

BUILD AN EFFICIENT DEEP LEARNING MODEL TO RECOGNIZE SKIN DISEASE BASED ON SELF-KNOWLEDGE DISTILLATION

Phung Thi Thu Trang*, Nguyen Pham Linh Chi, Nguyen Thi Ngoc Anh, Ho Thi Thuy Dung

TNU - School of Foreign Languages

ARTICLE INFO	ABSTRACT
<p>Received: 28/10/2022</p> <p>Revised: 22/11/2022</p> <p>Published: 22/11/2022</p>	<p>Skin cancer is currently one of the most common diseases with an increasing incidence. Therefore, early prediction or recognition of skin diseases is currently of great interest to researchers around the world, especially in the ISIC skin disease classification contests of 2017, 2018, 2019 and 2020. In this paper, we propose an effective new approach to solve the problem of skin disease identification based on self-knowledge distillation. Our method exploits and minimizes the difference between two probability distributions from two different versions of the same input image. The experiment results performed with the ResNet-50 network have shown that our proposed approach outperforms the state of the art proposed methods on standard datasets such as HAM10000, ISIC 2017 and ISIC 2019. Specifically, our method achieves 0.987 in terms of AUC on the HAM10000 dataset and 0.960 in terms of AUC, 0.901 in terms of accuracy, 0.910 in terms of sensitivity, and 0.866 in terms of specificity on the ISIC 2017 dataset.</p>
<p>KEYWORDS</p> <p>Skin disease</p> <p>Deep learning</p> <p>Knowledge distillation</p> <p>Self-knowledge distillation</p> <p>Classification</p>	

XÂY DỰNG MÔ HÌNH HỌC SÂU HIỆU QUẢ ĐỂ NHẬN DẠNG BỆNH NGOÀI DA DỰA TRÊN TỰ CHUNG CÁT KIẾN THỨC

Phùng Thị Thu Trang*, Nguyễn Phạm Linh Chi, Nguyễn Thị Ngọc Anh, Hồ Thị Thùy Dung

Trường Ngoại ngữ - ĐH Thái Nguyên

THÔNG TIN BÀI BÁO	TÓM TẮT
<p>Ngày nhận bài: 28/10/2022</p> <p>Ngày hoàn thiện: 22/11/2022</p> <p>Ngày đăng: 22/11/2022</p>	<p>Ung thư da hiện đang là một trong các loại bệnh phổ biến nhất với tỉ lệ người mắc bệnh ngày càng cao. Chính vì vậy, việc dự đoán sớm các loại bệnh ngoài da hiện đang được các nhà nghiên cứu trên thế giới quan tâm đặc biệt là trong các cuộc thi phân loại bệnh ngoài da ISIC của các năm 2017, 2018, 2019 và 2020. Trong bài báo này, chúng tôi đề xuất một cách tiếp cận mới hiệu quả nhằm giải quyết bài toán nhận dạng bệnh ngoài da dựa trên quá trình tự chung cát kiến thức. Phương pháp của chúng tôi khai thác và tối thiểu hóa sự khác biệt giữa hai phân bố xác suất từ hai phiên bản khác nhau của cùng một ảnh đầu vào. Kết quả thử nghiệm được thực hiện với mạng ResNet-50 cho thấy cách tiếp cận chúng tôi đề xuất vượt trội hơn so với các phương pháp hiện đại được đề xuất gần đây trên các bộ dữ liệu chuẩn chẳng hạn như HAM10000, ISIC 2017 và ISIC 2019. Cụ thể, phương pháp của chúng tôi đạt 0,987 AUC trên bộ dữ liệu HAM10000 và 0,960 AUC, 0,901 độ chính xác, 0,910 độ nhạy và 0,866 độ đặc hiệu trên bộ dữ liệu ISIC 2017.</p>
<p>TỪ KHÓA</p> <p>Bệnh ngoài da</p> <p>Học sâu</p> <p>Chung cát kiến thức</p> <p>Tự chung cát kiến thức</p> <p>Phân lớp</p>	

DOI: <https://doi.org/10.34238/tnu-jst.6803>

* Corresponding author. Email: phungthustrang.sfl@tnu.edu.vn

1. Giới thiệu

Ung thư da là bệnh phổ biến nhất trong số các loại ung thư, với số người được chẩn đoán mắc bệnh mỗi năm nhiều hơn tất cả các loại ung thư khác cộng lại [1]. Tỷ lệ mắc ung thư da ở người da trắng là cao nhất với khoảng 200 ca trên 100.000 dân. Ở Việt Nam, tỷ lệ mắc ung thư da mặc dù không cao như ở người da trắng nhưng lại có số lượng mắc ngày càng gia tăng, cụ thể với nam giới chiếm khoảng 3,2 ca trên 100.000 dân và 3,1 ca trên 100.000 dân đối với nữ giới. Ung thư tế bào hắc tố, dạng ung thư da nguy hiểm nhất, được dự báo sẽ đạt gần nửa triệu ca vào năm 2040. Con số này tăng 62% kể từ năm 2018. Cứ 4 phút lại có một người chết vì ung thư da. Do đó, sự gia tăng tỷ lệ mắc bệnh ung thư da được nhiều bác sĩ da liễu coi là một đại dịch toàn cầu.

Can thiệp sớm đối với ung thư da, đặc biệt là khối u ác tính, là điều cần thiết để đảm bảo tỷ lệ sống sót cao khi đối mặt với số lượng ca bệnh ngày càng tăng [3]. Nguyên nhân chính có thể xác định được của ung thư da là tiếp xúc quá nhiều với bức xạ tia cực tím (UV). Nồng độ ôzôn cạn kiệt dẫn đến sự gia tăng bức xạ UV ở mặt đất, có thể làm tăng nguy cơ phơi nhiễm dưới ánh sáng mặt trời tự nhiên.

Soi da là một kỹ thuật hình ảnh được sử dụng rộng rãi cho phép nhìn thấy bề mặt da bằng cách khuếch đại ánh sáng sử dụng chất lỏng ngậm [4]. Tuy nhiên, độ chính xác chẩn đoán của nó phụ thuộc nhiều vào kinh nghiệm của bác sĩ da liễu [5], [6]. Sự khan hiếm nguồn lực chuyên gia ở các nước đang phát triển hoặc các nước nghèo đói có thể ảnh hưởng đáng kể đến việc điều trị kịp thời các bệnh ung thư da. Do đó, nhu cầu về các giải pháp chuẩn đoán tự động từ xa ngày càng trở nên quan trọng. Điều này đặc biệt thích hợp với các nước đang phát triển nơi bệnh nhân không được tiếp cận với các thiết bị y tế mới nhất và bác sĩ chuyên môn cần thiết để chuẩn đoán chính xác.

Phân loại tổn thương da đã trở thành một lĩnh vực nghiên cứu phổ biến trong những năm gần đây sau khi áp dụng ngày càng nhiều các kỹ thuật học sâu trong lĩnh vực phân tích hình ảnh y tế. Khác với các phương pháp gần đây đề xuất các kiến trúc mạng khác nhau nhằm tăng độ chính xác khi phân loại các loại bệnh ngoài da khác nhau. Trong bài báo này, chúng tôi đề xuất một cách tiếp cận dựa trên việc tự chung cất kiến thức nhằm tăng khả năng tổng quát hóa và hiệu suất của mạng trong khi vẫn giữ được sự đơn giản kiến trúc mạng và chi phí tính toán. Đóng góp chính của bài báo này có thể được liệt kê như sau:

(1) Chúng tôi đề xuất một phương pháp nhận dạng bệnh ngoài da mới thông qua tự chung cất kiến thức.

(2) Phương pháp tự chung cất được chúng tôi đề xuất giúp làm tăng khả năng học hỏi của mạng trong quá trình đào tạo. Từ đó giúp làm tăng độ chính xác tổng thể của mạng trong khi vẫn giữ được sự đơn giản và linh hoạt của mạng nhẹ mà không cần sự hướng dẫn của mạng giáo viên.

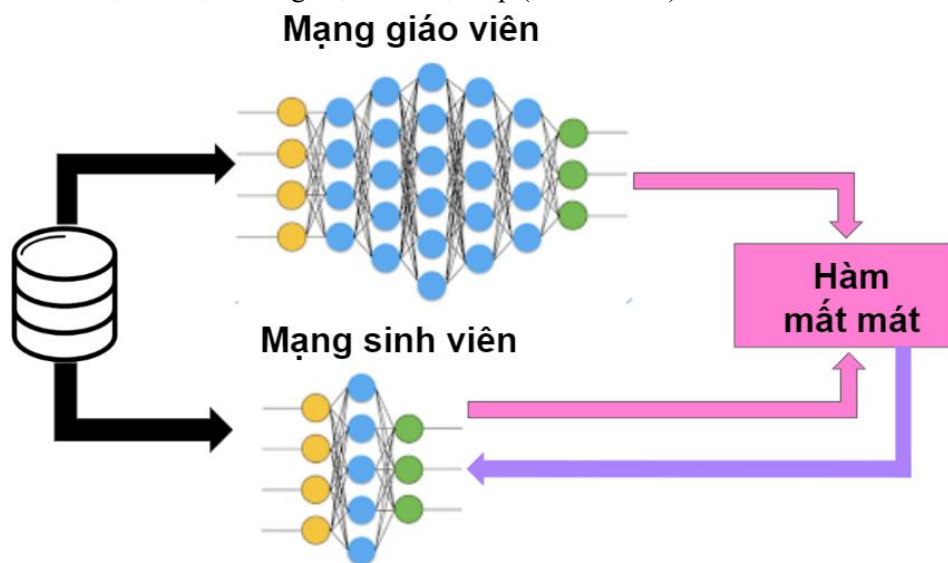
1.1. Bài toán nhận dạng bệnh ngoài da

Hầu hết các phương pháp được đề xuất giải quyết bài toán nhận dạng bệnh ngoài da hiện nay đều dựa trên học sâu và thị giác máy tính. Trong đó, hầu hết sử dụng đầu vào là ảnh RGB soi da được chụp tại các vùng da bị tổn thương với độ phóng đại lớn và độ nét cao. Năm 2020, Valle cùng các cộng sự [7] đã điều tra 10 yếu tố có thể ảnh hưởng đến hiệu suất của mô hình học sâu cho bài toán nhận dạng bệnh ngoài da bao gồm: sử dụng học tập chuyển giao (transfer learning), xây dựng kiến trúc mô hình mới, sử dụng các bộ dữ liệu khác nhau để huấn luyện, sử dụng các phương pháp chuẩn hóa khác nhau, sử dụng phân đoạn ảnh, tăng thời gian đào tạo, kết hợp với mô hình máy vector hỗ trợ (SVM), sử dụng tăng cường dữ liệu, sử dụng nhiều độ phân giải ảnh khác nhau, và sử dụng tăng cường dữ liệu cho tập dữ liệu kiểm tra. Kết quả điều tra cho thấy rằng tăng cường dữ liệu có ảnh hưởng lớn nhất đến hiệu suất của các mô hình. Jianpeng cùng các cộng sự [8] đã đề xuất một mô hình học sâu dựa trên sự chú ý (attention) giải quyết bài toán. Trong đó, nhóm tác giả đã giới thiệu một cơ chế học tập chú ý mới để cải thiện khả năng đại diện phân biệt

của mạng nơ-ron tích chập (CNN) bằng cách sử dụng các bản đồ đối tượng được học bởi lớp cao để tạo bản đồ chú ý cho một lớp thấp.

1.2. Chung cất kiến thức

Chung cất kiến thức là một trong những phương pháp phổ biến nhất trong học sâu được giới thiệu lần đầu bởi Hinton cùng các cộng sự năm 2015 [9]. Chung cất kiến thức nhằm mục đích chuyển giao kiến thức từ một mô hình nặng (thường được gọi là giáo viên – teacher) sang một mô hình nhẹ (thường được gọi là sinh viên – student) nhằm nâng cao hiệu suất của các mô hình nhẹ lên so với việc đào tạo chúng một cách độc lập (xem Hình 1).



Hình 1. Tổng quan về chung cất kiến thức

Đã có nhiều biến thể khác nhau dựa trên chung cất kiến thức được đề xuất và khai thác các khía cạnh khác nhau của quá trình chung cất kiến thức chẳng hạn như chung cất bản đồ đặc trưng [10], chung cất bản đồ gradient [11], chung cất mối quan hệ đặc trưng [6],... Chung cất kiến thức được chứng minh là phương pháp hiệu quả giúp cải thiện hiệu suất của mô hình so với quá trình đào tạo độc lập. Tuy nhiên, chung cất kiến thức vẫn còn tồn tại một số nhược điểm chẳng hạn như: Tốn nhiều chi phí về mặt thời gian và bộ nhớ do phải huấn luyện mạng giáo viên nặng, khoảng cách năng lực giữa mạng giáo viên và mạng sinh viên lớn dẫn đến sinh viên không “tiếp thu” được kiến thức từ mạng giáo viên truyền tới [12]. Chính vì vậy, phương pháp tự chung cất kiến thức được đề xuất để khắc phục hạn chế này, trong phương pháp tự chung cất kiến thức, không tồn tại mạng giáo viên, do đó, mạng sinh viên sẽ tự chất lọc và chuyển giao kiến thức cho chính bản thân mình [13], [14].

Có thể thấy rằng, hầu hết các phương pháp được đề xuất hiện nay cho bài toán nhận dạng bệnh ngoài da đều nhằm vào mục đích xây dựng một kiến trúc mạng mới hoặc các phép tăng cường dữ liệu mới. Khác với các phương pháp trên, trong bài báo này, chúng tôi đề xuất một cách tiếp cận mới lạ sử dụng phương pháp tự chung cất kiến thức giải quyết bài toán nhận dạng bệnh ngoài da. Chi tiết về phương pháp và kết quả được trình bày trong mục 2 và 3 của bài báo.

2. Tự chung cất kiến thức cho bài toán nhận dạng bệnh ngoài da

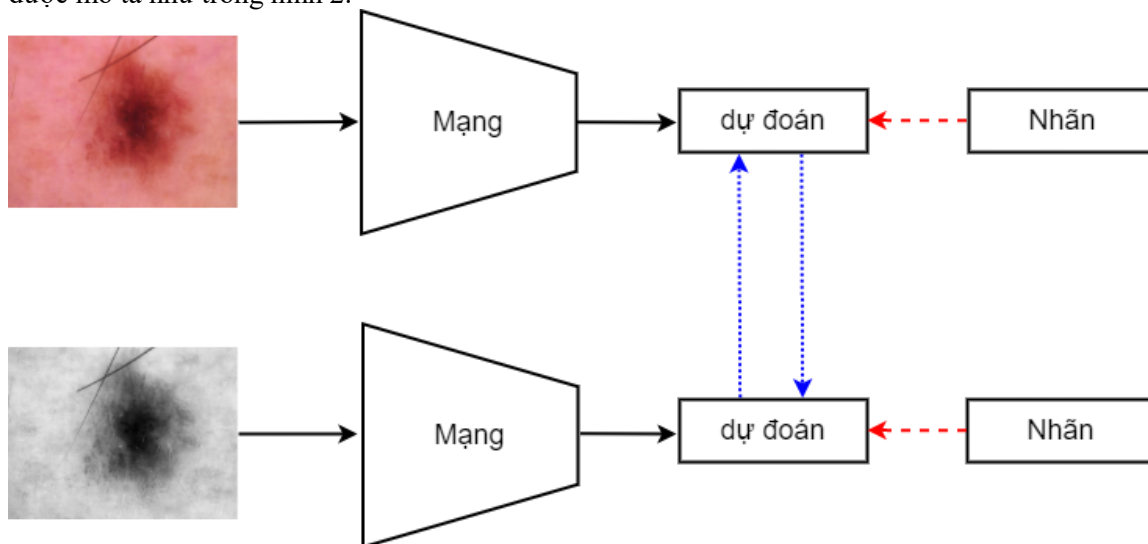
2.1. Mô tả bài toán

Cho bộ dữ liệu bệnh ngoài da $\mathcal{D} = \{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\}$ trong đó x_i với $i = 1, 2, \dots, N$ biểu diễn ảnh đầu vào dưới dạng RGB và y_i tương ứng với nhãn của chúng, N là số lượng mẫu trong tập dữ liệu. Chúng ta cần xây dựng mô hình \mathcal{F}_θ nhận đầu vào là x_i và đưa ra dự

đoán càng gần với y_i càng tốt, tức là $y_i \approx \hat{y}_i = \mathcal{F}_\theta(x_i)$ trong đó θ là bộ tham số cần huấn luyện của mô hình \mathcal{F} , \hat{y}_i là nhãn dự đoán (đầu ra) từ mô hình \mathcal{F}_θ .

2.2. Phương pháp tự chưng cất kiến thức

Lấy cảm hứng từ chưng cất kiến thức [9] và tự chưng cất kiến thức [12] - [14], chúng tôi đề xuất một mô hình tự chưng cất kiến thức nhằm giải quyết bài toán nhận dạng bệnh ngoài da như được mô tả như trong hình 2.



Hình 2. Mô hình tự chưng cất kiến thức cho nhận dạng bệnh ngoài da. Trong đó, đường nét đứt thể hiện hàm mất mát cross-entropy và đường chấm thể hiện hàm mất mát KLD.

Mô hình tự chưng cất kiến thức gồm 3 thành phần chính như sau:

- Tập các phép tăng cường dữ liệu giúp cho bộ dữ liệu trở nên đa dạng hơn mà không ảnh hưởng đến nhãn của mỗi hình ảnh.

- Một mạng sâu cho phép nhận đầu vào từ các ảnh chụp bệnh ngoài da và đưa ra dự đoán. Chúng ta có thể sử dụng nhiều kiến trúc mạng khác nhau trong phần này như VGGNet, ResNet, EfficientNet,...

- Hai hàm mất mát bao gồm hàm cross-entropy và Kullback–Leibler divergence (KLD). Trong đó, hàm cross-entropy dùng để tính toán độ sai khác giữa nhãn chân lý và kết quả dự đoán của mô hình. Hàm KLD được sử dụng để chuyển giao kiến thức giữa 2 dự đoán từ 2 phiên bản khác nhau của cùng một bức ảnh đầu vào.

Gọi $p_i^1 = \mathcal{F}_\theta(x_i^1)$ và $p_i^2 = \mathcal{F}_\theta(x_i^2)$ là đầu ra dự đoán của mô hình \mathcal{F} với x_i^1 và x_i^2 là hai biến thể đầu vào khác nhau của cùng một ảnh x_i . Các biến thể này được tạo ra bằng cách áp dụng các phép tăng cường dữ liệu khác nhau một cách ngẫu nhiên. Hàm mất mát cross-entropy tính độ lệch giữa phân bố xác suất đầu ra p_i^1 và p_i^2 so với nhãn đúng y_i như trong công thức sau:

$$\mathcal{L}_{CE} = -\frac{1}{n} \sum_{i=1}^n y_i \log p_i^1 - \frac{1}{n} \sum_{i=1}^n y_i \log p_i^2 \quad (1)$$

Chú ý y_i được mã hóa dưới dạng one-hot encoding. Ví dụ, giả sử số lượng bệnh ngoài da trong bộ dữ liệu là 4 ($K=4$), chúng ta có nhãn cho bức ảnh i là 2 thì dạng one-hot encoding của y_i sẽ tương ứng là vector $[0 \ 1 \ 0 \ 0]$, trong đó vị trí thứ 2 trong vector nhận giá trị 1 còn các giá trị khác trong K giá trị của vector sẽ nhận giá trị 0. Bên cạnh việc sử dụng hàm mất mát cross-entropy, chúng tôi đề xuất sử dụng hàm KLD để tối thiểu hóa khoảng cách giữa 2 lần dự đoán như công thức sau:

$$\mathcal{L}_{KLD} = \frac{1}{n} \sum_{i=1}^n p_i^1 \log \left(\frac{p_i^1}{p_i^2} \right) + \frac{1}{n} \sum_{i=1}^n p_i^2 \log \left(\frac{p_i^2}{p_i^1} \right) \quad (2)$$

Trong công thức (2), hàm mất mát \mathcal{L}_{KLD} nhận đầu vào là 2 phân bố xác suất p_i^1 và p_i^2 , đầu ra là độ tương đồng giữa hai phân bố xác suất này. Nếu 2 phân bố xác suất giống nhau thì $KL = 0$, hai phân bố xác suất càng khác nhau thì KL càng lớn. Mục tiêu của hàm mất mát này nhằm tối thiểu hóa sự khác nhau giữa 2 phân bố xác suất từ 2 biến thể khác nhau từ cùng một ảnh đầu vào. Cuối cùng, hàm mất mát của toàn bộ mô hình tự chưng cất kiến thức được đề xuất sẽ được tính toán như trong công thức (3).

$$\mathcal{L}_{net} = \alpha \times \mathcal{L}_{CE} + \beta \times \mathcal{L}_{KLD} \quad (3)$$

Trong đó, α và β tương ứng hai hệ số cho hai hàm mất mát.

2.3. Kiến trúc mạng

Bảng 1 mô tả kiến trúc mạng ResNet-50 [15] được sử dụng trong mô hình tự chưng cất kiến thức mà chúng tôi đề xuất. Mạng gồm 4 khối tích chập Conv Block 2, 3, 4, 5. Mỗi khối được lặp với số lần khác nhau và với độ sâu tăng dần. Các khối của lớp sau được cộng thêm giá trị từ khối trước đó thông qua skip connection [15].

Bảng 1. Kiến trúc mạng ResNet-50 cho bài toán nhận dạng bệnh ngoài da

Tên khối	Tham số	Lặp	Kích thước đầu ra
Input			(224,224,3)
Conv 1	7×7, 64	1	(112,112,64)
MaxPool	k = (3,3) s = (2,2)	1	(56, 56, 64)
Conv Block 2	1×1, 64 3×3, 64 1×1, 256	3	(56, 56, 256)
Conv Block 3	1×1, 128 3×3, 128 1×1, 512	4	(28, 28, 512)
Conv Block 4	1×1, 256 3×3, 256 1×1, 1024	6	(14, 14, 1024)
Conv Block 5	1×1, 512 3×3, 512 1×1, 2048	3	(7,7, 2028)
Ave Pool		1	(2048,)
FC	K units	1	(K,)

Đầu ra của tất cả các tầng tích chập mặc định đều được đưa vào tầng Batch Normalization và ReLU. Ở cột tham số, 7×7; 64 thể hiện rằng tầng tích chập đó có kích thước hạt nhân là 7×7 và số lượng bộ lọc là 64. Với khối MaxPool, k đại diện cho kích thước hạt nhân và s là bước nhảy. Khối FC đại diện cho tầng Fully Connected, trong tầng này chúng tôi sử dụng hàm softmax để đưa ra dự đoán phân lớp cho hình ảnh đầu vào.

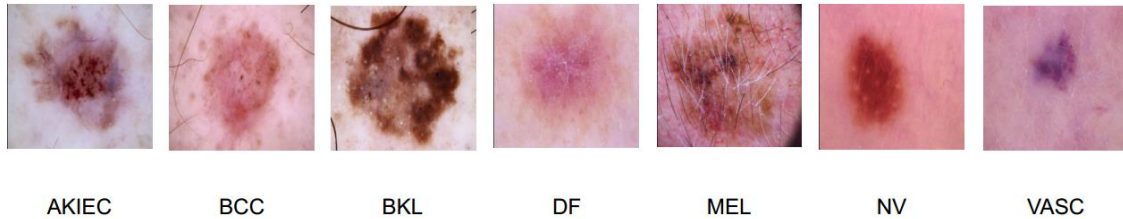
3. Thử nghiệm

3.1. Bộ dữ liệu và chi tiết cài đặt

Thử nghiệm được thực hiện trên ba tập dữ liệu riêng biệt là bộ dữ liệu HAM [16], bộ dữ liệu ISIC 2017 và bộ dữ liệu ISIC 2019.

Bộ dữ liệu HAM10000 [16] bao gồm 10015 hình ảnh soi da có kích thước 450 × 600. Nó bao gồm 7 phân loại chẩn đoán như sau: U hắc tố (MEL), Tế bào hắc tố Nevi (NV), Ung thư biểu mô

tế bào đáy (BCC), vảy sừng kích hoạt, và Ung thư biểu mô trong (AKIEC), vảy sừng lành tính (BKL), U da (DF), Tồn thương mạch máu (VASC). Tất cả các hình ảnh được thay đổi kích thước thành $224 \times 224 \times 3$ trước khi đưa vào mô hình dự đoán. Một số ví dụ về ảnh soi da của 7 lớp được thể hiện như trong hình 2. Bộ dữ liệu gồm 9187 hình ảnh cho tập huấn luyện và 828 hình ảnh cho tập thử nghiệm.



Hình 2. Một số ví dụ về 7 lớp trong bộ dữ liệu HAM10000

Tập dữ liệu ISIC 2017 bao gồm 2600 hình ảnh kích thước 767×1022 . Trong tập dữ liệu đào tạo có 2000 hình ảnh của 3 bệnh lý như sau: u lành tính, dày sừng tiết bã và u ác tính. Bộ dữ liệu thử nghiệm bao gồm 600 hình ảnh.

Bộ dữ liệu ISIC 2019 bao gồm số lượng hình ảnh lớn nhất thế giới về các ảnh soi tổn thương da được kiểm soát chất lượng và công khai để nghiên cứu. Bộ dữ liệu gồm 25.331 hình ảnh được chia thành 8 danh mục khác nhau gồm: U ác tính (MEL), Melanocytic nevus (NV), ung thư biểu mô tế bào (BCC), vảy sừng hoạt tính (AK), vảy sừng lành tính (BKL), Dermatofibroma (DF), tồn thương mạch máu (VASC), ung thư biểu mô tế bào vảy (SCC) và hình ảnh da bình thường (UNK). Trong khuôn khổ bài báo, bộ dữ liệu ISIC 2019 được chia với 80% cho tập huấn luyện và 20% cho tập thử nghiệm. Dữ liệu trong các tập dữ liệu sau đó được làm sạch để loại bỏ sự mất cân bằng của lớp. Điều này được thực hiện bằng quá trình over-sampling và under-sampling để có số lượng hình ảnh trên mỗi lớp bằng nhau. Sau đó, các hình ảnh được chuẩn hóa bằng cách chia mỗi pixel cho 255 để giữ các giá trị pixel trong phạm vi 0 đến 1. Kích thước lô (batch size) được thiết lập bằng 128. Chúng tôi sử dụng trình tối ưu hóa Stochastic gradient descent (SGD) với tốc độ học (learning rate) được khởi tạo bằng 0,01 và sẽ giảm dần 10 lần nếu trong 10 kỷ nguyên liên tiếp mà mô hình không cải thiện được kết quả trên tập kiểm tra. Theo [12], giá trị α và β được chúng tôi thiết lập tương ứng là 1 và 0,1. Các phép tăng cường dữ liệu được sử dụng dựa trên phương pháp RandAugment [17]. Tất cả các kết quả thử nghiệm đều được đánh giá trên tập thử nghiệm.

3.2. Kết quả thử nghiệm

Đầu tiên, chúng tôi so sánh phương pháp tự chưng cất được đề xuất với một số phương pháp hiện đại được đề xuất gần đây như [15], [18]-[20]. Kết quả được thể hiện trong Bảng 2, chúng tôi thấy rằng mô hình tự chưng cất chúng tôi đề xuất đạt hiệu suất tốt nhất 5/7 lớp gồm AKIEC, BKL, DF, MEL, và VASC.

Bảng 2. Kết quả thử nghiệm với bộ dữ liệu HAM10000 và so sánh phương pháp được đề xuất với các phương pháp hiện đại được công bố gần đây. Trong đó, SA là Soft Attention. Kết quả được báo cáo bằng độ đo AUC [21] trên tập kiểm tra.

Lớp	AKIEC	BCC	BKL	DF	MEL	NV	VASC
[18]	0,993	0,997	0,970	0,973	0,965	0,984	1,000
[18] + SA	0,981	0,998	0,982	0,982	0,974	0,984	1,000
[19]	0,975	0,993	0,960	0,851	0,963	0,975	0,993
[20]	0,949	0,977	0,93	0,847	0,925	0,954	0,972
[15]	0,980	0,997	0,948	0,973	0,961	0,974	0,995
Của chúng tôi	0,994	0,997	0,988	0,970	0,981	0,980	1,000

Chúng tôi tiếp tục thử nghiệm trên bộ dữ liệu ISIC 2017 và so sánh với một số phương pháp được đề xuất gần đây như [15], [22]-[25]. Từ kết quả ở Bảng 3, chúng tôi thấy rằng phương pháp được đề xuất đạt hiệu suất tốt nhất với 0,960 cho độ đo AUC, 0,901 cho độ chính xác, và 0,910 cho độ nhạy.

Bảng 3. Kết quả thử nghiệm với bộ dữ liệu ISIC 2017 và so sánh với các phương pháp được đề xuất gần đây trên 4 độ đo gồm AUC, độ chính xác, độ nhạy và độ đặc hiệu trên tập kiểm tra

Mạng	AUC	Độ chính xác	Độ nhạy	Độ đặc hiệu
ResNet-50 [15]	0,948	0,842	0,867	0,837
RAN50 [22]	0,942	0,862	0,878	0,859
SENet50 [23]	0,952	0,863	0,856	0,865
ARL-CNN50 [24]	0,958	0,868	0,878	0,867
Irv2+SA [25]	0,935	0,898	0,945	0,711
Của chúng tôi (ResNet-50)	0,960	0,901	0,910	0,866

Mở rộng thử nghiệm trên bộ dữ liệu ISIC 2019 và so sánh với phương pháp trong [26]. Bảng 4 cho thấy phương pháp được chúng tôi đề xuất cho kết quả tốt hơn phương pháp [26] ở hầu hết các lớp như MEL, NV, BCC, AK, BKL,... trên các độ đo chuẩn như AUC, độ nhạy và độ đặc hiệu.

Bảng 4. Kết quả thử nghiệm với bộ dữ liệu ISIC 2019 và so sánh với phương pháp [26] bằng ba độ đo gồm AUC, độ nhạy và độ đặc hiệu trên tập kiểm tra

Lớp	[26]			Của chúng tôi		
	AUC	Độ nhạy	Độ đặc hiệu	AUC	Độ nhạy	Độ đặc hiệu
MEL	92,8	59,4	96,2	93,4	62,1	96,8
NV	96,0	71,0	97,5	97,5	72,6	98,1
BCC	94,9	72,1	94,0	95,1	74,2	95,6
AK	91,4	48,4	96,5	92,9	56,6	97,1
BKL	90,4	39,4	98,5	92,4	45,6	99,0
DF	97,9	57,8	99,2	98,7	60,2	99,2
VASC	95,6	64,4	99,1	96,1	65,6	99,5
SCC	93,8	43,9	98,6	93,2	50,3	99,2
UNK	77,5	0,3	99,9	80,5	40,5	98,8

4. Kết luận

Trong bài báo này chúng tôi đề xuất một mô hình tự chưng cất kiến thức đơn giản nhưng hiệu quả nhằm giải quyết bài toán nhận dạng bệnh ngoài da. Phương pháp tự chưng cất được chúng tôi đề xuất nhằm mục đích tối thiểu hóa khoảng cách giữa hai phân phối xác suất của hai khung nhìn khác nhau của cùng một ảnh đầu vào. Kết quả thử nghiệm được tiến hành trên các bộ dữ liệu chuẩn cho thấy cách tiếp cận chúng tôi đề xuất cho kết quả tốt nhất trên nhiều độ đo khác nhau như AUC, độ chính xác, độ nhạy và độ đặc hiệu. Trong tương lai, chúng tôi sẽ tiếp tục cải tiến phương pháp và kết hợp chúng với học tập tương phản và/hoặc học đại diện để nâng cao hiệu suất của mô hình hơn nữa. Bên cạnh đó, việc nghiên cứu lý thuyết về tự chưng cất cũng như học tập hợp tác cũng là hướng tiếp cận mà chúng tôi quan tâm.

TÀI LIỆU THAM KHẢO/ REFERENCES

- [1] The Skin Cancer Foundation, "Skin Cancer Facts & Statistics," May 2022. [Online]. Available: <https://www.skincancer.org/skin-cancer-information/skin-cancer-facts/>. [Accessed Sept. 1, 2022].
- [2] Melanoma UK, "2020 Melanoma skin cancer report," May 2020. [Online]. Available: <https://www.melanomauk.org.uk/2020-melanoma-skin-cancer-report>. [Accessed Sept. 1, 2022].
- [3] R. K. Voss, T. N. Woods, K. D. Cromwell, K. C. Nelson, and J. N. Cormier, "Improving outcomes in patients with melanoma: strategies to ensure an early diagnosis," *Patient related outcome measures*, vol. 6, pp. 229-242, 2015.

- [4] H. Kittler, H. Pehamberger, K. Wolff, and M. J. T. I. O. Binder, "Diagnostic accuracy of dermoscopy," *The lancet oncology*, vol. 3, no. 3, pp. 159-165, 2022.
- [5] T. J. Brinker, A. Hekler, A. H. Enk, J. Klode, A. Hauschild, C. Berking, and P. Schrüfer, "A convolutional neural network trained with dermoscopic images performed on par with dermatologists in a clinical melanoma image classification task," *European Journal of Cancer*, vol. 111, pp. 148-154, 2019.
- [6] T. J. Brinker, A. Hekler, A. H. Enk, J. Klode, A. Hauschild, C. Berking, and P. Schrüfer, "Deep learning outperformed dermatologists in a head-to-head dermoscopic melanoma image classification task," *European Journal of Cancer*, vol. 113, pp. 47-54, 2019.
- [7] E. Valle, M. Fornaciali, A. Menegola, J. Tavares, F. V. Bittencourt, L. T. Li, and S. Avila, "Data, depth, and design: Learning reliable models for skin lesion analysis," *Neurocomputing*, vol. 383, pp. 303-313, 2020.
- [8] J. Zhang, Y. Xie, Y. Xia, and C. Shen, "Attention residual learning for skin lesion classification," *IEEE transactions on medical imaging*, vol. 38, no. 9, pp. 2092-2103, 2019.
- [9] G. Hinton, O. Vinyals, and J. Dean, "Distilling the knowledge in a neural network," *Annual Conference on Neural Information Processing Systems*, 2015, pp. 1-9.
- [10] R. Adriana, B. Nicolas, K. S. Ebrahimi, C. Antoine, G. Carlo, and B. Yoshua, "Fitnets: Hints for thin deep nets," *Proc. International Conference on Learning Representation (ICLR)*, 2015, pp. 1-13.
- [11] T. Guo, C. Xu, S. He, B. Shi, C. Xu, and D. Tao, "Robust student network learning," *IEEE transactions on neural networks and learning systems*, vol. 31, no. 7, pp. 2455-2468, 2019.
- [12] D. Q. Vu, N. Le, and J. C. Wang, "Teaching yourself: A self-knowledge distillation approach to action recognition," *IEEE Access*, vol. 9, pp. 105711-105723, 2021.
- [13] D. Q. Vu and J. C. Wang, "A novel self-knowledge distillation approach with siamese representation learning for action recognition," *International Conference on Visual Communications and Image Processing (VCIP)*, 2021, pp. 1-5.
- [14] Q. V. Duc, T. Phung, M. Nguyen, B. Y. Nguyen, and T. H. Nguyen, "Self-knowledge Distillation: An Efficient Approach for Falling Detection," *International Conference on Artificial Intelligence and Big Data in Digital Era*, 2022, pp. 369-380.
- [15] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770-778.
- [16] P. Tschandl, C. Rosendahl, and H. Kittler, "The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions," *Scientific data*, vol. 5, no. 1, pp. 1-9, 2018.
- [17] E. D. Cubuk, B. Zoph, J. Shlens, and Q. V. Le, "Randaugment: Practical automated data augmentation with a reduced search space," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*, 2020, pp. 702-703.
- [18] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. A. Alemi, "Inception-v4, inception-resnet and the impact of residual connections on learning," *Thirty-first AAAI conference on artificial intelligence*, 2017, pp. 1-12.
- [19] G. Huang, Z. Liu, L. V. D. Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," *IEEE conference on computer vision and pattern recognition*, 2017, pp. 4700-4708.
- [20] K. Simonyan, and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," *The 3rd International Conference on Learning Representations (ICLR)*, 2015, pp. 1-14.
- [21] J. Huang and C. X. Ling, "Using AUC and accuracy in evaluating learning algorithms," *IEEE Transactions on knowledge and Data Engineering*, vol. 17, no. 3, pp. 299-310, 2005.
- [22] F. Wang, M. Jiang, C. Qian, S. Yang, C. Li, H. Zhang, and X. Tang, "Residual attention network for image classification," *IEEE conference on computer vision and pattern recognition*, 2017, pp. 3156-3164.
- [23] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," *IEEE conference on computer vision and pattern recognition*, 2018, pp. 7132-7141.
- [24] J. Zhang, Y. Xie, Y. Xia, and C. Shen, "Attention residual learning for skin lesion classification," *IEEE transactions on medical imaging*, vol. 38, no.9, pp. 2092-2103, 2019.
- [25] S. K. Datta, M. A. Shaikh, S. N. Srihari, and M. Gao, "Soft Attention Improves Skin Cancer Classification Performance," *Interpretability of Machine Intelligence in Medical Image Computing, and Topological Data Analysis and Its Applications for Medical Data*, 2021, pp. 13-23.
- [26] N. Gessert, M. Nielsen, M. Shaikh, R. Werner, and A. Schlaefel, "Skin lesion classification using ensembles of multi-resolution EfficientNets with meta data," *MethodsX*, vol. 7, pp. 1-8, 2020.