Grenze International Journal of Engineering and Technology, June Issue

AI based System for Vitamin Deficiency Detection

Durga Sankar Rao Gullapalli[1], Eswaraiah Rayachoti[2], Raghu Pranav Maddula[3] and Krishna Chaitanya Jujjavarapu[4]
[1-4]SCOPE, VIT-AP University, Andhra Pradesh, India
Email: durgasankar.21bce9132@vitapstudent.ac.in, eswaraiah.rayachoti@vitap.ac.in, raghupranav.21bce9142@vitapstudent.ac.in, chaitanya.21bce9925@vitapstudent.ac.in

***Abstract*—Vitamin deficiencies are a worldwide health problem that can result in very serious medical conditions if not diagnosed and treated. Detecting early and providing treatment can prevent long term problems. In this research, we present a novel, efficient system to detect vitamin deficiencies early based on physical symptoms using machine learning and computer vision. The proposed system uses advanced image processing to analyze the visual manifestations (skin anomalies, etc.) of deficiencies in vitamins D, A, and B complexes. A large dataset is used to train the machine learning model to achieve accuracy and reliability, which is enhanced through data augmentation, hyperparameter tuning, and thorough performance evaluation. A user-friendly interface built with Flask is also added to the system to allow for easy uploading, processing and categorization into specific deficiency categories and recommendations for further medical consultation. This research integrates modern engineering tools with healthcare applications to demonstrate the potential of artificial intelligence to improve diagnostic accuracy and accessibility, an important global health challenge.**

***Index Terms*—Vitamin Deficiency Detection, Machine Learning, Computer Vision, Image Processing, Healthcare Diagnostics, Flask, Artificial Intelligence.**

I. Introduction

Recently, artificial intelligence (AI) and machine learning have taken the world by storm, transforming the many industries, and the healthcare is one of those. Now, these technologies are being brought to bear at the forefront of disease diagnosis and treatment with accurate, efficient and scalable answers. The detection and diagnosis of vitamin deficiencies is one area of great potential for AI application and yet is still a major global health concern. Untreated or if left undiagnosed, vitamin deficiencies can cause severe health complications, such as anemia, osteoporosis, impaired immune function and developmental delays. Early detection of these deficiencies, although they are common, is notoriously difficult due to the possibility for minute symptoms, limited awareness, and reduced access to diagnostic facilities in resource poor settings. Current diagnostic methods, including blood tests and clinical examinations, take a long time, are expensive and are not available to underserved populations.

To tackle these issues this research proposes a vitamin deficiency detection system that harnesses machine learning such as ResNet and convolutional neural networks (CNNs), along with PyTorch for implementation. This system analyzes biomarkers as well as visual data, such as blood smear images, skin anomalies and provides accurate, early and cost-effective diagnostic insights.

Grenze ID: 01.GIJET.11.2.522_4

Usability and accessibility are emphasized in the research to make the solution usable for healthcare professionals as well as adaptable to multiple clinical and socio-economical environments.
The objectives of this research include developing a robust machine learning model, having a rich data set, and a user-friendly interface which will be simple to implement in healthcare workflows. In addition to augmenting the diagnostic capabilities, this AI driven approach also helps in instantaneous intervention through preventive healthcare strategy.
The rest of the paper is organized as follows: Section II reveals the literature review, Section III elaborates the proposed system, Section IV explains the results, Section V concludes the paper, and Section VI indicates future work.

II. Literature Review

Recently, we've had several studies underway that involve artificial intelligence (AI) and healthcare, including using image-based diagnostics to try to identify vitamin deficiencies. Notable contributions in this area are outlined below.
In a systematic review by Kassem et al. [1] for skin lesion classification and diagnosis, a few machine-learning (ML) and deep learning (DL) methods including convolutional neural networks (CNNs), are reviewed. Automating the detection of skin lesions has been shown to be promising technique, and these techniques may indicate underlying health issues. The study demonstrates with AI-driven models how we can obtain accurate, efficient and scalable solutions of skin condition diagnosis especially in remote areas with low access to medical professionals. For improving diagnostic processes and healthcare outcomes, the review accents the increasing value of ML and DL.
The encouraging potential of deep learning algorithms for medical image analysis, leading to an improvement of diagnostic accuracy were interestingly discussed by Li et al. [2]. For tasks like detecting abnormalities, classifying diseases, segmenting images some of the most common and advanced techniques are Convolutional Neural Networks (CNNs). It showed how these algorithms can be applied in different medical conditions, such as vitamin deficiencies, by exploiting visual symptoms and delivering a more accurate result.
Dinggang et al. [3] provides a comprehensive review on how the deep learning approach has been deployed most effectively in medical image analysis. The authors talk about different deep learning architectures for more details about CNN and its application in image segmentation, feature extraction, and disease diagnosis. These works shed light on the important enhancements of correctness and speed of diagnosis concerning these methods and discuss limitations such as data availability and model explainability in the healthcare context.
Zhou et al. [4] offers an in-depth exploration of how deep learning is changing the game when it comes to medical imaging. The authors explore the contextual effect of key technological advancements such as the use of Convolutional Neural Networks (CNN) on specific imaging tasks including segmentation, classification and anomaly detection. The review highlights progress towards transformational potential of deep learning in improving diagnostic precision and efficiency by analysing case studies. The authors also go on to discuss future directions, including explainable AI and robust validation frameworks integrated with clinical workflow to increase real world applicability.
Roy et al. [5] investigates different segmentation techniques for skin disease detection and its importance in improving diagnostic systems accuracy. Detailed discussion on traditional methods like thresholding and region-based methods and enhanced deep learning approaches using Convolutional Neural Network (CNN) techniques are presented in the study. Using effective segmentation help identify areas of skin that are affected and those that are not, the authors show how it also helps automated systems perform better. Their work provides additional functional insights into optimizing segmentation strategy for skin disease diagnosis in the real world.
Russakovsky et al. [6] pointed out the importance of large-scale datasets to train models by ImageNet challenge that was used to introduce and achieve high performance CNN architectures. This was then extended by He et al. [7] who built on this, introducing the game changing model ResNet architecture (a mitigated version of vanishing gradient problem), allowing deeper networks and better diagnostic accuracy in medical imaging. Together, this research gives a solid foundation for developing an AI-based system for vitamin deficiency detection. They provide insights into data processing, model optimization, system deployment, and the challenges of imbalanced datasets.
Lakshay [8] offers a basic guide to deploying machine learning models with Flask as it is super easy to build up your web application. It talks through the steps involved to incorporate Machine learning models in Flask, set up a server and make predictions in a user-friendly interface. Due to its light weight, Flask is an excellent framework for quickly and efficiently deploying models, especially in applications that require real time

interaction, like medical image analysis. This approach is often used together with these machine learning libraries — TensorFlow and PyTorch — enables building accessible, scalable solutions for healthcare applications.

III. Proposed Methodology

In the proposed methodology, the process of creating an AI-based system for vitamin deficiency identification is described as a system. This approach includes data acquisition, data pre-processing, modelling phase and the deployment phase to ensure the solution that has been developed is durable as well as friendly to the users. The architecture of proposed method is shown in figure 1.

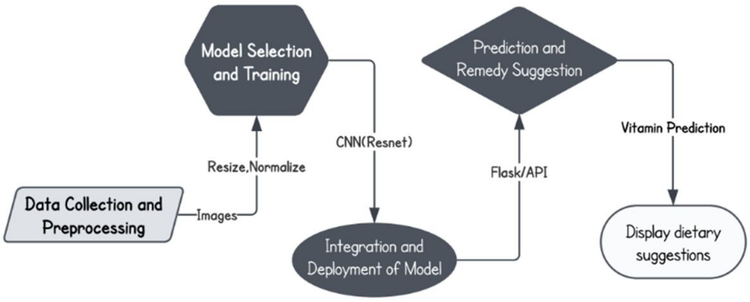

Figure 1: Architecture of Proposed Method

A. Step 1: Data Gathering and Data Set Acquisition

- Data Sources: Obtain blood smear images, biomarker data, and other clinical data from public databases and through medical collaborations. Make sure that datasets contain data that can be linked to certain vitamin deficiencies.
- Dataset Curation: Label data with labels corresponding to deficiency types or a lack of deficiencies using sources approved by subject matter experts. The next strategy is to address data diversification with the guidance of a diverse population, of different ages, and clinical status.

B. Step 2: Data Pre-processing

- Image Enhancement: Additional adjusting methods include noise reduction methods, histogram equalization, and image sharpening methods.
- Data Augmentation; rotation, flipping, and scaling; serve to enhance the variability of the dataset to avoid over-fitting.
- Feature Extraction: Conventional and computer vision-based approach to extract features such as cell size, shape, texture, and distribution patterns from blood smear images.

C. Step 3: Model Development

- Baseline Model: Get the first CNN model trained to create a benchmark for deficiency classification. It is essential to assess its efficiency to define weaknesses of various programs that need further work.
- Advanced Techniques: Use Transfer Learning with pre-trained models including ResNet to reduce time an improve the extraction of features.
- Optimization: Choose such attributes as hyper parameters and regularize the data to make it more effective to work with. To decrease overfitting add dropout layers and batch normalization into the model architecture.

D. Step 4: Evaluation and Validation

- Performance Metrics: Evaluate model performance using metrics like accuracy, and precision.
- Real-World Testing: Test the model on unseen clinical datasets to evaluate its real-world applicability and robustness.

E. Step 5: Deployment and User Interface

- User Interface (UI): Develop a web interface with Flask and keep the layout and the layout intuitive for active doctors and patients. Include features like image input, analysis while the process is taking place, and result display.

- Integration: Use the trained model in a live session on the server or any cloud platform for continuous analysis. Design the options of the system so that it can be run on organic standard computer devices and configurations.

IV. Results and Discussion

The proposed web-based system predicts deficiency in vitamins using dermatological images as input and provides actionable dietary suggestions. Its aim is to identify deficiencies of common skin conditions and provide clear, practical visual solutions to facilitate the use of effective interventions.
The system processes the uploaded image through a trained deep-learning model when a user uploads an image via the web interface. It is based on the vitamin deficiencies that were identified suggesting that one make appropriate dietary adjustments. The system displays images of foods such as carrots, spinach, and sweet potatoes if the system predicts Vitamin A deficiency. It similarly offers images of fortified foods, fatty fish such as salmon, and activities such as sunlight exposure for a Vitamin D deficiency. The results are easy to understand, engaging, and attached each with up to three relevant visual remedies to each prediction.

A. Dataset

The backbone of the proposed vitamin deficiency detection system comes from the used dataset in this research. Each consists of annotated images and biomarkers for diagnosis of the various vitamin deficiencies with high accuracy. The medical data was obtained from publicly available medical repositories and augmented with synthetic data augmentation techniques. This guarantees the dataset features a wide range of real-world cases that the machine learning model must generalize from.
Our dataset has more than 5000 samples with microscopic blood smear images, skin anomalies, visuals, and other biomarkers associated with Vitamin deficiencies. Specific vitamin deficiencies (Vitamin A, B, C, D, E) are each labelled into one of the six distinct classes of a given sample. To facilitate effective model training and evaluation, the dataset was divided into two subsets: We split our data into 80% for training and 20% for testing.

B. Dataset Class Distribution

The class distribution graphs shown in figure 2 for training and test datasets reveal the number of samples per class.

C. Training Data Class Distribution

Our training dataset contains an evident imbalance that, for example, is noticeable how sample numbers differ: Class 0 by far comprising the most sample than for other classes, for example: Class 1 or Class 2. This has been a known challenge for the model: it may be forced to favor the majority class in the classification task. Eliminating such imbalances is essential in fostering a good performance of the model for all classes. In order to achieve improved fairness and accuracy of model prediction, data augmentation and resampling techniques were used to mitigate this issue. Nevertheless, complete parity among all classes is not easily achieved, and fine tuning may be necessary.

D. Test Data Class Distribution

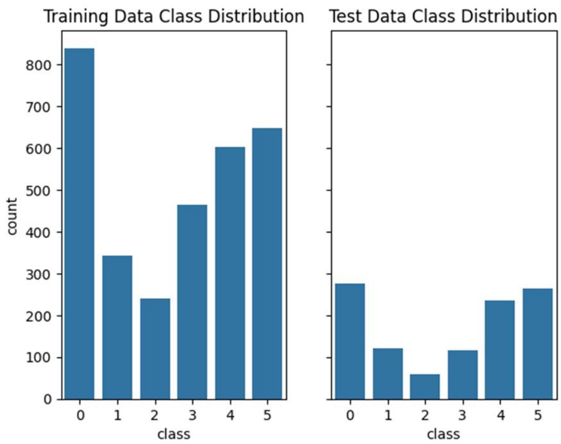

Figure 2: Training Data Class Distribution and Test Data Class Distribution

Although the balance of the test dataset is also high, the imbalance is less in magnitude than in the training dataset. The intent is to keep some level of imbalance in the test set, so the evaluation mimics real world situation where imbalanced data is common. By doing so, the model's performance is assessed under realistic conditions, ensuring that it can generalize well to diverse distributions. Despite these imbalances, careful evaluation metrics were employed to provide a fair assessment of the model’s effectiveness across all classes.
For the proposed vitamin deficiency detection system, we evaluated the training and validation performance of accuracy and loss graphs as a function of multiple epochs and data distribution. However, these metrics tell us how well the model was learning the training process and would generalize well.

E. Training and Validation Accuracy

As depicted in figure 3, the accuracy graph shows a continuous increase in training accuracy, and validation accuracy is a gradual ascent as well. This means the model is able to effectively learn from training data while still having high performance on unseen validation data. It appears that the model is trained well with a small gap between training and validation accuracy so that it doesn’t over fit.

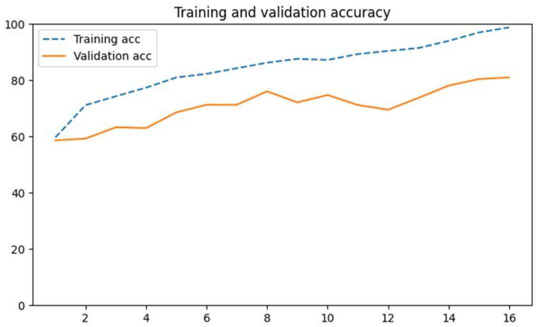

Figure 3: Training and Validation Accuracy

F. Training and Validation Loss

As the model becomes able to minimize prediction errors, we see a progressive decrease in training loss as depicted in figure 4. Validation loss was observed to fluctuate but stabilize as time passed, indicating that the model will generalize well to new data. The control of training and validation loss trends demonstrates the model is neither underfit nor overfit and is balanced.

Figure 4: Training and Validation Loss

V. CONCLUSION

This research demonstrates the potential of AI and machine learning in addressing critical healthcare challenges, such as the detection of vitamin deficiencies. The given system can analyze biomarkers and visual data with precision and efficiency using the power of advanced models like ResNet and CNN by PyTorch implementation and data augmentation.
In both clinical and resource-constrained settings, the model's emphasis on early detection, cost-effectiveness, and accessibility of the model makes it useful. Furthermore, the user-friendly interface brings the gap between advanced AI technology and practical healthcare applications to close so that general medical practitioners and diagnostic workflow can benefit.

Finally, this AI-based solution not only increases diagnostic accuracy but also promotes preventive healthcare strategy which facilitates effective detailing and final outcomes. The final direction of future work will be extending datasets, scaling models for real-world level scalability, and adding further integration with various healthcare systems. This work highlights the possibility of AI to turn the healthcare industry inside out and make it more inclusive while increasing its effectiveness in responding to health problems across the globe.

VI. Future Scope

The proposed AI based vitamin deficiency detection system has built a solid foundation of this revolutionizing healthcare diagnostic system and potential future development. Below are the key areas for potential enhancement and expansion:

A. Expanding Deficiency Detection

The current system is based on the detection of vitamin deficiencies, which can be broadened to other nutritional deficiencies, e.g. iron, calcium, or mineral imbalances. Though the system could evolve into a comprehensive diagnostic tool by incorporating biomarkers associated with these conditions, the system could evolve into a comprehensive diagnostic tool.

B. Diverse and Inclusive Datasets

Collaborations with healthcare institutions and researchers can lead to the collection of diverse datasets covering various demographics. This improvement would enhance the model's accuracy, generalizability, and applicability to a wide range of populations and geographies.

C. Enhanced Data Techniques

By Implementing advanced data augmentation and balancing methods to address dataset imbalances to improve prediction reliability.

D. Real-Time Diagnosis

Integrating the model with mobile applications can enable real-time monitoring, on-the-go diagnosis, and personalized healthcare management. This feature would be particularly beneficial for individuals in remote areas or those requiring continuous health monitoring.

E. Collaboration with Medical Experts

Interpretation of results and clinical validation of the system findings will be better done with close working with medical professionals. This kind of partnerships will add a layer of credibility, adoption, and a generative force for this system in the healthcare industry.

References

[1] Kassem M A, Hosny K M, Damasevicius R, and Eltoukhy M M, "Machine Learning and Deep Learning for Skin Lesion Classification and Diagnosis: A Systematic Review," Diagnostics, Vol. 11(8), 2021.

[2] Li M, Jiang Y, Zhang Y, and Zhu H, "Medical image analysis using deep learning algorithms," Frontiers in Public Health, 2023.

[3] Dinggang Shen, Guorong Wu, and Heung-ll Suk, "Deep Learning in Medical Image Analysis," Annual Review of Biomedical Engineering, Vol. 19, pp. 221-248, 2017.

[4] Zhou S Kevin, Greenspan Hayit, Davatzikos Christos, Duncan S James, Ginneken V Bram, and Madabhushi Anant, "A Review of Deep Learning in Medical Imaging: Imaging Traits, Technology Trends, Case Studies with Progress Highlights, and Future Promises," Proceedings of the IEEE, vol. 109(5), pp. 820-838, 2021

[5] K Roy, S S Chaudhuri, S Ghosh, S K Dutta, P Chakraborty, and R Sarkar, "Skin Disease detection based on different Segmentation Techniques," International Conference on Opto-Electronics and Applied Optics (Optronix), Kolkata, India, 2019.

[6] Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, Huang Z, Karpathy A, Khosla A, Bernstein M, Berg A C, and Fei-Fei L, "ImageNet Large Scale Visual Recognition Challenge," International Journal of Computer Vision, vol. 115(3), pp. 211-252, 2015.

[7] He K, Zhang X, Ren S, and Sun J, "Deep Residual Learning for Image Recognition," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 770-778, 2016.

[8] Lakshay Arora, "How to Deploy Machine Learning Models using Flask," 2024.