Abstract—Fire presents a serious hazard, potentially causing extensive damage to properties and endangering human lives. Traditional fire detection systems, which rely on temperature, air concentration, and particle movement, often fail to detect fires in early stages and are costly to install in existing structures. These limitations can be mitigated by visual-based fire detection systems utilizing existing CCTV cameras, offering a more practical and cost-effective solution. Such systems are possible with advancements in computer vision and machine learning to enhance fire detection capabilities, circumventing the need for significant structural modifications. This paper describes the methodology of a combined approach using HSV-based Harris Corner extraction and Vision Transformer classification for fire detection research. HSV color conversion was used to filter objects with fire-like color properties, then Harris Corner Detection was applied to filter objects with fire-like shape attributes. By using both rules, potential fire region of interests could be identified, these regions are then classified by a Vision Transformer model to determine whether they contain fire or not. The dataset used in this research was made up of 2,640 images of both indoor and outdoor lightning conditions, 1,200 fire images and 1,000 non-fire images for the training set, and a separate 240 fire and 200 non-fire images for the evaluation set. The effectiveness of this approach in accurately detecting fire patterns in images reached an accuracy score of 89.53%, recall score of 88.08%, precision score of 90.97%, and an overall F1-score of 89.50%.

Index Terms—fire detection, computer vision, hsv color conversion, harris corner detection, vision transformer