A
Synopsis
On

Web Meeting Suite

Submitted in partial fulfillment of the requirements for
completion of AI Lab
Of
TY COMP
In
Computer Engineering
by

Sarvesh Anand Mankar 142203013
Avdhut Aakash Kamble 142203010

Under the guidance of

Prof. Suraj Sawant
Department of Computer Engineering

Department of Computer Engineering
COEP Technological University (COEP Tech)
(A Unitary Public University of Gov. of Maharashtra)
Shivajinagar, Pune – 411005, Maharashtra, India

31st October 2023

Introduction and Motivation:

A decade ago, people used to work at their offices and have their meetings and conferences physically, but due to the pandemic, people were forced to work out of their homes, which demanded them to come up with new ways of interacting with their coworkers. While specific industries required a visual exchange of information, most could get by with a few voice conversations, so we started employing video conferencing using Zoom, Google Meet, and other platforms. Most of the issues were resolved by this, but one remained. Everyone on the call had to turn on their webcams so that others could see them because people liked to look at each other as they were conversing. A plethora of issues arise when cameras are enabled. One major problem is the security of people or things in the environment.

The webcam may record private information in the background during a video conference between staff members and a third party, or it may record an individual who wishes to remain anonymous. Video calls in the home raise additional security concerns. These might be intimate and private even if they aren't secret. It could be individuals or personal belongings. In this section, we'll talk about safeguarding others' identities when holding online meetings. Although we frequently discuss the "work from home" scenario, this is certainly applicable in various contexts.

When working from home, other environments needed to exercise caution when approaching someone on a video conversation for several reasons, including protecting their identity and not disturbing the other participants. Until now, the user had to be conscious of their surroundings, ask people to stay away from them and switch off the video when they walked in. This frequently causes the user to lose focus on the conference. Our solution handles this automatically, allowing the user to concentrate on the call without getting sidetracked. When the technology senses movement, the background is automatically blurred till the user leaves. As an added precaution, it pixelates their face as well. All of this takes place without compromising the user's image.

Abstract:

A number of people have started working from home as a result of the pandemic. And one of the main concerns these days is privacy. In contrast to business settings, residences and other public areas are populated by people who would like not to inadvertently be photographed during video chats. They want to preserve their identity for a variety of reasons. Therefore, we created a system that, when a video call is in progress, instantly recognizes and censors any unknown individuals that enter the frame. The result was a seamless background operation that doesn't interfere with the topic or other participants in the video conversation and doesn't significantly reduce the video output's frame rate or CPU burden. Now, anyone can walk near a person having a video call without having to worry about their identity being revealed.

Problem Statement:

The increasing use of video conferencing and online meeting tools has created new privacy and security problems. Attendees in virtual meetings run the risk of unintentionally disclosing private information or running into unwelcome visitors. In a variety of settings, including private video conferences and professional remote business meetings, these problems can be very troubling. Our project intends to create a robust solution that automatically protects the identities and privacy of participants during online sessions without interfering with their experience to solve these privacy and security concerns.

Objective:

Automatic Privacy Preservation: Develop an AI-based system that can automatically detect and respond to movement and unwanted intrusions during online meetings, safeguarding the privacy of participants.

- Face and Object Blurring: Implement advanced image processing techniques to blur faces and objects in the background, when necessary, while keeping the designated face in sharp focus.

- Real-time Monitoring: Ensure that the system operates in real-time, offering a seamless and uninterrupted meeting experience for users.

- User-Friendly Interface: Create an intuitive and user-friendly interface that allows participants to easily initiate and customize the privacy protection features as needed.

- Performance Metrics: Measure the accuracy, precision, recall, and Kappa coefficient of the developed system to ensure it effectively detects and responds to privacy concerns.

- Dataset Integration: Incorporate relevant datasets or sources to train and validate the system, enhancing its ability to recognize privacy threats.

- Advantages: Identify the advantages of the developed solution, such as the automatic handling of privacy concerns and improved user focus during online meetings.

- Limitations and Ethical Considerations: Analyze the limitations of the system, including hardware requirements, and address ethical considerations related to privacy preservation and potential misuse.

- Future Expansion: Explore the potential for extending the system's functionality to different contexts and applications, ensuring its adaptability and relevance beyond online meetings.

Methodology:

1. Data Collection:
   - Collect a diverse dataset of images and videos containing multiple faces, with annotations for face recognition and segmentation.

2. Preprocessing:
   - Face detection: Use Haar cascades or deep learning-based models like MTCNN to detect faces in images and frames of videos.
   - Data augmentation: Apply geometric and photometric transformations to augment the dataset.

3. Face Recognition:
   - Algorithm: Implement Local Binary Patterns (LBP) for face recognition, as described in the paper by Ahonen et al.

4. Face Segmentation:
   - Algorithm: Implement the Mask R-CNN architecture, as detailed in the paper by He et al.

5. Super-Resolution:
   - Algorithm: Employ an Efficient Sub-Pixel Convolutional Neural Network for enhancing the quality of the designated face, following the approach in the paper by Ledig et al.

6. Blurring:
   - Algorithm: Apply Gaussian blur or other suitable methods to the background faces for privacy preservation.

7. Integration:
   - Develop a user-friendly interface that accepts images or video streams and allows users to specify the face to keep in focus.

8. Testing and Evaluation:
   - Evaluate the accuracy of face recognition and segmentation.
   - Measure the effectiveness of the super-resolution technique.
   - Assess the overall quality of the blurred background.

Hardware and Software Requirements:

- Programming Languages: Python
- Libraries and Frameworks: OpenCV, Numpy, Face_Recognition, dlib
- Hardware: A GPU-accelerated system for faster deep learning processing.
- Storage: Store datasets and trained models on cloud or local storage.
- Deployment: Choose a cloud platform for hosting the tool or deploy it on local servers as needed.

By implementing this plan, we aim to create a powerful AI tool capable of blurring background faces while maintaining the sharp focus on a designated face, enhancing privacy, and enabling creative applications in various domains.

Future Work:

We can expand on this concept of censorship to include voice. It is possible that a private conversation is taking place in the background, and it would not be secure to record it while on a video call. Similar to this, in a home setting, people may be having private or intimate talks that would be perfect for others to overhear over an internet call. Therefore, we devise a comparable strategy. To differentiate between the user's voice and the speech of an unknown individual, we preload the user's voice profile and employ voice diarisation. The basic concept is to mute the microphone upon detection of an unknown voice profile. Regretfully, real-time voice diarisation remains in the research stage and is not yet feasible for efficient implementation.

Reference:

1. Face Recognition with Local Binary Patterns (LBP)
   - Author: Timo Ahonen, Abdenour Hadid, and Matti Pietikäinen
   - Paper: "Face Recognition with Local Binary Patterns"
   - Citation: Ahonen, T., Hadid, A., & Pietikäinen, M. (2004). Face recognition with local binary patterns. In European conference on computer vision (pp. 469-481).
   - This paper introduces LBP-based face recognition, a fundamental component of our project.

2. Real-Time Single Image and Video Super-Resolution Using an Efficient Sub-Pixel Convolutional Neural Network
   - Author: Christian Ledig, Lucas Theis, et al.
   - Paper: "Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network."
   - Citation: Ledig, C., Theis, L., et al. (2017). Photo-realistic single image super-resolution using a generative adversarial network. In Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR) (pp. 4681-4690).
   - This paper presents super-resolution techniques that enhance the quality of the focused face.

3. Mask R-CNN
   - Author: Kaiming He, Georgia Gkioxari, et al.
   - Paper: "Mask R-CNN"
   - Citation: He, K., Gkioxari, G., et al. (2017). Mask R-CNN. In Proceedings of the IEEE international conference on computer vision (ICCV) (pp. 2961-2969).
   - We will incorporate the Mask R-CNN architecture for accurate face segmentation.