

**TRƯỜNG ĐẠI HỌC SƯ PHẠM KỸ THUẬT TP. HỒ CHÍ MINH**

**KHOA ĐIỆN - ĐIỆN TỬ  
BỘ MÔN KỸ THUẬT ĐIỆN TỬ**



**HCMUTE**

**BÁO CÁO CHUYÊN ĐỀ 1**

**KỸ THUẬT NHẬN ĐIỆN PHỔ TÍN HIỆU CHO 5G VÀ  
LTE DỰA TRÊN CÁC PHƯƠNG PHÁP HỌC SÂU**

**NGÀNH KỸ THUẬT ĐIỆN TỬ**

Học viên: **Nguyễn Duy Huân**  
MSHV: 2390703

Hướng dẫn: TS. **Huỳnh Thế Thiện**

TP. HỒ CHÍ MINH, 08/2024

TRƯỜNG ĐẠI HỌC SƯ PHẠM KỸ THUẬT TP. HỒ CHÍ MINH  
KHOA ĐIỆN - ĐIỆN TỬ  
**BỘ MÔN KỸ THUẬT ĐIỆN TỬ**



**HCMUTE**

**BÁO CÁO CHUYÊN ĐỀ 1**

**KỸ THUẬT NHẬN ĐIỆN PHỔ TÍN HIỆU CHO 5G VÀ  
LTE DỰA TRÊN CÁC PHƯƠNG PHÁP HỌC SÂU**

**NGÀNH KỸ THUẬT ĐIỆN TỬ**

Sinh viên: **Nguyễn Duy Huân**  
MSHV: 2390703

Hướng dẫn: TS. **Huỳnh Thế Thiện**

TP. HỒ CHÍ MINH, 08/2024

# LÝ LỊCH KHOA HỌC

## Thông Tin Cá Nhân

- Họ và Tên:** NGUYỄN DUY HUÂN
- Ngày Sinh:** 29/30/2001
- Địa chỉ:** Khối phố 1, An Sơn, Tam Kỳ, Quảng Nam
- Email:** huan2931@gmail.com
- Điện Thoại:** +84 8660 78421

## Học Vấn

- Bằng cấp cao nhất:** Kỹ sư
- Nơi đào tạo:** Đại học Sư Phạm Kỹ Thuật thành phố Hồ Chí Minh
- Chuyên ngành:** Kỹ thuật máy tính
- Thời gian hoàn thành:** 2019-2023

## Kinh Nghiệm Làm Việc

- Tên công ty hoặc tổ chức:** Tập đoàn Công nghiệp – Viễn thông Quân đội (Viettel)
- Chức vụ:** Nhân viên
- Thời gian làm việc:** 05/2024 đến nay

Người khai thông tin  
(Ký và ghi rõ họ tên)

Nguyễn Duy Huân

## LỜI CẢM ƠN

Trước tiên, xin được thể hiện lòng biết ơn sâu sắc đến TS. Huỳnh Thế Thiện, giảng viên tại trường đại học Sư Phạm Kỹ Thuật thành phố Hồ Chí Minh, vì những đề xuất và hỗ trợ quý báu của thầy đã giúp tôi có đủ điều kiện để thực hiện đề tài này. Hơn nữa, những kinh nghiệm và nghiên cứu khoa học của thầy trong lĩnh vực học máy và trí tuệ nhân tạo cung cấp cho tôi những ý tưởng và cơ hội để hoàn thành tốt đề tài này. Xin chân thành cảm ơn các quý thầy cô giảng viên tại trường đại học Sư Phạm Kỹ Thuật thành phố Hồ Chí Minh, đặc biệt là các thầy cô trong ngành kỹ thuật điện tử, khoa Điện - Điện Tử, đã giảng dạy và hỗ trợ tôi trong suốt quá trình học tập tại trường. Những đóng góp quý báu của các thầy cô đã góp phần quan trọng vào sự thành công của đề tài này.

Người thực hiện luận văn  
(Ký và ghi rõ họ tên)

Nguyễn Duy Huân

## **LỜI CAM ĐOAN**

Người thực hiện đề tài cam đoan đề tài thực hiện dựa vào một số tài liệu trước đó và không sao chép nội dung, kết quả của các nghiên cứu trước đó. Các nội dung tham khảo đã được trích dẫn đầy đủ.

Người thực hiện luận văn  
(*Ký và ghi rõ họ tên*)

Nguyễn Duy Huân

## TÓM TẮT NỘI DUNG

Trong lĩnh vực truyền thông không dây, việc cảm nhận phổ đóng một vai trò quan trọng trong việc phát triển các mạng khôn dây thể hệ mới, cho phép các thiết bị nhận diện và đáp ứng nhu cầu kết nối không dây. Trong nghiên cứu này, chúng tôi tập trung vào giải quyết những thách thức của cảm nhận phổ và giới thiệu một mô hình mạng sâu mới được đặt tên là SpecSenseNet, nhằm giải quyết vấn đề này trong việc cảm nhận thông tin truyền tải qua không khí. Mô hình của chúng tôi dựa trên kiến trúc của Unet nhưng tích hợp các kỹ thuật tiên tiến như depthwise separable convolution, recurrent residual block và Atrous Spatial Pyramid Pooling, nhằm giảm thiểu độ phức tạp của mạng mà vẫn duy trì hiệu suất dự đoán cao. Chúng tôi sử dụng một bộ dữ liệu tín hiệu tổng hợp gồm hai loại tín hiệu chính (5G new radio và LTE - long-term evolution) được điều chỉnh cho các mức độ nhiễu bổ sung khác nhau.

# Mục lục

<b>1</b>	<b>TỔNG QUAN</b>	<b>1</b>
1.1	Đặt vấn đề . . . . .	1
1.2	Các nghiên cứu liên quan . . . . .	3
1.3	Lý do chọn đề tài nghiên cứu . . . . .	4
1.4	Mục tiêu và nhiệm vụ của nghiên cứu . . . . .	5
1.5	Phương pháp nghiên cứu . . . . .	6
1.6	Bố cục chuyên đề 1 . . . . .	7
<b>2</b>	<b>CƠ SỞ LÝ THUYẾT</b>	<b>8</b>
2.1	Học sâu và thị giác máy tính . . . . .	8
2.1.1	Giới thiệu chung . . . . .	8
2.1.2	Mạng nơ-ron nhân tạo . . . . .	9
2.1.3	Sự khác nhau giữa học sâu, học máy và trí tuệ nhân tạo . . . . .	10
2.2	Mô hình mạng CNN . . . . .	12
2.2.1	Giới thiệu chung . . . . .	12
2.2.2	Kiến trúc mạng CNN . . . . .	13
2.2.3	Phép tích chập (convolution) . . . . .	14
2.3	Phương pháp phân vùng ảnh . . . . .	15
2.3.1	Giới thiệu về phương pháp phân vùng ảnh . . . . .	15
2.3.2	Một số phương pháp phân vùng ảnh . . . . .	16
2.3.3	Bài toán phân loại hình ảnh . . . . .	17

2.4	Mô hình tín hiệu 5G NR và LTE . . . . .	18
<b>3</b>	<b>GIỚI THIỆU MẠNG HỌC SÂU ỨNG DỤNG TRONG PHÂN VÙNG</b>	
	<b>ẢNH</b>	<b>21</b>
3.1	Giới thiệu về mạng học sâu U-net . . . . .	21
3.2	Mô hình mạng SegNet . . . . .	23
3.3	Mô hình mạng SqueezeSegV2 . . . . .	24
3.4	Đánh giá ưu và nhược điểm . . . . .	25
3.5	Tập dữ liệu khảo sát . . . . .	26
3.6	Kết luận . . . . .	27
<b>4</b>	<b>KẾT LUẬN VÀ ĐỊNH HƯỚNG PHÁT TRIỂN CHO CHUYÊN ĐỀ 2</b>	<b>29</b>
4.1	Kết Luận . . . . .	29
4.2	Hướng phát triển . . . . .	30



# Danh sách bảng

1.1	Các thách Thức đặt ra trong lĩnh vực Cảm Biến Phổ . . . . .	2
1.2	Các nghiên cứu liên quan . . . . .	3
3.1	Thông tin về dữ liệu đầu vào cho mỗi phân loại . . . . .	27

# Danh sách hình vẽ

2.1	Hình ảnh minh họa về Deep learning (nguồn Internet). . . . .	8
2.2	Mô hình mạng Nơ-Ron (nguồn Internet). . . . .	10
2.3	Mô hình phân cấp AL, ML, DL (nguồn Internet). . . . .	11
2.4	Deep learning và Machine learning (nguồn Internet). . . . .	12
2.5	Kiến trúc mạng CNN (nguồn Internet). . . . .	13
2.6	Phép tính tích chập trên ma trận ảnh (nguồn Internet). . . . .	15
2.7	Ví dụ về phân đoạn ảnh (nguồn Internet). . . . .	16
2.8	Phân loại vật thể trong ảnh (nguồn Internet). . . . .	17
2.9	Mẫu phổ tín hiệu: (a) LTE, (b) 5G NR, (c) Chồng phổ 5G NR và LTE . . . . .	19
3.1	Kiến trúc mạng U-net. . . . .	22
3.2	Kiến trúc mạng SegNet (nguồn Internet). . . . .	23
3.3	Kiến trúc mạng SqueezeSegV2 (nguồn Internet). . . . .	24
3.4	Dữ liệu giao thoa phổ 5G NR và LTE với SNR khác nhau: (a) 0 dB, (b) 5 dB, (c) 10 dB, (d) 15 dB, (e) 20 dB, (f) 25 dB, (g) 30 dB. . . . .	26
3.5	Dữ liệu phổ 5G với SNR khác nhau: (a) 0 dB, (b) 5 dB, (c) 10 dB, (d) 15 dB, (e) 20 dB, (f) 25 dB, (g) 30 dB. . . . .	26
3.6	Dữ liệu phổ LTE với SNR khác nhau: (a) 0 dB, (b) 5 dB, (c) 10 dB, (d) 15 dB, (e) 20 dB, (f) 25 dB, (g) 30 dB. . . . .	26

# Danh mục các từ viết tắt

Dưới đây là danh mục các từ viết tắt được sử dụng trong luận văn.

Các từ viết tắt	Định nghĩa
NR	new radio
LTE	long-term evolution
DL	deep learning
SpecSenNet	Spectrum Sensing Network
DSC	depthwise separable convolution
RRC	recurrent residual convolution
ASPP	Atrous Spatial Pyramid Pooling
ML	machine learning
AI	Artificial Intelligence
CNN	convolutional neural network
FCN	fully connected network
RX	received signal
TX	transmitted signal
AWGN	Additive white Gaussian noise
STFT	short-time Fourier transform
SNR	signal-to-noise ratio
Deconv	Deconvolution

# Chương 1

## TỔNG QUAN

### 1.1 Đặt vấn đề

Trong kỉ nguyên công nghệ 4.0, trí tuệ nhân tạo đã mở ra nhiều hướng nghiên cứu mới trong lĩnh vực khoa học và công nghệ. Nhờ vào sự phát triển liên tục và mang tính cách mạng của lĩnh vực phần cứng, ngày càng nhiều phần cứng có khả năng tính toán siêu việt và giải quyết triệt để các giới hạn còn tồn đọng trong quá khứ. Điều này đã góp phần không nhỏ cho sự phát triển của các lĩnh vực nghiên cứu ứng dụng dựa trên phần cứng nói chung và lĩnh vực trí tuệ nhân tạo nói riêng. Cụ thể hơn, trong các ứng dụng sử dụng thị giác máy tính để xử lý và nhận diện hình ảnh ngày càng phổ biến hơn, phong trào nghiên cứu về lĩnh vực học sâu đang dần phát triển và được ứng dụng hiệu quả trong các lĩnh vực như xe tự hành, y tế, kết nối vạn vật, và các lĩnh vực liên quan về viễn thông.

Trong lĩnh vực truyền thông không dây, cảm biến phổ (spectrum sensing) được định nghĩa như quá là quá trình giám sát định kì trong một dải tần số cụ thể, nhằm xác định sự hiện diện của các loại sóng viễn thông và đưa ra thông tin dải tần số nào chưa được sử dụng. Khi mà các phương pháp phân bổ tần số truyền thông không thể cung cấp cho nhu cầu truyền tải dữ liệu cao một cách liên tục, điều này dẫn đến tình trạng lãng phí nguồn tài nguyên băng tần mà đáng lẽ ra có thể được

sử dụng để cấp phát cho các tín hiệu viễn thông khác như sóng radar, 5G NR, hoặc LTE. Do đó, cảm biến phổ trong hệ thống viễn thông trở thành đề tài thu hút lượng lớn các nhà nghiên cứu nhằm mục đích tăng cường tính hiệu quả trong việc nhận diện phổ tín hiệu.

<b>Thách Thức</b>	<b>Mô Tả</b>
<b>Yêu Cầu Phần Cứng</b>	Sự cần thiết của phần cứng tiên tiến và hiệu quả để thực hiện các nhiệm vụ cảm biến phổ một cách chính xác và nhanh chóng [1].
<b>Vấn Đề Người Dùng Chính Bị Ẩn</b>	Khó khăn trong việc phát hiện người dùng chính có thể bị ẩn hoặc có tín hiệu yếu, dẫn đến kết quả cảm biến không chính xác [1].
<b>Người Dùng Phổ Rộng</b>	Thách thức trong việc cảm nhận người dùng sử dụng các kỹ thuật phổ rộng, có thể gây khó khăn trong việc phát hiện và phân biệt [1].
<b>Hợp Nhất Quyết Định</b>	Sự phức tạp trong việc kết hợp các quyết định từ nhiều nút hoặc thuật toán cảm biến để cải thiện độ chính xác [1].
<b>Bảo Mật</b>	Đảm bảo an ninh của quá trình cảm biến phổ để ngăn chặn các cuộc tấn công ác ý và truy cập trái phép [1].
<b>Tần Số và Thời Gian Cảm Biến</b>	Xác định tần số và thời gian cảm biến tối ưu để cân bằng giữa độ chính xác và tiêu thụ tài nguyên [1].

Bảng 1.1: Các thách Thức đặt ra trong lĩnh vực Cảm Biến Phổ

Trong quá khứ, có rất nhiều kỹ thuật nhận diện phổ đã được giới thiệu được kể đến như cảm biến phổ dựa trên thuật toán, nhận diện phổ tín hiệu đa chiều, ước lượng kênh, cảm biến kết hợp. Tuy nhiên, các phương pháp này đều gặp những thử thách lớn khi mà phần cứng không thể đáp ứng về mặt hiệu năng, hầu hết các kỹ thuật đều yêu cầu phải lấy mẫu tín hiệu ở tần số cao cũng như độ phân giải cao, điều này dẫn đến yêu cầu về mặt phần cứng cũng cao hơn như là cần sử dụng bộ giải mã từ tín hiệu tương tự sang tín hiệu số với độ phân giải cao hơn, yêu cầu về

tốc độ của bộ xử lý để có thể lấy mẫu được tín hiệu ở tần số cao [1]. Theo hướng khác, cảm biến truyền thông không dây có thể được xác định dựa theo việc theo dõi thời lượng và tần số hoạt động [2]. Tuy nhiên, phương pháp này có một sự đánh đổi giữa hiệu suất và các thuật toán cảm nhận. Kể từ khi được giới thiệu, các nhà nghiên cứu đã dành nhiều nỗ lực để giải quyết các vấn đề cảm nhận phổ cải tiến, đối phó với nhiều thách thức và giới thiệu các giải pháp sáng tạo để tăng độ chính xác và hiệu suất trong các mạng truyền thông không dây nhận thức. Hơn nữa, các cuộc khảo sát nổi bật về chủ đề này được giới thiệu trong [3] và [4], trong đó hướng nghiên cứu và các giải pháp dựa trên máy học cho cảm nhận phổ thông minh được nghiên cứu và thảo luận một cách toàn diện.

## 1.2 Các nghiên cứu liên quan

Nghiên cứu liên quan	Hạn chế
DeepLabV3+ [5]	Trong đề tài nghiên cứu này, tác giả đã đưa ra được mô hình mạng học sâu nâng cấp từ mô hình DeepLabV3++ với độ chính xác nhận diện phổ trên 98%. Tuy nhiên chỉ duy trì được độ chính xác cao khi tín hiệu đầu vào có mức nhiễu SNR lớn hơn 60 dB.
ConvNet [6]	Trong đề tài nghiên cứu về tăng cường nhận diện phổ, tác giả Huỳnh Thế Thiện đã đưa ra cải tiến mạng học sâu ConvNet. Tuy nhiên độ chính xác của mạng học sâu khi nhận tín hiệu đầu vào với mức nhiễu SNR thấp thì không đáng kể.

Bảng 1.2: Các nghiên cứu liên quan

Trong những năm gần đây, nhiều công trình về truyền thông không dây nhận thức và mạng lưới radio đã được giới thiệu nhằm nâng cao hiệu suất cảm nhận. Nói chung, cảm nhận dựa trên bộ đo năng lượng là một phương pháp phổ biến

trong cảm nhận phổ vì nó cung cấp ít tính toán và độ phức tạp trong việc triển khai, cảm nhận dựa trên hình dạng sóng tăng cường hiệu suất cảm nhận, và cảm nhận dựa trên chu kỳ dành cho việc phát hiện và khớp bộ lọc với các tín hiệu của người dùng chính. Tuy nhiên, một số phương pháp học sâu (DL- deep learning) sáng tạo đã được giới thiệu để giải quyết hiệu suất cảm nhận phổ.

Việc áp dụng mạng nơ-ron tích chập để phân loại hình dạng sóng được cải thiện bởi tác giả Huỳnh Thế Thiện [7], công trình [6] giới thiệu một kiến trúc DL sáng tạo để cải thiện tỷ lệ dự đoán cảm nhận phổ cho tín hiệu 5G NR và LTE. Các công trình khác sử dụng deeplabv3+ [5] và DetectNet [8] để tăng độ chính xác cho cảm nhận phổ. Tuy nhiên, trong khi các thuật toán cảm nhận có sự đánh đổi giữa hiệu suất phần cứng và độ phức tạp tính toán, các mô hình DL cung cấp các tham số cao hơn để đạt được một mức độ chính xác cao đòi hỏi tài nguyên phần cứng đáng kể.

### **1.3 Lý do chọn đề tài nghiên cứu**

Lý do chọn đề tài này phản ánh sự nhận thức sâu sắc về vai trò quan trọng của trí tuệ nhân tạo và các công nghệ liên quan trong kỷ nguyên công nghệ 4.0. Trong một thời đại mà công nghệ đang phát triển mạnh mẽ và mang lại nhiều tiềm năng mới, việc tận dụng trí tuệ nhân tạo để giải quyết các vấn đề phức tạp là một hướng nghiên cứu cực kỳ hấp dẫn.

Cụ thể, việc áp dụng trí tuệ nhân tạo trong cảm biến phổ (spectrum sensing) trong hệ thống truyền thông không dây là một lựa chọn đáng chú ý. Điều này là do cảm biến phổ đóng vai trò quan trọng trong việc giám sát và quản lý tài nguyên tần số một cách hiệu quả, đặc biệt là trong bối cảnh ngày càng tăng của việc sử dụng tần số để hỗ trợ các dịch vụ truyền thông đa dạng như 5G NR, LTE, và các ứng dụng khác.

Bằng cách áp dụng các phương pháp và công nghệ mới nhất trong trí tuệ nhân

tạo và học sâu, ta có thể nâng cao khả năng cảm nhận và quản lý tài nguyên tần số, từ đó tối ưu hóa việc sử dụng và phân bổ tài nguyên một cách thông minh và hiệu quả. Việc này không chỉ giúp giảm thiểu lãng phí tài nguyên, mà còn tạo ra cơ hội phát triển mới trong các lĩnh vực ứng dụng của truyền thông không dây.

Do đó, việc nghiên cứu và phát triển các phương pháp cảm biến phổ thông minh và hiệu quả là một đề tài rất đáng quan tâm và có tiềm năng ứng dụng cao trong thời đại hiện nay.

## **1.4 Mục tiêu và nhiệm vụ của nghiên cứu**

Trong đề tài này, tác giả khai thác các kỹ thuật học sâu để giải quyết các nhiệm vụ cảm nhận và đề xuất một kiến trúc DL sáng tạo dựa trên U-net, được gọi là Spectrum Sensing Network (SpecSenseNet), giúp giảm đáng kể độ phức tạp mạng và nâng cao hiệu suất cảm nhận cho tín hiệu 5G NR và LTE [9, 10].

Với U-net là một mạng nơ-ron tích chập đầy đủ cho phân đoạn hình ảnh ngữ nghĩa hiệu quả, nó được thiết kế với một bộ mã hóa-giải mã song song và được kết nối với nhau thông qua các kết nối bỏ qua. Gần đây, U-net++ và các mô hình biến thể khác đã cho thấy sự cải thiện đáng kể về độ chính xác của phân đoạn; tuy nhiên, những mô hình này tiêu tốn bộ nhớ lớn và có chi phí tính toán đắt đỏ, không phù hợp cho các thiết bị có tài nguyên hạn chế. Để giảm độ phức tạp mạng mà không làm giảm độ chính xác của phân đoạn, tác giả giới thiệu SpecSenseNet, trong đó kiến trúc của nó tích hợp các depthwise separable convolutions [11] (DSC) và recurrent residual convolution [12–14] (RRC) phần vào từng lớp của đường dẫn mã hóa và giải mã. Mặt khác, các mô-đun Atrous Spatial Pyramid Pooling (ASPP) [15] được điều chỉnh để tăng cường việc học đặc điểm liên quan ở nhiều tỷ lệ khác nhau.

Tóm lại, trong đề tài này chủ yếu đóng góp vào hai điểm chính sau:

- Giải quyết vấn đề cảm biến phổ cho các vấn đề về 5G NR và LTE, đồng thời



sử dụng các phương pháp DL để giải quyết vấn đề cảm biến phổ.

- Giới thiệu kiến trúc SpecSenseNet cải tiến dựa trên U-net giúp giảm đáng kể độ phức tạp của mạng và nâng cao hiệu suất cảm biến cho tín hiệu 5G NR và LTE.

## 1.5 Phương pháp nghiên cứu

Trong đề tài này, tác giả chọn phương pháp nghiên cứu khoa học phân tích tổng kết dựa trên kinh nghiệm và thực nghiệm thông qua việc mô phỏng bằng công cụ MatLab. Việc này mang lại một cơ sở vững chắc để hiểu rõ hơn về hiệu suất và tính khả thi của mô hình đề xuất.

- Phương pháp phân tích và đánh giá dựa trên kinh nghiệm:
  - Đánh giá độ hiệu quả và so sánh các mô hình DL trong bài toán phân loại vật thể dựa trên việc phân vùng điểm ảnh.
  - Dựa vào các kết quả đã phân tích được trên các mô hình DL trước đó, tác giả tiến hành phân tích về ưu nhược điểm của các mô hình DL, đánh giá các điểm chưa được tối ưu và đề xuất ra các phương pháp mới để thiết kế lại mô hình DL sao cho tăng tối đa tốc độ tính toán và đạt được sự chính xác cao.
- Phương pháp thực nghiệm dựa trên mô phỏng sử dụng công cụ MatLab:
  - Tiến hành thiết kế lại các mô hình DL đã được giới thiệu trước đó trên công cụ mô phỏng MatLab và đánh giá hiệu năng của từng mô hình.
  - Phân tích ưu nhược điểm giữa các mô hình và đưa ra cải tiến đối mới trong thiết kế mô hình DL nhằm tăng hiệu suất của mô hình DL dựa trên các phương pháp thiết kế mạng DL tiên tiến.

- Tiến hành đánh giá mô hình đề suất dựa trên tập dữ liệu phổ tín hiệu được sinh ra từ công cụ MatLab với tỉ lệ nhiễu cao nhằm đánh giá độ hiệu quả của mô hình đề suất, cùng với đó so sánh với các mô hình trước đó.

## **1.6   Bố cục chuyên đề 1**

Báo cáo chuyên đề 1 sẽ bao gồm 4 chương:

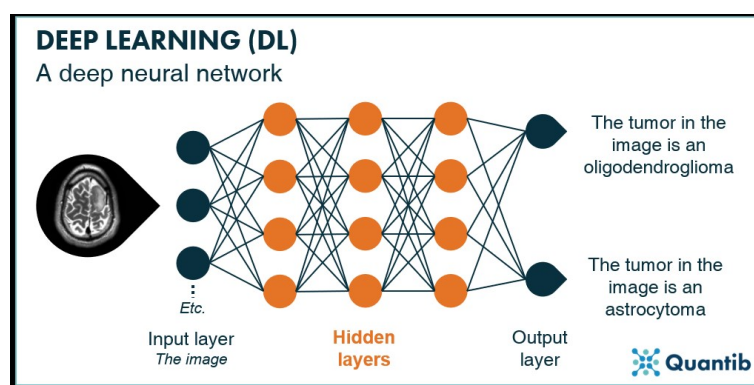
- Chương 1: Tổng Quan
- Chương 2: Cơ sở lý thuyết
- Chương 3: Giới thiệu mạng học sâu ứng dụng trong phân vùng ảnh
- Chương 4: Kết luận và định hướng phát triển cho chuyên đề 2

## Chương 2

# CƠ SỞ LÝ THUYẾT

## 2.1 Học sâu và thị giác máy tính

### 2.1.1 Giới thiệu chung



Hình 2.1: Hình ảnh minh họa về Deep learning (nguồn Internet).

Trong thời đại hiện nay, khả năng trí tuệ nhân tạo được cung cấp bởi các hệ thống thông minh dựa trên máy học (ML - machine learning) và học sâu đang ngày càng phổ biến. Máy học miêu tả khả năng của các hệ thống học tập từ lượng dữ liệu đào tạo để tự động hóa các quá trình phân tích mô hình và giải quyết các công việc liên quan [16]. Học sâu là một khái niệm dựa trên mạng Nơ-Ron nhân tạo và cho phép mô hình hóa các ứng dụng có yêu cầu cao hơn bằng cách sử dụng các cấu trúc mạng Nơ-Ron phức tạp hơn so với các mô hình học máy thông

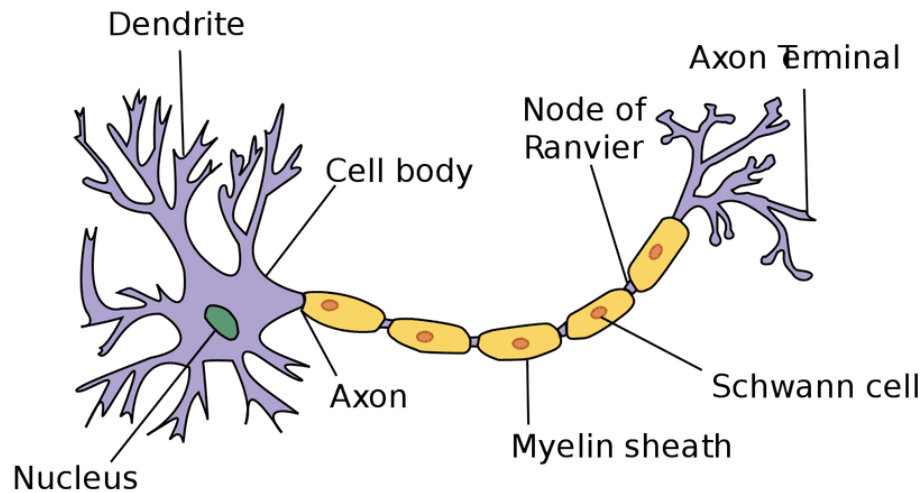
thường.

Học sâu (Deep Learning) là một phân nhánh của máy học và đã mang lại những tiến bộ đáng kể trong nhiều lĩnh vực nhờ vào khả năng xử lý lượng lớn dữ liệu và tự động hóa quá trình học tập. Các mô hình học sâu sử dụng mạng Nơ-Ron nhân tạo với nhiều lớp (multi-layer neural networks), cho phép chúng học được các đặc trưng phức tạp từ dữ liệu. Điều này giúp học sâu vượt trội trong các ứng dụng như nhận dạng hình ảnh, xử lý ngôn ngữ tự nhiên, và dự đoán xu hướng. Sự khác biệt quan trọng của học sâu so với các phương pháp khác là sử dụng mạng Nơ-Ron nhân tạo để tự động hóa việc đào tạo. Bằng cách sử dụng số lượng lớn dữ liệu đầu vào, học sâu có khả năng đào tạo mô hình tự động và tạo ra kết quả chính xác hơn so với phương pháp truyền thống của học máy.

Sự phát triển của AI, đặc biệt là học sâu, đã thay đổi cách chúng ta tiếp cận và giải quyết các vấn đề phức tạp. Từ chăm sóc sức khỏe, tài chính, đến giao thông và giải trí, AI đang dần trở thành một phần không thể thiếu, đem lại những cải tiến và tối ưu hóa vượt bậc cho nhiều lĩnh vực khác nhau. Nhờ vào AI, chúng ta có thể kỳ vọng vào một tương lai nơi công nghệ không chỉ hỗ trợ mà còn nâng cao chất lượng cuộc sống con người theo những cách chưa từng có.

### **2.1.2 Mạng nơ-ron nhân tạo**

Khái niệm Nơ-Ron là một đơn vị thần kinh được liên kết với hàng triệu đơn vị Nơ-Ron khác tạo nên một hệ thống thần kinh phức tạp và hoạt động bằng cách nhận dữ liệu từ đầu vào của nó và truyền đi cho các Nơ-Ron khác trong hệ thống. Ngày nay, DL dựa trên cách hoạt động của mạng Nơ-Ron thần kinh để xây dựng lên một hệ thống mạng thần kinh gồm các lớp Nơ-Ron nhân tạo và bắt chước theo các tính năng của nó để áp dụng vào trong lĩnh vực trí tuệ nhân tạo (AI – Artificial Intelligence). Từ đó mô hình mạng Nơ-Ron ra đời và giúp cho lĩnh vực AI phát triển một cách vượt bậc nhờ khả năng suy luận gần giống như hành vi của con người. Nói cách khác mô hình mạng Nơ-Ron nhân tạo cố gắng mô phỏng lại gần



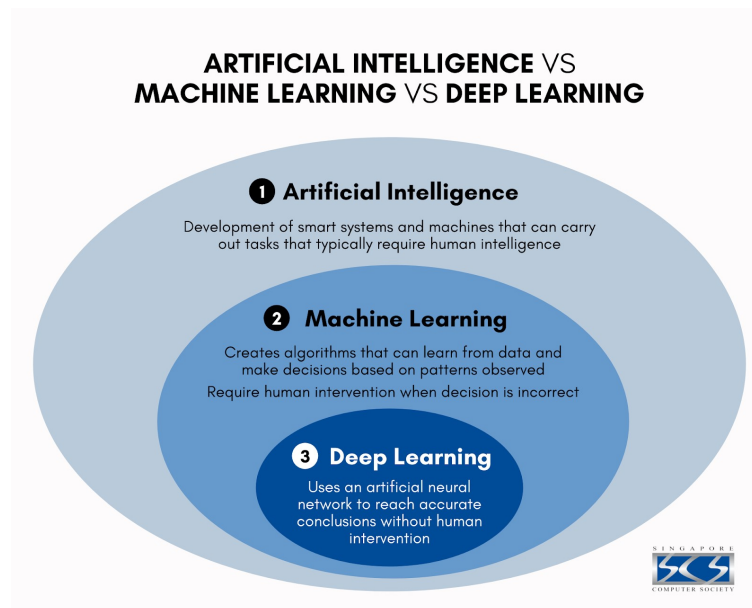
Hình 2.2: Mô hình mạng Nơ-Ron (nguồn Internet).

chính xác với quy trình xử lý thông tin như là bộ não của con người.

Trong mạng nơ-ron nhân tạo, các nơ-ron được tổ chức thành các lớp bao gồm lớp đầu vào (input layer), lớp ẩn (hidden layers) và lớp đầu ra (output layer). Mỗi lớp có nhiệm vụ chuyển đổi và truyền tải thông tin theo một cách thức nhất định. Nhờ vào quá trình học (training), trong đó mạng nơ-ron điều chỉnh các trọng số dựa trên dữ liệu đầu vào và đầu ra mong muốn, mạng nơ-ron nhân tạo có thể học hỏi từ dữ liệu và đưa ra các dự đoán hoặc quyết định giống như cách con người học từ kinh nghiệm. Điều này giúp chúng có khả năng giải quyết các vấn đề phức tạp như nhận dạng hình ảnh, xử lý ngôn ngữ tự nhiên và chơi trò chơi, đưa lĩnh vực trí tuệ nhân tạo tiến xa hơn bao giờ hết.

### 2.1.3 Sự khác nhau giữa học sâu, học máy và trí tuệ nhân tạo

Thay vì hệ thống hóa kiến thức vào máy tính, ML sẽ tìm cách tự động học các mối quan hệ và mẫu có ý nghĩa từ các ví dụ và quan sát. Những tiến bộ trong ML đã cho phép sự gia tăng gần đây của các hệ thống thông minh với khả năng nhận thức giống như con người, thâm nhập vào công việc kinh doanh và cuộc sống cá nhân của chúng ta, đồng thời định hình các tương tác được kết nối mạng trên thị trường điện tử theo mọi cách có thể.

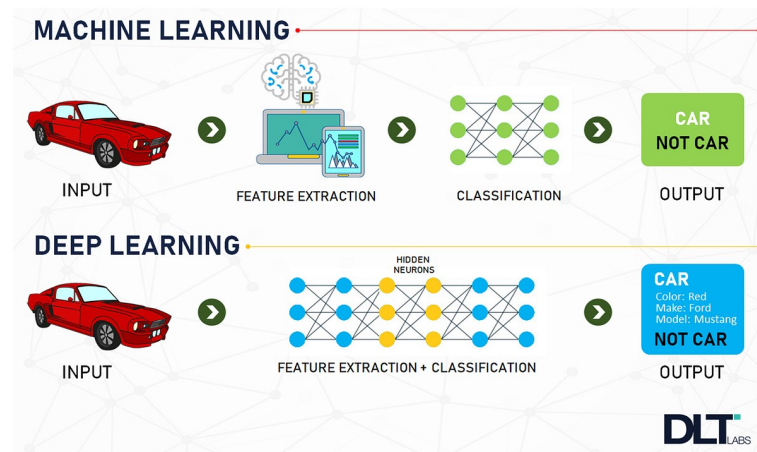


Hình 2.3: Mô hình phân cấp AL, ML, DL (nguồn Internet).

Các hệ thống ML và DL phức tạp hay còn được gọi là trí tuệ nhân tạo, nhờ khả năng xử lý và giải quyết các vấn đề phức tạp, phụ thuộc vào các mô hình phân tích và dự đoán các quy tắc để đưa ra câu trả lời hoặc đề xuất kết quả tương tự như con người. AI đã đóng góp vai trò quan trọng dựa trên việc giảm gánh nặng tính toán và dự đoán cho con người trong nhiều ứng dụng cụ thể. Tóm lại lĩnh vực trí tuệ nhân tạo phát triển phụ thuộc vào các mô hình DL và ML trong đó khái niệm về DL là một phương pháp được nâng cấp từ phương pháp học của ML trước đó bằng việc áp dụng mô hình mạng Nơ-Ron nhân tạo vào để xử lý các bài toán với cách thức tương tự như con người.

Từ đó lĩnh vực AI đã tạo ra nhiều tiến bộ trong ngành máy tính nói riêng và các lĩnh vực liên quan khác nói chung. Một trong những tiến bộ này là sự phát triển và ra đời của các mạng thần kinh Nơ-Ron nhân tạo từ đó khái niệm DL được ra đời [16]. Đối với các ứng dụng cụ thể như nhận diện vật thể hoặc phân loại đối tượng, việc ứng dụng DL đem lại nhiều hiệu quả vượt trội nhờ vào cách huấn luyện đặc biệt thông qua mạng thần kinh Nơ-Ron nhân tạo. Bằng cách trích xuất ra những đặc trưng của vật thể tại từng lớp mạng Nơ-Ron nhân tạo và điều chỉnh các trọng số sau mỗi chu kỳ huấn luyện cho đến khi tỷ lệ dự đoán đạt được là cao

nhất.



Hình 2.4: Deep learning và Machine learning (nguồn Internet).

Trong DL nguồn dữ liệu đầu vào là cơ sở giúp cho mô hình có thể bắt đầu quá trình học. Dữ liệu có thể là các đoạn văn bản, hình ảnh, biểu tượng và những dữ liệu đó được gọi là các dữ liệu chưa được xử lý. Khi dữ liệu đã được xử lý hay còn gọi là gán nhãn thì nó được gọi là các thông tin và DL sẽ dựa vào các thông tin đó để ghi nhớ hình dạng cũng như các đặc điểm của vật thể từ đó điều chỉnh các trọng số phù hợp để đưa ra các dự đoán chính xác. Từ sự phát triển của DL đã mở ra khả năng xử lý được nhiều vấn đề thực tế và đưa các hành vi của máy tính trở lại gần giống với cách con người suy luận. Từ đó lĩnh vực AL đã được mở rộng và bao quát hơn nhờ DL.

## 2.2 Mô hình mạng CNN

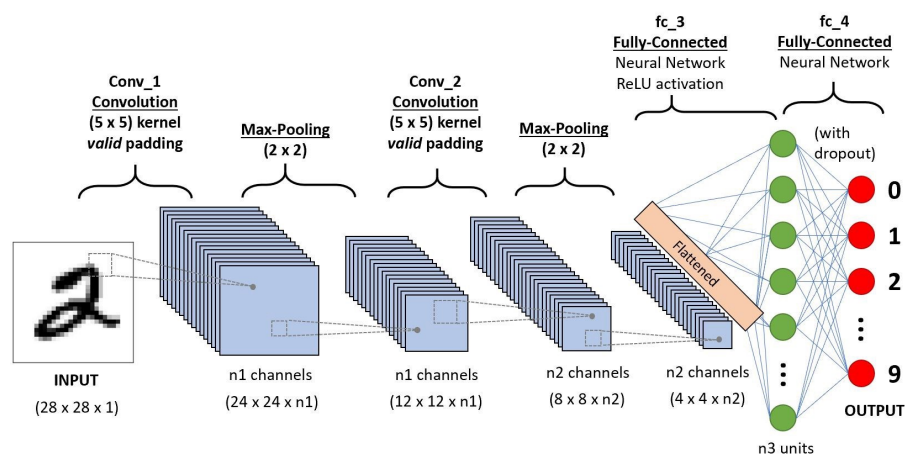
### 2.2.1 Giới thiệu chung

Mạng nơ-ron tích chập, hay còn gọi là CNN (convolutional neural networks), là một loại mô hình mạng thần kinh nhân tạo đặc biệt được thiết kế để xử lý dữ liệu dạng lưới như hình ảnh. Được áp dụng rộng rãi trong lĩnh vực nhận diện và phân loại hình ảnh, CNN đã trở thành công cụ quan trọng và không thể thiếu trong nhiều ứng dụng thực tế.

Điểm mạnh của CNN nằm ở khả năng học và trích xuất các đặc trưng quan trọng của hình ảnh thông qua việc sử dụng các lớp tích chập (convolutional layers). Mỗi lớp tích chập trong mạng CNN có khả năng phát hiện và nắm bắt các mẫu (patterns) khác nhau trong dữ liệu đầu vào, từ các đặc trưng cơ bản như cạnh và góc, đến các đặc trưng phức tạp hơn như hình dạng và kết cấu. Quá trình này diễn ra thông qua việc sử dụng các bộ lọc (filters) và phép tích chập, giúp CNN có khả năng phát hiện các đặc trưng không gian trong ảnh một cách hiệu quả. Ngoài ra, CNN còn sử dụng các lớp pooling để giảm kích thước không gian của các đặc trưng đã được trích xuất, giúp giảm thiểu số lượng tham số và tính toán trong mô hình, đồng thời tăng khả năng tổng quát hóa của mô hình. Các lớp kết nối toàn phần (fully connected) ở cuối mạng đảm nhận vai trò phân loại dựa trên các đặc trưng đã được trích xuất và tổng hợp qua các lớp trước đó.

Nhờ vào cấu trúc đặc biệt và khả năng học hỏi mạnh mẽ, CNN đã được ứng dụng trong nhiều lĩnh vực khác nhau, từ y tế, như phát hiện và chẩn đoán bệnh qua hình ảnh y học, đến các ứng dụng trong xe tự lái, nhận diện khuôn mặt, phân loại đối tượng trong hình ảnh, và thậm chí là trong nghệ thuật số và truyền thông.

### 2.2.2 Kiến trúc mạng CNN



Hình 2.5: Kiến trúc mạng CNN (nguồn Internet).



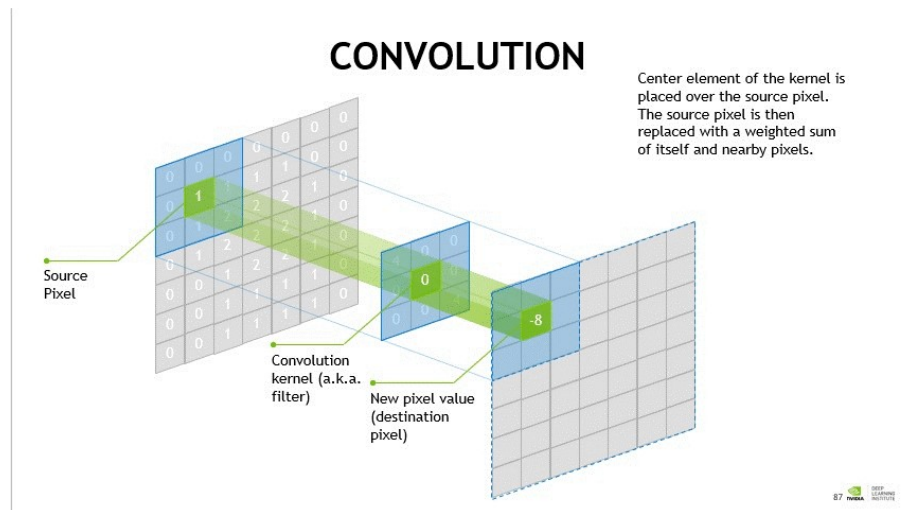
Kiến trúc của mạng CNN được xây dựng dựa trên các lớp cơ bản như lớp tích chập (Convolution layer), lớp kích hoạt (RELU layer), lớp tổng hợp (Pooling layer) và lớp kết nối đầy đủ (Fully connected layer).

- **Lớp tích chập (Convolution layer):** Lớp tích chập được coi là có vai trò quan trọng nhất trong kiến trúc mạng CNN, ở lớp này sẽ thực hiện các quá trình tính toán như nhân chập ảnh với mặt nạ (kernels), bộ lọc (filter), trích xuất đặc trưng (feature map). Trích xuất đặc trưng là yếu tố quan trọng nhất trong lớp tích chập.
- **Lớp Kích hoạt (RELU layer):** Lớp kích hoạt dùng để tính toán và mô phỏng lại tỷ lệ truyền xung qua axon của hệ Nơ-Ron thần kinh dùng các hàm kích hoạt (activation function) như leaky, mish, swish, v.v.
- **Lớp tổng hợp (Pooling layer):** Lớp tổng hợp sẽ lược bỏ đi các thông tin không cần thiết từ lớp tích chập từ đó cho ra kết quả mong muốn.
- **Lớp kết nối đầy đủ (Fully connected layer):** Tại lớp kết nối đầy đủ sẽ lấy kết quả đầu vào từ lớp trước đó và tiến hành nén thông tin ảnh về dưới dạng một véc tơ và dùng kết quả đó để làm đầu vào cho lớp kế tiếp.

### 2.2.3 Phép tích chập (convolution)

$$f(x, y) * k(x, y) = \sum_{u=-m/2}^{m/2} \sum_{v=-n/2}^{n/2} k(u, v) f(x - u, y - v), \quad (2.1)$$

Phép tích chập là một trong những phép toán quan trọng trong lĩnh vực xử lý ảnh, bằng cách sử dụng các mặt nạ để lướt trên từng điểm ảnh và tính tích chập tại từng điểm ảnh. Công thức tính phép tích chập tại một điểm ảnh  $f(x, y)$  với mặt nạ  $k(x, y)$  được biểu diễn tại (2.1). Phép tích chập có thể được hình dung như bằng cách dùng mặt nạ  $K$  dịch chuyển trên từng điểm ảnh từ góc bên của trái ảnh theo



Hình 2.6: Phép tính tích chập trên ma trận ảnh (nguồn Internet).

hàng ngang và theo hàng dọc khi đi tới điểm ngoài cùng bên phải của ảnh. Tương ứng với mỗi lần dịch chuyển sẽ thực hiện tính toán giá trị tại điểm ảnh đang xét bằng công thức tính phép tích chập.

## 2.3 Phương pháp phân vùng ảnh

### 2.3.1 Giới thiệu về phương pháp phân vùng ảnh

Phân đoạn hình ảnh [17] là một nhiệm vụ quan trọng trong thị giác máy tính và xử lý hình ảnh, liên quan đến việc chia nhỏ các hình ảnh thành các đoạn hoặc đối tượng riêng biệt. Quá trình này là nền tảng cho nhiều ứng dụng khác nhau như hiểu cảnh, phân tích hình ảnh y tế, nhận thức robot, giám sát video, thực tế tăng cường và nén hình ảnh.

Mục tiêu cơ bản của phân đoạn hình ảnh là đơn giản hóa hoặc thay đổi đại diện của một hình ảnh thành một cái gì đó có ý nghĩa hơn và dễ dàng phân tích hơn. Nó có thể được phân loại thành ba loại chính: phân đoạn ngữ nghĩa, phân đoạn đối tượng và phân đoạn tổng hợp. Phân đoạn ngữ nghĩa liên quan đến việc gán nhãn cho từng điểm ảnh của hình ảnh với một lớp tương ứng như 'xe hơi', 'cây cối' hoặc 'đường phố'. Phân đoạn đối tượng tiến xa hơn bằng cách xác định các đối tượng



Hình 2.7: Ví dụ về phân đoạn ảnh (nguồn Internet).

riêng lẻ, trong khi phân đoạn tổng hợp kết hợp cả phân đoạn ngữ nghĩa và phân đoạn đối tượng, cung cấp một sự hiểu biết toàn diện về cảnh.

### 2.3.2 Một số phương pháp phân vùng ảnh

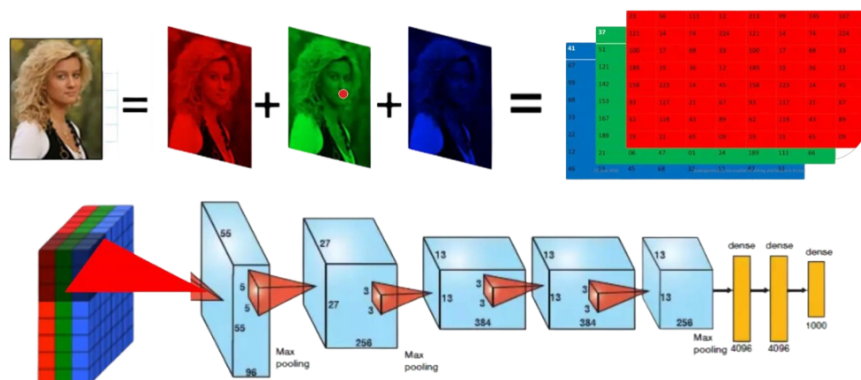
Trong quá khứ, nhiều phương pháp khác nhau đã được sử dụng cho phân đoạn hình ảnh [17], bao gồm phân ngưỡng, phát triển vùng, các phương pháp phân cụm như k-means, và các kỹ thuật tinh vi hơn như các đường viền hoạt động và cắt đồ thị. Tuy nhiên, sự xuất hiện của học sâu đã thay đổi đáng kể lĩnh vực này, giới thiệu các mô hình tận dụng các mạng nơ-ron tích chập (CNN), các kiến trúc mã hóa-giải mã, mạng nơ-ron hồi tiếp và mạng đối kháng tạo sinh (GAN) để đạt được kết quả tiên tiến nhất.

Các phương pháp phân đoạn hình ảnh dựa trên học sâu đã vượt trội hơn các phương pháp truyền thống, đạt được độ chính xác và hiệu quả cao hơn. Ví dụ, các mạng tích chập hoàn toàn (FCN - fully connected network) đã có vai trò then chốt trong cuộc cách mạng này, cho phép huấn luyện và dự đoán từ đầu đến cuối cho phân đoạn ngữ nghĩa. Các mạng này thay thế các lớp hoàn toàn kết nối trong các

CNN truyền thống bằng các lớp tích chập duy trì thông tin không gian, cho phép mạng tạo ra các dự đoán dày đặc cho từng điểm ảnh trong hình ảnh đầu vào. Ngoài FCN, nhiều mô hình tiên tiến khác đã được phát triển để giải quyết các thách thức cụ thể trong phân đoạn hình ảnh. Chúng bao gồm các mạng đa tỉ lệ và kim tự tháp để thu thập thông tin ngữ cảnh ở các độ phân giải khác nhau, các mạng nơ-ron hồi tiếp để mô hình hóa các phụ thuộc tuần tự, và các mô hình chú ý để tập trung vào các phần có liên quan của hình ảnh. Hơn nữa, GAN đã được sử dụng để nâng cao tính hiện thực và độ chính xác của các đầu ra phân đoạn bằng cách học cách tạo ra các phân đoạn chất lượng cao không phân biệt được so với dữ liệu thực tế.

Mặc dù đã có nhiều tiến bộ, phân đoạn hình ảnh vẫn đối mặt với các thách thức như cần có các bộ dữ liệu lớn được gán nhãn, sự kém hiệu quả về tính toán cho các ứng dụng thời gian thực, và khó khăn trong việc phân đoạn các đối tượng trong các cảnh phức tạp hoặc lộn xộn. Nghiên cứu đang diễn ra tiếp tục khám phá các giải pháp cho những vấn đề này, nhằm phát triển các phương pháp phân đoạn mạnh mẽ, hiệu quả và có thể tổng quát hơn.

### 2.3.3 Bài toán phân loại hình ảnh



Hình 2.8: Phân loại vật thể trong ảnh (nguồn Internet).

Bài toán phân đoạn hình ảnh (image segmentation) là một nhiệm vụ quan trọng

trong lĩnh vực thị giác máy tính và xử lý hình ảnh, và nó có vai trò nền tảng trong việc thực hiện bài toán phân loại (classification). Bài toán phân loại sẽ thực hiện nhiệm vụ quyết định đối tượng sẽ thuộc phân lớp nào trong tổng số phân loại đối tượng được nhận diện. Tuy nhiên, để có thể thực hiện chính xác bài toán phân loại, cần phải thực hiện phân đoạn hình ảnh trước đó.

Phân vùng hình ảnh chia hình ảnh thành các vùng hoặc đối tượng riêng biệt dựa trên các đặc điểm nhận dạng của chúng. Quá trình này giúp xác định vị trí và ranh giới của từng đối tượng trong hình ảnh. Sau khi phân đoạn, mỗi vùng hoặc đối tượng được nhận diện rõ ràng, từ đó có thể áp dụng các kỹ thuật phân loại để đưa ra dự đoán chính xác về phân lớp của từng đối tượng.

Bài toán phân loại sẽ dựa trên đặc điểm nhận dạng của các lớp khác nhau nhằm đưa ra dự đoán chính xác về vật thể [18]. Điều kiện tiên quyết để thực hiện được bài toán phân loại vật thể là phải nhận diện được các đối tượng trước từ đó mới dựa trên các đặc điểm của đối tượng để tiến hành phân loại đối tượng.

## 2.4 Mô hình tín hiệu 5G NR và LTE

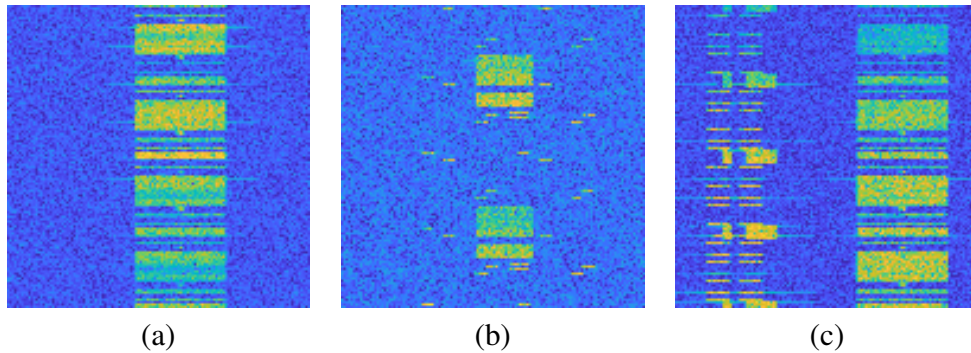
Ngày nay, các hệ thống liên lạc qua không trung đóng một vai trò vô cùng quan trọng trong nhiều kịch bản và ứng dụng đa dạng [19]. Công trình nghiên cứu này chủ yếu tập trung vào các hệ thống thông tin liên lạc 5G NR và LTE. Trên thực tế, các tín hiệu nhận được, hay còn gọi là tín hiệu RX (received signal), từ các nguồn khác nhau có thể được định nghĩa bằng phương trình (2.2). Trong đó,  $y(t)$  biểu thị cho tín hiệu RX,  $x(t)$  là tín hiệu được truyền đi, gọi là tín hiệu TX (transmitted signal),  $h(t)$  đại diện cho đáp ứng kênh, và  $n(t)$  biểu thị cho nhiễu trắng Gaussian cộng thêm (AWGN - Additive white Gaussian noise)

$$y(t) = x(t) * h(t) + n(t), \quad (2.2)$$

$$Y(\tau, w) = \int_{-\infty}^{\infty} y(t) \cdot w(t - \tau) \cdot e^{-j2\pi ft} dt, \quad (2.3)$$

$$E = \sum_{f=f_{\min}}^{f_{\max}} |Y(\tau, w)|^2. \quad (2.4)$$

Biến đổi Fourier ngắn hạn (STFT - short-time Fourier transform) là một kỹ thuật xử lý tín hiệu được sử dụng rộng rãi để phân tích nội dung tần số của tín hiệu theo thời gian. Tín hiệu RX (2.2) có thể được biểu diễn trong miền tần số bằng STFT, được định nghĩa trong (2.3). Cụ thể,  $Y(\tau, w)$  biểu diễn giản đồ phổ của tín hiệu RX tại tần số  $f$  và thời gian  $t$ ,  $y(t)$  là tín hiệu RX đầu vào,  $w(t - \tau)$  biểu thị hàm cửa sổ được sử dụng để phân đoạn tín hiệu thành các khung ngắn. Hơn nữa, mật độ năng lượng của các mô hình tín hiệu được minh họa bằng phương trình (2.4), với  $E$  biểu thị mật độ năng lượng trong khoảng  $[f_{\min}, f_{\max}]$ .



Hình 2.9: Mẫu phổ tín hiệu: (a) LTE, (b) 5G NR, (c) Chồng phổ 5G NR và LTE

Tỷ lệ tín hiệu trên nhiễu (SNR - signal-to-noise ratio) trong cả hệ thống 5G và LTE có thể được tính toán bằng cách so sánh công suất của tín hiệu với công suất của nhiễu. Trong các hệ thống thông tin liên lạc không dây, SNR thường được định nghĩa là tỷ lệ giữa công suất tín hiệu nhận trung bình và công suất nhiễu trung bình trên một băng thông xác định. Tỷ lệ nhiễu cao hơn có nghĩa là mức độ của các tín hiệu nhiễu không mong muốn so với tín hiệu mong muốn là đáng kể, điều này có thể làm giảm chất lượng và độ tin cậy của việc truyền dữ liệu. Công suất tín hiệu

<sup>1</sup>Toán tử \* biểu diễn cho phép nhân chập

phụ thuộc vào các yếu tố như công suất truyền, suy hao đường truyền, độ lợi của ăng-ten và các hiệu ứng fading. Đặc biệt, có một số mẫu giản đồ phổ trong Hình 2.9 trình bày về 5G NR, LTE, và sự chồng lấn giữa 5G và LTE tương ứng.

## Chương 3

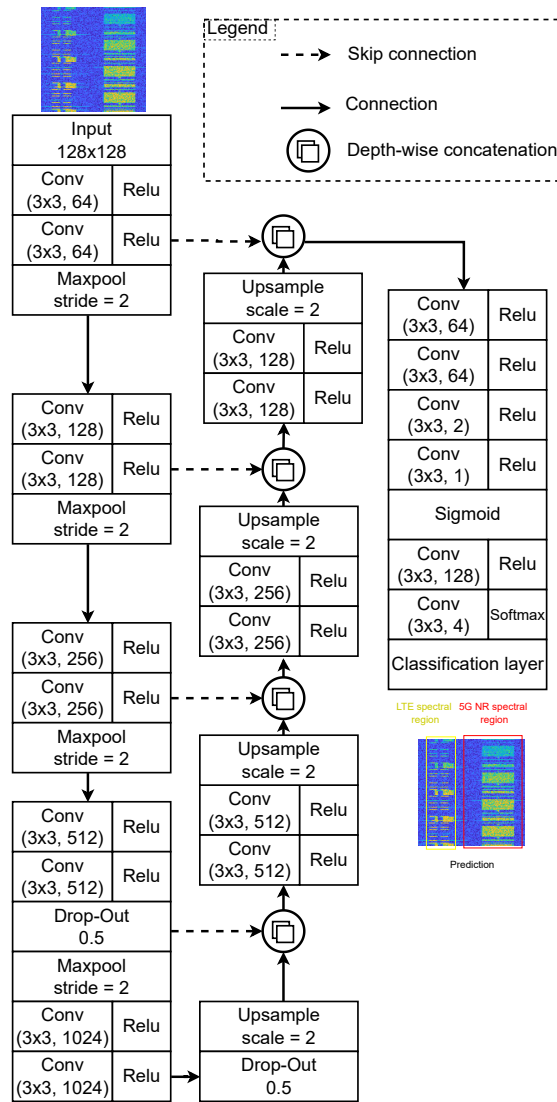
# GIỚI THIỆU MẠNG HỌC SÂU ỨNG DỤNG TRONG PHÂN VÙNG ẢNH

### 3.1 Giới thiệu về mạng học sâu U-net

Mô hình mạng U-net được giới thiệu bởi tác giả Ronneberger [14] cùng với các cộng sự nhằm ứng dụng vào lĩnh vực nhận diện và phân loại ảnh chụp quang học. U-net là một mô hình mạng nơ-ron tích chập toàn phần đã được cải tiến và phát triển dựa trên cấu trúc của mạng tích chập truyền thống. Đặc trưng của U-net là cấu trúc đối xứng hình chữ "U", được minh họa ở Hình 3.1. Phần mã hóa của mạng U-net tương tự như các mạng tích chập khác, bao gồm các lớp tích chập và tổng hợp tối đa để trích xuất đặc trưng từ ảnh. Trong Hình 3.1, các lớp tích chập sử dụng cửa sổ trượt kích thước  $3 \times 3$ , và các lớp tổng hợp tối đa sử dụng cửa sổ trượt kích thước  $2 \times 2$ , giảm kích thước đầu vào đi một nửa cả về chiều dài và chiều rộng.

Thông qua các lần lặp của tích chập và tổng hợp tối đa, kích thước của ảnh sẽ





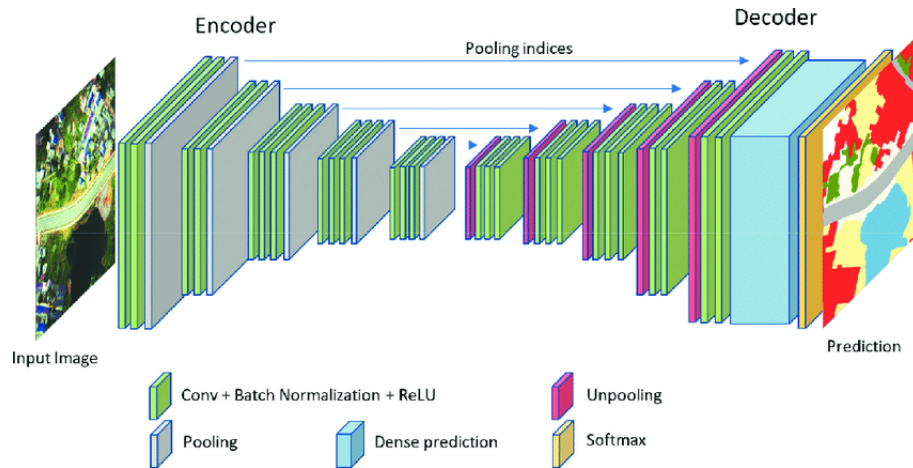
Hình 3.1: Kiến trúc mạng U-net.

giảm dần, tăng cường sự phức tạp của các đặc trưng được học. Để đảm bảo rằng thông tin chiều sâu đủ đáng kể, việc tăng số lớp tích chập ở cuối phần mã hóa là rất quan trọng để cải thiện hiệu suất học.

Một điểm đặc biệt của U-net là cấu trúc đối xứng của phần mã hóa và giải mã. Trong phần giải mã, việc mở rộng kích thước của ảnh tương đương với việc giảm kích thước của ảnh trong phần mã hóa tương ứng. Ngoài việc mở rộng kích thước, các lớp trong phần giải mã còn được kết nối đối xứng với các lớp tương ứng trong phần mã hóa, giúp khôi phục lại thông tin bị mất trong quá trình tổng hợp. Sau

mỗi lần mở rộng trong phần giải mã, kết quả được kết hợp với kết quả của lớp giảm kích thước tương ứng trong phần mã hóa, trước khi đi qua các lớp tích chập và tiếp tục quá trình mở rộng.

## 3.2 Mô hình mạng SegNet



Hình 3.2: Kiến trúc mạng SegNet (nguồn Internet).

SegNet là một mô hình mạng nơ-ron tích chập được công bố bởi tác giả Badri-narayanan và các cộng sự [20], đặc biệt được thiết kế để giải quyết bài toán phân đoạn ảnh. Mô hình này có nhiều điểm tương đồng với U-net, bao gồm một phần mã hóa được xây dựng theo kiến trúc mạng tích chập thông thường, với các lớp tích chập và tổng hợp được lấy cảm hứng từ các mô hình nổi tiếng như VGG16, VGG19 và ResNet. Tuy nhiên, SegNet loại bỏ lớp kết nối đầy đủ ở cuối của mạng.

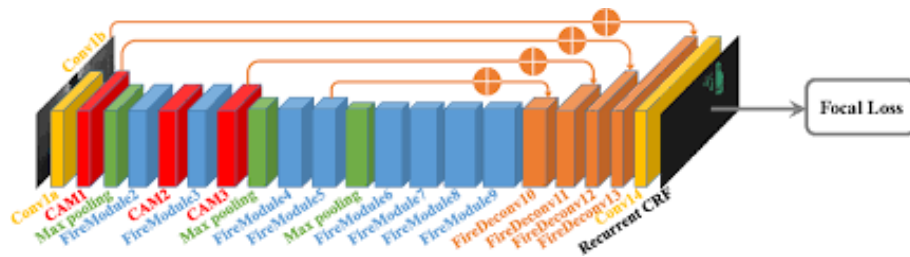
Như được minh họa trong Hình 3.3, phần mã hóa của SegNet sử dụng cấu trúc của mạng VGG16, bao gồm 13 lớp tích chập và 5 lớp giảm kích thước. Phần giải mã của SegNet cũng bao gồm các lớp mở rộng và tích chập, nhưng điểm đặc biệt và làm nên hiệu quả của mô hình nằm ở các lớp mở rộng.

Trong SegNet, đầu ra của các lớp mở rộng được mở rộng dựa trên vị trí ban đầu của các pixel tương ứng ở các lớp giảm kích thước trước đó. Để thực hiện điều này, vị trí của các pixel lớn nhất ở đầu vào của các lớp giảm kích thước được lưu giữ

lại. Khi đi qua các lớp tích chập, các pixel đầu ra của lớp giảm kích thước sẽ được đưa về vị trí ban đầu của chúng trước khi thực hiện giảm kích thước.

So với U-net, SegNet không có các kết nối với các biểu đồ đặc trưng của các lớp trước, nhưng thay vào đó, thông tin quan trọng được lưu trữ và chuyển tiếp tới các lớp sau là vị trí của các pixel. Điều này làm cho SegNet tiêu tốn ít bộ nhớ hơn trong các ứng dụng và thiết bị nhúng. Với phần giải mã phức tạp và lớn với nhiều tầng tích chập, SegNet đã cho thấy kết quả khả quan trong nhiều thử nghiệm thực tế.

### 3.3 Mô hình mạng SqueezeSegV2



Hình 3.3: Kiến trúc mạng SqueezeSegV2 (nguồn Internet).

SqueezeSegV2 [21], là một mạng nơ-ron tích chập được phát triển dựa trên mô hình giải mã - mã hóa, đặc biệt thiết kế để phân loại dữ liệu điểm thu thập từ LiDAR trong ứng dụng phân vùng ngữ nghĩa. Đây là phiên bản nâng cấp của mô hình gốc có tên là SqueezeSeg. SqueezeSeg được tạo ra để phát hiện và phân loại các đối tượng quan trọng trong không gian 3D bằng cách biểu diễn phân vùng ngữ nghĩa 3D dưới dạng phân loại cho từng điểm dữ liệu. Dữ liệu điểm từ LiDAR được chuyển đổi thành hình ảnh 2D thông qua cơ chế phóng chiếu hình cầu, tạo nên một sự cân đối giữa độ chính xác và độ phức tạp.

Kiến trúc của SqueezeSeg được xây dựng bằng cách sử dụng các mô-đun Fire và mô-đun Deconvolution (Deconv), nhằm khắc phục việc mất mát thông tin chi tiết ở mức thấp do các phép giảm kích thước dữ liệu. Đặc biệt, SqueezeSegV2

cải tiến bằng việc giới thiệu một mô-đun tổng hợp ngữ cảnh mới (CAM) để giảm thiểu tác động tiêu cực của nhiều dropout trong dữ liệu LiDAR, cũng như sử dụng hàm mất mát trọng tâm (focal loss) thay thế cho hàm mất mát cross-entropy ban đầu. Hàm mất mát trọng tâm giải quyết vấn đề mất cân bằng lớp bằng cách nhấn mạnh các mẫu âm trong quá trình huấn luyện.

So với U-net, SqueezeSegV2 tập trung vào việc phân loại dữ liệu điểm từ LiDAR thay vì phân đoạn hình ảnh. Điều này làm cho SqueezeSegV2 phù hợp hơn trong các ứng dụng cụ thể như tự lái xe, nơi mà dữ liệu LiDAR là quan trọng. SqueezeSegV2 cũng có một kiến trúc mạng phức tạp và các cải tiến trong hàm mất mát, giúp cải thiện độ chính xác và hiệu suất của mô hình. Tuy nhiên, U-net vẫn được ưa chuộng trong các ứng dụng phân đoạn hình ảnh chung do tính linh hoạt và khả năng áp dụng rộng rãi của nó.

### 3.4 Đánh giá ưu và nhược điểm

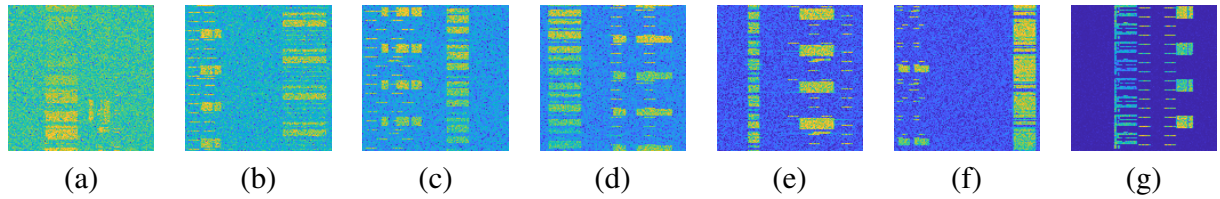
U-Net, SegNet và SqueezeSegV2 là ba mô hình quan trọng trong lĩnh vực phân vùng ngữ nghĩa ảnh, mỗi một trong số chúng đều có những đặc điểm và ưu điểm riêng.

U-Net nổi tiếng với cấu trúc đối xứng hình chữ "U", giúp nó trích xuất và tái tạo các đặc trưng cục bộ và toàn cục từ ảnh. Kiến trúc của U-Net giúp nó phù hợp với các tác vụ phân đoạn hình ảnh và đã được chứng minh hiệu quả trong nhiều ứng dụng y tế và địa lý. Tuy nhiên, U-Net có thể gặp khó khăn trong việc xử lý các đối tượng lớn và trong môi trường có nhiều nhiễu.

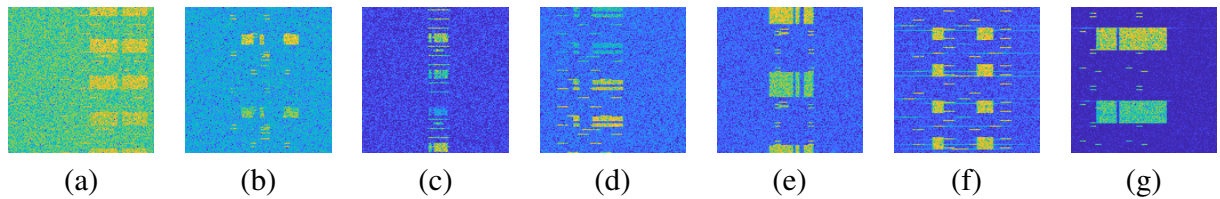
SegNet, trên cơ sở của mô hình tích chập truyền thống như VGG16 hoặc VGG19, chú trọng vào việc tái tạo vị trí của các pixel thông qua các lớp mở rộng và tích chập. Điều này giúp giữ lại thông tin về vị trí, đặc biệt quan trọng trong các ứng dụng nhúng và y tế. SegNet cũng có khả năng xử lý tốt các đối tượng lớn và nhiều nhiễu hơn so với U-Net.

SqueezeSegV2 là một mô hình tập trung vào việc phân loại dữ liệu điểm từ LiDAR, thích hợp cho các ứng dụng tự lái xe. Với cải tiến trong kiến trúc mạng và hàm mất mát, SqueezeSegV2 giúp cải thiện độ chính xác và hiệu suất trong việc phân loại các đối tượng trong không gian 3D từ dữ liệu LiDAR.

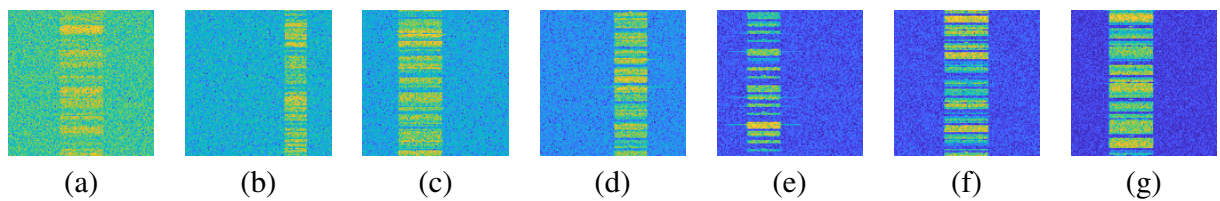
### 3.5 Tập dữ liệu khảo sát



Hình 3.4: Dữ liệu giao thoa phổ 5G NR và LTE với SNR khác nhau: (a) 0 dB, (b) 5 dB, (c) 10 dB, (d) 15 dB, (e) 20 dB, (f) 25 dB, (g) 30 dB.



Hình 3.5: Dữ liệu phổ 5G với SNR khác nhau: (a) 0 dB, (b) 5 dB, (c) 10 dB, (d) 15 dB, (e) 20 dB, (f) 25 dB, (g) 30 dB.



Hình 3.6: Dữ liệu phổ LTE với SNR khác nhau: (a) 0 dB, (b) 5 dB, (c) 10 dB, (d) 15 dB, (e) 20 dB, (f) 25 dB, (g) 30 dB.

Trong dự án nghiên cứu này, chúng tôi đã phải đối mặt với những thách thức thực tế, từ việc quản lý chi phí cho đến bảo vệ quyền riêng tư. Để nắm bắt bức tranh toàn diện về mạng 5G và LTE, chúng tôi đã sử dụng bộ công cụ 5G trong

Bảng 3.1: Thông tin về dữ liệu đầu vào cho mỗi phân loại

Dataset Information			
Category	No.Samples	Image size	SNR (dB)
LTE	5,000	$128 \times 128$	$[0, 30]$
5G	5,000	$128 \times 128$	$[0, 30]$
5G and LTE	5,000	$128 \times 128$	$[0, 30]$

Matlab để tạo ra một tập dữ liệu phổ tổng hợp, bao gồm ba loại chính: 5G, LTE và các khung giao thoa giữa 5G và LTE.

Tập dữ liệu của chúng tôi bao gồm tổng cộng 5,000 mẫu cho mỗi loại, mỗi mẫu có kích thước hình ảnh là  $128 \times 128$ . Đặc biệt, chúng tôi đã khám phá một loạt các mức độ tín hiệu tạp âm-tỷ lệ tín hiệu (SNR) từ 0 đến 30 dB, để mô phỏng các tình huống khác nhau trong môi trường thực tế.

### 3.6 Kết luận

Trong đề tài này, học viên đã đề xuất đề tài liên quan đến giám sát phổ tín hiệu dựa trên các mô hình học sâu, nhằm mục đích phát hiện các dải băng tần chưa được sử dụng và cấp phát cho các loại sóng di động cũng như ra-da khác. Dựa trên sự so sánh giữa các mô hình học sâu và những kinh nghiệm đã được tích lũy từ những đề tài nghiên cứu trước đó, học viên quyết định chọn mô hình U-net để tiếp tục tìm hiểu chi tiết và tìm ra giải pháp tăng cường hiệu năng của mô hình học sâu U-net sao cho đáp ứng được nhu cầu thực tế của đề tài. Việc lựa chọn U-Net cho việc cảm biến phổ có nhiều ưu điểm quan trọng:

- **Xử lý hình ảnh 2D:** U-Net được thiết kế đặc biệt cho việc phân đoạn hình ảnh, phù hợp cho dữ liệu từ cảm biến phổ trong không gian 2D.
- **Kiến trúc đặc biệt:** Kiến trúc đối xứng hình chữ "U" của U-Net giúp nó trích xuất đặc trưng cục bộ và toàn cục từ ảnh một cách hiệu quả. Điều này rất hữu

ích trong việc xử lý dữ liệu từ cảm biến phổ, nơi mà việc trích xuất các đặc trưng quan trọng từ dữ liệu là cần thiết.

- **Độ linh hoạt:** U-Net có thể dễ dàng được điều chỉnh và tinh chỉnh cho các yêu cầu cụ thể của ứng dụng cảm biến phổ. Cấu trúc mạng linh hoạt và khả năng đào tạo hiệu quả giúp nó thích hợp cho nhiều loại dữ liệu và nhiều tác vụ khác nhau.
- **Hiệu suất và hiệu quả:** U-Net đã được chứng minh hiệu quả trong nhiều ứng dụng phân đoạn hình ảnh, bao gồm cả việc xử lý dữ liệu từ cảm biến phổ. Với khả năng học và trích xuất đặc trưng tốt, U-Net có thể giúp tăng cường hiệu suất của hệ thống cảm biến phổ và giảm thiểu sai số.
- **Sử dụng trong các ứng dụng thực tế:** U-Net đã được sử dụng rộng rãi trong nhiều lĩnh vực, từ y tế đến xe tự lái và nông nghiệp. Sự phổ biến và khả năng thích nghi của nó làm cho U-Net trở thành một lựa chọn hợp lý cho việc xử lý dữ liệu từ cảm biến phổ trong các ứng dụng thực tế.

## Chương 4

# KẾT LUẬN VÀ ĐỊNH HƯỚNG PHÁT TRIỂN CHO CHUYÊN ĐỀ 2

### 4.1 Kết Luận

Trong quá trình nghiên cứu, học viên đã tiếp cận và hiểu biết sâu hơn về các mô hình mạng nơ-ron, đặc biệt là mạng nơ-ron tích chập và lý thuyết học sâu. Họ cũng đã tìm hiểu về bài toán phân vùng ảnh và cách áp dụng học sâu vào bài toán này. Dựa trên kiến thức thu thập được, học viên đã phát hiện ra rằng mô hình U-Net là sự lựa chọn tốt nhất cho dự án của mình.

U-Net đã chứng minh được hiệu quả và linh hoạt của mình trong việc phân vùng ảnh. So với SqueezeSegV2 và SegNet, U-Net có những ưu điểm nổi bật. Đặc biệt, U-Net có khả năng xử lý tốt các đối tượng nhỏ và linh hoạt trong việc điều chỉnh và tinh chỉnh cho các nhiệm vụ cụ thể. Mô hình này cũng không đối mặt với các hạn chế như SqueezeSegV2, như khả năng nhận diện các vật thể nhỏ và yêu cầu tài nguyên tính toán cao.



Do đó, việc sử dụng U-Net sẽ giúp học viên giải quyết các yêu cầu của dự án một cách hiệu quả và linh hoạt hơn, đồng thời cũng giảm thiểu được các hạn chế mà các mô hình khác như SqueezeSegV2 và SegNet đang đối mặt.

## 4.2 Hướng phát triển

Trong tương lai, việc nghiên cứu sâu hơn về việc tối ưu hóa cấu trúc mạng và các hàm mất mát có thể giúp cải thiện hiệu suất của mô hình U-Net. Các nghiên cứu này có thể tập trung vào việc tinh chỉnh kiến trúc mạng để tối ưu hóa việc trích xuất đặc trưng và dự đoán. Ngoài ra, cải tiến các hàm mất mát cũng có thể giúp tăng độ chính xác của mô hình và giảm thiểu thời gian dự đoán.

Thêm vào đó, việc thu thập một tập dữ liệu đa dạng hơn cũng là một phần quan trọng trong quá trình nghiên cứu. Dữ liệu này nên bao gồm các vị trí đối tượng lớn hơn và các trường hợp bổ sung của các lớp thiểu số, để mô hình có khả năng học được từ những trường hợp đa dạng và phong phú hơn.

Khám phá các độ phân giải dữ liệu khác nhau cũng là một hướng nghiên cứu tiềm năng. Việc này có thể giúp phát hiện và giải quyết các nhược điểm trong việc phân đoạn ảnh và cung cấp những hiểu biết có giá trị để cải thiện hiệu suất của mô hình U-Net trong các ứng dụng thực tế.

# Tài liệu tham khảo

- [1] T. Yucek and H. Arslan, “A survey of spectrum sensing algorithms for cognitive radio applications,” *IEEE Communications Surveys & Tutorials*, vol. 11, no. 1, pp. 116–130, Mar. 2009.
- [2] A. Kumar, N. Gaur, S. Chakravarty, M. H. Alsharif, P. Uthansakul, and M. Uthansakul, “Analysis of spectrum sensing using deep learning algorithms: CNNs and RNNs,” *Ain Shams Engineering Journal*, vol. 15, no. 3, p. 102505, Mar. 2024.
- [3] A. Ali and W. Hamouda, “Advances on spectrum sensing for cognitive radio networks: Theory and applications,” *IEEE Communications Surveys & Tutorials*, vol. 19, no. 2, pp. 1277–1304, Jun. 2017.
- [4] S. D. Liyanaarachchi, T. Riihonen, C. B. Barneto, and M. Valkama, “Optimized waveforms for 5G–6G communication with sensing: Theory, simulations and experiments,” *IEEE Transactions on Wireless Communications*, vol. 20, no. 12, pp. 8301–8315, Dec. 2021.
- [5] G.-V. Nguyen, C. V. Phan, and T. Huynh-The, “Accurate spectrum sensing with improved DeepLabV3+ for 5G-LTE signals identification,” in *Proc. 2023 of the 12th International Symposium on Information and Communication Technology*, HoChiMinh, Vietnam, Dec. 2023, pp. 221–227.

- [6] T. Huynh-The, Q.-V. Pham, T.-H. Vu, D. B. da Costa, and V.-P. Hoang, “Intelligent spectrum sensing with convnet for 5G and LTE signals identification,” in *Proc. 2023 IEEE Statistical Signal Processing Workshop (SSP)*, Hanoi, Vietnam, Jul. 2023, pp. 140–144.
- [7] T. Huynh-The, N. C. Luong, H. Phan, D. B. da Costa, and Q.-V. Pham, “Improved waveform classification for integrated radar-communication 6 G systems via convolutional neural networks,” *IEEE Transactions on Vehicular Technology*, Apr. 2024.
- [8] J. Gao, X. Yi, C. Zhong, X. Chen, and Z. Zhang, “Deep learning for spectrum sensing,” *IEEE Wireless Communications Letters*, vol. 8, no. 6, pp. 1727–1730, Dec. 2019.
- [9] W. Weng and X. Zhu, “INet: Convolutional networks for biomedical image segmentation,” *IEEE Access*, vol. 9, pp. 16 591–16 603, Jan. 2021.
- [10] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, and J. Liang, “Unet++: Redesigning skip connections to exploit multiscale features in image segmentation,” *IEEE transactions on medical imaging*, vol. 39, no. 6, pp. 1856–1867, Jun. 2019.
- [11] F. Chollet, “Xception: Deep learning with depthwise separable convolutions,” in *Proc. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Nov. 2017, pp. 1800–1807.
- [12] M. Z. Alom, C. Yakopcic, T. M. Taha, and V. K. Asari, “Nuclei segmentation with recurrent residual convolutional neural networks based U-Net (R2U-Net),” in *Proc. 2018 - IEEE National Aerospace and Electronics Conference*, Dec. 2018, pp. 228–233.

- [13] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proc. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Dec. 2016, pp. 770–778.
- [14] M. Aghalari, A. Aghagolzadeh, and M. Ezoji, “Brain tumor image segmentation via asymmetric/symmetric unet based on two-pathway-residual blocks,” *Biomedical signal processing and control*, vol. 69, p. 102841, Aug. 2021.
- [15] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, “DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 4, pp. 834–848, Apr. 2018.
- [16] C. Janiesch, P. Zschech, and K. Heinrich, “Machine learning and deep learning,” *Electronic Markets*, vol. 31, no. 3, pp. 685–695, 2021.
- [17] S. Minaee, Y. Boykov, F. Porikli, A. Plaza, N. Kehtarnavaz, and D. Terzopoulos, “Image segmentation using deep learning: A survey,” *IEEE transactions on pattern analysis and machine intelligence*, vol. 44, no. 7, pp. 3523–3542, 2021.
- [18] J. K. Leonard, “Image classification and object detection algorithm based on convolutional neural network,” *Science Insights*, vol. 31, no. 1, pp. 85–100, 2019.
- [19] X. Lin and N. Lee, “5G and beyond,” *Cham, Switzerland: Springer Nature Switzerland AG*, 2021.
- [20] V. Badrinarayanan, A. Kendall, and R. Cipolla, “Segnet: A deep convolutional encoder-decoder architecture for image segmentation,” *IEEE transactions on pattern analysis and machine intelligence*, vol. 39, no. 12, pp. 2481–2495, 2017.

- [21] B. Wu, X. Zhou, S. Zhao, X. Yue, and K. Keutzer, “Squeezesegv2: Improved model structure and unsupervised domain adaptation for road-object segmentation from a lidar point cloud,” in *2019 international conference on robotics and automation (ICRA)*. IEEE, 2019, pp. 4376–4382.