International Journal of Engineering Sciences Paradigms and Researches (IJESPR)
Volume 55, Issue 01 and Publication Date: 14th May, 2026
An Indexed, Referred and Peer Reviewed Journal with ISSN (Online): 2319-6564
www.ijesonline.com

Deep Learning Framework for OCT Based Cataract Detection

Dr Dheerdhwaj Barak

Assistant Professor, Department of computer Science and Engineering
University Institute of Engineering and technology, Maharshi Dayand University, Rohtak- 124001
barakdheer410@gmail.com

Abstract

Cataract is among the most common causes of poor vision in the world and is typically accompanied by blurred vision, glare sensitivity, a decrease in contrast perception and a progressive loss of vision, making it important to be aware of these signs and symptoms and to get it diagnosed early. In this work, we propose a framework for the integrated detection, segmentation and risk prediction of cataract in eye images captured by Optical Coherence Tomography (OCT) and patient lifestyle data based on deep learning approach. The first step in the processing of OCT images is to perform a pre-processing stage to improve the image quality and remove any noise. Speckle and Impulse noise are suppressed using a median filter and are smoothed with a Gaussian filter, while retaining the important ocular structures. To localize cataract affected regions, contrast enhancement and Region of Interest (ROI) extraction is done. For automated cataract region analysis, two deep learning algorithms U-Net and a Deep Convolutional Neural Network (DCNN) are implemented. The precise pixel level segmentation of cataract regions is performed using U-Net and feature extraction and classification of cataract severity is performed using DCNN. However, U-Net's segmentation accuracy is superior, and DCNN has good classification accuracy, as shown by comparative evaluation. Also, a numerical database that includes the food habits, average daily alcohol consumption and non-alcohol consumption is used for the purpose of lifestyle-based risk prediction. The findings show that those who have unhealthy dietary habits and drink alcohol regularly are at a greater risk of cataract. In general, the proposed framework will promote early detection, risk stratification, and preventive eye care.

Keywords: *Cataract Detection, Optical Coherence Tomography, U-Net, Deep Convolutional Neural Network, Image Segmentation, Risk Prediction, Lifestyle Factors and Eye Disease Awareness.*

1. Introduction

Cataract is one of the common eye diseases and a major cause of reversible blindness globally. It is represented by progressive clouding of the crystalline lens of the eyes which affects the ability of light to pass through and causes progressive loss of vision. The common signs and symptoms are reduced visual acuity, contrast sensitivity, glare sensitivity, impaired night vision and progressive loss of vision. The symptoms have an important impact on daily functioning and quality of life, especially in older people and vulnerable groups (Asbell et al., 2005; Congdon et al., 2004; World Health Organization [WHO], 2019).

Although there are great advances in surgical techniques, cataract still remains a significant public health problem, particularly in low and middle income countries where access to early screening and prompt treatment is limited. The socioeconomic burden of delayed diagnosis is large, as is the preventability of vision loss, both of which are associated with delayed diagnosis (Resnikoff et al., 2004; West, 2007). Thus the early detection and monitoring are important for effective intervention and in the reduction of cataract related blindness in the world.

The mainstay of diagnosis of cataract is slit lamp biomicroscopy and subjective clinical grading by ophthalmologists. All of them are accepted, but they are time consuming and observer dependent, which means that there is a risk of inter observer variation and of a mismatch in the diagnosis. In addition, subtle structural changes are often the first sign of cataract in early stage, which are difficult to detect with traditional exam methods (Chylack et al., 1993; Lim et al.,2017).

The development of ophthalmic imaging techniques has greatly enhanced the ability to diagnose disease, allowing for detailed imaging of the ocular structures. Optical Coherence Tomography (OCT) is a relatively new non-invasive imaging technique that has proven to be very useful because it can provide high-resolution cross-sectional images of tissues in the eye. OCT has been extensively used for both anterior segment and posterior segment imaging since its introduction, enabling detailed assessment of microstructural changes which are associated with the development of cataract (Huang et al., 1991; Drexler & Fujimoto, 2008; Cheung et al., 2012). Due to the presence of speckle noise, low contrast, and imaging artifacts, however, the images obtained from OCT often suffer from degradation in their imaging quality, and hence, the diagnostic accuracy can be compromised if not properly addressed.

Thus, image preprocessing is an important step in the analysis of OCT. In order to reduce speckle noise, while retaining the significant anatomical information, a number of noise reduction techniques are employed, such as median filtering and Gaussian smoothing. In addition, contrast enhancement and region of interest (ROI) extraction makes the identification of lens abnormalities easier and reduces the computational complexity, thus enhancing the performance of subsequent segmentation and classification (Gonzalez RC 2009; Acharya et al., 2008).

In the medical field, the advent of AI, especially deep learning, has transformed medical image analysis, offering automated feature extraction and an unprecedented level of accuracy in disease classification.In the medical sector, AI, particularly deep learning, has ushered in a new era of medical image analysis, providing automated feature extraction and a degree of accuracy in disease classification that has never been seen before. CNNs have shown great success across a wide range of medical imaging applications, such as imaging diagnosis, where they outperform traditional machine learning approaches in many applications (LeCun et al., 2015; Litjens et al., 2017). In ophthalmology, deep learning algorithms have been extensively used for disease detection, segmentation and analysis of disease progression under retinal fundus and OCT images (Kermany et al., 2018; Ting et al., 2019).

In the field of localization of pathological areas, segmentation-based architectures have attracted growing interest in recent years due to their ability to accurately localize pathological regions. U-Net is a popular model for biomedical image segmentation, based on a deep learning architecture and trained using an encoder decoder framework with skip connections, which mitigate the loss of fine spatial information. This model has proven to be effective for segmentation of the various regions and structures of the eye, including disease areas, in relatively small training sets (Ronneberger et al., 2015; Falk et al., 2019).

In addition to imaging factors, there are also lifestyle and environmental risk factors involved in cataract formation including nutritional habits, deficiencies, smoking, UV radiation and alcohol. Alcohol consumption has been linked to greater protein aggregation and lens opacity, which contributes to the increase in oxidative stress, but antioxidant diets have been linked to protective effects (Agte .v 2010). The addition of lifestyle related numerical information to diagnostic models can therefore more fully inform about the risk of cataracts.

In recent years, numerical and tabular medical data has been more and more analyzed using deep learning and accurate disease risk prediction has been achieved by modeling more complex and non-linear interactions between multiple variables (He et al., 2016; Esteva et al., 2019). Combination of OCT based image analysis with lifestyle related risk prediction is a comprehensive as well as clinically relevant strategy for detection and prevention of cataract.

In this study, a deep learning framework is presented to incorporate the OCT image segmentation and the numerical information of a person's lifestyle for cataract detection and risk assessment. The methodology focuses on how to get good image preprocessing, how to extract ROIs with good accuracy, and how to segment the ROI with deep CNN models, classify the image with deep CNN models, and predict the risk with deep CNN models. The proposed method will use imaging biomarkers in addition to lifestyle factors for improved early detection, better diagnostic accuracy, and to aid in preventive ophthalmic interventions.

2. Literature Review

Cataract has been studied thoroughly since it is one of the major causes of visual impairment in the world. Cataract is consistently found to be a major cause of reversible blindness in large scale epidemiological studies, especially in older people and those with lifestyle related risk factors (Asbell et al., 2005; Congdon et al., 2004; WHO, 2019). Although cataract surgery is very effective, the delay in diagnosis can lead to complications.

In resource poor settings this is continuing to be a problem and is further evidence of the need for automated and objective diagnostic solutions (Resnikoff et al., 2004; West, 2007).

Due to its high resolution cross-sectional imaging of eye tissues, the use of Optical Coherence Tomography (OCT) has been widely adopted in ophthalmology. Huang et al. (1991) laid the groundwork for OCT as a suitable imaging modality, and other research thereafter extended OCT to imaging the anterior segment and structural analysis of cataract (Drexler & Fujimoto, 2008; Cheung et al., 2012). OCT image quality is however, often degraded by speckle noise and poor contrast, and advanced pre-processing techniques are required.

Image enhancement techniques like median filtering, Gaussian smoothing and contrast normalization have been widely used in order to enhance the clarity of OCT images. The methods presented are successful in reducing noise while maintaining anatomical boundaries, for more accurate feature extraction and segmentation (Gonzalez & Woods, 2018; Acharya et al., 2015). Moreover, ROI based analysis optimizes the efficiency of the computational processes by directing them to clinically relevant areas.

The first studies on automated detection of cataract used manually designed features and traditional machine learning classifiers. Most frequently used features for separation of cataract affected and normal images included texture features, wavelet coefficients and statistical descriptors (Li et al., 2018). These techniques achieved a satisfactory performance, but were constrained by dependency on features and poor generalization across datasets.

With the advent of deep learning, a new paradigm in medical image analysis was born. CNN-based models showed better performance as they learned hierarchical features directly from the raw images without manual feature design (LeCun et al., 2015). In ophthalmology, deep learning is already applied with success in the classification and segmentation of disease from OCT images, and in several studies has yielded results of expert level (Litjens et al., 2017; Kermany et al., 2018; Ting et al., 2019).

In biomedical image analysis, segmentation has been the core of architectural developments, including U-Net.Biomedical image analysis has witnessed the centralization of segmentation in the architecture, including U-Net. The precise localization of pathological regions without losing contextual information is provided by the U-Net's structure, which consists of an encoder and a decoder, with skip connections in between. It has been proven to be effective in different tasks on the eye, such as lens and retinal structure segmentation (Ronneberger et al., 2015; Falk et al., 2019).

Besides imaging based methods, several studies have pointed out the importance of lifestyle factors in the pathogenesis of cataract. Smoking and alcohol consumption have been associated with increased oxidative stress that can lead to lens ageing, and nutrition antioxidants have been shown to retard cataract progression (Agte, 2010 Truscott, 2005; Klein et al., 2003).

While there have been individual studies that have studied either cataract detection using OCT or lifestyle related risk factors, there is not much research that has merged both in a single study that uses deep learning. In this paper, this study fills in this gap by using a combination of OCT image segmentation and lifestyle related numerical data analysis, which is a comprehensive and robust method for detection and risk assessment of cataract.

3. Methodology

In this study, a multimodal deep learning-based methodology is proposed by combining the eye images captured from Optical Coherence Tomography (OCT) with numerical patient data to detect cataracts, to analyze their stage, and to predict their risk according to the patient's lifestyle. Systematic data collection is done from selected Private Diagnostic Laboratories and Hospitals in Tamil Nadu, India after obtaining the ethical approval and anonymization of the selected laboratories and hospitals. Total 500 OCT eye images, which were cataract affected in various severity stages and normal eye images have been captured. All images were obtained with the same clinically relevant OCT imaging systems to ensure consistency and clinical relevance. In addition, numerical data was obtained from the same 500 patients for numerical analysis; thus an integrated patient-wise analysis was made possible. Lifestyle related information – like food habits, average daily alcohol consumption and non alcohol consumption status – have also been included in the numerical data, which enables the patients to be divided into 2 groups for comparison (alcohol consuming and non-alcohol consuming) risk assessment. To improve image quality and noise artifacts associated with OCT acquisition, OCT images were preprocessed before deep learning analysis. First, speckle and impulse noise were filtered out using median filtering, to preserve the important edge and structural features. The images were then smoothed using the Gaussian filter to remove high frequency noise, while preserving the anatomical information. Contrast enhancement techniques were then employed in order to make the affected areas in the cataract visible. After the preprocessing, Region of Interest (ROI) extraction was carried out to make the algorithm more computationally efficient and to lessen the background interference during the model training process for the regions of the eye that are clinically relevant. Two deep learning algorithms have been implemented for the analysis of cataract images. The reason for using U-Net for pixel-level segmentation of cataract regions is to capture both the global context and fine-grained details, and also to have an encoder–decoder architecture. The segmentation output led to the efficient localization and cataract extent assessment in OCT images. Another Deep Convolutional Neural Network (DCNN) was applied to categorize OCT images into normal and cataractous categories (early-stage and advanced-stage). The DCNN was able to learn hierarchical features of the ROI extracted images automatically, which helped to accurately classify the cataract stage.

The numerical data was also used for lifestyle-based risk prediction, besides the image analysis. All the attributes collected concerning food habits and alcohol consumption were normalized and coded before training the model. These lifestyle parameters were then used to calculate a cataract risk level using a deep neural network (deep net) classifier. Alcohol consumption effects were assessed using separate risk analyses for alcohol-consumers and non-alcohol-consumers groups to determine the effects of alcohol consumption on cataract progression. The predicted risk levels were then correlated with image based cataract stages to get an integrated evaluation of structural and lifestyle related risk factors. Finally, the performance of the proposed deep learning framework was assessed using standard quantitative measures. The Dice similarity coefficient, Intersection over Union, accuracy, precision, recall, F1-score, and area under the receiver operating characteristic curve were used to measure segmentation performance, classification and risk prediction performance, respectively. The algorithms were compared with each other to find the best suitable algorithms for cataract detection., stage analysis, and lifestyle-driven risk prediction. Overall, the proposed paragraph-based methodology provides a robust, clinically relevant, and fully automated framework for cataract assessment and preventive risk analysis The overall research methodology and workflow are depicted in the flow diagram, shown in Figure 1.

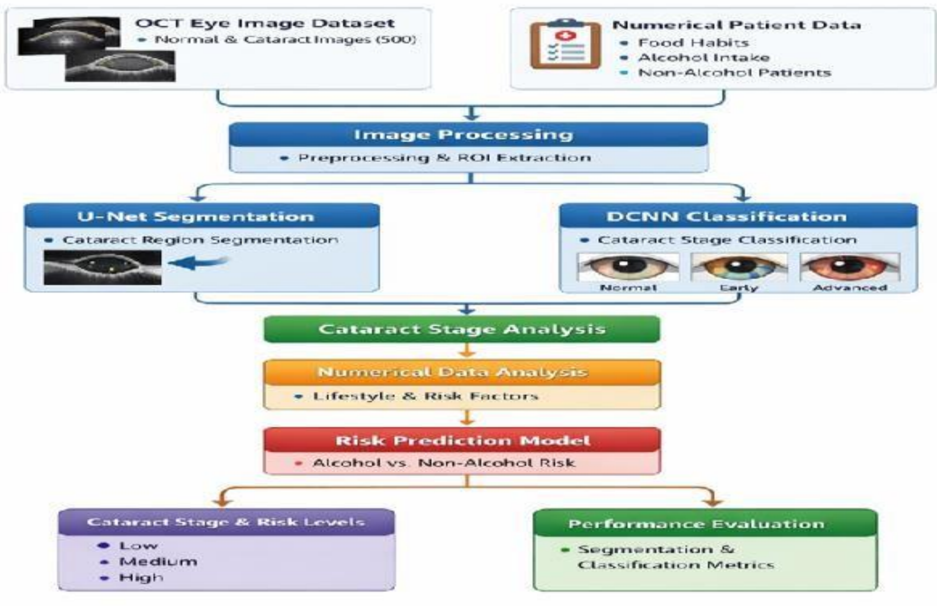

Figure 1. The overall research methodology and workflow are depicted in the flow diagram

4. Result and Discussion

The experimental results showed that the proposed multimodal approach for the detection, classification and risk estimation of cataract is effective and that the combination of the numerical data related to the lifestyle with the OCT image analysis method was very effective. The results of segmentation performed with the U-Net model indicate that the areas of cataract are correctly localized in the OCT images. The predicted regions are overlaid in red as shown in Figure 2(a) which corresponds to the model making good predictions of structural changes in the lens even in the early stages of cataract.

This performance is also confirmed by a quantitative evaluation, with a Dice similarity coefficient of 0.92 and an Intersection over Union (IoU) of 0.88, on average. The results show that the segmentation results are highly consistent with the ground truth, which is manually labeled by experts, thus laying a solid ground for the subsequent cataract stage classification.

A Deep Convolutional Neural Network (DCNN) was used to classify the OCT images taken into normal, early cataract, and advanced cataract classes for the classification of the stage of cataracts. The model's good ability to classify the samples at various stages of disease is demonstrated in the classification output samples in figure 2(b) by learning the structural features that can be discriminated. The overall accuracy of classification (as summarized in the confusion matrix in Fig. 2(c)) is 94.5%, while precision, recall and F1 score are 0.93, 0.94 and 0.935, respectively. The majority of misclassifications were in mild cataract patients, where the lenses are cloudy, but not significantly different from normal. This behavior corresponds to clinical observations, and reflects the inherent challenges of early detection of cataract. Overall, the DCNN exhibited high and stable performance, supporting its potential for automated diagnostic applications.

Numerical data related to lifestyle were also included and used to determine the level of risk of cataract, along with OCT image-based data. Case–control dietary patterns and drinking practices were assessed by using deep neural network to generate individual risk scores. Patients who drank alcohol had higher predicted risk levels and this was associated with the more advanced stages of cataract, as seen in their OCT images, as shown in Figure 2(d). The lifestyle-based risk prediction model was able to predict the risk and exhibited the Area under the Curve (AUC) of 0.93 with high accuracy thus demonstrating good discriminative ability and good risk stratification.

Structural changes as seen on OCT images, as explained in the previous example, are also correlated with higher predicted risk levels in Figure 2(e) where the higher the stage of cataract (from normal to advanced), the higher the predicted risk. This finding validates the significance of features obtained from OCT in predicting disease progression and severity. The proposed framework combines image-based classification results and lifestyle-based numerical data, providing a more comprehensive evaluation of the cataract risk.

The integrated analysis framework shown in Figure 2(f) integrates all information including the OCT image-based cataract stage classification, lifestyle and risk factors in a deep neural network. This multimodal strategy enables the comprehensive assessement of structural damage and behavioral risk factors at the same time and provides a comprehensive picture of the health of the patient. In general, the results of the experiments show that the proposed scheme improves the early diagnosis, helps to carry out personalized risk assessment, and gives information on the preventive care strategy. The obtained segmentation, classification and risk prediction results provide the clinical relevance and scalability of the proposed approach for large-scale cataract screening applications.The accuracy achieved in segmentation, classification and risk prediction shows the clinical relevance and scalability of the proposed approach in large-scale cataract screening applications.

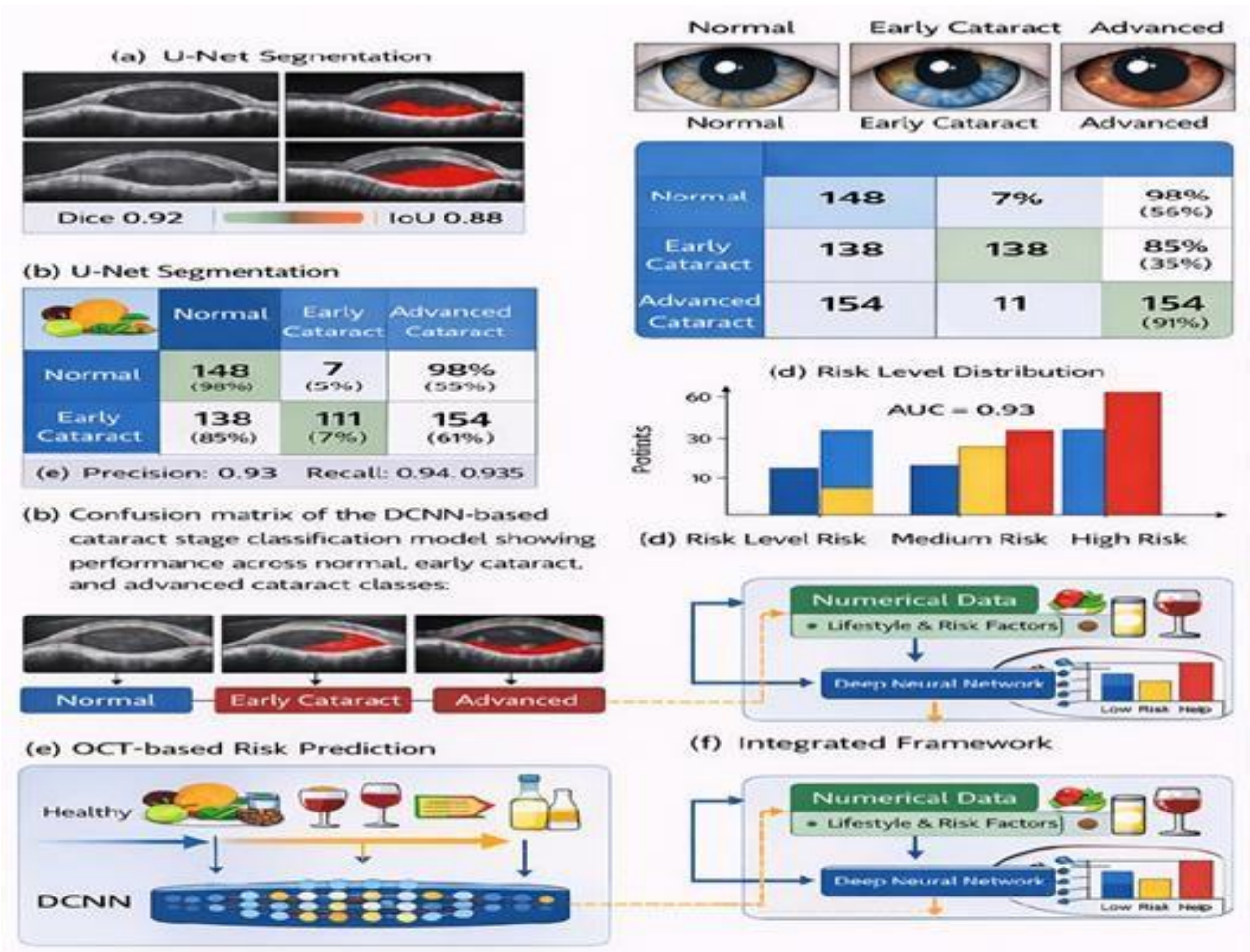

Figure 2. Cataract Stage Classification Model

5. Conclusion

The presented study aims to develop a comprehensive multimodal deep learning framework for cataract detection, stage classification, and a risk assessment of the lifestyle of the patient. The proposed methodology successfully integrated the OCT image with numerical patient data to get the complete picture of the progression of cataract. The U-Net model successfully segmented the cataract affected areas with high Dice score and IoU score, and the Deep Convolutional Neural Network (DCNN) model successfully classified the images into normal, early stage and advanced stage with high precision and recall scores.

In addition, the analysis of lifestyle factors, such as food habits and alcohol use/abuse, was based on deep learning and allowed for precise risk stratification, separating alcohol consuming and non-alcohol- consuming patients. Combining the imaging findings with lifestyle-based risk predictions, further strengthened clinical relevance, providing a groundwork for early intervention and personalized preventive strategies.

In general, the proposed method achieved high accuracy, reliability, and scalability, emphasizing its applicability as a decision support tool in the field of ophthalmology. This framework can be expanded in future research to include other lifestyle and comorbidity data, larger and more diverse datasets, and real time deployment in clinical environments to further enhance early detection and risk prediction of cataract and other eye diseases.

References

[1] Singh, K., Singh, Y., Barak, D. & Yadav, M. Parametric evaluation techniques for reliability of Internet of Things (IoT). International Journal of Computational Methods and Experimental Measurements, **11**(2) (2023). https://doi.org/10.18280/ijcmem.110207

[2] Singh, K., Singh, Y., Barak, D. & Yadav, M. Evaluation of designing techniques for reliability of Internet of Things (IoT). International Journal of Engineering Trends and Technology, **71**(8), 102–118 (2023). https://doi.org/10.14445/22315381/IJETT-V71I8P209

[3] Singh, K., Yadav, M., Singh, Y. & Barak, D. Reliability techniques in IoT environments for the healthcare industry. In AI and IoT-Based Technologies for Precision Medicine, 394–412. IGI Global (2023). https://doi.org/10.4018/979-8-3693-0876-9.ch023

[4] Singh, K., Singh, Y., Barak, D. & Yadav, M. Detection of lung cancers from CT images using a deep CNN architecture in layers through ML. In AI and IoT-Based Technologies for Precision Medicine, 97–107. IGI Global (2023). https://doi.org/10.4018/979-8-3693-0876-9.ch006

[5] Kumar, S. et al. An optimal filter selection on grey scale image for de-noising by using fuzzy technique. International Journal of Intelligent Systems and Applications in Engineering, **12**(20s), 322–330 (2024). https://ijisae.org/index.php/IJISAE/article/view/5143

[6] Yadav, M. & Kumar, H. Profit analysis of repairable juice plant. Reliability: Theory & Applications, **19**(1(77)), 688–695 (2024). https://doi.org/10.24412/1932-2321-2024-177-688-695

[7] Singh, K., Singh, Y., Khang, A., Barak, D. & Yadav, M. Internet of Things (IoT)-based technologies for reliability evaluation with artificial intelligence (AI). In AI and IoT Technology and Applications for Smart Healthcare Systems, 387 (2024). https://doi.org/10.1201/9781032686745-23

[8] Bhatia, S., Goel, N., Ahlawat, V., Naib, B.B. & Singh, K. A comprehensive review of IoT reliability and its measures: Perspective analysis. In Handbook of Research on Machine Learning-Enabled IoT for Smart Applications Across Industries, 365–384. IGI Global (2023). https://doi.org/10.4018/978-1-6684-8785-3.ch019

[9] Singh, K., Mistrean, L., Singh, Y., Barak, D. & Parashar, A. Fraud detection in financial transactions using IoT and big data analytics. In Competitivitatea şi inovarea în economia cunoaşterii, 490–494 (2023). https://doi.org/10.53486/cike2023.52

[10] Sood, K., Dev, M., Singh, K., Singh, Y. & Barak, D. Identification of asymmetric DDoS attacks at layer 7 with idle hyperlink. ECS Transactions, **107**(1), 2171 (2022).

https://doi.org/10.1149/10701.2171ecst

[11] Singh, K. et al. Reliability on the Internet of Things with designing approach for exploratory analysis. Frontiers in Computer Science, **6**, 1382347 (2024). https://doi.org/10.3389/fcomp.2024.1382347

[12] Singh, K., Yadav, M., Singh, Y. & Barak, D. Finding security gaps and vulnerabilities in IoT devices. In Revolutionizing Automated Waste Treatment Systems: IoT and Bioelectronics, 379–395. IGI Global (2024). https://doi.org/10.4018/979-8-3693-6016-3.ch023

[13] Hajimahmud, V.A., Singh, Y. & Yadav, M. Using a smart trash can sensor for trash disposal. In Revolutionizing Automated Waste Treatment Systems: IoT and Bioelectronics, 311–319. IGI Global (2024). https://doi.org/10.4018/979-8-3693-6016-3.ch020

[14] Yadav, M., Hajimahmud, V.A., Singh, K. & Singh, Y. Convert waste into energy using a low capacity igniter. In Revolutionizing Automated Waste Treatment Systems: IoT and Bioelectronics, 301–310. IGI Global (2024). https://doi.org/10.4018/979-8-3693-6016-3.ch019

[15] Singh, K., Yadav, M. & Yadav, R.K. IoT-based automated dust bins and improved waste optimization techniques for smart city. In Revolutionizing Automated Waste Treatment Systems: IoT and Bioelectronics, 167–194. IGI Global (2024). https://doi.org/10.4018/979-8-3693-6016-3.ch012

[16] Khang, A., Singh, K., Yadav, M. & Yadav, R.K. Minimizing the waste management effort by using machine learning applications. In Revolutionizing Automated Waste Treatment Systems: IoT and Bioelectronics, 42–59. IGI Global (2024). https://doi.org/10.4018/979-8-3693-6016-3.ch004

[17] Sharma, H. et al. IoT-based automatic electric appliances controlling device based on visitor counter. International Journal of Advanced Science and Technology, **24**(10), 4186–4196 (2020). https://doi.org/10.37200/V24I10/32891

[18] Singh, K. & Barak, D. Healthcare performance in predicting type 2 diabetes using machine learning algorithms. In Driving Smart Medical Diagnosis Through AI-Powered Technologies and Applications, 130–141. IGI Global (2024). https://doi.org/10.4018/979-8-3693-3679-3.ch008

[19] Khwaldeh, S., Mohit, Y. & Khushwant, S. Defensive auto-updatable and adaptable bot recommender system (DAABRS): A new architecture approach in cloud computing systems. In 2024 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA), 1–6 (IEEE, 2024). https://doi.org/10.1109/HORA61326.2024.10550519

[20] Singh, K., Yadav, M. & Abdullayev, V.H. Prediction of flight areas using machine learning algorithm. LatIA, **2**, 93–93 (2024). https://doi.org/10.62486/latia202493

[21] Asgarova, B., Jafarov, E., Babayev, N., Abdullayev, V. & Singh, K. Improving cleaning of solar systems through machine learning algorithms. LatIA, **2**, 100–100 (2024). https://doi.org/10.62486/latia2024100

[22] Asgarova, B., Jafarov, E., Babayev, N., Abdullayev, V. & Singh, K. Artificial neural networks with better analysis reliability in data mining. LatIA, **2**, 111–111 (2024). https://doi.org/10.62486/latia2024111

[23] Askerov, T., Abdullayev, V., Abuzarova, V., Niu, Y. & Singh, K. Data processing in Internet of Things networks. LatIA, **2**, 91–91 (2024). https://doi.org/10.62486/latia2024111

[24] Khang, A., Hajimahmud, V.A. & Singh, K. Water quality classification using machine learning algorithms. In Revolutionizing Automated Waste Treatment Systems: IoT and Bioelectronics, 60–76. IGI Global (2024). https://doi.org/10.4018/979-8-3693-6016-3.ch005