Guest Editorial: Computer Vision in Cancer Data Analysis

Recent progress in imaging hardware, acquisition techniques, and algorithmic processing of data have led to advances in detection, diagnosis, staging, treatment, and follow-up in cancer-related clinical workflows as well as fundamental understanding of cancer modelling and dynamics. Computer Vision presents a promising approach to process the ever-increasing amount of cancer-related data acquired and available through data repositories. The seven papers that constitute this Special Issue present a variety of algorithmic techniques, from texture analysis to adversarial autoencoders, which in turn are applied to different imaging modalities, including microscopy, mammography, ultrasound, computed tomography and dermoscopy.

The Papanicolau test, or cervical smear, is a widely-used procedure to observe potentially cancerous processes in the cervix. The procedure is relatively simple and inexpensive and allows the early detection of cervical cancer. Whilst the test itself is simple, the analysis of the microscopic images acquired requires an experienced pathologist to distinguish between cases, and in lowincome countries, the analysis may be a bottleneck, especially in rural populations. Arya et al. present a feature extraction pipeline in which they compared several texture-based approaches and different classifiers. The combination of features from many techniques provided a significant advantage over shape-based features.

Breast cancer is a common case of cancer, particularly among females, and also one in which survival is high if diagnosed early. This is partly due to screening programmes using different imaging technologies, two of which are presented in this Special Issue. Suhail et al. describe the classification of mass abnormalities in mammographic images using texture-based features extracted with a filter-bank, which generates a texton dictionary to be classified with a Naïve Bayes classifier. The classification results of benign and malignant patches are comparable with the literature in terms of classification accuracy. Ultrasound presents an alternative to the x-ray-based mammogram. Whilst it avoids the use of ionising radiation, the images produced by the ultrasound tend to be more difficult to interpret due speckle, reduced contrast and shadowing effects. Panigrahi et al. in turn analysed malignant and benign breast masses observed with ultrasound and confirmed with a biopsy. Their hybrid algorithm combined clustering and level set methods to segment the ultrasound images with high accuracy.

Computed Tomography (CT) is an x-ray-based imaging technique, which provides three-dimensional, high resolution and high-contrast images of the human body. CT is widely used to inspect cancer in the abdomen, pelvis, and chest. Some CT studies use contrast agents, like barium or iodine, that when injected into a patient provide enhanced contrast in the region of interest. These agents may provide additional risks to some patients, and thus techniques that can provide enhancement without contrast agents are of interest. Lakshmipriya et al. analysed one such image enhancement approach to study liver lesions observed with CT using fuzzy histogram equalization in the non sub-sampled contourlet transform (NSCT) domain followed by decorrelation stretching. The method demonstrates how liver CT diagnosis can be performed with high accuracy without the need for contrast agents.

Skin cancer is another important area for cancer data analysis, and there is considerable interest in both melanoma and nonmelanoma skin cancer. Two possible reasons are a significant increase in skin cancers, possibly linked to increase sun exposure and tanning beds, and the ubiquity of smartphones that allow the easy acquisition of skin lesion photographs. The last three papers of this Special Issue analyse skin images with different approaches. Jaisakhti et al. present an automated skin lesion segmentation, which is a first step towards a classification of the lesions. The authors combine a GrabCut algorithm with a refinement of the boundaries using colour features and k-means clustering. Sultana et al. take a different approach and describe a deep convolutional neural network for the detection of melanoma. The authors compared their results with several publicly available data sets, which provide enough cases for training and testing their algorithm. The network derives 2048 discriminative features that provided a high accuracy of detection of melanoma cases. Deep learning approaches often require a very large number of labelled training samples through which the networks learn the specific characteristics of each class. In the final paper of this issue, Creswell et al. discuss the use of a semi-supervised deep learning model using a denoising adversarial auto-encoder that can be used in cases where labelled data is limited and unlabelled data is available in large amounts.