ZANCO Journal of Pure and Applied Sciences

TYPE Original article
DOI http://dx.doi.org/10.21271/ZJPAS.38.1.15

OPEN ACCESS

***Corresponding author**

Alaa J. Albaqal
Mardin A. Anwar

RECEIVED :12 /06 /2025
ACCEPTED :18/08/ 2025
PUBLISHED :28/ 08/ 2026

KEYWORDS:

Type 2 Diabetes Mellitus, C-reactive Protein, Image encoding, Transfer learning, Vision Transformer.

Leveraging Transfer Learning for Accurate CRP Level Prediction in Diabetic Patients

Alaa J. Albaqal, Mardin A. Anwar

Department of Software and Informatics Department, College of Engineering, Salaheddin University-Erbil, Erbil, Kurdistan Region, Iraq.

ABSTRACT

This study addresses the challenge of predicting C-reactive protein (CRP) levels in patients with type 2 diabetes mellitus by integrating tabular-to-image transformation techniques with transfer learning architectures. Utilizing a dataset of 838 clinical records, three encoding methods (Zero-padded Grid, Recurrence Plot or Gramian Angular Field) were applied to convert tabular clinical data into two-dimensional images. ResNet50 and Vision Transformer, two pre-trained models, were employed. Notably, the findings of this study suggested that predictive accuracy is highly influenced by both the model architecture and image encoding technique. The integration of spatial encoding with Vision Transformer-based models presents a promising direction for non-invasive, data-driven inflammation monitoring in diabetic care. The Vision Transformer model paired with Zero-Padded Grid achieved the highest performance, with a test accuracy of 97.62% and an F1-score of 0.90, demonstrating strong discriminative capacity with AUC values exceeding 0.99 across all classes, the Gramian Angular Field encoding image technique with the Vision Transformer model combination also demonstrated strong results, with an average AUC of 0.98, slightly higher than Zero-Padded Grid with Vision Transformer model.

1.Introduction

Diabetes mellitus is a long-term metabolic disorder that affects the body's ability to regulate blood glucose levels (ElSayed et al., 2023). It is one of the most common chronic diseases globally and is associated with various complications, including cardiovascular disease, kidney damage and impaired wound healing (Yang et al., 2024, Majeed et al., 2021). A key underlying factor in these complications is chronic inflammation. In this context, C-reactive Protein (CRP), a substance produced by the liver in response to inflammation, is widely recognized as a biomarker for assessing systemic inflammation in patients, including those with diabetes (Ridker et al., 2000, Azeez and Gallaly, 2023). Elevated levels of CRP, particularly when measured using high-sensitivity assays (hs-CRP), have been linked to increased cardiovascular risk and accelerated progression of diabetic complications (Pradhan et al., 2001). Despite the clinical importance of monitoring CRP levels, traditional laboratory-based measurements are limited by accessibility, cost and processing time. As a result, there is a critical need for predictive tools that can estimate CRP levels using routinely collected clinical and laboratory data. However, conventional statistical approaches often struggle with the complex, non-linear interactions typical of biomedical datasets, particularly when these datasets are small and imbalanced.

In recent years, artificial intelligence (AI) models, particularly machine learning (ML) and deep learning (DL), have emerged as a powerful approach for analyzing medical data (Anwer et al., 2024). These techniques can identify patterns and generate accurate predictions from complex datasets. However, traditional ML models typically required large training sets and often perform poorly when applied on small or incomplete clinical datasets, which are common in healthcare settings. To overcome this limitation, researchers have adopted transfer learning, a technique that adapts models pre-trained on large datasets to new smaller clinical datasets (Amin et al., 2021, Yu et al., 2022). In addition, advancements in data transformation techniques have been developed to convert tabular data such as clinical tables into image format. Common techniques include Zero-Padded Grid, Recurrence Plot (RP) and Gramian Angular fields (GAF), which allow deep learning models such as ResNet50 and Vision Transformers (ViT) to be applied to non-image clinical data, both models of which have demonstrated strong performance in image classification tasks (Wang and Oates, 2015, He et al., 2016, Dosovitskiy et al., 2020).

Given these developments, this study proposes a hybrid framework that combines tabular-to-image encoding methods with transfer for accurate CRP level prediction in type 2 diabetic patient. This approach has potential to improve prediction accuracy, support early diagnosis and enhance clinical decision-making in diabetes management. The remainder section of this paper is organized as follow, section two reviews related work on CRP biomarkers and machine learning and deep learning methodologies relevant to this research, section three details the materials and methods, encompassing dataset preparation, data transformation technique and the design of deep learning models, section four presents the results and discussion, offering performance metrics and interpretations, finally section, section five concludes the thesis with a summary of the main findings, study limitations and recommendations for future research directions. The main contributions in this study are focus on developing deep learning-based approach for CRP level classification in type 2 diabetes patients' inflammation monitoring. First, a cleaned and preprocessed dataset of diabetic patient's records were prepared, ensuring suitability for visual encoding. Structured clinical data were transformed into 2D image representations using Zero-Padded Grid, Recurrence Plot (RP) and Gramian Angular Field (GAF) encoding techniques. Transfer learning was applied using pre-trained models ResNet50 and Vision Transformer (ViT) for CRP level classification tasks. Finally, comparative evaluation of six model-encoding combinations were conducted using accuracy, F1-score, AUC and loss curves. The proposed framework was assessed for its potential as a non-invasive decision-support tool for monitoring systemic inflammation in diabetic care.

2.Related work

C-reactive protein (CRP) has been widely investigated due to its clinical significance in metabolic and cardiovascular conditions. The association between elevated CRP levels and increased cardiovascular risk in type 2 diabetes patients was established by (Kuppa et al., 2023), that emphasizing CRP's roles in early cardiovascular disease detection. However, their findings supported the diagnostic value of CRP, the study was observational and did not provide a predictive framework suitable for clinical use.

The significance of assessing C-reactive protein (CRP) levels is contingent upon the categorization of individuals into one of three district cardiovascular risk classification (Lee et al., 2019, Johns et al., 2018, Luthra et al., 2023) low risk (CRP <1.0 mg/dL), C-reactive protein (CRP) levels below 10 mg/dL are typically considered within the normal range and are associated with low-grade or no systemic inflammation. This range is often observed in healthy individuals and may also reflect minor elevations due to benign factors such as obesity, smoking or a sedentary lifestyle. In the context of high-sensitivity CRP, values in this range indicate a low risk for future cardiovascular events(Lee et al., 2019, Johns et al., 2018). While Moderate risk (CRP 1.0-3.0 mg/dL), CRP levels between 1.0 and 3.0 mg/dL represents a moderate increase in systemic inflammation these values may be seen in patients with chronic inflammatory conditions such as diabetes, depression or autoimmune disorders. From a cardiovascular perspective, CRP values in this range are associated with moderate risk of developing cardiovascular disease and may warrant closer monitoring and early intervention strategies (Lee et al., 2019, Johns et al., 2018). High risk (CRP >3.0 mg/dL), CRP levels above 3.0 mg/dL are indicative of significant systemic inflammation. Elevated levels in this range have been linked to acute and chronic inflammatory conditions, including infection, myocardial infraction and autoimmune diseases. In cardiac risk assessment, CRP levels above this range are strongly associated with a high risk of adverse cardiovascular events and suggest the need for comprehensive risk factor modification and medical managements (Lee et al., 2019, Johns et al., 2018).

The complexity of interpreting CRP level due to confounding factors such as infections and acute events responses was further discussed by (Mouliou, 2023). the study highlighted the limitations of conventional CRP test and called algorithmic support tools, particularly in cases where high-sensitivity CRP assays are required, it did not propose or validate a predictive model. However, no computational or AI-based models were proposed to address these challenges.

A systematic review was presented by (Islam et al., 2023) analyzing 42 studies on ML and DL applications for early sepsis detection using EHR data. Models using vitals, lab results and demographics achieved high AUC up to (0.97). Despite strong predictive performance, major limitations were noted, including imbalanced datasets, lack of external the need for standardization, explain ability and integration with clinical biomarkers.

A comparative study was conducted by (El Massari et al., 2022) to evaluate six classical machine learning classifiers SVM, KNN, ANN, LR, NB and DT alongside on ontology-based classification model using the Pima Indian diabetes dataset. Models were assessed using accuracy, precision, recall, F-score and ROC area, with the ontology classifier achieving the highest accuracy (77.5%). Although the approach outperformed conventional ML methods, it did not consider biomarker prediction, deep learning integration or image-based data representation.

Comparisons between traditional regression models with ensemble machine learning models were presented by (Huang et al., 2022), who investigated predictors of diabetic nephropathy. The study found that models such as (CART, RF, SGB and XGBoost) outperformed linear regression in term of predictive power. However, interpretability remained a concern, especially when modeling nonlinear and high-dimensional clinical data.

A systematic review was conducted by (Fregoso-Aparicio et al., 2021), covering 90 studies between 2017-2021 that applied ML and DL techniques to type 2 diabetes prediction. Eighteen modeling strategies were identified, with tree-based algorithms achieving the often lower. The review highlighted that feature selection, class balancing and dataset structure greatly influenced results. However, no study explored image transformation or predictive modeling of inflammatory biomarkers.

Deep learning models have also been investigated in this domain. (Mo et al., 2025) developed a binary classification model using deep neural network (DNN) to distinguish between normal (≤ 10 mg/L) and abnormal (>10 mg/L) CRP levels based on complete blood count (CBC) feature. The model was trained on large dataset. The used large dataset (53,834) for development and (20,723) samples for external validation, their model achieving moderate performance (AUC = 0.82) and showed potential in guiding CRP test decision. The study identified the CRP10-C2 model as best perform, noting its potential to guide clinicians in ordering CRP tests. However, their binary classification approach lacks granularity and does not reflect clinical stratification (low, moderate, high).

Transfer learning is increasingly applied to overcome small dataset challenges (Yu et al., 2022). (Shishehbori and Awan, 2024) discussed the potential of machine learning and deep learning models including transfer learning for cardiovascular risk prediction, emphasizing the need for model interpretability and external validation. Despite its growing use, transfer learning approaches were often underexplored in the context of tabular biomedical data lacked consistent performance metrics across studies.

A hybrid CNN-LSTM model was explored by (Soltanizadeh and Naghibi, 2024), applied the model to diabetes prediction task and reporting good performance achieving 85% accuracy on small clinical dataset. However, the results confirmed the ability of hybrid deep learning in constrained data environments, the study did not explore image-based transformation of tabular data.

The application of neural networks for cardiovascular risk prediction in T2DM patients was reviewed by (Kee et al., 2023), reporting an AUC of 0.91 for several deep learning models. However, their study emphasized concerns related to generalizability, data imbalance and lack of external validation, particularly in studies with limited sample sizes. These challenges suggested a need for robust validation strategies and advanced model regularization techniques.

Limitations in transferability scoring methods were further discussed by (Chaves et al., 2023), that standard evaluation metrics do not adequately reflect performance in clinical applications. The need for context-aware validation frameworks was emphasized, particularly in resource-limited settings where model deployment recommended to align specific diagnostic thresholds and operational constraints.

Vision Transformer (ViT) developed by (Dosovitskiy et al., 2020), an attention-based deep learning architecture initially proposed for computer vision tasks, has recently been explored in biomedical contexts. (Hambarde et al., 2023) applied ViT to brain tumor and lung cancer classification, reporting high performance over traditional convolutional neural network (CNNs) with accuracy of value 85.71% and 89.02%, respectively. These results validate ViT's capacity to learn from small and complex datasets. However, Non-image clinical data was not addressed by the application, which was limited to imaging domains.

In sperate investigation, (Lee et al., 2022) applied deep learning to predict deep learning prediction for predicting cardiovascular mortality in hypertensive patients, demonstrating significant accuracy improvements over logistic regression baselines. However, they noted a lack of external validation and challenges in model interpretability were identified as major drawbacks, limiting the model practical utility in clinical settings.

A hybrid approach combine CNN with Support Vector Machine (SVM) was introduced by (Ahsan et al., 2022), where ReliefF is one of the feature selections based on distance (Tuncer et al., 2020) and data augmentation applied to the PIMA Indian Diabetes dataset. The resulting CNN-SVM model achieved classification accuracy (92.19%) and AUC of (0.96). while their work showed strong results, it was limited to binary diabetes classification and did not address biomarker prediction or image-based encoding.

From the existing literature, it is evident that conventional statistical models, standalone machine learning techniques and even deep learning trained on tabular data often face challenges in accurately predicting CRP levels, particularly in multi- class classification scenarios. Most studies either rely on binary classification frameworks or required large, clean datasets to achieved acceptable performance. Furthermore, limited generalizability, lack of external validation and insufficient integration of special feature representations remain critical issues. Therefore, in the present study, a hybrid approach is investigated by integrating tabular-to-image encoding techniques with advanced transfer learning architectures such as ResNet50 or Vision Transformer (ViT). By transforming structured clinical data into image-based formats and leveraging the feature extraction capabilities of pretrained vision models, the proposed framework aims to achieve performance and clinical applicability compared to existing methods.

3. Methodology

This study aimed to develop a predictive model for CRP levels in type 2 diabetic patient using deep learning and transfer learning techniques. A custom dataset was constructed from private clinical records, contain laboratory tests, biochemicals and lifestyle data overcome the limitation of existing datasets. The overall workflow of methodology is illustrated in Figure 1, providing an overview of sequential steps followed in this study.

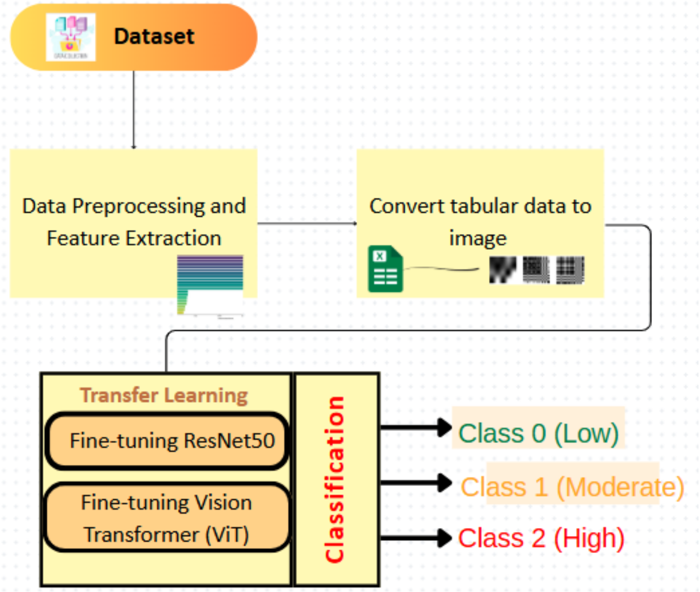

Figure 1. Overview of Proposed Method for CRP Levels Classification

3.1 Data Collection

Clinical and biochemical data were collected from 838 type 2 diabetic patients in the Kurdistan region. The dataset includes 22 features covering demographics, metabolic markers, renal and lipid profiles, lifestyle factors, and medical history as shown in Table 1 Data were anonymized and validated by specialists. Only type 2 diabetic patients were included, type 1 diabetic and pregnant women were excluded. Several publicly available datasets were considered but found unsuitable for the objectives of this study.

Table 1. The Clinical and Biochemical Feature of The Dataset

Category	Features
Demographic	Age, Gender
Metabolic Markers	FBG (Fasting Blood Glucose), HbA1C, BMI, BP (Blood Pressure), BS (Blood Sugar)
Renal Profile	Creatinine, Uric Acid
Lipid Profile	Cholesterol, Triglyceride, HDL, LDL
Urinalysis	Urine Protein
Medical History	Family History, Obesity, Sleep Apnea
Lifestyle and Behavior	Smoking, Alcohol Consumption, Physical Activity, Diet

The Pima Indians Diabetes dataset, although widely used in diabetes research, includes only eight general features (e.g., glucose, BMI, insulin) and lacks critical data such as CRP measurements, renal or lipid biomarkers and lifestyle factors. The MIMIC-IV (Medical Information Mart for Intensive Care) database is limited due to ethical and legal requirements, the database's focus on critically ill patients receiving intensive care does not accurately reflect the outpatient diabetic population that this study is intended to target. The Kaggle Diabetes Dataset by Tushar Potdar includes some relevant clinical features like HbA1c, HDL and LDL but its lacks demographic variable such as age, gender and has no pre-reviewed validation or clear documentation. Due to these limitations, none of these datasets aligned with the study's specific focus on CRP prediction in type 2 diabetic patients as shown in Table 2.

Table 2. Summary of Dataset Characteristics Relevant to CRP Prediction in Diabetic Patients

Dataset Name	Year Created / Published	No. of Features	CRP Levels	Salivary Biomarkers	Lifestyle and Behavior	Access Type
Custom CRP dataset (this study)	2024-2025	22	Classified (0-2)	Yes	Yes	Privat/ Institutional
Pima Indians diabetes	1990s (UCI Repository)	8	No	No	No	Public (open Access)
Kaggle diabetes (Tushar Potdar)	2021 (Kaggle)	34	No	Yes	Yes	Public (Kaggle Access)
MIMIC_IV	2020 (Physion et, MIT)	160+	Numeric (some patients)	No	No	Restricted (credentialed)

3.2 Data Preprocessing

Data preprocessing involved multiple steps. Missing values were handled by removal, encoding categorical variables such that binary and multiclass categorical variable were numerically encoded (e.g., male=0, female=1, smoking=1 and non-smoking =0) to facilitate deep learning training and target variables CRP levels were encoded as (0 for low, 1 for moderate and 2 for high) for classification tasks. Feature selection was performed using Recursive Feature Elimination (RFE) with Random Forest (RF)classifier that proposed by (Guyon et al., 2002), identifying the top 20 most important features as show in Figure 2 enhancing model accuracy and reducing overfitting risks, feature importance was calculated using Gini index and the least important features were iteratively removed until the top 20 remained. Data normalization was applied to all numerical features, scaling them to the range (0,1) for stable training. To prepare the data for deep learning, the tabular datasets was transformed into 2D image using three techniques such as Zero-Padded Grid, Recurrecet Plot and Gramian Angular Field methods. Each resulting image was resized to 224x224x3 pixel to match model input dimensions. In this study, a comprehensive data augmentation strategy was applied to the generated images to improve classification performance apply data augmentation such as (random rotation, horizontal flipping, color jitter, Gaussian blur and perspective distortion) to improve model generalization and prevent overfitting (Khalifa et al., 2022).

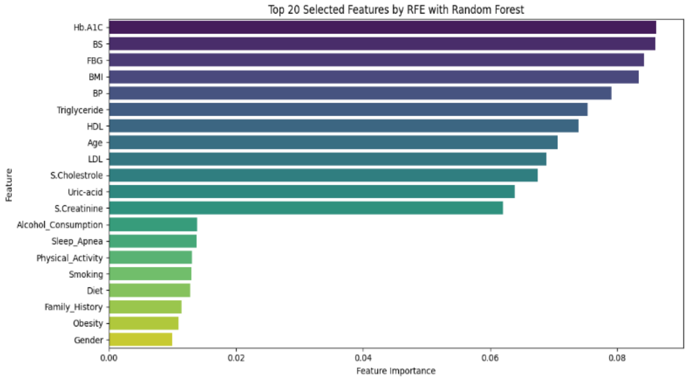

Figure 2. Random Forest (RF) RFE of Top 20 Feature Selection Algorithm

3.3 Data Splitting

The dataset was first split into training, validation and test sets using stratified split approach before applying image transformation or augmentation to prevent data leakage. A hold-out test set comprising 15% of the data was reserved for final performance evaluation the remaining 85% was used for training and validation. Stratified 5-fold cross-validation was used to evaluate the training samples for the ResNet50 model, guaranteeing that all data was used for training and validation in various rounds. This approach reduced overfitting and enabled accurate prediction by validating the model over multiple data subsets. However, for ViT model, a traditional train validation split was used on the 85% training subset. Particularly, applying cross-validation to ViT reduce performance due to its sensitivity to small-scale fold variations. As a result, a standard split approach was adopted for ViT to maintain higher accuracy and stable training dynamics.

3.4 Proposed Model and Transfer Learning Strategy

This section presents the comprehensive modeling framework developed for predicting C-reactive protein (CRP) levels in patients with type 2 diabetes mellitus using deep learning and transfer learning. the pipeline comprises data transformation, model architecture adaption, transfer learning, training strategies and evaluation procedures ensuring both methodological and reproducibility.

3.4.1 Data Transformation and Preparation

Given the tabular nature of clinical dataset, it was necessary to convert the structured numerical data into two-dimensional image representations to facilitate compatibility with convolutional neural network (CNNs) such as REsNet50 and Vision Transformer (ViT) architectures. Three transformation technique were applied, first the Zero-Padded Grid , where each patients feature vector was reshaped into a fixed-size square grid unoccupied cells filled with zeros, second the Recurrence Plot (RP), which visualizes pairwise recurrence within feature values highlighting structural similarities (Manjurul Ahsan and Siddique, 2021) and the last technique is Gramian Angular Field (GAF), which encodes temporal dependencies through angular summation (Altunkaya et al., 2023), capturing the holistic interrelationships among features as show in Figure 3, each resulting image was resized to 224x224 pixels with three channels, ensuring consistency with pre-trained model input requirements.

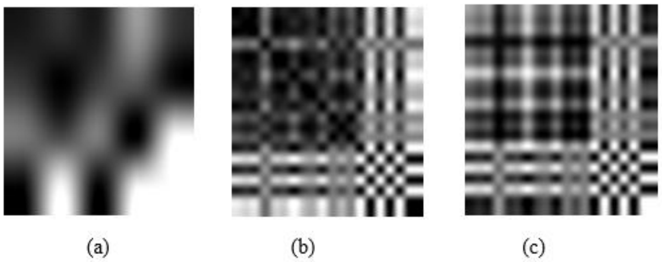

Figure 3. Image Representation of Tabular Data using Three Encoding Technique (a) Zero-Padded Grid, (b) Recurrence plot (RP) and (c) Gramian

Angular Field (GAF)

3.4.2 Transfer Learning Strategy

To leverage prior knowledge from large-scale image dataset such as (ImageNet) and to address the limited size of the biomedical dataset Transfer learning was employed. Specifically, lower-level layers of the pre-trained models, which capture general visual features were frozen during training, while higher-level layers were fine tuned to learn domain- specific representations. This hierarchical training strategy aimed to improve convergence and prevent overfitting. Additionally, the AdamW optimizer employed alongside dynamic learning rate scheduling to optimize model performance.

3.4.3 ResNet50 Architecture

The ResNet50 model, originally proposed by (He et al., 2016) was selected for its residual learning framework, which effectively mitigates vanishing gradient issues and enable the training deep networks as part of the residual network family. It leverages residual learning via shortcut connection, which allows the network to train very deep architectures without vanishing gradients. Per-trained on ImageNet, ResNet50 captures hierarchical spatial features such as edges, texture and patterns. In this study, after the complete dataset was divided into subset using stratified split 85% was allocated to training and validation and remaining 15% was reserves as hold-out test set to evaluate final model performance. The test set remained completely unseen throughout the training process to ensure unbiased evaluation. To enable the application of CNN, all tabular patient record were transformed into two-dimensional images using the encoding techniques as mention in 3.4.1 section. The images were saved in RGB format with size 224x224 and organized into class-specific folders based on CRP levels. To improve generalization and reduce overfitting, data augmentation was applied only to training images. Validation and test image were processed using only resize and normalization without any augmentation. The model architecture was based on a pre-trained ResNet50 model. Transfer learning was implemented by freezing the majority of convolutional layers, while selectively fine-tuning (layer3, layer4 and fully connected head) to adapt the model to the CRP classification task as show in Figure 4. The final classification to output three class corresponding to CRP levels (Low, Moderate, High). Training was performed using Adam optimizer (learning rate=0.0001) with early stopping to mitigate overfitting. To address class imbalance in the CRP level distribution, the model was trained using Focal Loss, which emphasizes harder-to-classify samples by dynamically scaling the loss contribution of each example. ResNet50-based models were trained for a maximum of 10 epochs with early stopping based on validation loss, using patience of 5 epochs. Additionally, a ReduceLROnPlateau scheduler was used to lower the learning rate when the validation performance plateaued. After completing the five-fold cross-validation, the model from the fold with highest validation accuracy was selected as the final model. This best-performing model was then evaluated on the independent 15 % test set. The final evaluation included overall accuracy, and visualization of training and validation loss and accuracy curve across epochs.

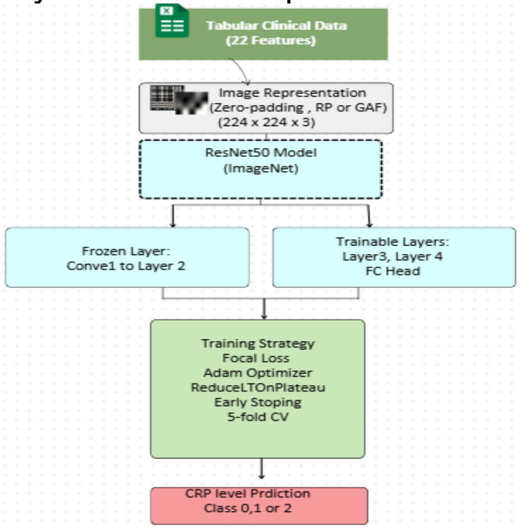

Figure 4. (Zero-Padded Grid, Recurrence Plot (RP) or Gramian Angular Field (GAF) with ResNet50 Workflow

3.4.4 Vision Transformer (ViT) Architecture

The Vision Transformer (ViT), as introduced by (Dosovitskiy et al., 2020) was adopted for its ability to model global dependencies within image representation via self-attention mechanisms, which are particularly advantageous in capturing complex feature interactions. The dataset was preprocessed and formatted for use the ViT model and complete dataset was divided into subset split 85% for training and validation and remaining 15% held-out test set to evaluate final model performance. The ViT model splits images into fixed size patches and encoding their positions, ViT treats image analysis as a sequence modeling problem similar to natural language processing. Pretrained on large datasets such as ImageNet-21k, ViT capture global dependencies more effectively than CNN. Vision Transformer model is robust with large datasets but is advantageous for small dataset when large-scale fine-tuning or global context modeling is essential. In this study ViT-based models, which required more than 10 epochs to converge, set a maximum of 30 epochs and applied early stopping with patience, the pre-trained ViT model was customized by replacing the original classification head with a dropout layer (dropout rate = 0.4) and fully connected layer tailored for three class CRP level prediction. Unlike the ResNet50 configuration, as shown in Figure 5 all layers of the ViT were unfrozen to enable comprehensive fine-tuning on encoded image data. The training employed the AdamW optimizer (learning rate = 1e-4) with weight decay of 0.01, accompanied by learning rate reduction upon validation loss plateauing. Early stopping was implemented after five epochs without improvement in validation performance. This allowed the ViT models sufficient capacity to stabilize, as reflected in their improved performance on the validation and test sets.

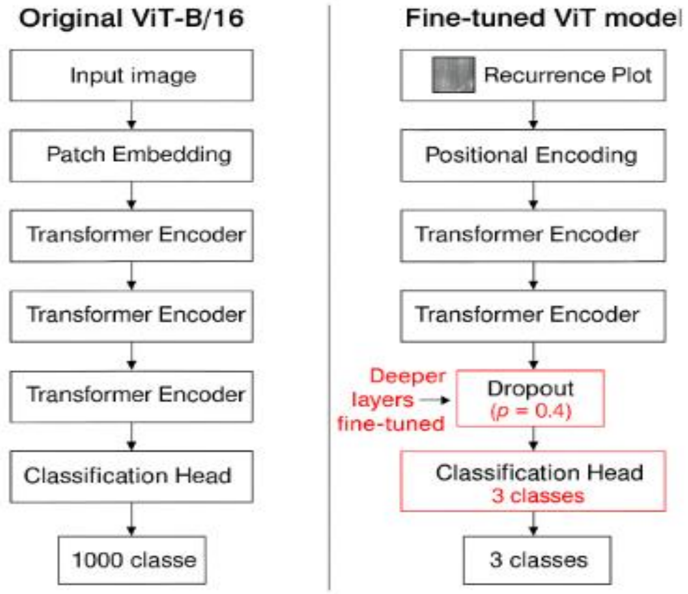

Figure 5. The Fine-Tuning ViT Model

4. Experimental Results

The subsection summarized the predictive performance of six model configurations was evaluated for predicting C-reactive protein (CRP) levels in patients with type 2 diabetes mellitus. Each configuration consists of a tabular-to-image transformation method (Zero-Padded Grid, Recurrence Plot or Gramian Angular Field) combine with two-model backbone architecture ResNet50 or Vision Transformer (ViT). Evaluation metrics such as accuracy, precision, recall, F1-score and AUC were used to assess performance across all CRP classes. Overall, ViT-based models consistently outperformed ResNet50-based models. Notably, the GAF-ViT model achieved the highest classification accuracy of 92.06% on test set, with strong AUC scores across all classes (97% - 98%), confirming its robustness and generalizability. The performance of all model is provided with further details in the following subsections. demonstrated best result across most encodings.

4.1 ResNet50 Based Models

4.1.1 Zero-Padded Grid with ResNet50

The Zero-Padded Grid with ResNet50 configuration demonstrated a mean validation accuracy of 83.43% as shown in Table 3 and achieved a final test accuracy of 87.30%. The training accuracy increased steadily as shown in Figure 6 from 51.58% to 92.11%, showing effective learning. Validation accuracy followed the training increasing starting low rising to peak of 85.92%, which indicate that the model generalized well. The close alignment between the training and validation curves in the later epochs suggests good convergence with minimal overfitting. The training loss decreased consistently across epochs, reflecting improved model confidence. The validation loss initially remains stable during the first epoch (0.4529) then dropped sharply from epoch 4 onward, eventually stabilizing near (0.1131). this pattern suggest that the model required several epochs to generalized well, after which it maintains low error in unseen data, supporting the overall robustness of the training process as show in Figure 7. The model exhibited strong discriminative ability across all CRP classes as evidenced by AUC score of 0.95, 0.93 and 0.99 for class 0, 1 and 2 respectively as shown in Figure 8.

Table 3. Cross-Validation Performance Metrics of Zero-Padded Grid with ResNet50 Model for CRP Classification

Fold	Training Accuracy	Training Loss	Validation Accuracy	Validation Loss
1	87.35	0.1070	83.92	0.116
2	91.21	0.0769	82.52	0.1350
3	89.12	0.0936	78.87	0.1655
4	92.11	0.0563	85.92	0.1131
5	91.75	0.0688	85.92	0.1309
mean ± std	90.31 ± 2.02	0.08 ± 0.02	83.43 ± 2.93	0.13 ± 0.02

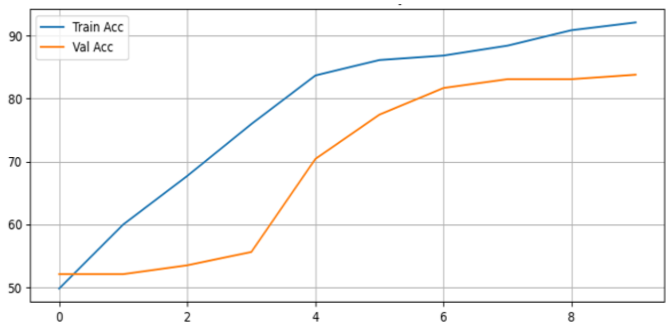

Figure 6. Training and Validation Accuracy Across Epoch for Fold-4 Using ZP-ResNet50

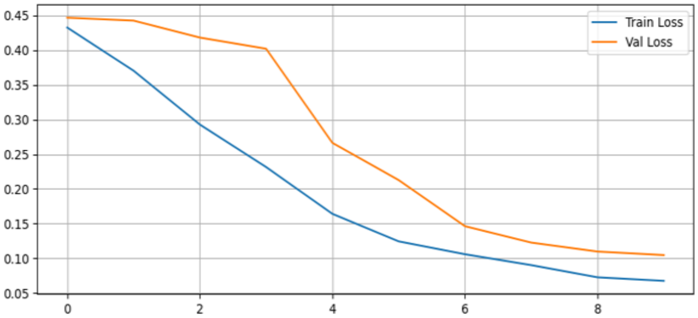

Figure 7. Training and Validation Loss Across Epoch for fold-4 ZP-ResNet50

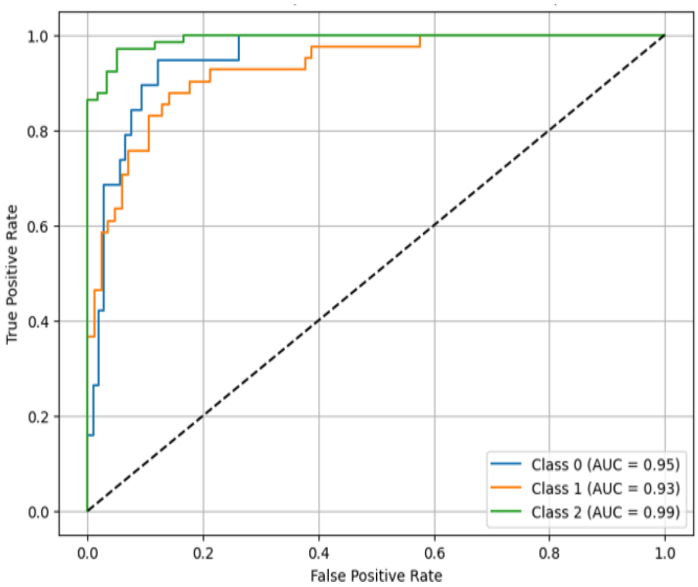

Figure 8. The ROC Curve of ZP with ResNet50 for CRP level Prediction

4.1.2 Recurrence Plot (RP) with ResNet50

The Recurrence Plot (RP) with ResNet50 model achieved the mean validation accuracy of 86.39% and a test accuracy of 90.48% as shown in Table 4. The stratified 5-fold cross-validation on fold 3 achieved the highest validation accuracy (88.05% at epoch 8 out of 10) and lowest validation loss (0.1124), indicate superior performance compared to fold 5 (85.05% accuracy at epoch 9 out of 10 and 0.12 loss) as shown in Figure 9 and 10. However, fold 5 display smooth learning curves, narrower training-validation gaps and more stable convergence, suggesting better generalization. These findings emphasize the need to consider both peak performance and training dynamic when interpreting cross-validation results. Notably, class 2 demonstrated the highest predictive performance, with a precision of 0.93, recall of 0.94 and F1-score of 0.93, while class 1 remained more challenging F1-score of 0.85, The class 0 with AUC 0.84, class 1 with AUC 0.81 and class 2 with AUC 0.92. The curve for class 2 remains strong showing high discriminative ability. However, the curve for class 0 and class 1 exhibit more gradual slopes, indicate less precise class separation compared to ZP variants shown in Figure 11.

Table 4 Cross Validation Performance Metrics of Recurrent Plot (RP) with ResNet50 Model for CRP Classification

Fold	Training Accuracy	Training Loss	Validation Accuracy	Validation Loss
1	87.48	0.0972	89.11	0.1042
2	88.53	0.0962	86.22	0.1179
3	89.47	0.0821	86.85	0.1124
4	85.69	0.1264	85.49	0.1650
5	89.05	0.0892	84.27	0.1148
mean± std	88.04 ± 1.51	0.10 ± 0.02	86.39 ± 1.80	0.12 ± 0.02

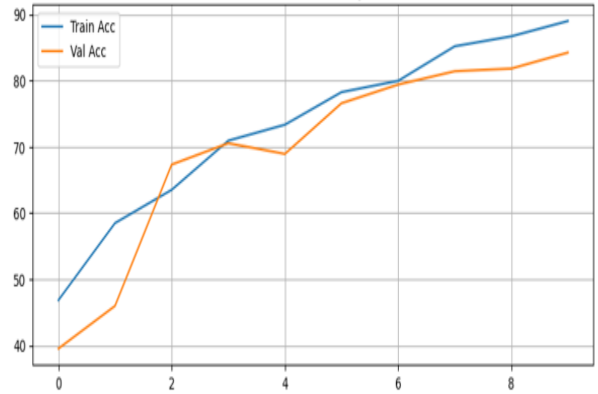

Figure 9. Training and Validation Accuracy Across Epoch for Fold-5 RP-ResNet50

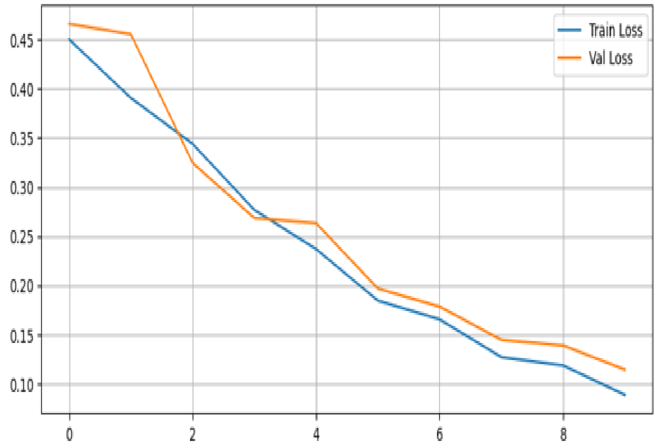

Figure 10. Training and Validation Loss Across Epoch for Fold -5 RP-ResNet50

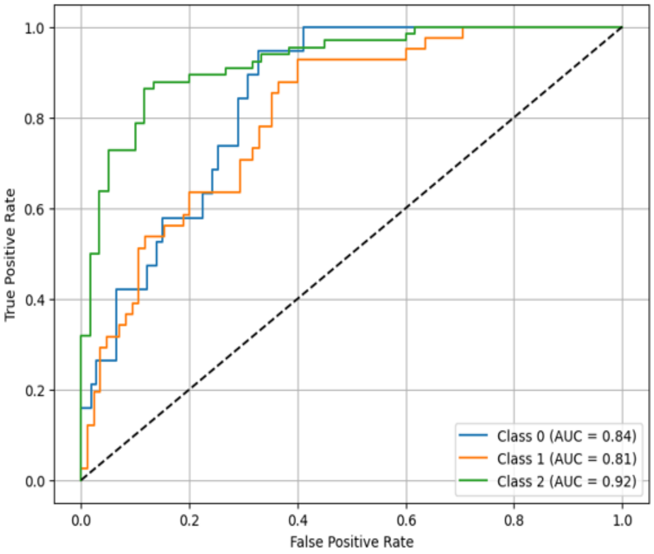

Figure 11. The ROC Curve of RP with ResNet50 for CRP Level Prediction

4.1.3 The Gramian Angular Field (GAF) with ResNet50

The GAF with ResNet50 model as shown is Table 5 was produced a mean validation accuracy of 82,17% and test accuracy of 84.92%. Figure 12 shows consistent improvements in training and validation accuracy reach over 86% and 87%, respectively, with minimal performance gaps suggesting effective learning. Figure 13 show that training and validation loss steadily declining to approximately 0.12 and closely aligning, indicating model stability and convergence. The small gap between training and validation metrics through training suggests no significant overfitting or underfitting, affirming the model’s robustness and generalizability. The model achieved excellent AUC scores (class 0 = 97%, class 1 = 96% and class 2 = 98%), reflecting its capacity to effectively separate CRP levels, particularly for class 2 as shown in Figure 14.

Table 5. Cross Validation Performance Metrics of GAF with ResNet50 Model for CRP Classification

Fold	Training Accuracy	Training Loss	Validation Accuracy	Validation Loss
1	86.47	0.1162	83.92	0.1170
2	85.06	0.1165	75.52	0.2007
3	89.30	0.0925	78.87	0.1665
4	90.35	0.0940	86.62	0.1417
5	88.07	0.0916	85.92	0.1405
mean ± std	87.85± 2.13	0.10 ± 0.01	82.17 ± 4.80	0.15 ± 0.03

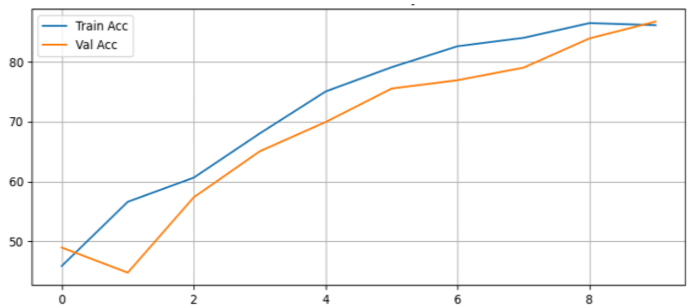

Figure 12. Training and Validation Accuracy Across Epoch for Fold-1 GAF-ResNet50

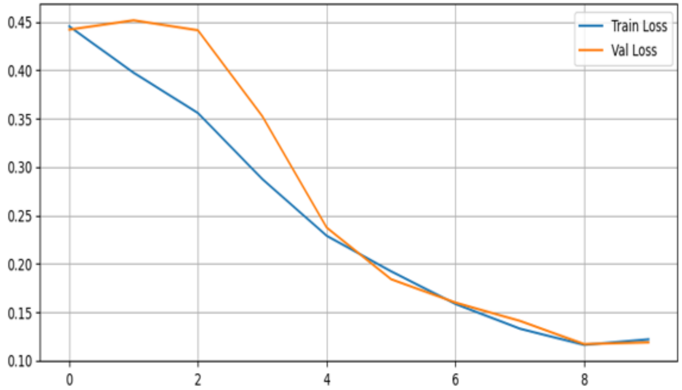

Figure 13. Training and Validation Loss Across Epoch for Fold -1 GAF-ResNet50

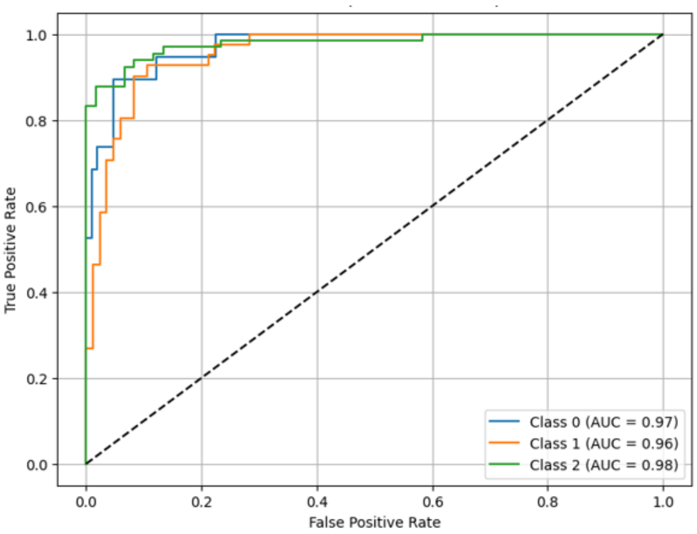

Figure 14. The ROC Curve of GAF with ResNet50 for CRP Level Prediction

4.2 Vision Transformer-Based Models

4.2.1 Zero-Padded Grid with ViT

The Zero-Padded Grid with ViT model was trained for a maximum of 30 epochs with early stopping triggered after epoch 29. The training and validation performance over time is visualized in Figure 15 and 16. The model exhibited steady learning progression as show in Figure 15 the final training accuracy is 92.02% and validation accuracy is 97.90% and the Figure 16 Illustrate the final training and validation loss which is 0.0791 and 0.0436. the best performing model checkpoint based on validation loss was evaluated on the held-out test set. The model achieved a final test accuracy of 97.62% and a test loss of 0.0878 which confirming its robust performance on unseen data. The model achieved perfect classification for class 2 (precision of 0.97, recall of 1.00 and F1-score of 0.99) and as show in Figure 17 the model demonstrated outstanding discriminative capacity AUC (class 0 = 99%, class 1 = 99 % and class 2 = 100%).

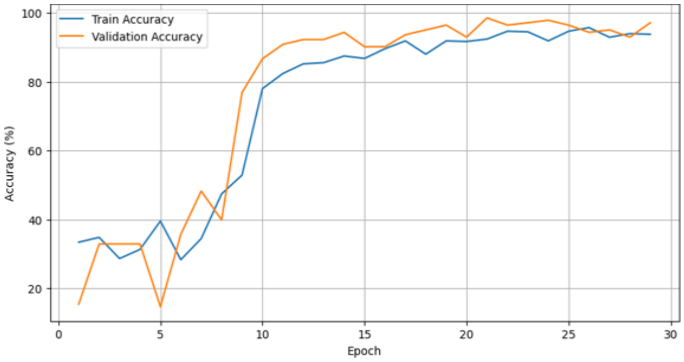

Figure 15. Training and Validation Accuracy Per Epochs ZP-ViT

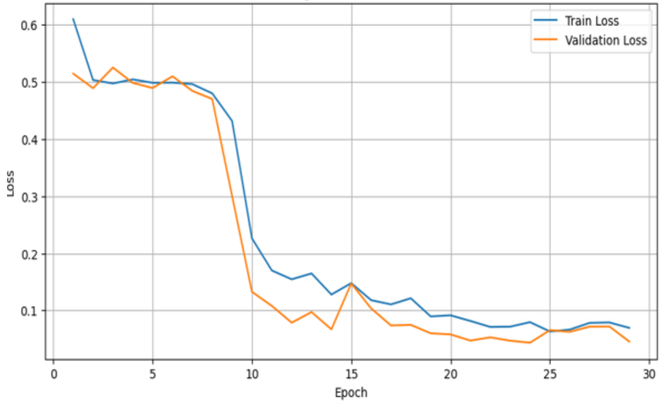

Figure 16. Training and Validation Loss Per Epochs ZP-ViT

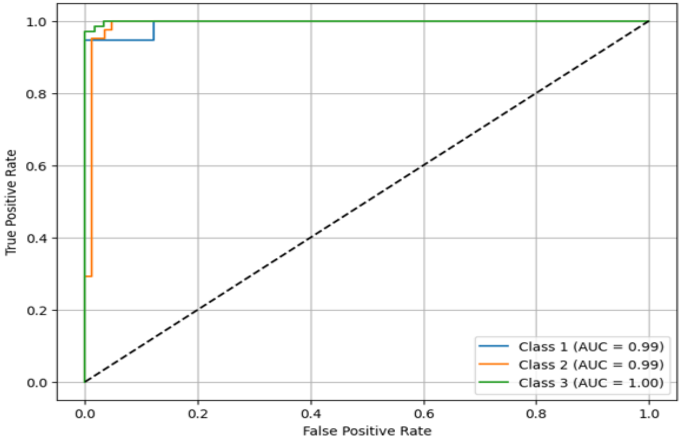

Figure 17. The ROC Curve of Zero-Padded Grid with ViT for CRP Level Prediction

4.2.2 Recurrence Plot (RP) Grid with ViT

The RP with ViT model attained a test accuracy of 97.62% with balanced class performance. The training and validation accuracy and loss curves over 30 epochs are shown in Figure 18 and 19, respectively. Early epochs exhibited fluctuating accuracy and loss indicating initial instability. However, as retraining progressed validation accuracy improved substantially with convergence observed after epoch15. The final training accuracy reached 91.74% and the corresponding validation accuracy was 93.01% indicating good generalization without signs of overfitting and a macro F1-score of 0.90. Despite minor class confusion, particularly in class 1, the model's ROC-AUC score high (≥ 0.98) as show in Figure 20.

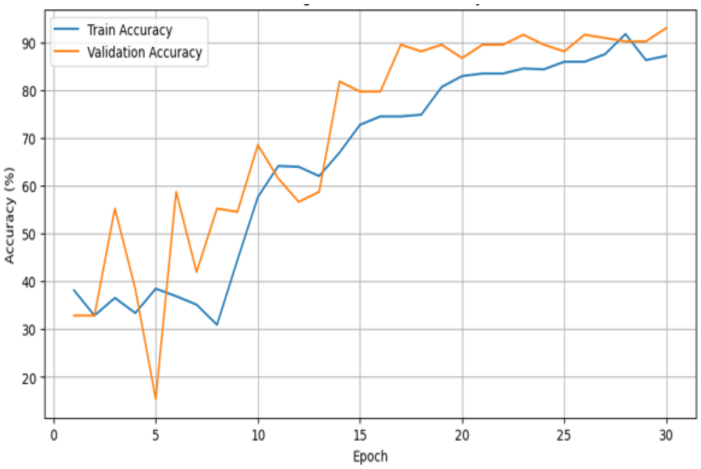

Figure 18. Training and Validation Accuracy Per Epochs RP-ViT

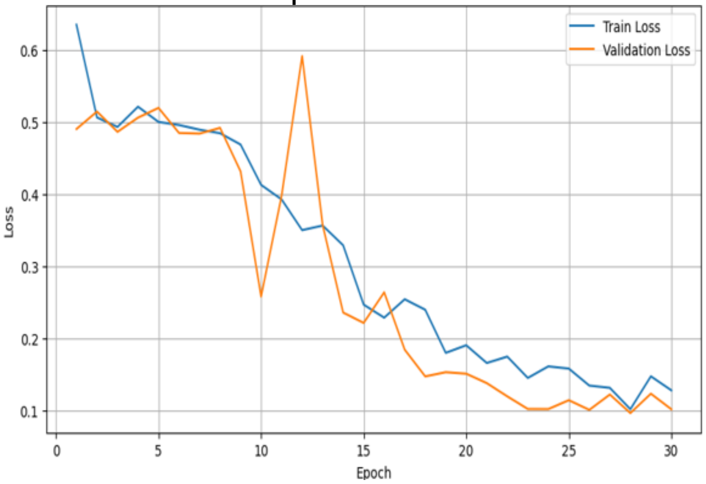

Figure 19. Training and Validation Loss Per Epochs RP-ViT

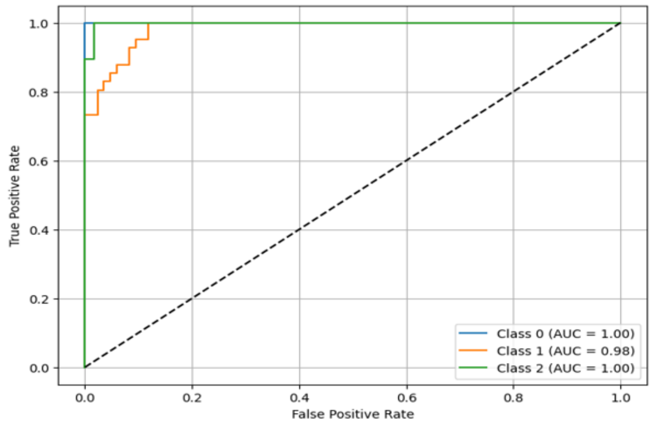

Figure 20. The ROC Curve of RP with ViT for CRP Level Prediction

4.2.3 The Gramian Angular Field (GAF) with ViT

The ViT model trained on GAF-encoded features achieved consistent convergence after initial fluctuations with training and validation accuracy stabilizing and reaching 89.28% and 92.31% by epoch 20 as shown in Figure 21, respectively. Validation loss steadily decreased as show in Figure 22 from 0.4877 to 0.0751, indicating effective learning. The model at epoch 26 (95.10% validation accuracy) was selected for final evaluation, highlighting its robust generalization without overfitting. The GAF with ViT models also performed well, achieving a test accuracy of 92.06% and the Figure 23 shows that the model demonstrating strong discriminative ability across all classes AUC (class 0 = 97%, class 1 = 96% and class 2 = 98%).

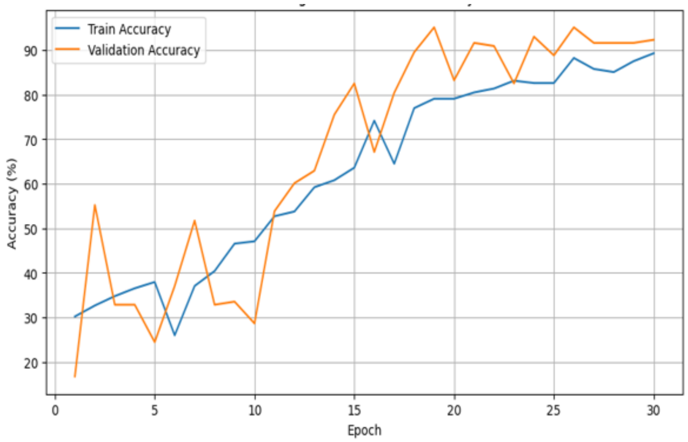

Figure 21. Training and Validation Accuracy Per Epochs GAF-ViT

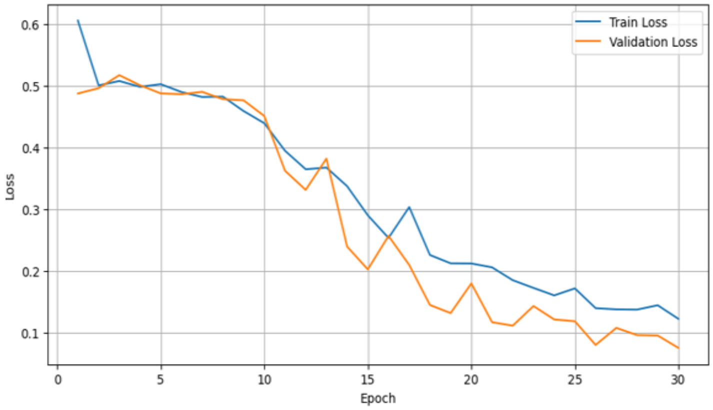

Figure 22. Training and Validation Loss Per Epochs GAF-ViT

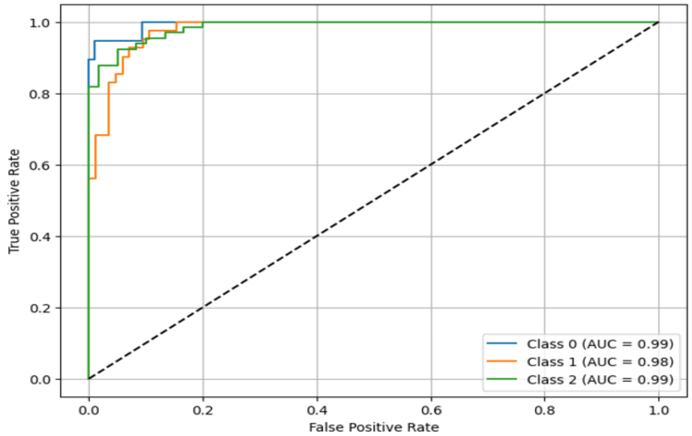

Figure 23. ROC Curve and AUC for Multi-Class CRP Levels Prediction

5. Discussion

This study demonstrates that the integration of tabular-to-image transformation with transfer-learning-based deep learning architectures offers significant promise for the prediction of CRP levels in type 2 diabetic patients, the findings consistently show that the ViT model outperforms the ResNet50 models across all image encoding strategies. A comparative analysis across all models revealed that ViT-based architectures consistently outperformed their ResNet50 counterparts as shown in Figure 24 and Table 6, A comparative summary table was constructed to visualize and quantify each model's performance across four key matrices precision, recall, F1-score and average area under the ROC curve (AUC). These indicators provided a comprehensive assessment of each of the three CRP classes classification ability. Best overall performance was the ZP encoding image technique with ViT model, which achieved the highest score across all metrics, with precision and recall both reaching 0.95 %, an F1-score of 0.96% and average score of approximately 0.97. this suggests that the combination of spatially structured grid encoding and transformer-based modeling is highly effective in capturing the underlying CRP-related patterns in clinical data. Strong alternative was the GAF encoding image technique with ViT model also give excellent results, with an average AUC of 0.98, slightly better that ZP with ViT model , meanwhile the RP image encoding with ViT showed competitive average AUC (0.99) but lower F1-score, likely due to reduce sensitivity in minority classes and the lowest performing model was the RP ending image with ResNet50 exhibited the weakest performance with the lowest precision (0.76) and recall (0.81), indicating a less effective ability to generalize from the recurrent-base encoding using CNN-based backbone. These results highlighted the impact of both encoding techniques and model architectures on classification performance. The observed improvement may be partially related to the ViT's self-attention mechanisms, which enables it to focus on structured and informative regions with GAF-transformed input. As shown in Figure 25, the attention heatmap reveals selective attention across multiple key patches, suggesting that the model identifies discriminative area relevant to CRP level. Among all methods the encoding methods, the Zero-Padded Grid transformation emerged as the most effective, likely due to its preservation of feature order and relative magnitude, facilitating more structured learning by the models. The consistently high AUC and F1-scores for class 2 across all models suggest that the models were particularly adept at identify patients with elevated CRP levels, a clinically significant finding given the association of high CRP with cardiovascular risk. However, all models exhibited relatively lower performance on class 1, indicating that moderate CRP levels are more challenging to classify due to potential overlaps in feature patterns with neighboring classes. This highlights the need for additional feature engineering, incorporation of domain knowledge or exploration of ensemble strategies to enhance model sensitivity for this class. The application of focal loss and data augmentation technique effectively mitigated class imbalance and improved model generalization as evidenced by the stable training dynamic and consistent performance across validation folds. These result underscore the importance of addressing imbalanced datasets in clinical predictive modeling. In summary, the study validates the feasibility of using deep learning models, particularly Vision Transformers in conjunction with innovative data encoding strategies for CRP level prediction. As Table 7 illustrate the comparative analysis of the performance metrics such as accuracy, F1-score and average AUC from three different studies. This study achieved an accuracy of 92.31%, an F1-score of 0.95 and an average AUC of 0.98, indicating strong predictive performance for CRP levels on proposed dataset.

Figure 24. Comparative Performance of Image Encoding Technique and Transfer Learning Models

Table 6. Performance Comparison of All Models Variants

Model Variant	Accuracy	F1-score	AUC
Zero-Padded+ ResNet50	92.11	0.95	0.95
RP + ResNet50	89.47	0.93	0.86
GAF+ ResNet50	86.47	0.90	0.97
Zero-Padded + ViT	92.02	0.99	0.99
RP + ViT	91.74	0.94	0.99
GAF+ ViT	92.31	0.95	0.98

Table 7. Comparative Performance Evaluation with Other Works

Research	Methodology	Accuracy	F1-score	AUC
(Hambarde et al., 2023)	Dataset: two medical image datasets Method: ViT models for classification	89.02	0.83	-
(Aslan and Sabanci, 2023)	Dataset: PIMA Indian Diabetes Method: ReliefF-selected CNN-SVM	92.19	0.94	0.96
(Mo et al., 2025)	Dataset: medical record dataset. Method:DNN model	0.75	0.64	0.82
This study	Dataset: clinical dataset Method: GAF encoding image with ViT model	92.31	0.95	0.98

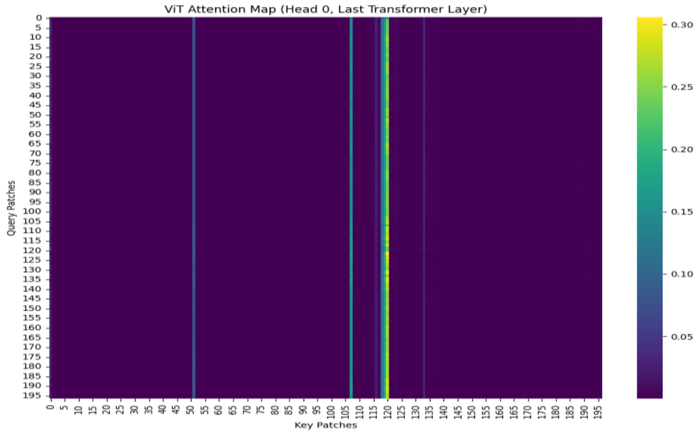

Figure 25. Attention Heatmap for Vision Transformer Model in this study

6.Conclusion

Accurate Classification of CRP levels in patients with type 2 diabetes mellitus is clinically important, as evaluated CRP is a key biomarker of systemic inflammation and is associated with increased risk of cardiovascular complications, poor glycemic control and disease progression. The insights of this research pave the way for integration of advanced AI Frameworks into clinical practices, potentially transforming the monitoring and management of inflammation in diabetic patients. This research demonstrates the efficacy of combining tabular-to-image transformation with transfer learning Architectures for CRP level prediction in type 2 diabetic patients. The findings reveal that Vision Transformer models, particularly when paired with structured image representation like Zero-Padded Grid, achieved best classification performance compared with traditional convolution networks. The result highlighted the potential of these models to capture complex feature relationships inherent in clinical data as evidenced by consistently high accuracy, F1-score and AUC values. However, challenging remain in classifying moderate-risk CRP levels indicating a need for refined feature engineering and model calibration. Future research recommended to prioritize external validation, incorporation of additional clinical features and exploration of ensemble methods to further enhance predictive performance and clinical applicability.

Acknowledgements

At the end of this work the author would like to take this opportunity for expression of gratitude for whom their efforts were very helpful to get this results, to whom help me directly academically and scientifically, the college of engineering-Software \department in Salaheddin university, and also a very thankful for Dr Aziz rawanduzy as specialist physician for his efforts is thankful for medical data collection and scientific supporting in his field that make the result more complete and integrated.

References

Ahsan, M. M., Luna, S. A. & Siddique, Z. Machine-learning-based disease diagnosis: A comprehensive review. Healthcare, 2022. MDPI, 541.

Altunkaya, D., Okay, F. Y. & Ozdemir, S. 2023. Image transformation for IoT time-series data: A review. *arXiv preprint arXiv:2311.12742*.

Amin, R., Al Ghamdi, M. A., Almotiri, S. H. & Alruily, M. 2021. Healthcare techniques through deep learning: issues, challenges and opportunities. *IEEE Access,* 9**,** 98523-98541.

Anwer, M. A., Qattan, G. A. & Ali, A. M. 2024. Ocular disease classification using different kinds of machine learning algorithms. *Zanco Journal of Pure and Applied Sciences,* 36**,** 25-34.

Aslan, M. F. & Sabanci, K. 2023. A novel proposal for deep learning-based diabetes prediction: converting clinical data to image data. *Diagnostics,* 13**,** 796.

Azeez, J. Q. & Gallaly, D. Q. 2023. Association of Adiponectin With biochemical parameters in patients with chronic kidney diseases on Hemodialysis process in Erbil city. *Zanco Journal of Pure and Applied Sciences,* 35**,** 189-199.

Chaves, L., Bissoto, A., Valle, E. & Avila, S. The performance of transferability metrics does not translate to medical tasks. MICCAI Workshop on Domain Adaptation and Representation Transfer, 2023. Springer, 105-114.

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G. & Gelly, S. 2020. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*.

El Massari, H., Sabouri, Z., Mhammedi, S. & Gherabi, N. 2022. Diabetes prediction using machine learning algorithms and ontology. *Journal of ICT Standardization,* 10**,** 319-337.

Elsayed, N. A., Aleppo, G., Aroda, V. R., Bannuru, R. R., Brown, F. M., Bruemmer, D., Collins, B. S., Gaglia, J. L., Hilliard, M. E. & Isaacs, D. 2023. 2. Classification and diagnosis of diabetes: standards of care in diabetes—2023. *Diabetes care,* 46**,** S19-S40.

Fregoso-Aparicio, L., Noguez, J., Montesinos, L. & García-García, J. A. 2021. Machine learning and deep learning predictive models for type 2 diabetes: a systematic review. *Diabetology & metabolic syndrome,* 13**,** 148.

Guyon, I., Weston, J., Barnhill, S. & Vapnik, V. 2002. Gene selection for cancer classification using support vector machines. *Machine learning,* 46**,** 389-422.

Hambarde, B., Hussain, M. Z., Hussain, M. S., Gupta, S., Tomar, R. & Gupta, S. From Pixels to Predictions: Medical Image Analysis and Prediction of Brain Tumor and Chest Cancer with Vision Transformers. International Conference on Advances and Applications of Artificial Intelligence and Machine Learning, 2023. Springer, 371-386.

He, K., Zhang, X., Ren, S. & Sun, J. Deep residual learning for image recognition. Proceedings of the IEEE conference on computer vision and pattern recognition, 2016. 770-778.

Huang, L.-Y., Chen, F.-Y., Jhou, M.-J., Kuo, C.-H., Wu, C.-Z., Lu, C.-H., Chen, Y.-L., Pei, D., Cheng, Y.-F. & Lu, C.-J. 2022. Comparing Multiple Linear Regression and Machine Learning in Predicting Diabetic Urine Albumin–Creatinine Ratio in a 4-Year Follow-Up Study. *Journal of Clinical Medicine,* 11**,** 3661.

Islam, K. R., Prithula, J., Kumar, J., Tan, T. L., Reaz, M. B. I., Sumon, M. S. I. & Chowdhury, M. E. 2023. Machine learning-based early prediction of sepsis using electronic health records: a systematic review. *Journal of clinical medicine,* 12**,** 5658.

Johns, I., Moschonas, K. E., Medina, J., Ossei-Gerning, N., Kassianos, G. & Halcox, J. P. 2018. Risk classification in primary prevention of CVD according to QRISK2 and JBS3 'heart age', and prevalence of elevated high-sensitivity C reactive protein in the UK cohort of the EURIKA study. *Open Heart,* 5.

Kee, O. T., Harun, H., Mustafa, N., Abdul Murad, N. A., Chin, S. F., Jaafar, R. & Abdullah, N. 2023. Cardiovascular complications in a diabetes prediction model using machine learning: a systematic review. *Cardiovascular Diabetology,* 22**,** 13.

Khalifa, N. E., Loey, M. & Mirjalili, S. 2022. A comprehensive survey of recent trends in deep learning for digital images augmentation. *Artificial Intelligence Review,* 55**,** 2351-2377.

Kuppa, A., Tripathi, H., Al-Darraji, A., Tarhuni, W. M. & Abdel-Latif, A. 2023. C-reactive protein levels and risk of cardiovascular diseases: a two-sample bidirectional Mendelian randomization study. *International Journal of Molecular Sciences,* 24**,** 9129.

Lee, S.-J., Lee, S.-H., Choi, H.-I., Lee, J.-Y., Jeong, Y.-W., Kang, D.-R. & Sung, K.-C. 2022. Deep learning improves prediction of Cardiovascular Disease-related mortality and admission in patients with hypertension: analysis of the Korean National Health Information Database. *Journal of Clinical Medicine,* 11**,** 6677.

Lee, Y., Mckechnie, T., Doumouras, A. G., Handler, C., Eskicioglu, C., Gmora, S., Anvari, M. & Hong, D. 2019. Diagnostic value of C-reactive protein levels in postoperative infectious complications after bariatric surgery: a systematic review and meta-analysis. *Obesity surgery,* 29**,** 2022-2029.

Luthra, S., Orlandi, M., Hussain, S. B., Leira, Y., Botelho, J., Machado, V., Mendes, J. J., Marletta, D., Harden, S. & D'aiuto, F. 2023. Treatment of periodontitis and C-reactive protein: a systematic review and meta-analysis of randomized clinical trials. *Journal of Clinical Periodontology,* 50**,** 45-60.

Majeed, H. J., Ibrahim, G. I., Yousif, P. A. & Abdulkareem, S. M. 2021. Association between some serum oxidative stress biomarkers and lipid profile in type 2 diabetic patients in Erbil City. *Zanco Journal of Pure and Applied Sciences,* 33**,** 107-112.

Manjurul Ahsan, M. & Siddique, Z. 2021. Machine learning based disease diagnosis: A comprehensive review. *arXiv e-prints***,** arXiv: 2112.15538.

Mo, D., Xiong, S., Ji, T., Zhou, Q. & Zheng, Q. 2025. Predicting abnormal C-reactive protein level for improving utilization by deep neural network model. *International Journal of Medical Informatics,* 195**,** 105726.

Mouliou, D. S. 2023. C-reactive protein: pathophysiology, diagnosis, false test results and a novel diagnostic algorithm for clinicians. *Diseases,* 11**,** 132.

Pradhan, A. D., Manson, J. E., Rifai, N., Buring, J. E. & Ridker, P. M. 2001. C-reactive protein, interleukin 6, and risk of developing type 2 diabetes mellitus. *jama,* 286**,** 327-334.

Ridker, P. M., Hennekens, C. H., Buring, J. E. & Rifai, N. 2000. C-reactive protein and other markers of inflammation in the prediction of cardiovascular disease in women. *New England journal of medicine,* 342**,** 836-843.

Shishehbori, F. & Awan, Z. 2024. Enhancing cardiovascular disease risk prediction with machine learning models. *arXiv preprint arXiv:2401.17328*.

Soltanizadeh, S. & Naghibi, S. S. 2024. Hybrid CNN-LSTM for predicting diabetes: a review. *Current diabetes reviews,* 20**,** 77-84.

Tuncer, T., Dogan, S. & Ozyurt, F. 2020. An automated Residual Exemplar Local Binary Pattern and iterative ReliefF based COVID-19 detection method using chest X-ray image. *Chemometrics and Intelligent Laboratory Systems,* 203**,** 104054.

Wang, Z. & Oates, T. Encoding time series as images for visual inspection and classification using tiled convolutional neural networks. Workshops at the twenty-ninth AAAI conference on artificial intelligence, 2015. Austin, TX, 1-7.

Yang, T., Qi, F., Guo, F., Shao, M., Song, Y., Ren, G., Linlin, Z., Qin, G. & Zhao, Y. 2024. An update on chronic complications of diabetes mellitus: from molecular mechanisms to therapeutic strategies with a focus on metabolic memory. *Molecular Medicine,* 30**,** 71.

Yu, F., Xiu, X. & Li, Y. 2022. A survey on deep transfer learning and beyond. *Mathematics,* 10**,** 3619.