image classification, and object detection on the image [42, 43]. This is because it can generate a rich feature map using the raw image. Generally, it consists of 3 layers. These layers are convolution, pooling, fully connected layers [44]. The feature is extracted in the Convolution layer. Reducing the size of the extracted feature maps is done with the pooling layer. Image features that pass through many convolution and pooling stages are flattened in the flatten layer and sent to the fully connected layer. This layer has a kind of neural network structure. Classification is made in this layer [45].

The structure of CNN models with high classification success emerges as a result of long trials [46]. The network weights obtained as a result of these trained models can be retrained using

So, this process is repeated k times. The overall classification success of the model is obtained by taking the arithmetic average of the classification successes obtained as a result of these processes. In our study, the k value was determined as 10. It is the k value at which the highest classification success is achieved after many trials. Figure 3 shows how the cross-validation method works.