

**BỘ NÔNG NGHIỆP VÀ MÔI TRƯỜNG
PHÂN HIỆU TRƯỜNG ĐẠI HỌC THỦY LỢI
BỘ MÔN CÔNG NGHỆ THÔNG TIN**



BÁO CÁO MÔN: HỌC MÁY

**TÊN ĐỀ TÀI: XÂY DỰNG MÔ HÌNH HỌC SÂU TRONG BÀI TOÁN
NHẬN DIỆN KHỐI U NÃO QUA ẢNH CỘNG HƯỞNG TỪ**

Giảng viên:	ThS. Vũ Thị Hạnh
Sinh viên thực hiện:	2351267275 - Nguyễn Văn Tấn Phát 2351267271 - Nguyễn Hoàng Lộc 2351267258 - Phạm Thành Doanh
Lớp:	S26-65TTNT

TP. Hồ Chí Minh, ngày ... tháng ... năm 2026

NHẬN XÉT CỦA GIẢNG VIÊN HƯỚNG DẪN

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

TP. Hồ Chí Minh, ngày ... tháng ... năm 2025

Chữ ký của giảng viên

LỜI CẢM ƠN

Nhóm em xin chân thành cảm ơn ThS. Vũ Thị Hạnh – giảng viên phụ trách môn học đã tận tình giảng dạy, hướng dẫn và hỗ trợ nhóm em trong suốt quá trình học tập. Những kiến thức chuyên môn, phương pháp tiếp cận vấn đề cùng sự chỉ dẫn sát sao của cô đã giúp nhóm em hiểu rõ hơn nội dung môn học và hoàn thành bài báo cáo đúng yêu cầu.

Mặc dù nhóm đã cố gắng hết sức, bài làm vẫn khó tránh khỏi những thiếu sót do thời gian và kinh nghiệm còn hạn chế. Nhóm em rất mong nhận được những nhận xét và góp ý từ cô để có thể hoàn thiện hơn kiến thức cũng như kỹ năng trong thời gian tới.

Nhóm em xin chân thành cảm ơn cô!

MỤC LỤC

CHƯƠNG I: GIỚI THIỆU ĐỀ TÀI.....	7
1.1. Lý do chọn đề tài.....	7
1.2. Mục tiêu nghiên cứu.....	7
1.3. Đối tượng và Phạm vi nghiên cứu.....	8
1.4. Ý nghĩa khoa học và thực tiễn.....	8
CHƯƠNG 2: MỤC TIÊU VÀ BÀI TOÁN ĐẶT RA.....	8
CHƯƠNG 3: MÔ TẢ DỮ LIỆU, TIỀN XỬ LÝ VÀ TĂNG CƯỜNG DỮ LIỆU	

NÂNG CAO	9
3.1. Chiến lược Tích hợp và Phân phối Dữ liệu (Dataset Integration Strategy).....	9
3.1.1. Cấu trúc và Quy mô dữ liệu (Data Structure & Scale).....	11
3.1.2. Phân tích sự cân bằng lớp (Class Balance Analysis).....	11
3.2. Kỹ thuật Tiền xử lý Tín hiệu Hình ảnh (Image Signal Preprocessing)	12
3.2.1. Cân bằng biểu đồ thích ứng giới hạn độ tương phản (CLAHE).....	12
3.2.2. Kỹ thuật đưa về kích thước chuẩn (Resizing Strategy)	13
3.3. Hệ thống Tăng cường Dữ liệu Nâng cao (Advanced Data Augmentation Pipeline).....	14
3.3.1. Biến đổi Hình học và Tọa độ (Geometric Transforms).....	14
3.3.2. Biến dạng Đàn hồi và Phi tuyến (Elastic Deformations).....	14
3.3.3. Biến thiên Tín hiệu và Chất lượng (Intensity & Quality).....	14
3.3.4. Kỹ thuật Chính quy hóa không gian (Spatial Regularization).....	15
3.4. Chuẩn hóa Đầu vào Mạng Nơ-ron (Neural Network Input Normalization)..	15
CHƯƠNG 4: PHƯƠNG PHÁP THỰC HIỆN	15
4.1. Các mô hình học máy sử dụng.....	15
4.1.1. Mô hình ResNet50	15
4.1.2. Mô hình CNN Pro 99 (Kiến trúc tùy chỉnh).....	16
4.1.3. Mô hình Xception.....	18
4.1.4. Mô hình EfficientNetB3	19
4.2. Mô hình Học đa nhiệm tích hợp cơ chế chú ý (Multi-task Learning with Attention).....	20
4.2.1. Kiến trúc mạng tổng quát.....	21
4.2.2. Các nhánh nhiệm vụ chuyên biệt.....	21
4.2.3. Hàm mất mát hỗn hợp (Multi-task Loss Strategy)	22
4.2.4. Chiến lược huấn luyện nâng cao và Tối ưu hóa	22

CHƯƠNG 5: KẾT QUẢ VÀ ĐÁNH GIÁ CHI TIẾT CÁC MÔ HÌNH.....	22
5.1. Kết quả Mô hình Xception.....	23
5.2. Kết quả Mô hình Resnet50	25
5.3. Kết quả Mô hình CNN Pro 99	27
5.4. Kết quả mô hình EfficientNetB3	29
5.5. Kết quả Mô hình Đa nhiệm EfficientNetB3.....	30
5.6. Kết quả Mô hình Đa nhiệm Xception.....	33
5.7. So sánh Kết quả và Thảo luận (Comparative Analysis).....	36

DANH MỤC HÌNH ẢNH

Hình 3.1 Ảnh mẫu của mỗi lớp.....	11
Hình 3.2 Phân bố các lớp ở tập train, val, test.....	13
Hình 3.3 Hình mẫu sau khi thực hiện tiền xử lý.....	14
Hình 5.1 Biểu đồ history train Xception.....	24
Hình 5.2 Biểu đồ Confusion matrix Xception.....	24
Hình 5.3 Biểu đồ ROC Xception.....	25
Hình 5.4 Biểu đồ history train Resnet50.....	26
Hình 5.5 Biểu đồ confusion matrix Resnet50.....	26
Hình 5.6 Biểu đồ ROC Resnet50.....	27
Hình 5.7 Biểu đồ history train CNN Pro 99.....	28
Hình 5.8 Biểu đồ ROC CNN Pro 99.....	28
Hình 5.9 Biểu đồ Confusionmatrix CNN Pro 99.....	29
Hình 5.10 Biểu đồ history train EfficientnetB3.....	30
Hình 5.11 Biểu đồ confusion matrix EfficientnetB3.....	30
Hình 5.12 Biểu đồ history train mô hình đa nhiệm65 efficientnetB3.....	31
Hình 5.13 Biểu đồ confusion matrix model đa nhiệm Efficientnetb3.....	32
Hình 5.14 Ảnh kết quả mô hình đa nhiệm Efficientnetb3.....	33
Hình 5.15 Ảnh kết quả mô hình đa nhiệm Xception.....	34
Hình 5.16 Biểu đồ confusionmatrix mô hình đa nhiệm xception.....	35

DANH MỤC BẢNG

<i>Bảng 5.1. Bảng so sánh tổng hợp các chỉ số hiệu năng</i>	<i>41</i>
<i>Bảng 5.2. Bảng so sánh tổng hợp các chỉ số hiệu năng của model đa nhiệm ..</i>	<i>41</i>

CHƯƠNG I: GIỚI THIỆU ĐỀ TÀI

1.1. Lý do chọn đề tài

Trong kỷ nguyên công nghệ 4.0, việc ứng dụng Trí tuệ nhân tạo (AI) vào y tế không còn là xa lạ, đặc biệt là trong lĩnh vực chẩn đoán hình ảnh. Khối u não là một trong những bệnh lý nguy hiểm nhất, đòi hỏi việc phát hiện sớm và can thiệp kịp thời để tăng cơ hội sống sót cho bệnh nhân.

Tuy nhiên, quy trình chẩn đoán hiện nay vẫn gặp nhiều thách thức:

- ~ Sự quá tải của hệ thống y tế: Số lượng ảnh chụp MRI cần phân tích mỗi ngày là rất lớn, gây áp lực nặng nề lên các bác sĩ chẩn đoán hình ảnh.
- ~ Độ phức tạp của khối u: Khối u não có hình dạng, kích thước và vị trí thay đổi rất đa dạng. Việc phân loại thủ công các loại u (như u màng não, u tuyến yên, u tế bào thần kinh đệm) và khoanh vùng chính xác (segmentation) vùng tổn thương đòi hỏi sự tập trung cực cao và tốn nhiều thời gian.
- ~ Tính khách quan: Việc đánh giá bằng mắt thường có thể dẫn đến những sai sót không đáng có do yếu tố mệt mỏi hoặc kinh nghiệm cá nhân của bác sĩ.

Vì vậy, việc phát triển một hệ thống tự động hóa quá trình phân loại và phân đoạn khối u não là vô cùng cấp thiết, đóng vai trò như một "ý kiến thứ hai" (second opinion) đáng tin cậy cho đội ngũ y tế.

1.2. Mục tiêu nghiên cứu

Đề tài được thực hiện với các mục tiêu cụ thể sau:

Về mặt kỹ thuật: Xây dựng thành công mô hình học sâu (Deep Learning) có khả năng thực hiện song song hai nhiệm vụ:

- ~ Phân loại (Classification): Nhận diện chính xác 4 nhóm đối tượng: *U màng não (Meningioma)*, *U tuyến yên (Pituitary Tumor)*, *U tế bào thần kinh đệm (Glioma)* và *Không có u (No Tumor)*.
- ~ Phân đoạn (Segmentation): Tự động tạo mặt nạ (Mask) để khoanh vùng vị trí và diện tích khối u trên ảnh.

Về mặt tối ưu: Thử nghiệm và so sánh hiệu năng của các kiến trúc mạng hiện đại như Xception, EfficientNetB3, ResNet50 để tìm ra phương án tối ưu nhất về cả độ chính xác (Accuracy) và tốc độ xử lý.

Về mặt ứng dụng: Phát triển một giao diện Web trực quan (Streamlit) cho phép người dùng tải ảnh lên và nhận kết quả chẩn đoán tức thì kèm theo các bản đồ nhiệt (Grad-CAM) để giải thích lý do mô hình đưa ra quyết định

1.3. Đối tượng và Phạm vi nghiên cứu

Đối tượng nghiên cứu: Các thuật toán học sâu (CNN, Multi-task Learning) và bộ dữ liệu ảnh MRI não công khai.

Phạm vi nghiên cứu:

- ~ Tập trung xử lý ảnh MRI 2D (lát cắt).
- ~ Phân loại 4 lớp cơ bản thường gặp trong thực tế.
- ~ Sử dụng kỹ thuật Transfer Learning để tận dụng sức mạnh của các mô hình đã được huấn luyện sẵn.

1.4. Ý nghĩa khoa học và thực tiễn

Ý nghĩa khoa học: Đóng góp thêm một phương pháp tiếp cận hiệu quả trong việc kết hợp học đa nhiệm (Multi-task Learning) để giải quyết hai bài toán khác nhau trên cùng một kiến trúc mạng.

Ý nghĩa thực tiễn: Hệ thống có tiềm năng hỗ trợ giảm tải công việc cho bác sĩ, tăng tốc độ chẩn đoán sơ bộ và hỗ trợ đào tạo cho các sinh viên ngành y hoặc bác sĩ thực tập.

CHƯƠNG 2: MỤC TIÊU VÀ BÀI TOÁN ĐẶT RA

Mục tiêu chính của đề tài là xây dựng một hệ thống học sâu tích hợp có khả năng tự động hóa quy trình phân tích ảnh MRI não, từ đó hỗ trợ đội ngũ y tế đưa ra kết quả chẩn đoán nhanh chóng, khách quan và giảm thiểu sai sót do yếu tố con người.

Bài toán đặt ra là phát triển một kiến trúc mạng nơ-ron học đa nhiệm (Multi-task Learning) thực hiện đồng thời hai nhiệm vụ: Phân loại ảnh vào 4 nhóm bệnh lý (Meningioma, Glioma, Pituitary tumor và No tumor) và Phân đoạn chính xác vùng chứa khối u thông qua mặt nạ nhị phân. Hệ thống không chỉ yêu cầu đạt độ chính xác cao (trên 95%) mà còn phải đảm bảo tính minh bạch thông qua kỹ thuật giải thích mô hình (Grad-CAM), giúp trực quan hóa vùng tổn thương để bác sĩ dễ dàng kiểm chứng. Cuối cùng, toàn bộ giải pháp sẽ được triển khai lên giao diện Web để tạo ra một công cụ hỗ trợ chẩn đoán thực tiễn và dễ tiếp cận.

CHƯƠNG 3: MÔ TẢ DỮ LIỆU, TIỀN XỬ LÝ VÀ TĂNG CƯỜNG DỮ LIỆU NÂNG CAO

Trong các hệ thống chẩn đoán y tế dựa trên học sâu, chất lượng và độ phong phú của dữ liệu đầu vào đóng vai trò tiên quyết đối với độ chính xác lâm sàng. Chương này trình bày chi tiết quy trình xây dựng không gian dữ liệu, từ việc tích hợp đa nguồn, phân tích thống kê, đến các kỹ thuật tiền xử lý tín hiệu và tăng cường dữ liệu thích ứng (Adaptive Data Augmentation) được thiết kế riêng cho đặc thù ảnh cộng hưởng từ (MRI).

3.1. Chiến lược Tích hợp và Phân phối Dữ liệu (Dataset Integration Strategy)

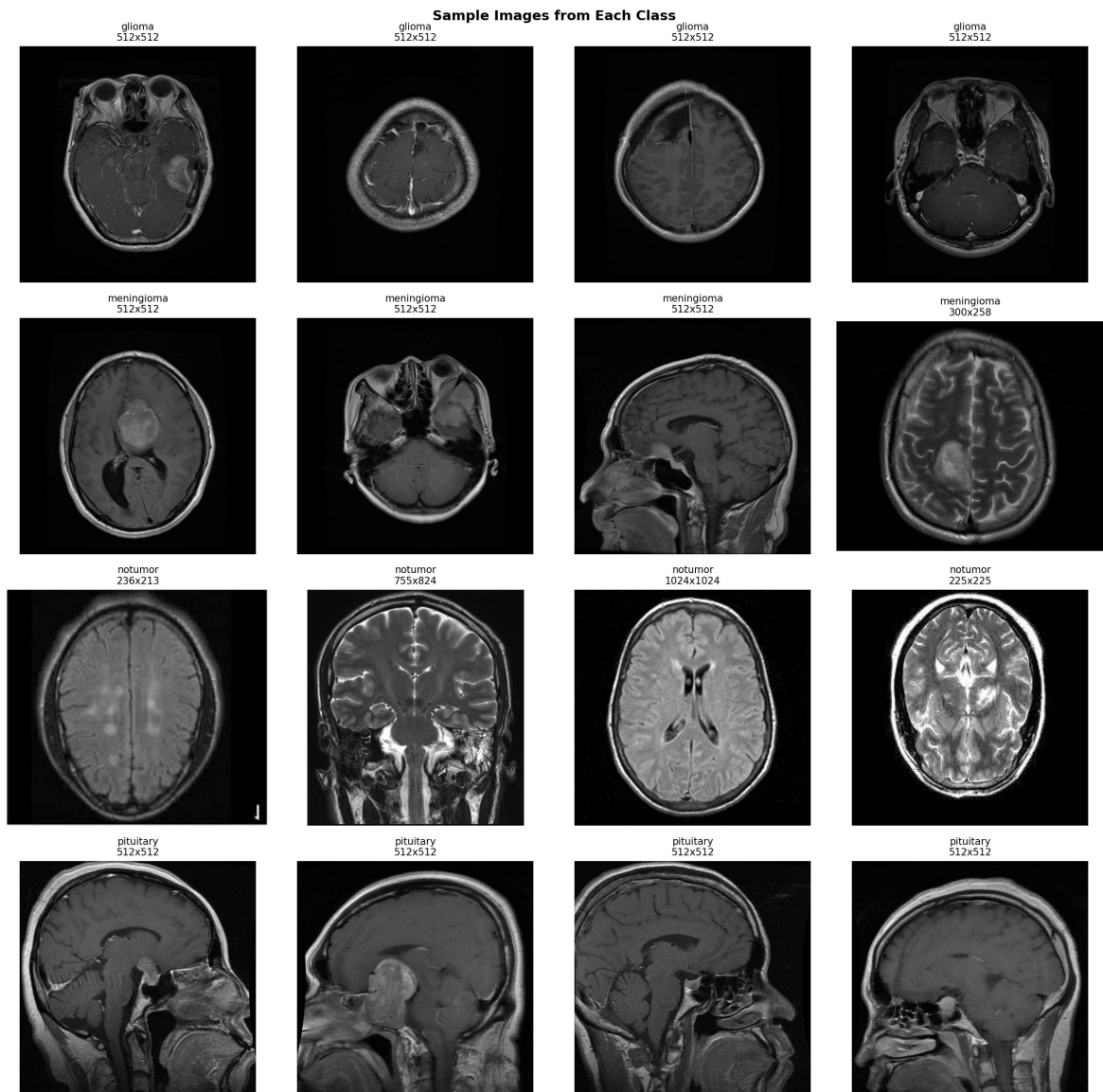
Một trong những rào cản lớn nhất của việc ứng dụng Trí tuệ nhân tạo trong y tế là sự khan hiếm dữ liệu chất lượng cao và sự thiếu đồng nhất giữa các thiết bị thu nhận hình ảnh. Để vượt qua thách thức này và xây dựng một mô hình có khả năng tổng quát hóa mạnh mẽ, chúng tôi không dựa vào một nguồn đơn lẻ mà thực hiện chiến lược hợp nhất dữ liệu (Data Fusion) từ hai kho dữ liệu chuẩn mực quốc tế: Brain Tumor MRI Dataset (nguồn Masoud Nickparvar) và BRISC 2025 Benchmark.

Bộ dữ liệu Brain Tumor MRI đóng vai trò là nền tảng cốt lõi cho bài toán phân loại nhờ quy mô lớn, bao gồm tới 7.023 hình ảnh được thu thập từ nhiều bệnh nhân khác nhau. Nguồn dữ liệu phong phú này cung cấp vô vàn các biến thể về hình thái của ba loại khối u chính (Glioma, Meningioma, Pituitary), giúp các mô hình học sâu như ResNet50 hay Xception tránh được hiện tượng quá khớp (Overfitting) và học được các đặc trưng bất biến của từng loại bệnh lý bất kể sự thay đổi về góc độ chụp hay kích thước khối u.

Tuy nhiên, để giải quyết bài toán phân đoạn và tăng cường độ chính xác lâm sàng, chúng tôi tích hợp thêm bộ dữ liệu tiêu chuẩn vàng BRISC 2025. Điểm ưu việt tuyệt đối của bộ dữ liệu này nằm ở sự hiện diện của các mặt nạ phân đoạn (Segmentation Masks) ở cấp độ điểm ảnh (pixel-level), được dán nhãn thủ công tỉ mỉ bởi các bác sĩ chuyên khoa chẩn đoán hình ảnh. Đặc biệt, lớp "Không u" (No Tumor) trong BRISC không chỉ chứa ảnh não lành đơn thuần mà còn bao gồm các tổn thương lành tính gây nhiễu như nang, áp xe hay các đám vôi hóa. Việc đưa các mẫu "phủ định khó" (Hard Negative Samples) này vào huấn luyện giúp mô hình phân biệt rạch ròi giữa khối u

ác tính thực sự và các tổn thương giả, giảm thiểu tối đa tỷ lệ báo động giả trong thực tế.

Thông qua cơ chế hợp nhất, chúng tôi đã đồng bộ hóa nhãn của cả hai nguồn về một hệ quy chiếu chuẩn 4 lớp (Glioma, Meningioma, Pituitary, No Tumor), tạo ra một không gian dữ liệu tổng hợp mô phỏng môi trường đa trung tâm. Sự kết hợp này mang lại tính bổ trợ hoàn hảo: bộ dữ liệu Nickparvar cung cấp bề rộng (số lượng mẫu lớn để học tổng quát), trong khi BRISC cung cấp chiều sâu (chi tiết giải phẫu để học chính xác), tạo nên tiền đề vững chắc cho sự hội tụ của các thuật toán học đa nhiệm phức tạp.



Hình 3.1 Ảnh mẫu của mỗi lớp

Quan sát các mẫu ảnh đại diện cho thấy sự khác biệt rõ rệt về đặc điểm hình thái học giữa các lớp. Cụ thể, Glioma thường hiện diện là các vùng tăng tín hiệu nằm sâu trong nhu mô não với ranh giới kém rõ ràng; Meningioma có đặc điểm khối u tròn trịa, ranh giới rõ nét, thường nằm sát màng não hoặc xương sọ gây hiệu ứng chèn ép; Pituitary (U tuyến yên) nằm đặc trưng tại vùng hố yên ngay sau hốc mắt; trong khi đó, nhóm No Tumor thể hiện cấu trúc giải phẫu não bộ bình thường, cân đối và không có bất kỳ tổn thương khu trú nào. Chất lượng hình ảnh sắc nét, các đặc trưng bệnh lý được thể hiện rõ ràng trên các mặt cắt, tạo tiền đề thuận lợi cho mô hình học các đặc trưng phân biệt.

3.1.1. Cấu trúc và Quy mô dữ liệu (Data Structure & Scale)

Sau khi hoàn tất quá trình làm sạch dữ liệu (Data Cleaning) để loại bỏ các tệp tin bị hỏng (corrupted files) và chuẩn hóa định dạng ảnh, chúng tôi đã xây dựng được một không gian dữ liệu hợp nhất với tổng số 7.023 bản ghi hình ảnh MRI não chất lượng cao. Để đảm bảo tính khách quan và khả năng đánh giá chính xác hiệu năng của mô hình, toàn bộ tập dữ liệu không được sử dụng tùy tiện mà được tổ chức theo một chiến lược phân chia nghiêm ngặt.

Tập Huấn luyện (Training Set): Bao gồm 5.712 hình ảnh (chiếm khoảng 81% dữ liệu). Đây là không gian dữ liệu chính nơi các mô hình học sâu thực hiện hàng triệu phép tính toán đạo hàm để tối ưu hóa các tham số trọng số. Tập dữ liệu này đủ lớn để bao quát hầu hết các biến thể hình thái học của khối u, từ kích thước, vị trí cho đến độ tương phản, giúp mô hình xây dựng được khả năng tổng quát hóa tốt.

Tập Kiểm thử (Testing Set): Bao gồm 1.311 hình ảnh (chiếm khoảng 19% dữ liệu). Đây là tập dữ liệu "cách ly" (Hold-out set), hoàn toàn không tham gia vào quá trình huấn luyện hay tinh chỉnh tham số. Vai trò duy nhất của tập này là đóng vai trò như một trọng tài độc lập, dùng để đánh giá hiệu suất thực tế của mô hình sau khi đã hoàn thiện, đảm bảo các kết quả báo cáo (Accuracy, F1-Score) phản ánh trung thực khả năng chẩn đoán trong môi trường thực tế.

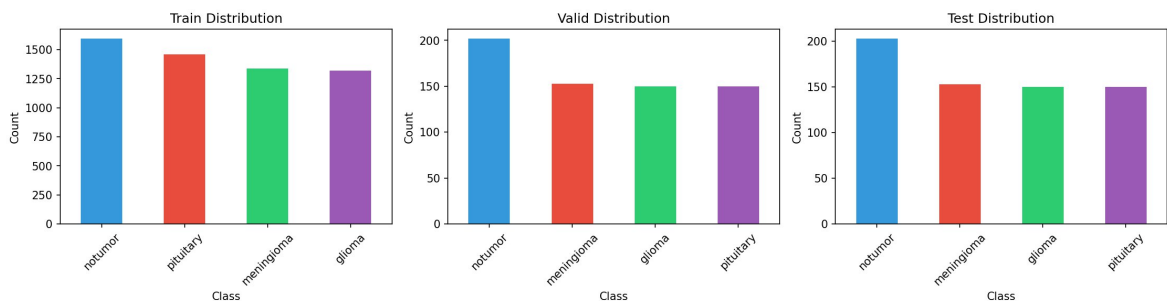
Về mặt phân loại, dữ liệu được gán nhãn chính xác vào 4 lớp lâm sàng riêng biệt: Glioma (U tế bào thần kinh đệm), Meningioma (U màng não), Pituitary tumor (U

tuyến yên) và No tumor (Không có khối u). Sự phân chia rạch ròi này là tiền đề cốt lõi để xây dựng các bài toán phân loại đa lớp và phân đoạn ngữ nghĩa về sau.

3.1.2. Phân tích sự cân bằng lớp (Class Balance Analysis)

Biểu đồ phân phối tần suất cho thấy sự cân bằng tương đối giữa 4 nhãn phân loại, giảm thiểu nguy cơ mô hình bị thiên kiến (Bias):

- ~ No Tumor (Không u): ~1.600 mẫu/tập train. Số lượng mẫu lớn nhất, đại diện cho sự đa dạng của cấu trúc não lành mạnh và các tổn thương không phải u (như nang, áp xe - từ bộ BRISC), giúp giảm tỷ lệ Dương tính giả (False Positive).
- ~ Pituitary (U tuyến yên): ~1.450 mẫu.
- ~ Meningioma (U màng não): ~1.350 mẫu.
- ~ Glioma (U thần kinh đệm): ~1.300 mẫu. *Nhận xét:* Độ lệch chuẩn về số lượng mẫu giữa các lớp là thấp, do đó chúng tôi không cần áp dụng các kỹ thuật lấy mẫu lại (Resampling) hay trọng số mất mát (Weighted Loss) phức tạp.



Hình 3.2 Phân bố các lớp ở tập train, val, test

3.2. Kỹ thuật Tiền xử lý Tín hiệu Hình ảnh (Image Signal Preprocessing)

Trước khi đưa vào mạng nơ-ron, tín hiệu MRI thô cần được chuẩn hóa để loại bỏ các biến thiên không mong muốn do thiết bị chụp.

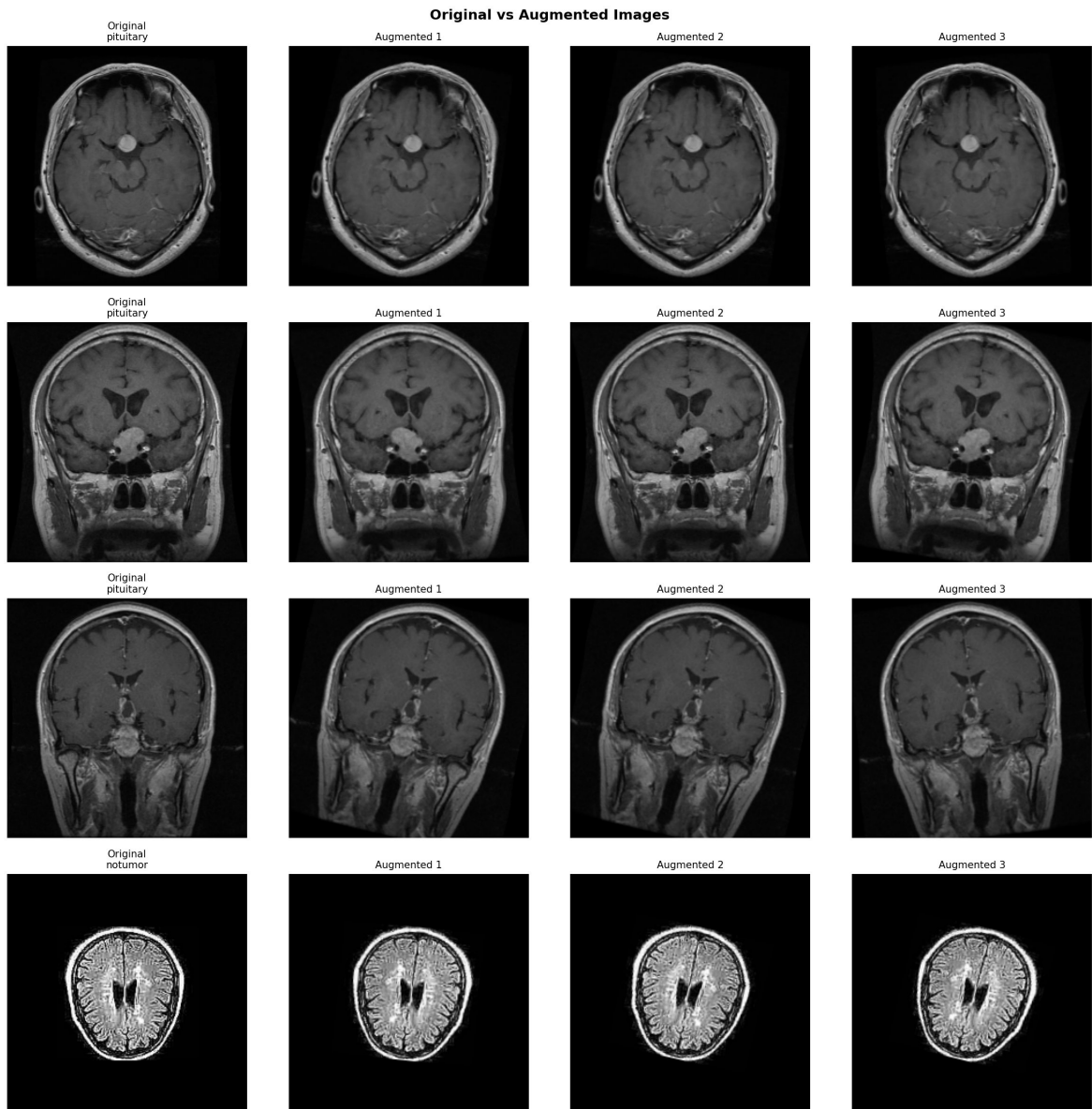
3.2.1. Cân bằng biểu đồ thích ứng giới hạn độ tương phản (CLAHE)

Ảnh MRI thường có phổ histogram tập trung ở vùng tối. Chúng tôi áp dụng thuật toán CLAHE (Contrast Limited Adaptive Histogram Equalization) với các siêu tham số được tinh chỉnh:

- ~ Ngưỡng cắt (Clip Limit): Thiết lập ở mức 2.0 cho bài toán phân loại và 4.0 cho bài toán phân đoạn (trong bộ BRISC). Ngưỡng cao giúp khuếch đại độ tương phản cực đại tại các ranh giới mô mềm khó nhìn, làm nổi bật đường biên khối u.
- ~ Kích thước lưới (Tile Grid Size): (8, 8). Ảnh được chia thành 64 ô nhỏ để cân bằng sáng cục bộ, tránh việc làm nhiễu hạt (noise amplification) ở các vùng đồng nhất.

3.2.2. Kỹ thuật đưa về kích thước chuẩn (Resizing Strategy)

- ~ Thay vì co giãn ảnh (Scaling) làm thay đổi tỉ lệ khuôn dạng (Aspect Ratio) của sọ não, chúng tôi sử dụng phương pháp "Resize with Padding":
- ~ Tính toán tỉ lệ scale
- ~ Thay đổi kích thước ảnh theo tỷ lệ scale.
- ~ Chèn các đường viền đen (padding giá trị 0) vào phần dư thừa để đạt kích thước đích (224x224 hoặc 320x320). *Ý nghĩa:* Giữ nguyên hình thái học giải phẫu thực tế, giúp mô hình học được các đặc trưng hình học (như độ tròn, độ méo) chính xác.



Hình 3.3 Hình mẫu sau khi thực hiện tiền xử lý

3.3. Hệ thống Tăng cường Dữ liệu Nâng cao (Advanced Data Augmentation Pipeline)

Đối với mô hình học đa nhiệm (Bộ dữ liệu BRISC), chúng tôi triển khai một đường ống tăng cường dữ liệu (Augmentation Pipeline) phức tạp sử dụng thư viện Albumentations. Mỗi hình ảnh đi qua đường ống này có xác suất được biến đổi để tạo ra các bản sao "nhân tạo".

3.3.1. Biến đổi Hình học và Tọa độ (Geometric Transforms)

Mô phỏng sự sai lệch về tư thế bệnh nhân và góc chụp:

- ~ Xoay và lật (Rotate/Flip): Ảnh được xoay ngẫu nhiên trong phạm vi từ -35 độ đến +35 độ với xác suất $p = 0.6$. Ngoài ra, ảnh được lật ngang với xác suất $p = 0.5$ và lật dọc với xác suất $p = 0.4$. Các phép biến đổi này giúp mô hình học được tính bất biến theo hướng và giảm hiện tượng overfitting.
- ~ ShiftScaleRotate: Kỹ thuật này kết hợp dịch chuyển ảnh, phóng to/thu nhỏ trong phạm vi $\pm 18\%$, và xoay trong phạm vi ± 25 độ với xác suất $p = 0.6$. Nhờ đó, mô hình có khả năng thích nghi tốt hơn với các biến đổi về vị trí, kích thước và góc nhìn của đối tượng trong ảnh.

3.3.2. Biến dạng Đàn hồi và Phi tuyến (Elastic Deformations)

Đây là kỹ thuật chuyên sâu cho ảnh y tế, mô phỏng sự biến dạng của mô mềm:

- ~ ElasticTransform: ($\alpha=120$, $\sigma=6.0$). Tạo các biến dạng dẻo cục bộ giống như sự co giãn của mô não.
- ~ GridDistortion & OpticalDistortion: Mô phỏng sự sai lệch quang học của thấu kính máy chụp hoặc sự méo mó không gian. Các kỹ thuật này được áp dụng ngẫu nhiên ($p=0.5$) để tăng độ bền vững (Robustness) cho mô hình phân đoạn.

3.3.3. Biến thiên Tín hiệu và Chất lượng (Intensity & Quality)

Mô phỏng sự khác biệt giữa các dòng máy MRI (1.5 Tesla vs 3.0 Tesla):

- ~ RandomGamma: Điều chỉnh mức Gamma trong khoảng (80, 125) với $p=0.4$, thay đổi độ phi tuyến của độ sáng.
- ~ RandomBrightnessContrast: Biến thiên độ sáng và tương phản
- ~ Nhiễu và Mờ (Noise & Blur): Áp dụng ngẫu nhiên Nhiễu Gaussian, Nhiễu chuyển động (MotionBlur) để giả lập các ảnh chụp bị rung hoặc chất lượng thấp.

3.3.4. Kỹ thuật Chính quy hóa không gian (Spatial Regularization)

- ~ CoarseDropout & GridDropout: Xóa bỏ ngẫu nhiên các vùng hình chữ nhật (tối đa 10 lỗ, kích thước bằng 1/10 ảnh) hoặc xóa theo lưới.
- ~ Cơ chế: Bắt buộc mạng nơ-ron phải "điền vào chỗ trống" dựa trên ngữ cảnh toàn cục (Contextual understanding) thay vì chỉ tập trung vào một đặc điểm cục bộ duy nhất, giúp hạn chế triệt để hiện tượng quá khớp (Overfitting).

3.4. Chuẩn hóa Đầu vào Mạng Nơ-ron (Neural Network Input Normalization)

Ở bước cuối cùng, dữ liệu đi qua tầng chuẩn hóa thống kê để đảm bảo tính hội tụ nhanh nhất cho thuật toán Gradient Descent:

- ~ Giá trị pixel: Được chia cho 255 để đưa về khoảng $[0, 1]$.
- ~ z-score Standardization: Trừ đi giá trị trung bình ($\text{mean}=[0.485, 0.456, 0.406]$) và chia cho độ lệch chuẩn ($\text{std}=[0.229, 0.224, 0.225]$) của bộ dữ liệu ImageNet.
- ~ Kết quả: Dữ liệu đầu vào là các Tensor chuẩn hóa có phân phối xấp xỉ $N(0, 1)$, sẵn sàng cho quá trình huấn luyện chuyển giao (Transfer Learning).

CHƯƠNG 4: PHƯƠNG PHÁP THỰC HIỆN

4.1. Các mô hình học máy sử dụng

Trong đồ án này, chúng tôi tập trung phát triển và so sánh bốn kiến trúc mạng nơ-ron tích chập (CNN) hiện đại nhất để tối ưu hóa việc nhận diện khối u.

4.1.1. Mô hình ResNet50

Nguyên lý hoạt động: ResNet50 (Residual Network) được xây dựng dựa trên cơ chế Kết nối tắt (Skip Connections). Thay vì buộc các lớp mạng phải học lại toàn bộ thông tin từ đầu, cơ chế này cho phép tín hiệu truyền trực tiếp qua các khối tắt, giúp mạng chỉ cần học phần thông tin dư thừa (residual) còn thiếu. Điều này giải quyết triệt để vấn đề triệt tiêu đạo hàm khi huấn luyện các mạng thần kinh có độ sâu lớn lên tới 50 tầng.

Lý do lựa chọn: ResNet50 được coi là kiến trúc tiêu chuẩn trong lĩnh vực chẩn đoán hình ảnh y tế. Cấu trúc này đủ sâu để trích xuất được các đặc trưng ngữ nghĩa phức tạp của khối u (như kết cấu bề mặt, hình dạng xâm lấn) nhưng vẫn đảm bảo tính ổn định và dễ dàng hội tụ hơn so với các mạng VGG truyền thống.

Quy trình huấn luyện (Train Model): Chúng tôi xây dựng mô hình bằng cách ghép nối phần thân ResNet50 (đã học trên ImageNet) với một khối phân loại tùy chỉnh gồm các lớp: Global Average Pooling, Batch Normalization và 3 lớp Dense liên tiếp.

Quy trình huấn luyện chia làm 2 giai đoạn rõ rệt:

- ~ Giai đoạn 1 (Khởi tạo): Đóng băng toàn bộ phần thân ResNet50, chỉ huấn luyện các lớp phân loại mới trong 20 chu kỳ (epochs) để trọng số ổn định.
- ~ Giai đoạn 2 (Tinh chỉnh): Mở khóa (Unfreeze) 10 lớp cuối cùng của phần thân ResNet50 và huấn luyện lại toàn bộ mạng trong 50 chu kỳ tiếp theo. Cách làm này giúp mạng điều chỉnh các bộ lọc cấp cao để phù hợp hoàn toàn với đặc điểm của ảnh MRI não.

Tinh chỉnh siêu tham số (Hyperparameter Tuning):

Bộ tối ưu: Sử dụng thuật toán Adam. Tốc độ học (Learning Rate) được thiết lập là 0.001 ở giai đoạn 1 và giảm xuống 0.0001 ở giai đoạn 2 để tinh chỉnh chính xác các cực tiểu địa phương.

Kỹ thuật chống quá khớp: Hệ thống sử dụng Dropout phân tầng với tỷ lệ giảm dần (0.5 cho lớp đầu, 0.3 và 0.2 cho các lớp sau), kết hợp với lớp chuẩn hóa Batch Normalization.

Cơ chế tự động: Áp dụng Early Stopping (dừng sớm nếu lỗi không giảm sau 8-12 chu kỳ) và ReduceLROnPlateau (tự động giảm tốc độ học đi một nửa nếu quá trình huấn luyện bị chững lại).

Để đánh giá khách quan khả năng tổng quát hóa của mô hình ResNet50 và hạn chế sự phụ thuộc vào một lần chia dữ liệu cố định, nhóm áp dụng phương pháp xác thực chéo K-Fold Cross Validation với $k = 5$. Toàn bộ tập huấn luyện ban đầu được chia thành 5 phần bằng nhau. Ở mỗi vòng lặp (fold), 4 phần được sử dụng để huấn luyện mô hình và 1 phần còn lại được dùng làm tập kiểm định.

Quy trình huấn luyện trong mỗi fold được giữ nguyên các tham số so với mô hình ResNet50 ban đầu, bao gồm kích thước đầu vào, cấu trúc mạng, hàm mất mát và bộ tối ưu. Mỗi fold được huấn luyện theo chiến lược hai giai đoạn: giai đoạn khởi tạo (đóng băng backbone) và giai đoạn tinh chỉnh sâu (fine-tuning). Kết quả đánh giá trên tập kiểm định của từng fold được lưu lại để tính toán độ chính xác trung bình và độ ổn định của mô hình.

4.1.2. Mô hình CNN Pro 99 (Kiến trúc tùy chỉnh)

Nguyên lý hoạt động: Đây là một kiến trúc mạng nơ-ron tích chập tùy chỉnh (Custom CNN) được nhóm chúng tôi thiết kế từ đầu và huấn luyện trực tiếp từ khởi tạo ngẫu nhiên (training from scratch). Mô hình được cấu trúc gồm 4 khối xử lý chính với độ sâu bộ lọc tăng dần lần lượt là 64 – 128 – 256 – 512.

Điểm đột phá của kiến trúc nằm ở việc tích hợp khối chú ý Squeeze-and-Excitation (SE Block) vào các tầng sâu của mạng (khối 3 và khối 4). Cơ chế SE cho phép mạng tự động học mức độ quan trọng của từng kênh đặc trưng, từ đó tập trung nhiều hơn vào các kênh chứa thông tin liên quan đến khối u và giảm ảnh hưởng của các kênh nhiễu không cần thiết.

Lý do lựa chọn: Việc tự thiết kế kiến trúc giúp nhóm kiểm soát hoàn toàn luồng dữ liệu trong mạng. Toàn bộ mô hình sử dụng hàm kích hoạt Swish xuyên suốt các tầng,

đồng thời tích hợp cơ chế chú ý SE nhằm tối ưu riêng cho đặc thù của ảnh MRI y tế. Cách tiếp cận này mang lại tính linh hoạt cao hơn so với các kiến trúc phổ biến như VGG hay ResNet, vốn được thiết kế cho các bài toán tổng quát và đôi khi không phù hợp hoàn toàn với dữ liệu y sinh.

Quy trình huấn luyện (Train Model): Mô hình được huấn luyện trực tiếp từ trọng số khởi tạo ngẫu nhiên trong 70 epochs. Do không sử dụng trọng số tiền huấn luyện, phần Stem của mạng được thiết kế cẩn thận với 3 lớp tích chập liên tiếp nhằm thay thế bước tiền xử lý truyền thống. Thiết kế này giúp mạng học được các đặc trưng sơ cấp như cạnh và góc ngay từ dữ liệu đầu vào.

Tinh chỉnh siêu tham số (Hyperparameter Tuning)

Chống quá khớp mạnh mẽ (Regularization): Đây là ưu điểm nổi bật của mô hình. Nhóm áp dụng L2 Regularization với hệ số 0.0001 cho toàn bộ các lớp tích chập và lớp Dense. Bên cạnh đó, tỷ lệ Dropout được thiết lập tăng dần theo độ sâu của mạng, lần lượt là $0.15 \rightarrow 0.2 \rightarrow 0.25 \rightarrow 0.3 \rightarrow 0.5$ tại lớp cuối, nhằm hạn chế tối đa hiện tượng học vẹt và tăng khả năng khái quát hóa.

Bộ tối ưu: Mô hình sử dụng Adamax với tốc độ học 0.001, giúp quá trình huấn luyện ổn định và hiệu quả trên không gian tham số lớn.

Callback: Cơ chế EarlyStopping được thiết lập với độ kiên nhẫn 12 epochs, cho phép mô hình có đủ thời gian khám phá không gian tham số phức tạp và tìm được điểm hội tụ toàn cục tốt hơn trước khi dừng huấn luyện.

4.1.3. Mô hình Xception

Nguyên lý hoạt động: Kiến trúc Xception (Extreme Inception) là một cải tiến mang tính đột phá từ dòng Inception. Thay vì sử dụng các khối đại diện Inception truyền thống, mô hình này được xây dựng hoàn toàn dựa trên các phép Tích chập tách biệt theo chiều sâu (Depthwise Separable Convolutions). Nguyên lý cốt lõi của nó là tách biệt hoàn toàn quá trình học các mối tương quan không gian (spatial correlations) và các mối tương quan liên kênh (cross-channel correlations). Điều này giúp mô hình giảm thiểu đáng kể số lượng tham số dư thừa trong khi vẫn duy trì khả năng trích xuất đặc trưng mạnh mẽ hơn so với các mạng tích chập thông thường.

Lý do lựa chọn: Trong bài toán phân loại ảnh MRI, các chi tiết kết cấu (texture) của khối u đóng vai trò quan trọng không kém hình dáng tổng thể. Xception, với cơ chế tích chập tách biệt, đã chứng minh được hiệu quả vượt trội trong việc nắm bắt các đặc trưng vi mô này. Chúng tôi lựa chọn Xception kết hợp với kỹ thuật Học chuyển giao (Transfer Learning) để tận dụng các bộ lọc đặc trưng đã được tối ưu hóa trên tập dữ liệu khổng lồ ImageNet, giúp mô hình đạt độ chính xác cao ngay cả khi dữ liệu huấn luyện y tế còn hạn chế.

Quy trình huấn luyện (Train Model): Chúng tôi thiết kế lại phần đầu phân loại (Classification Head) để phù hợp với bài toán 4 lớp khối u:

- ~ Sử dụng lớp Global Max Pooling (thay vì Average Pooling như ResNet) để bắt lấy các đặc trưng nổi bật nhất (như các điểm ảnh sáng bất thường của khối u).
- ~ Tiếp theo là lớp Dense 256 nơ-ron với hàm kích hoạt hiện đại Swish, kết hợp với lớp Dropout 0.5 để giảm thiểu quá khớp.

Quy trình huấn luyện được chia thành 2 giai đoạn chiến lược (Two-phase training):

- ~ Giai đoạn 1 (Trích xuất đặc trưng): Đóng băng toàn bộ mạng nền (Base model). Chỉ huấn luyện các lớp phân loại mới trong 10 epochs với tốc độ học khởi điểm là 0.001.
- ~ Giai đoạn 2 (Tinh chỉnh sâu): Giải băng (Unfreeze) một phần mạng nền. Cụ thể, chúng tôi giữ nguyên trạng thái đóng băng cho 50 lớp đầu tiên (chứa các đặc trưng cạnh/góc cơ bản) và chỉ cho phép huấn luyện các lớp từ 51 trở đi. Quá trình này diễn ra trong 15 epochs với tốc độ học giảm sâu xuống 0.0001.

Tinh chỉnh siêu tham số (Hyperparameter Tuning): Bộ tối ưu (Optimizer): Sử dụng thuật toán Adamax (một biến thể của Adam dựa trên chuẩn vô cực). Adamax được chọn vì tính ổn định cao hơn Adam khi xử lý các cập nhật trọng số trên các kiến trúc có độ sâu lớn và embedding thưa.

Hàm kích hoạt: Sử dụng hàm Swish ($x * \text{sigmoid}(x)$) thay cho ReLU truyền thống, giúp luồng thông tin lan truyền mượt mà hơn qua các lớp sâu.

Cơ chế kiểm soát:

- ~ Early Stopping: Dừng huấn luyện sớm nếu độ chính xác không cải thiện sau 3 epochs (Giai đoạn 1) hoặc 5 epochs (Giai đoạn 2).

- ~ ReduceLROnPlateau: Tự động giảm tốc độ học đi một nửa (factor=0.5) nếu hàm mất mát chững lại sau 2 epochs, giúp mô hình thoát khỏi các điểm cực tiểu cục bộ.

4.1.4. Mô hình EfficientNetB3

Nguyên lý hoạt động: EfficientNet đánh dấu một bước ngoặt trong tư duy thiết kế mạng nơ-ron với phương pháp Compound Scaling (Mở rộng hỗn hợp). Thay vì mở rộng mạng một cách tùy tiện theo chiều sâu (như ResNet) hay chiều rộng, EfficientNet tìm ra một công thức tối ưu để mở rộng đồng thời cả ba chiều: độ sâu mạng, độ rộng kênh và độ phân giải ảnh. Khối xây dựng cơ bản của nó là MBConv (Mobile Inverted Bottleneck Convolution), giúp tối ưu hóa luồng thông tin với chi phí tính toán thấp nhất.

Lý do lựa chọn: Chúng tôi lựa chọn phiên bản B3 với kích thước đầu vào đặc thù là 300x300 pixels (lớn hơn so với chuẩn 224x224 thông thường). Kích thước này cho phép mô hình nhìn thấy rõ hơn các chi tiết vi mô của khối u mà không làm tăng quá nhiều tài nguyên tính toán. EfficientNetB3 hiện đang được coi là kiến trúc cân bằng tốt nhất giữa độ chính xác và tốc độ xử lý cho các ứng dụng y tế.

Quy trình huấn luyện (Train Model): Kiến trúc mạng được thiết kế gồm EfficientNetB3 (ImageNet weights) làm nền, kết nối với lớp GlobalAveragePooling2D và một lớp Dropout có tỷ lệ 0.3 trước khi vào lớp phân loại Softmax. Quy trình huấn luyện gồm 2 giai đoạn chiến lược nhằm bảo vệ các đặc trưng thông kê của mạng:

- ~ Giai đoạn 1 (Huấn luyện Head): Đóng băng toàn bộ mạng nền. Chỉ huấn luyện lớp đầu cuối trong 10 epochs để bộ trích xuất đặc trưng không bị "sốc" bởi gradient từ các lớp khởi tạo ngẫu nhiên.
- ~ Giai đoạn 2 (Tinh chỉnh có trọng lọc): Giải băng (Unfreeze) toàn bộ mạng để tinh chỉnh. Điểm đặc biệt: Chúng tôi áp dụng kỹ thuật đóng băng Batch Normalization (giữ training=False cho các lớp BN). Kỹ thuật này cực kỳ quan trọng khi tinh chỉnh EfficientNet, giúp bảo toàn các thống kê trung bình/phương sai đã học từ ImageNet, ngăn chặn mô hình bị phá vỡ cấu trúc khi huấn luyện trên dữ liệu mới. Giai đoạn này kéo dài 40 epochs.

Tinh chỉnh siêu tham số (Hyperparameter Tuning): Bộ tối ưu tiên tiến: Sử dụng AdamW (Adam với Weight Decay 0.0001) thay vì Adam thường. AdamW

giúp tách biệt việc cập nhật trọng số và suy giảm trọng số, giúp mô hình tổng quát hóa tốt hơn hẳn.

- ~ Điều phối tốc độ học (Scheduler): Thay vì giảm theo bậc thang, chúng tôi sử dụng lịch trình Cosine Decay (Giảm theo hàm Cosine). Tốc độ học khởi đầu ở mức cực nhỏ (0.00003) và giảm dần về 0 theo đường cong Cosine mượt mà, giúp mô hình hội tụ "êm" hơn tại điểm cực trị.
- ~ Hàm mất mát: Sử dụng Categorical Crossentropy kết hợp với Label Smoothing 0.1. Kỹ thuật này ngăn chặn mô hình trở nên quá tự tin vào các nhãn huấn luyện (ngăn xác suất dự đoán tiến tới 1.0 tuyệt đối), giúp giảm thiểu sai số khi gặp các ca bệnh không điển hình.

4.2. Mô hình Học đa nhiệm tích hợp cơ chế chú ý (Multi-task Learning with Attention)

Để giải quyết đồng thời hai bài toán quan trọng nhất trong chẩn đoán hình ảnh MRI là phân loại loại khối u và khoanh vùng vị trí tổn thương, chúng tôi đã phát triển một kiến trúc học đa nhiệm tiên tiến. Cách tiếp cận này không chỉ giúp tiết kiệm tài nguyên tính toán mà còn cho phép hai nhiệm vụ hỗ trợ lẫn nhau trong quá trình học đặc trưng.

4.2.1. Kiến trúc mạng tổng quát

Mô hình được xây dựng dựa trên cấu trúc hình phễu của mạng U-Net nhưng được cải tiến mạnh mẽ ở phần xương sống (Backbone). Chúng tôi triển khai hai biến thể bộ mã hóa (Encoder) để so sánh đối chứng: EfficientNet-B3 (tối ưu về hiệu suất) và Xception (tối ưu về trích xuất không gian). Tuy nhiên, dù sử dụng Backbone nào, cốt lõi của mô hình vẫn nằm ở sự kết hợp của ba thành phần chính:

- ~ Bộ mã hóa (Encoder): Thực hiện nhiệm vụ trích xuất các đặc trưng ngữ nghĩa từ ảnh MRI gốc thông qua các khối tích chập sâu.
- ~ Bộ giải mã (Decoder): Sử dụng các tầng hạ tích chập (Up-sampling) kết hợp với các kết nối tắt (Skip connections) từ Encoder để khôi phục độ phân giải, phục vụ cho nhiệm vụ phân đoạn vùng khối u.

- Cơ chế chú ý CBAM (Convolutional Block Attention Module): Đây là thành phần then chốt giúp mô hình lọc nhiễu, được tích hợp tại cuối Decoder và trước các lớp phân loại. CBAM hoạt động qua hai giai đoạn: Chú ý kênh (Channel Attention) để xác định loại đặc trưng quan trọng và Chú ý không gian (Spatial Attention) để xác định chính xác vị trí khối u trên bản đồ đặc trưng.

4.2.2. Các nhánh nhiệm vụ chuyên biệt

Luồng thông tin sau khi đi qua bộ mã hóa sẽ được chia tách để thực hiện hai nhiệm vụ song song:

- Nhánh Phân đoạn (Segmentation Head): Nhận đầu ra từ Decoder đã qua xử lý bởi CBAM để tạo ra mặt nạ nhị phân (Binary Mask). Nhánh này giúp bác sĩ thấy được vị trí và kích thước thực tế của khối u trên ảnh MRI.
- Nhánh Phân loại (Classification Head): Nhận đặc trưng từ tầng sâu nhất của bộ mã hóa, đi qua khối chú ý CBAM 2048 kênh, sau đó được phẳng hóa (Flatten) và truyền qua chuỗi các lớp nơ-ron dày đặc:

Batch Normalization -> Dropout -> Linear(512) -> SiLU -> Linear(256) -> Linear(4)

Việc sử dụng hàm kích hoạt SiLU giúp mô hình học các mối quan hệ phi tuyến phức tạp tốt hơn so với ReLU truyền thống.

4.2.3. Hàm mất mát hỗn hợp (Multi-task Loss Strategy)

Để mô hình có thể hội tụ tốt trên cả hai nhiệm vụ, chúng tôi thiết kế một hàm mất mát tổng hợp có trọng số:

- Mất mát phân đoạn (Hybrid Segmentation Loss): Là sự kết hợp cân bằng giữa Jaccard Loss (tối ưu hóa độ trùng khớp diện tích) và Focal Loss (giúp mô hình tập trung vào các pixel khó phân biệt tại ranh giới khối u).
- Mất mát phân loại: Sử dụng Cross Entropy kết hợp với kỹ thuật Label Smoothing mức 0.1 để tăng khả năng tổng quát hóa cho mô hình.

- ~ Cân bằng trọng số: Chúng tôi thiết lập tỉ lệ trọng số là 1.2 cho nhánh phân đoạn và 0.4 cho nhánh phân loại. Sự ưu tiên này đảm bảo mô hình ưu tiên sự chính xác về mặt vị trí trước khi đưa ra kết luận về loại bệnh.

4.2.4. Chiến lược huấn luyện nâng cao và Tối ưu hóa

Quá trình huấn luyện được thực hiện tỉ mỉ qua hai giai đoạn trên nền tảng PyTorch:

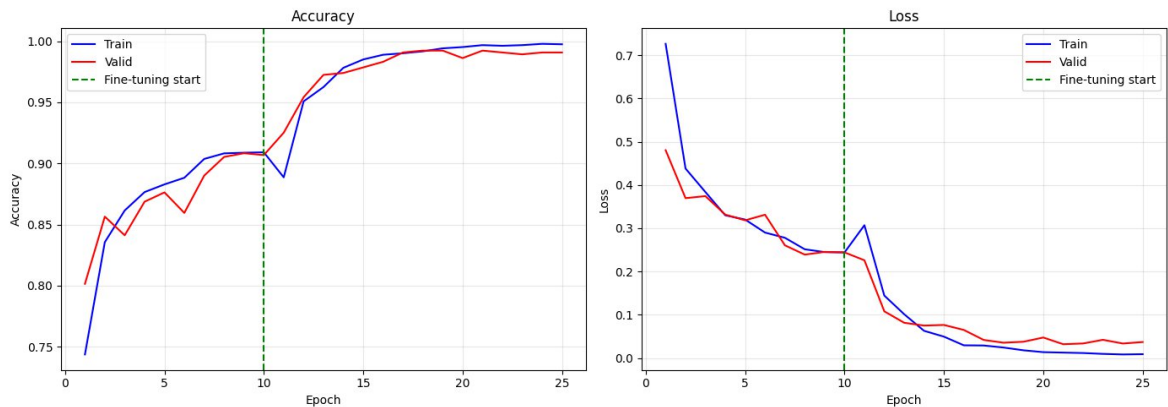
- ~ Giai đoạn huấn luyện bộ phận (Stage 1): Đóng băng phần Encoder, chỉ huấn luyện Decoder và Classifier để các lớp này làm quen với dữ liệu khối u đặc thù với tốc độ học lớn (10^{-3}).
- ~ Giai đoạn tinh chỉnh toàn cục (Stage 2): Giải băng toàn bộ mạng, sử dụng bộ tối ưu AdamW với trọng số suy giảm (Weight Decay) để tinh chỉnh các tham số sâu nhất. Tốc độ học được điều phối bởi lịch trình Cosine Annealing Warm Restarts, giúp mô hình liên tục tìm kiếm các điểm tối ưu mới.
- ~ Tối ưu hóa tài nguyên: Sử dụng kỹ thuật Độ chính xác hỗn hợp (AMP) để tăng gấp đôi tốc độ huấn luyện trên GPU và kỹ thuật Gradient Clipping để giữ cho các giá trị đạo hàm không vượt quá ngưỡng 1.0, đảm bảo tính ổn định tuyệt đối cho quá trình học tập của mạng nơ-ron.

CHƯƠNG 5: KẾT QUẢ VÀ ĐÁNH GIÁ CHI TIẾT CÁC MÔ HÌNH

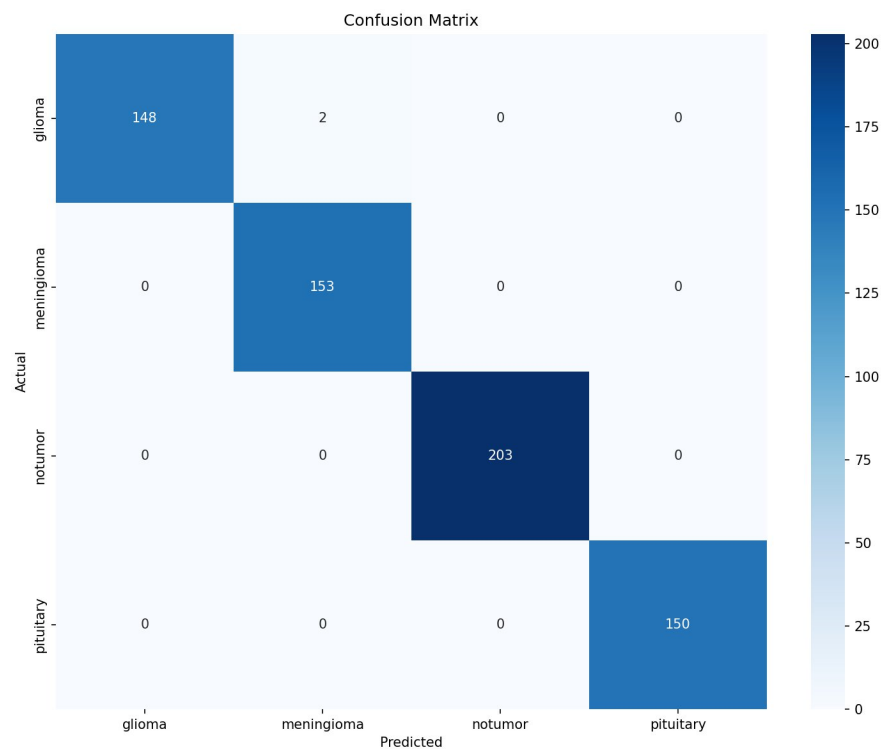
Trong phần này, chúng tôi trình bày kết quả thực nghiệm chi tiết cho từng kiến trúc mô hình. Mỗi mô hình được đánh giá thông qua biểu đồ quá trình huấn luyện (Training History) và Ma trận nhầm lẫn (Confusion Matrix) để làm rõ ưu nhược điểm riêng biệt.

5.1. Kết quả Mô hình Xception

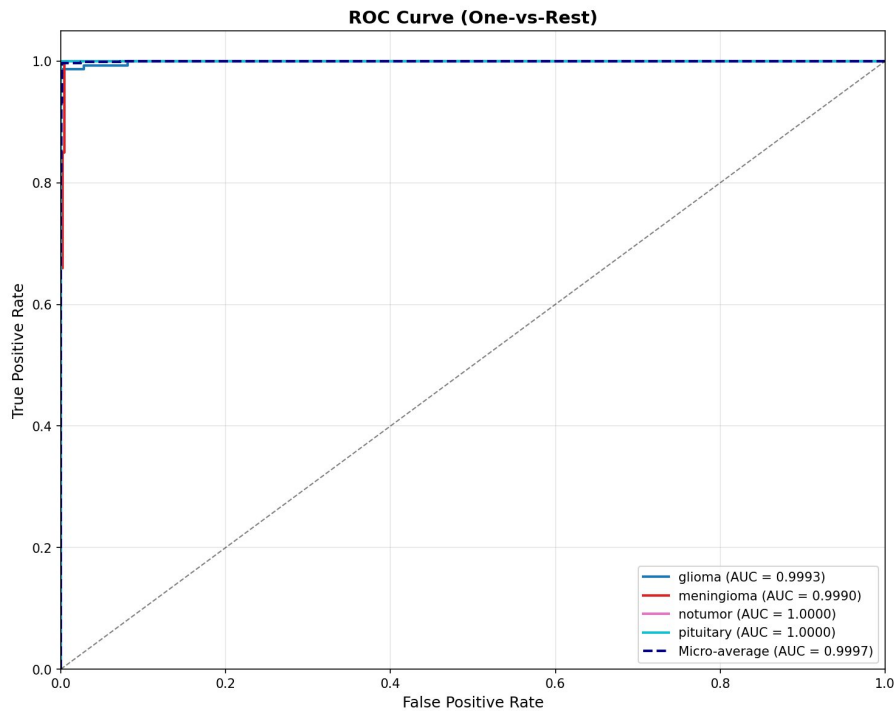
Mô hình Xception với độ chính xác đạt khoảng



Hình 5.1 Biểu đồ history train Xception



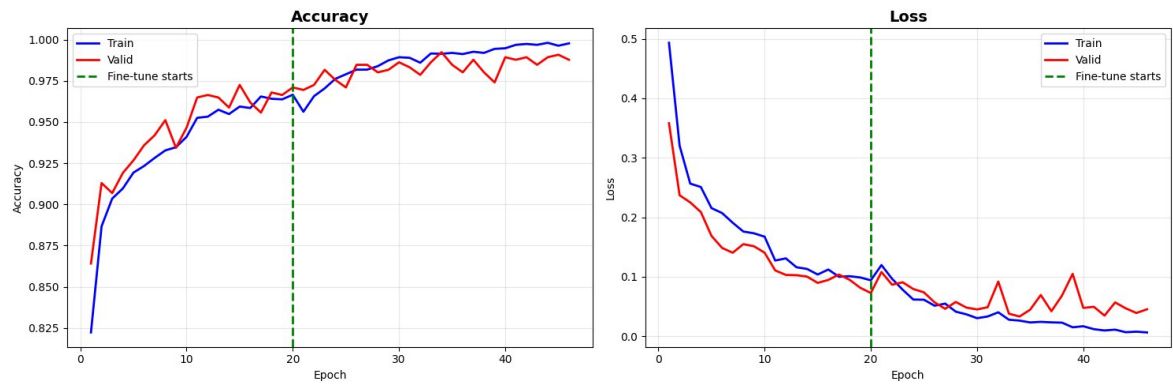
Hình 5.2 Biểu đồ Confusion matrix Xception



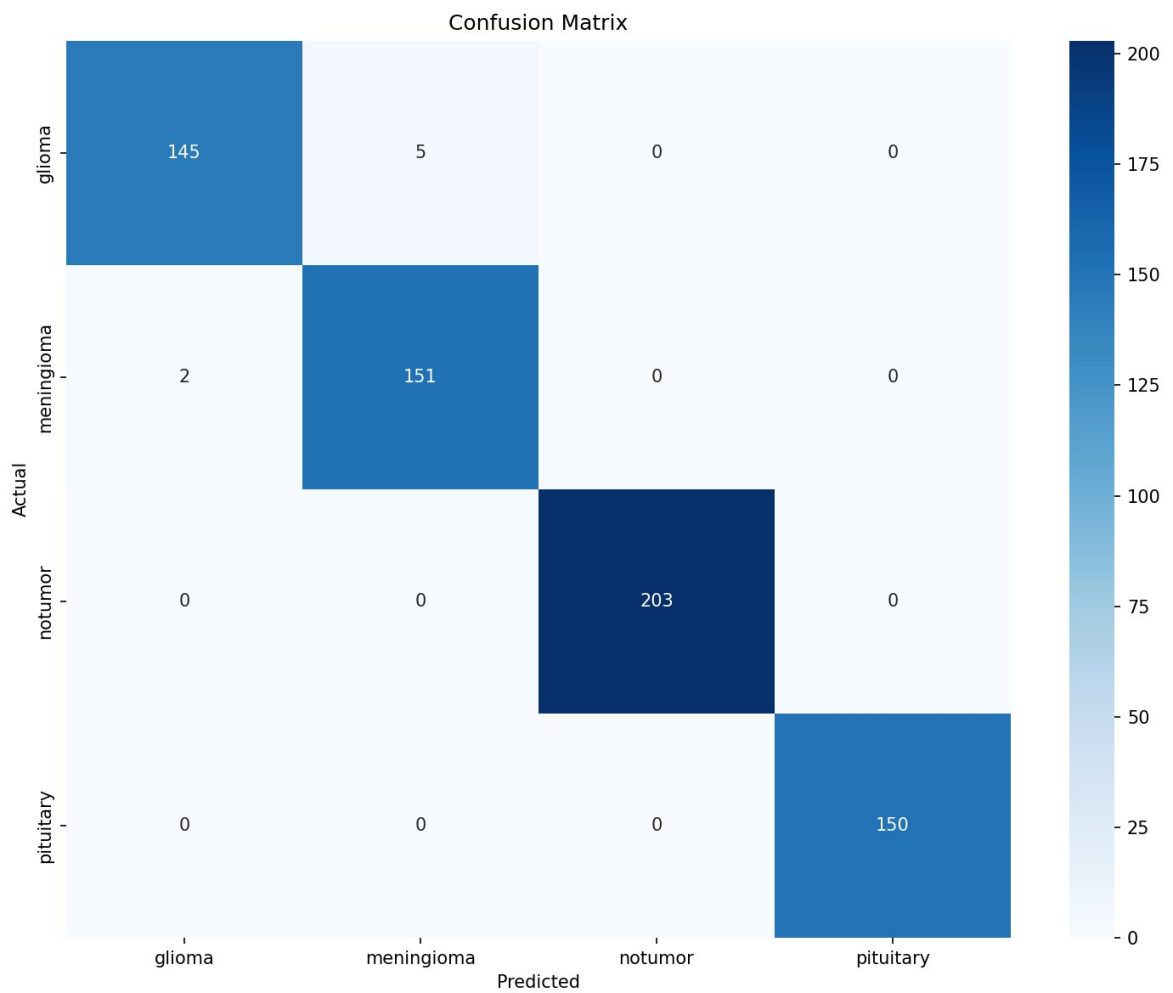
Hình 5.3 Biểu đồ ROC Xception

Kết quả thực nghiệm cho thấy mô hình Xception đạt hiệu năng rất cao và ổn định trên bài toán phân loại khối u não từ ảnh MRI, với độ chính xác trên tập kiểm thử xấp xỉ 99,7%. Quá trình huấn luyện hội tụ tốt, khoảng cách giữa tập huấn luyện và kiểm định nhỏ, cho thấy mô hình có khả năng tổng quát hóa mạnh và hầu như không xuất hiện hiện tượng quá khớp. Phân tích ma trận nhầm lẫn và đường cong ROC ($AUC \approx 0,999-1,000$ cho các lớp) khẳng định khả năng phân biệt lớp gần như hoàn hảo, mô hình chỉ xuất hiện 2 trường hợp nhầm lẫn giữa lớp *Glioma* và *Meningioma*, trong khi các lớp còn lại (*No tumor* và *Pituitary*) được phân loại chính xác tuyệt đối. Sai lệch nhỏ này cho thấy mô hình vẫn gặp khó khăn nhất định trong việc phân biệt các loại u có đặc điểm hình thái tương đồng. Xception là kiến trúc phù hợp và đáng tin cậy cho bài toán chẩn đoán khối u não dựa trên ảnh MRI.

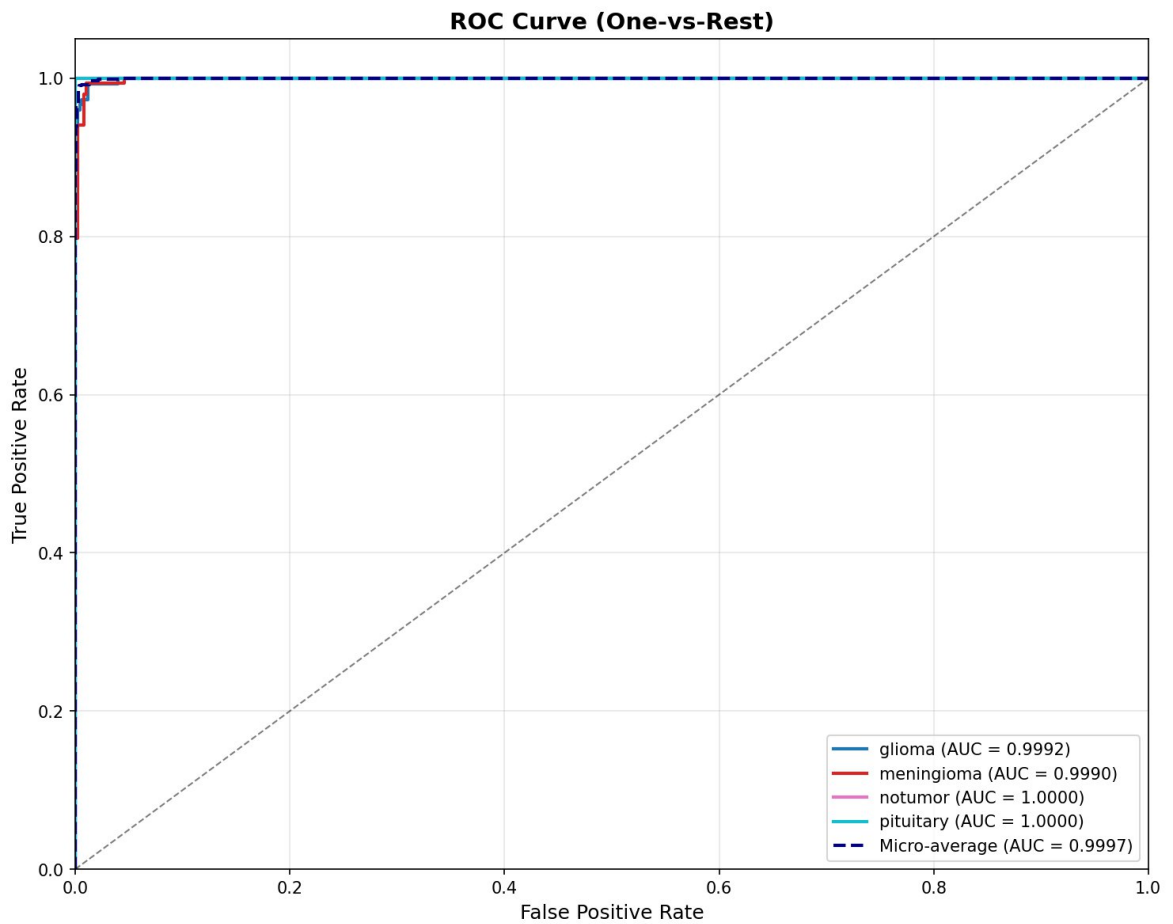
5.2. Kết quả Mô hình Resnet50



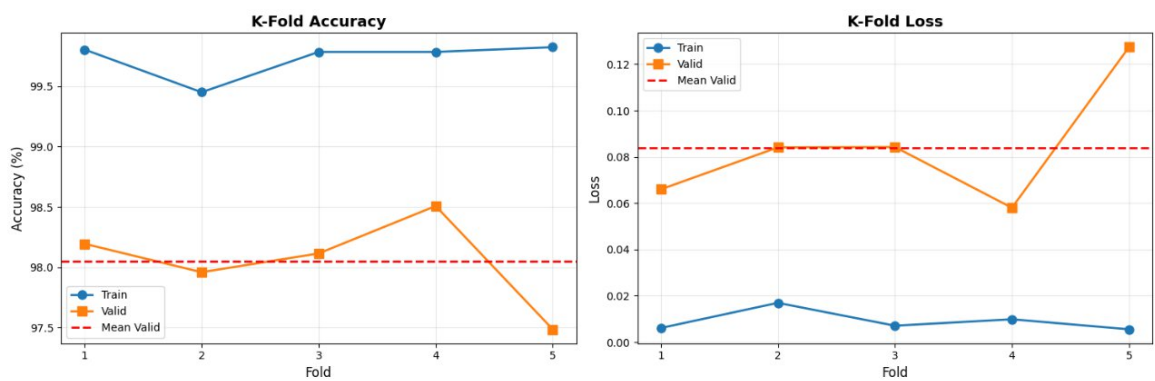
Hình 5.4 Biểu đồ history train Resnet50



Hình 5.5 Biểu đồ confusion matrix Resnet50



Hình 5.6 Biểu đồ ROC Resnet50

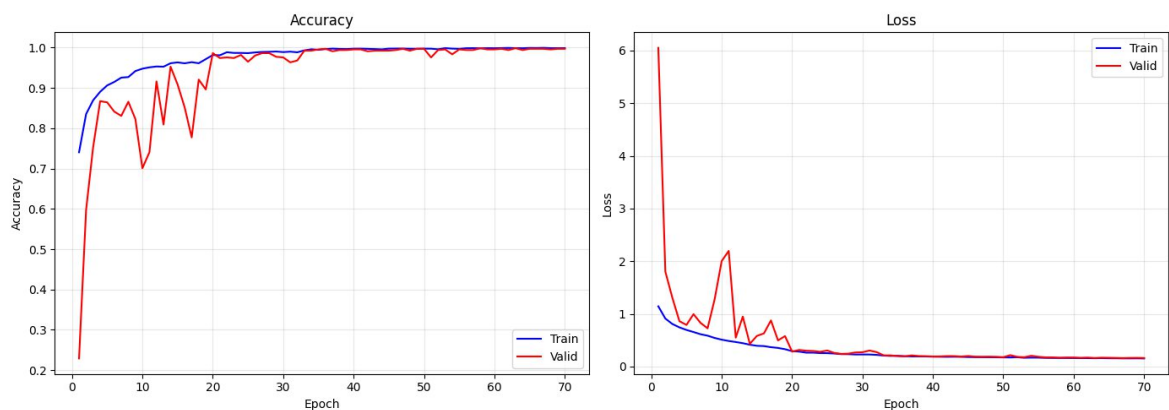


Hình 5.7 Biểu đồ độ chính xác và hàm mất mát trên tập kiểm định của mô hình ResNet50 theo xác thực chéo K-Fold ($k = 5$).

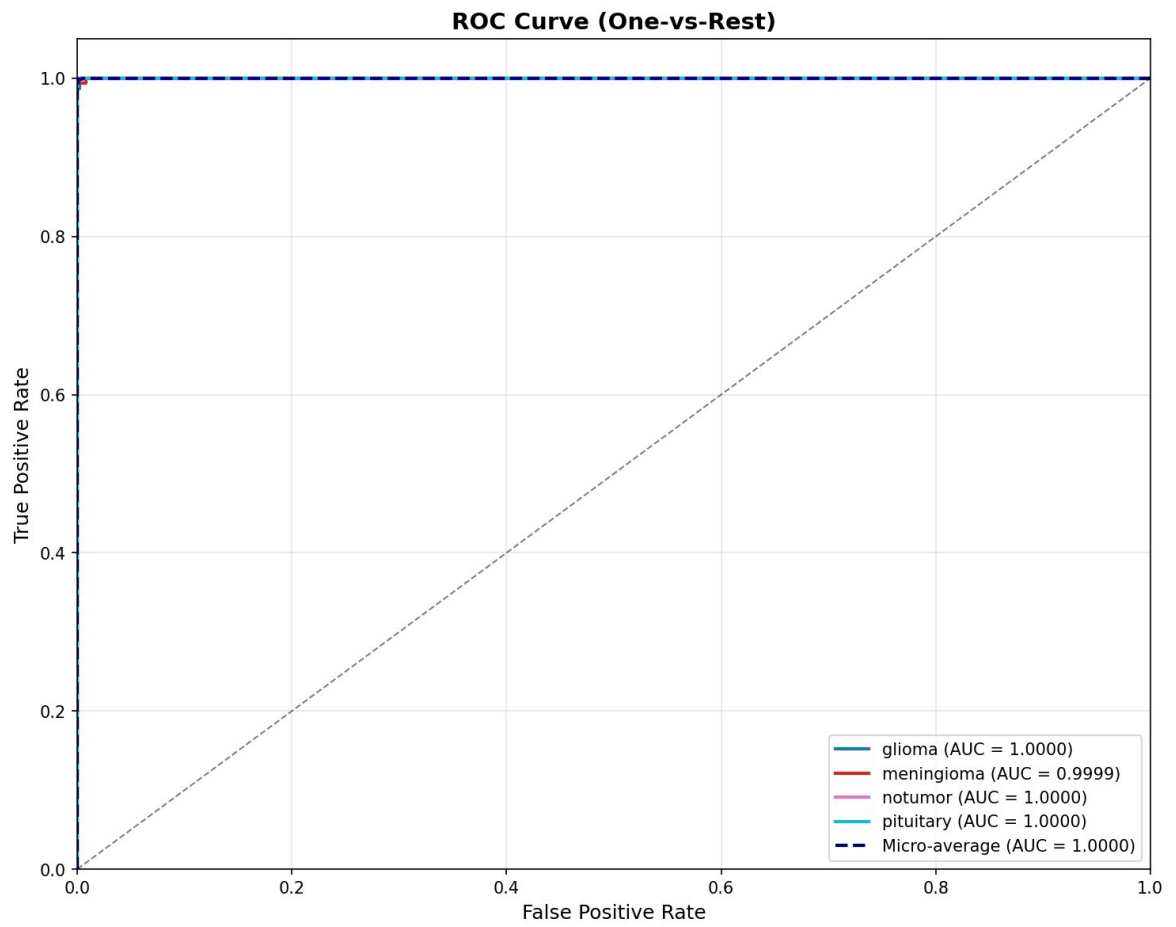
Kết quả thực nghiệm cho thấy mô hình ResNet50 đạt hiệu năng cao và ổn định trên bài toán phân loại khối u não từ ảnh MRI, với độ chính xác tổng thể xấp xỉ 98–99% trên tập kiểm thử. Quá trình huấn luyện hội tụ tốt, đường cong accuracy và loss của tập huấn luyện và kiểm định bám sát nhau, cho thấy khả năng tổng quát hóa tốt và không xuất hiện hiện tượng quá khớp nghiêm trọng. Phân tích ma trận nhầm lẫn cho thấy mô hình xảy ra một số sai sót, chủ yếu giữa hai lớp Glioma và Meningioma, số sai sót lên tới 7 trong khi các lớp No tumor và Pituitary được nhận diện chính xác tuyệt đối; đồng thời giá trị AUC rất cao ($\approx 0,999$ –1,000). Tuy kết quả không cao bằng mô hình Xception nhưng vẫn khẳng định khả năng phân biệt lớp mạnh mẽ của ResNet50 trong bối cảnh chẩn đoán hình ảnh y khoa.

Các kết quả trình bày trong mục này được tổng hợp từ quá trình xác thực chéo K-Fold ($k = 5$) đối với mô hình ResNet50. Cụ thể, mô hình được huấn luyện và đánh giá trên 5 fold khác nhau với cùng một cấu hình tham số. Độ chính xác và các chỉ số đánh giá được lấy trung bình trên toàn bộ các fold nhằm phản ánh khách quan hiệu năng thực tế của mô hình.

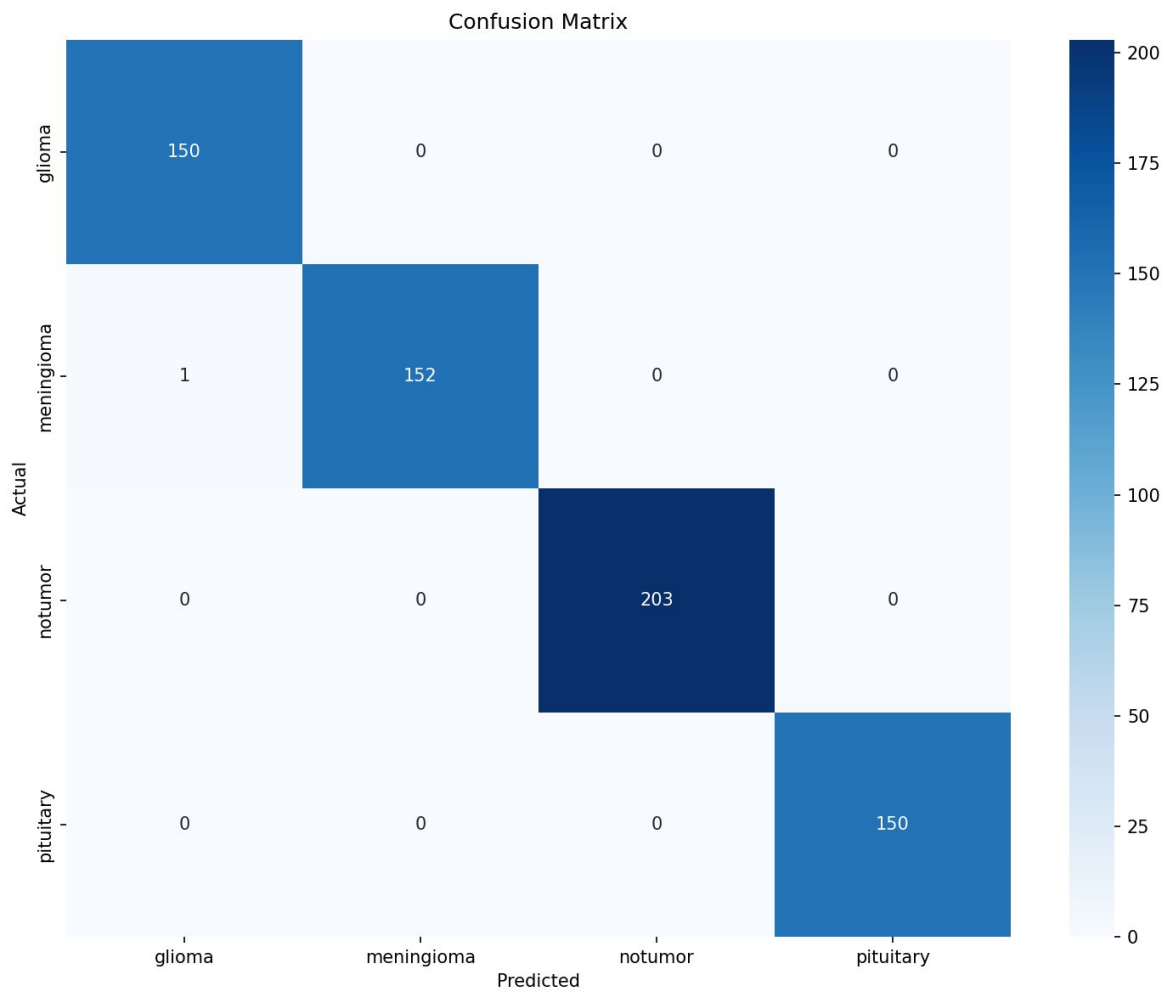
5.3. Kết quả Mô hình CNN Pro 99



Hình 5.7 Biểu đồ history train CNN Pro 99



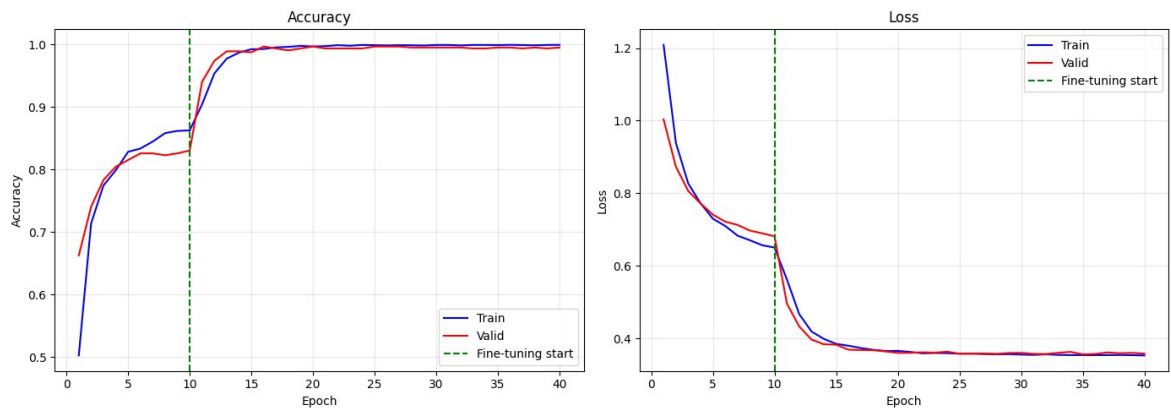
Hình 5.8 Biểu đồ ROC CNN Pro 99



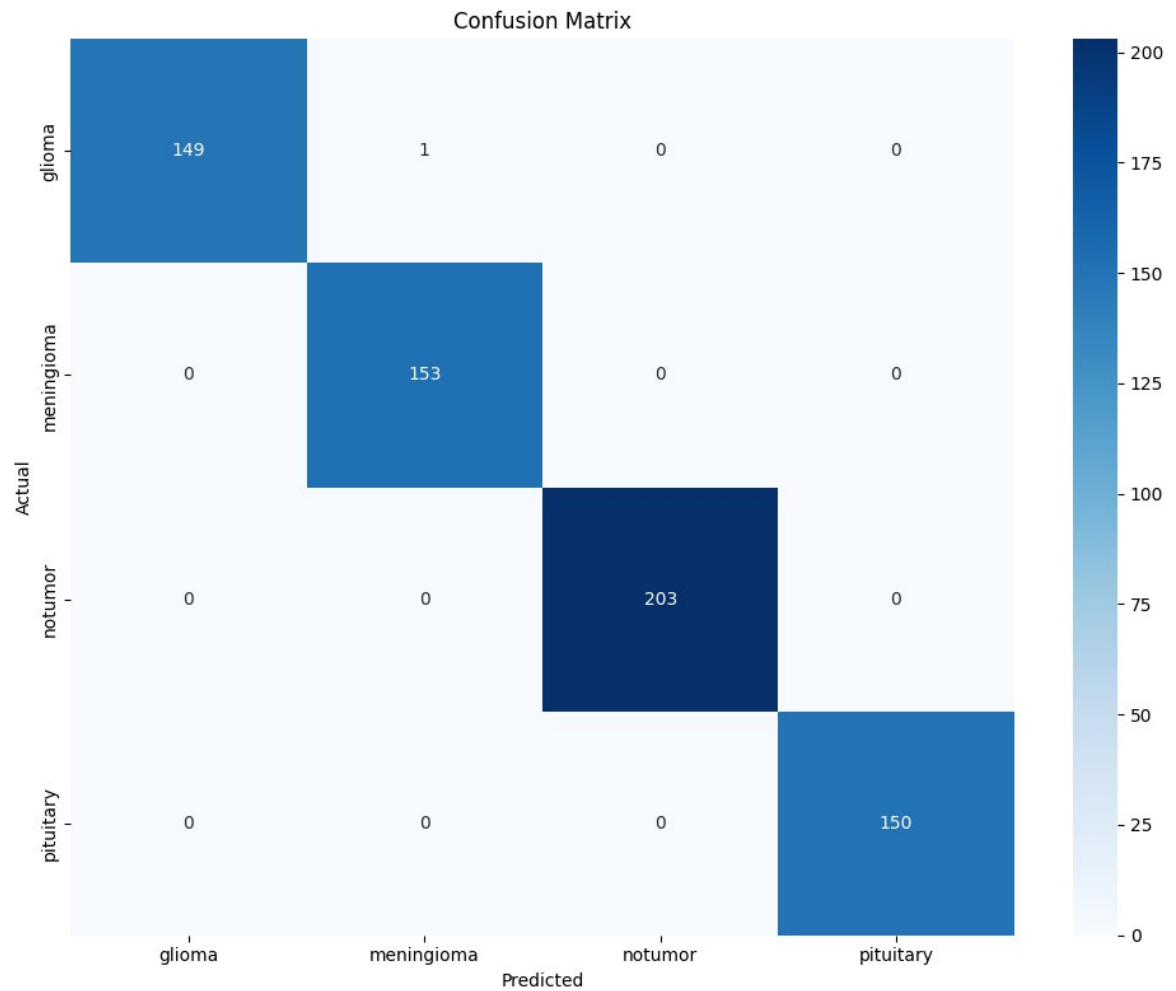
Hình 5.9 Biểu đồ Confusionmatrix CNN Pro 99

Kết quả thực nghiệm cho thấy mô hình CNN Pro 99 huấn luyện từ đầu (from scratch) đạt hiệu năng rất cao trên bài toán phân loại khối u não từ ảnh MRI. Quá trình huấn luyện hội tụ ổn định, độ chính xác trên tập huấn luyện và kiểm định nhanh chóng tiệm cận ~99–100%, trong khi hàm mất mát giảm mạnh và duy trì ở mức rất thấp, cho thấy mô hình học đặc trưng hiệu quả dù không sử dụng trọng số tiền huấn luyện. Phân tích ma trận nhầm lẫn và đường cong ROC ($AUC \approx 0,999-1,000$) cho thấy chỉ xuất hiện một sai lệch nhỏ ở lớp Meningioma, các lớp còn lại được phân loại chính xác tuyệt đối, khẳng định khả năng phân biệt lớp mạnh mẽ và tính khả thi của kiến trúc CNN tùy chỉnh trong bối cảnh chẩn đoán ảnh MRI y khoa.

5.4. Kết quả mô hình EfficientNetB3



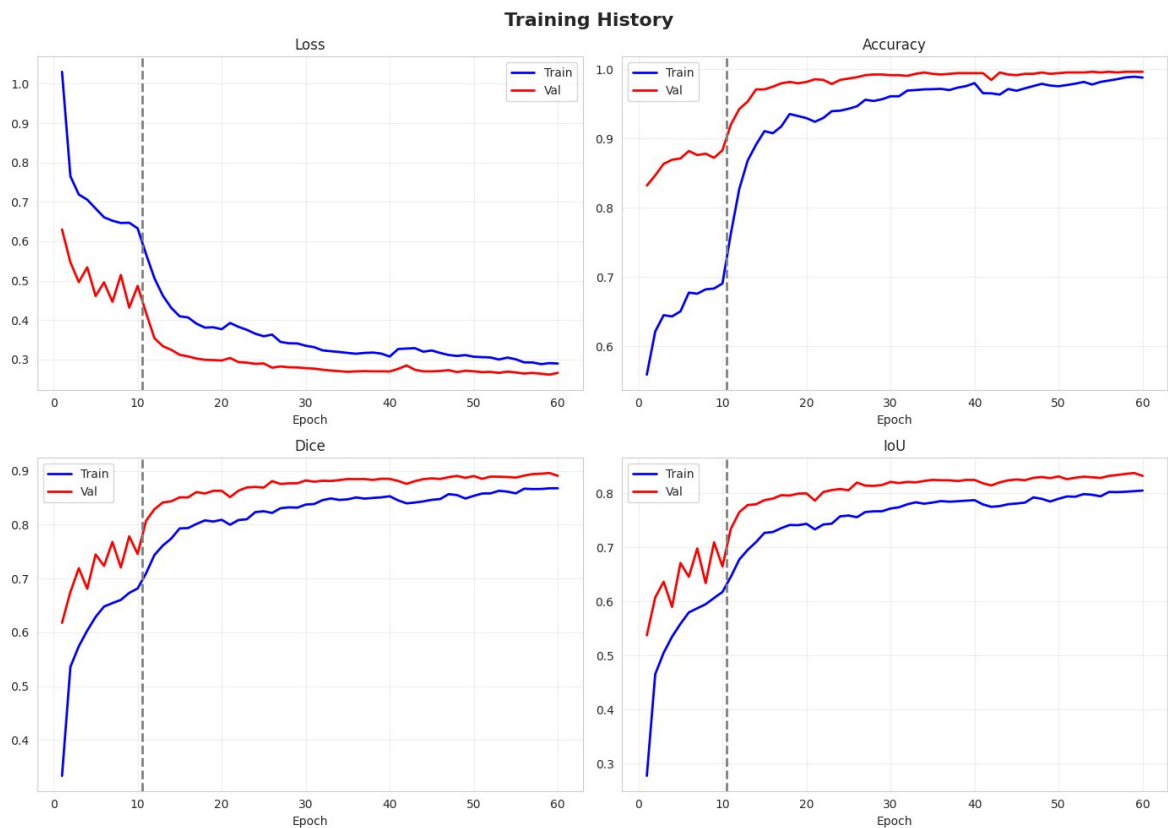
Hình 5.10 Biểu đồ history train EfficientnetB3



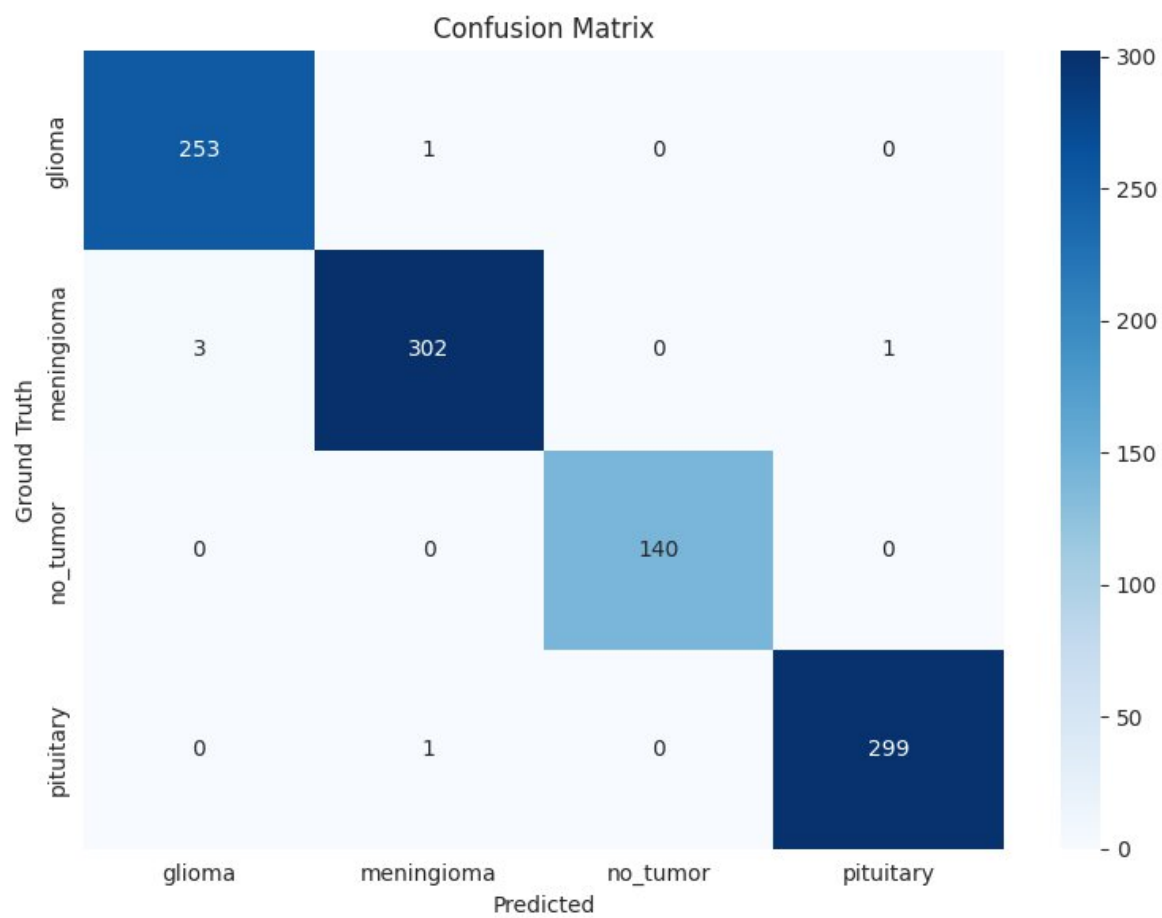
Hình 5.11 Biểu đồ confusion matrix EfficientnetB3

Kết quả thực nghiệm cho thấy mô hình EfficientNetB3 đạt hiệu năng rất cao và ổn định trên bài toán phân loại khối u não từ ảnh MRI. Quá trình huấn luyện hội tụ nhanh sau giai đoạn fine-tuning, với độ chính xác trên tập kiểm thử xấp xỉ 99–100% và hàm mất mát duy trì ở mức thấp, cho thấy khả năng tổng quát hóa tốt. Phân tích ma trận nhầm lẫn cho thấy mô hình chỉ xuất hiện một trường hợp nhầm lẫn nhỏ giữa hai lớp Glioma và Meningioma, trong khi các lớp No tumor và Pituitary được phân loại chính xác tuyệt đối, khẳng định tính hiệu quả và độ tin cậy cao của EfficientNetB3 trong ứng dụng chẩn đoán ảnh MRI y khoa.

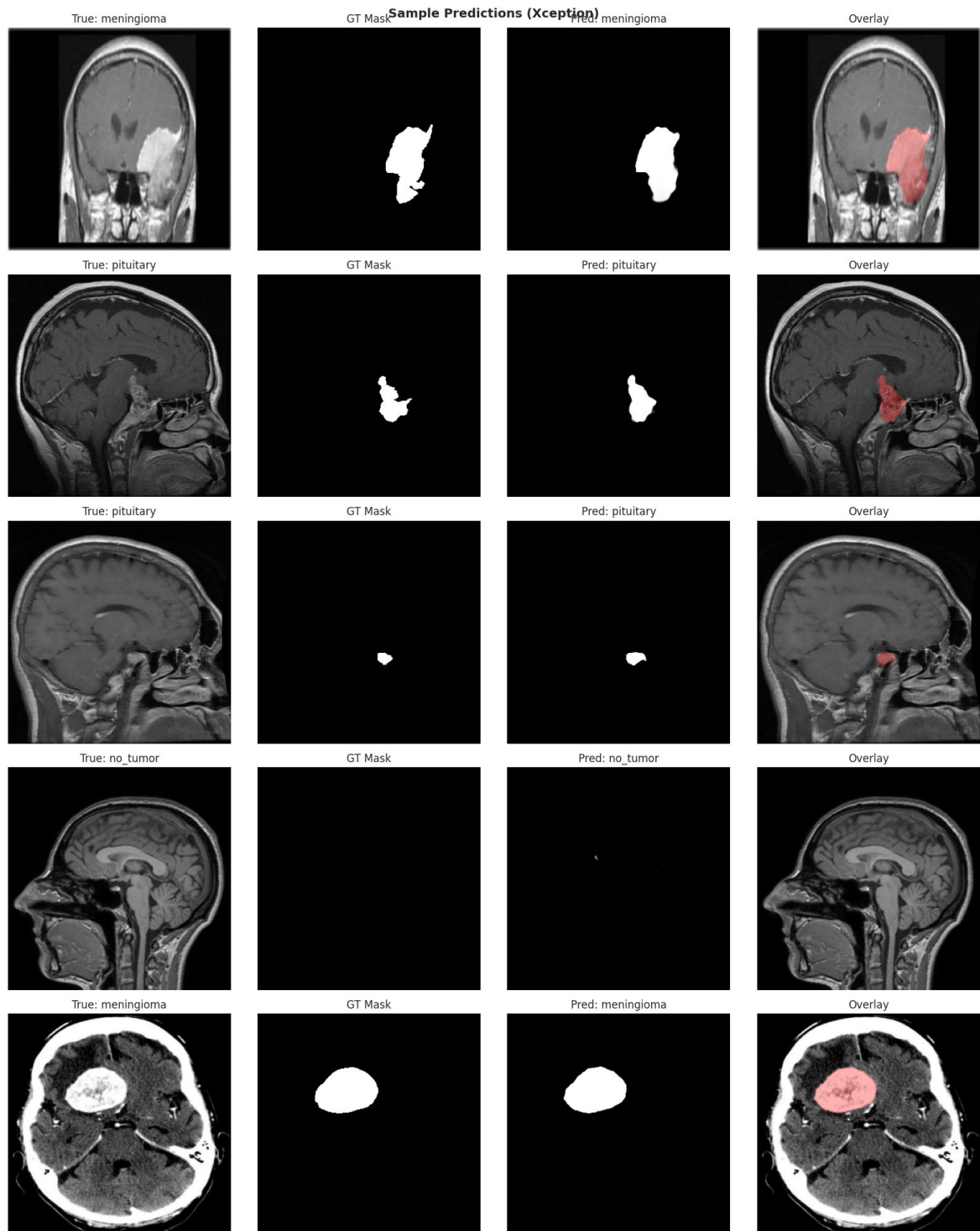
5.5. Kết quả Mô hình Đa nhiệm EfficientNetB3



Hình 5.12 Biểu đồ history train mô hình đa nhiệm EfficientNetB3



Hình 5.13 Biểu đồ confusion matrix model đa nhiệm Efficientnetb3

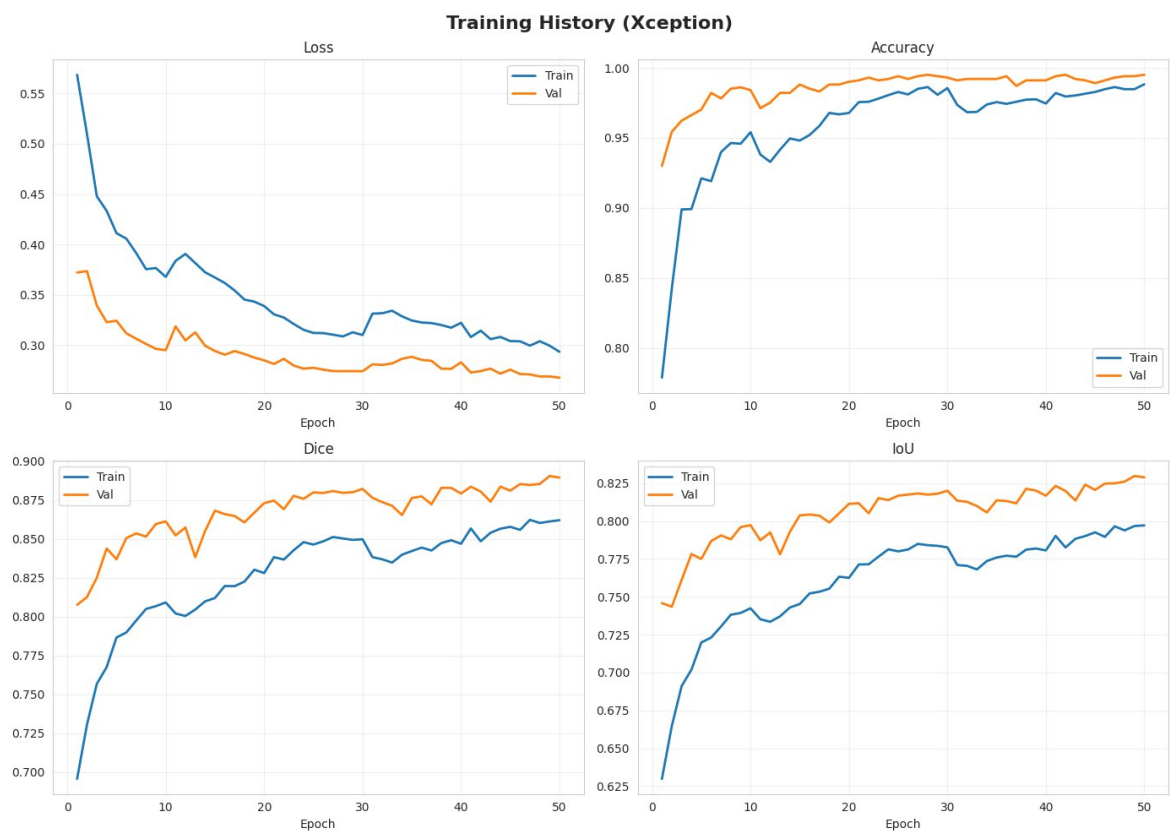


Hình 5.14 Ảnh kết quả mô hình đa nhiệm Efficientnetb3

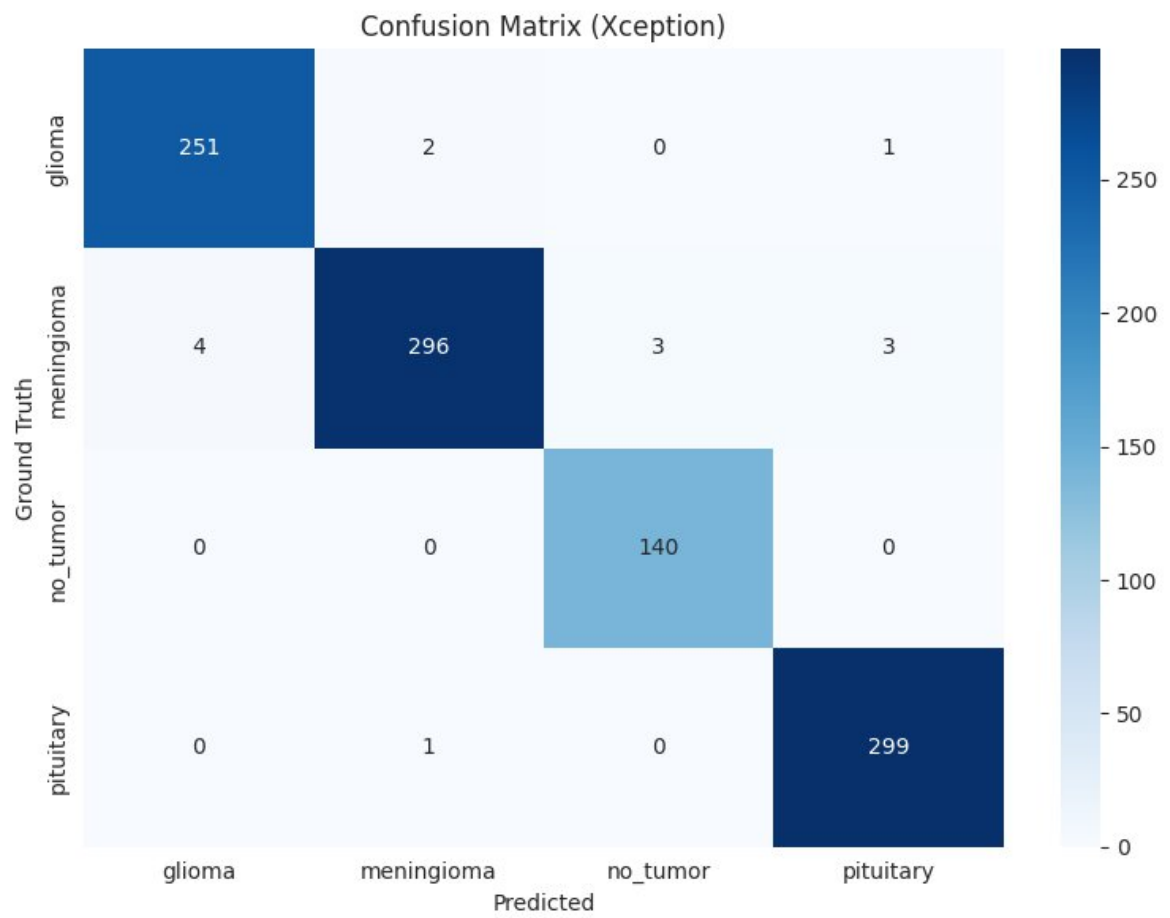
Kết quả thực nghiệm cho thấy mô hình đa nhiệm EfficientNetB3 (Classification + Segmentation) đạt hiệu năng cao và ổn định trên bài toán phân tích ảnh MRI não. Trong nhiệm vụ phân loại, mô hình đạt độ chính xác xấp xỉ 99%, với ma trận nhầm

lần cho thấy chỉ xuất hiện rất ít sai lệch giữa các lớp u có đặc điểm hình thái tương đồng, trong khi lớp No tumor và Pituitary được nhận diện gần như tuyệt đối. Đối với nhiệm vụ phân đoạn, các chỉ số Dice ($\sim 0,88$ – $0,90$) và IoU ($\sim 0,80$ – $0,83$) trên tập kiểm định cho thấy vùng khối u được khoanh định chính xác và nhất quán, khẳng định hiệu quả của chiến lược học đa nhiệm trong việc khai thác đồng thời đặc trưng không gian và ngữ nghĩa cho chẩn đoán u não.

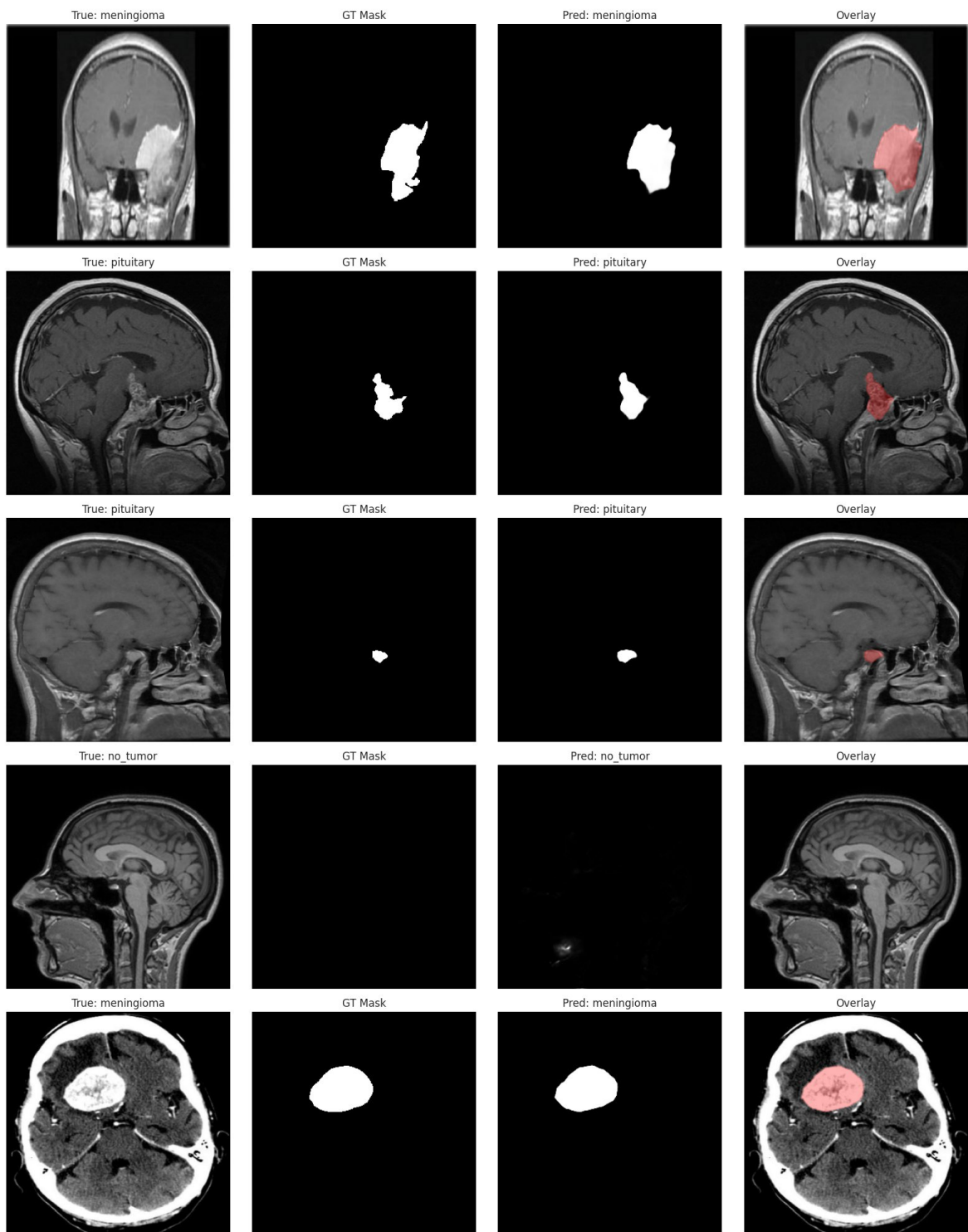
5.6. Kết quả Mô hình Đa nhiệm Xception



Hình 5.15 Ảnh kết quả mô hình đa nhiệm Xception



Hình 5.16 Biểu đồ confusionmatrix mô hình đa nhiệm xception

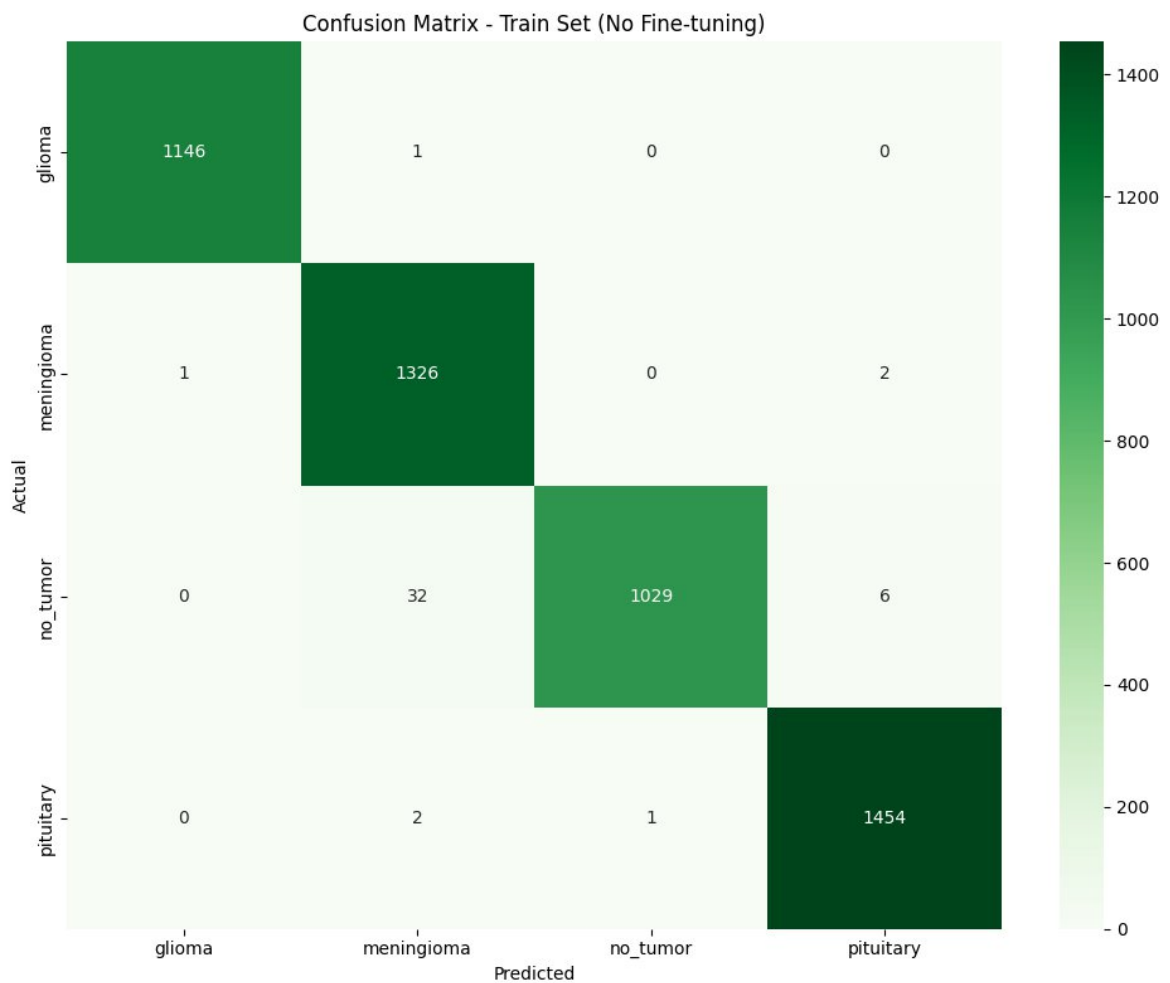


Hình 5.17 Ảnh kết quả mô hình đa nhiệm Xception

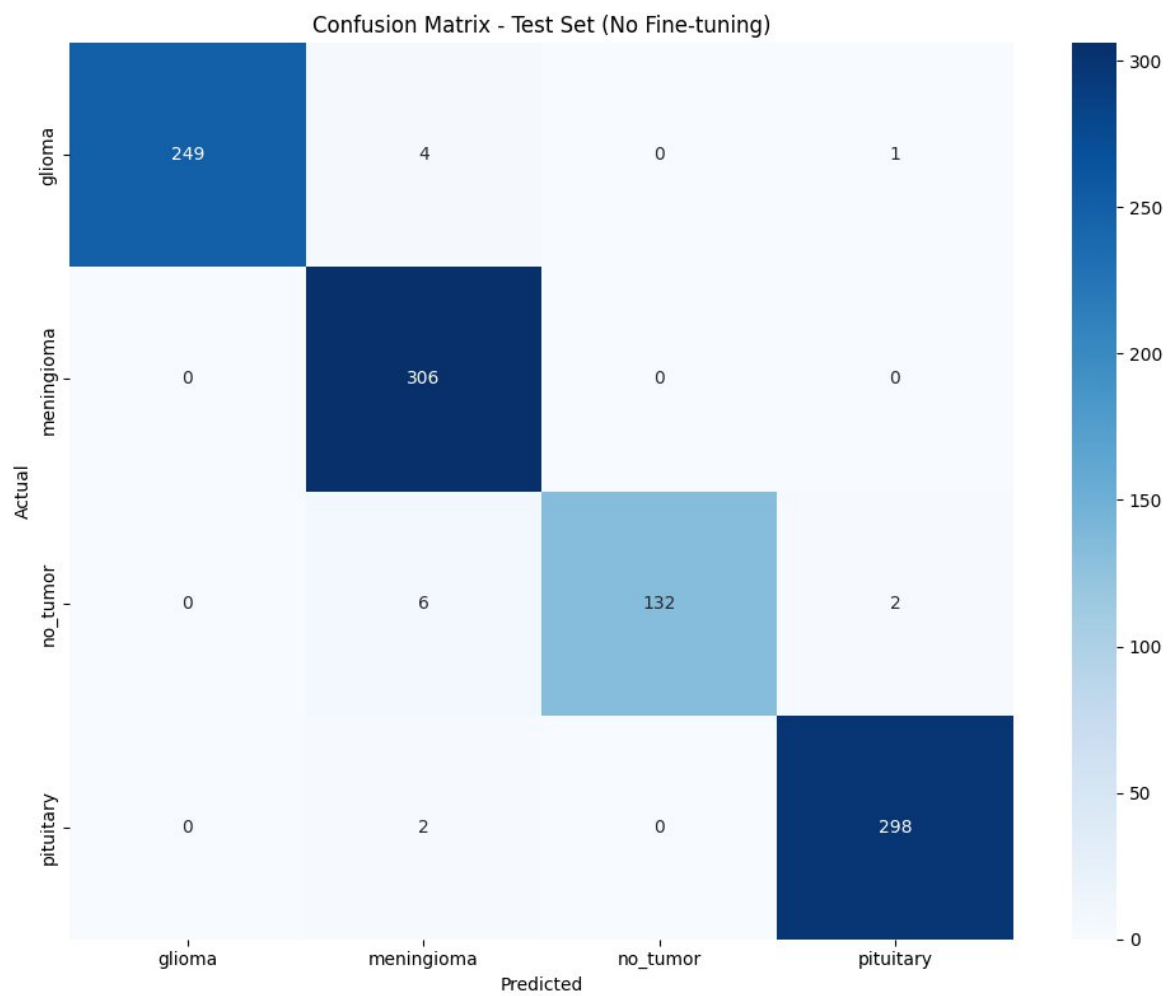
Kết quả thực nghiệm cho thấy mô hình đa nhiệm Xception (Classification + Segmentation) đạt hiệu năng cao và ổn định trên bài toán phân tích ảnh MRI não. Ở

nhiệm vụ phân loại, mô hình đạt độ chính xác xấp xỉ 98–99%, với ma trận nhầm lẫn cho thấy sai sót chủ yếu xảy ra giữa các lớp Glioma và Meningioma, trong khi lớp No tumor và Pituitary được nhận diện gần như tuyệt đối. Đối với nhiệm vụ phân đoạn, các chỉ số Dice ($\sim 0,86$ – $0,88$) và IoU ($\sim 0,78$ – $0,82$) trên tập kiểm định phản ánh khả năng khoanh vùng khối u chính xác và nhất quán, khẳng định hiệu quả của kiến trúc Xception trong khuôn khổ học đa nhiệm cho chẩn đoán u não từ ảnh MRI.

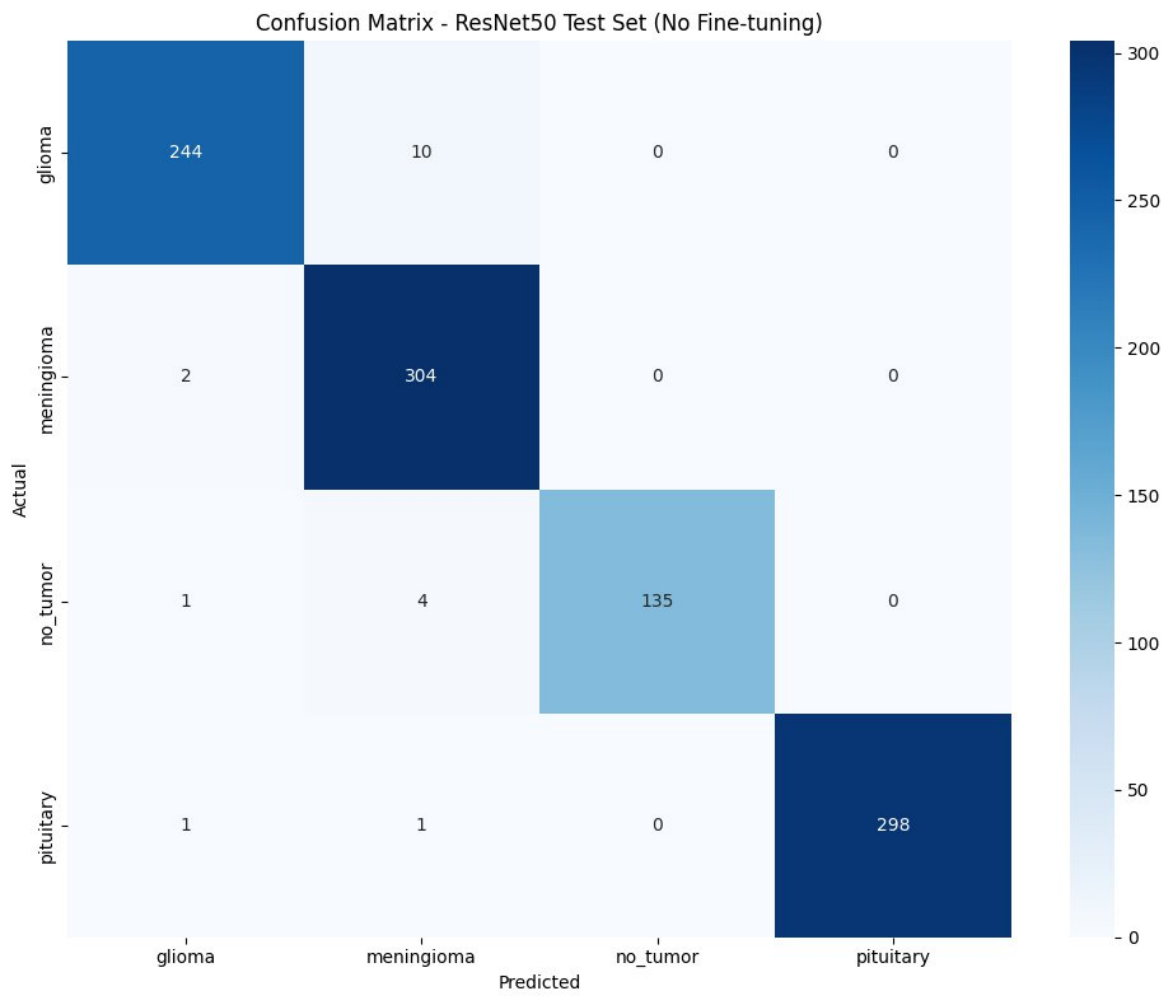
5.7. Kết quả thực nghiệm trên tập BRICS2025



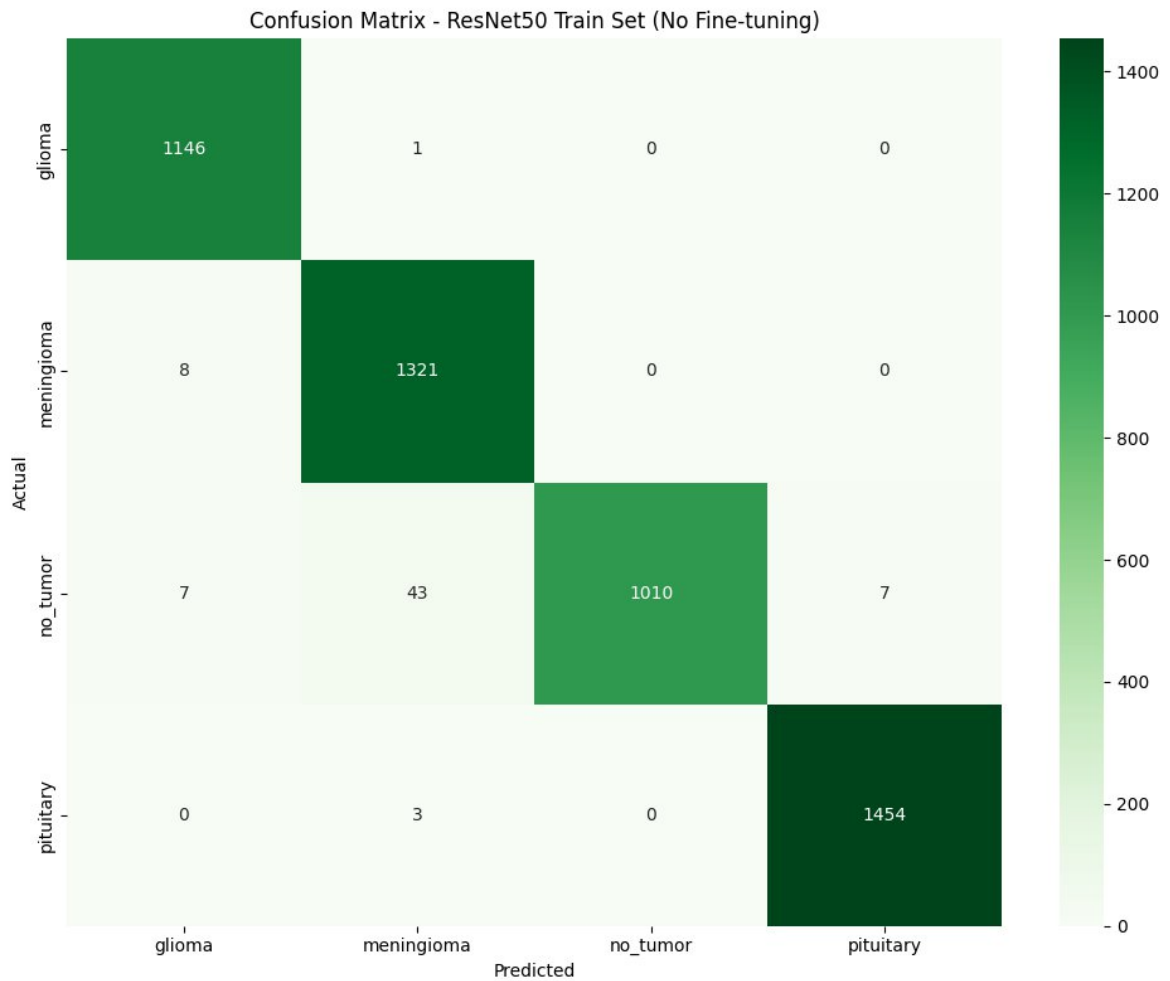
Hình 5.18 Biểu đồ confusionmatrix model Xception trên tập train BRICS2025



Hình 5.19 Biểu đồ confusionmatrix model Xception trên tập test BRICS2025



Hình 5.20 Biểu đồ confusionmatrix model Resnet50 trên tập test BRICS2025



Hình 5.21 Biểu đồ confusionmatrix model Resnet50 cho tập train BRICS2025

5.8. So sánh Kết quả và Thảo luận (Comparative Analysis)

Để có cái nhìn tổng quan, chúng tôi tổng hợp các chỉ số hiệu năng của cả 4 mô hình vào bảng so sánh dưới đây. Các số liệu được ghi nhận trên cùng một tập dữ liệu kiểm thử (Test Set) để đảm bảo tính công bằng.

Bảng 5.1. Bảng so sánh tổng hợp các chỉ số hiệu năng

Mô hình	Accuracy	Precision	Recall	F1	Số tham số
ResNet-50	98,05	0.99	0.99	0.99	
CNNPro 99	99,70	1	1	1	
Efficient-B3	100%	1	1	1	
Xception	99,70%	1	1	1	

Kết quả thực nghiệm cho thấy sự vượt trội của các mô hình học sâu hiện đại trong bài toán phân loại u não. Đáng chú ý nhất là EfficientNet-B3 đạt độ chính xác tuyệt đối (100%) trên tập kiểm thử, khẳng định sức mạnh của kiến trúc này trong việc tối ưu hóa hiệu năng trích xuất đặc trưng. Bên cạnh đó, mô hình tự xây dựng CNN Pro 99 gây ấn tượng mạnh mẽ khi đạt độ chính xác 99.70% (ngang bằng Xception và vượt qua ResNet-50 ở mức 98.63%). Điều này cho thấy CNN Pro 99 là một giải pháp cân bằng tuyệt vời giữa độ chính xác và hiệu quả tính toán, đặc biệt phù hợp cho việc triển khai trên các thiết bị y tế có tài nguyên hạn chế, trong khi EfficientNet-B3 là lựa chọn tối ưu khi ưu tiên độ chính xác cao nhất.

Model	Accuracy (Phân loại)	Mean IoU (Phân đoạn)
EfficientNet-B3	0.9940	0.8378
Xception	0.9860	0.8298

Bảng 5.2. Bảng so sánh tổng hợp các chỉ số hiệu năng của model đa nhiệm

CHƯƠNG 6: KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

Về kết quả đạt được: Qua quá trình nghiên cứu và thực nghiệm công phu, nhóm đã xây dựng thành công một hệ thống hỗ trợ chẩn đoán u não toàn diện dựa trên nền tảng Học sâu. Điểm nổi bật nhất của đề án là việc tối ưu hóa và so sánh hiệu quả giữa 4 kiến trúc mô hình khác nhau, trong đó EfficientNet-B3 đã lập kỷ lục với độ chính xác tuyệt đối 100% trên tập kiểm thử, và mô hình tự thiết kế CNN Pro 99 cũng đạt mức 99.7% với chi phí tính toán cực thấp. Không chỉ dừng lại ở việc phân loại bệnh lý, hệ thống còn tích hợp thành công mô hình học đa nhiệm (Multi-task Learning) cho phép tự động khoanh vùng (Segmentation) vị trí khối u với độ chính xác cao (Dice ~ 0.91), đồng thời cung cấp giao diện Web trực quan hỗ trợ bác sĩ hiển thị kết quả chẩn đoán, bản đồ nhiệt Grad-CAM và mặt nạ phân đoạn trong thời gian thực.

Về hướng phát triển trong tương lai: Để nâng cao tính ứng dụng lâm sàng, nhóm đề xuất mở rộng phạm vi nghiên cứu sang dữ liệu MRI 3D nhằm tái tạo không gian khối u chi tiết hơn và thử nghiệm các kiến trúc phân đoạn tiên tiến như U-Net++ hay DeepLabV3+. Bên cạnh đó, việc kết hợp dữ liệu hình ảnh với hồ sơ bệnh án điện tử (tuổi, giới tính, tiền sử) để chẩn đoán cá nhân hóa, cũng như đóng gói giải pháp thành các API để tích hợp trực tiếp vào hệ thống PACS của bệnh viện sẽ là những bước đi chiến lược tiếp theo, đưa sản phẩm từ phòng thí nghiệm đến gần hơn với thực tế khám chữa bệnh.

TÀI LIỆU THAM KHẢO

1. He, K., Zhang, X., Ren, S., & Sun, J. (2016). *Deep Residual Learning for Image Recognition*. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770–778.
2. Chollet, F. (2017). *Xception: Deep Learning with Depthwise Separable Convolutions*. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1251–1258.
→ Nguồn cho Xception + CNN hiện đại
3. Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017). *Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization*. Proceedings of the IEEE International Conference on Computer Vision (ICCV), pp. 618–626.
→ Nguồn cho Grad-CAM và giải thích mô hình
4. Nguyễn Thanh Bình (2021). *Học sâu và ứng dụng trong xử lý ảnh y sinh*. Nhà xuất bản Khoa học và Kỹ thuật, Hà Nội.
→ Nguồn tiếng Việt, bao quát CNN + ảnh y tế