

Department of Computer Science and Engineering (2025-26)

Project Stage-2 Report for B.E Final Year (Project Phase -1)

Batch No: 49	Guide Name: Amaresh A M
Project Title: Thermal Foot Image Analysis using Machine Learning for predicting Diabetic Neuropathy Risk	

USN	Student Name
02JST22UCS006	Agasthya Samyak Jnan
02JST22UCS043	Jayanth S
02JST22UCS017	Arvind Ningombam
01JST21CS099	Prajwal T P

Signature of Guide

(Name of Guide)

Signature of HOD

(Dr. Anil Kumar K.M)

Contents

1. Introduction	3
2. Literature review	4 - 5
3. System Requirements and Analysis	6
4. Tools and Technologies	7
5. System Design	8
6. System Implementation	9

1. Introduction

Diabetes mellitus is a chronic metabolic disorder that affects millions of people worldwide and poses a significant burden on the global healthcare system. One of the most highly prevalent complications that resulted due to diabetes is *Diabetic Neuropathy*, a condition that is characterised by progressive nerve damage caused by *hyperglycemia* (high sugar level in blood). *Diabetic neuropathy* primarily affects the peripheral nervous system and can lead to symptoms such as numbness, tingling, pain, muscle weakness and loss of protective sensation, particularly in the lower extremities.

Among the various forms of Diabetic neuropathy, *Diabetic Peripheral Neuropathy* (DPN) is the most chronic complication of diabetes affecting over 50% of its patients over their lifetime. If left undiagnosed or unmanaged, DPN can result in serious complications including foot ulcers, infections, gait abnormalities and lower limb amputations. Also, conventional diagnostic approaches are not always suitable for large scale screening.

In recent years, *infrared thermography* has emerged as a promising non-invasive imaging modality for medical diagnostics. Plantar foot thermograms capture the thermal distribution of the feet, which reflects underlying physiological processes such as blood circulation, inflammation and nerve dysfunction. In patients with diabetic neuropathy, abnormal temperature patterns and asymmetry between the left and right feet are commonly observed due to impaired autonomic nerve function and microvascular abnormalities.

This project presents a *thermogram-based diabetic neuropathy* risk classification system that utilizes deep learning techniques to analyze plantar foot thermal images. The proposed system aims to categorise patients into distinct risk groups: such as low, moderate and high risk, based on thermal abnormalities observed in foot thermograms. By providing an automated, non-invasive and objective screening tool, this approach has the potential to assist clinicians in early diagnosis, improve preventive care and reduce the incidence of diabetes-related foot complications.

2. Literature Review

Author(s), Year	Methodology	Key Findings	Limitations
Panamonta <i>et al.</i> , 2025	CNN-based risk classification	Lightweight CNNs enable potential clinical deployment	Limited clinical validation
Ma Y., 2025	Systematic review of ML techniques for DPN	Reviewed ML models, datasets, limitations, and research trends; identified lack of imaging and thermal datasets	No experimental validation; conclusions depend on availability and quality of existing studies
Shig G. <i>et al.</i> , 2025	ML risk prediction (RF, DT, etc.) with SHAP explainability	Provided comparative performance analysis of ML models with explainability	Does not incorporate medical imaging; limited external validation
Evangeline <i>et al.</i> , 2024	Deep CNN models	High accuracy in diabetic foot risk prediction	Small dataset size; limited generalization
Wu <i>et al.</i> , 2024	Hybrid ML–DL approaches	Combining thermal features with deep models improves classification accuracy	Lack of standardized acquisition protocols
Frontiers in Neuroscienc e, 2024	ML classifiers (AdaBoost, SVM)	Demonstrated potential for non-invasive, continuous neuropathy monitoring	Requires wearable hardware; lacks imaging-based biomarkers

Author(s), Year	Methodology	Key Findings	Limitations
Frontiers in Neuroscienc e, 2024	Transformer-based deep learning classifier using CCM	Improved classification performance using transformer architecture	Computationally expensive; requires large datasets
Cao <i>et al.</i> , 2023	Plantar region segmentation + feature extraction	Region-wise analysis improves sensitivity to neuropathic changes	Segmentation errors affect downstream classification
Khandakar <i>et al.</i> , 2022	Thermal Change Index (TCI) + ML classifiers (SVM, RF)	Handcrafted thermal indices effectively discriminate diabetic and non-diabetic feet; high interpretability	Relies on engineered features; limited deep feature learning
Preston F. G. <i>et al.</i> , 2022	End-to-end deep learning model using CCM	Achieved high diagnostic accuracy without manual segmentation	Limited to a single imaging modality; lacks cross-modal validation
Alam U. <i>et al.</i> , 2022	AI-assisted DPN diagnosis	Highlighted automated nerve damage metrics and AI sensitivity in subclinical neuropathy	Conceptual focus; lacks thermogram-based experimental validation
Khandakar A. <i>et al.</i> , 2021	CNN-based classification using dual-foot images	Demonstrated superior performance using bilateral foot thermograms	Limited dataset size; focused on ulcer prediction rather than explicit neuropathy risk

Table 1

From the literature summarised above, it is evident that infrared thermography is an effective non-invasive modality for assessing diabetic foot complications. Early studies primarily relied on handcrafted thermal features and classical machine learning classifiers, demonstrating that temperature distribution and asymmetry between plantar regions are strong indicators of diabetic neuropathy. Subsequent research shifted toward deep learning approaches, especially convolutional neural networks, which are capable of learning discriminative features directly from thermographic images. More recent works have focused on region-wise analysis, plantar segmentation, and lightweight CNN architectures to improve sensitivity and enable practical clinical deployment.

Despite these advancements, most existing approaches are limited by small dataset sizes, lack of standardized preprocessing pipelines, and reduced model interpretability. Furthermore, many deep learning models are computationally expensive, making them less suitable for real-time or resource-constrained healthcare environments. These limitations highlight the need for an efficient, accurate, and interpretable thermogram-based diabetic neuropathy risk classification system.

Based on the review of existing literature, the following research gaps have been identified:

- Most thermogram-based studies focus on binary classification (diabetic vs non-diabetic) rather than multi-level neuropathy risk classification.
- Many approaches rely on large and complex CNN architectures, making them unsuitable for deployment in low-resource clinical settings.
- Limited work has explored the use of lightweight transfer learning models, such as MobileNet variants, specifically for plantar thermogram-based neuropathy risk assessment.
- The lack of dataset available for training the model.

To address these gaps, the proposed work focuses on developing a thermogram-based diabetic neuropathy risk classification system using a lightweight deep learning architecture, enabling efficient, accurate, and clinically relevant risk prediction.

3. System Requirements and Analysis

The proposed system is designed to analyze thermal foot images and produce a risk classification output. From a functional perspective, the system must be capable of loading thermal image datasets, performing preprocessing and augmentation, training deep learning models, extracting relevant thermal features, and computing a final risk score.

The primary functional requirements include:

- Importing and organizing thermal foot image datasets.
- Applying data augmentation techniques to enhance dataset diversity.
- Preprocessing images by resizing and normalization.
- Training binary CNN classifiers to distinguish between control and diabetic thermal patterns.
- Extracting global thermal features from images.
- Computing a continuous risk score and mapping it to discrete risk levels.

In addition to functional requirements, several non-functional requirements guide system design. The system must be computationally efficient, as it is implemented and tested on limited-resource environments such as Google Colab. Reproducibility and modularity are essential to ensure that experiments can be repeated and extended. Interpretability is also a key requirement, as the outputs must be understandable and clinically meaningful without claiming diagnostic authority.

From an analytical standpoint, the system emphasizes ethical deployment by explicitly avoiding medical diagnosis. Instead, it frames outputs as risk indicators, aligning with responsible AI practices in healthcare. This design choice ensures that the system serves as a decision-support tool rather than a replacement for clinical expertise.

4. Tools and Technologies

The system is implemented using Python due to its extensive ecosystem for machine learning and scientific computing. TensorFlow and Keras are used as the primary deep learning frameworks, offering flexibility, scalability, and support for pretrained models. These frameworks enable rapid experimentation and efficient training of CNN architectures.

Image preprocessing and numerical computations are handled using NumPy and OpenCV, which provide efficient operations for image manipulation and feature extraction. Visualization of training progress and evaluation metrics is performed using Matplotlib. Scikit-learn is used for computing evaluation metrics such as precision, recall, F1-score, confusion matrices, and AUC.

Several pretrained CNN architectures are employed, including EfficientNet-B0 and MobileNetV2. These models are selected for their balance between performance and computational efficiency, making them suitable for large-scale experimentation and potential deployment. Transfer learning is utilized to leverage knowledge from large-scale image datasets.

Google Colab serves as the development and training platform, providing access to GPU acceleration and a cloud-based environment. Versioned experiment management is implemented through structured directory organization, ensuring that models, logs, plots, and metrics are systematically stored and documented.

5. System Design

The system design follows a modular pipeline architecture that processes thermal images from acquisition to final risk classification. The first stage involves dataset acquisition and augmentation, where various transformations are applied to improve robustness and generalization.

Next, image preprocessing standardizes all inputs by resizing images to a fixed resolution of the size 224 x 224 and normalizing pixel values where necessary. The preprocessed images are then fed into a binary pre-trained CNN classifier trained to identify abnormal thermal patterns associated with diabetic conditions.

In parallel, global thermal features are extracted from the same images. These features capture statistical and spatial properties of temperature distributions, providing interpretable indicators of physiological abnormalities. Feature normalization ensures consistency and prevents scale dominance during risk score computation.

The CNN output probability and normalized thermal features are combined using a weighted risk score formulation. This hybrid design allows the system to leverage both deep learning-based pattern recognition and physiologically meaningful statistics. Finally, the continuous risk score is mapped into low, moderate, or high-risk categories.

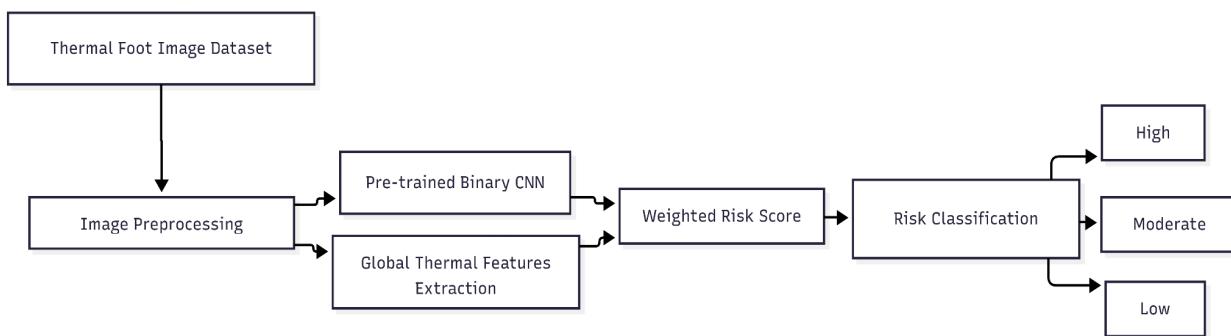


Fig 5.1 : System Architecture of proposed risk classifier model

The modular nature of the design allows each component to be independently tested, modified, or extended, making the system flexible and scalable.

6. System Implementation

The system implementation follows the designed pipeline using structured Python scripts and Jupyter notebooks. Dataset loading with batching, shuffling, and label assignment. Preprocessing steps like resizing and normalization are applied across training and validation datasets.

CNN training is conducted in two phases: an initial feature extraction phase where pre-trained layers are frozen, followed by a fine-tuning phase where selected deeper layers are unfrozen. This approach stabilizes training and improves generalization. Performance is monitored using multiple evaluation metrics, and training history is visualized through plots.

All trained models, logs, plots, hyperparameters, and evaluation metrics are automatically saved in a structured, version-controlled format.

	Accuracy	Precision	Recall	F1-Score	AUC	NPV	Loss
EfficientNet-B0 (N)	0.79510	0.79246	0.88995	0.83838	0.85461	0.80048	0.49847
EfficientNet-B0	0.89747	0.91625	0.91164	0.91394	0.96825	0.86999	0.25848
EfficientNet-B1	0.88705	0.93036	0.87646	0.90260	0.95431	0.83135	0.26242
MobileNetV2	0.91453	0.92340	0.93439	0.92886	0.97824	0.90100	0.25391

Table 6.1 : Comparison of performance of different Pre-trained CNN backbones.

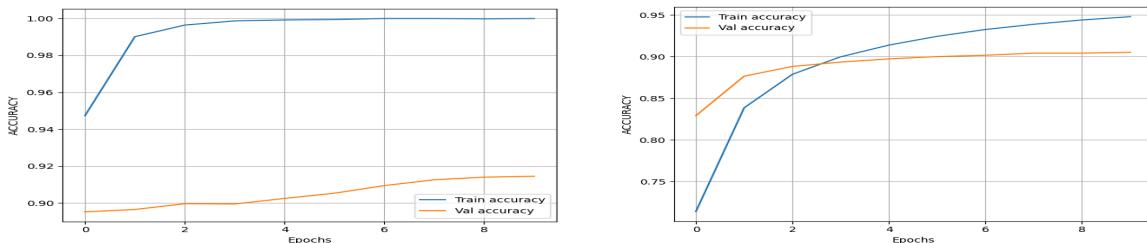


Fig 6.2 : Accuracy Plots of Training and Fine Tuning of MobileNetV2 based Classifier Model.