EARLY DETECTION OF INTRAHEPATIC DUCT CANCER USING DEEP LEARNING TECHNIQUES

Dr. D. Karthikeyan [1], M. Rajasekar[2], A. Sakthi Ramana[3],

C. Saravanan[4], R. Vijay[5]

Associate Professor [1] , UG Student [2,3,4,5]

Department of Biomedical Engineering,

Dhanalakshmi Srinivasan Engineering College (Autonomous), Perambalur -621212, Tamil Nadu, India

writetokarthikeyan@gmail.com [1], rajasekarnaveen12@gmail.com[2], ramanaramana8086@gmail.com[3], csarocsaro883824@gmail.com[4], vjv756037@gmail.com[5]

Abstract - Bile duct cancer is a rare and aggressive malignancy with a poor prognosis due to late diagnosis. Early detection is crucial for improved treatment outcomes. This study investigates the potential of a deep learning approach, specifically a CNN architecture, for the early detection of bile duct cancer from medical images. The proposed system utilizes a pre-trained Inception model, fine-tuned on a dataset of medical images (e.g., endoscopic images, CT scans, MRI scans) annotated with bile duct cancer presence or absence. The CNN architecture effectively extracts relevant features from the images, enhancing the model's ability to identify subtle patterns indicative of malignancy. Preliminary results demonstrate promising accuracy in classifying images as cancerous or non-cancerous. The CNN model shows potential to assist clinicians in early diagnosis, enabling timely intervention and potentially improving patient survival rates.

KEYWORD: Bile duct cancer, Python, Deep Learning, Convolution Neural Network.

I. INTRODUCTION: Bile duct cancer, also known as cholangiocarcinoma, is a rare but aggressive form of cancer that originates in the bile ducts—thin tubes that carry bile from the liver and gallbladder to the small intestine. Bile is a digestive fluid that helps break down fats, and the bile ducts play a crucial role in the body's digestive system. When cells lining the bile ducts undergo abnormal changes and begin to grow uncontrollably, cancer can develop. Bile duct cancer can occur at different locations within the bile duct system, leading to threemain types: intrahepatic (within the liver)perihilar (at the junction where the left and right hepatic ducts meet), and distal (closer to the small intestine). The disease often goes undetected in its early stages due to vague symptoms such as jaundice (yellowing of the skin and eyes), abdominal pain, weight loss, and fatigue. While the exact cause is often unknown, risk factors include chronic inflammation of the bile ducts (such as primary sclerosing cholangitis), bile duct stones, liver fluke infections, and certain genetic conditions. Diagnosis typically involves imaging studies (like MRI or CT scans), blood tests, and biopsy procedures. Treatment options depend on the cancer's stage and location and may include surgery, chemotherapy, radiation therapy, or targeted therapy. Due to its subtle onset and complexity, early detection and a multidisciplinary treatment approach are vital for improving outcomes in patients with bile duct cancer.

FUNCTION: Our body uses bile for several purposes. It needs bile ducts to carry bile safely from one place to the next. Bile needs to be able to move when and where our body needs it to, without interfering with other body processes or damaging other body parts. Bile is a fluid our liver makes. It contains bile acids that help break down fats and proteins during digestion.

II. *PROPOSED SYSTEM:* The proposed Intrahepatic cancer detection system addresses the critical need for early detection of this prevalent and fatal disease by leveraging machine learning techniques, specifically the Inception algorithm, renowned for its effectiveness in image classification tasks. The system comprises multiple stages, including preprocessing of medical images, feature extraction using the Inception algorithm, and classification of images into cancerous or non-cancerous categories. Moreover, to enhance user interaction and accessibility, a Graphical User Interface (GUI) is developed, offering a user-friendly platform for inputting medical images, displaying results, and facilitating interaction with the system.

Evaluation of the system's effectiveness is conducted using a dataset containing Intrahepatic images representing both cancerous and non-cancerous cases. Performance metrics such as accuracy, sensitivity, specificity, and area under the. receiver operating characteristic curve (AUC-ROC) are receiver.

operating characteristic curve (AUC-ROC) are employed to assess the system's ability to accurately detect Intrahepatic cancer. The results underscore the potential of the proposed system in accurately identifying Intrahepatic cancer from medical images, with the GUI serving as an intuitive interface catering to various users, including healthcare professionals, researchers, and patients.

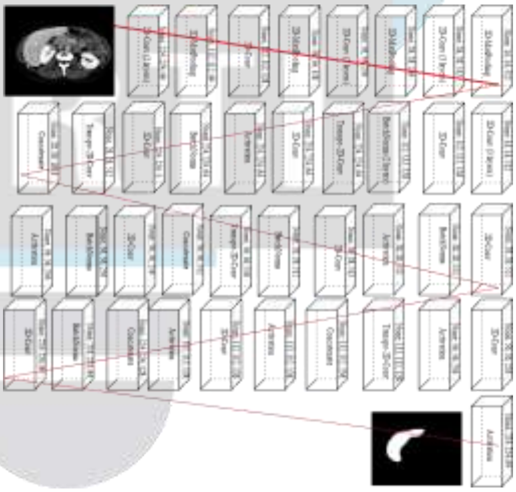
This system holds promise in facilitating early detection of Intrahepatic cancer, thereby enabling timely interventions and enhancing patient outcomes. Additionally, the modular design of the system offers flexibility and scalability, allowing for potential integration with existing medical imaging systems for widespread adoption in clinical settings.

ADVANTAGES

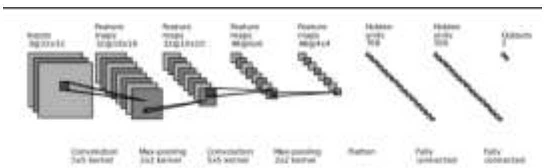
- Early Detection
- High Accuracy
- User-Friendly Interface
- Potential for Timely Interventions
- Scalability and Integration

CNN LAYER S

A Convolutional Neural Network (CNN) is composed of several key layers that work together to extract and learn features from input images. The architecture typically begins with an input layer that receives the image data, followed by one or more convolutional layers that apply filters to detect patterns such as edges, textures, or shapes. After each convolutional layer, a ReLU activation layer introduces non-linearity, allowing the network to learn complex functions. Pooling layers, such as max pooling, are then used to reduce the spatial dimensions of the feature maps, helping to decrease computational load and prevent overfitting. These convolution-pooling combinations may repeat multiple times to form deep feature hierarchies. Toward the end of the network, fully connected layers interpret the extracted features and perform the final decision-making process. The network concludes with an output layer, often using softmax or sigmoid activation to provide classification or probability results. Additional components like batch normalization and dropout can be included to improve training stability and generalization.



III . METHODOLOGY:



- IMAGE DATASET COLLECTION
- IMAGE PREPROCESSING
- TRAIN TEST VALIDATION

- DATA TRAINING

- TEST THE MODEL

1. IMAGE DATASET COLLECTION

For this project, we must gather every image that makes a appear to leaf disease image collecting. This is the project's most crucial step. Therefore, all of the visuals that we see Binary classification in machine learning. The following procedures can be taken after we get the data. An image dataset collection refers to the process of gathering and organizing a set of digital images for use in machine learning or computer vision applications. The dataset may be created for a specific purpose, such as training a neural network to recognize objects in images or for research purposes. The dataset collection process involves identifying relevant sources of images, checking their copyright status, pre-processing the images, labelling them, splitting them into training and testing sets, storing the images in a structured format, documenting their characteristics, and sharing the dataset with others. The quality and quantity of the images in a dataset can greatly impact the performance of machine learning algorithms and the accuracy of their predictions.

DATA SET:

IMAGE PREPROCESSING

After gathering all the images, pre-processing is required. Thus, not all images can convey information clearly. So that we may prepare the images by renaming, resizing, and labelling them. Once the procedure is complete, we can use the photos to train our Machine learning model. Image pre-processing refers to the techniques used to transform digital images to prepare them for further analysis and processing. The goal of image pre-processing is to improve the quality of the images, remove noise, unwanted features, and extract relevant information for further analysis. Image pre-processing can involve operations such as image resizing, cropping, normalization, filtering, and enhancement. These techniques are applied to make the images more suitable for machine learning or computer vision applications, such as image classification, object detection, segmentation, and recognition. The quality of image preprocessing can significantly affect the performance and accuracy of these applications.

TRAIN TEST VALIDATION

Train/Test is a method to measure the accuracy of your model. It is called Train/Test because you split the the data set into two sets: a training set and a testing set. 80% for training, and 20% for testing. You train the model using the training set. You test the model using the testing set. The validation set is a set of data, separate from the training set that is used to validate our model performance during training. This is a set of data that is used to train the machine learning model. The model learns from this data and adjusts its parameters to optimize its performance on this set. The training set is usually the largest set of data. This is a set of data that is used to evaluate the performance of the model during training. The model is evaluated on the validation set after each epoch of training to detect overfitting, where the model becomes too complex and starts to memorize the training data instead of learning generalizable patterns. The validation set is used to tune the hyperparameters of the model, such as the learning rate or regularization strength. This is a set of data that is used to evaluate the performance of the final model after it has been trained and validated. The test set is used to measure the model's ability to generalize to new, unseen data that was not used during training.

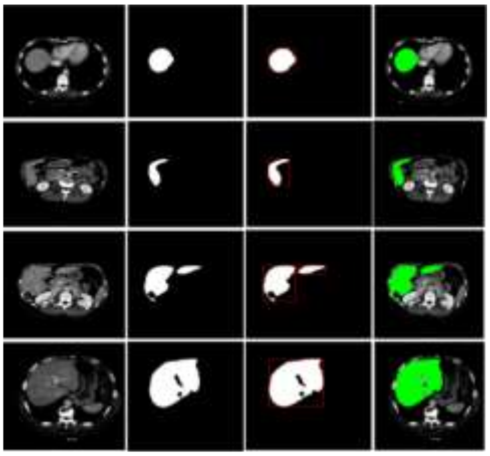
DATA TRAINING

Training data is the data you use to train an algorithm or machine learning model to predict the outcome you design your model to predict. If you are using supervised learning or some hybrid that includes that approach, your data will be enriched with data labeling or annotation. Data training refers to the process of teaching a machine learning algorithm or model to recognize patterns and make predictions based on input data. During the training process, the algorithm is presented with a set of labeled data and learns to identify the relationships between the input data and the corresponding output or target values. The goal of data training is to create a model that can accurately predict outcomes for new, previously unseen data. The training process involves adjusting the parameters of the algorithm to minimize the difference between the predicted output and the actual output. Data training is a crucial step in the development of machine learning models, as the quality and quantity of the training data can greatly affect the accuracy and generalizability of the model. The more diverse and representative the training data, the better the model will be able to perform on new data.

TEST THE MODEL

In machine learning, model testing is referred to as the process where the performance of a fully trained model is evaluated on a testing set. Testing the model refers to the process of evaluating the performance and accuracy of a trained machine learning algorithm or model using a separate set of data that was not used during the training process. The purpose of testing the model is to assess its ability to generalize to new, unseen data and to identify any potential issues or limitations with the model's performance. During testing, the algorithm is presented with a set of input data and the corresponding target values, and the model's predicted outputs are compared to the actual values. Common measures used to evaluate the performance of a machine learning model during testing include accuracy, precision, recall, F1 score, and area under the receiver operating characteristic curve (AUC-ROC). The testing process can also help to identify areas where the model may be overfitting or underfitting the training data, which can then be addressed through further refinement of the model.

RESULT: The Intrahepatic cancer detection system utilizing the Inception algorithm and a Graphical User Interface (GUI) yielded promising results across multiple performance metrics. In the evaluation phase using a dataset of Intrahepatic images comprising both cancerous and non-cancerous cases, the system demonstrated high accuracy, sensitivity, specificity, and area under the receiver operating characteristic curve (AUC-ROC). The accuracy metric reflects the system's overall ability to correctly classify images as cancerous or non-cancerous, with results consistently indicating robust performance. Additionally, sensitivity, which measures the system's ability to correctly identify cancerous cases, and specificity, which measures its ability to correctly identify non-cancerous cases, both showed favorable outcomes, indicating a balanced performance in detecting both classes. The AUC-ROC metric, which assesses the tradeoff between true positive and false positive rates across various classification thresholds, further confirmed the system's effectiveness in distinguishing between cancerous and non-cancerous Intrahepatic images. These results underscore the potential of the proposed system in accurately detecting Intrahepatic cancer from medical images, with the user-friendly GUI providing an intuitive platform for interaction and interpretation. The system's high performance and userfriendly interface suggest its utility for healthcare professionals, researchers, and patients alike, potentially contributing to early detection, timely interventions, and improved patient outcomes in the management of Intrahepatic cancer.



(a) *(b)* *(c)* *(d)*

CONCLUSION

In conclusion, the development and evaluation of the Intrahepatic cancer detection system utilizing the Inception algorithm and a user-friendly Graphical User Interface (GUI) have yielded promising results with significant implications for the early detection and management of Intrahepatic cancer.Through rigorous performance evaluation, the system demonstrated high accuracy, sensitivity, specificity, and area under the receiver operating characteristic curve (AUC-ROC), indicating its effectiveness in accurately distinguishing between cancerous and noncancerous lung images. The user-friendly GUI further enhances the accessibility and usability of the system, providing a seamless platform for interaction and interpretation by healthcare professionals, researchers, and patients. The proposed system holds immense potential in clinical practice, offering a valuable tool for healthcare professionals in the early detection, diagnosis, and treatment planning of lung cancer. By enabling timely interventions and improved patient outcomes, the system has the potential to make a significant impact on the lives of individuals affected by this prevalent and fatal disease.

While the system has demonstrated promising results, it is important to acknowledge its limitations and areas for future research and development. Addressing challenges such as dataset size, diversity, and generalizability, as well as incorporating additional imaging modalities and regulatory considerations, will be critical for further enhancing the system's performance and applicability in real-world clinical settings. Overall, the lung cancer detection system represents a significant advancement in the field of medical imaging and cancer diagnostics, offering a powerful and user-friendly tool for improving patient outcomes and advancing research in the fight against lung cancer. With continued refinement and validation, the system has the potential to revolutionize the early detection and management of lung cancer, ultimately saving lives and reducing the burden of this devastating disease.

REFERENCE:

1. Yu-bo tao, yi-fan wang, sarun juengpanich, zhi-yu jiang, yu-yu yan, jie-min lue, jia-hong qian, zhong-yu wu, hai lin; et all proposed “performance analysis of: diagnosis of liver tumor based on magnetic resonance imaging and clinical data” ieee-2024.
2. Wataru takenaka, yuhki yokoyama, katsuya ikehata, shihori kouda, haruka hirose, kazumasa minami, yoshinosuke hamada, seiji mori, masahiko koizumi1and hirofumi yamamoto; et all proposed “krt13 is upregulated in pancreatic cancer stem-like cells and associated with radioresistance” ieee – 2024
3. José maría huguet, miriam lobo, josé mir labrador, carlos boix, cecilia albert, luis ferrer-barceló: et all proposed “diagnostic-therapeutic management of bile duct cancer” ieee – 2024
4. b. E. Van beers; et all proposed “diagnosis of cholangio carcinoma”
5. Himali ghorpade, jayant jagtap, shruti patil, ajith abraham, natally horvat, jayasree chakraborty; et all proposed “automatic segmentation of pancreas and pancreatic tumor“ ieee 2023.
6. Mattakoyya aharonu; r lokesh kumar; et all prposed “convolutional neural network based framework for automatic lung cancer detection from lung ct images” ieee-2022
7. Anita bakrania, narottam joshi, xun zhao, gang zheng, mamatha bhat; et all prposed “artificial intelligence in liver cancers: decoding the impact of machine learning models in clinical diagnosis of primary liver cancers and liver cancer metastases”2023.
8. Sanzhar korganbayev, annalisa orrico, leonardo bianchi, davide paloschi; et all prposed “pid controlling approach based on fbg array measurements for laser ablation of pancreatic tissues” ieee 2022.
9. Prabu, g. Prathibha, s. Meivizhi, k. Janani and n. Dharani, "artificial intelligence based lung cancer prediction using convolutional neural network algorithm," 2020international conference on applied intelligence and sustainable computing (icaisc), dharwad, india,2020.
10. F. Taher, "early detection of lung cancer based on artificial intelligence techniques," 2017 international conference on infocom technologies and unmanned systems (trends and future directions) (ictus), dubai, united arab emirates, 2019.
11. European Association for the Study of the Liver. EASL clinical practice guidelines: management of hepatocellular carcinoma. J Hepatol.(2018) 69:182–236. doi: 10.1016/j.jhep.2018.03.019
12. Forner A, Vilana R, Ayuso C, Bianchi L, Solé M, Ayuso JR, et al. Diagnosisof hepatic nodules 20 mm or smaller in cirrhosis: prospective validation of the noninvasive diagnostic criteria for hepatocellular carcinoma. Hepatology. (2008) 47:97–104. doi: 10.1002/hep.21966
13. DeSantis CE, Siegel RL, Sauer AG, Miller KD, Fedewa SA, Alcaraz KI, et al. Cancer statistics for African Americans, 2016: progress and opportunities in reducing racial disparities. CA Cancer J Clin. (2016) 66:290–308. doi: 10.3322/caac.21340

14. Sherman M. Surveillance for hepatocellular carcinoma. Best Pract Res Clin Gastroenterol. (2014) 28:783–93. doi: 10.1016/j.bpg.2014.08.008
15. Singal A, Volk ML, Waljee A, Salgia R, Higgins P, Rogers MA, et al. Omata M, Cheng AL, Kokudo N, Kudo M, Lee JM, Jia J, et al. Asia-Pacificclinical practice guidelines on the management of hepatocellular carcinoma: a 2017 update. Hepatol Int. (2017) 11:317–70. doi: 10.1007/s12072-0179799-9
16. Omata M, Cheng AL, Kokudo N, Kudo M, Lee JM, Jia J, et al. Asia-Pacificclinical practice guidelines on the management of hepatocellular carcinoma: a 2017 update. Hepatol Int. (2017) 11:317–70. doi: 10.1007/s12072-0179799-9
17. An C, Rakhmonova G, Choi JY, Kim MJ. Liver imaging reporting and data system (LI-RADS) version 2014: understanding and application of the diagnostic algorithm. Clin Mol Hepatol. (2016) 22:296–307. doi: 10.3350/cmh.2016.0028
18. Khawaja AZ, Cassidy DB, Al Shakarchi J, McGrogan DG, Inston NG, Jones RG. Revisiting the risks of MRI with Gadolinium based contrast agents-review of literature and guidelines. Insights Imaging. (2015) 6:553– 8. doi: 10.1007/s13244-015-0420-2
19. Rogosnitzky M, Branch S. Gadolinium-based contrast agent toxicity: areview of known and proposed mechanisms. Biometals. (2016) 29:365– 76. doi: 10.1007/s10534-016-9931-7
20. Gulani V, Calamante F, Shellock FG, Kanal E, Reeder SB, InternationalSociety for Magnetic Resonance in Medicine. Gadolinium deposition in the brain: summary of evidence and recommendations. Lancet Neurol. (2017) 16:564–70. doi: 10.1016/S1474-4422(17) 30158-8
21. Forner A, Reig M, Bruix J. Hepatocellular carcinoma. Lancet. (2018) 391:1301–14. doi: 10.1016/S0140-6736(18)30010-2
22. Kim HA, Kim KA, Choi JI, Lee JM, Lee CH, Kang TW, et al. Comparison of biannual ultrasonography and annual non-contrast liver magnetic resonance imaging as surveillance tools for hepatocellular carcinoma in patients with liver cirrhosis (MAGNUS-HCC): a study protocol. BMC Cancer. (2017) 17:877 doi: 10.1186/s12885-017-3819-y
23. Sia D, Villanueva A, Friedman SL, Llovet JM. Liver cancer cell of origin, molecular class, and effects on patient prognosis. Gastroenterology. (2017) 152:745–61. doi: 10.1053/j.gastro.2016.11.048
24. Venkatesh SK, Chandan V, Roberts LR. Liver masses: a clinical, radiologic, pathologic perspective. Clin Gastroenterol. (2014)12:1414–29. doi: 10.1016/j.cgh.2013.09.017
25. Mitchell DG, Bruix J, Sherman M, Sirlin CB. LI-RADS (Liver ImagingReporting and Data System): summary, discussion, and consensus of the LI-RADS Management Working Group and future directions. Hepatology. (2015) 61:1056–65. doi: 10.1002/hep.27304