

The proposed approach trains and ensembles the deep learning-based models using multi-scale training images at the following four steps: (1) resizing the initial images to different scales(448×336 and 296×224), and extracting image patches(224×224) respectively for training network models; (2) Training sub-models using two CNN structures (DenseNet161 and ResNet152) on dataset of 448×336 scales and using DenseNet161 on dataset of 296×224 scales; (3) Inputting test data to the three trained sub-models to get the category; (4) Ensemble the results by using maximum voting. The overall flow of the algorithm as Fig. 1.

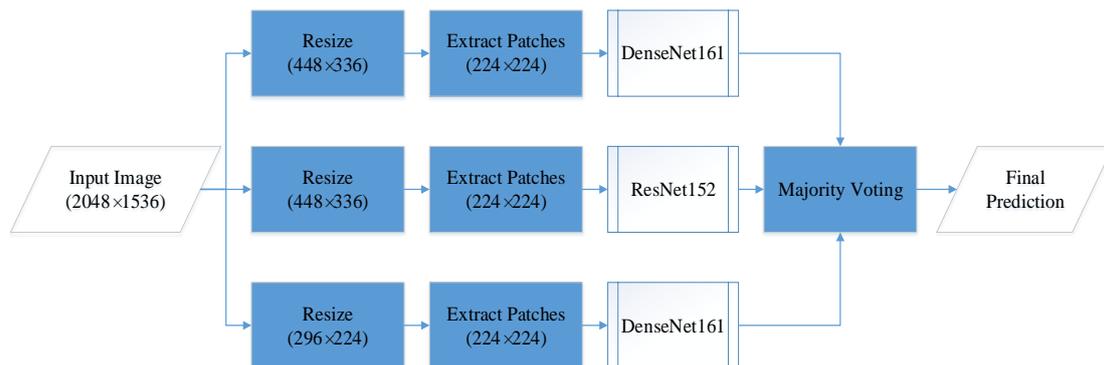


Fig. 1. The overall flow of the algorithm