International Journal of Engineering Technology and Management Science
Website: ijetms.in Issue: 2 Volume No.9 March - April – 2025
DOI:10.46647/ijetms.2025.v09i02.061 ISSN: 2581-4621

Liver Disease Prediction using GA features selection, Social Spider Optimization and CNN Classification

Gopi Krishna T L[1],Mrs.P.Maneesha[2]
[1]*M.Tech Student, Department of CSE, Golden Valley Integrated Campus (GVIC), Madanapalli, Andhra Pradesh, India*
[2]*Associate Professor, Department of CSE, Golden Valley Integrated Campus (GVIC), Madanapalli, Andhra Pradesh, India*

ABSTRACT
Liver disease prediction is a critical task in healthcare, enabling early diagnosis and intervention to improve patient outcomes. In this study, we propose a novel approach integrating genetic algorithm (GA) feature selection, social spider optimization (SSO), and convolutional neural network (CNN) classification for accurate and efficient prediction of liver diseases.
Keywords: Liver Diseases, Nature Inspired Algorithms, SSA, ResNet50, VGG16, GoogleNet, Liver Image Classification

I.INTRODUCTION

For a variety of illnesses, magnetic resonance imaging (MRI) is a crucial diagnostic and monitoring method. It provides the detailed anatomical information required for accurate medical evaluations. When diagnosing liver disease, magnetic resonance imaging (MRI) can help identify liver issues early on by thoroughly examining the histology of the liver. Unfortunately, the enormous dimensionality and complexity of the obtained data sometimes render MRI imaging useless for disease diagnosis. The enormous quantity of information that can be extracted from these images is computationally expensive, which may make it more difficult to categorize diseases effectively. It is computationally demanding to extract the vast amount of information from these photos, which could complicate the process of accurately classifying diseases. Consequently, there is growing interest in capturing important distinguishing features and minimizing computing burden without compromising diagnostic accuracy with advanced data reduction approaches. Examining the efficacy of the SocialSpider Algorithm (SSA) as a feature reduction method for liver image interpretation on MRI images is the aim of the proposed study. Social spider cooperation serves as the model for SSA, a metaheuristic optimization algorithm. It aims to imitate the cooperative hunting techniques and social dynamics displayed by social spiders. By using SSA, it is possible to retain crucial diagnostic information related to the diagnosis of liver illness while condensing the vast array ofMRIfeaturesintoamoremanageableandinformativesubset.Thestudytalksontwotopics. Firstly, to investigate the extent to which the SSA preserves the crucial discriminative informationrequiredfortheclassificationofliverdisorderswhilereducingthedimensionality of MRI characteristics, Secondly, to evaluate the impact of the SSA-preprocessed decreased feature set on the CNN liver image classification performance. Three CNN architectures are GoogleNet, ResNet50, and VGG16 that are well-known for performing well on image classification tasks will be used in the current study. These CNN models try to improve early liverdiseasediagnosisbycombiningSSAforMRIfeaturereductionandlookingintosuitable CNN topologies for improved classification. The decreased feature sets from SSA-processed MRI images will be used to train and fine-tune them for liver disease classification. When compared to other models with SSA, GoogleNet accuracy outperforms in the study.
NatureInspired Algorithm:

DOI:10.46647/ijetms.2025.v09i02.061 ISSN: 2581-4621

Nature-inspired optimization algorithms, as the name suggests, are algorithms that draw inspirationfromnaturalphenomenaincludingswarmintelligence,biologicalsystems,physical systems, and chemical systems. (Wang , Qin , Wan , & Song , 2021)

SocialSpider Algorithm:

The cooperative behavior of social spiders serves as the foundation for a novel swarm algorithm known as Social Spider Optimization. Similar to a swarm of spiders, search agents in algorithmmove in harmony with the biological activity of the colonies (Luque-Chang , Cuevas , Fausto , Zaldivar , & Pérez , 2018).It would be interesting to find SSA real-world applications that can be managed well and affordably (James & Li VO, 2015).

Featureselection:

Feature selection is the process of identifying which features are necessary for the model to perform as intended. Machine learning procedures is feature engineering, which primarily consists of two steps are feature extraction and feature selection.

Classifiers:

The Classification method, a Supervised Learning technique, establishes the category of new finds based on training data. When a program utilizes classification, it classifies new findings intodifferentclassesorcategoriesafterfirstlearningfromthegivendataset.(JavaTpoint,n.d.)

GoogleNet:

The Inception design serves as the foundation for convolutional neural networks of the GoogLeNetvariety.Thenetworkcanchoosefromarangeofconvolutionalfiltersizesforeach frame by using Inception modules.

ResNet50:

As a result of their architecture, which resolved the vanishing gradient problem and made it possible to build networks with hundreds or even millions of convolutional layers, convolutional neural networks outperform shallower networks in terms of performance.

II. LITERATURE SURVEY

[1] Joel Jacob et al., "Diagnosis of Liver Disease Using Machine Learning Techniques", International Research Journal of Engineering and Technology (IRJET), Volume: 05, Issue: 04,2018, pp 4011-4014.

Health is Wealth. Though the medical field has grown rapidly with highly effective technologies, chronic diseases such Heart and Liver diseases are life-menacing. Various life factors such as alcohol, smoking, stress, food, lifestyle, etc causes imbalance and add toxics to the human body leading to the occurrence of assorted diseases and disorders. The medical records of the patients as a vast source of data are applied to the data mining techniques to extract the valid dataset to predict the liver disease. The classification algorithms have been widely used in the decision- making process. RNN being a text classifier of deep learning technique with the advantage of processing in multiple loops in a sequential manner to obtain best performances measured by the factor of accuracy has been proposed in this study.

2] Pragati Bhagat et al.," System for diagnosis of Liver Disease Using Machine Learning Technique", International Research Journal of Creative Research Thoughts", ISSN: 2320-2882, pp25-30.

Many people suffer from liver disease but they don't have an idea about it. It is difficult to diagnosis of liver disease at high level. Before treatment of liver disease doctors first diagnose whether patient has liver disease or not, basis on different parameter. The system for diagnosis of liver disease using machine learning algorithms is an initiative towards better diagnosis of this disease as early as possible. Various algorithms are being studied in order to select the best algorithm which can give the best accurate results. According to the four parameters Accuracy, Precision, Sensitivity and Specificity the algorithm is being selected. After the study ANN algorithm turned out to be the best algorithm to implement and provides more accuracy than other algorithms. So, ANN is

DOI:10.46647/ijetms.2025.v09i02.061 ISSN: 2581-4621

implemented using MATLAB platform and the user interface is also constructed with the help of MATLAB.

[3] Nazmun Nahar and Ferdous Ara, "Liver Disease Prediction Using Different Decision Tree Techniques", International Journal of Data Mining & Knowledge Management Process Vol.8 No.2(March 2018).

To diagnose and forecast liver disease, a variety of machine learning algorithms are widely employed in the medical industry. We analysed several research publications in which we focused on various Data mining approaches for making use of data to support the study of high and multi-dimensional data in the health-care industry. In this regard, we have publications that are relevant to this topic in terms of methodology, algorithms, and outcomes. For selected publications, results and assessment techniques are examined, and a detailed summary of the findings is offered at the end. As a result, the purpose of this research is to use machine learning algorithms to improve the diagnosis and prediction of liver illness.

[4] A Saranya, G.Seenuvasan, "A Comparative Study Of Diagnosing Liver Disorder Disease Using Classification Algorithm", International Journal Of Computer Science and Mobile Computing, Vol. 6 Issue 8mpage no 49-54(August 2017).

Liver disease is the major cause of death every year. Liver diseases is the fifth big killer in England after cancer, stroke and respiratory disease. The most common causes of liver disease worldwide are chronic hepatitis B and C, alcohol and non alcoholic. Machine Learning has a strong potential in automated diagnosis of various diseases. With the recent upscale in various liver diseases, it is necessary to identify the liver disease at a preliminary stage. In this we propose a new classifier by extending the XGBoost classifier with genetic algorithm. This compares various classification models and visualization techniques used to predict liver disease with feature selection. Outlier detection is used to find out the extreme deviating values and they are eliminated using isolation forest. The performance is measured in terms of accuracy, precision, recall f-measure and time complexity.

[5] S. Dhamodharan, Liver Disease Prediction Using Bayesian Classification, National Conference on Advanced Computing, Application & Technologies, 2014.

In recent years in healthcare sectors, data mining became an ease of use for disease prediction. Data mining is the process of dredge up information from the massive datasets or warehouse or other repositories. It is a very challenging task to the researchers to predict the diseases from the voluminous medical databases. To overcome this issue the researchers use data mining techniques such as classification, clustering, association rules and so on. The main objective of this research work is to predict liver diseases using classification algorithms. The algorithms used in this work are Naïve Bayes and support vector machine (SVM). These classifier algorithms are compared based on the performance factors i.e. classification accuracy and execution time. From the experimental results it is observed that the SVM is a better classifier for predict the liver diseases.

III. MODULES

Service Provider

In this module, the Service Provider has to login by using valid user name and password. After login successful he can do some operations such as

- Login,
- Browse and Train & Test Data Sets
- View Trained and Tested Accuracy in Bar Chart
- View Trained and Tested Accuracy Results
- View Travel Prediction
- View Travel Prediction Type Ratio
- Download Predicted Data Sets

DOI:10.46647/ijetms.2025.v09i02.061 ISSN: 2581-4621

- View All Remote Users
- Logout

Registration Module

In this module, the new remote user can register by entering he/his details i.e.,

- Username
- Password
- Email
- Country
- Signup

Remote User Module

In this module, there are n numbers of users are present. User should register before doing any operations. Once user registers, their details will be stored to the database. After registration successful, he has to login by using authorized user name and password. Once Login is successful user will do some operations like

- Register and login
- Browse and Train & Test Data Sets
- View Trained and Tested Accuracy in Bar Chart
- View Trained and Tested Accuracy Results
- View Travel Prediction
- View Travel Prediction Type Ratio
- Download Predicted Data Sets

Materials and methods

Training dataset

In this project , to tune up the object detection model for human detection under various low light conditions, a recently released ExDARK dataset is considered which specifically focuses on a low-light environment. In this dataset, 12 different classes of objects are labeled, out of which we fetched data of our desired class for training. This dataset contains different indoor and outdoor low light images; furthermore, the data is subdivided for low light environment into 10 classes ambient, object, strong, twilight, low, weak, screen, window, shadow, and single. Sample images of various indoor-outdoor low-light environments from the datase

Testing dataset

A custom dataset is used for the evaluation of the proposed model. The dataset is collected from the market of Rawalpindi, Pakistan during the night in the days of COVID-19. Pakistan is one of the most urbanized countries in South Asia with a 3% yearly urban population growth rate. The large population and congested streets make it a riskier place in the growth of COVID-19 and it is very difficult to maintain safety distance in such narrow places. Hence, the monitoring system should need to have high accuracy in terms of the detection and location of the people. Evaluation of the proposed framework in such a highly-populated area will help us to better analyze the performance of the model. Test dataset is the collection of 346 RGB frames. Frames are collected with motionless ToF camera of Samsung galaxy note 10+ installed 4.5 feet above the ground where a 0° regular camera view calibration is adopted. Sample images of low-light conditions from the custom dataset

IV. RESULT AND DISCUSSION

CCTVs and Drones can be used for human detection. Closed Circuit television (CCTV) are being used as a means of surveillance from a long time but due to its limitations it is not completely reliable. The drone thus has a better communication with the rest of the swarm in a particular area to follow the human while also dividing the areas between the drones dynamically so as to not lose

International Journal of Engineering Technology and Management Science
Website: ijetms.in Issue: 2 Volume No.9 March - April – 2025
DOI:10.46647/ijetms.2025.v09i02.061 ISSN: 2581-4621

track of the human. OpenCV, computer vision and deep learning are used to monitor social distancing across the region. Initially, object detection is applied to detect pedestrians in a video stream. In the next step, the pairwise distances between all detected people are calculated and finally these distances are compared with the standard distance that should be maintained (6 feet or 2 meters) and are represented by red frame if they are violated and green frame otherwise. So, if 5-6 people gather around in a particular area, the local authorities or the local police stations will be immediately notified. Recently, after the outbreak of this virus, the police Authorities need to patrol across the city and are bound to invest time unnecessarily. Using this concept of social distancing detection, the police will be able to monitor and reach the exact location and control the scenario immediately. Thus, social distancing can be controlled and indirectly the spread of COVID-19 be prevented. The below figure shows the steps for implementation of a social distancing detector.

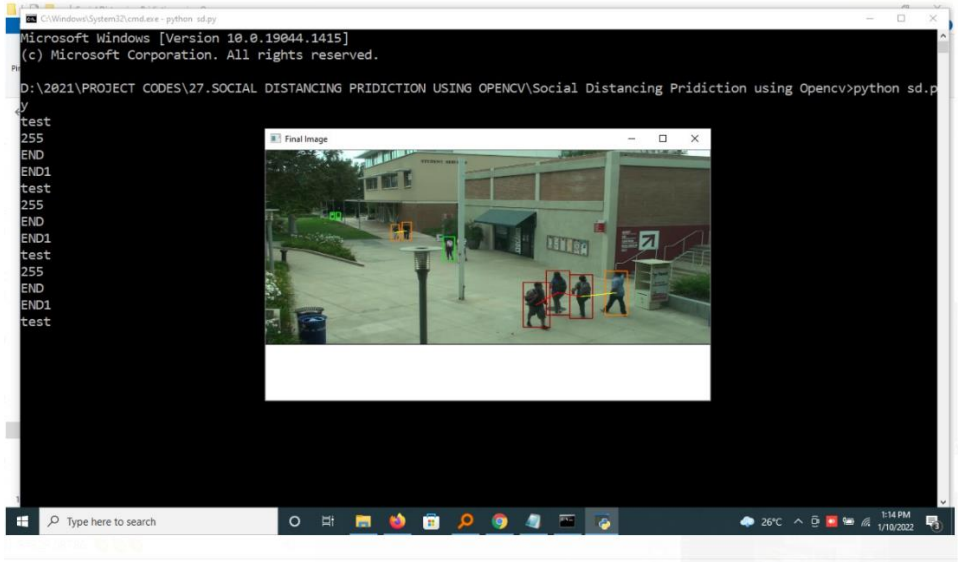

Fig: Red Indication showing people not maintain distance

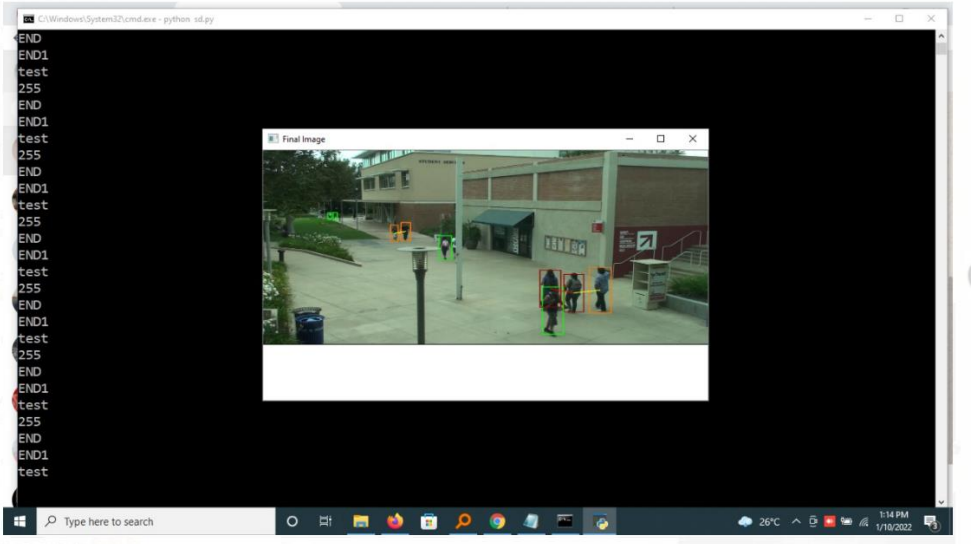

Fig: 1.5 Green Indication showing People maintaining distance

V. CONCLUSION

The article proposes an efficient real-time deep learning based framework to automate the process of monitoring the social distancing via object detection and tracking approaches, where each individual is identified in the real-time with the 9 help of bounding boxes. The generated bounding

DOI:10.46647/ijetms.2025.v09i02.061 ISSN: 2581-4621

boxes aid in identifying the clusters or groups of people satisfying the closeness property computed with the help of pair wise vectorized approach. The number of violations are confirmed by computing the number of groups formed and violation index term computed as the ratio of the number of people to the number of groups. The extensive trials were conducted with popular state-of-the-art object detection models: Faster RCNN, SSD, and YOLO v3, where YOLO v3 illustrated the efficient performance with balanced FPS and mAP score. Since this approach is highly sensitive to the spatial location of the camera, the same approach can be fine tuned to better adjust with the corresponding field of view.

VI. REFERENCES

[1] K. Graves, Ceh: Official certified ethical hacker review guide: Exam 312-50. John Wiley & Sons, 2007.

[2] R. Christopher, "Port scanning techniques and the defense against them," SANS Institute, 2001.

[3] M. Baykara, R. Das¸, and I. Karado ˘gan, "Bilgi g ¨uvenli ˘gi sistemlerinde kullanilan arac¸larin incelenmesi," in 1st International Symposium on Digital Forensics and Security (ISDFS13), 2013, pp. 231–239.

[4] RashmiT V. "Predicting the System Failures Using Machine Learning Algorithms".International Journal of Advanced Scientific Innovation, vol. 1, no. 1, Dec. 2020, doi:10.5281/zenodo.4641686.

[5] S. Robertson, E. V. Siegel, M. Miller, and S. J. Stolfo, "Surveillance detection in high bandwidth environments," in DARPA Information Survivability Conference and Exposition, 2003. Proceedings, vol. 1. IEEE, 2003, pp. 130–138.

[6] K. Ibrahimi and M. Ouaddane, "Management of intrusion detection systems based-kdd99: Analysis with lda and pca," in Wireless Networks and Mobile Communications (WINCOM), 2017 International Conference on. IEEE, 2017, pp. 1–6.

[7] Girish L, Rao SKN (2020) "Quantifying sensitivity and performance degradation of virtual machines using machine learning.",Journal of Computational and Theoretical Nanoscience, Volume 17, Numbers 9-10, September/October 2020, pp.4055-4060(6) https://doi.org/10.1166/jctn.2020.9019

[8] L. Sun, T. Anthony, H. Z. Xia, J. Chen, X. Huang, and Y. Zhang, "Detection and classification of malicious patterns in network traffic using benford's law," in Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC), 2017. IEEE, 2017, pp. 864–872.

[9] S. M. Almansob and S. S. Lomte, "Addressing challenges for intrusion detection system using naive bayes and pca algorithm," in Convergence in Technology (I2CT), 2017 2nd International Conference for. IEEE, 2017, pp. 565–568