**NHẬN XÉT CỦA GIẢNG VIÊN HƯỚNG DẪN**

*Trà Vinh, ngày ….. tháng …… năm ……*

**Giáo viên hướng dẫn**

*(Ký tên và ghi rõ họ tên)*

*Trà Vinh, ngày ….. tháng …… năm ……*

**Giáo viên hướng dẫn**

*(Ký tên và ghi rõ họ tên)*

**LỜI CẢM ƠN**

**NHẬN XÉT CỦA THÀNH VIÊN HỘI ĐỒNG**

*Trà Vinh, ngày ….. tháng …… năm ……*

**Thành viên hội đồng**

*(Ký tên và ghi rõ họ tên)*

Em xin gửi lời cảm ơn chân thành và sâu sắc đến thầy Nguyễn Mộng Hiền, người đã luôn tận tình chỉ bảo, hướng dẫn và hỗ trợ em trong suốt quá trình thực hiện bài báo cáo này. Sự kiên nhẫn và tâm huyết của thầy đã giúp em rất nhiều trong việc nghiên cứu và hoàn thiện bài viết.

Em cũng xin cảm ơn các thầy cô trong khoa đã dành thời gian quý báu để chia sẻ kiến thức, góp ý và động viên em. Sự giúp đỡ nhiệt tình của thầy cô đã tạo điều kiện thuận lợi để em hoàn thành công việc một cách tốt nhất.

Đặc biệt, em muốn gửi lời biết ơn sâu sắc đến thầy Nguyễn Mộng Hiền, không chỉ là người hướng dẫn mà còn là người đồng hành, luôn sát cánh bên em trong suốt hành trình này. Những kinh nghiệm quý báu mà thầy truyền đạt đã giúp em vượt qua nhiều khó khăn và nâng cao kỹ năng nghiên cứu của bản thân.

Dù em đã nỗ lực hết sức, nhưng bài báo cáo này vẫn không tránh khỏi những thiếu sót do thời gian và kiến thức có hạn. Em rất mong nhận được những góp ý quý báu từ thầy cô để có thể tiếp tục hoàn thiện và phát triển bản thân trong tương lai.

Em xin chân thành cảm ơn!

**MỤC LỤC**

[CHƯƠNG 1: TỔNG QUAN 10](#_Toc187186128)

[CHƯƠNG 2: NGHIÊN CỨU LÝ THUYẾT 13](#_Toc187186129)

[2.1 Cơ sở lý thuyết 13](#_Toc187186130)

[2.1.1 Phân đoạn ảnh y khoa (trong chương này phải có trích dẫn) 13](#_Toc187186131)

[2.1.2 Mạng Nơ-ron tích chập (CNN) 17](#_Toc187186132)

[2.2 Các khái niệm chính trong CNN 23](#_Toc187186133)

[2.2.1 Stride 23](#_Toc187186134)

[2.2.2 Padding 24](#_Toc187186135)

[2.2.3 Filters (Kernels) 25](#_Toc187186136)

[2.2.4 Feature Maps 25](#_Toc187186137)

[2.3 Ứng dụng thực tế của CNN 26](#_Toc187186138)

[2.3.1 Image Recognition & Classification 26](#_Toc187186139)

[2.3.2 Optical Character Recognition – OCR 26](#_Toc187186140)

[2.3.3 Image Segmentation 26](#_Toc187186141)

[2.3.4 Autonomous Driving 27](#_Toc187186142)

[2.3.5 Action Recognition & Video Analysis 27](#_Toc187186143)

[2.4 Ngôn ngữ và công cụ phát triển 27](#_Toc187186144)

[2.4.1 Ngôn ngữ hỗ trợ 27](#_Toc187186145)

[2.4.2 Môi trường và công cụ phát triển 28](#_Toc187186146)

[CHƯƠNG 3: HIỆN THỰC HÓA NGHIÊN CỨU 30](#_Toc187186147)

[3.1 Mô tả bài toán 30](#_Toc187186148)

[3.2 Sơ đồ khối 31](#_Toc187186149)

[3.3 Xử lý dữ liệu 32](#_Toc187186150)

[3.3.1 Tải dữ liệu và xử lý hình ảnh 32](#_Toc187186151)

[3.3.2 Tiền xử lý dữ liệu 32](#_Toc187186152)

[3.3.3 Chia dữ liệu thành tập huấn luyện và kiểm thử 33](#_Toc187186153)

[3.4 Xây dựng mô hình CNN 34](#_Toc187186154)

[3.5 Huấn luyện mô hình CNN 35](#_Toc187186155)

[3.6 Đánh giá trên tập kiểm tra 35](#_Toc187186156)

[3.7 Hiển thị kết quả phân đoạn và dự đoán 36](#_Toc187186157)

[CHƯƠNG 4: KẾT QUẢ NGHIÊN CỨU 37](#_Toc187186158)

[4.1 Môi trường cài đặt 37](#_Toc187186159)

[4.2 Tập dữ liệu thực hiện 37](#_Toc187186160)

[4.3 Kết quả huấn luyện 38](#_Toc187186161)

[4.3.1 Quá trình xây dựng mô hình CNN 38](#_Toc187186162)

[4.3.2 Quá trình huấn luyện mô hình CNN 39](#_Toc187186163)

[4.3.3 Phân đoạn và hiển thị kết quả dự đoán 44](#_Toc187186164)

[CHƯƠNG 5: KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN 46](#_Toc187186165)

[5.1 Kết luận 46](#_Toc187186166)

[5.2 Hướng phát triển 46](#_Toc187186167)

[DANH MỤC TÀI LIỆU THAM KHẢO 47](#_Toc187186168)

**DANH MỤC HÌNH ẢNH**

[Hình 1: *So sánh giữa phân đoạn thủ công và phân đoạn tự động trên ảnh MRI [1]* 14](#_Toc187186197)

[Hình 2: *Kiến trúc cơ bản của Mạng Nơ-ron Tích chập (CNN) [2]* 18](#_Toc187186198)

[Hình 3: *Phân đoạn khối u não trên ảnh MRI bằng mô hình CNN [3]* 19](#_Toc187186199)

[Hình 4: *Phân đoạn mô mềm trên ảnh CT bằng CNN [4]* 20](#_Toc187186200)

[Hình 5: *Phân đoạn tổn thương phổi ảnh CT để chẩn đoán COVID-19 [5]* 21](#_Toc187186201)

[Hình 6: *Phân đoạn khối u não trên ảnh MRI [6]* 22](#_Toc187186202)

[Hình 7: *Convolution với các bước nhảy (Stride) [8]* 23](#_Toc187186203)

[Hình 8: *Minh họa áp dụng Padding trên ma trận 6x6 [8]* 24](#_Toc187186204)

[Hình 9: *Phép tích chập trên ma trận đầu vào với Padding và Kernel [8]* 25](#_Toc187186205)

[Hình 10: *Hình ảnh đầu vào (Input)* 30](#_Toc187186206)

[Hình 11: *Hình ảnh đầu ra (Output)* 30](#_Toc187186207)

[Hình 12: *Sơ đồ khối* 31](#_Toc187186208)

[Hình 13: *Cấu trúc mô hình CNN cho phân đoạn khối u* 38](#_Toc187186209)

[Hình 14: *Kết quả huấn luyện mô hình CNN* 39](#_Toc187186210)

[Hình 15: *Biểu đồ Accuracy thể hiện độ chính xác trong quá trình huấn luyện* 40](#_Toc187186211)

[Hình 16: *Biểu đồ thể hiện hàm mất mát (Loss) trong quá trình huấn luyện* 41](#_Toc187186212)

[Hình 17: *Biểu đồ thể hiện hệ số Dice trong quá trình huấn luyện mô hình* 42](#_Toc187186213)

[Hình 18: *Biểu đồ thể hiện chỉ số IoU trong quá trình huấn luyện mô hình* 43](#_Toc187186214)

[Hình 19: *Ảnh đầu vào (Input)* 44](#_Toc187186215)

[Hình 20: *Mặt nạ thực tế (Mask)* 44](#_Toc187186216)

[Hình 21: *Mặt nạ dự đoán (Mask)* 45](#_Toc187186217)

**TÓM TẮT ĐỒ ÁN CHUYÊN NGÀNH**

Đồ án chuyên ngành tập trung nghiên cứu bài toán phân đoạn khối u não từ ảnh cộng hưởng từ (MRI), một vấn đề quan trọng trong lĩnh vực y học hiện đại nhằm hỗ trợ bác sĩ chẩn đoán và điều trị hiệu quả. Với mục tiêu tự động hóa quy trình phân đoạn, đồ án ứng dụng mô hình học sâu dựa trên mô hình CNN, nổi bật với khả năng xử lý dữ liệu hình ảnh phức tạp và phân đoạn chính xác.

Quy trình nghiên cứu bao gồm các bước:

* Thu thập và tiền xử lý dữ liệu: Bộ dữ liệu BRATS Custom Dataset từ Kaggle, bao gồm một tập nhỏ của BRATS 2018 nhằm đơn giản hóa việc xử lý ban đầu. Bộ dữ liệu này chứa ảnh MRI của não và được sử dụng cho bài toán phân đoạn khối u. Các bước tiền xử lý bao gồm chuẩn hóa ảnh và loại bỏ nhiễu để cải thiện chất lượng đầu vào, đảm bảo phù hợp với mô hình CNN.
* Xây dựng mô hình CNN: Thiết kế và huấn luyện trên nền tảng Google Colab, tận dụng GPU để tăng tốc xử lý.
* Kiểm thử và đánh giá: Mô hình được kiểm tra trên tập dữ liệu chưa từng gặp để đánh giá hiệu suất thông qua các chỉ số quan trọng như Loss, Accuracy, Dice Coefficient, và IoU.

Các hướng phát triển trong tương lai bao gồm: tích hợp các cơ chế chú ý (attention mechanisms), thử nghiệm trên dữ liệu đa chiều (3D MRI), mở rộng quy mô dữ liệu, và đánh giá thực nghiệm tại các cơ sở y tế. Đồ án không chỉ góp phần khẳng định tiềm năng của công nghệ học sâu trong lĩnh vực y khoa mà còn mở ra triển vọng ứng dụng thực tế, hỗ trợ nâng cao chất lượng chẩn đoán và chăm sóc sức khỏe cho bệnh nhân.

**MỞ ĐẦU**

**1. Lý do chọn đề tài**

Trong y học hiện đại, việc phát hiện sớm và chẩn đoán chính xác các khối u là rất quan trọng để điều trị kịp thời và hiệu quả cho bệnh nhân. Phân đoạn khối u trong ảnh y khoa, đặc biệt là các ảnh CT và MRI, hỗ trợ đắc lực cho các bác sĩ trong việc xác định vị trí và kích thước của khối u. Tuy nhiên, quá trình phân đoạn thủ công thường tốn nhiều thời gian và có thể dẫn đến những sai lệch do chủ quan. Với sự phát triển của trí tuệ nhân tạo, đặc biệt là các mô hình Mạng Nơ-ron Tích chập (CNN), việc tự động hóa phân đoạn khối u trở nên khả thi và chính xác hơn. Do đó, đề tài này tập trung vào việc ứng dụng mô hình CNN để phân đoạn khối u trên ảnh y khoa nhằm nâng cao độ chính xác và hiệu quả trong chẩn đoán.

**2. Mục tiêu**

Đề tài này hướng đến việc xây dựng một mô hình học sâu có khả năng phân đoạn khối u trên ảnh y khoa với độ chính xác cao. Cụ thể, đề tài sẽ sử dụng mô hình CNN để tự động xác định và tách biệt các vùng có khối u trên các ảnh CT và MRI. Mục tiêu cuối cùng là tạo ra một hệ thống hỗ trợ đắc lực cho bác sĩ, giúp giảm thời gian phân tích ảnh và tăng tính chính xác trong chẩn đoán, từ đó góp phần vào việc nâng cao chất lượng chăm sóc sức khỏe.

**3. Nội dung**

Đề tài bao gồm các nội dung chính:

* Thu thập dữ liệu: Tìm kiếm và thu thập ảnh y khoa, bao gồm cả ảnh CT và MRI từ nhiều nguồn, phục vụ cho việc huấn luyện và kiểm tra mô hình.
* Xử lý dữ liệu: Chuẩn hóa và phân chia dữ liệu thành các tập huấn luyện, đánh giá và kiểm tra để đảm bảo tính chính xác và hiệu quả của mô hình.
* Xây dựng và huấn luyện mô hình: Thiết kế mô hình CNN phù hợp cho nhiệm vụ phân đoạn, sau đó tiến hành huấn luyện trên tập dữ liệu đã chuẩn bị.
* Đánh giá hiệu suất: Kiểm tra và đánh giá mô hình trên tập kiểm tra để đảm bảo rằng mô hình đạt được độ chính xác cao trong việc phân đoạn khối u.
* Phân tích và báo cáo kết quả: Đánh giá các kết quả đạt được, phân tích những hạn chế và đề xuất hướng phát triển trong tương lai.

**4. Đối tượng và phạm vi nghiên cứu**

Đối tượng nghiên cứu bao gồm các hình ảnh y khoa, đặc biệt là ảnh CT và MRI của các bệnh nhân có và không có khối u. Phạm vi nghiên cứu giới hạn ở việc xây dựng một mô hình phân đoạn khối u trên dữ liệu ảnh đã thu thập, không bao gồm các vấn đề liên quan đến chẩn đoán lâm sàng.

**5. Phương pháp nghiên cứu**

Đề tài sẽ áp dụng phương pháp học sâu với CNN cho quá trình phân đoạn. Phương pháp này bao gồm các bước như thu thập dữ liệu, tiền xử lý dữ liệu, huấn luyện mô hình, và đánh giá hiệu suất. Các dữ liệu sau khi thu thập sẽ được xử lý và đưa vào mô hình CNN để huấn luyện. Mô hình này sẽ được đánh giá trên các ảnh chưa qua huấn luyện để xác định độ chính xác và khả năng nhận diện khối u. Quá trình này được thực hiện qua các công cụ hỗ trợ như Google Colab, giúp đơn giản hóa và tối ưu hóa quá trình thực hiện.

# CHƯƠNG 1: TỔNG QUAN

* 1. **Giới thiệu về phân đoạn khối u trong ảnh y khoa**

Trong y học, hình ảnh y khoa là một công cụ không thể thiếu để hỗ trợ chẩn đoán và theo dõi điều trị. Các kỹ thuật hình ảnh như cắt lớp vi tính (CT) và cộng hưởng từ (MRI) đã giúp các bác sĩ có thể quan sát chi tiết bên trong cơ thể bệnh nhân mà không cần can thiệp xâm lấn. Đặc biệt, việc phát hiện và đánh giá các khối u dựa trên ảnh y khoa đóng vai trò quan trọng trong việc điều trị các bệnh lý nguy hiểm như ung thư. Tuy nhiên, để xác định chính xác các vùng khối u trên ảnh y khoa, cần một quá trình phân đoạn phức tạp, đòi hỏi kinh nghiệm và kỹ năng của các chuyên gia.

Phân đoạn khối u trong ảnh y khoa là quá trình tách riêng các vùng có dấu hiệu bất thường, như khối u, khỏi các cấu trúc bình thường xung quanh. Điều này giúp bác sĩ xác định kích thước, hình dạng, vị trí và mức độ phát triển của khối u, từ đó đưa ra phương pháp điều trị phù hợp. Tuy nhiên, các phương pháp phân đoạn thủ công không chỉ mất nhiều thời gian mà còn dễ gặp sai sót do phụ thuộc vào yếu tố con người. Để giải quyết những hạn chế này, việc ứng dụng trí tuệ nhân tạo (AI) trong phân đoạn hình ảnh y khoa đã được chú trọng và mang lại nhiều triển vọng.

* 1. **Các phương pháp phân đoạn khối u**

Hiện nay, các phương pháp phân đoạn khối u trong ảnh y khoa có thể chia thành hai nhóm chính: các phương pháp truyền thống và các phương pháp dựa trên học máy và học sâu.

* Phương pháp truyền thống: Phương pháp này bao gồm các kỹ thuật như phân ngưỡng (thresholding), phân cụm (clustering), và phát hiện biên (edge detection). Phương pháp phân ngưỡng dựa trên việc chọn một mức ngưỡng để tách vùng có cường độ cao (thường là khối u) khỏi nền. Phân cụm sử dụng các thuật toán như K-means để chia ảnh thành các vùng đồng nhất về màu sắc hoặc cường độ. Tuy nhiên, các phương pháp truyền thống thường gặp khó khăn khi xử lý các ảnh phức tạp và không thể hiện được sự linh hoạt trong việc nhận diện các khối u có hình dạng và kích thước đa dạng.
* Phương pháp dựa trên học máy (Machine Learning): Một số thuật toán học máy như Random Forest, SVM (Support Vector Machine) cũng được sử dụng trong phân đoạn ảnh y khoa. Tuy nhiên, các phương pháp này vẫn đòi hỏi bước tiền xử lý và trích xuất đặc trưng thủ công, khiến quá trình không được tối ưu và độ chính xác chưa cao.
* Phương pháp dựa trên học sâu (Deep Learning): Trong những năm gần đây, các mô hình học sâu, đặc biệt là Mạng Nơ-ron Tích chập (CNN), đã chứng minh khả năng vượt trội trong việc xử lý các dữ liệu hình ảnh phức tạp. CNN không chỉ tự động học các đặc trưng từ dữ liệu mà còn có khả năng phát hiện các vùng có khối u với độ chính xác cao. Nhờ khả năng tự động trích xuất và học các đặc trưng phức tạp từ ảnh đầu vào, CNN đã trở thành công cụ phổ biến trong phân đoạn ảnh y khoa và mang lại kết quả ấn tượng trong nhiều nghiên cứu.
  1. **Ứng dụng của CNN trong phân đoạn ảnh y khoa**

CNN là một loại mô hình học sâu có cấu trúc đặc biệt, được thiết kế để xử lý dữ liệu hình ảnh. Mô hình này bao gồm nhiều lớp tích chập (convolutional layers) giúp tự động học các đặc trưng như biên, hình dạng, và cấu trúc từ dữ liệu. Mỗi lớp của CNN thực hiện các phép tính tích chập trên dữ liệu đầu vào, từ đó trích xuất các thông tin quan trọng mà không cần can thiệp thủ công. CNN có thể được áp dụng cho nhiều tác vụ khác nhau trong y học, bao gồm nhận diện bệnh lý, phân loại hình ảnh, và đặc biệt là phân đoạn.

Đối với bài toán phân đoạn khối u trong ảnh y khoa, CNN có khả năng phân tách vùng khối u với độ chính xác cao bằng cách học các đặc trưng về hình dạng và cường độ. CNN có thể nhận diện và tách biệt các vùng khối u khỏi nền ảnh, giúp bác sĩ dễ dàng đánh giá tình trạng của bệnh nhân mà không cần tốn nhiều thời gian cho các bước xử lý thủ công. Hiệu quả của CNN trong phân đoạn ảnh y khoa đã được kiểm chứng qua nhiều nghiên cứu quốc tế, và đề tài này sẽ tận dụng ưu điểm của CNN để xây dựng một mô hình có khả năng phân đoạn khối u trên ảnh CT và MRI.

* 1. **Các nghiên cứu liên quan**

Trên thế giới, đã có nhiều nghiên cứu ứng dụng CNN trong phân đoạn khối u trên ảnh y khoa. Một số nghiên cứu nổi bật có thể kể đến:

* Nghiên cứu của Ronneberger và cộng sự (2015) với mô hình U-Net, một kiến trúc CNN được thiết kế đặc biệt cho các tác vụ phân đoạn ảnh y khoa. Mô hình U-Net đã đạt được những kết quả ấn tượng trong việc phân đoạn tế bào và khối u, trở thành một tiêu chuẩn phổ biến trong nghiên cứu phân đoạn y khoa.
* Nghiên cứu của Kamnitsas và cộng sự (2017) với mô hình 3D DeepMedic, áp dụng CNN 3D để phân đoạn các khối u não trên ảnh MRI. Mô hình này đã chứng minh khả năng phân đoạn khối u hiệu quả trên các ảnh y khoa có độ phức tạp cao.

Các nghiên cứu này đã chứng minh tiềm năng của CNN trong việc tự động hóa và nâng cao hiệu quả phân đoạn ảnh y khoa, mở ra cơ hội áp dụng các mô hình này trong thực tế.

* 1. **Hướng tiếp cận**

Dựa trên những thành tựu từ các nghiên cứu trước đây, đề tài này sẽ ứng dụng CNN để xây dựng mô hình phân đoạn khối u trong ảnh y khoa. Các bước thực hiện bao gồm:

* Thu thập và tiền xử lý dữ liệu: Dữ liệu bao gồm ảnh CT và MRI của các bệnh nhân có và không có khối u, được thu thập từ nhiều nguồn khác nhau. Dữ liệu sẽ được chuẩn hóa và phân chia thành các tập huấn luyện, đánh giá và kiểm tra để đảm bảo tính chính xác và đa dạng.
* Xây dựng mô hình CNN: Dựa trên các kiến trúc CNN đã thành công như U-Net, đề tài sẽ thiết kế một mô hình CNN phù hợp để xử lý dữ liệu và phân đoạn các vùng có khối u.
* Huấn luyện và đánh giá mô hình: Mô hình sẽ được huấn luyện trên tập huấn luyện, sau đó đánh giá trên tập kiểm tra để đảm bảo mô hình đạt được độ chính xác cao. Để tối ưu hiệu quả, các tham số của mô hình sẽ được điều chỉnh trong quá trình huấn luyện.
* Phân tích kết quả: Cuối cùng, các kết quả phân đoạn sẽ được so sánh và đánh giá, từ đó rút ra các kết luận về hiệu quả của mô hình trong việc phân đoạn khối u trên ảnh y khoa.

# CHƯƠNG 2: NGHIÊN CỨU LÝ THUYẾT

## Cơ sở lý thuyết

### Phân đoạn ảnh y khoa

Phân đoạn ảnh y khoa là một bước quan trọng trong quá trình phân tích và chẩn đoán bệnh lý từ các ảnh y tế như CT (cắt lớp vi tính), MRI (cộng hưởng từ), X-quang, và siêu âm. Quá trình này giúp phân chia các đối tượng có ý nghĩa y học, chẳng hạn như khối u, mạch máu, mô mềm hay xương, khỏi các vùng khác trong ảnh, từ đó tạo ra các mặt nạ phân đoạn (segmentation masks). Những mặt nạ này sẽ chỉ ra rõ ràng vị trí và kích thước của các đối tượng cần quan tâm, giúp bác sĩ đưa ra quyết định chẩn đoán và điều trị chính xác hơn (ResearchGate, 16 November 2024).

Quá trình phân đoạn ảnh y khoa có thể được chia thành ba giai đoạn chính:

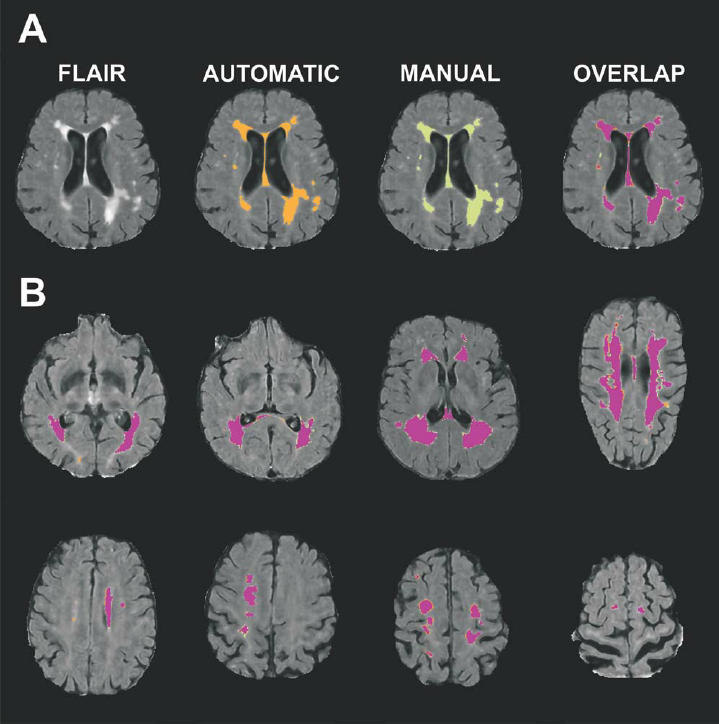
Tiền xử lý dữ liệu: Trước khi bắt đầu phân đoạn, dữ liệu ảnh y khoa cần được tiền xử lý để đảm bảo tính nhất quán và độ chính xác của quá trình phân đoạn. Các bước tiền xử lý có thể bao gồm:

* Chuẩn hóa ảnh: Các ảnh y khoa có thể có độ sáng và độ tương phản khác nhau. Việc chuẩn hóa giúp đưa các giá trị pixel về một phạm vi thống nhất để mô hình dễ dàng học.
* Cân bằng ảnh: Dữ liệu ảnh có thể không đồng đều, ví dụ như có quá nhiều ảnh không có khối u. Các phương pháp như data augmentation (tăng cường dữ liệu) có thể được sử dụng để tạo ra nhiều mẫu dữ liệu hơn từ các ảnh hiện có.
* Loại bỏ nhiễu: Ảnh y khoa, đặc biệt là ảnh MRI và CT, có thể chứa nhiễu làm giảm độ chính xác của phân đoạn. Các kỹ thuật như lọc Gaussian có thể được áp dụng để làm mịn ảnh trước khi phân đoạn.

Phân đoạn ảnh y khoa thường sử dụng các phương pháp học máy hoặc học sâu (deep learning). Các mô hình phân đoạn hiện đại như Mạng Nơ-ron Tích chập (CNN), có khả năng tự động trích xuất các đặc trưng phức tạp từ ảnh mà không cần trích xuất đặc trưng thủ công.

Phân đoạn thủ công: Trong quá khứ, phân đoạn thường được thực hiện thủ công bởi các bác sĩ hoặc chuyên gia y tế, nơi họ vẽ các đường viền quanh khối u hoặc đối tượng quan tâm trên ảnh. Tuy nhiên, phương pháp này không chỉ mất thời gian mà còn dễ mắc phải sai sót do yếu tố con người và phụ thuộc nhiều vào kinh nghiệm của bác sĩ.

Phân đoạn tự động: Các mô hình học sâu hiện nay có thể tự động phát hiện và phân đoạn các đối tượng từ ảnh mà không cần sự can thiệp của con người. Các mô hình CNN học cách nhận diện các đối tượng đặc biệt trong ảnh, như khối u, thông qua việc huấn luyện trên bộ dữ liệu có nhãn.



Hình 1: *So sánh giữa phân đoạn thủ công và phân đoạn tự động trên ảnh MRI [1]*

Hình ảnh trên bao gồm hai phần (A và B) minh họa sự khác biệt giữa phân đoạn thủ công và phân đoạn tự động [1].

Hình ảnh minh họa sự so sánh chi tiết giữa các phương pháp phân đoạn thủ công và tự động trên ảnh MRI sử dụng kỹ thuật FLAIR: phần A gồm bốn hình từ trái sang phải, trong đó hình đầu tiên là ảnh FLAIR gốc thể hiện rõ cấu trúc nội sọ và các vùng bất thường có thể liên quan đến bệnh lý; hình thứ hai là kết quả phân đoạn tự động bằng mô hình học sâu (CNN hoặc U-Net), trong đó các vùng tăng tín hiệu được đánh dấu bằng màu cam, cho thấy khả năng của mô hình trong việc tự động phát hiện các khu vực bất thường mà không cần sự can thiệp thủ công, hình thứ ba là phân đoạn thủ công do các chuyên gia thực hiện, với các vùng đáng chú ý được đánh dấu bằng màu xanh lá dựa trên kinh nghiệm chuyên môn; cuối cùng, hình thứ tư thể hiện sự chồng chéo giữa kết quả phân đoạn tự động và thủ công, với các vùng trùng khớp được tô màu hồng, chứng minh rằng phương pháp tự động có độ chính xác tương đương với chuyên gia trong việc xác định các tổn thương tiềm tàng. Phần B của hình bao gồm các lát cắt khác nhau của não từ trên xuống dưới, minh họa quá trình phân đoạn trên nhiều mặt phẳng khác nhau, với các vùng màu sắc thể hiện mức độ tương đồng giữa hai phương pháp, đặc biệt hữu ích trong chẩn đoán lâm sàng để tối ưu hóa độ chính xác. Hình ảnh này nhấn mạnh tiềm năng của mô hình tự động trong việc tiết kiệm thời gian và công sức so với phương pháp thủ công, đồng thời cải thiện độ nhạy và độ chính xác trong phát hiện các tổn thương não, từ đó hỗ trợ bác sĩ trong chẩn đoán và ra quyết định điều trị, giúp nâng cao hiệu quả chăm sóc sức khỏe cho bệnh nhân.

Xử lý cuối cùng: Sau khi phân đoạn, kết quả cần được xử lý để cải thiện độ chính xác và tính hợp lý của các phân đoạn. Các kỹ thuật như smoothing (làm mịn), morphological operations (phép toán hình học), và connected component analysis (phân tích thành phần liên kết) thường được sử dụng để làm sạch và tối ưu hóa kết quả phân đoạn.

Các phương pháp phân đoạn ảnh y khoa:

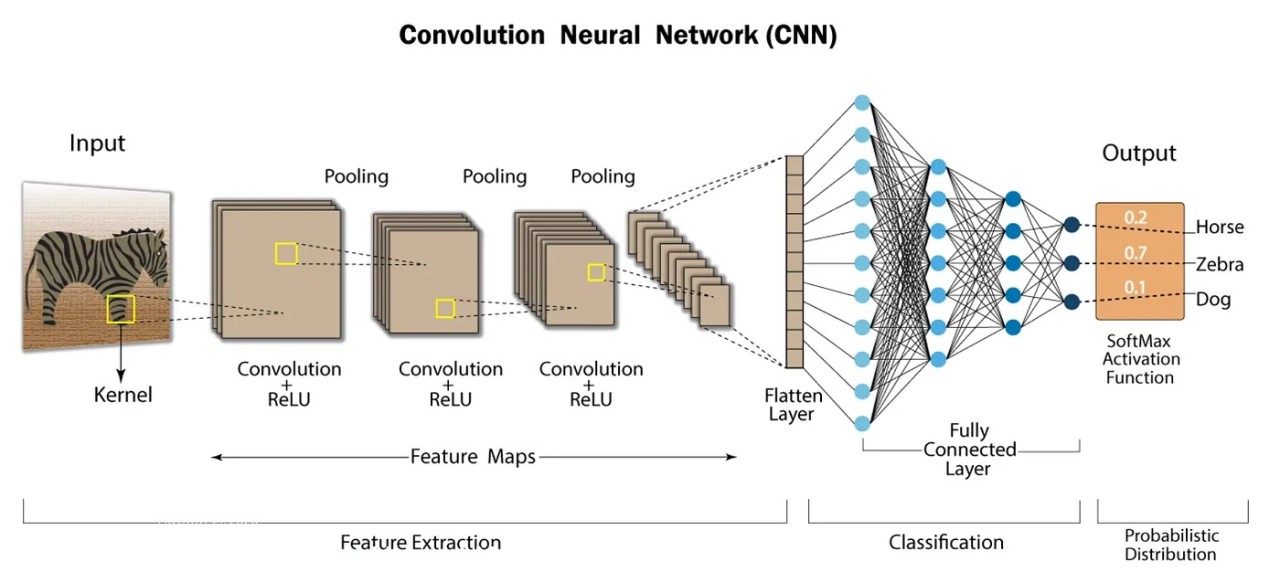
* Phương pháp truyền thống:
* Phân ngưỡng: Đây là phương pháp đơn giản nhất, trong đó các giá trị pixel được phân loại thành hai nhóm dựa trên một ngưỡng cường độ cố định. Phương pháp này hiệu quả khi có sự phân biệt rõ ràng giữa đối tượng và nền, nhưng khó áp dụng khi ảnh có độ tương phản thấp hoặc nhiễu cao.
* Phát hiện biên: Sử dụng các bộ lọc như Sobel, Canny để phát hiện các cạnh của đối tượng. Phương pháp này giúp xác định các biên trong ảnh, nhưng khó khăn trong việc phân đoạn chính xác khi các cạnh không rõ ràng.
* Phân cụm: Các thuật toán như K-means hoặc Fuzzy C-means được sử dụng để nhóm các điểm ảnh có đặc trưng tương tự vào cùng một cụm. Phương pháp này có thể hoạt động tốt với ảnh có độ tương phản rõ ràng, nhưng kém hiệu quả với ảnh có cấu trúc phức tạp.
* Phương pháp học máy:
* Support Vector Machine (SVM): SVM được sử dụng để phân loại các điểm ảnh dựa trên các đặc trưng đã trích xuất. Phương pháp này đòi hỏi trích xuất đặc trưng thủ công, do đó tốn nhiều thời gian và phụ thuộc vào kỹ năng của người thực hiện.
* Random Forest: Một phương pháp dựa trên các cây quyết định để phân loại các điểm ảnh. Random Forest có khả năng kháng nhiễu tốt, nhưng vẫn đòi hỏi trích xuất đặc trưng từ trước.
* Active Contour Model (Snakes): Phương pháp này sử dụng các đường cong điều khiển để tự động điều chỉnh biên của đối tượng. Tuy nhiên, nó dễ bị ảnh hưởng bởi nhiễu và yêu cầu khởi tạo tốt để đạt kết quả chính xác.
* Phương pháp học sâu: Phương pháp học sâu đã tạo ra bước đột phá trong việc phân đoạn ảnh y khoa, đặc biệt là với các mô hình CNN (Convolutional Neural Networks):
* Convolutional Neural Networks (CNN): CNN có khả năng tự động trích xuất đặc trưng từ ảnh đầu vào mà không cần các bước tiền xử lý thủ công. CNN rất hiệu quả trong các bài toán như phân loại và phát hiện đối tượng, nhưng cần được tùy chỉnh cho các bài toán phân đoạn.
* Mask R-CNN: Đây là một mở rộng của Faster R-CNN, cho phép thực hiện phân đoạn đối tượng cùng với nhận diện và phát hiện. Mask R-CNN thường được sử dụng để phân đoạn đa lớp trong ảnh y khoa.
* Fully Convolutional Networks (FCN): FCN là một kiến trúc mạng nơ-ron không sử dụng các lớp fully connected, mà thay vào đó, mọi lớp đều là tích chập, giúp tạo ra các mặt nạ phân đoạn có cùng kích thước với ảnh gốc.
* Lợi ích của các phương pháp phân đoạn ảnh y khoa: Phân đoạn ảnh y khoa đóng vai trò quan trọng trong lĩnh vực chẩn đoán hình ảnh, đặc biệt là trong việc phát hiện và phân tích khối u, tổn thương, và các bất thường khác trên ảnh MRI hoặc CT. Việc áp dụng các phương pháp hiện đại như học máy (Machine Learning) và học sâu (Deep Learning) mang lại nhiều lợi ích quan trọng, bao gồm:
* Tăng độ chính xác và hiệu quả trong chẩn đoán
* Tiết kiệm thời gian và công sức cho bác sĩ
* Khả năng ứng dụng trong điều trị và theo dõi bệnh lý
* Tối ưu hóa quy trình phẫu thuật và xạ trị
* Hỗ trợ nghiên cứu và phát triển y học
* Khả năng mở rộng và áp dụng trên các loại ảnh y khoa khác nhau

### Mạng Nơ-ron tích chập (CNN)

Mạng Nơ-ron Tích chập, hay còn gọi là Convolutional Neural Network (CNN), là một trong những mô hình học sâu tiên tiến nhất, được thiết kế đặc biệt để xử lý và phân tích dữ liệu dạng hình ảnh. CNN được lấy cảm hứng từ cách hoạt động của hệ thống thị giác của con người, nơi các tế bào thần kinh trong vỏ não thị giác phản ứng mạnh mẽ với các đặc trưng cụ thể trong ảnh, chẳng hạn như cạnh, góc, và các họa tiết phức tạp. Khả năng vượt trội của CNN trong việc nhận dạng và phân tích hình ảnh đã biến nó thành một công cụ không thể thiếu trong nhiều lĩnh vực như thị giác máy tính, xử lý ảnh, nhận diện khuôn mặt, và thậm chí cả xử lý ngôn ngữ tự nhiên khi dữ liệu được biểu diễn dưới dạng không gian hai chiều (N. Shahria, "What is Convolutional Neural Network — CNN (Deep Learning)," LinkedIn, November 2024).

Cấu trúc cơ bản của CNN:

* CNN bao gồm một chuỗi các lớp, mỗi lớp thực hiện một vai trò cụ thể trong quá trình trích xuất và phân loại đặc trưng:
* Lớp tích chập (Convolutional Layer): Đây là lớp nền tảng và quan trọng nhất trong CNN, nơi các bộ lọc (filters) được sử dụng để trích xuất các đặc trưng từ dữ liệu đầu vào. Bộ lọc thực hiện phép tích chập (convolution) bằng cách quét qua các vùng nhỏ của ảnh và tạo ra các bản đồ đặc trưng (feature maps). Quá trình này giúp nhận diện các đặc trưng cơ bản như cạnh, góc, và các mẫu phức tạp hơn. Mỗi bộ lọc trong lớp tích chập có khả năng học cách nhận diện một loại đặc trưng nhất định. Ví dụ, một bộ lọc có thể học cách phát hiện cạnh dọc, trong khi bộ lọc khác tập trung vào các họa tiết tròn. Các bản đồ đặc trưng này giúp mô hình xây dựng một biểu diễn có ý nghĩa về cấu trúc của ảnh, từ đó phân biệt các đối tượng khác nhau.
* Lớp kích hoạt (Activation Layer): Sau khi trích xuất đặc trưng, các hàm kích hoạt như ReLU (Rectified Linear Unit) được áp dụng để thêm tính phi tuyến cho mô hình. Điều này rất quan trọng vì nếu không có tính phi tuyến, CNN chỉ có thể học được các mối quan hệ tuyến tính, làm giảm khả năng biểu diễn của mô hình. Hàm ReLU hoạt động bằng cách loại bỏ các giá trị âm trong các bản đồ đặc trưng, giữ lại các giá trị dương. Quá trình này không chỉ cải thiện hiệu quả tính toán bằng cách loại bỏ các giá trị không cần thiết mà còn giúp mô hình học được các đặc trưng phức tạp hơn.
* Lớp pooling (Pooling Layer): Để giảm kích thước của map đặc trưng, Max Pooling hoặc Average Pooling được áp dụng, giúp giảm số lượng tham số và tăng tốc độ huấn luyện mô hình; Max Pooling chọn giá trị lớn nhất trong một vùng nhỏ của map đặc trưng, giúp giữ lại các đặc trưng quan trọng trong khi làm giảm độ phức tạp và hạn chế hiện tượng overfitting, đồng thời tăng khả năng chống nhiễu của mô hình.
* Lớp hoàn toàn kết nối (Fully Connected Layer): Các lớp này liên kết tất cả các neuron từ lớp trước đó với nhau, tạo ra một mạng lưới dày đặc; mục tiêu là sử dụng các đặc trưng đã trích xuất từ các lớp trước đó để phân loại hoặc dự đoán đầu ra cuối cùng. Các lớp hoàn toàn kết nối thường được đặt ở cuối mô hình CNN để tổng hợp tất cả các thông tin đã học được từ các lớp trước đó và đưa ra dự đoán chính xác dựa trên các đặc trưng đã học, chẳng hạn như nhận diện khối u trên ảnh MRI.

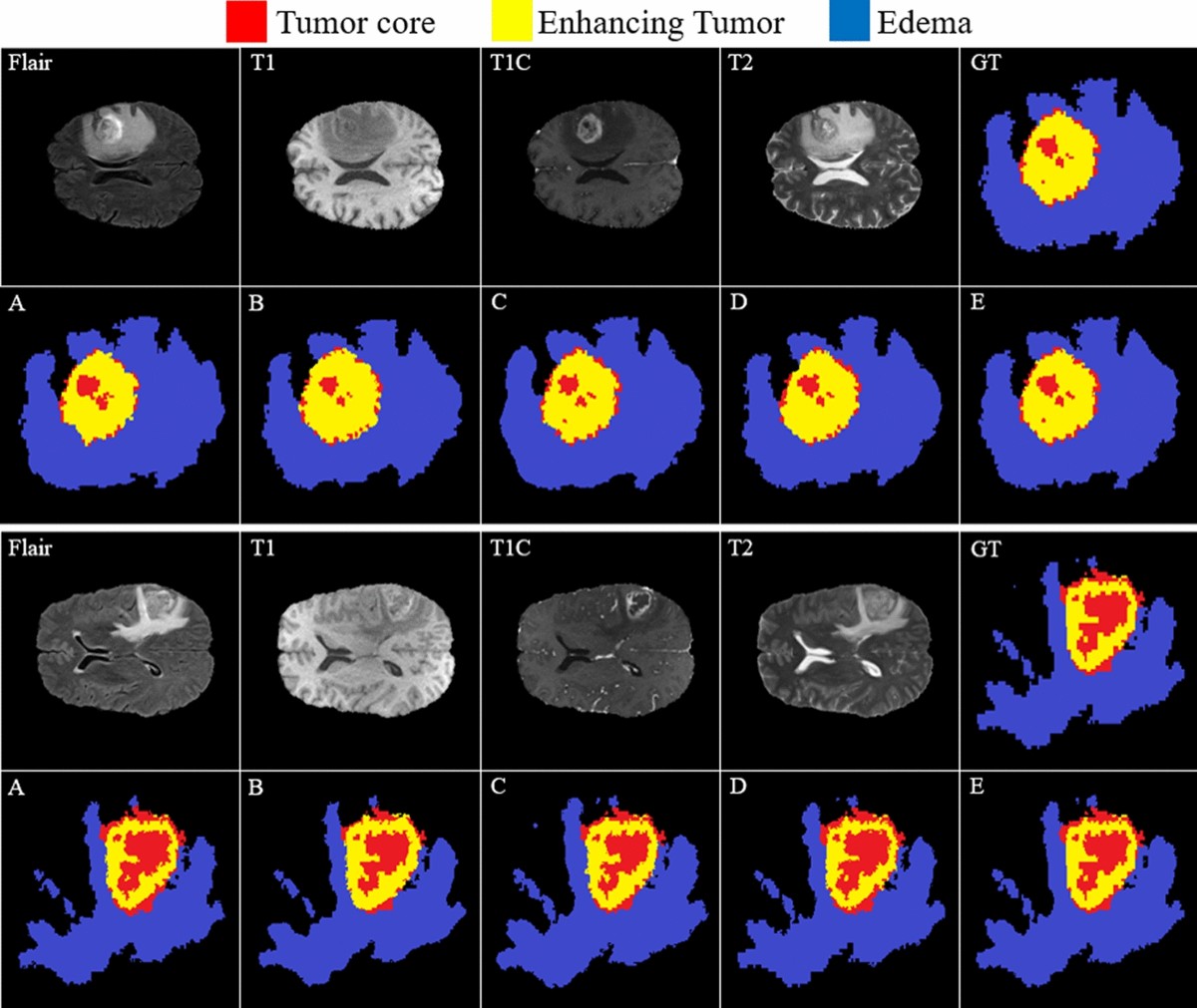


Hình 2: *Kiến trúc cơ bản của Mạng Nơ-ron Tích chập (CNN) [2]*

Hình ảnh trên thể hiện kiến trúc CNN gồm các lớp tích chập, lớp kích hoạt, lớp pooling và lớp hoàn toàn kết nối, được sử dụng cho các nhiệm vụ phân loại ảnh y khoa và các ứng dụng phân đoạn ảnh [2].

Mạng Nơ-ron Tích chập (CNN) đã chứng minh hiệu quả vượt trội trong việc phân tích và xử lý ảnh y khoa, đặc biệt trong các nhiệm vụ phân đoạn các cấu trúc y học phức tạp như khối u, mô mềm, mạch máu, và các tổn thương trên ảnh MRI, CT hoặc X-quang. CNN không chỉ cải thiện độ chính xác mà còn giảm thiểu thời gian và công sức của các chuyên gia y tế trong việc chẩn đoán và theo dõi bệnh lý. Dưới đây là một số ứng dụng chính của CNN trong phân đoạn ảnh y khoa:

* Phân đoạn khối u trong ảnh MRI và CT:



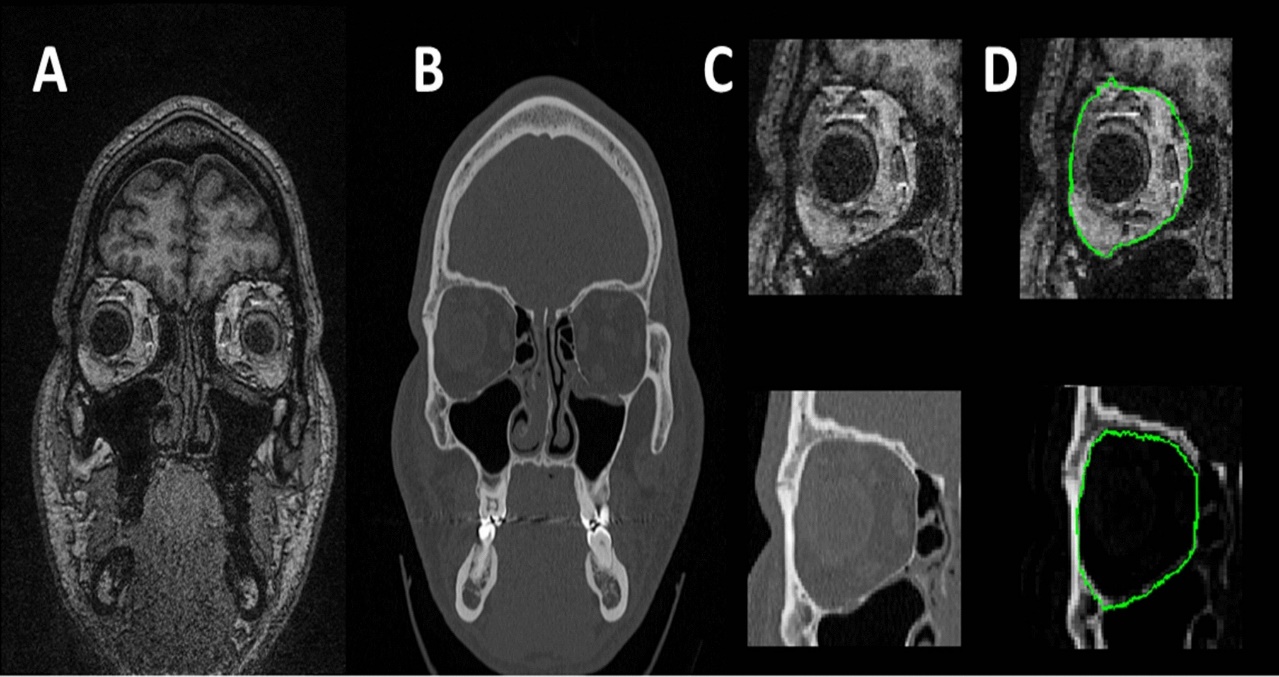
Hình 3: *Phân đoạn khối u não trên ảnh MRI bằng mô hình CNN [3]*

Hình ảnh trên thể hiện phân đoạn khối u não trên ảnh MRI sử dụng mô hình CNN. Hình ảnh bao gồm các loại ảnh MRI khác nhau (Flair, T1, T1C, T2) cùng với kết quả phân đoạn các vùng lõi khối u, vùng phát triển và vùng phù nề [3].

CNN được sử dụng rộng rãi để phân đoạn khối u não, giúp xác định ranh giới chính xác của khối u trên ảnh MRI. Mô hình CNN có thể tự động phát hiện và phân đoạn các vùng bất thường, giúp bác sĩ chẩn đoán sớm và đưa ra kế hoạch điều trị hiệu quả hơn.

Trong ảnh CT, CNN có thể được sử dụng để phân đoạn khối u gan, khối u phổi và các khối u ở các cơ quan khác, giúp đánh giá kích thước và mức độ lan rộng của khối u.

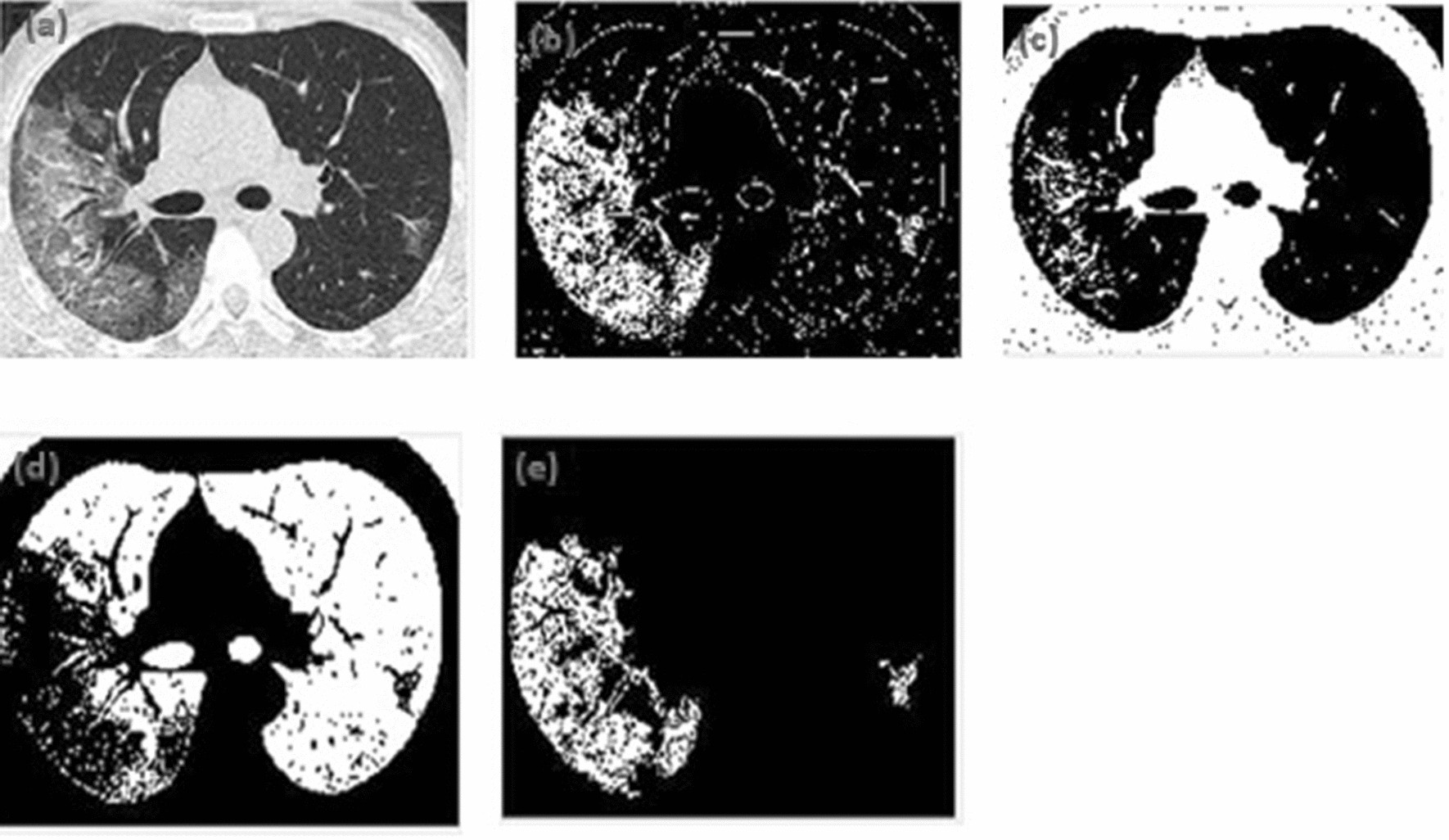
* Phân đoạn mô và các cấu trúc giải phẫu:
* CNN có khả năng phân đoạn các mô mềm, như tim, phổi, gan và các cấu trúc khác trên ảnh CT và MRI. Điều này giúp bác sĩ dễ dàng xác định và theo dõi sự thay đổi của các mô theo thời gian, đặc biệt trong các trường hợp bệnh lý mãn tính.
* Phân đoạn mạch máu trên ảnh chụp mạch (angiography) để phát hiện các vùng tắc nghẽn hoặc hẹp mạch máu, hỗ trợ trong việc lập kế hoạch phẫu thuật và điều trị.



Hình 4: *Phân đoạn mô mềm trên ảnh CT bằng CNN [4]*

Hình trên thể hiện phân đoạn mô mềm trên ảnh CT và MRI bằng mạng nơ-ron tích chập (CNN). Hình (A) và (B) là ảnh gốc, trong khi hình (C) và (D) cho thấy kết quả phân đoạn tự động các vùng mô mềm [4].

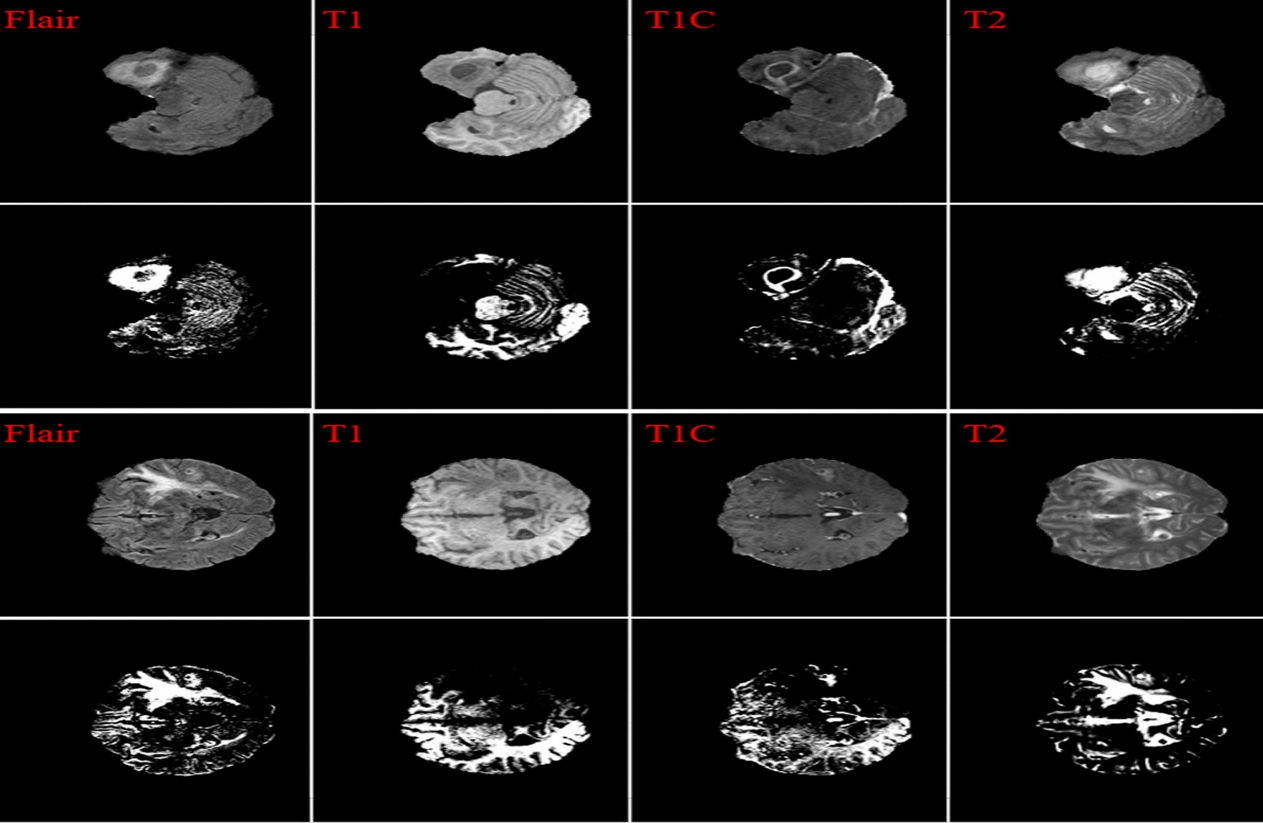
* Hỗ trợ chẩn đoán bệnh lý và theo dõi tiến triển của bệnh:
* CNN đã được ứng dụng hiệu quả trong việc phân đoạn các vùng viêm phổi trên ảnh CT, đặc biệt là trong chẩn đoán COVID-19. Các mô hình CNN có thể tự động phát hiện các vùng tổn thương như "mờ kính mờ" hoặc tổn thương lan tỏa, giúp xác định mức độ nghiêm trọng của bệnh và hỗ trợ bác sĩ trong việc lập kế hoạch điều trị. Khả năng theo dõi tiến triển của bệnh qua các giai đoạn cũng là một lợi thế lớn, giúp đánh giá hiệu quả điều trị và điều chỉnh phác đồ phù hợp.Phân đoạn các vùng tổn thương trên ảnh X-quang giúp phát hiện các bệnh như lao phổi, tràn dịch màng phổi, và các bệnh lý khác liên quan đến hệ hô hấp.
* Phân đoạn các vùng tổn thương trên ảnh X-quang giúp phát hiện các bệnh như lao phổi, tràn dịch màng phổi, và các bệnh lý khác liên quan đến hệ hô hấp.



Hình 5: *Phân đoạn tổn thương phổi ảnh CT để chẩn đoán COVID-19 [5]*

Hình ảnh trên thể hiện quy trình phân đoạn tổn thương phổi trên ảnh CT để chẩn đoán COVID-19 bằng mô hình học sâu. Các bước từ (a) đến (e) thể hiện quá trình từ ảnh gốc đến phân đoạn tự động [5].

* Phân đoạn tổn thương và theo dõi tiến triển của bệnh lý thần kinh:
* CNN đặc biệt hữu ích trong việc phân đoạn các tổn thương thần kinh trên ảnh MRI của não, giúp phát hiện các vùng bị tổn thương do đột quỵ, chấn thương sọ não, hoặc bệnh Alzheimer.
* Việc phân đoạn tự động giúp theo dõi sự tiến triển của bệnh lý thần kinh theo thời gian, từ đó điều chỉnh phác đồ điều trị cho bệnh nhân.



Hình 6: *Phân đoạn khối u não trên ảnh MRI [6]*

Hình ảnh trên thể hiện sự phân đoạn khối u não trên ảnh MRI sử dụng các chế độ hình ảnh khác nhau (FLAIR, T1, T1C, T2). Các ảnh thể hiện sự khác biệt về độ tương phản của khối u não khi sử dụng từng chế độ. Hàng trên cùng là ảnh gốc, hàng giữa là ảnh sau khi xử lý bằng các bộ lọc, và hàng dưới cùng là các kết quả phân đoạn khối u thu được từ mô hình CNN [6].

Lợi ích của việc ứng dụng CNN trong phân đoạn ảnh y khoa:

* Tự động hóa và tăng tốc độ chẩn đoán: CNN có khả năng phân tích hàng nghìn ảnh y khoa trong thời gian ngắn, giúp bác sĩ tiết kiệm thời gian và tập trung vào việc ra quyết định điều trị.
* Độ chính xác cao: Với khả năng học từ dữ liệu lớn, CNN có thể đạt được độ chính xác cao hơn so với các phương pháp thủ công truyền thống, giúp giảm thiểu sai sót trong chẩn đoán.
* Hỗ trợ cá nhân hóa điều trị: Việc phân đoạn chính xác các vùng bất thường giúp cá nhân hóa phác đồ điều trị, từ đó cải thiện hiệu quả điều trị và chất lượng cuộc sống của bệnh nhân.

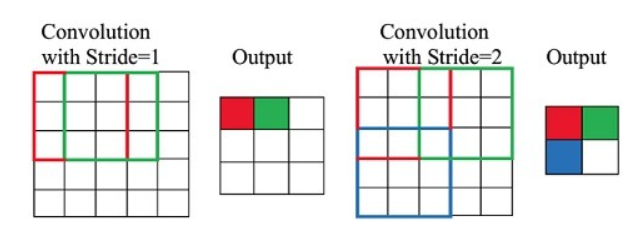
Hạn chế của việc ứng dụng CNN trong phân đoạn ảnh y khoa:

* Yêu cầu dữ liệu lớn: CNN cần một lượng lớn dữ liệu để huấn luyện nhằm đạt được độ chính xác cao.
* Tốn tài nguyên tính toán: Quá trình huấn luyện mô hình CNN đòi hỏi GPU mạnh để giảm thời gian xử lý.
* Nguy cơ overfitting: Nếu dữ liệu không đủ đa dạng, mô hìx`nh có thể dễ dàng bị quá khớp (overfitting) và không hoạt động tốt trên dữ liệu mới.

## Các khái niệm chính trong CNN

### Stride

Stride là bước nhảy của cửa sổ tích chập khi nó di chuyển qua ảnh đầu vào. Giá trị stride quyết định tốc độ di chuyển của cửa sổ. Nếu stride bằng 1, cửa sổ tích chập sẽ di chuyển từng bước một qua các pixel, còn nếu stride bằng 2, cửa sổ sẽ di chuyển cách 2 pixel một lần.



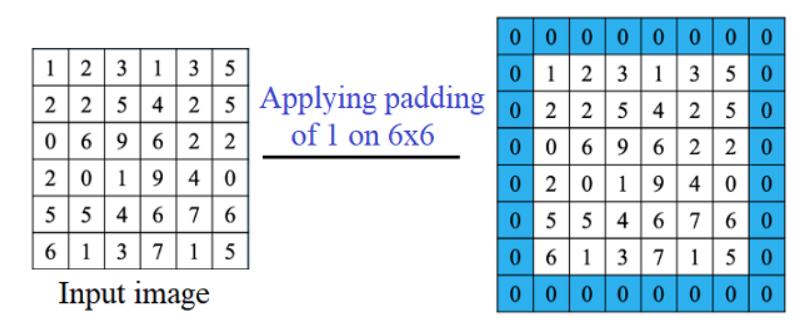
Hình 7: *Convolution với các bước nhảy (Stride) [8]*

Giá trị stride càng lớn, kích thước của đầu ra sẽ nhỏ hơn vì cửa sổ tích chập sẽ bỏ qua nhiều pixel hơn. Điều này có thể làm giảm độ phân giải của đầu ra, nhưng đồng thời giảm thiểu khối lượng tính toán.

Stride là một tham số linh hoạt, đóng vai trò quan trọng trong việc kiểm soát quá trình trích xuất đặc trưng và tối ưu hóa hiệu suất mạng CNN. Hiểu rõ tác động của stride sẽ giúp bạn xây dựng mô hình hiệu quả hơn và đạt được kết quả mong muốn.

### Padding

Padding là kỹ thuật thêm các pixel giả (thường là giá trị 0, gọi là zero padding) xung quanh biên của ảnh đầu vào. Điều này giúp duy trì kích thước đầu ra sau khi tích chập.



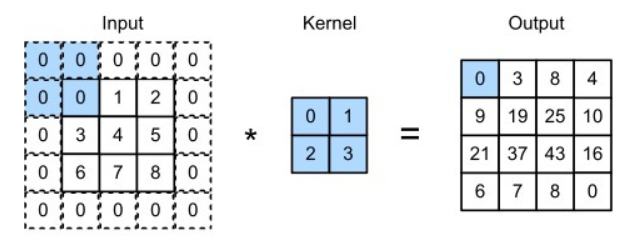
Hình 8: *Minh họa áp dụng Padding trên ma trận 6x6 [8]*

Trong nhiều trường hợp, người ta thêm padding để kích thước đầu ra của lớp tích chập không bị giảm. Ví dụ, nếu không có padding, mỗi lần tích chập có thể làm giảm kích thước không gian của ảnh đầu ra.

Padding là một kỹ thuật quan trọng trong Convolutional Neural Networks (CNN), giúp bảo toàn kích thước đầu ra và hạn chế mất mát thông tin, đặc biệt tại các biên của ảnh đầu vào. Kỹ thuật này thường được sử dụng để thêm các giá trị giả (thường là 0, gọi là zero padding) xung quanh ảnh, cho phép kernel quét qua toàn bộ dữ liệu mà không bỏ sót các pixel ở rìa. Padding giúp duy trì kích thước không gian của đầu ra, đặc biệt quan trọng trong các bài toán yêu cầu bảo toàn chi tiết hoặc khi mạng có nhiều lớp tích chập. Ngoài ra, việc sử dụng padding hợp lý giúp tăng khả năng trích xuất đặc trưng mà không làm tăng đáng kể khối lượng tính toán. Do đó, padding là công cụ không thể thiếu, đảm bảo hiệu suất và độ chính xác của mạng CNN.

### Filters (Kernels)

Filters hay còn gọi là kernels là các ma trận nhỏ được áp dụng lên ảnh đầu vào trong quá trình tích chập. Các bộ lọc này thực hiện việc quét qua toàn bộ ảnh đầu vào, tính toán các giá trị mới dựa trên phép nhân tích chập giữa bộ lọc và các phần tương ứng của ảnh.



Hình 9: *Phép tích chập trên ma trận đầu vào với Padding và Kernel [8]*

Mỗi bộ lọc sẽ phát hiện các đặc trưng cụ thể, chẳng hạn như cạnh, đường nét, hoặc chi tiết phức tạp hơn ở các lớp sâu. Mỗi lớp tích chập trong mạng CNN có thể sử dụng nhiều bộ lọc để phát hiện nhiều đặc trưng khác nhau.

Tóm lại, việc sử dụng nhiều filters trong các lớp tích chập cho phép mạng CNN trích xuất được nhiều đặc trưng đa dạng và phong phú, góp phần quan trọng vào hiệu quả và độ chính xác của mô hình trong các bài toán về thị giác máy tính. Filters là thành phần không thể thiếu, giúp CNN phân tích và hiểu được các đặc trưng phức tạp trong dữ liệu hình ảnh.

### Feature Maps

Feature Maps (Bản đồ đặc trưng) là kết quả đầu ra của một lớp tích chập sau khi áp dụng các bộ lọc lên ảnh đầu vào. Đây là nơi lưu trữ các đặc trưng đã được phát hiện bởi các bộ lọc trong quá trình tích chập.

Feature maps thể hiện sự hiện diện của các đặc trưng (như cạnh, góc) trong một bức ảnh tại các vị trí không gian khác nhau. Các feature maps càng sâu trong mạng CNN thì chứa các đặc trưng càng phức tạp, trừu tượng hơn.

Feature maps không chỉ là kết quả của các phép tích chập mà còn là biểu diễn quan trọng giúp CNN hiểu các đặc trưng trong dữ liệu hình ảnh. Chúng thể hiện sự hiện diện của các đặc trưng tại các vị trí khác nhau trong ảnh và là cơ sở để mạng CNN xây dựng các biểu diễn trừu tượng hơn ở các lớp tiếp theo. Đây là yếu tố cốt lõi giúp CNN đạt hiệu quả cao trong các bài toán xử lý và nhận diện hình ảnh.

## Ứng dụng thực tế của CNN

Mạng Nơ-ron tích chập (CNN) là một trong những công nghệ quan trọng và được sử dụng rộng rãi trong nhiều lĩnh vực liên quan đến xử lý hình ảnh và dữ liệu không gian. Dưới đây là các ứng dụng thực tế phổ biến của CNN:

### Image Recognition & Classification

CNN đã đạt được nhiều thành công vượt trội trong các bài toán nhận diện và phân loại hình ảnh. Các mô hình CNN có thể phân loại các đối tượng trong ảnh với độ chính xác rất cao.

Ứng dụng: Phân loại động vật, nhận diện các loại phương tiện giao thông, phân loại hoa và thực phẩm,….

### ****Optical Character Recognition – OCR****

CNN không chỉ ứng dụng cho hình ảnh mà còn được áp dụng trong xử lý ngôn ngữ tự nhiên và nhận diện ký tự quang học (OCR), giúp máy tính nhận diện và chuyển đổi văn bản từ hình ảnh sang dạng văn bản số.

Ứng dụng: Quét và nhận diện văn bản trong các tài liệu giấy, biển báo giao thông, sách, hoặc hóa đơn.

### Image Segmentation

Phân đoạn ảnh là quá trình chia một bức ảnh thành các phần hoặc đối tượng khác nhau dựa trên các đặc điểm. CNN giúp phân đoạn các vùng khác nhau trong ảnh và xác định các đặc trưng riêng biệt của từng vùng.

Ứng dụng: Phân đoạn trong y tế để xác định các tế bào ung thư hoặc các mô bất thường trong ảnh y khoa, phân đoạn đường trong hình ảnh vệ tinh.

### Autonomous Driving

CNN đóng vai trò quan trọng trong các hệ thống lái xe tự động, giúp xe nhận diện và phân tích môi trường xung quanh, từ đó đưa ra quyết định về điều hướng và tránh vật cản.

Ứng dụng: Nhận diện biển báo giao thông, làn đường, người đi bộ, phương tiện khác và các vật thể nguy hiểm.

### Action Recognition & Video Analysis

CNN cũng được mở rộng sang phân tích video, giúp nhận diện hành động, phân tích chuyển động và trích xuất thông tin từ các đoạn video.

Ứng dụng: Phân tích chuyển động trong các đoạn video thể thao, nhận diện các hành động bất thường trong video giám sát.

## Ngôn ngữ và công cụ phát triển

### Ngôn ngữ hỗ trợ

Python được lựa chọn làm ngôn ngữ chính cho dự án này nhờ tính linh hoạt, cú pháp dễ hiểu, và hỗ trợ mạnh mẽ cho các ứng dụng học máy và học sâu. Python có hệ sinh thái phong phú với nhiều thư viện hỗ trợ cho các tác vụ phân đoạn ảnh, đặc biệt là học sâu.

Python có một hệ sinh thái mạnh mẽ với hàng ngàn thư viện phục vụ cho các ứng dụng học máy và học sâu. Các thư viện chính được sử dụng trong dự án bao gồm:

* TensorFlow và Keras: Đây là hai thư viện học sâu hàng đầu, cung cấp các công cụ và mô hình tiên tiến giúp xây dựng, huấn luyện và đánh giá các mạng nơ-ron tích chập (CNN).
* NumPy: Hỗ trợ thao tác với mảng số liệu (arrays) và ma trận, giúp tối ưu hóa hiệu suất khi xử lý dữ liệu lớn.
* OpenCV: Được sử dụng cho các thao tác tiền xử lý hình ảnh như cắt, lọc nhiễu và tăng cường dữ liệu (data augmentation).
* Matplotlib và Seaborn: Các thư viện trực quan hóa dữ liệu, giúp vẽ đồ thị và biểu đồ để theo dõi quá trình huấn luyện mô hình, đánh giá kết quả.

Python tương thích tốt với nhiều hệ điều hành và nền tảng, bao gồm Linux, macOS và Windows. Điều này giúp dễ dàng triển khai mã nguồn trên các môi trường khác nhau như máy chủ đám mây (Google Colab, AWS, Azure).

Ứng dụng thực tế của Python trong lĩnh vực y tế: Python đã được chứng minh là một công cụ hữu ích trong các dự án y tế, đặc biệt là trong phân tích hình ảnh y khoa. Các mô hình học sâu như U-Net được xây dựng bằng Python đã giúp cải thiện độ chính xác trong phân đoạn ảnh MRI, CT và siêu âm, hỗ trợ chẩn đoán và điều trị bệnh nhân hiệu quả hơn.

### Môi trường và công cụ phát triển

Google Colab (viết tắt của Google Colaboratory) là một nền tảng trực tuyến do Google phát triển, cho phép người dùng viết và chạy mã Python ngay trên trình duyệt mà không cần cài đặt bất kỳ phần mềm nào. Colab đặc biệt hữu ích cho các dự án học máy (Machine Learning) và học sâu (Deep Learning) vì nó cung cấp GPU và TPU miễn phí, giúp tăng tốc quá trình huấn luyện mô hình.

Lý do sử dụng Google Colab:

* Tích hợp tốt với Python và các thư viện học sâu: Colab hỗ trợ hầu hết các thư viện phổ biến trong Python như TensorFlow, Keras, PyTorch, OpenCV, và NumPy. Điều này giúp dễ dàng phát triển và thử nghiệm các mô hình học sâu như CNN.
* Không cần cài đặt phần mềm: Tất cả các công cụ và thư viện đã được cài đặt sẵn trên Colab. Người dùng không cần phải lo lắng về việc cấu hình môi trường hay cài đặt thư viện thủ công.

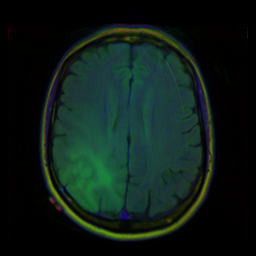
Ứng dụng của Google Colab trong dự án:

* Tiền xử lý dữ liệu: Sử dụng OpenCV và NumPy để chuẩn bị các ảnh MRI trước khi đưa vào mô hình CNN.
* Huấn luyện mô hình: Sử dụng TensorFlow và Keras để xây dựng và huấn luyện mô hình CNN với GPU miễn phí trên Colab, giúp giảm thời gian huấn luyện từ hàng giờ xuống chỉ còn vài phút.
* Trực quan hóa kết quả: Sử dụng Matplotlib để hiển thị kết quả phân đoạn và đánh giá hiệu suất mô hình.
* Lưu trữ và quản lý mô hình: Kết nối với Google Drive để lưu trữ các mô hình huấn luyện, cho phép truy cập dễ dàng và an toàn.

# CHƯƠNG 3: HIỆN THỰC HÓA NGHIÊN CỨU

## 3.1 Mô tả bài toán

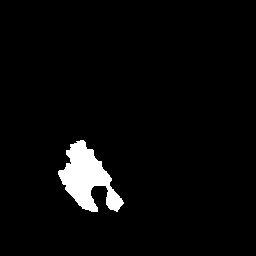
Bài toán phân đoạn ảnh não được dựa trên mô hình CNN, với mục tiêu chính là xác định vị trí và phân đoạn chính xác các vùng bị ảnh hưởng bởi khối u trong ảnh cộng hưởng từ (MRI).



Hình 10: *Hình ảnh đầu vào (Input)*

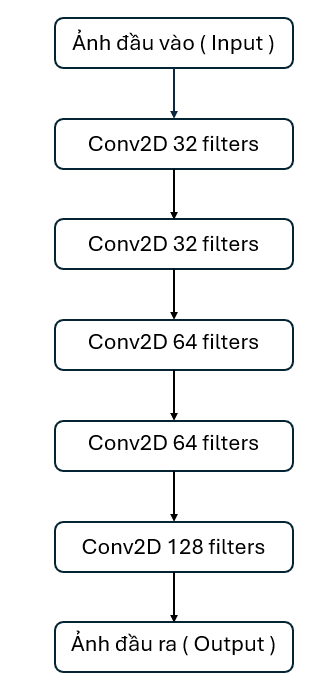
Hình ảnh đầu vào (Input) của bài toán là các ảnh MRI của não chưa được phân đoạn, thể hiện lát cắt ngang qua não, như được minh họa ở hình 10.

Đầu ra của bài toán (Output) là ảnh phân đoạn tương ứng, trong đó các vùng bị ảnh hưởng bởi khối u được đánh dấu rõ ràng, như được minh họa ở hình 11.



Hình 11: *Hình ảnh đầu ra (Output)*

## 3.2 Sơ đồ khối



Hình 12: *Sơ đồ khối*

Sơ đồ trên (hình 12) thể hiện mô hình CNN đơn giản được xây dựng cho bài toán phân đoạn khối u trong ảnh y khoa (MRI). Mô hình này gồm các khối lớp tích chập (Conv2D) với các bộ lọc (filters) khác nhau được thể hiện như sau:

Ảnh đầu vào (Input): Ảnh đầu vào là một hình ảnh có kích thước 256x256 với 3 kênh màu (RGB), tương ứng với các thông tin hình ảnh MRI chưa được phân đoạn.

Các lớp Conv2D:

* Conv2D (32 filters): Lớp này thực hiện phép tích chập với 32 bộ lọc, kích thước mỗi bộ lọc là 3x3. Lớp này được lặp lại 2 lần để trích xuất đặc trưng từ ảnh đầu vào.
* Conv2D (64 filters): Lớp tích chập thứ ba sử dụng 64 bộ lọc để tăng cường việc học các đặc trưng phức tạp hơn trong ảnh.
* Conv2D (128 filters): Các lớp tích chập sau sử dụng 128 bộ lọc, giúp mô hình học được các đặc trưng chi tiết và phức tạp hơn từ ảnh.

Ảnh đầu ra (Output): Ảnh đầu ra của mô hình là một bản phân đoạn với 1 kênh màu (chỉ gồm hai giá trị: vùng khối u và nền). Kết quả phân đoạn này giúp xác định vùng ảnh có khối u.

## 3.3 Xử lý dữ liệu

## 3.3.1 Tải dữ liệu và xử lý hình ảnh

Quá trình tải dữ liệu và xử lý hình ảnh ban đầu đóng vai trò quan trọng trong việc chuẩn bị dữ liệu đầu vào cho mô hình CNN. Trong nghiên cứu này, dữ liệu được tải từ hai thư mục chính: ảnh gốc (ở thư mục tumor) và mặt nạ tương ứng (ở thư mục mask).

Các bước thực hiện bao gồm:

* Xác định đường dẫn thư mục hình ảnh:

image\_dir = ‘/content/dataset\_extracted/brain\_tumor\_custom/Tumor’.

* Xác định đường dẫn thư mục mặt nạ:

mask\_dir = ‘/content/dataset\_extracted/brain\_tumor\_custom/Tumor’.

* Kiểm tra tính đồng bộ: Đảm bảo rằng mỗi ảnh trong thư mục image đã có mặt nạ tương ứng trong thư mục mask.

Quá trình tải dữ liệu là nền tảng cho các bước tiếp theo trong nghiên cứu, bao gồm tiền xử lý dữ liệu và huấn luyện mô hình. Trong bước này, dữ liệu đã được đọc đúng và đảm bảo chất lượng cho các giai đoạn xử lý và huấn luyện tiếp theo.

## 3.3.2 Tiền xử lý dữ liệu

Quá trình tiền xử lý dữ liệu là bước quan trọng trong việc chuẩn bị dữ liệu đầu vào cho mô hình. Mục tiêu của bước này là đảm bảo chất lượng dữ liệu và đảm bảo rằng các ảnh đầu vào và mặt nạ tương ứng có thể được mô hình hóa một cách hiệu quả. Các bước thực hiện tiền xử lý dữ liệu trong nghiên cứu này bao gồm:

* Đọc và xử lý ảnh và mặt nạ (Image and Mask Processing): Ảnh đầu vào với dữ liệu gốc nằm trong thư mục Tumor và mặt nạ tương ứng với dữ liệu hình ảnh nằm trong thư mục mask. Ảnh và mặt nạ được đọc và xử lý đồng thời với nhau, giúp đảm bảo tính đồng bộ giữa ảnh và mặt nạ.
* Kiểm tra tính đồng bộ: Dữ liệu được kiểm tra tính đồng bộ giữa ảnh và mặt nạ. Điều này được thực hiện bằng cách so khớp tên tệp ảnh với tên tệp mặt nạ. Tệp ảnh có định dạng .tif và tệp mặt nạ sẽ có tên tương tự nhưng có thêm hậu tố \_mask.tif. Việc này giúp đảm bảo rằng mỗi ảnh trong thư mục Tumor có một mặt nạ tương ứng trong thư mục Mask.
* Xử lý ảnh: Các ảnh được mở từ thư mục Tumor, chuyển sang định dạng RGB và thay đổi kích thước về kích thước chuẩn (256, 256). Đây là bước quan trọng giúp chuẩn hóa ảnh, làm cho chúng có cùng kích thước và dễ dàng đưa vào mô hình CNN. Ảnh sau khi được xử lý sẽ được chuẩn hóa bằng cách chia giá trị pixel cho 255, giúp đưa các giá trị pixel về phạm vi từ 0 đến 1.
* Xử lý mặt nạ: Các mặt nạ được mở từ thư mục Mask, chuyển sang định dạng grayscale (đơn sắc) và cũng thay đổi kích thước để khớp với kích thước ảnh đầu vào. Sau đó, mặt nạ được nhị phân hóa, tức là các pixel có giá trị lớn hơn 0 được gán là 1, còn các pixel còn lại được gán là 0. Điều này giúp đơn giản hóa việc phân đoạn ảnh, khi chỉ cần xác định các vùng có giá trị khác 0 (vùng có khối u).
* Lưu dữ liệu: Sau khi hoàn thành các bước tiền xử lý, ảnh và mặt nạ được lưu vào các mảng images và masks. Dữ liệu này sẽ được sử dụng làm đầu vào cho mô hình học sâu trong các bước huấn luyện tiếp theo.
* Kiểm tra kết quả: Cuối cùng, sau khi xử lý xong dữ liệu, số lượng ảnh và mặt nạ được in ra để xác nhận rằng dữ liệu đã được tải và xử lý đúng cách.

## 3.3.3 Chia dữ liệu thành tập huấn luyện và kiểm thử

Sau khi hoàn tất việc xử lý và chuẩn hóa dữ liệu, bước tiếp theo là chia dữ liệu thành các tập huấn luyện, đánh giá và kiểm tra. Quá trình này giúp phân chia dữ liệu thành ba phần riêng biệt, mỗi phần sẽ phục vụ một mục đích khác nhau trong quá trình huấn luyện và đánh giá mô hình:

* Chia dữ liệu thành tập huấn luyện (Training Set): Dữ liệu được chia thành 70% cho tập huấn luyện. Tập huấn luyện này sẽ được sử dụng để huấn luyện mô hình CNN, cho phép mô hình học cách phân loại ảnh và phân đoạn các khối u từ ảnh đầu vào.
* Chia dữ liệu thành tập đánh giá (Validation Set): 15% dữ liệu sẽ được sử dụng cho tập đánh giá. Tập đánh giá sẽ được sử dụng trong quá trình huấn luyện để đánh giá hiệu quả của mô hình sau mỗi lần huấn luyện, giúp điều chỉnh các tham số của mô hình như learning rate, số lượng epochs,….
* Chia dữ liệu thành tập kiểm tra (Test Set): 15% dữ liệu còn lại được sử dụng làm tập kiểm tra. Tập kiểm tra này được sử dụng sau khi huấn luyện để đánh giá độ chính xác của mô hình khi áp dụng vào dữ liệu chưa thấy trước.

Quá trình chia dữ liệu được thực hiện bằng cách sử dụng hàm train\_test\_split từ thư viện sklearn.model\_selection. Cụ thể, dữ liệu ảnh và mặt nạ được chia theo tỷ lệ 70% cho tập huấn luyện, 15% cho tập đánh giá và 15% cho tập kiểm tra. Phương pháp này đảm bảo rằng mỗi phần dữ liệu được chọn một cách ngẫu nhiên và không bị trùng lặp, giúp mô hình huấn luyện và kiểm tra chính xác hơn.

## 3.4 Xây dựng mô hình CNN

Sau khi dữ liệu đã được tiền xử lý và chia thành các tập huấn luyện, đánh giá và kiểm tra, bước tiếp theo là xây dựng mô hình CNN. Mô hình CNN trong nghiên cứu này được xây dựng với các lớp tích chập (Convolutional Layers) để thực hiện phân đoạn khối u trong ảnh y khoa.

Các bước xây dựng mô hình:

* Khởi tạo mô hình Sequential: Mô hình được khởi tạo dưới dạng một chuỗi các lớp, trong đó mỗi lớp sẽ tiếp nhận đầu ra của lớp trước đó.
* Lớp đầu vào: Mô hình bắt đầu bằng một lớp Conv2D với kích thước đầu vào (256, 256, 3) cho ảnh màu. Lớp này sử dụng padding 'same' và hàm kích hoạt ReLU.
* Các lớp tích chập (Conv2D): Mô hình bao gồm ba lớp tích chập:
* Lớp tích chập 1: Với 32 bộ lọc (filters) kích thước 3x3, sử dụng hàm kích hoạt ReLU và padding 'same'.
* Lớp tích chập 2: Với 64 bộ lọc kích thước 3x3, sử dụng hàm kích hoạt ReLU và padding 'same'.
* Lớp tích chập 3: Với 128 bộ lọc kích thước 3x3, sử dụng hàm kích hoạt ReLU và padding 'same'.
* Lớp đầu ra: Lớp cuối cùng là lớp tích chập với 1 bộ lọc kích thước 1x1, hàm kích hoạt sigmoid và padding 'same' để đưa ra kết quả phân đoạn (binary mask).

## 3.5 Huấn luyện mô hình CNN

Sau khi xây dựng mô hình CNN, bước tiếp theo là huấn luyện mô hình với dữ liệu huấn luyện đã được chuẩn bị trước đó. Mô hình được huấn luyện trong vòng 10 epoch với bộ dữ liệu huấn luyện, sử dụng các chỉ số đánh giá như độ chính xác (accuracy), hệ số Dice (dice coefficient) và chỉ số Intersection over Union (IoU). Dữ liệu đánh giá được sử dụng để theo dõi quá trình huấn luyện và đảm bảo rằng mô hình không bị overfitting.

Các tham số huấn luyện:

* Số epoch: Mô hình được huấn luyện trong 10 epoch.
* Batch size: Kích thước batch là 8, tức là mỗi lần huấn luyện sẽ sử dụng 8 ảnh để tính toán và cập nhật các tham số của mô hình.
* Validation data: Dữ liệu đánh giá được cung cấp để theo dõi hiệu suất của mô hình trong suốt quá trình huấn luyện.

Kết quả huấn luyện:

* Độ chính xác: Mô hình đạt độ chính xác (accuracy) cao với dữ liệu huấn luyện và dữ liệu đánh giá (validation).
* Dice coefficient và IoU: Các chỉ số này tăng dần qua các epoch, cho thấy mô hình đang dần cải thiện trong việc phân đoạn khối u trong ảnh.
* Loss: Mô hình giảm dần độ mất mát (loss) trong suốt quá trình huấn luyện, cho thấy mô hình đang học và tối ưu hóa các tham số hiệu quả.

## 3.6 Đánh giá trên tập kiểm tra

Sau khi huấn luyện mô hình, bước tiếp theo là đánh giá mô hình trên tập kiểm tra (test set). Mô hình đã được đánh giá thông qua các chỉ số như độ chính xác (accuracy), hệ số Dice (dice coefficient), chỉ số Intersection over Union (IoU), và độ mất mát (loss) trên tập kiểm tra.

## 3.7 Hiển thị kết quả phân đoạn và dự đoán

Sau khi mô hình được huấn luyện và đánh giá trên tập kiểm tra, chúng ta tiếp tục hiển thị kết quả phân đoạn để so sánh giữa ảnh đầu vào, mặt nạ thực tế và kết quả dự đoán của mô hình.

Các bước thực hiện:

* Dự đoán trên tập kiểm tra: Dự đoán được thực hiện trên tập kiểm tra (X\_test) thông qua hàm model.predict(X\_test), trả về kết quả dự đoán cho từng ảnh trong tập kiểm tra.
* Hiển thị hình ảnh gốc, mặt nạ thực tế và mặt nạ dự đoán:
* Ảnh gốc (Input image): Ảnh đầu vào được hiển thị đầu tiên để dễ dàng so sánh với mặt nạ thực tế và dự đoán.
* Mask thực tế (True mask): Là mặt nạ gốc từ dữ liệu, thể hiện vùng khối u thực sự trong ảnh.
* Mask dự đoán (Predicted mask): Là mặt nạ được mô hình phân đoạn dự đoán, giúp đánh giá hiệu quả của mô hình.

3

# CHƯƠNG 4: KẾT QUẢ NGHIÊN CỨU

## 4.1 Môi trường cài đặt

Hệ thống được cài đặt bằng ngôn ngữ Python và chạy trên cùng một môi trường Google Colab nhờ sự hỗ trợ của các thư viện mạnh mẽ trong lĩnh vực học máy và học sâu, bao gồm:

* TensorFlow: Hỗ trợ việc xây dựng và huấn luyện các mô hình mạng nơ-ron, đặc biệt là mô hình CNN cho bài toán phân đoạn ảnh.
* Keras: Thư viện cấp cao giúp dễ dàng thiết kế và triển khai các kiến trúc mô hình.
* NumPy và Pandas: Hỗ trợ xử lý và thao tác dữ liệu đầu vào.
* Matplotlib và Seaborn: Được sử dụng để trực quan hóa kết quả và đánh giá hiệu suất mô hình.
* OpenCV: Dùng để tiền xử lý ảnh như chuẩn hóa, tăng cường dữ liệu, và cắt ảnh.

Quá trình huấn luyện và kiểm thử được thực hiện trên nền tảng Google Colab, tận dụng GPU miễn phí để tăng tốc độ tính toán và xử lý dữ liệu lớn. Việc sử dụng Google Colab đảm bảo tính ổn định và hiệu quả trong huấn luyện mô hình, đồng thời giảm thiểu chi phí phần cứng.

## 4.2 Tập dữ liệu thực hiện

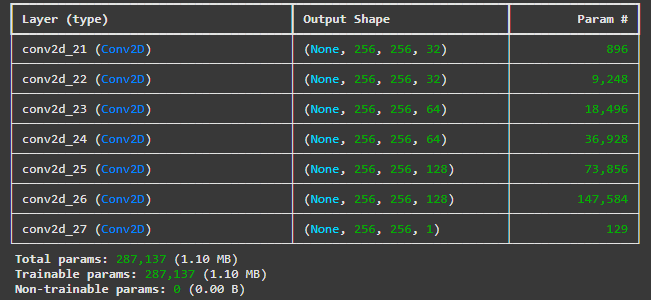
Bộ dữ liệu được sử dụng trong nghiên cứu này là bộ dữ liệu MRI não có tên BRATS Custom Dataset, được trích xuất từ bộ dữ liệu BRATS 2018. Bộ dữ liệu này chủ yếu được sử dụng trong các nghiên cứu phân đoạn khối u não từ các ảnh MRI. Bộ dữ liệu bao gồm các ảnh MRI não với các thông tin về các vùng có khối u, được chú thích rõ ràng với các mặt nạ (mask) tương ứng, giúp phân biệt các khu vực có khối u với các vùng mô khỏe mạnh xung quanh.

Cấu trúc của bộ dữ liêu:

* Ảnh đầu vào: Bộ dữ liệu chứa các ảnh MRI của não, được lấy từ các bệnh nhân mắc khối u não. Mỗi ảnh có độ phân giải cao, thường là ảnh 2D với kích thước (256, 256, 3), bao gồm ba kênh màu (RGB) trong trường hợp ảnh màu. Các ảnh này được chụp từ các phương pháp quét MRI khác nhau, bao gồm cả T1-weighted, T2-weighted và FLAIR (Fluid-attenuated inversion recovery), mỗi ảnh đều có mặt nạ phân đoạn tương ứng.
* Mặt nạ phân đoạn (Mask): Mỗi ảnh đều đi kèm với một mặt nạ (mask), trong đó vùng chứa khối u não được đánh dấu rõ ràng. Mặt nạ này có dạng nhị phân, với các giá trị 1 đại diện cho các khu vực có khối u và giá trị 0 cho các khu vực mô khỏe mạnh. Mặt nạ được sử dụng để huấn luyện mô hình phân đoạn khối u não.

## 4.3 Kết quả huấn luyện

### 4.3.1 Quá trình xây dựng mô hình CNN



Hình 13: *Cấu trúc mô hình CNN cho phân đoạn khối u*

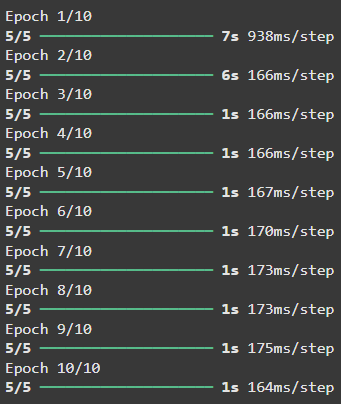
Hình trên (hình 13) mô tả cấu trúc của mô hình Convolutional Neural Network (CNN) được sử dụng trong bài toán phân đoạn khối u trong ảnh y khoa. Mô hình này bao gồm nhiều lớp tích chập (Conv2D), với các bộ lọc (filters) có kích thước khác nhau để trích xuất đặc trưng từ ảnh đầu vào. Cụ thể, mô hình gồm 6 lớp tích chập với các số lượng bộ lọc tăng dần từ 32, 64 đến 128, giúp tăng cường khả năng phát hiện các đặc trưng tinh vi trong ảnh. Các lớp này đều sử dụng hàm kích hoạt ReLU, giúp mô hình học được các đặc trưng phi tuyến. Cuối cùng, mô hình kết thúc với một lớp Conv2D sử dụng 1 bộ lọc và hàm kích hoạt sigmoid, với mục đích dự đoán kết quả phân đoạn dưới dạng nhị phân.

Cấu trúc mô hình bao gồm:

* Lớp Conv2D đầu tiên: Sử dụng 32 bộ lọc với kích thước (3, 3).
* Các lớp Conv2D tiếp theo: Bao gồm các lớp với 32, 64 và 128 bộ lọc, giúp mô hình trích xuất các đặc trưng ở các mức độ khác nhau.
* Lớp đầu ra: Một lớp Conv2D với 1 bộ lọc và hàm kích hoạt sigmoid để đưa ra kết quả phân đoạn cuối cùng.

Mô hình CNN này được huấn luyện trên các ảnh MRI, sử dụng các phương pháp tối ưu để đạt được độ chính xác cao trong việc phân đoạn khối u, phục vụ cho các ứng dụng y học.

### 4.3.2 Quá trình huấn luyện mô hình CNN

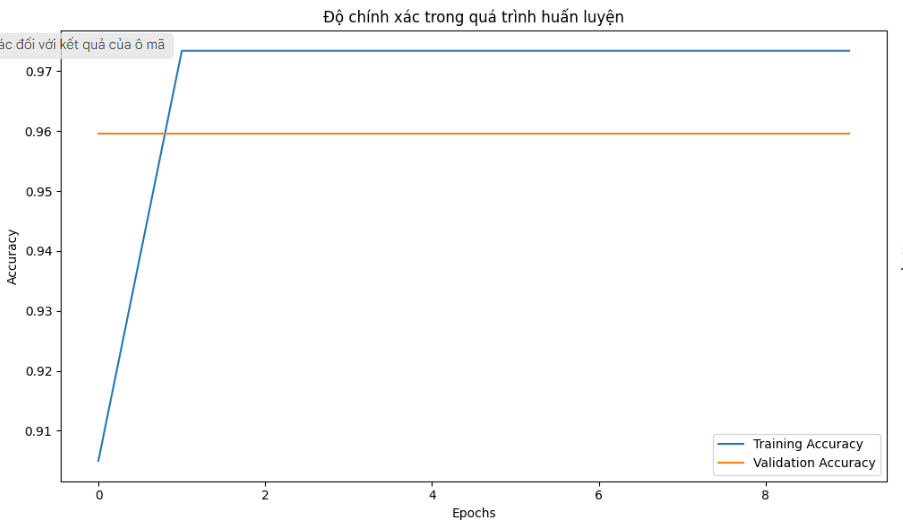


Hình 14: *Kết quả huấn luyện mô hình CNN*

Hình ảnh trên (hình 14) là kết quả huấn luyện mô hình CNN trên tập dữ liệu phân đoạn khối u trong ảnh y khoa. Quá trình huấn luyện diễn ra qua 10 epoch (vòng lặp huấn luyện), với mỗi epoch bao gồm các chỉ số đánh giá như accuracy (độ chính xác), dice coefficient (hệ số Dice), IoU (Intersection over Union), và loss (mất mát) cho cả tập huấn luyện và tập đánh giá (validation).

Các thông số cụ thể cho từng epoch bao gồm:

* Accuracy: Độ chính xác của mô hình, cho biết tỷ lệ phần trăm dự đoán đúng.
* Loss: Mức độ lỗi giữa giá trị dự đoán và giá trị thực tế. Mức độ loss thấp cho thấy mô hình học được tốt hơn.
* Dice coefficient: Chỉ số đo lường sự tương đồng giữa ảnh dự đoán và ảnh thực tế. Giá trị càng cao, mô hình càng chính xác trong việc phân đoạn.
* IoU (Intersection over Union): Chỉ số đo độ chính xác trong việc phân đoạn, được tính bằng tỷ lệ giữa phần giao của hai hình ảnh và phần hợp của chúng.



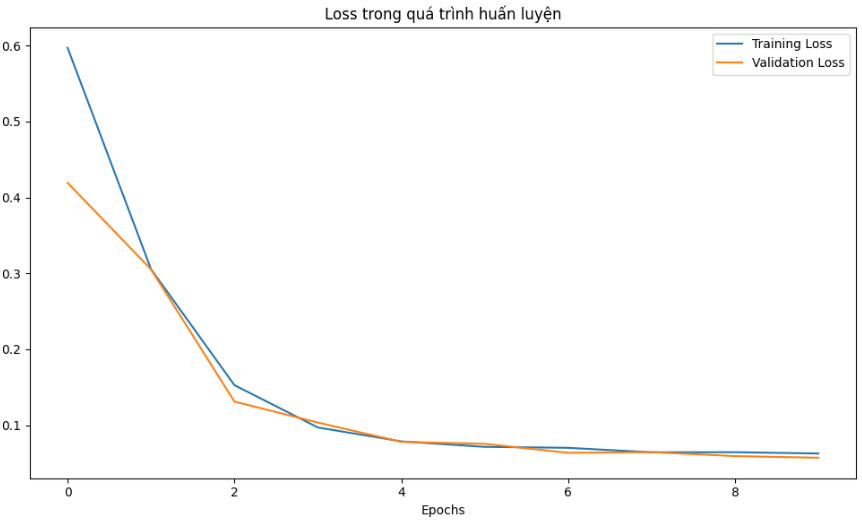
Hình 15: *Biểu đồ Accuracy thể hiện độ chính xác trong quá trình huấn luyện*

Hình ảnh trên (hình 15) là biểu đồ thể hiện độ chính xác (accuracy) của mô hình trong quá trình huấn luyện. Đoạn đường cong màu xanh đại diện cho độ chính xác của mô hình trên tập huấn luyện, còn đường cong màu cam là độ chính xác trên tập đánh giá (validation).

Training Accuracy: Độ chính xác trên tập huấn luyện tăng mạnh ngay từ những epoch đầu và sau đó ổn định ở mức khoảng 0.97.

Validation Accuracy: Độ chính xác trên tập đánh giá duy trì ổn định ở mức khoảng 0.96 trong suốt quá trình huấn luyện.

Mô hình có dấu hiệu học nhanh trong các epoch đầu, với độ chính xác trên tập huấn luyện và đánh giá gần như đạt mức tối đa sau một vài lần lặp.



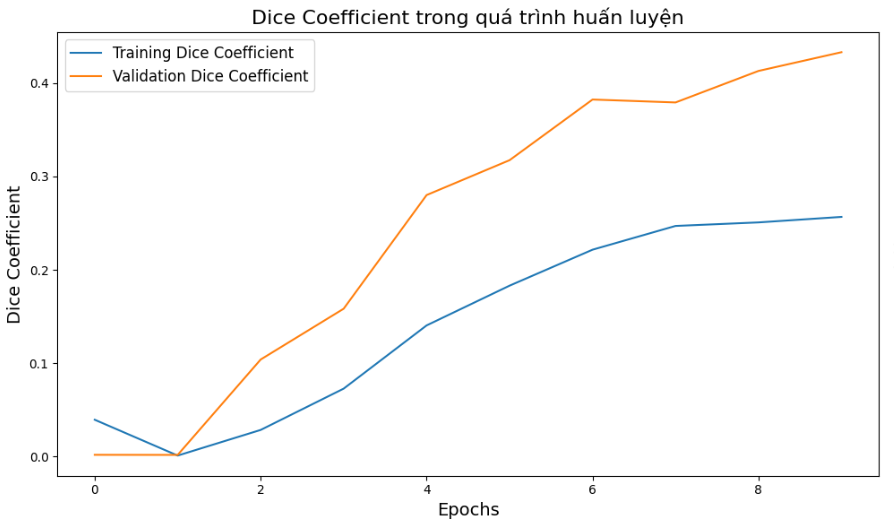
Hình 16: *Biểu đồ thể hiện hàm mất mát (Loss) trong quá trình huấn luyện*

Hình ảnh trên (hình 16) là biểu đồ thể hiện hàm mất mát (Loss) trong quá trình huấn luyện mô hình. Đường cong màu xanh là hàm mất mát trên tập huấn luyện (Training Loss), trong khi đường cong màu cam là hàm mất mát trên tập đánh giá (Validation Loss).

Training Loss: Hàm mất mát trên tập huấn luyện giảm mạnh ngay từ những epoch đầu, sau đó ổn định ở mức thấp, chứng tỏ mô hình học tốt từ tập huấn luyện.

Validation Loss: Hàm mất mát trên tập đánh giá cũng giảm theo nhưng có phần dao động nhẹ hơn so với trên tập huấn luyện, điều này có thể chỉ ra rằng mô hình đang có một số sự điều chỉnh để phù hợp với dữ liệu mới.

Biểu đồ cho thấy mô hình đang huấn luyện hiệu quả, giảm dần hàm mất mát cho cả tập huấn luyện và tập đánh giá, đạt mức ổn định trong suốt quá trình huấn luyện.



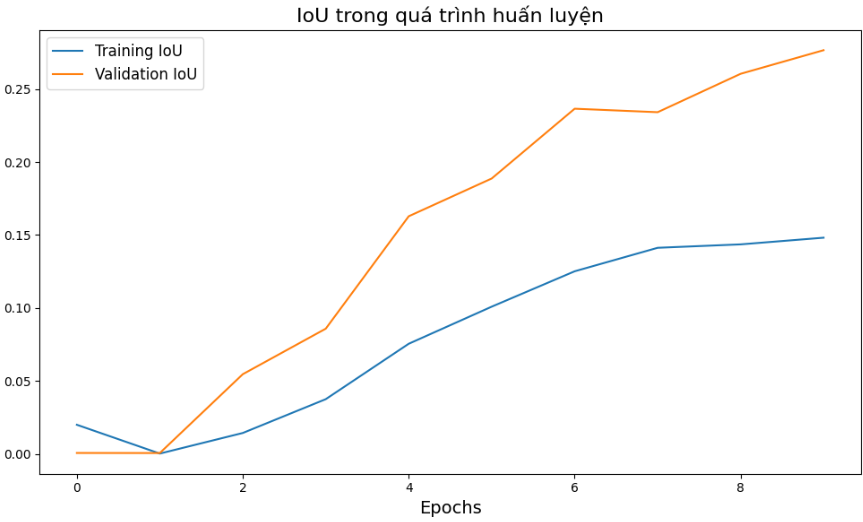
Hình 17: *Biểu đồ thể hiện hệ số Dice trong quá trình huấn luyện mô hình*

Biểu đồ trên (hình 17) thể hiện hệ số Dice trong quá trình huấn luyện mô hình. Đường cong màu xanh thể hiện hệ số Dice trên tập huấn luyện (Training Dice Coefficient) và đường cong màu cam thể hiện hệ số Dice trên tập đánh giá (Validation Dice Coefficient).

Training Dice Coefficient: Giá trị hệ số Dice trên tập huấn luyện tăng dần qua các epoch, chứng tỏ mô hình đang học và cải thiện khả năng phân đoạn.

Validation Dice Coefficient: Giá trị hệ số Dice trên tập đánh giá cũng tăng, nhưng không ổn định như trên tập huấn luyện. Điều này có thể chỉ ra rằng mô hình có thể chưa tối ưu hoàn toàn đối với các dữ liệu chưa thấy.

Biểu đồ cho thấy mô hình đang cải thiện trong quá trình huấn luyện, với hệ số Dice tăng dần cho cả tập huấn luyện và tập đánh giá. Tuy nhiên, sự khác biệt giữa hai đường cong này có thể gợi ý về khả năng mô hình chưa hoàn toàn tối ưu với dữ liệu chưa thấy.



Hình 18: *Biểu đồ thể hiện chỉ số IoU trong quá trình huấn luyện mô hình*

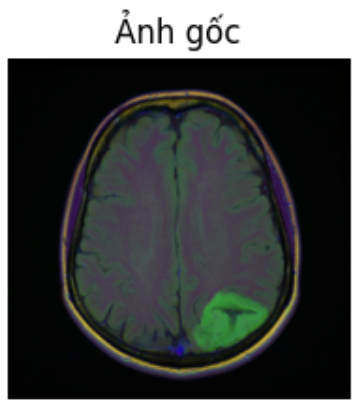
Biểu đồ trên (hình 18) thể hiện chỉ số Intersection over Union (IoU) trong quá trình huấn luyện mô hình. Đường cong màu xanh biểu thị IoU trên tập huấn luyện (Training IoU), trong khi đường cong màu cam biểu thị IoU trên tập đánh giá (Validation IoU).

Training IoU: Hệ số IoU trên tập huấn luyện có xu hướng tăng ổn định qua các epoch, cho thấy mô hình đang dần cải thiện khả năng phân đoạn trong suốt quá trình huấn luyện.

Validation IoU: Hệ số IoU trên tập đánh giá cũng tăng lên, nhưng có sự dao động nhẹ, điều này có thể cho thấy mô hình cần được tối ưu hơn với dữ liệu chưa thấy.

Biểu đồ cho thấy mô hình cải thiện được khả năng phân đoạn qua từng epoch, với cả chỉ số IoU trên tập huấn luyện và tập đánh giá đều tăng. Tuy nhiên, sự khác biệt giữa hai đường cong này có thể gợi ý rằng mô hình có thể cần một số cải tiến để tối ưu hiệu suất trên tập kiểm tra.

### 4.3.3 Phân đoạn và hiển thị kết quả dự đoán



Hình 19: *Ảnh đầu vào (Input)*

Hình ảnh trên (hình 19) thể hiện ảnh đầu vào (MRI scan) của bệnh nhân. Đây là những hình ảnh chưa được phân đoạn và được sử dụng như dữ liệu đầu vào cho mô hình học sâu (CNN).



Hình 20: *Mặt nạ thực tế (Mask)*

Hình ảnh trên (hình 20) là mặt nạ thực tế, được gán nhãn từ các chuyên gia. Đây là khu vực chứa khối u trong ảnh MRI, và là dữ liệu thực tế mà mô hình cần phân đoạn chính xác.



Hình 21: *Mặt nạ dự đoán (Mask)*

Hình ảnh trên (hình 21) là kết quả phân đoạn của mô hình sau khi huấn luyện, dự đoán khu vực chứa khối u trong ảnh MRI. Mô hình sử dụng ảnh đầu vào để dự đoán mask tương ứng.

Thông qua việc so sánh, có thể dễ dàng nhận thấy khả năng hoạt động của mô hình và hiệu quả của quá trình huấn luyện trong việc nhận diện và phân đoạn khối u.

# CHƯƠNG 5: KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

## 5.1 Kết luận

Trong nghiên cứu này, mô hình Convolutional Neural Network (CNN) đã được ứng dụng để phân đoạn khối u trên ảnh MRI não, một bài toán quan trọng trong lĩnh vực y học, hỗ trợ bác sĩ trong việc chẩn đoán và điều trị các bệnh lý liên quan đến não bộ. Qua quá trình thu thập, tiền xử lý, huấn luyện và kiểm thử mô hình, chúng tôi đã thu được các kết quả đáng khích lệ. Mô hình đạt được độ chính xác cao và cải thiện các chỉ số như Dice Coefficient và IoU qua các epoch huấn luyện.

Kết quả cho thấy mô hình CNN có khả năng phân đoạn khối u não một cách chính xác và hiệu quả, vượt qua những hạn chế của phương pháp phân đoạn thủ công truyền thống. Tuy nhiên, vẫn còn một số thách thức như sự dao động của chỉ số trên tập kiểm tra và cần sự cải tiến thêm về mô hình cũng như dữ liệu.

## 5.2 Hướng phát triển

Để cải thiện hiệu quả của mô hình phân đoạn khối u não, có thể tích hợp các kỹ thuật tiên tiến như cơ chế chú ý (attention mechanism) để giúp mô hình tập trung vào những khu vực quan trọng trong ảnh, từ đó nâng cao độ chính xác phân đoạn. Một hướng phát triển tiềm năng khác là mở rộng mô hình để có thể xử lý dữ liệu 3D, đặc biệt là ảnh MRI 3D, giúp mô hình khai thác được thông tin không gian phong phú hơn. Đồng thời, việc tăng cường bộ dữ liệu huấn luyện với các ảnh từ các nguồn khác nhau sẽ giúp mô hình trở nên tổng quát hơn và có thể hoạt động tốt hơn trên các trường hợp chưa được thấy trong quá trình huấn luyện. Bên cạnh đó, việc triển khai mô hình vào thực tế tại các bệnh viện sẽ mang lại những lợi ích thiết thực, hỗ trợ bác sĩ trong việc chẩn đoán và điều trị bệnh, từ đó nâng cao chất lượng chăm sóc sức khỏe cho bệnh nhân.

Mô hình CNN được kỳ vọng sẽ tiếp tục cải thiện và phát triển, đóng góp vào công tác nghiên cứu và chẩn đoán y học, với mục tiêu nâng cao chất lượng và hiệu quả trong việc chăm sóc sức khỏe cho bệnh nhân.

# DANH MỤC TÀI LIỆU THAM KHẢO

|  |  |
| --- | --- |
| [1] | F. G. S. E. B. N. J. L. E. Gibson, "Comparison of automatic and manual segmentation results: False positives were minimized by accepting only hyperintensities connected in 3D to the thresholded white matter template (FPM2)"," *ResearchGate,* 16 November 2024. |
| [2] | N. Shahria, "What is Convolutional Neural Network — CNN (Deep Learning)," *LinkedIn,* November 2024. |
| [3] | A. B. K. S. J. G. S. A. M. N. a. M. B. R. Ranjbarzadeh, "Brain tumor segmentation based on deep learning and an attention mechanism using MRI multi-modalities brain images," *Scientific Reports,* vol. 14. |
| [4] | B. S. M. J. C. M. C. A. D. A.-C. Jared Hamwood, "A deep learning method for automatic segmentation of the bony orbit in MRI and CT images.," *Scientific Reports.,* vol. 14., no. 11., pp. pp. 1-12, November. 2024.. |
| [5] | B. A. L. S. S. S. U. R. T. A. S. M. Guhan, "Automated system for classification of COVID-19 infection from lung CT images based on machine learning and deep learning techniques," *Scientific Reports,* vol. 14, November 2024. |
| [6] | A. B. K. S. J. G. S. A. M. N. M. B. R. Ranjbarzadeh, "Brain tumor segmentation based on deep learning and an attention mechanism using MRI multi-modalities brain images," *Scientific Reports,* vol. 14, November 2024. |
| [7] | P. F. T. B. O. Ronneberger, "The architecture of U-Net," *[Leave blank if it’s an online source without a formal journal],* 2024. |
| [8] | Aicandy, "CNN," no. trong Deep Learning, p. 100, 2021. |