MULTIMEDIA COMMUNICATIONS TECHNICAL COMMITTEE
http://www.comsoc.org/~mmc

MMTC Communications-Frontiers

Vol. 19, No. 1, Jan 2024

CONTENTS

SPECIAL ISSUE on Advanced Deep Learning Applications in Image Analysis and Recognition

Guest Editor: Chandan Kumar[1], and Sheng-Lung Peng [2]
[1]*School of Computing, Amrita Vishwa Vidyapeetham, Amaravati, Andhra Pradesh, India*
[2]*National Taipei University of Business,* Taiwan
[1]*k_chandan@av.amrita.edu,* [2]slpeng@ntub.edu.tw

The Advanced Deep Learning Applications in Image Analysis and Recognition explores the forefront of deep learning methodologies applied to image processing. This special issue delves into the intricacies of advanced algorithms and techniques, showcasing their utility in tackling complex tasks such as object detection, segmentation, and image classification. From innovative approaches in medical imaging for disease diagnosis to advancements in surveillance and autonomous systems, the issue provides a comprehensive overview of the diverse applications of deep learning in image analysis and recognition. By highlighting recent innovations and emerging trends, this collection serves as a valuable resource for researchers, practitioners, and industry professionals seeking to harness the power of deep learning for solving real-world image-related challenges and driving technological innovation.

The first paper proposed a novel architecture to enhance retinal images, with a comprehensive approach that integrates multichannel analysis and Particle Swarm Optimization (PSO), specifically tailored for implementation on the Raspberry Pi 4 platform. The utilization of Raspberry Pi 4 brings a practical dimension to the research, making the proposed enhancement methodology not only effective but also deployable in real-time scenarios with portability, aligning with the demand for efficient solutions in medical settings. The focus of the research is on enhancing fundus images for ophthalmological assessments, targeting contrast, color balance, and visual clarity.

In the second paper, shows the use of deep learning techniques for early detection of diseases in the Curcuma longa leaf to be used in medicine. This paper also discusses various methods of deep learning, including the three types of deep learning networks used in the Curcuma longa disease dataset. It also talks about the importance of research and development in vaccines and providing access to safe and effective medicines and vaccines.

The third paper proposes an integrated Face and ID Card Image Detection System using YOLO V8 Deep Learning Methods, which aims to promote responsible behavior among students by identifying and discouraging violations of college rules and enhancing campus security. The proposed methodology uses Yolo V8 deep learning algorithm on a custom dataset that comprises of two classes with Face and ID cards, and it achieves a mean precision accuracy of 93.4% in classifying the two classes.

Dr. Chandan Kumar is currently working as an Assistant Professor (Sl. Gr.) in the Department of Computer Science and Engineering, School of Computing, Amrita Vishwa Vidyapeetham, Amaravati Campus, Andhra Pradesh, India. Earlier, he worked as a Professional Consultant at University Grants Commission (UGC) in the Institution of Eminence (IoE) bureau, New Delhi from July 2019 to September 2021. Before joining UGC, he worked as Lecturer at NIT Jamshedpur from April 2011 to May 2018. He received his Doctor of Philosophy (Ph.D.) from NIT Jamshedpur in 2018. Dr. Chandan Kumar serves as an editorial member for esteemed journals, a conference chair, and a reviewer. He is an active member of professional bodies like IEEE, ACM, IAENG and IACSIT. Dr. Kumar is a well experienced academician & administrator with a strong commitment to education, research, and professional growth. His research interests are Machine Learning, Software Engineering, IoT, Steganography, and Scientometrics Analysis.

Dr. Sheng-Lung Peng is a professor at the National Taipei University of Business, Taiwan, Honorary Professor at Beijing Information Science and Technology University, China, and Visiting Professor at Ningxia Institute of Science and Technology, China. He completed his Doctor of Philosophy in Computer Science from National Tsing Hua University, Taiwan in 1999. His research interests include Designing and analyzing algorithms for Bioinformatics, Combinatorics, Data Mining, and Networks.

Fundus Image Enhancement through Multichannel Analysis and Particle Swarm Optimization using Raspberry Pi 4

Harish Gavel[1], Chandrashekhar Azad[2]
[1,2]Dept of CSE, National Institute of Technology Jamshedpur, India
2021rsca004@nitjsr.ac.in, csazad.cse@nitjsr.ac.in

Abstract

The accurate diagnosis and management of eye diseases rely heavily on high-quality retinal images. This study proposes a novel architecture to enhance retinal images, focusing on improving contrast, color balance, and visual clarity for ophthalmological assessments.

The architecture features: Targeted Blue Channel Processing: Prioritizes the blue channel due to its sensitivity to retinal vasculature, vital for disease detection., Selective CLAHE Application: Enhances low-contrast areas without overprocessing well-contrasted regions., PSO-based Global Refinement: Further optimizes overall image quality across all color channels.

Evaluation results demonstrate significant improvement in image quality metrics like mean color values, PSNR, and SSIM. While further optimizations are needed to improve execution speed, especially on devices like Raspberry Pi 4, the proposed architecture presents a promising approach for enhancing retinal images with clinically relevant improvements. This paves the way for potentially broader adoption in clinical settings and improved diagnosis and management of eye diseases.

Keywords: Retinal image enhancement, contrast improvement, CLAHE, PSO, Raspberry pi, ophthalmology.

1. Introduction

Fundus imaging, an essential diagnostic tool for ocular diseases like diabetic retinopathy [1], faces inherent challenges, including uneven illumination and low contrast [2] in acquired images. To overcome these limitations and contribute to the advancement of medical imaging, our research focuses on innovatively enhancing fundus images. We propose a comprehensive approach that integrates multichannel analysis and Particle Swarm Optimization (PSO)[3], specifically tailored for implementation on the Raspberry Pi 4 platform. The utilization of Raspberry Pi 4 brings a practical dimension [4] to our research, making the proposed enhancement methodology not only effective but also deployable in real-time scenarios with portability, aligning with the demand for efficient solutions in medical settings.

Acknowledging the increasing role of embedded systems in healthcare, our research places a particular emphasis on the practical implementation potential of Raspberry Pi 4. This compact and cost-effective platform offers computational capabilities suitable for real-time processing of fundus images. The integration of our proposed multichannel analysis and PSO-based enhancement [5] approach with Raspberry Pi 4 addresses the need for accessible and efficient solutions, ensuring adaptability to diverse medical environments.

The background section highlights the critical role of fundus imaging in ocular disease diagnostics and emphasizes the shortcomings of conventional approaches. Recognizing the demand for more sophisticated and practical strategies, our research leverages the synergies between multichannel analysis [6], PSO, and the unique capabilities of Raspberry Pi 4.

Multichannel analysis taps into complementary information from different color channels within fundus images, enriching the features for more effective image enhancement. Simultaneously, the integration of PSO optimizes the enhancement process dynamically, adapting parameters to ensure a robust and adaptive approach [5]. The combined utilization of multichannel analysis, PSO, and Raspberry Pi 4 presents a holistic solution to address the challenges faced in fundus image enhancement.

Our study objectives guide the research, aiming to investigate the efficacy of multichannel analysis [14], evaluate the impact of PSO optimization [7], and provide a comprehensive analysis of the proposed approach considering the practical implementation on Raspberry Pi 4. Subsequent sections delve into the methodology, present results, and discuss broader implications within the realm of fundus image enhancement [8], emphasizing the significance of technological integration for real-world applications.

2. Background

Several studies explore metaheuristic algorithms and deep learning for analyzing ocular diseases in medical images. Deep learning approaches excel in detecting and classifying diabetic retinopathy [3], achieving near-perfect accuracy on benchmark datasets. Hybrid models combining deep learning with metaheuristic feature selection (e.g., Simulated Annealing, XGBoost)[5] also show promise in early diabetes diagnosis. New metaheuristic algorithms prove effective in vessel segmentation and optic disc localization in fundus images, outperforming traditional methods with improved accuracy and stability. Bacterial Foraging Optimization emerges as a top performer for precise optic disc localization [9].

While metaheuristics are proving fruitful in various aspects of medical research, their impact shines in two key areas: diagnosing diseases and enhancing medical images. Kaur et al. [10] provide a comprehensive analysis of 10 different metaheuristics and their effectiveness in predicting diseases like heart disease and diabetes, offering researchers a data-driven guide to choosing the optimal approach for each case. On the image enhancement front, Mukhopadhyay et al. [11] showcase an innovative method that leverages a metaheuristic algorithm for automatic parameter tuning in contrast enhancement, while Mohd Azmi et al. [12] unveil NUCE, a swarm-intelligence driven technique that enhances underwater images with exceptional quality and minimal distortion. These advancements demonstrate the vast potential of metaheuristics in revolutionizing both disease diagnosis and medical image analysis.

Retinal image analysis thrives on advancements in image enhancement, and swarm intelligence algorithms are buzzing with promise. From Acharya and Kumar's robust SIAGC [4] using background subtraction and swarm-tuned parameters, to Ghosh et al.'s fuzzy logic [5] and PSO combo for low-contrast enhancement, diverse approaches are revolutionizing clarity. Aurangzeb et al.'s MPSO-optimized CLAHE [2], feeding improved images to deep learning models, boosts vessel segmentation accuracy. Even beyond deep learning, Alhussein et al.'s unsupervised method, weaving CLAHE, PSO, and filtration techniques [13], delivers efficient and precise vessel extraction. Swarm intelligence is clearly swarming with potential, enhancing images and ultimately, our ability to diagnose vision-threatening diseases.

3. Proposed Architecture

This study proposes a novel architecture as shown in fig 1 for enhancing retinal images to facilitate accurate diagnosis and monitoring of eye diseases. The core goal is to improve image quality, particularly in areas with low contrast, while preserving vital details essential for ophthalmological assessment.
The proposed architecture comprises four key components:

1. Channel Splitting and Blue Variance Calculation: The input RGB retinal image is first split into its red, green, and blue channels. Subsequently, the variance of the blue channel is calculated. This choice prioritizes the blue channel due to its sensitivity to retinal vasculature, vital for disease detection.

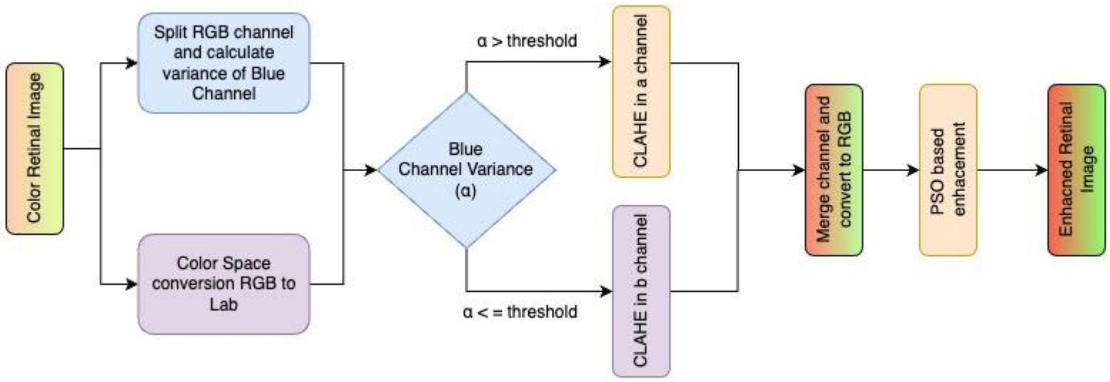

Figure 1. Proposed Architecture

2. Threshold-based Processing: The computed blue channel variance is compared to a pre-determined threshold which was found to be 1500 after careful examination. Images with variance below the threshold are deemed adequately contrasted. Conversely, images exceeding the threshold, typically low-contrast ones, both proceed to the next stage.
3. CLAHE for Contrast Enhancement: For images exceeding the variance threshold, Contrast Limited Adaptive Histogram Equalization (CLAHE) is applied to the blue channel. CLAHE operates on small image tiles, enhancing local contrast within each tile while mitigating noise amplification. This targeted approach preserves crucial details while improving overall image clarity.
4. Particle Swarm Optimization (PSO) for Global Refinement: Finally, both original and CLAHE-processed images are subjected to PSO-based enhancement. PSO optimizes image quality across the entire RGB spectrum, adjusting parameters like brightness, contrast, and color balance. This final step further refines the image and promotes overall visual appeal and diagnostic value.

4. Result

4.1 Dataset Description

Retinal Fundus Multi-disease Image Dataset (RFMiD) [15] is the publicly available dataset which came into the public domain in the year 2021. From then many researchers have used this dataset in their research. To the best of our knowledge, RFMiD is the only dataset that is made available to the public and covers such a broad range of diseases that frequently arise in clinical settings. From the dataset 10 images were chosen for this experiment.

4.2 Experimental Setup

Three machines are used in this experiment as shown in table 1. Their respective configuration is also shown in the table along with its operating system and python version.

Table 1: average execution times of different operating systems on different machines.

Machine	Operating System	Python version	Avg Execution Time(ms)
Raspberry pi 4 (Broadcom SoC, 8GB Ram)	Raspbian	3.9	850ms
Workstation (intel i7 8th gen, 24 GB ram)	Windows 10	3.9	420ms
Macbook Air (M1, 8 GB ram)	Mac OS 14.1	3.8.8	520ms

4.3 Experimental Results

The table presented here in table 2 delves into the quantitative effects of the proposed retinal image enhancement architecture on color distribution. By comparing the mean values of each color channel (Red, Green, and Blue) in both the input and output images, we gain valuable insights into how the processing impacts overall tone and balance. Analyzing these shifts in mean values, particularly for the diagnostically crucial blue channel, allows us to assess the effectiveness of the architecture in restoring lost color information and improving visualization for ophthalmological examinations.

Table 2: Quantitative Analysis of Retinal Image Enhancement.

ID & Class	Mean values (Input Image)	Mean values (Output Image)	PSNR	SSIM	MSE
583 (ODP)	R: 9.81, G: 41.44, B: 84.48	R: 67.27, G:66.52, B: 84.05	15.22	0.769	1955.9
949 (OR)	R: 10.67, G: 36.61, B: 68.11	R: 55.04, G: 55.81, B: 68.11	17.27	0.787	1217.3
918 (MYA)	R: 23.98, G: 61.27, B: 98.8	R: 80.74, G: 79.61, B: 97.47	15.78	0.857	1718.7
133 (EDN)	R: 26.01, G: 67.39, B: 92.0	R: 75.10, G: 78.56, B: 91.13	17.24	0.890	1226.4
665 (ODE)	R: 30.20, G: 53.88, B: 77.11	R: 65.04, G: 66.40, B: 77.00	19.91	0.916	662.9
259 (RS)	R: 40.39, G: 59.44, B: 78.3	R: 66.88, G: 68.50, B: 77.89	22.11	0.956	399.1
345 (CRS)	R: 43.77, G: 49.27, B: 55.4	R:51.09, G: 52.23, B: 55.08	31.01	0.991	51.4
401 (RS)	R: 54.59, G: 70.92, B: 86.5	R: 75.83, G: 77.47, B: 85.77	24.02	0.978	257.4
213 (MS)	R: 63.09, G: 93.81, B: 118.3	R: 98.60, G: 102.26, B: 116.02	19.55	0.961	721.4
464 (ERM)	R: 66.05, G: 78.68, B:100.6	R: 87.07, G: 87.58, B: 100.60	323.90	0.982	265.1

5. Conclusion

The proposed architecture for enhancing retinal images appears to be effective in improving image quality, particularly in terms of contrast and color balance. The mean values for each class in the output image are generally closer to the reference values than the mean values in the input image. This suggests that the proposed architecture is able to successfully restore lost color information in low-contrast areas. The PSNR and SSIM values also indicate that the proposed architecture is able to improve image quality without introducing significant distortion. The PSNR values for all classes are above 15 dB, which is considered to be good quality for medical images. The SSIM values are also all above 0.75, which indicates that the structural similarity between the input and output images is high.

Further optimization efforts are necessary to improve the architecture's execution speed, particularly on lower-powered devices like the Raspberry Pi 4, to enable potential real-world deployment and near-real-time image analysis in clinical settings.

Overall, the proposed architecture presents a promising approach for enhancing retinal image quality while maintaining structural fidelity. The trade-off between execution time and image quality necessitates further investigation and optimization for broader clinical adoption. This could have important implications for the diagnosis and management of eye diseases, as it could allow ophthalmologists to see more detail in retinal images, which could help them to detect subtle abnormalities.

6. Reference

1. Al-Fahdawi S, Al-Waisy AS, Zeebaree DQ, Qahwaji R, Natiq H, Mohammed MA, Nedoma J, Martinek R, Deveci M (2024) Fundus-deepnet: Multi-label deep learning classification system for enhanced detection of multiple ocular

2. Aurangzeb K, Aslam S, Alhussein M, Naqvi RA, Arsalan M, Haider SI (2021) Contrast enhancement of fundus images by employing modified pso for improving the performance of deep learning models. IEEE Access 9:47930–47945, DOI 10.1109/ACCESS.2021.3068477

3. Cetinkaya MB, Duran H (2022) Performance comparison of most recently proposed evolutionary, swarm intelligence, and physics-based metaheuristic algorithms for retinal vessel segmentation. Mathematical Problems in Engineering 2022:1–25

4. Acharya UK, Kumar S (2021) Swarm intelligence based adaptive gamma corrected (siagc) retinal image enhancement technique for early detection of diabetic retinopathy. Optik 247:167904, DOI https://doi.org/10.1016/j.ijleo.2021.167904

5. Ghosh SK, Biswas B, Ghosh A (2020) A novel approach of retinal image enhancement using pso system and measure of fuzziness. Procedia Computer Science 167:1300–1311, DOI https://doi.org/10.1016/j.procs.2020.03.446, international Conference on Computational Intelligence and Data Science

6. Sule OO, Ezugwu AE (2023) A two-stage histogram equalization enhancement scheme for feature preservation in retinal fundus images. Biomedical Signal Processing and Control 80:104384, DOI https://doi.org/10.1016/j.bspcdiseases through data fusion of fundus images. Information Fusion 102:102059, DOI https://doi.org/10.1016/jnffus.2023.102059

7. Dayana AM, Emmanuel WS (2023) A comprehensive review of diabetic retinopathy detection and grading based on deep learning and metaheuristic optimization techniques. Archives of Computational Methods in Engineering pp 1–35

8. Khanna JPSAK (2018) Metaheuristic techniques for detection of optic disc in retinal fundus images

9. Kaur S, Kumar Y, Koul A, Kumar KambojS (2023) A systematic review on metaheuristic

optimization techniques for feature selections in disease diagnosis: open issues and challenges.

10. Archives of Computational Methods in Engineering 30(3):1863–1895
11. Mukhopadhyay S, Hossain S, Malakar S, Cuevas E, Sarkar R (2023) Image contrast improvement through a metaheuristic scheme. Soft Computing 27(18):13657–13676
12. Mohd Azmi KZ, Abdul Ghani AS, Md Yusof Z, Ibrahim Z (2019) Natural-based underwater image color enhancement through fusion of swarm-intelligence algorithm. Applied Soft Computing 85:105810, DOI https://doi.org/10.1016/j.asoc.2019.105810
13. Alhussein M, Aurangzeb K, Haider SI (2020) An unsupervised retinal vessel segmentation using hessian and intensity based approach. IEEE Access 8:165056–165070, DOI 10.1109/ACCESS.2020.3022943
14. Huang SC, Cheng FC, Chiu YS (2012) Efficient contrast enhancement using adaptive gamma correction with weighting distribution. IEEE transacions on image processing 22(3):1032–1041
15. Pachade S, Porwal P, Thulkar D, Kokare M, Deshmukh G, Sahasrabuddhe V, Giancardo L, Quellec G, M´eriaudeau F (2021) Retinal fundus multidisease image dataset (rfmid): A dataset for multidisease detection research. Data 6(2), DOI 10.3390/data6020014, URL https://www.mdpi.com/2306-5729/6/2/14

AuthorsBio-data

Harish Gavel is a Research Scholar at the Department of Computer Science and Engineering, National Institute of Technology, Jamshedpur, India. He completed Master of Technology (Mtech) specialising in Information Systems and Security Engineering from the same institute where he was the topper of his batch and received a silver medal in 2021. In 2014, he completed his Bachelor's in Electrical Engineering from the Bhilai Institute of Technology, Durg, India. Harish's research interests revolve around Machine Learning, Deep Learning, Pattern Recognition, Image Processing, and Medical Image Enhancement.

Dr. Chandrashekhar Currently working as an Assistant Professor (Grade I). in the Department of Computer Science & Engineering at National Institute of Technology, Jamshedpur. He has published more than 50 research papers in journals and conferences of international repute. He serves as a reviewer for various international journals and conferences. His primary area of research includes Intrusion Detection System, Meta Heuristic Algorithms, Swarm Intelligence, Pattern Recognition, and Machine Learning**.**

Training Issues in Early Disease Detection in Curcuma longa Leaf using Deep Learning Techniques

Ranjith Kumar Banala[1] and Rajesh Duvvuru[2]
[1,2]*School of computer Science and Engineering, VIT-AP University, Guntur, India*
[1]*ranjith.22phd7037@vitap.ac.in*, [2]*rajesh.duvvuru@vitap.ac.in*

Abstract

After the COVID-19 plague, there is a huge need for plants that can be used as medicine. Students these days depend on medicines based on AYUSH because they work better and have fewer side effects. Curcuma longa, or turmeric, is a well-known herb that is used in both Ayurvedic and allopathic medicine. India grows a lot of Curcuma longa, which is also known as turmeric. But because of changes in the weather, food production has been slowly going down lately. At the moment, Curcuma longa (Turmeric) needs to be able to spot diseases early on for both home and medical use. We looked into the best Deep Learning (DL) training network that could get 100% accuracy on the Curcuma longa (Turmeric-Duggirala variant) disease dataset, which has pictures of leaf blotch, Colletotrichum leaf spot, Cercospora leaf spot, and leaf blight. There are three different kinds of deep learning networks used: SqueezNet, GoogleNet, and ResNet-50. The training settings (epochs and Learning Rate (LR)) and optimizers (SGDM, ADAM, and RMSProp) are all fine-tuned.

1. Introduction

In this case of world climate change, technology-based methods are thought to be the best way to get better crop yields. The Internet of Things (IoT) and artificial intelligence (AI) are two famous areas that help traditional farming methods with health information and food security. India, China, the United States, and Brazil, which depend a lot on farmland, need smart, efficient ways to make food and medicine last. One of the seventeen global goals of the 2030 Agenda for Sustainable Development [1] is to improve health and well-being. This is written in the United Nations' Sustainable Development Goal (SDG) three. SDG #3, which is about health, wants to make sure that everyone is healthy and happy [2]. As part of this, they have promised to end outbreaks of infectious illnesses like AIDS, tuberculosis, and malaria by 2030. It also aims to provide Universal Health Coverage (UCH) and make sure that everyone has access to safe and effective medicines and vaccines [3]. For this process to work, it is important to fund research and development on vaccines and make sure that people can get medications at low prices.

2. Background Works

One important area of study is early disease identification, and people are working hard to get 100% classification accuracy in this area. Orchi et al. did a thorough study on how AI and IoT can be used to find field diseases. Their poll showed that 97% of the time, these methods were able to correctly classify the diseases [4]. The disease detection work by Thangaraj et al. was done on the Tomato leaf disease dataset. Their research showed that 52% of current leaf disease detection methods are accurate to 90%, while 48% of current studies are less than 90% [5]. The study also found that AlexNet was 98.6% accurate at classifying diseases. It was also looked at on GoogleNet (99.18%), ResNet (98.8%), Xception (98.13%), and VGG16 (99.25%) to see which had the highest accuracy. However, it has been seen that none of the deep learning architectures are 100% accurate. Math et al. used a deep convolution network to find grape disease datasets early on, and the classification model got 99.34% of the time [6]. Naeem et al.

used machine learning to identify medical plant leaves based on multispectral and textural features. Here are the percentages they got: On the other hand, 99.80% of peppermint, 99.10% of tulsi, 98.40% of bael, 99.20% of stevia, and 99.90% of lemon balm were found to be safe [7]. Sai et al. suggested using Convolutional Neural Network (CNN) methods like ResNet101, InceptionV3, and VGG16 to find diseases in healing plants. On the Ayur Bharat dataset, InceptionV3 achieved 97.32% classification accuracy.

3. Methodology

The planned work is mostly about putting diseases in turmeric leaves into groups. At first, pictures of the leaves are gathered for three different diseases: leaf spot, leaf blotch, and rhizome rot. These pictures are then saved in a database so they can be used later. The database has 800 pictures of leaves, with 200 pictures of sick leaves in each disease group and 200 pictures of healthy leaves in the other category. The suggested method is made up of the following steps.The given picture is resized and turned into HSI photos during the pre-processing step. Only the hue part is taken for further processing. The K-means segmentation method is used to separate the diseased part of the original picture during the segmentation process. After the segmentation is done, color, texture, and shape features are added to create feature vectors. For correct classification using the ranking method, a group of the best features is chosen from the set of features.

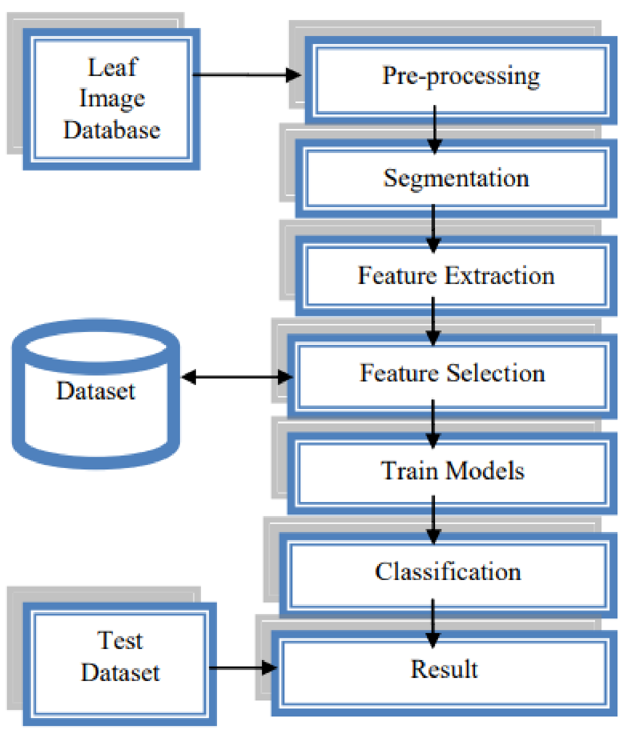

Figure 1. Methodology of Curcuma longa leaf disease system

Pre-Processing

The pictures of turmeric leaves are gathered by hand in DUGGIRALA,KOLLIPARA, and Mandadam. A high-resolution camera with 12 mega pixels was used to take the pictures. Since the pictures that were gathered are of different sizes, they need to be changed to the same size so that pre-processing can go smoothly. The approximation method is used to change the size of the pictures to 256*256.

Segmentation

An important part of digital picture processing is image segmentation. It could be thought of as the process of giving pixels to areas that are similar but not connected. These areas make up a partition of the picture that shares some visual traits. The main purpose of segmentation is to clear up or change the way a picture is shown so that it is more meaningful and easy to look into. The main idea behind segmentation is that it divides a picture into close-up items that can be moved around. The kmeans segmentation method is used to do the segmentation in this suggested work.The K-Means segmentation method divides the input data points into several groups based on how far apart they are. The program thinks of the data features as being in a vector space and tries to group them in a way that makes sense.

4. Results and Discussion

Overall, the mean epochs (25–50) and mean LR (0.01–1) scored well. The 'L2Norm' vector measuring method is used to find out how well ResNet-50 (SGDM) works. It's 48.8% better than SqueezNet (SGDM) and 24.55% better than GoogleNet (SGDM). ResNet-50 (ADAM) is 36.74% more accurate at training than SqueezNet (ADAM) and 28.1% more accurate than GoogleNet (ADAM) when ADAM is used. Last but not least, ResNet-50 (RMSProp) performs 37.13% better than SqueezNet (RMSProp) and 40.88% better than GoogleNet (RMSProp).

The 'Global_L2Norm' vector measuring method is used to find out how well ResNet-50 (SGDM) works. It is 49.16.8% better than SqueezNet (SGDM) and 24.47% better than GoogleNet (SGDM). ResNet-50 (ADAM) is 38.29% more accurate at training than SqueezNet (ADAM) and 29.88% more accurate than GoogleNet (ADAM) when ADAM is used. Last but not least, ResNet-50 (RMSProp) performs 32.28% better than SqueezNet (RMSProp) and 36.63% better than GoogleNet (RMSProp).

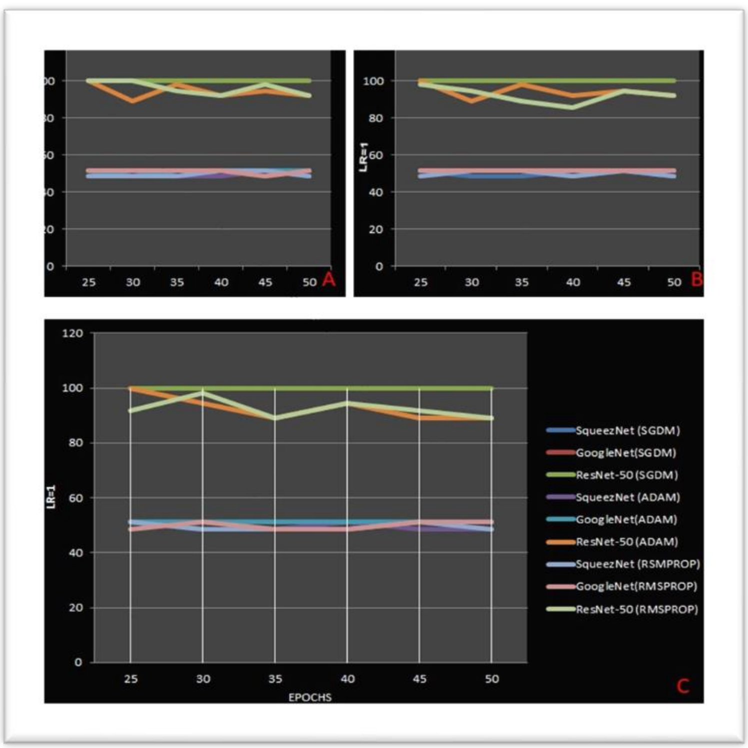

Figure 2. Performance analysis of overall DeepLearning Optimizers on Curcuma longa dataset

5. Conclusion

The study created a model for finding diseases and sorting turmeric leaves into different groups. Several image analysis and machine learning methods are used in the work. It was done with the MATLAB software and machine learning methods like Support Vector Machine (SVM), Decision Tree (DT), and Naive Bayes (NB). The finding of 10-fold cross validation showed that this study was correct. The study finds that Support Vector Machine (SVM) is more accurate than other methods at classifying diseases on turmeric leaves (99.75%). In the future, the study can be expanded by using different ways to cut out the sick part of the original leaf.

6. Reference

1. Tsalis, T. A., Malamateniou, K. E., Koulouriotis, D., & Nikolaou, I. E. (2020). New challenges for corporate sustainability reporting: United Nations' 2030 Agenda for sustainable development and the sustainable development goals. Corporate Social Responsibility and Environmental Management, 27(4), 1617-1629.
2. Khetrapal, Sonalini, and Rajesh Bhatia. "Impact of COVID-19 pandemic on health system & Sustainable Development Goal 3." The Indian journal of medical research 151.5 (2020): 395.
3. Martinez-Millana, Antonio, et al. "Artificial intelligence and its impact on the domains of universal health coverage, health emergencies and health promotion: An overview of systematic reviews." International Journal of Medical Informatics 166 (2022): 104855.
4. Mattihalli, Channamallikarjuna, et al. "Plant leaf diseases detection and auto-medicine." Internet of Things 1 (2018): 67-73.
5. Modak, Manisha, et al. "Indian herbs and herbal drugs used for the treatment of diabetes." Journal of clinical biochemistry and nutrition 40.3 (2007): 163-173.
6. Shankar, Darshan, and Bhushan Patwardhan. "AYUSH for New India: Vision and strategy." Journal of Ayurveda and integrative medicine 8.3 (2017): 137.
7. Villena-Tejada, Magaly, et al. "Use of medicinal plants for COVID-19 prevention and respiratory symptom treatment during the pandemic in Cusco, Peru: A cross-sectional survey." PloS one 16.9 (2021): e0257165.

Author’s Bio-data

Ranjith kumar Banala receive the Bachelor degree in Information Technology From Acharya Nagarjuna University;A.P,INDIA, and the Master of technology in Computer Science and Engineering From JNTUK University, India. He is currently pursuing the Ph.D. degree in Computer Science at VIT-AP University, INDIA. He worked in various engineering colleges in Andhra Pradesh. His Research Interests include Artificial Intelligence, Machine learning, Deeplearning and Disaster Management. He can be contacted at email: ranjith.22phd7037@vitap.ac.in.

Rajesh Duvvuru received the Bechlore of Technology from S.V.University, Tirupati. He pursued Master of Technology from A.N.University, Guntur and Ph.D. in Geospatial Intelligence from Andhra University, India. He is a former Assistant Professor in Department of Computer Science and Engineering, National Institute of Technology, Jamshedpur, INDIA. He worked as a software trainee in Virtuoso Soft Technologies, Chennai. He worked in various engineering colleges in Andhra Pradesh. Presently he is Assistant Professor (Sr. Grade-1) at the School of Computer Science and Engineering, VIT-AP University. His research interests include Deep Learning, IoT, Computer Vision, Disaster Management, Computer, and communication Networks. He can be contacted at email: rajesh.duvvuru@vitap.ac.in.

Integrated Face and ID Card Image Detection System using YOLO V8 Deep Learning Methods

***Rajesh Duvvuru*[1], *Eswaraiah Rayachoti* [2], *Rajiv Jetson*[3]**

[1,2]*School of computer Science and Engineering, VIT-AP University, Guntur, India*

[3]*Department of Computer Science and Engineering, KHIT, Guntur, India*

[1]*rajesh.duvvuru@vitap.ac.in*, [2]*eswaraiah.rayachoti@vitap.ac.in*, [3]*drrajivjetson@gmail.com*

Abstract

Integrated Face and ID Card Image Detection System recognizes that cultivating good habits during college not only enhances the educational environment but also prepares students for professional settings where adherence to organizational rules is paramount. The present work promotes adherence to college regulations but also enhances security measures while mitigating conflicts between students and security personnel. The proposed methodology uses Yolo V8 deep learning algorithm on student custom dataset that comprises of two classes with Face and ID cards. The mean precision accuracy (mPA) accuracy is 93.4% in classifying the two classes, that is classified as good accuracy.

1. Introduction

In today's fast-paced world, instilling discipline and adherence to organizational rules and regulations is essential, even during the early stages of one's academic journey. The 'Integrated Face-ID Detection System' (IFDS) work aims to promote responsible behavior among students by identifying and discouraging violations of college rules, particularly the requirement to wear identification (ID) cards [5]. This project recognizes that cultivating good habits during college not only enhances the educational environment but also prepares students for professional settings where adherence to organizational rules is paramount. Colleges and educational institutions have established various rules and regulations to maintain discipline, safety, and security on their campuses. One common requirement is for students and staff to wear identification (ID) cards at all times while on college premises. Security guards are often tasked with the responsibility of checking and ensuring that every individual within the premises is wearing a valid ID card. This manual verification process is not only time-consuming but can also lead to confrontations and disagreements between students and security personnel. The present study proposes a novel student ID card detection mechanism using deep learning studies [6]. Detection models are always a challenging task, especially, wherever there are multiple people present in one scenario. Develop an automated system for educational institutions to enforce ID card compliance and promote responsible behaviour. The system should combine advanced face recognition, ID card detection, and automated messaging to identify individuals without ID cards and send fines or reminders, accordingly, reducing the burden on security personnel and enhancing campus security.

2. Background Works

This paper is the first to provide an in-depth review of the YOLO evolution from the original YOLO to the recent release (YOLO-v8) from the perspective of industrial manufacturing. The review explores the key architectural advancements proposed at each iteration, followed by examples of industrial deployment for surface defect detection endorsing its compatibility with industrial requirements [1]. Deep learning-based face recognition models have achieved state-of-the-art performance on a variety of benchmark datasets. Schroff et al. (2015) proposed a

'FaceNet: A Unified Embedding for Face Recognition and Clustering' that introduces a deep learning-based face embedding algorithm that can be used for both face recognition and clustering [2]. The algorithm achieves an accuracy of 99.63% on the Labeled Faces in the Wild (LFW) dataset, which is a popular benchmark dataset for face recognition. But it is limited to face class working two different classes i.e., id cards with faces is challenging task. Redmon et al. (2018) proposed improved a YOLOv3 for a real-time object detection algorithm that can detect a variety of objects, including ID cards, with high accuracy [3]. The algorithm achieves an average precision of 80.6% on the Microsoft Common Objects in Context (COCO) dataset, which is a popular benchmark dataset for object detection. Behaviour promotion: There are a variety of mechanisms that can be incorporated into integrated face detection and recognition systems to promote responsible behaviour. Wang et al. (2018) proposed a Face Recognition-Based System for School Attendance Monitoring and Behaviour Analysis" to recognize the facial based system for school attendance monitoring and behaviour analysis [4]. The work primarily focused on system is able to identify and track students in real time and to generate reports on student attendance and behavior.

3. Methodology

The key components of your system, such as face recognition, ID card detection, behaviour promotion, database integration, automated messaging, and user-friendly interface, are all well-established and widely used in existing systems (Figure 1).

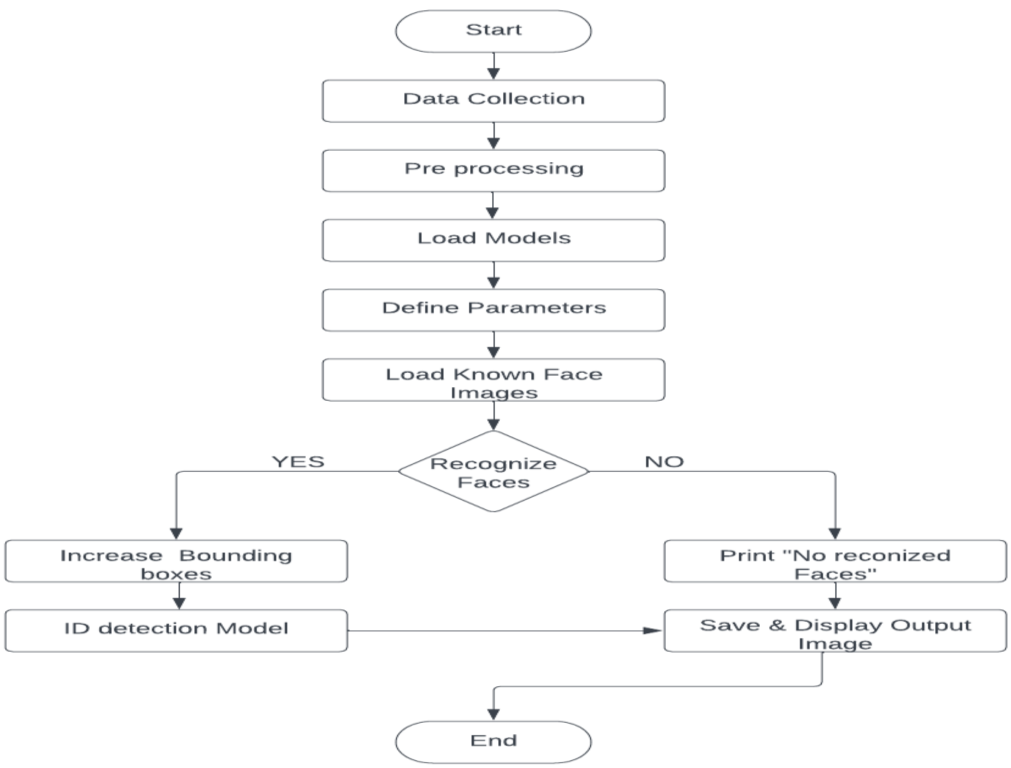

Figure 1. Methodology of Facial Student Authentication Recognition System using YOLO-V8

The Face Recognition module deals with state-of-the-art deep learning techniques, specifically the YOLOv8 architecture and the face recognition library, to accurately detect and recognize faces within a given frame. *Data Collection and Preparation*: Gather a dataset of images containing students, collect images of student ID cards for training the ID card detection model.

Organize student data, including names, registration numbers, and phone numbers, in an Excel spreadsheet. *Face Recognition Model:* Utilize the Face Recognition library to build a deep learning model capable of recognizing faces. Train the model on the student image dataset to create face embeddings. *ID Card Detection Model:* Employ the YOLOv8 architecture to create a custom model for detecting ID cards. Train the model on the ID card image dataset to identify ID cards in images. *Automated Processing:* Implement a script to process input images. Use face recognition to detect faces in the frame. If faces are recognized, send bounding box coordinates to the "Increase Bounding Box" function. Crop the region of interest based on expanded bounding boxes. Use the ID card detection model to check for ID cards in cropped images. If an ID card is detected, mark the bounding box with green; otherwise, mark it with red. Extract the name of the person in the bounding box. *Database Integration:* Connect the script to the Excel spreadsheet containing student data. When an ID card is not detected, compare the person's name with the database.

4. Results and Discussion

The project focuses on addressing a common challenge faced by educational institutions worldwide - ensuring that students comply with identification card (ID card) policies. College campuses often have a large student population, making it challenging for security personnel to manually verify that every student is wearing their ID card. This challenge can result in breaches of security and conflicts between students and security personnel. To tackle this issue, the project integrates cutting-edge technologies such as YOLOv8 for object detection and the face recognition library for facial recognition. By leveraging these technologies, the system automates the process of identifying individuals without ID cards. One of the significant improvements over traditional manual methods is the use of bounding box enhancement. The system dynamically adjusts the bounding boxes around detected faces, allowing for precise detection of ID cards. This ensures that even if an ID card is partially obscured or not prominently displayed, it can still be detected.The screenshot of the student images detected one with wearing ID card and other one with no ID card as shown in Figure 2. The test results accuracy of Recall-confidence curve is 98% and mean Precision Accuracy (mPA) is 92.4%, which is considered to be good classification results (Figure 3).

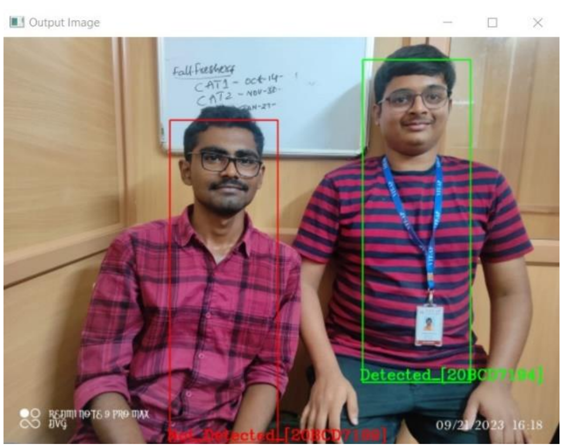

Figure 2. Output screenshot of Integrated Face and ID Card Image Detection System detected images using YOLOv8

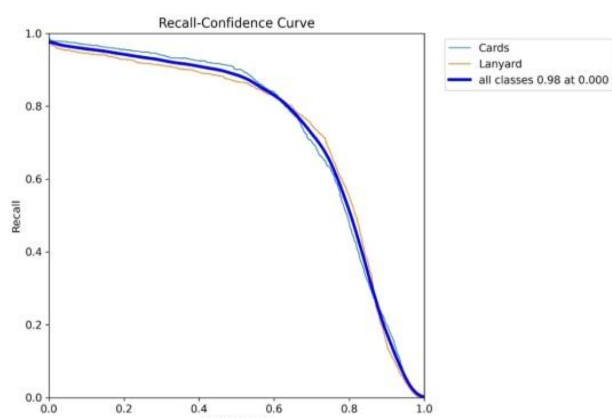

Figure 3. Recall-confidence curve classification accuracy is 0.98 at 0.0

5. Conclusion

The integrated Face-ID detection system developed for enhancing college security and compliance with organizational rules has proven to be a significant technological advancement. The key highlight of this project is Automated Detection. The system automates the process of identifying individuals without ID cards, relieving the burden on college security personnel. Bounding Box Enhancement: It dynamically adjusts bounding boxes around faces, allowing for precise ID card detection. Database Integration: The system integrates seamlessly with an Excel database, matching detected individuals with student records. The test results mPA is 93.4 %, which is considered to be good classification results.

6. References

1. Al-obidi, Dhufr, and Seydi Kacmaz. "Facial Features Recognition Based on Their Shape and Color Using YOLOv8." In 2023 7th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), pp. 1-6. IEEE, 2023.
2. Schroff, F., Kalenichenko, D., & Philbin, J. (2015). Facenet: A unified embedding for face recognition and clustering. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 815-823).
3. Redmon, J., & Farhadi, A. (2018). Yolov3: An incremental improvement. arXiv preprint arXiv:1804.02767.
4. Cao, F., Wang, M., & Wang, K. (2018, November). The classroom attendance management system of face recognition based on LBS. In 2018 5th International Conference on Education, Management, Arts, Economics and Social Science (ICEMAESS 2018) (pp. 955-960). Atlantis Press.
5. Pratama, M. O., Satyawan, W., Fajar, B., Fikri, R., & Hamzah, H. (2018, October). Indonesian ID card recognition using convolutional neural networks. In 2018 5th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI) (pp. 178-181). IEEE.
6. Kushal, M., BV, K. K., MJ, C. K., & Pappa, M. (2020, July). ID Card Detection with Facial Recognition using Tensorflow and OpenCV. In 2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA) (pp. 742-746). IEEE.

AuthorsBio-data

Rajesh Duvvuru received the Bechlore of Technology from S.V.University, Tirupati. He pursued Master of Technology from A.N.University, Guntur and Ph.D. in Geospatial Intelligence from Andhra University, India. He is a former Assistant Professor in Department of Computer Science and Engineering, National Institute of Technology, Jamshedpur, INDIA. He worked as a software trainee in Virtuoso Soft Technologies, Chennai. He worked in various engineering colleges in Andhra Pradesh. Presently he is Assistant Professor (Sr. Grade-1) in School of Computer Science and Engineering, VIT-AP University. His research interest includes, Deep Learning, IoT, Computer Vision, Disaster Management, Computer and communication Networks. He can be contacted at email: rajesh.duvvuru@vitap.ac.in.

Dr. Eswaraiah Rayachoti received Bachelor of Technology from Sri Venkateswara University, Andhra Pradesh. He received Master of Technology from JNTUH, Hyderabad and Ph.D. in Computer Science and Engineering from Acharya Nagarjuna University, Andhra Pradesh. He worked in various engineering colleges in Andhra Pradesh. Presently he is Professor in School of Computer Science and Engineering, VIT-AP University. His research interests include, Digital Image Processing and Computer Vision. He can be contacted at email: eswaraiah.rayachoti@vitap.ac.in

Dr. Rajiv Jetson Vemula received his double Masters in Computer Science from Nagarjuna University & BIHER and his Research Degrees M.Phil & Ph.D in Computer Science from Bharathi Dasan University, Trichy & ANU, AP respectively. He is currently working as Professor & Head in the Department of Computer Science & Engineering in Kallam Haranadha Reddy Institute of Technology, Guntur, AP, India since 2012. He has published more than 20 research papers in reputed international journals and conferences and it's also available online. His main research work focuses on Sftware Engineering & other emerging areas like IOT, Cloud etc. He can be contacted at email: drrajivjetson@gmail.com