

ABSTRACT

Criminality is increasing every day. CCTV cameras are intentionally set up to track and control illegal activity using video surveillance technology effectively. However, the effectiveness of video analysis accomplished through direct visual inspection by security personnel and traditional image processing in correctly evaluating video surveillance footage is restricted, particularly when humans wear several facial masks such as glasses, face masks, caps, distinct styles of hair, and others. FaceNet is a facial recognition system that can identify criminals in CCTV footage. FaceNet is built on CNN, which frequently struggles with generalizations based on unseen data. Using special pooling techniques, such as hybrid pooling can help increase CNN generalization. The goal of this study is to demonstrate how the system uses hybrid pooling in FaceNet embedding to identify suspects and enhance generalization CNN. Hybrid pooling is a probabilistic blend of max and average pooling. The suggested model surpassed the model that utilizes traditional pooling (max and average pooling) in practically every measurement metric when hybrid pooling replaces max pooling in the Reduction-B block in FaceNet embedding, with a probability value of 0.5, accuracy of 85.96%, f1-score of 84.56%, recall of 85.16%, and AUC score of 90%. Applying max pooling to all blocks yields an accuracy of 85.67%, an F1-Score of 82.64%, a recall of 80.19%, and an AUC score of 87%. By contrast, hybrid pooling results in a performance improvement ranging from 0.29% to 4.97%. Given the experiment's findings, the probability value and the selection of hybrid pooling in a particular block have an important impact on the performance of the model. As a result, the suggested approach has been shown to increase the generalization of FaceNet embedding for suspect identification.

Keywords: FaceNet embedding, suspect identification, Hybrid Pooling, CCTV footage, generalization CNN