FaceTrack AI
Video Frame Analysis Software

https://github.com/mayank-cse/FaceTrackAI

CSE4019 - Image Processing

VIT
Vellore Institute of Technology
CHENNAI

School of Computer Science Engineering
(S C O P E)

20BCE1538 – MAYANK GUPTA

Dr. Jagadeesh Kannan Raju | Professor Grade 2, SCOPE

Abstract:
The use of Closed-Circuit Television (CCTV) cameras for security surveillance has become ubiquitous in modern society. The extensive use of Closed-Circuit Television (CCTV) cameras for security surveillance in contemporary society. An enormous amount of information is captured by these cameras, making manual analysis difficult. In order to derive valuable information from these data streams, there is a growing demand for automated CCTV footage analysis systems.

In this report, we introduce a novel CCTV footage analyser that employs Convolutional Neural Networks (CNNs) for face recognition and detection, followed by image enhancement and a combination of CNN and Long Short-Term Memory (LSTM) models to generate textual descriptions of the identified individuals.

Keywords: Image quilting, texture synthesis, computer vision, computer graphics, Markov random fields, CCTV analysis.

Introduction:
Today, closed-circuit television (CCTV) is employed for a range of daily activities. CCTV has evolved from a straightforward passive surveillance system to an integrated intelligent control system.

In this project, CCTV video frames are used as the basis for decision-making to automatically, effectively, and efficiently identify the images that contain people.

Our suggested system, which uses photographs of a person, succeeds in detecting and marking the presence of the person in a cctv video or images of a busy place, but earlier methods have difficulty recovering accurate facial details or extracting face identity. Additionally, it enables us to combine colour enhancement and restoration, resulting in a nice mix of realism and realness with less artefacts.

Proposed Methodology:
Algorithms Implemented: *Texture Synthesis*
The following outlines the algorithm for Image Quilting :
Step 1: Face Detection When a face is present in a picture, the bounding box that contains it must be found. The face detector in this programme is a Multi-task Cascaded Convolutional Network (MTCNN). Our solution offers a flawless bounding box that properly encloses the face without chopping out significant facial forms and characteristics or adding more background than is required.
Step 2: Feature Extraction The crucial phase of the work of face identification is feature extraction. In this stage, we extract facial component features such landmark points (like eyes, nose, and mouth) and the relationship between them from a human face image that was collected in the Face Detection step above. Our VGG Neural Network of choice is the Resnet-50-based VGGFace2 model created by Oxford University's Visual Geometry Group.
Step 3: Classification In this phase, a classifier evaluates the information given to it to determine whether the face in the image matches the identifier face. Our use case is best suited for cosine similarity, which works by calculating the cosine of the angle between the feature points in the photos to determine how near the images are to one another in an N-dimensional feature space.
Step 4: Face Attribute Extraction and Matching Face attribute matching with ResNet50 involves using a pre-trained deep neural network called ResNet50 to identify and match specific attributes in facial images.
ResNet50 is a convolutional neural network architecture that has been trained on large image datasets, such as ImageNet, and is capable of extracting high-level features from images. This makes it a suitable

\

Result:
The results of our study show that image detection and recognition are effective techniques for generating high-quality images with minimal manual intervention. The proposed CCTV footage analyzer system was successfully implemented and evaluated using actual CCTV footage. The project's image enhancement techniques effectively removed noise and enhanced the overall quality of the footage, allowing for more precise object identification. In addition, the system's image captioning component could generate textual descriptions of the CCTV footage.

Image Recognition

Conclusion:
In conclusion, The project has demonstrated the potential of deep learning techniques to improve the accuracy and efficiency of CCTV surveillance systems, which can have a significant impact on security measures in various sectors. The successful implementation of our project, which included face identification, detection, image enhancement, and image captioning, has made substantial contributions to the improvement of the efficiency and effectiveness of CCTV surveillance systems. Deep learning techniques such as Convolutional Neural Networks (CNN) have enabled the precise detection and classification of various objects, including faces, in real-time CCTV footage.

Contact Details: mayank.gupta2020@vitstudent.ac.in

References:
[1] Li, C., Li, C., Guo, J., Porikli, F., & Pang, Y. (2018). LightenNet: A Convolutional Neural Network for Weakly Illuminated Image Enhancement. IEEE Transactions on Image Processing, 27(8), 4018-4030.
[2] A., V., N., Kameswari. (2021). Image Caption Generator Using Deep Learning. International Journal for Research in Applied Science & Engineering Technology (IJRASET), ISSN: 2321-9653
[3] Palak Kabra, Mihir Gharat , Dhiraj Jha , Shailesh Sangle(2022) Image Caption Generator Using Deep Learning. International Journal for Research in Applied Science & Engineering Technology (IJRASET), ISSN: 2321-9653.