

Fire and Smoke Detection using Vision Transformer

Introduction:

The "Fire and Smoke Detection using Vision Transformer" project aims to enhance indoor safety by employing advanced deep learning techniques. Utilizing Vision Transformers, this system accurately identifies the presence of fire and smoke within a household environment. By leveraging the transformative power of these state-of-the-art models, the project ensures rapid and reliable detection, significantly improving response times and potentially saving lives and property.

Components:

- Python environment
- Visual Studio Code
- OpenCV
- Pre-trained transformer from Huggingface
- Vit-base-patch-16-224

Methodology:

First, I learned about Vision Transformers (ViT), including their structure, working principles, the type of input they accept, and their output methods. After acquiring this foundational knowledge, I downloaded a relevant dataset from the Roboflow website. Initially, I attempted to build the model from scratch, but found this approach to be inefficient. Consequently, I explored how to utilize pre-trained transformers from the Hugging Face platform and integrated the “vit-base-patch-16-224” model, which is pre-trained on a large dataset, into my Python script.

You can view the Python scripts for this project on my GitHub repository: [Fire and Smoke Detection using ViT](https://github.com/AhmedShafique313/fireandsmoke_ViT).

Upon running the code, I evaluated the model's accuracy based on my image inputs, achieving an accuracy of approximately 0.40. Since this pre-trained model was not specifically trained for detailed fire and smoke classification, it was able to detect the presence of fire and smoke but did not distinguish between different types of fire or smoke.