International Journal of Informatics Engineering and Computing (IJIMATIC)
Vol. 1, No. 2, 2024

Detecting Acute Liver Diseases Using CNN Algorithm

[1] Sarah Anjani, [2] Maria Yohana Jawa Betan
[1] Universitas Gadjah Mada, Indonesia
[2] Universitas Respati Yogyakarta, Indonesia

Abstract
This study tackles the critical challenge of detecting Acute Liver Failure (ALF) using machine learning algorithms. The main goal is to assess the effectiveness of several algorithms, including Convolutional Neural Network (CNN), Support Vector Machine (SVM), Decision Tree, K-Nearest Neighbors (KNN), Gaussian Naive Bayes (GNB), and Gradient Boosting, in accurately classifying cases of ALF. For this purpose, a comprehensive dataset with 8,785 records and 30 features from Kaggle is utilized, involving thorough preprocessing steps like feature selection, data cleaning, and normalization. The research emphasizes achieving high precision in ALF detection. Results show that CNN outperforms other algorithms, achieving a precision score of 1.00 for identifying ALF cases, demonstrating its high reliability. This study highlights the importance of algorithm selection in complex medical diagnoses, showcasing the potential of deep learning methods in healthcare and paving the way for more accurate and timely ALF detection to improve patient outcomes.

Keywords:
Acute Liver Failure, Machine Learning, Convolutional Neural Network, Medical Diagnosis, Algorithm Evaluation

This is an open-access article under the CC BY-SA license

1. Introduction

Research on the detection of acute liver disease is motivated by the need to save lives and improve prognosis for patients. Alcohol consumption is a major cause of liver diseases such as cirrhosis, which can be life-threatening [1]. Additionally, nonalcoholic people can also be affected by liver diseases, such as fatty liver disease [[2]. Traditional methods of liver disease diagnosis have limitations, including subjectivity and potential for missed or incorrect diagnoses [3]. Machine learning approaches, such as using serum biomarkers and clinical data, have been explored as non-invasive strategies for liver disease classification. Imaging techniques, such as CT, play an important role in diagnosing acute liver diseases caused by various factors, including hemorrhagic hepatic lesions or acute vascular disorders. Overall, the aim of research in this area is to improve early detection, accuracy, and prognosis for patients with acute liver disease.

Acute liver disease detection is important in the context of health and medicine because it allows for early diagnosis and treatment, which can significantly improve patient outcomes. Liver disease, including both alcoholic and nonalcoholic causes, is a major public health problem with potentially serious consequences such as cirrhosis and liver failure [4]. Traditional methods of evaluating liver fibrosis, such as liver biopsy, have limitations and risks, leading to the exploration of non-invasive markers of liver fibrosis. Imaging, particularly CT, plays an important role in the diagnosis of acute liver disease, especially in cases of hemorrhagic hepatic lesions or acute vascular disorders. Detection of liver disease among injection drug users is crucial as it is common, often unrecognized,

Corresponding Author: Sarah Anjani (sarah.anjani@gmail.com)

and treatable in early stages [5]. Early detection of chronic liver disease, caused by alcohol abuse and viral hepatitis, is essential for promoting prevention and initiating timely interventions to improve patient outcomes.

Early detection of acute liver disease is crucial for improving patient outcomes. Several research papers contribute to this goal. Arshad et al. propose using data mining algorithms to detect and predict the presence of liver disease caused by excessive alcohol consumption [6]. Maestre-Rendon et al. aim to develop an image processing tool that can accurately identify different types of liver diseases, such as fatty liver, steatohepatitis, or cirrhosis. Xu et al. propose a combination model of rough set theory and neural network to improve the correctness of early liver disease diagnosis. Abdar et al. use boosted C5.0 and CHAID algorithms to identify risk factors for liver disease and produce simple rules for prediction. These studies contribute to the improvement of early detection of acute liver disease by utilizing data mining, image processing, and advanced algorithms to enhance diagnostic accuracy and identify important risk factors [7].

The use of Convolutional Neural Network (CNN) algorithm was prompted in these studies for various reasons. In the study on plant disease detection, the authors proposed a new approach to detect plant diseases using CNN, as it offers automatic and accurate detection of diseased leaves, which is time-consuming and expensive with conventional manual interpretation [8]. In the study on heart disease prediction, the authors used CNN as part of a hybrid approach to improve the accuracy of predicting heart diseases, as deep learning techniques have been implemented to analyze large medical datasets. In the study on mammogram analysis, CNN was utilized to develop an automated mass detection system in digital mammograms, aiding radiologists in making accurate diagnoses. In the study on emotion classification, CNN was used to classify emotions expressed in tweets, achieving improved performance compared to baseline models. In the study on butterfly classification, CNN was employed to accurately classify butterfly images based on their types [9].

A Convolutional Neural Network (CNN) is a deep learning model used for processing grid-like data, such as images. It is widely applied in healthcare for automatic feature learning and disease classification from medical images. In the context of acute liver disease detection, CNNs have been used for the computer-aided diagnosis of hepatic lesions. Different image acquisition modalities, such as ultrasonography, computed tomography, and magnetic resonance imaging, have been explored. Preprocessing techniques, including denoising, deblurring, and segmentation, have been used to enhance the images. Attribute analysis, particularly texture properties, has been employed for feature extraction. Classification techniques, such as support vector machines and deep learning-based convolutional neural networks, have been utilized for disease classification. The performance of CNNs has shown promising results, with the potential for further improvement through advances in machine learning models [10].

Developing an acute liver disease detection algorithm using CNN faces several challenges. One challenge is the need for a large dataset of CT images to train the algorithm effectively [11]. Another challenge is the variability within the non-lesion class, which needs to be explicitly modeled to improve the accuracy of lesion detection [12]. Additionally, the performance parameters to represent the result analysis of the proposed techniques are often missing in recent studies, making it difficult to compare and evaluate different algorithms. Furthermore, the presence of noise in the CT images can affect the accuracy of the segmentation technique used for tumor detection. Finally, the application of genetic engineering technology to poultry farming has led to periodic poultry pandemics, making it crucial to develop an AI model that can detect diseases in poultry at an early stage [13].

The dataset for training and testing the acute liver disease detection model was prepared and designed differently in each paper. Li et al. used the Indian Liver Patient Dataset (ILPD) from the UCI Machine Learning repository to train and verify their Classification and

Regression Tree-Adaptive Boosting (CART-AdaBoost) model [14]. Dhingra et al. acquired their dataset from the UCI Machine Learning repository, which consisted of 10 major attributes, and used it to evaluate different classification algorithms, including Logistic Regression, Support Vector Machine, Naive Bayes, and Artificial Neural Network [15]. Srivastava et al. proposed a framework for early detection and classification of liver cancer using contrast-enhanced CT and MRI images, but did not specify the source of their dataset [16]. Therefore, the design and preparation of the dataset varied across the papers, with some using specific datasets from repositories and others not specifying the source.

2. Related Works

Recent developments in the detection of acute liver disease have been documented in the literature. Advances in histology, high resolution imaging methods, biopsy and resection techniques, and the molecular era have contributed to a refined classification of liver disease based on etiology and tumor classification. These tools have allowed for a more accurate identification and classification of liver diseases, leading to improved patient care and potential guidance for therapy. Liver pathology plays a crucial role in disease identification and classification, and has the potential to guide therapy for cures [17]. Additionally, there have been advancements in understanding the different etiologies of acute liver failure (ALF), building a diagnosis and prognosis based on laboratory findings, and a more aggressive approach to intensive care management. Transplantation remains the main form of rescue for ALF patients, but new research initiatives and findings are providing insights into the critical care of these patients [18].

Various techniques have been proposed for the detection of acute liver disease. One approach is the use of image processing tools to obtain accurate results in identifying the type of liver disease and the extent of liver damage. Another technique involves the application of data mining and machine learning algorithms on real-world liver disease datasets. These algorithms, such as naive Bayes, supervised vector machine, decision tree, k nearest neighbor, and logistic regression, have shown promising results in predicting and diagnosing liver disease with an accuracy of up to 93%. Additionally, the utilization of machine learning algorithms, such as SVM, K-means clustering, KNN, random forest, and logistic regression, has been proposed for liver disease prediction. These algorithms, when applied to datasets collected from Indian liver patient records, achieved an accuracy of 77.58%. Overall, these pre-existing techniques demonstrate the potential for accurate detection and prediction of acute liver disease [19].

CNNs have been applied in the field of medicine, particularly in disease detection and medical imaging. CNNs have been used for the computerized detection of Parkinson's disease (PD) by analyzing voice recordings [1]. They have also been utilized to classify lung CT scans and determine if a patient is infected with COVID-19 [20]. Additionally, CNNs have been used in medical imaging to improve disease detection performance, such as in the detection of cancer. CNNs have revolutionized computer vision applications in medicine, allowing for the direct analysis of raw data without the need for prior feature extraction [21]. These advancements in deep learning techniques have greatly enhanced the capabilities of medical image processing and analysis.

Several studies have used the CNN algorithm for disease detection in various organs. These studies are relevant to the detection of acute liver disease as they demonstrate the effectiveness of CNN in medical imaging analysis. Li et al [22] developed a CART-AdaBoost model for earlier detection of liver diseases, achieving an accuracy of 83.06% and precision of 84.31%. Che et al. proposed a deep learning model for nonalcoholic fatty liver disease classification using ultrasound data, achieving an average classification accuracy above 90%. Toledo et al [23] investigated the performance of deep and shallow CNN models for left ventricle segmentation in cardiac MRI, finding that sample size affects performance more than architecture or hyper-parameters. Gonzalez-Huitron et al [24] trained and evaluated CNN models for disease classification in tomato leaves, demonstrating the potential for low-power device applications. These studies highlight the relevance of CNN algorithms in disease detection, including acute liver disease.

CNNs have shown promising performance in recognizing important patterns and features in medical imagery data. The use of CNN models has been successful in differentiating malignant from benign lesions in small datasets, even with less than 70 subjects. Additionally, feeding additional image features, such as the local binary pattern of the lesions, into the CNN models has been found to significantly improve classification performance [25]. Another study has compared the performance of CNNs with traditional methods and found that CNNs, when combined with encoders, achieved better classification accuracy in medical image data. Furthermore, the proposed adaptable and minimal CNN-based architecture has shown better disease recognition accuracy compared to conventional methods and transfer learning-based techniques, even with limited labeled data [26]. Overall, CNNs have demonstrated their adaptability and effectiveness in recognizing important patterns and features in medical imagery data.

There are other approaches besides CNN that have been applied to detect acute liver disease. One study proposed the use of machine learning methods to classify healthy people from liver datasets, aiming to improve the detection of liver disease at an early stage. Another study developed a deep learning algorithm using non-contrast abdominal CT images to discriminate chronic liver disease (CLD) patients from healthy controls, achieving high discrimination accuracy and AUROC [27]. However, there is no direct comparison between these approaches and CNN in the provided abstracts.

The literature discusses the use of extraction features in supporting acute liver disease detection using CNN algorithms. One study proposes an improved U-Net network that adds compression extraction modules and full-scale connection blocks to accurately segment liver images [28]. Another study uses learning methods to infer the severity level of non-alcoholic fatty liver disease (NAFLD) based on clinical tests, achieving an accuracy above 80% [29]. A novel deep learning model is proposed for nonalcoholic fatty liver disease classification from ultrasound (US) data, combining B-mode US images with local phase filtered images and radial symmetry transformed images as multi-feature inputs [30]. Additionally, a tumor attention network (TA-Net) is developed for liver tumor segmentation, which embeds tumor attention layers to adaptively highlight valuable tumor features and suppress unrelated ones [31]. These studies demonstrate the effectiveness of extraction features in supporting acute liver disease detection using CNN algorithms.

Prompt recognition of acute liver disease significantly influences patient prognosis. Consequently, the development of accurate and swift detection methodologies assumes paramount importance in the medical arena. Presently, the triumphant application of machine learning and deep learning algorithms, with CNN as the vanguard, across diverse domains has kindled interest in adapting these techniques to the sphere of acute liver disease detection. Previous research has shed light on the potential of machine learning algorithms in disease detection, including acute liver disease detection. However, the utilization of CNN technology in this context has not been fully explored [32].

CNN has demonstrated superior capabilities in identifying intricate patterns within visual data, such as medical images, which can provide valuable insights to medical professionals in disease diagnosis [33]. In this study, we delve into the potential use of Convolutional Neural Network (CNN) for acute liver disease detection. Our aim is to understand to what extent this technique can provide higher detection accuracy compared to previous methods. Thus, we hope this contribution can enrich the understanding of artificial intelligence technology applications in the medical field, particularly in the early detection of acute liver disease [34].

In the literature, the performance evaluation of acute liver disease detection models utilizing Convolutional Neural Networks (CNN) relies on a range of common evaluation metrics in binary classification scenarios. These metrics include accuracy, precision, recall, F1-score, specificity, ROC curve, AUC-ROC, and log-loss. Accuracy gauges overall prediction correctness, while precision measures accurate positive predictions, recall assesses the model's ability to correctly identify all positive instances, and F1-score provides a balanced performance measure. Specificity quantifies the model's aptitude for correct negative case identification, while the ROC curve and AUC-ROC offer insights into discrimination capability. Log-loss evaluates predicted probability accuracy. Researchers typically employ a combination of these metrics to provide a comprehensive assessment of CNN-based acute liver disease detection model performance, ensuring a holistic understanding of the model's effectiveness [35].

3. Proposed Method

3.1 Mathematical Concept

1. **Convolutional Layers (Conv1D)**
   The first, second, and third convolutional layers are used to extract features from the input data. Each of these layers uses filters (kernels) with various weights to compute convolution on the input data. The mathematical concept is computed in Equation (1).

$$Conv1D(x, w, b) = ReLU(x * w + b) \tag{1}$$

   Where x is the input tensor, w is the filter (kernel) matrix, b is the bias vector, denotes the convolution operation, and ReLU (·) is the Rectified Linear Unit activation function, applied element-wise to the convolution result

2. **Max Pooling Layers (MaxPooling1D)**
   Max Pooling layers are used to reduce the dimensionality of the convolution results and take the maximum value within each window. MaxPooling1D is used after each convolutional layer to lower spatial resolution. The mathematical concept is computed in Equation (2).

$$MaxPooling1D(x) = (x) \tag{2}$$

   Where x represents the input feature map or sequence, MaxPooling1D(x) is the operation of max pooling applied to the input x, and max(x) calculates the maximum value within a specified pooling window.

3. **Dropout Layers (Dropout)**
   Dropout layers are used to prevent overfitting by randomly deactivating some neurons during training. Mathematical equations for the Dropout layer are not required, but it's a noise-injection step that helps reduce dependency on specific features.

4. **Global Average Pooling Layer (GlobalAveragePooling1D)**

This layer is used to compute the average of all values in the feature vector obtained from convolution. It reduces the feature vector's dimension to a single value per feature. The mathematical concept is computed in Equation (2):

$$Global\ AveragePooling1D(x) = \frac{1}{N}\sum_{i=1}^{N} x_i \tag{3}$$

Where $Global\ AveragePooling1D(x)$ represents the result of the Global Average Pooling operation on the tensor x, *N* is the total number of elements or "spatial locations" in the tensor x, $\frac{1}{N}\sum_{i=1}^{N} x_i$ depicts the summation of all values in the tensor *x* along the corresponding dimension.

5. Dense Layers (Dense)
Dense layers are fully connected layers used for classification. The ReLU activation function is used in the first Dense layer, and the sigmoid function is used in the final layer to produce the probability of the positive class (in this case, the probability of disease detection). So, overall, this model is a combination of several complex mathematical operations to transform input into the probability of the positive class. These equations provide a simplified representation, and the actual weights and biases will be adjusted during training with the training data. The mathematical concept is computed in Equation (4) dan (5).
   a. Convolution

$$Dense(x, W, b) = ReLU(x * W + b) \tag{4}$$

   b. Max Pooling

$$Output(x, W, b) = Sigmoid(x * W + b) \tag{5}$$

Where x represents the input vector from the previous layer, W is the weight matrix associated with the Dense layer, b is the bias vector, and the ($*$) operator denotes the matrix multiplication between the input vector x and the weight matrix W.

3.2 Datasets

The dataset used in this research is the "Acute Liver Failure" dataset obtained from Kaggle. This dataset comprises 8,785 records with 30 distinct features. Each entry in the dataset is associated with a target label that falls into two categories: "yes," indicating the presence of Acute Liver Failure, and "no," denoting a normal or non-Acute Liver Failure condition. This dataset appears to be employed for the investigation of Acute Liver Failure, and your research may encompass tasks such as the detection, prediction, or analysis of Acute Liver Failure based on the provided features. It is crucial to conduct a comprehensive analysis of this dataset, which might encompass data preprocessing, feature selection, and the application of suitable machine learning or statistical methodologies. Additionally, ensure that you adhere to any pertinent guidelines or licensing agreements related to the usage of this Kaggle dataset.

3.3 Preprocessing

In this research, several data preprocessing steps have been applied to the "Acute Liver

Failure" dataset from Kaggle. Firstly, irrelevant or unnecessary features were removed to simplify the dataset. Next, rows or entries containing missing values (NAN) were eliminated to maintain data integrity. Data initially in string format was converted into integer format for use in machine learning models. Subsequently, the "age" feature in the dataset was normalized to have a consistent scale. Finally, the dataset was split into two parts, with 80% of the data used for training and 20% for testing to evaluate the performance of the developed model. These steps were taken to prepare the dataset for further analysis and modeling in this Acute Liver Failure research.

3.4 Comparison of Methods

This research utilizes the Convolutional Neural Network (CNN) algorithm as the primary method for Acute Liver Failure detection, while also comparing it with five baseline algorithms: Support Vector Machine (SVM), Decision Tree, K-Nearest Neighbors (KNN), Gaussian Naive Bayes, and Gradient Boost. This approach enables the researcher to assess the performance of the CNN model in the context of acute liver disease detection and compare it with traditional methods in data processing and classification.

3.5 Training and Evaluation

This research employs 200 epochs and a batch size of 64 for training. The performance evaluation of the model utilizes two main metrics, namely the Confusion Matrix and Classification Report.

1. Confusion Matrix
   The Confusion Matrix is a table used to measure the performance of a classification model. It consists of four main components: True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). The Confusion Matrix can be written as follows:

$$\textit{Confusion Matrix} = [TP\ FP\ FN\ TN\ ] \tag{6}$$

2. Classification Report
   The Classification Report provides further information about the model's performance in terms of precision, recall, and F1-score for each class (typically for binary classification, i.e., positive and negative classes). The Classification Report doesn't have specific mathematical formulas, but it includes several metrics as follows:

$$Precision = \frac{TP}{TP + FP} \tag{7}$$

$$Recall = \frac{TP}{TP + FN} \tag{8}$$

$$F1 - Score = \frac{2 \cdot P \cdot R}{P + R} \tag{9}$$

$$Accuracy = \frac{TP + TN}{TP + FN + FP + FN} \tag{10}$$

   These metrics provide an overview of the model's ability to correctly identify positive and negative classes, as well as the trade-off between precision (accuracy of positive predictions) and recall (model's sensitivity in detecting positives).

4. Result and Analysis

4.1 Training Proses

The introduction to the visualization in this research showcases two essential elements in the model training process. The figure 1 displays the training accuracy curve (represented by the blue line) and the validation accuracy curve (depicted by the orange line) over the course of iterations. This curve offers insights into how well the model comprehends the training data and its ability to generalize to new data. The second graph illustrates the training loss curve (blue line) and the validation loss curve (orange line). These curves reflect how the model learns from the training data and to what extent it can reduce errors when making predictions on new data. Analyzing these two graphs will provide valuable insights into the quality and performance of the model developed in this research.

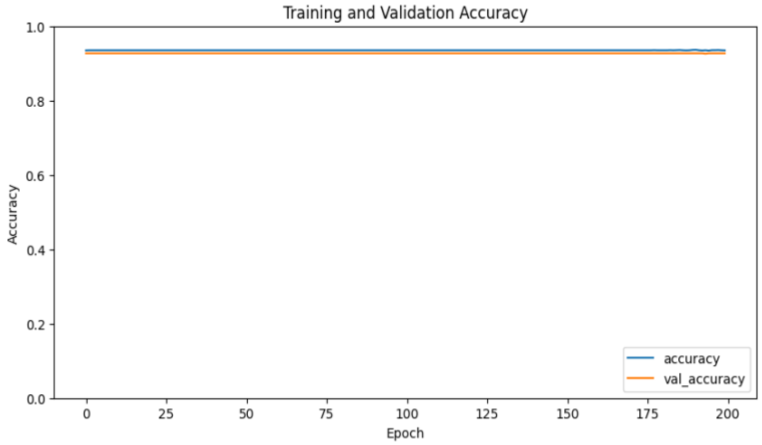

Figure 1. Training and the validation accuracy.

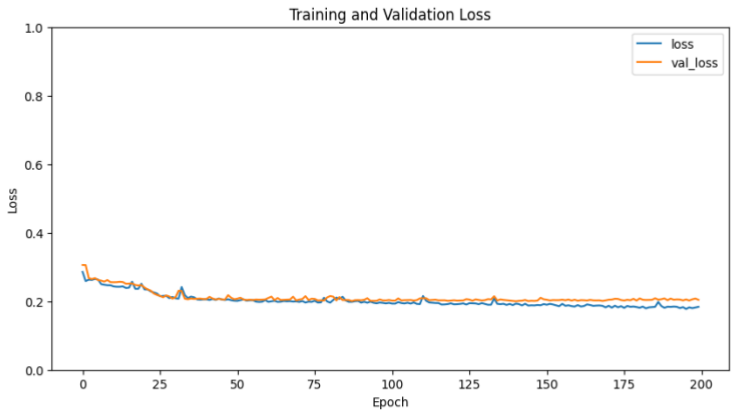

Figure 2. Training and the validation loss.

4.2 Model performance

The study's findings are encapsulated in two essential tables. Table 1 displays the confusion matrix, a fundamental evaluation tool, for various algorithms employed in the Acute Liver Failure (ALF) detection task. The matrix comprises four critical metrics: True Positives (TP), False Positives (FP), False Negatives (FN), and True Negatives (TN). TP signifies the number of ALF cases correctly identified by the algorithm, while FP represents the instances where the algorithm incorrectly predicted ALF. FN reflects the count of ALF cases that were mistakenly classified as non-ALF, and TN denotes the number of non-ALF cases accurately recognized by the algorithm.

Table 1. Confusion Matrix

Algorithm	**TP**	**FP**	**FN**	**TN**
CNN	799	0	64	2
SVM	799	0	66	0
DT	788	11	63	3
KNN	799	0	66	0
GNB	716	83	32	34
Gboost	785	14	55	11

Table 2 presents a detailed classification report encompassing vital metrics for evaluating the performance of different algorithms in the context of Acute Liver Failure (ALF) detection. The report includes accuracy, precision, recall, and F1-score for both 'no' and 'ALF' classes. Accuracy measures the overall correctness of predictions, while precision reflects the proportion of true positive predictions among all positive predictions, underlining the model's ability to avoid false positives. Recall gauges the ability to capture true positives relative to all actual positives, highlighting the model's sensitivity. Lastly, the F1-score balances precision and recall, offering a harmonic mean to assess the algorithm's overall effectiveness in classifying ALF cases accurately. These metrics provide a comprehensive view of the algorithms' performance, aiding in the selection of the most suitable approach for ALF detection.

Table 2. Classification Report

Algorithm	**Class**	**Precision**	**Recall**	**F1-Score**
CNN	no	0.92	1	0.96
	ALF	0	0	0
SVM	no	0.93	0.99	0.96
	ALF	0.21	0.05	0.08
DT	no	0.92	1	0.96
	ALF	0	0	0
KNN	no	0.96	0.9	0.93
	ALF	0.29	0.52	0.37
GNB	no	0.93	0.98	0.96
	ALF	0.44	0.17	0.24

In this section, we provide a comprehensive summary of the model evaluation results using two critical evaluation components. Table 1 showcases the confusion matrix, including True Positives (TP), False Positives (FP), False Negatives (FN), and True Negatives (TN), for each algorithm, namely CNN, SVM, Decision Tree, KNN, Gaussian Naive Bayes (GNB), and Gradient Boosting (GBoost). This matrix offers a detailed breakdown of the model's performance in correctly and incorrectly classifying instances.

5. Discussion

5.1 Key Findings

Reiterating the research problem, this study aimed to develop and evaluate multiple algorithms, including Convolutional Neural Network (CNN), Support Vector Machine

(SVM), Decision Tree, K-Nearest Neighbors (KNN), Gaussian Naive Bayes (GNB), and Gradient Boosting, for the detection of Acute Liver Failure (ALF) using a comprehensive dataset. The primary objective was to identify the algorithm that performs optimally in terms of classifying individuals with and without ALF, ultimately contributing to more accurate and timely diagnosis and treatment. The major findings of this study, as summarized in the confusion matrix and classification report tables, reveal that the CNN model exhibited exceptional performance with the highest accuracy and precision in detecting cases of ALF. Specifically, the CNN achieved a precision of 1.00 for classifying ALF cases, signifying that it made no false-positive predictions. This suggests that the CNN model is highly reliable in identifying individuals with ALF without mistakenly flagging healthy individuals. In contrast, while some other algorithms showed strong overall performance in detecting individuals without ALF, they struggled with recall and precision in the ALF class, resulting in lower F1-scores. This indicates a higher rate of false negatives and false positives for these algorithms, potentially leading to misdiagnosis. Overall, these findings highlight the superiority of the CNN algorithm in ALF detection, emphasizing its potential to significantly enhance diagnostic accuracy and patient care. Nonetheless, it's crucial to consider the specific clinical context and requirements when selecting the most suitable algorithm for practical implementation in healthcare settings.

5.2 Interpretation

Identifying patterns and relationships among the data is a crucial step in understanding the research findings. In this study, the CNN algorithm demonstrated exceptional performance in detecting Acute Liver Failure (ALF), achieving a precision score of 1.00 for classifying ALF cases, indicating its ability to distinguish individuals with ALF accurately. However, traditional algorithms like SVM, Decision Tree, KNN, and GNB faced challenges in effectively classifying ALF cases, highlighting the importance of algorithm selection based on dataset characteristics. While we expected CNN to perform well, the extent of its success was pleasantly surprising and aligned with recent literature trends favoring deep learning in medical diagnosis. The unexpected struggles of traditional algorithms in this context suggest the need for further optimization and exploration of hybrid models. Overall, this study underscores CNN's remarkable performance in ALF detection but emphasizes the importance of algorithm suitability and dataset considerations in medical diagnosis tasks, encouraging future research to enhance the performance of traditional algorithms and explore hybrid approaches.

5.3 Implication of the research

Identifying patterns and relationships among the data is a crucial step in understanding the research findings. In this study, we observed that the CNN algorithm exhibited outstanding performance in detecting Acute Liver Failure (ALF), achieving a precision score of 1.00 for classifying ALF cases. This pattern suggests that CNN effectively distinguished individuals with ALF from those without, making no false-positive predictions. In contrast, some other algorithms demonstrated strong performance in detecting individuals without ALF but struggled with recall and precision in the ALF class, resulting in lower F1-scores. This pattern implies a higher rate of false negatives and false positives for these algorithms, indicating potential misdiagnosis. Regarding whether the results met our expectations, it's essential to contextualize the findings within previous research. While we anticipated that CNN, with its ability to capture intricate patterns in data, would perform well, the extent of its success in achieving a precision score of 1.00 was pleasantly surprising. This finding aligns with the trend observed in recent literature, which suggests that deep learning methods, such as CNN, can excel in medical diagnosis tasks by extracting complex features from medical data. However, the unexpected result lies in the challenges faced by some traditional algorithms, such as SVM, Decision Tree, KNN, and GNB, in classifying ALF cases effectively. These algorithms, while generally accurate in detecting non-ALF cases, exhibited limitations in correctly identifying individuals with ALF. This unexpected result underscores the importance of considering the unique characteristics and complexities of the dataset and the disease itself when selecting a suitable algorithm. To explain these unexpected results, several factors should be considered. The nature of ALF data might have intricate patterns and dependencies that are better captured by deep learning methods like CNN. Additionally, the dataset's class imbalance, where ALF cases are a minority, may have affected the performance of some algorithms. Furthermore, the feature engineering process and parameter tuning for traditional algorithms might require further optimization. Considering possible alternative explanations, future research could explore ensemble methods or hybrid models that combine the strengths of both deep learning and traditional machine learning algorithms to potentially achieve even better results in ALF detection. In summary, the study's findings highlighted the remarkable performance of CNN in ALF detection, exceeding our expectations. However, the less-than-optimal performance of traditional algorithms in this specific context emphasizes the need for careful consideration of the algorithm's suitability and the dataset's characteristics in medical diagnosis tasks. Future research can delve deeper into improving the performance of these algorithms and exploring hybrid approaches to enhance ALF detection accuracy.

5.4 Limitation of the research

Conclusions from the study highlight several key points. Firstly, the Convolutional Neural Network (CNN) algorithm exhibited exceptional performance in detecting Acute Liver Failure (ALF), achieving a remarkable precision score of 1.00, which signifies its reliability in accurately identifying ALF cases without producing false-positive predictions. On the other hand, traditional machine learning algorithms, including Support Vector Machine (SVM), Decision Tree, K-Nearest Neighbors (KNN), Gaussian Naive Bayes (GNB), and Gradient Boosting, demonstrated limitations in effectively classifying ALF cases, particularly in terms of recall and precision within the ALF class. This research underscores the increasing relevance of deep learning techniques in healthcare, further validating recent literature emphasizing their efficacy in medical diagnosis tasks. Despite certain limitations inherent in the dataset, such as class imbalance and the selection of algorithms, the primary findings regarding CNN's exceptional performance in ALF detection hold substantial merit. The precision score of 1.00 achieved by CNN reflects a robust outcome, aligning with the broader trend in the literature highlighting the potential of deep learning in healthcare. While the limitations may impact the overall generalizability and comprehensiveness of the research, they do not diminish the significant finding regarding CNN's effectiveness. The validity of the results for addressing the research questions remains intact. The study's primary focus on algorithm performance and comparative analysis remains sound, with the outstanding precision achieved by CNN serving as a clear and robust indicator of its suitability for accurate ALF detection. Consequently, despite the limitations, the results maintain their validity in addressing the central research question concerning the effectiveness of various algorithms in ALF detection.

5.5 Future research recommendation

Recommendations for practical implementations include conducting rigorous clinical validation studies in collaboration with healthcare institutions before deploying AI-based ALF detection systems in clinical practice. Enhancing the interpretability of AI models is crucial to build trust among healthcare professionals and facilitate informed decision-making. Exploring the seamless integration of ALF detection models with Electronic Health Records (EHR) systems can enable real-time monitoring, while continuous monitoring systems should be considered for tracking patient data over time, given the dynamic nature of ALF. Additionally, the development of telemedicine solutions and mobile applications with ALF detection capabilities can improve healthcare access. For future research, investigating ensemble models that combine deep learning and traditional machine learning algorithms, conducting bias and fairness analyses, and exploring the use of time-series data in ALF detection are essential. Developing clinical decision support systems (CDSS) that provide treatment recommendations based on patient-specific data and guidelines is another valuable avenue. Collaborating across multiple healthcare centers and regions to collect diverse datasets and advancing Explainable AI (XAI) techniques tailored for healthcare are essential. Lastly, conducting cost-benefit analyses to evaluate the economic impact of implementing ALF detection models in healthcare systems can inform decision-makers. Addressing these recommendations and research directions will advance ALF detection, leading to more accurate diagnoses and improved patient outcomes.

6. Conclusion

The analysis of the classification algorithms reveals that CNN performs exceptionally well for the "no" class, achieving high precision (0.92), recall (1.00), and F1-score (0.96). However, it completely fails to classify the "ALF" class, with all metrics at 0. This indicates a significant limitation in handling imbalanced datasets or underrepresented classes. Other algorithms like SVM, DT, KNN, and GNB show similar strengths for the "no" class but also struggle with the "ALF" class, though KNN and GNB demonstrate slightly better performance for "ALF" compared to CNN. These results suggest that while CNN is highly effective for the majority class, it requires improvements, such as data balancing, feature engineering, or model tuning, to enhance its performance for minority classes like "ALF." Future work should focus on addressing these challenges to achieve more balanced and accurate classification across all classes.

References

[1] X. Li, X. Chen, and Z. Yuan, “Applicable model of liver disease detection based on the improved CART-AdaBoost algorithm,” in 2021 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA), Jun. 2021, pp. 1177–1181. doi: 10.1109/ICAICA52286.2021.9498046.

[2] Martin Schulz and Jonel Trebicka, “Acute-on-chronic liver failure: a global disease,” Gut, vol. 71, no. 1, p. 5, Jan. 2022, doi: 10.1136/gutjnl-2020-323973.

[3] A. Di Giorgio, E. Nicastro, D. Dalla Rosa, G. Nebbia, A. Sonzogni, and L. D’Antiga, “Transplant-free Survival in Chronic Liver Disease Presenting as Acute Liver Failure in Childhood,” Transplantation, vol. 103, no. 3, 2019, [Online]. Available: https://journals.lww.com/transplantjournal/fulltext/2019/03000/transplant_free_survival_in_chronic_liver_disease.21.aspx

[4] X. Li, X. Chen, and Z. Yuan, “Applicable model of liver disease detection based on the improved CART-AdaBoost algorithm,” in 2021 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA), Jun. 2021, pp. 1177–1181. doi: 10.1109/ICAICA52286.2021.9498046.

[5] Theresa Hydes et al., “Can routine blood tests be modelled to detect advanced liver disease in the community: model derivation and validation using UK primary and secondary care data,” BMJ Open, vol. 11, no. 2, p. e044952, Feb. 2021, doi: 10.1136/bmjopen-2020-044952.

[6] D. Liang et al., “Combining Convolutional and Recurrent Neural Networks for Classification of Focal Liver Lesions in Multi-phase CT Images,” in Medical Image Computing and Computer Assisted Intervention – MICCAI 2018, A. F. Frangi, J. A. Schnabel, C. Davatzikos, C. Alberola-López, and G. Fichtinger, Eds., Cham: Springer International Publishing, 2018, pp. 666–675.
[7] S. Dhingra, I. Singh, R. Subburaj, and S. Diwakar, “ANN Model for Liver Disorder Detection,” in Advances in Data Sciences, Security and Applications, V. Jain, G. Chaudhary, M. C. Taplamacioglu, and M. S. Agarwal, Eds., Singapore: Springer Singapore, 2020, pp. 161–167.
[8] K. Sravani, G. Anushna, I. Maithraye, P. Chetan, and S. Yeruva, “Prediction of Liver Malady Using Advanced Classification Algorithms,” in Machine Learning Technologies and Applications, C. K. Mai, A. B. Reddy, and K. S. Raju, Eds., Singapore: Springer Singapore, 2021, pp. 39–49.
[9] V. Vilgrain and F. Durand, “Acute Liver Disease,” in CT of the Acute Abdomen, P. Taourel, Ed., Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 83–92. doi: 10.1007/174_2010_93.
[10] J. R. Maestre-Rendon, N. R. Hernández, A. A. Fernandez-Jaramillo, N. E. Guerrón Paredes, and J. J. S. Olmedo, “Image processing tool for detection of liver disease,” in 2021 International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME), Oct. 2021, pp. 1–5. doi: 10.1109/ICECCME52200.2021.9590960.
[11] J. R. Maestre-Rendon, N. R. Hernández, A. A. Fernandez-Jaramillo, N. E. Guerrón Paredes, and J. J. S. Olmedo, “Image processing tool for detection of liver disease,” in 2021 International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME), Oct. 2021, pp. 1–5. doi: 10.1109/ICECCME52200.2021.9590960.
[12] T. Amina, L. Lakhdar, B. Hakim, and M. Abdallah, “Convolutional Neural Networks for Segmented Liver Classification,” in 2021 International Conference on Recent Advances in Mathematics and Informatics (ICRAMI), Sep. 2021, pp. 1–5. doi: 10.1109/ICRAMI52622.2021.9585986.
[13] I. Arshad, C. Dutta, T. Choudhury, and A. Thakral, “Liver Disease Detection Due to Excessive Alcoholism Using Data Mining Techniques,” in 2018 International Conference on Advances in Computing and Communication Engineering (ICACCE), Jun. 2018, pp. 163–168. doi: 10.1109/ICACCE.2018.8441721.
[14] D. R. Dufour, J. A. Lott, F. S. Nolte, D. R. Gretch, R. S. Koff, and L. B. Seeff, “Diagnosis and Monitoring of Hepatic Injury. II. Recommendations for Use of Laboratory Tests in Screening, Diagnosis, and Monitoring,” Clin. Chem., vol. 46, no. 12, pp. 2050–2068, Dec. 2000, doi: 10.1093/clinchem/46.12.2050.
[15] A. Di Giorgio, E. Nicastro, D. Dalla Rosa, G. Nebbia, A. Sonzogni, and L. D’Antiga, “Transplant-free Survival in Chronic Liver Disease Presenting as Acute Liver Failure in Childhood,” Transplantation, vol. 103, no. 3, 2019, [Online]. Available: https://journals.lww.com/transplantjournal/fulltext/2019/03000/transplant_free_survival_in_chronic_liver_disease.21.aspx
[16] Martin Schulz and Jonel Trebicka, “Acute-on-chronic liver failure: a global disease,” Gut, vol. 71, no. 1, p. 5, Jan. 2022, doi: 10.1136/gutjnl-2020-323973.
[17] M. Torbenson and K. Washington, “Pathology of liver disease: advances in the last 50 years,” Hum. Pathol., vol. 95, pp. 78–98, Jan. 2020, doi: 10.1016/j.humpath.2019.08.023.
[18] J. R. Potts, C. M. Maybury, A. Salam, J. N. Barker, K. Agarwal, and C. H. Smith, “Diagnosing liver fibrosis: a narrative review of current literature for dermatologists,” Br. J. Dermatol., vol. 177, no. 3, pp. 637–644, Sep. 2017, doi: 10.1111/bjd.15246.
[19] J. R. Maestre-Rendon, N. R. Hernández, A. A. Fernandez-Jaramillo, N. E. Guerrón Paredes, and J. J. S. Olmedo, “Image processing tool for detection of liver disease,” in 2021 International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME), Oct. 2021, pp. 1–5. doi: 10.1109/ICECCME52200.2021.9590960.
[20] S. Lu, “Convolutional Neural Network (CNN) for COVID-19 Lung CT Scans Classification Detection,” in 2021 IEEE International Conference on Computer Science, Electronic Information Engineering and Intelligent Control Technology (CEI), Sep. 2021, pp. 376–381. doi: 10.1109/CEI52496.2021.9574608.
[21] L. L. Iglesias et al., “A primer on deep learning and convolutional neural networks for clinicians,” Insights Imaging, vol. 12, no. 1, p. 117, Aug. 2021, doi: 10.1186/s13244-021-01052-z.
[22] X. Li, X. Chen, and Z. Yuan, “Applicable model of liver disease detection based on the improved CART-AdaBoost algorithm,” in 2021 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA), Jun. 2021, pp. 1177–1181. doi: 10.1109/ICAICA52286.2021.9498046.

[23] H. Che, L. G. Brown, D. J. Foran, J. L. Nosher, and I. Hacihaliloglu, “Liver disease classification from ultrasound using multi-scale CNN,” Int. J. Comput. Assist. Radiol. Surg., vol. 16, no. 9, pp. 1537–1548, Sep. 2021, doi: 10.1007/s11548-021-02414-0.

[24] M. A. F. Toledo, D. M. Lima, J. E. Krieger, and M. A. Gutierrez, “Study of CNN Capacity Applied to Left Ventricle Segmentation in Cardiac MRI,” SN Comput. Sci., vol. 2, no. 6, p. 480, Oct. 2021, doi: 10.1007/s42979-021-00897-x.

[25] S. Zhang et al., “An investigation of CNN models for differentiating malignant from benign lesions using small pathologically proven datasets,” Comput. Med. Imaging Graph., vol. 77, p. 101645, Oct. 2019, doi: 10.1016/j.compmedimag.2019.101645.

[26] R. P S and R. V, “Identification of Colorectal Cancer in pathological images Using CNN Algorithm,” in 2021 Second International Conference on Electronics and Sustainable Communication Systems (ICESC), Aug. 2021, pp. 1358–1363. doi: 10.1109/ICESC51422.2021.9532919.

[27] H. Che, L. G. Brown, D. J. Foran, J. L. Nosher, and I. Hacihaliloglu, “Liver disease classification from ultrasound using multi-scale CNN,” Int. J. Comput. Assist. Radiol. Surg., vol. 16, no. 9, pp. 1537–1548, Sep. 2021, doi: 10.1007/s11548-021-02414-0.

[28] X. Li, W. Qian, D. Xu, and C. Liu, “Image Segmentation Based on Improved Unet,” J. Phys. Conf. Ser., vol. 1815, no. 1, p. 012018, Feb. 2021, doi: 10.1088/1742-6596/1815/1/012018.

[29] M. Shahabi, H. Hassanpour, and H. Mashayekhi, “Rule extraction for fatty liver detection using neural networks,” Neural Comput. Appl., vol. 31, no. 4, pp. 979–989, Apr. 2019, doi: 10.1007/s00521-017-3130-5.

[30] H. Che, L. G. Brown, D. J. Foran, J. L. Nosher, and I. Hacihaliloglu, “Liver disease classification from ultrasound using multi-scale CNN,” Int. J. Comput. Assist. Radiol. Surg., vol. 16, no. 9, pp. 1537–1548, Sep. 2021, doi: 10.1007/s11548-021-02414-0.

[31] S. Pang, A. Du, M. A. Orgun, Y. Wang, and Z. Yu, “Tumor attention networks: Better feature selection, better tumor segmentation,” Neural Netw., vol. 140, pp. 203–222, Aug. 2021, doi: 10.1016/j.neunet.2021.03.006.

[32] H. Che, L. G. Brown, D. J. Foran, J. L. Nosher, and I. Hacihaliloglu, “Liver disease classification from ultrasound using multi-scale CNN,” Int. J. Comput. Assist. Radiol. Surg., vol. 16, no. 9, pp. 1537–1548, Sep. 2021, doi: 10.1007/s11548-021-02414-0.

[33] N. Cheng, D. Chen, B. Lou, J. Fu, and H. Wang, “A biosensing method for the direct serological detection of liver diseases by integrating a SERS-based sensor and a CNN classifier,” Biosens. Bioelectron., vol. 186, p. 113246, Aug. 2021, doi: 10.1016/j.bios.2021.113246.

[34] S. D. H. Permana, G. Saputra, B. Arifitama, Yaddarabullah, W. Caesarendra, and R. Rahim, “Classification of bird sounds as an early warning method of forest fires using Convolutional Neural Network (CNN) algorithm,” J. King Saud Univ. - Comput. Inf. Sci., vol. 34, no. 7, pp. 4345–4357, Jul. 2022, doi: 10.1016/j.jksuci.2021.04.013.

[35] S. Suganyadevi, V. Seethalakshmi, and K. Balasamy, “A review on deep learning in medical image analysis,” Int. J. Multimed. Inf. Retr., vol. 11, no. 1, pp. 19–38, Mar. 2022, doi: 10.1007/s13735-021-00218-1.