

ĐẠI HỌC ĐÀ NẴNG
TRƯỜNG ĐẠI HỌC BÁCH KHOA

VÕ HOÀNG THÀNH

**ỨNG DỤNG KỸ THUẬT HỌC SÂU NHẬN DẠNG TIỀN
TRONG CÁC HỆ THỐNG GIAO DỊCH TỰ ĐỘNG**

Chuyên ngành: Khoa học máy tính

Mã số: 8480101

TÓM TẮT LUẬN VĂN THẠC SĨ KỸ THUẬT

Đà Nẵng - Năm 2018

Công trình được hoàn thành tại
TRƯỜNG ĐẠI HỌC BÁCH KHOA

Người hướng dẫn khoa học: TS. HOÀNG VĂN DŨNG

Phản biện 1: TS. LÊ THỊ MỸ HẠNH

Phản biện 2: TS. PHẠM XUÂN HẬU

Luận văn được bảo vệ trước Hội đồng chấm Luận văn tốt nghiệp thạc sĩ kỹ thuật họp tại Trường Đại học Bách khoa Đà Nẵng vào ngày 05 tháng 01 năm 2019

Có thể tìm hiểu luận văn tại:

- Trung tâm Học liệu và Truyền thông Trường Đại học Bách khoa
Đại học Đà Nẵng
- Thư viện Khoa Công nghệ thông tin, Trường Đại học Bách khoa
Đại học Đà Nẵng

MỞ ĐẦU

1. Lý do chọn đề tài

Ngày nay, kỹ thuật học sâu (deep learning) đang tạo nên sự phát triển mạnh mẽ về công nghệ ở trên toàn thế giới, nó ảnh hưởng tới rất nhiều lĩnh vực trong cuộc sống hiện đại (robot, IoT, tài chính, y tế...).

Kỹ thuật học sâu thể hiện tính ưu việt hơn các kỹ thuật khác trong việc nhận diện hình ảnh cũng như nhiều lĩnh vực khác, những kết quả nghiên cứu gần đây áp dụng kỹ thuật này cho độ chính xác cao hơn. Nhiều công ty hoạt động về lĩnh vực công nghệ trên thế giới đang đầu tư rất nhiều nguồn lực vào nghiên cứu và phát triển nhằm tạo ra các sản phẩm ứng dụng sử dụng kỹ thuật học sâu. Những công ty lớn như Google, Baidu... cũng sử dụng kỹ thuật học sâu dùng trong các thuật toán nhận dạng giọng nói, hình ảnh và đã mang kết quả tốt khi đem vào ứng dụng trong thực tế. Các lĩnh vực về công nghệ đã dùng kỹ thuật học sâu có thể kể đến là: công nghệ nhận diện giọng nói (các trợ lý ảo giọng nói như Siri của Apple, Cortana của Microsoft, Alexa của Amazon, DeepVoice của Baidu...), dịch tự động và các công nghệ xử lý ngôn ngữ tự nhiên, công nghệ nhận diện hình ảnh (Google, Microsoft, Facebook, Baidu,...).

Công nghệ nhận diện hình ảnh sử dụng kỹ thuật học sâu đang được nhiều nhà nghiên cứu tìm hiểu và đưa vào ứng dụng thực tế; nhóm nghiên cứu về trí tuệ nhân tạo của Facebook đã tạo ra phần mềm DeepFace[1] có khả năng nhận dạng khuôn mặt tốt như con người với độ chính xác khoảng 97.35%, công trình được công bố năm 2014, sử dụng khoảng 4 triệu ảnh khuôn mặt của hơn 4000 người để huấn luyện cho mạng nơron nhiều lớp, kết quả thu được đã vượt qua các kỹ thuật được nghiên cứu đề xuất trước đó. Trong lĩnh

vực y tế có nhiều công ty khởi nghiệp (Enlitic, Freenome, Merck, Atomwise ...) về y sinh đã đưa ra những sản phẩm điện toán có khả năng đọc X-quang, MRI và phim chụp CT một cách nhanh chóng và chính xác bệnh chuẩn hơn cả bác sỹ điều trị[34].

Ở Việt Nam, việc nghiên cứu về thị giác máy tính (Computer Vision) và các kỹ thuật xử lý ảnh, trí tuệ nhân tạo (AI) để áp dụng giải quyết các bài toán nhận dạng tiền trong các hệ thống giao dịch tự động chưa được nghiên cứu nhiều. Ứng dụng Money Reader về nhận dạng Tiền giấy dựa trên hình ảnh và tự động chuyển đổi tỷ giá (Trường Đại học Công nghệ - Hà Nội)[35] là một trong số ít về nghiên cứu này.

Cuộc cách mạng công nghiệp lần thứ 4 bước đầu đã có tác động sâu sắc đến định hướng phát triển xã hội. Các nhà lãnh đạo đang hồi thúc phát triển chính quyền điện tử, đô thị thông minh, y tế thông minh,... Như vậy, một tất yếu cần phải có các hệ thống tương tác giao dịch thông minh giữa con người và máy tính từ các hệ thống lớn như trung tâm hành chính công giao dịch tự động, các siêu thị mua bán tự động đến các hệ thống nhỏ như dịch vụ đổi tiền tự động, hệ thống thu tiền tự động, hệ thống in các loại văn bản chứng thực tự động,... Các giao dịch tự động sẽ giúp con người tiết kiệm thời gian, chi phí về nhân lực làm tăng hiệu quả dịch vụ, kinh doanh. Bên cạnh tính chính xác, nó còn phải đáp ứng yêu cầu về tính linh động, đa dạng và mềm dẻo trong yêu cầu giao dịch, ví dụ như công việc thu phí, tiền trong mua bán hàng tự động thì ngoài những cách giao dịch qua thẻ ngân hàng, chuyển khoản; cách tiếp nhận tiền mặt vẫn được người tiêu dùng lựa chọn nhiều. Trong giao dịch tiền mặt, yêu cầu hệ thống phải nhận dạng được nhiều mệnh giá, loại tiền và độ chính xác

cao là các yêu cầu hết sức quan trọng nhằm đáp ứng các yêu cầu thực tế trong hệ thống giao dịch tự động.

Với những lý do như trên, tôi xin đề xuất đề tài luận văn cao học: “*Ứng dụng kỹ thuật học sâu nhận dạng tiền trong các hệ thống giao dịch tự động*”

2. Mục đích và ý nghĩa đề tài

2.1. Mục đích

Nghiên cứu các kỹ thuật trí tuệ nhân tạo, đặc biệt là kỹ thuật học sâu và các ứng dụng của nó trong thực tế.

Nghiên cứu và đề xuất mô hình nhận dạng tiền với nhiều mệnh giá, loại tiền khác nhau và xác thực chúng để ứng dụng trong các hệ thống giao dịch tự động.

2.2. Ý nghĩa khoa học

- Kế thừa những nghiên cứu về các hệ thống nhận dạng tiền trước đây.
- Tìm hiểu, mô tả về mô hình hệ thống nhận dạng tiền mặt.
- So sánh giữa các kỹ thuật được sử dụng trong hệ thống nhận dạng.
- Đồng thời đánh giá hiệu suất của phương pháp bằng các thực nghiệm trên tập dữ liệu thực.
- Vận dụng trí tuệ nhân tạo trong việc giải quyết bài toán nhận dạng tiền mặt với khả năng xử lý được với nhiều mệnh giá tiền, loại tiền một cách tự động.

2.3. Ý nghĩa thực tiễn

Dùng kỹ thuật học sâu để xây dựng mô hình nhận dạng tiền để ứng dụng trong các hệ thống giao dịch tự động.

3. Mục tiêu và nhiệm vụ

3.1. Mục tiêu

- Nắm vững được kỹ thuật học sâu để áp dụng trong bài toán nhận dạng tiền.
- Đề xuất được mô hình nhận diện tiền và xác thực với các loại tiền, mệnh giá khác nhau để ứng dụng trong hệ thống giao dịch tự động.

3.2. Nhiệm vụ

- Nghiên cứu kỹ thuật trí tuệ nhân tạo, mạng neural học sâu.
- Phát biểu bài toán
- Phân tích và cài đặt giải thuật cho bài toán nhận dạng tiền.
- Đánh giá kết quả theo yêu cầu của đề tài.

4. Đối tượng và phạm vi nghiên cứu

4.1. Đối tượng nghiên cứu

- Nghiên cứu các kỹ thuật trí tuệ nhân tạo, kỹ thuật mạng neural học sâu để ứng dụng trong phân tích hình ảnh, nhận dạng mệnh giá tiền, loại tiền.
- Nghiên cứu hệ thống nhận dạng tiền (tiền giấy) để ứng dụng cho các giao dịch tự động

4.2. Phạm vi nghiên cứu

Nghiên cứu và đề xuất mô hình nhận diện tiền (tiền giấy) cho hệ thống giao dịch tự động.

5. Phương pháp nghiên cứu

5.1. Phương pháp lý thuyết

- Phương pháp phân tích điều tra số liệu: thu thập và nghiên cứu các tài liệu có liên quan đến đề tài.
- Phương pháp nghiên cứu tài liệu: các kỹ thuật xử lý ảnh, trí tuệ nhân tạo và đặc biệt là kỹ thuật học sâu.

5.2. Phương pháp thực nghiệm

- Nghiên cứu và khai thác các mô hình nhận diện hình ảnh đã được đề xuất.
- Xây dựng chương trình ứng dụng vào nhận dạng mệnh giá và các điểm đặc trưng của tờ tiền.
- Kiểm tra, thử nghiệm, nhận xét và đánh giá kết quả.

6. Cấu trúc của luận văn

Luận văn được tổ chức thành ba chương chính với các nội dung cụ thể như sau:

Chương 1: Trong chương cơ sở lý thuyết này, chúng tôi trình bày những kiến thức liên quan đến những nghiên cứu về cách tiếp cận bài toán nhận dạng tiền giấy ứng dụng cho các hệ thống giao dịch tự động. Từ đó, có cái nhìn tổng quan về các hướng tiếp cận và giải pháp để ứng dụng trong bài toán nhận dạng tiền giấy, từ các phương pháp xử lý ảnh đơn giản và học máy truyền thống đến các phương pháp học sâu.

Chương 2: Ở chương này sẽ trình bày một số mạng neural tích chập điển hình được ứng dụng nhiều trong thực tế. Qua đó để vận dụng vào thiết kế mô hình hệ thống nhận dạng tiền giấy, từ việc mô tả bài toán đến việc phân tích hệ thống, xây dựng hoàn chỉnh mô hình nhận dạng tiền. Đưa ra một giải pháp để giải pháp để giải quyết bài toán nhận dạng tiền, phương pháp đề xuất là sự kết hợp của mạng neural sâu DNN dùng để huấn luyện dữ liệu trích xuất đặc trưng và máy vector hỗ trợ SVM để phân loại.

Chương 3: Trong chương này, chúng tôi sẽ cài đặt chương trình thực nghiệm và đánh giá kết quả đạt được. Cơ sở dữ liệu ảnh sẽ được mô tả cụ thể, quá trình huấn luyện của giải pháp và kết quả đạt được của luận văn sẽ được trình bày cụ thể tại đây.

Chương 1: CƠ SỞ LÝ THUYẾT

1.1. Tổng quan về xử lý ảnh

1.1.1. Các khái niệm về ảnh số, điểm ảnh, màu sắc

1.1.2. Các dạng ảnh

1.1.3. Các thành phần trong hệ thống xử lý ảnh

1.2. Cơ bản về học máy

1.2.1. Khái niệm học máy

Học máy (machine learning) là một lĩnh vực nghiên cứu trong trí tuệ nhân tạo, nó liên quan đến việc nghiên cứu, xây dựng các kỹ thuật để các hệ thống “học” tự động từ dữ liệu từ đó giải quyết những vấn đề cụ thể nào đó.

1.2.2. Phân loại phương pháp học máy

- Học có giám sát (supervised learning) là cách học với tập dữ liệu huấn luyện ban đầu được gán nhãn hoàn toàn từ trước.

- Học không có giám sát (unsupervised learning) là cách học với tập dữ liệu huấn luyện ban đầu hoàn toàn chưa được gán nhãn.

- Học bán giám sát (semi-supervised learning) là cách học với tập dữ liệu huấn luyện gồm cả dữ liệu đã được gán nhãn và dữ liệu chưa được gán nhãn.

- Học tăng cường (reinforcement learning) là cách học để giải quyết các bài toán giúp cho một hệ thống tự động xác định hành vi dựa trên hoàn cảnh để đạt được lợi ích cao nhất.

1.2.3. Ứng dụng của học máy

Học máy ngày càng có nhiều ứng dụng rộng khắp trong các ngành khoa học và công nghiệp đang làm việc với hàm lượng lớn dữ liệu, tầm quan trọng của công nghệ Machine Learning được nâng lên ở vị thế mới, đặc biệt những ngành cần phân tích khối lượng dữ liệu khổng lồ.

1.3. Một số phương pháp học máy truyền thống

1.3.1. Mạng neural

Mạng neural (nơ-ron) hay còn gọi mạng neural nhân tạo (Artificial Neural Networks) là một mô hình toán học được xây dựng dựa theo mô phỏng của các mạng neural sinh học. Mạng neural bao gồm một nhóm các neural nhân tạo (còn gọi là các nút) được nối với nhau và xử lý thông tin bằng cách truyền theo các kết nối, tính giá trị mới tại các nút.

Kiến trúc mạng neural

Kiến trúc mạng neural gồm tập các nút gọi là các neural, tập các cung để kết nối các neural lại với nhau. Tập các nút được tổ chức thành các lớp, gọi là các layer. Trong đó có một lớp đầu vào (input layer) và một lớp đầu ra (output layer), các lớp giữa gọi là các lớp ẩn (hidden layer). Mỗi cung kết nối 2 cặp neural với nhau, trong đó có một nút vào và một nút ra. Mỗi liên hệ giữa các nút được thể hiện thông qua hàm biến đổi (hàm truyền) cùng với bộ trọng số tương ứng cho hàm truyền.

Quá trình học của mạng neural

Mạng ANN có thể là một trong ba dạng: học có giám sát, học không giám sát và học tăng cường.

Mạng Perceptron

Perceptron là mô hình mạng neural được phát triển bởi nhà khoa học Frank Rosenblatt[5]. Ngày nay nó trở nên phổ biến trong việc sử dụng các model khác nhau của Neural nhân tạo.

1.3.2. Máy vector hỗ trợ (SVM)

Giới thiệu

Thuật toán SVM là một thuật toán phân loại, thuộc phương pháp học có giám sát. Nó nhận dữ liệu vào và phân loại chúng thành

hai lớp khác nhau. Thuật toán huấn luyện SVM xây dựng một mô hình để phân loại hai lớp dữ liệu cho trước. Thuật toán SVM được tìm ra bởi Corinna và Vapnik[7] năm 1995.

Phân loại tuyến tính

Phân loại tuyến tính lề mềm (Soft Margin Classifiers)

Hàm nhân

Phân loại phi tuyến (nonlinear classification)

1.4. Kỹ thuật học sâu

1.4.1. Tổng quan về học sâu

Học sâu là một lĩnh vực nghiên cứu của ngành máy học, nó dựa trên một tập các thuật toán hướng tới giải quyết mô hình dữ liệu trừu tượng ở mức cao bằng cách sử dụng nhiều lớp xử lý hoặc bằng cách kết hợp nhiều biến đổi phi tuyến với kiến trúc phức tạp hơn nhiều so với phương pháp học máy truyền thống.

1.4.2. Mạng neural sâu

Mạng Neural sâu (Deep Neural Network - DNN) là một mạng neuron nhân tạo (Artificial Neural Network - ANN) với nhiều đơn vị lớp ẩn kết nối với nhau từ lớp đầu vào cho tới lớp đầu ra. DNN tìm kiếm và sử dụng các phép toán chuẩn xác để biến dữ liệu đầu vào thành dữ liệu đầu ra, cho dù mạng có mối quan hệ tuyến tính hay mối quan hệ phi tuyến tính phức tạp. Các mạng neural sâu khác với những mạng neural đơn giản ở mức độ sâu của mạng, nghĩa là số nút trong mỗi lớp và số lớp ẩn lớn hơn nhiều so với mạng neural đơn giản.

1.4.3. Mạng neural tích chập (CNN)

Mạng CNN phát triển dựa trên ý tưởng cải tiến cách thức của các mạng neural nhân tạo truyền thẳng (Feedforward Neural

Network) bị hạn chế học với các dữ liệu hình ảnh bởi sự liên kết quá đầy đủ giữa các điểm ảnh vào node.

Lớp Convolutional (lớp tích chập)

Lớp tích chập là lớp quan trọng nhất của mạng CNN, nó thể hiện ý tưởng của mạng neural tích chập. Thay vì kết nối toàn bộ điểm ảnh, phép biến đổi tích chập là phép kết nối cục bộ, với một bộ trọng số được gọi là mặt nạ lọc (filter kernel) hay mặt nạ tích chập.

Lớp pooling (lớp giảm kích thước)

Lớp pooling là lớp thực hiện các tính toán chính trong cấu trúc mạng CNN nhằm mục đích làm giảm kích thước các neural ở lớp trước, qua đó làm giảm độ phức tạp tính toán trong xử lý và biểu diễn đặc trưng mà vẫn giữ được thông tin quan trọng nhất của dữ liệu.

Lớp ReLU (lớp hiệu chỉnh hay lớp sửa dữ liệu)

Lớp ReLU về cơ bản là chỉ áp dụng hàm truyền ReLU nhằm hiệu chỉnh giá trị các nút trong mạng, nó có vai trò quan trọng trong quá trình xử lý của mạng CNN nhằm loại bỏ những phần tử âm và thay thế tất cả các phần tử âm bởi giá trị không (0).

Lớp normalization (lớp chuẩn hóa)

Lớp chuẩn hóa thường được sử dụng giữa lớp tích chập và lớp ReLU nhằm mục đích tăng tốc độ huấn luyện và làm giảm sự nhạy cảm việc khởi tạo tham số của mạng.

Lớp fully connection (lớp kết nối đầy đủ)

Lớp kết nối đầy đủ có kiến trúc mạng tương tự kết nối trong mạng neural truyền thẳng truyền thống. Các node trong lớp kết nối đầy đủ được kết với toàn bộ các node ở lớp tiếp theo.

Chương 2: MẠNG NEURAL TÍCH CHẬP TRONG NHẬN DẠNG TIỀN GIẤY

2.1. Một số mạng neural tích chập điển hình trong nhận dạng đối tượng

2.1.1. Mạng LeNet

Mạng LeNet là một trong những kiến trúc mạng CNN được phát triển bởi Yann Lecun, được sử dụng rất thành công việc nhận dạng chữ viết tay, đặc biệt là với kiến trúc LeNet-5 và mô hình pretrain[12]. Mạng Lenet5 chỉ có 60.000 tham số.

2.1.2. Mạng AlexNet

Mạng AlexNet phát triển bởi nhóm tác giả: Alex Krizhevsky, Ilya Sutskever và Geoff Hinton. Năm 2012, mô hình mạng này được giới thiệu với cấu trúc tương tự như LeNet nhưng với số lượng neuron, filter và layer lớn hơn. Mô hình pretrain của mạng AlexNet được sử dụng rất thành công và phổ biến trong lĩnh vực nhận dạng đối tượng ảnh. Được xem là mạng neural đầu tiên phổ biến rộng rãi khả năng ứng dụng của mạng CNN.

2.1.3. Mạng ZFNet

Mạng ZFNet (ZF viết tắt của Zeiler và Fergus) là một mô hình kiến trúc mạng CNN được phát triển bởi Matthew Zeiler và Rob Fergus[15]. Mạng ZFNet có kiến trúc chiến thắng cuộc thi trên ImageNet 2013 (ILSVRC 2013). Mạng neural này được phát triển từ nền tảng mạng AlexNet với việc tinh chỉnh các hyperparameter (filter size, stride,...) với những layer gần input layer, filter size và stride nhỏ hơn.

2.1.4. Mạng GoogLeNet

Mạng GoogLeNet được đánh giá là mạng học sâu tốt nhất năm 2014, là mạng có kiến trúc chiến thắng cuộc thi trên ImageNet 2014

(ILSVRC 2014). Mạng này được đề xuất bởi nhóm nghiên cứu của C.Szegedy của công ty Google, kiến trúc mạng gồm 22 lớp sâu. Nó làm giảm số lượng tham số từ 60 triệu của mạng AlexNet xuống còn 4 triệu.

Ở mạng GoogLeNet đã đề cập đến một khái niệm mới Inception Module. Trong Module sử dụng các bộ lọc convolution nhiều mức, tỷ lệ kích thước khác nhau như 5×5 , 3×3 và 1×1 .

2.1.5. Mạng VGGNet

Mạng VGGNet là mạng CNN đạt giải nhì cuộc thi trên ImageNet 2014 (ILSVRC 2014). Mạng VGGNet được nghiên cứu và đề xuất bởi Simonyan và Zisserman [16]. Mạng VGGNet được xem là mạng tốt nhất năm 2015, nó bao gồm 16 lớp tích chập và có kiến trúc rất thống nhất. Mạng sử dụng các bộ lọc tích chập 3×3 và bộ lọc pooling 2×2 , kích thước này được thống nhất từ lớp đầu đến lớp cuối của mạng.

2.2. Mô tả bài toán nhận dạng tiền giấy

Trong bài toán nhận dạng tiền giấy này, chúng tôi tập trung nghiên cứu xây dựng mạng neural sâu để khai thác các đặc trưng, cung cấp các đặc trưng này cho phương pháp học máy truyền thống như SVM để giải quyết các vấn đề nhiều loại dữ liệu tiền tệ khác nhau. Những vân (mosaic) của thông tin tiền giấy được phân tích và tìm ra các mối quan hệ để phân biệt các đặc trưng của vân (mosaic) tiền giấy tương ứng. Ngoài ra, nhiệm vụ hỗ trợ dữ liệu tăng cường để làm cho dữ liệu huấn luyện đủ lớn và bao gồm điều kiện cho các tình huống trong việc “học” của mạng CNN, giúp cải thiện đáng kể về kết quả nhận dạng tiền giấy. Trong nghiên cứu này, chúng tôi hy vọng sẽ cải thiện kết quả dựa trên việc phân biệt đặc trưng được tách rút bằng cách sử dụng kết hợp mạng DNN và máy vector hỗ trợ (SVM).

2.3. Phân tích và thiết kế hệ thống

2.3.1. Phân tích tổng quan hệ thống

Hệ thống nhận dạng tiền giấy được xây dựng dựa trên ý tưởng kết hợp sử dụng mạng neural sâu và học máy truyền thống SVM. Kiến trúc mạng neural sâu được xây dựng để trích xuất các đặc trưng phân biệt bằng kích hoạt đầu ra lớp kết nối đầy đủ. Vector đặc trưng của mẫu được cấp cho SVM cho quyết định cuối cùng của nhận dạng tiền giấy.

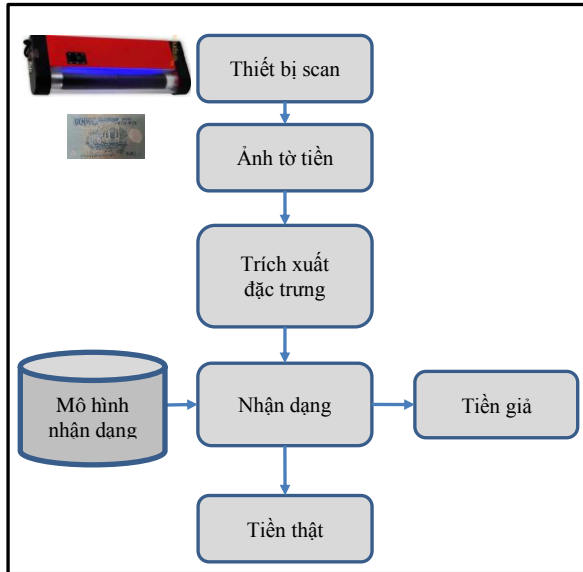
2.3.2. Phân tích chức năng của hệ thống

Trong phương pháp nghiên cứu của chúng tôi, thông tin vân tiền được sử dụng như một dấu hiệu chính để nhận dạng. Những thông tin về vân tiền được kiểm tra để cho phép nhận dạng giá trị đúng của tờ tiền, phát hiện tiền giả và mệnh giá của tờ tiền. Trong hệ thống nhận dạng tiền giấy này, nó có thể được mở rộng với số lượng lớn các mệnh giá khác nhau và có thể nâng cao chất lượng bằng cách thêm một tập dữ liệu để học máy mà không cần bất kỳ sự can thiệp thủ công nào. Hệ thống tổng quan để kiểm tra tiền giấy được trình bày trong hình 2.8.

Mô hình nhận dạng tiền giấy được xây dựng dựa trên một số bước sau:

- Tiền giấy cần nhận dạng đưa vào thiết bị scan.
- Ảnh của tờ tiền được lấy ra từ thiết bị scan.
- Từ ảnh của tờ tiền đưa vào mạng DNN huấn luyện và trích xuất đặc trưng.
- Sau khi có được đặc trưng của tờ tiền thì đem vào máy phân loại SVM để nhận dạng.
- Kết quả thu được sẽ thông báo mệnh giá tiền thật hoặc thông báo tiền giả.

Hệ thống bao gồm những module chính sau: mô hình “học” tiền giấy, trích xuất đặc trưng, nhận dạng tiền giấy.

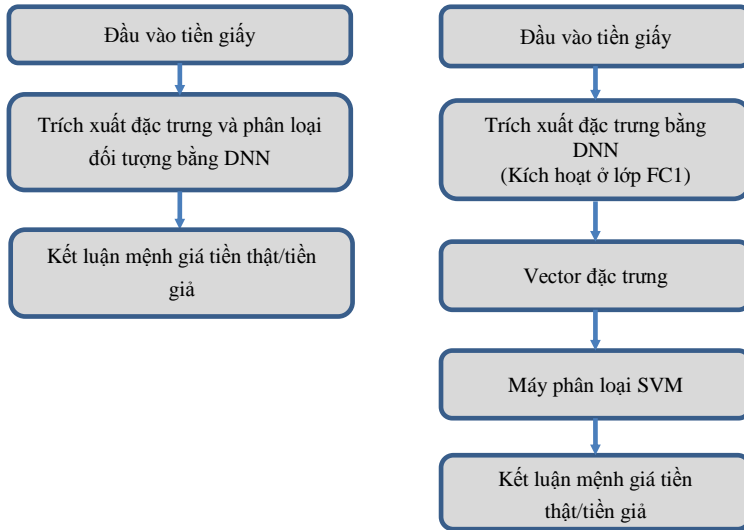


Hình 2.8. Tổng quan kiến trúc của quá trình nhận dạng tiền giấy

2.3.3. Sơ đồ giải pháp nhận dạng tiền giấy

Trong luận văn này, chúng tôi không trình bày tất cả các kỹ thuật để xây dựng một hệ thống đầy đủ, thay vào đó chúng tôi chỉ tập trung vào một số kỹ thuật nhận dạng các giá trị, mệnh giá tiền giấy cho các loại tiền và phát hiện tiền giả. Để cải thiện độ chính xác của hệ thống, giai đoạn nhận diện được xây dựng dựa trên mạng CNN sử dụng một số khối Inception cho mục đích khai thác đặc trưng và SVM để phân loại xem sơ đồ ở Hình 2.9 (b). Chúng tôi cũng triển khai hệ thống chỉ sử dụng mạng CNN sâu cho việc trích xuất đặc trưng và nhận dạng cho mệnh giá tiền giấy thể hiện qua Hình 2.9 (a). Kết quả thực nghiệm đã chứng minh rằng phương pháp CNN + SVM với phương pháp CNN là tương đương nhau.

Tổng quan về hệ thống nhận dạng tiền có thể được thực hiện bằng cách chỉ sử dụng kỹ thuật học sâu hoặc kết hợp giữa học sâu và học nông, được trình bày trong Hình 2.9.



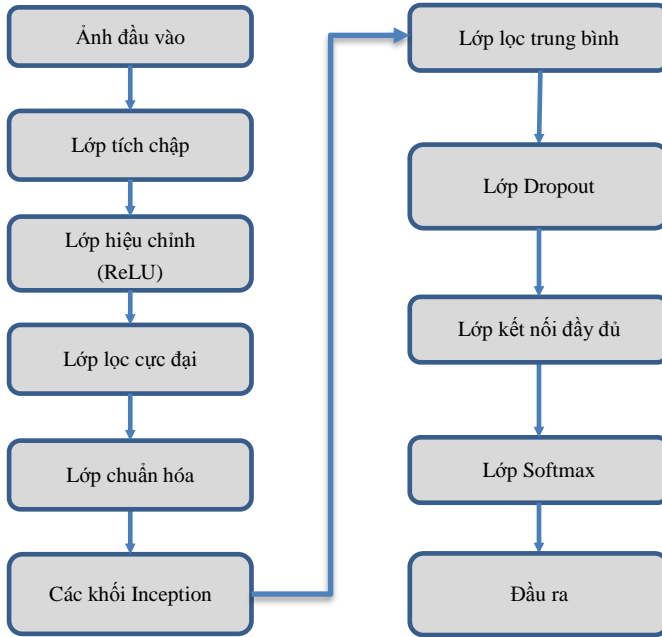
(a) Nhận dạng tiền bằng kỹ thuật học sâu (b) Nhận dạng tiền bằng cách kết hợp DNN và SVM

Hình 2.9. Kiểm tra tiền giấy dựa trên các kỹ thuật DNN và SVM

2.3.4. Thiết kế mạng neural sâu trong bài toán nhận dạng tiền giấy

Trong nghiên cứu này, để đơn giản hóa trong ứng dụng chúng tôi đã đề xuất một phương pháp tiếp cận mới để nhận dạng tiền giấy dựa trên sự kết hợp của mô hình DNN để trích xuất đặc trưng và máy phân loại SVM nhận dạng mệnh giá tiền và phân biệt tiền giả. Mô hình DNN cơ bản dựa trên nhiều kênh của 2DCNN và bốn khối Inception, hỗ trợ cho việc tận dụng lợi thế của đặc trưng cục bộ và

khai thác đặc trưng toàn cục. Mô hình 2DCNN tổng thể để khai thác đặc trưng được minh họa trong Hình 2.10.

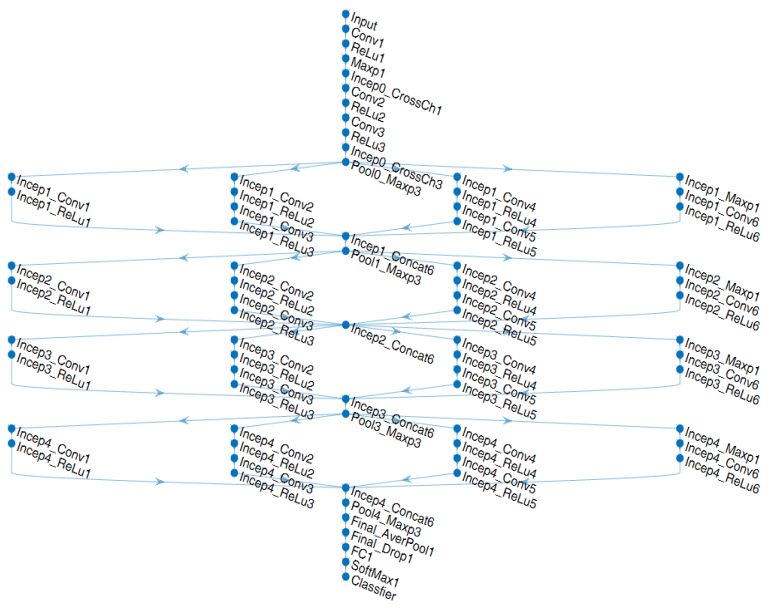


Hình 2.10. Sơ đồ tổng thể của kiến trúc mạng neural sâu trong nhận dạng tiền giấy

Đầu vào của mô hình kiến trúc mạng này là hình ảnh của tờ tiền được lấy từ các thiết bị scan, được đưa vào lớp tích chập để giảm kích thước ảnh, sau đó đưa vào lớp hiệu chỉnh (ReLU) để loại bỏ những giá trị âm (đưa về giá trị 0) giúp các hệ thống xử lý giữ chắc chắn mô hình bằng cách làm cho các giá trị đã được học khỏi việc bị mắc kẹt gần giá trị 0. Dữ liệu tiếp tục được đưa xuống lớp lọc cực đại để giảm kích thước, sau đó chuyển qua lớp chuẩn hóa dữ liệu, tiếp theo ta đưa vào khối Inception để làm cho mạng có độ sâu và

rộng hơn, tiếp theo ta đưa dữ liệu vào lớp lọc trung bình để giảm kích thước, sau đó ta cho dữ liệu qua lớp Dropout để giải quyết vấn đề quá khớp với dữ liệu (overfitting). Tiếp theo, dữ liệu được đưa vào lớp kết nối đầy đủ với mục đích là để nhận dạng đối tượng với độ chính xác cao hơn. Cuối cùng ta cho dữ liệu vào lớp đầu ra Softmax để phân lớp dữ liệu và chuyển dữ liệu ra ở đầu ra.

Hệ thống nhận dạng tiền giấy được xây dựng dựa trên cơ sở kiến trúc Inception để thực hiện quy trình trích xuất đặc trưng, như minh họa trong Hình 2.11.



Hình 2.11. Lược đồ của từng mạng Inception.

Chương 3: THỰC NGHIỆM, ĐÁNH GIÁ KẾT QUẢ

3.1. Cài đặt chương trình

3.1.1. Môi trường và công cụ lập trình

3.1.2. Thiết lập các lớp của mạng DNN trong matlab

3.2. Cơ sở dữ liệu thực nghiệm

3.2.1. Nguồn dữ liệu và mô tả

Nguồn dữ liệu để đưa vào mô hình huấn luyện nhận dạng tiền giấy được tạo ra bằng cách thu thập các mệnh giá tiền thật VND gồm 6 mệnh giá tiền giấy polymer: 10.000đ (10k), 20.000đ (20k), 50.000đ (50k), 100.000đ (100k), 200.000đ (200k), 500.000đ (500k), các tờ tiền này đã qua sử dụng. Tiền giả được thu thập trên chất liệu khác nhau và các hình ảnh không phải tiền. Ở trong một hệ thống giao dịch tự động khi đưa tiền vào máy kiểm tra thì sẽ không phân biệt mặt trước, mặt sau, quay ngược hay thuận chiều (nghĩa là chỉ cần đưa vào tiền 1 mặt với hướng bất kỳ). Sau khi có được các mệnh giá, chúng tôi tiến hành đưa các mẫu này đưa qua thiết bị scan để quét lấy hình ảnh của từng tờ tiền, dữ liệu thu được lưu vào từng thư mục khác tương ứng với các mệnh giá ở trong máy vi tính huấn luyện. Hình ảnh thu được của tờ tiền có kích thước tương ứng với mẫu thật. Sau đó, ở tập dữ liệu huấn luyện các ảnh tờ tiền được chuẩn hóa về cùng độ phân giải 200x450 pixel.

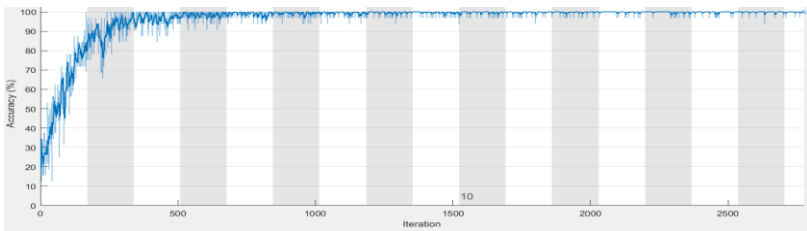
3.2.2. Tăng cường dữ liệu

Do vấn đề mất cân đối về dữ liệu huấn luyện, nên điều quan trọng là phải tăng cường dữ liệu, điều này giúp đủ số dư và đủ dữ liệu để xây dựng các tham số mạnh mẽ của hệ thống nhận dạng tiền tệ. Vấn đề tăng cường dữ liệu có nghĩa là tăng số lượng tập dữ liệu hình ảnh. Trong mạng DNN, càng có nhiều dữ liệu thì sẽ tốt hơn cho

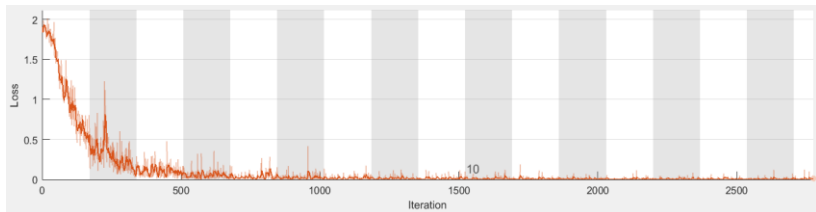
mô hình hệ thống, điều đó hỗ trợ để giảm thiểu gặp phải vấn đề của overfitting. Tuy nhiên, quá trình xử lý dữ liệu liên quan đến thời gian tiêu thụ và tài nguyên tính toán cho việc huấn luyện. Có nhiều cách để tăng thêm dữ liệu bao gồm xoay, thay đổi điều kiện ánh sáng, hướng khác nhau, lật, làm mờ với các ảnh mẫu. Vì vậy đối với một hình ảnh có thể được tạo ra các mẫu phụ khác nhau.

3.3. Quá trình huấn luyện

Hệ thống nhận dạng tiền giấy hiển thị đầy đủ ở sơ đồ trong hình 2.10. Hệ thống này được áp dụng theo một cách tiếp cận truyền thống trong DNN. Tuy nhiên, nó gặp vấn đề quá khớp với dữ liệu (overfitting) đáng kể so với một dữ liệu huấn luyện như được minh họa trong Hình 3.2.



(a)



(b)

Hình 3.2. Quá trình huấn luyện trên tập dữ liệu tiền trong mạng DNN hội tụ với các giá trị dự kiến sau một vài epochs: (a) độ chính xác; (b) giá trị lỗi.

Kết quả huấn luyện cho thấy hệ thống đạt được độ chính xác gần 99,99% và tiêu chí đánh giá chỉ ra một tỷ lệ học hơi quá nhỏ dựa trên tốc độ phân rã của nó sau một số epochs.

Sự kết hợp dựa trên mạng DNN để tách đặc trưng và SVM để nhận dạng tiền giấy. Chúng biến đổi hình ảnh đầu vào thành một mạng đồ thị theo chu kỳ (DAG) để học sâu để trích xuất mô tả đặc trưng thay vì phân loại. Trong phương pháp này, kích hoạt đầu ra ở lớp kết nối đầy đủ để trích xuất đặc trưng. Dữ liệu đầu vào là tập hợp cường độ điểm ảnh của hình ảnh được đưa vào mạng neural sâu. Các dữ liệu huấn luyện và thử nghiệm là đồng nhất với hình ảnh khung hình 200×450 pixel. Dữ liệu đầu vào bao gồm hình ảnh gốc $200 \times 450 \times 3$ và kích hoạt lớp FC đầu tiên dẫn đến 1.440.000 vector đặc trưng. Trong giai đoạn huấn luyện, các vector đặc trưng, được trích xuất từ tập dữ liệu huấn luyện, được cấp cho SVM để học. Trong giai đoạn đánh giá, mô hình DNN được áp dụng cho tập dữ liệu thử nghiệm để trích xuất các vector đặc trưng, được nhập vào mô hình SVM để nhận dạng tiền giấy.

Trong mô hình này, các bộ lọc tích chập ở lớp đầu vào đầu tiên là 3 kênh tương ứng với 3 màu của ảnh R-G-B. Bộ lọc độc lập kết nối với nhau, liên quan đến 3 kênh của hình ảnh đầu vào. Quá trình lớp cuối cùng trên vector đặc trưng được trích xuất sẽ được kích hoạt bằng kết nối đầy đủ, trao đổi với SVM để xử lý. Các chi tiết của kiến trúc nhận dạng tiền giấy được thể hiện trong Hình 2.9.

3.4. Kết quả thực nghiệm

Để đánh giá phương pháp được đề xuất, các tập dữ liệu huấn luyện và đánh giá bao gồm ba kịch bản. Kịch bản đầu tiên là tập hợp các hình ảnh thu được từ tiền giấy thật với nhiều mệnh giá như 10k, 20k, 50k, 100k, 200k, 500k VND. Thứ hai là, tập hợp các hình ảnh

thu được từ tiền giấy giả đã sử dụng chất liệu khác so với tiền giấy thật của cùng một tờ giấy bạc cùng mệnh giá. Thứ ba là, các hình ảnh không phải là tờ tiền. Tổng cộng, có 9.739 mẫu tiền thật, 1.083 mẫu tiền giấy giả và 1.548 mẫu ảnh không phải là tờ tiền. Tất cả các mẫu được chuẩn hóa với cùng độ phân giải 200x450 pixel. Tập dữ liệu được tăng cường dữ liệu 2 lần, trong đó dùng 50% tập dữ liệu cho huấn luyện, 50% tập dữ liệu cho đánh giá.

Kết quả kiểm tra chéo cho thấy độ chính xác của hệ thống khi đánh giá qua hai phương pháp CNN và CNN + SVM đều đạt được tỷ lệ chính xác trung bình đến 97,85%. Kết quả cho thấy rằng cách tiếp cận dựa trên mạng CNN và sự kết hợp của mạng CNN và máy phân loại SVM là tương đương nhau. Hai cách tiếp cận này thích hợp cho ứng dụng trong điều kiện chung của tiền giấy trong thực tế. Kết quả cuối cùng được minh họa như sau:

- Kết quả kiểm tra của phương pháp chỉ dùng CNN:

Bảng 3.5. Ma trận kiểm tra kết quả dự đoán trên tập dữ liệu đánh giá của phương pháp CNN (b) Đánh giá theo tỷ lệ %

Mẫu dữ liệu	10k	20k	50k	100k	200k	500k	Tiền giả	Không phải tiền
10k	99.72%	0.00%	0.10%	0.00%	2.05%	0.09%	0.09%	3.88%
20k	0.05%	97.01%	0.00%	0.24%	0.00%	0.00%	0.37%	1.10%
50k	0.00%	0.00%	98.76%	0.00%	0.00%	0.00%	1.39%	0.78%
100k	0.05%	0.25%	0.00%	99.37%	0.00%	0.35%	0.00%	0.00%
200k	0.05%	0.00%	0.00%	0.00%	97.86%	0.00%	0.00%	0.00%
500k	0.05%	2.09%	0.00%	0.39%	0.00%	99.48%	0.00%	0.00%
Tiền giả	0.05%	0.65%	1.14%	0.00%	0.00%	0.00%	97.60%	1.42%
Không phải tiền	0.05%	0.00%	0.00%	0.00%	0.09%	0.09%	0.55%	92.83%

- Kết quả kiểm tra của phương pháp dùng CNN + SVM:

Bảng 3.6. Ma trận kiểm tra kết quả dự đoán trên tập dữ liệu đánh giá của phương pháp CNN + SVM (b) Đánh giá theo tỷ lệ %

Mẫu dữ liệu	10k	20k	50k	100k	200k	500k	Tiền giả	Không phải tiền
10k	99.82%	0.25%	0.15%	0.00%	2.05%	0.00%	0.28%	3.75%
20k	0.00%	95.62%	0.00%	0.08%	0.00%	0.00%	0.37%	1.03%
50k	0.00%	0.00%	98.61%	0.00%	0.00%	0.00%	0.55%	0.26%
100k	0.05%	0.60%	0.00%	99.53%	0.00%	0.26%	0.00%	0.00%
200k	0.05%	0.00%	0.00%	0.00%	97.95%	0.00%	0.00%	0.00%
500k	0.05%	2.84%	0.00%	0.39%	0.00%	99.65%	0.00%	0.52%
Tiền giả	0.05%	0.70%	1.24%	0.00%	0.00%	0.09%	98.80%	0.90%
Không phải tiền	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	93.54%

3.5. Chương trình minh họa kết quả kiểm tra

Chương trình minh họa được xây dựng trên phần mềm Matlab, được chúng tôi cho nhận dạng các mệnh giá tiền với dữ liệu tờ tiền đưa vào theo các hướng khác nhau (mặt trước, mặt sau, ngược và xuôi tờ tiền) và kết quả được thể hiện trong các Hình 3.6., Hình 3.7



Nhận dạng mệnh giá tiền 100.000 VND



Nhận dạng mệnh giá tiền 200.000 VND



Nhận dạng mệnh giá tiền 500.000 VND

Hình 3.6. Kết quả nhận dạng mệnh giá 100.000, 200.000, 500.000 VND



Chương trình nhận dạng với các mẫu ảnh bất kỳ



Hình 3.7. Kết quả nhận dạng khả nghi tiền giả và không phải mẫu tiền

KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

1. Kết luận

Dựa vào những nghiên cứu các vấn đề về bài toán nhận dạng sử dụng kỹ thuật xử lý ảnh, trí tuệ nhân tạo và đặc biệt là mạng tích chập học sâu CNN. Luận văn này trình bày hai giải pháp cho việc nhận dạng tiền giấy và xác thực tiền thật dựa trên mạng DNN và sự kết hợp giữa mạng DNN và SVM truyền thống.

Luận văn đã nghiên cứu kỹ thuật học sâu, kỹ thuật phân loại SVM để đề xuất và xây dựng kiến trúc mạng neural sâu, tổ chức huấn luyện mô hình mạng và sử dụng trong nhận dạng tiền giấy thật. Cách tiếp cận với phương pháp mạng DNN và kết hợp giữa mạng DNN và máy phân loại SVM với dữ liệu sẽ tạo ra kết quả tốt trong cả ba tình huống nhận dạng mệnh giá tiền thật, tiền giả và không phải tờ tiền từ các loại dữ liệu đầu vào khác nhau.

Trong hệ thống kiểm tra này có sự khác biệt đối với phương pháp truyền thống, khi chúng tôi sử dụng máy ảnh để quét tiền giấy và sử dụng mạng neural tích chập sâu để huấn luyện trên hình ảnh màu. Mô hình CNN được sử dụng như trích xuất đặc trưng đối với phương pháp CNN+SVM (hoặc nhận dạng phát hiện tiền thật và tiền giả với phương pháp CNN). Mẫu ảnh màu này được phân thành tiền thật, tiền giả hay không phải tiền với các loại tiền tệ khác nhau bằng cách sử dụng kỹ thuật SVM. Cách tiếp cận này được đề xuất được thực hiện trên các tập dữ liệu thực (với 9.736 mẫu ảnh tiền thật với 6 mệnh giá, 1.083 mẫu ảnh tiền giả và 1.548 mẫu ảnh không phải là tờ tiền)

Về chương trình minh họa, mặc dù đã thực hiện tương đối tốt nhưng độ chính xác nhận dạng phân biệt tiền thật và tiền giả vẫn còn là thách thức vì số mẫu huấn luyện còn ít so với số mẫu dùng trong

ứng dụng thực tế, chất lượng dữ liệu hình ảnh mẫu chưa được tốt, số lượng dữ liệu mẫu cho các loại ngoại tệ chưa có nên để đánh giá mô hình hệ thống nhận dạng tiền để ứng dụng trong các hệ thống giao dịch tự động chưa được toàn diện.

2. Hướng phát triển

Để đánh giá và cải thiện hệ thống nhận dạng tiền tệ trong tương lai, chúng tôi đang tập trung vào thu thập thêm dữ liệu của nhiều loại tiền khác nhau, trang bị thiết bị scan chất lượng cao để nâng cao chất lượng dữ liệu mẫu. Đồng thời việc giảm chi phí tính toán và cải thiện độ chính xác cho nhiều loại tiền tệ khác nhau và nhận dạng mệnh giá tiền thật, tiền giả để áp dụng cho hệ thống kiểm tra tiền giấy trong thời gian thực.