TRƯỜNG ĐẠI HỌC AN GIANG

**KHOA CÔNG NGHỆ THÔNG TIN**

**ĐỀ CƯƠNG KHOÁ LUẬN TỐT NGHIỆP**

**NGÀNH KỸ THUẬT PHẦN MỀM**

# NGHIÊN CỨU VÀ XÂY DỰNG HỆ THỐNG CAMERA GIÁM SÁT THÔNG MINH PHỤC VỤ CHĂM SÓC SỨC KHỎE TẠI NHÀ

Họ tên sinh viên : **Thái Trường An**

Mã số sv : **DPM175008**

Khóa : **18**

Giảng viên hướng dẫn: **TS. Đoàn Thanh Nghị**

## **AN GIANG, THÁNG 11 NĂM 2020**

**ĐỀ CƯƠNG KHÓA LUẬN TỐT NGHIỆP**

Họ và tên sinh viên: Thái Trường An Mã số sinh viên: DPM175008

Điện thoại: 0914518169 Email: thaitruonganlxag91@gmail.com

Là sinh viên lớp: DH18PM ngành Kỹ thuật phần mềm

thuộc Khoa Công nghệ Thông tin.

Tôi bắt đầu nhận đề tài khóa luận tốt nghiệp từ ngày 23 tháng 11 năm 2020 đến ngày 23 tháng 4 năm 2020

**Tên đề tài:**  Nghiên cứu và xây dựng hệ thống camera giám sát thông minh phục vụ chăm sóc sức khỏe tại nhà.

dưới sự hướng dẫn của Thầy/Cô: TS. Đoàn Thanh Nghị

**CHƯƠNG I**

**TỔNG QUAN**

* 1. **TÍNH CẤP THIẾT CỦA ĐỀ TÀI**

Ngày nay, khi kinh tế ngày càng phát triển cuộc sống con người cũng vì thế mà được nâng cao. Nhu cầu về chăm sóc sức khỏe trở nên là ưu tiên hàng đầu. Cùng với đó, là sự phát triển của khoa học và công nghệ trong cuộc cách mạng công nghiệp 4.0, lĩnh vực về chăm sóc sức khỏe lại được nâng lên tầm cao mới. Việc chăm sóc sức khỏe thông qua việc phát hiện các triệu chứng bất thường như: té xỉu, đau bụng, nhức đầu,… và cảnh báo các triệu chứng cho người thân trong gia đình là vô cùng cần thiết. Từ những thực tiễn đó là một sinh viên chuyên ngành Kỹ thuật phần mềm (KTPM) tôi mong muốn tìm hiểu, nghiên cứu và ứng dụng thị giác máy tính cũng như kỹ thuật học sâu vào tự động phát hiện các triệu chứng bất thường thông qua hành động con người.

Lịch sử giải quyết vần đề: nhận diện hành động con người đã từng được ứng dụng trong tóm tắt, truy vấn video và phân tích thể thao. Nhưng trên lĩnh vực camera giám sát chăm sóc sức khỏe tại nhà thì rất ít nghiên cứu trước đó. Vì vậy, nghiên cứu về camera giám sát thông minh có ý nghĩa khoa học và thực tiễn rất cao.

* 1. **LÍ DO CHỌN ĐỀ TÀI**

Cùng với sự phát triển của khoa học công nghệ trong cuộc cách mạng công nghiệp 4.0, các ứng dụng về trí tuệ nhân tạo (TTNT) và học máy ngày càng phát triển và là lĩnh vực được quan tâm nhiều nhất hiện nay, trong đó lĩnh vực học sâu (Deep Learning) là một loại phổ biến của máy học đã có những thành tựu phát triển vượt bậc. Mở ra một bước ngoặc mới trong việc giải quyết các bài toán về TTNT trước đây đã gặp khó khăn như nhận diện sự vật (Object Detection), nhận diện hình ảnh, hành động, … Trong đó phải kể đến sự phát triển của lĩnh vực thị giác máy tính (Computer Vision), đặt nền tảng cho nhiều ứng dụng thực tiển như các hệ thống giám sát thông minh, hệ thống xe tự lái, rôbôt thông minh,… Lĩnh vực học sâu là một kỹ thuật trong máy học sử dụng mạng nơ ron nhiều lớp để giải quyết các bài toán phức tạp dựa trên phương pháp tích chập và trích chọn các đặt trưng từ các tập dữ liệu lớn và đem lại kết quả chính xác cao trong giải quyết các bài toán TTNT trong thực tế. Cụ thể là nó tập trung giải quyết các vần đề về mạng nơ ron về thị giác máy tính, xử lý ngôn ngữ tự nhiên, xử lý giọng nói,… Hiện nay, cũng có nhiều nghiên cứu về nhận diện hành động con người thông qua video để xác định hành vi của con người. Tuy nghiên đây là một lĩnh vực tương đối rộng và hiện có nhiều kỹ thuật khác nhau để giải quyết bài toán này Vì vậy phạm vi luận văn này tập trung nghiên cứu lý thuyết và sử dụng các kỹ thuật mạng học sâu 3D CNN vào việc giải quyết bài toán nhận diện hành động con người đưa ra các dự đoán về các triệu chứng bất thường.

Trên cơ sở đó, tôi thực hiện khóa luận này để nghiên cứu ứng dụng kỹ thuật học sâu vào bài toán nhận diện hành động của con người trong video với đề tài “Nghiên cứu và xây dựng hệ thống camera giám sát thông minh phục vụ chăm sóc sức khỏe tại nhà”. Thông qua đề tài này góp phần phát triển hệ thống camera giám sát thông minh.

* 1. **MỤC TIÊU ĐỀ TÀI**

Sử dụng OpenCV, Python, Pytorch, Twilio để xây dựng mô hình camera giám sát thông minh với độ chính xác cao trong thực tế giúp nâng cao hiệu quả và chất lượng của quá trình giám sát đảm bảo sức khỏe cho các thành viên trong gia đình.

Sử dụng trực tiếp thuật toán xử lý trên bộ mạch Raspberry Pi 4 và tự động gửi tin nhắn thông báo TXT/MMS, hình ảnh, video clip khi camera được kích hoạt và phát hiện các hành động, triệu chứng bất thường của các thành viên trong gia đình.

* 1. **MỤC TIÊU NGHIÊN CỨU**

Mục tiêu chính của đề tài là nghiên cứu nhận dạng hành động người, tập trung vào các hành động, triệu chứng bất thường dựa trên kỹ thuật học sâu mạng tích chập 3D – CNN trong lĩnh vực thị giác máy tính.

Thực nghiệm áp dụng kỹ thuật mạng tích chập 3D-CNN trong trích xuất đặt trưng hình ảnh từ video để dự đoán hành động con người phục vụ chăm sóc sức khỏe tại nhà. Từ đó phân tích, đánh giá hiệu quả và độ chính xác trên các thư viện có sẵn.

Xây dựng tập dữ liệu và thực nghiệm dự đoán một số hành động, triệu chứng bất thường trong hệ thống camera giám sát thông minh.

* 1. **PHƯƠNG PHÁP NGHIÊN CỨU**

Phương pháp nghiên cứu tài liệu: Cơ sở lý thuyết về các kỹ thuật nhận diện hành động, các nghiên cứu tương tự về đề tài.

Phương pháp thực nghiệm: Xây dựng mô hình camera và cài đặt chương trình chạy thử nghiệm với một số video sẳn có và đánh giá kết quả. Từ đó xây dựng mô hình camera xử lý video thời gian thực.

**CHƯƠNG II**

**CƠ SỞ LÝ THUYẾT**

**2.1. HỆ THỐNG THÔNG MINH**

**2.1.1. Khái niệm**

Hệ thống thông minh (Intelligent Systems - IS) là một hệ thống thực thi một mục đích hoạt động xã hội nào đó, nó được xem như một máy tính kết nối với các máy tính khác qua hệ thống mạng internet, có khả năng thu thập và phân tích dữ liệu và giao tiếp các hệ thống khác. Các tiêu chí của hệ thống thông minh bao gồm khả năng học hỏi từ kinh nghiệm, bảo mật, kết nối, khả năng thích ứng theo dữ liệu hiện tại và khả năng giám sát,quản lý từ xa.

Trong nghành công nghệ thông tin, một hệ thống thông minh được định nghĩa là một tập hợp các phần tử (Elements) hoặc các thành phần (Components) tạo nên một hệ thống thực hiện một mục đích chung. Như vậy, ngoài việc nó cấu thành từ các thiết bị thông minh, nó còn bao gồm các hệ thống kết nối các thiết bị lại với nhau: bao gồm mạng và các hệ thống lớn hơn, các hệ thống phần mềm thông minh (AI) như chatbots, hệ thông chăm sóc sức khỏe,… và những loại phần mềm khác.

Như vậy các hệ thống thông minh có các đặc điểm sau:

* Cung cấp cách tiếp cận về mặt phương pháp luận chuẩn để giải quyết các vấn đề quan trong, phức tạp và nhận được các kết quả nhất quán, đáng tin cậy qua thời gian
* Có khả năng kết nối, giao tiếp với các hệ thống khác, quản lý từ xa.
* Có trí tuệ bao gồm khả năng lĩnh hội, khả năng hiểu, khả năng rút kinh nghiệm, khả năng thu lượm và duy trì tri thức, khả năng đáp ứng nhanh và thành công trong các tình huống mới.
* Trí tuệ nhân tạo là cơ sở cốt yếu để xây dựng các hệ thống thông minh. Các hệ thống thông minh hiện nay được ứng dụng rất rộng rãi trong tất cả các lĩnh vực như y tế (hệ thống chẫn đoán bệnh), giáo dục (trường học thông minh), sản xuất Công nghiệp (Công nghệ in 3D), hệ thống xe tự hành, thành phố thông minh… là các hệ thống thông minh phổ biến hiện nay ở các nước phát triển.

**2.1.2. Hệ thống hỗ trợ chăm sóc sức khỏe bệnh nhân, người già**

Cùng với sự phát triển của trí tuệ nhận tạo, hệ thống robot chăm sóc sức khỏe cho người già ra đời và phát triển mạnh ở các nước phát triển như Nhật Bản, rôbôt có thể thay thế con người trong một số nhiệm vụ chăm sóc người già, tuy nhiên nó vẫn mang tính hỗ trợ, giải phóng các lao động đơn giản của y tá chưa hoàn toàn thay thế con người trong các dịch vụ chăm sóc y tế cho người già.

**2.2. THỊ GIÁC MÁY TÍNH**

Thị giác máy tính là một lĩnh vực gồm các phương pháp thu nhận, phân tích, xử  
lý, nhận dạng hình ảnh và dữ liệu đa chiều từ thực tế để cho ra thông tin số hoặc biểu  
tượng dùng cho các mục đích nhất định.

Ý tưởng được dựa trên thị giác của con người về khả năng nhận dạng và hiểu  
một hình ảnh. Để làm được điều đó, thị giác máy tính áp dụng các mô hình xây dựng  
dựa trên các ngành lý thuyết học, vật lý, thống kê và hình học.

Trong các ứng dụng thị giác máy tính thường được lập trình để giải quyết một  
vấn đề cụ thể nhưng trên cơ bản các phương pháp đều dựa trên sự học (learning) trở  
nên ngày càng phổ biến. Ví dụ các ứng dụng thị giác máy tính như:

* Điều khiển tiến trình (Controlling processes) như rô bốt công nghiệp.
* Nhận diện khuôn mặt (Recognition Face).
* Sự kiểm tra tự động (Automation inspection) như kiểm tra chất lượng sản phẩm trong các nhà máy xí nghiệp.

**2.3. NHẬN DIỆN HÀNH ĐỘNG CỦA CON NGƯỜI**

Nhận diện hành động của con người là một lĩnh vực nghiên cứu rộng rãi trong thị giác máy tính. Các ứng dụng của nó bao gồm các hệ thống giám sát, phân tích video, robot và một loạt các hệ thống có liên quan đến sự tương tác giữa con người và các thiết bị điện tử như giao diện người - máy tính.

Nhận diện hành động của con người được phát triển bắt đầu vào năm 1980. Hiện nay, các nghiên cứu chủ yếu tập trung vào việc học và nhận biết các hành động từ chuỗi video.

Một hành động là một chuỗi các chuyễn động cơ thể đồng thời. Từ quan điểm của thị giác máy tính, việc nhận dạng hành động này là để phù hợp với các quan sát (ví dụ: video) với các mẫu được xác định trước đó và sau đó gán cho nó một nhãn là loại hành động. Tùy thuộc vào độ phức tạp, hoạt động của con người có thể được phân loại thành bốn cấp độ cử chỉ, hành động, tương tác và hoạt động của nhóm và nhiều nghiên cứu theo hướng một cấu trúc từ dưới lên về nhận dạng hoạt động của con người. Các phần chính của hệ thống như vậy bao gồm trích xuất đặc trưng, học tập hành động, phân loại, nhận dạng hành động và phân loại. Một quy trình đơn giản gồm ba bước, cụ thể là phát hiện của con người hoặc các bộ phận cơ thể, theo dõi, và sau đó nhận bằng cách sử dụng kết quả theo dõi. Ví dụ, để nhận ra hành động “bắt tay”, cánh tay và bàn tay của hai người được phát hiện trước tiên và theo dõi để tạo ra một mô tả không gian – thời gian của chuyển động của họ. Mô tả này được so sánh với các mẫu hiện có trong dữ liệu huấn luyện để xác định hành loại hành động. Mô hình này dựa rất nhiều vào tính chính xác của việc theo dõi, điều này không đáng tin cậy trong những cảnh lộn xộn.

Nhiều phương pháp đã được đề xuất, và có thể được phân loại theo nhiều tiêu chí khác nhau như trong bài luận điều tra hiện có. Poppe thảo luận nhận dạng hành động của con người từ biễu diễn hình ảnh và phân loại hành động riêng rẽ. Weinland khảo sát các phương pháp cho biểu diễn hành động, phân đoạn và nhận dạng. Turaga chia vấn đề nhận dạng hành động và hoạt động theo độ phức tạp, và các hướng tiếp cận phân loại theo khả năng của mình để xử lý các mức độ phức tạp khác nhau. Có nhiều tiêu chí phân loại khác nhau. Trong số đó, Aggarwal và Ryoo là một trong những tổng kết toàn diện mới nhất và sự so sánh của sự tiến bộ quan trọng nhất trong lĩnh vực này. Dựa vào hành động nhận dạng từ hình ảnh đầu vào trực tiếp, Aggarwal và Ryoo phân chia phương pháp nhận dạng thành hai loại chính: phương pháp tiếp cận đơn lớp và phương pháp tiếp cận phân cấp. Cả hai đều là thêm loại con phụ thuộc vào các phương pháp biễu diễn đặc trưng và học.

**2.4. THƯ VIỆN OPENCV**

**2.4.1. Giới thiệu**

OpenCV (Open Source Computer Vision) là một thư viện mã nguồn mở về  
thị giác máy với hơn 500 hàm và hơn 2500 các thuật toán đã được tối ưu về xử lý ảnh,  
và các vấn đề liên quan tới thị giác máy. Là thư viện được thiết kế để chạy trên nhiều  
nền tảng khác nhau (cross-patform) như là hệ điều hành Windows, Linux, Mac, IOS...  
Việc sử dụng thư viện OpenCV tuân theo các quy định về sử dụng phần mềm mã  
nguồn mở BSD do đó bạn có thể sử dụng thư viện này một cách miễn phí cho cả mục  
đích phi thương mại lẫn thương mại.

Dự án về OpenCV được khởi động từ những năm 1999, đến năm 2000 nó được  
giới thiệu trong một hội nghị của IEEE về các vấn đề trong thị giác máy và nhận dạng,  
tuy nhiên bản OpenCV 1.0 mãi tới tận năm 2006 mới chính thức được công bố và năm  
2008 bản 1.1 pre-release mới được ra đời. Tháng 10 năm 2009, bản OpenCV thế hệ thứ  
hai ra đời thường gọi là phiên bản 2.0, phiên bản này có giao diện của C++ (khác với  
phiên bản rước có giao diện của C) và có khá nhiều điểm khác biệt so với phiện bản  
thứ nhất.

**2.4.2. Đặc trưng**

* Là thư viện mã nguồn mở dùng trong nhiều ngôn ngữ lập trình như: C/C++, Java, Python
* Được sử dụng rộng rãi, thu hút, được một lượng lớn người dùng trong đó có các công ty lớn như Microsoft, IBM, Sony, v.v.
* Chạy độc lập với các nền tảng phần cứng và phần mềm.
* Giúp cho công việc xây dựng các ứng dụng xử lý ảnh, thị giác máy tính một cách nhanh chóng hơn.
* OpenCV tập trung vào thu thập ảnh, xử lý ảnh và các thuật toán phân tích dữ liệu từ ảnh.

**2.4.3. Tổ chức**

Thư viện OpenCV bao gồm các module:

* core: gồm các kiểu dữ liệu và các hàm cơ bản để sử dụng các module khác.
* improc: bao gồm các giải thuật phân tích xử lý ảnh như:
  + Lọc ảnh tuyến tính (linear) và phi tuyến (non-linear).
  + Chuyển đổi hình học (geomatrical image transformation).
  + Biểu đồ tầng số (histograms).
* highgui: gồm các hàm về giao diện, sử dụng để đọc, ghi, hiển thị hình ảnh và video; các giải thuật phân tích, xử lý video như:
  + Ước lượng chuyển động (motion estimation).
  + Trích xuất nền sau (background subtraction).
  + Theo dõi đối tượng (object tracking).
* Features2d: gồm các hàm dò tìm đặt trưng nổi bật (salient feature detectors), mô tả (descriptors) và so khớp (descriptor matchers).
* Objdetect: gồm các giải thuật phát hiện đối tượng và các tập dữ liệu để phát  
  hiện một số đối tượng như khuôn mặt, mắt, miệng, con người, xe hơi, …
* ml: gồm các giải thuật máy học như mô hình thống kê, hồi quy, phân loại,  
  gom cụm.

Ngoài ra, còn rất nhiều module khác như flann, gpu, photo, stitching,  
nonfree, contrib, legacy, ocl, superres, viz.

**2.5. ỨNG DỤNG DEEP LEARNING TRONG NHẬN DIỆN HÀNH ĐỘNG CON NGƯỜI**

**2.5.1. Giới thiệu Convolutional Neural Network**

**2.5.1.1. Tổng quan**

Convolution Neural Network (CNN) bao gồm một hoặc nhiều lớp chập (thường với một bước lấy mẫu con) và sau đó theo sau bởi một hoặc nhiều hơn các lớp kết nối như trong một mạng nơ ron đa lớp chuẩn. Kiến trúc của một CNN được thiết kế để tận dụng lợi thế của cấu trúc 2 chiều của một hình ảnh đầu vào (hoặc đầu vào 2 chiều khác như một tín hiệu tiếng nói). Điều này đạt được với các kết nối cục bộ và trọng số rang buộc theo một số hình thức tổng hợp mà kết quả là các đặc trưng không thay đổi. Một lợi ích khác của CNN là dễ dàng huấn luyện hơn và có ít thông số so với các mạng kết nối đầy đủ với cùng một số đơn vị ẩn.

**2.5.1.2. Kiến trúc CNN**

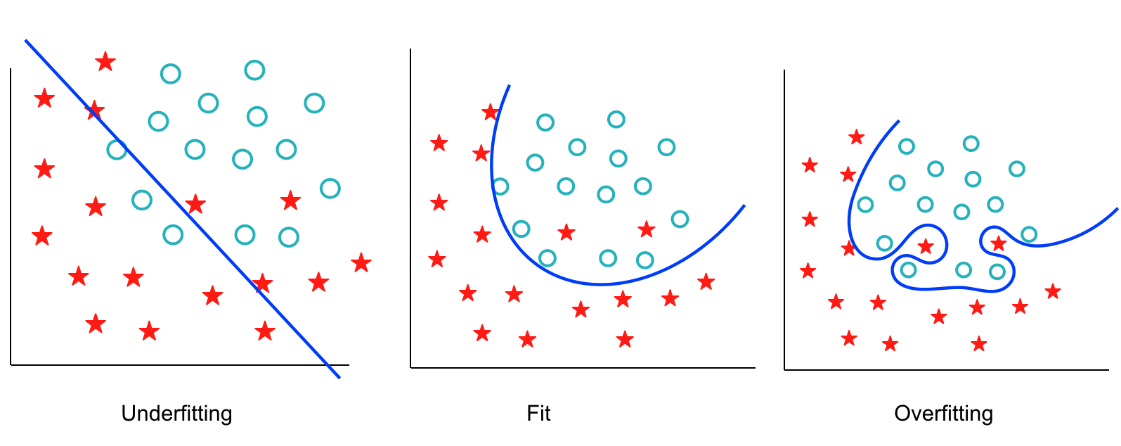
CNN gồm một số lớp chập và lớp lấy mẫu con tùy chọn, tiếp theo là các lớp kết nối. Đầu vào cho một lớp chập là một hình ảnh m \* m \* r, với m là chiều cao và chiểu rộng của hình ảnh và r là số kênh (ví dụ RGB có r=3). Lớp chập sẽ có k bộ lọc có kích thước n \* n \* q, trong đó n nhỏ hơn kích thước hình ảnh và q có thể bằng hoặc nhỏ hơn số kênh r. Kích thước của các bộ lọc tăng đến cấu trúc kết nối cục bộ, trong đó mỗi bộ lọc xoắn với hình ảnh để tạo ra k bản đồ đặc trưng có kích thước m – n + 1. Mỗi bản đồ được lấy mẫu con đặc trưng với việc hợp nhất trung bình hoặc lớn nhất p \* p khu vực lân cận, trong đó phạm vi của p là giữa 2 (cho các hình ảnh nhỏ) và 5 (chi các đầu vào lớn hơn). Trước hoặc sau khi lấy mẫu con, một bias bổ sung và xích ma phi tuyến được áp dụng cho mỗi bản đồ đặc trưng.



Sau các lớp chập, có một số lớp kết nối đầy đủ. Các lớp kết nối dày đặc giống với các lớp trong mạng nơ ron đa lớp chuẩn.

**2.5.2. Vấn đề Overfitting và Underfitting**

Sau khi lựa chọn một mô hình phù hợp chẳng hạn Logistic Regression cho Classification Problem, hay Linear Regression cho Regression Problem ta xây dựng dự đoán h(x), hàm chi phí C() và tối thiểu hóa hàm chi phí nhiều nhất có thể. Đến bước này chúng ta sẽ gặp vấn đề vơi Overfitting và Underfitting.



Underfitting là khi mô hình quá đơn giản, không thể giảm thiểu được đáng kể hàm chi phí nên cũng không thể mô tả được xu hướng của dữ liệu (còn được gọi là High Bias). Ngược lại Overfitting lại là khi mô hình quá phức tạp, tuy giảm thiểu được đáng kể , thậm chí toàn bộ sai số nhưng cũng không thể mô tả được xu hướng của dữ liệu (còn được gọi là High Variance).

Cách giải quyết:

Với Underfitting thì chỉ đơn thuần là mô hình quá đơn giản, nên ta cần thêm những thành phần đa thức để nó phức tạp hơn. Nên khi giá trị hàm chi phí lớn ta sẽ đẩy bậc của hàm số lên. Dĩ nhiên đẩy lên cao quá ta sẽ gặp vấn đề về hiệu năng và Overfitting.

Với Overfitting thì công việc lại ngược lại, do mô hình quá phức tạp nên ta cần giảm bậc của hàm số hay giảm số lượng đặc trưng. Việc giảm đặc trưng ngoài bỏ bớt những thành phần đa thức, ta có còn có thể bỏ bớt những đặc trưng không cần thiết.

Một cách khác là áp dụng Regularization.

**2.5.3. Các giải thuật Optimization**

Thuật toán huấn luyện của rất nhiều mô hình máy học được phát biểu hình thức dưới dạng dạng một bài toán Optimization, chẳng hạn các mô hình như linear/logistic regression, ANN, SVM, k-means… Ý tưởng chung là mỗi mô hình đều có một tập hợp tham số nào đó cần phải được xác định giá trị thông qua quá trình huấn luyện với dữ liệu và ta cần tìm để cực tiêu hóa hàm chi phí C(). Mặc dù lí thuyết optimization về vần đề này rất đầy đủ với nhiều phương pháp khác nhau, trong nghiên cứu này trình bày tổng quan về một số giải thuật Optimization.

Các giải thuật cơ bản:

* Gradient Descent
* Stochastic Gradient Descent
* Momentum

Các giải thuật dựa vào chiến lược lựa chọn Learning Rate

* AdaGrad
* RMSPro
* Adam

Lựa chọn giải thuật Optimization nảo?

Hiện tại, các giải thuật Optimization hóa phổ biến nhất được sử dụng bao gồm SGD, SGD với momentum, RMSProp, RMSProp với momentum, AdaDelta và Adam.

Sự lựa chọn giải thuật để sử dụng, dường như phụ thuộc phần lớn vào sự quen thuộc của người dùng với giải thuật (để dễ dàng điều chỉnh siêu số).

**2.6. THUẬT TOÁN CNN TRONG NHẬN DIỆN HÀNH ĐỘNG**

**2.6.1. Thuật toán**

Một CN, gồm nhiều giai đoạn huấn luyện được xếp chồng lên nhau, được sử dụng để trích xuất các đặc trưng phân cấp.

Chập hình ảnh đầu vào với các bộ lọc huấn luyện khác nhau và các bias bổ sung, nhiều bản đồ đặc trưng được tạo trong lớp C1. Mỗi bản đồ đặc trưng trong S2 thu được bởi thao tác tổng hợp các bản đồ đặc trưng tương ứng trong lớp C1. Chập và tổng hợp cực đại trong lớp C3 và S4 thì giống trong lớp C1 và S2. Trong bước nhận diện cuối cùng, các đặc trưng thu được sau khi tổng hợp cực đại trong lớp S4 thì được mã hóa thành một vector 1 chiều.

Quá trình chập: sử dụng bộ lọc huấn luyên fx = 1+e-x, để ảnh chập đầu vào, sau đó them một bias bx, tạo thành lớp xoắn Cx. Quá trình tổng hợp cực đại: lấy pixel lơn nhất của 4 pixel lân cận, sau đó tính trọng số Wx+1 và thêm bias bx+1, sau đó truyền một hàm kích hoạt xích ma , tạo thành một bản đồ đặc trưng có kích thước được giảm đi 4 lần.

**2.6.2. Kiến trúc CNN được xem xét nghiên cứu**

Mô hình mạng được sử dụng trong luận văn gồm có 2 lớp Convolution đều có 64 filter với kích thước 5x5, xen kẽ 2 lớp Convolution là 2 lớp Max-pool với filter có kích thước 2x2. Tiếp sau đó là 2 lớp Fully-connected với kích thước lần lượt là 256 và 128. Cuối cùng, tập phân loại softmax cho phân bố xác suất trên các nhãn trong đầu ra, Nhãn có xác suất cao nhất được chọn làm dự đoán cuối cùng.

Tôi hứa sẽ hoàn thành nội dung báo cáo khóa luận tốt nghiệp đúng theo đề cương đã viết và đúng thời gian theo kế hoạch thực hiện khóa luận mà Khoa Công nghệ thông tin đã phổ biến.

*An Giang, ngày … tháng ….. năm 202…*

**Xác nhận của Giảng viên hướng dẫn Sinh viên thực hiện**

*(Ký và ghi rõ họ tên) (Ký và ghi rõ họ tên)*