Savitribai Phule Pune University
R. H. Sapat College Of Engineering, Management Studies And Research,
Nashik - 422 005
DEPARTMENT OF COMPUTER ENGINEERING
A. Y. [2022-23]

Intelligent Video Inference System

Guided by

Dr. N. A. Deshpande

Presented By

Mayur Patil
Omkar Divate
Suyash Jadhavar
Prashant Bagul

Contents

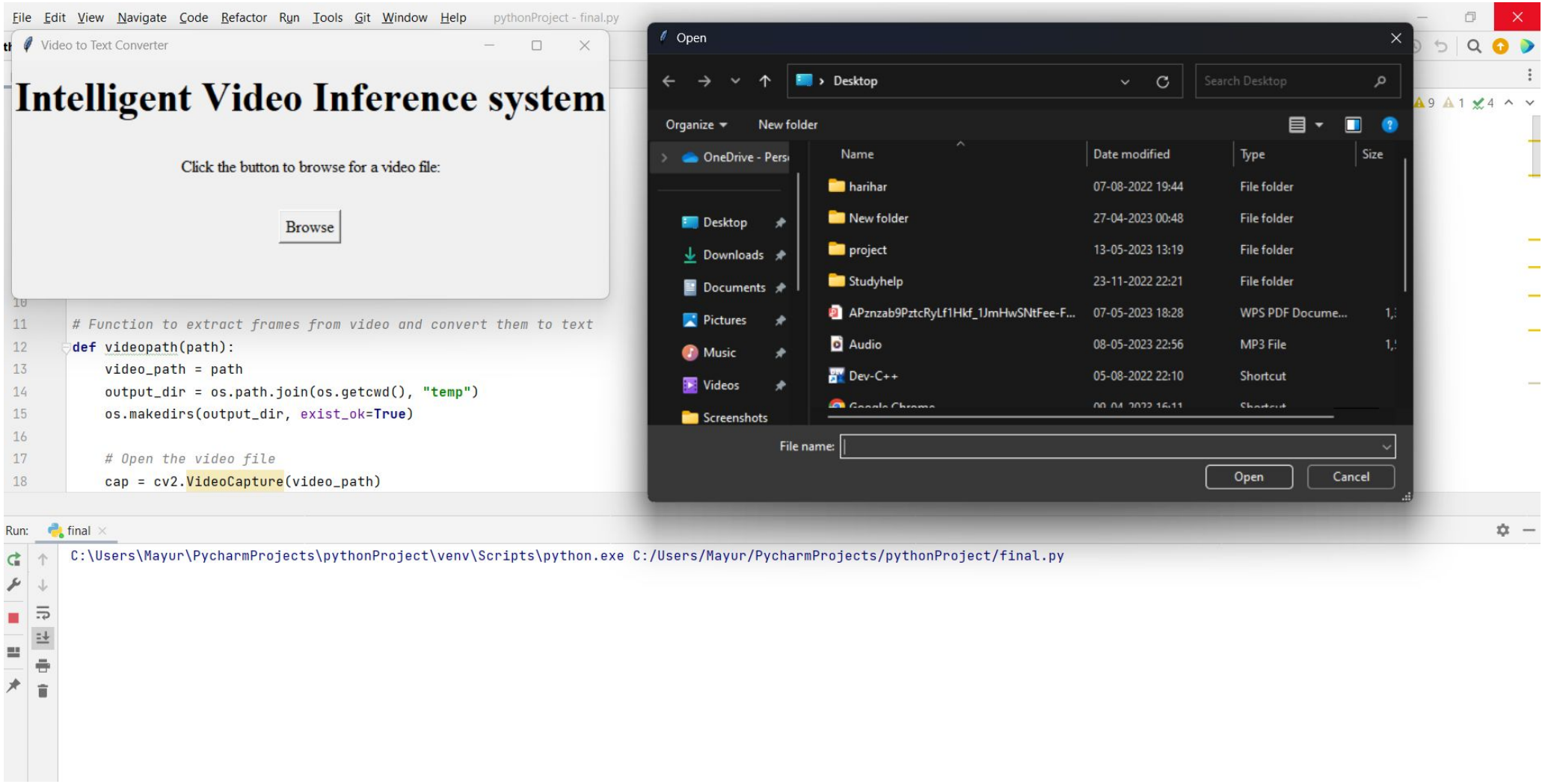
Introduction

Video inference is a task of localizing interesting events from an untrimmed video and producing textual description (captions) for each localized event. An intelligent video inference system is a computer vision technology that is designed to automatically analyze video data and extract useful information. This system can be used to convert video to frames and then analyze those frames using advanced deep learning techniques such as transformer-based image captioning , which involves using a neural network to generate natural language descriptions of the visual content in each frame.

Problem Statement

Video understanding has become increasingly important as surveillance, social, and informational videos weave themselves into our everyday lives, to overcome such problems we can make use of '**Intelligent Video Inference System**'.

Implementation




Video to Text Converter
Intelligent Video Inference system
Click the button to browse for a video file:
Browse
"generated_text": "a person standing in front of a counter with a lot of items "
"generated_text": "a man in a hat is standing in front of a store "
"generated_text": "a person is standing with a shoe on "
"generated_text": "a room with a lot of shelves and a lot of hanging items "
"generated_text": "a man and woman are standing in a store "
"generated_text": "a person standing in the street with a car "

Image	O/P With Blip Small	Accuracy	O/P With Blip Large	Accuracy	O/P With Our Model	Accuracy
	two people on a beach with a dog	rouge1': Score(precision=0.77, recall=0.58, fmeasure=0.66)	a photography of a woman and her dog	rouge1': Score(precision=0.75, recall=0.5, fmeasure=0.6)	a woman sitting on the beach with her dog	rouge1': Score(precision=0.75, recall=0.5, fmeasure=0.6)
	a man standing next to another man holding a cell phone	rouge1': Score(precision=0.855711, recall=0.66666, fmeasure=0.75)	two men are walking down the street	rouge1': Score(precision=0.857271, recall=0.666666, fmeasure=0.75)	there are two men standing next to each other on a sidewalk	rouge1': Score(precision=0.85932, recall=0.59333, fmeasure=0.7599)
	a row of boats are docked at a pier	rouge1': Score(precision=0.66666666, recall=0.4, fmeasure=0.5)	boats are parked on a dock with a boat in the water	rouge1': Score(precision=0.75, recall=0.6, fmeasure=0.66665)	a photography of a boat in the water	rouge1': Score(precision=1.0, recall=0.5333333, fmeasure=0.696)