CMPE 491
Senior Project 1

Coffee Machine Assistance for the Visually Impaired: A Computer Vision-Based Mobile Application

Project Supervisor
Dr. Hakki Gökhan Ilk

Project Team Members
Murad Huseynov
Bedir Esen
Hazem Mustafa
Alperen Dalgıç

Course Coordinator
Dr. Gökçe Nur Yilmaz

Table Of Contents

Contents

1. Introduction

Visually impaired individuals face numerous challenges in their daily lives, especially when interacting with modern electronic devices and appliances that heavily rely on visual interfaces. One such appliance that presents significant accessibility barriers is the coffee machine. Traditional coffee machines with mechanical buttons and simple interfaces have been gradually replaced by touch-sensitive panels, digital displays, and multifunctional controls, which are designed primarily for sighted users. This shift, while enhancing functionality and aesthetic appeal, has unintentionally excluded a large segment of the population who rely on tactile feedback and non-visual cues to operate devices independently.

Despite the advancements in assistive technologies, many visually impaired individuals still struggle with everyday tasks such as making coffee, which most people take for granted. Existing accessibility solutions, such as Braille labels, third-party assistance, and voice-controlled smart appliances, provide some level of support but are often limited in scope and effectiveness. For instance, Braille labels require users to memorize locations, and voice assistants, while helpful, do not always integrate seamlessly with standalone appliances. Additionally, many smart coffee machines are expensive, limiting their accessibility to a wider audience.

To address these challenges, this project proposes a computer vision-based mobile application that enables visually impaired individuals to operate coffee machines independently and efficiently. By leveraging deep learning and real-time object recognition, the application will allow users to interact with their coffee machines using a smartphone camera and receive step-by-step auditory guidance. Unlike traditional accessibility solutions, this approach does not require modifications to the coffee machine, making it a more practical and scalable solution.

The system will utilize EfficientDet, a state-of-the-art deep learning-based object detection model, to recognize and label various components of a coffee machine, such as buttons, levers, and displays. Once identified, the system will generate real-time voice feedback, instructing the user on which button to press or what action to take next. This will be powered by Android and iOS Text-to-Speech (TTS) engines to ensure clear, responsive auditory instructions. The application will be developed using React Native, a popular cross-platform framework, ensuring that it is accessible on both Android and iOS devices.

A key advantage of this approach is its adaptability to different coffee machine models. While the initial implementation will focus on a specific model, the underlying methodology can be extended to other coffee machines through additional training of the deep learning model. This scalability ensures that the solution can benefit a wide range of users, making coffee preparation an accessible and enjoyable experience for individuals with visual impairments.

Beyond the technical aspects, this project also aims to address broader issues related to digital accessibility and inclusivity. By demonstrating the feasibility of AI-driven accessibility solutions, it encourages manufacturers to consider universal design principles in future appliances. Furthermore, the project highlights the role of emerging technologies in enhancing the independence and quality of life of visually impaired individuals.

This document provides an in-depth discussion of the project's objectives, implementation strategy, system design, constraints, ethical considerations, and future improvements. Through this comprehensive analysis, we aim to present a practical and innovative solution to an everyday problem, ultimately promoting greater accessibility and inclusivity in modern technology.

1.1. Project Description

The core objective of this project is to develop a mobile application that empowers visually impaired individuals to interact with coffee machines through computer vision and voice guidance. This solution leverages AI-driven object detection to identify buttons, labels, and machine interfaces, providing an intuitive, accessible method for users to operate their coffee machines without requiring visual input.

The system consists of several key components that work together to create a seamless user experience:

- **Camera-based detection module:** The smartphone camera captures real-time images of the coffee machine's interface. The system processes these images to detect and recognize buttons, levers, and displays.
- **Deep learning model (EfficientDet):** A machine learning model trained to identify key components of coffee machines accurately. The model undergoes rigorous training with a dataset containing various coffee machine layouts, ensuring robust recognition across different lighting conditions and angles.
- **Voice guidance system (Android/iOS TTS):** The application provides step-by-step auditory instructions, guiding users on which buttons to press and in what sequence. The system ensures that the feedback is clear, concise, and responsive to user actions.
- **React Native-based UI:** A cross-platform mobile application that integrates all the system's functionalities. It supports gesture-based controls and compatibility with screen readers to enhance accessibility.

Unlike existing assistive technologies, which often require physical modifications or expensive proprietary devices, this approach is cost-effective and universally applicable. The system does not require any alterations to the coffee machine, making it a flexible solution that can be adopted by a wide range of users.

Features and Functionalities

To ensure ease of use and a high degree of accessibility, the application incorporates several features designed specifically for visually impaired individuals:

1. **Automatic Coffee Machine Recognition:** The app can identify the specific model of a coffee machine, adjusting its guidance based on the recognized layout.
2. **Real-Time Object Detection:** Using EfficientDet, the system accurately detects buttons and machine components, even under varying lighting conditions.
3. **Step-by-Step Voice Instructions:** The user receives verbal instructions guiding them through the coffee-making process, including button locations and sequence of operations.

4. **Offline Functionality:** The app will work without an internet connection, ensuring accessibility in all environments.

Target Users

The primary users of this application include:

- **Visually impaired individuals:** The app will enable them to operate coffee machines independently without relying on external assistance.
- **Elderly users:** Many older adults experience declining vision and may find modern coffee machines challenging to operate. This app can serve as an assistive tool for them as well.
- **Organizations and institutions:** Disability support centers, rehabilitation programs, and accessibility-focused organizations can benefit from adopting this solution for their clients.

Scalability and Future Applications

Although the initial development will focus on a single coffee machine model, the methodology used can be expanded to support a variety of applications, such as bus stops, office and home navigation. Future versions of the application could incorporate:

- **Cloud-based updates** that allow the app to learn and adapt to new machine models dynamically.
- **Integration with smart home ecosystems** for seamless operation alongside other accessibility tools.

This project represents a significant step towards making everyday appliances more accessible to visually impaired individuals, promoting digital inclusivity through the power of CV and mobile technology. By addressing a common yet overlooked problem, the proposed solution has the potential to enhance independence and improve the quality of life for millions of people worldwide.

2. Constraints

2.1. Implementation Constraints

As we develop this system, we will make sure that real-time processing is achieved so that visually impaired users receive immediate feedback while operating the coffee machine. Since relying on cloud processing introduces latency and privacy concerns, we will ensure that all computations are performed directly on the mobile device using TensorFlow Lite. To make this work efficiently, we will optimize the EfficientDet model, balancing accuracy and speed to ensure smooth performance.

We will also ensure that the user interface (UI) is fully accessible. Our app will strictly follow accessibility guidelines, such as WCAG (Web Content Accessibility Guidelines), to support screen readers, gesture navigation, and other assistive features. Additionally, we will carefully synchronize the object detection module with the text-to-speech engine to make sure the voice guidance system is responsive and clear.

2.2. Hardware Constraints

Since our system relies on a smartphone camera, we will make sure that the app works well with mid-range devices that have decent cameras. We will test the application on multiple devices to ensure it performs well under various lighting conditions.

Battery consumption is another factor we will consider. Running real-time object detection can drain the battery quickly, so we will work on optimizing power efficiency. The app must also handle different environmental conditions, so we will implement image processing techniques to compensate for low-light or glare scenarios.

2.3. Software Constraints

We will make sure the app is compatible with Android 8.0+ and iOS 12+ to support a wide range of users. To achieve this, we will use React Native, allowing us to develop a single codebase for both platforms. However, we will conduct extensive testing to ensure all features work seamlessly across different devices.

Our voice guidance system will rely on the built-in text-to-speech (TTS) engines of Android and iOS. Since speech synthesis quality varies between devices, we will include adjustable settings for speech rate, volume, and language preferences so users can personalize their experience.

2.4. Health and Safety Constraints

Safety is a key priority since users will be interacting with a coffee machine that involves hot surfaces and liquids. To prevent accidents, we will implement safety warnings in the voice guidance system, such as alerting users when the coffee is ready and advising them to be cautious around heated surfaces.

2.5. Ethical and Accessibility Considerations

We are committed to protecting user privacy. Our app will not store, transmit, or share any user data. All processing will be done on the device, ensuring users' personal information remains secure.

Additionally, we want our CV Object Detection model to work fairly for all users. To prevent biases, we will train EfficientDet on a diverse dataset of coffee machines captured under different conditions. We will also gather user feedback to refine the model and improve its performance.

Since accessibility is a core focus of this project, we will make sure that every feature is designed with visually impaired users in mind. This means supporting screen readers, gesture controls, and other assistive technologies to make the app as intuitive as possible.

3. Professional and Ethical Issues

3.1. User Privacy and Data Protection

Since our system processes images of coffee machines, we will make sure that no sensitive data is stored or transmitted. We will implement all object recognition and voice guidance features locally on the device, eliminating the need for internet connectivity.

3.2. Fair AI Training and Bias Mitigation

We will train our model on a wide range of coffee machine images to ensure it works well under various lighting conditions and angles. We will continuously collect user feedback to refine the model and reduce any biases that could affect usability.

3.3. Accessibility-First Design

We will prioritize accessibility from the start. We will focus on non-visual interaction methods, making sure all features are optimized for screen readers and voice commands. We will also ensure that our voice instructions are clear and easy to follow.

4. System Requirements

4.1. User Requirements

- Users should be able to operate the coffee machine entirely through voice commands or gestures.
- The app should provide clear and intuitive voice guidance.
- The system must work offline without requiring an internet connection.

4.2. Functional Requirements

- The app must recognize and label coffee machine buttons in real time.
- The voice guidance system should provide step-by-step instructions for making coffee.
- The app should support gesture-based controls for non-visual interaction.

4.3. Non-Functional Requirements

- The system should provide real-time feedback with minimal delay.
- The app should be lightweight and optimized for mobile devices.
- The UI must follow WCAG accessibility standards.

5. System Design

5.1. Architecture Overview

Our system will have three main components:

1. **Camera Module:** Captures images of the coffee machine and processes them for analysis.
2. **Object Recognition Module:** Uses EfficientDet to identify machine buttons, displays, and labels.
3. **Voice Guidance System:** Provides real-time instructions using Android/iOS TTS engines.

5.2. Machine Learning Model Design

- We will train EfficientDet using TensorFlow on a dataset of coffee machine images.
- To ensure smooth performance, we will optimize the model for on-device processing using TensorFlow Lite.
- We will use data augmentation techniques to improve accuracy under different conditions.

6. Development and Implementation

6.1. Development Tools & Technologies

- **React Native:** To build a cross-platform mobile app.
- **TensorFlow Lite & OpenCV:** For object detection and image processing.
- **Android/iOS TTS Engines:** To provide voice guidance.

6.2. Dataset Collection and Training

- We will collect images of different coffee machines under various lighting conditions.
- We will apply data augmentation techniques to improve the model's ability to recognize buttons and labels.

7. Testing and Evaluation

7.1. Performance Testing

- We will test model accuracy, response time, and efficiency on various devices.
- We will ensure the app provides real-time feedback with minimal delay.

7.2. Usability Testing with Visually Impaired Users

- We will conduct real-world testing with visually impaired participants.
- We will collect feedback from users to refine the system and improve accessibility features.

8. Limitations and Future Work

8.1. Current Limitations

- The initial version of the app will support only one coffee machine model.
- The voice feedback system will be predefined and not fully conversational.

8.2. Future Improvements

- We plan to expand support for multiple coffee machine models.
- We will integrate additional accessibility features based on user feedback.

9. References

1. Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.
2. Tan, M., Pang, R., & Le, Q. V. (2020). *EfficientDet: Scalable and Efficient Object Detection*. IEEE Conference on Computer Vision and Pattern Recognition (CVPR).
3. Krishnan, R. (2021). *Computer Vision-Based Assistive Technologies for the Visually Impaired*. Springer.
4. Lazar, J., Feng, J. H., & Hochheiser, H. (2017). *Research Methods in Human-Computer Interaction*. Morgan Kaufmann.
6. TensorFlow Lite Documentation: https://www.tensorflow.org/lite
7. React Native Accessibility Guide: https://reactnative.dev/docs/accessibility
8. World Wide Web Consortium (W3C): *Web Content Accessibility Guidelines (WCAG)*
    - https://www.w3.org/WAI/standards-guidelines/wcag/