Computerized Medical Imaging and Graphics xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Computerized Medical Imaging and Graphics

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

# Automatic detection and classification of regions of FDG uptake in whole-body PET-CT lymphoma studies

Lei Bi[a], Jinman Kim[a,*], Ashnil Kumar[a], Lingfeng Wen[a,b], Dagan Feng[a,c], Michael Fulham[a,b,d]

[a] School of Information Technologies, University of Sydney, NSW, Australia
[b] Department of Molecular Imaging, Royal Prince Alfred Hospital, NSW, Australia
[c] Med-X Research Institute, Shanghai Jiao Tong University, Shanghai, China
[d] Sydney Medical School, University of Sydney, NSW, Australia

ARTICLE INFO

*Article history:*
Received 30 April 2016
Received in revised form
29 November 2016
Accepted 30 November 2016

*Keywords:*
Classification
Thresholding
PET-CT
CNN

ABSTRACT

[$^{18}$F]-Fluorodeoxyglucose (FDG) positron emission tomography–computed tomography (PET–CT) scans of lymphoma patients usually show disease involvement as foci of increased radiotracer uptake. Existing methods for detecting abnormalities, model the characteristics of these foci; this is challenging due to the inconsistent shape and localization information about the lesions. Thresholding the degree of FDG uptake is the standard method to separate different sites of involvement. But may fragment sites into smaller regions, and may also incorrectly identify sites of normal physiological FDG uptake and normal FDG excretion (sFEPU) such as the kidneys, bladder, brain and heart. These sFEPU can obscure sites of abnormal uptake, which can make image interpretation problematic. Identifying sFEPU is therefore important for improving the sensitivity of lesion detection and image interpretation. Existing methods to identify sFEPU are inaccurate because they fail to account for the low inter-class differences between sFEPU fragments and their inconsistent localization information. In this study, we address this issue by using a multi-scale superpixel-based encoding (MSE) to group the individual sFEPU fragments into larger regions, thereby, enabling the extraction of highly discriminative image features via domain transferred convolutional neural networks. We then classify there regions into one of the sFEPU classes using a class-driven feature selection and classification model (CFSC) method that avoids overfitting to the most frequently occurring classes. Our experiments on 40 whole-body lymphoma PET-CT studies show that our method achieved better accuracy (an average F-score of 91.73%) compared to existing methods in the classification of sFEPU.



## 1. Introduction

[$^{18}$F]Fluorodeoxyglucose positron emission tomography–computed tomography (FDG PET–CT) is regarded as the imaging modality of choice for the evaluation, staging and assessment of response in many malignancies including the lymphomas (Hong et al., 2007; Freudenberg et al., 2004; Wahl, 2009). The combination of PET and CT in one device combines the sensitivity of PET to detect regions of abnormal function and the anatomical localization of CT (Wahl, 2009). Sites of disease usually display greater FDG uptake than normal structures. The standardized uptake value (SUV) is a semi-quantitative measure of FDG uptake or glucose metabolism and is extensively used in clinics to measure the degree of FDG uptake in sites of disease (Wahl, 2009). Different malignancies have varying degrees of FDG uptake and lymphomas are one of the most glucose-avid cancers that are routinely staged and re-staged with PET-CT. SUV thresholding of PET images is the main approach used to detect sites of abnormal FDG uptake before and after treatment (Yu et al., 2009; Vauclin et al., 2009). Regions with an SUV value higher than a specified limit (called the 'threshold value') are identified as regions of interest (ROIs) e.g., tumors (Hong et al., 2007; Wahl, 2009; Vauclin et al., 2009; Hirata et al., 2014; Nestle et al., 2005; Paulino and Johnstone, 2004). Common SUV threshold values include an SUV of ≥2.5 (Hellwig et al., 2007), 4.4 (Vansteenkiste et al., 1998), 5.3 (Bryant et al., 2006), and a value above the average SUV of a background reference region (Francis et al., 2007; Zasadny et al., 1998), such as the liver (Wahl, 2009; Paquet et al., 2004) and the mediastinal blood pool in the thoracic aorta (Paquet et al., 2004; Ghosh and Kelly, 2010). However, (global) SUV thresholding

* Corresponding author.
*E-mail address:* jinman.kim@sydney.edu.au (J. Kim).

http://dx.doi.org/10.1016/j.compmedimag.2016.11.008
0895-6111/

does not take local SUV variations into account and so can include normal tissue.

We define sites of FDG excretion and physiologic uptake (sFEPU) as the globally thresholded sites of expected normal FDG uptake that are thresholded alongside tumors and other abnormal regions in whole-body PET studies. These sFEPU predominately belong to the excretion uptake in the kidneys and both ureters, normal physiological uptake in the brain and the heart, and pooling of FDG in the bladder. A single sFEPU is often split into many smaller fragments, which is a byproduct of global thresholding on heterogeneous structures such as the kidneys, which have varying degrees of FDG present in different locations. Global thresholding can therefore make the image-driven assessment of disease problematic as it can obscure disease in adjacent structures, in particular, in the paravertebral regions, in the mid and lower abdomen where involve lymph nodes lie adjacent to the ureters. The automatic identification and labeling of sFEPU, and their separation from sites of disease would thus therefore improve lesion detection and computer aided diagnosis. The automated detection and labelling of sFEPU is challenging because: (1) there are low inter-class differences, as some sFEPU fragments may only partially represent a class/structure. e.g., a kidney fragment only represents a portion of the whole kidney which makes some image features ineffective for differentiation (see Fig. 1b); (2) there is inconsistent localization information about sFEPU fragments due to the random localization of abnormal sites with the body; and (3) there is a large variation in the degree of FDG uptake among different patients where a structure (e.g., heart) may not have been thresholded (due to being under the threshold value) and thus appears 'absent'.

There have been many different approaches that attempt to separate and label different structures on PET-CT studies: (a) abnormality classification/detection, which attempts to classify/detect one type of abnormality, e.g., liver tumors; (b) multi-structure classification where the aim is to detect or semantically label multiple anatomical structures that excludes abnormalities; and (c) abnormalities and multi-structure classification, which attempts to label different types of structures and abnormalities within the same framework. Existing research in abnormality detection is mainly limited to detecting only a single type of abnormality e.g., liver tumors (Pescia et al., 2008), lung nodules (Zhang et al., 2014), lung tumors (Ballangan et al., 2010). The underlying assumption is that there is only single lesion type in the image. These methods typically require prior knowledge to model the abnormality and to constrain the detection e.g., lung segmentation is usually required for lung tumor detection and the classification accuracy will rely on the segmentation performance (Ballangan et al., 2010). In addition, these methods are unproven for the simultaneous detection of abnormalities on whole-body images as they depend on the segmentation of anatomical structures and a priori knowledge about specific abnormalities. The majority of multi-structure classification approaches are optimized to localize normal structures: Zhan et al. (2008) used an active scheduling approach to detect multiple organs, Criminisi et al. (2009) used relative spatial features with random forest to localize different organs on CT volumes, and Linguraru and Summers (2008) used template matching to detect abdominal organs on CT. Methods using probabilistic atlases with deformable registration, geometric transformation and probabilistic averaging have also been used to identify multiple organs (Han et al., 2008; Zhuang et al., 2008; Fenchel et al., 2008; Shimizu, 2006; Yao et al., 2006). The focus on normal structures, however, means that these methods struggle in the presence of the deformations introduced by disease, which affect the size and shape of the involved structure in variable and inconsistent ways (Li et al., 2012).

There has been limited work on the simultaneous classification (detection and separation) of abnormalities and multiple normal structures. In general, this work has involved in localizing individual regions, extracting discriminative features, and then using supervised classification algorithms to label each region. Lartizien et al. (2014) used a combination of texture features, filter based feature selection, and support vector machines (SVMs) to separate several types of lymphoma and non-lymphoma regions in PET-CT images. However, input ROIs required manual delineation, which is highly operator dependent, time-consuming, and is poorly reproducible across different user groups. Wu et al. (2012) used region growing to detect potential abnormalities and then used SVMs to classify these regions into different classes. Similarly, Song et al. (2012) used a multi-stage classification framework that combined SVM with conditional random field (CRF) for detecting the lungs, mediastinum, lung tumors and lymph nodes. In a later work, Song et al. (2013) used a weighted sparse representation with image patches for classification. However, all these works were designed to work with specific anatomical regions in PET-CT images such as the thorax (Song et al., 2012; Song et al., 2013) and head and neck area (Wu et al., 2012). Furthermore, these methods relied on contextual features to separate different structures and were dependent on the accurate localization of the normal structures. Such methods are not suitable for whole-body PET-CT lymphoma studies where there can be innumerable sites of disease seen across the region examined.

In previous work, we conducted preliminary studies to address the simultaneous classification of abnormalities and multiple normal structures on whole-body PET-CT studies (Bi et al., 2014a,b; Bi et al., 2015a). Our approach was to detect all the potential abnormalities e.g., thresholding and then iteratively filtering out normal structures rather than model lesions that can have inconsistent shapes and localization information. Abnormalities can be detected in a reverse manner through the filtering (removal) of normal structures. In the initial work, we used PET-CT features (Bi et al., 2014a) to classify and separate the sFEPU fragments, where we selectively used PET, CT or PET-CT features based on the image characteristics of different structures. We extended this work to cluster the thresholded fragments thereby increasing the discriminative power of the features derived from clustered fragments when compared to using individual fragments (Bi et al., 2014b). We also investigated the optimal feature representation to individual structures using a structure based feature selection strategy together with a SVM for classification (Bi et al., 2015a). These previous approaches relied on using individual thresholded fragments which lack discriminative power especially for small fragments (as shown in Fig. 1b) (Bi et al., 2014a; Bi et al., 2015a). The clustering based method assembled the fragments of the same structure (see Fig. 1c) but was not able to describe the structures since the cluster only partially represented the actual structure and left large semantic differences between the clusters and the actual structure (see Fig. 1a, c and e)

In this study, we propose a novel algorithm that uses a multi-scale superpixel-based encoding method (MSE) and a class-driven feature selection and classification model (CFSC) for sFEPU classification in whole-body PET-CT lymphoma studies. We derived class-driven features from multi-scale superpixel regions encoded with domain transferred deep convolutional neural networks (CNN) features for classification. Our algorithm differs from other methods as follows:

(1) Our MSE approach enables the grouping of the sFEPU fragments which then permits the extraction of optimal features on multi-scale superpixels, thereby increasing the discriminative power compared to using individual fragments. When compared with traditional methods reliant on sliding windows, our approach minimizes the risk of merging unrelated pixels by aggregating pixels conservatively into superpixels to capture local redundancy in the data. The use of multi-scale superpixels allows

Please cite this article in press as: Bi, L., et al., Automatic detection and classification of regions of FDG uptake in whole-body PET-CT lymphoma studies. Comput Med Imaging Graph (2016), http://dx.doi.org/10.1016/j.compmedimag.2016.11.008