TRƯỜNG ĐẠI HỌC NÔNG LÂM TP. HỒ CHÍ MINH

**KHOA CÔNG NGHỆ THÔNG TIN  
  
  
BÁO CÁO**

**ĐỀ TÀI**: Ứng dụng ResNet-50 trong phân loại hình ảnh tế bào ung thư cổ tử cung

Môn học: Data Mining

Học kỳ: 1/2024-2025

Giảng viên phụ trách: Trần Quốc Việt

Danh sách sinh viên thực hiện:

1. Trần Thị Thiện Nhân 21130461
2. Nguyễn Thanh Bình 21130291
3. Trịnh Trần Sỹ Đông 21130319
4. Lâm Hồng Phong 21130473

MỤC LỤC

[ĐẶT VẤN ĐỀ 3](#_Toc189686798)

[CHƯƠNG 1 NỘI DUNG NGHIÊN CỨU 4](#_Toc189686799)

[1.1 Tổng quan về tập dữ liệu [2] 4](#_Toc189686800)

[1.2 Mô hình CNN 5](#_Toc189686801)

[1.2.1 Định nghĩa CNN 5](#_Toc189686802)

[1.2.2 Cấu trúc mạng CNN 5](#_Toc189686803)

[1.3 Mô hình ResNet50 [3] 6](#_Toc189686804)

[1.4 Pre-trained model 7](#_Toc189686805)

[1.5 Fine-tuning 8](#_Toc189686806)

[1.5.1 Tổng quan 8](#_Toc189686807)

[1.5.2 Lợi ích của fine-tuning 8](#_Toc189686808)

[1.5.3 Những khó khăn khi sử dụng fine-tuning 8](#_Toc189686809)

[1.6 Tiền xử lý dữ liệu 9](#_Toc189686810)

[1.7 Đánh giá mô hình 10](#_Toc189686811)

[CHƯƠNG 2 PHÂN TÍCH KẾT QUẢ 12](#_Toc189686812)

[2.1 Kết quả 12](#_Toc189686813)

[2.2 Phân tích 14](#_Toc189686814)

[CHƯƠNG 3 KẾT LUẬN 16](#_Toc189686815)

[3.1 Ý nghĩa của bài nghiên cứu 16](#_Toc189686816)

[3.2 Hạn chế 16](#_Toc189686817)

[TÀI LIỆU THAM KHẢO 17](#_Toc189686818)

[PHỤ LỤC 18](#_Toc189686819)

[PHÂN CÔNG CÔNG VIỆC 19](#_Toc189686820)

**DANH SÁCH HÌNH ẢNH**

[Hình 1.1: Cấu trúc cơ bản của CNN 5](#_Toc189685575)

[Hình 1.2: Cấu trúc mô hình Resnet50 6](#_Toc189685576)

[Hình 1.3: Confusion Matrix 9](#_Toc189685577)

[Hình 2.1: Biểu đồ theo dõi hiệu suất mô hình ResNet50 11](#_Toc189685578)

[Hình 2.2: Confusion Matrix của mô hình ResNet50 12](#_Toc189685579)

[Hình 2.3: Biểu đồ ROC và chỉ số AUC từng lớp của mô hình ResNet50 13](#_Toc189685580)

# ĐẶT VẤN ĐỀ

Ung thư cổ tử cung là loại ung thư phổ biến thứ tư ở phụ nữ trên thế giới. Theo thống kê của BGI [1], trong năm 2023 ước tính có 604,000 ca mới và 342,000 ca tử vong trên toàn thế giới. Gần 90% số ca tử vong vào năm 2022 xảy ra ở các quốc gia thu nhập thấp và trung bình, những quốc gia tụt hậu về tiêm chủng, sàng lọc và chữa trị. Ung thư cổ tử cung là một bệnh nguy hiểm nhưng nó được chữa trị hoàn toàn nếu phát hiện sớm và điều trị kịp thời với tỷ lệ lên đến 100%.

Xét nghiệm Pap smear hiện là công cụ sàng lọc đáng tin cậy để phát hiện các tế bào ung thư và tổn thương cổ tử cung. Cùng với sự phát triển của mạng nơ-ron tích chập, việc xây dựng mô hình để phân loại các loại tế bào với mục đích cuối cùng là phát hiện ung thư cổ tử cung có thể giúp cải thiện độ chính xác trong chuẩn đoán, giảm thiểu thời gian và chi phí cho người bệnh.

Vì vậy, bài nghiên cứu này tập trung nghiên cứu và áp dụng mô hình ResNet50 vào bài toán chuẩn đoán bệnh ung thư cổ tử cung.

# NỘI DUNG NGHIÊN CỨU

## Tổng quan về tập dữ liệu [2]

Tập dữ liệu SIPaKMeD bao gồm 4049 hình ảnh của các tế bào đơn lẻ được cắt thủ công từ 966 hình ảnh của các tiêu bản Pap smear. Những hình ảnh này được thu nhận bằng máy ảnh CCD (Infinity 1 Lumenera) được gắn vào kính hiển vi quang học (OLYMPUS BX53F). Có 5 loại tế bào sau:

* **Parabasal**: là các tế bào vảy chưa trưởng thành và là các tế bào biểu mô nhỏ nhất được quan sát thấy trong một tiêu bản phết tế bào âm đạo điển hình. Bào tương của chúng thường có ái tính với màu cyan (xanh lam) và thường chứa một nhân dạng túi lớn. Các tế bào parabasal có đặc điểm hình thái tương tự với các tế bào metaplastic và rất khó phân biệt giữa chúng.
* **Superficial-Intermediate**: Chiếm phần lớn trong các tế bào được tìm thấy trong xét nghiệm Pap smear. Thông thường, chúng có dạng phẳng với bào tương hình tròn, hình bầu dục hoặc hình đa giác. Bào tương chủ yếu nhuộm màu ái eosin (hồng) hoặc ái cyan (xanh lam). Chúng có một nhân pycnotic. Bào tương có hình đa giác lớn, rõ ràng và dễ nhận biết với giới hạn nhân (nhân pycnotic nhỏ ở tế bào bề mặt và nhân dạng túi ở tế bào trung gian). Loại tế bào này có thể biểu hiện các thay đổi hình thái đặc trưng (koilocytic atypia) do các tổn thương nghiêm trọng hơn.
* **Dyskeratotic**: là các tế bào vảy đã trải qua quá trình sừng hóa bất thường sớm trong từng tế bào đơn lẻ hoặc thường thấy hơn là trong các cụm ba chiều. Chúng có bào tương màu ái orangeophilic (cam sáng). Đặc điểm nhận dạng là sự hiện diện của các nhân dạng túi, tương tự như nhân của các tế bào koilocytotic. Chúng là một đặc điểm nổi bật của nhiễm HPV và đôi khi ngay cả khi không có tế bào koilocyte, chúng cũng có thể là bằng chứng đặc trưng của nhiễm HPV. Chúng thường xuất hiện trong các cụm dày, ba chiều và khó phân biệt được nhân hoặc giới hạn của bào tương.
* **Koilocytic**: Tế bào koilocytotic thường tương ứng với các tế bào vảy trưởng thành (trung gian và bề mặt) và đôi khi là các tế bào dị sản kiểu koilocytotic. Chúng thường có ái tính với màu cyan (xanh lam), được nhuộm rất nhạt và đặc trưng bởi một không bào lớn xung quanh nhân. Phần viền bào tương có màu nhuộm đậm đặc. Nhân của các tế bào koilocyte thường to hơn, nằm lệch tâm, đậm màu (hyperchromatic) và có đường viền màng nhân không đều. Trong nhiều trường hợp, có sự xuất hiện của các tế bào có hai nhân (binucleated) và/hoặc đa nhân (multinucleated). Tế bào koilocyte là các tế bào đặc trưng cho nhiễm HPV và nhân của các tế bào này thường có các mức độ thoái hóa khác nhau, tùy thuộc vào giai đoạn nhiễm trùng và loại virus gây nhiễm khác nhau.
* **Metaplastic**: Tế bào dị sản về bản chất là các tế bào loại parabasal nhỏ hoặc lớn với đường viền tế bào nổi bật, thường có nhân lệch tâm và đôi khi chứa một không bào lớn bên trong. Phần nhuộm màu ở trung tâm thường có màu nâu nhạt và thường khác với phần viền ngoài. Bào tương của chúng có màu nhuộm đậm hơn và chúng thể hiện sự đồng đều lớn về kích thước và hình dạng so với các tế bào parabasal, đặc trưng bởi hình dạng bào tương gần như tròn và rõ ràng. Sự hiện diện của các tế bào này trong xét nghiệm Pap có liên quan đến tỷ lệ phát hiện cao hơn của các tổn thương tiền ung thư (HSIL).

## Mô hình CNN

### Định nghĩa CNN

Convolutional Neural Network (CNNs – Mạng nơ-ron tích chập) là một mạng nơ-ron nhân tạo thường được dùng để nhận dạng và phân loại hình ảnh.

### Cấu trúc mạng CNN



Hình 1.1: Cấu trúc cơ bản của CNN

**Lớp đầu vào – Input Layer:**

* Lớp chứa hình ảnh đầu vào với kích thước (W, H, C). Trong đó W là chiều rộng, H là chiều cao, C là số kênh màu (thường là 3 với RGB hoặc ảnh màu xám).
* Một số đầu vào phổ biến là 224x224x3, 229x229x3, 32x32x3, 28x28x1, tùy theo đặc trưng của bài toán mà sử dụng kích thước đầu vào thích hợp.

**Lớp tích chập - Convolution Layer:**

* Đây là lớp đầu tiên được dùng để trích xuất đặc trưng (feature) của hình ảnh đầu vào.
* Kernel (filter) là một ma trận nhỏ, sẽ trượt trên hình ảnh để tính toán các giá trị.
* Stride (bước nhảy) là số pixel mà kernel di chuyển sau mỗi bước.
* Padding được dùng để xử lý các cạnh của ảnh, trong trường hợp số pixel trên ảnh không chia hết cho kernel thì padding được thêm vào để xử lý vấn đề này.

**Lớp kích hoạt - Activation Layer:**

* Là lớp được đặt sau convolution layer để tăng tính phi tuyến tính, fully connected layer để tối ưu hóa đầu ra.
* Lớp sử dụng các hàm kích hoạt như ReLu, Sigmoid, Tanh, … để tăng tính phi tuyến tính để mạng có thể học được các đặc trưng phức tạp, đồng thời giúp chuẩn hóa đầu ra sao cho phù hợp với bài toán.

**Lớp Pooling - Pooling Layer:**

* Được sử dụng để giảm kích thước đầu ra từ các lớp convolution, giúp giảm số lượng tham số đồng thời giữ lại các đặc trưng quan trọng. Lớp sử dụng một cửa sổ (window) có kích thước k x k, để di chuyển trên feature sau đó thực hiện tính toán lại giúp giảm kích thước của feature, giữ lại các đặc trưng quan trọng và tăng khả năng tổng quát hóa mô hình. Với:
  + **Max Pooling**: lấy giá trị lớn nhất trong từng vùng window trong feature map.
  + **Average Pooling**: Lấy giá trị trung bình trong từng cùng window trong feature map.
  + **L2 Pooling**: lấy giá trị tổng bình phương các giá trị trong window, sau đó căn bậc hai để lấy kết quả.

**Lớp normalization - Batch Normalization hoặc Layer Normalization:**

* **Batch Normalization**: Chuẩn hóa dữ liệu trong từng batch giảm vấn đề vanishing/exploding gradients và giúp huấn luyện ổn định hơn bằng cách tính trung bình độ lệch chuẩn của mỗi vector trong mỗi lớp ẩn.
* **Layer Normalization**: Chuẩn hóa dữ liệu theo từng nơ-ron, giảm phụ thuộc vào kích thước batch, thường được dùng trong các mô hình RNN.

**Lớp Dropout - Dropout Layer:**

* Lớp này có chức năng giảm overfitting trong quá trình huấn luyện mô hình bằng cách tắt ngẫu nhiên một vài nơ-ron với một xác suất cho trước (dropout rate). Điều này bắt buộc mô hình phải học đặc trưng từ các tổ hợp nơ-ron khác nhau, tăng khả năng tổng quát hóa mô hình.

**Lớp Fully Connected - Dense Layer:**

* Đây là lớp được kết nối với tất cả các nơ-ron của lớp trước, tiến hành tổng hợp các đặc trưng kết hợp với hàm softmax để phân loại đa lớp hoặc có thể sử dụng sigmoid cho phân loại nhị phân và thường được đặt ở cuối mô hình.

**Lớp Đầu Ra - Output Layer:**

* Đối với phân loại nhị phân: Đầu ra bao gồm 1 nơ-ron với hàm kích hoạt sigmoid. Giá trị đầu ra nằm trong khoảng [0;1].
* Đối với phân loại đa lớp: Đầu ra có n nơ-ron tương ứng với n lớp cần phân loại với softmax để đưa ra xác suất cho mỗi lớp.
* Đối với hồi quy: Đầu ra là một nơ-ron.

## Mô hình ResNet50 [3]

A diagram of a diagram

Description automatically generated

Hình 1.2: Cấu trúc mô hình Resnet50

Resnet50 là kiến trúc CNN thuộc họ ResNet, ResNet có nhiều kiến trúc với các độ sâu khác nhau như ResNet18, ResNet32, trong đó ResNet50 là một biến thể cỡ trung bình.

Vấn đề chính mà ResNet giải quyết là suy giảm hiệu suất trong mạng nơ-ron. Khi các mạng trở nên sâu hơn, độ chính xác đạt đến mức bão hòa và sau đó giảm xuống nhanh chóng. Sự suy giảm này do khó khăn trong quá trình tối ưu hóa quá trình huấn luyện.

ResNet50 được xây dựng với 50 lớp bao gồm 48 lớp tích chập và 1 lớp Max Pooling với 1 lớp Average Pooling trước khi đi qua lớp Fully Connected.

Resnet50 có lớp đầu vào là hình ảnh có kích thước 224x224x3, 7 lớp tiếp theo sẽ tiền xử lý dữ liệu để giảm kích thước dữ liệu đồng thời giữ được các thông tin quan trọng hỗ trợ cho các block tiếp theo của mô hình. Các lớp tiếp theo được chia thành các block, mỗi block bao gồm các lớp Convolution để học các đặc trưng từ ảnh, Batch Normalization giúp chuẩn hóa các đặc trưng được học từ lớp Convolution, Activation (ReLU) dùng hàm kích hoạt ReLU để tăng tính phi tuyến tính trong mô hình, cuối cùng sử dụng Residual Connections để giữ lại các đặc trưng của các lớp trước đó, mạng sẽ tập trung học các đặc trưng mới mà không làm mất các đặc trưng trước đó.

## Pre-trained model

Pre-trained model: Là một mô hình học máy hoặc học sâu đã được huấn luyện trước trên một tập dữ liệu lớn và có cấu trúc. Thay vì bắt đầu từ đầu với mô hình chưa được huấn luyện (tức là với các trọng số được khởi tạo ngẫu nhiên) thì có thể sử dụng trực tiếp một mô hình đã được huấn luyện trước hoặc tinh chỉnh (fine-tuning) nếu cần thiết để phù hợp với nhu cầu hoặc bài toán cụ thể.

Một số đặc điểm của pre-trained model:

* Trọng số (weight) đã được tối ưu: Các trọng số trong mô hình đã được tối ưu thông qua quá trình huấn luyện với một lượng dữ liệu lớn
* Sử dụng lại kiến thức: Mô hình đã học được các đặc trưng phổ quát từ dữ liệu ban đầu, chẳng hạn như các đặc trưng cơ bản trong nhận diện hình ảnh hoặc hiểu ngữ nghĩa từ ngữ.
* Tiết kiệm tài nguyên: Giảm thời gian và chi phí tính toán cần thiết để huấn luyện mô hình từ đầu

Ứng dụng của pre-trained model:

* Xử lý ngôn ngữ tự nhiên (NLP): các mô hình GPT, BERT, T5 được huấn luyện trên dữ liệu văn bản lớn và có thể được sử dụng trong các tác vụ như phiên dịch, phân loại văn bản hoặc tạo văn bản.
* Thị giác máy tính (Computer Vision): các mô hình ResNet, VGG, EfficientNet đã được huấn luyện trên ImageNet và thường được dùng để nhận diện hình ảnh hoặc phân đoạn ảnh.
* Xử lý âm thanh: các mô hình Wav2Vec, DeepSpeech để nhận diện giọng nói hoặc chuyển đổi âm thanh thành văn bản.

Các cách tận dụng pre-trained model:

* Feature Extraction (Rút trích đặc trưng): Sử dụng các đặc trưng đã học từ mô hình mà không huấn luyện lại
* Fine-tuning (Tinh chỉnh): Huấn luyện lại toàn bộ hoặc một phần mô hình với dữ liệu riêng của bạn để phù hợp với nhiệm vụ cụ thể

## Fine-tuning

### Tổng quan

Fine-tuning trong học máy là quá trình điều chỉnh một mô hình đã được huấn luyện sẵn (pre-trained) để phù hợp với các tác vụ hoặc trường hợp sử dụng cụ thể. Đây là một kỹ thuật cơ bản trong học sâu, đặc biệt phổ biến khi làm việc với các mô hình nền tảng (foundation models) trong AI sinh tạo.

Fine-tuning được xem là một phần của transfer learning: tận dụng kiến thức mà mô hình đã học trước đó làm nền tảng để học các tác vụ mới.

Thay vì huấn luyện từ đầu, fine-tuning giúp điều chỉnh một mô hình lớn đã được học các đặc trưng chung, giảm chi phí tính toán và dữ liệu nhãn cần thiết. Phương pháp này đặc biệt hữu ích với các mô hình lớn (như LLMs trong xử lý ngôn ngữ tự nhiên hoặc CNNs/Vision Transformers trong xử lý hình ảnh).

Fine-tuning có thể được sử dụng để:

* Thay đổi giọng điệu giao tiếp của mô hình ngôn ngữ lớn (LLM).
* Tích hợp dữ liệu độc quyền hoặc kiến thức chuyên ngành vào mô hình đã có.
* Điều chỉnh phong cách minh họa hoặc cải thiện hiệu năng trên các tác vụ cụ thể.

Nhờ đó, fine-tuning đóng vai trò quan trọng trong việc áp dụng mô hình học máy vào thực tế, giúp cá nhân hóa và mở rộng khả năng tiếp cận các mô hình hiện đại.

### Lợi ích của fine-tuning

**Học chuyển giao (Transfer Learning):** Tận dụng mô hình đã huấn luyện trên tập dữ liệu lớn để áp dụng cho nhiệm vụ tương tự, giúp cải thiện hiệu suất đặc biệt khi dữ liệu hạn chế.

**Huấn luyện nhanh hơn:** Mô hình đã học được các đặc trưng hữu ích, nên quá trình fine-tuning trên tập dữ liệu nhỏ hội tụ nhanh hơn, tiết kiệm tài nguyên và thời gian.

**Thích nghi theo miền dữ liệu (Domain Adaptation):** Hiệu quả trong việc điều chỉnh mô hình từ một miền dữ liệu sang miền khác, như từ hình ảnh của một loại camera sang loại camera khác.

**Giảm yêu cầu về dữ liệu:** Fine-tuning hoạt động tốt với tập dữ liệu nhỏ, phù hợp khi việc thu thập dữ liệu gắn nhãn quy mô lớn gặp khó khăn.

### Những khó khăn khi sử dụng fine-tuning

**Quá khớp (Overfitting):**

* Mô tả: Khi fine-tuning trên tập dữ liệu nhỏ, mô hình có thể trở nên quá phù hợp với dữ liệu huấn luyện, dẫn đến hiệu suất kém trên dữ liệu mới.
* Hậu quả: Mô hình thiếu khả năng tổng quát hóa, không thể áp dụng tốt cho các trường hợp thực tế.

**Không tương thích nhiệm vụ (Task Mismatch):**

* Mô tả: Nếu nhiệm vụ của tập dữ liệu tiền huấn luyện và tác vụ cần fine-tuning khác biệt quá lớn, các đặc trưng đã học có thể không phù hợp hoặc không thể chuyển giao.
* Hậu quả: Hiệu quả của fine-tuning bị giảm sút, thậm chí mô hình có thể hoạt động tệ hơn.

**Quên lãng nghiêm trọng (Catastrophic Forgetting):**

* Mô tả: Fine-tuning có thể làm mô hình "quên" đi các đặc trưng quan trọng từ tác vụ tiền huấn luyện, dẫn đến mất khả năng tổng quát hóa.
* Hậu quả: Mô hình trở nên kém hiệu quả khi áp dụng cho các nhiệm vụ liên quan.

**Phụ thuộc vào chất lượng mô hình tiền huấn luyện:**

* Mô tả: Thành công của fine-tuning phụ thuộc rất nhiều vào chất lượng và sự liên quan của mô hình tiền huấn luyện. Nếu mô hình không phù hợp, fine-tuning có thể không mang lại kết quả tốt.
* Hậu quả: Lãng phí tài nguyên mà không đạt được cải thiện đáng kể.

**Hạn chế trong việc giải thích (Limited Interpretability):**

* Mô tả: Các mô hình fine-tuned thường phức tạp hơn và khó giải thích hơn so với mô hình được huấn luyện từ đầu.
* Hậu quả: Khó hiểu và đánh giá cách mô hình đưa ra quyết định, gây trở ngại trong các ứng dụng yêu cầu tính minh bạch.

## Tiền xử lý dữ liệu

**ImageDataGenerator**: Là kĩ thuật dùng để tăng cường hình ảnh bằng cách sử dụng các phép biến đổi khác nhau để tạo ra các biến thể của ảnh gốc, từ đó có thể tăng thêm số lượng mẫu cho tập dữ liệu.

**Resize**: Là kĩ thuật dùng để đưa hình ảnh về một kích thước nhất định cho phù hợp với yêu cầu bài toán.

**Normalization**: Chuẩn hóa dữ liệu về trong một khoảng nhất định. Cụ thể trong đề tài sẽ chuẩn hóa điểm ảnh về khoảng [0,1] giúp tăng tốc độ học và cải thiện mô hình.

**Switch** **RGB**: Đây là quá trình chuyển đổi màu của ảnh từ không gian màu BGR sang RGB hoặc ngược lại. Việc chuyển đổi này giúp hình ảnh được đưa về hệ màu phù hợp với yêu cầu của các mô hình học sâu.

Quá trình tiền xử lý dữ liệu thực hiện qua nhiều giai đoạn với sự hỗ trợ của các thư viện OpenCV, keras và scikit-learn với các trình tự sau:

* Tiến hành chia dữ liệu thành 3 tập: train, validation và test.
* Sử dụng thư viện OpenCV để đọc ảnh từ Google Drive. Trong quá trình đọc ảnh sẽ tiến hành resize lại tất cả các ảnh về kích thước 224x224x3 đối với mô hình ResNet50. Tiến hành chuyển kênh màu từ RGB về BGR để phù hợp với thư viện OpenCV. Ảnh tiếp tục được chuẩn hóa pixel để đưa về khoảng [0,1] giúp cho mô hình dễ huấn luyện hơn.
* Sau khi hình ảnh được chuẩn hóa sẽ được tiến hành tăng cường bằng thư viện ImageDataGenerator được cung cấp bởi thư viện Keras. Thực hiện các kỹ thuật xoay (rotation), dịch chuyển ngang hoặc dọc (width\_shift\_range, height\_shift\_range) 10 %, biến dạng ảnh (shear\_range) tối đa 10%, phóng to hoặc thu nhỏ (zoom\_range) 10%, lật ảnh theo chiều ngang (horizontal\_flip). Sau đó dùng fill\_mode để lấp vùng ảnh bị mất sau khi biến đổi. Chỉ tăng cường cho tập train, không sử dụng với tập validation và test. Dùng preprocessing\_function để chuẩn hóa dữ liệu đầu vào sao cho phù hợp với mô hình được chọn để fine-tune.
* Sử dụng Label Encoder để chuyển nhãn từ chuỗi sang số, sau đó chuyển thành dạng one-hot cho tập train, validation và test.

## Đánh giá mô hình

Mô hình được đánh giá dựa trên các chỉ số được tính từ confusion matrix (ma trận nhầm lẫn).

Confusion matrix là một bảng được dùng để đánh giá hiệu suất của mô hình phân loại nhị phân hoặc đa lớp.



Hình 1.3: Confusion Matrix

Với:

* **TP**: số lượng mẫu dương được dự đoán là dương.
* **TN**: số lượng mẫu âm được dự đoán là âm.
* **FP**: số lượng mẫu âm được dự đoán là dương.
* **FN**: số lượng mẫu dương được dự đoán là âm.

Mô hình được đánh giá dựa trên:

* **Accuracy**: đo tỷ lệ dự đoán đúng trên toàn bộ mẫu dữ liệu.
* **Precision**: đo tỷ lệ dự đoán dương đúng trên tổng số dự đoán dương.
* **Recall**: đo tỷ lệ dự đoán dương đúng trên tổng số thực tế dương.
* **F1-Score**: còn gọi là thước đo F, cho biết sự cân bằng giữa precision và recall.
* **Đường cong ROC**: là biểu đồ được vẽ bằng cách tính tỷ lệ dương thực tế (recall) và tỷ lệ dương sai:
* **AUC**: là diện tích dưới đường cong ROC. AUC đo lường khả năng phân biệt giữa các lớp của mô hình. AUC nằm trong khoảng từ 0 đến 1, mô hình có thể được xem là tốt khi AUC từ 0.8 trở lên.

Vẽ biểu đồ huấn luyện: Sử dụng Matplotlib để vẽ biểu đồ theo dõi độ chính xác (accuracy) và độ mất mát (loss) trên tập train và tập validation. Biểu đồ trực quan hóa quá trình huấn luyện của mô hình, giúp phát hiện các vấn đề trong quá trình huấn luyện như overfitting

# PHÂN TÍCH KẾT QUẢ

## Kết quả

Bảng 2.1: Hyperparameter của Resnet50

|  |  |
| --- | --- |
| **Hyperparameter** | |
| **Số lớp trainable** | 10 |
| **Learning Rate** | 0.001 (trước khi fine-tune), 0.0001 (khi fine-tune) |
| **Batch Size** | 32 |
| **Số Epochs** | 14 |
| **Optimizer** | Adam |
| **Dropout** | 0.5 |
| **Accuracy (Train)** | 0.9918 |
| **Accuracy (Val)** | 0.9538 |
| **Thời gian huấn luyện (Epochs)** | 728s - 804s |

**A graph of different colored lines

Description automatically generated with medium confidence**

Hình 2.1: Biểu đồ theo dõi hiệu suất mô hình ResNet50

**A screenshot of a computer

Description automatically generated**

Hình 2.2: Confusion Matrix của mô hình ResNet50

**A graph with a line

Description automatically generated with medium confidence**

Hình 2.3: Biểu đồ ROC và chỉ số AUC từng lớp của mô hình ResNet50

## Phân tích

Dựa vào chỉ số Accuracy, Precision, Recall, F1-Score của mô hình rút ra được kết luận sau:

* Các chỉ số đánh giá Accuracy, Precision, Recall, F1-Score có giá trị tương đương nhau.
* Resnet50 mô hình cho kết quả tốt với các chỉ số hơn 96%. Điều này cho thấy mô hình này có khả năng xử lý tốt trên tập dữ liệu hình ảnh.
* Cùng với biểu đồ theo dõi hiệu suất, ta có thể thấy độ chính xác và độ mất mát của mô hình trong quá trình huấn luyện khá tốt, tuy nhiên trong khi sử dụng tập kiểm thử song song quá trình huấn luyện thì kết quả có dao động lên xuống qua từng epoch, từ đó có thể thấy có overfitting nhẹ nhưng không đáng lo ngại.

Kết quả của Confusion Matrix:

* Mô hình phân loại tốt ở các lớp Dyskeratotic, Superficial-Intermediate và Parabasal. Tuy nhiên lớp Parabasal trong một số trường hợp vẫn dễ bị nhầm lẫn với lớp khác, đặc biệt là Metaplastic.
* Thông qua kết quả của ma trận nhầm lẫn, có thể thấy được mô hình đều khó khăn trong việc phân loại giữa lớp Koilocytotic và Metaplastic.
* Lý do cho 2 vấn đề trên là, theo lý thuyết việc phân biệt 2 lớp Parabasal và Metaplastic khó khăn vì màu sắc khi tế bào được nhuộm. Tuy nhiên, trong mô hình học máy, việc phân biệt màu sắc (Metaplastic có màu đậm hơn, Parabasal nhạt hơn) giúp mô hình học máy phân loại hiệu quả.
* Giữa Koilocytotic và Metaplastic lại có sự tương đồng lớn về cấu trúc của tế bào (không bào lớn và bào tương viền đậm, cấu trúc nhân của tế bào trong hai lớp đều có các đặc điểm khó phân biệt như nhân lệch tâm và hình thái không đồng đều).

Quan sát đồ thị ROC và chỉ số AUC trong biểu đồ ta có thể thấy, như kết quả trong confusion matrix, các mô hình nhầm lẫn nhiều nhất trong 2 lớp Koilocytotic và Metaplastic.

# KẾT LUẬN

## Ý nghĩa của bài nghiên cứu

Qua bài nghiên cứu cho thấy mô hình ResNet50 phù hợp cho bài toán phân loại hình ảnh tế bào. Mô hình càng hiện đại, càng cho kết quả tốt hơn trong thời gian ngắn hơn.

Việc ứng dụng mô hình ResNet-50 giúp giảm thiểu nhu cầu về dữ liệu lớn, rút ngắn thời gian huấn luyện, đồng thời tận dụng được các đặc trưng tổng quát đã học từ dữ liệu trước đó.

Nghiên cứu có tiềm năng hỗ trợ bác sĩ trong việc phát hiện sớm các bất thường ở tế bào cổ tử cung, góp phần nâng cao hiệu quả trong công tác phòng ngừa và điều trị ung thư cổ tử cung.

## Hạn chế

Nghiên cứu còn phụ thuộc vào chất lượng và số lượng của tập dữ liệu hình ảnh tế bào cổ tử cung.

Mô hình có thể gặp khó khăn khi phân loại những hình ảnh không rõ ràng hoặc chứa nhiều nhiễu.

Hình ảnh trong tập dữ liệu hiện tại có kích thước không được ổn định do quá trình cắt ảnh được làm một cách thủ công từ ảnh phết tế bào cổ tử cung. Nếu có thể cải thiện được vấn đề này, thì mô hình xây dựng dựa trên tập dữ liệu đó sẽ cho kết quả khả quan hơn.

# TÀI LIỆU THAM KHẢO

[1] “BGI Genomics 2023 Global State of Cervical Cancer Awareness Report.” Accessed: Jul. 20, 2023. [Online]. Available: https://www.bgi.com/global/news/bgi-genomics-global-2023-state-of-cervical-cancer-awareness-report

[2] M. E. Plissiti, P. Dimitrakopoulos, G. Sfikas, C. Nikou, O. Krikoni, and A. Charchanti, “Sipakmed: A New Dataset for Feature and Image Based Classification of Normal and Pathological Cervical Cells in Pap Smear Images,” in *Proceedings - International Conference on Image Processing, ICIP*, IEEE Computer Society, Aug. 2018, pp. 3144–3148. doi: 10.1109/ICIP.2018.8451588.

[3] L. Zhang, Y. Bian, P. Jiang, and F. Zhang, “A Transfer Residual Neural Network Based on ResNet-50 for Detection of Steel Surface Defects,” *Applied Sciences (Switzerland)*, vol. 13, no. 9, May 2023, doi: 10.3390/app13095260.

[4] “Papanicolaou stain.” Accessed: Sep. 26, 2024. [Online]. Available: https://en.wikipedia.org/wiki/Papanicolaou\_stain

[5] “Papanicolaou Stain (Pap Stain) Protocol.” Accessed: Sep. 26, 2024. [Online]. Available: https://ihcworld.com/2024/01/26/papanicolaou-stain-pap-stain-protocol/

# PHỤ LỤC

Ái eosin (hồng)/ cyan (xanh lam)/ orangeophilic (cam sáng): Trong xét nghiệm Pap, sử dụng phương pháp nhuộm Papanicolaou để nhuộm tế bào. Một số thành phần của tế bào, tùy theo tính chất hóa học, sẽ bắt và giữ lại các loại thuốc nhuộm khác nhau, cho phép chúng hiển thị rõ ràng dưới kính hiển vi [4], [5]. Trong đó:

* Ái eosin (ưa hồng): Eosin nhuộm màu hồng cho bào tương của tế bào không bị sừng hóa và hiển thị màu hồng nhạt dưới kính hiển vi (Dùng để nhận dạng loại tế bào Superficial-Intermediate).
* Ái cyan (ưa xanh lam): Nhân tế bào sẽ bắt và giữ thuốc nhuộm Hematoxylin và sẽ có màu xanh lam (Dùng để nhận dạng loại tế bào Parabasal và Superficial-Intermediate, Koilocytic).
* Ái orangeophilic (ưa cam sáng): Orange G nhuộm màu cam sáng cho bào tương của các tế bào đã sừng hóa hoặc có chứa nhiều keratin. (Dùng để nhận dạng loại tế bào Dyskeratotic).

Sự bắt màu này của tế bào giúp phân biệt các loại tế bào khác nhau trong xét nghiệm Pap.

# PHÂN CÔNG CÔNG VIỆC

|  |  |
| --- | --- |
| Thành viên | Nhiệm vụ |
| Trần Thị Thiện Nhân | 1. Format word nhóm 2. Chương 1 – 1.1 Tổng quan về tập dữ liệu 3. Chương 2: Phân tích kết quả 4. Chương 3: Kết luận 5. Code ứng dụng |
| Nguyễn Thanh Bình | 1. Powerpoint nhóm 2. Chương 1 – 1.2.1: Định nghĩa CNN 3. Chương 1 – 1.5: Fine-tuning |
| Trịnh Trần Sỹ Đông | 1. Chương 1 – 1.2.2: Cấu trúc mạng CNN 2. Chương 1 – 1.4: Pre-trained model |
| Lâm Hồng Phong | 1. Chương 1 – 1.3 Mô hình ResNet-50 2. Chương 1 – 1.6: Tiền xử lý dữ liệu 3. Chương 1 – 1.7: Đánh giá mô hình |