

ĐẠI HỌC QUỐC GIA THÀNH PHỐ HỒ CHÍ MINH  
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN



## BÁO CÁO ĐỒ ÁN CUỐI KỲ

CS117.Q11.KHTN - TƯ DUY TÍNH TOÁN

---

### IMAGE-TO-INSTRUCTIONS FOR WATER HYACINTH WEAVING

---

*Giảng viên :*

TS. Ngô Đức Thành

*Sinh Viên :*

Nguyễn Văn Hồng Thái - 23521418

Nguyễn Văn Minh - 23521418

Hoàng Đức Dũng - 23520328

Phan Nhật Tân - 23521405

Nguyễn Đình Thiên Quang - 23521285



---

Ngày 24 tháng 1 năm 2026



# Mục lục

## 2) Vận dụng các kỹ thuật Computational/AI Thinking (CAT)

Quá trình xây dựng giải pháp từ bài toán thực tế được thực hiện tuần tự qua các bước áp dụng tư duy máy tính (Computational Thinking) nhằm chuyển đổi từ dữ liệu thô sang hướng dẫn có cấu trúc, cụ thể như sau:

### Step 1: Abstraction (Trừu tượng hóa) trên Bài toán thực tế

- **Input thực tế:** Một hình ảnh chụp bề mặt đan lục bình (Water Hyacinth Weaving) chứa nhiều chi tiết nhiễu (màu sắc, ánh sáng, background, độ cong vật lý).
- **Quá trình:** Loại bỏ các chi tiết không cần thiết (màu sắc sọc, khiếm khuyết vật liệu nhỏ), chỉ tập trung vào cấu trúc topo học của các mối đan. Bài toán được mô hình hóa thành việc tìm kiếm hàm ánh xạ  $F$ .
- **Kết quả:** Bài toán thực tế trở thành bài toán tính toán **Root Problem: Image-to-Instruction Synthesis**.
  - *Input:* Ảnh RGB đơn  $I \in \mathbb{R}^{H \times W \times 3}$ .
  - *Output 1:* Biểu diễn cấu trúc đan (Weaving Structure Representation) dưới dạng Ma trận  $M$  (Weaving Matrix).
  - *Output 2:* Chuỗi hướng dẫn (Ordered Instruction Sequence)  $S$ .
  - *Ràng buộc:*  $M$  phải thỏa mãn tính đúng đắn về cấu trúc (Structural Fidelity - R1) và  $S$  phải đảm bảo tính khả dụng cho con người (Usability - R2).

### Step 2: Decomposition (Phân rã) trên Root Problem

- **Quá trình:** Nhận thấy việc ánh xạ trực tiếp từ  $I \rightarrow S$  (End-to-End) là quá phức tạp và thiếu tính giải thích (interpretable). Bài toán được phân rã dựa trên chức năng cốt lõi: “Nhìn/Hiểu ảnh” và “Suy luận quy trình”.
- **Kết quả:** Root Problem được phân rã thành 2 module độc lập nối tiếp nhau:
  - **Module 1: Perception (X1)** - Chuyển đổi ảnh sang dữ liệu cấu trúc.
    - *Input:* Ảnh  $I$ .
    - *Output:* Ma trận  $M$ .
  - **Module 2: Reasoning (X2)** - Lập kế hoạch từ dữ liệu cấu trúc.
    - *Input:* Ma trận  $M$ .
    - *Output:* Hướng dẫn  $S$ .

### Step 3: Decomposition (Phân rã) trên Module 1 (Perception)

- **Quá trình:** Để tạo ra ma trận  $M$ , hệ thống cần biết vị trí các nút giao và mối quan hệ trên/dưới tại đó. Module 1 tiếp tục được chia nhỏ thành các tác vụ xử lý thị giác máy tính cụ thể.
- **Kết quả:** Module 1 được phân rã thành 3 bài toán con nối tiếp:
  - **Sub-problem 1.1 (Local Feature Extraction):** Xác định tọa độ các điểm giao nhau.
    - *Input:* Ảnh RGB gốc  $I$ .
    - *Output:* Tập hợp tọa độ các nút giao  $N = \{(x_i, y_i)\}$ .
  - **Sub-problem 1.2 (Local Relation Classification):** Xác định trạng thái đan tại từng điểm giao.
    - *Input:* Tọa độ nút  $N$  và các vùng ảnh cục bộ (local patches) quanh  $N$  trích xuất từ  $I$ .
    - *Output:* Tập hợp các nút đã gán nhãn quan hệ  $N' = \{(x_i, y_i, r_i)\}$  (với  $r_i$  là nhãn over/under).
  - **Sub-problem 1.3 (Topology Encoding):** Mã hóa các quan hệ cục bộ thành ma trận toàn cục.
    - *Input:* Tập hợp nút đã gán nhãn  $N'$ .
    - *Output:* Ma trận cấu trúc đan  $M$  (Weaving Matrix).

### Step 4: Pattern Recognition - Matching (Nhận diện mẫu) với Sub-problem 1.1 & 1.2

- **Quá trình:**
  - Đối với Sub-problem 1.1: Nhận diện đây là bài toán tìm điểm đặc trưng (key-points) trên ảnh.
  - Đối với Sub-problem 1.2: Nhận diện đây là bài toán phân loại hình ảnh dựa trên các đặc điểm cục bộ (visual cues).
- **Kết quả:**
  - Sub-problem 1.1 khớp với bài toán **Keypoint Detection / Object Detection**.
    - *Giải pháp:* Sử dụng mạng CNN hoặc ViT-based để trích xuất đặc trưng.
  - Sub-problem 1.2 khớp với bài toán **Image Classification**.
    - *Giải pháp:* Cắt các local patches quanh điểm giao và đưa qua mô hình phân loại.

## Step 5: Decomposition (Phân rã) trên Module 2 (Reasoning)

- **Quá trình:** Từ ma trận  $M$ , cần tạo ra hướng dẫn  $S$ . Việc này đòi hỏi phải tìm ra một đường đi hợp lý (logic flow) trước khi chuyển đổi thành ngôn ngữ tự nhiên.
- **Kết quả:** Module 2 được phân rã thành:
  - **Sub-problem 2.1 (Procedure Planning):** Tìm thứ tự thực hiện các mối đan.
    - *Input:* Ma trận cấu trúc đan  $M$ .
    - *Output:* Kế hoạch thực hiện (Execution Plan)  $P = \{p_1, p_2, \dots, p_n\}$ .
  - **Sub-problem 2.2 (Generate Instruction):** Chuyển đổi kế hoạch  $P$  thành ngôn ngữ tự nhiên.
    - *Input:* Kế hoạch thực hiện  $P$ .
    - *Output:* Chuỗi văn bản hướng dẫn  $S$  (Ordered Textual Instructions).

## Step 6: Pattern Recognition - Matching (Nhận diện mẫu) với Sub-problem 2.1 & 2.2

- **Quá trình:**
  - Đối với Sub-problem 2.1: Ma trận  $M$  thực chất là một biểu diễn đồ thị (Graph) hoặc lưới (Grid). Việc tìm thứ tự đan tương ứng với việc duyệt qua các nút.
  - Đối với Sub-problem 2.2: Việc chuyển đổi dữ liệu cấu trúc sang văn bản hướng dẫn tuân theo các quy tắc ngữ pháp cố định.
- **Kết quả:**
  - Sub-problem 2.1 khớp với bài toán **Graph Traversal (Duyệt đồ thị)**.
    - *Giải pháp:* Sử dụng thuật toán **DFS (Depth-First Search)** kết hợp Cycle Detection để xác định đường đi hợp lệ.
  - Sub-problem 2.2 khớp với bài toán **Rule-based Text Generation**.
    - *Giải pháp:* Sử dụng phương pháp **Template-based** (điền vào mẫu câu) để đảm bảo tính chính xác và dễ hiểu.