

ỦY BAN NHÂN DÂN TP HỒ CHÍ MINH

TRƯỜNG ĐẠI HỌC SÀI GÒN

KHOA CÔNG NGHỆ THÔNG TIN

~~~~~\*~~~~~



**BÁO CÁO CUỐI KÌ**

**CÁC CÔNG NGHỆ LẬP TRÌNH HIỆN ĐẠI  
TRONG LĨNH VỰC XỬ LÝ HÌNH ẢNH**

**Sinh viên thực hiện:**

Trần Đồng Gia Hân

3121411066

**Lớp: DCT121C3**

Giảng viên hướng dẫn: TS. Đỗ Như Tài

*TP. Hồ Chí Minh, tháng 05 năm 2025*

# MỤC LỤC

|                                                                         |    |
|-------------------------------------------------------------------------|----|
| MỤC LỤC.....                                                            | i  |
| DANH MỤC HÌNH ẢNH .....                                                 | ii |
| CHƯƠNG 1: GIỚI THIỆU .....                                              | 1  |
| 1.1 Đặt vấn đề .....                                                    | 1  |
| 1.2 Khó khăn của phương pháp thủ công truyền thống .....                | 1  |
| 1.3 Mục tiêu của đề tài.....                                            | 1  |
| 1.4 Khái niệm nhận diện hình ảnh.....                                   | 2  |
| 1.4.1 Các bài toán phổ biến trong nhận diện hình ảnh .....              | 2  |
| 1.5 Các công nghệ và mô hình phổ biến trong nhận diện hình ảnh.....     | 3  |
| 1.5.1 Mạng nơ-ron tích chập (CNN - Convolutional Neural Networks) ..... | 3  |
| 1.5.2 Mô hình phát hiện đối tượng (Object Detection Models).....        | 3  |
| 1.5.3 Mô hình phân vùng ảnh (Segmentation Models).....                  | 3  |
| 1.5.4 Thư viện và Framework hỗ trợ.....                                 | 4  |
| CHƯƠNG 2: CƠ SỞ LÝ THUYẾT .....                                         | 4  |
| 2.1 Tổng quan về Deep Learning.....                                     | 4  |
| 2.1.1 Các mạng nơ-ron trong Deep Learning.....                          | 4  |
| 2.1.2 Quá trình huấn luyện trong Deep Learning .....                    | 5  |
| 2.2 Các mô hình học sâu thường dùng trong thị giác máy tính.....        | 5  |
| 2.2.1 Mạng nơ-ron tích chập (CNN - Convolutional Neural Networks) ..... | 5  |
| 2.2.2 Mô hình phân vùng ảnh (Segmentation Models).....                  | 6  |
| 2.3 Kiến thức về xử lý ảnh y tế .....                                   | 7  |
| 2.3.1 Các loại ảnh y tế phổ biến .....                                  | 7  |
| 2.3.2 Tiền xử lý ảnh y tế .....                                         | 7  |
| 2.4 Kiến trúc U-Net và/hoặc DeepLabV3+ cho phân vùng khối u.....        | 7  |
| 2.4.1 Kiến trúc U-Net .....                                             | 8  |
| 2.4.2 Kiến trúc DeepLabV3+ .....                                        | 8  |
| CHƯƠNG 3: THIẾT KẾ HỆ THỐNG .....                                       | 13 |
| 3.1 Kiến trúc hệ thống .....                                            | 13 |

|                                                               |                                     |
|---------------------------------------------------------------|-------------------------------------|
| 3.2 Phân chia theo Clean Architecture.....                    | 14                                  |
| Domain Layer.....                                             | 14                                  |
| Use Case Layer .....                                          | 14                                  |
| Interface/Adapter Layer .....                                 | 14                                  |
| External Layer .....                                          | 15                                  |
| 3.3 Mô hình AI.....                                           | 15                                  |
| 3.3.1 Kiến trúc mô hình.....                                  | 15                                  |
| 3.3.2 Tiền xử lý ảnh.....                                     | 15                                  |
| 3.4 Phân vùng khối u trong ảnh sinh thiết .....               | 16                                  |
| 3.5 Hệ thống phân tầng.....                                   | 16                                  |
| CHƯƠNG 4: DỮ LIỆU.....                                        | 16                                  |
| 4.1 Nguồn Dữ Liệu .....                                       | 16                                  |
| 4.2 Tiền Xử Lý Dữ Liệu .....                                  | 17                                  |
| 4.2.1 Chuyển đổi kích thước (Resize) .....                    | 17                                  |
| 4.2.2 Chuẩn hóa ảnh (Normalization) .....                     | 17                                  |
| 4.2.3 Tăng cường dữ liệu (Data Augmentation) .....            | 17                                  |
| 4.2.4 Tạo mask (cho bài toán phân vùng) .....                 | 18                                  |
| 4.3 Chia Dữ Liệu .....                                        | 18                                  |
| 4.4 Các Thách Thức trong Quá Trình Làm Việc với Dữ Liệu ..... | 19                                  |
| 4.5 Định Dạng Dữ Liệu .....                                   | 19                                  |
| 4.6 Kết Luận .....                                            | <b>Error! Bookmark not defined.</b> |
| CHƯƠNG 5: HUẤN LUYỆN MÔ HÌNH.....                             | 19                                  |
| 5.1 Cấu Hình Huấn Luyện.....                                  | 19                                  |
| 5.1.1 Số Epochs .....                                         | 20                                  |
| 5.1.2 Batch Size.....                                         | 20                                  |
| 5.1.3 Learning Rate.....                                      | 20                                  |
| 5.1.4 Optimizer .....                                         | 20                                  |
| 5.1.5 Loss Function .....                                     | 21                                  |
| 5.1.6 Metric Đánh Giá.....                                    | 21                                  |
| 5.2 Quy Trình Huấn Luyện.....                                 | 22                                  |

|                                                          |                                     |
|----------------------------------------------------------|-------------------------------------|
| 5.2.1 Chuẩn Bị Dữ Liệu.....                              | 22                                  |
| 5.2.2 Huấn Luyện Mô Hình .....                           | 22                                  |
| 5.2.3 Điều Chỉnh Tham Số.....                            | 22                                  |
| 5.2.4 Đánh Giá Mô Hình .....                             | 23                                  |
| 5.3 Biểu Đồ Huấn Luyện.....                              | 23                                  |
| 5.4 Thử Nghiệm Nhiều Mô Hình .....                       | 23                                  |
| 5.5 Kết Luận .....                                       | <b>Error! Bookmark not defined.</b> |
| <b>CHƯƠNG 6: KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN .....</b>      | <b>23</b>                           |
| 6.1 Tổng Kết .....                                       | 23                                  |
| Những điểm nổi bật của dự án: .....                      | 24                                  |
| 6.2 Hướng Phát Triển.....                                | 24                                  |
| 6.2.1 Cải Tiến Mô Hình .....                             | 24                                  |
| 6.2.2 Mở Rộng Dữ Liệu .....                              | 24                                  |
| 6.2.3 Tích Hợp Công Nghệ AI Vào Quy Trình Chẩn Đoán..... | 25                                  |
| 6.2.4 Mở Rộng Ứng Dụng.....                              | 25                                  |
| 6.2.5 Tối Ưu Hóa và Triển Khai Thực Tế.....              | 25                                  |
| 6.3 Kết Luận Chung .....                                 | 26                                  |

## DANH MỤC MÌNH ẢNH

|                                             |                                     |
|---------------------------------------------|-------------------------------------|
| 1.1. Minh họa bài toán Phân loại ảnh 1..... | <b>Error! Bookmark not defined.</b> |
|---------------------------------------------|-------------------------------------|

# CHƯƠNG 1: GIỚI THIỆU

## 1.1. Đặt vấn đề

Ung thư xương là một trong những loại ung thư hiếm gặp nhưng rất nguy hiểm, chủ yếu ảnh hưởng đến trẻ em và thanh thiếu niên. Việc phát hiện ung thư xương sớm và chính xác là yếu tố quyết định trong quá trình điều trị và tỷ lệ sống sót của bệnh nhân. Tuy nhiên, hiện nay, việc chẩn đoán ung thư xương chủ yếu dựa vào các xét nghiệm hình ảnh sinh thiết mô (biopsy images). Các bác sĩ phải phân tích hàng trăm mẫu ảnh mỗi ngày, dẫn đến một số vấn đề như sai sót trong chẩn đoán do sự mệt mỏi và thiếu công cụ hỗ trợ.

Để giải quyết vấn đề này, công nghệ trí tuệ nhân tạo (AI) và học sâu (Deep Learning) có thể đóng vai trò quan trọng trong việc tự động phân tích và nhận diện các dấu hiệu ung thư từ hình ảnh sinh thiết. Hệ thống này sẽ giúp tăng độ chính xác, giảm thời gian chẩn đoán và hỗ trợ bác sĩ trong quá trình ra quyết định.

## 1.2. Khó khăn của phương pháp thủ công truyền thống

Các phương pháp thủ công trong việc phân tích ảnh sinh thiết gặp phải một số vấn đề như:

- **Tốn thời gian:** Việc phân tích hàng nghìn bức ảnh mỗi ngày rất tốn thời gian và đòi hỏi nhiều công sức từ các bác sĩ.
- **Phụ thuộc vào kinh nghiệm:** Quyết định cuối cùng trong việc chẩn đoán rất dễ bị ảnh hưởng bởi kinh nghiệm của bác sĩ, dẫn đến sự thay đổi kết quả giữa các chuyên gia.
- **Khó phát hiện giai đoạn đầu:** Sự khác biệt giữa mô ung thư và mô lành rất mờ hồ trong giai đoạn đầu, khiến cho việc phát hiện kịp thời trở nên khó khăn.
- **Thiếu công cụ hỗ trợ:** Những bệnh viện tuyến dưới hoặc các cơ sở y tế chưa có đủ trang thiết bị và chuyên gia để phân tích ảnh một cách chính xác.

## 1.3. Mục tiêu của đề tài

Mục tiêu của đề tài là xây dựng một hệ thống ứng dụng học sâu để nhận diện ung thư xương từ hình ảnh sinh thiết, giúp bác sĩ trong quá trình chẩn đoán bệnh. Hệ

thống này sẽ sử dụng các mô hình học sâu như CNN, U-Net và DeepLabV3+ để phân tích và phát hiện các dấu hiệu ung thư, cung cấp thông tin chính xác và nhanh chóng.

#### **1.4. Khái niệm nhận diện hình ảnh**

Nhận diện hình ảnh (Image Recognition) là một nhánh quan trọng của thị giác máy tính (Computer Vision) và trí tuệ nhân tạo (AI). Đây là quá trình tự động phân tích, nhận dạng và phân loại các đối tượng hoặc đặc điểm có trong hình ảnh kỹ thuật số. Nhận diện hình ảnh giúp máy tính hiểu nội dung của hình ảnh bằng cách sử dụng các thuật toán xử lý ảnh và học sâu (Deep Learning). Công nghệ này đã được ứng dụng rộng rãi trong nhiều lĩnh vực như y tế, an ninh, giao thông, thương mại điện tử và nhận diện khuôn mặt.

#### **Các bài toán phổ biến trong nhận diện hình ảnh**

1. **Phân loại ảnh (Image Classification):** Là bài toán xác định ảnh đầu vào thuộc về một hoặc nhiều danh mục đã biết. Ví dụ: phân loại hình ảnh động vật thành các nhóm như chó, mèo, ngựa, v.v. Các mô hình phổ biến cho bài toán này bao gồm:
  - LeNet-5
  - AlexNet
  - VGGNet
  - ResNet
2. **Phát hiện đối tượng (Object Detection):** Là bài toán xác định và định vị các đối tượng trong một hình ảnh bằng cách vẽ các hộp giới hạn (bounding box). Các ứng dụng phổ biến của phát hiện đối tượng bao gồm giám sát an ninh, xe tự lái và phát hiện dị vật trong y tế. Các mô hình phổ biến:
  - YOLO (You Only Look Once)
  - SSD (Single Shot MultiBox Detector)
  - Faster R-CNN
3. **Phân vùng ảnh (Image Segmentation):** Là bài toán phân chia hình ảnh thành nhiều vùng khác nhau để xác định đối tượng một cách chính xác hơn. Các loại phân vùng ảnh bao gồm:

- Phân vùng theo pixel (Semantic Segmentation)
- Phân vùng theo đối tượng (Instance Segmentation)

Các mô hình phổ biến:

- U-Net
- DeepLabV3+
- Mask R-CNN

4. **Nhận diện khuôn mặt (Face Recognition):** Là bài toán nhận diện và phân tích các đặc điểm khuôn mặt để xác định danh tính của một người. Các mô hình sử dụng trong nhận diện khuôn mặt bao gồm các mạng học sâu như FaceNet, ArcFace, v.v.

### 1.5. Các công nghệ và mô hình phổ biến trong nhận diện hình ảnh

#### 1.5.1. Mạng nơ-ron tích chập (CNN - Convolutional Neural Networks)

Mạng nơ-ron tích chập (CNN) là một loại mạng nơ-ron nhân tạo được thiết kế đặc biệt để xử lý dữ liệu có cấu trúc lưới, chẳng hạn như hình ảnh. CNN sử dụng các lớp convolution để tự động trích xuất các đặc trưng từ ảnh mà không cần phải dựa vào các đặc trưng thủ công. Các mô hình CNN nổi bật có thể kể đến như LeNet, AlexNet, VGG, ResNet, và EfficientNet.

#### 1.5.2. Mô hình phát hiện đối tượng (Object Detection Models)

Phát hiện đối tượng trong hình ảnh là một trong những bài toán quan trọng trong thị giác máy tính. Các mô hình phát hiện đối tượng phổ biến bao gồm YOLO, SSD, Faster R-CNN, RetinaNet. Các mô hình này không chỉ xác định được đối tượng mà còn vẽ các hộp giới hạn (bounding boxes) xung quanh đối tượng, hỗ trợ rất nhiều trong các ứng dụng an ninh, xe tự lái và nhận diện bệnh lý y tế.

#### 1.5.3. Mô hình phân vùng ảnh (Segmentation Models)

Phân vùng ảnh là quá trình chia hình ảnh thành nhiều vùng có ý nghĩa để dễ dàng phân tích các đối tượng. Các mô hình phân vùng ảnh như U-Net và DeepLabV3+ đặc



biệt hữu ích trong các ứng dụng y tế, nơi mà việc phân vùng các vùng mô bị ung thư là rất quan trọng để chẩn đoán chính xác.

#### 1.5.4. Thư viện và Framework hỗ trợ

Các thư viện và framework hỗ trợ xây dựng các mô hình nhận diện hình ảnh trong thị giác máy tính bao gồm:

- **TensorFlow** và **Keras**: Các thư viện phổ biến để xây dựng và huấn luyện các mô hình học sâu.
- **PyTorch**: Một framework khác rất mạnh mẽ và linh hoạt, đặc biệt được sử dụng trong nghiên cứu học sâu.
- **OpenCV**: Một thư viện mạnh mẽ trong xử lý ảnh và video, hỗ trợ nhiều thuật toán thị giác máy tính.
- **Scikit-learn**: Thư viện cho học máy truyền thống, cũng có thể tích hợp với các mô hình học sâu.

## CHƯƠNG 2: CƠ SỞ LÝ THUYẾT

### 2.1. Tổng quan về Deep Learning

Deep Learning (học sâu) là một nhánh con của học máy, dựa trên các mạng nơ-ron nhân tạo (Artificial Neural Networks - ANN) để mô phỏng các quá trình học tự động. Khác với các thuật toán học máy truyền thống, học sâu có khả năng tự động học và trích xuất đặc trưng từ dữ liệu mà không cần phải có sự can thiệp của con người trong việc thiết kế đặc trưng. Deep Learning đã chứng minh được khả năng vượt trội trong việc xử lý các vấn đề phức tạp như nhận diện hình ảnh, nhận diện giọng nói, và xử lý ngôn ngữ tự nhiên.

#### 2.1.1 Các mạng nơ-ron trong Deep Learning

- **Mạng nơ-ron đa lớp (MLP - Multi-Layer Perceptron):** Là kiểu mạng nơ-ron đơn giản nhất, bao gồm các lớp đầu vào, lớp ẩn và lớp đầu ra. MLP chủ yếu được sử dụng trong các bài toán phân loại đơn giản.
- **Mạng nơ-ron tích chập (CNN - Convolutional Neural Networks):** CNN là một kiến trúc đặc biệt trong Deep Learning được tối ưu hóa để xử lý hình ảnh. CNN sử dụng các lớp chập (convolutional layers) để trích xuất các đặc trưng hình ảnh và các lớp pooling (max pooling hoặc average pooling) để giảm số lượng tham số trong mô hình.
- **Mạng nơ-ron tái phát (RNN - Recurrent Neural Networks):** RNN được sử dụng cho các dữ liệu có tính chất chuỗi, như văn bản hoặc âm thanh, với khả năng nhớ lại thông tin từ quá khứ.

### 2.1.2 Quá trình huấn luyện trong Deep Learning

Quá trình huấn luyện một mô hình học sâu bao gồm ba bước chính:

1. **Tiền xử lý dữ liệu:** Dữ liệu đầu vào được chuẩn hóa và chuyển đổi thành định dạng phù hợp cho việc huấn luyện. Trong bài toán nhận diện hình ảnh, ảnh sẽ được chia nhỏ, làm sạch và chuẩn hóa kích thước.
2. **Huấn luyện mô hình:** Dựa trên các thuật toán tối ưu (như Gradient Descent, Adam Optimizer), mô hình sẽ học từ dữ liệu huấn luyện. Quá trình này kéo dài cho đến khi mô hình đạt được độ chính xác đủ cao.
3. **Đánh giá và tinh chỉnh:** Sau khi mô hình huấn luyện xong, nó được kiểm tra trên bộ dữ liệu kiểm tra để đánh giá khả năng tổng quát hóa. Các tham số của mô hình có thể được điều chỉnh để cải thiện hiệu suất.

## 2.2. Các mô hình học sâu thường dùng trong thị giác máy tính

Trong thị giác máy tính, các mô hình học sâu đã được áp dụng rộng rãi để giải quyết các bài toán nhận diện hình ảnh, phát hiện đối tượng, và phân vùng ảnh. Dưới đây là các mô hình nổi bật thường được sử dụng:

### 2.2.1. Mạng nơ-ron tích chập (CNN - Convolutional Neural Networks)

Mạng CNN được thiết kế đặc biệt để xử lý các dữ liệu có cấu trúc hình ảnh, giúp máy tính nhận diện được các đối tượng trong ảnh thông qua các lớp chập. Các lớp chập sẽ tự động trích xuất các đặc trưng như cạnh, góc và hình dạng từ ảnh đầu vào.

Các mô hình CNN phổ biến:

- **LeNet-5:** Một trong những mô hình CNN đầu tiên, được sử dụng cho nhận diện ký tự viết tay trong các hệ thống như OCR.
- **AlexNet:** Một trong những mô hình CNN nổi bật, giúp CNN trở nên phổ biến trong cộng đồng nghiên cứu và ứng dụng. Nó sử dụng nhiều lớp chập và có thể nhận diện các đối tượng trong ảnh với độ chính xác cao.
- **VGGNet:** Đặc trưng bởi số lượng lớp chập rất sâu, giúp trích xuất các đặc trưng rất chi tiết của ảnh.
- **ResNet:** Với các lớp Residual Connection, ResNet giải quyết vấn đề suy giảm hiệu suất khi số lớp của mô hình quá sâu. Nó giúp tăng hiệu quả huấn luyện và độ chính xác.

### 2.2.2. Mô hình phân vùng ảnh (Segmentation Models)

Phân vùng ảnh là một bài toán quan trọng trong thị giác máy tính, giúp chia một ảnh thành các vùng tương ứng với các đối tượng hoặc phần quan trọng trong ảnh. Phân vùng ảnh giúp phân tích chi tiết hơn các đối tượng trong hình ảnh, đặc biệt là trong các ứng dụng y tế như phân tích ảnh sinh thiết.

Các mô hình phân vùng ảnh phổ biến:

- **U-Net:** Một mô hình phân vùng ảnh được thiết kế đặc biệt cho các ứng dụng y tế. U-Net có kiến trúc đặc biệt với các phần encoder (phần giảm kích thước) và decoder (phần phục hồi kích thước) giúp tách rời các đối tượng trong ảnh. U-Net là mô hình rất hiệu quả trong việc phân vùng các mô hình y tế như tế bào ung thư, mô u, v.v.
- **DeepLabV3+:** Đây là một mô hình phân vùng ảnh tiên tiến được Google phát triển, sử dụng mạng CNN để phân vùng hình ảnh với độ chính xác cao.

DeepLabV3+ sử dụng các kỹ thuật như atrous convolution để nâng cao độ phân giải của phân vùng.

### 2.3. Kiến thức về xử lý ảnh y tế

Trong y tế, xử lý ảnh đóng một vai trò quan trọng trong việc hỗ trợ chẩn đoán và điều trị. Các loại ảnh y tế, đặc biệt là ảnh sinh thiết và ảnh chụp cắt lớp, giúp bác sĩ xác định tình trạng bệnh lý và theo dõi sự phát triển của bệnh.

#### 2.3.1. Các loại ảnh y tế phổ biến

- **Ảnh X-quang:** Dùng để phát hiện các bệnh lý về xương, phổi, tim, v.v.
- **MRI (Magnetic Resonance Imaging):** Dùng trong việc chụp các cấu trúc mềm của cơ thể như não, tủy sống, và cơ bắp.
- **CT Scan (Computed Tomography):** Cung cấp hình ảnh 3D của cơ thể, đặc biệt hữu ích trong việc xác định vị trí và kích thước của khối u.
- **Ảnh sinh thiết (biopsy images):** Là những hình ảnh mô được chụp sau khi sinh thiết tế bào hoặc mô từ cơ thể bệnh nhân. Những bức ảnh này thường được dùng để phân tích các tế bào ung thư.

#### 2.3.2. Tiền xử lý ảnh y tế

Trước khi đưa vào mô hình học sâu, các ảnh y tế thường cần được xử lý qua một số bước tiền xử lý như:

- **Tiêu chuẩn hóa:** Điều chỉnh kích thước ảnh về một chuẩn chung, chẳng hạn như 224x224 pixels.
- **Phân vùng:** Dùng các mô hình phân vùng ảnh (như U-Net) để tách biệt các vùng chứa đối tượng quan trọng, ví dụ như tế bào ung thư.
- **Augmentation:** Các kỹ thuật tăng cường dữ liệu giúp cải thiện khả năng tổng quát của mô hình, như xoay ảnh, làm mờ, thay đổi độ sáng.

---

### 2.4. Kiến trúc U-Net và/hoặc DeepLabV3+ cho phân vùng khối u

Kiến trúc **U-Net** và **DeepLabV3+** là hai mô hình phân vùng ảnh đặc biệt hiệu quả trong các bài toán y tế, bao gồm phân vùng khối u trong ảnh sinh thiết.

#### 2.4.1. Kiến trúc U-Net

U-Net là một kiến trúc mạng nơ-ron tích chập được thiết kế đặc biệt cho phân vùng ảnh y tế. Với đặc điểm là một mạng đối xứng, U-Net gồm hai phần:

- **Encoder:** Phần này giảm dần kích thước của ảnh đầu vào để trích xuất các đặc trưng.
- **Decoder:** Phục hồi kích thước ảnh để tạo ra các phân vùng chính xác với kích thước ban đầu.

Ưu điểm của U-Net là khả năng phân vùng chính xác với số lượng dữ liệu huấn luyện nhỏ và hiệu quả trong phân tích ảnh sinh thiết.

#### 2.4.2. Kiến trúc DeepLabV3+

DeepLabV3+ là một mô hình phân vùng ảnh sử dụng các kỹ thuật tiên tiến như atrous convolution để cải thiện độ phân giải của phân vùng. DeepLabV3+ đã được áp dụng thành công trong các bài toán nhận diện khối u trong ảnh y tế nhờ khả năng phân vùng chi tiết và hiệu quả.

### 2.5. Khảo sát và phân tích các mô hình học sâu tiêu biểu

Để có cái nhìn trực quan và sâu sắc hơn về các kiến trúc học sâu, nhóm đã tiến hành khảo sát, thực hành và phân tích một số mô hình học sâu tiêu biểu thông qua các notebook mẫu trong thư mục intro\_deep. Các mô hình này không chỉ giúp nhóm làm quen với công cụ và quy trình huấn luyện mô hình học sâu, mà còn cung cấp nền tảng vững chắc để từ đó phát triển mô hình nhận diện ung thư xương dựa trên ảnh sinh thiết trong các chương tiếp theo.

Những notebook này trải dài từ các mô hình cơ bản như MLP (Multilayer Perceptron), CNN (Convolutional Neural Network) cho đến mô hình kết hợp CNN–LSTM và các

ứng dụng thực tế như nhận diện cảm xúc từ ảnh khuôn mặt. Dưới đây là phân tích chi tiết từng notebook:

### **2.5.1. Giới thiệu TensorFlow thông qua mô hình cơ bản**

**(getting\_started.ipynb)**

#### **Mục tiêu:**

Notebook này giới thiệu một pipeline đơn giản để huấn luyện mô hình học máy sử dụng TensorFlow, phù hợp với những người mới bắt đầu.

#### **Nội dung chính:**

- Nhập các thư viện như TensorFlow, NumPy, và matplotlib.
- Tạo dữ liệu giả hoặc sử dụng dataset nhỏ có sẵn từ Keras.datasets.
- Xây dựng mô hình bằng API Sequential gồm các lớp Dense.
- Biên dịch mô hình với hàm mất mát, bộ tối ưu và chỉ số đánh giá.
- Huấn luyện mô hình với model.fit(), sau đó đánh giá bằng model.evaluate() và dự đoán bằng model.predict().

#### **Ý nghĩa:**

Notebook này minh họa đầy đủ quy trình phát triển mô hình học sâu: từ khâu xử lý dữ liệu, xây dựng mô hình, huấn luyện, cho đến đánh giá và dự đoán. Đây là nền tảng cần thiết để hiểu cách các mô hình học sâu vận hành trước khi tiếp cận các mô hình phức tạp hơn.

### **2.5.2. Ứng dụng mạng MLP trong phân loại ảnh phẳng**

**(Func\_apxm\_MLP.ipynb)**

#### **Mục tiêu:**

Áp dụng mạng Perceptron đa tầng (MLP) để giải quyết bài toán phân loại ảnh, chẳng hạn như ảnh chữ số MNIST đã được làm phẳng (flattened).

#### **Cấu trúc mô hình:**

- Lớp Flatten: Biến ảnh 2D thành vector 1D.

- Lớp ẩn Dense(128, activation='relu').
- Lớp ẩn tiếp theo Dense(64, activation='relu').
- Lớp đầu ra Dense(10, activation='softmax') để phân loại 10 chữ số.

#### **Thông số huấn luyện:**

- Optimizer: Adam.
- Loss: sparse\_categorical\_crossentropy.
- Chỉ số đánh giá: Accuracy.

#### **Đánh giá:**

- MLP dễ cài đặt và phù hợp cho dữ liệu có chiều thấp.
- Tuy nhiên, mô hình không tận dụng được cấu trúc không gian trong ảnh, do đó hiệu quả sẽ kém hơn CNN khi áp dụng vào bài toán xử lý ảnh.

### **2.5.3. Mạng CNN trong phân loại ảnh MNIST (1\_MNIST\_CNN.ipynb)**

#### **Mục tiêu:**

Xây dựng mô hình CNN để phân loại ảnh chữ số viết tay từ tập dữ liệu MNIST, sử dụng các lớp tích chập và pooling để trích xuất đặc trưng.

#### **Cấu trúc mô hình:**

- Conv2D(32, kernel\_size=(3,3), activation='relu')
- MaxPooling2D(pool\_size=(2,2))
- Conv2D(64)
- Flatten
- Dense(64, activation='relu')
- Dense(10, activation='softmax')

#### **Kết quả huấn luyện:**

- Với cấu hình phù hợp và số epoch đủ lớn (khoảng 10), mô hình đạt độ chính xác trên 98% trên tập test.

### **Ý nghĩa:**

Đây là ví dụ điển hình cho thấy sức mạnh của CNN trong xử lý ảnh nhờ khả năng học đặc trưng cục bộ và giảm nhiễu. Kinh nghiệm từ mô hình này rất hữu ích cho việc phát triển hệ thống nhận diện ảnh sinh thiết sau này.

#### **2.5.4. Kết hợp mô hình CNN và LSTM trong xử lý ảnh tuần tự (1\_MNIST\_CNN\_LSTM.ipynb)**

### **Mục tiêu:**

Tận dụng đồng thời khả năng trích xuất đặc trưng của CNN và khả năng học thông tin chuỗi của LSTM để cải thiện hiệu quả phân loại ảnh MNIST.

### **Chi tiết mô hình:**

- CNN để trích xuất đặc trưng không gian từ ảnh  $28 \times 28$ .
- Sau đó reshape đầu ra CNN để đưa vào LSTM.
- LSTM học mối liên hệ giữa các vùng đặc trưng theo trật tự hàng ảnh.

### **Cấu trúc:**

- Conv2D  $\rightarrow$  MaxPooling  $\rightarrow$  Flatten  $\rightarrow$  Reshape  $\rightarrow$  LSTM  $\rightarrow$  Dense(10, softmax)

### **Ý nghĩa:**

Kết hợp CNN và LSTM là một hướng tiếp cận hiệu quả trong các bài toán ảnh động hoặc có tính chuỗi, ví dụ như nhận diện chữ viết tay thời gian thực hoặc video ảnh y tế.

#### **2.5.5. Ứng dụng CNN trong nhận diện cảm xúc từ ảnh khuôn mặt (2\_Emotion\_Reg\_CNN.ipynb)**

### **Mục tiêu:**

Phát triển mô hình học sâu phân loại cảm xúc (vui, buồn, giận, ngạc nhiên...) từ ảnh khuôn mặt người.



### **Đặc điểm dữ liệu:**

- Ảnh grayscale (thường  $48 \times 48$ ).
- Số lớp tương ứng với số cảm xúc.

### **Cấu trúc mô hình:**

- Conv2D(32)  $\rightarrow$  MaxPooling
- Conv2D(64)  $\rightarrow$  MaxPooling
- Flatten  $\rightarrow$  Dense(128)  $\rightarrow$  Dropout  $\rightarrow$  Dense(output, softmax)

### **Chiến lược huấn luyện:**

- Sử dụng Adam optimizer và categorical crossentropy.
- Đánh giá hiệu quả mô hình qua độ chính xác và confusion matrix.

### **Ứng dụng thực tế:**

- Nhận diện cảm xúc người dùng qua camera.
- Giao tiếp người – máy (robot, trợ lý ảo).

#### **2.5.6. Các module phụ trợ hỗ trợ huấn luyện mô hình**

Ngoài các notebook chính, thư mục còn chứa nhiều module phụ trợ nằm trong các thư mục như utils, callbacks, hoặc layers. Các module này đóng vai trò quan trọng trong quá trình huấn luyện và bảo trì mô hình.

### **Một số chức năng chính:**

- Tiền xử lý dữ liệu: resize ảnh, augment dữ liệu, chuẩn hóa (normalize).
- Callbacks: EarlyStopping, ModelCheckpoint, giúp quá trình huấn luyện tối ưu hơn.
- Custom layers: xây dựng các khối mạng riêng (như ResNet block, Inception module).

- Cấu hình các siêu tham số: learning rate, batch size, số epoch...

**Ý nghĩa:** Việc tách riêng các chức năng phụ trợ giúp mô hình dễ mở rộng, bảo trì và tái sử dụng trong các project khác – đây cũng là tinh thần của kiến trúc sạch (Clean Architecture) mà nhóm sẽ áp dụng ở các chương sau.

### 2.5.7. Tổng kết

Việc khảo sát và thực hành các mô hình học sâu thông qua các notebook trên đã giúp nhóm:

- Làm quen với quy trình xây dựng, huấn luyện và đánh giá mô hình học sâu.
- Hiểu rõ đặc điểm, ưu – nhược điểm của từng kiến trúc (MLP, CNN, LSTM).
- Tích lũy kinh nghiệm trong tiền xử lý dữ liệu, lựa chọn kiến trúc, và điều chỉnh tham số.

Từ những kiến thức và kinh nghiệm tích lũy trong quá trình phân tích này, nhóm đã sẵn sàng để áp dụng vào xây dựng hệ thống nhận diện tế bào ung thư xương từ ảnh sinh thiết trong các chương tiếp theo.

## CHƯƠNG 3: THIẾT KẾ HỆ THỐNG

### 3.1. Kiến trúc hệ thống

Hệ thống nhận diện ung thư xương thông qua hình ảnh sẽ được thiết kế dựa trên mô hình **Clean Architecture**, giúp đảm bảo tính mở rộng và dễ bảo trì. Hệ thống bao gồm nhiều lớp, mỗi lớp đảm nhiệm một nhiệm vụ riêng biệt:

- **Domain Layer (Lớp miền):** Đây là lớp chứa logic nghiệp vụ cốt lõi của hệ thống. Các đối tượng trong lớp này không phụ thuộc vào bất kỳ công nghệ hay công cụ nào. Trong bài toán nhận diện ung thư xương, các lớp này sẽ bao gồm các đối tượng như mô hình AI, các thuật toán phân tích và xử lý ảnh.
- **Use Case Layer (Lớp trường hợp sử dụng):** Lớp này bao gồm các chức năng nghiệp vụ cụ thể mà hệ thống cần thực hiện. Các use case trong hệ thống sẽ bao

gồm các thao tác như tải ảnh, tiền xử lý ảnh, huấn luyện mô hình, dự đoán kết quả, v.v.

- **Interface/Adapter Layer (Lớp giao diện/adapter):** Đây là lớp chịu trách nhiệm giao tiếp với người dùng. Ứng dụng sẽ sử dụng **Streamlit** hoặc **Django** để xây dựng giao diện người dùng, giúp người dùng tải ảnh, nhận dự đoán và hiển thị kết quả.
- **External Layer (Lớp bên ngoài):** Lớp này quản lý các hệ thống bên ngoài như cơ sở dữ liệu, mô hình AI, hệ thống file, v.v. Trong bài toán này, lớp này sẽ bao gồm các module liên quan đến việc lưu trữ ảnh, lưu trữ kết quả dự đoán và mô hình học sâu.

### 3.2. Phân chia theo Clean Architecture

#### 3.2.1. Domain Layer

- **Mô hình AI (Model):** Đây là lớp thực hiện nhiệm vụ nhận diện ung thư xương thông qua việc phân tích ảnh sinh thiết. Các mô hình học sâu như **U-Net** hoặc **DeepLabV3+** sẽ được sử dụng trong lớp này để phân vùng các vùng bị ung thư trong ảnh.
- **Thuật toán phân tích (Analysis):** Lớp này sẽ đảm nhiệm việc tiền xử lý ảnh, xử lý mask và các bước khác cần thiết cho việc phân tích ảnh.

#### 3.2.2. Use Case Layer

- **Chức năng huấn luyện mô hình (Training):** Lớp này sẽ tổ chức quá trình huấn luyện mô hình với các tham số như số epoch, batch size, optimizer, hàm loss, và metric.
- **Chức năng dự đoán (Prediction):** Lớp này nhận đầu vào là một ảnh mới và đưa ra kết quả dự đoán, cho biết ảnh đó có bị ung thư hay không, đồng thời cung cấp phân vùng khối u.

#### 3.2.3. Interface/Adapter Layer

- **Giao diện người dùng:** Giao diện này sẽ được xây dựng bằng **Streamlit** hoặc **Django**, cho phép người dùng tải ảnh sinh thiết lên và nhận kết quả dự đoán từ mô hình AI.
- **API và tích hợp:** Nếu hệ thống cần tích hợp với các dịch vụ khác, lớp này sẽ tạo các API cho việc giao tiếp giữa hệ thống và các dịch vụ bên ngoài.

#### 3.2.4. External Layer

- **Lưu trữ ảnh và kết quả:** Lớp này quản lý việc lưu trữ ảnh và kết quả dự đoán trong cơ sở dữ liệu hoặc hệ thống file.
- **Mô hình học sâu:** Mô hình AI đã được huấn luyện sẽ được lưu trữ và sử dụng trong lớp này. Khi người dùng tải ảnh lên, mô hình sẽ thực hiện phân tích và trả về kết quả.

### 3.3. Mô hình AI

#### 3.3.1. Kiến trúc mô hình

Mô hình AI sẽ được xây dựng trên nền tảng của **Deep Learning**, sử dụng các mô hình **U-Net** hoặc **DeepLabV3+** để phân vùng và phát hiện các tế bào ung thư trong ảnh sinh thiết. Cả hai mô hình này đều sử dụng cấu trúc CNN để tự động trích xuất đặc trưng và phân vùng các vùng có khả năng chứa khối u.

- **U-Net:** Đây là một mô hình phổ biến trong phân tích ảnh y tế, với một kiến trúc đối xứng bao gồm phần encoder (giảm kích thước ảnh) và decoder (tái tạo lại ảnh ở kích thước ban đầu). U-Net rất hiệu quả trong các bài toán phân vùng với dữ liệu ảnh y tế.
- **DeepLabV3+:** Đây là mô hình phân vùng ảnh tiên tiến của Google, sử dụng atrous convolution để mở rộng trường nhận diện và giúp phân vùng ảnh với độ chi tiết cao.

#### 3.3.2. Tiền xử lý ảnh

Quá trình tiền xử lý ảnh sẽ bao gồm các bước:

- **Chuẩn hóa ảnh:** Đảm bảo ảnh có kích thước đồng nhất và chuẩn hóa giá trị pixel trong phạm vi  $[0, 1]$ .
- **Tạo mask:** Nếu làm bài toán phân vùng, các vùng bị ung thư trong ảnh sẽ được đánh dấu bằng mask. Đây là các đối tượng cần được mô hình phân vùng chính xác.

### 3.4. Phân vùng khối u trong ảnh sinh thiết

Các kỹ thuật phân vùng ảnh như **Semantic Segmentation** (phân vùng theo pixel) và **Instance Segmentation** (phân vùng theo đối tượng) sẽ được áp dụng để xác định các khu vực khối u trong ảnh sinh thiết. Phân vùng chính xác sẽ giúp các bác sĩ xác định vị trí và kích thước của khối u, từ đó đưa ra phác đồ điều trị chính xác hơn.

### 3.5. Hệ thống phân tầng

Kiến trúc hệ thống được chia thành các tầng để đảm bảo tính modular và dễ bảo trì:

- **Tầng giao diện:** Dùng Streamlit hoặc Django để tạo giao diện người dùng, giúp dễ dàng tải ảnh lên và nhận kết quả dự đoán.
- **Tầng logic nghiệp vụ:** Chứa các chức năng xử lý ảnh, huấn luyện và dự đoán mô hình AI.
- **Tầng lưu trữ:** Lưu trữ ảnh, kết quả dự đoán và các mô hình đã huấn luyện.

## CHƯƠNG 4: DỮ LIỆU

### 4.1. Nguồn Dữ Liệu

Dữ liệu sử dụng trong hệ thống nhận diện ung thư xương thông qua hình ảnh được lấy từ các bộ sưu tập ảnh sinh thiết xương. Bộ dữ liệu này bao gồm các ảnh chụp từ các mẫu sinh thiết, trong đó một số mẫu có khối u và một số mẫu không có khối u. Bộ dữ liệu này có định dạng ảnh thường là **PNG**, **JPEG** hoặc **TIFF**, với kích thước và độ phân giải phù hợp để phân tích bằng các mô hình học sâu.

Ngoài ảnh, bộ dữ liệu còn kèm theo các file **annotation** hoặc **mask** (đối với bài toán phân vùng), chứa thông tin về vùng bị ung thư trong ảnh. Mỗi ảnh sẽ có một file

**JSON** hoặc **XML** chứa thông tin về các vùng khối u, như vị trí, diện tích, và hình dạng của khối u.

Các bộ dữ liệu có thể được lấy từ các nguồn như:

- Các nghiên cứu y tế công cộng hoặc các tổ chức nghiên cứu ung thư.
- Các kho dữ liệu mở như **Kaggle**, **The Cancer Imaging Archive (TCIA)**.
- Các bệnh viện hoặc phòng khám hợp tác trong nghiên cứu y tế.

## 4.2. Tiền Xử Lý Dữ Liệu

Trước khi đưa vào huấn luyện mô hình, dữ liệu ảnh sẽ phải trải qua một số bước tiền xử lý để làm sạch và chuẩn bị sẵn sàng cho mô hình học sâu. Các bước tiền xử lý bao gồm:

### 4.2.1. Chuyển đổi kích thước (Resize)

Vì các ảnh trong bộ dữ liệu có thể có kích thước và tỷ lệ khác nhau, bước đầu tiên trong tiền xử lý là thay đổi kích thước các ảnh về một kích thước cố định. Trong bài toán này, chúng ta sẽ thay đổi kích thước tất cả các ảnh thành **224x224 pixel** hoặc **224x294 pixel** để đảm bảo đồng nhất và phù hợp với mô hình học sâu.

### 4.2.2. Chuẩn hóa ảnh (Normalization)

Để mô hình học sâu có thể học tốt hơn, giá trị pixel của ảnh sẽ được chuẩn hóa về phạm vi từ 0 đến 1. Điều này giúp quá trình huấn luyện diễn ra nhanh hơn và ổn định hơn. Chuẩn hóa được thực hiện bằng cách chia mỗi giá trị pixel cho 255 (giá trị tối đa của một pixel màu RGB).

### 4.2.3. Tăng cường dữ liệu (Data Augmentation)

Để tăng độ đa dạng của dữ liệu và giúp mô hình học tốt hơn, các kỹ thuật tăng cường dữ liệu sẽ được áp dụng. Các kỹ thuật phổ biến bao gồm:

- **Lật ảnh (Flip):** Lật ảnh theo chiều ngang hoặc dọc để tạo ra các biến thể khác của ảnh.

- **Xoay ảnh (Rotation):** Xoay ảnh một góc nhỏ để tạo sự đa dạng.
- **Cắt ảnh (Crop):** Cắt một phần ngẫu nhiên của ảnh để mô hình học được các đặc trưng ở các vùng khác nhau.
- **Thay đổi độ sáng (Brightness):** Điều chỉnh độ sáng và độ tương phản của ảnh.

#### 4.2.4. Tạo mask (cho bài toán phân vùng)

Đối với bài toán phân vùng, chúng ta cần tạo các **mask** cho mỗi ảnh, chỉ ra chính xác khu vực có khối u. Mỗi pixel trong mask có thể có giá trị 0 (không phải khối u) hoặc 1 (đây là khối u). Các mask này thường được lưu dưới dạng ảnh nhị phân (binary image) với cùng kích thước và vị trí tương ứng với ảnh gốc.

Các kỹ thuật tạo mask có thể bao gồm:

- **Annotation thủ công:** Các chuyên gia y tế có thể vẽ khối u lên ảnh và lưu lại dưới dạng mask.
- **Sử dụng thuật toán:** Các mô hình học sâu có thể giúp tự động tạo mask cho những ảnh chưa được gán nhãn.

### 4.3. Chia Dữ Liệu

Để huấn luyện, đánh giá và kiểm tra mô hình, bộ dữ liệu sẽ được chia thành ba phần chính:

- **Training set:** Bộ dữ liệu dùng để huấn luyện mô hình. Đây là phần lớn nhất của bộ dữ liệu và sẽ được sử dụng để cập nhật trọng số của mô hình.
- **Validation set:** Bộ dữ liệu dùng để điều chỉnh các tham số mô hình và tránh overfitting (quá khớp với dữ liệu huấn luyện). Validation set giúp xác định khi nào mô hình đã học đủ và không cần huấn luyện thêm.
- **Test set:** Bộ dữ liệu dùng để đánh giá hiệu suất của mô hình sau khi đã huấn luyện xong. Bộ dữ liệu này không được sử dụng trong quá trình huấn luyện và điều chỉnh tham số.

Tỉ lệ chia bộ dữ liệu phổ biến là 70% cho huấn luyện, 15% cho kiểm tra (validation), và 15% cho thử nghiệm (test).

#### 4.4. Các Thách Thức trong Quá Trình Làm Việc với Dữ Liệu

Trong quá trình xử lý và làm việc với dữ liệu y tế, có một số thách thức mà chúng ta cần lưu ý:

- **Dữ liệu không cân bằng:** Số lượng ảnh chứa khối u và không chứa khối u có thể không đồng đều. Điều này có thể dẫn đến việc mô hình học không chính xác cho các lớp ít dữ liệu. Các kỹ thuật như **undersampling**, **oversampling**, hoặc **loss weighting** có thể được sử dụng để giải quyết vấn đề này.
- **Dữ liệu thiếu:** Trong một số trường hợp, dữ liệu có thể thiếu hoặc bị hỏng. Các phương pháp thay thế như **interpolation** hoặc **data imputation** có thể giúp khắc phục vấn đề này.
- **Đặc trưng phức tạp của ảnh y tế:** Những đặc trưng trong ảnh y tế có thể rất phức tạp, đòi hỏi các mô hình phân tích và học sâu phải được thiết kế đặc biệt để xử lý tốt các chi tiết tinh vi.

#### 4.5. Định Dạng Dữ Liệu

Dữ liệu ảnh và annotation thường sẽ được lưu trữ dưới các định dạng sau:

- **Định dạng ảnh:** PNG, JPEG, TIFF.
- **Annotation:** JSON, XML, hoặc các định dạng đặc biệt cho phân vùng như Mask R-CNN (PNG hoặc TIFF với độ sâu màu).

Thông tin annotation sẽ bao gồm các thông số như:

- **Vị trí và diện tích khối u:** Các tọa độ của bounding box hoặc mask của khối u.
- **Loại khối u:** Nếu có, có thể bao gồm các loại ung thư khác nhau.

## CHƯƠNG 5: HUẤN LUYỆN MÔ HÌNH

### 5.1. Cấu Hình Huấn Luyện



Cấu hình huấn luyện là một yếu tố quan trọng giúp mô hình học tốt hơn và đạt được kết quả tốt nhất. Để huấn luyện mô hình nhận diện ung thư xương từ ảnh sinh thiết, các thông số sau sẽ được thiết lập:

#### 5.1.1. Số Epochs

Số **epoch** xác định số lần mà mô hình sẽ lặp qua toàn bộ bộ dữ liệu huấn luyện. Mỗi lần qua bộ dữ liệu (1 epoch) sẽ giúp mô hình cập nhật trọng số để cải thiện khả năng dự đoán. Việc chọn số epochs cần cân nhắc kỹ lưỡng, vì số epochs quá ít có thể khiến mô hình chưa học đủ, trong khi số epochs quá nhiều có thể dẫn đến **overfitting** (quá khớp với dữ liệu huấn luyện).

- **Ví dụ:** 50 epochs có thể là một lựa chọn hợp lý đối với bài toán này, nhưng cần điều chỉnh tùy vào kích thước bộ dữ liệu và độ phức tạp của mô hình.

#### 5.1.2. Batch Size

**Batch size** là số lượng mẫu dữ liệu mà mô hình sẽ xử lý trong một lần cập nhật trọng số. Việc chọn **batch size** quá nhỏ có thể làm chậm quá trình huấn luyện, trong khi chọn quá lớn có thể gây tốn nhiều bộ nhớ và làm giảm hiệu quả huấn luyện.

- **Ví dụ:** Batch size phổ biến cho bài toán nhận diện ảnh có thể là 16, 32 hoặc 64.

#### 5.1.3. Learning Rate

**Learning rate** là tham số điều khiển bước nhảy của mô hình trong quá trình huấn luyện. Một learning rate quá nhỏ sẽ khiến quá trình huấn luyện chậm, trong khi learning rate quá lớn có thể làm mô hình không hội tụ hoặc hội tụ đến một điểm cực trị không tối ưu.

- **Ví dụ:** Learning rate có thể được thử nghiệm trong khoảng từ **1e-4** đến **1e-2** và điều chỉnh theo từng giai đoạn huấn luyện.

#### 5.1.4. Optimizer

**Optimizer** là thuật toán dùng để cập nhật trọng số của mô hình trong quá trình huấn luyện. Một số optimizers phổ biến bao gồm:

- **SGD (Stochastic Gradient Descent):** Cập nhật trọng số theo tỷ lệ gradient, dễ thực hiện nhưng có thể gặp khó khăn khi xử lý các hàm mất mát phức tạp.
- **Adam (Adaptive Moment Estimation):** Là một trong các optimizers phổ biến hiện nay, Adam tự động điều chỉnh learning rate cho mỗi tham số và làm việc rất hiệu quả trong việc huấn luyện các mô hình phức tạp như mạng nơ-ron tích chập (CNN).
- **RMSprop:** Cải tiến từ SGD với việc sử dụng bình quân động, giúp xử lý tốt trong các trường hợp có sự thay đổi lớn về gradient.
- **Ví dụ:** Adam là lựa chọn phổ biến cho huấn luyện các mô hình học sâu như CNN hoặc U-Net.

#### 5.1.5. Loss Function

**Loss function** đo lường sự khác biệt giữa giá trị dự đoán của mô hình và giá trị thực tế. Một số hàm loss phổ biến cho bài toán nhận diện ung thư xương bao gồm:

- **Cross Entropy Loss:** Phù hợp cho các bài toán phân loại, đo lường độ sai lệch giữa phân phối xác suất của mô hình và phân phối xác suất thực tế.
- **Dice Loss:** Là một hàm loss được sử dụng phổ biến trong phân vùng ảnh y tế, đặc biệt cho các bài toán phân vùng khối u, giúp tính toán độ tương đồng giữa mask dự đoán và mask thực tế.
- **Mean Squared Error (MSE):** Thường dùng cho bài toán hồi quy nhưng cũng có thể được áp dụng trong phân vùng ảnh.
- **Ví dụ:** Đối với bài toán phân vùng khối u, **Dice Loss** có thể là lựa chọn tốt vì nó tối ưu hóa được sự giống nhau giữa vùng khối u thực tế và dự đoán.

#### 5.1.6. Metric Đánh Giá

Các **metric** đánh giá giúp theo dõi sự tiến bộ của mô hình trong suốt quá trình huấn luyện và đánh giá hiệu quả của mô hình sau khi huấn luyện. Một số metric phổ biến bao gồm:

- **Accuracy:** Tỷ lệ chính xác của các dự đoán so với tổng số mẫu trong tập dữ liệu.
- **IoU (Intersection over Union):** Đo lường độ chính xác của phân vùng khối u, tính tỷ lệ diện tích giao nhau giữa mask dự đoán và mask thực tế so với diện tích hợp nhất của chúng.
- **Dice Score:** Đo lường sự tương đồng giữa hai tập hợp, dùng trong phân vùng ảnh để đánh giá độ chính xác của mô hình trong việc phân chia khối u.
- **Confusion Matrix:** Bảng tóm tắt số lượng dự đoán đúng và sai của các lớp.
- **Ví dụ: IoU và Dice Score** là các metric rất quan trọng đối với bài toán phân vùng, vì chúng đo lường chất lượng của mask dự đoán.

## 5.2. Quy Trình Huấn Luyện

Quá trình huấn luyện mô hình nhận diện ung thư xương từ ảnh sinh thiết có thể được chia thành các bước chính sau:

### 5.2.1. Chuẩn Bị Dữ Liệu

Trước khi huấn luyện, dữ liệu cần được chia thành ba bộ: **Training set**, **Validation set**, và **Test set**. Dữ liệu sẽ được tiền xử lý bao gồm các bước như resize ảnh, chuẩn hóa giá trị pixel và tăng cường dữ liệu (data augmentation).

### 5.2.2. Huấn Luyện Mô Hình

Mô hình sẽ được huấn luyện với bộ dữ liệu huấn luyện (training set) thông qua các epochs. Trong mỗi epoch, mô hình sẽ học từ dữ liệu và cập nhật trọng số bằng cách sử dụng hàm loss và optimizer đã chọn. Các metric sẽ được tính toán để theo dõi sự tiến bộ của mô hình trong suốt quá trình huấn luyện.

### 5.2.3. Điều Chỉnh Tham Số

Sau mỗi vài epoch, mô hình sẽ được đánh giá trên bộ validation set. Nếu mô hình có dấu hiệu overfitting (quá khớp với dữ liệu huấn luyện) hoặc underfitting (học không đủ), các tham số như learning rate, batch size, hoặc số epoch có thể được điều chỉnh.

#### 5.2.4. Đánh Giá Mô Hình

Sau khi huấn luyện xong, mô hình sẽ được kiểm tra với bộ dữ liệu test để đánh giá hiệu quả của mô hình. Các metric như accuracy, IoU, Dice score sẽ giúp đánh giá chất lượng của mô hình trong việc nhận diện và phân vùng khối u.

#### 5.3. Biểu Đồ Huấn Luyện

Trong quá trình huấn luyện, các biểu đồ loss, accuracy và các metric khác sẽ được vẽ để theo dõi sự tiến triển của mô hình. Các biểu đồ này giúp nhận diện các vấn đề như overfitting hoặc học quá nhanh.

- **Biểu đồ loss** giúp theo dõi sự giảm dần của hàm mất mát trong quá trình huấn luyện.
- **Biểu đồ accuracy** giúp đánh giá độ chính xác của mô hình trong quá trình huấn luyện và validation.

#### 5.4. Thử Nghiệm Nhiều Mô Hình

Khi huấn luyện mô hình, có thể thử nghiệm với nhiều kiến trúc mô hình khác nhau như **U-Net**, **ResNet**, hoặc **VGGNet**. Việc so sánh hiệu suất của các mô hình khác nhau sẽ giúp chọn ra mô hình tối ưu cho bài toán nhận diện ung thư xương.

## CHƯƠNG 6: KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

### 6.1. Tổng Kết

Ứng dụng nhận diện ung thư xương thông qua ảnh sinh thiết đã thể hiện tiềm năng to lớn trong việc hỗ trợ quá trình chẩn đoán và điều trị bệnh. Các mô hình học sâu như **U-Net**, **DeepLabV3+** cùng với các kỹ thuật phân vùng ảnh đã cho thấy khả

năng phân tích và phát hiện các khối u trong ảnh sinh thiết một cách chính xác và nhanh chóng. Việc áp dụng Deep Learning trong y tế, đặc biệt trong việc nhận diện các tổn thương ung thư, không chỉ giúp bác sĩ tăng cường độ chính xác trong chẩn đoán mà còn giảm thiểu thời gian và công sức trong quá trình phân tích.

### **Những điểm nổi bật của dự án:**

- **Mô hình học sâu (Deep Learning)** đã được ứng dụng hiệu quả trong nhận diện và phân vùng các khối u trong ảnh sinh thiết, với các mô hình như U-Net, DeepLabV3+.
- **Xử lý ảnh y tế:** Các kỹ thuật tiền xử lý và phân vùng ảnh đã được áp dụng đúng cách, mang lại kết quả đáng tin cậy cho bài toán phân loại và phát hiện các khối u.
- **Clean Architecture** đã giúp xây dựng hệ thống linh hoạt và dễ duy trì, từ đó có thể mở rộng và tích hợp với các công nghệ khác trong tương lai.

## **6.2. Hướng Phát Triển**

Mặc dù mô hình hiện tại đã cho kết quả khả quan, vẫn còn nhiều hướng phát triển để nâng cao hiệu quả và mở rộng ứng dụng của hệ thống trong chẩn đoán ung thư xương. Dưới đây là một số hướng phát triển tiềm năng:

### **6.2.1. Cải Tiến Mô Hình**

- **Kết hợp nhiều mô hình:** Một trong những hướng phát triển là kết hợp các mô hình phân vùng và nhận diện khác nhau, chẳng hạn như U-Net với các mô hình CNN tiên tiến hơn như ResNet, để cải thiện độ chính xác.
- **Mô hình học sâu mới:** Các kiến trúc mới như Transformer hoặc Hybrid Models có thể được thử nghiệm để tìm ra mô hình tối ưu hơn cho bài toán này.
- **Tối ưu hóa mô hình cho môi trường sản xuất:** Việc tối ưu hóa mô hình để giảm thiểu tài nguyên tính toán mà vẫn duy trì độ chính xác cao sẽ là một yếu tố quan trọng khi triển khai trong môi trường thực tế.

### **6.2.2. Mở Rộng Dữ Liệu**

- **Tăng cường dữ liệu:** Mặc dù đã có một bộ dữ liệu lớn, nhưng việc mở rộng bộ dữ liệu bằng cách thu thập thêm ảnh sinh thiết từ các nguồn khác hoặc sử dụng phương pháp tăng cường dữ liệu (data augmentation) sẽ giúp cải thiện độ chính xác của mô hình.
- **Dữ liệu đa dạng hơn:** Sử dụng dữ liệu từ nhiều bệnh viện và vùng miền khác nhau để mô hình có thể làm việc chính xác trên các dữ liệu đa dạng.

### 6.2.3. Tích Hợp Công Nghệ AI Vào Quy Trình Chẩn Đoán

- **Tích hợp với bác sĩ:** Mô hình có thể được tích hợp trực tiếp vào quy trình làm việc của bác sĩ, giúp họ đưa ra quyết định nhanh chóng và chính xác hơn. Một hệ thống hỗ trợ ra quyết định (decision support system) sẽ giúp bác sĩ trong việc xác định và phân loại các khối u.
- **Phát triển giao diện người dùng (UI):** Xây dựng các giao diện người dùng đơn giản, dễ sử dụng để bác sĩ và nhân viên y tế có thể tương tác với hệ thống một cách thuận tiện hơn. Điều này có thể bao gồm các giao diện web hoặc ứng dụng di động, giúp bác sĩ có thể truy cập và sử dụng hệ thống mọi lúc mọi nơi.

### 6.2.4. Mở Rộng Ứng Dụng

- **Mở rộng sang các loại ung thư khác:** Mô hình hiện tại có thể được phát triển thêm để nhận diện các loại ung thư khác, chẳng hạn như ung thư vú, ung thư phổi, hay ung thư gan, thông qua việc sử dụng các dữ liệu ảnh sinh thiết từ các bộ phận khác của cơ thể.
- **Ứng dụng trong y học cá nhân hóa:** Hệ thống có thể được mở rộng để hỗ trợ chẩn đoán ung thư cho từng bệnh nhân cụ thể, kết hợp với các dữ liệu di truyền, lịch sử bệnh tật và các yếu tố khác để đưa ra các khuyến nghị điều trị chính xác.

### 6.2.5. Tối Ưu Hóa và Triển Khai Thực Tế

- **Triển khai trên nền tảng đám mây:** Việc triển khai mô hình trên các nền tảng đám mây sẽ giúp mở rộng khả năng tiếp cận, đồng thời giảm thiểu yêu cầu về tài nguyên máy tính tại các cơ sở y tế.

- **Đảm bảo bảo mật dữ liệu:** Khi làm việc với dữ liệu y tế, việc bảo mật và tuân thủ các quy định về bảo vệ thông tin cá nhân (HIPAA, GDPR) là rất quan trọng. Các biện pháp bảo mật và mã hóa dữ liệu sẽ cần được triển khai để đảm bảo rằng hệ thống đáp ứng các yêu cầu về quyền riêng tư và bảo mật.

### 6.3. Kết Luận Chung

Ứng dụng học sâu trong nhận diện ung thư xương thông qua ảnh sinh thiết là một bước tiến quan trọng trong việc cải thiện chất lượng chẩn đoán và điều trị bệnh. Hệ thống đã được xây dựng dựa trên kiến trúc Clean Architecture, giúp việc phát triển và bảo trì hệ thống trở nên linh hoạt và dễ dàng. Mặc dù mô hình đã cho kết quả tốt, nhưng việc tối ưu hóa và mở rộng mô hình cũng như cải tiến quy trình chẩn đoán sẽ giúp mang lại những cải tiến lớn trong việc phát hiện sớm ung thư xương, từ đó hỗ trợ các bác sĩ trong việc đưa ra quyết định điều trị chính xác và nhanh chóng hơn.

Các hướng phát triển tiếp theo sẽ tập trung vào việc cải thiện mô hình, mở rộng bộ dữ liệu, tích hợp công nghệ vào quy trình làm việc của bác sĩ, và đảm bảo bảo mật và hiệu quả khi triển khai thực tế.

/