

MỘT MÔ HÌNH DEEP LEARNING NHẸ CHO BÀI TOÁN NHẬN DẠNG TUỔI VÀ GIỚI TÍNH SỬ DỤNG MẠNG CNN

Phùng Thị Thu Trang^{1*}, Ma Thị Hồng Thu²

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

TÓM TẮT

Bài toán nhận dạng tuổi và giới tính đang thu hút được nhiều sự chú ý từ các nhà nghiên cứu đặc biệt là khi mạng xã hội và mạng truyền thông ngày càng phổ biến. Các phương pháp được công bố gần đây cho kết quả khá tốt về độ chính xác nhưng còn tỏ ra kém hiệu quả trong vấn đề nhận diện thời gian thực bởi vì các mô hình này được thiết kế quá phức tạp. Trong bài báo này, chúng tôi đề xuất một mô hình nhẹ mang tên lightweight CNN thực hiện song song 2 nhiệm vụ là phân lớp tuổi và giới tính. Về độ chính xác trong nhận diện tuổi thì lightweight CNN tốt hơn 5.1% so với mô hình tốt nhất đã được công bố gần đây. Về thời gian chạy và số lượng tham số được sử dụng thì lightweight CNN sử dụng ít hơn nhiều so với các mô hình khác trên bộ dữ liệu Adience, đáp ứng được yêu cầu về nhận dạng trong thời gian thực.

Từ khóa: Học sâu, Mạng CNN, Phân lớp tuổi, phân lớp giới tính, Mạng nơ-ron

Ngày nhận bài: 09/4/2019; Ngày hoàn thiện: 26/4/2019; Ngày duyệt đăng: 07/5/2019

A LIGHTWEIGHT DEEP LEARNING MODEL FOR AGE AND GENDER IDENTITY PROBLEM USING THE CNN NETWORK

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

¹School of Foreign Language – TNU, ²Tan Trao University

ABSTRACT

Age and gender identification problems are gaining a lot of attention from researchers since social and multimedia networks are becoming more popular nowadays. Recently published methods have yielded quite good results in terms of accuracy but also proved ineffective in real-time identification because these models were designed too complicated. In this paper, we propose a lightweight model called lightweight CNN that performs parallel tasks of age and gender classification. In terms of accuracy in identifying age, lightweight CNN is 5.1% better than the best model recently published. About runtime and the number of parameters used, lightweight CNN uses much less than other models on the Adience dataset, meet the identification requirements in real time.

Keywords: Deep learning, CNN Network, Age Classification, Gender Classification, Neural Network

Received: 09/4/2019; Revised: 26/4/2019; Approved: 07/5/2019

* Corresponding author: Tel: 0395 314806, Email: phungthutrang.sfl@tnu.edu.vn

1. Giới thiệu

Xử lý ảnh và thị giác máy tính đang là những lĩnh vực được quan tâm nhiều nhất trong trí tuệ nhân tạo với nhiều bài toán thực tế. Bên cạnh đó, sự phát triển vượt bậc của các thuật toán học sâu đặc biệt là mạng lưới thần kinh tích chập (convolutional neural network – CNN) đã cho những kết quả vượt bậc trong các bài toán điển hình. Ví dụ Alex cùng các cộng sự [1] đã đề xuất một mô hình sử dụng mạng CNN và giành chiến thắng trong cuộc thi ImageNet với tỷ lệ lỗi đạt 15.3% vào năm 2012. Đây là cuộc thi có quy mô lớn nhất thế giới về bài toán nhận diện đối tượng trong ảnh. Năm 2013, Zeiler và Fergus [2] đã đề xuất một mô hình có tên ZFNet và giảm lỗi từ 15,3% xuống còn 14,8%. GoogleNet (Inception) và VGGNet đã được đề xuất năm 2014 [3] với tỷ lệ lỗi lần lượt là 6,67% và 7,32%. Năm 2015, Kaiming He [4] đã đề xuất kiến trúc mạng ResNet và đạt tỷ lệ lỗi 3,57%, tỷ lệ lỗi này còn tốt hơn cả hiệu suất của con người. Ngoài trừ bài toán nhận diện đối tượng trong ảnh, CNN thường được áp dụng cho nhiều bài toán khác như: Phát hiện đa đối tượng trong ảnh, đặt tiêu đề cho ảnh, phân đoạn ảnh,... Thậm chí, Yoo Kim [5] đã áp dụng mạng CNN cho bài toán phân lớp câu và đạt hiệu quả cao trong nhiều bộ cơ sở dữ liệu về văn bản khác nhau.

Khuôn mặt là một đối tượng trong cơ thể con người và hình ảnh khuôn mặt mang rất nhiều thông tin quan trọng như: tuổi tác, giới tính, trạng thái cảm xúc, dân tộc,... Trong đó, việc xác định tuổi tác và giới tính là hết sức quan trọng, đặc biệt trong giao tiếp, chúng ta cần sử dụng những từ ngữ phù hợp với giới tính của người nghe ví dụ trong tiếng Việt chúng ta có: anh/chị, chú/cô... Hay với nhiều ngôn ngữ khác nhau trên thế giới, chẳng hạn như tiếng Việt thì lời chào hỏi dành cho người lớn tuổi khác với người trẻ tuổi. Do đó, việc xác định tuổi và giới tính dựa trên khuôn mặt là một bài toán hết sức quan trọng, có ý nghĩa thực tế to lớn.

Bài toán ước lượng tuổi và giới tính đã được quan tâm nhiều trong suốt 20 năm gần đây, đã có rất nhiều các công trình được công bố với nhiều kỹ thuật khác nhau chẳng hạn như: AGing pattErn Subspace (AGES), Gaussian Mixture Models (GMM), Hidden-Markov-Model (HMM), Support Vector Machines (SVM), ... Từ khi các mô hình học sâu được áp dụng cho bài toán này đã cải thiện đáng kể kết quả về mặt hiệu suất cũng như tốc độ. Độ chính xác của mô hình khi ước lượng tuổi đạt 62,8% và đối với giới tính đạt 92,6% [6].

Tuy nhiên, để đạt được hiệu suất cao thì các mô hình thường được xây dựng càng phức tạp với số lượng tham số lớn (từ 10 triệu đến hơn 100 triệu tham số), do đó gây khó khăn trong vấn đề nhận dạng trong thời gian thực. Trong bài báo này, chúng tôi đề xuất một mô hình nhẹ sử dụng CNN với khoảng 1 triệu tham số nhưng đạt kết quả nhận diện tuổi lên đến 67,9% và nhận diện giới tính lên đến 88,8%. Với số lượng tham số nhỏ này thì mô hình của chúng tôi hoàn toàn có thể chạy được trên các thiết bị nhúng và thiết bị di động một cách dễ dàng đảm bảo vấn đề thời gian thực. Sự đóng góp của chúng tôi trong bài báo này là: (1) Xây dựng một mô hình nhẹ để giải quyết bài toán đa nhiệm vụ (dự đoán tuổi và giới tính từ ảnh chụp khuôn mặt). (2) Từ kết quả của mô hình cho thấy rằng thuật toán không chỉ tốt về mặt hiệu suất mà còn giảm thiểu số lượng tham số được sử dụng từ đó giúp cải thiện tốc độ của mô hình và đáp ứng được yêu cầu về nhận diện trong thời gian thực.

2. Các nghiên cứu gần đây

Như đã được đề cập ở phần Giới thiệu, bài toán ước lượng tuổi và giới tính đã được nghiên cứu từ rất lâu. Nhưng hầu như chúng chỉ được nghiên cứu tách rời nhau. Các mô hình được xây dựng riêng biệt cho từng nhiệm vụ. Cho đến năm 2016, Linnan Zhu cùng các cộng sự [7] đề xuất một mô hình đa nhiệm vụ giải quyết cả hai bài toán cùng một lúc. Trước tiên chúng ta xét lần lượt các bài toán để có cái nhìn tổng quan.

2.1 Bài toán phân lớp tuổi: Nhiệm vụ của bài toán này là đưa ra ước lượng tuổi của một người từ bức ảnh chụp khuôn mặt của họ. Bài toán này được giới thiệu lần đầu tiên bởi Kwon và Lobo [8] trong đó, họ sử dụng phương pháp phát hiện và tính toán tỷ lệ của các nếp nhăn trên khuôn mặt để có thể dự đoán độ tuổi và sau đó nó được cải tiến bởi Ramanathan và Chellappa [9]. Tuy nhiên, phương pháp này có thể phân biệt được độ tuổi giữa người lớn và trẻ em, nhưng rất khó có thể phân biệt được độ tuổi giữa những người lớn với nhau. Một cách tiếp cận khác do Geng cùng các cộng sự [10] trình bày là sử dụng AGES cho hiệu quả cao hơn nhưng thuật toán này cần một lượng lớn hình ảnh khuôn mặt của từng người và đặc biệt hình ảnh đầu vào này cần phải ở chính giữa, mặt hướng thẳng và được căn chỉnh đúng kích thước. Tuy nhiên, trên thực tế thì các bức ảnh chụp lại rất ít khi thỏa mãn điều kiện như vậy do đó cách tiếp cận này không được phù hợp với nhiều ứng dụng thực tế.

Một cách tiếp cận khác dựa trên các thuật toán thống kê đã được sử dụng như GMM [11] và HMM, super-vectors [12] được sử dụng để làm đại diện cho từng phần của khuôn mặt. Trong thập kỷ qua, khi các thuật toán học máy dần được cải tiến và đạt được thành tựu to lớn đặc biệt là học sâu, thì một loạt các công trình nghiên cứu về phân lớp tuổi được công bố cho kết quả khả quan, có thể kể đến như: Eidinger cùng các cộng sự [13] đã sử dụng SVM kết hợp với dropout cho bài toán nhận diện tuổi và nhận diện giới tính. Năm 2015, Gil Levi và Tal Hassner [14] đã đưa ra mô hình Deep Neural Network đầu tiên cho bài toán phân lớp tuổi và giới tính. Sau đó, Zhu cùng các cộng sự [7] đã xây dựng một mô hình đa nhiệm vụ cho phép chia sẻ và tìm hiểu các tính năng tối ưu để cải thiện hiệu suất nhận dạng cho cả hai nhiệm vụ. Đây là bài báo đầu tiên áp dụng mô hình tối ưu hóa bài toán nhận diện tuổi và giới tính cùng nhau để thấy được mối quan hệ giữa 2 bài toán.

2.2 Bài toán phân lớp giới tính: Cùng với sự phát triển của bài toán nhận dạng tuổi, bài toán nhận biết giới tính đã được đề xuất và giải quyết từ những năm 1990. Tổng quan về các phương pháp phân lớp giới tính bạn đọc có thể được tìm thấy trong [15]. Sau đây, chúng tôi sẽ tóm tắt một số phương pháp liên quan. Cottrell [16] là người đầu tiên đề xuất mô hình mạng nơron giải quyết bài toán nhận dạng giới tính, tuy nhiên các khuôn mặt đầu vào phải đảm bảo nhiều yêu cầu nhất định, gây ra nhiều hạn chế cho mô hình. Sau đó, Lyons cùng các cộng sự [17] đã sử dụng thuật toán PCA (Principal Component Analysis) và LDA (Linear Discriminant Analysis) để nhận diện ra giới tính. SVM và AdaBoost được sử dụng trong [18] và [19]. Trong [20], Ullah đã sử dụng Bộ mô tả kết cấu cục bộ Webers để nhận dạng giới tính. Hầu hết các phương pháp được thảo luận ở trên đã sử dụng bộ cơ sở dữ liệu FERET để đánh giá hiệu suất của mô hình. Tuy nhiên, các hình ảnh trong bộ dữ liệu FERET được chụp trong điều kiện tốt, hình ảnh các khuôn mặt không bị che phủ, và hướng thẳng. Hơn nữa, kết quả thu được trên bộ dữ liệu này cho thấy nó đã bão hòa và không thách thức đối với các phương pháp hiện đại. Do đó, những năm gần đây bộ cơ sở dữ liệu Adience thường được sử dụng để so sánh kết quả giữa các mô hình. Bởi vì bộ dữ liệu này chứa hình ảnh thách thức hơn so với bộ dữ liệu FERET và được thiết kế để khai thác tốt hơn các thông tin từ các ảnh dữ liệu đào tạo [14]. Cũng tương tự như bài toán phân lớp tuổi, các mô hình như SVM, Deep Neural Network bao gồm AdienceNet [14], CaffeNet, VGG-16, và GoogleNet [6] cũng được áp dụng cho bài toán nhận diện giới tính.

Tuy nhiên, các phương pháp được nêu ở trên đều tồn tại những hạn chế nhất định. Với các phương pháp gần đây sử dụng mạng neural network thì đã khắc phục được những hạn chế đó nhưng số lượng tham số được sử dụng còn rất lớn, gây khó khăn cho vấn đề nhận diện trong thời gian thực và cho các thiết bị nhúng.

3. Đề xuất thuật toán

Trong phần này, chúng tôi sẽ trình bày một mô hình đa tác vụ nhẹ mang tên lightweight CNN để giải quyết bài toán phân lớp tuổi và giới tính. Mô hình của chúng tôi được trình bày thành 3 phần bao gồm: Mạng tích chập nhẹ, kiến trúc mô hình và cuối cùng là huấn luyện và thử nghiệm. Sau đây, chúng tôi sẽ giới thiệu về mạng tích chập nhẹ.

3.1 Mạng tích chập nhẹ: là sử dụng mạng CNN để xây dựng ra mô hình với số lượng tham số ít, nhưng vẫn đảm bảo hiệu quả về mặt hiệu suất. Hay nói cách khác là làm thế nào để xây dựng một mô hình CNN với số lượng tham số ít nhất có thể nhưng lại đạt hiệu quả tốt nhất có thể, đây cũng là thách thức khó khăn nhất được đặt ra đối với các mô hình nhẹ nói chung. Khác với các mô hình như VGG Net hay ResNet sử dụng hơn 40 triệu hoặc thậm chí hơn 100 tham số, các mô hình nhẹ chỉ sử dụng vài triệu hoặc thậm chí chỉ hơn 1 triệu tham số. Ví dụ: với phân loại độ tuổi, mô hình AdienceNet từ [14] đã sử dụng hơn 10 triệu tham số và độ chính xác là 50,7%, mô hình VGG-16 từ [6] đã sử dụng hơn 100 triệu tham số và độ chính xác là 62,8%, nhưng mô hình nhẹ từ [7] chỉ sử dụng 10 triệu tham số và độ chính xác lên tới 46,0%.

3.2 Kiến trúc mô hình: mô hình lightweight CNN được mô tả như trong hình 1. Phần đầu tiên của mô hình, chúng tôi sử dụng mạng CNN để trích chọn ra các đặc trưng từ dữ liệu ảnh đầu vào. Các hoạt động trong tầng CNN này bao gồm: Convolution (Conv) + Batch Normalization (BN) + Rectified Linear Unit (ReLU) + Max Pooling (MaxPool) với kích thước cửa sổ trượt là 2x2, bước nhảy bằng 2 + Drop out (Dropout) với tỷ lệ drop là 0,25. Ở phần sau của mô hình, chúng tôi sử dụng mạng Fully Connected (FC) với tỷ lệ dropout là 0,25.

3.3 Huấn luyện và thử nghiệm: Đầu vào của mô hình là các hình ảnh RGB được thay đổi kích thước xuống còn 64x64, đầu ra của mô hình là vector y bao gồm 2 giá trị tương ứng với

ước lượng tuổi và ước lượng giới tính của người trong ảnh đầu vào. Hàm mất mát của mô hình $\mathcal{L}$ được thiết kế như trong công thức (1).

$$\mathcal{L}(w, b) = \frac{1}{N} \sum_{i=1}^N (y_i^t - \hat{y}_i^t)^2 \quad (1)$$

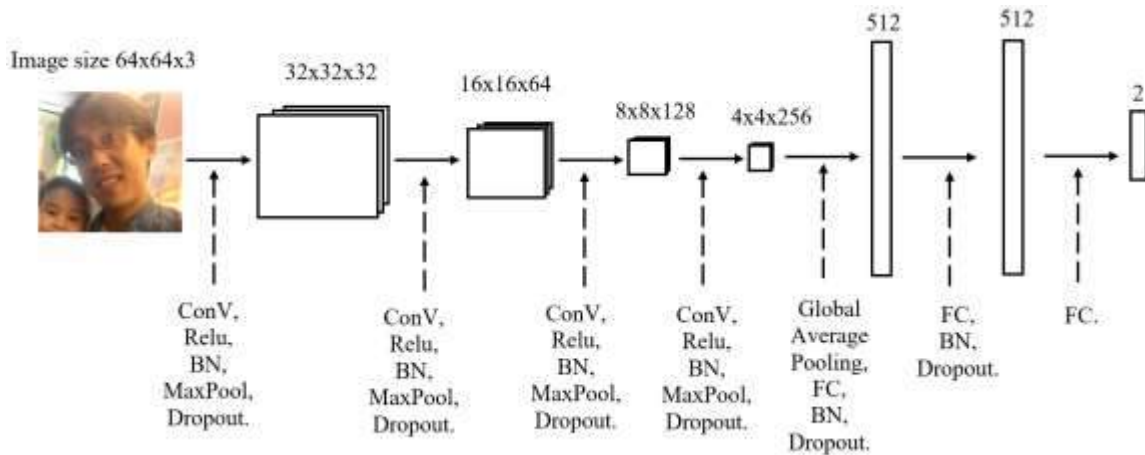
Trong đó, N là số mẫu đưa vào mô hình huấn luyện, T là số lượng nhiệm vụ (với bài toán này $T = 2$). Chúng ta có $\hat{y}$ là kết quả đầu ra của mô hình và y là kết quả thực tế của dữ liệu. Hàm mất mát $\mathcal{L}$ được xây dựng dựa trên công thức MSE và áp dụng cho bài toán đa nhiệm vụ.

4. Thử nghiệm

4.1 Bộ cơ sở dữ liệu Adience: Như đã được đề cập ở mục trước, chúng tôi sử dụng bộ cơ sở dữ liệu Adience từ [21] để tiến hành huấn luyện và đánh giá mô hình. Bộ cơ sở dữ liệu Adience chủ yếu được xây dựng để nhận biết độ tuổi và giới tính dựa vào ảnh chụp khuôn mặt. Adience chứa hơn 26 nghìn hình ảnh với độ phân giải 816×816 của hơn 2 nghìn người khác nhau. Hầu hết các hình ảnh từ bộ dữ liệu được tự động tải xuống từ Flickr và chúng được thu thập trực tiếp từ các thiết bị di động mà không qua lọc thủ công trước đó.

Có 8 nhóm đại diện cho độ tuổi của các đối tượng bao gồm 0-2, 4-6, 8-13, 15-20, 25-32, 38-43, 48-53, 60-. Hình 2 là một ví dụ về các hình ảnh với chất lượng điều kiện ánh sáng kém, bị che một phần khuôn mặt, các tư thế đầu khác nhau, ... cho thấy sự thách thức từ bộ cơ sở dữ liệu này.

Chúng tôi không sử dụng bất kỳ dữ liệu bên ngoài nào trong giai đoạn huấn luyện. Mô hình lightweight CNN được đào tạo từ đầu với hàm tối ưu hóa là Adam. Các hình ảnh huấn luyện được chia thành nhiều phần với kích thước là 32 hình ảnh trên mỗi batch và tỷ lệ học tập chúng tôi sử dụng là 0,001. Để đánh giá chính xác hiệu suất của mô hình, chúng tôi sử dụng five-fold cross validation và so sánh kết quả của mô hình với các phương pháp đã được đề xuất gần đây trong [14], [6], [7], [22] về cả độ chính xác, lượng tham số sử dụng cũng như thời gian thực hiện.



Hình 1. Kiến trúc mô hình lightweight CNN



Hình 2. Một số ảnh trong bộ dữ liệu Adience

4.2 Kết quả và so sánh:

Từ bảng 1, có thể thấy rằng mô hình của chúng tôi cho kết quả cao nhất trong việc ước lượng độ tuổi (đạt 67.9% cao hơn 5.1% so với kết quả tốt nhất hiện tại là VGG-16), về mặt dự đoán giới tính, mô hình của chúng tôi kém hơn 5% so với các mô hình học sâu khác như VGG-16.

Bảng 1. So sánh độ chính xác giữa các mô hình

Mô hình	Tuổi	Giới tính
AdienceNet	50,7% $\pm$ 5,1%	86,8% $\pm$ 1,4%
Best from	46,0% $\pm$ 0,6%	86,0% $\pm$ 1,2%
CaffeNet	54,3%	90,6%
GoogleNet	58,5%	91,7%
VGG-16	62,8%	92,6%
CNN-ELM	52,3% $\pm$ 5,7%	88,2% $\pm$ 1,7%
Lightweight CNN	67,9% $\pm$ 1,9%	88,8% $\pm$ 1,8%

Bảng 2, cho thấy số lượng tham số được sử dụng của các mô hình. Mô hình light weight CNN chỉ sử dụng khoảng 1 triệu tham số, trong khi các mô hình khác sử dụng vài triệu thậm chí hơn 100 triệu tham số chẳng hạn như VGG-16 sử dụng tới 138 triệu tham số.

Bảng 2. So sánh số lượng tham số được sử dụng giữa các mô hình

Mô hình	Số lượng tham số sử dụng
AdienceNet [14]	12 triệu
Best from [7]	7 triệu
CaffeNet [6]	61 triệu
GoogleNet [6]	4 triệu
VGG-16 [6]	138 triệu
CNN-ELM [22]	11 triệu
Lightweight CNN	1 triệu

Về thời gian thực hiện, chúng tôi so sánh với Best from [7] bởi vì đây là mô hình nhẹ duy nhất và cũng là mô hình duy nhất có công bố thời gian chạy. Chúng tôi xây dựng lại mô hình của họ và chạy chúng trên cùng một máy tính có cấu hình 3.6GHz CPU và 20GB RAM. Mô hình trong [7] mất 0.4 giây để dự đoán ra tuổi và giới tính từ một bức ảnh đầu vào, trong khi đó mô hình light weight CNN chỉ mất 0.08 giây để làm việc tương tự.

5. Kết luận

Trong bài báo này, chúng tôi đã đề xuất một mô hình học sâu nhẹ sử dụng mạng CNN để nhận diện tuổi và giới tính dựa vào hình ảnh khuôn mặt. Mô hình mới này cho phép sử dụng một số lượng nhỏ các tham số nhưng đạt hiệu suất tốt hơn các mô hình đã được công bố gần đây, đồng thời góp phần giải quyết vấn đề nhận diện trong thời gian thực.

Trong tương lai gần, chúng tôi đang có kế hoạch cải thiện độ chính xác của mô hình, đặc biệt là đối với ước lượng giới tính. Mặt khác, chúng tôi sẽ áp dụng mô hình của chúng tôi

cho các bài toán khác trong lĩnh vực thị giác máy tính và xử lý hình ảnh.

TÀI LIỆU THAM KHẢO

- [1]. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in neural information processing systems*, pp. 1097-1105, 2012.
- [2]. M. D. Zeiler and R. Fergus, "Visualizing and understanding convolutional networks", *European conference on computer vision*. Springer, pp. 818-833, 2014.
- [3]. C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions", *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1-9, 2015.
- [4]. He Kaiming et al. "Deep residual learning for image recognition", *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016.
- [5]. Y. Kim, "Convolutional neural networks for sentence classification," *arXiv preprint arXiv*, 1408.5882, 2014.
- [6]. S. Lapuschkin, A. Binder, K. R. Muller, and W. Samek, "Understanding and comparing deep neural networks for age and gender classification", *the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1629-1638, 2017.
- [7]. L. Zhu, K. Wang, L. Lin, and L. Zhang, "Learning a lightweight deep convolutional network for joint age and gender recognition", *Pattern Recognition (ICPR)*, 2016 23rd International Conference on. IEEE, pp. 3282-3287, 2016.
- [8]. Y. H. Kwon and da Vitoria Lobo, "Age classification from facial images", in *1994 Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, June 1994, pp. 762-767, 1994.
- [9]. N. Ramanathan and R. Chellappa, "Modeling age progression in young faces", in *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, vol. 1. IEEE, pp. 387-394, 2006.
- [10]. X. Geng, Z. H. Zhou, and K. Smith-Miles, "Automatic age estimation based on facial aging patterns", *IEEE Transactions on pattern analysis and machine intelligence*, Vol. 29, No. 12, pp. 2234-2240, 2007.
- [11]. S. Yan, M. Liu, and T. S. Huang, "Extracting age information from local spatially flexible patches", in *Acoustics, Speech and Signal Processing, ICASSP 2008. IEEE International Conference on*, pp. 737-740, 2008.
- [12]. X. Zhuang, X. Zhou, M. Hasegawa-Johnson, and T. Huang, "Face age estimation using patch-based hidden markov model supervectors", in *Pattern Recognition, 2008. ICPR 2008. 19th International Conference on. IEEE*, pp. 1-4, 2008.
- [13]. E. Eiding, R. Enbar, and T. Hassner, "Age and gender estimation of unfiltered faces", *IEEE Transactions on Information Forensics and Security*, Vol. 9, No. 12, pp. 2170-2179, 2014.
- [14]. G. Levi and T. Hassner, "Age and gender classification using convolutional neural networks", in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 34-42, 2015.
- [15]. D. A. Reid, S. Samangooei, C. Chen, M. S. Nixon, and A. Ross, "Soft biometrics for surveillance: an overview", in *Handbook of statistics*. Elsevier, Vol. 31, pp. 327-352, 2013.
- [16]. G. W. Cottrell and J. Metcalfe, "Empath: Face, emotion, and gender recognition using holons", in *Advances in neural information processing systems*, pp. 564-571, 1991.
- [17]. M. J. Lyons, J. Budynek, A. Plante, and S. Akamatsu, "Classifying facial attributes using a 2-d gabor wavelet representation and discriminant analysis", *Automatic Face and Gesture Recognition, Proceedings. Fourth IEEE International Conference on. IEEE*, pp. 202-207, 2000.
- [18]. B. Moghaddam and M.-H. Yang, "Learning gender with support faces", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 24, No. 5, pp. 707-711, 2002.
- [19]. S. Baluja and H. A. Rowley, "Boosting sex identification performance", *International Journal of computer vision*, Vol. 71, No. 1, pp. 111-119, 2007.
- [20]. I. Ullah, M. Hussain, G. Muhammad, H. Aboalsamh, G. Bebis, and A. M. Mirza, "Gender recognition from face images with local wld descriptor", in *Systems, Signals and Image Processing (IWSSIP)*, 2012 19th International Conference on. IEEE, pp. 417-420, 2012.
- [21]. E. Eiding, R. Enbar, and T. Hassner, "Age and gender estimation of unfiltered faces", *IEEE Transactions on Information Forensics and Security*, Vol. 9, No. 12, pp. 2170-2179, 2014.
- [22]. M. Duan, K. Li, C. Yang, and K. Li, "A hybrid deep learning cnn- elm for age and gender classification", *Neurocomputing*, Vol. 275, pp. 448-461, 2018.