Multimodal AI-Powered Triage System for Oral Health Conditions: A Distributed Architecture Approach

Pedro Gabriel Amorim Soares[1]
[1]PUC Minas, Brazil
pedrogabrielbhz@gmail.com

September 22, 2025

Abstract

This work presents the design and implementation of a multimodal artificial intelligence system for triaging patients with oral health conditions. The proposed system integrates natural language processing through Large Language Models (LLMs), computer vision for image analysis, and multimodal fusion techniques to provide comprehensive patient assessment. The architecture employs a distributed message queue system for handling healthcare professional inputs through a versioned questionnaire interface, enabling scalable and priority-based patient triage. By combining textual patient history with visual examination data, the system aims to improve triage accuracy and reduce waiting times in oral healthcare settings. The research contributes to the growing field of AI-assisted medical diagnosis by addressing the specific challenges of oral pathology detection and prioritization in resource-constrained healthcare environments.

1 Introduction

Oral diseases affect nearly 3.5 billion people worldwide, making them among the most common health conditions globally. In Brazil, access to specialized oral healthcare remains limited, particularly in underserved regions, creating significant delays in diagnosis and treatment of potentially serious conditions including oral cancer, periodontal diseases, and infectious pathologies.

The integration of artificial intelligence (AI) in healthcare has shown remarkable potential for improving diagnostic accuracy and streamlining clinical workflows. Recent advances in multimodal AI, which combines multiple data modalities such as text and images, offer unprecedented opportunities for comprehensive patient assessment. However, the application of these technologies to oral health triage remains underexplored, particularly in developing integrated systems that can handle both clinical questionnaires and visual examination data.

This research proposes a novel multimodal AI system designed specifically for oral health triage, combining Large Language Models (LLMs) for processing patient history and symptoms with computer vision models for analyzing oral cavity images. The system employs a distributed architecture with message queuing to ensure scalability and implements priority-based routing to optimize patient care delivery.

2 Literature Review

2.1 AI in Oral Health Diagnosis

Traditional approaches to oral disease detection have relied primarily on visual examination and manual screening. Recent studies have demonstrated the effectiveness of deep learning models in detecting oral cancer with sensitivity rates exceeding 85%. Convolutional Neural Networks (CNNs) have been particularly successful in identifying malignant lesions, with architectures such as ResNet and EfficientNet showing promising results.

2.2 Large Language Models in Healthcare

The emergence of LLMs has revolutionized natural language processing in medical contexts. Models like GPT-4, Claude, and specialized medical variants such as Med-PaLM have demonstrated capabilities in clinical reasoning, symptom assessment, and treatment recommendation. However, their application in structured triage systems, particularly for oral health, remains limited.

2.3 Multimodal Fusion Techniques

Multimodal fusion in medical AI involves combining different data types to improve diagnostic accuracy. Early fusion, late fusion, and hybrid approaches each offer distinct advantages. Recent work on Vision-Language models has shown particular promise for medical applications.

2.4 Healthcare System Architecture and Message Queuing

Distributed systems in healthcare must balance scalability, reliability, and compliance with data protection regulations. Message queue architectures provide asynchronous processing capabilities essential for handling varying loads in clinical settings.

3 Motivation

The Brazilian public health system (SUS) faces significant challenges in providing timely oral healthcare, with waiting times for specialist consultations often exceeding several months. This delay is particularly critical for conditions such as oral cancer, where early detection significantly impacts survival rates.

Furthermore, the COVID-19 pandemic has accelerated the adoption of digital health solutions, creating an opportune moment for implementing AI-assisted triage systems. The integration of multimodal AI can address several critical gaps:

1. **Standardization Gap**: Variability in triage criteria across different healthcare facilities

2. **Expertise Gap**: Limited availability of specialized oral health professionals in remote areas

3. **Efficiency Gap**: Manual triage processes creating bottlenecks in patient flow

4. **Data Integration Gap**: Disconnected systems preventing comprehensive patient assessment

4 Problem Statement and Objectives

4.1 Problem Statement

Current oral health triage systems in Brazilian healthcare facilities rely predominantly on manual assessment by general practitioners or nurses who may lack specialized training in oral pathology. This leads to:

- Inconsistent prioritization of urgent cases
- Delayed diagnosis of serious conditions
- Inefficient resource allocation
- Limited integration of patient history with visual examination findings

The core research question is: *How can multimodal AI techniques be leveraged to create an accurate, scalable, and clinically viable triage system for oral health conditions that integrates both textual patient data and visual examination findings?*

4.2 Objectives

4.2.1 General Objective

To design, implement, and evaluate a multimodal AI-powered triage system for oral health conditions that combines natural language processing and computer vision techniques within a distributed, scalable architecture.

4.2.2 Specific Objectives

1. **Develop a versioned questionnaire system** that can adapt to changing clinical protocols while maintaining data consistency
2. **Implement a robust message queue architecture** for handling concurrent patient submissions with priority-based processing
3. **Design and train specialized AI models** for:
    - Text analysis of patient symptoms and history using LLMs
    - Image classification of oral cavity photographs
    - Multimodal fusion for comprehensive assessment
4. **Create an ensemble verdict system** that aggregates multiple model outputs into clinically meaningful priority levels
5. **Validate the system** through comparative studies with healthcare professionals' assessments
6. **Develop a user-friendly interface** for healthcare professionals to interact with the system

5 Methodology

5.1 System Architecture Design

The proposed system follows a microservices architecture pattern with the following core components:

5.1.1 Data Collection Layer

- **Chatbot Interface**: Python-based REST API using FastAPI
- **Conversation Manager**: Redis-backed session management with conversation versioning
- **Schema Versioning**: JSON Schema with semantic versioning for questionnaires

5.1.2 Message Queue System

- **Primary Queue**: RabbitMQ with priority queues (5 levels)
- **Dead Letter Queue**: For handling processing failures
- **Monitoring**: Prometheus metrics for queue health

5.1.3 AI Processing Pipeline

- **Text Processing**: Fine-tuned BioBERT or ClinicalBERT model
- **Image Analysis**: EfficientNet-B4 or Vision Transformer (ViT)
- **Multimodal Fusion**: Late fusion with attention mechanisms

5.2 Data Collection and Preparation

5.2.1 Dataset Compilation

5.2.2 Annotation Protocol

5.3 Model Development

5.3.1 Text Model

1. Pre-training on Portuguese medical texts
2. Fine-tuning on oral health specific corpus
3. Prompt engineering for symptom extraction

5.3.2 Vision Model

1. Transfer learning from ImageNet pre-trained models
2. Data augmentation strategies (rotation, color adjustment)
3. Class imbalance handling through weighted loss functions

5.3.3 Multimodal Integration

5.4 Evaluation Methodology

5.4.1 Technical Metrics

- **Classification Performance**: Accuracy, Precision, Recall, F1-Score, AUROC
- **System Performance**: Throughput, Latency, Queue processing time
- **Robustness**: Performance under data drift, adversarial examples

5.4.2 Clinical Validation

- **Expert Agreement**: Cohen's Kappa with specialist assessments
- **Time-to-Diagnosis**: Comparison with traditional triage
- **User Satisfaction**: System Usability Scale (SUS) scores

5.5 Implementation Timeline

1. **Months 1-2**: Literature review and system design
2. **Months 3-4**: Data collection and annotation
3. **Months 5-7**: Model development and training
4. **Months 8-9**: System integration and testing
5. **Months 10-11**: Clinical validation
6. **Month 12**: Analysis and documentation

6 Expected Results

6.1 Technical Outcomes

We expect the proposed system to achieve:

- **Triage Accuracy**: ¿90% agreement with specialist assessment for high-priority cases
- **Processing Time**: ¡30 seconds per patient from submission to priority assignment
- **System Throughput**: Capability to handle 1000+ concurrent users
- **Model Performance**:
  - Sensitivity ¿85% for urgent conditions
  - Specificity ¿90% for routine cases
  - AUROC ¿0.95 for binary urgent/non-urgent classification

6.2 Clinical Impact

- Reduction in median time-to-specialist-consultation for urgent cases
- Improved early detection rates for oral cancer and pre-cancerous lesions
- Standardization of triage criteria across participating institutions
- Enhanced documentation and tracking of patient presentations

6.3 Research Contributions

This work is expected to contribute:

1. **Novel Architecture**: First integrated multimodal triage system specifically designed for oral health in Portuguese-speaking populations
2. **Benchmark Dataset**: Annotated dataset of oral health cases with corresponding questionnaires and images
3. **Methodological Framework**: Reusable architecture pattern for medical triage systems in resource-constrained settings
4. **Clinical Guidelines**: Evidence-based recommendations for AI-assisted triage implementation

6.4 Limitations and Future Work

- Initial focus on common conditions (expand to rare diseases in future)
- Portuguese language specificity (plan for multilingual support)
- Static image analysis (future: video, 3D imaging integration)
- Limited to triage (future: treatment recommendation, follow-up scheduling)