifestyle after a cancer diagnosis. With this system in place, healthcare providers can provide effective treatment plans and improve the quality of life for patients with lung cancer, ultimately saving lives.

II. LITERATURE REVIEW

Here, we provide a brief overview of the most recent relevant works for identifying and manage of lung cancer occurrence using ML techniques and models.

By utilizing CT scan pictures, the authors of [9] demonstrated an effective method for the identification and categorization of lung cancer. A decision tree, a random forest, a support vector machine, a naive Bayes model, k-nearest neighbors, stochastic gradient descent, and a multi-layer perceptron were just a few of the seven classification models they used. A dataset of 15,750 clinical data encompassing 6910 benign and 8840 malignant lung cancer-related images was taken into consideration for the training and testing of these classifiers. The multi-layer perceptron classifier outperformed the other classifiers in the collected results, with an accuracy value of 88.55%.

The paper [11] introduces a novel approach to the detection and classification of lung cancer in CT images. It combines the Mask R-CNN algorithm, which is capable of object detection and instance segmentation, with a generated mask method specifically designed for lung cancer detection. The generated masks aim to accurately delineate the boundaries of cancerous regions within the images. The proposed method is evaluated using a dataset of lung cancer CT images and achieves high accuracy in both lung cancer detection and classification.[12]

The study has proposed a novel deep learning-based framework for segmenting lung nodules in CT images. Segmenting lung nodules, tiny, rounded growths on the lung, is an important first step in identifying and treating lung cancer. The state-of-the-art object recognition and segmentation model, Mask R-CNN architecture, is the foundation of the suggested framework.[22]

The authors of the paper improve upon Mask R-CNN in two ways.. The weighted binary focal loss, which concentrates on the challenging-to-segment regions of the nodule boundary, is the first new loss function they introduce. Second, they suggest a brand-new approach for fusing multi-scale characteristics that enhances the model's capacity to recognize and segment nodules of various sizes.

Similar techniques were used by the authors in [10] to predict lung cancer, including a neural network, radial basis function network, support vector machine, logistic regression, random forest, J48, naive Bayes, and K-nearest neighbours. They demonstrated that the radial basis function network had a greater accuracy of 81.25% when applied to data related to lung cancer. The primary goal of [15] is to examine the effectiveness of classification algorithms to make an early diagnosis of lung cancer. The authors used decision trees, logistic regression, support vector machines, naive Bayes, and other classification algorithms. Support vector machine (SVM) achieved a greater accuracy of 99.2% in the lung cancer dataset from the data. the world compared to the lung cancer dataset from UCI, where logistic regression had a higher accuracy of 96.9%.

By integrating well-established metabolomics mechanisms with machine learning techniques, the authors of [13] developed a mechanism to find the right biomarkers for early diagnosis of lung cancer. Their investigation was based on a dataset that included 43 healthy people and 110 patients with lung cancer. Following ROC analysis with an AUC, Sensitivity, and Specificity equal to 0.989, 0.981, and 1, respectively, six biomarkers were chosen to enable the separation of first-stage lung cancer patients from healthy persons. The top 5 relative relevance metabolic biomarkers were identified using the fast correlation-based filter (FCBF). The Nave Bayes model is the one that is recommended for the early diagnosis of lung cancers among the studied models.

The Lalaby app addresses the limitations of current mobile apps and wearable devices used for monitoring lung cancer patients' quality of life (QoL) and symptoms. It offers tailored designs and objective data collection, providing remote data acquisition, symptom management, and correlation with QoL dimensions. This comprehensive solution empowers patients to conveniently track their health progress, manage symptoms, and enhance their well-being during treatment. By bridging patient needs with advanced technology, the app aims to offer a personalized and user-friendly tool, ultimately improving the monitoring and management of lung cancer patients' health throughout their treatment journey. [14]

ISSN (online): 2581-3048
Volume 7, Issue 11, pp 79-86, November-2023
https://doi.org/10.47001/IRJIET/2023.711011

Recent evidence in oncology highlights the link between patients' reported outcomes and cancer-free survival, underscoring the significance of psychological and physical support. Studies on lung cancer surgery patients found that declines of ≥10% in physical and mental components postoperatively increased death risks. Higher physical component scores in pulmonary lobectomy patients correlated with extended overall and cancer-specific survival. In a multicenter study, poor baseline QoL scores didn't predict worse survival for high-risk operable patients over 24 months. VATS showed early improvement in physical functioning and dyspnea scores. While NSCLC surgical treatment advances survival, its impact on daily life warrants investigation. Tailored surgical questionnaires can aid thoracic surgeons in future QoL research, enriching patient care and treatment strategies.[5]

III. METHODOLOGY

This proposed system contains four main components as

- Identify the cancer risk and take the appropriate decision based on the cancer risk status.
- Detect and identify the cancer stage and the affected areas using CT scans.
- Generate the treatment priority of the patient.
- Displaying changes in a patient's life cycle after a cancer diagnosis.

Each component uses different Machine Learning algorithms and techniques to predict and visualize raw datasets.

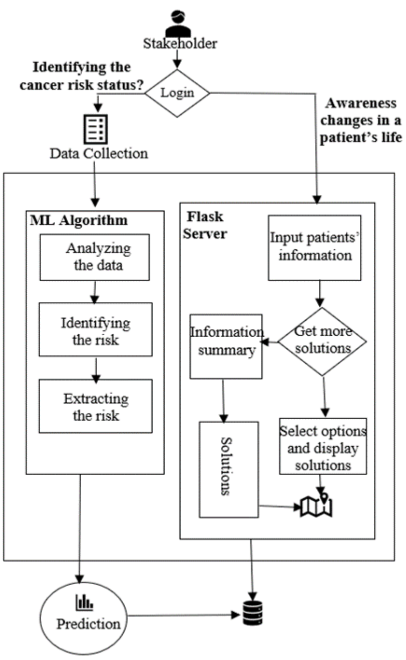

Figure 1: Overall system diagram of the proposed mobile application

When a stakeholder visits the application and answers the questionnaire, the process will start. Otherwise, the user should manually input the data where he/she wants to get the prediction. The existing trained model will analyze the data and it will predict the cancer risk status. This model can output a probability score between 0 and 1, representing the likelihood that a patient has cancer risk or not based on their input data.

If the user is a lung cancer patient, the user can access and review the patient's physical changes, which are regularly updated based on their previous visits, providing valuable insights into their health progress. a dataset was utilized for NLP training, incorporating tokenization and lemmatization techniques to preprocess the text. The tokenized text was then parsed, enabling efficient analysis of questions and answers, and facilitating a deeper understanding of user inquiries. This prioritizes mental health support, addressing psychological aspects of the patient's journey with a focus on comprehensive well-being. Lastly, this diagram provides economic propositions, including suggested locations, accessible to users who seek additional information. This feature enhances user convenience by offering valuable insights into potential healthcare services and resources available in specific locations.

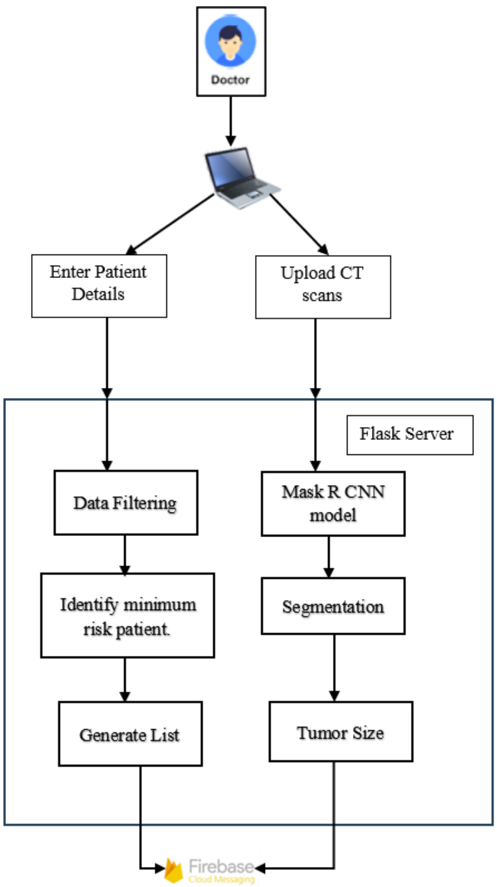

Figure 2: Overall system diagram of the proposed web application

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 7, Issue 11, pp 79-86, November-2023

https://doi.org/10.47001/IRJIET/2023.711011

This web application is designed to facilitate the efficient and user-friendly process of uploading CT scans for lung tumor identification. The system diagram is structured around three main components: the front-end interface, the back-end server, and the deep learning mode which is Mask R-CNN, to enable seamless CT scan upload, lung tumor identification, and precise size measurement. By offering such functionality, the system aims to aid in understanding lung tumors with greater efficiency and accuracy.

The system design comprises a sequential procedure beginning with the entry of patient details into the system. And using a prioritization system for the patients, the data is then filtered before being analyzed for risk levels. The doctor receives this list of patients with the highest priority for decision-making. Healthcare systems can streamline the process of patient prioritization. By utilizing an advanced algorithm correlation analysis, healthcare systems can streamline the process of patient prioritization.

Data collection, processing, and training of the detection models

A) Identify the cancer risk prediction

Machine learning (ML) models were trained using a dataset of information about patients and various factors related to lung cancer. Each row in the dataset represents a different patient and the columns represent different attributes. The gender of the patient can be either M (male) or F (female). With a value of 1 indicating no and 2 indicating yes. From the dataset, 75% of the data is used for training the model and the remaining is used for testing. The lung cancer column and the gender column are converted to numeric values between 0 and 1. Remove duplicates and pre-process the data set. Preprocessing techniques were used to improve accuracy and reduce the complexity of the dataset.

Logistic Regression, SVC, Gaussian naive Bayes, Decision Tree Classifier, and RandomForestClassifier algorithm were used to train the models. Finally, based on the testing accuracies the most appropriate model was selected. RandomForestClassifier was given the best accuracy. the RandomForestClassifier is imported, and an instance is created with 100 estimators specified by the 'n_estimators' parameter. the classifier is fitted to the training data (xtrain and ytrain), which means it learns patterns and relationships in the data. The accuracy of the model on the training data is 99.54% and the accuracy score of the model on the test data (ytest) is 91.07%. The model is used to make predictions on new data. With a set of input features, the predict() function is used to make predictions. The predictions are shown for each instance, indicating the predicted class (0 or 1).

B) Detect and identify the cancer stage and affected areas

This study utilizes the COCO dataset as the primary dataset for evaluating and benchmarking the proposed methodology. The dataset offers comprehensive annotations, diverse object categories, and suitability for various computer vision tasks. The researchers train deep learning models using object bounding boxes, segmentation masks, key points, and captions, enabling precise training and accurate object detection. The COCO dataset also serves as a benchmark for comparing the proposed methodology against existing state-of-the-art approaches. The evaluation metrics, such as Average Precision (AP) and Average Recall (AR), quantitatively measure the performance of the models in tasks like object detection, instance segmentation, or keypoint detection.

Mask R-CNN is a state-of-the-art deep learning model, for Instance segmentation, which combines object detection and pixel-level segmentation into a single architecture. It builds upon the Faster R-CNN framework, which is widely used for object detection tasks. The main innovation of Mask R-CNN is the addition of a mask prediction branch to the Faster R-CNN architecture. This branch enables the model to generate pixel-wise masks for each detected object, providing detailed segmentation information. Mask R-CNN has displayed exceptional performance in a variety of instance segmentation tasks, including the COCO dataset.

This study demonstrates the training of a model, for Instance segmentation using the Detectron2 library. Instance segmentation involves identifying and delineating individual objects within an image. The specific architecture employed in this code is the Mask R-CNN, which combines object detection with pixel-level segmentation. The chosen architecture utilizes a ResNet-101 backbone, which is a deep convolutional neural network known for its strong feature extraction capabilities. A feature pyramid network (FPN) is also incorporated to enhance multi-scale feature representation.

The subsequent code demonstrates the inference process for individual images. It utilizes the glob module to iterate over a specified directory containing test images. Each image is loaded using cv2.imread() and passed to the predictor to obtain the predicted instances.

C) Generate the treatment priority of the patients

This section provides an overview of the research objective, which is to generate lung cancer patients' treatment priority and support clinical decision-making for doctors.

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048
Volume 7, Issue 11, pp 79-86, November-2023
https://doi.org/10.47001/IRJIET/2023.711011

Collecting appropriate datasets that can be used to train and test a machine learning model for generating treatment priorities. The source and composition of the dataset with diagnoses, laboratory tests, medications, and provider notes from the relevant parties. Data gathered for the dataset mainly with the factors considering Age, Gender, and Cancer level, checks for whether it has spread to Lymph nodes inside the lung and whether it is spread to the bone. The dataset was obtained and rearranged the factors from the dataset with proper citations and referrals the data was preprocessed. Handling the missing values and removing the duplicated values from the dataset before training the model.

The use of logistic regression to determine patient treatment priorities. It addresses the complexities of resource allocation in healthcare and advocates for data-driven approaches. The logistic regression model was trained using the features (age, gender, chest pain type, spread to lymph nodes, and spread to bone) to predict the cancer level of each patient. A train-test split of 80:20 was applied to assess the model's performance, with an emphasis on accuracy evaluation. The calculated accuracy was 83.86%, indicating the model's predictive capability. To facilitate the doctor's decision-making process, a doctor's input of minimum risk level and a patient's age were utilized to filter the dataset. The resultant filtered data provided tailored recommendations for patients meeting the specified criteria. The recommendations included patient details such as name, ID, cancer level, age, gender, chest pain type, and the status of cancer spread to lymph nodes and bones.

This approach empowers healthcare professionals to make informed decisions on treatment priority, ensuring that patients with the most critical needs receive immediate attention, optimizing resource utilization, and ultimately enhancing patient outcomes.

D) Displaying changes in a patient's life cycle after a cancer diagnosis

Using Natural Language Processing (NLP) techniques, a dataset of questions and answers related to the changes in the life cycle of a lung cancer patient was curated and used for training purposes. This dataset primarily focused on mental health questions and answers to gain insights into the psychological aspects of a patient's journey with lung cancer. To create this dataset, various sources were explored, including medical literature which provided valuable information about the mental health challenges faced by lung cancer patients.

Additionally, problems concerning the physical health of lung cancer patients and their corresponding appropriate answers were also sourced from the internet. These questions were related to the physical symptoms, side effects of treatments, and other health-related concerns faced by individuals diagnosed with lung cancer.

Here NLTK is used. NLTK is the Natural Language Toolkit, a library for working with human language data in Python. It provides various tools for tasks like tokenization, stemming, lemmatization, part-of-speech tagging, and more. Here tokenization is used. The text was tokenized and parsed to break it down into smaller units for analysis, allowing the system to understand the structure of the questions and answers effectively.

Furthermore, stopwords and special characters were removed to focus on the essential content of the questions and answers. NLTK library's lemmatization reduces words to their base form, aiding analysis by simplifying derived words and enhancing text comprehension. The curated and preprocessed dataset was then used to train the NLP models, such as deep learning models or transformer-based models like BERT. The models were optimized during training to identify patterns, relationships, and context within the questions and corresponding answers. By learning from this dataset, the NLP models gained the capability to generate accurate and informative responses to user queries related to the changes in the life cycle of a lung cancer patient. The primary focus on mental health questions and answers in the dataset allowed the system to address the emotional and psychological challenges patients face, while also providing information related to physical health concerns. This comprehensive approach ensures that the assistive system can offer valuable insights and support to lung cancer patients and their caregivers throughout their treatment journey.

IV. RESULTS AND DISCUSSION

The models for identifying lung cancer risk were trained using several algorithms, and the best model was selected among them by considering the accuracy. The accuracy indicates which model is best at identifying relationships and patterns between variables in a dataset based on the training data. The highest testing accuracies were considered to select the best machine learning models. Using the Logistic Regression, the lung cancer risk prediction classification model achieved 94% training accuracy and 89.3% testing accuracy. Furthermore, the SVC model achieved training and testing accuracies of 86.82% and 83.93%, respectively. Gaussian NB model was achieved, resulting in 92.73% training and 83.93% testing accuracy and Decision Tree Classifier achieved 96.82% training accuracy and 89.29% testing accuracy. Similarly, for the Lung cancer risk prediction classification model, Random Forest Classifier was selected as

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 7, Issue 11, pp 79-86, November-2023

https://doi.org/10.47001/IRJIET/2023.711011

the best model with 99.55% training and 91.07 % testing accuracy.

The Mask R CNN model was used to identify lung cancer tumors from the CT scans. In computer vision and object detection, the IOU and mAP metrics play a crucial role in evaluating the performance of object detection models. The IOU is a measurement that quantifies the overlap between the predicted bounding boxes and the ground truth bounding boxes. A higher IOU indicates a better match between the expected and ground truth bounding boxes. This metric is calculated as the ratio of the two masks' intersection area to the union area.

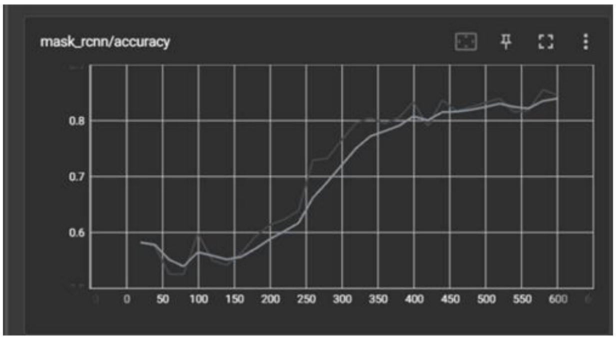

Figure 3: Model Accuracy after training model

Accuracy is a measure of overall correct predictions made by the model. Higher accuracy indicates better performance. However, accuracy alone may not be sufficient for evaluating the performance of Mask R-CNN.

As a result of the training model for generating a Patient Treatment Priority introduces an innovative approach using The logistic regression model, trained with patient data, achieved an accuracy of 83.86%, supporting informed clinical decision-making for lung cancer treatment priorities. By filtering data based on a doctor's input and patient age, tailored recommendations were provided, enhancing patient outcomes and resource allocation. Additionally, we explored the application of the K-nearest neighbors (KNN) algorithm as an alternative approach. However, it did not surpass the predictive capability of logistic regression, and its accuracy was suboptimal. Future research may delve into KNN's performance more comprehensively. Our data-driven approach empowers healthcare professionals to optimize treatment allocation, prioritizing critical cases, and ultimately improving patient outcomes.

Tokenization is utilized to identify and recognize users' questions, where queries are identified and broken down into individual tokens. During this process, punctuation marks like dots, commas, and question marks are excluded from the tokens.

The preprocessing steps for user queries mirror those used during knowledge base preparation. These steps involve tokenization, slang word verification, and morphological analysis. For instance, consider the query: "What does it mean to have a mental illness??" After tokenization, the query is transformed into individual tokens. These tokens are then cross-referenced with a dictionary to identify and replace any slang words with standard equivalents. Lastly, morphology is examined to determine the basic form and type of each word. The result of this preprocessing yields the following tokens and their corresponding word types: what(WH), does(VB), it(NN), mean(VB), to(PRE), have(VB), a, mental(NN), illness(NN).

While doing the research due to the rarity of lung cancer in the general population, it was difficult to gather a large enough sample size for the research study. There were ethical limitations in conducting research on lung cancer risk status. These limitations were hinder the ability to gather comprehensive and accurate data. Research studies are often conducted in specific populations, which cannot be representative of the broader population. This was limit the external validity of the findings and make it difficult to generalize the results to other populations. And also for limitation of Mask R-CNN is its computational cost and resource requirements. This limitation arises due to the complex nature of the model, which consists of multiple stages including region proposal, object detection, and mask prediction. This complexity leads to longer training and inference times, making it less suitable for real-time applications or deployment on resource-constrained devices.

Another limitation of Mask R-CNN is its dependency on labeled data for training. Acquiring large amounts of accurately labeled data can be challenging and time-consuming, especially for niche or specialized domains. Additionally, the process of manually labeling the masks for each instance in the dataset can be labor-intensive and prone to errors.

V. CONCLUSION AND FUTURE WORK

In conclusion, the development of an assistive system for identifying and managing lung cancer has the potential to greatly improve diagnosis, treatment, and research. By utilizing NLP and machine learning algorithms, this system can analyze large amounts of data, provide valuable insights, and aid in early detection and personalized treatments.

As the future of this research project, enriching the knowledge base with diverse patterns and sentence structures can enhance response quality. It could include creating a user-friendly mobile app to predict lung cancer risk based on individual health records and manage the lifecycle of a patient.

ISSN (online): 2581-3048
Volume 7, Issue 11, pp 79-86, November-2023
https://doi.org/10.47001/IRJIET/2023.711011

Data collected from app users can contribute to a larger database, improving the predictive model's accuracy over time. Additionally, enhancing the web application with multi-modal data integration, real-time analysis, and expanded cancer detection can lead to comprehensive patient profiles and improved personalized treatment plans. Strengthening privacy and security measures will protect patient data.

ACKNOWLEDGMENT

We would also want to express our gratitude to Dr. Senaka Kandegedra as our External Supervisor and to Dr.A.J .Hilmi Cancer oncologist for their guidance and support throughout the research work.

REFERENCES

[1] Kunkler, I. H., Audisio, R., Belkacemi, Y., Betz, M., Gore, E., … Hoffe, S. (2014). Review of current best practice and priorities for research in radiation oncology for elderly patients with cancer: the International Society of Geriatric Oncology (SIOG) task force. Annals of Oncology, 25(11), 2134–2146. doi:10.1093/annonc/mdu104.

[2] Passaro, A., Addeo, A., Von Garnier, C., Blackhall, F., Planchard, D., Felip, E., Dziadziuszko, R., De Marinis, F., Reck, M., Bouchaab, H., & Peters, S. (2020, January 1). ESMO Management and treatment adapted recommendations in the COVID-19 era: Lung cancer. ESMO Open; Elsevier BV. https://doi.org/10.1136/esmoopen-2020-000820.

[3] Ellis J. The impact of lung cancer on patients and carers. Chron Respir Dis. 2012 Feb;9(1):39-47. doi: 10.1177/1479972311433577. PMID: 22308553.

[4] Polanski J, Jankowska-Polanska B, Rosinczuk J, Chabowski M, Szymanska-Chabowska A. Quality of life of patients with lung cancer. Onco Targets Ther. 2016 Feb 29;9:1023-8. doi: 10.2147/OTT.S100685. PMID: 27013895; PMCID: PMC4778772.

[5] Pompili C. Quality of life after lung resection for lung cancer. J Thorac Dis. 2015 Apr;7(Suppl 2):S138-44. doi: 10.3978/j.issn.2072-1439.2015.04.40. PMID: 25984359; PMCID: PMC4419027.

[6] M. Maambo, J. Phiri, M. Kalumbilo, and L. Jaganathan, "Assisted Artificial Intelligence Medical Diagnosis System for Heart Disease", zictjournal, vol. 6, no. 1, pp. 38–43, Dec. 2022.

[7] Margaret R. Spitz, Carol J. Etzel, Qiong Dong, Christopher I. Amos, Qingyi Wei, Xifeng Wu, Waun Ki Hong; An Expanded Risk Prediction Model for Lung Cancer. Cancer Prev Res (Phila) 1 September 2008; 1 (4): 250–254. https://doi.org/10.1158/1940-6207.CAPR-08-0060.

[8] Muller DC, Johansson M, Brennan P. Lung Cancer Risk Prediction Model Incorporating Lung Function: Development and Validation in the UK Biobank Prospective Cohort Study. J Clin Oncol. 2017 Mar 10;35(8):861-869. doi: 10.1200/JCO.2016.69.2467. Epub 2017 Jan 17. PMID: 28095156.

[9] Dritsas, E. and Trigka, M. (2022) Lung cancer risk prediction with machine learning models, MDPI. Available at: https://www.mdpi.com/2504-2289/6/4/139.

[10] G. A. Singh and P. K. Gupta, "Performance analysis of various machine learning-based approaches for detection and classification of lung cancer in humans," Neural Computing and Applications, vol. 31, no. 10, pp. 6863–6877, 2018. doi:10.1007/s00521-018-3518-x.

[11] Patra, R. (2020). Prediction of Lung Cancer Using Machine Learning Classifier. In: Chaubey, N., Parikh, S., Amin, K. (eds) Computing Science, Communication and Security. COMS2 2020. Communications in Computer and Information Science, vol 1235. Springer, Singapore. https://doi.org/10.1007/978-981-15-6648-6_11.

[12] S. Bhardwaj et al., "Detection and classification of lung cancer CT images using mask R-CNN based generated mask method," 2023 1st International Conference on Innovations in High-Speed Communication and Signal Processing (IHCSP), 2023. doi:10.1109/ihcsp56702.2023.10127160.

[13] M. Liu, J. Dong, X. Dong, H. Yu, and L. Qi, "Segmentation of Lung Nodule in CT Images Based on Mask R-CNN," 2018 9th International Conference on Awareness Science and Technology (iCAST), Fukuoka, Japan, 2018, pp. 1-6, doi: 10.1109/ICAwST.2018.85.

[14] Ying Xie, Wei-Yu Meng, Run-Ze Li, Yu-Wei Wang, Xin Qian, Chang Chan, Zhi-Fang Yu, Xing-Xing Fan, Hu-Dan Pan, Chun Xie, Qi-Biao Wu, Pei-Yu Yan, Liang Liu, Yi-Jun Tang, Xiao-Jun Yao, Mei-Fang Wang, Elaine Lai-Han Leung, Early lung cancer diagnostic biomarker discovery by machine learning methods, Translational Oncology, Volume 14, Issue 1,2021,100907,ISSN 1936-5233,https://doi.org/10.1016/j.tranon.2020.100907.

[15] Sabina Asensio-Cuesta, Ángel Sánchez-García, Teresa Soria Comes, Inmaculada Maestu, Maria Martín Ureste, J. Alberto Conejero & Juan M. García-Gómez (2022) Testing Lung Cancer Patients' and Oncologists' Experience with the Lalaby App for Monitoring the Quality of Life through Mobile Sensors and Integrated Questionnaires, International Journal of Human–Computer Interaction, DOI: 10.1080/10447318.2022.2121561.

ISSN (online): 2581-3048
Volume 7, Issue 11, pp 79-86, November-2023
https://doi.org/10.47001/IRJIET/2023.711011

[16] R. P.R., R. A. S. Nair, and V. G., "A Comparative Study of Lung Cancer Detection using Machine Learning Algorithms," 2019 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT), Coimbatore, India, 2019, pp. 1-4, doi: 10.1109/ICECCT.2019.8869001.

[17] Lu, T. et al. (2019) 'Trends in the incidence, treatment, and survival of patients with lung cancer in the last four decades', Cancer Management and Research, Volume 11, pp. 943–953. doi:10.2147/cmar.s187317.

[18] TAS, F. et al. (2013) 'Age is a prognostic factor affecting survival in lung cancer patients', Oncology Letters, 6(5), pp. 1507–1513. doi:10.3892/ol.2013.1566.

[19] Pourhomayoun, M. and Shakibi, M. (2021) 'Predicting mortality risk in patients with COVID-19 using machine learning to help medical decision-making', Smart Health, 20, p. 100178. doi:10.1016/j.smhl.2020.100178.

[20] Wang, Y. et al. (2021) 'A machine learning-based investigation of gender-specific prognosis of lung cancers', Medicina, 57(2), p. 99. doi:10.3390/medicina57020099.

[21] Elcholiqi, Abidah & Musdholifah, Aina. (2020). Chatbot in Bahasa Indonesia using NLP to Provide Banking Information. IJCCS (Indonesian Journal of Computing and Cybernetics Systems). 14. 91. 10.22146/ijccs.41289.

[22] H. Yan, H. Lu, M. Ye, K. Yan, Y. Xu, and Q. Jin, "Improved Mask R-CNN for Lung Nodule Segmentation," 2019 10th International Conference on Information Technology in Medicine and Education (ITME), Qingdao, China, 2019, pp. 137-141, doi: 10.1109/ITME.2019.00041.Motorola Semiconductor Data Manual, Motorola Semiconductor Products Inc., Phoenix, AZ, USA, 1989.

Citation of this Article:

D.A.H.K. Rathnayaka, E.M.V.Y. Ekanayake, K.K. Paranavithana, K.B.A.S.M. Dissanayake, Wishalya Tissera, Dasuni Nawinna, "Assistive System to Identify and Manage Lung Cancer" Published in *International Research Journal of Innovations in Engineering and Technology - IRJIET*, Volume 7, Issue 11, pp 79-86, November 2023. Article DOI https://doi.org/10.47001/IRJIET/2023.711011

****