Deep Learning-Based Liver Histopathology Image Classification: State-of-the-Art Techniques and Emerging Trends

E Pavan Kumar¹; Habibur Rahaman²; Chityala Vishnuvardhan Reddy³; Rokkam Sahil⁴; Dr. Shwetha Buchanalli⁵; Bharani Kumar Depuru⁶ ¹Research Associate, AiSpry, Hyderabad, India. ²Research Associate, AiSpry, Hyderabad, India. ³Research Associate, AiSpry, Hyderabad, India. ⁴Research Associate, AiSpry, Hyderabad, India. ⁵Team Leader, Research and Development, AiSpry, Hyderabad, India.

Abstract:- This research investigates the application of deep learning techniques to enhance the diagnostic accuracy of liver tumour classification in collaboration with a prominent hospital in South India. By leveraging a carefully curated dataset of histopathological images, we evaluated the performance of several advanced deep learning architectures, including DenseNet 121, ResNet50, and VGG16. Our findings reveal that DenseNet121 outperformed the other models, achieving the highest accuracy in both training and testing phases, thus exceeding our predefined accuracy benchmarks. The superior performance of DenseNet121 is attributed to its dense connectivity, which facilitates improved feature and gradient propagation throughout the network. This study highlights the significant potential of AI-driven diagnostics in enhancing liver tumour classification. thereby optimizing the diagnostic workflow and providing substantial benefits for patient care and healthcare system efficiency.

Keywords:- Deep Learning Models, Classification, Cholangiocarcinoma (CC), Hepatocellular Carcinoma (HCC).

I. INTRODUCTION

The liver, a vital organ in animals, particularly vertebrates, plays a crucial role in supporting various bodily functions. Liver diseases can develop insidiously in humans, often without noticeable symptoms, making early diagnosis critical for improving patient outcomes. Among these diseases, liver tumors are particularly dangerous due to their high mortality rates. Radiologists face significant challenges in the early detection and accurate staging of liver tumors. Early detection significantly improves survival rates, highlighting the importance of prompt and precise diagnostic methods[19].

Histopathological image analysis(HIA) is the gold standard for tumor qualitative and clinical opinion. HIA is a crucial step in achieving the pretensions of early discovery, opinion, and treatment of liver tumor, which is generally performed by pathologists through visual observation, but this process is time-consuming, tedious, and fluently limited by the experience of pathologists themselves. Thus, it's further necessary to apply an automated HIA for liver tumor which can ameliorate the delicacy and effectiveness of opinion [7].

The adoption of AI in the medical industry is significantly enhancing diagnostic practices, especially in the intricate area of liver tumor identification [3]. This project, in collaboration with a prominent South Indian hospital, seeks to transform the traditional approach of relying on manual image interpretation by clinicians into one that utilises AI for more accurate and quicker diagnostics. However, manual liver pathology examination requires considerable time and labour[4].

The liver, a critical organ second only in size to the skin, typically weighs between 1.2 and 1.5 kilograms and is divided into two main lobes [12]. It lies adjacent to the gallbladder, pancreas, and intestines, functioning primarily through hepatocytes, which account for 70-85% of its mass. These cells are essential for various functions, including metabolism and detoxification. Liver tumors can range from benign to malignant, with Hepatocellular carcinoma (HCC) being the most prevalent type of primary liver tumor[1], frequently associated with chronic liver diseases like cirrhosis. Another notable but rarer malignancy is intrahepatic cholangiocarcinoma, which forms within the bile ducts of the liver [12].

For challenging classification tasks that may be difficult for human experts or require the quick review of a large volume of images, AI offers significant benefits, including time efficiency, high accuracy, and consistent performance[8]. Moreover, recent uses of AI in pathology have predominantly concentrated on models that operate on whole-slide images (WSI) in a fully automated manner, without human involvement, until the post-analytical phase[2]. By applying deep learning technology, our project categorises liver imaging into three classifications: Cholangiocarcinoma, Hepatocellular carcinoma tumor

International Journal of Innovative Science and Research Technology

ISSN No:-2456-2165

(HCC), and normal liver [2]. Hepatocellular carcinoma (HCC) is a prevalent form of cancer and one of the leading causes of cancer-related mortality globally[5].

This methodology not only expedites the diagnostic process but also increases its precision, thus significantly improving over traditional methods.

Our project aims to markedly enhance the accuracy of liver tumor diagnostics, reduce the costs associated with

these diagnoses, and improve patient outcomes. Our objectives include raising the liver tumor diagnosis rate by at least 10%, achieving a classification accuracy of at least 98%, and attaining substantial savings in diagnostic expenses. These advancements are expected to benefit patient care and provide significant economic benefits by optimizing resource utilization in healthcare settings.

https://doi.org/10.38124/ijisrt/IJISRT24JUN1368

\succ CRISP-ML(Q)



Fig 1 This figure Illustrates the CRISP-ML(Q) Architecture Implemented in Our Research Project. (Source:- Mind Map - 360DigiTMG)

CRISP-ML(Q) provides a structured approach for machine learning projects with a focus on quality assurance. The framework guides the development process from initial planning to final deployment and ongoing monitoring.

Here is how CRISP-ML(Q) can be applied to our project "Liver Tumour Classification" in the healthcare industry:

The first step is **business understanding**[Fig.1], which involves defining the project's objectives and goals, such as creating a deep learning models for classifying liver tumours. This stage requires a deep understanding of the challenges within the health sector, including the potential for improved diagnostic accuracy and personalized treatment plans.

In the **data understanding phase** [Fig.1], relevant data such as medical images and associated patient data is collected and analyzed. Exploratory data analysis helps identify data quality issues, completeness, and patterns that will inform subsequent steps.

Data preparation[Fig.1] involves cleaning and transforming the data to make it suitable for modeling. This includes addressing missing values, normalizing and segmenting images, and using augmentation techniques to enhance data quality and diversity.

During the **modelling phase**[Fig.1], deep learning models such as convolutional neural networks are employed for liver tumor classification. The selection and design of models are guided by project objectives and data characteristics. Hyperparameter tuning and model optimization are conducted to improve performance.

• DenseNet121:

DenseNet121 is a convolutional neural network architecture renowned for its dense connectivity pattern, where each layer is connected to every other layer. With 121 layers, it efficiently learns hierarchical representations of visual features and has shown remarkable performance in image classification tasks. In this research, we harness the power of DenseNet121 for the precise identification of liver tumours, leveraging its deep learning capabilities to achieve accurate and reliable classification results.[18]

ISSN No:-2456-2165

International Journal of Innovative Science and Research Technology

https://doi.org/10.38124/ijisrt/IJISRT24JUN1368

• ResNet50:

Introduced by Microsoft Research in 2015. Features 50 layers and innovative residual blocks. Residual connections mitigate the vanishing gradient problem.[17]

• VGG16:

Developed by the Visual Geometry Group at the University of Oxford. Consists of 16 layers (13 convolutional and 3 fully connected). Known for its simplicity and uniformity. Captures fine-grained details due to small filter sizes.[16]

Model evaluation[Fig.1] assesses the trained model's performance using metrics such as accuracy, precision, recall, and F1-score. Cross-validation and testing on holdout datasets ensure the model's reliability and ability to generalize.

Once the model is validated, it is deployed into a medical diagnostic system for liver tumor classification. This allows healthcare professionals to use the model's predictions to support diagnostic decisions and treatment planning.

Ongoing monitoring and maintenance[Fig.1] ensure the model continues to perform well over time. This includes tracking key performance indicators, updating the model with new data, and adjusting the model as medical imaging technology and standards change.

By implementing CRISP-ML(Q) in the liver tumor classification project, the project benefits from a comprehensive approach to machine learning development that prioritizes high-quality outcomes and improved patient care.

➢ Architecture Diagram

An architecture diagram provides the initial clear picture of any project, serving as a visual blueprint that outlines the overall structure and flow of data and processes. This diagram captures key components such as Business and Data collection, Data collection, Data preprocessing, Model building, Model deployment, monitoring and maintenance and offers a high-level overview that guides decisionmaking and planning. By establishing a clear architectural vision, teams can align on objectives and design scalable, efficient solutions from the outset.



Fig 2 Architecture Diagram Representing Liver Tumor Classification

The first step is to work closely with Business and Data Understanding[Fig.2]. Accomplishing this requires us to identify important indicators, and objectives, as well as understand where data is derived and its quality.

Data collection [Fig.2] takes at the heart of information retrieval from the image collection system capturing a wide range of pictures as data sources for the subsequent study. First of all, data is collected, [Fig.2] then stored accordingly and afterwards undergoes the pre-processing [Fig.2] part, i.e. image resizing and augmentation which assures quality and diversity. These pretreatment steps are aimed at gathering knowledge that is further useful for model building.

Having pre-processed data, CNN [Fig.2] is used as a model to exactly deal with the issue in our hand. The model is trained and blends with the pre-processed data, making it able to learn and it recognizes the patterns. When the model is developed, it goes through the process of rigorous testing against metrics that have already been established, for the purpose of determining its efficiency in quantitative terms. best-doing[Fig.2] This model uploads for deployment[Fig.2], into taking account these evaluations[Fig.2] as criteria.

ISSN No:-2456-2165

Streamlit[Fig.2], the web application framework, helps with the deployment of the model we have selected as it provides the means of establishing such interactive interfaces. This feature gives users an opportunity to get visually engaged in the model's prediction in actual time.

Regular monitoring and maintenance[Fig.2] are essential in sustaining the higher efficiency of the model. The operations of the system revolve around monitoring the main components, updating the model if needed, and conducting regular quality control checks to ensure that accuracy and reliability remain unaffected. By continuously monitoring the model's performance and adjusting it according to new data and business requirements, the research team can ensure that the model remains up-to-date and in sync with the commercial establishment's objectives.

II. METHODOLOGY AND TECHNIQUES

https://doi.org/10.38124/ijisrt/IJISRT24JUN1368

> Data Collection

For this research project, the dataset provided by the client focuses on categorizing patients into three differentiation types: Normal Liver, Hepatocellular Carcinoma (HCC), and Cholangiocarcinoma. The dataset includes a total of 21000 histopathology images: 7000 images each representing HCC, Normal Liver, and Cholangiocarcinoma tumor types. This dataset will be utilized to develop and train a model for automated classification of liver cancer types based on histopathological images.

> Data Preprocessing

Medical image analysis plays a pivotal role in diagnosing diseases, monitoring treatment progress, and aiding clinical decision-making. However, before feeding these images into deep learning models, a robust preprocessing pipeline is essential to enhance data quality, mitigate noise, and ensure optimal model performance.



Fig 3 a: Cholangiocarcinoma Liver Patch b: Normal Liver Patch and c: HCC Liver Patch. The Patch Slice Size was Set to 256×256 Pixels.

(Source: https://www.researchgate.net/figure/Left-Normal-patch-and-right-HCC-patch-The-patch-slice-size-was-set-to-448448-pixels_fig1_349508422)

Noise Reduction

✓ Noise Filtering:

Medical images often suffer from noise due to acquisition imperfections or artefacts. Applying filters (e.g., Gaussian or median filters) can suppress noise while preserving relevant features.[12]

• Shift:

To introduce diverse perspectives, random alterations in width and height are applied to enhance the dataset, thus creating an impression of shifts in perspective. The maximum shift allowed is 20%.[14].

• Cropping and Padding

✓ *Region of Interest (ROI):*

Cropping the image to focus on the relevant anatomical region (e.g., liver or tumor) reduces computational load and enhances model efficiency[15].

✓ Padding:

Adding padding to ensure uniform image dimensions is essential for batch processing within neural networks.[15]

Normalization

Data normalisation aims to convert numerical columns in a dataset into more readable scales without losing their original variability. While not mandatory, it becomes particularly important when dealing with data that has wide value ranges. By applying this process, we can ensure each record stores only relevant information and prevent unnecessary duplications of data, leading ultimately to significant space savings compared to alternatives. Managing storage needs is crucial for efficient data collection and preservation; hence the implementation of flexible database architecture designed against redundancies caused by duplicated records. In essence, normalising values on different scales effectively merges them onto one commonly shared scale before averaging, thus conceptually defining data normalization [11].

ISSN No:-2456-2165

Data Augmentation

Deep learning in medical data analysis faces challenges due to limited datasets. To address this, augmentation techniques are employed to artificially expand the data. The most common augmentations include 90degree rotation, transposition, horizontal flip, and vertical flip. These techniques help prevent overfitting by diversifying the training data.

For preprocessing, TensorFlow's Keras API and its ImageDataGenerator class are used for real-time normalization and augmentation. Normalization adjusts

International Journal of Innovative Science and Research Technology

https://doi.org/10.38124/ijisrt/IJISRT24JUN1368

pixel values to a 0-1 range, enhancing model training. Augmentation introduces variations through random rotations (up to 10 degrees), shifts in width and height (up to 20%), shear transformations (0.2 range), zooming (0.2 range), and horizontal flips. The fill mode parameter determines how new pixels are filled, with options like nearest neighbor interpolation.

The approach uses a pre-trained model, which benefits from the augmented data to develop adaptable features and improve convergence speed during training through effective weight updates.[11,15].



Horizontal Flip

Vertical Flip

Fig 4 Data augmentation of Image (Source: Data augmentation of BreakHis image | Download Scientific Diagram (researchgate.net))

Model Approach \geq

We evaluated several models, including DenseNet121, ResNet50, ResNet101, ResNet152, and VGG16, to classify liver histopathology images into three categories: Normal Hepatocellular Carcinoma Liver, (HCC), and Cholangiocarcinoma. Our target was to achieve at least 98% accuracy on the testing dataset.

- Model Performance Summary:
- DenseNet121:
- Training Accuracy: 98.21%
- Testing Accuracy: 96 %
- Analysis:

DenseNet121 exceeded our accuracy benchmark. Its architecture, characterized by dense connectivity, promotes feature reuse and mitigates the vanishing-gradient problem, enhancing both learning efficiency and generalization capability.

ResNet50:

This model resulted in overfitting and didn't work well on the unseen data

- ResNet101 and ResNet152:
- Performance:

Both models did not meet our accuracy benchmark, indicating that further optimization or hyperparameter tuning might be necessary.

- *VGG16:*
- Testing Accuracy: 72%
- Analysis:

VGG16 had the lowest testing accuracy, highlighting its limitations for this particular classification task. This may be due to its relatively shallow architecture and lack of advanced connectivity features compared to DenseNet and ResNet models.Comparison with Other Models:

ISSN No:-2456-2165

• Key Insights on DenseNet121:

DenseNet121's distinctive feature is its dense connectivity pattern, where each layer is connected to every other layer in a feed-forward fashion. This design facilitates improved information flow and gradient propagation throughout the network, leading to better learning outcomes. DenseNet121 not only achieved a training accuracy of 98.21% and a testing accuracy of 97%, surpassing our benchmark, but also demonstrated robustness against overfitting. This can be attributed to its architectural design, which effectively addresses the vanishing-gradient problem and promotes efficient feature reuse.

https://doi.org/10.38124/ijisrt/IJISRT24JUN1368

In conclusion, DenseNet121 was the clear winner in our study. It's not just efficient, but also effective for our image classification tasks. And the best part? It's robust against overfitting, as shown by its consistent accuracy across the training and testing phases.

> Research Findings

The primary objective of this research was to ascertain the efficacy of various deep learning models in the classification of liver tumours using image data. Among the models tested, DenseNet121 was found to be particularly effective, as evidenced by its high accuracy rates on both training (98.21%) and testing datasets (97%).

Model name	Train Accuracy	Test Accuracy
Densenet121	98.21%	97%
ResNet50	85%	67%
ResNet150	94.00%	92%
ResNet151	88.10%	86%
VGG16	72%	65%

> Implications for Future Research:

The findings from this study underscore the potential of DenseNet121 in medical imaging tasks. Future research could explore the integration of DenseNet121 into clinical workflows, assessing its real-world applicability and impact on diagnostic processes.

Additionally, comparative studies involving larger datasets and varied tumour types could provide deeper insights into the model's generalizability and scalability. This section should be tailored with specific statistical analyses, dataset details, and any unique observations from your research to enhance its relevance and impact. If you have specific results or data you would like to include, please let me know, and I can incorporate them into this section.

III. RESULTS AND DISCUSSION

In this study, we evaluated the performance of several deep learning models—DenseNet121, ResNet50, ResNet101, ResNet152, and VGG16—on the task of classifying liver histopathology images into three categories: Normal Liver, Hepatocellular Carcinoma (HCC), and Cholangiocarcinoma. The goal was to achieve at least 96% accuracy on the test dataset.

> DenseNet121:

DenseNet121 demonstrated superior performance, achieving a training accuracy of 98.21% and a test accuracy of 97%, slightly below our benchmark but significantly higher than the other models. Its dense connectivity pattern, which facilitates feature reuse and mitigates the vanishing-gradient problem, likely contributed to its high accuracy and robust generalization.

ResNet Models:

• ResNet50:

Achieved a training accuracy of 85% and a test accuracy of 67%. The substantial drop in accuracy from training to testing indicates overfitting, suggesting that the model struggled to generalize from the training data to unseen data.

• ResNet152:

Achieved a training accuracy of 88.10% and a test accuracy of 86%. This model also exhibited a gap between training and testing accuracy, indicating overfitting.

• VGG16:

VGG16 had the lowest performance, with a training accuracy of 72% and a test accuracy of 65%. This significant performance drop suggests that VGG16 is less suitable for this specific task, likely due to its relatively shallow architecture compared to DenseNet and ResNet models.

IV. CONCLUSION

Among the evaluated models, DenseNet121 emerged as the most effective for classifying liver histopathology images, achieving a training accuracy of 98.21% and a test accuracy of 97%. Although it fell slightly short of the 96% test accuracy benchmark, its performance was significantly better than the other models. DenseNet121's dense connectivity pattern facilitated efficient learning and robust generalization, making it a strong candidate for this classification task. ISSN No:-2456-2165

ResNet models, particularly ResNet101 and ResNet152, showed promise but did not meet the benchmark accuracy. These models may benefit from further optimization and fine-tuning to improve their generalization capabilities.

VGG16's performance was the lowest, indicating that it may not be suitable for this specific task due to its architectural limitations.

Future work could explore further optimization of DenseNet121 and the ResNet models, as well as the application of advanced techniques such as data augmentation and ensemble learning to enhance model performance.

> Declarations

REFERENCES

- [1]. Chen, M., Zhang, B., Topatana, W. et al. Classification and mutation prediction based on histopathology H&E images in liver cancer using deep learning. npj Precis. Onc. 4, 14 (2020). https://doi.org/10.1038/s41698-020-0120-3
- [2]. Kiani, A., Uyumazturk, B., Rajpurkar, P. et al. Impact of a deep learning assistant on the histopathologic classification of liver cancer. npj Digit. Med. 3, 23 (2020). https://doi.org/10.1038/ s41746-020-0232-8Classification of multidifferentiated liver cancer pathological images based on deep learning attention
- [3]. Chen, C., Chen, C., Ma, M. et al. Classification of multi-differentiated liver cancer pathological images based on deep learning attention mechanism. BMC Med Inform Decis Mak 22, 176 (2022). https://doi.org/10.1186/s12911-022-01919-1
- [4]. Y. -S. Lin, P. -H. Huang and Y. -Y. Chen, "Deep Learning-Based Hepatocellular Carcinoma Histopathology Image Classification: Accuracy Versus Training Dataset Size," in IEEE Access, vol. 9, pp. 33144-33157, 2021, doi: 10.1109/ACCESS. 2021.3060765
- [5]. Du, L., Yuan, J., Gan, M. et al. A comparative study between deep learning and radiomics models in grading liver tumors using hepatobiliary phase contrast-enhanced MR images. BMC Med Imaging 22, 218 (2022). https://doi.org/10.1186/s12880-022-00946-8
- [6]. Sailer, Maria, Schil ler, Florian, Falk, Thorsten, Jud, Andreas, Arke Lang, Sven, Ruf, Juri and Mix, Michael. "Prediction of the histopathological tumor type of newly diagnosed liver lesions from standard abdominal computer tomography with a machinelearning classifier based on convolutional neural networks" Current Directions in Biomedical Engineering, vol. 7, no. 1, 2021, pp. 150-153. https://doi.org/10.1515/cdbme-2021-1032

[7]. Dong, X., Li, M., Zhou, P. et al. Fusing pre-trained convolutional neural networks features for multidifferentiated subtypes of liver cancer on histopathological images. BMC Med Inform Decis Mak 22, 122 (2022). https://doi.org/10.1186/s12911-022-01798-6

https://doi.org/10.38124/ijisrt/IJISRT24JUN1368

- [8]. Chen, W.-M., Fu, M., Zhang, C.-J., Xing, Q.-Q., Zhou, F., Lin, M.-J., Dong, X., Huang, J., Lin, S., Hong, M.-Z., Zheng, Q.-Z., & Pan, J.-S. (2022). Deep Learning-Based Universal Expert-Level Recognizing Pathological Images of Hepatocellular Carcinoma and Beyond. Frontiers in Medicine, 9, 853261. https://doi.org/10.3389/fmed.2022.853261
- [9]. Bhaskar, Nuthanakanti & Sasi Kiran, Jangala & Satyanarayan, Suma & Divya, Gaddam & Raju, Srujan & Kanthi, Murali & Patra, Raj. (2024). An approach for liver cancer detection from histopathology images using hybrid pre-trained models. TELKOMNIKA (Telecommunication Computing Electronics and Control). 22. 401-412. 10.12928/TELKOMNIKA.v22i2.25588.
- [10]. Kavitha, V.R., Hussain, F.B.J., Chillakuru, P., Shanmugam, P. (2024). Automated classification of liver cancer stages using deep learning on histopathological images. Traitement du Signal, Vol. 41, No. 1, pp. 373-381. https://doi.org/10.18280/ts. 410131
- [11]. Sridhar K, C K, Lai W-C, Kavin BP. Detection of Liver Tumour Using Deep Learning Based Segmentation with Coot Extreme Learning Model. Biomedicines. 2023; 11(3):800. https://doi.org/10. 3390/biomedicines11030800
- [12]. Rahman H, Bukht TFN, Imran A, Tariq J, Tu S, Alzahrani A. A Deep Learning Approach for Liver and Tumor Segmentation in CT Images Using ResUNet. Bioengineering (Basel). 2022;9(8):368. Published 2022 Aug 5. doi:10.3390/bioengineering 9080368
- [13]. C. Sun, A. Xu, D. Liu, Z. Xiong, F. Zhao and W. Ding, "Deep Learning-Based Classification of Liver Cancer Histopathology Images Using Only Global Labels," in IEEE Journal of Biomedical and Health Informatics, vol. 24, no. 6, pp. 1643-1651, June 2020, doi: 10.1109/JBHI.2019.2949837.
- [14]. da Nóbrega, R.V.M., Rebouças Filho, P.P., Rodrigues, M.B. et al. Lung nodule malignancy classification in chest computed tomography images using transfer learning and convolutional neural networks. Neural Comput & Applic 32, 11065– 11082 (2020). https://doi.org/10.1007/s00521-018-3895-1
- [15]. Albelwi, Saleh. (2022). Deep Architecture based on DenseNet-121 Model for Weather Image Recognition. International Journal of Advanced Computer Science and Applications. 13. 10.14569/ IJACSA.2022.0131065.

- [16]. Tammina, Srikanth. (2019). Transfer learning using VGG-16 with Deep Convolutional Neural Network for Classifying Images. International Journal of Scientific and Research Publications (IJSRP). 9. p9420. 10.29322/IJSRP.9.10.2019.p9420.
- [17]. Liang, Jiazhi. (2020). Image classification based on RESNET. Journal of Physics: Conference Series. 1634. 012110. 10.1088/1742-6596/1634/1/012110
- [18]. Li X, Shen X, Zhou Y, Wang X, Li T-Q (2020) Classification of breast cancer histopathological images using interleaved DenseNet with SENet (IDSNet). PLoS ONE 15(5): e0232127. https:// doi.org/10.1371/journal. pone.0232127
- [19]. Qiu C, Wang X, Batson SA, Wang B, Casiano CA, Francia G, Zhang J-Y. A Luminex Approach to Develop an Anti-Tumor-Associated Antigen Autoantibody Panel for the Detection of Prostate Cancer in Racially/Ethnically Diverse Populations. Cancers. 2023; 15(16):4064. https://doi.org/10.3390/ cancers15164064