Computerized Medical Imaging and Graphics xxx (2017) xxx–xxx

Contents lists available at ScienceDirect

Computerized Medical Imaging and Graphics

journal homepage: www.elsevier.com/locate/compmedimag

Transfer learning on fused multiparametric MR images for classifying histopathological subtypes of rhabdomyosarcoma

Imon Banerjee*, Alexis Crawley, Mythili Bhethanabotla, Heike E Daldrup-Link, Daniel L. Rubin

Department of Radiology, Stanford University School of Medicine, Stanford, CA, United States of America

ARTICLE INFO

Article history:
Received 31 January 2017
Received in revised form 30 April 2017
Accepted 4 May 2017

Keywords:
Rhabdomyosarcoma
Computer aided diagnosis
Image fusion
Transfer learning
Deep neural networks

ABSTRACT

This paper presents a deep-learning-based CADx for the differential diagnosis of embryonal (ERMS) and alveolar (ARMS) subtypes of rhabdomyosarcoma (RMS) solely by analyzing multiparametric MR images. We formulated an automated pipeline that creates a comprehensive representation of tumor by performing a fusion of diffusion-weighted MR scans (DWI) and gadolinium chelate-enhanced T1–weighted MR scans (MRI). Finally, we adapted transfer learning approach where a pre-trained deep convolutional neural network has been fine-tuned based on the fused images for performing classification of the two RMS subtypes. We achieved 85% cross validation prediction accuracy from the fine-tuned deep CNN model. Our system can be exploited to provide a fast, efficient and reproducible diagnosis of RMS subtypes with less human interaction. The framework offers an efficient integration between advanced image processing methods and cutting-edge deep learning techniques which can be extended to deal with other clinical domains that involve multimodal imaging for disease diagnosis.

1. Introduction

Rhabdomyosarcoma (RMS) represents the most common extracranial solid malignancy in children and adolescents with an age of less than 20 years (Ognjanovic et al., 2009). The majority of RMS is of embryonal (ERMS) and alveolar (ARMS) subtypes. Patient outcomes vary considerably, with 5 years survival rates between 35% and 95% depending on the type of RMS involved, tumor grade, tumor stage and patient age, among other factors (Malempati and Hawkins, 2012). Most ARMS are more aggressive than ERMS and require more intense treatment. A diagnosis of the histopathological subtype is critical for effective personalized treatment and survival. In the clinic, RMS subtypes are classified based on specific morphological and genetic characteristics, obtained from the biopsy specimens.

Medical imaging can contribute to the classification of RMS subtypes based on tumor location, but traditional imaging findings are non-specific. A few clinical studies (Baum et al., 2011; Brenner et al., 2004) have shown that the extent of 18F-FDG uptake (representing tumor metabolism) on PET images and the degree of restricted diffusion on MR images (representing tumor cell density) can be linked to prognostic information. While these studies showed some correlation between tumor metabolism/diffusion and patient survival, the degree or distribution of 18F-FDG uptake or diffusion restriction in the tumor tissue is not yet established as an identifier for differentiating ERMS and ARMS. New imaging signs to the categorization of RMS tumors into high and low risk groups could potentially improve assignment of treatment options and outcomes. To the best of our knowledge, no computerized diagnosis system exists that can classify ERMS from ARMS by analyzing only the organ-level scans (e.g. MRI, DWI, PET).

The purpose of our study is to differentiate ARMS and ERMS by analyzing a fusion of diffusion weighted MR and T1 weighted contrast enhanced MR images with less manual intervention. The extraction of effective image features for the differentiation of RMS subtypes is the most crucial component of this study. However, it is an extremely complicated task due to the need of hand crafted descriptor design/selection which requires much manual effort and a deep investigation of the data. We hypothesize that deep convolutional neural network (CNN) based RMS classifier that learns automatically informative features from the fused multiparametric MR images, can provide an effective and convenient solution for the differentiation of ERMS and ARMS subtypes. We also believe that transfer learning approach can be more suitable for our study. Mainly due to the fact that the current dataset is restricted in size, and therefore inappropriate to train a deep-

* Corresponding author.
E-mail address: imonb@stanford.edu (I. Banerjee).

http://dx.doi.org/10.1016/j.compmedimag.2017.05.002
0895-6111/

CNN from scratch. But, the CNNs comprehensively trained on the large scale well-annotated ImageNet dataset (contains 1.2 million images labeled with 1000 categories) may be transferred and fine-tuned on the small medical dataset for performing the RMS classification, regardless the disparity between natural images and fused MR scan images.

We developed a radiomics framework that classifies ARMS from ERMS tumors by exploiting a tight integration of advanced image processing methods and cutting-edge deep learning techniques. First, we independently segmented the tumor from two multiparametric MR images (contrast enhanced T1 weighted MR image and diffusion weighted MR image) using a completely automatic segmentation pipeline. Afterwards, we registered the segmented tumor images using a sophisticated non-rigid registration technique and generated the fused RGB color images from the registered data. We applied standard data augmentation technique to obtain a sufficiently large training dataset to train the CNN. Finally, the fused RGB images were used to train a deep neural network in which we adapted a transfer learning approach with AlexNet model for the RMS classification task. During the training phase, we fine-tuned the ImageNet pre-trained AlexNet model and achieved 85% cross validation classification accuracy. The following sections (2, 3, 4, 5) give further details on the dataset, procedure and explain how the segmentation, registration, data augmentation, and deep learning aspects were practiced focusing our scientific contributions in each step.

2. Related works

Tumor classification by integrating texture analysis of medical images and standard machine learning techniques, is a common approach in literature. In Othman et al. (2011) authors performed classification of brain tumor using Daubechies (db4) wavelet texture analysis and Support Vector Machine (SVM) and 65% accuracy was obtained, where, only 39 images were successfully classified from 60 images. It was concluded that classification using Support Vector Machine resulted in a limited precision, since it cannot work accurately for a large data due to training complexity. In Othman and Basri (2011), a Probabilistic Neural Network (PNN) for tumor classification was proposed to classify brain tumor using Principal Component Analysis for feature extraction and PNN for classification. They concluded that PNN is a promising tool for brain tumor classification, based on its fast speed and its accuracy which ranges from 73 to 100% for spread values (smoothing factor) from 1 to 3. Classification of brain MRI using the LH and HL wavelet transform sub-bands was performed in Lahmiri and Boukadoum (2011) that shows that feature extraction from the LH (Low-High) and HL (High-Low) sub-bands using first order statistics has higher performance than features from LL (Low-low) bands. A few studies (Mayerhoefer et al., 2008; Juntu et al., 2010) showed that texture analysis can also be informative for discrimination between benign and malignant soft-tissue sarcomas in MRI images. However, such standard machine learning techniques need very specific feature extractors for each type of tumor classification task and this requires much manual data analysis, and it becomes difficult to extend the approaches for a new dataset.

In the last few years, deep convolutional neural networks (Deep-CNNs) that try to learn high level features from the given data, has been successfully applied to a wide range of applications, including natural language processing, image classification, semantic tagging (LeCun et al., 2015). Deep-CNN is reducing the task of making new feature extractor for each type of data (speech, image, etc.). Recently, the Deep-CNNs have also been introduced to the medical domain with promising results in various areas, like organ segmentations and detection, image standard plane selection, computerized diagnosis and prognosis, etc. (Greenspan et al., 2016). The Deep-CNN could potentially change the design paradigm of the computerized diagnosis and prognosis framework due to several advantages over the standard machine learning. First, deep learning can directly uncover features from the training data, and hence the effort of explicit elaboration on feature extraction can be significantly alleviated (Bengio et al., 2007). The neuron-crafted features may compensate and even surpass the discriminative power of the conventional feature extraction methods. Second, feature interaction and hierarchy can be exploited jointly within the intrinsic deep architecture of a neural network (Lee et al., 2011). Consequently, the feature selection process will be significantly simplified. Third, the three steps of feature extraction, selection and supervised classification can be realized within the optimization of the same deep architecture (Krizhevsky et al., 2012). With such a design, the performance can be tuned more easily in a systematic fashion.

Recently, ImageNet pre-trained CNNs have been used for chest pathology identification and detection in X-ray and CT modalities (van Ginneken et al., 2015; Bar et al., 4140; Ciompi et al., 2015) and have yielded the best performance results. However, the fine-tuning of a pre-trained CNN model on multimodal RMS tumor image datasets has not yet been exploited. In this work, we propose to use the transfer leaning approach where we use pre-trained AlexNet Deep-CNN for classifying aggressive ARMS from less aggressive ERMS brain tumor by using fusion of T1 weighted contrast enhance MR and diffusion weighted MR images.

3. Materials and method

Fig. 1 illustrates a simplified schematic diagram of the proposed pipeline where we show how an unseen dataset of multiparametric MR images will be automatically analyzed and classified within our proposed framework. In the following sub-sections, we detail each of the core components including description of the dataset used in the study and the model training methodology.

3.1. Dataset

In a retrospective, *Institutional Review Board* (IRB) approved study, we evaluated diffusion-weighted MR scans and contrast enhanced T1-weighted MR scans of 21 children and adolescents (age 1–20 years) with newly diagnosed intermediate-risk ARMS ($n = 6$) and ERMS ($n = 15$) in the head and neck. The cohort included 7 girls and 14 boys with ages between 1–20 years (10.04 ± 5.42 SD). To our knowledge, this is the largest multi-institutional cohort of children with RMS imaged with these two modalities to date. The imaging work-flow consisted of contrast enhanced T1 weighted MR scans (MRI) and diffusion weighted MR scans (DWI) before initiation of chemotherapy. The scans are acquired in the axial plane for all patients. An intravenous infusion of gadolinium chelate was used as contrasting agent. The dose of the contrast agent was decided based on the infant's weight to get a similar distribution in the patient's blood. The MR scans' slice thickness was 3–5 mm and in-plane resolution range was 0.4–0.5 mm^2. The diffusion weighted MR scans' slice thickness was 2.5–5 mm and in plane resolution range was 1–1.2 mm^2. The shorter b-value in the range of 500–700 s/mm has been used.

Tumor was outlined as 2D region of interest (ROI) on single slice of each scan by the certified radiologists via an interactive web-based software-epad (Rubin et al., 2014). In Fig. 2, we show an epad snapshot of how the annotation has been done. Through visual inspection of each scan volume, the radiologist also determined the slice range where the tumor is visible. For the current dataset, there were between 5–8 slices. For each patient data, RMS

Please cite this article in press as: Banerjee, I., et al., Transfer learning on fused multiparametric MR images for classifying histopathological subtypes of rhabdomyosarcoma. Comput Med Imaging Graph (2017), http://dx.doi.org/10.1016/j.compmedimag.2017.05.002