Tumor Size and Lymph Node Metastasis Detection in Colon Cancer Patients Using MR Images

Authors : Mohammadreza Hedyehzadeh, Mahdi Yousefi

Abstract : Colon cancer is one of the most common cancer, which predicted to increase its prevalence due to the bad eating habits of peoples. Nowadays, due to the busyness of people, the use of fast foods is increasing, and therefore, diagnosis of this disease and its treatment are of particular importance. To determine the best treatment approach for each specific colon cancer patients, the oncologist should be known the stage of the tumor. The most common method to determine the tumor stage is TNM staging system. In this system, M indicates the presence of metastasis, N indicates the extent of spread to the lymph nodes, and T indicates the size of the tumor. It is clear that in order to determine all three of these parameters, an imaging method must be used, and the gold standard imaging protocols for this purpose are CT and PET/CT. In CT imaging, due to the use of X-rays, the risk of cancer and the absorbed dose of the patient is high, while in the PET/CT method, there is a lack of access to the device due to its high cost. Therefore, in this study, we aimed to estimate the tumor size and the extent of its spread to the lymph nodes using MR images. More than 1300 MR images collected from the TCIA portal, and in the first step (pre-processing), histogram equalization to improve image qualities and resizing to get the same image size was done. Two expert radiologists, which work more than 21 years on colon cancer cases, segmented the images and extracted the tumor region from the images. The next step is feature extraction from segmented images and then classify the data into three classes: T0N0， T3N1 ， T3N2. In this article, the VGG-16 convolutional neural network has been used to perform both of the above-mentioned tasks, i.e., feature extraction and classification. This network has 13 convolution layers for feature extraction and three fully connected layers with the softmax activation function for classification. In order to validate the proposed method, the 10-fold cross validation method used in such a way that the data was randomly divided into three parts: training (70% of data), validation (10% of data) and the rest for testing. It is repeated 10 times, each time, the accuracy, sensitivity and specificity of the model are calculated and the average of ten repetitions is reported as the result. The accuracy, specificity and sensitivity of the proposed method for testing dataset was 89/09%, 95/8% and 96/4%. Compared to previous studies, using a safe imaging technique (MRI) and non-use of predefined hand-crafted imaging features to determine the stage of colon cancer patients are some of the study advantages.