

NHẬN DẠNG CẢM XÚC KHUÔN MẶT SỬ DỤNG CÔNG NGHỆ HỌC SÂU

Nguyễn Thị Thu Hoà¹, Ngô Hữu Huy^{2,*}, Giáp Mạnh Tuyên²,
Nông Văn Dương², Nguyễn Thị Kiều Oanh³, Nguyễn Bá Bằng⁴

¹Trường Cao đẳng Lào Cai, Việt Nam

²Trường Đại học Công nghệ thông tin và Truyền thông, Đại học Thái Nguyên, Việt Nam

³Trường Trung học phổ thông Nguyễn Huệ, Quỳnh Phụ, Thái Bình, Việt Nam

⁴Trường Trung học phổ thông Số 1, Bảo Thắng, Lào Cai, Việt Nam

ARTICLE INFORMATION TÓM TẮT

Journal: Vinh University
Journal of Science
Natural Science, Engineering
and Technology
p-ISSN: 3030-4563
e-ISSN: 3030-4180

Volume: 53

Issue: 3A

***Correspondence:**
nhuhuy@ictu.edu.vn

Received: 09 April 2024

Accepted: 28 June 2024

Published: 20 September 2024

Citation:
Nguyen Thi Thu Hoa, Ngo Huu
Huy, Giap Manh Tuan, Nong
Van Duong, Nguyen Thi Kieu
Oanh, Nguyen Ba Bang (2024).

Facial emotion recognition
using deep learning technology

Vinh Uni. J. Sci.

Vol. 53 (3A), pp. 23-31

doi: 10.56824/vujs.2024a049a

OPEN ACCESS

Copyright © 2024. This is an
Open Access article distributed
under the terms of the [Creative
Commons Attribution License
\(CC BY NC\)](#), which permits non-
commercially to share (copy and
redistribute the material in any
medium) or adapt (remix,
transform, and build upon the
material), provided the original
work is properly cited.

Bài toán nhận dạng cảm xúc trên mặt người luôn nhận được sự quan tâm và thu hút tới từ các nhà khoa học. Cùng với sự phát triển của trí tuệ nhân tạo, các camera nhận dạng cảm xúc khuôn mặt đang càng ngày càng được đưa vào nhiều lĩnh vực khác nhau như y tế, giáo dục, thương mại. Nhờ sự hỗ trợ đặc lực của camera tích hợp với công nghệ trí tuệ nhân tạo, các lĩnh vực trên đang ngày một phát triển mạnh mẽ. Do đó, nghiên cứu này sẽ trình bày một mô hình nhận dạng cảm xúc khuôn mặt hiệu quả và đơn giản dựa trên công nghệ học sâu. Việc nhận dạng cảm xúc khuôn mặt được thực hiện một cách tự động và chính xác, qua đó có thể rút ngắn thời gian thực hiện và tăng hiệu suất công việc. Trong nghiên cứu này, mô hình VGG-Face được huấn luyện bằng tập dữ liệu FER-2013 và ứng dụng trong bài toán đánh giá chất lượng phục vụ khách hàng. Dữ liệu đầu vào là video có khuôn mặt con người để thực hiện nhận dạng cảm xúc và nghiên cứu này thực hiện nhận dạng 7 trạng thái biểu cảm khác nhau. Các kết quả thực nghiệm cho thấy tính hiệu quả của mô hình này.

Từ khóa: DeepFace; nhận dạng cảm xúc khuôn mặt; MTCNN; FER-2013; VGG-Face.

1. Giới thiệu

Ngày nay, việc ứng dụng trí tuệ nhân tạo rộng rãi đã góp phần lớn giúp đơn giản hóa công việc, giảm chi phí cũng như gia tăng hiệu suất lao động. Có rất nhiều công nghệ tích hợp với trí tuệ nhân tạo đã và đang được nghiên cứu, phát triển như nhận dạng chữ, nhận dạng hành vi con người, v.v. Một trong số đó phải kể đến công nghệ đang được nghiên cứu và phát triển mạnh mẽ gần đây là ứng dụng trí tuệ nhân tạo trong việc nhận dạng cảm xúc khuôn mặt của con người. Trên thực tế, trong các cách biểu đạt ngôn ngữ cơ thể, thì cảm xúc là một cách thể hiện rõ nhất những gì mà người nghe cảm thấy, thông qua những câu từ được thể hiện qua cảm xúc trên gương mặt. Do đó, các nhà nghiên cứu về lĩnh vực thị giác máy tính đã thực hiện nghiên cứu sâu rộng cho phép máy tính có thể nhận dạng cảm xúc trên khuôn mặt con người, qua đó sự tương tác giữa người và máy ngày càng chặt chẽ hơn [1].



Hình 1: Nhận dạng cảm xúc khuôn mặt khách hàng

Bài toán nhận dạng khuôn mặt vẫn gặp rất nhiều thách thức, ví dụ như hệ thống camera công cộng, ảnh chụp hình không chính diện hay chất lượng ảnh không tốt, v.v. Những yếu tố này ảnh hưởng trực tiếp tới các thuật toán nhận dạng khuôn mặt. Do đó, nhiều mô hình nhận dạng khuôn mặt đã được phát triển (VGG-Face [2], FaceNet [3], OpenFace [4]...), trong đó mô hình DeepFace [5] đã đưa ra một cấu trúc mạng nơ-ron nhận dạng khuôn mặt có độ chính xác cao, hoạt động ổn định và tính thích nghi cao. Qua đó, công nghệ nhận dạng cảm xúc khuôn mặt trở thành một công cụ mới trong việc đánh giá mức độ hài lòng của khách hàng với các dịch vụ (Hình 1).

Được nghiên cứu và sử dụng rộng rãi trên các lĩnh vực khác nhau như y tế, giáo dục, thương mại, DeepFace có những điểm vượt trội hơn đáng kể so với những công nghệ khác, đó là khả năng nhận diện khuôn mặt với độ chính xác 97,53% với các khuôn mặt được gắn nhãn trong tập dữ liệu Wild (LFW), có khả năng xử lý khuôn mặt ở các góc độ, điều kiện ánh sáng khác nhau [5]-[6]. Mô hình này được phát triển với kiến trúc mạng nơ-ron sâu (DNN) hiệu quả. Phương pháp học tận dụng tập dữ liệu khuôn mặt được gắn nhãn rất lớn để có được biểu diễn khuôn mặt, từ đó có thể khái quát hóa tốt cho các tập dữ liệu khác. Hệ thống căn chỉnh khuôn mặt hiệu quả dựa trên mô hình khuôn mặt 3D rõ ràng.

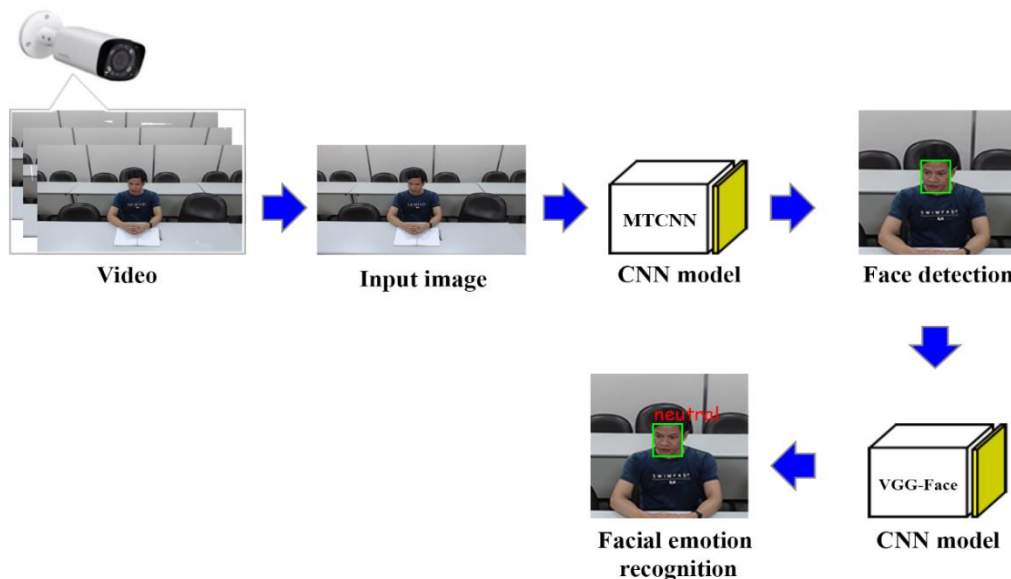
Nhận dạng cảm xúc trên khuôn mặt (Facial Emotion Recognition) là một lĩnh vực nghiên cứu thuộc trí tuệ nhân tạo (AI) sử dụng các mô hình học sâu để tự động phân tích biểu cảm khuôn mặt và xác định cảm xúc của con người. Mô hình học sâu đóng vai trò quan trọng trong nhận dạng cảm xúc trên khuôn mặt, giúp nâng cao độ chính xác và hiệu quả của việc nhận dạng cảm xúc so với các phương pháp truyền thống. Nhóm tác giả Khaireddin đã nghiên cứu và xác định 6 trạng thái cảm xúc bằng công nghệ VGGNet [7]. Bên cạnh đó, các phương pháp nhận dạng cảm xúc khuôn mặt sử dụng mô hình mạng nơ-ron cũng được quan tâm bởi nhiều nhà khoa học [8]-[13]. Phương pháp nhận dạng này được sử dụng rất nhiều bởi có độ chính xác cao.

Bài báo này trình bày mô hình nhận dạng cảm xúc khuôn mặt trên mạng nơ-ron sử dụng công nghệ học sâu một cách đơn giản và hiệu quả. Mô hình này có thể được ứng dụng trong bài toán đánh giá chất lượng phục vụ khách hàng dựa trên nhận dạng cảm xúc khuôn mặt. Trong đó, mô hình này có thể phát hiện khuôn mặt một cách tự động, chính xác và nhận dạng 7 trạng thái biểu cảm, bao gồm: tức giận, chán ghét, sợ hãi, hạnh phúc, buồn bã, ngạc nhiên, trung tính.

2. Mô hình nhận dạng cảm xúc khuôn mặt

2.1. Cấu trúc hệ thống

Tổng quan của mô hình nhận dạng khuôn mặt được thể hiện như Hình 2. Dữ liệu đầu vào là video các khuôn mặt để thực hiện nhận dạng. Các thuật toán học máy và trí tuệ nhân tạo sẽ được áp dụng để nhận dạng các biểu hiện khuôn mặt liên quan đến cảm xúc, chẳng hạn như mỉm cười, biểu lộ sự bất mãn hoặc hứng thú. Cơ sở dữ liệu huấn luyện được sử dụng là bộ dữ liệu FER-2013 (Facial Expression Recognition 2013) [14]. Đây là tập dữ liệu được sử dụng nhiều trong lĩnh vực nhận diện cảm xúc khuôn mặt. Đầu tiên, hệ thống thực hiện trích xuất các khung ảnh liên tiếp từ video đầu vào để thực hiện xử lý trên các khung ảnh riêng biệt đó. Tiếp đến, mô hình mạng nơ-ron tích chập xếp tầng đa nhiệm (MTCNN) được sử dụng để phát hiện khuôn mặt. Sau khi phát hiện được khuôn mặt, mô hình VGG-Face được sử dụng để nhận dạng cảm xúc khuôn mặt. Cuối cùng, hệ thống đưa ra kết quả thống kê cảm xúc của khách hàng.



Hình 2: Tổng quan mô hình nhận dạng cảm xúc khuôn mặt

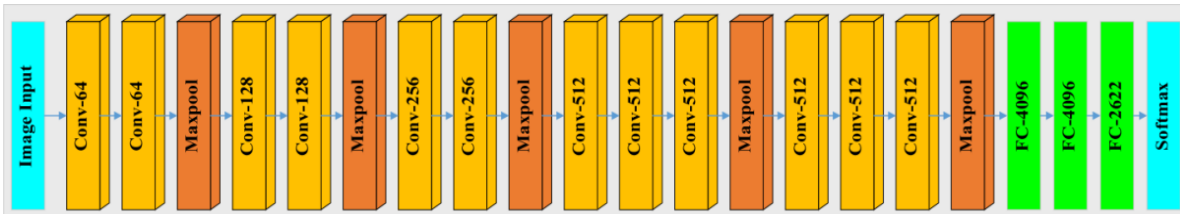
2.2. Mô hình mạng nơ-ron tích chập xếp tầng đa nhiệm

Mô hình mạng nơ-ron tích chập xếp tầng đa nhiệm (MTCNN) [15] được thiết kế nhằm thực hiện phát hiện khuôn mặt trong ảnh. Ý tưởng của mô hình này là sử dụng phương pháp tiếp cận tỷ lệ bằng mạng nơ-ron tích chập (CNN) gồm 3 giai đoạn là P-Net, R-Net và O-Net. Giai đoạn đầu tiên (P-Net) lấy đầu vào là một kim tự tháp hình ảnh được tạo thành từ các bản sao có tỷ lệ khác nhau của hình ảnh đầu vào. Điều này cung cấp cho mô hình một loạt các kích thước cửa sổ để lựa chọn và giúp mô hình luôn bất biến về tỷ lệ. Giai đoạn thứ hai là mạng lưới lọc CNN (R-Net). R-Net khi nhận được các đề xuất thì sẽ tiến hành thực hiện các bước tương tự như P-Net, tuy nhiên mạng còn sử dụng thêm phương pháp đệm (Padding) thực hiện chèn thêm các pixel 0 vào các phần còn thiếu của hộp giới hạn nếu hộp giới hạn bị vượt quá biên giới của hình ảnh. Đầu ra của R-Net cũng tương tự như P-Net, nó gồm tọa độ của hộp giới hạn mới, chính xác hơn. R-Net tiếp tục giảm số lượng và hợp nhất các hộp giới hạn chồng chéo bằng cách sử dụng thuật toán Non-Maximum Suppression (NMS). Cuối cùng, mạng đầu ra trong giai đoạn thứ ba (O-Net) thực hiện tương tự như R-

Net và O-Net bổ sung thêm việc phát hiện 5 điểm mốc trên khuôn mặt bao gồm 2 điểm mắt, 1 điểm mũi, 2 điểm góc miệng trong hộp giới hạn cuối cùng chứa khuôn mặt được phát hiện.

2.3. Mô hình VGG-Face

Mô hình VGG-Face [2] trên nền tảng DeepFace [16] được sử dụng để xây dựng hệ thống và được huấn luyện trên bộ dữ liệu FER-2013. Mô hình VGG-Face được thiết kế dựa trên kiến trúc VGG16, một kiến trúc mạng nơ-ron sâu được phát triển bởi nhóm nghiên cứu Visual Geometry Group (VGG) tại Đại học Oxford. Cấu trúc mô hình VGG-Face được thể hiện ở Hình 3. Mô hình này sử dụng một kiến trúc mạng nơ-ron tích chập gồm 16 lớp chồng chéo, trong đó có 13 lớp tích chập và 3 lớp kết nối đầy đủ (Fully connected layers). Mỗi lớp tích chập sẽ sử dụng các bộ lọc nhỏ tìm một đặc trưng khác nhau của khuôn mặt từ hình ảnh. Lớp kết nối đầy đủ cuối cùng sẽ thực hiện nhận dạng dựa trên các đặc trưng đã học từ các lớp trước đó.



Hình 3: Cấu trúc mô hình VGG-Face

3. Kết quả thực nghiệm

3.1. Kết quả đào tạo mô hình

Thiết bị phần cứng: Nghiên cứu này được triển khai trên máy tính có CPU core i7 3.70GHz, hệ điều hành Windows 64bit, RAM 32GB, GPU (NVIDIA TITAN V).

Phần mềm: Nền tảng Anaconda được sử dụng để phát triển chương trình. Đây là một nền tảng mã nguồn mở được phát triển để hỗ trợ việc quản lý môi trường lập trình, phát triển ứng dụng và thực hiện các dự án khoa học dữ liệu.

Cơ sở dữ liệu huấn luyện: Việc thu thập cơ sở dữ liệu huấn luyện rất quan trọng vì nó ảnh hưởng trực tiếp đến chất lượng hệ thống và kết quả đầu ra. Do đó, tập dữ liệu FER-2013 được chọn làm cơ sở dữ liệu huấn luyện. Trong tập dữ liệu bao gồm 35.887 hình ảnh khuôn mặt và mỗi hình ảnh có kích thước 48x48 pixels. FER-2013 chứa hình ảnh khuôn mặt con người với 7 lớp trạng thái cảm xúc khác nhau gồm: tức giận, chán ghét, sợ hãi, hạnh phúc, buồn bã, ngạc nhiên, trung tính. Tập dữ liệu này được chia làm 3 phần là tập huấn luyện (Training Set), tập kiểm thử (Testing Set) và cuối cùng là tập xác thực (Validation Set). Việc chia dữ liệu như vậy giúp đảm bảo mô hình không chỉ học từ tập huấn luyện mà còn có khả năng tổng quát hóa trên dữ liệu mới.



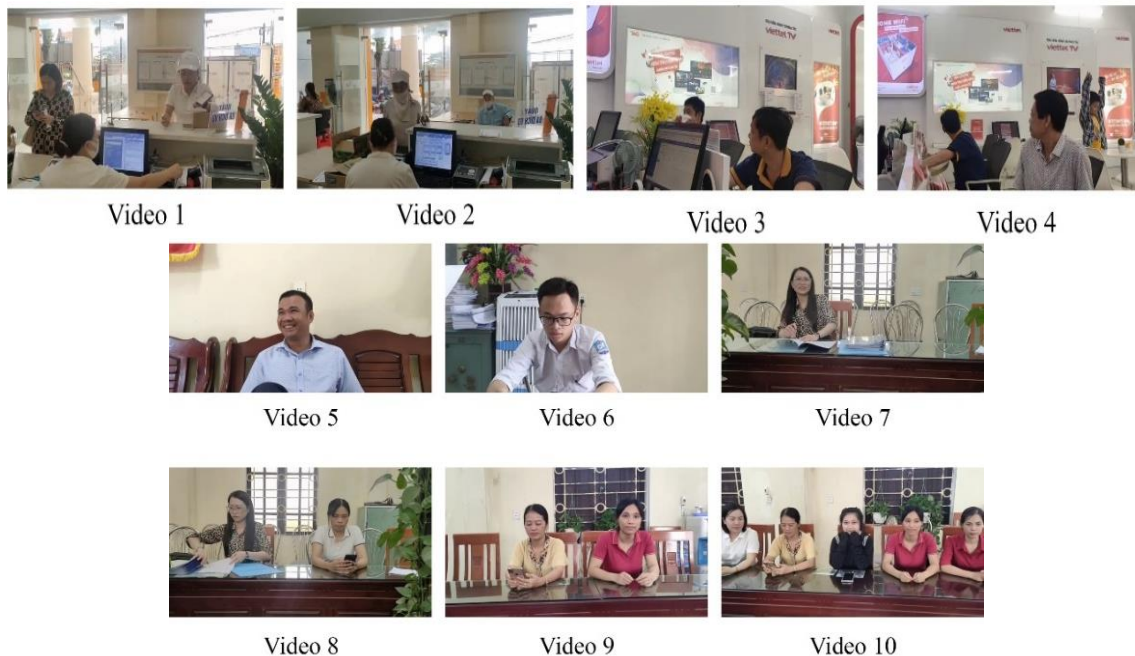
Hình 4: Ví dụ một số tập dữ liệu trong FER-2013

3.2. Đánh giá kết quả trong ứng dụng thực tế

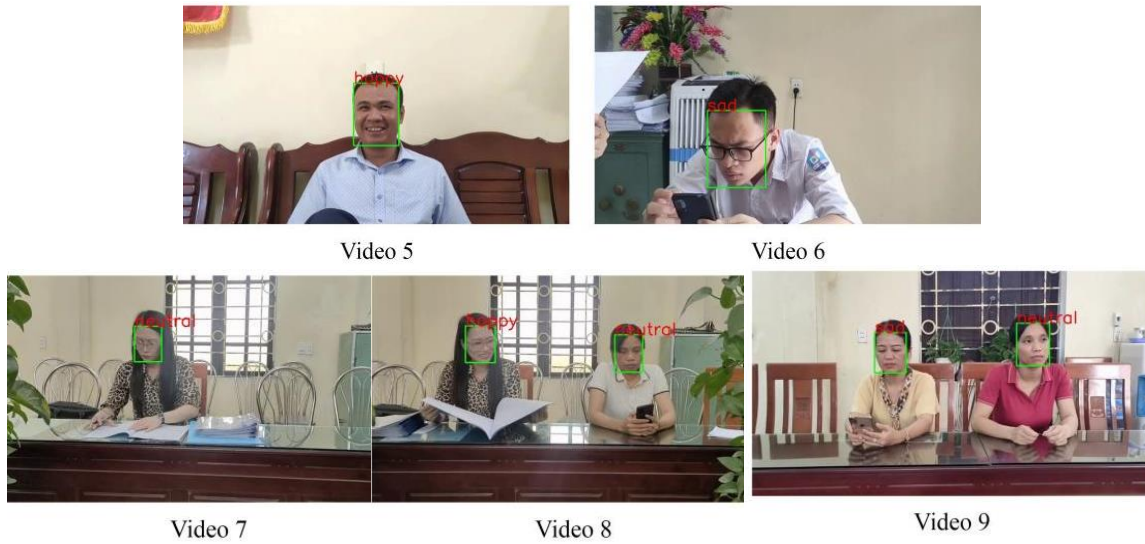
Kết quả của hệ thống nhận dạng cảm xúc được đánh giá trên cơ sở dữ liệu đã thu thập. Tập dữ liệu này được xây dựng bằng việc tách các frame (khung hình) từ các video. Các video này được quay tại một số địa điểm, cửa hàng như: Bưu điện, cửa hàng Viettel... Tổng số video đã thu thập là 10 và tương ứng với 2.000 ảnh phục vụ cho việc kiểm tra và đánh giá mô hình.

Với các video chỉ có 1-2 người trên một khung hình và khuôn mặt nhìn thẳng như các video 5-9 trong Hình 5 thì việc nhận dạng khuôn mặt sẽ dễ dàng thực hiện hơn. Hình 6 hiển thị ảnh minh họa kết quả nhận dạng cảm xúc khuôn mặt trên các video 5-9. Trong đó các cảm xúc hạnh phúc (Happy) được thể hiện ở video 5 và 8. Đây là một trạng thái cảm xúc tích cực của khách hàng về dịch vụ. Ngược lại, cảm xúc buồn bã (Sad) được xác định ở video 6 và 8, trong cả hai video này khách hàng đều tập trung vào điện thoại và cảm xúc là không tốt. Ở các video 7, 8 và 9 thì cảm xúc được xác định là trung tính (Neutral). Đây là cảm xúc phổ biến nhất được xác định ở khách hàng.

Trong trường hợp, các video có nhiều người trên 1 khung hình và các khuôn mặt được theo dõi ở các góc khác nhau như video 1-4, 10 trong Hình 5. Việc nhận dạng cảm xúc trên các video này sẽ khó khăn hơn. Hình 7 thể hiện ảnh minh họa kết quả nhận dạng cảm xúc khuôn mặt trên các video 1-4 và 10. Trong video 1 và 2, các khuôn mặt có kích thước nhỏ do ở xa camera hay khách hàng đeo khẩu trang. Tuy nhiên, hệ thống vẫn có thể phát hiện khuôn mặt và nhận dạng cảm xúc bao gồm: buồn bã (Sad) và trung tính (Neutral). Các video 3,4 và 10 gồm các nhóm khách hàng và được nhận dạng cảm xúc một cách chính xác bao gồm: hạnh phúc (Happy), buồn bã (Sad) và trung tính (Neutral). Việc nhận dạng cảm xúc gặp một khó khăn do một số yếu tố như: góc nhìn nghiêng, cúi mặt xuống, hoặc đeo khẩu trang...



Hình 5: Ảnh chụp từ các video trong cơ sở dữ liệu

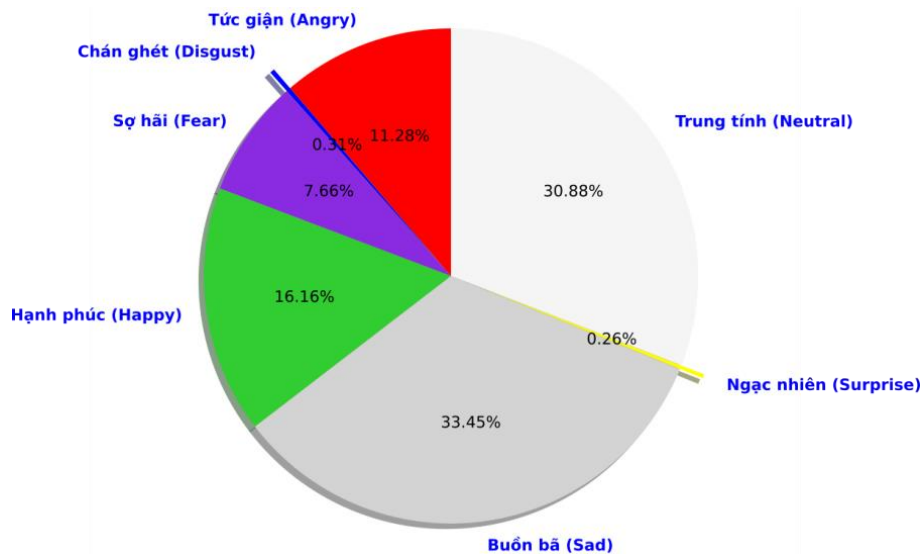


Hình 6: Kết quả nhận dạng cảm xúc trong video 5-9

Thực hiện thống kê và phân tích kết quả trên 10 video kiểm thử, tương ứng với 2.000 khung ảnh được trích xuất. Hình 8 thống kê lại kết quả nhận dạng cảm xúc khuôn mặt đã thu được. Qua biểu đồ, chúng ta có thể đánh giá phần lớn cảm xúc khách hàng là trung tính. Cảm xúc tích cực (Hạnh phúc) chiếm tỉ lệ thấp hơn so với cảm xúc tiêu cực (Buồn bã, sợ hãi, giận dữ).



Hình 7: Kết quả nhận dạng cảm xúc trong các video từ 1-4 và 10



Hình 8: Thống kê kết quả nhận dạng cảm xúc khuôn mặt

4. Kết luận và định hướng phát triển nghiên cứu

Nghiên cứu này đã trình bày một mô hình nhận dạng cảm xúc khuôn mặt một cách đơn giản và hiệu quả, sử dụng công nghệ học sâu. Hệ thống này sẽ thực hiện nhận dạng cảm xúc khuôn mặt một cách tự động và chính xác, qua đó có thể rút ngắn thời gian tìm hiểu và tăng hiệu suất công việc. Trong đó, mô hình VGG-Face được huấn luyện bằng tập dữ liệu FER-2013 và sử dụng để phát triển hệ thống. Các kết quả thực nghiệm cho thấy tính hiệu quả của mô hình này.

Trong các nghiên cứu tiếp theo, cần tiếp tục cải thiện mô hình bằng cách thu thập thêm nhiều dữ liệu về khuôn mặt. Từ đó, mô hình này có thể gia tăng độ chính xác khi nhận dạng cảm xúc khuôn mặt. Qua đó, có thể sử dụng tích hợp mô hình này với camera có chất lượng ảnh tốt để có góc quan sát tốt hơn. Kết hợp nhận dạng cảm xúc với một số bài toán khác như nhận dạng các hành động, giới tính, dự đoán độ tuổi để hệ thống có thêm nhiều tính năng và hoàn thiện hơn trong việc hỗ trợ phân tích và đánh giá chất lượng phục vụ khách hàng một cách thông minh và tự động.

Lời cảm ơn: Nghiên cứu này là sản phẩm của đề tài KH&CN cấp Cơ sở có mã số T2024-07-03, được tài trợ bởi kinh phí của Trường Đại học Công nghệ Thông tin và Truyền thông, Đại học Thái Nguyên.

TÀI LIỆU THAM KHẢO

- [1] H. V. Dung, *Image Recognition and Processing, Computer Vision, Science and Technics Publisher*, 2018.
- [2] O. M. Parkhi, A. Vedaldi and A. Zisserman, "Deep Face Recognition," in *Proc. British Machine Vision Conference (BMVC)*, Swansea, United Kingdom, pp. 1-12, 2015. DOI: 10.5244/C.29.41

- [3] F. Schroff, D. Kalenichenko and J. Philbin, “FaceNet: A Unified Embedding for Face Recognition and Clustering,” in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, Boston, Massachusetts, USA, pp. 815-823, 2015. DOI: 10.1109/CVPR.2015.7298682
- [4] T. Baltrušaitis, P. Robinson and L. P. Morency, “OpenFace: An Open Source Facial Behavior Analysis Toolkit,” in *Proc. IEEE Winter Conference on Applications of Computer Vision (WACV)*, Lake Placid, NY, USA, pp. 1-10, 2016. DOI: 10.1109/WACV.2016.7477553
- [5] Y. Taigman, M. Yang, M. Ranzato and L. Wolf, “DeepFace: Closing the Gap to Human-Level Performance in Face Verification,” in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, Ohio, USA, pp. 1701-1708, 2014. DOI: 10.1109/CVPR.2014.220
- [6] S. I. Serengil and A. Ozpinar, “LightFace: A Hybrid Deep Face Recognition Framework,” in *Proc. Innovations in Intelligent Systems and Applications Conference (ASYU)*, Istanbul, Turkey, pp. 1-5, 2020. DOI: 10.1109/ASYU50717.2020.9259802
- [7] Y. Khairuddin and Z. Chen, “Facial Emotion Recognition: State of the Art Performance on FER2013,” arXiv, 2021.
- [8] M. A. H. Akhand, S. Roy, N. Siddique, Md A. S. Kamal and T. Shimamura, “Facial Emotion Recognition Using Transfer Learning in the Deep CNN,” *Electronics*, vol. 10, no. 9, p. 1036, 2021. DOI: 10.3390/electronics10091036
- [9] E. Pranav, S. Kamal, C. Satheesh Chandran and M. H. Supriya, “Facial Emotion Recognition Using Deep Convolutional Neural Network,” in *Proc. 6th International Conference on Advanced Computing and Communication Systems (ICACCS)*, Coimbatore, India, pp. 317-320, 2020. DOI: 10.1109/ICACCS48705.2020.9074302
- [10] I. Lasri, A. R. Solh and M. E. Belkacemi, “Facial Emotion Recognition of Students using Convolutional Neural Network,” in *Proc. Third International Conference on Intelligent Computing in Data Sciences (ICDS)*, Marrakech, Morocco, pp. 1-6, 2019. DOI: 10.1109/ICDS47004.2019.8942386
- [11] C. K. Lee, S. K. Yoo, Y. Park, N. Kim, K. Jeong and B. Lee, “Using Neural Network to Recognize Human Emotions from Heart Rate Variability and Skin Resistance,” in *Proc. Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Shanghai, China, pp. 5523-5525, 2006.
- [12] K. Dai, H. Fell and J. Macauslan, “Recognizing Emotion in Speech Using Neural Networks,” *Telehealth and Assistive Technologies*, vol. 31, pp. 38-43, 2008.
- [13] M. Wang and W. Deng, “Deep Face Recognition: A Survey,” *Neurocomputing*, vol. 429, pp. 215-244, 2021. DOI: 10.1016/j.neucom.2020.10.081
- [14] Manas sambare, “FER-2013, Learn facial expressions from an image,” *Kaggle*, 2020. Available: <https://www.kaggle.com/datasets/msambare/fer2013>. Accessed 26/3/2024.

- [15] K. Zhang, Z. Zhang, Z. Li and Y. Qiao, "Joint Face Detection and Alignment Using Multitask Cascaded Convolutional Networks," *IEEE Signal Processing Letters*, vol. 23, no. 10, pp. 1499-1503, 2016. DOI: 10.1109/LSP.2016.2603342
- [16] S. I. Serengil, "DeepFace," *GitHub*. Available: <https://github.com/serengil/deepface>. Accessed 26/3/2024.

ABSTRACT

FACIAL EMOTION RECOGNITION USING DEEP LEARNING TECHNOLOGY

Nguyen Thi Thu Hoa¹, Ngo Huu Huy², Giap Manh Tuyen²,
Nong Van Duong², Nguyen Thi Kieu Oanh³, Nguyen Ba Bang⁴

¹Lao Cai College, Vietnam

²Thai Nguyen University of Information and Communication Technology, Thai Nguyen, Vietnam

³Nguyen Hue High School, Quynh Phu District, Thai Binh, Vietnam

⁴High school No. 1, Bao Thang District, Lao Cai, Vietnam

Received on 09/4/2024, accepted for publication on 28/6/2024

Recognizing emotions on human faces has always received attention and attraction from researchers. Along with the development of artificial intelligence, facial emotion recognition cameras are increasingly being introduced into various fields such as healthcare, education, and commerce. Thanks to the effective support of cameras integrated with artificial intelligence technology, the above fields are growing strongly. Therefore, this study will present an effective and simple facial emotion recognition model based on deep learning technology. Facial emotion recognition is performed automatically and accurately, thereby shortening implementation time and increasing work efficiency. In this study, the VGG-Face model is trained using the FER-2013 data set and applied to evaluate customer service quality. The input data are videos of human faces to perform emotion recognition and 7 different facial expressions have been performed recognition. Experimental results have shown the effectiveness of the model.

Keywords: DeepFace; facial emotion recognition; FER-2013; MTCNN; VGG-Face.