COMP5425 Multimedia Retrieval

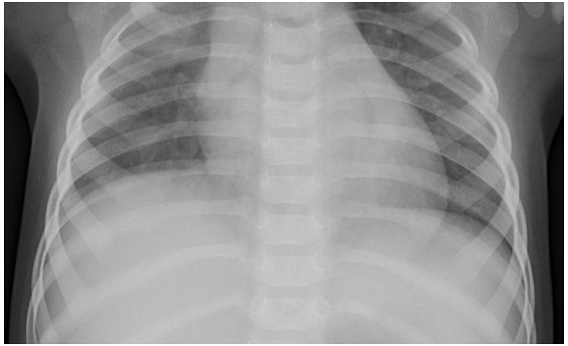
Project Final Presentation

- Project Number: Graoup39
- Project Title: PNEUMONIA CLASSIFICATION USING CHEST X-RAY IMAGES
- Presenters: Min Huang, Meiyi Chen,
  Junyang Feng, Yayuan Zheng

THE UNIVERSITY OF SYDNEY

Pneumonia X-ray Image Retrieval System

📌 Background



- Pneumonia is especially dangerous for children and the elderly.
- Manual diagnosis is slow and expertise-dependent.
- Misdiagnosis is common in under-resourced areas.

🚀 Our Solution

- We developed an *X-ray image retrieval system* to assist diagnosis.

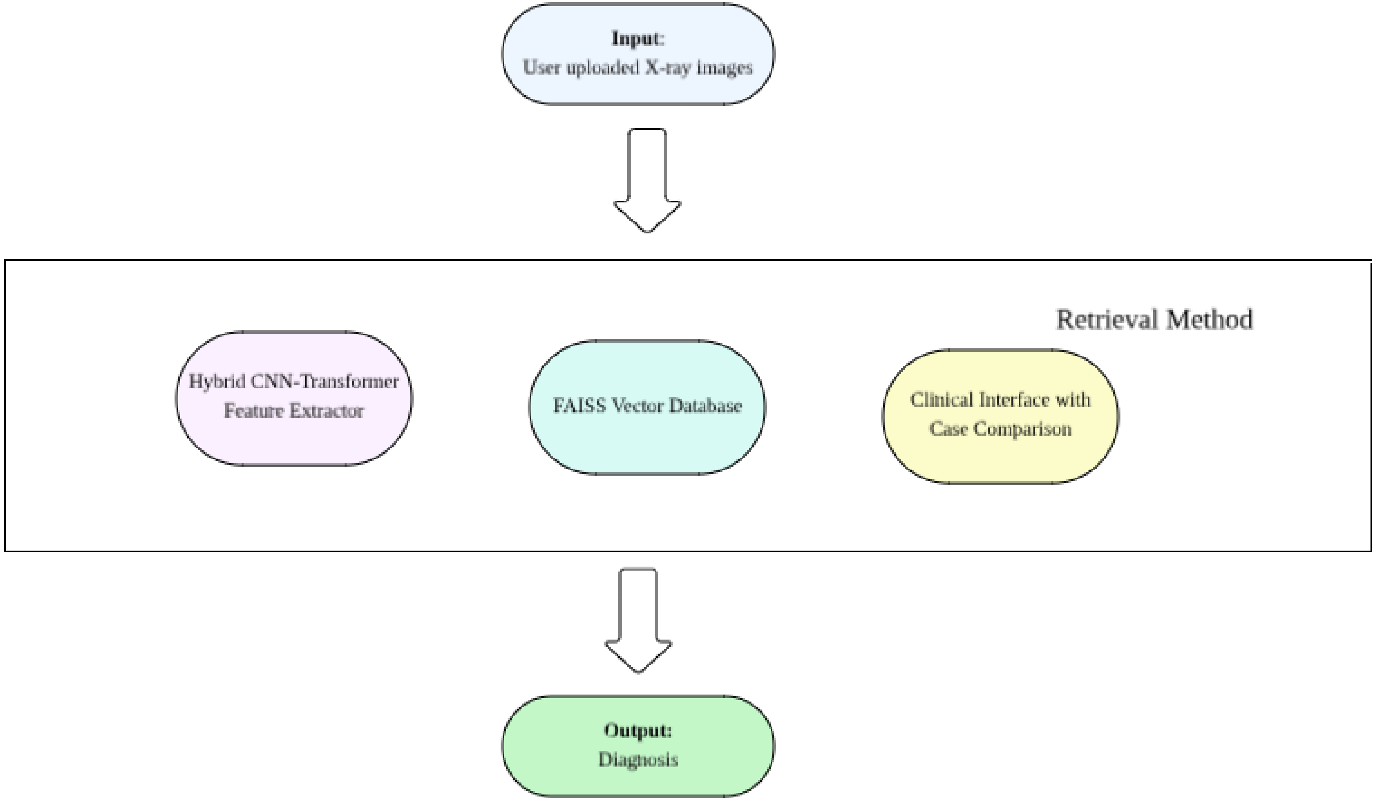
📌 Key features:

- 🔍 Retrieve visually similar chest X-rays
- 🧠 AI-based classification
- 🌐 User-friendly web interface

📌 Benefit

- Faster access to relevant cases
- Supports diagnosis and review
- Designed for clinical and research use

System architecture diagram



Methodology

✅ **Dual-path feature extraction:** EfficientNet-B3 (local features) + ViT (global features)

✅ **FAISS vector index:** supports fast similarity retrieval (retrieval time: ~1.2 seconds)

✅ **Interactive interface:** supports uploading images, displaying Top-K similar cases and heat map interpretation.

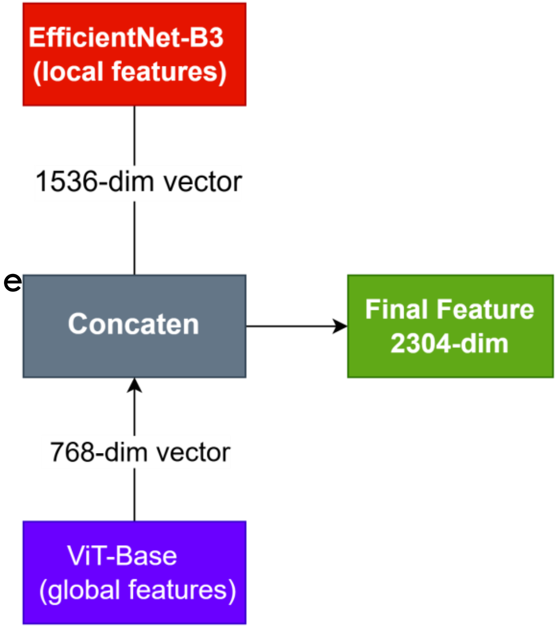
📁 **Reason for selection:** Local + global features complement each other to improve the ability to distinguish pathological patterns (such as pneumonia exudate area vs normal vascular texture)

Technical Solution – 1.Feature extraction module

THE UNIVERSITY OF SYDNEY

Key technologies:

- **EfficientNet-B3** (extract local features such as texture and edges):
  - Implementation: load the ImageNet pre-trained model, remove the classification layer, and output a 1536-dimensional vector.
- **Vision Transformer /ViT-Base** (capture global structure, such as lung lobe distribution):
  - Implementation: segment the image into 16 × 16 patches, perform self-attention calculations, and output a 768-dimensional vector.



Key steps:

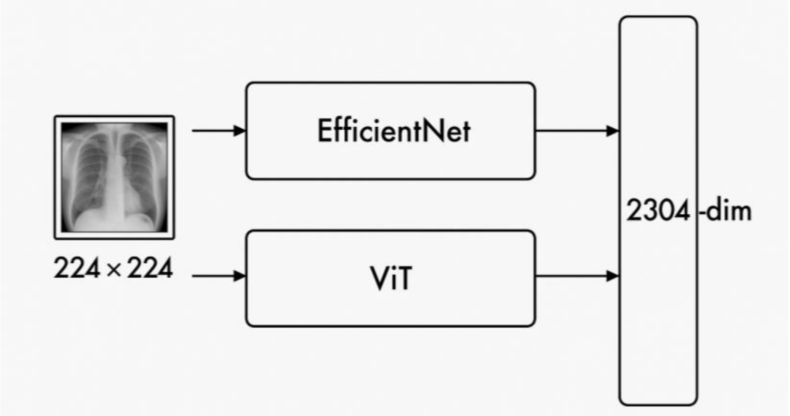
- **L2 normalization**: Normalize the output vectors of the two models separately to ensure the stability of cosine similarity calculation.
- **Multi-source feature concatenation**: Concatenate the two vectors into 2304-dimensional features to retain the difference between local and global information.

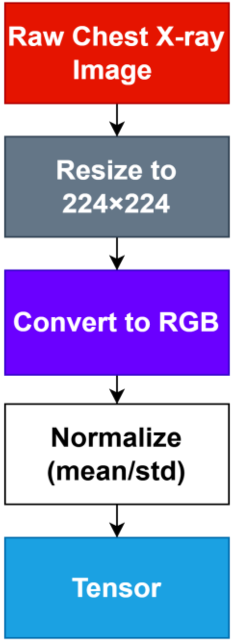
2. Feature processing and storage module

Image preprocessing pipeline:

- **Size normalization:** scale to 256px → center crop to 224 × 224.
- **Pixel normalization:** RGB values are normalized by ImageNet mean ([0.485, 0.456, 0.406]) and standard deviation ([0.229, 0.224, 0.225]).

Feature Vector Structure





Preprocessing Pipline

Feature management design (structured storage):

- Store feature vectors by "comparison image group" and "test image group".
- Pair indexes with labels to support fast retrieval (e.g. {image_id: (feature_vector, label)}).

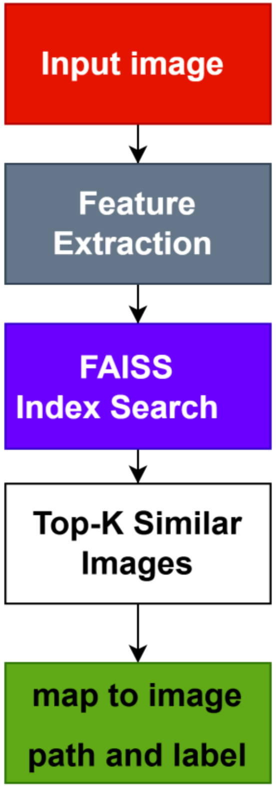
3. Image retrieval module

📥 **Core components:**

- **FAISS index:**
  - Type: IndexFlatIP (cosine similarity search based on inner product).
  - Function: pre-calculate training set feature vectors, build efficient indexes, and support Top-K retrieval (single time ~1.2 seconds).

📥 **Result decoding:**

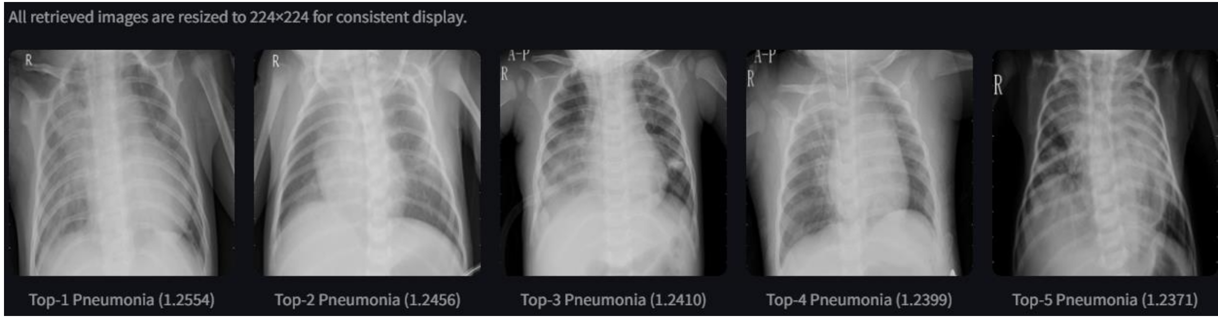
Input query image → extract features → FAISS returns Top-K index → map to image path and label.



4. Feature processing and storage module

Interactive display:

- **Top-K visualization:** grid display of query images and similar cases (including labels and feature distances).
- **Grad-CAM heat map:** heat map generated by gradient back propagation, highlighting the model's focus area (such as the middle and lower lung fields).

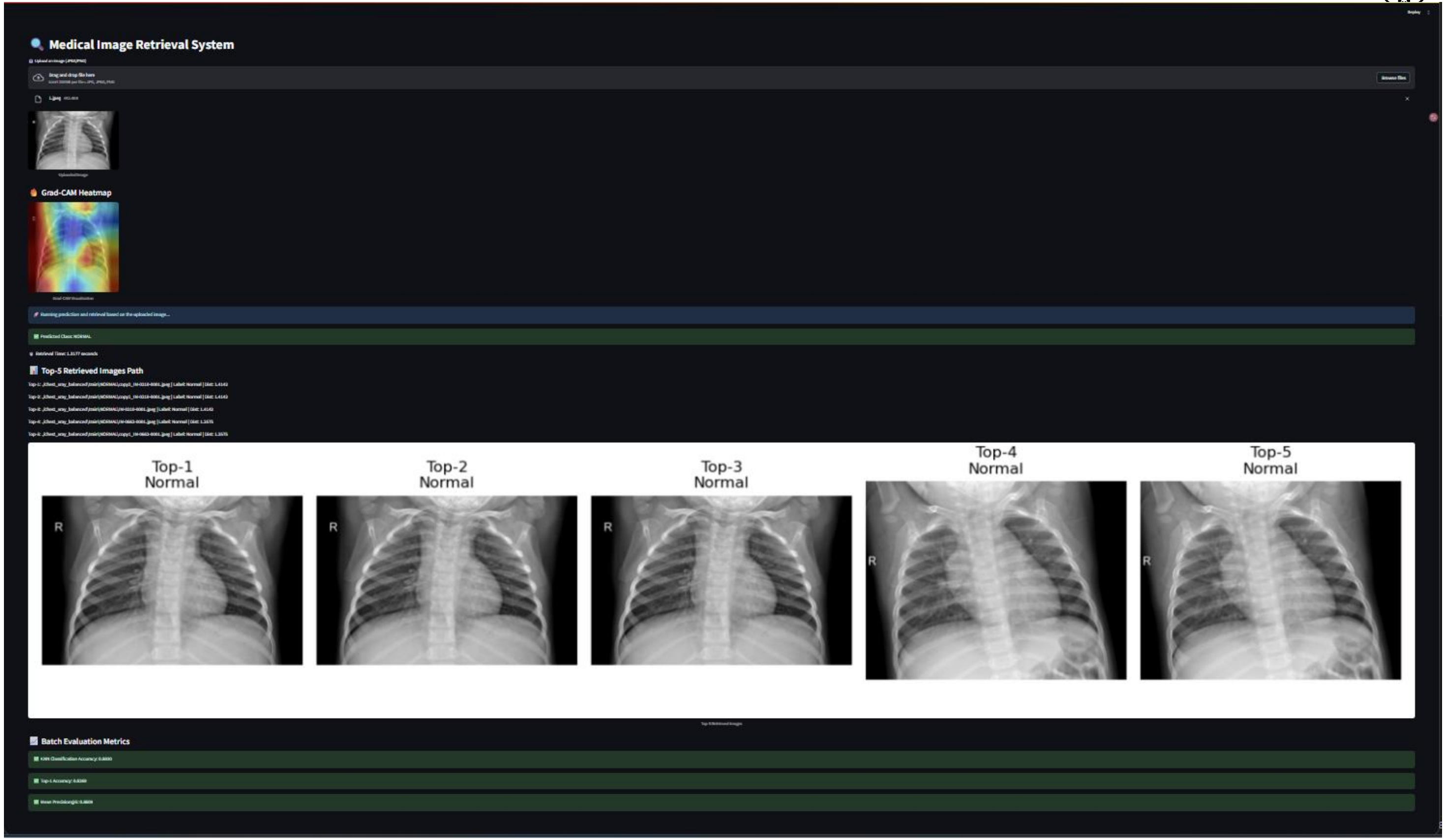


Evaluation indicators:

- **Precision@5 (86.09%)**: the proportion of the same category in the top 5 results.
- **Top-1 accuracy (82.69%)**: the matching rate of the most similar image category.
- **KNN Precision（88.3%）**: the model's overall classification capacity.

Result

THE UNIVERSITY OF SYDNEY



Result

⏱️ Grad-CAM heatmap interpretative analysis

- Red/yellow area: high response value (lesions in the middle and lower lung fields, consistent with the pathological characteristics of pneumonia).

- Blue area: irrelevant anatomical structure (such as ribs, chest wall soft tissue).

⏱️ Clinical consistency:

- The model's focus area is consistent with traditional medical diagnostic rules, which enhances the credibility of the results.

Result

⏱️ Retrieval time: 1.1988 seconds ✅ Predicted class: Pneumonia

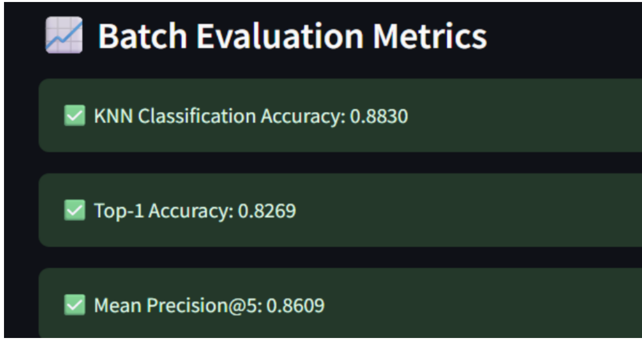
📸 Top-5 similar image

Retrieved Images	**Label**	**Distance**
Top-1	Pneumonia	1.3143
Top-2	Pneumonia	1.3004
Top-3	Pneumonia	1.3001
Top-4	Pneumonia	1.3000
Top-5	Pneumonia	1.2991

Result

✅ Quantitative Performance Evaluation

Evaluation Metrics	ResNet-50 + Faiss	Explanation
KNN classification accuracy	88.30%	Overall classification capability of the model
Top-1 retrieval accuracy	82.69%	Single image retrieval accuracy
Mean Precision@5	86.09%	The overall ranking quality of relevant images in the top 5 results



Discussion

System Advantages

- **Multi-granularity feature fusion (combined with EfficientNet-B3 + ViT)**
  - Effect: KNN classification accuracy increased to 88.30% (3.2% higher than single model).
  - Reason: Local and global features complement each other (such as alveolar exudation vs. lung lobe distribution).
- **FAISS efficient search (based on IndexFlatIP and L2 normalization)**
  - Effect: A single search takes only 1.2 seconds, and Precision@5 reaches 86.09%.
  - Reason: Optimize vector alignment, taking into account both speed and accuracy.
- **Heatmap interpretability**
  - Method: Grad-CAM visualization model focuses on the area (middle and lower lung fields).
  - Effect: It complies with clinical diagnosis rules and improves doctors' trust.

Discussion

📥 System Disadvantages

- **Aesthetic Customisation of the User Interface:** The Streamlit-based interface has a default colour scheme and layout that may look unduly basic to certain clinical users who are used to specialised medical software.

- **Limited Multilingual Interface Support:** the web interface currently operates exclusively in English. While this does not hinder core functionality, offering additional language options (e.g., Spanish or Mandarin) could improve accessibility for non-English-speaking medical staff in global healthcare settings.

Thank You !

THE UNIVERSITY OF
SYDNEY