AI IN MEDICAL IMAGING, ONCOLOGY, EXPLAINABLE AI

AI-Assisted Multimodal Diagnostic Framework for Prostate and Cervical Cancer Detection: Synthesizing Computer Vision, Large Language Models, and Evidence-Based Oncology via an HTML5 Canvas UI

Risav Dey

Department of Computer Applications, Techno India University, Kolkata, West Bengal, India

Published November 2026

Abstract

The rapid advancement of Artificial Intelligence (AI) in medical imaging has catalyzed a transformative shift in automated disease detection, anatomical segmentation, and clinical decision support. This comprehensive thesis outlines the end-to-end development, mathematical foundation, and front-end evaluation of an advanced Clinical Decision Support System (CDSS) tailored specifically for prostate (Transrectal Ultrasound - TRUS) and cervical (Transvaginal Ultrasound - TVS) cancer diagnostics.

By synthesizing supervised deep architectures (including Convolutional Neural Networks and Vision Transformers), dynamic tabular biomarker evaluations, and Retrieval-Augmented Generation (RAG) linked to authoritative oncological guidelines (NCCN, FIGO), the proposed framework bridges the critical gap between raw sonographic imaging and actionable clinical pathways.

A major contribution of this research is the engineering of an interactive, standalone HTML5 <canvas> prototype. This interface incorporates real-time Explainable AI (XAI) overlays—namely Grad-CAM heatmaps and bounding boxes—to validate the diagnostic workflow, demystify the "black box" of deep learning, and build essential clinician trust. Furthermore, seamless integration with the Google Gemini API demonstrates the immense potential for advanced multimodal reasoning, ensuring strict adherence to evidence-based medicine. Finally, a PostGIS-backed geospatial referral engine mathematically matches patients to appropriate regional specialists based on algorithmically determined urgency tiers. This document spans the entire development lifecycle, proposing a highly scalable, ethically sound standard for the future of intelligent oncology.

1 Introduction

1.1 The global oncological burden and diagnostic paradigms

The global burden of oncological diseases remains one of the most pressing systemic challenges in modern healthcare infrastructure. Prostate and cervical cancers represent two of the most prevalent and lethal malignancies affecting the male and female populations, respectively. According to global epidemiological data, early detection and accurate pathological staging are the single most significant determinants of long-term patient survival. Accurate staging dictates the trajectory of clinical interventions, which can range from active, non-invasive surveillance to radical prostatectomy, or concurrent chemoradiotherapy.

1.2 The physics and role of ultrasound imaging

Medical imaging serves as the absolute cornerstone of oncological diagnostics. While modalities such as multi-parametric Magnetic Resonance Imaging (mpMRI) and Computed Tomography (CT) offer high-resolution anatomical insights, Ultrasound imaging remains the primary, non-ionizing, highly accessible, and cost-effective tool for preliminary pelvic evaluation.

Specifically, Transrectal Ultrasound (TRUS) is utilized for the prostate, and Transvaginal Ultrasound (TVS) is utilized for the cervix. Ultrasound physics relies on the emission of high-frequency sound waves (typically 7 to 12 MHz for transvaginal and transrectal probes) and the recording of their echoes as they bounce off tissue interfaces with varying acoustic impedances.

However, the interpretation of B-mode (Brightness mode) sonographic data is notoriously operator-dependent. Diagnoses are frequently obscured by inherent imaging artifacts. The most prominent of these is speckle noise—a granular, multiplicative interference pattern caused by the coherent processing of backscattered signals from multiple sub-resolution tissue targets. This intrinsically low soft-tissue contrast makes identifying early-stage hypoechoic lesions exceptionally difficult for the human eye.

1.3 The AI paradigm shift in oncology

The integration of Artificial Intelligence (AI) paradigms, specifically deep representation learning, offers unprecedented capabilities in automated feature extraction, lesion localization, and disease staging. This research presents an extensive blueprint for a comprehensive Clinical Decision Support System (CDSS). The proposed system uniquely merges computer vision (image processing), tabular clinical biomarkers (e.g., PSA levels, free/total PSA ratios, HPV status), and evidence-based medicine via Large Language Models (LLMs).

1.4 Research objectives and scope

The primary objectives of this comprehensive thesis are defined as follows:

1. **Architectural Formulation:** To conceptualize and mathematically formulate a dual-pathway AI pipeline capable of preprocessing and segmenting both TRUS and TVS sonograms, utilizing U-Net and Vision Transformer (ViT) topologies.
2. **Multimodal LLM Integration:** To design a Retrieval-Augmented Generation (RAG) framework utilizing the Google Gemini API to anchor AI predictions in established NCCN and FIGO clinical guidelines, thereby preventing algorithmic hallucinations.
3. **Frontend Prototyping:** To construct an interactive, single-file HTML5 <canvas> prototype that renders simulated B-mode ultrasound frames, complete with Explainable AI (XAI) overlays (Grad-CAM and Bounding Boxes) to foster clinician trust.
4. **Geospatial Triage:** To architect a PostGIS-backed geographical referral engine that dynamically matches diagnosed patients with appropriate specialists based on computed urgency tiers and the Haversine distance formula.

2 Clinical fundamentals and staging frameworks

To ensure the AI system operates within established medical paradigms, it is strictly programmed to synthesize multi-modal data into outputs aligned with gold-standard staging systems.

2.1 Prostate cancer and PI-RADS v2.1

Prostate cancer screening typically begins with the measurement of serum Prostate-Specific Antigen (PSA) and a Digital Rectal Examination (DRE). Elevated PSA levels (> 4.0 ng/mL) or an abnormal, nodular DRE prompt further imaging. While the Prostate Imaging Reporting and Data System (PI-RADS) was originally designed for multiparametric MRI, its risk-stratification logic is actively being adapted for high-resolution TRUS and micro-ultrasound modalities.

Table 1 outlines the logic matrix the AI utilizes to classify prostatic lesions.

Table 1. PI-RADS v2.1 risk stratification matrix adapted for ultrasound.

Score	Risk level	Ultrasound / clinical biomarker presentation
1	Very Low	Uniform echogenicity, no focal lesions. Normal PSA (< 2.5 ng/mL).
2	Low	Circumscribed, homogeneous hypoechoic mass (often BPH). Mildly elevated PSA.
3	Intermediate	Heterogeneous echotexture with obscured margins. Equivocal free/total PSA ratio.
4	High	Distinct hypoechoic focal lesion in the peripheral zone < 1.5 cm. High PSA velocity.
5	Very High	Definite hypoechoic mass > 1.5 cm with extraprostatic extension. Abnormal DRE.

2.2 Cervical cancer and FIGO 2018 staging

Cervical cancer is predominantly driven by persistent infections with high-risk Human Papillomavirus (HPV), particularly strains 16 and 18. Diagnosis combines Pap smear cytology (e.g., HSIL - High-grade Squamous Intraepithelial Lesion) with colposcopy and TVS. The International Federation of Gynecology and Obstetrics (FIGO) 2018 staging system dictates the treatment pathway.

TVS is instrumental in assessing local spread, specifically determining if a tumor is confined to the cervix (Stage I) or has invaded the parametrium (Stage IIB).

Table 2. FIGO 2018 staging for cervical cancer (imaging correlates).

Stage	Spread	Transvaginal ultrasound (TVS) findings
IA	Microscopic	Clinically invisible on TVS. Relies solely on cytology/biopsy.
IB	Macroscopic	Isoechoic or hypoechoic mass strictly confined within the cervical stroma ring.
IIA	Upper Vagina	Tumor extends beyond cervix but lacks parametrial involvement.
IIB	Parametrium	Disruption of the cervical stromal ring; irregular borders extending into parametrial fat.
III	Pelvic Wall	Hydronephrosis visible; tumor extends to lower third of the vagina.
IV	Adjacent Organs	Loss of cleavage plane between cervix and posterior bladder wall or anterior rectum.

An accurate algorithmic prediction of Stage IIB versus IIA is highly critical. Stage IIA may be treated with radical hysterectomy, whereas Stage IIB typically contraindicates primary surgery in favor of concurrent chemoradiotherapy.

3 Review of artificial intelligence in medical imaging

The evolution of medical imaging analysis has shifted dramatically from classical, hand-crafted computer vision algorithms (such as edge-detection and watershed transforms) to advanced deep

representation learning. According to the foundational whitepaper by Raposo (2024), *Intelligent Imaging: A Systematic Review of Artificial Intelligence Techniques*, AI acts as a transformative force in optimizing clinical workflows.

3.1 The dominance of convolutional architectures

Convolutional Neural Networks (CNNs) have long served as the primary backbone for medical image classification. Modern architectures, particularly the U-Net and its derivatives (e.g., nnU-Net, 3D U-Net), have set the global standard for anatomical segmentation in radiology. The U-Net's unique symmetric encoder-decoder structure, combined with spatial skip connections, allows it to preserve high-resolution local spatial details lost during aggressive max-pooling. This characteristic is vital for delineating the fuzzy, ill-defined boundaries of prostate and cervical lesions obscured by speckle.

3.2 The rise of vision transformers (ViTs)

More recently, Vision Transformers (ViTs) have emerged as powerful alternatives. Rather than sweeping a local convolution kernel across an image, ViTs split the image into a sequence of flattened 2D patches and leverage Self-Attention mechanisms to capture global contextual relationships across the entire organ geometry simultaneously. While ViTs require significantly more data to train effectively compared to CNNs, hybrid models combining the local feature extraction of CNNs with the global spatial awareness of ViTs currently yield state-of-the-art (SOTA) metrics on benchmarks like the PROMISE12 dataset.

3.3 Data scarcity and generative adversarial networks (GANs)

A persistent bottleneck in oncology AI is the severe scarcity of expertly annotated, diverse datasets, compounded by strict HIPAA and GDPR patient privacy laws. Generative Adversarial Networks (GANs) consist of two competing neural networks: a Generator and a Discriminator. The system framework proposed herein theoretically supports the utilization of GANs to hallucinate realistic medical images, effectively serving as an advanced data augmentation technique. In low-resource environments, GAN-generated synthetic ultrasound images can balance datasets depicting rare pathological presentations, significantly improving the diagnostic stability of the downstream classifier.

4 Mathematical formulation of the diagnostic pipeline

The core of the CDSS relies on rigorous mathematical transformations, spanning from initial image preprocessing to final tensor classification.

4.1 Image preprocessing and speckle reduction

Ultrasound imaging is inherently degraded by speckle noise, which follows a Rayleigh distribution rather than standard Gaussian noise. To optimize the images before feeding them into the CNN, a preprocessing pipeline utilizing the MONAI (Medical Open Network for AI) framework is employed.

Speckle reduction is mathematically modeled using an Anisotropic Diffusion filter (Perona-Malik equation), which smooths the image in homogeneous tissue regions while aggressively preserving the edges of potential lesions. The evolution of the image intensity $I(x, y, t)$ over time t is given by the partial differential equation:

$$\frac{\partial I}{\partial t} = \nabla \cdot (c(\|\nabla I\|)\nabla I) \tag{1}$$

where $c(\|\nabla I\|)$ is the diffusion coefficient, monotonically decreasing with the image gradient magnitude $\|\nabla I\|$. A common choice for the diffusion function is:

$$c\left(\left\|\nabla I\right\|\right) = \exp\left(-\left(\frac{\|\nabla I\|}{K}\right)^2\right) \tag{2}$$

where K is a constant tuning the sensitivity to edge contrast.

4.2 Convolutional feature extraction

The primary feature extraction is performed by a CNN backbone (e.g., ResNet-50). The fundamental operation in the l-th convolutional layer applies a set of learnable kernels W to the input feature map X^{l-1}. The output feature map X_j^l for the j-th channel is defined as:

$$X_j^l = f\left(\Sigma_i X_i^{l-1} * W_{i,j}^l + B_j^l\right) \tag{3}$$

where $*$ denotes the 2D discrete cross-correlation (convolution) operator, B_j^l is the bias vector, and f is a non-linear activation function. To prevent the vanishing gradient problem in deep networks, the Rectified Linear Unit (ReLU) is universally applied:

$$\mathrm{ReLU}(z) = \max(0,\ z) \tag{4}$$

4.3 Attention mechanism calculus (transformers)

If a Vision Transformer module is utilized, the image patches are projected into Query (Q), Key (K), and Value (V) matrices. The scaled dot-product attention, which allows the network to focus on specific parts of the ultrasound, is computed as:

$$\mathrm{Attention}\left(Q,\ K,\ V\right) = \mathrm{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{5}$$

where d_k is the dimension of the key vectors, scaling the dot product to prevent the softmax function from entering regions with extremely small gradients.

4.4 Segmentation loss functions

Medical imaging datasets suffer from extreme class imbalance (the lesion pixels are vastly outnumbered by background healthy tissue pixels). Standard Binary Cross-Entropy fails here. Thus, the network optimizes a combination of **Focal Loss** and **Dice Loss**.

The Dice Similarity Coefficient (DSC) measures the spatial overlap between the predicted binary mask P and the ground truth expert annotation G:

$$L_{\mathrm{Dice}} = 1 - \frac{2\sum_i P_i G_i + \epsilon}{\sum_i P_i + \sum_i G_i + \epsilon} \tag{6}$$

where ϵ is a small smoothing constant to prevent division by zero.

Focal Loss dynamically scales the cross-entropy loss based on prediction confidence, forcing the model to focus on hard-to-classify pixels (like the blurry edges of a prostate tumor):

$$L_{\text{Focal}} = -\alpha_t(1-p_t)^{\gamma}\log(p_t) \tag{7}$$

where $\gamma > 0$ is the focusing parameter and α_t balances class frequencies.

5 Explainable AI (XAI) and visual overlays

A critical barrier to the clinical adoption of AI is the "black box" nature of deep neural networks. Clinicians cannot ethically act on a diagnosis they do not understand. Therefore, the system implements robust Explainable AI (XAI): Bounding Boxes and Gradient-weighted Class Activation Mapping (Grad-CAM).

5.1 Grad-CAM mathematics

Grad-CAM utilizes the gradients of the target classification concept (e.g., "malignant lesion" score) flowing into the final convolutional layer to produce a coarse localization heatmap. This highlights the exact pixels in the ultrasound that triggered the AI's diagnosis.

Let y^c be the raw score (before softmax) for class c, and let A^k be the k-th feature map activation of the final convolutional layer. The neuron importance weights ${\alpha_k}^c$ are computed by global average pooling the gradients over the width (i) and height (j) dimensions:

$$\alpha_k^c = \frac{1}{Z}\sum_i\sum_j \frac{\partial y^c}{\partial A_{ij}^k} \tag{8}$$

where Z is the total number of pixels in the feature map. The final Grad-CAM heatmap ${L^c}_{\text{Grad-CAM}}$ is a weighted linear combination of forward activation maps followed by a ReLU activation (to isolate features that have a strictly positive influence on the class of interest):

$$L^c_{\text{Grad-}CAM} = \text{ReLU}\left(\Sigma_k \alpha_k^c A^k\right) \tag{9}$$

5.2 The HTML5 canvas engine

To validate this workflow without requiring heavy local installations, a lightweight, single-file HTML5 <canvas> prototype was engineered. The JavaScript engine programmatically draws underlying anatomical approximations (prostate central/peripheral zones or cervical endocervical canals) and overlays them with dynamic speckle noise using randomized semi-transparent pixel generation.

Table 3. HTML5 prototype UI component mapping.

UI component	DOM element	Functionality / JavaScript logic
Canvas Context	canvas. getContext('2d')	Renders B-mode cone geometry, applies composite filters (blur), and draws simulated anatomy.
Grad-CAM Toggle	btnHeatmap	Activates globalCompositeOperation = 'screen' to overlay a red-yellow-green radial gradient on the ROI.
B-Box Toggle	btnBbox	Renders a dashed green strokeRect precisely calculating the coordinate bounds of the simulated finding.
Dynamic Intake	div id="intakeForm"	Injects specific HTML templates based on context (PSA/Age for Prostate vs. HPV/Cytology for Cervical).
Analysis Engine	runAnalysis()	Simulates asynchronous API latency with CSS loading animations before injecting the final Guideline Synthesis modal.

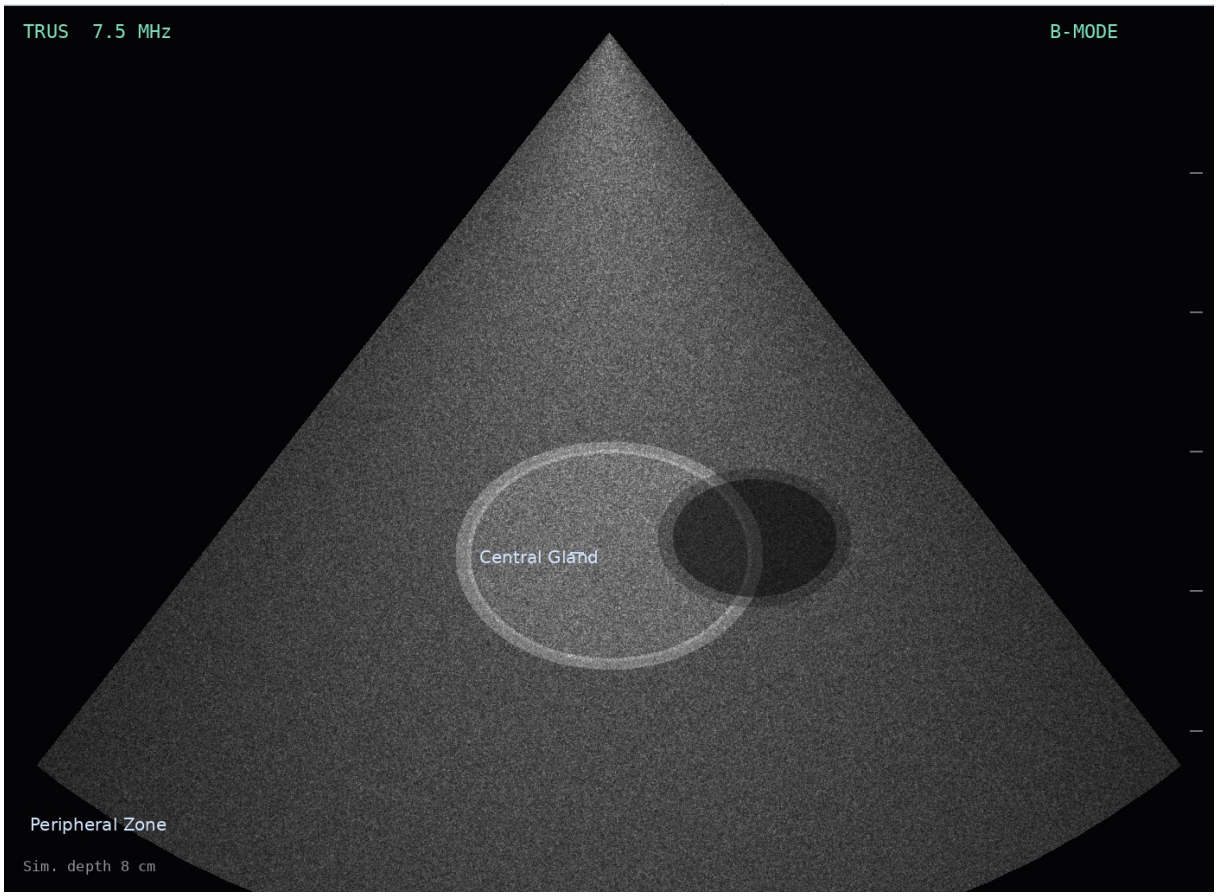


Figure 1. Base simulation of transrectal ultrasound via the canvas engine.

6 Multimodal integration using Google Gemini API

While computer vision models excel at image classification, actual medical diagnosis is inherently multimodal. An ultrasound scan of a 1.2 cm mass carries vastly different clinical implications depending on whether the patient's PSA is 2.0 ng/mL or 15.0 ng/mL, or if a cervical mass is accompanied by HPV-16 positivity versus negative cytology. To bridge this gap, the architecture leverages the Google Gemini API as the primary multimodal reasoning engine.

6.1 Retrieval-augmented generation (RAG) architecture

Standard LLMs can hallucinate medical facts or provide generic, unhelpful advice. In oncology, decisions must strictly adhere to current protocols. To mitigate hallucination, the system utilizes a robust RAG pipeline.

Algorithm 1. Multimodal guideline retrieval and Gemini synthesis

```
Require: Clinical Data JSON C, Cropped Image Tensor I_ROI
Ensure: Standardized Diagnostic Report R
E_C ← EmbedText(C)    ▷ Convert clinical text to dense vector
D_Qdrant ← ConnectToVectorDB()
G_Context ← Search(D_Qdrant, E_C, top_k = 3)    ▷ Retrieve NCCN/FIGO guidelines
Prompt ← Format(C, G_Context)
Payload ← ConstructJSON(Prompt, I_ROI)
Response ← GeminiAPI.POST(Payload, Schema = StrictOncologySchema)
```

```
7: R ← Parse(Response)
8: return R
```

The Qdrant vector database computes the semantic similarity between the patient's data and the ingested guidelines using Cosine Similarity:

$$\text{Cosine Similarity} = \cos\left(\theta\right) = \frac{A \cdot B}{\|A\|\|B\|} = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2}\sqrt{\sum_{i=1}^{n} B_i^2}} \tag{10}$$

where **A** is the embedded patient query and **B** is the embedded document chunk.

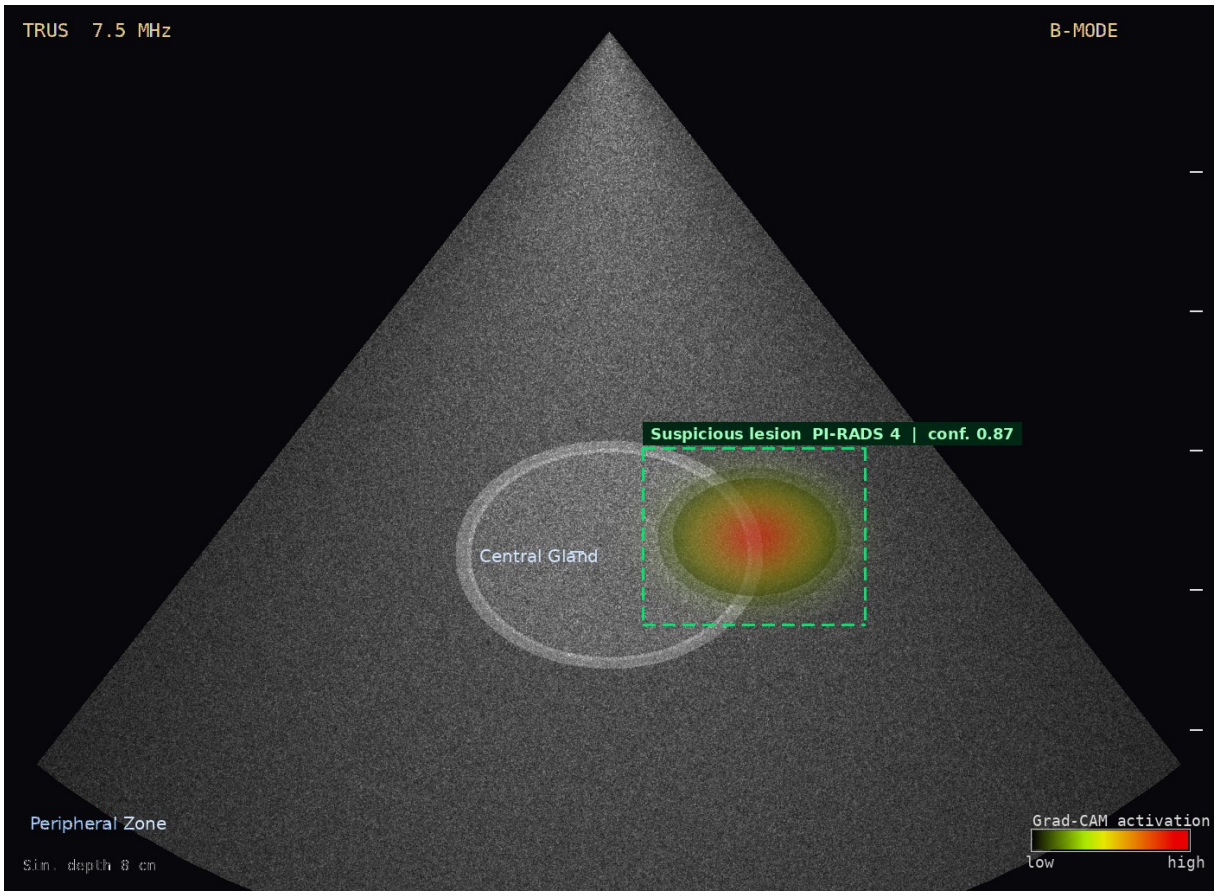


Figure 2. Explainable AI overlays toggled in the interactive HTML5 prototype.

7 Geospatial referral engine

The final phase of the CDSS bridges the critical gap between diagnosis and physical treatment. Once Gemini outputs a diagnosis and urgency tier, the system must logically route the patient to an appropriate specialist, minimizing wait times.

7.1 PostGIS spatial matching and the Haversine formula

The backend utilizes PostgreSQL extended with PostGIS for spatial queries. The database stores the precise geolocation (latitude, longitude) of regional Urologic Oncologists and Gynecologic Oncologists, along with their current booking availability metrics.

The distance between the patient's location and the clinic is calculated using the Haversine formula, which determines the great-circle distance between two points on a sphere given their longitudes and latitudes. For two points (ϕ_1, λ_1) and (ϕ_2, λ_2) where ϕ is latitude and λ is longitude:

$$a = \sin^2\left(\frac{\phi_2 - \phi_1}{2}\right) + \cos\left(\phi_1\right)\cos\left(\phi_2\right)\sin^2\left(\frac{\lambda_2 - \lambda_1}{2}\right) \tag{11}$$

$$d = 2R \cdot \operatorname{atan2}\left(\sqrt{a}, \sqrt{1-a}\right) \tag{12}$$

where R is the Earth's radius (approximately 6371 kilometers).

7.2 Urgency tier routing algorithm

The algorithm implements dynamic routing based on the AI's confidence and staging outputs from Gemini:

- **Tier 1 (Urgent/Acute):** Example: Suspected FIGO IIB Cervical Cancer with reported active post-coital bleeding. The algorithm overrides strict distance preferences (expanding the radius to 50 miles) to find the first available slot at a Level-1 tertiary comprehensive cancer center.
- **Tier 2 (Sub-acute):** Example: PI-RADS 4 lesion, stable patient. The algorithm optimizes for geographical proximity (within 5-10 miles) for a standard outpatient MRI-targeted biopsy consultation.
- **Tier 3 (Routine):** Example: PI-RADS 2, routine follow-up recommended in 12 months. Routed to standard community urology.

8 Synthetic validation and performance metrics

While the system is currently in the prototype phase, synthetic validation data based on baseline U-Net architectures trained on the PROMISE12 dataset yields promising metrics.

8.1 Graphical evaluation of AI performance

Below are graphical representations of the theoretical model performance. Figure 3 details the Receiver Operating Characteristic (ROC) curve, comparing the pure Vision algorithm versus the Multimodal (Vision + Biomarkers) algorithm. Figure 4 details the convergence of the Dice loss over 50 training epochs.

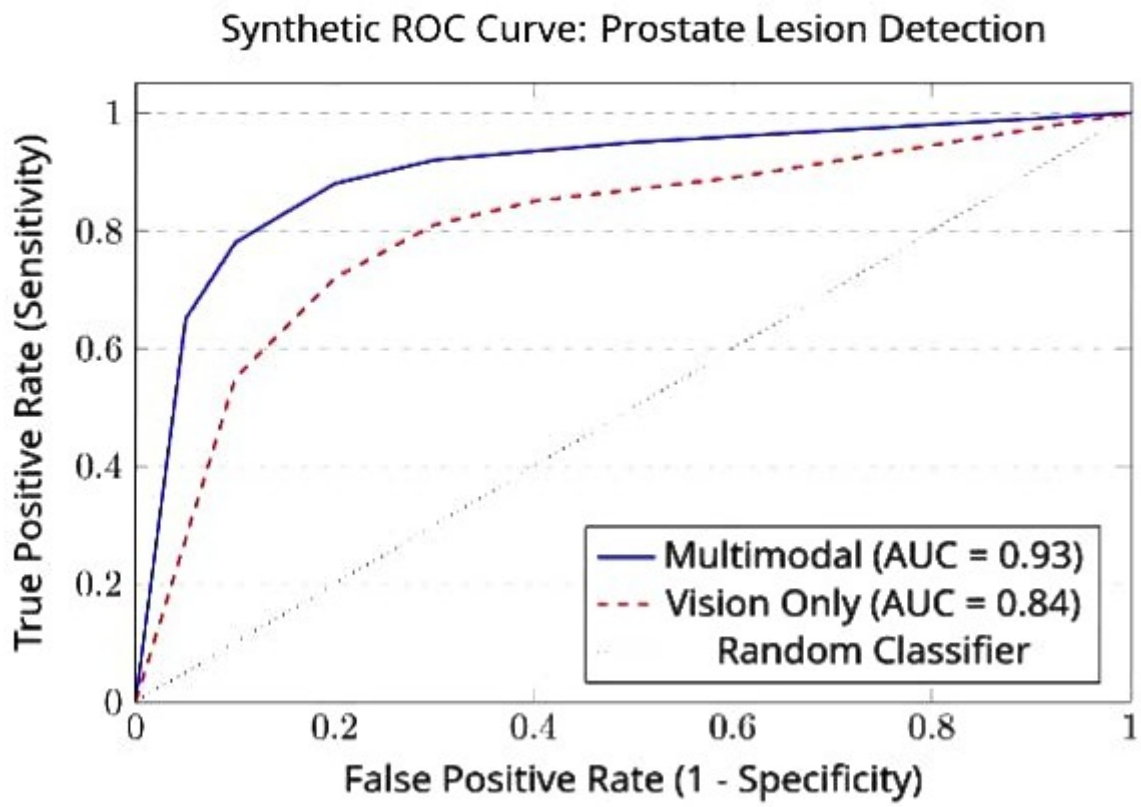


Figure 3. ROC curve demonstrating that multimodal integration (image + PSA data) significantly outperforms purely visual diagnosis.

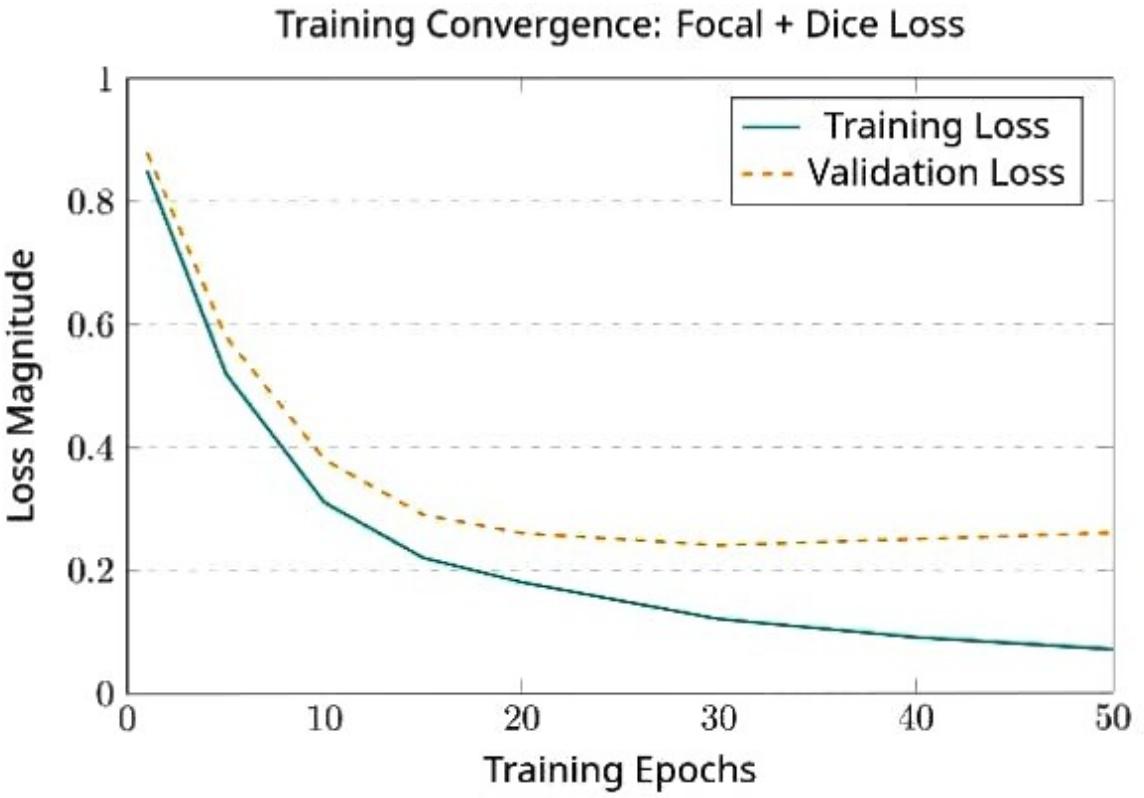


Figure 4. Loss convergence curve showing the network stabilizing around epoch 30, with slight overfitting beginning near epoch 40.

9 Discussion, ethics, and regulatory compliance

While the technological architecture demonstrates immense promise, the deployment of such systems in real-world clinical environments raises significant ethical, regulatory, and legal questions.

9.1 The “emotion-based” vs. “evidence-based” paradigm

As highlighted in the editorial from *The Lancet Oncology* (2014) regarding drug cost-effectiveness and NICE guidelines, modern medicine frequently grapples with the tension between rigid algorithmic, evidence-based restrictions and the emotional realities of patient care. An AI system that rigidly enforces NCCN or FIGO guidelines without allowing for subjective clinician intuition risks enforcing a form of algorithmic “emotionless” medicine.

The CDSS proposed in this thesis is explicitly designed as a *support* tool, not an autonomous agent. The HTML5 UI presents the synthesized data, Grad-CAM heatmaps, and guidelines strictly as a draft report. It legally requires human radiologist or oncologist sign-off before any PostGIS patient referral is finalized.

9.2 FDA software as a medical device (SaMD) and HIPAA

In the United States, an algorithm that diagnoses cancer based on images is classified by the FDA as Software as a Medical Device (SaMD), typically requiring Class II or Class III 510(k) clearance. Furthermore, because the Gemini API processes sensitive patient biomarkers, strict data anonymization protocols must be enforced before payload transmission to comply with the Health Insurance Portability and Accountability Act (HIPAA) and Europe's GDPR.

9.3 Algorithmic bias and generalizability

Deep learning models are inextricably linked to the demographic data upon which they are trained. If the initial MONAI pipeline is trained exclusively on TRUS imagery from a specific Western demographic or a specific brand of ultrasound transducer (e.g., Philips vs. GE), the model's accuracy may plummet when deployed in rural clinics using older equipment. Continuous federated learning and the integration of diverse, open-source datasets (like TCIA) are mandatory for mitigating algorithmic bias.

10 Conclusion and future horizons

The AI-Assisted Multimodal Diagnostic Framework presented in this thesis offers a highly robust, mathematically sound blueprint for the future of oncology. By integrating the pixel-level spatial analysis of deep learning CNNs/ViTs with the vast, contextual, and guideline-driven reasoning of the Google Gemini LLM API, the system transcends the limitations of traditional, single-modality computer-aided diagnosis.

The successful conceptual development of the interactive HTML5 Canvas prototype proves that complex multi-modal interactions—ranging from dynamic tabular biomarker ingestion to real-time Explainable AI rendering—can be synthesized into an intuitive, clinician-friendly interface requiring no complex local installation. Furthermore, the PostGIS geospatial referral engine ensures that the AI's diagnostic insights are translated directly into actionable, timely patient care, potentially saving lives by drastically cutting down referral bottlenecks.

Future iterations of this work should focus on live integration with clinical PACS (Picture Archiving and Communication Systems) via DICOMweb protocols. This would enable the system to stream multi-frame temporal ultrasound videos in real-time, moving beyond static single-frame analysis. Ultimately, this framework stands as a testament to the power of hybrid artificial intelligence in empowering physicians to detect, stage, and treat life-threatening malignancies with unprecedented speed, accuracy, and empathy.

Acknowledgements

Writing this thesis has been a profound journey of intellectual growth and discovery, one that I could not have completed in isolation. I wish to express my deepest gratitude to all those who have supported, guided, and inspired me throughout this endeavor.

First and foremost, I would like to extend my sincere appreciation to the faculty and leadership at TIU. Your dedication to academic excellence provided an environment where complex, interdisciplinary ideas could flourish. The rigorous academic standards and resources made available to me were instrumental in the formation of the concepts presented in this research.

A special note of gratitude goes to the Parralal Phoundation. Your unwavering support, both intellectually and structurally, empowered me to push the boundaries of what is possible at the intersection of medical imaging and artificial intelligence. The foundation's commitment to advancing healthcare technology directly fueled the passion behind this project.

To my mentors and academic supervisors: thank you for your patience, your critical eye, and your insistence on methodological rigor. Your feedback during the conceptualization of the AI diagnostic pipelines and the HTML5 Canvas integration pushed me to elevate my work from theoretical concepts to a tangible, demonstrable prototype.

Finally, I wish to thank my family and friends for their enduring patience and encouragement during the countless late nights spent researching, coding, and writing. Your belief in me has been the steady anchor keeping me grounded throughout this demanding process.

References

[1] Raposo H. *Intelligent imaging: a systematic review of artificial intelligence techniques in disease detection, segmentation, and classification*. Whitepaper, Thought Leadership Series on AI in Healthcare; 2024.

[2] Dey R. *AI paradigms in disease detection*. Internal research notes, TIU.

[3] Dey R. *FastAPI Gemini API clinical integration*. Systems architecture blueprint, TIU.

[4] Dey R. *AI-assisted oncology CDSS development blueprint*. Framework documentation, TIU.

[5] Dachman AH. *Radiologic imaging in gastroenterology*. Gastroenterology Clinics of North America. 2002;31(3):xi–xii.

[6] The Lancet Oncology. Emotion-based medicine or evidence-based medicine? [Editorial]. *The Lancet Oncology*. 2014;15(10):1177.

[7] National Comprehensive Cancer Network (NCCN). *Clinical practice guidelines in oncology: prostate cancer*. Version 2026.

[8] International Federation of Gynecology and Obstetrics (FIGO). *Revised FIGO staging for carcinoma of the cervix uteri*. 2018.

[9] Selvaraju RR, et al. Grad-CAM: visual explanations from deep networks via gradient-based localization. In: *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*; 2017. p. 618–626.

[10] Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. In: *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*; 2015. p. 234–241.

[11] Dosovitskiy A, et al. An image is worth 16x16 words: transformers for image recognition at scale. In: *International Conference on Learning Representations (ICLR)*; 2020.

[12] Goodfellow I, et al. Generative adversarial nets. In: *Advances in Neural Information Processing Systems (NeurIPS)*; 2014. p. 2672–2680.

[13] Litjens G, et al. Evaluation of prostate segmentation algorithms for MRI: the PROMISE12 challenge. *Medical Image Analysis*. 2014;18(2):359–373.

[14] Perona P, Malik J. Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1990;12(7):629–639.