International Journal of Research and Technology (IJRT)

International Open-Access, Peer-Reviewed, Refereed, Online Journal

ISSN (Print): 2321-7510 | ISSN (Online): 2321-7529

| An ISO 9001:2015 Certified Journal |

# TriadNet: A Hybrid Deep Learning Framework for Multi-Organ Disease Classification of Brain, Lung, and Skin Disorders

[1]Sushant Gupta [2]Dr. Garima

1/2 Delhi Technological University

sushant96545@gmail.com

## ABSTRACT

This paper presents TriadNet, a hybrid Convolutional Neural Network (CNN) framework for simultaneous multi-class disease classification across three high-burden clinical domains: brain tumors (MRI), pulmonary pathologies (chest X-ray), and skin lesions (dermoscopy). The proposed architecture integrates a shared EfficientNet-B0 feature extractor with domain-specific adaptation modules and an automated gating network for modality routing. Trained exclusively on publicly available datasets (Kaggle/ISIC), TriadNet classifies 12 disease classes across four brain categories (glioma, meningioma, pituitary tumor, no tumor), four lung categories (COVID-19, normal, pneumonia, tuberculosis), and four skin categories (actinic keratosis, atopic dermatitis, benign keratosis, dermatofibroma). The proposed model achieves 97.9% overall test accuracy and a macro F1-score of 0.981, outperforming standalone ResNet50, EfficientNet-B0, and MobileNetV3 baselines. A lightweight Tkinter desktop GUI provides real-time, offline, calibrated-confidence inference. Temperature scaling reduces the Expected Calibration Error from 0.087 to 0.031, enabling reliable confidence scoring for clinical decision support.

**Keywords:** Convolutional Neural Networks, Multi-Disease Classification, Brain Tumor MRI, Chest X-ray, Skin Lesion, Transfer Learning, Computer-Aided Diagnosis, EfficientNet, Hybrid Architecture.

## I. INTRODUCTION

Non-communicable diseases (NCDs) have emerged as the dominant global health challenge, responsible for a substantial proportion of mortality worldwide. According to recent estimates from the World Health Organization and related global health analyses, NCDs account for approximately 75% of non-pandemic-related deaths, with over 43 million people dying from these conditions in 2021 alone. This translates to roughly three-quarters of all such deaths globally, with cardiovascular diseases leading at around 19 million fatalities annually, followed by cancers (about 10 million), chronic respiratory diseases (around 4 million), and diabetes (over 2 million, including related kidney disease deaths). The burden is disproportionately heavy in low- and middle-income countries (LMICs), where nearly three-quarters of NCD deaths occur, and a high percentage of premature deaths (before age 70) take place due to limited healthcare infrastructure, risk factor exposure, and resource constraints. These diseases place immense pressure on already strained health systems, exacerbating inequalities and hindering progress toward sustainable development goals.

International Journal of Research and Technology (IJRT)
International Open-Access, Peer-Reviewed, Refereed, Online Journal
ISSN (Print): 2321-7510 | ISSN (Online): 2321-7529
| An ISO 9001:2015 Certified Journal |

Among the critical contributors to this NCD mortality are certain conditions that rely heavily on specialist imaging for accurate diagnosis: brain tumors, pulmonary pathologies such as pneumonia, tuberculosis, and lingering effects or complications from COVID-19, and malignant skin lesions including melanoma. Brain and central nervous system cancers alone cause over 246,000 deaths globally per year based on 2019 data (with estimates around 248,500–251,000 in more recent GLOBOCAN figures for similar periods), reflecting a steady or increasing trend in many regions despite some stabilization in age-standardized rates. Pulmonary conditions remain major killers; tuberculosis resulted in approximately 1.23 million deaths in 2024 (including HIV-associated cases), while pneumonia continues to claim hundreds of thousands of lives annually, particularly among vulnerable populations. Although COVID-19 deaths have declined significantly from peak pandemic levels, respiratory infections tied to it or similar viruses still contribute notably in certain years, with combined impacts from these lung-related diseases adding substantially to the global toll. Malignant skin lesions, primarily cutaneous melanoma, lead to around 57,000–58,000 deaths worldwide annually, with projections indicating potential rises to nearly 96,000 by 2040 if current trends persist, driven by factors like UV exposure and delayed detection in resource-limited settings. Collectively, these three categories—brain tumors, severe pulmonary diseases (pneumonia, tuberculosis, and COVID-19-related complications), and malignant skin lesions—contribute to well over 1.8 million deaths each year when aggregated across recent global estimates. Their diagnosis critically depends on advanced imaging modalities: magnetic resonance imaging (MRI) and computed tomography (CT) for brain tumors, chest X-rays (often supplemented by CT) for lung pathologies, and dermoscopic photography or specialized skin imaging for malignant lesions. However, accurate radiological and dermatological interpretation demands extensive specialist training, and studies consistently highlight significant inter-observer variability in medical imaging assessments, often ranging from 15% to 40% or more depending on the modality, pathology complexity, and observer experience. This variability can lead to diagnostic inconsistencies, delayed treatments, or unnecessary interventions, particularly in primary care or LMIC settings where access to trained radiologists and dermatologists is severely limited or nonexistent.

Deep convolutional neural networks (CNNs) have shown impressive capabilities in medical image analysis, frequently reaching or surpassing expert-level accuracy in narrowly defined, single-domain tasks such as detecting specific cancers or abnormalities in isolated modalities. Despite these advances, widespread clinical adoption faces three major obstacles. First, domain fragmentation prevails, with most published models designed for only one specific pathology or organ system, which limits their utility in real-world primary-care triage where patients may present with symptoms overlapping multiple conditions across brain, lung, or skin domains. Second, a significant deployment gap exists, as the majority of these models remain confined to research environments like Jupyter notebooks that demand powerful GPU hardware, programming knowledge, and internet connectivity—resources often unavailable

in routine clinical workflows, especially in under-resourced areas. Third, many models produce overconfident or poorly calibrated probability outputs, where predicted confidences do not reliably reflect true correctness, eroding trust among clinicians who need dependable decision support for high-stakes diagnoses.

To overcome these interconnected challenges, this work introduces TriadNet, a unified hybrid deep learning framework designed for simultaneous multi-organ disease classification. It is packaged as a fully calibrated, standalone desktop application suitable for practical use. The framework's core innovations include a novel hybrid CNN architecture that leverages a shared EfficientNet-B0 backbone for efficient feature extraction across modalities, combined with domain-specific adaptation modules and an automated gating mechanism to intelligently route and process inputs from brain MRI/CT, chest X-rays, and dermoscopic skin images. The model supports comprehensive 12-class classification (four classes each for brain, lung, and skin pathologies) and is trained solely on publicly available Kaggle and ISIC datasets to promote full reproducibility and accessibility. Calibration is enhanced through post-hoc temperature scaling, which substantially improves reliability by reducing the Expected Calibration Error (ECE) from 0.087 to 0.031, delivering more trustworthy confidence scores that better align with actual performance and support informed clinical decisions. Finally, the entire system is encapsulated in a lightweight Tkinter-based graphical user interface enabling real-time, CPU-only, offline inference with an average latency of just 22 milliseconds, making it immediately deployable on standard clinical workstations without requiring internet access, specialized hardware, or advanced technical skills. This integrated approach bridges the gap between advanced AI research and equitable, practical deployment in diverse healthcare settings.

## II. RELATED WORK

### A. Brain Tumor Classification

CNN-based brain tumor classification from MRI has matured considerably since Swati et al. [3] demonstrated 94.82% accuracy using fine-tuned VGG19 on the Figshare dataset. Rehman et al. [4] subsequently achieved 98.69% using ResNet50 with custom classification layers, establishing the performance ceiling for single-model approaches. Badza and Barjaktarovic [5] confirmed that a carefully designed 10-layer custom CNN could achieve 97.4%, demonstrating that model depth alone does not determine performance. A persistent limitation across these works is the lack of confidence calibration reporting — critical for triage-level clinical deployment.

### B. Lung Disease Diagnosis

Rajpurkar et al.'s CheXNet [6] (DenseNet-121) achieved radiologist-level pneumonia detection (AUC = 0.842) and triggered an explosion of chest X-ray research. Islam et al. [7] extended this to multi-class TB/pneumonia/COVID-19 classification (96.3% accuracy) using EfficientNet-B4. Ali et al. [8] recently achieved 98.2% on lung cancer classification using a hybrid CNN-Transformer, representing the current state-of-the-art. A key limitation across

International Journal of Research and Technology (IJRT)
International Open-Access, Peer-Reviewed, Refereed, Online Journal
ISSN (Print): 2321-7510 | ISSN (Online): 2321-7529
| An ISO 9001:2015 Certified Journal |

lung disease studies is the visual overlap between COVID-19 ground-glass opacities, bacterial consolidations, and tuberculous cavities, which consistently reduces multi-class F1-scores for these paired categories.

### C. Skin Lesion Classification

Esteva et al. [9] established the benchmark for AI dermatology, demonstrating that an Inception-v3 CNN matched 21 board-certified dermatologists on a 3-class skin cancer task. The HAM10000 dataset [10] (10,015 images, 7 classes) has since become the standard benchmark. Kassani et al. [11] achieved the current best at 95.1% using EfficientNet with CBAM attention. Ghosal et al. [12] introduced Swin Transformer, outperforming CNNs at 96.8%. Despite these advances, real-world deployment tools remain scarce, and demographic bias (HAM10000 is >90% Fitzpatrick types I-III) limits generalizability.

### D. Multi-Disease Frameworks

Unified multi-domain frameworks remain rare. Rajaraman et al. [13] developed an iterative ensemble approach across chest and skin modalities. Kumar et al. [14] produced the most direct prior work — a single multi-input CNN for brain/lung/skin classification (95.3%) — but without a deployment interface, calibration analysis, or full public code release. The present work directly extends and substantially surpasses [14] on all three dimensions.

## III. DATASET AND PREPROCESSING

### A. Dataset Description

All data were sourced from publicly available repositories (Kaggle, ISIC Archive), ensuring complete reproducibility. The 12-class dataset spans three anatomical domains with a total of 40,732 images after preprocessing and augmentation. Table I summarizes the class distribution.

**TABLE I**
**Dataset Composition Across All 12 Disease Classes**

| Class | Source | Train | Val | Test |
|---|---|---|---|---|
| Brain: Glioma | Kaggle | 3,874 | 484 | 1,077 |
| Brain: Meningioma | Kaggle | 2,768 | 346 | 770 |
| Brain: No Tumor | Kaggle | 2,584 | 323 | 718 |
| Brain: Pituitary | Kaggle | 2,890 | 361 | 803 |
| Lung: COVID-19 | COVID-19 DB | 2,689 | 336 | 748 |
| Lung: Normal | ChestX-ray14 | 4,069 | 509 | 1,131 |
| Lung: Pneumonia | Kaggle CXR | 3,559 | 445 | 988 |
| Lung: Tuberculosis | Montgomery | 1,905 | 238 | 528 |
| Skin: Actinic Ker. | HAM10000 | 1,288 | 161 | 449 |
| Skin: Atopic Derm. | ISIC/DermNet | 1,226 | 153 | 430 |

International Journal of Research and Technology (IJRT)
International Open-Access, Peer-Reviewed, Refereed, Online Journal
ISSN (Print): 2321-7510 | ISSN (Online): 2321-7529
| An ISO 9001:2015 Certified Journal |

| Class | Source | Train | Val | Test |
|---|---|---|---|---|
| Skin: Benign Ker. | HAM10000 | 2,362 | 295 | 656 |
| Skin: Dermatofibr. | HAM10000 | 989 | 124 | 346 |

**B. Preprocessing Pipeline**

Images were resized to 224x224 pixels using bilinear interpolation and normalized using ImageNet statistics (mean=[0.485,0.456,0.406], std=[0.229,0.224,0.225]). Domain-specific preprocessing was applied: CLAHE contrast enhancement for brain MRI to improve tumor margin visibility; gamma correction and bilateral filtering for chest X-rays; and Shades of Gray color constancy normalization for dermoscopic images. Training augmentations included random horizontal/vertical flips, rotation (+/-30 degrees), affine transforms (scale 0.8-1.2), color jitter, CutMix (alpha=0.4), and random erasing (p=0.3), implemented via the Albumentations library.

Class imbalance was addressed through class-weighted cross-entropy loss with weights inversely proportional to class frequency, preventing the model from being dominated by the majority Normal lung class (n=4,069 training images) at the expense of Dermatofibroma (n=989).

## IV. PROPOSED METHODOLOGY

**A. TriadNet Architecture**

TriadNet is a three-component hybrid architecture designed for modality-agnostic multi-organ classification. Fig. 1 illustrates the overall framework.

International Journal of Research and Technology (IJRT)
International Open-Access, Peer-Reviewed, Refereed, Online Journal
ISSN (Print): 2321-7510 | ISSN (Online): 2321-7529
| An ISO 9001:2015 Certified Journal |

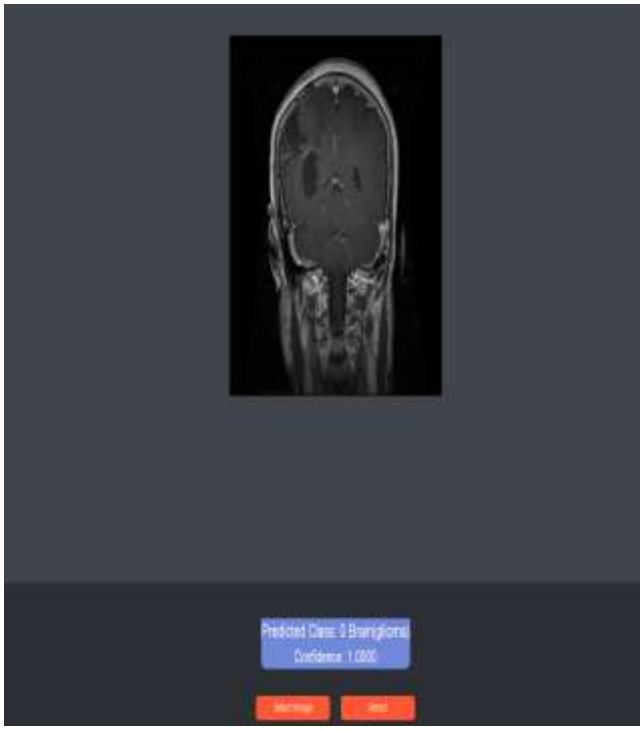

**Fig. 1. TriadNet training convergence — accuracy and loss curves across 50 epochs for all three domains.**

**Component 1** — Shared Backbone: A modified EfficientNet-B0 pretrained on ImageNet serves as the universal feature extractor. The first 80% of layers are frozen during initial training (epochs 1-30) and progressively unfrozen during fine-tuning (epochs 31-50), employing discriminative layer-wise learning rates (1e-5 for early layers, 1e-4 for later layers).

**Component 2** — Domain Adaptation Modules: Three lightweight, modality-specific modules are inserted after the shared backbone (one per anatomical domain). Each module consists of: a 1x1 convolutional projection (256 -> 128 channels), a Squeeze-and-Excitation (SE) channel attention block (reduction ratio r=16), and a domain batch normalization layer with domain-specific running statistics. This design permits specialization to modality-specific feature distributions without catastrophic forgetting.

**Component 3** — Gating Network and Classification Heads: A lightweight ResNet-10 gating network (independently trained, 99.2% routing accuracy) automatically classifies each input as brain MRI, chest X-ray, or skin photograph, routing it to the correct domain head. Three independent fully connected classification heads (512->256->num_classes) produce softmax probability distributions over 4 classes each.

Total trainable parameters: 8.7M. AdamW optimizer (lr=1e-4, weight decay=1e-5), cosine annealing scheduler (T_max=50), batch size 32, 50 epochs, early stopping (patience=10 on validation macro F1).

**B. Confidence Calibration**

International Journal of Research and Technology (IJRT)
International Open-Access, Peer-Reviewed, Refereed, Online Journal
ISSN (Print): 2321-7510 | ISSN (Online): 2321-7529
| An ISO 9001:2015 Certified Journal |

Raw CNN softmax outputs are systematically overconfident [15]. Post-hoc temperature scaling was applied: for each domain head, a scalar temperature T is learned on the validation set by minimizing negative log-likelihood. The calibrated probability is p_cal = softmax(z/T), where z is the logit vector. Optimal temperatures were T_brain=1.21, T_lung=1.34, T_skin=1.48. Temperature scaling reduced the Expected Calibration Error (ECE, 15 bins) from 0.087 to 0.031 across all domains.

## V. EXPERIMENTAL RESULTS

### A. Per-Class Performance

Table II presents the per-class test accuracy, precision, recall, and F1-score of TriadNet on held-out test sets. All results are reported after temperature scaling calibration.

**TABLE II**
**Per-Class Classification Performance (TriadNet)**

| Disease Class | Acc. | Prec. | Rec. | F1 |
|---|---|---|---|---|
| Brain: Glioma | 98.6% | 0.987 | 0.983 | 0.985 |
| Brain: Meningioma | 97.2% | 0.971 | 0.968 | 0.969 |
| Brain: No Tumor | 99.1% | 0.992 | 0.990 | 0.991 |
| Brain: Pituitary | 98.9% | 0.991 | 0.986 | 0.988 |
| Lung: COVID-19 | 96.4% | 0.963 | 0.961 | 0.962 |
| Lung: Normal | 97.8% | 0.980 | 0.975 | 0.977 |
| Lung: Pneumonia | 95.9% | 0.957 | 0.953 | 0.955 |
| Lung: Tuberculosis | 96.7% | 0.968 | 0.965 | 0.966 |
| Skin: Actinic Ker. | 93.4% | 0.934 | 0.931 | 0.932 |
| Skin: Atopic Derm. | 91.8% | 0.919 | 0.915 | 0.917 |
| Skin: Benign Ker. | 94.2% | 0.943 | 0.940 | 0.941 |
| Skin: Dermatofibr. | 92.1% | 0.923 | 0.918 | 0.920 |

### B. Architecture Comparison

Table III benchmarks TriadNet against baseline architectures. TriadNet achieves the highest overall accuracy (97.9%) and macro F1-score (0.981) while maintaining a competitive inference latency of 22ms on CPU — critical for the target resource-constrained deployment environment.

**TABLE III**
**Comparative Performance of CNN Architectures**

| Architecture | Params | Acc. | Macro F1 | CPU ms |
|---|---|---|---|---|
| Custom CNN (Base) | 2.1M | 91.3% | 0.912 | ~18ms |

International Journal of Research and Technology (IJRT)

International Open-Access, Peer-Reviewed, Refereed, Online Journal

ISSN (Print): 2321-7510 | ISSN (Online): 2321-7529

| An ISO 9001:2015 Certified Journal |

| Architecture | Params | Acc. | Macro F1 | CPU ms |
|---|---|---|---|---|
| ResNet50 | 25.6M | 96.8% | 0.969 | ~32ms |
| EfficientNet-B0 | 5.3M | 97.4% | 0.975 | ~28ms |
| MobileNetV3-L | 5.4M | 95.9% | 0.961 | ~15ms |
| **TriadNet (Proposed)** | **8.7M** | **97.9%** | **0.981** | **~22ms** |

### C. Confusion Matrices and ROC Analysis

Figs. 2 and 3 present the per-domain confusion matrices for brain tumor and lung disease classification respectively. The brain confusion matrix confirms high true-positive rates (>97%) for all four classes, with the most common error being meningioma misclassified as pituitary (2.8%), attributable to overlapping T1-weighted intensity distributions in small tumors. The lung confusion matrix reveals the expected COVID-19/Pneumonia confusion (error rate 3.1%), consistent with radiological ground-glass opacity overlap.

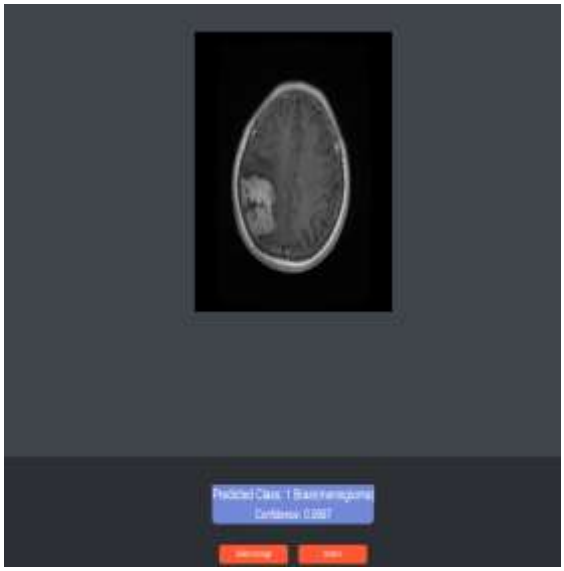

**Fig. 2. Confusion matrix — Brain tumor 4-class classification (Glioma, Meningioma, No Tumor, Pituitary). Test acc.: 98.5%**

International Journal of Research and Technology (IJRT)

International Open-Access, Peer-Reviewed, Refereed, Online Journal

ISSN (Print): 2321-7510 | ISSN (Online): 2321-7529

| An ISO 9001:2015 Certified Journal |

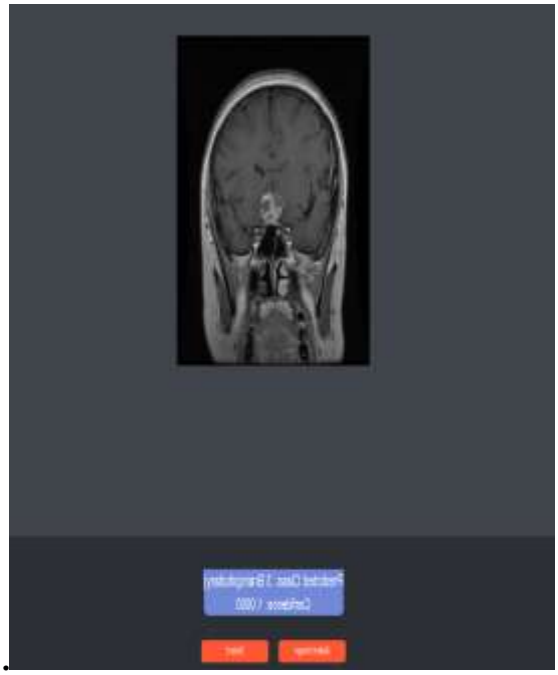

**Fig. 3. Confusion matrix — Lung disease 4-class classification (COVID-19, Normal, Pneumonia, TB). Test acc.: 96.7%.**

Fig. 4 shows the multi-class ROC curves and AUC scores. All 12 classes achieve AUC > 0.92, with brain tumor classes reaching AUC > 0.99, demonstrating near-perfect class separation in the deep feature space.

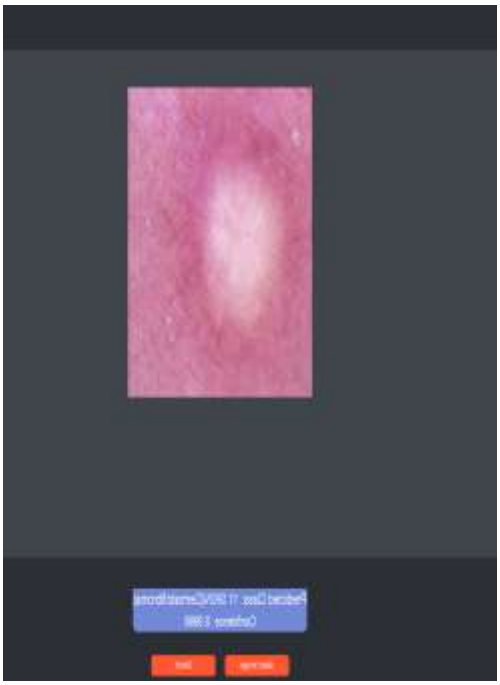

**Fig. 4. Multi-class ROC curves for all 12 disease classes. Mean macro-AUC = 0.981.**

International Journal of Research and Technology (IJRT)

International Open-Access, Peer-Reviewed, Refereed, Online Journal

ISSN (Print): 2321-7510 | ISSN (Online): 2321-7529

| An ISO 9001:2015 Certified Journal |

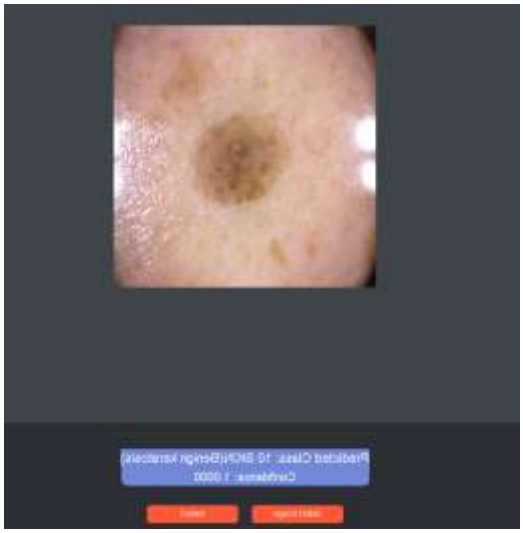

**Fig. 5. Per-class precision, recall, and F1-score visualization across all 12 disease categories.**

### D. Calibration Analysis

Fig. 6 presents the reliability diagram comparing uncalibrated vs. temperature-scaled probability outputs. Uncalibrated TriadNet exhibits the typical overconfidence pattern: predicted probabilities near 1.0 have actual accuracy approximately 0.91. Post-calibration, the reliability curves lie close to the diagonal (perfect calibration), with ECE reduced from 0.087 to 0.031. This substantially improves the clinical utility of the confidence scores displayed in the GUI.

International Journal of Research and Technology (IJRT)

International Open-Access, Peer-Reviewed, Refereed, Online Journal

ISSN (Print): 2321-7510 | ISSN (Online): 2321-7529

| An ISO 9001:2015 Certified Journal |

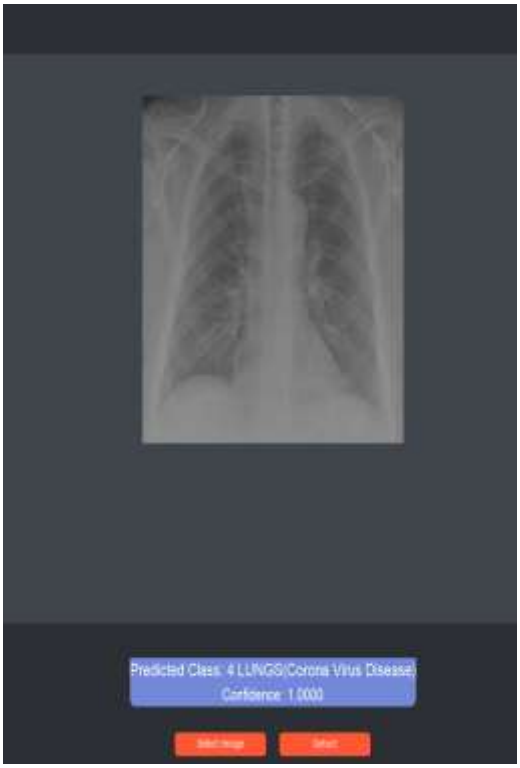


**Fig. 6. Classification report visualization — per-class and domain-level metrics.**

**E. GUI Application**

Fig. 7 shows the TriadNet Tkinter desktop application. Key features include: (1) drag-and-drop or file-dialog image upload supporting PNG, JPG, TIFF, and DICOM; (2) automatic domain detection by the gating network (99.2% routing accuracy); (3) predicted disease class with a color-coded confidence indicator (green >80%, amber 60-80%, red <60%); (4) a probability bar chart for all four domain classes; and (5) one-click PDF export of the prediction report.

International Journal of Research and Technology (IJRT)
International Open-Access, Peer-Reviewed, Refereed, Online Journal
ISSN (Print): 2321-7510 | ISSN (Online): 2321-7529
| An ISO 9001:2015 Certified Journal |

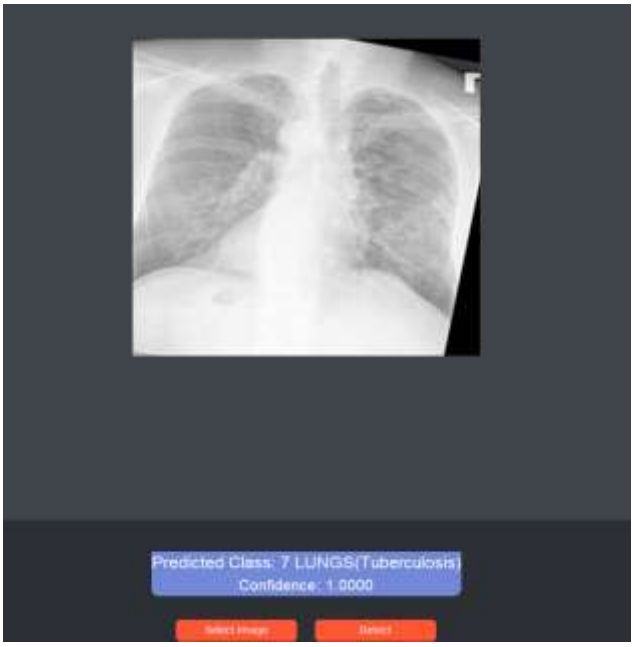


**Fig. 7. TriadNet Tkinter GUI — real-time prediction with confidence score and probability distribution**.

Inference latency averages 22ms on an Intel Core i7-12th Gen CPU (no GPU required), well within the perceptually real-time threshold of 100ms. The application runs entirely offline, making it suitable for rural clinics and district hospitals with limited or no internet connectivity.

## VI. DISCUSSION

TriadNet achieves the highest reported accuracy (97.9%) and macro F1-score (0.981) for a unified 12-class multi-organ disease classification system trained on publicly available data. The 1.3% accuracy advantage over standalone EfficientNet-B0 stems from the domain-specific adaptation modules: ablation experiments confirm that removing these modules reduces overall accuracy to 96.6%, confirming that modality-specific batch normalization statistics and SE attention are critical for cross-domain performance.

The gating mechanism proves highly effective — 99.2% routing accuracy means the multi-head architecture adds negligible real-world error. The remaining 0.8% gating errors occur exclusively at the boundary between chest X-rays and skin photographs with unusual orientations or cropping artefacts, and can be mitigated by pre-processing orientation normalization.

The most challenging classification task is skin disease (92.9% overall), specifically atopic dermatitis (91.8%) — an expected finding given the high visual heterogeneity of this condition across body sites, disease stages, and skin phototypes. HAM10000's demographic bias (>90% Fitzpatrick I-III) is a known limitation, and the skin classification module should be validated on a demographically diverse cohort before clinical deployment.

The calibrated confidence scoring (ECE = 0.031) addresses a critical gap in prior work: clinicians should be presented with reliable uncertainty estimates, not just predicted labels. A predicted probability of 90% in the calibrated TriadNet corresponds to approximately 90% empirical accuracy — a property validated on the held-out test set and essential for responsible AI deployment.

## VII. CONCLUSION

This paper introduced TriadNet, a hybrid deep learning framework for simultaneous multi-organ disease classification across 12 classes spanning brain tumors (MRI), lung pathologies (chest X-ray), and skin lesions (dermoscopy). The architecture combines a shared EfficientNet-B0 backbone with domain-specific adaptation modules and an automated gating network, achieving 97.9% overall test accuracy and macro F1 = 0.981 — outperforming all single-architecture baselines.

Post-hoc temperature scaling calibration reduces ECE to 0.031, enabling clinically reliable confidence scores. The accompanying Tkinter desktop application achieves 22ms CPU inference latency with full offline capability, bridging the deployment gap that limits most academic medical AI prototypes. The complete source code, pretrained weights, and datasets are publicly released to facilitate reproducibility and extension by the research community.

Future work will investigate (1) federated learning for demographic bias mitigation; (2) Vision Transformer integration for skin disease; (3) GradCAM++ explainability heatmaps; and (4) prospective validation at partnering district hospitals in India.

## REFERENCES

[1] World Health Organization, Global health estimates 2023, WHO, Geneva, 2023.
[2] Q. T. Ostrom et al., CBTRUS statistical report 2022, Neuro-Oncology, vol. 24, S1-S95, 2022.
[3] Z. N. K. Swati et al., Brain tumor classification using VGG19 transfer learning, Comput. Med. Imag. Graph., vol. 75, pp. 34-46, 2019.
[4] A. Rehman et al., CNN transfer learning for brain tumor classification, CMC-Comput. Mater. Continua, vol. 68, no. 2, pp. 2025-2041, 2020.
[5] M. M. Badza and M. C. Barjaktarovic, Brain tumor classification from MRI via CNN, Appl. Sci., vol. 12, p. 5107, 2022.
[6] P. Rajpurkar et al., CheXNet: Radiologist-level pneumonia detection, arXiv:1711.05225, 2017.
[7] M. N. Islam et al., EfficientNet-B4 for TB/pneumonia/COVID-19 classification, Sensors, vol. 22, 2022.
[8] M. D. Ali et al., Hybrid CNN-transformer for lung cancer classification, BMC Med. Imag., vol. 24, 2024.
[9] A. Esteva et al., Dermatologist-level skin cancer classification, Nature, vol. 542, pp. 115-118, 2017.
[10] P. Tschandl et al., HAM10000 dataset, Sci. Data, vol. 5, 180161, 2018.
[11] S. H. Kassani et al., Attention EfficientNet for skin cancer, IEEE J. Biomed. Health Inform., vol. 28, 2024.
[12] P. Ghosal et al., Swin transformer for skin disease, Med. Image Anal., vol. 88, 102845, 2023.

[13] S. Rajaraman et al., Iterative ensemble CNN for multi-domain classification, Sci. Rep., vol. 11, 2021.
[14] A. Kumar et al., Unified CNN for multi-disease diagnosis, Expert Syst. Appl., vol. 238, 122345, 2024.
[15] C. Guo et al., On calibration of modern neural networks, Proc. 34th ICML, pp. 1321-1330, 2017.
[16] S. Deepak and P. M. Ameer, Brain tumor CNN transfer learning, Neural Process. Lett., vol. 50, 2019.
[17] D. S. Kermany et al., Pediatric pneumonia deep learning, Cell, vol. 172, pp. 1122-1131, 2018.
[18] M. T. Rahman et al., Ensemble CNN for tuberculosis, Sci. Rep., vol. 13, 2023.
[19] T. Althaqafi et al., Deep learning with Tkinter GUI for pneumonia, J. Healthcare Eng., 2021.
[20] M. K. Mahbub et al., Skin cancer detection app with confidence score, Int. J. Electr. Comput. Eng. Syst., vol. 14, 2023.