Data augmentation on Medical Images to Deep Learning usage

I. Fantini[1], L. Rittner[1], C. Yasuda[2], R. A. Lotufo[1]
[1]Medical Image Computing Lab, FEEC, UNICAMP, [2]Neuroimaging Laboratory, FCM,UNICAMP

Introduction: Deep Learning is considered the most effective tool for image classification, but its use requires a considerable amount of annotated data [1]. Medical images annotation is costly and time-consuming, becoming scarce. On the other hand, medical image datasets are increasing due to noticeable benefits of supporting clinical analysis. The use of small datasets on Deep Convolutional Neural Networks (CNNs) is allowed by transfer learning and data augmentation techniques [2,3]. This study explored these techniques to automatically identify MRI corrupted by motion artifacts.

Materials and Methods: The Magnetic Resonance Image (MRI) data were acquired at the University of Campinas, on a 3T Phillips Achieva scanner. A T1-weighted volumetric sequence was acquired in the sagittal plane (thickness = 1 mm, flip angle = 8 degrees, TR = 7.1 msec, TE = 3.2 msec, matrix 240x240x180, FOV 24x24 isotropic voxels of 1 mm) from 37 healthy volunteers. The dataset is divided into two classes: control images, which contains 24 acquisitions (matched by gender and age); and images corrupted by motion, which comprises 13 acquisitions (8 females, 5 males, aged from 21 to 53).
An InceptionV3 [4] was used to perform the transfer learning experiment, which consists of fine-tuning the pre-trained model using the specific dataset. Since MRI is a 3D sequence while the architecture is a 2D CNN, one model for each MRI axis was trained using their 40 central slices. As the dataset contains 37 acquisitions, we performed data augmentation by rotating, translating and normalizing image intensity. The transformations were combined and applied to each slice. Two intensity normalizations were applied: maximum of 3-sigma or a range from 0.8-1.2 of original maximum. The other parameters obeyed realistic conditions: maximum rotation of 15 degrees; maximum translation of 15 pixels; and were selected randomly per acquisition. As a result, each slice has a different rotation, translation, and intensity range. We opted to extract 128x128 patches size from slices, thereby the InceptionV3 was adapted to the new input size removing the multiclass classifier and the last Inception block layer, and then attaching the new binary classifier. Since the control group was bigger than the motion corrupted group, it was required a different number of patches from each group to balance the data. Also, the positions of the patches were selected randomly, covering the four quadrants of the brain area. Combining the cited transformations, we extracted 40 different patches from each slice from motion corrupted acquisition group and the half for slices from control acquisitions group. The dataset was separated, on acquisition level, into 60% training data (14 control, 8 motion) and 40% validation data (10 control, 5 motion).

Results: The adapted InceptionV3 model detect if the patch has motion artifact or not. The results are shown in Table 1. Checking the original data, without transformations, the trained networks achieve similar performance.

	Axial	Coronal	Sagittal
Accuracy	95.06	91.03	88.60
Sensitivity	99.15	94.50	86.99
Specificity	89.21	86.08	90.92

Table 1: Accuracy, Sensitivity and Specificity, in percentage, for each axis.

Discussion: Although the dataset contains only 37 acquisitions the Deep CNN results reported good accuracy. This indicate that the data augmentation generated the necessary extra data to perform the deep network fine-tuning.

Conclusion: The present work confirmed that data augmentation and transfer learning are useful techniques to train Deep CNNs using limited annotated medical images data.

References: [1] doi:10.1038/nature14539; [2] doi:10.1146/annurev-bioeng-071516-044442; [3] doi:10.1109/TMI.2016.2535302; [4] SZEGEDY, C. et al., Computer Vision and Pattern Recognition: 2818-2826, 2016