**Tìm kiếm ảnh dựa trên đồ thị tri thức và đối tượng láng giềng**

**1. Xác định yêu cầu và mục tiêu**

* **Mục tiêu chính**: Tìm kiếm ảnh từ COCO dựa trên mối quan hệ giữa các đối tượng trong ảnh, sử dụng đồ thị tri thức (KG) và thông tin về các đối tượng láng giềng.
* **Ứng dụng thực tế**: Truy vấn ảnh theo ngữ cảnh, tìm kiếm ảnh thông minh, hỗ trợ hệ thống gợi ý hình ảnh.

**2. Tổng quan phương pháp**

**Bước 1: Tiền xử lý dữ liệu COCO**

* COCO chứa các **chú thích ảnh** (captions), **danh mục đối tượng** (80 danh mục), **tọa độ bounding box**, và **mối quan hệ giữa các đối tượng**.
* **Xây dựng KG từ dữ liệu COCO**:
  + Nút (Nodes): Các đối tượng trong ảnh (ví dụ: “dog”, “car”, “tree”).
  + Cạnh (Edges): Quan hệ giữa các đối tượng (ví dụ: “dog next to car”).
  + Thuộc tính: Vị trí, kích thước, số lượng đối tượng.

**Bước 2: Xây dựng đồ thị tri thức từ COCO**

* **Trích xuất thông tin đối tượng**: Dùng mô hình **Detectron2** hoặc **YOLO** để nhận diện đối tượng trong ảnh.
* **Biểu diễn dưới dạng KG**:
  + Dùng **RDF triplets** hoặc **Knowledge Graph Embeddings** (TransE, TransH) để lưu trữ quan hệ.
  + Ví dụ: (dog, next\_to, car), (person, holding, umbrella).
* **Lưu trữ KG**:
  + **Neo4j** (cơ sở dữ liệu đồ thị)

**Bước 3: Xây dựng mô hình tìm kiếm dựa trên láng giềng**

* **Biểu diễn mối quan hệ trong KG** bằng cách dùng:
  + **Graph Neural Networks (GNN)** để học embedding của các nút.
  + **HITS / PageRank** để xác định các nút quan trọng trong truy vấn.
* **Khai thác thông tin láng giềng**:
  + Khi truy vấn “dog next to car”, tìm các ảnh có đối tượng dog có láng giềng là car.
  + **Tính điểm tương đồng** giữa các ảnh bằng **cosine similarity trên vector embedding**.

**Bước 4: Xây dựng công cụ tìm kiếm ảnh**

* **Truy vấn ngôn ngữ tự nhiên**: Chuyển câu truy vấn thành đồ thị con và tìm kiếm trong KG.
* **Giao diện tìm kiếm**:
  + Xây dựng API bằng **FastAPI** hoặc **Flask**.
  + Hiển thị kết quả bằng React hoặc Streamlit.

**3. Công nghệ sử dụng**

* **Xử lý ảnh**: Detectron2, OpenCV, PyTorch
* **Đồ thị tri thức**: Neo4j, NetworkX, DGL
* **Mô hình tìm kiếm ảnh**: Graph Neural Networks (GNN), Knowledge Graph Embeddings
* **Triển khai hệ thống**: FastAPI, Streamlit

**4. Mở rộng và tối ưu**

* **Bổ sung ngữ cảnh từ captions COCO** để cải thiện tìm kiếm.
* **Học sâu hơn với Graph Attention Networks (GAT)** để tăng độ chính xác.
* **Tối ưu truy vấn với Neo4j Cypher Query Language** để tăng tốc tìm kiếm.