

# **XÂY DỰNG HỆ THỐNG TRUY VẤN HÌNH ẢNH SỬ DỤNG VECTOR ĐẶC TRƯNG ĐƯỢC TRÍCH XUẤT BỞI MÔ HÌNH RESNET50 VÀ CLIP**

**Nguyễn Tam Điệp - 19521360**

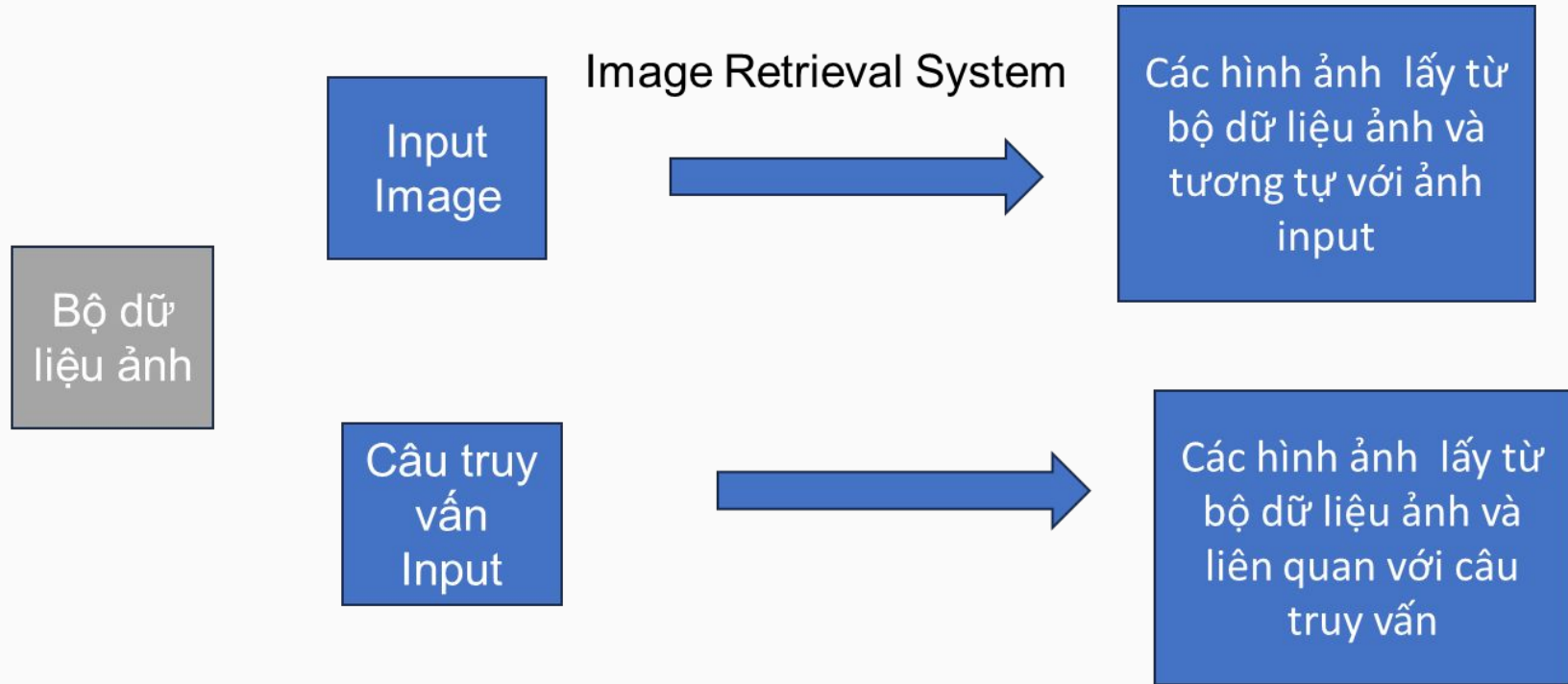
# Tóm tắt

- Lớp: CS519.011
- Link Github của nhóm:  
<https://github.com/nguyentamdiep/CS519.011>
- Link YouTube video: <https://youtu.be/mH4XrvWTR0s>
- Ảnh + Họ và Tên của các thành viên



Nguyễn Tam Điệp

# Giới thiệu



# Mục tiêu

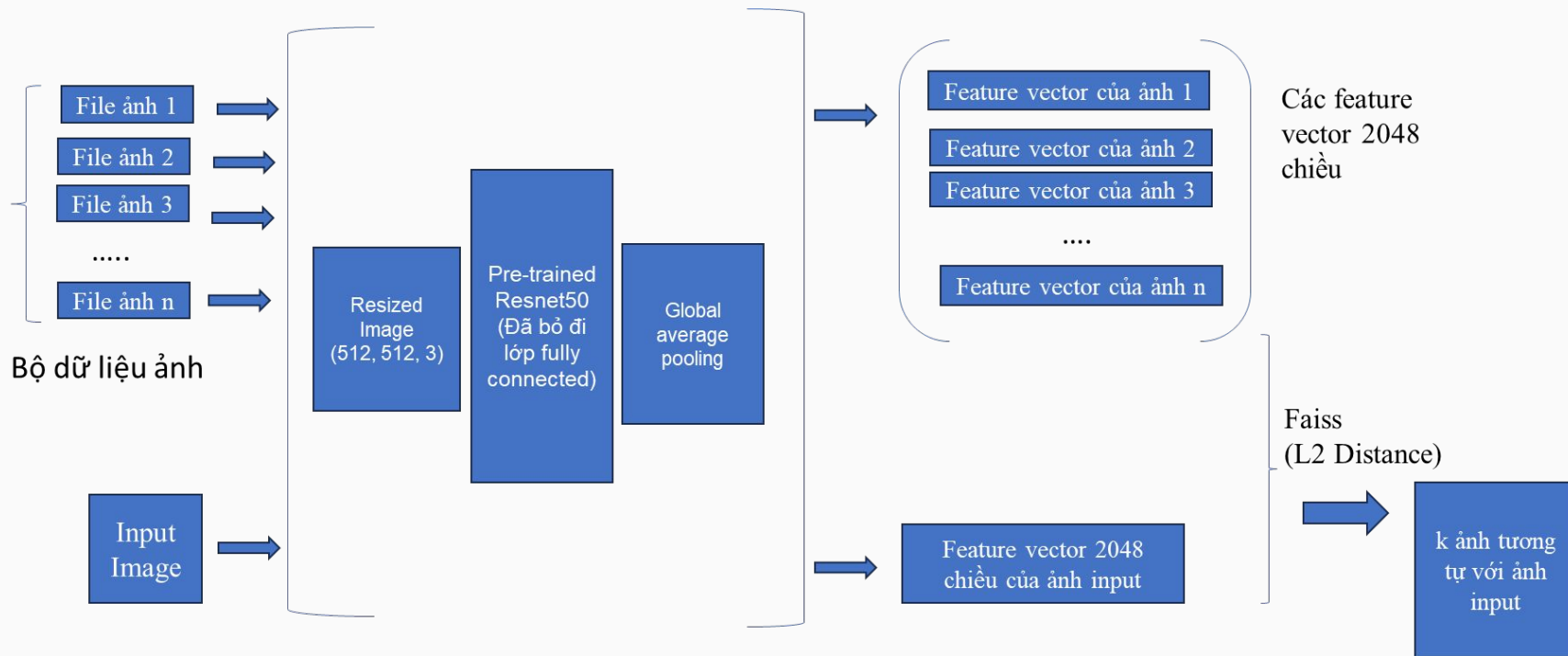
Xây dựng hệ thống truy xuất hình ảnh với bộ dữ liệu Flickr 8k.  
Đây là bộ dữ liệu có 8091 ảnh.

Hệ thống truy xuất hình ảnh thực hiện 2 tác vụ:

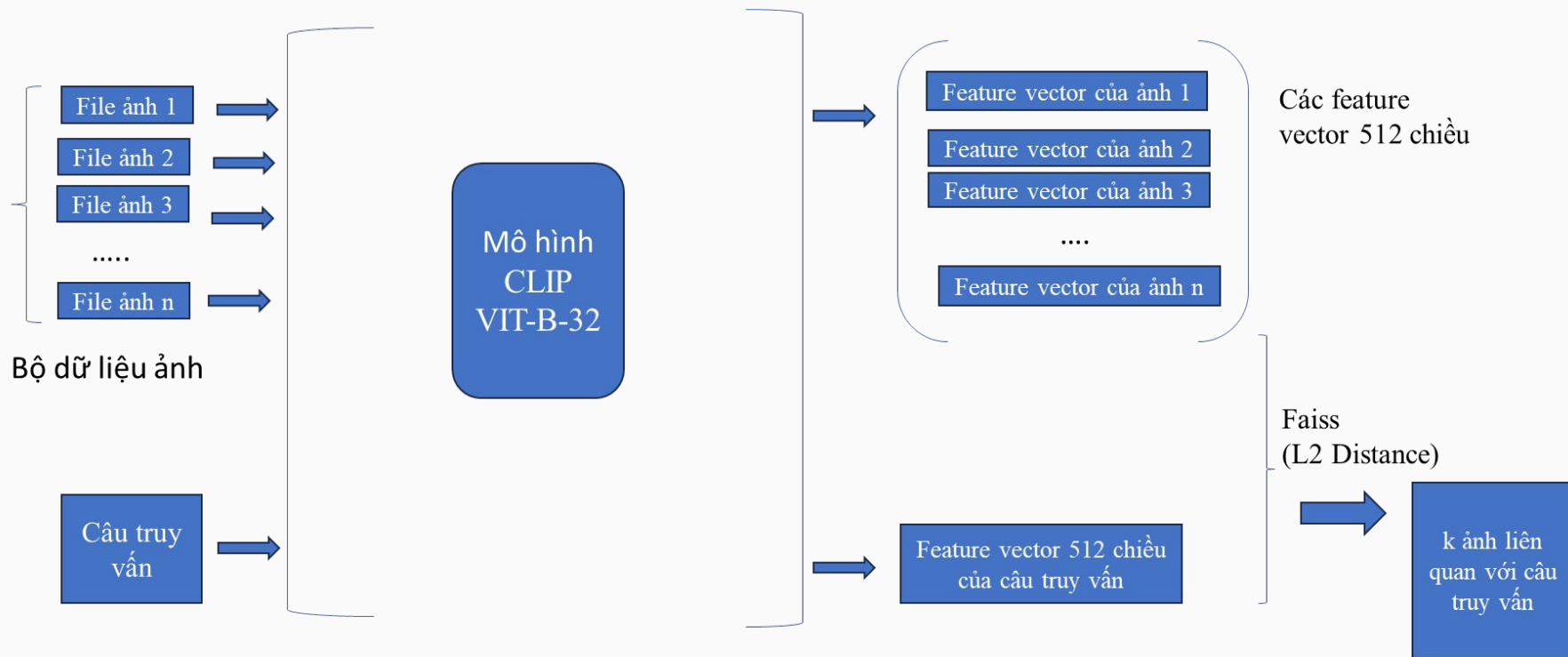
- + Truy xuất ảnh tương tự với ảnh đầu vào
- + Truy xuất ảnh liên quan với câu truy vấn

# Nội dung và Phương pháp

## 1 Truy xuất các ảnh tương tự với ảnh đầu vào



## 2. Truy xuất các ảnh có liên quan với câu truy vấn



# Kết quả dự kiến

Nhập số lượng kết quả muốn trả về  
100

Upload an image

Drag and drop file here  
Limit 200MB per file • JPG, PNG, JPEG

bongda.jpg 0.9MB

search image

enter query

search text query



uploaded image

các ảnh tương đồng:



flickr8k/images/2947453228\_08f2d2a467.jpg   flickr8k/images/2678612999\_893e5071f8.jpg   flickr8k/images/3326545455\_960e5442e9.jpg



enter query

A child plays with a fountain

search text query

Kết quả:

Xây dựng hệ thống truy xuất ảnh với bộ dữ liệu ảnh Flickr 8k

# Tài liệu tham khảo

He, K. *et al.* (2015a) *Deep residual learning for image recognition*, *arXiv.org*. Available at:  
<https://arxiv.org/abs/1512.03385>

Radford, A. *et al.* (2021a) *Learning transferable visual models from Natural Language Supervision*, *arXiv.org*.  
Available at: <https://arxiv.org/abs/2103.00020>