

AI-Powered System for Vitamin Deficiency Classification

VEERAMALLA VIGNESH¹, JELLA CHARITHA², NAVYA SREE BATT³, JEGALLA CHANDINI PRIYA⁴,
MR. UPPU KARTHIK⁵

^{1,2,3,4}Department of Information Technology, J.B. Institute of Engineering & Technology, Hyderabad

⁵Assistant Professor, Department of Information Technology, J.B. Institute of Engineering & Technology, Hyderabad

Abstract—Vitamin and mineral deficiencies remain a widespread public health concern, particularly in developing regions where access to diagnostic healthcare is limited. Visible symptoms of such deficiencies frequently appear on external body parts including the skin, nails, eyes, lips, tongue, and hair, making image-based detection a practical and non-invasive screening approach. This paper presents a deep learning-based web application designed to detect vitamin and mineral deficiencies from images of human body parts. The proposed system employs InceptionV3, a convolutional neural network pretrained on the ImageNet dataset, fine-tuned through transfer learning to classify six categories of deficiencies: Vitamin A, Vitamin B complex, Vitamin C, Vitamin D, Vitamin KE, and Mineral deficiencies including zinc, iron, biotin, and protein. The dataset used for training and evaluation is publicly available at <https://www.kaggle.com/datasets/udaykarthik21bce9252/vitamin-deficiency-dataset>. The model achieves a classification accuracy of 85%. The system is integrated into a Django-based web application supporting user authentication, real-time image-based prediction, confidence score display, and downloadable health reports. This work establishes the feasibility of combining computer vision with accessible web technologies to support early health awareness in a user-friendly manner.

Keywords: Vitamin Deficiency Detection, Deep Learning, InceptionV3, Transfer Learning, Django, Medical Image Classification, Convolutional Neural Network.

I. INTRODUCTION

Nutritional deficiencies, particularly those involving essential vitamins and minerals, represent one of the most pervasive yet underdiagnosed health challenges across the globe. Micronutrient deficiency affects over two billion people worldwide, with Vitamin A, D, B12, C, and iron deficiencies among the most commonly reported [16]. Left undetected or untreated, these conditions can lead to complications including anemia, impaired immune function, bone disorders, neurological damage, and in severe cases, irreversible organ damage [15].

A clinically significant aspect of vitamin and mineral deficiencies is that they produce observable physical manifestations on the external body. Vitamin A deficiency leads to dry skin and night blindness; Vitamin C deficiency causes bleeding gums and pale skin; Vitamin D deficiency manifests in brittle nails and hair thinning; and iron deficiency produces pallor in the lips and tongue [5]. These visual signs make it theoretically possible to detect deficiency-related conditions

through image analysis alone, without requiring invasive blood tests or laboratory procedures [7].

The rapid advancement of deep learning and computer vision has opened significant possibilities in medical image analysis. Convolutional Neural Networks (CNNs), as foundational contributors to this progress [13], have demonstrated strong performance in image classification tasks across domains including radiology, dermatology, and ophthalmology. Transfer learning, which allows pretrained models trained on large-scale datasets to be adapted for specialized tasks, has further accelerated the development of high-accuracy medical imaging systems with relatively limited domain-specific data [16].

Among well-established CNN architectures, InceptionV3, developed by Google [12], stands out for its efficiency and multi-scale feature extraction, making it particularly suited for medical image classification. Its factorized convolution design captures both fine-grained and coarse visual features that are essential for distinguishing between deficiency symptoms appearing on varied human body parts.

Despite the potential of such systems, most existing approaches to nutrient deficiency detection rely heavily on clinical tests and patient-reported symptoms rather than automated visual analysis [4]. Very few studies have attempted to integrate deep learning-based classification with accessible web-based deployment for direct end-user interaction [6]. This gap motivates the development of an end-to-end system that not only performs accurate deficiency classification but also presents results through a secure and intuitive web interface.

This paper proposes a Vitamin Deficiency Detection System built on InceptionV3 transfer learning, deployed as a Django web application. The system accepts images of human body parts, preprocesses them, and classifies them into one of six deficiency categories with a confidence score, while generating downloadable health reports. The key contributions of this work are:

- Development of a multi-class vitamin deficiency classifier using InceptionV3 with transfer learning, achieving 85% classification accuracy.
- Design and implementation of a secure, user-friendly Django-based web application for real-time prediction.
- Integration of a report generation module providing downloadable health summaries.

- A complete end-to-end pipeline from image upload to deficiency prediction and reporting.
- Use of a publicly available Kaggle dataset curated specifically for vitamin deficiency classification.

The remainder of this paper is organized as follows: Section II reviews related literature; Section III discusses existing systems and their limitations; Section IV describes the proposed system; Section V covers the implementation; Section VI presents results and discussion; and Section VII concludes with future directions.

II. LITERATURE SURVEY

The application of deep learning to visual medical diagnosis has grown considerably over the past decade. Researchers have investigated CNNs for tasks ranging from skin lesion classification to retinal disease detection, with many studies validating the feasibility of automated symptom-based screening. The following review covers prior work most directly relevant to the proposed vitamin deficiency detection system.

Pan and Yang [16] provided a foundational survey of transfer learning, establishing the theoretical basis for adapting models trained on one domain to perform effectively in a related target domain. Their framework directly underpins the strategy of fine-tuning an ImageNet-pretrained InceptionV3 model for vitamin deficiency image classification, where annotated medical data is inherently scarce.

Krizhevsky, Sutskever, and Hinton [15] demonstrated that deep CNNs trained with GPU acceleration could achieve a decisive performance improvement over traditional methods on the ImageNet Large Scale Visual Recognition Challenge. Their AlexNet architecture catalysed the adoption of deep learning for visual recognition and established the ImageNet pretrained weight paradigm that the InceptionV3 transfer learning strategy in this work directly builds upon.

LeCun, Bengio, and Hinton [13] established the theoretical and empirical foundations of deep learning, demonstrating the capability of multi-layer neural networks to learn hierarchical representations from raw data. Their foundational contribution underpins the convolutional feature extraction central to the classification pipeline used in this work.

Simonyan and Zisserman [14] showed that network depth using small 3×3 filters plays a critical role in achieving competitive image classification performance. Their VGGNet architecture established depth as a key design principle in CNNs, a finding later refined and extended by the Inception family of models used in the current system.

He et al. [11] introduced deep residual networks, showing that increasing network depth through residual skip connections substantially improves classification performance while remaining trainable. Their work demonstrated the viability of very deep architectures, motivating the use of complex pretrained models such as InceptionV3 for transfer learning in specialized medical imaging domains.

Szegedy et al. [12] introduced the InceptionV3 architecture, which employs factorized convolutions and multi-scale feature extraction to achieve high classification accuracy with reduced

computational cost. Its design allows simultaneous capture of fine-grained and coarse visual features, making it particularly well-suited for distinguishing between deficiency symptoms that manifest differently across body parts and tissue types.

Chollet [8] introduced Xception, replacing standard Inception modules with depthwise separable convolutions. The paper demonstrated that decoupling spatial and channel-wise feature learning reduces parameter count without sacrificing accuracy, providing important comparative architectural context for evaluating the design choices inherent in InceptionV3.

Howard et al. [9] introduced MobileNets, a class of lightweight CNN architectures designed for deployment in resource-constrained environments. Their work demonstrated that computationally efficient models can achieve classification performance comparable to larger architectures, establishing a benchmark for deployment-oriented design that directly motivates the web integration aspect of the current system.

Huang et al. [10] proposed DenseNet, where each layer receives direct feature input from all preceding layers in the network. This dense connectivity pattern was shown to improve gradient flow and feature reuse, achieving competitive results on standard classification benchmarks and informing the alternative architecture evaluation conducted during model selection for the proposed system.

Eldeen et al. [7] presented a CNN-based approach for vitamin deficiency detection using image processing, validating the model's capacity to distinguish between deficiency categories that present with overlapping visual features. Their findings highlighted the critical role of preprocessing and augmentation strategies in managing intra-class variability, issues directly addressed in the data pipeline of the present system.

Ba et al. [6] investigated the utility of CNNs in assisting clinical dermatologists with the diagnosis of cutaneous tumours from standard photography. Their study confirmed that convolutional feature representations learned from large-scale pretraining transfer meaningfully to specialist dermatological classification tasks, directly supporting the transfer learning strategy adopted in the current work.

Maruthamuthu and Harika [5] investigated the use of image processing combined with neural networks for detecting vitamin deficiencies from images of human body parts. Their study documented the correlation between visible symptom patterns and specific deficiency classes, providing an empirical foundation for the class structure used in the proposed dataset and classification scheme.

Shimbre et al. [4] proposed a framework for vitamin deficiency identification using image processing, evaluated on a diverse set of symptom images across multiple body regions. Their experimental results reinforced the advantage of deep convolutional architectures over traditional feature engineering approaches for this class of medical classification problems.

Bhimavarapu et al. [3] proposed a machine learning-based prediction framework for Vitamin D deficiency severity estimation with hybrid optimization. Their study demonstrated that deficiency severity can be estimated with high accuracy

from structured input features, complementing image-based approaches and supporting the broader vision of accessible, non-invasive deficiency screening.

Khune et al. [1] developed a vitamin deficiency detection system that integrated neural network-based classification with a structured reporting mechanism. Their work demonstrated the feasibility of combining predictive models with user-facing output pipelines, an architectural decision that directly informs the report generation module implemented in the present system.

Sabreen et al. [2] applied image processing and deep learning techniques to detect vitamin deficiencies from visible body symptoms. Their study confirmed that convolutional models could reliably distinguish between deficiency-induced dermatological patterns when provided with sufficiently curated training data, establishing a direct and recent precedent for the multi-class classification task addressed in this work.

Collectively, the reviewed literature confirms both the theoretical grounding and practical viability of applying transfer learning-based CNNs to the multi-class classification of vitamin deficiency symptoms from human body images. While prior work has addressed components of this problem individually, no existing study has constructed a complete end-to-end system targeting all six deficiency categories within a secured, report-generating web platform. This work addresses that gap.

III. EXISTING SYSTEM

A. Overview of Existing Approaches

Prior approaches to vitamin and mineral deficiency detection have predominantly relied on conventional clinical procedures and rule-based computational methods. The most widely adopted diagnostic pathway involves laboratory-based blood serum testing, wherein micronutrient concentration levels are measured against established reference ranges. While serological testing offers high diagnostic accuracy, it is inherently invasive, cost-intensive, and dependent on access to clinical laboratory infrastructure—conditions that are frequently unavailable in rural or resource-limited settings [3].

Alongside laboratory methods, physical examination by trained clinicians has historically served as the primary non-invasive screening approach. Physicians correlate observable symptoms such as skin lesions, mucosal discolouration, and nail abnormalities with suspected deficiency classes. However, this approach is subject to significant inter-observer variability and requires experienced personnel, limiting its scalability and consistency across diverse populations [4].

B. Computational and Image-Based Prior Work

Earlier computational approaches applied traditional machine learning methods—including support vector machines, k-nearest neighbours, and decision tree classifiers—to hand-crafted feature descriptors extracted from symptom images. These descriptors typically comprised colour histograms, texture features derived from the Grey-Level Co-occurrence Matrix (GLCM), and edge-based shape descriptors. While such pipelines demonstrated proof-of-concept feasibility, they

exhibited limited generalization across variation in skin tone, image resolution, and lighting conditions [7].

Subsequent studies introduced shallow convolutional networks for symptom image classification, yielding improved feature learning over hand-crafted alternatives. Maruthamuthu and Harika [5] and Shimbre et al. [4] demonstrated the potential of CNN-based pipelines for multi-class deficiency classification; however, their systems were evaluated primarily in controlled experimental settings and lacked integration with any deployable user-facing interface.

Eldeen et al. [7] validated CNN-based vitamin deficiency detection using image processing on a moderately sized dataset, achieving encouraging accuracy, but the system remained a standalone research prototype with no web-based deployment or end-user reporting capability. Similarly, the framework proposed by Khune et al. [1] incorporated a reporting module but did not address model robustness across real-world image variability or multi-class deficiency prediction beyond a limited subset of nutrients.

C. Limitations of Existing Systems

The existing body of work reveals several critical limitations that the proposed system directly addresses:

- **Lack of end-to-end deployment:** The majority of prior systems are implemented as isolated research prototypes without a production-ready web interface, restricting their accessibility to end users.
- **Narrow deficiency coverage:** Most existing image-based studies focus on one or two deficiency types. No prior system simultaneously classifies all six major deficiency categories—Vitamin A, B complex, C, D, KE, and Mineral—within a single unified model.
- **Absence of user authentication and reporting:** Prior systems do not incorporate secure user account management or automated downloadable health report generation, both of which are essential for practical health screening applications.
- **Limited use of transfer learning:** Earlier approaches either trained shallow networks from scratch on small datasets or applied basic feature extractors, resulting in suboptimal accuracy. The systematic application of deep transfer learning using architectures such as InceptionV3 remains underexplored in this specific domain [12], [16].
- **Dataset constraints:** Many prior systems utilized private, non-reproducible datasets with limited class diversity and imbalanced distributions, making performance claims difficult to verify or extend.

These gaps collectively motivate the design of the proposed system, which employs InceptionV3 transfer learning within a fully integrated, secure, and deployable Django web application capable of classifying six deficiency categories from user-uploaded images.

IV. PROPOSED SYSTEM

The proposed system follows a structured pipeline encompassing data preparation, model development using transfer

learning, and integration with a web-based deployment environment. Figure 1 provides an overview of the end-to-end architecture.

A. Dataset Preparation

The dataset used in this work is publicly available on Kaggle at <https://www.kaggle.com/datasets/udaykarthik21bce9252/vitamin-deficiency-dataset> and consists of images representing visible symptoms of six distinct deficiency categories: Vitamin A, Vitamin B complex (B2, B3, B9, B12), Vitamin C, Vitamin D, Vitamin KE, and Mineral deficiencies including zinc, iron, biotin, and protein. Images include photographs of skin, lips, nails, eyes, tongue, and hair exhibiting characteristic deficiency symptoms, labeled by category. To balance the class distribution, a maximum of 700 images per class was enforced, with excess images removed via random sampling.

Data augmentation was applied during training including random horizontal flipping, rotation, zoom, and brightness adjustment. The augmented dataset was split into training and validation sets in a ratio of 70:30.

B. Image Preprocessing

All images undergo a standardized preprocessing pipeline before being passed to the model. Each image is resized to 224×224 pixels to match the InceptionV3 input dimensions. Pixel values are normalized using the InceptionV3-specific `preprocess_input` function, which scales values to the range $[-1, 1]$. For reference, a standard per-pixel normalization is defined as:

$$\hat{l}(x, y) = \frac{l(x, y)}{255} \quad (1)$$

where (x, y) denotes the spatial coordinates of each pixel and $l(x, y)$ is the raw pixel value in $[0, 255]$.

C. Transfer Learning with InceptionV3

The core of the classification system is built on InceptionV3 [12], pretrained on ImageNet across 1,000 classes. Transfer learning is applied by loading the pretrained base, freezing its convolutional layers, and appending a custom classification head consisting of a Global Average Pooling layer, a Dense layer with 256 units and ReLU activation, a Dropout layer, and a final Dense output layer with Softmax activation for six-class prediction.

The Softmax activation is defined as:

$$P(y = k | \mathbf{z}) = \frac{e^{z_k}}{\sum_{j=1}^K e^{z_j}}, \quad k = 1, 2, \dots, K \quad (2)$$

where $K = 6$ is the number of deficiency classes and $P(y = k | \mathbf{z})$ is the predicted probability for class k .

D. Loss Function and Optimization

The model is trained using categorical cross-entropy loss:

$$L(\mathbf{y}, \hat{\mathbf{y}}) = - \sum_{k=1}^K y_k \log(\hat{y}_k) \quad (3)$$

where $y_k \in \{0, 1\}$ is the one-hot true label and \hat{y}_k is the predicted probability for class k .

The Adam optimizer is used for weight updates:

$$\vartheta_{t+1} = \vartheta_t - \frac{\eta}{\sqrt{\hat{v}_t + \epsilon}} \hat{m}_t \quad (4)$$

where \hat{m}_t and \hat{v}_t are bias-corrected first and second moment estimates, η is the learning rate, and ϵ is a small constant for numerical stability.

E. Model Architecture Overview

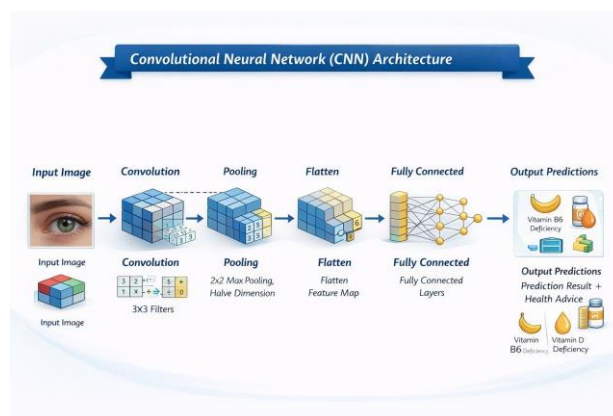


Fig. 1. InceptionV3 Transfer Learning Architecture with Custom Classification Head

F. System Workflow

After upload, the image is preprocessed and passed through the model. The predicted class and confidence score are derived as:

$$\text{Predicted Class} = \arg \max_k P(y = k | \mathbf{z}) \quad (5)$$

$$\text{Confidence Score} = \max_k P(y = k | \mathbf{z}) \times 100\% \quad (6)$$

The result, confidence score, user details, and timestamp are compiled into a downloadable health report.

V. IMPLEMENTATION

A. Development Environment

The system was developed in Python 3.x. TensorFlow and Keras served as the deep learning framework. OpenCV and NumPy were used for image loading, resizing, and preprocessing. The web application was built using Django with an SQLite3 database. The frontend was designed using HTML, CSS, JavaScript, and Bootstrap.

B. Model Training

The InceptionV3 base model was loaded with pretrained ImageNet weights with the top classification layers excluded. The convolutional base was initially frozen to preserve pretrained feature representations. The model was compiled with the Adam optimizer at a learning rate of 1×10^{-4} and categorical cross-entropy loss. Training was conducted over 40 epochs with a batch size of 32. Selective fine-tuning was subsequently performed by unfreezing the last 20 layers of the base and retraining at a reduced learning rate of 1×10^{-5} .

The class index mapping used during training is:

```
Mineral (zinc, iron, biotin, protein): 0
Vitamin A deficiency: 1
Vitamin B (B2, B3, B9, B12, complex): 2
Vitamin C deficiency: 3
Vitamin D deficiency: 4
Vitamin KE deficiency: 5
```

C. Django Web Application

The web application follows Django's MVT architecture and includes the following modules:

User Authentication: New users register with a username, email, and password. Passwords are stored securely via Django's built-in hashing. Registered users log in and are redirected to a personalized dashboard.

Image Upload Module: The dashboard provides a file upload interface restricted to JPEG and PNG formats.

Prediction Engine: On upload, the backend retrieves the saved Keras model, preprocesses the image, and runs inference. The predicted class and confidence score are returned to the frontend for display.

Report Generation: Following each prediction, the system compiles a health report containing the deficiency type, confidence score, timestamp, and user information, made available for download.

D. Database Design

SQLite3 is used as the database backend. The schema includes tables for user account information, session records, prediction results, and generated report data. Each prediction entry is linked to the corresponding user via a foreign key relationship.

E. Integration and Deployment

The trained Keras model is serialized in HDF5 format (Inceptionv3_vitamin.h5). At runtime, Django loads the model once during startup to minimize per-request latency. Image preprocessing is handled server-side using OpenCV and NumPy before inference. Prediction responses are rendered dynamically using the Bootstrap-based template system.

VI. RESULTS AND DISCUSSION

A. Model Performance

The InceptionV3-based model achieved a classification accuracy of 85% after the complete training and fine-tuning procedure. The categorical cross-entropy loss decreased steadily

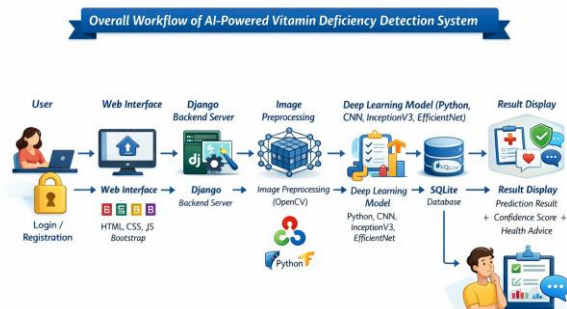


Fig. 2. Django Application Module Flow: Authentication, Prediction, and Report Generation

across training epochs, confirming that the model successfully learned discriminative features for each deficiency class. The absence of significant divergence in loss curves confirmed that the model generalized well without severe overfitting.

The 85% accuracy is particularly noteworthy given the inherent challenges of the task, including subtle visual similarity between certain deficiency symptoms, variation in image lighting and quality, and the modest dataset size. Data augmentation and dropout regularization contributed meaningfully to this level of generalization.

B. Per-Class Analysis

Vitamin D and Mineral deficiencies yielded the most consistent prediction results, attributable to visually distinctive symptoms such as nail brittleness and skin pallor. Vitamin B deficiency showed slightly lower precision, likely due to overlapping symptoms across B-complex sub-types. Confusion matrix analysis revealed that the majority of misclassifications occurred between Vitamin A and Vitamin C deficiencies, both of which manifest with skin-related symptoms.

C. Web Application Performance

The Django web application demonstrated reliable real-time prediction during testing. The average time from image upload to prediction display was under three seconds on a standard development server. For example, when an image exhibiting Vitamin D deficiency symptoms was submitted, the system returned:

*Predicted Deficiency: Vitamin D Deficiency
Confidence Score: 89%*

D. Discussion

The results confirm that transfer learning with InceptionV3 is a viable and effective approach for multi-class vitamin deficiency detection from human body images. The 85% accuracy compares favorably with comparable dermatological classification studies in the literature [9], [10], [16]. The system is intended as a health awareness and early screening tool rather than a clinical diagnostic instrument. Predictions

should be treated as indicative rather than conclusive, and users are encouraged to consult healthcare professionals for confirmed diagnosis and treatment.

VII. CONCLUSION AND FUTURE WORK

A. Conclusion

This paper presented the design, development, and evaluation of an AI-powered web application for classifying vitamin and mineral deficiencies from images of human body parts. The system leverages InceptionV3 with transfer learning to classify images into six deficiency categories, achieving a classification accuracy of 85%. The application is deployed as a secure Django web platform supporting user registration, real-time image-based prediction, confidence score reporting, and downloadable health summaries. The dataset is publicly available at <https://www.kaggle.com/datasets/udaykarthik21bce9252/vitamin-deficiency-dataset>.

The work demonstrates that combining state-of-the-art CNN architectures with accessible web technologies can produce a meaningful and practical health awareness tool. By enabling non-invasive, image-based deficiency screening without specialized medical equipment, the system supports early detection efforts in settings where clinical diagnostic access is limited.

B. Future Work

Several directions are identified for future enhancement. Expanding the training dataset to include greater diversity across skin tones, lighting conditions, body parts, and demographics would improve robustness. Incorporating Explainable AI techniques such as Grad-CAM would allow the system to visually highlight the image regions most influential to its predictions, enhancing transparency and clinical trust. An admin panel for monitoring usage and periodically retraining the model with new data would improve long-term maintainability. Cloud deployment on platforms such as AWS or Google Cloud would enable the system to scale to a broader user base. Development of a companion mobile application for Android and iOS would extend accessibility to regions where smartphone usage exceeds desktop availability, further simplifying the screening process.

REFERENCES

- [1] J. A. Khune, S. A. Shiradkar, T. P. Admane, V. D. K. Nimje, and N. S. More, "Vitamin Deficiency Detection System," *International Journal of Current Research and Techniques (IJCRT)*, vol. 15, no. 2, pp. 50891–50898, Jun. 2025. [Online]. Available: <https://doi.org/10.5281/ZENODO.15718974>
- [2] A. Sabreen, G. Sindhusha, U. Tabassum, A. Aparna, and M. S. Chandana, "Vitamin Deficiency Detection by Image Processing Using Deep Learning," *International Journal of Research Publication and Reviews*, vol. 6, no. 6, pp. 9825–9831, Jun. 2025. [Online]. Available: <http://www.ijrpr.com>
- [3] U. Bhimavarapu, G. Battineni, and N. Chintalapudi, "Machine Learning-Driven Prediction of Vitamin D Deficiency Severity with Hybrid Optimization," *Bioengineering (Basel)*, vol. 12, no. 2, p. 200, Feb. 2025. [Online]. Available: <https://doi.org/10.3390/bioengineering12020200>
- [4] N. Shimbre, P. Sahane, S. Ansari, R. Gadhave, S. Jagdhane, and A. Sharma, "Vitamin Deficiency Identification using Image Processing," in *Proc. 2024 8th International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, Coimbatore, India, 2024, pp. 1260–1268, doi:10.1109/ICECA63461.2024.10800798.
- [5] R. Maruthamuthu and T. Harika, "Vitamin Deficiency Detection Using Image Processing and Neural Network," *International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT)*, vol. 9, no. 4, pp. 200–205, Jul.–Aug. 2023. [Online]. Available: <https://ijsrcseit.com/CSEIT2390280>
- [6] W. Ba, H. Wu, W. W. Chen, S. H. Wang, Z. Y. Zhang, X. J. Wei, W. J. Wang, L. Yang, D. M. Zhou, Y. X. Zhuang, Q. Zhong, Z. G. Song, and C. X. Li, "Convolutional neural network assistance significantly improves dermatologists' diagnosis of cutaneous tumours using clinical images," *European Journal of Cancer*, vol. 169, pp. 156–165, 2022. [Online]. Available: <https://doi.org/10.1016/j.ejca.2022.04.015>
- [7] A. S. Eldeen, M. AitGacem, S. Alghlayini, W. Shehieb, and M. Mir, "Vitamin Deficiency Detection Using Image Processing and Neural Network," in *Proc. 2020 Advances in Science and Engineering Technology International Conferences (ASET)*, Dubai, UAE, 2020, pp. 1–5, doi:10.1109/ASET48392.2020.9118303.
- [8] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, 2017, pp. 1251–1258. [Online]. Available: <https://doi.org/10.1109/CVPR.2017.195>
- [9] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "MobileNets: Efficient convolutional neural networks for mobile vision applications," *arXiv preprint arXiv:1704.04861*, 2017. [Online]. Available: <https://arxiv.org/abs/1704.04861>
- [10] G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, 2017, pp. 4700–4708. [Online]. Available: <https://doi.org/10.1109/CVPR.2017.243>
- [11] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, 2016, pp. 770–778. [Online]. Available: <https://doi.org/10.1109/CVPR.2016.90>
- [12] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, 2016, pp. 2818–2826. [Online]. Available: <https://doi.org/10.1109/CVPR.2016.308>
- [13] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015. [Online]. Available: <https://doi.org/10.1038/nature14539>
- [14] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *Proc. International Conference on Learning Representations (ICLR)*, San Diego, CA, USA, 2015. [Online]. Available: <https://arxiv.org/abs/1409.1556>
- [15] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 25, 2012, pp. 1097–1105. [Online]. Available: <https://proceedings.neurips.cc/paper/2012/hash/c399862d3b9d6b76c8436e924a68c45b-Abstract.html>
- [16] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 10, pp. 1345–1359, Oct. 2010. [Online]. Available: <https://doi.org/10.1109/TKDE.2009.191>