Chapter 6

Teeth and Landmarks Detection and Classification Based on Deep Neural Networks

Lyudmila N. Tuzova  
Denti.AI, Russia

Dmitry V. Tuzoff  
Steklov Institute of Mathematics in St. Petersburg, Russia

Sergey I. Nikolenko  
Steklov Institute of Mathematics in St. Petersburg, Russia

Alexey S. Krasnov  
Dmitry Rogachev National Research Center of Pediatric Hematology, Oncology, and Immunology, Russia

ABSTRACT

In the recent decade, deep neural networks have enjoyed rapid development in various domains, including medicine. Convolutional neural networks (CNNs), deep neural network structures commonly used for image interpretation, brought the breakthrough in computer vision and became state-of-the-art techniques for various image recognition tasks, such as image classification, object detection, and semantic segmentation. In this chapter, the authors provide an overview of deep learning algorithms and review available literature for dental image analysis with methods based on CNNs. The present study is focused on the problems of landmarks and teeth detection and classification, as these tasks comprise an essential part of dental image interpretation both in clinical dentistry and in human identification systems based on the dental biometrical information.
INTRODUCTION

Over the last decade, computer vision models and algorithms based on deep learning models have been successfully applied to various health and medicine domains in a number of medical imaging tasks such as detection and staging of cancer, lung segmentation, diagnosis of colitis, detection and classification of brain diseases, and many others (Lee, et al., 2017; Rezaei, Yang, & Meinel, 2017; Liu, et al., 2017; Litjens, et al., 2017; Shen, Wu, & Suk, 2017). In dentistry, several works have applied deep learning models and algorithms for dental radiograph analysis (Miki, et al., 2016; Lee, Park, & Kim, 2017; Wang, et al., 2016; Ö. Arik, Ibragimov, & Xing, 2017; Tuzoff, et al., 2018). However, deep learning in dentistry still remains an underdeveloped area of research, even though deep neural networks provide state-of-the-art results in many kinds of image recognition tasks (Lee, et al., 2017; Litjens, et al., 2017; LeCun, Bengio, & Hinton, 2015) and problems such as teeth and landmark detection appear to be straightforward object detection problems that could be amenable to modern computer vision approaches based on deep learning methods.

In this chapter, the relevant literature on deep learning methods, specifically convolutional neural networks (CNNs), applied for the tasks of teeth and landmarks detection and classification by type or tooth number is reviewed. These tasks comprise an important part of dental X-ray image analysis. The results can be used to automatically fill a patient’s dental records for medical history and treatment planning, preprocess an image for further pathology detection, improve speed and accuracy of postmortem human identification, perform anatomical measurements, and other problems. The deep learning methods have previously been studied for pathology detection purposes (Wang, et al., 2016; Oliveira & Proença, 2011; Imangaliyev, et al., 2016) as well; however, a review of computer-aided disease diagnostics is out of scope of the present study.

CNNs represent state-of-the-art deep learning architectures commonly applied for image recognition tasks. Modern object classification, detection and segmentation approaches based on CNNs have shown promising results, often outperforming methods based on traditional computer vision or other machine learning techniques. An important advantage of deep learning techniques compared with traditional computer vision and other machine learning approaches is that deep learning algorithms do not rely on handcrafted feature extraction and can achieve high performance working with raw input such as pixel values for X-Ray image sources. These methods allow interpreting medical images even if the images are noisy, taken with a different equipment, or in a different setting than that of the data used for training the model. Despite the increasing popularity of the CNNs, the challenges of the application of such architectures still exist. One of the most significant limitation is the amount of annotated data required for the effective model training.

A number of other deep learning models have previously been studied for medical image analysis, including stacked auto encoders (SAEs) and deep belief networks (DBNs) (Shen, Wu, & Suk, 2017; Litjens, et al., 2017). However, CNNs currently represent the state-of-the-art architectures for image recognition tasks. In (Shen, Wu, & Suk, 2017) high performance of CNNs for the images interpretation is explained by the specific properties of the CNNs architectures that better utilize spatial information of images, when most of the other deep learning models process the input in the one-dimensional vector form. Moreover, training of the models, such as DBNs and SAEs, is a complex task combining unsupervised pre-training phase followed by the fine-tuning supervised step. A number of prior studies demonstrated that CNNs models outperformed other deep learning techniques for the image interpreta-