



Nhận dạng khuôn mặt bằng cách sử dụng mạng nơ-ron tích chập của Google

Lê Quang Vinh - 3120411176
Vũ thị thanh vân - 3120411173

Mục lục

1. Motivation

2. Phương pháp đề xuất

3. demo chương trình

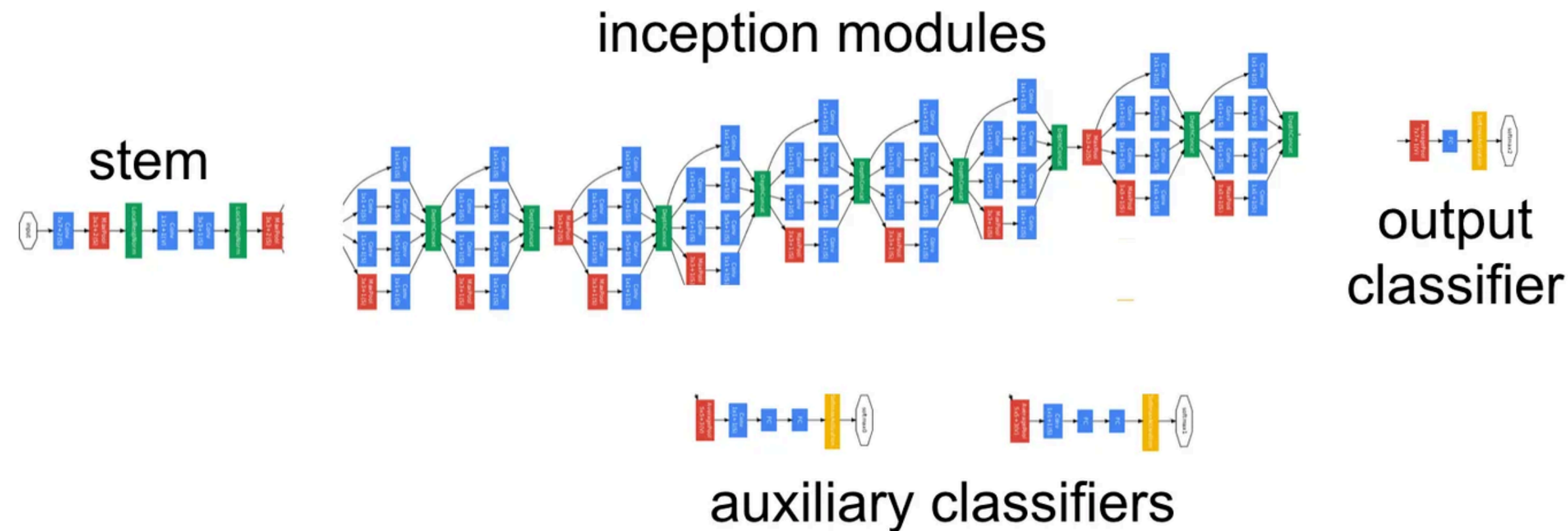
4. Kết luận

1. Motivation (động cơ nghiên cứu)

1. Motivation (Động cơ nghiên cứu)

1.1 Bối cảnh

- Nhận diện khuôn mặt là bài toán quan trọng trong lĩnh vực thị giác máy tính, có ứng dụng trong bảo mật, giám sát, và xác thực cá nhân.
- Mô hình CNN Inception V3 của Google đã đạt hiệu quả cao trong nhận dạng hình ảnh tổng quát, tuy nhiên chưa được tối ưu cho nhận diện khuôn mặt cụ thể.











Sơ đồ Inception V3

1. Motivation (Động cơ nghiên cứu)

1.2 Vấn đề

- Input: Tập ảnh khuôn mặt của nhiều cá nhân, trong đó mỗi ảnh có các biến thiên về góc chụp, ánh sáng, biểu cảm, và phong nền.
- Output: Nhãn (label) tương ứng với từng cá nhân trong tập dữ liệu, tức là mô hình cần phân loại chính xác từng ảnh thuộc về ai trong số 10 cá nhân nhiều nhất được chọn.

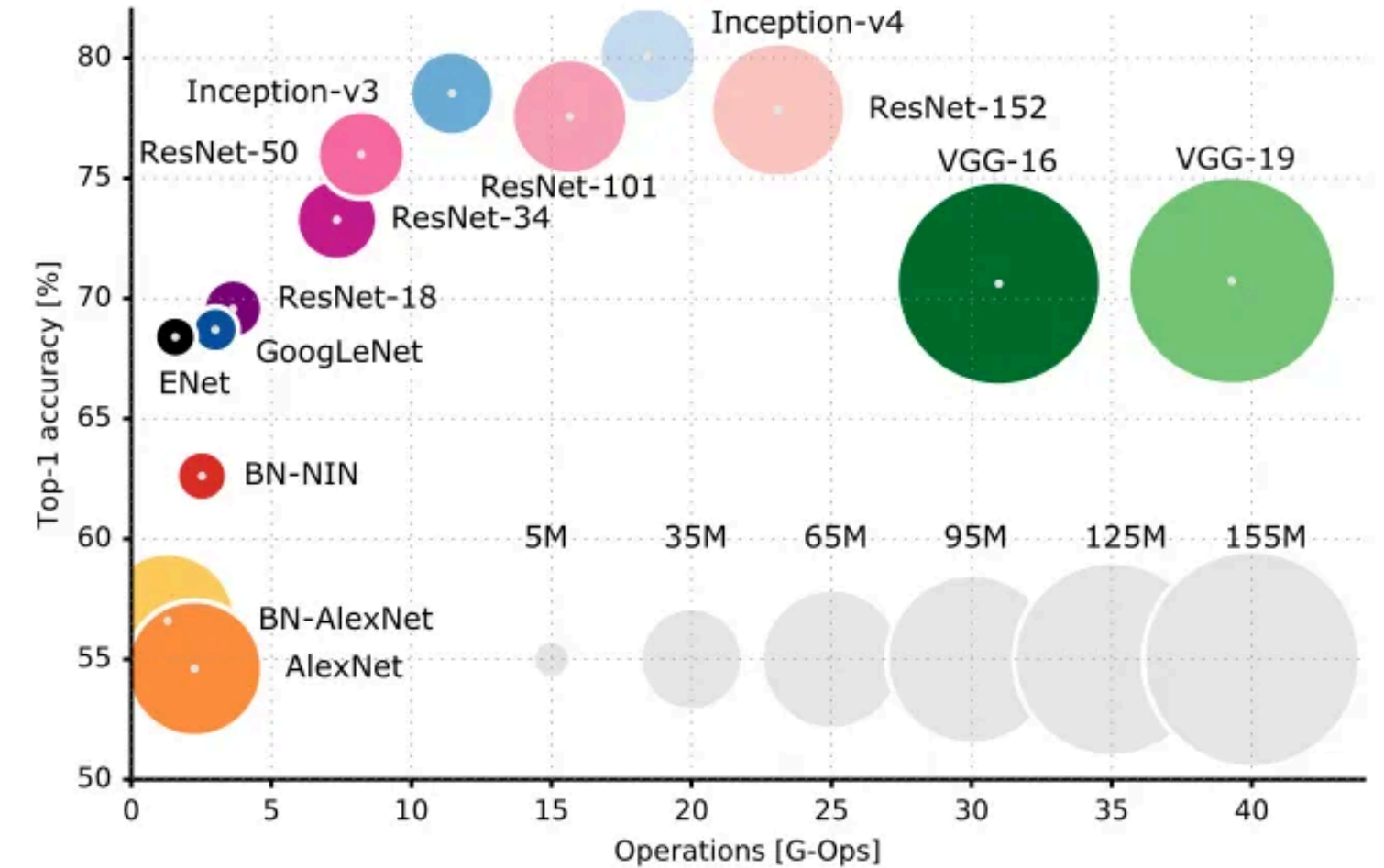
 530 George_W_Bush	8/4/2017 9:24 PM	File folder
 236 Colin_Powell	8/4/2017 9:24 PM	File folder
 144 Tony_Blair	8/4/2017 9:24 PM	File folder
 121 Donald_Rumsfeld	8/4/2017 9:24 PM	File folder
 109 Gerhard_Schroeder	8/4/2017 9:24 PM	File folder
 077 Ariel_Sharon	8/4/2017 9:24 PM	File folder
 071 Hugo_Chavez	8/4/2017 9:24 PM	File folder
 060 Junichiro_Koizumi	8/4/2017 9:24 PM	File folder
 055 Jean_Chretien	8/4/2017 9:24 PM	File folder
 053 John_Ashcroft	8/4/2017 9:24 PM	File folder

Thư mục với 10 cá nhân có nhiều hình ảnh nhất

1. Motivation (Động cơ nghiên cứu)

1.3 Thách thức

- Đa dạng và biến thiên của dữ liệu ảnh khuôn mặt
- Phát hiện khuôn mặt chính xác trong ảnh đa đối tượng
- Phân biệt chính xác các cá nhân (face verification/recognition)
- Chi phí tính toán và quá trình huấn luyện
- Yêu cầu về độ chính xác và bảo mật cao



So sánh độ chính xác và độ phức tạp của các kiến trúc CNN

2. Phương pháp đề xuất

2. Phương pháp đề xuất

2.1 Tổng quan

- Sử dụng mô hình Inception V3 đã được huấn luyện trước trên ImageNet, chỉ huấn luyện lại (fine-tune) lớp cuối cùng để phân biệt 10 cá nhân khác nhau dựa trên ảnh khuôn mặt.
- Giới hạn tập dữ liệu chỉ gồm 10 cá nhân có nhiều ảnh nhất trong bộ dữ liệu Labeled Faces in the Wild (LFW), mỗi cá nhân có ít nhất 50 ảnh.
- Ảnh được xử lý bằng kỹ thuật deep funneling để căn chỉnh khuôn mặt, giảm biến thiên do góc chụp, ánh sáng, giúp mô hình học đặc trưng khuôn mặt chính xác hơn.



Ví dụ hình ảnh được trích từ bộ dữ liệu LFW

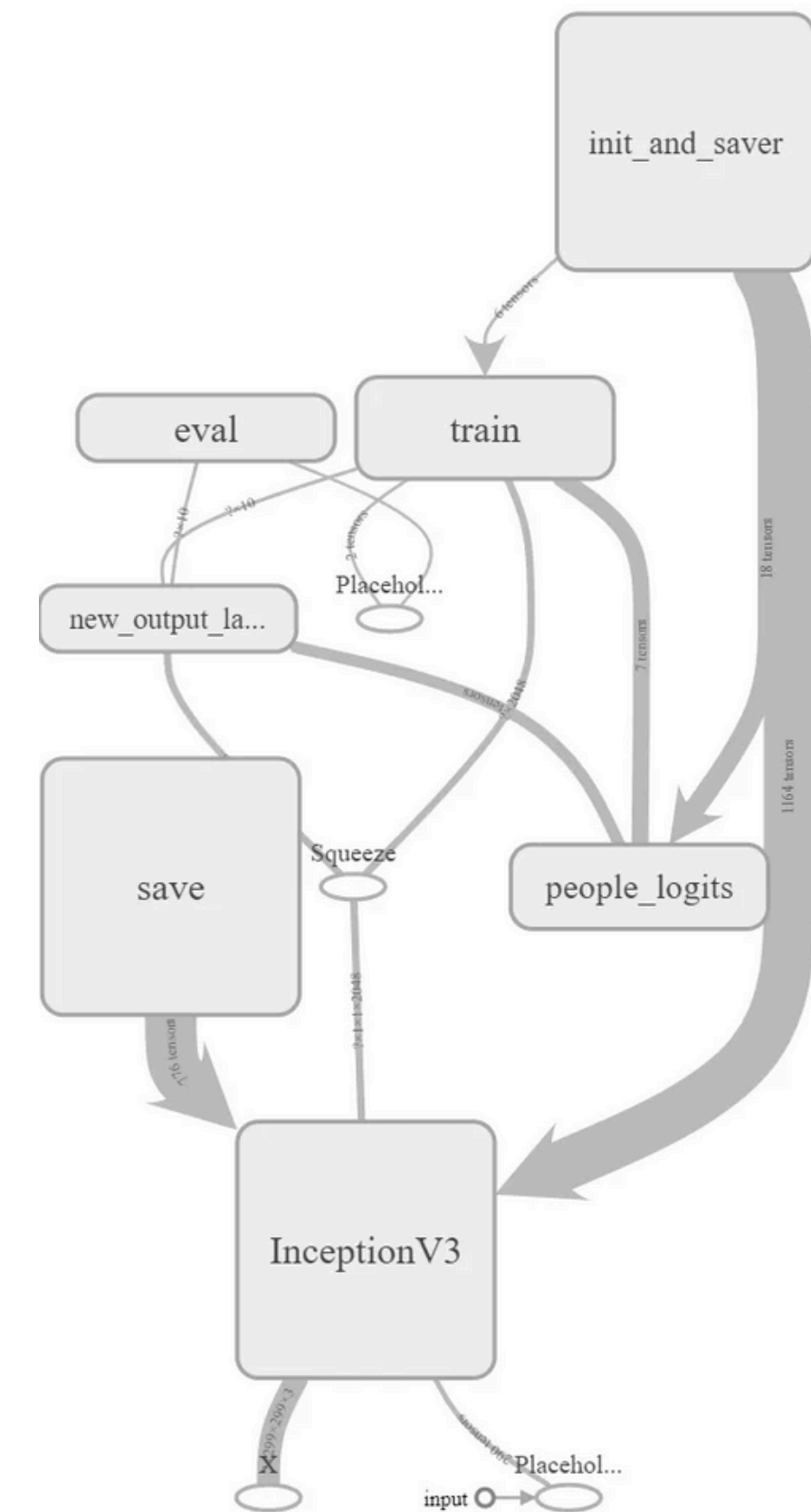
2.2 Quy trình huấn luyện lại mô hình Inception V3

2.2.1 Lớp ban đầu

Sơ đồ này mô tả pipeline huấn luyện lại mô hình nhận diện khuôn mặt:

- Ảnh đầu vào được trích xuất đặc trưng qua InceptionV3.
- Đặc trưng này được đưa vào lớp phân loại mới để xác định cá nhân.
- Quá trình huấn luyện, đánh giá, lưu mô hình được kiểm soát qua các node train, eval, save, init_and_saver.
- Mục tiêu là tối ưu hóa khả năng phân biệt giữa các cá nhân dựa vào ảnh khuôn mặt đầu vào.

Sơ đồ này là một biểu diễn trực quan cho pipeline fine-tuning CNN trong nhận diện khuôn mặt, thể hiện rõ các bước chính và luồng dữ liệu trong quá trình huấn luyện mô hình.

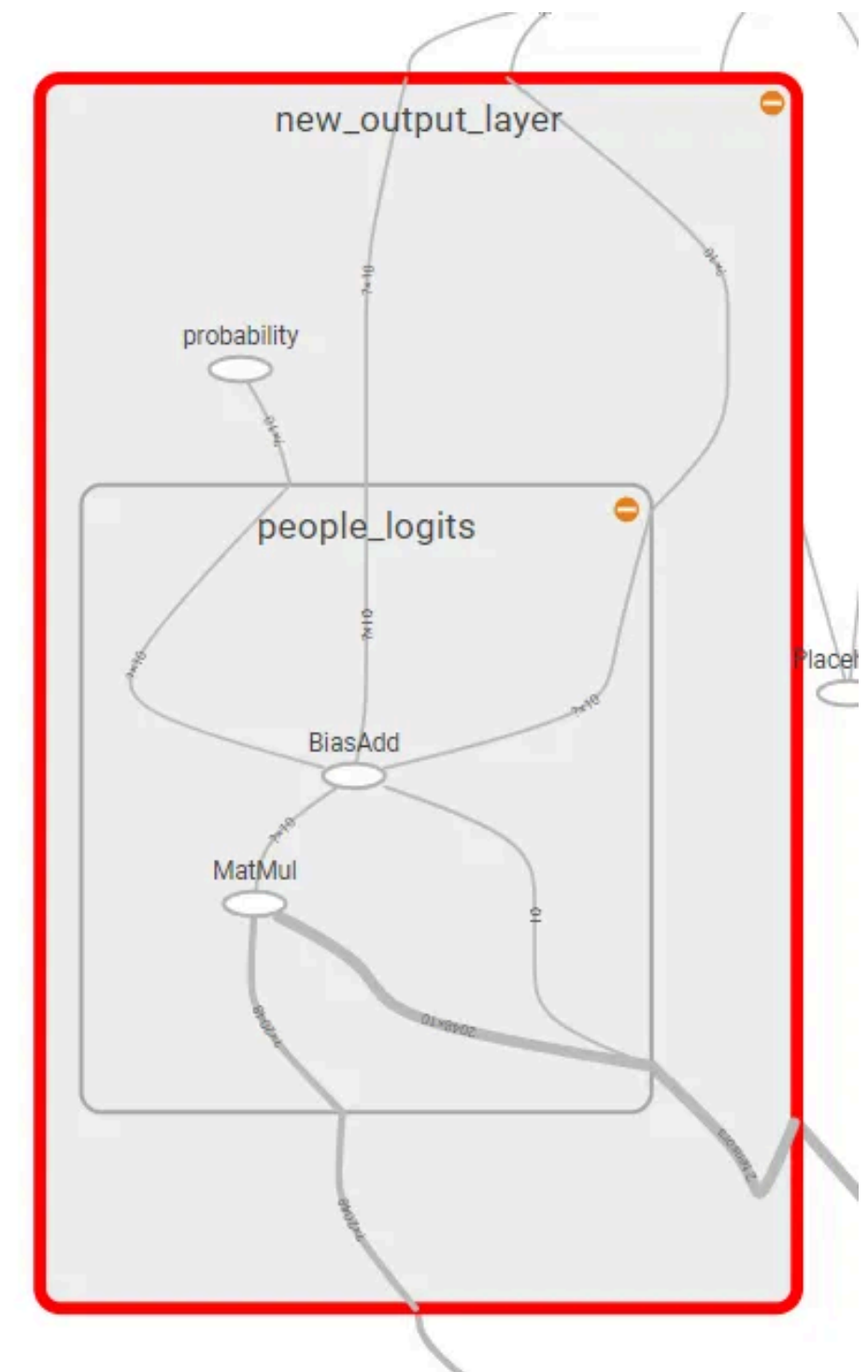


2.2 Quy trình huấn luyện lại mô hình Inception V3

2.2.2 Lớp đầu ra mới

- Luồng xử lý: Đặc trưng ảnh đầu vào → MatMul (nhân trọng số) → BiasAdd (cộng bias) → people_logits (logits) → probability (xác suất dự đoán).
- Chức năng: Lớp này đảm nhận nhiệm vụ phân loại cuối cùng, chuyển đổi đặc trưng trích xuất thành xác suất dự đoán cho từng cá nhân trong bài toán nhận diện khuôn mặt.

Sơ đồ này minh họa rõ cách lớp phân loại mới được xây dựng và hoạt động trong pipeline fine-tuning mô hình nhận diện khuôn mặt, đảm bảo mô hình có thể học và phân biệt các cá nhân dựa trên dữ liệu huấn luyện.

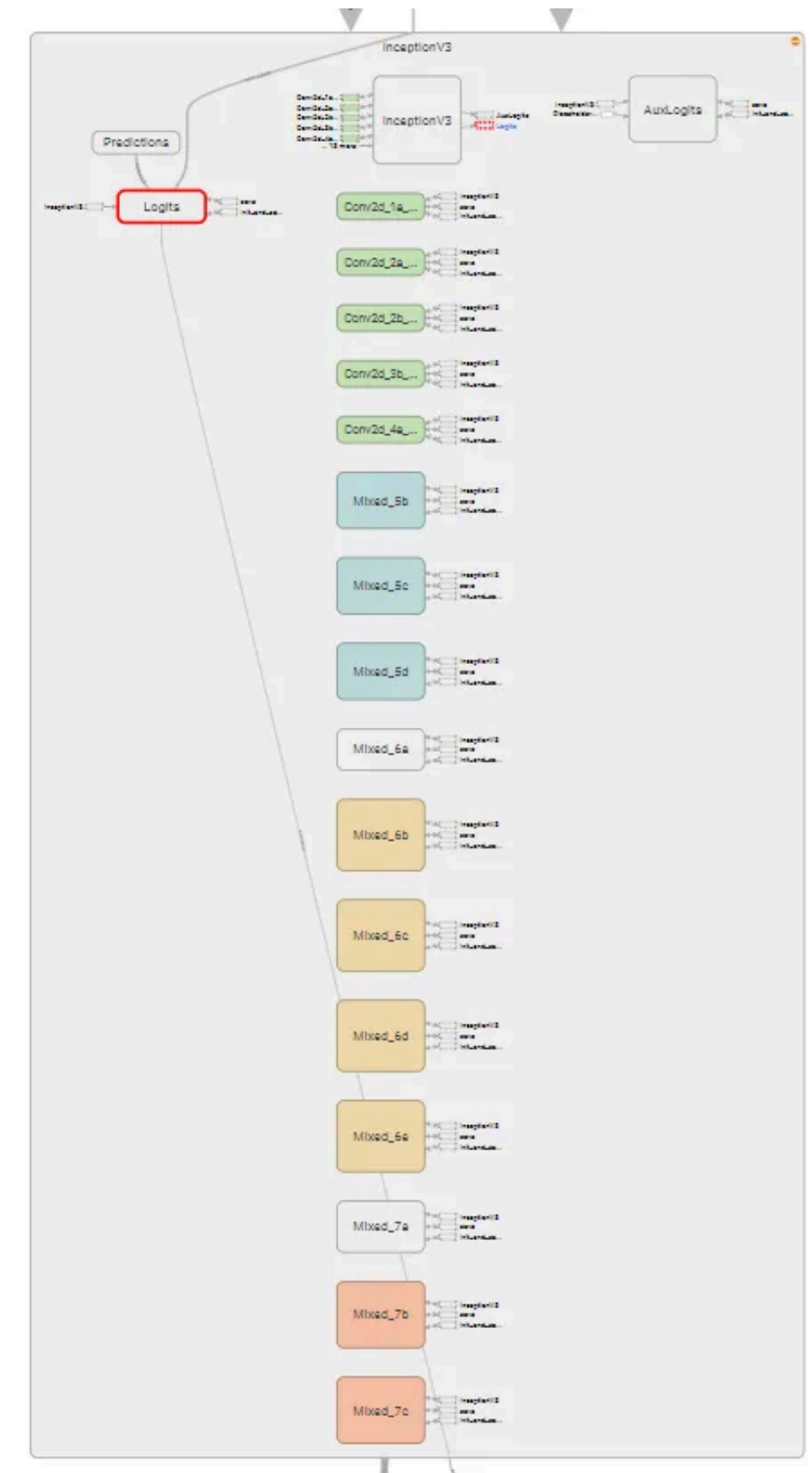


2.3 Sơ đồ tổng thể của mạng tích chập Inception V3

Sơ đồ này mô tả pipeline xử lý của mô hình InceptionV3:

- Ảnh đầu vào → qua nhiều lớp convolution và khối Inception → gom lại thành đặc trưng tổng hợp → lớp Logits → dự đoán cuối cùng (Predictions).
- Ứng dụng:
- Trong nhận diện khuôn mặt, đầu ra của lớp Logits sẽ được điều chỉnh (fine-tune) để phù hợp phân loại các cá nhân cụ thể dựa trên dữ liệu huấn luyện.

Sơ đồ này giúp hình dung rõ ràng kiến trúc sâu, phức tạp và khả năng trích xuất đặc trưng mạnh mẽ của InceptionV3, nền tảng cho các ứng dụng nhận diện khuôn mặt hiện đại.



3. demo chương trình

4. Kết luận

4. Kết luận

- Việc tinh chỉnh mô hình CNN Inception V3 đã được huấn luyện trước giúp tiết kiệm thời gian và tài nguyên so với xây dựng mô hình mới từ đầu.
- Kỹ thuật deep funneling giúp cải thiện chất lượng dữ liệu, giảm nhiễu do biến thiên ảnh, nâng cao hiệu quả nhận diện khuôn mặt.
- Phương pháp này có thể áp dụng cho các bài toán nhận dạng khuôn mặt với tập dữ liệu hạn chế và tài nguyên tính toán vừa phải.



Thank You !