



## ĐỀ CƯƠNG KHOÁ LUẬN TỐT NGHIỆP

# ÁP DỤNG MÔ HÌNH TẠO SINH ĐỐI NGỊCH VÀO PHẦN MỀM CHỤP ẢNH GIẢI TRÍ TRÊN ĐIỆN THOẠI ANDROID

*(APPLYING GENERATIVE ADVERSARIAL NETWORKS TO  
ANDROID ENTERTAINMENT PHOTOGRAPHY SOFTWARE)*

## 1 THÔNG TIN CHUNG

**Người hướng dẫn:**

– TS. Ngô Huy Biên (Khoa Công nghệ Thông tin)

**Nhóm sinh viên thực hiện:**

1. Huỳnh Kim Ninh (MSSV: 1612484)

2. Vũ Thừa Khang (MSSV: 1512239)

**Loại đề tài:** Ứng dụng

**Thời gian thực hiện:** Từ 09/2020 đến 03/2021

## 2 NỘI DUNG THỰC HIỆN

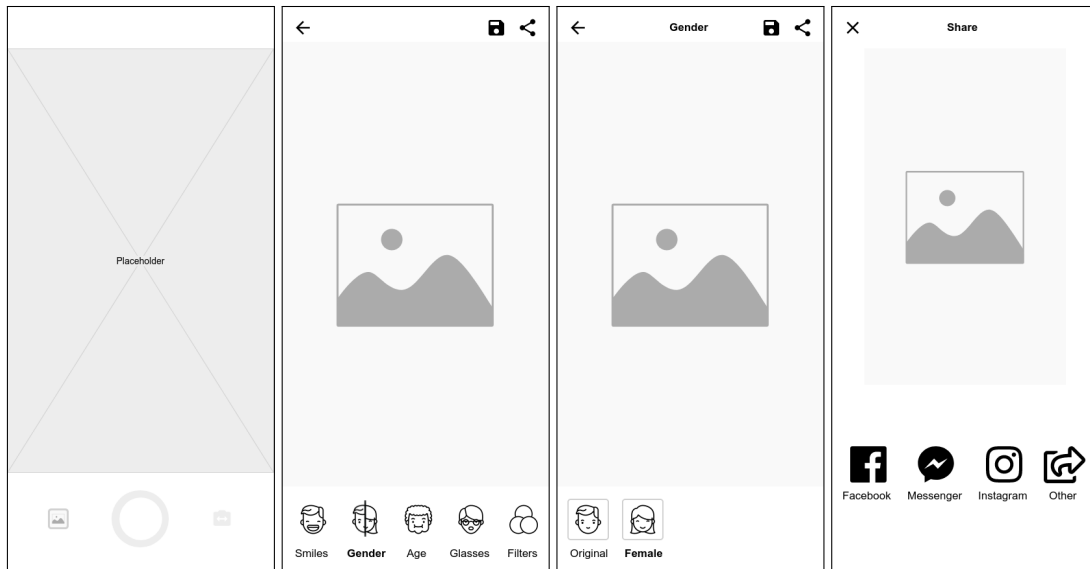
### 2.1 Giới thiệu về đề tài

- Xây dựng và thu thập dữ liệu đào tạo mô hình mạng đối nghịch tạo sinh để sửa một bức ảnh mặt người.

- Vai trò sinh viên: Data Scientist, Data Collector, Developer, Tester, Project Manager/Scrum Master.
- Kỹ năng yêu cầu: Kỹ năng lập trình Android, lập trình Python, các thuật toán học máy.
- Ứng dụng Mạng đối nghịch tạo sinh vào phần mềm chụp ảnh giải trí trên điện thoại Android giúp cho người dùng tạo nên những bức ảnh độc đáo, cá tính và vui tươi. Nhóm sử dụng mô hình CycleGAN, ưu điểm của việc này chính là không yêu cầu quá khắc khe đối với bộ dữ liệu để đào tạo mô hình. Bên cạnh đó, khác với các phần mềm hiện tại như Faceapp, 360,... việc ứng dụng Mạng đối nghịch tạo sinh sẽ chỉnh sửa ảnh một cách tự động, nhanh nhạy và thông minh, giúp cho người dùng thư giãn sau những giờ làm việc căng thẳng.

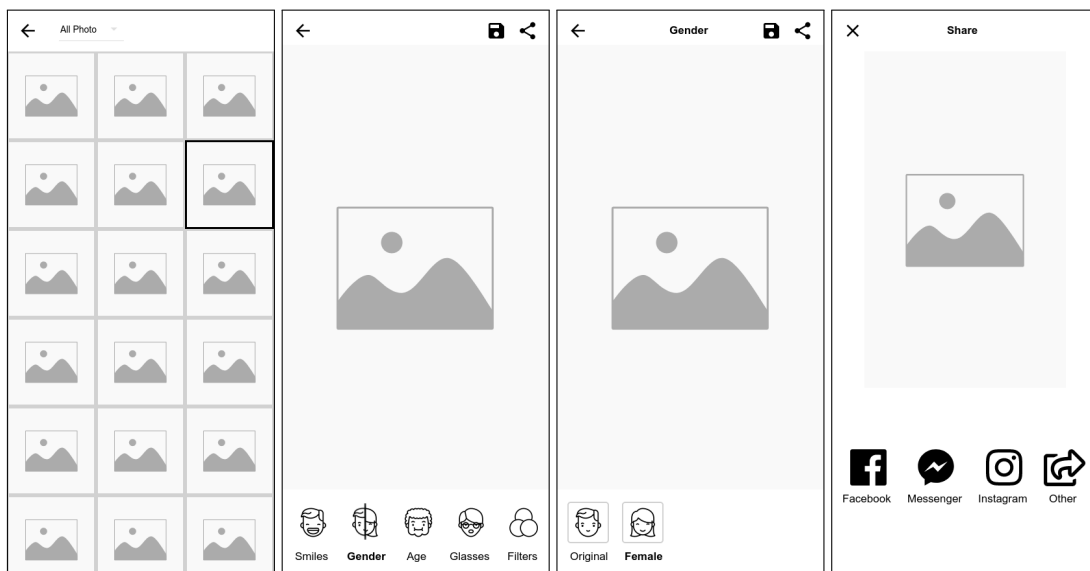
## 2.2 Mục tiêu đề tài

- Trình bày lý thuyết nền tảng và giải pháp để xử lý một bức ảnh bằng Mạng đối nghịch tạo sinh.
- Viết 120 trang luận văn theo đúng chuẩn yêu cầu và trích dẫn tài liệu tham khảo đầy đủ.
- Xây dựng một ứng dụng Android để minh họa các hiệu ứng, lưu ảnh kết quả và chia sẻ với bạn bè trên mạng xã hội.
- Các hiệu ứng của ứng dụng: :
  - Hiệu ứng già hóa, trẻ hóa khuôn mặt.
  - Hiệu ứng hoán đổi giới tính nam/nữ.
  - Hiệu ứng chuyển đổi sắc thái khuôn mặt (nụ cười).
- Bản mẫu (Prototype):



(a) Mở camera (b) Chọn hiệu ứng (c) Chọn hiệu ứng con (d) Chia sẻ ảnh

Hình 1: Chụp ảnh, áp dụng hiệu ứng và chia sẻ với bạn bè



(a) Mở thư viện ảnh (b) Chọn hiệu ứng (c) Chọn hiệu ứng con (d) Chia sẻ ảnh

Hình 2: Chọn ảnh từ thư viện, áp dụng hiệu ứng và chia sẻ với bạn bè

## 2.3 Phạm vi của đề tài

- Ứng dụng cần có kết nối Internet để có thể sử dụng hiệu ứng.
- Ứng dụng chỉ hỗ trợ điện thoại Android phiên bản 5.0 trở lên.

- Các hiệu ứng của ứng dụng chỉ hỗ trợ biến đổi một chiều:
  - Chuyển đổi mặt trẻ thành mặt già
  - Chuyển đổi nam thành nữ
  - Chuyển đổi khuôn mặt không cười thành có cười.

## 2.4 Cách tiếp cận dự kiến

- Mô hình GAN:

Mô hình giúp chúng ta tạo ra các ảnh mới từ một vector nhiễu bất kì  $z$ . Ảnh được tạo ra sẽ giống với tập ảnh đưa vào huấn luyện. Gồm có 2 mạng con: Generator và Discriminator có nhiệm vụ lần lượt là phát sinh ảnh mới và phân biệt ảnh thật với ảnh được phát sinh từ Generator.

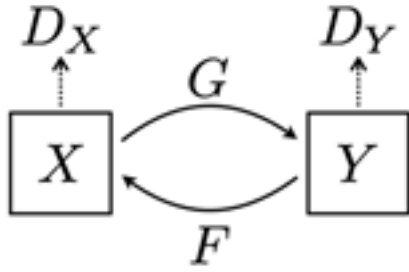
- Pix2pix

Đây là một dạng của mô hình GAN. Mô hình giúp chuyển đổi ảnh  $x$  thành ảnh  $y$  bằng cách chuyển đổi màu sắc trên các pixel. Do mô hình này chỉ đơn thuần học cách chuyển đổi màu sắc trên 2 tấm ảnh nên cặp dữ liệu đưa vào huấn luyện  $x, y$  phải là bức ảnh của cùng một đối tượng, chủ thể (Cụ thể trong đề tài của nhóm thì  $x, y$  phải là cặp ảnh chứa khuôn mặt trẻ/già của cùng một người)

- CycleGAN (Mô hình nhóm sẽ áp dụng)

Mô hình này cũng là một mô hình giúp chuyển đổi ảnh giống mô hình pix2pix. Tuy nhiên mô hình này giúp khắc phục được hạn chế trên mô hình pix2pix, giúp chúng ta có thể biến đổi ảnh từ khuôn mặt trẻ thành khuôn mặt già và không yêu cầu cặp ảnh đưa vào huấn luyện phải cùng khuôn mặt của một người.

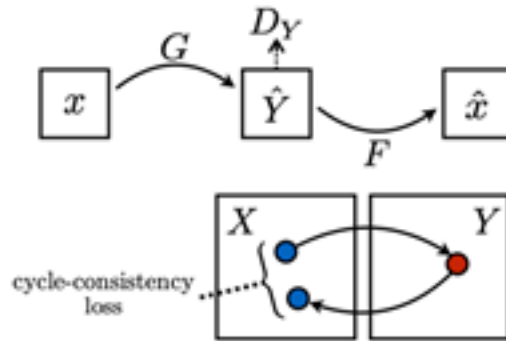
- Kiến trúc



Hình 3: Kiến trúc mô hình CycleGAN

Mô hình gồm có 4 mạng con cụ thể như sau: Hai mạng Generator  $G(X)$ ,  $F(Y)$  giúp biến đổi  $X$  thành  $Y$  và biến đổi  $Y$  thành  $X$ . Hai mạng Discriminator  $D_Y$ ,  $D_X$  giúp phân biệt ảnh thật/giả từ domain  $Y$  và phân biệt ảnh thật/giả từ domain  $X$ .

– Ý tưởng



Hình 4: Hàm chi phí của CycleGAN

$G$  sẽ biến đổi khuôn mặt trẻ  $x$  thành khuôn mặt già  $\hat{Y}$  sao cho  $\hat{Y}$  giống  $Y$  trong bộ dữ liệu nhất (đánh lừa được  $D_y$ ). Khuôn mặt già  $\hat{Y}$  sẽ được biến đổi ngược về khuôn mặt trẻ  $\hat{x}$  (thông qua  $F$ ) sao cho  $\hat{x}$  giống  $x$  ban đầu nhất. Mục đích của phép biến đổi  $F$  là đảm bảo  $\hat{Y}$  và  $x$  phải là cặp ảnh mặt già/trẻ của cùng một người.

## 2.5 Kết quả dự kiến của đề tài

- Máy chủ cung cấp các API để tải lên bức ảnh của người dùng và trả về tấm ảnh đã được xử lý.
- Phần mềm mẫu chạy trên điện thoại Android cho phép chụp hoặc chọn hình có sẵn, sau đó chọn hiệu ứng yêu thích và cuối cùng là lưu về máy hoặc chia sẻ với bạn bè trên mạng xã hội.

## 2.6 Các bộ dữ liệu dự kiến sử dụng

- Large-scale CelebFaces Attributes (CelebA) Dataset:

Gồm hai bộ ảnh khuôn mặt riêng biệt:

- CelebAMask-HQ<sup>1</sup>: Đây là tập dữ liệu hình ảnh khuôn mặt quy mô lớn có 30.000 hình ảnh khuôn mặt với độ phân giải cao được chọn từ tập dữ liệu CelebA.
- CelebA-Spoof<sup>2</sup>: Đây là tập dữ liệu hình ảnh khuôn mặt chống giả mạo quy mô rất lớn với 625.537 hình ảnh từ 10.177 đối tượng, bao gồm 43 thuộc tính phong phú về khuôn mặt, ánh sáng, môi trường.

- UTKFace<sup>3</sup>:

Tập dữ liệu UTKFace là một tập dữ liệu về khuôn mặt có quy mô lớn với độ tuổi đa dạng (từ 0 đến 116 tuổi). Tập dữ liệu bao gồm hơn 20.000 hình ảnh khuôn mặt với chú thích về tuổi, giới tính và sắc tộc. Các hình ảnh bao gồm sự thay đổi lớn về góc chụp, biểu cảm khuôn mặt, độ chiếu sáng, độ phân giải, v.v. Bộ dữ liệu này có thể được sử dụng cho nhiều tác vụ khác nhau, ví dụ: phát hiện khuôn mặt, ước tính tuổi, ước đoán tương lai, v.v.

- IMDB-WIKI<sup>4</sup> :

Hơn 500.000 khuôn mặt đã được đánh nhãn tuổi và giới tính. Đây là bộ thư viện khuôn mặt tốt nhất từng được công bố trên cộng đồng, các hình ảnh đã

---

<sup>1</sup><https://github.com/switchablenorms/CelebAMask-HQ>

<sup>2</sup><https://github.com/Davidzhangyuanhan/CelebA-Spoof>

<sup>3</sup><https://susanqq.github.io/UTKFace>

<sup>4</sup><https://data.vision.ee.ethz.ch/cvl/rrothe/imdb-wiki/>

được đánh nhãn rõ ràng để thuận tiện cho việc đào tạo mô hình máy học. Có thể sử dụng cho việc dự đoán tuổi và dự đoán giới tính.

- Cross-Age Celebrity Dataset<sup>5</sup>:

Bao gồm 163,446 hình ảnh từ hơn 2000 đối tượng trên Internet. Bộ dữ liệu này được sử dụng cho việc nghiên cứu khoa học. Tất cả hình ảnh đều được thu thập từ trên internet và quyền sở hữu thuộc về chủ nhân của chúng. Việc thu thập dữ liệu được thực hiện bằng cách tìm kiếm tên kèm theo ngày tháng năm. Tất cả các bức ảnh đều được lọc nhiễu. Tuy nhiên vẫn có một số không thể nhận diện ngay cả đối với con người.

---

<sup>5</sup><https://bcsiriuschen.github.io/CARC/>

## 2.7 Kế hoạch thực hiện

Mốc thời gian	Công việc
19/07/2020	Chuẩn bị kiến thức nền cho luận văn
23/09/2020	Làm quen với Python, Numpy, Tensorflow, Keras
28/09/2020	Thực hiện đề cương khóa luận
04/10/2020	Khảo sát các ứng dụng trên thị trường
05/10/2020	Viết chương 1 luận văn
11/10/2020	Chạy thử các mô hình mã nguồn mở Chuẩn bị các bộ dữ liệu khuôn mặt người
12/10/2020	Chuẩn bị mô hình, huấn luyện thử với bộ dữ liệu
18/10/2020	Hoàn chỉnh chương 1 luận văn
19/10/2020	Viết chương 5 luận văn
25/10/2020	Ghi nhận kết quả của mô hình
26/10/2020	Viết chương 4 luận văn
01/11/2020	Cải thiện mô hình
02/11/2020	Viết chương 3 luận văn
15/11/2020	Deploy mô hình lên server
16/11/2020	Viết chương 2 luận văn
06/12/2020	Tiếp tục cải thiện mô hình, ghi nhận kết quả.
07/12/2020	Viết các tài liệu hướng dẫn
13/12/2020	
14/12/2020	Hoàn thiện app Android để demo
20/12/2020	
21/12/2020	Quay video bảo vệ thử luận văn
27/12/2020	
28/12/2020	Cập nhật các sản phẩm đủ chất lượng đăng ký bảo vệ
17/01/2021	
25/01/2021	Cải thiện mô hình
28/02/2021	Cải thiện ứng dụng Android
	Cập nhật các chương luận văn



## Tài liệu

- [1] A. Géron, *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow*. O'Reilly Media, Inc, 2019.
- [2] M. Nielsen, “Neural networks and deep learning.” <http://neuralnetworksanddeeplearning.com/>.
- [3] “Generative models.” <https://paperswithcode.com/methods/category/generative-models>.
- [4] S. GUAN, “Generating custom photo realistic faces using ai.” <https://blog.insightdatascience.com/generating-custom-photo-realistic-faces-using-ai-d170b1b59255>.
- [5] A. Wang, “Hairstyle transfer semantic editing gan latent code.” <https://medium.com/swlh/hairstyle-transfer-semantic-editing-gan-latent-code-b3a6ccf91e82>.

XÁC NHẬN  
CỦA NGƯỜI HƯỚNG DẪN  
(Ký và ghi rõ họ tên)

TP. Hồ Chí Minh, ngày/tháng/năm  
NHÓM SINH VIÊN THỰC HIỆN  
(Ký và ghi rõ họ tên)