



«««< HEAD

=====

»»»> 26686f86fe95d75738ec5f8dd57d5c86a3f5f16d

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

«««< HEAD

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.

=====

Chương 1

Problem Definition / Clarification

1.0.1 Problem Description

The objective of this project is to develop an automated system capable of converting a single image of a woven product (specifically water hyacinth weaving) into detailed, step-by-step weaving instructions.

- **Input:** A single RGB image (JPG/PNG) of a water hyacinth weaving pattern.
- **Output:**
 1. *Weaving Structure Representation (Weaving Matrix M):* A matrix constructed based on cells with predefined rules.
 2. *Ordered Weaving Instruction Sequence (S):* A textual sequence of instructions guiding the user through the weaving process.

1.0.2 Scope

The project focuses on close-up images of complete woven surfaces featuring regular, repetitive patterns. The aim is to generate a valid and internally consistent weaving procedure that reproduces the observed over/under topology. Recovering the original or unique artisan procedure used to create the specific sample is out of scope.

1.0.3 Assumptions

- Strands are visually separable with sufficient contrast for reliable tracing.
- Over/under relationships at intersections are locally observable and unambiguous.
- Images are captured from an approximately top-down viewpoint with limited perspective distortion.
- All required weaving strands are available in advance, with quantities determined by the weaving matrix dimensions and sufficient lengths to execute the generated procedure.

1.0.4 Constraints

- The system processes a single RGB image per instance; video or multi-view inputs are not supported.
- Each image contains a single homogeneous weaving pattern.
- The pipeline is modular and interpretable, without end-to-end image-to-instruction learning (avoiding "black box" behavior).
- Image resolution must ensure at least 32×32 pixels per strand intersection.
- The weaving surface is approximately planar, with limited perspective distortion and no severe occlusion.
- The image is analyzed using local patches centered at candidate intersections, with moderate intersection density to allow unambiguous local processing.

1.0.5 Requirements

- **R1 (Structural Fidelity):** The generated Weaving Matrix (M) must seamlessly map to the physical topology of the input image.
- **R2 (Usability):** The system must generate valid, human-readable weaving instructions.

Chương 2

Evaluation

2.1 Evaluation

To ensure the efficacy and practicality of the solution, the evaluation is conducted based on two main pillars: Perception Verification and Usability Verification.

2.1.1 Evaluation Metrics

Bảng 2.1: Evaluation Metrics and Targets

Req. Ref	Verification Goal	Metric	Value	Formula
R1	Intersection Accuracy	F1-Score (@ IoU 0.5)	90%	$F1 = 2 \times \frac{Pre \times Rec}{Pre + Rec}$
R1	Relation Logic	Binary Accuracy	90%	$Acc = \frac{TP+TN}{Samples}$
R2	Human Usability	Human Success Rate	85%	$\frac{1}{M} \sum_{i=1}^M \mathbb{I}(Art_i = GT)$

Explanation of Formulas:

- **F1-Score:**

$$F1 = 2 \times \frac{Pre \times Rec}{Pre + Rec}$$

where:

- *Pre* (Precision): The proportion of correctly predicted positive intersections out of all predicted positive intersections.
- *Rec* (Recall): The proportion of correctly predicted positive intersections out of all actual positive intersections.

This metric balances precision and recall, ensuring both are optimized.

- **Binary Accuracy:**

$$Acc = \frac{TP + TN}{Samples}$$

where:

- TP (True Positives): Correctly predicted positive relations.
- TN (True Negatives): Correctly predicted negative relations.
- $Samples$: Total number of samples.

This metric measures the proportion of correct predictions (both positive and negative).

- **Human Success Rate:**

$$\frac{1}{M} \sum_{i=1}^M \mathbb{I}(Art_i = GT)$$

where:

- M : Total number of participants in the user study.
- $\mathbb{I}(Art_i = GT)$: Indicator function that equals 1 if the artifact produced by participant i matches the ground truth (GT), and 0 otherwise.

This metric evaluates the usability of the generated instructions by measuring how often participants successfully reproduce the target pattern.

2.1.2 Data & Verification Protocol

1. Perception Verification (Verifies R1):

- **Data Source:** Real-world Test Set ($N = 200$ self-collected images, with manual Ground Truth).
- **Usage:** This dataset is used strictly to calculate the Intersection F1-Score and Relation Accuracy, ensuring the system correctly identifies strand positions and their topological relationships.

2. Usability Verification (Verifies R2):

- **Protocol:** A User Study involving $N = 20$ participants with basic weaving skills.
- **Procedure:** Participants are asked to reproduce artifacts using only the reference image and the generated instruction text.
- **Usage:** Determines the *Human Success Rate* based on the structural correctness of the artifacts produced by the participants compared to the target pattern.

2.1.3 Error Attribution & Mitigation (Addressing User vs. System Error)

To address the potential validity threat where failure is caused by participant execution error rather than instruction quality:

1. **Pre-qualification:** Participants are screened to ensure they possess basic manual dexterity and weaving fundamentals to minimize purely motor-skill failures.
2. **Post-Failure Analysis:** In cases where the final artifact is incorrect ($Art_i \neq GT$), a root cause analysis interview is conducted:
 - **Type A (System Error):** If the participant followed the instruction but the instruction was logically wrong or ambiguous. → *Counted as System Failure.*
 - **Type B (Execution Error):** If the participant admits to understanding the instruction correctly but made a manual slip (e.g., missed a strand, clumsy handling). → *Excluded from System Failure rate or noted as Noise.*
3. **Instruction Validation:** Before the user study, the generated instructions are cross-verified against the Ground Truth topology. If the generated text is mathematically correct but the user fails, it suggests a need for better *Natural Language Generation* (clarity improvement) rather than a logic failure.

Chương 3

Ethical & Social Impacts

3.1 Ethical & Social Impacts

This project extends beyond technical implementation to address significant cultural and social values, particularly in the context of preserving traditional Vietnamese craftsmanship.

3.1.1 Cultural Preservation

The system supports the digitization and preservation of traditional craft knowledge. By converting physical weaving patterns into digital data and explicit instructions, the project helps prevent the loss of traditional water hyacinth weaving techniques, ensuring this knowledge remains accessible to future generations.

3.1.2 Human-Centered Design

The technology is designed to *assist* rather than *replace* artisans.

- **For Artisans:** The system serves as a documentation tool, helping them record and archive their creative patterns.
- **For Learners:** It provides visual and textual guides, shortening the learning curve and making weaving techniques more approachable.

3.1.3 Knowledge Ownership

The research team is acutely aware of intellectual property and data ethics. Care must be taken to respect local communities and avoid unauthorized exploitation of traditional patterns. The usage of weaving designs must ensure transparency and proper attribution of indigenous knowledge.

»»»> 26686f86fe95d75738ec5f8dd57d5c86a3f5f16d