



## BÁO CÁO PHƯƠNG PHÁP LUẬN NCKH

### A. THÔNG TIN CHUNG

#### A1. Tên đề tài

- Tên tiếng Việt: Dự đoán tuổi qua khuôn mặt
- Tên tiếng Anh: Age Prediction

#### A2. Danh sách thành viên

- Phạm Ngọc Trường : 18521571
- Nguyễn Văn Thịnh : 18521448
- Nguyễn Minh Quang : 18521299

### B. MÔ TẢ NGHIÊN CỨU

#### B1. Giới thiệu về đề tài

(Ghi các ý về tổng quan tình hình nghiên cứu liên quan đến đề tài, lí do thực hiện đề tài, các thách thức)

##### B1.1 Tính cấp thiết

Trong thời kì hiện nay, nhờ điều kiện sống được nâng cao nên trẻ em ngày càng phát triển và trưởng thành nhanh chóng, một số còn phát triển trông lớn hơn nhất về khuôn mặt so với lứa tuổi. Người lớn thì càng ngày càng già nhanh đi làm chúng ta khó có thể đoán đúng số tuổi để xưng hô, hay khó phân biệt độ tuổi.... Những khu vui chơi, giải trí, các phương tiện vận tải,... thường bán vé theo độ tuổi, trong trường hợp một số cá nhân muốn trốn vé thì khó lòng phân biệt được. Trong khi việc kiểm tra giấy tờ tùy thân được xem là một hành động thiếu tôn trọng. Vì vậy, việc dự đoán được độ tuổi của một người thông qua ảnh chụp, video về gương mặt người theo góc nhìn chính diện từ các camera quan sát là một yêu cầu quan trọng.

## B1.2 Tổng quan tình hình nghiên cứu

- Một số mạng học sâu như MobileNet, GoogleNet, VGG16, ResidualNet đã được cải tiến và phát triển mạnh trong việc nhận dạng hình ảnh. Tuy nhiên, việc đào tạo các mạng nơon học sâu gặp một số khó khăn, bao gồm việc mất mát gradients và tốc độ đào tạo.
- Một số kỹ thuật nhằm cải tiến, cho phép đào tạo mô hình mạng học sâu mạnh mẽ hơn như: initialization strategies, better optimizers, skip connection, knowledge transfer, layer-wise training giúp cải tiến đáng kể tốc độ và độ chính xác của các mạng học sâu có sẵn.

## B1.3 Đối tượng nghiên cứu

Đối tượng nghiên cứu bao gồm 9778 ảnh chụp gương mặt cận cảnh theo góc nhìn chính diện ở các độ tuổi trải dài từ 1 đến 110 tuổi.

## B1.4 Các thách thức

Để nhận biết được độ tuổi qua khuôn mặt thì yếu tố, các điểm trên khuôn mặt là rất quan trọng, các tác nhân như lão hóa, phẫu thuật thẩm mỹ, trang điểm, râu, các ứng dụng chỉnh sửa ảnh,... ảnh hưởng không nhỏ đến kết quả dự đoán.

Để loại bỏ sự ảnh hưởng của các yếu tố này vào kết quả dự đoán, cần phải sử dụng các phương pháp khác để nhận diện độ tuổi một cách chính xác hơn.



Hình 1. Mặt bị lão hóa và không bị lão hóa



Hình 2. Mặt không trang điểm và trang điểm

### B1.5 Định nghĩa của bài toán

- Input: là hình ảnh các khuôn mặt sau detect.



Hình 3. Ảnh input

- Output: Ảnh có gán nhãn về độ tuổi của khuôn mặt.



Hình 4. Ảnh output

## **B2. Mục tiêu, nội dung, kế hoạch nghiên cứu**

### **B2.1 Mục tiêu**

- Gắn nhãn độ tuổi cho khuôn mặt được truyền vào dưới dạng ảnh hoặc video một cách chính xác mà không phụ thuộc vào các yếu tố trên khuôn mặt.
- Tốc độ xử lý đủ nhanh để có khả năng chạy realtime.

### **B2.2 Nội dung và phương pháp nghiên cứu**

#### **Nội dung 1: Khảo sát các phương pháp giải quyết bài toán dự đoán tuổi qua khuôn mặt**

Phương pháp:

- Nghiên cứu mô hình chung để giải quyết bài toán phân lớp với input là ảnh
- Tìm hiểu các phương pháp để nhận diện độ tuổi dựa vào gương mặt

Bài toán này được giới thiệu lần đầu tiên bởi Kwon và Lobo [1] 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 [2]. 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 [3] 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 [4] và HMM, super-vectors [5] 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 [6] đã sử dụng SVM kết hợp với dropout cho bài toán nhận diện. Năm 2015, GilLevi và Tal Hassner [7] 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 [8] 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. Đây là bài báo đầu tiên áp dụng mô hình tối ưu hóa bài toán nhận diện tuổi.

- Tìm hiểu một số cấu trúc mạng nơ-ron: GoogleNet, ResNet, DenseNet, MobileNet, VGG16

## Available models

Model	Size	Top-1 Accuracy	Top-5 Accuracy	Parameters	Depth
Xception	88 MB	0.790	0.945	22,910,480	126
VGG16	528 MB	0.713	0.901	138,357,544	23
VGG19	549 MB	0.713	0.900	143,667,240	26
ResNet50	98 MB	0.749	0.921	25,636,712	-
ResNet101	171 MB	0.764	0.928	44,707,176	-
ResNet152	232 MB	0.766	0.931	60,419,944	-
ResNet50V2	98 MB	0.760	0.930	25,613,800	-
ResNet101V2	171 MB	0.772	0.938	44,675,560	-
ResNet152V2	232 MB	0.780	0.942	60,380,648	-
InceptionV3	92 MB	0.779	0.937	23,851,784	159
InceptionResNetV2	215 MB	0.803	0.953	55,873,736	572
MobileNet	16 MB	0.704	0.895	4,253,864	88
MobileNetV2	14 MB	0.713	0.901	3,538,984	88
DenseNet121	33 MB	0.750	0.923	8,062,504	121
DenseNet169	57 MB	0.762	0.932	14,307,880	169
DenseNet201	80 MB	0.773	0.936	20,242,984	201
NASNetMobile	23 MB	0.744	0.919	5,326,716	-
NASNetLarge	343 MB	0.825	0.960	88,949,818	-
EfficientNetB0	29 MB	-	-	5,330,571	-

Hình 5. Các mô hình và số parameters cần thiết

- 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.

Dự kiến kết quả:

- Tài liệu kỹ thuật về các nghiên cứu liên quan
- Tài liệu về các mạng nơ-ron



## Nội dung 2: Trình bày chi tiết các bước cơ bản để giải quyết bài toán

Phương pháp:

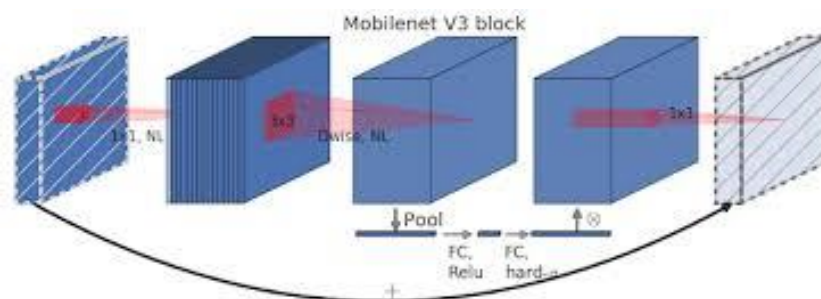
- So sánh các phương pháp: GoogleNet, MobileNet, VGG16

Model	ImageNet Accuracy	Million Mult-Adds	Million Parameters
1.0 MobileNet-224	70.6%	569	4.2
GoogleNet	69.8%	1550	6.8
VGG 16	71.5%	15300	138

Hình 6. MobileNet-224 với GoogleNet và VGG 16 trên tập ImageNet

Khi so sánh 1.0 MobileNet-224 với GoogleNet và VGG 16, chúng ta thấy rằng độ chính xác của cả 3 thuật toán là hầu như tương đương nhau. Nhưng 1.0 MobileNet-224 có số lượng tham số ít (75% so với GoogleNet) và số lượng phép toán nhỏ hơn rất nhiều => chạy nhanh hơn => lựa chọn MobileNet để triển khai

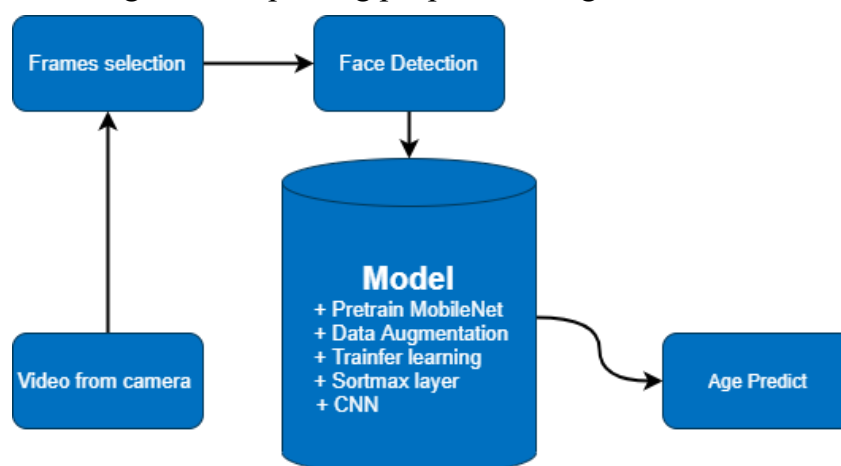
- Kiến trúc mạng MobileNet V3



## Nội dung 3: Trình bày chi tiết các bước cơ bản để giải quyết bài toán

Phương pháp:

- Cài đặt thử nghiệm các phương pháp ở nội dung 2 theo lưu đồ sau:



- Đánh giá các kết quả thu được

Dự kiến kết quả:

- Tài liệu kỹ thuật được sử dụng để cài đặt
- Báo cáo đánh giá hiệu năng của phương pháp

### **B3. Kết quả nghiên cứu dự kiến**

1. So sánh kết quả các phương pháp nghiên cứu và chọn ra được phương pháp tối ưu nhất trong các phương pháp GoogleNet, MobileNet, VGG16.
2. Demo Apk app/Web.

### **B4. Tài liệu tham khảo**

- [1]. 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.
- [2]. 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.
- [3]. 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.
- [4]. 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.
- [5]. 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.
- [6]. 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.
- [7]. 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.
- [8]. 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.