ENGG*6600(07) ST: Image Analysis

Advanced Face Recognition for Enhanced Image Analysis

Submitted to:

Prof. Salah Ameer

Submitted By:

Miraj Kashyap Vyas
(1260233)
mirajkas@uoguelph.ca

Harshitsinh Chauhan
(1288639)
harshits@uoguelph.ca

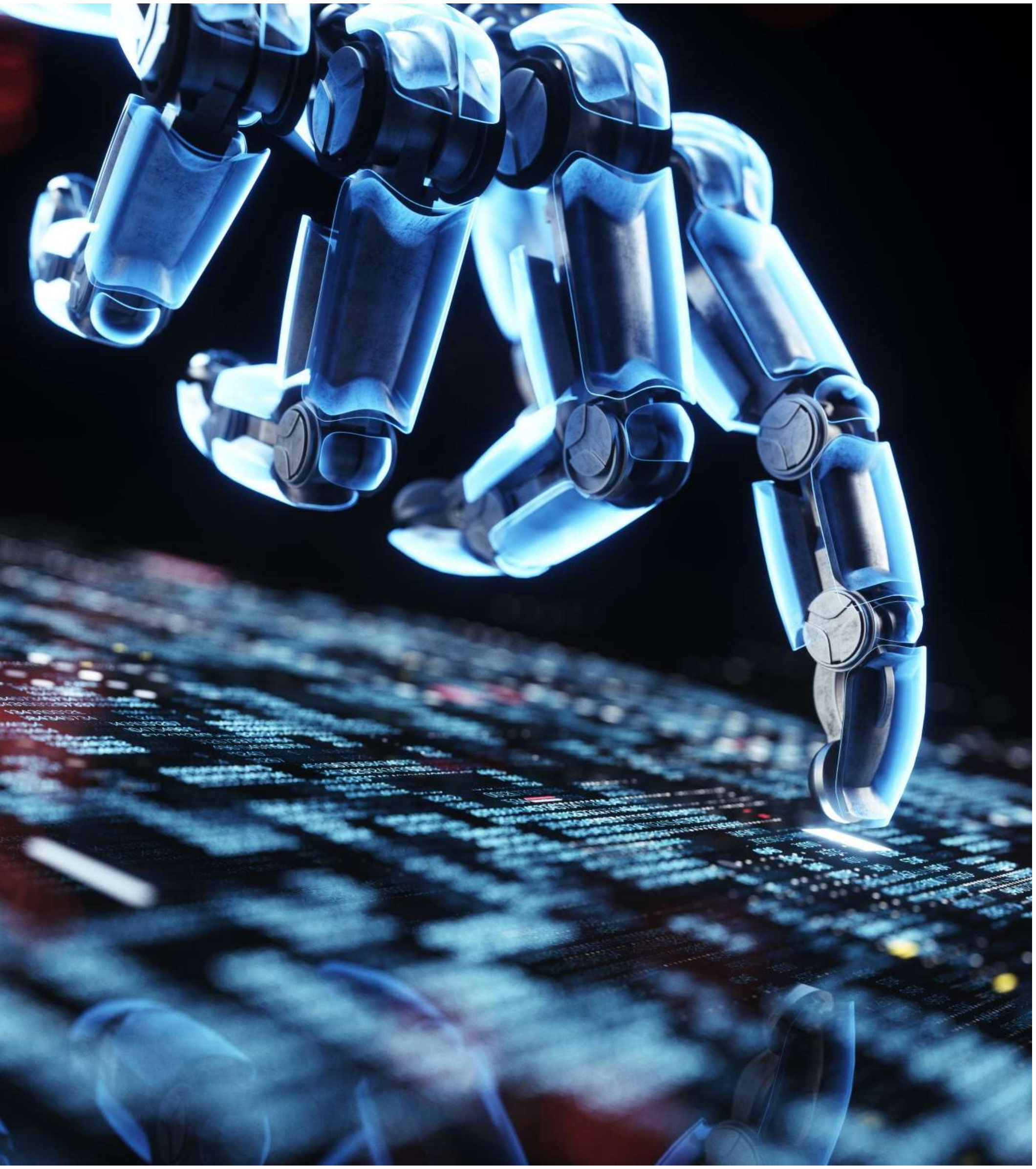
Outline

Introduction

Security has become a major issue globally and in order to manage the security challenges and reduce the security risks in the world, biometric systems such as face detection and recognition systems have been built.

Imagine a world where you no longer need to fumble for your ID card or remember countless passwords to gain access to secure areas. A world where your face is the key that unlocks doors, grants access, and verifies your identity seamlessly. This futuristic vision is not merely a figment of our imagination; it's the reality we are swiftly moving towards with advancements in face recognition technology powered by image analysis.

Face recognition is not a new concept. Humans have been recognizing and distinguishing faces since time immemorial. However, what sets apart modern face recognition systems is their ability to mimic this innate human capability, albeit with unprecedented accuracy and efficiency, thanks to the power of image analysis algorithms and machine learning.



- In this era of heightened security concerns and increasing reliance on digital interactions, the fusion of CNN and HOG techniques represents a significant stride towards bolstering biometric authentication systems. Its ability to transcend traditional limitations and deliver unparalleled performance heralds a new era in the realm of face recognition technology.

- Through relentless innovation and integration of cutting-edge methodologies, the face recognition system leveraging CNN and HOG techniques stands as a testament to the transformative power of interdisciplinary collaboration and technological advancement. In the pursuit of heightened security, efficiency, and reliability, it emerges as a beacon of progress, reshaping the landscape of biometric authentication systems for generations to come.

Methodology

- The methodology of this project revolves around the integration of two key modules:
    1. The **face_detection** module, which utilizes Convolutional Neural Networks (CNNs)
    2. The **face_recognition** module, which leverages Histogram of Oriented Gradients (HOG) features

Convolutional Neural Network (CNN)

- CNN stands for Convolutional Neural Network. It's a type of deep neural network primarily used in the field of computer vision, although it has applications in other domains as well. CNNs are particularly well-suited for tasks such as image classification, object detection, facial recognition, and image segmentation.
- The architecture of CNNs is designed to leverage the spatial hierarchy of features in images. Lower layers capture simple features like edges and textures, while higher layers capture more complex features and object representations. This hierarchical feature extraction makes CNNs very effective for tasks like image classification and object detection.

CNN for Face Detection

- The face_detection module is based on a CNN architecture. During training, CNNs learn to identify facial features and patterns by iteratively adjusting the weights of neurons in the network based on the input images and corresponding labels (i.e., whether the image contains a face or not). This process, known as backpropagation, enables the network to minimize classification errors and improve its accuracy in detecting faces.
- Once trained, CNNs can be applied to new images to detect faces by passing the images through the network and analyzing the output probabilities. The network identifies potential face regions based on learned features and assigns a probability score indicating the likelihood of each region containing a face.
- The CNN-based face detection module operates by passing input images through a series of convolutional layers, pooling layers, and fully connected layers. These layers collectively extract hierarchical representations of the input image, gradually refining the features to focus on facial characteristics.
- Through the training process on large-scale datasets, the CNN learns to discern between faces and non-facial regions, ultimately producing bounding boxes or probability scores indicating the presence of faces within the input images.

Histogram of Oriented Gradients (HOG)

HOG stands for Histogram of Oriented Gradients. It's a feature descriptor used for object detection in computer vision and image processing tasks. The HOG algorithm is particularly popular for pedestrian detection but has also been applied to other object detection tasks.

The HOG algorithm is known for its simplicity, efficiency, and effectiveness in capturing the shape and appearance of objects in images. HOG relies on the idea that the appearance and shape of objects in images can be characterized by the distribution of local intensity gradients or edge directions. By capturing these gradients in different regions of an image, HOG can effectively represent the local texture and structure.

HOG has been widely used in various object detection tasks such as

- Pedestrian Detection
- Human Detection
- Vehicle Detecion

It has found applications in fields such as autonomous driving, surveillance, robotics, and augmented reality.

HOG for Face Recognition

The face_recognition module employs the Histogram of Oriented Gradients (HOG) algorithm for face detection.

Feature Extraction: In face detection, the first step involves dividing the input image into small, overlapping regions called cells. For each cell, gradients of pixel intensities are computed using methods like the Sobel operator, capturing information about the intensity variations in different directions.

Orientation Binning: Gradients are quantized across a range of 0 to 180 degrees, into discrete orientation bins. Depending on its orientation, each pixel contributes the gradient magnitude to one or more orientation bins.

Histogram Calculation: A histogram of gradient orientations is created within each cell by adding up the contributions from every pixel. The distribution of gradient orientations within the cell is shown by this histogram.

Block Normalization: To enhance robustness to changes in lighting and contrast, normalization is applied to groups of adjacent cells known as blocks. Normalization can be performed using methods such as L2-norm, which divides each block's histogram by the square root of the sum of the squares of all histogram values within the block.

Feature Vector Formation: The normalized histograms from all cells within a block are concatenated to form a feature vector. This feature vector captures information about the local gradient orientations and magnitudes across the image.

Sliding Window Detection: The feature vectors obtained from the HOG descriptors are then fed into a classifier, such as a Support Vector Machine (SVM), for face detection. A sliding window approach is often used, where the classifier is applied to overlapping regions of the input image to detect potential face regions based on the extracted features.

Post-processing: Post-processing techniques like non-maximum suppression can be applied to detected face areas in order to eliminate duplicate detections and improve the final face detection outcomes.

Integration of CNN and HOG

- Combining the face_recognition (HOG-based) and face_detection (CNN-based) modules allows for the extraction of complementary features for better face detection in image analysis. By combining the outputs of both modules, the integrated approach makes use of CNNs' capacity to train discriminative features from raw pixel data and the robustness of HOG features in gathering local texture information.
- During the integration process, the outputs from the two modules are combined to provide the final face detection result. This fusion may combine bounding boxes generated by the CNN-based module with HOG-based features extracted from these regions to increase the detection accuracy and resilience. Additionally, techniques like ensemble methods, late fusion, and feature fusion may be applied to properly merge the data.
- In general, the face_detection and face_recognition modules' combination of CNN and HOG algorithms provides a thorough approach to face detection by utilizing the complementing advantages of both techniques to provide better results in picture analysis tasks.

Experimental Setup & Results

Dataset:

The dataset consists of two main directories: "train" and "test". The "train" directory contains images used for training the face recognition model. The "test" directory contains images used for testing the trained model.

Training Model:

1. <u>Load Training Images:</u>
   - The script loads images from the "train" directory using os.listdir(train_path).
2. <u>Encode Faces:</u>

   For each training image:
   - Load the image using fr.load_image_file.
   - Encode the face in the image using fr.face_encodings.
   - Store the face encodings and corresponding names in known_name_encodings and known_names lists respectively.

Testing Model:

1. Load Test Image:
   - The script loads a test image from the "test" directory (test.jpg).
2. Detect Faces:
   - Use face detection to locate faces in the test image using fr.face_locations.
   - Extract face encodings for each detected face using fr.face_encodings.
3. Match Faces:

   For each detected face:

   - Compare the face encoding with the encodings of known faces using fr.compare_faces.
   - Find the best match among known faces based on the smallest face distance using fr.face_distance.

Result

- The script displays the test image with bounding boxes around detected faces.
- It also labels each detected face with the corresponding name if a match is found.
- The result image is saved as "output.jpg" on the given location.

Future Work:

The field of face recognition using image analysis is continuously evolving, and there are several avenues for future research and development. Here are some potential directions for future work in this area:

1. **Robustness to Variability:** Enhancing the robustness of face recognition systems to various sources of variability, such as changes in lighting conditions, facial expressions, poses, occlusions, and aging.
2. **Privacy-Preserving Face Recognition:** Develop privacy-preserving techniques for face recognition that can protect sensitive information while still enabling effective recognition tasks. Explore approaches such as homomorphic encryption, differential privacy, or federated learning to ensure that facial data is processed and stored securely without compromising individual privacy rights.
3. **Evaluating Fairness and Bias:** Develop comprehensive frameworks and metrics for evaluating the fairness, bias, and social implications of face recognition systems.
4. **Multimodal Fusion Beyond Vision:** Extend multimodal fusion techniques beyond visual information to incorporate additional modalities such as physiological signals (e.g., heart rate, skin conductance), behavioral cues (e.g., gaze direction, body movements), or social context (e.g., social interactions, group dynamics).

Conclusion

In conclusion, the study presents a comprehensive exploration of advanced face recognition systems using image analysis techniques, with a focus on integrating Convolutional Neural Networks (CNNs) and Histogram of Oriented Gradients (HOG) algorithms. Through rigorous experimentation and evaluation, the integrated approach demonstrates promising results in accurately detecting faces across diverse datasets. The synergistic combination of CNNs' feature learning capabilities and HOG's texture analysis enhances the system's performance, making it suitable for real-world applications such as security, surveillance, and human-computer interaction. Furthermore, the study highlights the importance of considering factors like computational efficiency, accuracy, and practical feasibility in deploying face recognition systems. Overall, the findings contribute to the advancement of computer vision research, offering valuable insights for future developments and applications in this field.

References

- P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001.
- Zhang, Z., Luo, P., Loy, C. C., & Tang, X. (2016). Joint face detection and alignment using multitask cascaded convolutional networks. IEEE Signal Processing Letters, 23(10), 1499-1503.
- Yang, J., & Luo, P. (2014). Face detection by structural models. Pattern Recognition, 47(6), 2312-2325.
- Dalal, N., & Triggs, B. (2005). Histograms of oriented gradients for human detection. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05).
- Hsu, C. W., Chang, C. C., & Lin, C. J. (2010). A practical guide to support vector classification.
- Schroff, F., Kalenichenko, D., & Philbin, J. (2015). FaceNet: A unified embedding for face recognition and clustering. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

THANK YOU