Journal of Applied Data Sciences
Vol. 7, No. 3, September 2026, pp. 1804-1818

Diffusion2D and Anchored Inference for Asymptotic Stabilization of Diffusion-Convolutional Neural Networks in Multidomain Medical Image Classification

Hanna Willa Dhany[1], Sutarman[2,*], Poltak Sihombing[3], Mohammad Andri Budiman[4]

[1]*Doctoral Program in Computer Science and Information, Universitas Sumatera Utara, Indonesia*

[2]*Department of Mathematics and Natural Sciences, Universitas Sumatera Utara, Indonesia*

[3,4]*Department of Computer Science, Universitas Sumatera Utara, Indonesia*

(Received: March 10, 2026; Revised: May 5, 2026; Accepted: June 15, 2026; Available online: June 28, 2026)

Abstract

Medical image classification across heterogeneous domains remains challenging due to domain shift, spatial variability, and unstable inference behavior. This study proposes a diffusion-stabilized Diffusion-Convolutional Neural Network (DCNN) framework that integrates Diffusion2D and post-hoc Anchored Diffusion to improve inference stability, probabilistic consistency, and robustness in multidomain medical image classification. The main contribution of this work is the introduction of a two-stage stabilization mechanism in which Diffusion2D performs controlled intra-image diffusion on feature representations before graph construction, while Anchored Diffusion refines uncertain predictions in the logit space through a k-nearest neighbors graph without retraining. The framework was evaluated on heterogeneous medical imaging datasets consisting of brain MRI, leukemia microscopy, and COVID-19 chest radiographs. Experimental results show that the proposed approach maintained baseline classification performance with an accuracy of 64.70% while improving the Macro-F1 score from 0.7045 to 0.7061. The diffusion mechanism reduced the average Laplacian value from 0.864355 to 0.187525, corresponding to a 78.23% reduction in spatial gradient variability. Internal analysis further demonstrated stable diffusion coefficients with a mean value of 0.141734 and a standard deviation of 0.003757, indicating controlled diffusion behavior. Anchored Diffusion selectively refined uncertain predictions, affecting only 0.6% of evaluated samples while preserving overall decision consistency. Repeated inference experiments across 40 iterations also revealed highly stable confidence trajectories with no observable variance after diffusion stabilization. The novelty of this research lies in combining feature-level diffusion stabilization, post-hoc anchored inference, and asymptotic regularization within a unified DCNN framework, providing a theoretically grounded and uncertainty-aware approach for robust multidomain medical image classification.

Keywords: Diffusion-Convolutional Neural Networks, Medical Image Classification, Diffusion2D, Anchored Diffusion, Inference Stability, Probabilistic Calibration, Deep Learning, Multidomain Medical Imaging

---

1. Introduction

The rapid advancement of Artificial Intelligence (AI) has entered a transformative stage, characterized by the increasing ability of computational systems to autonomously learn complex data representations. This transformation has been largely driven by deep learning, which has emerged as a dominant paradigm due to its effectiveness in extracting hierarchical features from large-scale datasets. By enabling end-to-end representation learning, deep learning models have substantially advanced a wide range of applications, particularly in image and signal processing.

Among various deep learning architectures, Convolutional Neural Networks (CNNs) have played a pivotal role in this progress. Their success is primarily attributed to the exploitation of spatial locality and weight sharing, which makes them highly effective for data with regular Euclidean structures. As a result, CNNs have become the de facto standard for natural image analysis and structured signal processing. However, the performance of CNNs is inherently constrained by their reliance on assumptions of spatial regularity and stationarity. When these assumptions are violated, CNNs often fail to generalize effectively. This limitation becomes especially evident when CNNs are applied to non-

---

***Corresponding author: Sutarman (sutarman@usu.ac.id)**

DOI: https://doi.org/10.47738/jads.v7i3.1215

Euclidean data, such as graphs, relational networks, and multidomain medical images, which frequently exhibit heterogeneous spatial configurations and complex topological relationships. Medical imaging data further exacerbate these challenges due to intrinsic variability across imaging modalities, scanners, acquisition protocols, and patient populations. Such variability often leads to domain shift, where discrepancies between training and testing data distributions degrade model performance and reliability. Consequently, CNN-based approaches may produce unstable or inconsistent predictions when deployed across heterogeneous clinical environments, raising concerns regarding their robustness and suitability for real-world medical applications.

To address the limitations of CNNs in non-Euclidean settings, Atwood and Towsley introduced Diffusion-Convolutional Neural Networks (DCNNs), a representation learning framework that models relationships among nodes through diffusion processes on graphs [1]. By propagating information across graph structures, DCNNs provide a principled mechanism for capturing relational dependencies that extend beyond local spatial neighborhoods. Following their introduction, DCNNs have been applied to various domains, including bioinformatics [2], video analysis [3], and large-scale graph representation learning [4], demonstrating their potential for structured and relational data modeling.

Although Diffusion-Convolutional Neural Networks (DCNNs) have been explored in several medical imaging studies, their adoption remains limited and non-standardized, particularly in clinical settings that require stable inference behavior, calibrated uncertainty, and robustness under domain shift [5]. In particular, repeated diffusion steps may amplify noise or induce excessive smoothing, resulting in degraded feature representations. Moreover, most existing DCNN architectures do not provide mechanisms for estimating predictive uncertainty, a critical limitation in clinical contexts where poorly calibrated predictions can lead to severe consequences in decision-making. These limitations have motivated growing interest in establishing stronger mathematical foundations to regulate diffusion dynamics and control parameter complexity within DCNN frameworks. Architectural modifications alone are often insufficient to guarantee stability and generalization, especially in high-dimensional and heterogeneous data regimes. Instead, principled analytical tools are required to characterize model behavior under varying conditions and to impose constraints that prevent instability.

Asymptotic analysis offers such a framework. It has long been employed in algorithmic theory [6], probabilistic statistics [7], and modern computational modeling [8] to describe system behavior as data dimensionality increases, distributions deviate, or extreme conditions such as high noise levels are encountered. By characterizing limiting behavior and growth rates, asymptotic analysis provides theoretical insights into stability, scalability, and complexity control. However, despite its established role in other fields, asymptotic principles have not yet been systematically integrated into DCNN architectures to enhance diffusion stability and generalization, particularly in medical imaging scenarios characterized by substantial variability in texture and structural patterns.

This observation reveals a significant and unresolved research gap. Existing studies have predominantly focused on architectural adaptations of graph neural networks, refinements to message-passing mechanisms, or extensive data augmentation strategies. While these approaches offer incremental improvements, they do not provide a mathematically grounded framework for asymptotically stabilizing diffusion processes, controlling parameter growth, or ensuring consistent predictions across heterogeneous medical domains. As a result, the reliability and clinical applicability of DCNN-based models remain limited.

To address this gap, the present study proposes a diffusion-based DCNN architecture augmented with asymptotic stabilization mechanisms designed to improve numerical stability, generalization, and cross-domain consistency. The proposed framework introduces two complementary components that operate at different stages of the learning and inference process. The selection of brain MRI, leukemia microscopy, and COVID chest imaging datasets was driven by methodological considerations rather than disease-specific objectives. These domains were intentionally chosen to represent heterogeneous medical imaging characteristics across multiple axes of variability, including spatial complexity, textural consistency, acquisition modality, and inter-domain distribution shift. Brain MRI images exhibit high structural and textural variability influenced by scanner settings and anatomical differences, leukemia microscopy images present relatively homogeneous cellular morphology with low domain noise, and COVID chest images introduce subtle inter-class boundaries combined with significant inter-institutional variability. This diversity enables a controlled evaluation of diffusion-based stabilization mechanisms under contrasting representational conditions, allowing the proposed framework to be assessed for robustness, calibration, and inference stability rather than domain-specific accuracy optimization.

These diffusion mechanisms are further reinforced through asymptotic regularization using L1 and L2 constraints. By imposing theoretical bounds on parameter growth, this regularization strategy mitigates overfitting and enhances generalization, particularly under limited data conditions, a common challenge in medical imaging research. The integration of asymptotic regularization with stabilized diffusion provides a principled framework for controlling model complexity and ensuring robust behavior under domain shift.

The proposed framework is empirically evaluated across three heterogeneous medical imaging domains: brain magnetic resonance imaging (MRI), leukemia microscopy images, and COVID-19 chest radiographs. These datasets represent diverse imaging modalities with distinct structural and textural characteristics. Experimental results demonstrate that the proposed model not only achieves competitive classification accuracy but also produces predictions that are more stable, better calibrated, and more robust to domain shift than baseline approaches.

In summary, this study advances DCNN-based medical image classification by offering both technical enhancements and theoretical contributions that integrate diffusion modeling, asymptotic analysis, and graph-based learning. By establishing a mathematically grounded approach to diffusion stabilization and complexity control, the proposed framework provides a more cohesive, stable, and reliable foundation for AI-driven clinical decision support systems, where robustness, calibration, and accountability are essential for real-world deployment.

2. Literature Review

2.1. Deep Learning

Recent studies have consistently demonstrated the effectiveness of deep learning in learning hierarchical representations from complex data, particularly in computer vision, speech processing, and natural language understanding [9],[10]. Compared with conventional machine learning approaches that rely heavily on handcrafted features, deep neural networks automatically learn discriminative representations through multiple nonlinear layers, enabling superior performance across a wide range of tasks [11]. In medical imaging, deep learning has become the dominant paradigm due to its ability to capture subtle anatomical variations and complex visual patterns from heterogeneous imaging modalities. Convolutional Neural Networks (CNNs) have achieved state of the art results in image classification, segmentation, and detection applications. Nevertheless, recent research indicates that deep learning models remain sensitive to domain shifts and often require large-scale annotated datasets, limiting their robustness and generalizability in heterogeneous and data-scarce medical environments [12]. These limitations motivate the exploration of alternative architectures capable of incorporating structural and relational information beyond conventional grid-based representations.

2.2. Diffusion-Convolutional Neural Networks

Since the introduction of Diffusion-Convolutional Neural Networks (DCNNs) by Atwood and Towsley [1], diffusion-based graph learning has attracted increasing attention as an effective approach for modeling relational and non-Euclidean data. Existing studies have shown that diffusion mechanisms enable the propagation of information across graph structures while preserving important topological relationships. This capability allows DCNNs to capture structural dependencies that are difficult to represent using conventional convolutional architectures. Subsequent research has successfully applied DCNNs in bioinformatics [13], video analysis [14], and large-scale graph representation learning [15], demonstrating their effectiveness in handling relational information. Despite these advances, several studies continue to report challenges associated with graph construction, diffusion stability, computational complexity, and scalability, particularly when applied to complex medical imaging scenarios. These limitations suggest that further improvements are required to enhance the robustness and stability of DCNN-based learning frameworks.

2.2.1. Role of Diffusion-Convolutional Neural Networks (DCNN)

Previous studies have highlighted the ability of DCNNs to model relational information through diffusion-based propagation mechanisms that aggregate information across graph neighborhoods [1]. By integrating information from multiple diffusion steps, DCNNs can simultaneously capture local and global structural patterns, resulting in richer feature representations than those produced by conventional CNN architectures. In medical imaging applications, diffusion-based graph representations have been explored to model anatomical relationships, spatial dependencies, and inter-sample similarities that cannot be adequately represented through pixel-level features alone [16]. Such capabilities are particularly valuable in multidomain environments where significant variations exist across imaging devices, acquisition protocols, and patient populations. However, recent investigations have identified oversmoothing as a major challenge, where excessive diffusion causes node representations to become increasingly similar, thereby reducing discriminative power and classification accuracy [17].

2.2.2. DCNN Network Layers

Existing DCNN architectures generally consist of graph construction modules, diffusion-convolution layers, and task-specific output layers. Prior studies have demonstrated that diffusion-convolution operations effectively integrate node features across multiple diffusion steps, enabling the extraction of multiscale relational information from graph-structured data. This capability has contributed to improved representation learning performance in various graph-based applications. However, research findings indicate that increasing diffusion depth may introduce numerical instability, parameter sensitivity, and noise amplification, particularly in dense and heterogeneous graph structures [18]. These challenges may degrade feature quality and reduce model robustness. Consequently, recent studies emphasize the importance of controlling diffusion dynamics and stabilizing information propagation within DCNN architectures to achieve reliable performance across diverse application domains.

2.2.3. Comparison Between DCNN and Kernel-on-Graph Methods

Several studies have compared diffusion-based neural architectures with kernel-on-graph methods for graph representation learning. Kernel-based approaches offer strong theoretical foundations and interpretability by defining graph convolution operations through mathematically derived kernels [19]. These methods have demonstrated effectiveness in various graph analysis tasks but often suffer from high computational complexity and limited scalability when applied to large-scale datasets. In contrast, DCNNs integrate diffusion mechanisms directly into neural network architectures, enabling end-to-end optimization and improved scalability [20]. Empirical studies suggest that DCNNs generally achieve better performance in large and complex graph environments. Nevertheless, unlike kernel-based approaches, many DCNN frameworks lack explicit theoretical guarantees regarding diffusion stability and asymptotic behavior. This limitation highlights the need for enhanced diffusion control mechanisms capable of improving both theoretical robustness and practical performance.

2.3. Convolutional Layer

Recent studies continue to identify convolutional operations as essential components of deep learning architectures for medical image analysis [21]. Through local connectivity and weight-sharing mechanisms, convolutional layers efficiently extract discriminative spatial features while maintaining computational feasibility. Research in medical imaging has demonstrated that convolutional representations are effective for capturing anatomical structures, textures, and pathological patterns from high-resolution images [22]. However, several studies have reported that conventional convolutional operations exhibit a strong locality bias, limiting their ability to model long-range dependencies and global structural relationships [23]. This limitation has motivated the integration of graph-based learning and diffusion mechanisms to complement traditional convolutional feature extraction and enhance contextual representation learning.

2.4. Training of Convolutional Neural Networks

The optimization of convolutional neural networks remains a critical factor influencing model performance and generalization. Previous studies have shown that gradient-based optimization algorithms, including stochastic gradient descent and adaptive methods such as Adam, are effective for training deep neural architectures [24], [25], [26]. Despite these advances, research in medical imaging consistently reports challenges related to limited labeled data, class imbalance, and domain heterogeneity [27]. These factors often increase the risk of overfitting and unstable convergence, thereby reducing model reliability in real-world clinical settings. Consequently, recent investigations emphasize the importance of robust optimization strategies, regularization techniques, and domain-aware learning mechanisms to improve training stability and generalization performance [28].

2.5. Cross-Entropy Cost Function

Cross-entropy loss remains one of the most widely adopted objective functions in classification tasks due to its effectiveness in optimizing probabilistic predictions [29]. Numerous studies have demonstrated that cross-entropy-based learning enables deep neural networks to achieve high classification accuracy across various application domains. However, recent research has also revealed important limitations, particularly regarding predictive calibration and uncertainty estimation. Models trained solely with cross-entropy loss frequently generate overconfident predictions despite incorrect classifications, potentially reducing reliability in safety-critical applications such as medical diagnosis [30]. These findings have motivated ongoing research into calibration-aware learning strategies and uncertainty-sensitive optimization frameworks.

2.6. Softmax Activation Function

Softmax-based prediction mechanisms remain the standard approach for transforming network outputs into probability distributions over predefined classes [31]. Empirical studies demonstrate that softmax facilitates efficient optimization and interpretable probabilistic outputs in multiclass classification problems. Nevertheless, recent investigations have consistently reported that softmax-based classifiers tend to produce overconfident predictions, particularly under distribution shifts, domain mismatches, and adversarial perturbations [32]. Such behavior may compromise model reliability and clinical trustworthiness. Consequently, improving calibration quality and predictive confidence estimation has emerged as an important research direction in modern deep learning systems.

2.7. ReLU Activation Function

The Rectified Linear Unit (ReLU) remains one of the most widely adopted activation functions in deep learning due to its computational simplicity and effectiveness in mitigating vanishing gradient problems [33]. Numerous studies have demonstrated that ReLU-based architectures achieve efficient optimization and strong predictive performance across various computer vision applications. However, previous research has also identified limitations such as dead neuron phenomena, sensitivity to initialization, and unstable activation dynamics under certain training conditions [34]. Despite these challenges, ReLU continues to serve as a practical and effective activation function in medical image classification frameworks.

2.8. Asymptotic Analysis

Recent research increasingly recognizes the importance of asymptotic analysis for understanding the stability, convergence, and scalability of machine learning models [35],[36]. Asymptotic perspectives provide theoretical insights into model behavior under increasing data complexity, parameter growth, and high-dimensional learning conditions. Several studies have applied asymptotic principles to investigate generalization performance, optimization dynamics, and model complexity in deep learning systems [37]. However, explicit incorporation of asymptotic stabilization mechanisms remains relatively unexplored in diffusion-based neural architectures. In particular, existing DCNN-based medical imaging frameworks rarely address theoretical issues related to diffusion stability and long-term convergence behavior. This research gap motivates the development of diffusion-stabilized DCNN models that integrate asymptotic considerations to improve robustness, reliability, and generalization in multidomain medical imaging environments. Overall, prior studies demonstrate that diffusion-based neural architectures provide effective mechanisms for modeling relational and heterogeneous data. Nevertheless, existing approaches continue to face challenges related to diffusion stability, probabilistic calibration, robustness under domain shifts, and theoretical convergence guarantees. Furthermore, the integration of asymptotic stabilization principles within DCNN-based medical image classification remains insufficiently explored. These limitations establish a clear research gap and provide the primary motivation for the diffusion-stabilized DCNN framework proposed in this study.

3. Methodology

The research workflow is designed to explicitly demonstrate the interconnections among methodological components, from dataset preparation to final model validation. The process begins with the collection of multidomain datasets, including brain MRI, leukemia microscopy images, and chest radiographs for COVID-19 and non-COVID cases. These datasets then undergo preprocessing steps consisting of intensity normalization, image resizing, and stratified data splitting to ensure balanced class distributions. L1/L2 regularization in this study is employed as a theoretical asymptotic constraint to prevent uncontrolled parameter growth rather than as a standalone performance-enhancing component. Therefore, its contribution is interpreted analytically rather than through isolated ablation. The empirical focus of this work lies in inference-time stabilization mechanisms rather than training-time regularization effects.

In the subsequent stage, a stochastic DCNN architecture integrated with a stabilized Diffusion2D module is trained using the Adam optimizer, incorporating L1 and L2 regularization as well as an early stopping mechanism to prevent overfitting. The baseline model is evaluated using standard performance metrics, including accuracy, Macro-F1 score, confusion matrices, and embedding visualizations generated through PCA and t-SNE. The k-nearest neighbors (kNN) graph was constructed in the latent embedding space using Euclidean distance. The value of k was selected empirically to balance local neighborhood consistency and diffusion stability. Preliminary evaluations were conducted for $\kappa \in \{5, 7, 9\}$, and $\kappa = 7$ was chosen as it provided the most stable calibration behavior without inducing oversmoothing effects. Within this range, model performance and diffusion behavior were observed to be robust, indicating low sensitivity to moderate variations of the neighborhood size.

The resulting embeddings are then utilized to construct a k-nearest neighbors graph, which serves as the foundation for the post hoc Anchored Diffusion mechanism applied in the logit space. Following inter-sample diffusion, the model is further assessed in terms of probabilistic reliability using Expected Calibration Error (ECE), Monte Carlo (MC) Dropout, and reliability diagrams. The final stage involves statistical validation through McNemar's test and ablation studies to quantify the contribution of each architectural module. The overall methodological workflow of the proposed diffusion-stabilized DCNN framework is illustrated in figure 1, covering dataset preprocessing, diffusion-based stabilization, anchored inference, calibration analysis, and statistical validation.

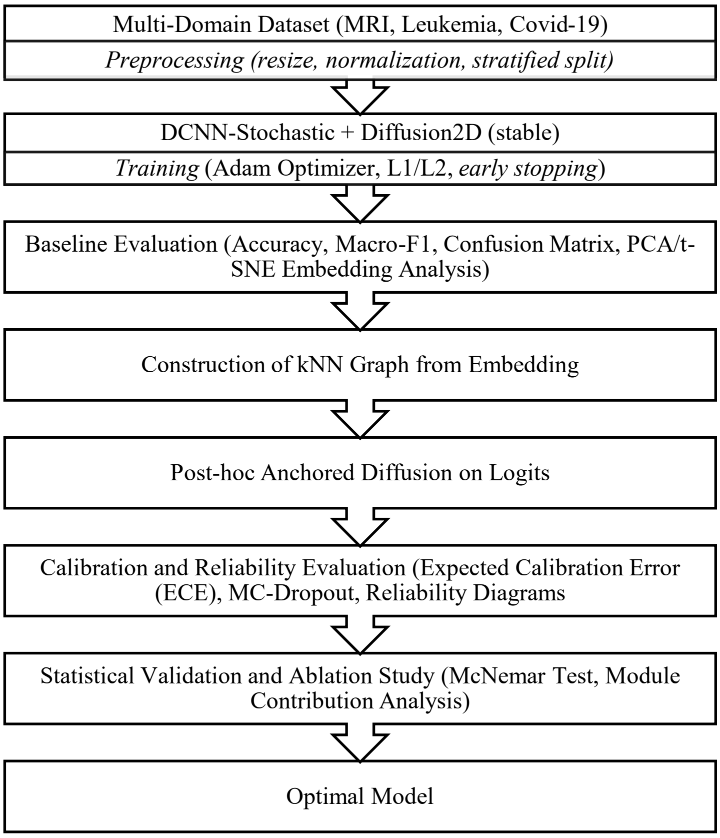

Figure 1. The research workflow

Diffusion2D is introduced as an intra-image diffusion operator specifically designed for medical images. Unlike conventional diffusion mechanisms, Diffusion2D maintains numerical stability under spatial and textural variability commonly observed across imaging modalities. By operating directly on two-dimensional image representations, this operator enhances structural coherence while preventing excessive smoothing. Experimental results demonstrate that Diffusion2D improves representational clarity, increases class separability, and reduces sensitivity to noise.

In this study, Diffusion2D is defined as a stabilized intra-image diffusion operator applied directly to two-dimensional feature representations prior to graph construction. Formally, Diffusion2D performs controlled spatial smoothing by iteratively updating feature maps according to a discrete Laplacian operator:

$$X^{(t+1)} = X^{(t)} - \propto \mathcal{L}\left(X^{(t)}\right) \tag{1}$$

where $\mathcal{L}(\cdot)$ denotes the two-dimensional Laplacian operator capturing local spatial gradients, and α is a bounded diffusion coefficient that regulates the magnitude of smoothing. Unlike graph-based diffusion in classical DCNNs, Diffusion2D operates at the intra-sample level and is designed to reduce representational oscillations caused by local intensity variations while preserving diagnostically relevant structures. This operator serves as a stabilization mechanism rather than a feature extractor and does not alter the backbone convolutional architecture.

4. Results and Discussion

Diffusion2D is integrated during training as a feature-level stabilization module, whereas Anchored Diffusion operates strictly as a post-hoc inference mechanism and does not require retraining. The claim of no retraining applies exclusively to the Anchored Diffusion stage.

4.1. Ablation Study Analysis and Contribution of Diffusion Mechanism

It is important to emphasize that Diffusion2D is not designed as an accuracy optimization module. Instead, its primary role is to stabilize latent representations under spatial and domain-induced variability, ensuring that predictive behavior remains consistent without degrading baseline accuracy. An ablation study was conducted to identify the contribution of Diffusion2D and noise injection to the DCNN architecture by disabling each component separately. The goal was to assess whether the improvement in prediction stability stems from the diffusion mechanism, the stochastic effects of noise, or a combination of both. The evaluation used accuracy, macro-F1, and performance per class. To evaluate the contribution of Diffusion2D and stochastic noise injection to inference stability, the comparative ablation results are summarized in table 1.

Table 1 Ablation study results using accuracy, Macro-F1, and prediction stability metrics.

Model Configuration	Diffusion2D	Noise	Accuracy	Macro-F1	Main Characteristics
DCNN Complete	√	√	64.70%	0.7045	The most stable prediction, decision change is only 0.6% of the total sample
Diffusion OFF (Noise ON)	-	√	64.95%	0.7073	Representation is more sensitive to distributional variations, decreased stability
Noise OFF (Diffusion ON)	√	-	65.04%	0.7085	Stability increases, prediction patterns are more controlled and less extreme.

The comparable accuracy values observed across ablation settings in table 1 indicate that the integration of Diffusion2D does not artificially inflate performance metrics. Instead, it preserves classification accuracy while significantly improving prediction stability and robustness, which are not captured by accuracy alone. Therefore, the effectiveness of Diffusion2D should be evaluated based on stability-related indicators, including: (i) reduction in spatial gradient variance, (ii) consistency of latent embeddings, and (iii) resistance to domain-induced representational drift, rather than raw accuracy gains.

4.2. Implementation of Diffusion2D in Model Architecture

The integration of Diffusion2D aims to strengthen the stability of spatial representations at the *feature encoding stage* without replacing the main feature extraction mechanism. This module works as a partial diffusion regularizer that suppresses local pixel fluctuations with extreme intensities, thus ensuring a more controlled distribution of internal activations even when the model is faced with different variations in medical image domains. This concept aligns with previous findings on baseline models that demonstrated high sensitivity to local gradient variations, particularly in MRI domains with complex textures. Formally, the diffusion operation follows the equation:

$$I_{t+1}(x,y) = I_t(x,y) + \lambda \cdot \Delta I_t(x,y) \tag{2}$$

ΔI represents the Laplacian operator that estimates local spatial gradient variations, while λ denotes the stabilization coefficient controlling the diffusion magnitude. As shown in Equation (2), the diffusion process performs controlled spatial smoothing without removing diagnostically relevant structures. Consequently, Diffusion2D reduces representational oscillations while preserving important anatomical and pathological characteristics. The proposed Diffusion2D stabilization mechanism begins by applying a fixed four-neighbor Laplacian operator independently to each feature channel to estimate local spatial variations. A learnable parameter vector is subsequently transformed into channel-wise diffusion coefficients constrained by a predefined stability threshold. The resulting diffusion response is iteratively incorporated into the original feature representation through residual updates, allowing local inconsistencies to be gradually reduced while preserving discriminative anatomical and pathological information. After K diffusion iterations, the stabilized feature maps are forwarded to the graph construction stage for subsequent relational learning. The proposed stabilization procedure applies a constrained diffusion process to intermediate feature maps. The learnable diffusion coefficient is restricted by α_{max} to ensure numerical stability while enabling adaptive smoothing across feature channels. Through iterative diffusion updates, local feature inconsistencies and high-frequency fluctuations are progressively reduced without significantly affecting discriminative anatomical and pathological information. As a result, the stabilized feature representations exhibit improved robustness to domain-specific variations while preserving the structural characteristics required for accurate classification.

Internal evaluation shows that the diffusion parameters $\alpha$are within a narrow and stable range, with a mean value of 0.141734 and a standard deviation of 0.003757, as generated in the model inference process. These data indicate that the changes occurring during the diffusion process are controlled and do not cause representational deviations. The Laplacian calculation results show an average decrease from 0.864355 to 0.187525, or a reduction of 78.23%, which indicates the diffusion function as an effective gradient softening mechanism without causing loss of diagnostic structures such as leukemia cell shape or brain tissue pattern.

Although the accuracy improvement is marginal ($\Delta Acc = + 0.0003$), Diffusion2D integration has a significant effect on the stability of cross-domain predictions. The impact of the proposed stabilization mechanism is reflected in the improved overall classification performance, as indicated by the increase in the Macro-F1 score from 0.7045 to 0.7061. Importantly, this improvement was achieved without causing any performance degradation in the leukemia and COVID-19 domains, demonstrating the consistency of the proposed approach across multiple datasets. In addition, the model exhibited a more robust predictive response to pixel-level shift variations in the MRI domain, suggesting that the Diffusion2D stabilization mechanism effectively enhances feature stability while preserving discriminative information.

A comparison of PCA and t-SNE reveals a more discrete and defined representation of some classes after Diffusion2D integration. This pattern indicates that the diffusion effect strengthens the shape of the latent manifold, resulting in a more consistent vectorial structure across samples within the same domain. This observation is consistent with the hypothesis that extreme gradient reduction helps improve *latent smoothness.* The quantitative impact of Diffusion2D is summarized in table 2.

Table 2. Quantitative impact of Diffusion2D on representation stability and prediction consistency.

Evaluation Aspects	Baseline Condition	After Diffusion2D	Implications
α stability	Not available	Stable (±0.0038)	representation *drift* occurs
Laplacian Mean	0.864355	0.187525	−78.23% spatial noise
Macro-F1	0.7045	0.7061	Prediction consistency increases
Prediction Changes	–	0.6% of cases changed	Probabilistic stability

These findings indicate that Diffusion2D should not be interpreted as a direct accuracy optimization mechanism. Instead, its primary contribution lies in improving representational stability, probabilistic consistency, and robustness under multidomain medical imaging conditions. The stabilized latent representations subsequently provide a more reliable foundation for the post-hoc Anchored Diffusion mechanism discussed in the following subsection.

4.3. Implementation of Anchored Diffusion in the Post-inference Stage

After stabilizing spatial representations through Diffusion2D, the next stage applies Anchored Diffusion as a post-hoc inference mechanism operating in the logit space. Unlike conventional retraining-based refinement strategies, Anchored Diffusion adjusts probabilistic decisions without modifying the learned feature extraction backbone or requiring additional gradient optimization. Consequently, the proposed mechanism functions as a decision refinement layer that improves probabilistic consistency while preserving the original latent representation learned during training.

Anchored Diffusion exploits the relational structure between samples through a k-nearest neighbors (kNN) graph constructed from the latent embedding space generated during inference. Logit values are iteratively updated using a combination of the initial logits and aggregated neighborhood logits derived from nearby samples within the latent manifold. The iterative process terminates once relative changes between consecutive diffusion steps satisfy the convergence criterion. This procedure ensures that decision refinement remains gradual and controlled, thereby preventing excessive correction that could disrupt the underlying representational structure.

To evaluate the selective decision behavior of Anchored Diffusion, a confidence-based prediction analysis was conducted using a threshold of $\tau = 0.75$. This threshold was selected to ensure that only predictions with sufficiently reliable confidence levels were accepted as final outputs. The resulting accept–reject distribution is illustrated in figure 2.

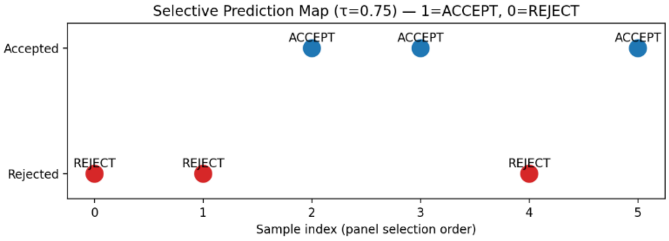

Figure 2. Selective Prediction Map based on probability threshold τ = 0.75 after Anchored Diffusion refinement.

Figure 2 illustrates the distribution of accepted (blue) and rejected (red) predictions after the Anchored Diffusion process. The horizontal axis represents sample indices within the evaluation panel, while the vertical axis indicates the final selection category. The observed pattern shows that rejected samples are concentrated near the confidence threshold region, whereas predictions with stronger latent representational consistency are consistently accepted. To ensure transparency in the evaluation process, representative samples exhibiting prediction changes before and after Anchored Diffusion were recorded and analyzed. Table 3 summarizes five representative cases selected from the 20 prediction adjustments observed during inference.

Table 3. Representative Prediction Changes After Anchored Diffusion Refinement

No	Sample Index	Initial Prediction (CNN Baseline)	Prediction After Anchored Diffusion	Final Confidence	Selection Status
1	1325	*brain_Training*	*leukemia_Segmented*	1,000	ACCEPT
2	1577	*brain_Training*	*leukemia_Segmented*	1,000	ACCEPT
3	3156	*sarscov2_non-COVID*	*sarscov2_COVID*	0.859	ACCEPT
4	3106	*sarscov2_COVID*	*sarscov2_non-COVID*	0.764	REJECT
5	2984	*brain_Testing*	*brain_Training*	0.742	REJECT

Importantly, the diffusion mechanisms do not alter the learned feature representations or force label changes. Instead, they refine decision consistency by repositioning uncertain samples within the latent neighborhood structure, allowing collective evidence to resolve ambiguous logits.

The analysis results show that prediction changes occur primarily in samples with inhomogeneous spatial distributions or those near the classification boundary. In contrast, the leukemia domain showed no decision changes, indicating that the latent representation for that domain has stabilized since the Diffusion2D stage. The fact that only 0.6% of samples experienced prediction changes suggests that the anchoring mechanism does not disrupt the model as a whole, but rather selectively corrects predictions with high uncertainty. An inspection of the samples affected by Anchored Diffusion (table 3) reveals that prediction changes are concentrated on borderline cases characterized by heterogeneous spatial patterns or low inter-class margins. In contrast, samples with well-separated latent representations remain unchanged, confirming that Anchored Diffusion operates as a targeted refinement mechanism rather than a global decision modifier. In high-risk medical applications, indiscriminate prediction changes are undesirable; therefore, a low but targeted correction rate is preferable to aggressive post-hoc adjustment. The misclassified samples observed prior to diffusion do not indicate a failure in feature extraction, but rather boundary-induced uncertainty, where samples reside near class decision margins within the latent embedding space. Decision deferral is operationalized through selective prediction, where samples with confidence below τ = 0.75 are explicitly rejected rather than forcibly classified.

4.4. Risk Coverage and Stability Trace Validation

The next evaluation focused on examining risk behavior characteristics *and* the relationship between prediction coverage *and* model decision reliability. This evaluation was conducted because, in the medical context, model performance is measured not only by predictive accuracy but also by the system's ability to manage uncertainty during inference. Therefore, the model must be able to reject unstable or highly uncertain predictions, as demonstrated through the selection mechanism in the previous subsection. To further evaluate inference reliability under repeated prediction scenarios, a stability trace analysis was conducted across 40 inference iterations. The resulting confidence trajectories are presented in figure 3.

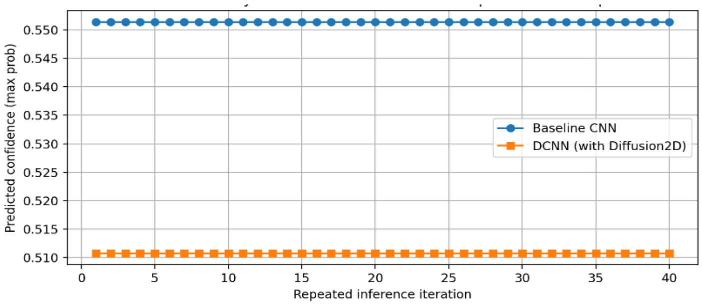

Figure 3. Stability Trace Confidence in Repeated Inference (n = 40)

This figure shows a comparison of the stability of prediction probabilities between the CNN baseline and the Diffusion2D–Anchored Diffusion model. The CNN baseline shows an average confidence level of around 0.551 with small fluctuations between iterations. However, the diffusion model shows a lower confidence level (0.510) but is very stable with no variation between iterations. While the magnitude of stabilization effects varies across domains, consistent improvements in inference stability and calibration are observed in all three datasets (MRI, leukemia, and COVID).

4.5. Visual Comparative Analysis of Baseline CNN vs DCNN with Diffusion

The visual analysis in this section evaluates the changes in the model's predictive characteristics before and after the integration of Diffusion2D and Anchored Diffusion. The evaluation focuses not on the accuracy level alone, but on changes in probabilistic belief patterns, interdomain consistency, and the model's response to multidomain medical image variations. Thus, this section reinforces the thesis that probabilistic stability and reliability are equally important indicators as numerical performance, especially in the context of clinical applications that are vulnerable to the risk of misdiagnosis. Representative prediction comparisons across leukemia, brain MRI, and COVID imaging domains are illustrated in figure 4.

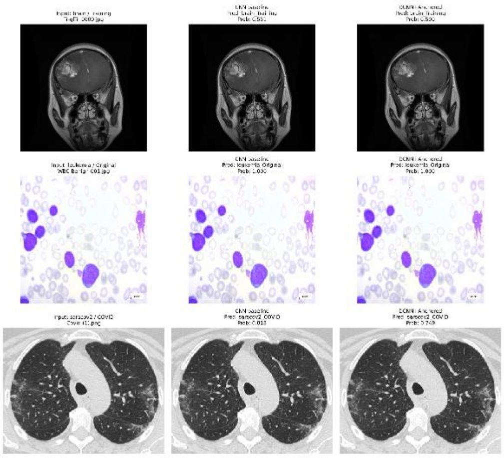

Figure 4. Comparison Panel of Baseline CNN vs Diffusion2D–Anchored DCNN Model Predictions

Overall, the comparison results show three main patterns, the first of which is the change in ambiguous cases (Near-Decision Threshold Samples). Samples with a narrow probability margin (especially MRI images with complex textures) show a clearer difference after diffusion. In the baseline model, the prediction decision is strongly dependent on local pixel fluctuations. This is evident from the probability values that are close to the decision threshold (≈0.51–0.56). Accordingly, figure 4 should be interpreted as visual confirmation of statistically and probabilistically validated trends, rather than as an isolated qualitative claim.

After diffusion integration, the predicted probabilities become *smoother* and approach the estimated collective latent distribution (*anchored state*). This phenomenon indicates that the model prioritizes a stable representation vector structure over instantaneous responses to micro-spatial variations. Following diffusion integration, prediction probabilities become smoother and converge toward more stable latent neighborhood distributions. Although the predicted labels remain unchanged in most samples, confidence values consistently move toward stabilized convergence regions with relatively small fluctuations of approximately 2–3%. Second, the case of Error Rectification Events (ERR) for several samples previously misclassified by the baseline CNN demonstrates label correction after diffusion. This case is characterized by a significant probability shift (greater than 8–14%), indicating that diffusion not only smooths out noise but also strengthens the manifold structure between classes.

4.6. Qualitative Validation of Diffusion-Stabilized Inference Behavior

This subsection provides a qualitative evaluation of inference behavior before and after the integration of Diffusion2D and Anchored Diffusion across multidomain medical imaging datasets. Unlike the numerical evaluation presented in previous subsections, this analysis focuses on probabilistic consistency, decision stability, and the visual characteristics of prediction behavior under heterogeneous domain conditions. Representative qualitative comparisons between the baseline CNN and the proposed diffusion-stabilized framework are presented in figure 5 across brain MRI, leukemia, and COVID imaging domains. Each panel includes the input image, predicted class, and confidence score, allowing inference behavior to be analyzed directly at the probabilistic output level rather than solely through final classification labels.

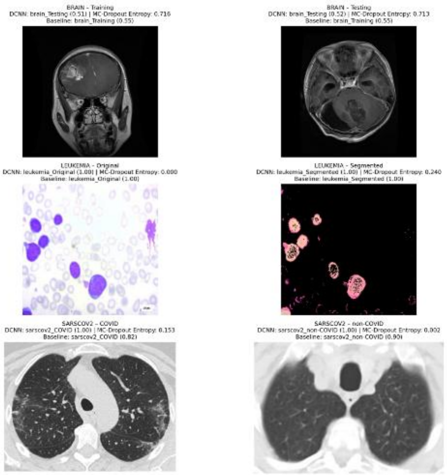

Figure 5. Qualitative comparison of prediction behavior between the baseline CNN and the Diffusion2D–Anchored DCNN framework across multidomain medical imaging datasets.

Overall, qualitative observations indicate that the proposed diffusion-stabilized framework produces more structured and probabilistically consistent predictions across heterogeneous medical imaging domains. Although prediction labels remain unchanged for most samples, the confidence distributions exhibit smoother and more stable behavior following diffusion refinement. In samples characterized by complex spatial textures, particularly within the MRI domain, the baseline CNN demonstrates high sensitivity to local pixel fluctuations, resulting in unstable confidence values concentrated near the decision threshold region. After the integration of Diffusion2D and Anchored Diffusion, prediction probabilities become smoother and converge toward more stable latent neighborhood distributions. This behavior suggests that the proposed framework prioritizes latent representational consistency rather than instantaneous responses to local spatial perturbations.

Several prediction correction cases were also observed in samples previously misclassified by the baseline CNN. These corrections are characterized by substantial probability shifts, indicating that the diffusion process not only suppresses representational noise but also strengthens the latent manifold structure between neighboring classes. Consequently, uncertain predictions located near class boundaries become more consistent after diffusion refinement.

Anchored Diffusion operates exclusively in the logit space as a post-hoc inference mechanism and therefore does not require retraining, gradient computation, or modification of the feature extraction backbone. The computational overhead introduced by the method primarily depends on the size of the k-nearest neighbors graph and the number of diffusion iterations rather than network depth or image resolution.

While the proposed framework is motivated by clinical decision support requirements, this study does not claim real-time deployment readiness. Detailed latency analysis and optimization for high-throughput clinical environments remain important directions for future research.

In this framework, interpretability is defined in terms of probabilistic transparency and decision consistency rather than feature attribution. Accordingly, the proposed approach emphasizes stable confidence trajectories, selective prediction behavior, and explicit rejection of uncertain samples instead of attribution-based explanation techniques such as SHAP or LIME. The observed reduction in inference volatility and stabilization of confidence trajectories empirically support the theoretical expectation of asymptotic stabilization, namely bounded representational behavior and reduced sensitivity to perturbations across heterogeneous medical imaging domains.

5. Conclusion

This study yields three principal conclusions that highlight its contribution to the development of machine learning systems for multidomain medical image classification. First, the integration of Diffusion2D and Anchored Diffusion is shown to fundamentally alter the inference dynamics of the model, shifting it from a purely deterministic decision-making process toward a more conservative and uncertainty-aware framework. This shift is reflected in the model's ability to regulate predictions based on uncertainty levels, maintain probabilistic stability across inference iterations, and reduce the tendency toward overconfidence that is commonly observed in conventional CNN-based models. Second, empirical results indicate that the primary impact of the diffusion mechanisms lies not in maximizing numerical accuracy, but in enhancing the consistency of latent representations and the stability of the inference process. The ablation results confirm that Diffusion2D does not aim to outperform baseline CNNs in terms of accuracy, but instead ensures that such performance is achieved with greater representational stability and reliability—an essential requirement for medical decision support systems.

Consequently, the proposed model performs classification while simultaneously treating decision-making as a calibrated probabilistic function that is measurable and interpretable—an essential property for medical domains characterized by high decision risk. Third, the proposed approach demonstrates strong potential as a more adaptive and reliable machine learning architecture when extended toward risk-aware decision systems. Rather than serving solely as a predictive tool, the model functions as an information filtering mechanism capable of deferring or rejecting decisions when elevated uncertainty is detected, thereby supporting safer and more accountable clinical decision-making. Although recent graph neural network architectures such as Graph Attention Networks (GAT) and GraphSAGE have demonstrated strong representation learning capabilities, this study does not aim to benchmark architectural expressiveness across GNN variants. Instead, the proposed framework focuses on inference-time stabilization through diffusion regulation and anchoring mechanisms that are applied independently of the backbone graph architecture. Consequently, comparisons with GAT or GraphSAGE are beyond the current scope and are left as promising directions for future work, particularly in the context of integrating attention-based message passing with diffusion-stabilized inference.

6. Declarations

6.1. Author Contributions

Conceptualization: H.W.D. and S.; Methodology: H.W.D.; Software: H.W.D.; Validation: H.W.D., S., P.S., and M.A.B.; Formal Analysis: H.W.D. and S.; Investigation: H.W.D.; Data Curation: H.W.D.; Writing Original Draft Preparation: H.W.D.; Writing Review and Editing: H.W.D., S., P.S., and M.A.B.; Visualization: H.W.D.; Supervision: S., P.S., and M.A.B. All authors have read and agreed to the published version of the manuscript.

6.2. Data Availability Statement

The corresponding author is available upon request for the data used in this study.

6.3. Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

6.4. Institutional Review Board Statement

Not applicable.

6.5. Informed Consent Statement

Not applicable.

6.6. Declaration of Competing Interest

All authors declare that there are no conflict of interest.

References

[1] J. Atwood and D. Towsley, "Diffusion-Convolutional Neural Networks," in *Advances in Neural Information Processing Systems, vol. 29, Daniel D. Lee, Masashi Sugiyama, Ulrike von Luxburg, Isabelle Guyon, and Roman Garnett, Eds., Curran Associates, Inc.*, vol. 2016, no. 1, pp. 1993–2001. Available: https://proceedings.neurips.cc/paper/2016/hash/390e982518a50e280d8e2b535462ec1f-Abstract.html

[2] D. Andreoletti, S. Troia, F. Musumeci, S. Giordano, G. Maier, and M. Tornatore, “Network Traffic Prediction Based on Diffusion Convolutional Recurrent Neural Networks,” in *Proc. IEEE INFOCOM 2019 – IEEE Conf. Comput. Commun. Workshops (INFOCOM WKSHPS), Paris, France*, vol. 2019, no. Apr., pp. 246–251, 2019, doi: 10.1109/INFCOMW.2019.8845132.

[3] K. Tamil Selvi, R. Thamilselvan, and S. Mohana Saranya, “Diffusion convolution recurrent neural network–a.K. Tamil Selvi, R. Thamilselvan, and S. Mohana Saranya, “Diffusion Convolution Recurrent Neural Network—A Comprehensive Survey,” *IOP Conference Series: Materials Science and Engineering*, vol. 1055, no. 1, Art. no. 012119, pp. 10-13, 2021, doi: 10.1088/1757-899X/1055/1/012119.

[4] O. Guendoul, M. Zobi, H. A. Abdelali, Y. Tabii, R. O. H. Thami, and O. Bourja, “Capturing Spatio-Temporal Patterns of Falls Individuals Using Efficient Graph Convolutional Network Model,” *Applied Intelligence*, vol. 55, no. 11, pp. 1-12, 2025, doi: 10.1007/s10489-025-06316-5.

[5] P. Wu, Xi. He, W. Dai, J. Zhou, Y. Shang, Y.Fan, and T. Hu, "A Review on Research and Application of AI-Based Image Analysis in the Field of Computer Vision," *IEEE Access*, vol. 13, no. 1, pp. 76684–76702, 2025, doi: 10.1109/ACCESS.2025.3565300.

[6] A. Wigderson and J. Zuiddam, "Asymptotic spectra: Theory, applications and extensions," *Bull. Amer. Math. Soc.*, vol. 63, no. 2, pp. 199–331, 2026, doi: 10.1090/bull/1880.

[7] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, and P. S. Yu, “A comprehensive survey on graph neural networks,” IEEE Trans. *Neural Netw. Learn. Syst.*, vol. 32, no. 1, pp. 4–24, 2020, doi: 10.1109/TNNLS.2020.2978386

[8] O. Kuznetsov, E. Frontoni, N. Kryvinska, O. Smirnov, and A. E. Imoize, “Computational Modeling of Enhanced Spread Spectrum Codes for Asynchronous Wireless Communication,” in *Computational Intelligence in Communication Systems and Software Engineering, Boca Raton, FL, USA: CRC Press*, vol. 2024, no. 1, pp. 403–447, 2024, doi: 10.1201/9781003457428-17.

[9] Mishra, R. Kumar, Reddy, G. Y. Sandesh, Pathak, Himanshu, “The Understanding of Deep Learning: A Comprehensive Review,” *Mathematical Problems in Engineering*, vol. 2021, no. 1, pp. 1-15, 2021, doi: 10.1155/2021/5548884

[10] I. Goodfellow, Y. Bengio, and A. Courville, “Softmax Units for Multinoulli Output Distributions,” in *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016, pp. 180–184, ISBN: 978-0-26203561-3.

[11] N. Rasool and J. I. Bhat, “Unveiling the complexity of medical imaging through deep learning approaches,” *Chaos Theory and Applications*, vol. 5, no. 4, pp. 267–280, 2023, doi: 10.51537/chaos.1326790.

[12] A. A. Abe and M. Nyathi, “SULBA: A Task-Agnostic Data Augmentation Framework for Deep Learning in Medical Image Analysis,” *Diagnostics*, vol. 16, no. 10, Art. no. 1546, pp. 1-12, 2026. doi: 10.3390/diagnostics16101546.

[13] Z. Guo, J. Liu, Y. Wang, M. Chen, D. Wang, D. Xu, and J. Cheng, “Diffusion models in bioinformatics and computational biology,” *Nat. Rev. Bioeng.*, vol. 2, no. 2, pp. 136–154, Feb. 2024, doi: 10.1038/s44222-023-00114-9.

[14] Z. Xing, Q. Feng, H. Chen, Q. Dai, H. Hu, H. Xu, Z. Wu, and Y.-G. Jiang, “A Survey on Video Diffusion Models,” *ACM Comput. Surv.*, vol. 57, no. 2, pp. 1–42, Feb. 2025, doi: 10.1145/3696415.

[15] M. Bistroń, J. M. Żurada, and Z. Piotrowski, "Deep Learning for Image Watermarking: A Comprehensive Review and Analysis of Techniques, Challenges, and Applications," *Sensors*, vol. 26, no. 2, Art. no. 444, pp. 1-12, 2026, doi: 10.3390/s26020444.

[16] Y. Liu, Y. Shi, F. Mu, J. Cheng, C. Li, and X. Chen, “Multimodal MRI Volumetric Data Fusion With Convolutional Neural Networks,” *IEEE Transactions on Instrumentation and Measurement*, vol. 71, no. 1, pp. 1–15, 2022, doi: 10.1109/TIM.2022.3184360.

[17] T. Konstantin Rusch, Michael M. Bronstein, and Siddhartha Mishra, “A Survey on Oversmoothing in Graph Neural Networks,” *arXiv preprint arXiv*:2303.10993, vol. 2023, no. 1, pp. 1-12, 2023, doi: 10.48550/arXiv.2303.10993.

[18] Z.Li, L. Xia, H. Hua, S. Zhang, S. Wang, and C. Huang, “DiffGraph: Heterogeneous Graph Diffusion Model,” in *Proceedings of the Eighteenth ACM International Conference on Web Search and Data Mining (WSDM ’25), Hannover, Germany*, vol. 2025, no. 1, pp. 40–49, 2025, doi: 10.1145/3701551.3703590.

[19] T. Xiao, T. Ai, H. Yu, M. Yang, and P. Liu, “A point selection method in map generalization using graph convolutional network model,” *Cartogr. Geogr. Inf. Sci.*, vol. 51, no. 1, pp. 20–40, 2024, doi: 10.1080/15230406.2023.2187886.

[20] B. B. Moser, A. S. Shanbhag, F. Raue, S. Frolov, S. Palacio, and A. Dengel, “Diffusion models, image super-resolution, and everything: A survey,” *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 36, no. 7, pp. 11793–11813, 2024, doi: 10.48550/arXiv.2401.00736.

[21] M. M. Taye, “Theoretical Understanding of Convolutional Neural Network: Concepts, Architectures, Applications, Future Directions,” *Computation*, vol. 11, no. 3, Art. no. 52, pp. 1-12, 2023, doi: 10.3390/computation11030052.

[22] D. R. Sarvamangala and R. V Kulkarni, “Convolutional neural networks in medical image understanding: a survey,” *Evol. Intell.*, vol. 15, no. 1, pp. 1–22, 2022, doi: 10.1007/s12065-020-00540-3.

[23] H. Chen, C. Xu, L. Zheng, Q. Zhang, and X. Lin, “Diffusion-based graph generative methods,” *IEEE Trans. Knowl. Data Eng.*, vol. 36, no. 12, pp. 7954–7972, 2024, doi: 10.1109/TKDE.2024.3466301.

[24] I. A. T. Hashem, F. A. Alaba, M. H. Jumare, A. O. Ibrahim, and A. W. Abulfaraj, “Adaptive Stochastic Conjugate Gradient Optimization for Backpropagation Neural Networks,” *IEEE Access*, vol. 12, no. 1, pp. 33757–33768, 2024, doi: 10.1109/ACCESS.2024.3370859.

[25] F.Mehmood, S. Ahmad, and T. K. Whangbo, “An Efficient Optimization Technique for Training Deep Neural Networks,” *Mathematics*, vol. 11, no. 6, Art. no. 1360, pp. 1-14, 2023, doi: 10.3390/math11061360.

[26] S. Nagendram, A. Singh, G. H. Babu, R. Joshi, S. D. Pande, S. K. H. Ahammad, D. Dhabliya, and A. Bisht, “Stochastic Gradient Descent Optimisation for Convolutional Neural Network for Medical Image Segmentation,” *Open Life Sciences*, vol. 18, no. 1, Art. no. 20220665, pp. 1-18, 2023, doi: 10.1515/biol-2022-0665.

[27] S. Dhanka, A. Sharma, A. Kumar, S. Maini, and H. Vundavilli, “Advancements in hybrid machine learning models for biomedical disease classification using integration of hyperparameter-tuning and feature selection methodologies: A comprehensive review,” *Archives of Computational Methods in Engineering*, vol. 33, no. 1, pp. 289–324, 2026., doi: 10.1007/s11831-025-10309-5.

[28] C. F. G. Dos Santos and J. P. Papa, “Avoiding overfitting: A survey on regularization methods for convolutional neural networks,” *ACM Computing Surveys (Csur)*, vol. 54, no. 10s, pp. 1–25, 2022, doi: 10.1145/3510413.

[29] G. Polat, Ü. M. Çağlar, and A. Temizel, “Class Distance Weighted Cross Entropy Loss for Classification of Disease Severity,” *Expert Systems with Applications*, vol. 269, Art. no. 126372, pp. 1-13, 2025, doi: 10.1016/j.eswa.2024.126372.

[30] M. J. Meni, R. T. White, M. L. Mayo, and K. R. Pilkiewicz, “Entropy-Based Guidance of Deep Neural Networks for Accelerated Convergence and Improved Performance,” *Information Sciences*, vol. 681, Art. no. 121239, pp. 1-16, 2024, doi: 10.1016/j.ins.2024.121239.

[31] I. Kouretas and V. Paliouras, “Simplified Hardware Implementation of the Softmax Activation Function,” in *2019 8th International Conference on Modern Circuits and Systems Technologies (MOCAST), Thessaloniki, Greece*, vol. 2019, no. 1, pp. 1–4, 2019, doi: 10.1109/MOCAST.2019.8741677.

[32] G. Xia and C.-S. Bouganis, “Augmenting the Softmax with Additional Confidence Scores for Improved Selective Classification with Out-of-Distribution Data,” *Int. J. Comput. Vis.*, vol. 132, no. 9, pp. 3714–3752, 2024, doi: 10.1007/s11263-024-02029-3.

[33] X. Hu, P. Niu, J. Wang, and X. Zhang, “A Dynamic Rectified Linear Activation Units,” *IEEE Access*, vol. 7, no. 1, pp. 180409–180416, 2019, doi: 10.1109/ACCESS.2019.2959036.

[34] A. R. Khan and S. Almuhaideb, “TRex: A Smooth Nonlinear Activation Bridging Tanh and ReLU for Stable Deep Learning,” *Electronics,* vol. 14, no. 23, Art. no. 4661, pp. 1-21, 2025, doi: 10.3390/electronics14234661.

[35] E. Melezhnikov, O. Shestakov, and E. Stepanov, “Asymptotic Analysis of a Thresholding Method for Sparse Models with Application to Network Delay Detection,” *Mathematics*, vol. 14, no. 1, Art. no. 148, pp. 1-12, 2026, doi: 10.3390/math14010148.

[36] T. Hussain, E. Villamor, M. Shakil, M. Ahsanullah, and B. M. G. Kibria, “A Novel Probabilistic Model for Streamflow Analysis and Its Role in Risk Management and Environmental Sustainability,” *Axioms*, vol. 15, no. 2, Art. no. 113, pp. 1-12, 2026, doi: 10.3390/axioms15020113.

[37] R. Sun, “Optimization for Deep Learning: An Overview,” *Journal of the Operations Research Society of China*, vol. 8, no. 2, pp. 249–294, 2020, doi: 10.1007/s40305-020-00309-6.