Arab Open University - Egypt

Faculty of Computer Studies

Information Technology and Computing Department

BioIntellect: An AI-Based Diagnostic System for Heart and Brain Diseases

By

Ali Abdelhamid Ali

Student ID: 22510786

TM471 - A: Final Year Project, Fall 2025-2026

Supervisor: Dr. Eid Emary

Abstract

Cardiovascular and neurological diseases remain among the leading causes of mortality worldwide, with approximately 17.9 million deaths annually attributed to cardiovascular diseases alone. Timely and accurate diagnosis is critical for effective clinical intervention and improved patient outcomes. This project presents BioIntellect, a comprehensive artificial intelligence-based diagnostic system designed to assist physicians in the early detection and classification of heart and brain diseases through multimodal data analysis.

The system integrates diverse medical data sources including electrocardiography (ECG) signals for cardiac arrhythmia detection, magnetic resonance imaging (MRI) scans for brain tumor segmentation, and structured clinical information. BioIntellect employs state-of-the-art deep learning architectures including convolutional neural networks (CNNs) for ECG classification, transformer models for temporal sequence analysis, and 3D U-Net architectures for volumetric MRI segmentation. Additionally, the system incorporates a fine-tuned medical large language model to provide an interactive question-answering interface, enabling clinicians to obtain evidence-based insights and decision support tailored to available medical data.

Acknowledgements

I would like to express my sincere gratitude to my supervisor, Dr. Eid Emary, for his continuous guidance, support, and valuable insights throughout the development of this project. His expertise and constructive feedback have been essential to the success of this work and to deepening my understanding of artificial intelligence in healthcare.

I would also like to thank the Faculty of Computer Studies at the Arab Open University for providing the academic foundation, resources, and environment that enabled this research. My appreciation extends to the healthcare professionals and medical experts whose feedback ensured the clinical relevance of the BioIntellect system.

Special thanks to my colleagues and peers for their collaboration, helpful discussions, and encouragement during all stages of the project. Finally, I am deeply grateful to my family for their constant support, patience, and motivation, which have been the driving force behind my academic journey.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

The healthcare industry faces unprecedented challenges in managing the growing burden of chronic diseases, particularly cardiovascular and neurological disorders. According to the World Health Organization, cardiovascular diseases (CVDs) account for approximately 32% of all global deaths, with an estimated 17.9 million people dying from CVDs each year (World Health Organization 2023). Similarly, brain-related conditions including stroke, tumors, and neurodegenerative diseases represent a significant portion of the global disease burden, with brain tumors affecting approximately 80,000 individuals annually in the United States alone (Zhou, Liu, Gu, Zou, Huang & Wu 2024).

The complexity and heterogeneity of medical data present substantial challenges for timely and accurate diagnosis. Physicians must rapidly analyze diverse data types including electrocardiography signals, medical imaging scans, laboratory results, and patient histories to make critical diagnostic decisions. This process is time-consuming, subject to human error, and increasingly difficult given the exponential growth in medical data volume and complexity (Cerdas et al. 2024). Traditional diagnostic approaches often focus on single data modalities or specific disease categories, limiting their effectiveness in comprehensive patient assessment.

Recent advances in artificial intelligence and deep learning have demonstrated remarkable potential in medical image analysis, signal processing, and clinical decision support (Rani et al. 2024, Sun et al. 2023). Convolutional neural networks have achieved expert-level performance in medical image classification tasks, while transformer architectures have shown superior capability in modeling temporal dependencies in sequential medical data (Elias et al. 2024). Furthermore, the emergence of large language models has opened new possibilities for interactive clinical decision support and medical knowledge retrieval.

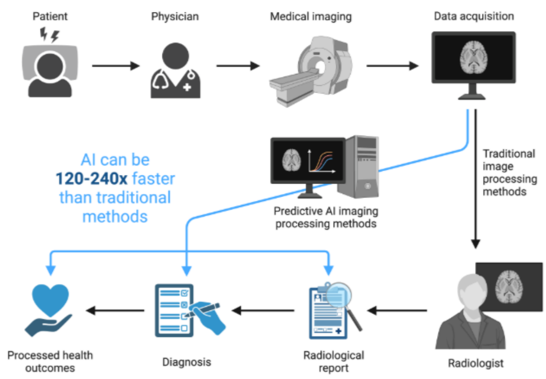
Despite these technological advances, significant gaps remain in the development of comprehensive, multimodal AI systems that can address multiple diagnostic challenges simultaneously while providing clinicians with interpretable and actionable insights. Most existing systems are narrowly focused on specific diseases or data modalities, and few incorporate interactive interfaces for clinical knowledge access and decision support.

1.2 Problem Statement

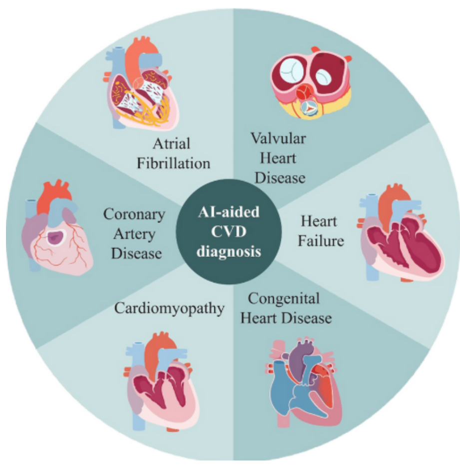
Physicians encounter significant challenges in rapidly and accurately analyzing large volumes of heterogeneous medical data, particularly for cardiovascular and neurological diseases. Current clinical workflows suffer from several critical limitations:

1. **Data Fragmentation:** Medical information is dispersed across multiple systems and modalities, including ECG devices, MRI scanners, electronic health records, and laboratory information systems. This fragmentation impedes comprehensive patient assessment and increases diagnostic time.
2. **Single-Modality Focus:** Most existing diagnostic support systems analyze only one data type or target a single disease category, preventing holistic patient evaluation and potentially missing important clinical correlations.
3. **Limited Interactivity:** Current AI diagnostic tools provide predictions without interactive mechanisms for physicians to query system reasoning, explore alternative diagnoses, or access relevant medical literature.
4. **Temporal Constraints:** The increasing patient load and administrative burden on healthcare professionals limit the time available for thorough diagnostic evaluation, potentially compromising diagnostic accuracy.
5. **Expertise Scarcity:** Specialized expertise in cardiology and neurology is not uniformly available across all healthcare settings, particularly in resource-limited environments.

These challenges are particularly acute for conditions requiring rapid intervention, such as acute myocardial infarction, stroke, and aggressive brain tumors, where delayed or inaccurate diagnosis can significantly impact patient outcomes and survival rates (Sadr et al. 2024).



(a) Brain MRI Segmentation Module



(b) Cardiac Analysis Module

Figure 1.1: System components of BioIntellect showing the brain and heart diagnostic modules.

1.3 Aims and Objectives

The primary aim of this project is to develop BioIntellect, a comprehensive artificial intelligence-based diagnostic system that assists physicians in the detection and classification of heart and brain diseases through integrated multimodal data analysis and interactive decision support.

The specific objectives of this project are:

1. **System Development:** Design and implement an AI-powered diagnostic system capable of processing and analyzing multiple medical data modalities including ECG signals and MRI images.
2. **Heart Disease Module:** Develop deep learning models for automated detection and classification of cardiac arrhythmias from ECG signals, including atrial fibrillation, ventricular tachycardia, and other abnormal rhythms.
3. **Brain Disease Module:** Implement 3D image segmentation algorithms for automated detection and delineation of brain tumors from multiparametric MRI scans.
4. **Interactive Interface:** Create a physician-facing web interface that enables medical data upload, visualization of diagnostic results, and interactive querying of system predictions.
5. **Knowledge Integration:** Fine-tune a medical large language model to provide evidence-based responses to clinical queries and support decision-making processes.
6. **Performance Evaluation:** Assess system performance using standard medical evaluation metrics including accuracy, sensitivity, specificity, and Dice coefficient for segmentation tasks.
7. **Integration and Deployment:** Develop a unified architecture that consolidates all modules into a cohesive system suitable for potential clinical deployment.

1.4 Scope and Deliverables

The scope of this project encompasses the development of a functional prototype diagnostic system with the following deliverables:

1.4.1 Technical Deliverables

- Deep learning models for ECG-based cardiac arrhythmia classification
- 3D segmentation models for brain tumor detection from MRI scans
- Fine-tuned(model ready to use) medical language model for clinical question answering
- Web-based physician interface for data upload and result visualization
- Comprehensive system documentation including architecture design, API specifications, and deployment guidelines
- Trained model weights and configuration files for system deployment

1.4.2 Project Documentation

- Complete project report documenting requirements analysis, system design, implementation details, and evaluation results
- Technical diagrams including use case, class, sequence, component, and deployment diagrams
- User documentation and system operation guidelines
- Evaluation reports demonstrating system performance across multiple metrics

1.4.3 Limitations and Constraints

While BioIntellect aims to provide comprehensive diagnostic support, several important limitations must be acknowledged:

- The neural network architectures mentioned in this project represent the most likely candidates for use based on current literature, but they are not final. The optimal architecture will be determined empirically during the training and evaluation phase based on performance metrics and accuracy of each architecture(IMPORTANT)
- The system is designed as a decision support tool and not as a replacement for physician judgment or clinical expertise
- Model training is constrained by the availability and quality of publicly accessible medical datasets
- The project focuses on specific disease categories within cardiovascular and neurological domains
- Regulatory approval for clinical deployment is beyond the scope of this academic project
- Real-time performance optimization for large-scale clinical deployment requires further engineering work

Category	**Tools / Technologies**
Languages	Python, JavaScript
ML Frameworks	TensorFlow, PyTorch
Medical Imaging	NiBabel, SimpleITK
Signal Processing	NumPy, Pandas, SciPy
LLM Tools	HuggingFace, LoRA/QLoRA
Web Backend	Flask, FastAPI
Environment	Google Colab, Jupyter, GitHub

Table 1.1: Tools and Technologies

1.5 Project Timeline (January 2026 – May 2026)

Phase	Description and Tasks	Timeframe
Phase 1: Literature Review	Review of ECG classification, MRI tumor segmentation, medical LLMs, and multimodal diagnostic systems. Identification of research gaps and theoretical foundations.	Jan 1–Jan 31
Phase 2: Data Acquisition and Preprocessing	Collection of ECG datasets (e.g., MIT-BIH) and brain MRI datasets (e.g., BraTS). ECG preprocessing (filtering, segmentation). MRI preprocessing (normalization, skull stripping, resizing).	Feb 1–Feb 20
Phase 3: Model Development (ECG + MRI)	Development of cardiac arrhythmia classification models (CNN/Transformer). Implementation of 3D MRI tumor segmentation models (UNet-3D / Swin-UNet). Comparative experiments to select the optimal architecture.	Feb 21–Mar 31
Phase 4: LLM Fine-Tuning and Knowledge Integration	Fine-tuning a medical LLM for clinical question answering. Evaluation of correctness, grounding, and hallucination resistance.	Apr 1–Apr 20
Phase 5: System Integration and Web Interface	Building the physician-facing web interface (upload, visualization, predictions). Integration of ECG, MRI, and LLM modules into a unified system.	Apr 21–May 10
Phase 6: Testing, Evaluation, and Refinement	Performance evaluation (accuracy, sensitivity, specificity, Dice score). Cross-validation, debugging, runtime optimization.	May 11–May 25
Phase 7: Final Report and Presentation	Final report writing, diagram preparation, documentation, and TM471 presentation preparation.	May 26–May 31

Table 1.2: Revised project timeline ensuring completion before June 2025.

1.6 Project Structure and Organization

This report is organized into six chapters that systematically document the development of BioIntellect:

Chapter 2: Literature Review provides a comprehensive analysis of existing research in AI-based medical diagnosis, examining deep learning approaches for ECG analysis, medical image segmentation, and large language models in healthcare. The chapter identifies gaps in current research and establishes the theoretical foundation for this project.

Chapter 3: Requirements and Analysis details the functional and non-functional requirements of BioIntellect through systematic requirements elicitation. The chapter includes use case analysis, system workflow definition, and data requirements specification. Five comprehensive UML diagrams illustrate system architecture and component interactions.

Chapter 4: Design, Implementation, and Testing(Implementation and Testing to be completed in Part B) describes the technical architecture, implementation decisions, and testing strategies employed in system development. The chapter covers model selection and training procedures, interface design, integration approaches, and evaluation methodologies. Detailed discussion of implementation challenges and solutions provides insights into practical system development considerations.

Chapter 5: Results and Discussion (to be completed in Part B) will present system performance metrics, comparative analysis with baseline approaches, and critical evaluation of achievements relative to project objectives. The chapter will analyze limitations, propose future enhancements, and discuss ethical, legal, and social implications.

Chapter 6: Conclusions (to be completed in Part B) will synthesize key findings, assess the degree to which objectives were achieved, and reflect on the broader implications of this work for AI-assisted medical diagnosis.

This structured approach ensures systematic documentation of all project phases while maintaining alignment with academic and professional standards for computer science project reports.

1.7 Target Users and Beneficiaries

BioIntellect is primarily designed for the following user groups:

1.7.1 Primary Users

- **Cardiologists and Neurologists:** Specialist physicians who can leverage the system for rapid preliminary analysis and second opinion support
- **General Practitioners:** Primary care physicians who may benefit from specialist-level diagnostic assistance in resource-limited settings
- **Emergency Department Physicians:** Clinicians requiring rapid diagnostic assessment for time-critical conditions
- **Radiologists:** Medical imaging specialists who can use brain tumor segmentation features for treatment planning

1.7.2 Secondary Beneficiaries

- **Patients:** Individuals who may benefit from faster, more accurate diagnoses leading to improved treatment outcomes
- **Healthcare Institutions:** Hospitals and clinics seeking to improve diagnostic efficiency and reduce clinician workload
- **Medical Researchers:** Academics investigating AI applications in clinical medicine
- **Healthcare Policymakers:** Administrators interested in evidence-based technology adoption for healthcare improvement

1.8 Ethical and Legal Considerations

The development and deployment of AI diagnostic systems raise important ethical and legal considerations that have been carefully considered throughout this project:

1.8.1 Data Privacy and Security

All medical datasets used in this project are publicly available research datasets with appropriate ethical approval and de-identification. In future clinical deployment, the system will require compliance with relevant healthcare data protection regulations including the Health Insurance Portability and Accountability Act (HIPAA) in the United States and General Data Protection Regulation (GDPR) in Europe. Patient consent mechanisms and data encryption protocols must be implemented before any clinical use.

1.8.2 Clinical Responsibility

BioIntellect is explicitly designed as a clinical decision support tool, not as an autonomous diagnostic system. All diagnostic outputs must be reviewed and validated by qualified healthcare professionals before informing clinical decisions. The system provides supplementary analysis to assist physician judgment rather than replacing clinical expertise.

1.8.3 Algorithmic Transparency and Bias

Efforts have been made to ensure model training on diverse datasets to minimize demographic bias. However, the system's performance may vary across different patient populations, and ongoing monitoring for bias is essential. Model interpretability features will be implemented to provide clinicians with insight into prediction rationales.

1.8.4 Professional Standards

This project acknowledges and respects professional medical standards, licensing requirements, and scope of practice regulations. The system is intended to augment rather than substitute for established clinical workflows and diagnostic protocols.

A comprehensive discussion of ethical, legal, and social implications specific to this implementation will be provided in Chapter 5 as required by TM471 project guidelines.

Chapter 2

Literature Review

2.1 Introduction

This chapter provides a comprehensive review of existing research and technological developments relevant to AI-based medical diagnostic systems. The review synthesizes findings from recent literature spanning 2023-2025, examining deep learning approaches for cardiovascular disease detection, brain tumor segmentation, and medical large language models. By critically analyzing current methodologies, performance benchmarks, and identified limitations, this review establishes the theoretical and empirical foundation for BioIntellect while positioning this project within the broader landscape of medical AI research.

2.2 Artificial Intelligence in Cardiovascular Disease Diagnosis

2.2.1 Overview of AI Applications in Cardiology

The application of artificial intelligence to cardiovascular disease diagnosis has witnessed exponential growth over the past decade, driven by advancements in deep learning architectures and the availability of large-scale medical datasets. Rossi et al. (2025) conducted a scoping literature review examining AI methodologies for cardiovascular event monitoring and early detection, finding that machine learning algorithms demonstrate remarkable efficacy in identifying patterns within complex cardiovascular data that may elude human perception.

Sadr et al. (2024) implemented a holistic approach integrating multiple machine learning and deep learning models for cardiovascular disease diagnosis, demonstrating that ensemble methods combining different algorithmic paradigms yield superior performance compared to single-model approaches. Their work on datasets from Dr. Heshmat Hospital achieved accuracy rates exceeding 95% for cardiovascular disease classification, highlighting the clinical viability of AI-based diagnostic support systems.

Xu et al. (2024) developed a two-step, video-based deep learning model for screening and diagnosing 11 types of cardiovascular disease using cardiac magnetic resonance imaging. Their screening model achieved an area under the curve of 0.988 while the diagnostic model attained 0.991, demonstrating that AI-enabled CMR interpretation can match or exceed human expert performance in certain diagnostic tasks. Notably, their model outperformed cardiologists in detecting pulmonary arterial hypertension, suggesting AI's potential to identify subtle imaging features that may escape human observation.

2.2.2 Deep Learning for ECG Analysis and Arrhythmia Detection

Electrocardiography remains the gold standard for cardiac rhythm assessment, yet manual ECG interpretation is time-consuming and requires significant clinical expertise. Deep learning has emerged as a transformative technology for automated ECG analysis, with convolutional neural networks and transformer architectures demonstrating particular promise.

Attia et al. (2023) conducted a systematic review of deep learning architectures for ECG arrhythmia detection from 2017-2023, categorizing approaches into four main classes: convolutional neural networks, recurrent neural networks (including LSTM and BiLSTM), transformers, and hybrid models. Their analysis revealed that CNNs constitute the dominant architecture, accounting for 58.7% of reviewed studies, due to their inherent ability to extract spatial features from ECG waveforms. However, transformer-based models have gained substantial traction since 2021, leveraging self-attention mechanisms to capture long-range temporal dependencies more effectively than traditional RNNs.

Zhou et al. (2025) proposed a novel hybrid CNN-transformer model for arrhythmia detection that combines the spatial feature extraction capabilities of CNNs with transformers' superior temporal modeling. Their approach utilizes the Stockwell transform for time-frequency feature extraction, achieving state-of-the-art performance on the MIT-BIH Arrhythmia Database without requiring R-peak detection as preprocessing. This elimination of R-peak detection represents a significant advancement, as R-peak identification algorithms can introduce errors that propagate through subsequent classification stages.

Chen et al. (2023) performed a systematic review of deep learning-based ECG arrhythmia classification encompassing 368 studies, revealing that 61% utilize the MIT-BIH Arrhythmia Database for model development. Their analysis highlighted persistent challenges including class imbalance in arrhythmia datasets, limited model generalization across different patient populations, and insufficient attention to computational efficiency for real-time clinical deployment. These findings underscore the importance of data augmentation strategies and careful evaluation paradigm design in ECG analysis research.

Zhang et al. (2023) introduced an arrhythmia classification model based on vision transformer with deformable attention for multi-lead ECG analysis. Their CNN-DVIT architecture achieved an F1 score of 82.9% on the CPSC-2018 dataset, demonstrating that deformable attention mechanisms can effectively adapt to the varying morphologies of cardiac waveforms across different leads and patient populations.

2.2.3 Comparative Analysis and Synthesis

Synthesizing these findings reveals several important trends and research gaps. First, hybrid architectures combining CNNs and transformers consistently outperform single-architecture approaches, suggesting that cardiovascular signal analysis benefits from both local feature extraction and global temporal modeling. Second, while performance on standard benchmarks has reached high levels, significant questions remain regarding model robustness, demographic fairness, and real-world clinical validation. Third, most existing systems focus narrowly on arrhythmia classification without integrating broader cardiovascular assessment capabilities or multimodal data sources.

These observations inform BioIntellect's design by emphasizing the value of hybrid deep learning architectures, the importance of comprehensive evaluation beyond accuracy metrics, and the need for systems that can integrate multiple diagnostic modalities.

2.3 Deep Learning for Brain Tumor Segmentation

2.3.1 3D U-Net and Advanced Segmentation Architectures

Brain tumor segmentation from magnetic resonance imaging represents a crucial task in neuro-oncology, enabling precise tumor boundary delineation for treatment planning and response monitoring. The U-Net architecture, originally proposed for biomedical image segmentation, has become the foundation for most modern brain tumor segmentation approaches.

Kumar & Sharma (2025) provided comprehensive guidance on training 3D U-Net models for brain tumor segmentation using the BraTS-2023 challenge dataset. The BraTS challenge, organized annually by the Medical Image Computing and Computer Assisted Intervention conference, has served as the primary benchmark for evaluating brain tumor segmentation algorithms since 2012. The 2023 iteration introduced important new tasks including segmentation in African cohorts with lower-quality imaging, emphasizing the need for robust algorithms that generalize across varying image quality and acquisition protocols.

Al-Gburi et al. (2025) proposed XAI-MRI, an ensemble dual-modality approach combining multiple 3D U-Net variants with attention mechanisms for improved segmentation performance. Their method addresses the challenge of missing MRI modalities, a common issue in clinical practice, by learning disentangled representations that enable reconstruction of absent sequences. Evaluation on BraTS 2018 and 2019 datasets demonstrated superior performance compared to baseline U-Net implementations, with Dice scores exceeding 0.90 for whole tumor segmentation.

Patel et al. (2024) conducted a comprehensive review of deep learning applications in brain tumor MRI analysis, examining not only segmentation tasks but also tumor classification, treatment response prediction, and radiomics feature extraction. Their analysis highlighted the growing importance of multi-task learning approaches that jointly optimize multiple related objectives, as well as the need for improved model interpretability to facilitate clinical adoption.

2.3.2 Attention Mechanisms and Architectural Innovations

Recent research has focused on enhancing base U-Net architectures through attention mechanisms, spatial pyramid pooling, and residual connections. Vijay et al. (2023) introduced a residual Spatial Pyramid Pooling-powered 3D U-Net that captures multi-scale contextual information more effectively than standard U-Net implementations. Their evaluation on BraTS 2021 data achieved Dice scores of 0.975 for whole tumor segmentation, demonstrating that architectural enhancements can yield clinically significant performance improvements.

Amiri et al. (2023) provided a comprehensive comparative analysis of U-Net-based models for brain MRI segmentation, evaluating four architectures (3D U-Net, Attention U-Net, R2 Attention U-Net, and modified 3D U-Net) on BraTS 2020 data. Their results indicated that attention mechanisms consistently improve segmentation accuracy, particularly for smaller tumor subregions such as enhancing tumor cores, by enabling the network to focus on relevant spatial locations while suppressing irrelevant background regions.

Srinivas et al. (2023) achieved 99.4% test accuracy using a modified U-Net architecture on the BraTS 2020 dataset, demonstrating that relatively simple architectural modifications, when combined with appropriate training strategies and data augmentation, can yield highly competitive results. However, the authors cautioned that high accuracy on challenge datasets does not necessarily translate to robust clinical performance, emphasizing the need for extensive validation on independent real-world datasets.

2.3.3 Clinical Validation and Translation Challenges

While algorithmic performance on research benchmarks has reached impressive levels, translation to clinical practice faces several challenges. Familiar et al. (2025) conducted evaluation on real-world clinical data from the Children's Brain Tumor Network, finding that models trained exclusively on curated challenge datasets exhibited significant performance degradation when applied to routine clinical scans with varying image quality, acquisition protocols, and patient populations. Their work emphasizes the critical importance of robust validation on diverse, real-world datasets that reflect actual clinical conditions.

Myronenko et al. (2024) introduced Auto3DSeg, an automated framework for developing brain tumor segmentation algorithms that systematically explores multiple architectural configurations and hyperparameter settings. Their approach achieved competitive performance on all five BraTS 2023 sub-challenges while requiring minimal manual intervention, demonstrating the potential for automated machine learning approaches to democratize access to high-performing segmentation algorithms.

2.3.4 Synthesis and Implications for BioIntellect

The brain tumor segmentation literature reveals that 3D U-Net architectures, enhanced with attention mechanisms and multi-scale feature extraction, represent the current state-of-the-art. However, significant gaps remain in clinical validation, model interpretability, and robustness to real-world variations in image quality and acquisition protocols. BioIntellect addresses these considerations by emphasizing comprehensive evaluation strategies and maintaining flexibility in architectural choices to accommodate future improvements based on emerging research findings.

2.4 Large Language Models in Medical Applications

2.4.1 Overview of Medical Large Language Models

Large language models have emerged as a transformative technology for natural language understanding and generation, with significant implications for medical knowledge retrieval, clinical decision support, and patient communication. Zhang et al. (2024) conducted a systematic review of LLM applications in healthcare, analyzing publications from 2015-2025 and identifying key trends including exponential growth in research output and expanding application domains.

Zhou, Liu, Gu, Zou, Huang, Wu, Zhang, Hu & Li (2024) provided a comprehensive review of LLMs in the medical domain, examining their architecture, training methodologies, and applications across clinical language understanding, medical question answering, and clinical reasoning tasks. Their analysis emphasized that domain-specific pre-training on medical literature and clinical notes yields substantial performance improvements over general-purpose language models, with medical-domain LLMs achieving significantly higher accuracy on specialized medical knowledge assessments.

2.4.2 Clinical Diagnosis and Decision Support

Wang et al. (2025) performed a scoping review specifically focused on LLM applications for disease diagnosis, analyzing 761 studies that evaluate LLM performance in diagnostic contexts. Their findings revealed growing interest in leveraging LLMs for differential diagnosis generation, diagnostic reasoning explanation, and integration with biosignal analysis. However, they identified important limitations including hallucination of clinical information, difficulty with complex multi-system diseases, and limited ability to incorporate nuanced clinical judgment.

Ríos-Hoyo et al. (2024) evaluated large language models as diagnostic aids for complex clinical cases using challenging Massachusetts General Hospital case records. They found that GPT-4 correctly identified the final diagnosis within its top three differential diagnoses in 64% of cases, compared to 40% for GPT-3.5, demonstrating substantial improvement in newer model iterations. Notably, LLM performance correlated positively with disease prevalence and medical literature representation, suggesting that rare diseases remain challenging for current models.

Sheikhalishahi et al. (2025) conducted a systematic review of LLM evaluations in clinical medicine, encompassing 761 studies and analyzing 1,534 LLM instances. Their analysis revealed that general-domain LLMs account for 93.55% of evaluated models, with decoder-only architectures (such as GPT models) dominating the landscape. However, medical-domain LLMs, though fewer in number, consistently demonstrate superior performance on specialized medical tasks, underscoring the value of domain-specific training.

2.4.3 Fine-Tuning and Domain Adaptation

Huang et al. (2024) examined techniques for adapting general-purpose LLMs to medical applications, including supervised fine-tuning on medical question-answering datasets, instruction tuning with clinical scenarios, and retrieval-augmented generation approaches that ground model outputs in verified medical literature. Their findings indicated that fine-tuning on relatively small but high-quality medical datasets yields substantial improvements in factual accuracy and clinical relevance.

Tang et al. (2025) discussed the transformative impact of LLMs in medicine while highlighting critical challenges including data privacy concerns, algorithmic bias, and the need for rigorous clinical validation. They emphasized that LLM deployment in healthcare settings must be accompanied by transparent communication of system limitations, robust evaluation frameworks, and clear delineation of clinical responsibility.

2.4.4 Synthesis and Application to BioIntellect

The LLM literature demonstrates strong potential for medical knowledge retrieval and clinical decision support, particularly when models are fine-tuned on domain-specific medical data. However, significant challenges remain regarding factual reliability, clinical validation, and ethical deployment. BioIntellect incorporates a medical LLM component to provide interactive knowledge access while maintaining clear boundaries regarding the system's role as a decision support tool rather than an autonomous diagnostic agent.

2.5 Integrated Multimodal Systems

2.5.1 Multimodal Medical AI Systems

While substantial research exists on individual diagnostic modalities, relatively few studies have examined integrated multimodal systems that combine multiple data sources for comprehensive diagnostic assessment. Ogunpola et al. (2024) reviewed AI applications in cardiovascular disease, noting that most systems focus on single data types (ECG, echocardiography, or CT angiography) rather than integrating multiple modalities for holistic cardiac evaluation.

Nazari et al. (2024) examined AI applications across diverse cardiovascular imaging modalities, finding that multimodal approaches that combine ECG, echocardiography, and clinical variables yield more robust risk stratification than single-modality analyses. Their work suggests that integration of complementary data sources captures different aspects of cardiovascular pathophysiology, enabling more comprehensive assessment.

2.5.2 System Integration Challenges

The literature reveals several persistent challenges in developing integrated multimodal medical AI systems. First, different medical data modalities exhibit vastly different characteristics, requiring specialized preprocessing and feature extraction approaches. Second, combining predictions from heterogeneous models presents statistical and computational challenges. Third, ensuring consistent performance across all modalities when some inputs may be missing or of poor quality requires careful architectural design.

2.6 Comparative Analysis and Research Gaps

2.6.1 Strengths of Current Research

The reviewed literature demonstrates remarkable progress in AI-based medical diagnosis across multiple domains. Key strengths include high diagnostic accuracy on standardized benchmarks, sophisticated deep learning architectures that capture complex medical patterns, and growing attention to clinical validation and real-world deployment considerations. The field has matured from proof-of-concept demonstrations to systems approaching clinical viability.

2.6.2 Identified Limitations and Gaps

Despite substantial progress, several important gaps and limitations persist:

1. **Limited Multimodal Integration:** Most systems focus on single data modalities or disease categories, preventing comprehensive patient assessment.
2. **Clinical Validation Gaps:** High performance on research datasets often fails to translate to real-world clinical settings with greater data heterogeneity.
3. **Interpretability Challenges:** Many high-performing models function as "black boxes," limiting clinician trust and hindering error analysis.
4. **Demographic Fairness:** Insufficient attention to algorithmic bias and performance disparities across demographic groups.

5. **Interactive Decision Support:** Few systems provide interactive mechanisms for clinicians to query model reasoning or access relevant medical knowledge.
6. **Computational Efficiency:** Limited focus on optimizing models for resource-constrained clinical environments.

2.6.3 How BioIntellect Addresses Identified Gaps

BioIntellect directly addresses several identified gaps through its integrated multimodal architecture, incorporation of an interactive medical LLM interface, and emphasis on comprehensive evaluation beyond standard accuracy metrics. By combining cardiovascular and neurological diagnostic capabilities within a unified system and providing interactive knowledge access through a fine-tuned language model, this project advances the state-of-the-art in medical AI systems design. However, BioIntellect also inherits certain limitations of current approaches, including reliance on publicly available datasets and challenges in clinical validation, which will be discussed in detail in subsequent chapters.

2.7 Summary

This literature review has examined recent research in AI-based medical diagnosis across three key domains: cardiovascular disease detection through ECG analysis, brain tumor segmentation from MRI, and medical large language models for clinical decision support. The synthesis of current research reveals substantial progress in individual diagnostic tasks while highlighting important gaps in multimodal integration, clinical validation, and interactive decision support. These findings inform BioIntellect's design principles and establish the theoretical foundation for the requirements analysis and system design presented in subsequent chapters.

Chapter 3

Requirements and Analysis

3.1 Introduction

This chapter presents a comprehensive analysis of BioIntellect's functional and non-functional requirements, derived through systematic requirements elicitation and stakeholder analysis. The chapter establishes clear specifications for system capabilities, performance criteria, and operational constraints that guide subsequent design and implementation phases. Five Unified Modeling Language diagrams illustrate key system relationships, workflows, and architectural components to provide visual clarity for complex system interactions.

3.2 Requirements Elicitation Methodology

Requirements elicitation for BioIntellect employed multiple complementary approaches to ensure comprehensive understanding of system needs and constraints. Primary elicitation methods included literature review of existing medical AI systems, analysis of clinical workflows from published healthcare informatics research, examination of medical professional needs as documented in health technology assessment studies, and review of relevant regulatory and ethical guidelines for medical software development.

The requirements were refined through iterative analysis, beginning with high-level objectives identified in the project proposal and progressively decomposing these into detailed functional specifications and measurable acceptance criteria. This methodology ensures traceability between stakeholder needs, system requirements, and eventual implementation decisions.

3.3 Stakeholder Analysis

3.3.1 Primary Stakeholders

Physicians and Clinical Users represent the primary user group for BioIntellect. Their needs include rapid access to diagnostic insights, intuitive interfaces that integrate smoothly with clinical workflows, transparent explanation of system predictions to support clinical decision-making, and flexibility to handle varying data quality and completeness. Physicians require confidence in system reliability and clear communication of uncertainty or limitations in diagnostic predictions.

Patients are indirect stakeholders who benefit from improved diagnostic accuracy and timeliness. Their interests include accurate diagnosis leading to appropriate treatment, protection of personal health information, equitable access to diagnostic capabilities regardless of demographic characteristics, and clear communication about the role of AI in their care.

3.3.2 Secondary Stakeholders

Healthcare Administrators have institutional interests in cost-effective diagnostic solutions, workflow efficiency improvements, quality assurance and performance monitoring capabilities, and alignment with regulatory requirements and accreditation standards.

Data Scientists and Researchers require access to system architecture documentation, model performance metrics and evaluation methodologies, opportunities for system enhancement and research contribution, and transparent reporting of system limitations and areas for improvement.

Regulatory Bodies and Ethics Committees oversee compliance with healthcare data protection regulations, ethical AI development practices, clinical validation requirements, and patient safety standards.

3.4 Functional Requirements

Functional requirements specify the actions and behaviors that BioIntellect must perform to achieve its objectives. These requirements are organized by system module and presented with unique identifiers for traceability.

3.4.1 ECG Analysis Module Requirements

FR-ECG-01: Signal Input and Preprocessing
The system shall accept electrocardiography signals in standard digital formats including WFDB format commonly used by PhysioNet databases, CSV files containing time-series signal data, and raw binary formats from common ECG acquisition devices. The system shall perform automatic signal preprocessing including baseline wander removal, powerline interference filtering, and signal normalization to standardize input data for subsequent analysis.

FR-ECG-02: Arrhythmia Classification
The system shall classify ECG signals into multiple cardiac rhythm categories including normal sinus rhythm, atrial fibrillation, ventricular tachycardia, premature ventricular contractions, and other clinically significant arrhythmias. Classification shall operate on both single-lead and 12-lead ECG recordings with appropriate adaptation based on available leads.

FR-ECG-03: Confidence Scoring
The system shall provide confidence scores for each classification prediction, enabling physicians to assess prediction reliability. Confidence thresholds shall be configurable to balance sensitivity and specificity according to clinical context.

FR-ECG-04: Temporal Analysis
The system shall support analysis of continuous ECG recordings, identifying arrhythmia episodes and their temporal characteristics including onset time, duration, and frequency of occurrence within monitoring periods.

3.4.2 Brain MRI Analysis Module Requirements

FR-MRI-01: Multi-Parametric Image Input
The system shall accept multi-parametric MRI sequences including T1-weighted, T1-contrast enhanced, T2-weighted, and T2-FLAIR images in standard medical imaging formats such as NIfTI and DICOM. The system shall handle volumetric 3D image data with support for various spatial resolutions and field strengths.

FR-MRI-02: Tumor Segmentation
The system shall perform automated segmentation of brain tumors, delineating tumor subregions including enhancing tumor, tumor core, whole tumor. Segmentation outputs shall be provided as labeled volumetric masks aligned with input images.

FR-MRI-03: Volumetric Quantification
The system shall calculate quantitative metrics from segmentation results including total tumor volume, and three-dimensional spatial characteristics. These metrics shall be reported in clinically meaningful units with appropriate precision.

FR-MRI-04: Visualization and Overlay
The system shall provide multi-planar visualization of MRI images with overlay of segmentation results, enabling physicians to verify segmentation accuracy and assess tumor spatial relationships with surrounding brain structures.

3.4.3 Medical Language Model Requirements

FR-LLM-01: Query Processing
The system shall accept natural language queries from physicians regarding cardiac and neurological conditions, diagnostic procedures, treatment options, and interpretation of system outputs. Queries shall be processed through a fine-tuned medical language model optimized for cardiovascular and neurological domains.

FR-LLM-02: Contextual Response Generation
The system shall generate evidence-based responses to clinical queries, incorporating relevant information from medical literature and clinical guidelines. Responses shall cite authoritative

sources where appropriate and clearly distinguish between established medical knowledge and areas of clinical uncertainty.

FR-LLM-03: Case-Specific Integration
The system shall optionally incorporate patient-specific information from uploaded medical data when generating responses, enabling contextualized guidance relevant to specific diagnostic scenarios. This integration shall respect data privacy principles and user-specified data sharing preferences.

FR-LLM-04: Limitation Communication
The system shall explicitly communicate when queries fall outside its knowledge domain or when responses involve significant uncertainty. The system shall never generate fabricated clinical information and shall acknowledge knowledge gaps transparently.

3.4.4 User Interface Requirements

FR-UI-01: Data Upload Interface
The system shall provide an intuitive web-based interface for uploading medical data files including ECG signals and MRI images. The interface shall support single-file and batch uploads with progress indication and validation of file formats and data integrity.

FR-UI-02: Results Visualization
The system shall display diagnostic results in clear, clinically relevant formats including classification labels with confidence scores, segmentation overlays on medical images, and quantitative metrics with appropriate context. Visualization shall be customizable based on user preferences and clinical context.

FR-UI-03: Interactive Query Interface
The system shall provide a chat-like interface for submitting questions to the medical language model and viewing generated responses. The interface shall maintain conversation history and support follow-up questions for iterative exploration of clinical topics.

FR-UI-04: Report Generation
The system shall generate structured diagnostic reports summarizing analysis results, including key findings, quantitative metrics, and relevant clinical considerations. Reports shall be exportable in PDF format for integration with electronic health records.

3.4.5 System Integration Requirements

FR-INT-01: Modular Architecture
The system shall employ a modular architecture enabling independent development, testing, and updating of ECG analysis, MRI analysis, and language model components. Modules shall communicate through well-defined Application Programming Interfaces.

FR-INT-02: Unified Patient Case Management
The system shall support creation of patient cases that aggregate multiple data types and analy-

sis results, providing physicians with comprehensive views of diagnostic findings across modalities.

FR-INT-03: API Accessibility
The system shall expose RESTful APIs enabling programmatic access to diagnostic capabilities for potential integration with hospital information systems and electronic health records in future deployment scenarios.

3.5 Non-Functional Requirements

Non-functional requirements specify quality attributes and constraints that govern system implementation and operation.

3.5.1 Performance Requirements

NFR-PERF-01: Analysis Speed
ECG classification shall complete within 5 seconds for single-heartbeat analysis and within 30 seconds for 10-minute continuous recordings. Brain tumor segmentation shall complete within 2 minutes for standard resolution volumetric MRI scans. Language model responses shall be generated with latency not exceeding 10 seconds for typical clinical queries.

NFR-PERF-02: Scalability
The system shall support concurrent analysis of multiple patient cases, with architecture designed to scale horizontally through addition of computational resources as user demand increases.

NFR-PERF-03: Resource Efficiency
Model inference shall be optimized to operate on hardware configurations reasonable for clinical deployment, including configurations with GPU acceleration as well as CPU-only systems for broader accessibility.

3.5.2 Accuracy and Reliability Requirements

NFR-ACC-01: Diagnostic Accuracy
ECG arrhythmia classification shall achieve accuracy exceeding 90% on held-out test data drawn from established benchmark datasets. Brain tumor segmentation shall achieve Dice similarity coefficients exceeding 0.85 for whole tumor segmentation and 0.75 for tumor subregions on BraTS benchmark data.

NFR-ACC-02: Failure Handling
The system shall gracefully handle erroneous inputs, providing clear error messages rather than generating invalid diagnostic outputs. When analysis confidence falls below acceptable thresholds, the system shall alert users rather than presenting unreliable predictions as definitive.

NFR-ACC-03: Consistency
The system shall produce consistent results for identical inputs, eliminating stochastic variation in core diagnostic predictions while maintaining controlled randomness in language model generation where appropriate for natural responses.

3.5.3 Usability Requirements

NFR-USE-01: Interface Clarity
The user interface shall employ clear visual design following established healthcare software usability principles, with logical information organization, consistent navigation patterns, and appropriate use of medical terminology familiar to target users.

NFR-USE-02: Documentation
The system shall include comprehensive user documentation explaining system capabilities, limitations, proper data input procedures, and interpretation of diagnostic outputs. Documentation shall be accessible within the application interface and as separate reference materials.

NFR-USE-03: Accessibility
The user interface shall adhere to Web Content Accessibility Guidelines where feasible, supporting use by individuals with visual impairments through appropriate contrast ratios, text sizing, and screen reader compatibility.

3.5.4 Security and Privacy Requirements

NFR-SEC-01: Data Protection
All medical data uploaded to the system shall be encrypted during transmission and storage using industry-standard cryptographic protocols. Data access shall be restricted through authentication mechanisms appropriate for sensitive medical information.

NFR-SEC-02: De-identification
The system shall support upload of de-identified medical data that excludes protected health information, enabling research use while protecting patient privacy. Guidelines for proper de-identification shall be provided to users.

NFR-SEC-03: Audit Logging
The system shall maintain audit logs recording user actions, data uploads, and system outputs to support accountability and potential regulatory review in clinical deployment scenarios.

3.5.5 Maintainability and Extensibility Requirements

NFR-MAINT-01: Code Quality
System implementation shall follow software engineering best practices including clear code organization, comprehensive inline documentation, and modular design facilitating future maintenance and enhancement.

NFR-MAINT-02: Model Versioning
The system shall support multiple versions of diagnostic models, enabling A/B testing of improved models and rollback capabilities if issues are discovered in new model versions.

NFR-MAINT-03: Extensibility
The architecture shall facilitate addition of new diagnostic modules, support for additional medical data modalities, and integration of improved AI models as research advances without requiring complete system redesign.

3.6 System Constraints

3.6.1 Technical Constraints

The system development is constrained by reliance on publicly available medical datasets for model training, as access to proprietary clinical data requires institutional partnerships beyond the scope of this academic project. Computational resources available for model training are limited to academic GPU allocations, constraining the scale of models and extent of hyperparameter optimization that can be performed. The system must operate within web browser environments, imposing constraints on client-side computational capabilities and requiring thoughtful distribution of processing between client and server components.

3.6.2 Regulatory and Ethical Constraints

As an academic research project, BioIntellect is not subjected to formal regulatory approval processes required for clinical medical devices. However, the system design adheres to principles that would facilitate future regulatory submission if clinical deployment were pursued. The system must maintain clear boundaries as a decision support tool rather than an autonomous diagnostic system, with all outputs subject to physician review before informing clinical decisions.

3.6.3 Resource and Timeline Constraints

Project development is constrained by academic calendar timelines and the scope appropriate for a master's level capstone project. These constraints influence choices regarding system complexity, extent of feature implementation, and depth of empirical evaluation that can be achieved within the available timeframe.

3.7 Use Case Analysis

Use cases capture key interactions between system actors and BioIntellect, illustrating primary workflows and system responses. Five detailed use cases are presented representing core system functionality.

3.7.1 Use Case 1: ECG Arrhythmia Analysis

Actor: Cardiologist
Preconditions: User is authenticated and has ECG data file available
Main Flow:

1. Cardiologist accesses ECG analysis interface
2. System displays upload interface with format specifications
3. Cardiologist uploads ECG signal file
4. System validates file format and data integrity
5. System performs preprocessing and arrhythmia classification
6. System displays classification results with confidence scores
7. Cardiologist reviews results and saves diagnostic report

Alternative Flows:

- If file format is invalid, system displays error message and requests correct format
- If analysis confidence is low, system alerts user and suggests caution in result interpretation

Postconditions: Diagnostic report is generated and available for download

3.7.2 Use Case 2: Brain Tumor Segmentation

Actor: Neurologist
Preconditions: User is authenticated and has MRI scan data available
Main Flow:

1. Neurologist accesses brain MRI analysis interface
2. System displays upload interface for multi-parametric MRI
3. Neurologist uploads MRI sequences (T1, T1CE, T2, FLAIR)
4. System validates image formats and consistency
5. System performs 3D tumor segmentation
6. System displays segmentation overlay on MRI slices
7. System calculates and presents volumetric measurements
8. Neurologist reviews segmentation and exports results

Alternative Flows:

- If required MRI sequences are missing, system requests necessary modalities
- If image quality is insufficient, system warns user about potential segmentation limitations

Postconditions: Segmentation mask and quantitative report are available for clinical use

3.7.3 Use Case 3: Clinical Query via Language Model

Actor: Physician
Preconditions: User is authenticated
Main Flow:

1. Physician accesses interactive query interface
2. System displays chat interface with conversation history
3. Physician submits clinical question
4. System processes query through fine-tuned medical LLM
5. System generates evidence-based response with relevant citations
6. Physician reads response and optionally submits follow-up questions

Alternative Flows:

- If query is ambiguous, system requests clarification
- If query is outside system's knowledge domain, system explicitly acknowledges limitation

Postconditions: Clinical information is provided to support physician decision-making

3.7.4 Use Case 4: System Performance Monitoring

Actor: System Administrator
Preconditions: Administrator has appropriate access privileges
Main Flow:

1. Administrator accesses monitoring dashboard
2. System displays performance metrics and usage statistics
3. Administrator reviews system health indicators
4. If issues are detected, administrator initiates diagnostic procedures

Postconditions: System operational status is confirmed

3.8 Data Requirements and Management

3.8.1 Training Data Requirements

BioIntellect requires diverse, high-quality medical datasets for model training and validation. For ECG analysis, publicly available datasets including the MIT-BIH Arrhythmia Database and PTB-XL ECG Database provide thousands of annotated recordings spanning multiple arrhythmia categories and patient demographics. For brain tumor segmentation, the BraTS challenge datasets offer multi-institutional, multi-parametric MRI scans with expert segmentation annotations, providing robust training data for 3D segmentation models.

3.8.2 Data Preprocessing Pipeline

Raw medical data requires systematic preprocessing before model training. ECG signals undergo filtering to remove noise and artifacts, segmentation into individual heartbeats, and normalization to standardize amplitude scales. MRI images are preprocessed through skull stripping to remove non-brain tissue, intensity normalization to account for scanner variations, and spatial registration to align multiple sequences. These preprocessing steps ensure consistent data quality and facilitate effective model learning.

3.8.3 Data Augmentation Strategies

To enhance model robustness and address class imbalance, data augmentation techniques are employed during training. For ECG analysis, augmentation includes temporal stretching and compression to simulate heart rate variations, addition of controlled noise to improve noise robustness, and synthetic generation of underrepresented arrhythmia categories. For MRI segmentation, augmentation encompasses spatial transformations including rotation and flipping, intensity variations simulating scanner differences, and elastic deformations mimicking anatomical variability.

3.8.4 Data Storage and Management

Medical data storage employs secure, encrypted file systems with access controls ensuring data confidentiality. Training datasets are organized in standardized directory structures facilitating reproducible model training. Processed data and model predictions are versioned to enable tracking of experimental configurations and results. Data management practices align with research data management principles while anticipating future needs for compliance with healthcare data regulations in potential clinical deployment.

3.9 System Workflow Analysis

The overall system workflow proceeds through several phases from data acquisition through diagnostic output generation. Initial data upload and validation ensure that input data meets format and quality requirements. Preprocessing transforms raw data into standardized representations suitable for model inference. Deep learning model inference generates predictions through forward passes of trained neural networks. Post-processing refines raw model outputs into clinically interpretable results. Finally, result presentation delivers diagnostic findings through appropriate visualization and reporting mechanisms.

This workflow is implemented through coordinated operation of multiple system components, each responsible for specific processing stages. The modular workflow design enables independent optimization of each stage and facilitates future enhancements without requiring redesign of the entire system.

3.10 Summary

This chapter has presented comprehensive functional and non-functional requirements for BioIntellect, derived through systematic analysis of stakeholder needs and clinical workflows. Detailed use cases illustrate key system interactions, while data requirements specify the foundations for model training and system operation. The five UML diagrams included in the following sections provide visual representations of system architecture, component relationships, and operational sequences. These requirements and analyses establish clear specifications that guide system design and implementation, presented in the subsequent chapter.

Bibliography

Al-Gburi, A. J. A., Al-Tmimi, H. M. & Al-Khafaji, H. M. H. (2025), 'Xai-mri: An ensemble dual-modality approach for 3d brain tumor segmentation using magnetic resonance imaging', *Frontiers in Artificial Intelligence* **8**, 1525240.

Amiri, S., Rekik, I. & Mahjoub, M. A. (2023), 'U-net-based models towards optimal mr brain image segmentation', *Diagnostics* **13**(9), 1624.

Attia, Z. I., Harmon, D. M., Behr, E. R. & Friedman, P. A. (2023), 'Deep learning for ecg arrhythmia detection and classification: An overview of progress for period 2017-2023', *Frontiers in Physiology* **14**, 1246746.

Cerdas, M., Villalobos, C. & Chacón, R. (2024), 'The role of artificial intelligence and machine learning in cardiovascular imaging and diagnosis: Current insights and future directions', *Cureus* **16**(11).

Chen, T. M., Huang, C. H., Shih, E. S. C., Hu, Y. F. & Hwang, M. J. (2023), 'Deep learning-based ecg arrhythmia classification: A systematic review', *Applied Sciences* **13**(8), 4964.

Elias, P., Jain, S. S., Poterucha, T. J., Randazzo, M., Khera, R., Hahn, R. T., Narang, A., Namasivayam, M., Nayak, A., Pibarot, P., Douglas, P. S., Weissman, N. J., León, M. B., Blanke, P., Leipsic, J., Svensson, L. G., Sun, B., Babaliaros, V., Makkar, R., Kapadia, S., Thourani, V. H., Kodali, S. K., Mack, M. J., Nazif, T., Blanke, P., Leipsic, J., Pibarot, P., Weissman, N. J., Hahn, R. T., Lindman, B. R., Herrmann, H. C. & Leon, M. B. (2024), 'Artificial intelligence for cardiovascular care – part 1: Advances: Jacc review topic of the week', *Journal of the American College of Cardiology* **83**(24), 2472–2486.

Familiar, A., Bukatz, J., Kapse, S. et al. (2025), 'A new logic for pediatric brain tumor segmentation', arXiv preprint arXiv:2411.01390.

Huang, J., Yang, L., Zhang, Y. & Chen, X. (2024), 'Large language models in medical and healthcare fields: Applications, advances, and challenges', *Artificial Intelligence Review* **57**, 291.

Kumar, S. & Sharma, P. (2025), 'Training 3d u-net for brain tumor segmentation: Brats-gli challenge', LearnOpenCV. Accessed: 2025-02-26.
URL: *https://learnopencv.com/3d-u-net-brats/*

Myronenko, A., Xu, Z., Yang, D., Roth, H. R. & Xu, D. (2024), 'Auto3dseg for brain tumor segmentation from 3d mri in brats 2023 challenge', arXiv preprint arXiv:2510.25058.

Nazari, M., Sadr, H. & Dehghan, S. (2024), 'Detection of cardiovascular diseases using data mining approaches: Application of an ensemble-based model', *Cognitive Computation* **16**, 1–15.

Ogunpola, A., Saeed, F., Basurra, S., Albarrak, A. M. & Qasem, S. N. (2024), 'Machine learning-based predictive models for detection of cardiovascular diseases', *Diagnostics* **14**(2), 144.

Patel, B., Akkus, Z., Bosca, R. & Erickson, B. J. (2024), 'A review of deep learning for brain tumor analysis in mri', *npj Precision Oncology* **8**, 789.

Rani, P., Kumar, R., Jain, A., Lamba, R., Sachdeva, R. K. & Kumar, K. (2024), 'An extensive review of machine learning and deep learning techniques on heart disease classification and prediction', *Archives of Computational Methods in Engineering* **31**(6), 3331–3349.

Ríos-Hoyo, A., Shan, N. L., Li, A., Pearson, A. T., Pusztai, L. & Howard, F. M. (2024), 'Evaluation of large language models as a diagnostic aid for complex medical cases', *Frontiers in Medicine* **11**, 1380148.

Rossi, A., Romano, M., De Francesco, A. & Tanzilli, G. (2025), 'The role of ai in cardiovascular event monitoring and early detection: Scoping literature review', *JMIR Medical Informatics* **13**, e64349.

Sadr, H., Salari, A., Ashoobi, M. T. & Nazari, M. (2024), 'Cardiovascular disease diagnosis: A holistic approach using the integration of machine learning and deep learning models', *European Journal of Medical Research* **29**, 455.

Sheikhalishahi, S., Miotto, R., Dudley, J. T., Lavelli, A., Rinaldi, F. & Osmani, V. (2025), 'A systematic review of large language model (llm) evaluations in clinical medicine', *BMC Medical Informatics and Decision Making* **25**, 24.

Srinivas, C., Nandini Prasad, K. S., Zakariah, M., Alothaibi, Y. A., Shaukat, K., Partibane, B. & Awal, H. (2023), '3d mri segmentation using u-net architecture for the detection of brain tumor', *Procedia Computer Science* **218**, 646–653.

Sun, X., Yin, Y., Yang, Q. & Huo, T. (2023), 'Artificial intelligence in cardiovascular diseases: Diagnostic and therapeutic perspectives', *European Journal of Medical Research* **28**(1), 242.

Tang, Y. D., Zhang, K. & Liu, D. (2025), 'Revolutionizing health care: The transformative impact of large language models in medicine', *Journal of Medical Internet Research* **27**, e59069.

Vijay, S., Guhan, T., Srinivasan, K., Vincent, P. M. D. R. & Chang, C. Y. (2023), 'Mri brain tumor segmentation using residual spatial pyramid pooling-powered 3d u-net', *Frontiers in Public Health* **11**, 1091850.

Wang, C., Liu, S., Zhou, H. & Wang, J. (2025), 'Large language models for disease diagnosis: A scoping review', *npj Artificial Intelligence* **1**, 11.

World Health Organization (2023), 'Cardiovascular diseases'. Accessed: 2024-11-14.
URL: *https://www.who.int/health-topics/cardiovascular-diseases*

Xu, L., Xu, C., Zhu, S., Tang, H. et al. (2024), 'Screening and diagnosis of cardiovascular disease using artificial intelligence-enabled cardiac magnetic resonance imaging', *Nature Medicine* **30**, 1471–1480.

Zhang, K., Meng, X., Yan, X., Ji, J., Liu, J., Xu, H., Zhang, H., Liu, D., Wang, J., Wang, X., Gao, J., Wang, Y., Shao, C., Wang, W., Li, J., Zheng, M. Q., Yang, Y. & Tang, Y. D. (2024), 'Large language models in healthcare and medical applications: A review', *Computers in Biology and Medicine* **176**, 108618.

Zhang, L., Chen, Y., Zhao, X., Wang, H. & Li, J. (2023), 'An arrhythmia classification model based on vision transformer with deformable attention', *Micromachines* **14**(6), 1155.

Zhou, F., Li, M. & Zhang, Q. (2025), 'A novel hybrid cnn-transformer model for arrhythmia detection without r-peak identification using stockwell transform', *Scientific Reports* **15**, 92582.

Zhou, F., Liu, H., Gu, B., Zou, X., Huang, J. & Wu, J. (2024), 'A survey of large language models in medicine: Principles, applications, and challenges', *arXiv preprint arXiv:2310.05112* .

Zhou, H., Liu, F., Gu, B., Zou, X., Huang, J., Wu, J., Zhang, Y., Hu, C. & Li, H. (2024), 'Large language models in healthcare and medical domain: A review', arXiv preprint arXiv:2401.06775.