Tạp chí Khoa học Trường Đại học Cần Thơ
Phần A: Khoa học Tự nhiên, Công nghệ và Môi trường
website: ctujsvn.ctu.edu.vn

DOI:10.22144/ctu.jvn.2022.242

PHÂN LOẠI ẢNH DỰA VÀO HÀM MẬT ĐỘ XÁC SUẤT ĐƯỢC TRÍCH XUẤT

Võ Văn Tài*, Trần Nam Hưng và Huỳnh Văn Nhạng
Khoa Khoa học Tự nhiên, Trường Đại học Cần Thơ
**Người chịu trách nhiệm về bài viết: Võ Văn Tài (email: vvtai@ctu.edu.vn)*

Thông tin chung:
Ngày nhận bài: 15/08/2022
Ngày nhận bài sửa: 21/09/2022
Ngày duyệt đăng: 26/09/2022

Title:
Classifying for image based on the extracted probability density function

Từ khóa:
Hàm mật độ xác suất, khoảng cách, phân loại, trích xuất ảnh

Keywords:
Classification, distance, extracting image, probability density function

ABSTRACT

In this study, an algorithm was proposed to classify the images based on the extracted one-dimensional probability density function. First, based on the resolution of grey level, each image was extracted into a representative one-dimensional probability density function and truncated on [0;1]. Next, the prior probability of the classified image was determined based on the fuzzy clustering problem. Finally, a principle to classify images was proposed. The image was assigned to the known group if it has the maximum value of the prior probability and a similar level to that group. The proposed algorithm was illustrated step by step, and applied to specific image sets. The application obtained a good result because it had fewer errors than other methods. This study also shows potential in practical application for many fields related to image recognition.

TÓM TẮT

Nghiên cứu nhằm đề xuất một thuật toán phân loại ảnh dựa vào hàm mật độ xác suất một chiều được trích xuất. Đầu tiên, dựa vào độ phân giải mức xám, mỗi ảnh được trích xuất thành một hàm mật độ xác suất một chiều đại diện và được chặt cụt trên đoạn $[0;1]$. *Tiếp theo, xác suất tiên nghiệm của một ảnh cần phân loại được xác định dựa vào bài toán phân tích chùm mờ. Cuối cùng, một nguyên tắc phân loại cho ảnh được đề xuất. Ảnh được xếp vào một nhóm đã biết nếu nó có xác suất tiên nghiệm và sự tương tự đến nhóm đó lớn nhất. Thuật toán đề nghị được minh hoạ từng bước và được áp dụng trên những tập ảnh cụ thể. Áp dụng đã cho kết quả tốt bởi vì nó có sai số thấp hơn nhiều phương pháp khác. Nghiên cứu này cũng cho thấy tiềm năng trong áp dụng thực tế của nhiều lĩnh vực liên quan đến nhận dạng ảnh.*

1. GIỚI THIỆU

Phân loại là việc gán một phần tử vào một tổng thể thích hợp dựa trên các biến quan sát của nó. Đây là một hướng phát triển quan trọng của thống kê nhiều chiều và khoa học dữ liệu ngày nay. Bài toán phân loại có rất nhiều ứng dụng trong thực tế, vì vậy nó nhận được sự quan tâm của rất nhiều nhà khoa học. Tuy nhiên, cho đến nay nó vẫn là bài toán chưa có lời giải cuối cùng (Ngọc et al., 2021, 2022). Đối tượng phân loại có thể là các phần tử số hoặc các hình ảnh. Với sự phát triển mạnh của các thiết bị ghi hình và trí tuệ nhân tạo, phân loại cho đối tượng ảnh ngày càng đóng vai trò quan trọng trong thực tế. Đối với đối tượng ảnh, việc phân loại được thực hiện qua hai bước chính: trích xuất đặc trưng cho ảnh để tìm phần tử đại diện và xây dựng mô hình phân loại cụ thể từ dữ liệu được trích xuất.

Trích xuất ảnh là bước đầu tiên rất quan trọng để thực hiện bài toán phân loại. Việc trích xuất ảnh được xem là tối ưu nếu nó đặc trưng được cho ảnh và dễ nhận dạng, dễ phân biệt. Thông thường ảnh được trích xuất thông qua màu sắc, kết cấu hoặc hình dạng (Thao & Tai, 2017; Koklu, 2021). Mặc dù có nhiều thảo luận khác nhau về vấn đề này, nhưng cho đến nay chưa có một phương pháp nào được xem là tối ưu cho tất cả các trường hợp (Zhu et al., 2000; Zhang et al., 2018). Khi các đặc trưng được trích xuất, để áp dụng vào bài toán phân loại, chúng ta phải chọn lựa phần tử đại diện cho ảnh. Việc chọn lựa này rất quan trọng bởi vì nhận dạng ảnh chính là nhận dạng các phần tử đại diện này. Hiện nay có ba đối tượng chính được sử dụng để đại diện cho ảnh: các phần tử rời rạc, hàm mật độ xác suất (PDF) và khoảng dữ liệu (Ha et al., 2020; Dinh & Tai, 2021). Việc sử dụng các phần tử rời rạc dạng ma trận đại diện cho ảnh để phân loại được nhiều nhà nghiên cứu quan tâm và được áp dụng phổ biến trong thực tế (Tai et al., 2021). Sử dụng khoảng đại diện cho ảnh, xây dựng thuật toán phân loại đã được đề xuất gần đây (Ngoc et al., 2022). Tuy nhiên nghiên cứu phân loại loại ảnh dựa vào PDF được trích xuất chưa được tìm thấy.

Về vấn đề phân loại, có nhiều phương pháp đã và đang được ứng dụng. Dựa vào thống kê, các phương pháp như Fisher, hồi quy Logistic và Bayes thường được đề cập. Phương pháp của Fisher có thể phân loại cho hai hoặc nhiều hơn hai tổng thể, nhưng phải giả sử ma trận hiệp phương sai của các tổng thể là bằng nhau. Vì vậy, nó có nhiều hạn chế khi áp dụng cho dữ liệu thực (Fisher, 1938). Phương pháp hồi quy Logistic được sử dụng rất phổ biến, nhưng chỉ hiệu quả khi dữ liệu có khả năng tách nhóm tốt và biến phụ thuộc là nhị phân (Ha et al., 2020). Phương pháp Bayes không yêu cầu điều kiện dữ liệu, có thể phân loại cho nhiều tổng thể. Do đó, phương pháp Bayes được đánh giá có nhiều tiềm năng trong việc mở rộng và ứng dụng lý thuyết vào thực tế (Tai et al., 2021). Với kỹ thuật học máy và học sâu, chúng ta có nhiều phương pháp phân loại khác như kNN, Máy vector hỗ trợ (SVM), Trees, ENS, Neural Networks, Optimal Autofit, SqueezeNet, GoogleNet, MobileNetv2, VGG-19 và Inceptionv. Các phương pháp này tận dụng được lợi thế phát triển của công nghệ thông tin, cho kết quả phân loại thực tế khá tốt, tuy nhiên chúng đòi hỏi tập dữ liệu lớn, chi phí thời gian cho huấn luyện nhiều và thiếu ổn định (Cinar, 2022; Huang et al., 2018; Imadoust & Bolandraftar, 2013; Pham et al., 2016).

Trong phân loại bằng phương pháp Bayes, hai vấn đề quyết định đến kết quả là tìm xác suất tiên nghiệm và ước lượng PDF. Xác suất tiên nghiệm thường được xác định dựa vào kinh nghiệm, sự hiểu biết của người thực hiện, hoặc cập nhật một kết luận thống kê trước đó về đối tượng được quan sát. Nếu chưa có bất cứ thông tin nào về phần tử phân loại, xác suất tiên nghiệm thường được chọn theo phân phối đều. Nếu dựa vào tập huấn luyện, phương pháp tiên nghiệm theo tỷ lệ đóng góp của các nhóm trong mẫu hoặc phương pháp Laplace thường được đề xuất. Tuy nhiên, ta có thể khẳng định rằng chưa có phương pháp nào được xem là tối ưu cho tất cả các trường hợp (Ngoc et al., 2022). Sự phù hợp chỉ được xem xét cho từng bộ dữ liệu cụ thể. Bên cạnh xác suất tiên nghiệm, khi thực hiện bài toán phân loại bằng phương pháp Bayes, chúng ta phải ước lượng PDF đại diện cho mỗi ảnh và PDF đại diện cho các nhóm. Mặc dù có nhiều cải tiến trong những năm gần đây về vấn đề này, nhưng cho đến nay nó vẫn còn nhiều hạn chế. Ngoài hai vấn đề trên, những phức tạp trong tính toán như tìm hàm cực đại, tính tích phân trong không gian nhiều chiều cũng là cản trở trong áp dụng thực tế của phương pháp này.

Tất cả các phương pháp được đề xuất ở trên đều thực hiện cho dữ liệu rời rạc. Do đó khi áp dụng phân loại cho đối tượng hình ảnh nó phải được trích xuất thành các phần tử rời rạc. Thực tế áp dụng cho thấy cách làm này có những hạn chế trong nhiều trường hợp. Mô hình phân loại ảnh dựa trên dữ liệu đầu vào liên tục là các hàm mật độ xác suất được trích xuất chưa được quan tâm và đề xuất từ các nhà nghiên cứu trong lĩnh vực phân loại hình ảnh.

2. CÁC VẤN ĐỀ LIÊN QUAN ĐẾN THUẬT TOÁN ĐỀ NGHỊ

2.1. Ước lượng hàm mật độ xác suất

Hiện nay có rất nhiều phương pháp tham số và phi tham số để ước lượng các PDF, tuy nhiên chưa có phương pháp nào được xem là tối ưu cho tất cả các trường hợp. Trong nghiên cứu này, phương pháp hàm hạt nhân, một phương pháp phổ biến nhất hiện nay được sử dụng. Hàm mật độ xác suất d-chiều được ước lượng bởi phương pháp hàm hạt nhân có dạng như sau:

$$\hat{p}(x) = \frac{1}{N}\frac{1}{h_1 \times \cdots \times h_d}\sum_{i=1}^{N}\prod_{j=1}^{d} K_j\left(\frac{x_j - x_{ij}}{h_j}\right), \quad (1)$$

trong đó

$x_j, j = 1,2, \dots, d$ là các biến,

x_{ij} là dữ liệu thứ i của biến thứ $j, i = 1, 2, \dots, N$,

h_j là tham số trơn của biến thứ j,

$K_j(\cdot)$ là hàm hạt nhân của biến thứ *j*. Hàm hạt nhân phải thoả mãn hai điều kiện $K(\cdot) \geq 0$ và $\int K(x)\mathrm{d}x = 1$.

Có nhiều hàm hạt nhân khác nhau như hàm tam giác, chữ nhật, song lượng và chuẩn tắc. Mặc dù có nhiều nghiên cứu về vấn đề này, nhưng chưa có hàm hạt nhân nào được xem là tốt nhất (Ha et al., 2020). Hầu hết các áp dụng hiện tại sử dụng hàm hạt nhân dạng chuẩn:

$$K(x) = \frac{1}{\sqrt{2\pi}}\exp(-x^2/2).$$

Một vấn đề khác trong ước lượng PDF theo (1) là việc chọn tham số trơn. Khi tham số trơn càng nhỏ, đồ thị của PDF ước lượng càng trơn và ngược lại. Việc chọn tham số trơn đã được nhiều nhà thống kê quan tâm, tuy nhiên nó vẫn là bài toán chưa có lời giải cuối cùng. Nghiên cứu này chọn tham số trơn theo Terrell (1992), một phương pháp đã được đề xuất từ rất sớm nhưng được sử dụng phổ biến hiện nay. Theo phương pháp này, tham số trơn của biến thứ *j* được tính như sau:

$$h_j = \left(4/\left(N(d+2)\right)\right)^{1/(d+4)} \times \sigma_j,$$

trong đó N là số phần tử trong dữ liệu, d là số chiều và $\sigma_j^2 = \frac{1}{d}\sum_{i=1}^{d} s_{ii}$, trong đó s_{ii} là phần tử đường chéo trong ma trận hiệp phương sai mẫu.

2.2. Hàm mật độ xác suất đại diện chùm

Cho tập $\mathcal{F}$ gồm các PDF với

$$\mathcal{F} = \{f_1(x), f_2(x), \dots, f_n(x)\}, n \geq 2$$

cần được xếp vào k nhóm:

$$C = \{C_1, C_2, \dots, C_k\}, (k \geq 2).$$

Hàm số đại diện của mỗi nhóm được định nghĩa bởi (2):

$$f_{v_i} = \frac{1}{\sum_{i=1}^{k}\left(\mu_{f_{v_i}}\right)^2} \times \sum_{i=1}^{k}\left(\mu_{f_{v_i}}\right)^2 f_i, \quad (2)$$

trong đó $\mu_{f_{v_i}}$ là xác suất xếp hàm mật độ f_i vào chùm C_i. Dựa vào Công thức (2), ta dễ dàng nhận được hàm đại diện nhóm f_{v_i} không âm và $\int_{\mathbb{R}^d} f_{v_i}(x)\mathrm{d}x = 1$. Như vậy hàm đại diện của một nhóm cũng là một PDF.

2.3. Hàm mật độ xác suất chặt cụt trên một đoạn

Cho hàm mật độ xác suất *f*(*x*) trong không gian $\mathbb{R}$, đặt $m = \int_a^b f(x)dx$, trong đó $a, b \in \mathbb{R}$. Từ *f*(*x*), ta thiết lập hàm *g*(*x*) như sau:

$$g(x) = \begin{cases} \dfrac{f(x)}{m} & \text{nếu } x \in [a;b], \\ 0 & \text{nếu } x \notin [a;b]. \end{cases}$$

Vì *f*(*x*) là một PDF, nên *f*(*x*) không âm và $\int_{-\infty}^{+\infty} \boldsymbol{f(x)dx = 1}$ nên ta có thể kiểm tra dễ dàng rằng *g*(*x*) không âm và

$$\int_{-\infty}^{+\infty} g(x)dx = 1.$$

Như vậy *g*(*x*) cũng là một PDF. Lúc này ta gọi *g*(*x*) là PDF chặt cụt của *f*(*x*) trên đoạn [a;b].

2.4. Khoảng cách của hai hàm mật độ xác suất

Trong không gian xác suất $(\Omega, \mathcal{F}, \mathbb{P})$, với Ω là không gian mẫu, họ $\mathcal{F}$ các tập con đo được của σ −đại số trong Ω, và hàm xác suất $\mathbb{P} : \mathcal{F} \to [0,1]$ gồm hai hàm mật độ xác suất f và g. Ta có các khoảng cách sau:

- *Khoảng cách Housdorff:*

$$\mathcal{D}_H(f \parallel g) = \max\left(h(f,g), h(g,f)\right),$$

trong đó $h(f,g) = \max_{p\in f}\left(\min_{q\in g}\left(d(p,q)\right)\right)$ và $h(g,f) = \max_{q\in g}\left(\min_{p\in f}\left(d(q,p)\right)\right)$ và $d(p,q)$ là khoảng cách $\mathcal{L}^2$ giữa hai vector.

Khoảng cách Kulback-Leirber:

$$\mathcal{D}_{KL}(f \parallel g) = \int_{-\infty}^{+\infty} \sqrt{f(x)\cdot g(x)}.$$

Khoảng cách Jesen-Shannon:

$$JSD(f \parallel g) = \frac{1}{2}\mathcal{D}_{KL}(f \parallel h) + \frac{1}{2}\mathcal{D}_{KL}(g \parallel h),$$

trong đó $h = \frac{1}{2}(f+g)$.

Khoảng cách Bhattacharyya**:**

$$BC(f,g) = \int_{-\infty}^{+\infty} \sqrt{f(x)\cdot g(x)}\, dx.$$

Các khoảng cách trên được sử dụng để đánh giá sự tương tự của hai PDF. Khi khoảng cách càng nhỏ thì sự tương tự của chúng càng lớn và ngược lại. Hiện chưa có một kết luận nào về sự tối ưu của việc chọn khoảng cách trong bài toán phân loại. Nghiên cứu này khảo sát các loại khoảng cách trên để tìm trường hợp phù hợp nhất khi xem xét cho dữ liệu cụ thể.

2.5. Phương pháp trích xuất hình ảnh

Điểm ảnh là một phần tử của hình ảnh kỹ thuật số tại một tọa độ với độ xám hoặc màu nhất định. Chẳn hạn, trong hình ảnh hai chiều, mỗi điểm ảnh bao gồm một cặp tọa độ (x, y), trong đó x, y là các giá trị đại diện cho độ sáng cụ thể (mức xám). Các cặp tọa độ (x, y) tạo nên độ phân giải của ảnh. Cụ thể, màn hình máy tính có độ phân giải 768×1024 có nghĩa chiều rộng màn hình là 768 pixel và chiều dài là 1024 pixel.

Trong nghiên cứu này, việc trích xuất hình ảnh thành PDF dựa trên độ phân giải của ảnh theo các bước sau:

Bước 1. Xây dựng ma trận điểm ảnh với kích thước của mỗi hình ảnh trong bộ dữ liệu.

Bước 2. Chuyển ma trận điểm ảnh thành vector cột.

Bước 3. Ước lượng phi tham số mật độ điểm ảnh dựa trên phương pháp hàm hạt nhân.

3. THUẬT TOÁN ĐỀ NGHỊ

Cho k tổng thể $w_1, w_2, \dots, w_k$, trong đó nhóm thứ i có n_i ảnh, $n_1 + n_2 + \cdots + n_k = n$. Để phân loại một ảnh I_0, ta thực hiện theo các bước sau:

Bước 1. Trích xuất mỗi ảnh thành một PDF đại diện theo các bước của Mục 2.5 và xây dựng các PDF chặt cụt trên [0;1]. Gọi F_i với $i = 1,2, \dots, k$ là tập các PDF đại diện cho các ảnh của nhóm w_i và f_0 là PDF đại diện cho ảnh cần phân loại I_0.

Bước 2. Tìm xác xuất tiên nghiệm cho ảnh I_0 bởi các bước sau:

Bước 2.1. Thiết lập ma trận xác suất ban đầu $\boldsymbol{u}^{(0)}$ với k hàng và $N + 1$ cột:

$$\boldsymbol{u}^{(0)} = [\mu_{ij}]_{k\times(N+1)} = \begin{bmatrix} \mu_{11} & \mu_{12} & \cdots & \mu_{1N} & 1/k \\ \mu_{12} & \mu_{22} & \cdots & \mu_{2N} & 1/k \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mu_{k1} & \mu_{k2} & \cdots & \mu_{kN} & 1/k \end{bmatrix},$$

trong đó N cột đầu tiên là ma trận phân vùng không mờ của các phần tử trong tập dữ liệu khi xếp vào k nhóm $w_1, w_2, \dots, w_k$. Cụ thể $\mu_{ij} = 1$ khi phần tử thứ j thuộc chùm thứ i và $\mu_{ij} = 0$ nếu phần tử thứ j không thuộc chùm thứ i. Cột cuối cùng $(N + 1)$ là xác suất ban đầu để I_0 xếp vào các chùm $w_1, w_2, \dots, w_k$. Ở bước này, ta cho xác suất tiên nghiệm của I_0 bằng nhau và bằng $1/k$.

Bước 2.2. Cập nhật PDF đại diện cho các chùm w_i theo công thức (2) để có f_{v_i}, i = 1, 2, …, k.

Bước 2.3. Cập nhật ma trận xác suất mới $\boldsymbol{u}^{(t)}$, trong đó các phần tử được xác định theo công thức sau:

$$\mu_{ik}^{(t)} = \frac{1}{\sum_{j=1}^{k}\left[d\left(f_{v_i}, f_j\right)/\left(d\left(f_{v_i}, f_j\right)\right)^2\right]},$$

trong đó $d(.)$ là khoảng cách giữa hai PDF.

Bước 2.3. Tính chuẩn

$$\left\|\boldsymbol{u}^{(t)} - \boldsymbol{u}^{(t-1)}\right\| = \max_{i,j}\left\{\left|\mu_{ik}^{(t)} - \mu_{ik}^{(t-1)}\right|\right\}.$$

Bước 2.4. Lặp lại các Bước 2.2, Bước 2.3 và Bước 2.4 t lần cho đến khi điều kiện sau được thỏa:

$$\left\|\boldsymbol{u}^{(t)} - \boldsymbol{u}^{(t-1)}\right\| < \varepsilon.$$

Khi thuật toán kết thúc, ta nhận được một ma trận cỡ $k \times (N + 1)$ mà tổng các cột của nó luôn bằng 1. Cột cuối cùng trong ma trận này chính là xác suất tiên nghiệm để xếp I_0 vào các nhóm.

Bước 3. Giả sử ta có cột cuối cùng trong ma trận khi kết thúc Bước 2 là:

$$q = \begin{bmatrix} q_1 \\ q_2 \\ \vdots \\ q_k \end{bmatrix}.$$

Tính $g_{i0} = q_i[1 - d(f_0, f_{v_i})]$, trong đó f_0 là PDF được trích xuất đại diện cho ảnh I_0. Thực hiện nguyên tắc phân loại như sau:

Nếu $\max\left\{q_i\left(1 - d\left(f_0, f_{v_i}\right)\right)\right\} = q_c\left(1 - d\left(f_0, f_{v_c}\right)\right)$, c = 1, 2, …, k thì ảnh thứ j được xếp vào Nhóm w_c.

4. VÍ DỤ MINH HỌA

Để minh họa cho thuật toán đề nghị, việc phân loại cho ba loại gạo dựa vào các ảnh của chúng được thực hiện. Ba loại gạo có tên gọi là Ipsala, Jasmine và Karacadag. Đây là những loại gạo được trồng rất phổ biến ở Thái Lan. Tập dữ liệu được xem xét gồm 36 ảnh với số lượng cho mỗi loại bằng nhau được lấy từ tập dữ liệu DATASETS (muratkoklu.com). Mẫu ảnh của ba loại được cho bởi Hình 1 và Hình 2.

Ta tiến hành thực hiện thuật toán đề nghị theo các bước cụ thể sau:

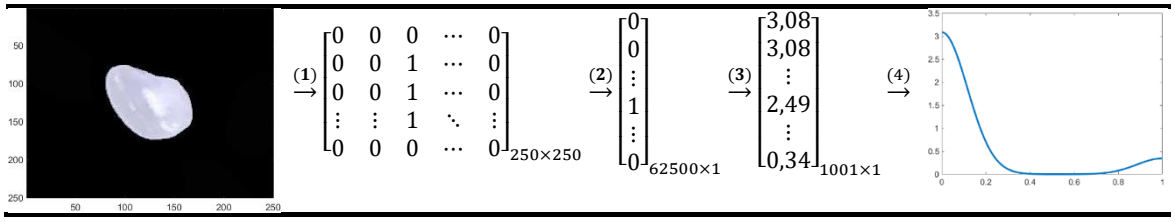
Trích xuất các PDF đặc trưng cho mỗi ảnh dựa trên tần suất điểm ảnh và phương pháp ước lượng hàm hạt nhân. Minh hoạ cho việc trích xuất một ảnh được cho bởi Hình 3.

(a) Gạo Ipsala (b) Gạo Jasmine (c) Gạo Karacadag

Hình 1. Hình ảnh minh họa các loại gạo được chụp theo chiều ngang

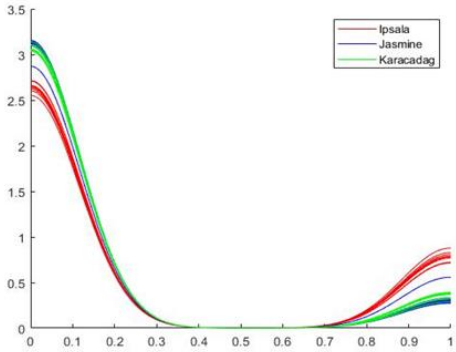
(a) Gạo Ipsala (b) Gạo Jasmine (c) Gạo Karacadag

Hình 2. Hình ảnh minh họa các loại gạo được chụp theo chiều dọc



Hình 3. Minh họa các bước trích xuất một ảnh gạo thành PDF

Thực hiện tương tự cho 36 ảnh, ta có được các PDF được cho bởi Hình 4.



Hình 4. Các PDF được trích xuất cho 36 ảnh gạo

Để đánh giá hiệu quả của các phương pháp, tập dữ liệu được chia thành tập huấn luyện và tập kiểm tra với số lượng lần lượt là 30 và 6. Tập huấn luyện được sử dụng để xây dựng thuật toán và tập kiểm tra dùng để đánh giá hiệu quả của thuật toán đã xây dựng từ tập huấn luyện.

Ta gán nhãn cho ba loại gạo Ipsala, Jasmine và Karacadag lần lượt là 1, 2 và 3. Với 30 ảnh của tập huấn luyện, ta có các nhãn của chúng được cho bởi ma trận có cấp 1×30 như sau:

$$S = [1 \quad 2 \quad 2 \quad 3 \quad \cdots \quad 1].$$

Giả sử ta cần phân loại ảnh I_0 thuộc loại gạo Ipsala. Trước tiên, ta đặt ma trận phân vùng có kích thước 3×31 như sau:

$$\boldsymbol{u}^{(0)}_{3\times 31} = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 1 & 1/3 \\ 0 & 1 & 1 & 0 & \cdots & 0 & 1/3 \\ 0 & 0 & 0 & 1 & \cdots & 0 & 1/3 \end{bmatrix}.$$

Trong ma trận trên, ta có

* Từ Cột 1 đến Cột 30, số 1 ở vị trí dòng ứng với ma trận S và số 0 ở các vị trí khác.

* Cột cuối cùng có chung giá trị $1/3$ ở mỗi dòng (xác suất tiên nghiệm xếp I_0 vào ba nhóm là như nhau).

Thực hiện các Bước 2.2 và Bước 2.3, sau 7 vòng lặp ta có chuẩn $\|\boldsymbol{u}^{(7)} - \boldsymbol{u}^{(6)}\| = 5{,}5 \times 10^{-6}$, nên

Bước 2 được dừng lại. Khi đó, cột cuối cùng của ma trận phân vùng xác suất $\boldsymbol{U}^{(7)}$ có kết quả như sau:

$$\begin{bmatrix} q_1 \\ q_2 \\ q_3 \end{bmatrix} = \begin{bmatrix} \mu_{1\times31}^{(7)} \\ \mu_{2\times31}^{(7)} \\ \mu_{3\times31}^{(7)} \end{bmatrix} = \begin{bmatrix} 0{,}9985 \\ 0{,}0006 \\ 0{,}0009 \end{bmatrix}.$$

Kết quả trên cho ta xác suất tiên nghiệm khi xếp ảnh I_0 vào các nhóm gạo Ipsala, Jasmine và Karacadag lần lượt là 0,9985; 0,0006 và 0,0009.

Thiết lập PDF đại diện cho nhóm w_i là $f_{v_i}, i = 1, 2, 3$ theo công thức (2) và tính khoảng City-Block cách từ f_0 đến các PDF đại diện của ba nhóm ta có các kết quả:

$$d(f_0, f_{v_1}) = 7{,}361,$$
$$d(f_0, f_{v_2}) = 148{,}744,$$
$$d(f_0, f_{v_3}) = 123{,}516.$$

Khi đó, xác suất hậu nghiệm khi xếp ảnh I_0 vào các lớp có các giá trị lần lượt là:

$$q_1 d(f_0, f_{v_1}) = 7{,}349,$$
$$q_2 d(f_0, f_{v_2}) = 0{,}090,$$
$$q_3 d(f_0, f_{v_3}) = 0{,}112.$$

Vì $\max\{q_i . d(f_0, f_{v_i})\} = q_1 d(f_0, f_{v_1}) = 7{,}349$ nên ta xếp ảnh I_0 vào lớp thứ nhất. Như vậy, hình ảnh hạt gạo đã được xếp vào đúng loại giống Ipsala bằng thuật toán đề xuất.

Trong các bài toán phân loại để so sánh hiệu quả của các thuật toán, các tham số như xác suất phân loại đúng toàn cục (ACC), chỉ số F_1, chỉ số Precison và chỉ số Kappa thường được sử dụng. ACC có giá trị từ 0% đến 100%, các chỉ số khác có giá trị [0;1] với 1 là kết quả tốt nhất của phân loại. Thực hiện phân loại cho tất cả các phần tử của tập kiểm tra bằng thuật toán đề nghị với khoảng cách giữa hai PDF theo các định nghĩa khác nhau ta có Bảng 1.

Bảng 1. So sánh các mô hình phân loại cho dữ liệu ba loại gạo

Thuật toán đề xuất	ACC (%)	Precision	F1	Kappa
Hausdorff	86,25	0,86	0,86	1,00
City-Block	**100**	**1,00**	**1,00**	**1,00**
Kullback-Leibler	97,25	0,97	0,97	1,00
Jesen Shannon	96,25	0,96	0,96	1,00
Bhattacharyya	91,25	0,91	0,91	1,00

Bảng 1 cho thấy thuật toán đề xuất cho kết quả khá tốt với tất cả các loại khoảng cách. ACC có giá trị thấp nhất là 86% và các chỉ số còn lại gần bằng 1. Trong các loại khoảng cách được khảo sát, khoảng cách City-Block cho kết quả tốt nhất với tất cả các chỉ số đều tối ưu.

5. ÁP DỤNG TRONG PHÂN LOẠI THUỐC BẨN VÀ KHÔNG BẨN

Nghiên cứu này sử dụng tập ảnh có tên là *PillsQC* trong nghiên cứu của Philipp et al. (2021) để thực hiện. Tập dữ liệu này gồm 287 ảnh được chia thành hai nhóm: thuốc không nhiễm bẩn (Nhóm 1) và thuốc bị nhiễm bẩn (Nhóm

2) với kích thước ảnh 225×225 pixels giống nhau cho tất cả ảnh. Mẫu ảnh thuốc của hai nhóm đại diện hai loại thuốc được cho trong Hình 5.

(a) Thuốc không nhiễm bẩn (b) Thuốc nhiễm bẩn

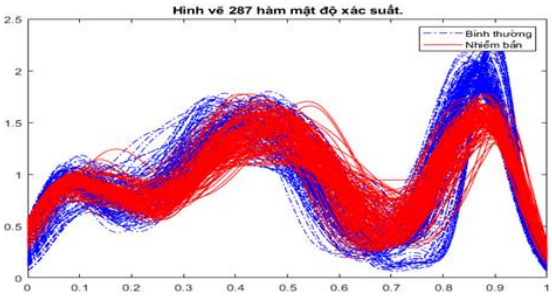
Hình 5. Mẫu ảnh thuốc nhiễm bẩn và không nhiễm bẩn

Trong ứng dụng này, tập dữ liệu được chia thành tập huấn luyện và tập kiểm tra với tỷ lệ lần lượt là 80% và 20%. Số lượng cụ thể mỗi nhóm được cho bởi Bảng 2.

Bảng 2. Số lượng ảnh cho tập huấn luyện và tập kiểm tra của mỗi nhóm

Nhóm	Tổng	Tập huấn luyện	Tập kiểm tra
Nhóm 1	149	122	27
Nhóm 2	138	108	30
Tổng	287	230	57

Trích xuất 287 ảnh thành 287 PDF, ta nhận được Hình 6.



Hình 6. Đồ thị của các PDF được trích xuất trong bộ dữ liệu hình ảnh *PillQC*

Hình 6 cho thấy sự chồng lấp các PDF đại diện cho 2 nhóm thuốc rất lớn. Do đó, việc phân loại có thể là một thách thức cho thuật toán đề nghị.

Thực hiện các bước của thuật toán đề nghị cho các khoảng cách khác nhau, ta nhận được Bảng 3.

Bảng 3. Kết quả phân loại của thuật toán đề nghị với các khoảng cách khác nhau

Thuật toán đề nghị	ACC (%)	Precision	Recall	F1-score	Kappa
Hausdorff	**100**	**1,00**	**1,00**	**1,00**	**1,00**
City-Block	94,74	0,95	0,95	0,95	1,00
Kullback-Leibler	77,19	0,77	0,77	0,77	0,57
Jesen Shannon	82,46	0,82	0,82	0,82	0,47
Bhattacharyya	84,21	0,84	0,84	0,84	0,66

Mặt dù PDF được trích xuất từ hai nhóm có sự chồng lấp lớn, khó phân biệt nhưng Bảng 3 cho thấy thuật toán đề nghị với những khoảng cách khác nhau đã cho kết quả khá tốt, trong đó trường hợp khoảng cách Hausdorff cho kết quả tốt nhất. Chúng ta cũng nhận kết quả tối ưu với tất cả các chỉ số trong trường hợp này. So sánh kết quả tốt nhất của Bảng 3 với các thuật toán phân loại dựa vào học máy và học sâu đang được sử dụng rộng rãi, ta có Bảng 4.

Bảng 4. Kết quả phân loại của các thuật toán cho 2 nhóm thuốc

Mô hình phân loại	ACC (%)	Precision	Recall	F1-score	Kappa
Đề xuất	**100**	**1,00**	**1,00**	**1,00**	**1,00**
Fisher	69,00	0,69	0,69	0,69	0,79
Logistic	94,83	0,95	0,95	0,95	0,90
kNNs	75,86	0,76	0,76	0,76	0,49
SVM	98,28	0,98	0,98	0,98	0,96
Trees	74,14	0,74	0,74	0,74	0.48
ENS	89,66	0,90	0,90	0,90	0,79
NB	91,38	0,92	0,92	0,92	0,83
Neural Networks	96,55	0,97	0,97	0,97	0,93
Optimal auto-fit	91,38	0,92	0,92	0,92	0,83
ConvNet	93,10	0,93	0,93	0,93	0,86
SqueezeNet	91,38	0,91	0,91	0,91	0,83
GoogleNet	51,72	0,52	0,52	0,52	0,00
MobileNetv2	94,83	0,95	0,95	0,95	0,90
VGG-19	52,17	0,52	0,52	0,52	0,00
Inceptionv3	98,28	0,98	0,98	0,98	0,97

Với rất nhiều phương pháp phân loại được thực hiện cho tập dữ liệu này, thuật toán đề nghị đã cho những kết quả nổi bậc với tất cả các chỉ số đánh giá.

6. KẾT LUẬN

Nghiên cứu đã đề xuất một phương pháp mới trong phân loại ảnh dựa trên độ phân giải của chúng được trích xuất thành PDF đại diện. Thuật toán này đã thể hiện được ưu điểm trong so sánh với nhiều phương pháp khác qua một tập ảnh cụ thể. Vì nhận dạng ảnh là tiền đề cho nhiều áp dụng trong cách mạng số ngày nay, nên nghiên cứu này có nhiều tiềm năng lớn để áp dụng vào nhiều lĩnh vực khác nhau của thực tế.

LỜI CẢM ƠN

Đề tài này được tài trợ bởi Trường Đại học Cần Thơ, Mã số: TSV2022-54.

TÀI LIỆU THAM KHẢO

Cinar, I. A. (2022). Identification of rice varieties using machine learning algorithms. *Journal of Agricultural Sciences, 28*(2), 307-325. https://doi.org/10.15832/ankutbd.862482

Dinh, P. T. & Tai, V. V. (2021). Automatic fuzzy genetic algorithm in clustering for images based on the extracted intervals. *Multimedia Tools and Applications, 80*(28), 35193-35215. https://doi.org/10.1007/s11042-020-09975-3.

Fisher, R. A. (1938). The statistical utilization of multiple measurements. *Annals of Eugenics,*

8(4), 376–386.https://doi.org/10.1007/s11042-020-09975-3.

Ha, C. N., Thao, N. T., Bao, T. N. Trung, N. T. & Tai, V. V. (2020). A new approach for face detection using the maximum function of probability density functions. *Annals of Operations Research*. https://doi.org/10.1007/s10479-020-03823-1.

Huang, S., N., Cai, P., Pacheco, P., Narrandes, S., Wang, Y. & Xu, W. (2018). Applications of support vector machine (SVM) learning in cancer geometrics. *Cancer Genomics-Proteomics*, *15*(1), 41–51. https:// doi.org/10.21873/cgp.20063

Koklu, M. C. (2021). Classification of rice varieties with deep learning methods. *Computers and Electronics in Agriculture, 187*, 106285. https://doi.org/10.1016/j.compag.2021.106285

Ngoc, L.T.K, Thao, N. T. & Tai, V. V. (2022). A new image classification method using interval texture feature and improved Bayesian classifier. *Multimedia Tools and Applications*. https://doi.org/10.1007/s11042-022-13531-6

Pham, B. T., Pradhan, B., Bui, D. T., Prakash, I. & Dholakia, M. (2016). A comparative study of different machine learning methods for landslide susceptibility assessment: A case study of uttarakhand area (India). *Environmental Modelling & Software, 84*, 240–250. https://doi.org/10.1016/j.envsoft.2016.07.005

Philipp, L., Lukas, R., Robert, A. V., Billy J. F., Marius, K. & Klaus-Robert M. (2021). Explainable deep one-class classification. *The conference ICLR 2021*. https://arxiv.org/abs/2007.01760.

Tai, V. V., Ha, C. N., Nghiep, L. D. & Thao, N. T. (2021). A new strategy for short-term stock in vestment using Bayesian approach. *Computational Economics.* https://doi.org/10.1007/s10614-021- 10115-8.

Thao, N. T. & Tai, V. V. (2017). A new approach for determining the prior probabilities in the classification problem by Bayesian method. *Advances in Data Analysis and Classification, 11*(3), 629–643. https://doi.org/10.1080/1351847X.2017.1419273

Terrell, G. R. (1992). Variable kernel density estimation. *The Annals of Statistics, 20*(3), 1236-1265 (September, 1992). https://doi.org/ 10.1214/aos/1176348768

Zhang, N., Ruan, S., Lebonvallet, S., Liao, Q., & Zhu, Y. (2018). Kernel feature selection to fuse multi-spectral MRI images for brain tumor segmentation. *Computer Vision and Image Understanding, 155*, 256–269

Zhu, X., Yang, J. & Waibel, A. (2000). Segmenting hands of arbitrary color in automatic face and gesture recognition. *The International Conference on Automatic Face and Gesture Recognition. IEEE*. https://doi.org/ 10.1109/AFGR.2000.84067