

MỘT MÔ HÌNH DEEP LEARNING HIỆU QUẢ CHO BÀI TOÁN NHẬN DẠNG MỔNG MẮT

Phùng Thị Thu Trang¹, Ma Thị Hồng Thu²

¹Khoa Ngoại ngữ - ĐH Thái Nguyên, ²Đại học Tân Trào

TÓM TẮT

Bài toán nhận dạng bộ phận của cơ thể người là một trong những bài toán quan trọng trong lĩnh vực tin sinh học. Chúng có nhiều ứng dụng thực tế, đặc biệt là trong lĩnh vực bảo mật. Các đặc điểm sinh trắc học đang được nghiên cứu hiện nay như: vân tay, mống mắt, DNA, giọng nói, dáng người,... Trong đó, mống mắt là một trong những đặc điểm sinh trắc học tốt nhất. Các phương pháp truyền thống thường dựa trên việc so khớp giữa các mống mắt để đưa ra kết luận và chúng rất hiệu quả trong các điều kiện thuận lợi. Tuy nhiên, với các điều kiện bất lợi thì độ chính xác của các phương pháp bị giảm đi đáng kể. Trong bài báo này, chúng tôi đề xuất một mô hình học sâu dựa trên mạng nơron nhân tạo để nhận diện và phân loại đối tượng dựa trên mống mắt đạt độ chính xác cao trong điều kiện không thuận lợi.

Từ khóa: Học sâu, Mô hình học sâu, Mống mắt, Nhận dạng mống mắt, Mạng nơron

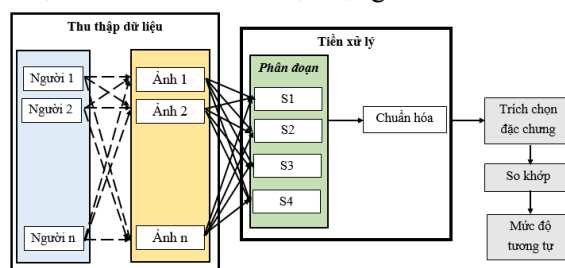
GIỚI THIỆU

Trong kỷ nguyên công nghệ thông tin hiện nay, các hệ thống nhận dạng dựa trên đặc điểm sinh học đang được sử dụng phổ biến và thu hút ngày càng nhiều sự chú ý trên toàn thế giới. Các hệ thống này khai thác các đặc điểm hành vi hoặc sinh lý học để phân biệt được các đối tượng với nhau. Các đặc điểm sinh lý như: DNA, mống mắt, mặt, vân tay, bàn tay, giọng nói, dáng người,... đang được sử dụng rộng rãi như là các tính năng sinh trắc học để phân biệt giữa người này với người khác.

Mống mắt được coi là một trong những tính năng sinh trắc học tốt nhất, dựa trên kết cấu của nó chúng ta hoàn toàn có thể xác định được thông tin của một người. Mống mắt được bảo vệ bởi mí mắt và giác mạc để tránh bị tổn thương từ môi trường bên ngoài. Nhận diện mống mắt được chia thành nhiều giai đoạn khác nhau (Hình 1), trong đó giai đoạn trích chọn mống mắt (phân đoạn) và giai đoạn so khớp đưa ra mức độ tương tự được coi là quan trọng hơn cả.

Đã có nhiều thuật toán được đề xuất để giải quyết hai vấn đề này (trong [3], [5-7]), chúng đều đạt hiệu suất lên đến 99% trong môi trường thuận lợi. Tuy nhiên, với các trường hợp như: người nhận dạng đeo kính, di

chuyển đầu hoặc đứng ở một khoảng cách xa với camera, thiếu ánh sáng,... thì hình ảnh thu được rất khó khăn để nhận dạng.



Hình 1. Mô hình một hệ thống nhận dạng mống mắt

Học sâu (deep learning) là một nhánh của học máy, hiện đang được nghiên cứu và áp dụng rộng rãi trên nhiều lĩnh vực. Trong bài báo này, chúng tôi xây dựng một mô hình học sâu giải quyết bài toán nhận dạng mống mắt đạt hiệu suất cao trong các tình huống không thuận lợi.

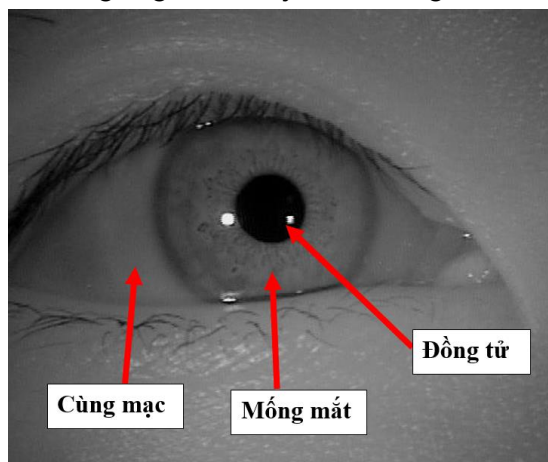
Phần tiếp theo của bài báo như sau: Trong Phần 2 chúng tôi mô tả tổng quan về bài toán nhận dạng mống mắt cùng các nghiên cứu đã được công bố gần đây. Ở Phần 3 chúng tôi đưa ra mô hình học sâu và phân tích mô hình. Chi tiết về các thử nghiệm cùng phân tích, đánh giá được trình bày trong Phần 4. Phần 5 là kết luận của bài báo.

CÁC NGHIÊN CỨU GẦN ĐÂY

Quá trình nhận dạng mống mắt được thực hiện lần lượt qua các bước sau: Từ một hay

* Tel: 01695 314806, Email: phungthutrang.sfl@tnu.edu.vn

nhiều bức ảnh mống mắt được chụp của một hay nhiều người (Hình 2), chúng ta tiến hành tiền xử lý. Trong bước này, viền trong và ngoài mống mắt được trích chọn dựa trên một thuật toán phân đoạn, sau khi có vùng mống mắt, chúng ta chuẩn hóa chúng bằng cách đưa về tọa độ cực. Cuối cùng, chúng ta sẽ trích chọn đặc trưng và so khớp để đánh giá mức độ tương tự giữa hai hay nhiều mống mắt.



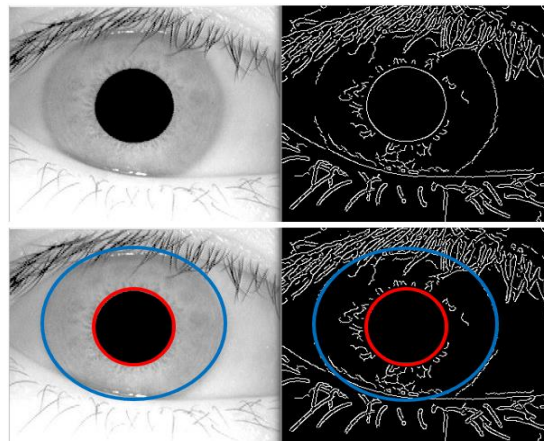
Hình 2. Mô tả vị trí của mống mắt

Thuật toán phân đoạn nhằm xác định vùng mống mắt đầu tiên được đề xuất bởi Daugman năm 1994 và năm 1997 Wildes cùng các cộng sự đã áp dụng biến đổi Hough tìm đường tròn đạt hiệu suất khá cao. Đã có rất nhiều các tác giả sử dụng biến đổi Hough này như trong [5-7] với biểu đồ cạnh của bức ảnh (canny edge) cho quá trình xác định đường viền mống mắt (Hình 3). Trong bài báo này, bán kính của đồng tử được duyệt trong khoảng (28, 75) pixel, bán kính mống mắt được duyệt trong khoảng (80, 150) pixel.

Contrast-Adjusted Hough Transform (CHT) là thuật toán dùng để phân chia mống mắt được Masek [3] đề xuất, với mỗi bức ảnh phương pháp này sẽ điều chỉnh độ tương phản nhằm làm rõ vùng biên giữa mống mắt với đồng tử và mống mắt với cùng mạc đồng thời loại bỏ các cạnh không cần thiết.

Weighted Adaptive Hough and Ellipsopolar Transform (WHT) [7] sử dụng phép biến đổi Hough ở nhiều độ phân giải khác nhau để ước lượng gần đúng tâm của mống mắt. Sau đó, sẽ tìm đường biên thứ nhất giữa mống mắt và

cùng mạc, biến đổi Ellipsopolar sẽ tìm đường biên thứ hai giữa mống mắt và đồng tử dựa trên đường biên thứ nhất.



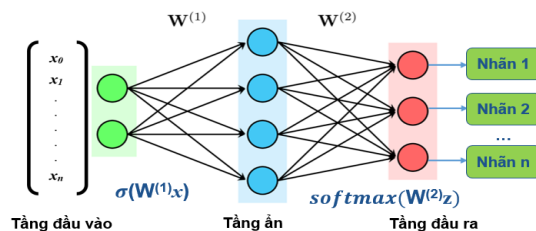
Hình 3. Quá trình xác định vùng mống mắt bằng thuật toán biến đổi Hough

Thuật toán Viterbi [8] được trình bày như sau: Đầu tiên chúng ta sẽ tiến hành xác định vùng đồng tử và mống mắt, đây là vùng rất dễ xác định vì chúng có màu tối hơn so với cùng mạc và da (nhưng cần lưu ý loại bỏ các điểm trắng trên vùng này do phản xạ ánh sáng), sau đó ta thu được tâm và đường biên bên ngoài mống mắt. Bước thứ hai, chúng ta dựa vào tâm đã thu được để tìm ra chính xác vùng đồng tử.

Đối với bước so khớp thì hầu hết các tác giả hiện này đều sử dụng phương pháp đo khoảng cách Hamming giữa hai mống mắt (sau khi đã chuẩn hóa). Nếu chúng nhỏ hơn một ngưỡng θ thì ta kết luận hai mống mắt đó của cùng một người và ngược lại.

ĐỀ XUẤT THUẬT TOÁN

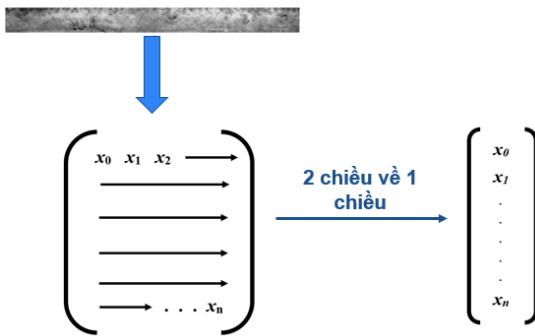
Trong phần này, chúng tôi đề xuất thuật toán DeepLearning-Iris dựa trên mô hình mạng nơron để giải quyết bài toán nhận dạng mống mắt. Kiến trúc của mô hình được mô tả trong Hình 4.



Hình 4. Mô hình mạng nơron

Đầu tiên với mỗi bức ảnh đầu vào chúng tôi áp dụng kỹ thuật WHT để trích chọn ra vùng mỏng mắt, sau đó chuẩn hóa chúng ta thu được một ma trận hai chiều, tiếp theo tiến hành giảm số chiều của ma trận thành một vector (Hình 5).

Mỏng mắt đã được chuẩn hóa



Hình 5. Giảm số chiều của ma trận ảnh mỏng mắt
Vector này chính là đầu vào của mô hình mạng nơron. Trên các tầng ẩn, chúng tôi sử dụng hàm sigmoid như trong biểu thức (1):

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (1)$$

Dễ dàng nhận thấy rằng khi $x \in (-\infty; +\infty)$ thì $\sigma(x) \in (0; 1)$. Đây là một hàm activation thông dụng nhất trong mạng nơron.

Ở tầng output, chúng tôi sử dụng hàm softmax để đưa ra phân bố xác suất của từng lớp với mỗi lớp ứng với mỗi đối tượng. Hàm softmax được mô tả như ở biểu thức (2).

$$a_i = \frac{\exp(z_i)}{\sum_{j=1}^C \exp(z_j)} \quad \forall i = 1, 2, \dots, C \quad (2)$$

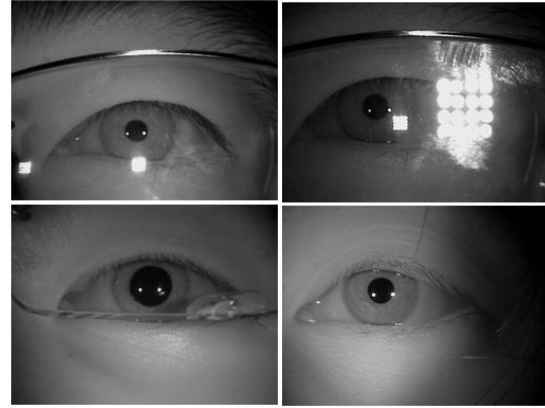
Trong đó a_i chính là phân bố xác suất của z_i trong C lớp.

THỬ NGHIỆM

Trong phần này, chúng tôi tiến hành thử nghiệm thuật toán Deeplearning-Iris trên bộ dữ liệu CASIA V4.

CASIA Iris là bộ cơ sở dữ liệu về mỏng mắt đầu tiên do các nhà nghiên cứu khoa học được Viện hàn lâm khoa học và Viện tự động hóa Trung Quốc thu thập. Đến thời điểm hiện nay, đã có 4 phiên bản dữ liệu được công bố. Trong đó, phiên bản thứ 4 là bộ cơ sở dữ liệu về mỏng mắt mà các hình ảnh được chụp khi

các đối tượng di chuyển, ở một khoảng cách xa, sử dụng kính và chất lượng kém. CASIA V4 được coi là bộ dữ liệu mới nhất và hiện đang được cộng đồng nhà nghiên cứu trên thế giới sử dụng nhiều nhất.



Hình 6. Một vài ví dụ về bộ cơ sở dữ liệu CASIA V4

Chúng ta có thể tải bộ dữ liệu này từ địa chỉ: <http://biometrics.idealtest.org/dbDetailForUse.r.do?id=4>.

Bộ tham số chúng tôi sử dụng như sau: Kích thước ảnh mỏng mắt sau khi chuẩn hóa là 64x512, suy ra vector đầu vào mô hình mạng nơron là 32.768, số lớp là 20, số tầng ẩn là 3, số nút trên mỗi tầng ẩn lần lượt là 100, 50 và 50. Bảng thử nghiệm chúng tôi thấy rằng nếu thiết lập số nút và số tầng ẩn lớn hơn cũng không cải thiện được độ chính xác thậm chí còn giảm thời gian chạy của thuật toán do số lượng tham số cần tối ưu là lớn hơn. 80% dữ liệu được sử dụng để huấn luyện mô hình, 10% được dùng để đánh giá và phần còn lại được dùng để kiểm thử. Mã nguồn của chương trình được viết trên Matlab 2017b và có thể tải về từ địa chỉ: <https://goo.gl/MCwPQ9>. Các thử nghiệm được tiến hành trên máy tính có cấu hình là: CPU Intel i7 3.6GHz, RAM 20GB, Windows 10 64bit.

Bảng 1. So sánh độ chính xác của thuật toán Deeplearning-Iris với các thuật toán khác.

Thuật toán	Độ chính xác
PCA [4]	92%
KELM [1]	98.6%
Deeplearning-Iris	99.5%

Từ Bảng 1 ta có thể thấy rằng, độ chính xác của thuật toán Deeplearning-Iris cao hơn 7.5% so với thuật toán PCA được Rana cùng các cộng sự công bố năm 2015 và cao hơn 0.9% so với thuật toán KELM do Ali công bố năm 2016.

Bảng 2. So sánh EER của thuật toán Deeplearning-Iris với các thuật toán khác.

Thuật toán	Mức độ lỗi cân bằng
PCA [2]	5.7
EM [2]	5.6
LP[2]	5.1
FV (iris) [9]	0.040
FV (eye) [9]	0.039
O ² PT (iris) [9]	0.040
O ² PT (iris) + FV (eye) [9]	0.021
Deeplearning-Iris	0.027

Bảng 2 so sánh kết quả của các thuật toán dựa trên độ đo mức độ lỗi cân bằng (EER - Equal Error Rate). Mức độ lỗi cân bằng là độ đo được sử dụng nhiều trong sinh trắc học, dùng để so sánh các loại thiết bị và công nghệ sinh trắc học khác nhau. Độ đo càng thấp chứng tỏ hệ thống càng tin cậy và chính xác.

Thuật toán Deeplearning-Iris chỉ xếp sau thuật toán O²PT (iris) + FV (eye) do Yang cùng các cộng sự đề xuất.

KẾT LUẬN

Trong bài báo này, chúng tôi đã đề xuất một thuật toán dựa trên mô hình mạng nơron để giải quyết bài toán nhận dạng mống mắt. Kết quả của thuật toán được so sánh với các kết quả khác được công bố gần đây trên cùng một bộ dữ liệu. Trong đó, thuật toán Deeplearning-Iris vẫn tỏ ra nổi trội hơn cả. Trong tương lai, chúng tôi sẽ thử nghiệm mô hình này với một số bộ dữ liệu khác và áp dụng giải các bài toán khác.

TÀI LIỆU THAM KHẢO

1. Ali, L.E., Luo, J. and Ma, J., 2016, November. Iris recognition from distant images based on multiple feature descriptors and classifiers. In Signal Processing (ICSP), 2016 IEEE 13th International Conference on (pp. 1357-1362). IEEE.
2. Llano, E.G., García-Vázquez, M.S., Zamudio-Fuentes, L.M., Vargas, J.M.C. and Ramírez-Acosta, A.A., 2017. Analysis of the Improvement on Textural Information in Human Iris Recognition. In VII Latin American Congress on Biomedical Engineering CLAIB 2016, Bucaramanga, Santander, Colombia, October 26th-28th, 2016 (pp. 373-376). Springer, Singapore.
3. Masek, L.: Recognition of human iris patterns for biometric identification. Technical report (2003)
4. Rana, H.K., Azam, M.S. and Akhtar, M.R., 2017. Iris Recognition System Using PCA Based on DWT. SM J Biometrics Biostat, 2(3), p.1015.
5. Sanchez-Gonzalez, Y., Chacon-Cabrera, Y. and Garea-Llano, E., 2014, November. A comparison of fused segmentation algorithms for iris verification. In Iberoamerican Congress on Pattern Recognition (pp. 112-119). Springer, Cham.
6. Sutra, G., Garcia-Salicetti, S. and Dorizzi, B., 2012, March. The Viterbi algorithm at different resolutions for enhanced iris segmentation. In Biometrics (ICB), 2012 5th IAPR International Conference on (pp. 310-316). IEEE.
7. Uhl, A. and Wild, P., 2012, March. Weighted adaptive hough and ellipsopolar transforms for real-time iris segmentation. In Biometrics (ICB), 2012 5th IAPR International Conference on (pp. 283-290). IEEE.
8. Wildes, R.P., Asmuth, J.C., Green, G.L., Hsu, S.C., Kolczynski, R.J., Matey, J.R. and McBride, S.E., 1994, December. A system for automated iris recognition. In Applications of Computer Vision, 1994., Proceedings of the Second IEEE Workshop on (pp. 121-128). IEEE.
9. Yang, G., Zeng, H., Li, P. and Zhang, L., 2015, September. High-order information for robust iris recognition under less controlled conditions. In Image Processing (ICIP), 2015 IEEE International Conference on (pp. 4535-4539). IEEE

SUMMARY
**AN EFFECTIVE DEEP LEARNING MODEL
FOR IRIS RECOGNITION PROBLEM**

Phung Thi Thu Trang^{1*}, Ma Thi Hong Thu²

¹*School of Foreign Languages - TNU*

²*Tan Trao University*

Recognition problem on human parts is one of the most important problems in bioinformatics, they have many practical applications especially in the field of security. Biometric characteristics are currently being researched as: Fingerprint, iris, DNA, voice, body,... The iris is one of the best property for recognition. Traditional methods are often based on iris matching to give good results in favorable conditions. However, the precision of the methods decreased significantly in unfavorable conditions. In this paper, we propose a deep learning model based on the artificial neural network to identify and classify objects by iris. The model can produce high accuracy in disadvantaged conditions.

Keywords: *Deep learning, Deep learning model, Iris, Iris recognition, Neural network*

Ngày nhận bài: 13/3/2018; Ngày phản biện: 06/4/2018; Ngày duyệt đăng: 31/5/2018

* Tel: 01695 314806, Email: phungthutrang.sfl@tnu.edu.vn