

---

## LIST OF FIGURES

2.1	General Architecture of CNN Model [29] . . . . .	6
2.2	Illustration of Calculation in Max Pooling . . . . .	7
2.3	Illustration of Calculation in Average Pooling . . . . .	8
2.4	Illustration of Calculation in Hybrid Pooling . . . . .	9
2.5	MTCNN’s architecture that consists of P-Net, R-Net, and O-Net. Conv: 3x3 means that the convolution layer has a size of 3x3, while MP: 3x3 means that the Max Pooling layer has a size of 3x3. . . . .	10
2.6	Architecture of FaceNet Model . . . . .	11
2.7	Illustration of Learning Process in Triplet Loss. A blue circle represents the anchor, a red circle represents negative data, and a green circle represents positive data. The Triplet Loss algorithm aims to reduce the distance between an anchor and a positive of the same identity while maximizing the distance between the anchor and a negative of an identity that is distinct. . . . .	11
3.1	Flowchart of Design System . . . . .	15
3.2	An example from the mugshot dataset in which each suspect consists of 12 images taken in various scenarios from three different angles: right, left, and center. . . . .	16
3.3	Example of CCTV Dataset . . . . .	17
3.4	Modification of FaceNet Embedding Architecture. 3x3 Conv2D (32 stride 2 V) means that on the convolution layer, there are 32 filters or kernels 3x3-sized which are transferred or stride 2 with padding equal to valid. . . . .	19
4.1	Example of the Output System . . . . .	23
4.2	Comparison ROC-AUC Curve between the Proposed Model and Baseline . . . . .	30
4.3	The example demonstrates error analysis, with the image on the left representing suspect C from CCTV footage and the image on the right representing suspect B from the mugshot dataset. The system determines these two persons to have the same identity, denoted as suspect B. . . . .	32