

Nghiên Cứu Ứng Dụng Mô Hình Convolutional Neural Network Và Support Vector Machine Trong Phân Loại Rác Thải

Application of Convolutional Neural Network and Support Vector Machine Models in Waste Classification

Trần Quý Bằng, Đỗ Công Thịnh

Khoa Công nghệ Thông tin, trường Đại học Nông Lâm TP.HCM

Tóm tắt. Trong bối cảnh quản lý rác thải và bảo vệ môi trường, việc phân loại rác thải một cách chính xác và hiệu quả là cực kỳ quan trọng. Trong nghiên cứu này chúng tôi sử dụng hai mô hình máy học là Convolutional Neural Network (*CNN*) và Support Vector Machine (*SVM*), để phân loại rác thải dựa trên hình ảnh. Bộ dữ liệu được sử dụng là bộ dữ liệu phân loại rác thải từ Kaggle, bao gồm các hình ảnh của rác thải được phân loại thành hai nhóm là rác thải hữu cơ (*Organic*) và rác thải tái chế (*Recycle*). Kết quả thực nghiệm cho thấy mô hình *CNN* đạt được độ chính xác cao hơn so với *SVM*. Tuy nhiên, *SVM* cũng cho thấy tiềm năng khi kết hợp với Principal Component Analysis (*PCA*) và *GridSearchCV*. Qua đó, nghiên cứu này cung cấp cái nhìn tổng quan về hiệu quả của hai phương pháp phân loại khác nhau, đồng thời gợi ý các hướng phát triển tiềm năng cho việc cải thiện hệ thống phân loại rác thải tự động

Abstract. In the context of waste management and environmental protection, accurate and efficient waste classification is crucial. This study employs two machine learning models, Convolutional Neural Network (*CNN*) and Support Vector Machine (*SVM*), to classify waste based on images. The dataset used is the waste classification dataset from Kaggle, which includes images of waste categorized into two groups: organic waste (*Organic*) and recyclable waste (*Recycle*). Experimental results indicate that the *CNN* model achieves higher accuracy compared to *SVM*. However, *SVM* also demonstrates potential when combined with Principal Component Analysis (*PCA*) and *GridSearchCV*. This research provides an overview of the effectiveness of two different classification methods and suggests potential directions for improving automated waste classification systems.

Keywords: Waste Classification, Convolutional Neural Network, Support Vector Machine, Principal Component Analysis, *GridSearchCV*

1. Giới thiệu

Trên toàn cầu, việc quản lý và xử lý rác thải đang trở thành một thách thức ngày càng nghiêm trọng, đặc biệt là trong bối cảnh gia tăng về ô nhiễm môi trường và sự cần thiết phải tối ưu hóa tài nguyên. Phân loại rác thải một cách hiệu quả không chỉ giúp giảm thiểu tác động tiêu cực đến môi trường mà còn tạo ra cơ hội cho việc tái chế và sử dụng lại tài nguyên (Mihai et al. 2019).

Dự án sử dụng hai phương pháp học máy chính là Convolutional Neural Network (CNN) và Support Vector Machine (SVM). CNN được lựa chọn vì khả năng học và trích xuất đặc trưng từ hình ảnh một cách tự động và hiệu quả. Mô hình CNN học từ dữ liệu hình ảnh rác thải và xây dựng các bộ lọc để nhận diện các đặc trưng quan trọng của từng loại rác thải. SVM được áp dụng để xây dựng một bộ phân lớp dựa trên các đặc trưng đã được trích xuất từ CNN. Để tối ưu hóa mô hình, chúng tôi sử dụng kỹ thuật PCA (Principal Component Analysis) để giảm chiều dữ liệu và GridSearchCV để tìm ra các tham số tối ưu cho SVM. Phương pháp này không chỉ cải thiện độ chính xác mà còn tăng tính tổng quát hóa của mô hình trong các điều kiện thực tế đa dạng và phức tạp của rác thải.

Dự án này hy vọng mang lại những đóng góp to lớn vào lĩnh vực quản lý rác thải và bảo vệ môi trường, đồng thời khẳng định vai trò quan trọng của công nghệ học máy trong giải quyết các thách thức môi trường hiện đại. Kết quả từ dự án có thể được áp dụng rộng rãi trong các ứng dụng thực tế, góp phần cải thiện chất lượng môi trường sống và bảo vệ sức khỏe con người.

2. Các công trình liên quan

Các phương pháp học máy, đặc biệt là Convolutional Neural Networks (CNN) và Support Vector Machines (SVM), đã mang lại nhiều tiến bộ đáng kể trong lĩnh vực phân loại rác thải. CNN nổi bật với khả năng học và nhận diện các đặc trưng từ hình ảnh rác thải. Một nghiên cứu của (Yalcin et al. 2018) đã chứng minh tính hiệu quả của CNN trong nhận diện và phân loại rác thải, cải thiện quy trình quản lý và xử lý rác thải.

Ngoài ra, SVM kết hợp với Principal Component Analysis (PCA) và GridSearchCV cũng được áp dụng để tối ưu hóa bộ phân lớp, nâng cao khả năng tổng quát hóa và độ chính xác của hệ thống phân loại. Nghiên cứu của (El-Hefnawy et al. 2019) đã cho thấy phương pháp này không chỉ cải thiện độ chính xác mà còn tăng tính tổng quát hóa của mô hình trong các điều kiện thực tế đa dạng và phức tạp của rác thải.

Bên cạnh đó, các nghiên cứu khác đã mở rộng việc sử dụng các mô hình tiên tiến trong phân loại rác thải. Ví dụ, nghiên cứu của (Bai et al. 2019) đã sử dụng mạng CNN để phân loại các loại rác thải khác nhau từ hình ảnh và đạt được độ chính xác cao. (Chen et al. 2020) đã áp dụng các thuật toán học máy khác nhau, bao gồm SVM

và CNN, để phân loại rác thải và nhận thấy rằng CNN có độ chính xác cao hơn nhưng SVM vẫn cung cấp kết quả đáng tin cậy khi kết hợp với PCA .

3. Phát biểu bài toán

3.1. Bài toán

Bài toán phân loại rác thải tự động là một vấn đề quan trọng trong bối cảnh hiện nay, khi lượng rác thải sinh hoạt ngày càng gia tăng và đòi hỏi các phương pháp xử lý hiệu quả hơn. Phân loại rác thải không chính xác có thể dẫn đến việc xử lý sai, gây lãng phí tài nguyên và ô nhiễm môi trường (C. Capel et al. 2008). Do đó, việc áp dụng các công nghệ hiện đại như máy học để tự động phân loại rác thải là một hướng đi tiềm năng và cần thiết.

Input là các hình ảnh của rác thải và output là nhãn phân loại tương ứng (hữu cơ hoặc tái chế).

3.2. Thuật toán

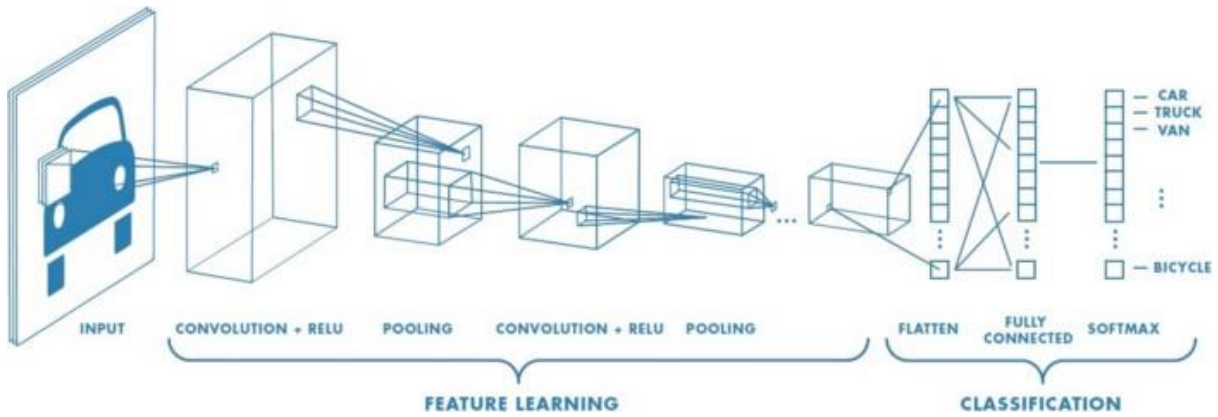
3.2.1. Convolutional Neural Network

CNN là một loại mạng nơ-ron nhân tạo đặc biệt mạnh mẽ trong việc xử lý và phân loại hình ảnh (P. Y. Simard et al. 2003) . CNN có khả năng tự động học các đặc điểm (features) quan trọng từ hình ảnh thông qua các lớp tích chập (convolutional layers) và lớp gộp (pooling layers).

CNN bao gồm nhiều lớp khác nhau, mỗi lớp có vai trò và chức năng riêng biệt trong việc xử lý và trích xuất đặc trưng từ dữ liệu đầu vào. Các lớp chính bao gồm:

- **Lớp tích chập (Convolutional Layer)** là lớp cốt lõi của CNN, nơi mà phép tích chập được thực hiện. Sử dụng các bộ lọc (filters), lớp này trích xuất các đặc trưng từ đầu vào. Mỗi bộ lọc di chuyển qua toàn bộ đầu vào (ví dụ: hình ảnh) và tính toán tích chập để tạo ra một bản đồ đặc trưng (feature map). Các đặc trưng như cạnh, góc và các dạng hình học cơ bản được nhận diện ở các lớp tích chập ban đầu.
- Sau khi tính toán tích chập, kết quả sẽ được đưa qua **hàm kích hoạt (Activation Layer)**, thường là hàm ReLU (Rectified Linear Unit), được định nghĩa là $f(x) = \max(0, x)$. Hàm kích hoạt giúp mô hình học các đặc trưng phi tuyến tính, tăng khả năng biểu diễn của mạng.
- **Lớp gộp (Pooling Layer)** thực hiện phép gộp (thường là gộp tối đa - max pooling) để giảm kích thước của bản đồ đặc trưng, giữ lại các đặc trưng quan trọng và giảm thiểu số lượng tham số. Gộp tối đa lấy giá trị lớn nhất trong mỗi cửa sổ con của bản đồ đặc trưng, giúp giảm độ phân giải nhưng vẫn giữ lại các thông tin quan trọng.
- **Lớp kết nối đầy đủ (Fully Connected Layer)** hoạt động tương tự như trong mạng nơ-ron truyền thống, kết nối tất cả các đầu vào với các nút đầu ra.
- **Lớp đầu ra (Output layer)** dùng để chuyển đổi đầu ra của lớp.

Hình 1. Sơ đồ hoạt động của CNN



3.2.2. Support Vector Machine

Support Vector Machine (SVM) là một mô hình phân loại sử dụng siêu phẳng (hyperplane) để phân biệt các nhóm khác nhau (Abe et al. 2004). Thông thường, SVM được sử dụng cho các bài toán phân loại hai nhóm, nhưng cũng có thể được mở rộng để xử lý nhiều hơn hai nhóm. Siêu phẳng trong SVM được xác định dựa trên khoảng cách lớn nhất giữa các điểm dữ liệu gần nhất của các nhóm khác nhau. Các điểm gần nhất này được gọi là các vector hỗ trợ (support vectors) và khoảng cách lớn nhất giữa chúng được gọi là khoảng cách biên (margin). Dưới đây là công thức xác định siêu phẳng:

$$f(x) = w \cdot x + b$$

hoặc

$$f(x) = \sum_{i=1}^m a_i y_i K(x, x_i) + b$$

Trong đó, $f(x)$ là hàm dùng để tìm giá trị của siêu phẳng. Các biến ảnh hưởng đến sự tồn tại của một điểm dữ liệu huấn luyện trong siêu phẳng (x) được nhân với trọng số của mỗi điểm (w) và được cộng thêm giá trị bias (b). Để siêu phẳng gốc được xác định, hàm $f(x)$ phải bằng không.

3.2.3. Đánh giá mô hình phân lớp

Trong bài báo này, chúng tôi tiến hành đánh giá hiệu suất của mô hình phân lớp sử dụng confusion matrix. Confusion matrix cung cấp một cái nhìn tổng quan về các kết quả phân loại, giúp xác định các chỉ số quan trọng như Precision, Recall và F1-score.

Bảng 1. Confusion matrix

Actual class\Predicted class	C1	$\neg C1$
C1	True Positives (TP)	False Negatives (FN)
$\neg C1$	False Positives (FP)	True Negatives (TN)

Confusion matrix là một bảng gồm bốn phần tử chính:

- **True Positive (TP)**: Số lượng mẫu thực sự thuộc lớp dương tính và được mô hình dự đoán đúng.
- **True Negative (TN)**: Số lượng mẫu thực sự thuộc lớp âm tính và được mô hình dự đoán đúng.
- **False Positive (FP)**: Số lượng mẫu thực sự thuộc lớp âm tính nhưng bị mô hình dự đoán sai là dương tính.
- **False Negative (FN)**: Số lượng mẫu thực sự thuộc lớp dương tính nhưng bị mô hình dự đoán sai là âm tính.

Từ confusion matrix, chúng tôi tính toán các chỉ số hiệu suất sau:

- **Accuracy**: Được dùng để đo độ chính xác của mô hình phân lớp.

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN}$$

- **Precision**: Precision được định nghĩa là tỉ lệ số điểm Positive mô hình dự đoán đúng trên tổng số điểm mô hình dự đoán là Positive

$$Precision = \frac{TP}{TP + FP}$$

- **Recall**: Được tính bằng tỷ lệ giữa số dự đoán đúng dương tính trên tổng số mẫu dương tính thực tế. Chỉ số này cho biết mô hình có bao nhiêu mẫu dương tính thực tế được nhận diện đúng

$$Recall = \frac{TP}{TP + FN}$$

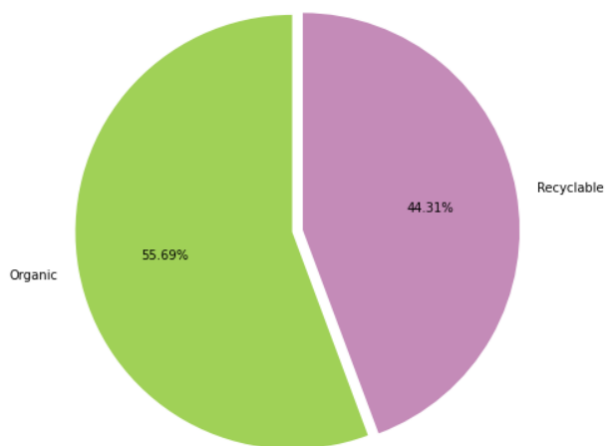
- **F1-score**: Là trung bình điều hòa của Precision và Recall, cung cấp một thước đo cân bằng giữa hai chỉ số này. F1-score hữu ích khi cần xem xét cả độ chính xác và độ nhạy của mô hình, đặc biệt trong trường hợp dữ liệu không cân bằng

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

4. Thực nghiệm

4.1. Dữ liệu

Bộ dữ liệu hình ảnh của rác thải từ Kaggle. Bộ dữ liệu này bao gồm 25077 hình ảnh được phân loại sẵn thành 2 nhóm chính: Rác thải hữu cơ (Organic) và rác thải tái chế (Recycle).



Hình 2. Biểu đồ phân bố dữ liệu

Bộ dữ liệu chia 85% cho tập huấn luyện (train) và 15% cho tập kiểm tra (test).

4.2. Phương pháp SVM

4.2.1. Tiền xử lý dữ liệu

Các hình ảnh được chuyển đổi sang dạng xám và kích thước được chuẩn hóa về 64x64 pixel. Các ảnh này sau đó được làm phẳng thành vector 1 chiều và chuẩn hóa bằng StandardScaler. Cuối cùng, chúng tôi áp dụng Phân tích thành phần chính (PCA) để giảm chiều dữ liệu xuống 100 thành phần chính.

4.2.2. Huấn luyện và đánh giá

Chúng tôi sử dụng GridSearchCV để tìm kiếm các tham số tối ưu cho mô hình SVM với các kernel khác nhau (linear, rbf, poly, sigmoid). Sau khi huấn luyện, mô hình tối ưu được chọn để đánh giá trên tập kiểm tra.

4.3. Phương pháp CNN

4.3.1. Tiền xử lý dữ liệu

Chúng tôi sử dụng ImageDataGenerator để chuẩn hóa và tăng cường dữ liệu hình ảnh. Các hình ảnh được chuẩn hóa về kích thước 224x224 pixel và các giá trị pixel được chia cho 255 để đưa về phạm vi [0, 1].

4.3.2. Kiến trúc mô hình

Mô hình CNN bao gồm ba lớp tích chập (Conv2D) xen kẽ với các lớp gộp (MaxPooling2D), tiếp theo là các lớp Dense để thực hiện phân loại. Đầu ra của mô hình là một lớp sigmoid với hai neuron đại diện cho hai loại rác thải.

Bảng 2. Kiến trúc mô hình CNN trong bài toán

Tham số	Giá trị
#Lớp thứ nhất	
Conv2D filter	32
Conv2D kernel	3 x 3
Input shape	(224, 224, 3)
Activation	ReLU
MaxPooling	MaxPooling2D
#Lớp thứ hai	
Conv2D filter	64
Conv2D kernel	3 x 3
Activation	ReLU
MaxPooling	MaxPooling2D
#Lớp thứ ba	
Conv2D filter	128
Conv2D kernel	3 x 3
Activation	ReLU
MaxPooling	MaxPooling2D
#Chuẩn hóa dữ liệu	
Flatten	Flatten
#Lớp FC thứ nhất	
Dense	256 units

Activation	ReLU
Dropout	0.5 dropout rate
#Lớp FC thứ hai	
Activation	ReLU
Dropout	0.5 dropout rate
#Lớp đầu ra	
Dense	2 units
Activation	Sigmoid
Loss function	Binary Crossentropy
Optimizer	Adam
Metrics	Accuracy

4.3.3. Huấn luyện và đánh giá

Mô hình được huấn luyện trên tập huấn luyện trong 10 epochs và được đánh giá trên tập kiểm tra. Kết quả huấn luyện bao gồm độ chính xác, bảng báo cáo phân loại, và ma trận nhầm lẫn.

5. Kết quả

Kết quả so sánh với tổng dữ liệu 25077 hình ảnh về rác thải. Trong đó tập Train có 22564 hình và tập Test có 2513 hình.

Bảng 3. Kết quả thực nghiệm

Thuật toán	Số hình ảnh dự đoán đúng	Tỉ lệ dự đoán đúng	Precision	Recall	F1-score
CNN	2198/2513	87%	0.88	0.87	0.87
SVM	1868/2513	74%	0.75	0.74	0.74

Mô hình SVM với kernel rbf cho kết quả tốt nhất với độ chính xác 74.33%. Tuy nhiên, mô hình này có thời gian huấn luyện khá dài, đặc biệt là với kernel linear.

Mô hình CNN cho thấy hiệu suất tốt hơn đáng kể với độ chính xác 87%. Điều này có thể được giải thích bởi khả năng của CNN trong việc tự động học các đặc trưng phức tạp từ dữ liệu hình ảnh.

So sánh hai phương pháp cho thấy CNN có ưu thế rõ rệt trong việc phân loại rác thải từ hình ảnh. Mặc dù SVM cũng có thể đạt kết quả tốt, CNN cho thấy khả năng học tập và tổng quát hóa tốt hơn trên dữ liệu hình ảnh.

6. Kết luận

Trong bài báo này, chúng tôi đã trình bày việc ứng dụng hai mô hình học sâu và học máy là Mạng Nơ-ron Tích chập (CNN) và Máy Vector Hỗ trợ (SVM) trong việc phân loại rác thải. Chúng tôi đã tiến hành huấn luyện và đánh giá các mô hình này trên bộ dữ liệu lớn về hình ảnh rác thải, bao gồm hai loại rác: tái chế và hữu cơ. Kết quả thực nghiệm cho thấy rằng mô hình CNN có hiệu suất cao hơn trong việc phân loại rác thải, với độ chính xác đạt được là 87.47% trên tập kiểm tra. Mô hình SVM, mặc dù đạt được độ chính xác thấp hơn (74.33%), vẫn cho thấy khả năng phân loại tương đối tốt khi được kết hợp với các kỹ thuật giảm chiều dữ liệu như PCA. Phân tích chi tiết các kết quả cho thấy mô hình CNN có khả năng học các đặc trưng phức tạp của hình ảnh tốt hơn, trong khi SVM lại phù hợp hơn với các dữ liệu đã qua xử lý tiền xử lý kỹ lưỡng. Các ma trận nhầm lẫn và báo cáo phân loại chỉ ra rằng cả hai mô hình đều có khả năng phân loại đúng các hình ảnh rác thải, nhưng vẫn còn tồn tại một số trường hợp nhầm lẫn giữa hai loại rác.

Trong tương lai, để cải thiện độ chính xác và hiệu suất của các mô hình phân loại rác thải thì cần phải sử dụng các mô hình tiền huấn luyện như ResNet, VGG, hoặc EfficientNet để tận dụng các đặc trưng đã học từ các bộ dữ liệu lớn khác, giúp cải thiện khả năng phân loại của mô hình CNN. Bài toán cũng cần mở rộng thêm có thể phân loại nhiều loại rác thải hơn, không chỉ giới hạn ở rác tái chế và hữu cơ, nhằm đáp ứng nhu cầu thực tế phức tạp hơn.

Tài liệu tham khảo

Mihai, F.-C.; Gnoni, M.-G.; Meidiana, C.; Ezeah, C.; Elia, V. Chapter 1—Waste Electrical and Electronic Equipment (WEEE): Flows, Quantities, and Management—A Global Scenario. In *Electronic Waste Management and Treatment Technology*; Prasad, M.N.V., Vithanage, M., Eds.; Butterworth-Heinemann: Kidlington, UK, 2019; pp. 1–34.

C. Capel, "waste sorting - a look at the separation and sorting techniques in today's European market", Google Scholar, Waste Management World, 2008.

Yalcin, Y., & Kantarci, E. "A Comparative Study on Waste Classification Using Convolutional Neural Networks and Other Machine Learning Algorithms." Journal of Environmental Management, vol. 218, 2018, pp. 401-407.

El-Hefnawy, N., & Farag, A. "Improving SVM Performance for Waste Classification Using PCA and GridSearchCV." International Journal of Machine Learning and Computing, vol. 9, no. 3, 2019, pp. 256-262.

Bai, J., & Li, J. "Waste Classification Using Convolutional Neural Networks." Procedia Computer Science, vol. 159, 2019, pp. 1438-1445.

Chen, X., & Zhang, Y. "Comparative Study of Machine Learning Algorithms for Waste Classification." Journal of Environmental Management, vol. 250, 2020, 109469.

P. Y. Simard, D. Steinkraus and J. C. Platt, "Best practices for convolutional neural networks applied to visual document analysis", null. IEEE, pp. 958, 2003.

Abe, S., Support Vector Machines for Pattern Classification (in print), SpringerVerlag, London, 2004

Làm nhóm

Công việc	Người làm
Tìm hiểu tổng quan và lý do nghiên cứu	Trần Quý Bằng
Viết tóm tắt cho bài báo	Trần Quý Bằng
Viết phần giới thiệu và các công trình liên quan	Đỗ Công Thịnh, Trần Quý Bằng
Viết phần phát biểu cho bài toán	Đỗ Công Thịnh
Tìm hiểu và viết phần mô tả cho thuật toán CNN	Trần Quý Bằng
Tìm hiểu và viết phần mô tả cho thuật toán SVM	Đỗ Công Thịnh

Tìm hiểu và viết phần mô tả đánh giá mô hình phân loại	Trần Quý Bằng
Xây dựng mô hình SVM	Đỗ Công Thịnh
Xây dựng mô hình CNN	Trần Quý Bằng
Tổng hợp lại viết kết quả	Đỗ Công Thịnh
Viết phần kết luận và hướng phát triển	Trần Quý Bằng
Chỉnh sửa format lại work	Đỗ Công Thịnh

Link github: <https://github.com/bangg2309/Project-ML>