MULTI-CLASS NUTRIENT DEFICIENCY DETECTION FROM SKIN IMAGES USING CLASSICAL MACHINE LEARNING AND DEEP LEARNING MODELS

Sharmeen Mahmood[1], Yousuf Iqbal[*2], Saiqa Khalid[3], Zunaria Mustafa[4]

[1]*COMSATS University Islamabad, Sahiwal Campus, Pakistan*
[*2]*School of Computer Science, Wuhan University, Hubei, China*
[3]*The University of Haripur, Khyber Pakhtunkhwa, Pakistan*
[4]*COMSATS University Islamabad, Sahiwal Campus, Pakistan*

[1]sharmeenmahmood02@gmail.com, [*2]yousufiqbal600@gmail.com

DOI: https://doi.org/10.5281/zenodo.21153031

Keywords
Vision transformer, deficiency detection, resnet50.

Article History
Received: 25 April 2026
Accepted: 04 June 2026
Published: 21 June 2026

Corresponding Author: *
Yousuf Iqbal
yousufiqbal600@gmail.com

Abstract
The purpose of deficiency detection is to provide a methodology that accurately detects deficiency in human body without using any laboratory reports. In the field of computer vision, different deep learning models have been used to detect deficiency in humans. These models can detect only one or two types of deficiencies. There is no method available that detects multiple types of deficiencies in human body. However, it is important to keep in mind that these methods have some limitations when it comes to detect different categories of deficiencies. In previous research, low accuracy is observed due to the use of insufficient data. However, in this study, we address this limitation by introducing an additional category of malnutrition. To enhance the dataset, we collect images of patients with malnutrition specifically from DHQ hospital. This approach aims to improve the accuracy of our research findings and contribute valuable insights to the field of deficiency detection. We exploit six machine learning models such as Decision Tree, Naive Bayes, SVM, ResNet50, CNN, and Vision Transformer, as we train and test them for image classification tasks. The Vision Transformer out-performs other models, achieving an exceptional accuracy of 98%. It stands out in finding detailed patterns in image data, showcasing impressive accuracy. The ongoing comparison of these models not only confirms the effectiveness of the Vision Transformer but also provides insights for utilizing transformer-based architectures in various computer vision applications.

1. INTRODUCTION

Anemia is a big health problem for kids and pregnant women all over the world. The World Health Organization (WHO) says 42% of kids under 6 and 40% of pregnant women have anemia. The main reason is not having enough iron, affecting about one-third (33%) of people globally. Anemia happens when there aren't enough red blood cells in the body or when these cells are damaged or weak [1]. Nutritional anemia is divided into three categories: iron deficiency anemia, B12 deficiency anemia, and foliate deficiency anemia. Iron deficiency anemia is the most common in both men and women, often caused by bleeding in the stomach and intestines. This type can also occur if you don't get enough iron from your food, your body struggles to absorb iron from your stomach and intestines, or it can't produce sufficient red blood cells due to an iron shortage. If not identified early, iron deficiency anemia can lead to various issues like headaches, numbness in hands and feet, irritability, hair loss, heart palpitations, and restless leg syndrome. Conversely, B12 deficiency anemia is more prevalent in less developed countries where people may not consume enough animal products. The primary cause is the body's inability to absorb sufficient vitamin B12 in the small intestine. Anemia has connections to leukemia and solid organ cancers, underscoring the importance of thorough studies to identify causes and ensure accurate diagnosis. Timely treatment with appropriate medication is critical once anemia is detected. Early diagnosis is especially crucial for anemia resulting from a lack of specific nutrients. HGB-anemia occurs when the body lacks enough healthy red blood cells to carry oxygen to organs and tissues. This can happen when the red blood cell count decreases due to other diseases or when red blood cells are destroyed or lost too rapidly. Symptoms may include fatigue, reduced energy, difficulty breathing, chest pain, sleep disturbances and leg cramps [2]. The Pakistan National Nutrition Survey 2018 (NNS 2018) is the largest nutrition survey conducted in Pakistan, covering children, adolescents, and women across the country. The survey found that 40% of children under five are stunted, 17.7% are wasted, and 28.9% are underweight, indicating a high burden of malnutrition. It also reported that 9.5% of children are overweight, showing a growing trend of childhood obesity. Underweight prevalence varies across regions, with the highest rate observed in Sindh. Additionally, 53.7% of children suffer from anemia, with higher rates in rural areas than urban areas. These findings highlight the serious nutritional challenges faced by children in Pakistan [9].

Malnutrition can cause many health problems including stunting growth, digestive problem, and poor bone development. This problem occurs in children aged 2 to 12 years. This affects the child's overall performance. Malnutrition can be fatal if left untreated. It kills 45% of children under the age of five years because people don't have awareness about it. It is important to diagnose these adverse effects so that appropriate medications can be taken to eliminate or overcome them. Some common symptoms of deficiencies are given below

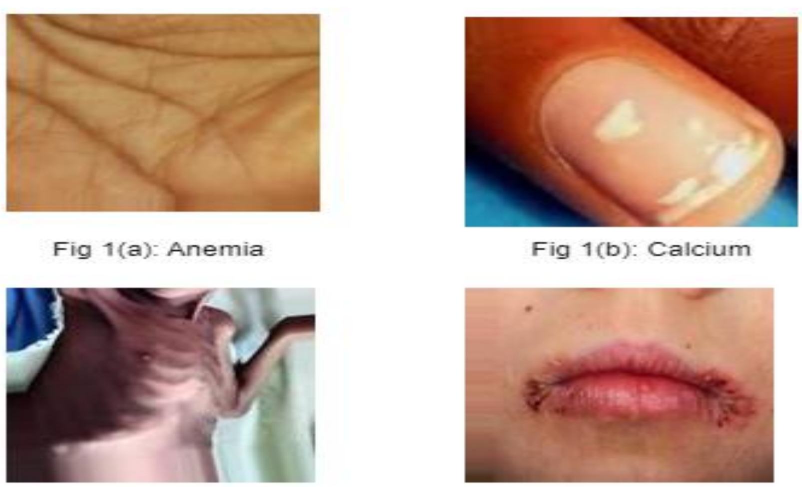


Fig 1(a): Anemia
Fig 1(b): Calcium
Fig 1(c): Malnutrition
Fig 1(d): Vitamin_B2

Previous studies on nutrient deficiency detection were limited by small datasets, a focus on only a few deficiency types, and reliance on blood sample analysis. Furthermore, low public awareness of deficiency symptoms makes early detection difficult. To address these gaps, this study employs computer vision techniques to detect multiple nutrient deficiencies from skin images, providing a more accessible and comprehensive approach for early diagnosis [8].

This study contributes to the application of artificial intelligence and computer vision in healthcare by addressing the challenge of nutrient deficiency detection through skin image analysis. To identify the most effective approach, six machine learning and deep learning models, including Convolutional Neural Network (CNN), ResNet50, Support Vector Machine (SVM), Naïve Bayes, Decision Tree, and Vision Transformer (ViT), are evaluated and compared using accuracy, precision, recall, and F1-score metrics. A key contribution of this work is the development of a custom nutrient deficiency image dataset collected from DHQ Hospital, providing a valuable resource for research in this domain. The findings of this study demonstrate the potential of computer vision-based techniques for early and non-invasive nutrient deficiency detection and provide a foundation for future advancements in AI-driven healthcare applications.

1.2 PROBLEM STATEMENT

Nutrient deficiencies pose a significant threat to human health, yet early detection remains challenging due to limited awareness of visual symptoms. Existing studies rely on blood-based analysis, and image-based approaches have been restricted to one or two deficiency categories, leaving a critical diagnostic gap. This study addresses these limitations by proposing a computer vision-based framework for multi-class classification of nutrient deficiencies through skin images. The objective is to develop a model capable of accurately identifying the deficiency type from a given skin image, enabling accurate, non-invasive, and accessible early detection across a broader diagnostic spectrum.

1.3 RESEARCH QUESTIONS

1. How can we enhance the accuracy of our model?
2. Can computer vision techniques accurately identify and differentiate between various nutrients deficiencies through the analysis of skin images?
3. What features and characteristics in skin images are more informative for detecting nutrients deficiencies?
4. How cost effective is the use of computer vision techniques for nutrients deficiency detection in comparison to traditional diagnostic methods?

1.4 RESEARCH OBJECTIVES

1. To explore and evaluate different techniques for enhancing the accuracy of deficiency detection.
2. To create a computer vision system capable of identifying nutrients deficiencies through the analysis of skin images.
3. To compare and analyze the strengths and limitations of existing deficiency detection techniques, and identify the most effective ones.
4. To ensure that the system can accurately classify and differentiate between various nutrients deficiencies, such as vitamin deficiency or iron deficiency.

1.5 SIGNIFICANCE OF THE STUDY

This study makes a meaningful contribution to the growing intersection of computer vision and healthcare, particularly in the underexplored domain of nutrient deficiency detection through dermatological imaging. The significance of this work is multifaceted. First, the research addresses a critical gap in clinical diagnostics by proposing an automated, non-invasive approach to identifying malnutrition-related skin manifestations, which holds considerable promise for resource-limited healthcare settings where laboratory testing may not be readily accessible. Second, the study presents a rigorous comparative analysis of six machine learning and deep learning methodologies — namely Convolutional Neural Networks (CNN), ResNet50, Support Vector Machine (SVM), Naive Bayes, Decision Tree, and Vision Transformer (ViT) — evaluated against four key performance metrics: accuracy, precision, recall, and F1-score. This systematic benchmarking provides researchers and practitioners with evidence-based guidance for model selection in similar medical imaging tasks.

Third, and most notably, this study introduces an original image dataset collected directly from DHQ Hospital, capturing real-world clinical cases of malnutrition-induced skin conditions. The development of this dataset is a significant scholarly contribution in itself, as the scarcity of domain-specific, labeled medical image data remains one of the foremost challenges in healthcare AI research. Collectively, this work advances the application of computer vision in preventive and diagnostic medicine, and serves as a foundational reference for future investigations into skin-based biomarker analysis and automated nutrient deficiency screening.

2. LITERATURE REVIEW

The increasing integration of computer vision and machine learning in healthcare has opened promising avenues for non-invasive diagnostic applications, particularly in the detection of nutrient deficiencies and malnutrition-related conditions. This section critically examines existing literature on machine learning-based deficiency detection, image-based diagnostic systems, deep learning models for dermatological and nail analysis, and nutritional assessment frameworks, with particular attention to methodological strengths, limitations, and identified research gaps.

2.1 Machine Learning Approaches for Nutrient Deficiency and Anemia Detection

Considerable research has focused on the automated identification of nutrient deficiencies and related blood disorders using machine learning techniques. [2] utilized an anemia dataset consisting of complete blood test results from 15,300 patients collected between 2013 and 2018. Noise removal and feature selection were performed with professional assistance, and a hybrid approach combining Genetic Algorithm with Stacked Autoencoder (GA-SAE) and GA-CNN was proposed to detect and classify anemia. Hyperparameter values were optimized using GA, and the dataset was divided under two configurations: a 60/40 and a 70/30 train-test split. While the approach demonstrated effective classification, a notable limitation is that it exclusively relies on blood samples and excludes data from children and pregnant women groups in which anemia is most prevalent [15] utilized a dataset sourced from healthcare centers to identify nutrient deficiencies in children, employing a hybrid algorithm combining AdaBoost and decision tree methodologies. The primary objective was to determine the number of meal features using decision tree algorithms. However, the approach exclusively relies on meal-related features, which may not provide sufficiently reliable information for accurate deficiency identification. The researchers advocate for a more comprehensive set of features beyond meal data to improve robustness in deficiency detection among children. Padmaja et al. [16] evaluated eleven different machine learning classification models using performance metrics including support, F1-score, recall, and precision to assess accuracy for vitamin D deficiency prediction. Models such as Gaussian Naive Bayes, Stochastic Gradient Descent, Multilayer Perceptron, Logistic Regression, K-Neighbors Classifier, and SVM showed subpar performance. The Decision Tree and AdaBoost classifiers achieved accuracies of 56% and 55%, respectively, while the Gradient Boosting Classifier, Extra Trees Classifier, and Random Forest achieved 68%, 73.3%, and 72.1%, respectively. The Extra Trees Classifier was ultimately integrated into a dashboard for predicting vitamin D deficiency, though the dataset does not encompass all age groups, limiting generalizability.

2.2 Image-Based and Mobile Diagnostic Systems

Several studies have explored the use of image-based inputs and mobile applications for deficiency and malnutrition detection. [8] proposed a system that detects nutrient deficiencies by capturing images of various body parts through an Android application. The system employs neural network training to recognize patterns and features indicative of nutrient deficiencies, with Natural Language Processing (NLP) incorporated for feature extraction. Despite the innovative use of mobile technology for health monitoring, the approach is acknowledged to provide less accuracy, which may arise from the complexity of training a neural network for diverse symptom patterns or challenges in feature extraction using NLP techniques. [20] focused on utilizing the Fuzzy C Means (FCM) method within a mobile application to evaluate a patient's nutritional status. The clustering parameters included height, weight, and age, categorizing nutritional status into three clusters: good nutrition, malnutrition, and improved nutrition. The application also serves as a reminder for nutritional information during food consumption. System testing revealed a validation rate of 80%, demonstrating that FCM-based nutritional assessment through mobile applications offers a practical tool for monitoring and improving nutritional status. [19] asserted that effectively addressing malnutrition requires a holistic digital approach extending beyond data provision to establish connections between program indicators, socio-economic conditions, and family demographics. The Integrated Child Development Services (ICDS) program encounters challenges such as limited skills among community health workers and a lack of detailed socio-economic data. A microservices-oriented digital framework was proposed to ensure data availability, integrity, connectivity, and comprehension of causality, with insights drawn from a field trial testing a digitization prototype.

2.3 Deep Learning Models for Dermatological and Nail-Based Diagnosis

Deep learning has been extensively applied to image-based diagnosis of skin and nail conditions, which are closely associated with nutrient deficiencies and systemic health disorders. [1] identified iron deficiency anemia by assessing the performance of CNN, Naïve Bayes, Decision Tree, k-NN, and SVM models. These models were trained, validated, and tested using images of the eyes' conjunctiva, palm texture, and fingernail color. CNN outperformed all other models, demonstrating higher accuracy across all image types. The results further indicate that the palpable palm is a reliable feature for anemia detection in children due to its superior accuracy. [22] examined four distinct nail conditions healthy nails, nail hyperpigmentation, nail clubbing, and nail fungus using five deep CNN models: AlexNet, VGG16, GoogleNet, ResNet50, and DenseNet201. Evaluation was based on six measures including accuracy, recall, specificity, precision, F-score, and time, with implementation carried out in MATLAB. AlexNet, VGG16, and GoogleNet achieved accuracies of 92.5%, 87.5%, and 93.98%, respectively, while ResNet50 and DenseNet201 attained an impressive accuracy of 96.39%, highlighting their efficacy in identifying subtle differences in nail health conditions. [11] created a web-based tool using the EfficientNet-B2 model to diagnose 17 different fingernail diseases, achieving a detection success rate of 72%. While this detection rate may appear relatively modest, the system is distinguished by its ability to classify images into 17 distinct disease categories in real time, in contrast to other methods that often focus on binary or limited classification tasks.

2.4 Nutritional and Hematological Studies

Beyond image-based approaches, several studies have examined nutritional and hematological dimensions of deficiency-related conditions. [21] explored the correlation between hemoglobin levels, anemia, age, and various health conditions in elderly individuals, finding that among those aged 90 and above, anemia rates rise to 41% for men and 21% for women. Contributing factors identified include age, race, body mass index, smoking habits, cancer history, hospitalization, kidney issues, and low albumin levels, providing important insights into the complexity of anemia in aging populations. [27] compared characteristics of patients with vitamin B12 deficiency who had undergone gastrectomy with those who had not, reviewing complaints, oral findings, blood test results, and medical histories. The most common oral findings included erythema and depapillation of the tongue, more prevalent in patients with gastrectomy history. The study concludes that vitamin B12 deficiency should be considered in patients experiencing glossodynia, even when oral mucosa appears normal.

2.5 Research Gaps and Identified Limitations

The reviewed literature reveals several critical gaps that this study seeks to address:

1. The majority of existing studies rely on blood samples or single-modality data for deficiency detection, excluding the potential of skin image analysis as a non-invasive diagnostic tool.
2. Limited research has undertaken a systematic comparative evaluation of multiple machine learning and deep learning models including CNN, ResNet50, SVM, Naive Bayes, Decision Tree, and Vision Transformer within the specific context of malnutrition detection.
3. There is a notable absence of clinically collected, domain-specific image datasets for malnutrition, particularly from developing country healthcare settings such as Pakistan.
4. Existing approaches frequently exclude vulnerable populations, particularly children and women, despite their heightened susceptibility to nutrient deficiencies.

The present study addresses these gaps by conducting a comprehensive multi-model comparison using an original dataset collected from DHQ Hospital, offering a clinically grounded and methodologically rigorous contribution to the field of computer vision-based malnutrition detection.

3. METHODOLOGY

3.1 Research Design

This study used a quantitative experimental research methodology in order to create and assess a computer vision-based framework for identifying vitamin deficiencies using skin images. The design enabled systematic collection, preprocessing, training, and validation of multiple machine learning and deep learning models using a custom-compiled image dataset. A model-comparison approach was employed to assess the performance of six classification architectures across standardized evaluation metrics.

3.2 Data Collection

The dataset comprises images representing five categories of deficiencies. These categories include anemia, calcium deficiency, malnutrition, iron deficiency and vitamin B2 deficiency. The images of anemia, iron deficiency and calcium deficiency are collected from kaggle datasets. Vitamin B2 deficiency images are collected from DermNet medical images. And images of malnutrition are collected from DHQ hospital Vehari as shown in table below:

Table 1: Dataset

Deficiency category	Number of images
Anemia	560
Calcium Deficiency	560
Iron Deficiency	560
Malnutrition	560
Vitamin B2 Deficiency	560

3.3 Data Preprocessing

The goal of data preprocessing is to enhance the quality of data, improving the performance of machine learning and deep learning models and ensure that the data is in usable and meaningful form. Our methodology consists of these data preprocessing steps:

3.3.1 Image Cropping

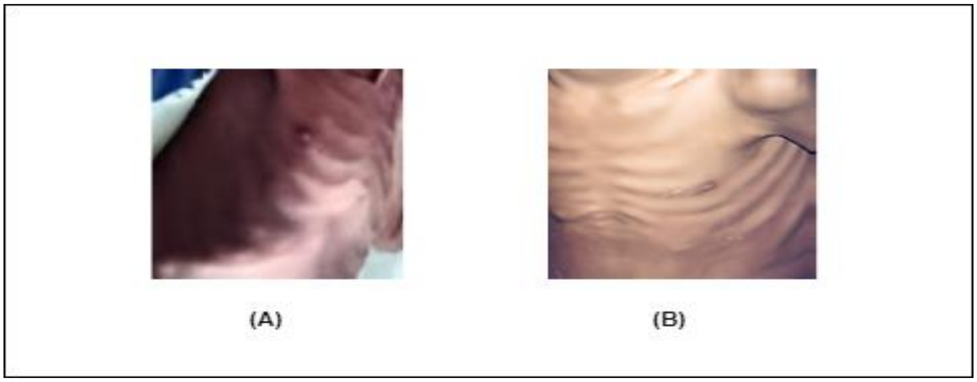
In image cropping process a portion or section of an image is selected and retained, while the rest of the image is removed. This involves adjusting the frame or boundaries of the image to focus on specific area of interest. Cropping is commonly used in photography, graphic design and image processing to improve composition, eliminate unwanted elements or resize image to fit specific dimensions. We can manually select any irregular shape to crop the image. The data of malnutrition collected from the hospital was initially in raw form. Image cropping has been employed on data to enhance its usability. As shown below

Raw images of malnourished child

Cropping is performed manually as we selectively define and extract the desired regions from the images. This hands-on approach allows us to precisely tailor the images, ensuring that relevant features are high-lighted while irrelevant and redundant information is excluded. As shown in Figure

Region of interest (ROI) of images

In Figure below, we're showing the pictures of ribs that we have taken out from the original images. Think of it like cutting out a specific part of a photo to focus only on the ribs. We did this carefully to make it easier to see and study the details of the ribs. Both pictures are like a close-up view of the ribs, helping us understand them better. We used a method to pick out the ribs parts from the raw images and this figure illustrate what we got after doing that.



Cropped image

3.3.2 Image Resizing

In this step, we made all the pictures in dataset the same size, specifically 128x128 pixels. We did this to make the images easier to work with and analyze. The size 128x128 was picked to balance keeping important details in the picture while not making the analysis too complicated.

3.3.3 Gray Scale Conversion

In this step, we turned all the images into black and white. This means we changed the RGB images to shades of gray. We did this to simplify the images and focus on their shapes and structures. Converting to gray scale makes it easier to work with the pictures. Especially when analyzing details or using certain computer programs. This step was carried out to enhance the clarity of the images and facilitate a straighter forward analysis. Using gray scale data can be advantageous in deep learning because it reduces the complexity of the information while retaining essential visual features.

3.3.4 Data Augmentation

To achieve greater accuracy, machine learning models should undergo training on an extensive dataset. A limited dataset size can lead to over fitting. This occurs because the smaller number of data samples in both training and testing stages makes it challenging for the model to generalize effectively. The following augmentation parameters were utilized:

Rotation Range: Images were randomly rotated by an angle within the range of -20 to +20 degrees. This introduced variations in object orientations, allowing the model to batter capture features from different perspectives.

Gray-scale conversion

Width Shift Range: Horizontal shifts, up to 10% of the image width, were applied randomly. This simulated changes in object positions along the x-axis, contributing to the model's ability to recognize objects at different locations.

Height Shift Range: Similar to width shifting, vertical shifts up to 10% of the image height were applied randomly. This helped the model learn to recognize the objects at various vertical positions.

Shear Range: Shearing transformations, with a shear intensity ranging from -20 to +20 degrees, were applied randomly. This introduced deformations to the images, aiding the model in handling distorted input.

Zoom Range: Random zooming with a range of up to 20% was applied to simulate variations in object sizes. This ensured the model's capability to recognize objects at different scales.

Horizontal Flip: Images were horizontally flipped with a 50% probability. This augmented the dataset by providing mirrored versions of the original images, contributing to the model's ability to handle horizontally inverted objects.

3.3.5 Data Splitting

The division of dataset ensures that the model learns from one portion of the data and is evaluated on another, unseen portion. The typical split ratio is 80/20, where 80% of the data is allocated for training, and the remaining 20% is reserved for testing. The training set is utilized to train the machine learning model. During this phase, the model learns patterns, relationships, and features within the data, adjusting its parameters to optimize its predictive capabilities. The testing set serves as an independent dataset that the model has not seen during the training phase. It is used to evaluate the model's performance and generalization to new, unseen data.

3.4 Implemented Models

3.4.1 Support Vector Machine (SVM)

SVM is a supervised learning method used for multi-class classification. A linear kernel was selected after testing polynomial, sigmoid, and RBF kernels, as it consistently achieved the highest accuracy. Images were standardized to 100×100 pixels and flattened into 1D array for training.

3.4.2 Naïve Bayes

A Gaussian Naïve Bayes classifier was employed due to its suitability for continuous-valued features. HOG (Histogram of Oriented Gradients) features were extracted from each image to capture local gradients and edge patterns, which were then used to train the probabilistic model.

3.4.3 Decision Tree

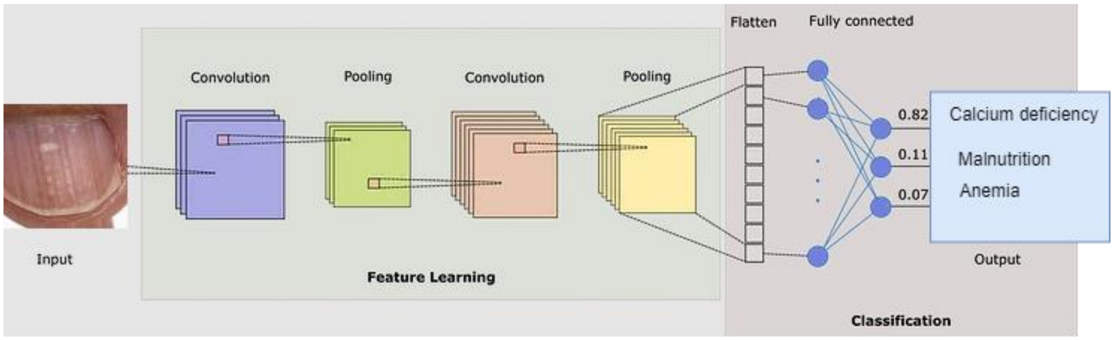
A Decision Tree classifier was configured using Gini impurity as the splitting criterion. HOG features were used as input vectors. Hyper-parameters such as maximum depth and minimum samples per leaf were tuned to prevent overfitting while maintaining generalization.

3.4.4 Convolution Neural Network (CNN)

The Convolutional Neural Network (CNN) used in this study is designed for image classification and plays a key role in identifying nutrient deficiency patterns from images. The model is trained on a dataset containing five classes, with data augmentation techniques such as rescaling and train-validation splitting applied to improve robustness and generalization. A sequential CNN architecture is employed, consisting of two convolutional layers with 32 and 64 filters, respectively, using 3×3 kernels and ReLU activation functions to extract hierarchical image features. Each convolutional layer is followed by a max-pooling layer to reduce spatial dimensions while preserving important information. The extracted feature maps are then flattened and passed through two fully connected dense layers with 128 neurons, followed by a softmax output layer for multi-class classification. The model is compiled using the Adam optimizer, categorical cross-entropy loss function, and accuracy metric, enabling efficient learning and accurate classification of nutrient deficiency categories.

3.4.5 Resnet50

The ResNet50 model is utilized in this study through transfer learning to improve image classification performance by leveraging knowledge learned from large-scale datasets. The pre-trained ResNet50 architecture, consisting of 50 layers and trained on ImageNet, serves as the base model, with its layers frozen to preserve previously learned features. The dataset is prepared using data augmentation techniques, including image rescaling and train-validation splitting, while parameters such as batch size, image size, and epochs are optimized for effective training. The extracted feature maps from ResNet50 are flattened and passed through a dense layer with ReLU activation to learn task-specific patterns. A dropout layer is incorporated to reduce overfitting and improve generalization by randomly deactivating neurons during training. Finally, a softmax output layer performs multi-class classification, and the model is compiled using the Adam optimizer, categorical cross-entropy loss function, and accuracy metric, ensuring efficient training and accurate prediction of the target classes.



Architectural Overview of CNN

3.4.6 Vision Transformer (ViT)

The Vision Transformer (ViT) treats an image as a sequence of fixed-size, non-overlapping patches that are converted into vector representations and processed using the Transformer architecture. Through the self-attention mechanism, the model captures global contextual information and learns relationships between different image regions, making it highly effective for image classification tasks. In this study, ViT is used to classify skin images and identify nutrient deficiencies by analyzing visual patterns and generating class probabilities through a softmax classifier.

3.5 Implemented Approaches

3.5.1 Embedding

The input skin image is divided into patches, which are flattened and transformed into embedding vectors through a dense layer. A learnable class token and positional encoding are added to these embeddings to preserve class-specific and spatial information. This process creates a meaningful representation of image patches that serves as input to the Transformer model.

3.5.2 Transformer Encoder

The Transformer Encoder consists of multiple identical blocks containing Multi-Head Attention and Feed-Forward layers, along with residual connections and layer normalization. These components enable the model to effectively learn contextual relationships among image patches and generate a rich context vector that captures important information for nutrient deficiency classification.

3.5.3 Multi-Head Attention

Multi-Head Attention transforms the input embeddings into Queries (Q), Keys (K), and Values (V) and computes attention scores to determine the importance of each image patch. Multiple attention heads operate simultaneously to capture different feature relationships, and their outputs are combined to produce a comprehensive representation of the image, improving the model's understanding of complex visual patterns.

3.5.4 MLP Head

The context vector produced by the Transformer Encoder is passed to a Multilayer Perceptron

(MLP) head for classification. The MLP head processes the class token and generates final probability scores for each nutrient deficiency category, enabling accurate multi-class prediction.

3.5.5 Training and Optimization

The Vision Transformer model is trained using the cross-entropy loss function and optimized with the AdamW optimizer. A Cosine Annealing Learning Rate Scheduler dynamically adjusts the learning rate during training, improving convergence and stability. The model learns from batches of labeled skin images over multiple epochs to minimize classification errors and enhance performance.

3.5.6 Model Evaluation

The Vision Transformer achieved excellent classification performance, with an accuracy of 98.21%, recall of 97.20%, precision of 96.28%, and F1-score of 96.49%. These results demonstrate the model's strong capability to accurately detect and classify nutrient deficiencies from skin images while maintaining a balanced trade-off between precision and recall.

3.6 Evaluation Matrices

In our research, we have utilized recall, precision, F1 score and accuracy as the primary metrics for thoroughly assessing the effectiveness of the model.

3.6.1 Recall

Recall is a vital measure that assesses how well the model captures the real cases of nutrient deficiencies in this research. In the context of image classification for nutrient deficiency detection in skin images, recall checks if the model can catch all the crucial patterns indicating nutrient deficiencies. A higher recall score suggests the model's effectiveness in recognizing and including relevant content, contributing to a more thorough and comprehensive detection of nutrient deficiencies within skin images.

$$Recall = \frac{Correctly\ Identified\ Deficiency\ Patterns}{Total\ Instances\ of\ Nutrient\ Deficiency\ (TP + FN)}$$

3.6.2 Precision

In the context of image classification for nutrient deficiency detection in skin images, precision measures the ratio of accurately identified true positive nutrient deficiency patterns among all the patterns identified by the model. This metric assesses how precisely the model avoids including irrelevant or redundant content in its classification results.

$$Precision = \frac{Correctly\ Identifiend\ Deficiency\ Patterns}{Total\ Identified\ Patterns\ as\ Nutrients\ Deficiencies}$$

Here, "Correctly Identified Nutrient Deficiency Patterns" represents the true positives (TP), and "Total Identified Patterns as Nutrient Deficiencies" includes both true positives (TP) and false positives (FP).

3.6.3 F1 Score

F1 score evaluates the model's accuracy in identifying patterns of nutrient deficit and capturing relevant data by combining precision and recall. A high F1 score means the model achieves this balance well, leading to an effective and accurate detection of nutrient deficiencies in skin images, aligning with my thesis goals.

$$F1 - Score = \frac{2 * Precision * Recall}{Precision + Recall}$$

3.6.4 Accuracy

In assessing the performance of a model designed for nutrient deficiency detection through skin images, accuracy holds significant importance. When it comes to image classification, accuracy is defined as the proportion of correctly classified nutritional deficiency patterns relative to the total number of skin image datasets. This metric becomes important since it assesses how well the model can distinguish and classify pertinent patterns linked to various nutrient deficits while eliminating irrelevant ones. In our specific study, where there are five categories of deficiency, the accuracy metric ensures that the model precisely identifies the presence of any of these nutrient deficiencies within the skin images.

$$Accuracy = \frac{Sum\ of\ Correctly\ Predicted\ Samples\ across\ all\ Categories}{Total\ Observations}$$

4. RESULTS

This section presents the performance evaluation of six models, Decision Tree, SVM, Naïve Bayes, CNN, ResNet50, and Vision Transformer for nutrient deficiency detection from skin images, assessed across dataset variations using accuracy, precision, recall, and F1-score.

Table 1. Comparative Model Performance Across Dataset Configurations

Model	Configuration	Recall (%)	Precision (%)	F1-Score (%)	Accuracy (%)
8.4.3 Decision Tree	RGB w/o Aug.	46.52	48.66	46.66	47.92
SVM	RGB w/o Aug.	86.04	87.65	86.66	87.50
8.4.2 Naïve Bayes	RGB w/o Aug.	81.90	82.82	82.24	83.33
CNN	RGB w/o Aug.	87.50	88.37	87.11	87.50
ResNet50	RGB w/o Aug.	81.25	80.71	80.59	81.25
8.4.3 Decision Tree	RGB w/ Aug.	55.57	56.74	56.05	55.86
SVM	RGB w/ Aug.	87.16	87.39	86.99	86.74
CNN	RGB w/ Aug.	88.50	88.53	88.09	88.50
ResNet50	RGB w/ Aug.	96.88	96.19	96.88	96.88
Vision Transformer	RGB w/ Aug.	97.20	96.20	96.49	98.21

Data augmentation consistently improved model performance, particularly for deep learning models. The Vision Transformer achieved the highest overall accuracy of 98.21%, followed by ResNet50 at 96.88%.

Table 2. Vision Transformer — Per-Class Performance

Deficiency Category	Recall (%)	Precision (%)	F1-Score (%)	Accuracy (%)
Anemia	90.24	91.94	90.22	90.24
Calcium Deficiency	94.74	94.74	94.74	94.74

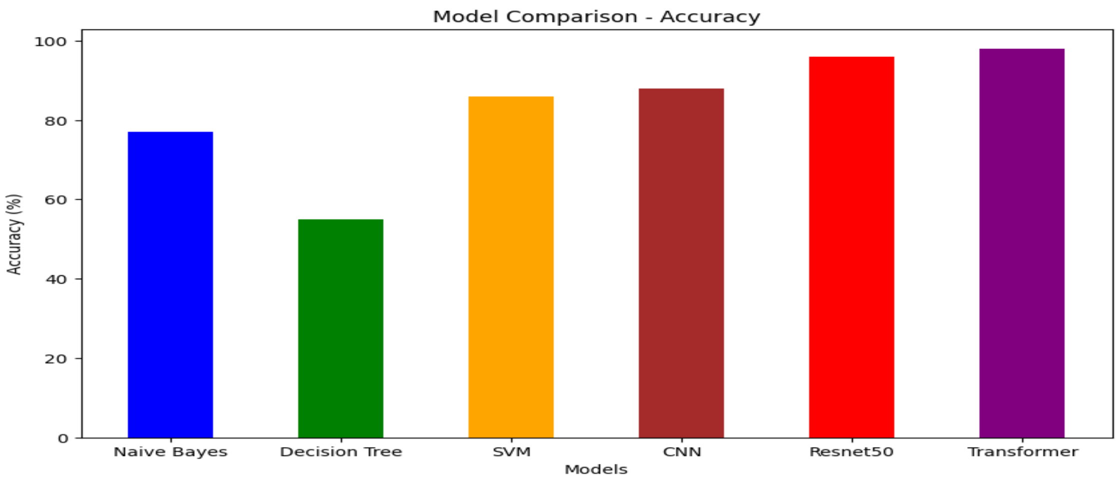
Deficiency Category	Recall (%)	Precision (%)	F1-Score (%)	Accuracy (%)
Malnutrition	95.45	95.72	95.32	95.45
Vitamin-B2 Deficiency	96.88	97.16	96.91	96.88
Iron Deficiency	80.00	80.63	79.70	80.00

Vitamin-B2 detection yielded the highest per-class accuracy (96.88%), while Iron deficiency proved the most challenging category (80.00%).

Table 3. Best Accuracy Summary — All Models

Model	Best Accuracy (%)
8.4.3 Decision Tree	56.25
8.4.2 Naïve Bayes	83.33
SVM	86.74
CNN	88.50
ResNet50	96.88
Vision Transformer	98.21

The progressive improvement from Decision Tree (56.25%) to Vision Transformer (98.21%) confirms the advantage of attention-based architectures in capturing complex visual patterns for dermatological deficiency detection.



5. CONCLUSION AND FUTURE WORK

This study presented a comprehensive comparative analysis of six machine learning and deep learning models Decision Tree, Naïve Bayes, SVM, CNN, ResNet50, and Vision Transformer for multi-class nutrient deficiency detection from skin images. A custom dataset comprising five deficiency categories (anemia, calcium deficiency, iron deficiency, vitamin B2 deficiency, and malnutrition) was compiled from DHQ Hospital Vehari, Kaggle, and DermNet. Among all evaluated models, the Vision Transformer achieved the highest classification accuracy of 98.21%, demonstrating the superiority of attention-based architectures in capturing complex visual patterns from dermatological images. These findings confirm the viability of computer vision as a non-invasive and accessible alternative to traditional blood-based diagnostic methods for nutrient deficiency detection.

Future research should focus on expanding and diversifying the dataset to improve model generalization across different populations and skin types. Further investigation into advanced model architectures, fine-tuning strategies, and explainability techniques would enhance both performance and clinical interpretability. Interdisciplinary collaboration between computer science and healthcare professionals is encouraged to facilitate the translation of these findings into deployable diagnostic tools for real-world healthcare settings.

REFERENCES

J. W. Asare, P. Appiahene, E. T. Donkoh, and G. Dimauro, "Iron deficiency anemia detection using machine learning models: A comparative study of fingernails, palm and conjunctiva of the eye images," *Engineering Reports,* p. e12667, 2023.

S. Kilicarslan, M. Celik, and Ş. Sahin, "Hybrid models based on genetic algorithm and deep learning algorithms for nutritional Anemia disease classification," *Biomedical Signal Processing and Control,* vol. 63, p. 102231, 2021.

M. Dithy and V. Krishnapriya, "Anemia selection in pregnant women by using random prediction (Rp) classification algorithm," *Int J Recent Technol Eng,* vol. 8, no. 2, pp. 2623-2630, 2019.

P. Appiahene, J. W. Asare, E. T. Donkoh, G. Dimauro, and R. Maglietta, "Detection of iron deficiency anemia by medical images: a comparative study of machine learning algorithms," *BioData Mining,* vol. 16, no. 1, pp. 1-20, 2023.

G. Dimauro, D. Caivano, and F. Girardi, "A new method and a non-invasive device to estimate anemia based on digital images of the conjunctiva," *Ieee Access,* vol. 6, pp. 46968-46975, 2018.

A. A. V. Cruz *et al.*, "Adult blindness secondary to vitamin A deficiency associated with an eating disorder," *Nutrition,* vol. 21, no. 5, pp. 630-633, 2005.

S. Muthayya, J. H. Rah, J. D. Sugimoto, F. F. Roos, K. Kraemer, and R. E. Black, "The global hidden hunger indices and maps: an advocacy tool for action," *PloS one,* vol. 8, no. 6, p. e67860, 2013.

J. Akram, N. Vakmiki, and A. Chamma, "A Survey Paper on Various Techniques used in Detection of Vitamin Deficiency using Image Processing and Neural Network," *Grenze International Journal of Engineering & Technology (GIJET),* vol. 9, no. 2, 2023.

G. o. Pakistan and UNICEF, "National Nutrition Survey 2018: Key Findings Report," ed: Ministry of National Health Services and UNICEF Islamabad, 2018.

I. ul Haq *et al.*, "Prevalence of micronutrient deficiencies among preschool and school-going children in flood-hit areas of Pakistan," *The American Journal of Tropical Medicine and Hygiene,* vol. 105, no. 6, p. 1638, 2021.

C. Zuhal and I. Şahin, "DIAGNOSING DISEASES FROM FINGERNAIL IMAGES," *Eskişehir Osmangazi Üniversitesi Mühendislik ve Mimarlık Fakültesi Dergisi,* vol. 30, no. 3, pp. 464-470, 2022.

O. Faust, Y. Hagiwara, T. J. Hong, O. S. Lih, and U. R. Acharya, "Deep learning for healthcare applications based on physiological signals: A review," *Computer methods and programs in biomedicine,* vol. 161, pp. 1-13, 2018.

A. Mitani *et al.*, "Detection of anaemia from retinal fundus images via deep learning," *Nature Biomedical Engineering,* vol. 4, no. 1, pp. 18-27, 2020.

S. Kasiviswanathan, T. Bai Vijayan, L. Simone, and G. Dimauro, "Semantic segmentation of conjunctiva region for non-invasive anemia detection applications," *Electronics,* vol. 9, no. 8, p. 1309, 2020.

S. Najaflou and M. Rabiei, "Recommended system for controlling malnutrition in Iranian children 6 to 12 years old using machine learning algorithms," *International Journal of Web Research,* vol. 4, no. 1, pp. 27-33, 2021.

B. Padmaja, B. R. Reddy, R. V. Sagar, and H. K. Pradhan, "Prognosis of Vitamin D deficiency severity using SMOTE optimized machine learning models," *Turkish Journal of Computer and Mathematics Education (TURCOMAT),* vol. 12, no. 6, pp. 4553-4567, 2021.

B. Priambodo, Y. Jumaryadi, S. Rahayu, D. Firdaus, M. Sobri, and Z. P. Putra, "Prediction of Drug Users Based on Facial Scratching Pattern," *Int. J. Online Biomed. Eng.,* vol. 17, no. 3, pp. 141-150, 2021.

R. Mente and S. Marulkar, "A review: Fingernail images for disease detection," *Int. J. Eng. Comput. Sci,* vol. 6, no. 11, pp. 22830-22835, 2017.

A. Maitra, R. E. Rote, and N. Kuntagod, "Managing child malnutrition via digital enablement: Insights from a field trial," in *2017 IEEE Global Humanitarian Technology Conference (GHTC),* 2017: IEEE, pp. 1-7.

S. Winiarti, S. Kusumadewi, I. Muhimmah, and H. Yuliansyah, "Determining the nutrition of patient based on food packaging product using fuzzy C means algorithm," in *2017 4th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI),* 2017: IEEE, pp. 1-6.

M. E. Salive *et al.*, "Anemia and hemoglobin levels in older persons: relationship with age, gender, and health status," *Journal of the American Geriatrics Society,* vol. 40, no. 5, pp. 489-496, 1992.

J. Abdulhadi, A. Al-Dujaili, A. J. Humaidi, and M. A.-R. Fadhel, "Human nail diseases classification based on transfer learning," *ICIC Express Letters,* vol. 15, no. 12, pp. 1271-1282, 2021.

T. S. Indi and D. D. Patil, "Nail feature analysis and classification techniques for disease detection," *Int. J. Comput. Sci. Eng,* vol. 7, no. 5, pp. 1376-1383, 2019.

V. Thahira Banu and M. R. Devi, "Hybrid Classifier to Classify The Finger Nail Abnormalities," *Information Technology In Industry,* vol. 9, no. 1, pp. 549-555, 2021.

M. Yani, S. Budhi Irawan, Si, MT, and S. Casi Setiningsih, MT, "Application of transfer learning using convolutional neural network method for early detection of terry's nail," in *Journal of Physics: Conference Series,* 2019, vol. 1201, no. 1: IOP Publishing, p. 012052.

P. Bose and S. K. Bandyopadhyay, "Facial spots detection using convolution neural network," *Asian Journal of Research in Computer Science,* vol. 5, no. 3, pp. 71-83, 2020.

J. Kim, M.-J. Kim, and H.-S. Kho, "Oral manifestations in vitamin B 12 deficiency patients with or without history of gastrectomy," *BMC Oral Health,* vol. 16, pp. 1-9, 2016.

E. Paillaud, I. Merlier, C. Dupeyron, E. Scherman, J. Poupon, and P.-N. Bories, "Oral candidiasis and nutritional deficiencies in elderly hospitalised patients," *British Journal of Nutrition,* vol. 92, no. 5, pp. 861-867, 2004.

N. Shahriar, "What is Convolutional Neural Network — CNN (Deep Learning)," 2019. [Online]. Available: https://nafizshahriar.medium.com/what-is-convolutional-neural-network-cnn-deep-learning-b3921bdd82d5.

Javatpoint. "Support Vector Machine Algorithm." https://www.javatpoint.com/machine-learning-support-vector-machine-algorithm (accessed)

S. Mukherjee, "The Annotated ResNet-50 Explaining how ResNet-50 works and why it is so popular," Aug 18, 2022. [Online]. Available: https://towardsdatascience.com/the-annotated-resnet-50-a6c536034758.

JavaTpoint, "Decision Tree Classification Algorithm," 2019. [Online]. Available: https://www.javatpoint.com/machine-learning-decision-tree-classification-algorithm.

S. Junawane, "ViT: Vision Transformer," Apr 24, 2022. [Online]. Available: https://medium.com/machine-intelligence-and-deep-learning-lab/vit-vision-transformer-cc56c8071a20.

P. D. Cameron R. Wolfe, "Using Transformers for Computer Vision," Oct 5, 2022. [Online]. Available: https://towardsdatascience.com/using-transformers-for-computer-vision-6f764c5a078b.