|  | TRƯỜNG ĐẠI HỌC THUỶ LỢI  **KHOA CÔNG NGHỆ THÔNG TIN**  BẢN TÓM TẮT ĐỀ CƯƠNG ĐỒ ÁN TỐT NGHIỆP |
| --- | --- |

**Tên đề tài: Ứng dụng Deep Learning trong việc tái tạo nhân vật hoạt hình dạng 3D từ các ảnh 2D**

*Sinh viên thực hiện*: Nguyễn Đức Minh

*Lớp*: 63HTTT1

*Mã sinh viên:* 2151163707

*Số điện thoại:* 0333550452

*Email:* ducminhxp.tp@gmail.com

*Giáo viên hướng dẫn*: Trần Mạnh Tuấn

**TÓM TẮT ĐỀ TÀI**

Trong đề tài này, em tập trung vào việc ứng dụng trí tuệ nhân tạo để tái tạo hình dạng 3D của các nhân vật hoạt hình từ ảnh 2D.

Đầu vào bài toán sẽ là một hình ảnh 2D và đầu ra bài toán sẽ là một mô hình 3D chính xác của nhân vật hoạt hình, được tái tạo dưới dạng lưới (mesh) 3D có độ phân giải cao. Mô hình này sẽ chứa đầy đủ các đặc điểm hình học và cấu trúc của nhân vật, có thể được sử dụng cho các ứng dụng như sản xuất phim hoạt hình, game, thực tế ảo (VR), và các hệ thống mô phỏng số.

Mục tiêu của nghiên cứu là sử dụng các phương pháp Deep Learning để phát triển một mô hình có khả năng tự động chuyển đổi hình ảnh nhân vật hoạt hình 2D thành mô hình 3D chính xác và chân thực.

Việc tái tạo mô hình 3D từ ảnh 2D là một bài toán phức tạp và có tính ứng dụng cao trong nhiều lĩnh vực như sản xuất phim hoạt hình, game, thực tế ảo (VR), in 3D, và mô phỏng nhân vật số. Việc phát triển một hệ thống có thể tái tạo 3D một cách chính xác sẽ giúp giảm thời gian và chi phí so với các phương pháp dựng hình thủ công truyền thống.

Để giải quyết bài toán này, em sử dụng các kỹ thuật tiên tiến trong Deep Learning, bao gồm **Generative Adversarial Networks (GANs), PIFuHD, và MeshCNN**. Các thuật toán này giúp mô hình học được thông tin chiều sâu từ ảnh 2D, tái dựng lại hình dạng nhân vật dưới dạng lưới 3D (mesh) có độ phân giải cao. Đồng thời, các phương pháp Remeshing và Topology Optimization cũng được áp dụng để cải thiện chất lượng mô hình sau khi sinh ra.

Hiện nay, nhiều nghiên cứu trên thế giới đã đạt được những kết quả đáng kể trong việc tái tạo mô hình 3D từ ảnh 2D. Tuy nhiên, các phương pháp hiện tại vẫn còn nhiều thách thức như mất chi tiết ở các vùng phức tạp, khó xử lý hình dạng đặc thù của nhân vật hoạt hình, hoặc đòi hỏi lượng dữ liệu huấn luyện lớn.

Trong đề tài này, em đề xuất một phương pháp kết hợp các mô hình học sâu tiên tiến với kỹ thuật tiền xử lý dữ liệu nhằm tăng độ chính xác và hiệu suất của mô hình tái tạo 3D. Kết quả nghiên cứu có thể được ứng dụng rộng rãi trong công nghiệp hoạt hình, phát triển trò chơi, và các hệ thống tương tác ảo.

**CÁC MỤC TIÊU CHÍNH**

* Mục tiêu 1: Phân tích và xử lý dữ liệu hình ảnh nhân vật hoạt hình 2D để chuẩn bị dữ liệu huấn luyện cho mô hình Deep Learning.
* Mục tiêu 2: Nghiên cứu và ứng dụng các kỹ thuật Deep Learning như GANs, PIFuHD, MeshCNN vào bài toán tái tạo hình dạng 3D từ ảnh 2D.
* Mục tiêu 3: Xây dựng và đánh giá hiệu suất mô hình trong việc tái tạo mô hình 3D có độ chính xác cao và tối ưu hóa chất lượng lưới (mesh).
* Mục tiêu 4: Phát triển một ứng dụng thử nghiệm cho phép nhập ảnh vẽ tay hoặc ảnh 2D của nhân vật hoạt hình và sinh ra mô hình 3D tương ứng.

**CÔNG CỤ VÀ PHƯƠNG PHÁP**

* Công cụ sử dụng:
* Ngôn ngữ lập trình: Python
* Cơ sở dữ liệu: Thư mục hệ thống
* Phần mềm quản lý mã nguồn: Git, Github
* Các công nghệ hỗ trợ:
* **TensorFlow, PyTorch**: Được sử dụng để xây dựng và huấn luyện mô hình Deep Learning (PiFuHD, MeshCNN, GANs).
* **Open3D, Trimesh:** Dùng để xử lý, phân tích và trực quan hóa mô hình 3D.
* **NumPy, Pandas:** Xử lý dữ liệu, đặc biệt là trong việc chuẩn bị và quản lý các bộ dữ liệu 2D và 3D.
* **Matplotlib, Seaborn:** Dùng để trực quan hóa kết quả mô hình và các số liệu phân tích.
* Phương pháp sử dụng:

Trong đề tài Tái tạo hình dạng 3D của các nhân vật hoạt hình từ ảnh 2D sử dụng Deep Learning, các phương pháp và thuật toán sau sẽ được áp dụng để xây dựng mô hình:

* Học sâu (Deep Learning): Sử dụng dữ liệu hình ảnh nhân vật hoạt hình 2D và áp dụng các mô hình học sâu (Deep Learning) như **PiFuHD**, **MeshCNN**, và **GANs** để tái tạo hình ảnh 3D. Các mô hình này sẽ học được các đặc trưng hình học và đặc điểm nhân vật hoạt hình từ hình ảnh 2D và tạo ra mô hình 3D tương ứng.
* **Mô hình PiFuHD:** PiFuHD (Pixel-Aligned Implicit Function for High-Resolution 3D Human Shape Reconstruction) là mô hình học sâu được thiết kế để tái tạo mô hình 3D từ hình ảnh 2D, với khả năng xử lý chi tiết cao và chính xác trong việc tái tạo các nhân vật hoạt hình. Mô hình này sẽ được sử dụng để tạo ra hình dạng 3D chính xác từ các ảnh 2D.
* **Mô hình MeshCNN:** MeshCNN sử dụng các mạng nơ-ron sâu để học và xử lý các thông tin hình học của mô hình 3D. MeshCNN có thể giúp cải thiện chất lượng mô hình 3D bằng cách học được các đặc trưng không gian của lưới tam giác (mesh) từ dữ liệu đầu vào.
* **Mô hình GANs (Generative Adversarial Networks):** GANs sẽ được sử dụng để cải thiện chất lượng và độ chân thực của các mô hình 3D. Bằng cách sử dụng một bộ sinh (generator) và bộ phân biệt (discriminator), GANs sẽ tối ưu hóa quá trình tái tạo hình ảnh 3D từ ảnh 2D, tạo ra các mô hình có độ phân giải và độ chi tiết cao.
* **Kết hợp các mô hình:** Việc kết hợp giữa PiFuHD, MeshCNN, và GANs tận dụng ưu điểm của từng mô hình: PiFuHD có thể tái tạo hình dạng 3D chi tiết từ ảnh 2D, MeshCNN giúp xử lý hình học của các mô hình 3D, còn GANs cải thiện chất lượng và sự chân thực của mô hình đầu ra. Trong mô hình kết hợp này, PiFuHD sẽ được sử dụng để trích xuất các đặc trưng hình học từ ảnh 2D, sau đó MeshCNN và GANs sẽ hoàn thiện mô hình 3D với độ chính xác và độ chi tiết cao hơn.

**NỘI DUNG CHÍNH**

* Tiền xử lý dữ liệu:
* **Tiền xử lý ảnh 2D**: Lọc ảnh mờ, nhiễu; phục hồi phần bị thiếu; chuẩn hóa kích thước, giá trị pixel; tăng cường dữ liệu bằng xoay, lật, điều chỉnh độ sáng.·
* **Tiền xử lý ảnh 3D**: Chuẩn hóa mô hình (mesh, voxel, point cloud); loại bỏ lỗi và điểm không hoàn chỉnh; chuyển đổi sang định dạng phù hợp với mô hình AI.
* **Đồng bộ dữ liệu 2D & 3D**: Ghép cặp chính xác ảnh 2D với mô hình 3D; áp dụng augmentation để tăng sự đa dạng dữ liệu.
* **Kiểm tra chất lượng**: Đánh giá và loại bỏ dữ liệu không đồng nhất để đảm bảo đầu vào chất lượng cao
* Xây dựng mô hình Deep Learning:
* Lựa chọn mô hình:
* **PiFuHD**: Mô hình tái tạo hình dạng 3D chi tiết từ ảnh 2D bằng cách sử dụng implicit functions.
* **MeshCNN**: Mô hình chuyên xử lý dữ liệu dưới dạng mesh để tối ưu hóa bề mặt 3D.
* **GANs**: Sử dụng để cải thiện tính chân thực và tạo chi tiết cho mô hình 3D.
* **Kết hợp mô hình**: PiFuHD để tái tạo hình dạng thô, MeshCNN để tinh chỉnh bề mặt, GANs để bổ sung chi tiết giúp cải thiện độ chính xác.
* Huấn luyện và tối ưu hóa mô hình:
* **Huấn luyện với bộ dữ liệu 2D-3D**: Sử dụng các tập dữ liệu đã xử lý để huấn luyện mô hình, đảm bảo mô hình học được cách chuyển đổi từ ảnh 2D sang 3D.
* **Điều chỉnh siêu tham số**: Tối ưu learning rate, batch size và kiến trúc mạng để cải thiện độ chính xác và tốc độ huấn luyện.
* **Sử dụng loss functions phù hợp**: Áp dụng các hàm mất mát như Chamfer Distance, L1 Loss để tối ưu kết quả mô hình 3D.
* Đánh giá và so sánh mô hình:
* **Đánh giá hiệu suất**: Sử dụng các chỉ số như Intersection over Union (IoU), Chamfer Distance để đo độ chính xác của mô hình 3D tái tạo.:
* **So sánh các mô hình**: Đánh giá PiFuHD, MeshCNN và GANs để xác định mô hình nào tái tạo chi tiết tốt hơn, mô hình nào nhanh hơn.
* **Phân tích lỗi và cải thiện**: Kiểm tra các lỗi của mô hình, tinh chỉnh hyperparameters hoặc kết hợp thêm các kỹ thuật như adversarial training để tăng độ chính xác.
* Triển khai và ứng dụng thực tế:
* **Xây dựng hệ thống ứng dụng**: Tích hợp mô hình vào một hệ thống cho phép người dùng tải ảnh 2D lên và nhận lại mô hình 3D.
* **Tối ưu hóa tốc độ xử lý**: Giảm kích thước mô hình, sử dụng TensorRT hoặc ONNX để tăng tốc độ inferencing.
* **Ứng dụng trong thực tế**: Hỗ trợ tạo mô hình 3D cho game, phim hoạt hình, hoặc thiết kế nhân vật số (digital avatars).

**KẾT QUẢ DỰ KIẾN**

* Báo cáo tổng kết đồ án
* Chương trình demo hệ thống tái tạo hình ảnh 3D của các nhân vật hoạt hình từ hình ảnh 2D bằng các thuật toán DeepLearning.

**KẾ HOẠCH THỰC HIỆN**

| **STT** | **Thời gian** | **Nội dung công việc** | **Kết quả dự kiến đạt được** |
| --- | --- | --- | --- |
| 1 | 31/03/2025 -13/04/2025 (2 tuần) | Tìm hiểu các mô hình tái tạo hình 3D từ ảnh 2D, đặc biệt là GANs, Autoencoder, PIFuHD và MeshCNN. | Hiểu rõ nguyên lý hoạt động và cách áp dụng của các mô hình này trong bài toán tái tạo 3D. |
| 2 | 14/04/2025 - 27/04/2025 (2 tuần) | Thu thập và xử lý dữ liệu hình ảnh 2D và mô hình 3D tương ứng từ các nguồn như dataset ShapeNet, RenderPeople, hoặc tự tạo dữ liệu. | Có tập dữ liệu đầy đủ để huấn luyện mô hình. |
| 3 | |  | | --- | | 28/04/2025- 11/05/2025 (1 tuần) | | | Tiền xử lý dữ liệu: làm sạch ảnh, loại bỏ nhiễu, chuẩn hóa kích thước, chuyển đổi định dạng cần thiết cho huấn luyện. | Dữ liệu sẵn sàng để đưa vào mô hình học sâu. |
| 4 | | 05/05/2025 - 11/05/2025 (1 tuần) | | | --- | --- | |  | | Trích xuất đặc trưng từ ảnh 2D, thử nghiệm các kỹ thuật tăng cường dữ liệu để cải thiện kết quả mô hình. | Dữ liệu tối ưu hơn, giúp mô hình học hiệu quả hơn. |
| 5 | 12/05/2025 - 25/05/2025 (2 tuần) | Xây dựng và huấn luyện mô hình Deep Learning (PIFuHD hoặc GAN-based) để tái tạo hình dạng 3D từ ảnh 2D. | Huấn luyện được mô hình chuyển đổi từ ảnh 2D sang 3D. |
| 6 | 26/05/2025 - 01/06/2025 (1 tuần) | Đánh giá mô hình dựa trên độ chính xác hình dạng, độ mịn của bề mặt và tính thực tế của mô hình 3D tạo ra. | Xác định mô hình tối ưu nhất để triển khai. |
| 7 | 02/06/2025 - 08/06/2025 (1 tuần) | Xây dựng API để cho phép tải ảnh 2D lên và trả về mô hình 3D tương ứng. | API hoạt động, có thể nhận ảnh 2D đầu vào và trả về kết quả 3D. |
| 8 | 09/06/2025 - 15/06/2025 (1 tuần) | Xây dựng giao diện demo bằng Qt để hiển thị kết quả mô hình 3D. | Ứng dụng có giao diện trực quan, hiển thị mô hình 3D được tạo ra. |
| 9 | 16/06/2025 - 22/06/2025 (1 tuần) | Kiểm tra lại toàn bộ hệ thống, tối ưu hiệu suất, cải thiện thời gian xử lý. | Hệ thống hoàn thiện, tối ưu hiệu suất sẵn sàng triển khai. |
| 10 | 23/06/2025 - 06/07/2025 (2 tuần) | Viết báo cáo tổng kết và chuẩn bị bài bảo vệ. | Báo cáo hoàn chỉnh, sẵn sàng bảo vệ đề tài. |