SPU SIT
SCHOOL OF INFORMATION TECHNOLOGY

2025 SHOWCASE

PILLAR OF INFORMATION TECHNOLOGY

DOCTOR OF PHILOSOPHY
(INFORMATION TECHNOLOGY)

https://www.spu.ac.th/fac/informatics/

SPU SIT
SCHOOL OF INFORMATION TECHNOLOGY

Cleft Lip and Palate Classification Through Vision Transformers and Siamese Neural Networks

Oraphan Nantha
PhD. IT Candidate

Published 2 papers in SCOPUS Q1

2 Papers in SCOPUS Q1

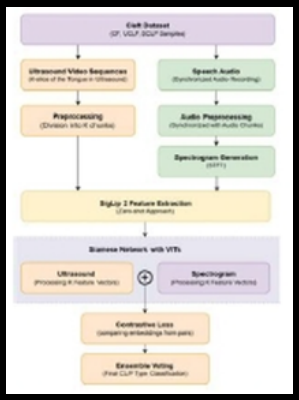
1) O. Nantha, B. Sathanarugsawait, P. Praneetpolgrang. Enhanced Cleft Lip and Palate Classification Through Vision Transformers and Siamese Neural Networks, Journal by Applied Sciences (ISSN 2076-3417) (SCOPUS Q1), 2025

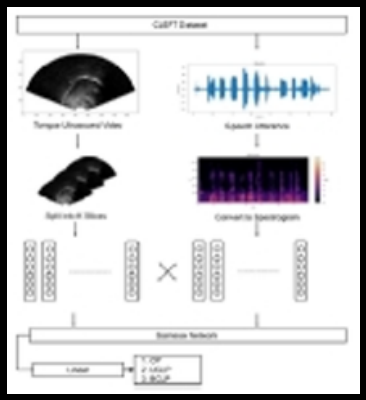
2) O. Nantha, B. Sathanarugsawait, P. Praneetpolgrang Cleft Lip and Palate Classification Through Vision Transformers and Siamese Neural Networks, Journal of Imaging (Q1 SCOPUS) ,2024

Abstract

This study introduces a novel approach for the diagnosis of Cleft Lip and/or Palate (CI/P) integrating Vision Transformers (ViTs) and Siamese Neural Networks. Our study is the first to employ this integration specifically for CL/P classification, leveraging the strengths of both models to handle complex, multimodal data and few shot learning scenarios.

Unlike previous studies that rely on single-modality data or traditional machine learning models, we uniquely fuse anatomical data from ultrasound images with functional data from speech spectrograms. This multimodal approach captures both structural and acoustic features critical for accurate CL/P classification. Employing Siamese Neural Networks enables effective learning from a small number of labeled examples, enhancing the model's generalization capabilities in medical imaging contexts where data scarcity is a significant challenge.





The models were tested on the UltraSuite CLEFT dataset, which includes ultrasound video sequences and synchronized speech data, across three cleft types: Bilateral, Unilateral, and Palate-only clefts. The two-stage model demonstrated superior performance in classification accuracy (82.76%), Fl-score (80.00-86.00%), precision, and recall, particularly distinguishing Bilateral and Unilateral Cleft Lip and Palate with high efficacy. This research underscores the significant potential of advanced Al techniques in medical diagnostics, offering valuable insights into their application for improving clinical outcomes in patients with
CL/P.

SPU SIT
SCHOOL OF INFORMATION TECHNOLOGY

Application of Generative Artificial Intelligence Models for Accurate Prescription Label Identification and Information Retrieval for the Elderly in Northern East of Thailand.

Parinya Thetbanthad
PhD. IT Candidate

Published 2 Papers in SCOPUS Q1

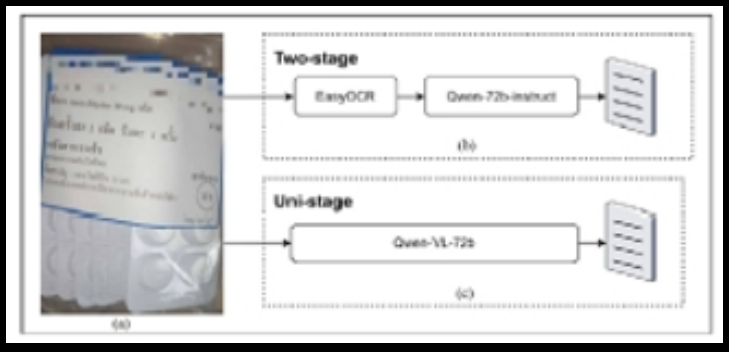
2 Papers in SCOPUS Q1

1) P. Thetbanthad, B. Sathanarugsawait, P.Praneetpolgrang. Enhanced Cleft Lip and Palate Classification Through Vision Transformers and Siamese Neural Networks, Journal by Applied Sciences (ISSN 2076-3417) (SCOPUS Q1), 2025

2) P. Thetbanthad, B. Sathanarugsawait, P. Praneetpolgrang .Cleft Lip and Palate Classification Through Vision Transformers and Siamese Neural Networks, Journal of Imaging (SCOPUS Q1),2024

Abstract

This study introduces an innovative Al-driven method designed to assist elderly patients in Thailand with medication management through precise interpretation of drug labels. research explores two model architectures: a Two-Stage Optical Character Recognition (OCR) and Large. Language Model (LLM) pipeline, which integrates EasyOCR with Qwen2-72b-instruct, and a Uni-Stage Visual Question Answering (VQA) model that utilizes Qwen2-72b-VL. Both models function in a zero-shot capacity, employing Retrieval-Augmented Generation (RAG) with DrugBank references to ensure contextually relevant and accurate outputs. The models were evaluated using a dataset of 100 diverse prescription labels from Thai healthcare facities, with performance assessed using RAGAs metrics focusing on Context Recall, Factual Correctness, Faithfulness, and Semantic Similarity.



The Two-Stage model demonstrated high accuracy (94%) and strong RAGAGAs scores, particularly in Context Recall (0.88) and Sernantic Similarity (0.91), making suitable for interpreting complex medication instructions. The Uni-Stage model offered quicker response times, proving advantageous for high-volume settings like pharmades. Overall, the study highlights the potential of zero-shot Al models in addressing medication management challenges for the elderly, delivering clear, accurate, and contextually relevant label interpretations and demonstrating A's adaptablity in healthcare to meet diverse real-world needs.