

NHẬN DẠNG BIỂU CẢM KHUÔN MẶT BẰNG PHƯƠNG PHÁP HỌC SÂU SỬ DỤNG KIẾN TRÚC RESNET

Nhóm 10
Đỗ Đình Thái
Đào Thành Phát
Nguyễn Minh Trí
Nguyễn Trọng Phú



INTRODUCTION

- Nhận diện biểu cảm khuôn mặt (FER) là lĩnh vực quan trọng trong AI, giúp máy tính nhận diện cảm xúc con người qua hình ảnh.
- Đề tài sử dụng học sâu với kiến trúc ResNet để tăng độ chính xác nhận diện biểu cảm. ResNet giải quyết vấn đề suy giảm độ chính xác ở mạng sâu nhờ "skip connections" (kết nối nhảy)

Mục tiêu chính của nghiên cứu

- Cải thiện độ chính xác nhận diện biểu cảm (vui, buồn, tức giận, ngạc nhiên, v.v.).
- Đánh giá hiệu quả của ResNet trong FER.
- So sánh ResNet với các kiến trúc khác (VGG, AlexNet, CNN).

Điều kiện	Phương pháp	Ưu điểm	Nhược điểm
Ánh sáng tốt	ResNet	Độ chính xác cao	Yêu cầu phần cứng mạnh
Dữ liệu đa dạng	ResNet + Data Aug.	Tăng khả năng tổng quát	Dữ liệu mất cân bằng
Dữ liệu hạn chế	Transfer Learning	Tận dụng mô hình pretrained	Phụ thuộc vào dữ liệu gốc

GIẢI PHÁP (SOLUTION)

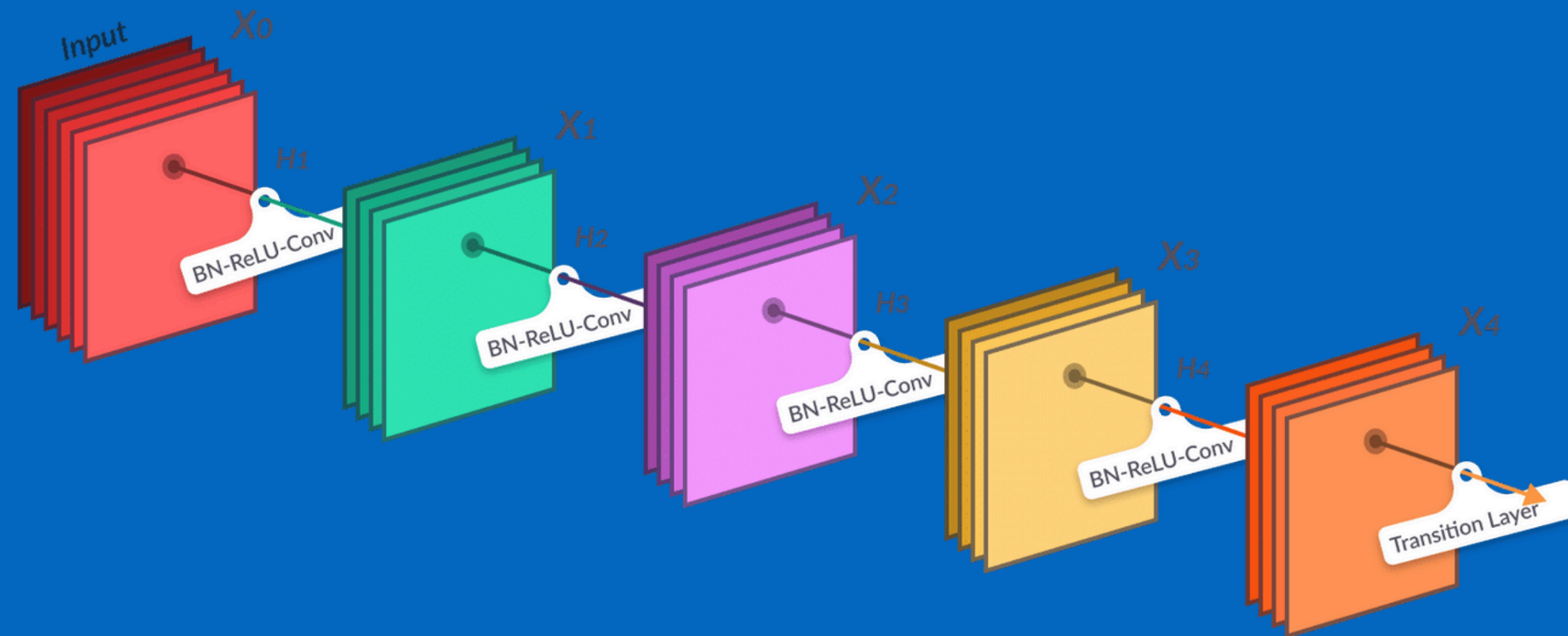
Kiến trúc ResNet

- ResNet (Residual Network) được sử dụng làm nền tảng chính trong nghiên cứu này.
- ResNet sử dụng các "skip connections" để cho phép gradient truyền ngược dễ dàng hơn, giúp huấn luyện các mạng sâu (có thể lên đến hàng trăm lớp) mà không gặp vấn đề vanishing gradient (gradient biến mất).
- Các phiên bản phổ biến của ResNet bao gồm ResNet18, ResNet50, ResNet101, v.v.

GIẢI PHÁP (SOLUTION)

- Cấu trúc cơ bản của ResNet:

- + Mỗi khối residual bao gồm các lớp convolution, batch normalization, và ReLU activation.
- + Skip connections: Bỏ qua một số lớp và kết nối trực tiếp đầu vào với đầu ra, giúp giảm thiểu mất mát thông tin.



Công thức: $y = F(x) + x$, trong đó $F(x)$ là hàm học được qua các lớp, x là đầu vào

- Ứng dụng trong FER:

- + ResNet được huấn luyện trên tập dữ liệu khuôn mặt (ví dụ: FER2013, CK+, RAF-DB) để nhận diện các biểu cảm.
- + Sử dụng transfer learning: Tập dụng ResNet đã được huấn luyện trước trên tập dữ liệu lớn như ImageNet, sau đó tinh chỉnh (fine-tune) trên dữ liệu khuôn mặt.

GIẢI PHÁP (SOLUTION)

Tăng cường dữ liệu (Data Augmentation)

Để cải thiện khả năng tổng quát hóa của mô hình, các kỹ thuật tăng cường dữ liệu được áp dụng:

- Xoay, lật, thay đổi độ sáng/tương phản của hình ảnh.
- Cắt ngẫu nhiên (random cropping) để mô hình học được các đặc trưng từ nhiều góc độ.
- Thêm nhiễu (noise) để tăng khả năng chống chịu với dữ liệu thực tế.

CÔNG TRÌNH LIÊN QUAN (RELATED WORK)

Tác giả	Mục tiêu	Kỹ thuật	Kết quả	Tập dữ liệu	Độ chính xác
Zhang et al. (2021)	Cải thiện độ chính xác	ResNet50 + Attention	Tăng độ chính xác	FER2013, CK+	73.5%, 95.2%
Liu et al. (2022)	Điều kiện thực tế	ResNet18 + Data Aug	Hiệu quả tốt	RAF-DB	85.6%
Kim et al. (2023)	Tối ưu cho thiết bị	ResNet18 + Pruning	Giảm kích thước	FER2013	70.8%
Tran et al. (2024)	Xử lý dữ liệu mất cân bằng	ResNet50 + Focal Loss	Cải thiện biểu cảm hiếm	FER2013, RAF-DB	74.2%, 87.3%

Xu hướng nghiên cứu (2020-2025):

- Tập trung vào cải thiện độ chính xác trên các tập dữ liệu thực tế (real-world datasets).
- Tối ưu hóa mô hình học sâu để triển khai trên thiết bị có tài nguyên hạn chế (như điện thoại, camera an ninh).
- Giải quyết các thách thức như dữ liệu mất cân bằng, ánh sáng yếu, và thiên lệch (bias) trong dữ liệu.

HẠN CHẾ (LIMITATIONS)

- Hiệu suất giảm trong ánh sáng yếu, cần kỹ thuật tiền xử lý.
- ResNet yêu cầu phần cứng mạnh, không phù hợp thiết bị nhúng.
- Dữ liệu mất cân bằng gây hiệu suất không đồng đều.
- Nhiều từ môi trường thực tế (che khuất, góc chụp).

The background is a solid blue color. It is decorated with a repeating pattern of circular icons. Each icon depicts a stylized landscape with green rolling hills at the bottom, a light blue sky in the middle, and a single white cloud at the top. The icons are arranged in a grid-like pattern, with some being larger than others, creating a sense of depth and repetition.

*Thank
You*