

ĐẠI HỌC BÁCH KHOA HÀ NỘI
LỚP B2CQ-CNTT03-K68

ỨNG DỤNG AI AVATAR THEO CẢM XÚC

Người thực hiện: Hoàng Kiên Quyết

Hà Nội 2024

ĐẠI HỌC BÁCH KHOA HÀ NỘI
LỚP B2CQ-CNTT03-K68

ỨNG DỤNG AI AVATAR THEO CẢM XÚC

Người thực hiện: Hoàng Kiên Quyết

NGƯỜI THỰC HIỆN

Hoàng Kiên Quyết

MỤC LỤC

NỘI DUNG BÁO CÁO	1
I. TỔNG QUAN ỨNG DỤNG	1
1.1. Giới thiệu ứng dụng	1
1.2. Mục tiêu ứng dụng	1
1.3. Phạm vi của ứng dụng.....	1
1.4. Yêu cầu đặt ra	1
II. THIẾT KẾ VÀ XÂY DỰNG ỨNG DỤNG	2
2.1. Công nghệ sử dụng	2
2.2. Cấu trúc ứng dụng.....	4
2.3. Quá trình xây dựng phần mềm.....	5
III. THỬ NGHIỆM VÀ ĐÁNH GIÁ.....	8
TÀI LIỆU THAM KHẢO	9

MỤC LỤC HÌNH ẢNH

Hình ảnh 1: Sơ đồ cấu trúc ứng dụng	4
Hình ảnh 2: Camera Component.....	5
Hình ảnh 3: Giao diện ứng dụng ở chế độ phát triển.....	6
Hình ảnh 4: Kết quả thử nghiệm.....	8

NỘI DUNG BÁO CÁO

I. TỔNG QUAN ỨNG DỤNG

1.1. Giới thiệu ứng dụng

Tên ứng dụng: Avatar theo cảm xúc

Phiên bản: 1.0.0

Mô tả: Ứng dụng cho phép tạo avatar hoạt họa cho người dùng bằng camera trong thời gian thực.

1.2. Mục tiêu ứng dụng

Mục tiêu của ứng dụng là sử dụng AI để nhận diện và biểu đạt cảm xúc của người dùng thông qua một avatar. Ứng dụng dựa trên mô hình nhận diện cảm xúc của thư viện face-api.js, một công cụ mã nguồn mở dành cho nhận diện khuôn mặt và phân tích cảm xúc. Khi người dùng thể hiện cảm xúc trên khuôn mặt, AI sẽ phân tích biểu cảm và phản hồi bằng cách thay đổi biểu cảm của avatar sao cho phù hợp. Ứng dụng cho phép những người hướng nội, ngại giao tiếp có thể thể hiện cảm xúc của mình thông qua avatar.

1.3. Phạm vi của ứng dụng

Ứng dụng nghiên cứu trong phạm vi thử nghiệm tính năng của AI cụ thể là model nhận diện cảm xúc có sẵn của thư viện face-api.js.

Giao diện được xây dựng dựa trên thư viện ReactJs.

Triển khai: Ứng dụng có thể triển khai trên bất kỳ máy chủ web, hosting nào. Phù hợp cho hầu hết các trình duyệt như: chrome, firefox, safari,...

1.4. Yêu cầu đặt ra

Ứng dụng phải có khả năng tương tác trong thời gian thực, có độ chính xác cao trên 85%.

II. THIẾT KẾ VÀ XÂY DỰNG ỨNG DỤNG

2.1. Công nghệ sử dụng

2.1.1. Các công nghệ sử dụng

Model AI được lấy từ thư viện Face API.js. Face API.js là một thư viện JavaScript mã nguồn mở, cung cấp các công cụ để nhận diện và phân tích khuôn mặt trong trình duyệt. Được xây dựng trên nền tảng TensorFlow.js, Face API.js hỗ trợ nhận diện khuôn mặt, phát hiện đặc điểm trên khuôn mặt, ước lượng cảm xúc và nhận diện danh tính. Thư viện này phù hợp cho các ứng dụng web tương tác, vì nó hoạt động trực tiếp trên trình duyệt mà không cần gửi dữ liệu đến máy chủ, giúp bảo vệ quyền riêng tư và tăng tốc độ xử lý.

Avatar được tạo ra dựa trên Dice Avatar: Là một công cụ tạo avatar ngẫu nhiên dựa trên hình ảnh xúc xắc (dice). Nó sử dụng các hình dạng và màu sắc khác nhau để tạo ra các avatar độc đáo, thường có phong cách đơn giản và dễ nhận diện. Dice Avatar thường được dùng trong các ứng dụng, trò chơi hoặc diễn đàn để gán avatar tạm thời cho người dùng một cách ngẫu nhiên, giúp tạo ra sự đa dạng mà không cần người dùng tự thiết kế hoặc tải lên hình ảnh cá nhân.

Xây dựng giao diện dựa trên React JS: Face API.js là một thư viện JavaScript mã nguồn mở, cung cấp các công cụ để nhận diện và phân tích khuôn mặt trong trình duyệt. Được xây dựng trên nền tảng TensorFlow.js, Face API.js hỗ trợ nhận diện khuôn mặt, phát hiện đặc điểm trên khuôn mặt, ước lượng cảm xúc và nhận diện danh tính. Thư viện này phù hợp cho các ứng dụng web tương tác, vì nó hoạt động trực tiếp trên trình duyệt mà không cần gửi dữ liệu đến máy chủ, giúp bảo vệ quyền riêng tư và tăng tốc độ xử lý.

Ngôn ngữ lập trình xử dụng Typescript: TypeScript là một ngôn ngữ lập trình mã nguồn mở, được phát triển bởi Microsoft, là một superset của JavaScript. TypeScript bổ sung hệ thống kiểu tĩnh (static typing) và các tính năng hiện đại cho JavaScript, giúp phát hiện lỗi sớm và cải thiện khả năng bảo

trì mã. Các tệp TypeScript (.ts) khi biên dịch sẽ chuyển thành mã JavaScript thuần để chạy trên mọi trình duyệt hay môi trường hỗ trợ JavaScript.

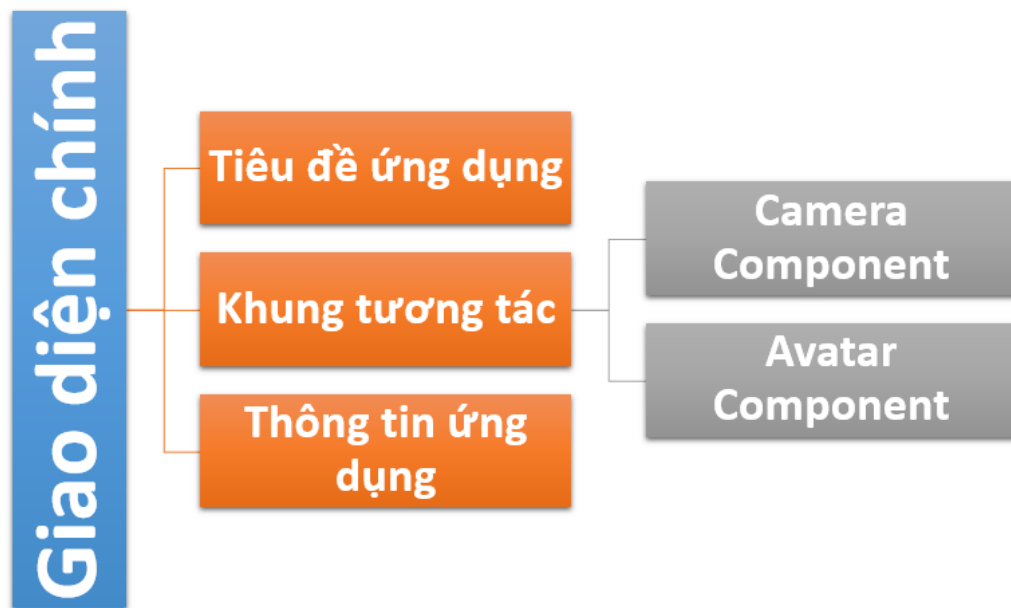
2.1.2. Quy trình phát triển và công cụ hỗ trợ

Xây dựng các component CameraComponent và AvatarComponent sau đó tích hợp vào App component để hoàn thiện ứng dụng.

Công cụ lưu trữ mã nguồn Github: Là một nền tảng lưu trữ mã nguồn và quản lý phiên bản dựa trên Git, cho phép các lập trình viên lưu trữ, chia sẻ và hợp tác phát triển phần mềm. Ngoài ra, GitHub cung cấp các công cụ để theo dõi thay đổi, quản lý dự án, và hỗ trợ làm việc nhóm, giúp dễ dàng quản lý và phát triển các dự án mã nguồn mở cũng như các dự án cá nhân.

Công cụ xây dựng mã nguồn Visual Studio Code: Là một trình soạn thảo mã nguồn mở, miễn phí, được phát triển bởi Microsoft. VS Code hỗ trợ nhiều ngôn ngữ lập trình, tích hợp Git, và có kho tiện ích mở rộng phong phú, giúp lập trình viên tùy chỉnh và mở rộng chức năng. Nhờ giao diện thân thiện và các tính năng mạnh mẽ như IntelliSense, debug và tích hợp terminal, VS Code trở thành công cụ phổ biến cho lập trình viên trên nhiều nền tảng (Windows, macOS, Linux).

2.2. Cấu trúc ứng dụng



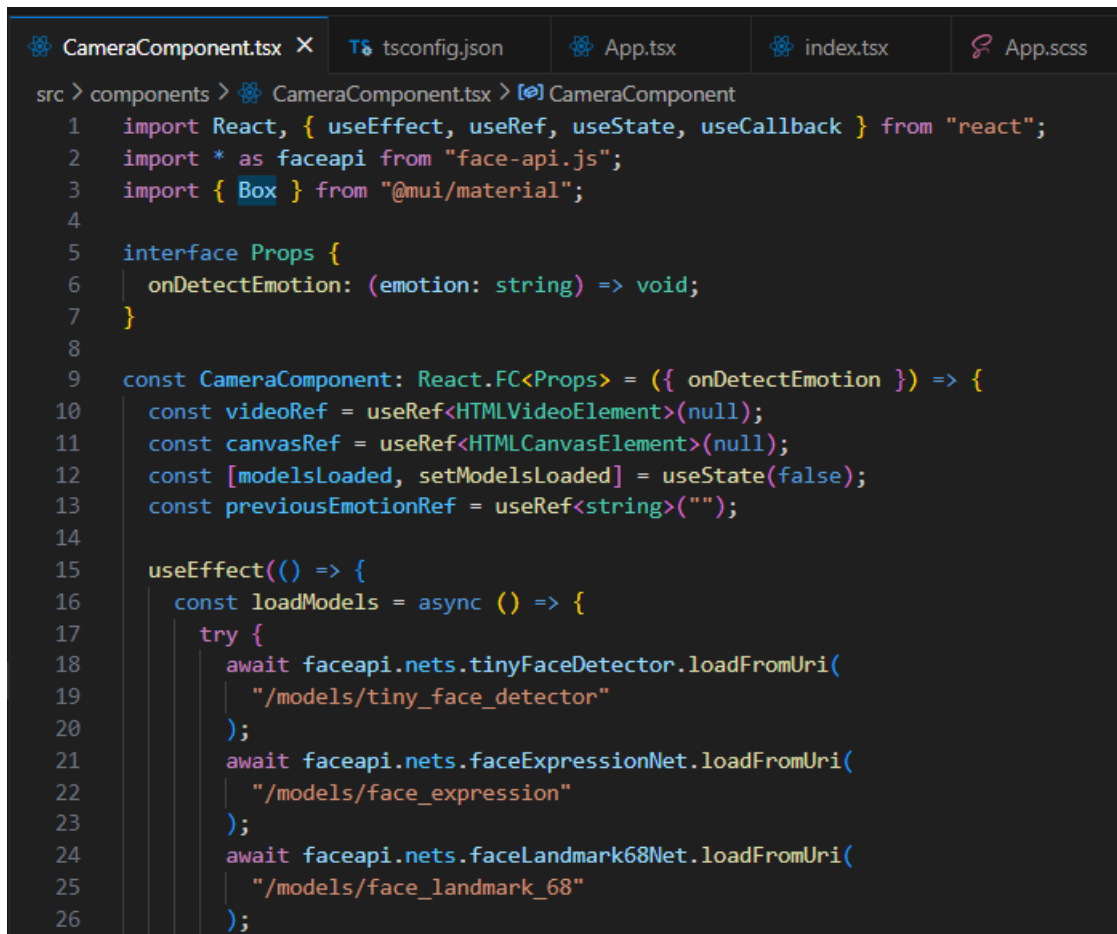
Hình ảnh 1: Sơ đồ cấu trúc ứng dụng

Ứng dụng được chia thành các thành phần chính sau:

- **Tiêu đề của ứng dụng:** Bao gồm tiêu đề và mô tả ngắn về ứng dụng.
- **Khung tương tác:** Bao gồm Camera Component và Avatar Component. Camera Component có nhiệm vụ lấy hình ảnh người dùng và nhận diện cảm xúc trong thời gian thực. Avatar Component có nhiệm vụ tạo Avatar theo cảm xúc đã nhận diện.
- **Thông tin ứng dụng:** Vị trí tại chân trang ghi thông tin của ứng dụng và người thực hiện.

2.3. Quá trình xây dựng phần mềm

2.3.1. Xây dựng Camera Component



```
src > components > CameraComponent.tsx > CameraComponent
1  import React, { useEffect, useRef, useState, useCallback } from "react";
2  import * as faceapi from "face-api.js";
3  import { Box } from "@mui/material";
4
5  interface Props {
6    onDetectEmotion: (emotion: string) => void;
7  }
8
9  const CameraComponent: React.FC<Props> = ({ onDetectEmotion }) => {
10    const videoRef = useRef<HTMLVideoElement>(null);
11    const canvasRef = useRef<HTMLCanvasElement>(null);
12    const [modelsLoaded, setModelsLoaded] = useState(false);
13    const previousEmotionRef = useRef<string>("");
14
15    useEffect(() => {
16      const loadModels = async () => {
17        try {
18          await faceapi.nets.tinyFaceDetector.loadFromUri(
19            "/models/tiny_face_detector"
20          );
21          await faceapi.nets.faceExpressionNet.loadFromUri(
22            "/models/face_expression"
23          );
24          await faceapi.nets.faceLandmark68Net.loadFromUri(
25            "/models/face_landmark_68"
26          );
```

Hình ảnh 2: Camera Component

Đầu tiên tải các model cần thiết vào thư mục public/models sau đó load các model có tên: tiny_face_detector, face_landmark_68 và face_landmark_68.

Tạo phần tử HTML tiếp nhận nguồn video đầu vào, xử lý bằng model AI để nhận diện cảm xúc và vẽ lại sơ đồ khuôn mặt theo thời gian thực bằng canvas. Đồng thời Camera Component nhận tham số đầu vào là hàm onDetectEmotion để truyền cảm xúc nhận diện được sang Avatar Component.

2.3.2. Xây dựng Avatar Component

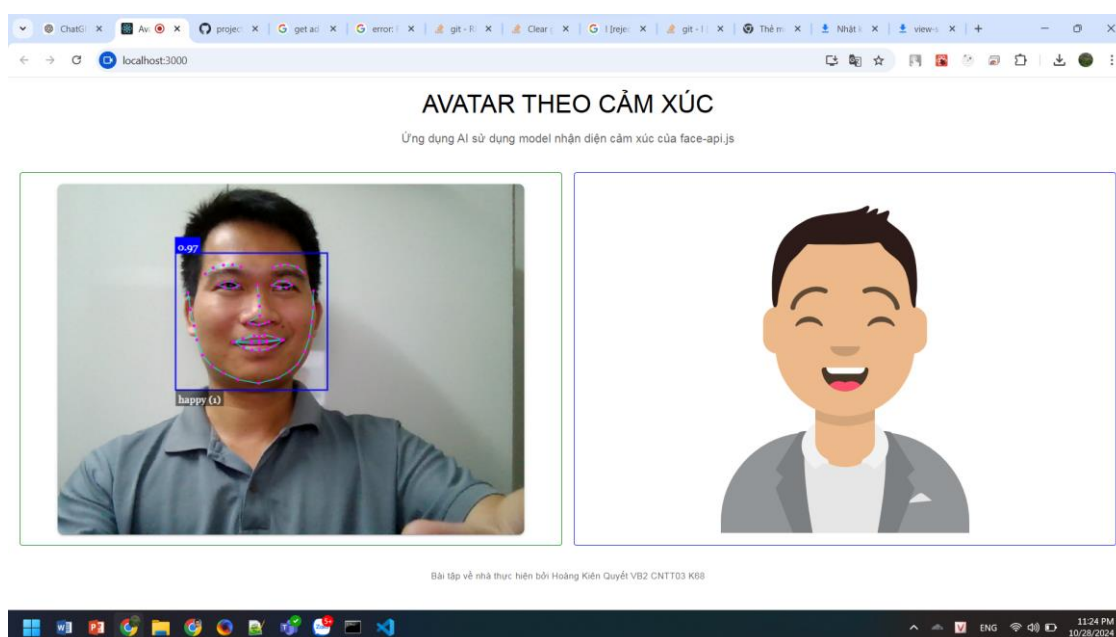
Avatar Component sẽ nhận tham số đầu vào là emotion với các giá trị như: angry, happy, sad,... tương ứng với mỗi giá trị của cảm xúc sẽ là avatar có tâm trạng tương ứng dưới dạng ảnh SVG.

2.3.3. Hoàn thiện ứng dụng

Trong App Component tiến hành tạo phần tử thông tin tiêu đề, thông tin ứng dụng (chân trang), lồng ghép Camera Component và Avatar Component trong khung tương tác, bố cục cân đối các phần tử.

Sau khi ứng dụng hoàn tất kết quả đạt được như sau:

Giao diện ứng dụng ở chế độ phát triển



Hình ảnh 3: Giao diện ứng dụng ở chế độ phát triển

Upload mã nguồn lên Github (địa chỉ mã nguồn: <https://github.com/hkienquyet/face2avatar>) để lưu trữ và chia sẻ ứng dụng. Build ứng dụng, đưa ứng dụng đã build lên địa chỉ <https://mausosanh.com/face2avatar> để tiếp cận đến người dùng.

2.3.4. Hướng phát triển và ứng dụng

Trong kỷ nguyên số, trải nghiệm khách hàng được cá nhân hóa không chỉ giúp cải thiện chất lượng dịch vụ mà còn thúc đẩy sự gắn kết và hài lòng của khách hàng. Ứng dụng công nghệ nhận diện cảm xúc vào các cuộc trò chuyện trực tuyến với khách hàng là một bước tiến mới, cho phép nhân viên chăm sóc và khách hàng thấy được phản ứng cảm xúc của nhau qua các avatar hoạt họa, dựa trên phân tích biểu cảm khuôn mặt thời gian thực. Đồng thời, giao tiếp

thông qua avatar cũng đảm bảo tính riêng tư nhất định cho khách hàng thay vì sử dụng hình ảnh thật của bản thân. Ngoài ra, tính năng ghi nhận lịch sử cảm xúc sẽ giúp doanh nghiệp phân tích và đánh giá hiệu quả của mỗi cuộc trò chuyện, từ đó tối ưu hóa quy trình tư vấn và chăm sóc.

Quy trình đưa module avatar theo cảm xúc vào chức năng chat của doanh nghiệp:

1. Phân tích và Thiết kế yêu cầu hệ thống:

- Đánh giá các yêu cầu về phần cứng và phần mềm cần thiết để xử lý video thời gian thực.
- Thiết kế UI/UX cho chức năng chat, đảm bảo module avatar tương thích với giao diện trò chuyện.

2. Tích hợp Module Tạo avatar theo cảm xúc:

- Từ ứng dụng avatar theo cảm xúc tạo Module hiển thị avatar trên khung chat (không hiển khuôn mặt được quay từ camera để đảm bảo quyền riêng tư).
- Thêm nút bật, tắt sử dụng avatar khi trò chuyện (giá trị mặc định là bật camera).

3. Ghi nhận và Phân tích Lịch sử Cảm xúc:

- Phát triển hệ thống ghi nhận cảm xúc của khách hàng trong suốt cuộc trò chuyện.
- Xây dựng các báo cáo và bảng điều khiển để phân tích xu hướng cảm xúc, hỗ trợ đánh giá kết quả tư vấn bằng AI. Các công nghệ AI có thể sử dụng gồm LSTM, GRU cho phân tích chuỗi thời gian cảm xúc, Transformer như BERT và GPT để phân tích cảm xúc phức tạp, cùng với các mô hình học máy và deep learning để dự đoán xu hướng và phát hiện sự thay đổi cảm xúc.

4. Tích hợp và Kiểm thử chức năng Chat:

- Kiểm thử tích hợp toàn bộ module vào hệ thống chat của doanh nghiệp.
- Chạy thử nghiệm thực tế để đánh giá độ mượt mà, độ chính xác và mức độ tương thích của module.

5. Đào tạo và Triển khai:

- Đào tạo nhân viên sử dụng module avatar và hướng dẫn cách phản hồi dựa trên biểu cảm của khách hàng.
- Triển khai hệ thống và tiếp tục theo dõi, cải tiến dựa trên phản hồi từ cả khách hàng và nhân viên.

Với quy trình này, doanh nghiệp có thể tận dụng công nghệ AI để nâng cao trải nghiệm tương tác, giúp khách hàng cảm thấy được thấu hiểu và chăm sóc chu đáo hơn.

III. THỬ NGHIỆM VÀ ĐÁNH GIÁ

Ứng dụng được triển khai tại địa chỉ <https://mausosanh.com/face2avatar> cho phép truy cập từ mọi nơi, mọi thiết bị thông qua trình duyệt web. Thử nghiệm với các trình duyệt: Chrome, safari, firefox, cốc cốc, microsoft edge đều hoạt động tốt trong thời gian thực.

STT	Loại cảm xúc	Số lần kiểm tra	Số lần chính xác	Tỉ lệ chính xác (%)
1	Neutral	100	100	100
2	Happy	100	97	97
3	Sad	100	80	80
4	Angry	100	81	81
5	Surprised	100	91	91
Tổng kết		500	449	89.8

Hình ảnh 4: Kết quả thử nghiệm

Thử nghiệm đối với từng loại cảm xúc khác nhau, mỗi loại tiến hành thử nghiệm 100 lần, kết quả nhận diện và thể hiện cảm xúc bằng avatar chính xác ở mức 89.8%, đáp ứng yêu cầu nghiên cứu đề ra.

TÀI LIỆU THAM KHẢO

1. Trang <https://www.wikipedia.org/>, Bách khoa toàn thư mở đa ngôn ngữ;
2. Trang <https://nodejs.org/docs/latest/api/>, tài liệu chính thức của NodeJS;
3. Trang <https://react.dev/learn>, tài liệu chính thức của ReactJS.