IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Diagnosis Of Vitamin Deficiency Detection Using Deep Learning And Image Processing

[1]Dr. Dnyanada Hire ,[2] Vishakha Makasare , [3]Shreya Pawar ,[4] Sakshi Hambir
[1] Professor, [2] Student, [3] Student , [4] Student
[1]Department of Semiconductor Engineering
D. Y. Patil International University, Pune, India

Abstract: In this work, we introduce an advanced intelligent system designed to use deep learning techniques to identify and distinguish between human tissue status and vitamin deficiencies effectively. Our approach begins with the use of image aggregation techniques to isolate and better focus lesion areas in images. The large amount of data available on the image databases is used to fine-tune the pre-trained networks to obtain color, texture, and structure features related to the pathologies. This approach is effective and provides an opportunity to detect vitamin deficiencies quickly and unequivocally even to expose patients to invasive procedures. The system aims to conduct an image analysis of physical characteristics associated with deficiency diseases, which include pale skin, smooth tongue, and dry eyes. Our primary objectives are to thoroughly evaluate the effectiveness of the proposed segmentation procedure, determine which aspects are most critical for exact classification, and compare the effectiveness of our grouping results with those achieved using alternative methods. This method aims to enhance the precision and reliability of identifying and distinguishing between various tissue types and deficiencies, thereby advancing diagnostic capabilities in medical imaging significantly.

Index Terms - **Vitamin Deficiency, Deep Learning, CNN, ResNet, Medical Imaging.**

I. Introduction

Among the most important health problems globally, vitamin deficiencies affect millions of people. Early detection improves outcomes. Formerly, the diagnosis was predicated upon test results and clinical examination that were invasive, time-consuming, and expensive, particularly in resource- poor environments. It shows breakthroughs in processing images and machine learning, most notably in Convolutional Neural Networks, bring forth the possibilities of non-invasive methods for the detection and monitoring of vitamin deficiencies. The research of this study is focused on the development of a computer vision system based on the analysis of images concerning specific parts of the body and comparison with visual markers related to different deficiency cases. The objective of this project is to develop a novel approach for vitamin deficiency detection using image processing techniques and CNNs.

The aim of this application is to assist in the diagnosis of how a potential vitamin deficiency can be detected on pictures made from the eyes, lips, tongue, and nails of the individual in question, sans the invasive blood tests. Traditionally, the process of diagnosing vitamin deficiency entails one carrying out blood tests and clinical examinations, which is very time- consuming and expensive to one, as well as sometimes painful and not welcoming to the patients. Most of these methods will be difficult to use in resource-poor environments where there is limited access to medical laboratories and healthcare professionals. This new application will solve the above concerns by providing a more accessible, non-invasive, and user-friendly approach to diagnostics for health.

Secondary data has been gathered for this study. The image dataset employed in this work was downloaded from the open Kaggle platform, namely the "Vitamin Deficiency Skin Signs" collection. The dataset comprises 15,575 high-resolution RGB images, classified according to manifest signs of different vitamin deficiencies over different anatomical regions including skin, hair, nails, and body. The dataset utilized ranges from 2020 to 2023 and consists of a varied set of patient-like images shared from different sources such as public health databases, dermatological research, and open medical databases. All the images were annotated by the type of deficiency it represents to ensure relevance and quality for the training of machine learning models. Furthermore, references and background information on vitamin symptoms and clinical indicators were obtained from verified health sources, such as WHO guidelines, dermatology research papers, and clinical case histories. This secondary data was used to develop the image categories, identify deficiency indicators, and train a model that can classify visual symptoms in real time

1) Type Style and Fonts

This report follows the IEEE report style. The default font is Times New Roman, and the body text is in 12-point, with headings in bold and suitable size hierarchies (14-point for primary headings, 12-point for sub-headings). All line spacing is 1.5 lines, and standard 1-inch margins are applied on all sides for readability and printing consistency. All tables, figures, and sections adhere to IEEE formatting standards for clarity and professional presentation.

2) Ease of Use

This template document provides uniformity in the formatting of the report. All the margins, fonts, line spacing, and section styles are standardized and should not be modified. The format is structured to provide uniformity when preparing project reports from multiple authors. Manual pagination should be avoided, and hard tabs or spacing should be avoided. The format is optimized for A4 printing, PDF export, and online review.

1.1Population and Sample

The subjects of this study are populations of digital images that are visibly apparent symptoms of vitamin deficiency, including hair loss, glossitis, nail pigmentation, and dermatological changes. These images mimic patient data in real-life healthcare settings where invasive diagnosis is not readily available. Sample utilized consists of 15,575 images taken from the publicly disseminated Kaggle dataset "Vitamin Deficiency Skin Signs.". These images were chosen to represent a broad range of symptoms over four regions of anatomy: skin, hair, nails, and body. A stratified sampling approach was used to have balance across categories for training, validation, and testing. The size and diversity of this sample allow the deep learning model to generalize well across populations with different skin tones, lighting, and intensities of symptoms.

1.2 Data and Sources of Data

In this project, secondary data have been utilized. The major image data source is the Kaggle dataset named "Vitamin Deficiency Skin Signs", which includes labeled image data showing several vitamin deficiencies in the form of physical symptoms.

Dataset Period: The images included in the dataset were curated and uploaded from 2020 to 2023 on Kaggle and taken from medical journals, public domain donations, and clinical stores.

Dataset.Size: 15,575 images of skin, hair, nails, and body visual markers.

Data Type: Time-independent RGB image data (static visual data).

Additional Sources: Supplemental macro-level references on symptoms, functions of vitamins, and signs of deficiencies were retrieved from public health databases, dermatological research, and academic research publications.

1.3 Theoretical framework

This research has dependent and independent variables:

Dependent Variable: Vitamin deficiency class predicted from the input image features (e.g., Vitamin A, B12, D, K).

Independent Variables: The visual characteristics in the images—i.e., skin pigmentation, nail dryness, glossitis, hair thinning, etc.—are used as the inputs to the deep model.

The model makes predictions from a multi-input convolutional neural network structure that separately processes features from every anatomical region. These visual features are the chief basis upon which the detection and classification of the deficiency type present can be achieved. The research seeks to project visible symptoms onto probable deficiencies through image-based AI interpretation, providing a non-invasive, scalable diagnostic option.

Equations

The operation extracts patches from the low-resolution image Y using a set of pretrained filters. These patches will represent as a high-dimensional vector. These vectors comprise a set of feature maps and have the same number of feature maps as its dimensional vectors. The first step is described by an operation X1:

$$X1(Y) = \max(0, W1 * Y + B1)$$

Let, W1 and B1 represent the filters and biases, respectively. W1 has dimension $C \times f1 \times f1 \times n1$ with g as the number of channels in the input image, f1 representing the spatial size of the filter, and k1 as the number of filters. Intuitively, W1 applies n1 convolutions on the image, each with a kernel size of $C \times f1 \times f1$. The output consists of n1 feature maps. B1 is an n1 -dimensional vector, where each entry is connected to a filter.

II. RESEARCH METHODOLOGY

The methodology section outline the plan and method that how the study is conducted. This includes Universe of the study, sample of the study, Data and Sources of Data, study's variables and analytical framework. The details are as follows;

2.1Population and Sample

The dataset used for this project was obtained from Kaggle, a well-known platform for open- access datasets. Specifically, the dataset titled "Vitamin Deficiency Skin Signs" was selected, which contains a comprehensive collection of images of body parts, categorized by Vitamin deficiencies. The dataset is divided into four directories, each representing a distinct Vitamins:

Dataset Distribution for Training and Validation

Class	**Training images count**	**Validation images count**	**Total images**
Body	1478	369	1847
Hair	191	48	239
Nail	832	208	1040
Skin	9945	2486	12,431
Total	12,446	3111	15,575

2.2 Data and Sources of Data

Secondary data has been gathered for this research. From the website of Kaggle, which is an open data platform, a labeled image dataset named "Vitamin Deficiency Skin Signs" was retrieved. The dataset comprises 15,575 RGB images of different parts of the body like skin, hair, nails, and body, each displaying observable symptoms of varied vitamin deficiencies.

Moreover, for validation of the clinical significance of the work, corroborative information regarding vitamin roles and corresponding deficiencies has been obtained from widely acclaimed health websites like the World Health Organization (WHO) and scholarly papers indexed in IEEE Xplore, Elsevier, and ResearchGate. These sources have been referenced to normalize image tags with medically confirmed symptoms.

The dataset of this project is not a particular time series because it has static visual data. The data collection and curation duration is around 2020 to 2023, as estimated from upload timestamps and publication references on Kaggle.

This secondary data is the basis of the study and is used as the input of the deep learning model that conducts image-based classification of vitamin deficiency in real-time and non-invasive terms.

2.3 Theoretical framework

The variables of the present study are divided into dependent and independent variables, which are aimed at studying how effective deep learning is in the classification of vitamin deficiencies from visual images.

Dependent Variable:

The class of predicted vitamin deficiency (e.g., Vitamin A, B12, D, K) serves as the dependent variable. It is the output result of the system and is obtained from the extracted features of the input images through a trained CNN model.

Independent Variables:

The independent variables for this research are the visual features in the image data. These are:

Skin pigmentation and discoloration

Nail brittleness and shape abnormalities

Tongue swelling or glossitis

Hair thinning or patchiness

Visibly apparent body signs (e.g., dryness or rashes)

These features are obtained from RGB images through EfficientNetB0-based CNN branches, and each branch takes a single anatomical area—skin, hair, nail, or body—to process independently. The extracted features are concatenated and mapped into one of the deficiency classes using a softmax output layer.

Throughout this work, preprocessing and augmentation are applied to each image, and this ensures that extracted visual cues are clean, boosted, and reflective of real-world conditions.

This is similar to how macroeconomic variables (such as interest rates or inflation) in financial modeling serve as predictors of stock returns. In our instance, visual manifestations are predictors (independent variables), and the label for vitamin deficiency is the outcome being predicted (dependent variable).

The research points out that image-based variables, if pulled out by a deep CNN model, can be good indicators for diagnosis purposes, offering a contrast to the conventional blood-based diagnostics.

2.4 Statistical Tools and Computational Models

This step elucidates the computational and statistical methods used to convert the obtained image data into useful diagnostic findings. The research utilizes a mix of deep learning, performance metrics, and monitoring tools for models to extract inferences from the processed dataset. The methodology is explained in depth below:

2.4.1 Descriptive Statistics and Performance Metrics

Whereas traditional statistical measures such as mean, standard deviation, and variance are used widely in economic models, performance measures in this study form the statistical basis for measuring model effectiveness. They are:

Accuracy

Refers to the overall percentage of accurately predicted vitamin deficiency classes in the entire dataset.

Precision

Reports the number of actually correct positive results among the predicted ones—necessary for knowing how precise the model is to each deficiency.

Recall

Demonstrates how well the model identifies all real positive cases—valuable in medical diagnostics to avoid false negatives.

F1 Score

A harmonic mean of precision and recall that gives a balanced measure of a model's performance.

AUC-ROC Curve

Estimates the model's sensitivity and specificity for various threshold levels, particularly useful in multi-class classification.

2.4.2 Validation and Normality Consideration

Unlike traditional econometrics, deep learning models rely less on normality assumptions and data distribution analysis and augmentation. Nevertheless, for robustness and generalizability purposes, the dataset went through preprocessing procedures such as normalization (rescaling pixel values to [0,1]) and data augmentation to mimic variability seen in reality.

Although statistical tests such as the Jarque-Bera test are employed in financial research to test for normality and identify arbitrage, in our image-based environment, variability introduced by augmentation is being used as a surrogate to mimic non-uniform conditions of the real world (e.g., lighting, texture, angle).
This diversity within the data limits overfitting and enhances model generalization—just as normality of economic data reduces bias and makes inferences more valid.

2.5 Model Evaluation through Epoch-Wise Performance and Multi-Input Fusion

Statistical regression techniques like Fama-MacBeth two-pass regression are replaced by epoch-wise training evaluation and multi-input feature fusion methods in deep learning-based classification issues such as detection of vitamin deficiency to evaluate model accuracy and dependability with respect to training iterations and class outputs.
Rather than validating whether "beta" factors can forecast asset returns (in financial models), the major question in this research is whether visually extracted features from images (e.g., skin, hair, nails, body) result in a correct forecast of vitamin deficiency.
First Pass – Feature Extraction (Similar to Time Series Regression)
In the initial model development stage, visual features from four individual anatomical areas are captured through EfficientNetB0 CNN branches. Every input image (224x224 RGB) passes through a specific pipeline that undergoes convolutional, pooling, and normalization processes to capture high-level patterns like color variations, textures, and shapes. These are equivalent to estimating separate "betas" for various risk factors in Fama-MacBeth first pass.
Second Pass – Feature Combination and Classification (Similar to Cross-Sectional Regression)
All the four branch outputs (feature maps) are subjected to Global Average Pooling and concatenated to create a complete vector. This vector is fed into a Dense Softmax Layer that makes the image go into one of the vitamin deficiency classes. This is analogous to the second pass in Fama-MacBeth, in which betas are employed to forecast returns, and in this case, combined features are employed to forecast vitamin class probabilities.

Epoch-Wise Model Performance Evaluation (Statistical Reliability Analysis)

Rather than regression diagnostics such as Durbin-Watson, the paper assesses model performance based on:
Epoch-wise loss and accuracy monitoring (for 20 epochs)
Confusion matrix to quantify class-wise prediction errors
F1-score and ROC curve to evaluate model generalization
Validation accuracy, which consistently rose from 82.55% to 87.03% during training

This training behavior guarantees the model learns generalizable patterns and prevents overfitting same as financial models need stable estimates of beta over time and across assets for reliability.

Model Robustness Criteria

Similar to omitting low R^2 or high p-value securities in financial regression:
Noisy or low-quality images were excluded during preprocessing
Images were normalized and resized
Augmentation provided robustness over lighting, angles, and variations
A post-processing correction layer was used to enhance the accuracy, particularly for Vitamin K classification

2.5.1 Model for Multi-Branch CNN

During the initial phase of model training, a four-branch Convolutional Neural Network (CNN) architecture is employed to process four images for different anatomical regions separately: skin, hair, nails, and body. Each image is fed into a pre-trained EfficientNetB0 model, which obtains deep features pertaining to identifying visual symptoms of vitamin deficiency.

The generalized transformation of image

$$F_i = f_{EffNetB0}(I_i)$$

Where F_i = Feature vector for region
I_i = Input image of region
$f_{EffNetB0}$ = Feature extraction function using EfficientNetB0 model

2.5.2 Model Fusion and Softmax Classification

Following feature extraction from every input stream, the results from all four EfficientNetB0 branches are combined to create a single feature vector. This aggregated representation is submitted to a fully connected Dense layer with Softmax activation for multi-class classification. The Softmax function returns class probabilities for deficiencies including Vitamin A, B12, D, or K.

The classification function can be represented as:

$$y = \text{Softmax}(W.[F_{skin}\ , F_{hair}\ , F_{nail}\ , F_{body}] + b)$$

Where y = Predicted class probabilities
W = Weights of the dense layer
b = Bias term
$[F\ldots]$= Concatenated feature vectors from each image region

2.5.3 Comparison of the Models

To compare the performance of the suggested model, a comparison was made between the simplest CNN architecture and

- Accuracy
- Precision
- Recall
- Prediction time

  The results evidently show that the multi-input EfficientNetB0-based model performs much better compared to simpler CNN baselines, with overall accuracy achieving up to 87.03%, and an F1-score of 0.87, with superior reliability and classification capabilities.

2.5.2.1 Davidson and MacKinnon-Inspired Comparative Evaluation

In classical finance, the Davidson and MacKinnon equation is employed for model comparison of non-nested models such as CAPM and APT. Likewise, in this study, two non-nested CNN architectures the baseline CNN model and the new multi-branch EfficientNetB0 model are compared to decide which model is a better image-based predictor of vitamin deficiencies.

We borrow the Davidson and MacKinnon evaluation method and use a weighted performance measure to quantify the predictive performance of the two models as follows:

$$Y_i = \alpha.Y_{proposed} + (1-\alpha) . Y_{baseline} + \epsilon_i$$

Where , Y_i = Final predicted class probabilities
$Y_{proposed}$ = Prediction from the EfficientNetB0-based model
$Y_{baseline}$ = Prediction from the baseline CNN
α = Weight indicating comparative model effectiveness
ϵ_i= Error term

If α is nearing 1, then it suggests that EfficientNetB0-based model outperforms the baseline CNN by a great extent and is thus a better and more reliable classifier.

2.5.2.2 Posterior Odds Ratio Analysis

To provide further confirmation of one model being superior to the other, a Posterior Odds Ratio (motivated by Zellner, 1979) is calculated based on Error Sum of Squares (ESS) of both models. Here, ESS stands for the total of misclassification errors over the validation set.

The formula applied to compare models is:

$$\mathbf{R = (ESS_{baseline} / ESS_{proposed})^{N/2} \cdot N^{(K_{baseline} - K_{proposed})/2}}$$

Where , $ESS_{baseline}$ = Total classification error from the baseline CNN model
$ESS_{proposed}$ = Total classification error from the proposed EfficientNetB0 model
N = Number of samples used in evaluation
K = Number of parameters (e.g., trainable weights) in each model

Interpretation:
If R>1: The baseline CNN is more supported by the validation data
If R<1: The EfficientNetB0-based model is more strongly supported

III. Results of Descriptive Statistics of Study Variables

Here, descriptive statistics were applied to describe the properties of the dataset used for training as well as validation of the vitamin deficiency detection model. RGB image pixel values and performance metrics like accuracy, precision, recall, and loss values over various training epochs are the main variables that were considered.

The data collection utilized has 15,575 annotated images spread over four classes and four anatomical categories of skin, hair, nail, and body. Preprocessed and augmented images were used, and their pixel distributions were normalized to enhance learning efficiency and generalizability.

3.1 Descriptive Statistics of Model Input and Training Variables

Table 3.1: Descriptive Statics

Variable	Minimum	Maximum	Mean	Std. Deviation	Jarque-Bera test	Sig
Training Accuracy	0.755	0.984	0.865	0.067	3.462	0.178
Validation Accuracy	0.825	0.870	0.858	0.016	1.902	0.387
Training loss	0.128	0.380	0.245	0.064	2.143	0.343
Validation loss	0.215	0.305	0.262	0.025	1.667	0.436
F1 Score	0.810	0.870	0.868	0.014	1.211	0.547

Table 4.1 shows the mean, standard deviation, minimum, and maximum of the major performance indicators during the training period. Furthermore, the Jarque-Bera test was employed to determine the normality of the model performance measures, which is paramount in ensuring statistical reliability.

Normality Testing

Null Hypothesis (H_0): The data is normally distributed

Alternative Hypothesis (H_1): The data is not normally distributed

For a 5% significance level, the null hypothesis of normality could not be rejected for any variable. This means that model training and validation metrics had a distribution pattern that was stable and normally distributed during the experiment.
The validation accuracy and loss consistency across epochs further indicate that the model was not sensitive to outliers or variance and that there were no overfitting phenomena. Consequently, the suggested system is statistically reliable and stable for vitamin deficiency prediction from visual symptoms.

IV Acknowledgment

The authors gratefully acknowledge the great guidance, persistent support, and motivation given by Dr. D. N. Hire, under whose guidance this project was successfully submitted. The contributors — Vishakha Makasare, Shreya Pawar, and Sakshi Hambir thank the open-source community, particularly the dataset contributors of the Kaggle dataset "Vitamin Deficiency Skin Signs", which was the central dataset for training and testing. The group also credits the use of platforms and libraries like TensorFlow, Keras, and Flask that facilitated the successful deployment and implementation of the deep learning-based diagnosis system. The project was carried out strictly for academic reasons without financial sponsorship or institutional support.

References

[1] Antonio Augusto Velasco Cruz M.D., Ph.D., Fla´via A. Attie´-Castro M.D., Sandra L. Fernandes M.D., Jussara Fialho F. Cortes M.D., Paulo de Tarso P. Pierre-Filho M.D, Eduardo Melani Rocha M.D., Ph.D., Ju´lio Se´rgio Marchini M.D., Ph.D. ”Adult blindness secondary to vitamin A deficiency associated with an eating disorder”, International Journal of Applied and Basic Nutritional Science, vol. 21, no.5, pp.630-633 2005, DOI: 10.1016/j.nut.2004.12.003

[2] Alfred Sommer, ”Vitamin a deficiency and clinical disease: an historical overview”, Journal of Nutrition, vol. 138, no.10, pp.1835 to 1839. 2008, DOI: 10.1093/jn/138.10.1835.

[3] Jordi Graells MD, Rosa Maria Ojeda MD, Cristina Muniesa MD, Jesus Gonzalez MD, ”Glossitis with linear lesions: An early sign of vitamin B12 deficiency”, Journal of the American Academy of Dermatology, vol. 60, no.3, pp. 498-500. 2009, DOI:10.1016/j.jaad.2008.09.011

[4] Archana Ajith, Vrinda Goel, ”Digital Dermatology Skin Disease De- tection Model using Image Processing”, International Conference on Intelligent Computing and Control Systems, 2017, DOI: 10.1109/IC-CONS.2017.8250703

[5] C Pe´tavy-Catala, V Fonte`s, N Gironet, B Hu¨ttenberger, G Lorette, L Vaillant, ”Clinical manifestations of the mouth revealing Vitamin B12 deficiency before the onset of anemia”, Journal of the European Academy of Dermatology and Venereology vol.17, no. 1, 2003 Feb, DOI: 10.1046/j.1468-3083.2003.00545.x.

[6] Harshavardhan J R, Vaishnavi M, K R Sahana, Sneha A S, Sanjana G ”Vitamin deficiency detection using image processing”, International Journal for Research in Applied Science and Engineering Technology, 2023, DOI: 10.22214/ijraset.2023.56822.