ĐẠI HỌC THÁI NGUYÊN
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

LÊ CẨM HÀ

NGHIÊN CỨU MẠNG NƠRON CNN VÀ ỨNG DỤNG TRONG BÀI TOÁN PHÂN LOẠI ẢNH

LUẬN VĂN THẠC SĨ KHOA HỌC MÁY TÍNH

THÁI NGUYÊN - 2020

ĐẠI HỌC THÁI NGUYÊN
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

LÊ CẨM HÀ

NGHIÊN CỨU MẠNG NƠRON CNN VÀ ỨNG DỤNG TRONG BÀI TOÁN PHÂN LOẠI ẢNH

Chuyên ngành: Khoa học máy tính
Mã số: 8 48 01 01

LUẬN VĂN THẠC SĨ KHOA HỌC MÁY TÍNH

Giáo viên hướng dẫn: TS.Nguyễn Đình Dũng

THÁI NGUYÊN - 2020

LỜI CẢM ƠN

Luận văn này được hoàn thành tại Trường Đại học Công nghệ Thông tin và Truyền thông dưới sự hướng dẫn của TS. Nguyễn Đình Dũng. Tác giả xin bày tỏ lòng biết ơn tới các thầy cô giáo thuộc Trường Đại học Công nghệ Thông tin và Truyền thông, các thầy cô giáo thuộc Viện Công nghệ Thông tin – Viện Hàn lâm Khoa học và Công nghệ Việt Nam đã tạo điều kiện, giúp đỡ tác giả trong quá trình học tập và làm luận văn tại Trường, đặc biệt tác giả xin bày tỏ lòng biết ơn tới TS. Nguyễn Đình Dũng đã tận tình hướng dẫn và cung cấp nhiều tài liệu cần thiết để tác giả có thể hoàn thành luận văn đúng thời hạn.

Xin chân thành cảm ơn anh chị em học viên cao học và bạn bè đồng nghiệp đã trao đổi, khích lệ tác giả trong quá trình học tập và làm luận văn tại Trường Đại học Công nghệ Thông tin và Truyền thông – Đại học Thái Nguyên.

Cuối cùng tác giả xin gửi lời cảm ơn đến gia đình, những người đã luôn bên cạnh, động viên và khuyến khích tôi trong quá trình thực hiện đề tài.

Thái Nguyên, tháng 10 năm 2020

Học viên cao học

Lê Cẩm Hà

LỜI CAM ĐOAN

Tôi xin cam đoan luận văn này do chính tôi thực hiện, dưới sự hướng dẫn khoa học của TS. Nguyễn Đình Dũng, các kết quả lý thuyết được trình bày trong luận văn là sự tổng hợp từ các kết quả đã được công bố và có trích dẫn đầy đủ, kết quả của chương trình thực nghiệm trong luận văn này được tác giả thực hiện là hoàn toàn trung thực, nếu sai tôi hoàn toàn chịu trách nhiệm.

Thái Nguyên, tháng 10 năm 2020

Học viên

Lê Cẩm Hà

MỤC LỤC

LỜI CẢM ƠN i
LỜI CAM ĐOAN ii
DANH MỤC CÁC HÌNH ẢNH vii
DANH MỤC BẢNG BIỂU ix
MỞ ĐẦU 1
1. Tính khoa học và cấp thiết của đề tài 1
2. Đối tượng và phạm vi nghiên cứu của đề tài 2
3. Phương pháp luận nghiên cứu 2
4. Nội dung và bố cục của luận văn 2
CHƯƠNG 1TỔNG QUAN BÀI TOÁN PHÂN LOẠI ẢNH SỐ 3
1.1 Tổng quan xử lý ảnh số 3
1.1.1 Một số khái niệm cơ bản trong xử lý ảnh 3
1.1.2 Tổng quan về một hệ thống xử lý ảnh 4
1.1.3 Một số thao tác cơ bản trong xử lý ảnh 5
1.2 Biểu diễn ảnh trong máy tính 7
1.2.1 Ảnh màu 7
1.2.2 Ảnh xám 10
1.3 Phép tích chập trong xử lý ảnh 10
1.4 Lý thuyết phân loại ảnh số 13
1.4.1 Các khái niệm cơ bản 13
1.4.2 Phương pháp số phân loại ảnh 15
1.4.3 Phương pháp phân loại theo cấu trúc: 17
1.5 Một số thuật toán tiêu biểu trong phân loại ảnh 19
1.5.1 Thuật toán KNN 19
1.5.2 Thuật toán sử dụng mạng Nơ ron 20
1.5.3 Thuật toán SVM 21
1.6 Kết luận chương 1 21
CHƯƠNG 2 MẠNG NƠ RON CNN VÀ ỨNG DỤNG TRONG PHÂN LOẠI ẢNH 23

2.1 Các khái niệm chung về mạng nơron 23
2.1.1 Mạng nơron sinh học 23
2.1.2 Mạng nơron nhân tạo 24
2.1.3 Mô hình toán học và kiến trúc mạng nơron 27
2.1.4 Phân loại mạng nơ ron 30
2.1.5 Huấn luyện mạng nơron 31
2.2 Mạng nơron CNN 32
2.2.1 Giới thiệu 32
2.2.2 Kiến trúc mạng CNN 33
2.2.3 Ứng dụng CNN trong phân loại ảnh 37
2.3 Xây dựng mạng CNN cho phân loại ảnh 38
2.3.1 Trường tiếp nhận cục bộ (Local receptive fields) 38
2.3.2 Trọng số chia sẻ và độ lệch (Shared weights and biases) 42
2.3.3 Lớp chứa hay lớp tổng hợp (Pooling layer) 42
2.3.4 Cách chọn tham số cho CNN 45
2.4 Cập nhật một số hướng nghiên cứu về bài toán phân loại ảnh sử dụng mạng nơ ron CNN 45
2.4.1 Các nghiên cứu trên thế giới 45
2.4.2 Các nghiên cứu trên trong nước 46
2.5 Kết luận chương 48
CHƯƠNG 3 XÂY DỰNG CHƯƠNG TRÌNH MÔ PHỎNG ỨNG DỤNG MẠNG CNN TRONG PHÂN LOẠI ẢNH 49
3.1 Đặt vấn đề 49
3.2 Bài toán nhận dạng chữ viết tay 50
3.2.1 Mô tả bài toán 50
3.2.2 Các bước thực hiện 51
3.2.3 Một số kết quả đạt được 57
3.3 Bài toán giải mã Capcha 61
3.3.1 Mô tả bài toán 61
3.3.2 Các bước thực hiện 65

3.3.3 Một số kết quả đạt được 67
3.4 Kết luận chương 68
KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN 70
TÀI LIỆU THAM KHẢO 72

DANH MỤC CÁC TỪ VIẾT TẮT

Từ hoặc cụm từ	Từ tiếng Anh	Từ tiếng Việt
AI	Artificial Intelligence	Trí tuệ nhân tạo
ANN	Artificial Neural Network	Mạng nơron nhân tạo
CV	Computer Vision	Thị giác máy tính
CNN	Convolutional Neural Network	Mạng nơron tích chập
DL	Deep Learning	Học sâu
CAPCHA	Completely Automated Public Turing test to tell Computers and Humans Apart	Phép thử Turing công cộng hoàn toàn tự động để phân biệt máy tính với người
MCR	Miss Classification Rate	Tỷ lệ nhận dạng sai
RMSE	Root Mean Square Error	Sai số bình phương trung bình
MLP	Multilayer Neural Network	Mạng nơron nhiều lớp
MNIST	Modified National Institute of Standards and Technology database	Cơ sở dữ liệu về chữ số viết tay
ReLU	Rectified Linear Units	Hàm tinh chỉnh các đơn vị tuyến tính

DANH MỤC CÁC HÌNH ẢNH

Hình 1.1. Các giai đoạn chính trong xử lý ảnh 5
Hình 1.2. Minh họa hệ màu RGB 8
Hình 1.3. Ví dụ về ảnh màu 8
Hình 1.4. Biểu diễn ảnh theo tensor 3 chiều 9
Hình 1.5. Ví dụ về ảnh xám 10
Hình 1.6. Minh họa phép tích chập trong xử lý ảnh 11
Hình 1.7. Ma trận đầu ra Y khi chập ảnh X với kernel W 11
Hình 1.8. Stride=1, padding=1 12
Hình 1.9. Stride=2, padding=1 12
Hình 1.10. Một số bộ lọc Kerne1 trong xử lý ảnh 13
Hình 1.11. Phương pháp lưới 16
Hình 1.12. Phương pháp cung 16
Hình 1.13. Biểu diễn mẫu bằng tập kí hiệu 18
Hình 1.14. Minh họa thuật toán KNN 19
Hình 2.1. Cấu trúc cơ bản của nơron sinh học 23
Hình 2.2. Nơron nhân tạo 25
Hình 2.3. Mô hình toán học mạng nơron nhân tạo 27
Hình 2.4. Nơron 1 đầu vào với hàm hoạt hoá là hàm hardlimit 29
Hình 2.5. Phân loại mạng nơ ron 30
Hình 2.6. Học có giám sát 31
Hình 2.7. Học không có giám sát 31
Hình 2.8. Học tăng cường 32
Hình 2.9. Cách máy tính “nhìn” một hình [16] 32
Hình 2.10. Mạng nơ-ron thông thường (trái) và CNN (phải) 34
Hình 2.11. Kiến trúc mạng CNN 34
Hình 2.12. Max pooling kích thước 2×2 36
Hình 2.13. Lớp kết nối đầy đủ 36
Hình 2.14. Các bước phân loại ảnh sử dụng mạng CNN 37

Hình 2.16. Lớp input gồm 28x28 nơ ron cho nhận dạng chữ từ tập dữ liệu MNIST 38
Hình 2.17. Kết nối vùng 5x5 nơ ron input với nơ ron lớp ẩn 39
Hình 2.18. Vị trí bắt đầu của trường tiếp nhận cục bộ.....39
Hình 2.19. Vị trí thứ 2 của trường tiếp nhận cục bộ và nơ ron lớp ẩn.....40
Hình 2.20. Trường tiếp nhận cục bộ với ba bản đồ đặc trưng 40
Hình 2.21. Trường tiếp nhận cục bộ với 20 bản đồ đặc trưng.....41
Hình 2.22. Ví dụ về Max pooling 2x2 43
Hình 2.23. Max pooling với ba bản đồ đặc trưng.....43
Hình 2.24. Một kiến trúc mạng CNN cho nhận dạng chữ viết từ dữ liệu MNIST ...44
Hình 3.1. Giao diện chính của chương trình mô phỏng.....49
Hình 3.2. Chữ viết tay số “5” từ bộ dữ liệu MNIST.....50
Hình 3.3. Giao diện thiết kế mạng CNN.....55
Hình 3.4. Mạng CNN cơ bản 55
Hình 3.5. Tiến trình luyện mạng với kernel 7 x 7 and 8 bản đồ đặc trưng. 56
Hình 3.6. Giao diện chương trình nhận dạng chữ viết tay. 60
Hình 3.7. Một số mẫu captcha 62
Hình 3.8. Một số kết quả tấn công captcha.....63
Hình 3.9. Hai cách tiếp cận để nhận dạng captcha bằng CNN 64
Hình 3.10. Kiểu dữ liệu captcha dùng trong bài toán nhận dạng.....65
Hình 3.11. Kí tự W và Q bị dính với nhau.....65
Hình 3.12. Giãn nở ký tự trong captcha để dễ phát hiện vùng liên thông 66
Hình 3.13. Phát hiện thành phần liên thông.....66
Hình 3.14. Một mẫu captcha có 2 ký tự dính liền nhau.....66
Hình 3.15. Vùng nhận dạng liên tục nhận 2 ký tự vào 1 ảnh cắt, chưa tốt.....66
Hình 3.16. Kết quả sau khi dùng thủ thuật cắt đôi vùng nhận các ký tự liền nhau ..66
Hình 3.17. Ví dụ tập các ảnh kí tự đã được cắt và xếp theo thư mục 67
Hình 3.18. Chương trình mô phỏng nhận dạng mã Captcha 68