TDLH : Tumor Detection in Lymph Node Metastases from Histo-pathology Biopsy Images

Abstract

The advent of deep learning (DL) in the field of healthcare, especially histo-pathological images, helps significantly in improving the efficacy of diagnostic systems. The state-of-the-art methods of deep learning are capable of generating equivalent or even better results in comparison to human experts for the development of medical disease diagnostic systems. One such diagnosis, namely metastasis detection, is labor intensive and error prone because of its imaging complexity. The proposed approach TDLH is inspired from characteristics of deep learning and thus evaluates the performance of TDLH in detecting metastases in hematoxylin and eosin-stained tissue sections of lymph nodes. TDLH utilized DL-based model VGG19 to detect tumour by exploiting transfer learning. The variant of VGG19 namely VGG19_x2 outperforms the recent baseline methods and has achieved an accuracy of 95.94% on the Patch-Camelyon test dataset.

Keywords: Histopathology analysis, convolutional neural network, optical biopsy

1. Introduction

Histopathology includes studying a biopsy or surgical specimen by a pathologist following diagnosis of the patient and placing of histological sections on glass slides. Generally, it seems that pathologists have been using conventional microscopes as a sample observational tool and microscopic slides to prognosticate and diagnose illness during last few decades. The traditional procedure includes analysis of several microscopic slides manually that exploits significant time and effort of a pathologist (Nakhleh, 2006). The advent of anatomical imaging makes these microscopic slide-based imaging devices capable of producing high-resolution electronic microscopic images which has put the modern pathology into the digital era of computation providing numerous contributions to the processes in the pathology (Gurcan et al., 2009).

Medical imaging helps pathologist using Computer Aided Diagnosis (CAD) to aid automated image analysis in the examination and quantification of anatomical imaging, thus reduces the stipulated time required for manual screening and improves the performance in terms of detection, prediction and prognostication. Medical imaging is a field which is revolutionized using machine learning (ML) and DL methods whose basis is convolutional neural networks (CNNs). Deep learning (DL) methods based on CNNs have shown enormous potential in the detection, prediction and prognostication of clinical objects (Ghaznavi et al., 2013; Litjens et al., 2016) such as medical image classification, lesion detection, medical image segmentation, grading of prostate cancer and glioma, tumor cell detection in primary breast cancer and many more (Xu et al., 2016).

2. RELATED WORK

Deep CNNs managed to successfully classify and segment clinical objects (Ronneberger et al., 2015) such as lymph node identification (Litjens et al., 2016). In order to detect nuclei, (Sirinukunwattana et al., 2016) designed two CNNs and categorized them into layered
sub-categories. (Bejnordi et al., 2017) has performed study of whole slide images (WSI) of breast cancer using DL, with or without lymph node metastases for identification of metastases with eosine-stained tissue and hematoxyline and comparing the ground truth (GT) recorded by 11 pathologists and these findings are influenced by quality, time and interpretation of the slides by pathologists. (Veeling et al., 2018) has proposed a model to enhance tumor localization over lymph node metastases dataset using rotation equivariant CNNs and achieved an accuracy of 89.80%. (Wang et al., 2016) suggested a DL-based model for auto-detecting metastatic cancer from full-slide images of sentinel lymph nodes over Camelyon16 dataset.

3. PROPOSED METHOD: TDLH

3.1. Dataset and its pre-processing

Patch Camelyon (PCam) dataset (Veeling et al., 2018), a subset of the modified version of Camelyon16 dataset (Bejnordi et al., 2017) is used in TDLH consisting of 327,680 histopathologic scans of lymph node sections in 2D color images with dimension 96 × 96 in pixels. A binary label indicating the existence of metastatic tissue is annotated by medical expert for each image where 0 and 1 represents no tumor and presence of tumor tissue(s), respectively. A total of 160k images from both labels in equal proportion is considered along with holdout scheme (80 : 20) to validate TDLH. Few sample images of both binary labels are depicted in Figure 1.

![Sample Images from Dataset](image-url)

Figure 1: Sample images from dataset (Veeling et al., 2018). The top row includes healthy specimens with labels “0” whereas tumor samples in the lower rows with labels “1”.

3.2. Proposed Network Architecture: VGG19_v2

The proposed network structure VGG19_v2 is initialized from a pre-trained VGG19 model parameters and weights exploiting transfer learning with exclusion of average pooling and fully connected layers. For classification task, the output of VGG19 model is fed to flatten layer, then to a stack of dense layers and eventually, this output is fed into a sigmoid activated layer. The model is compiled using Adam optimizer, ReLU activation and binary cross-entropy loss for binary predictions. The significant layers used in VGG19_v2 are discussed. The first one, Input, is first layer which inputs training data into model and consists of shape 96 × 96 × 3. Next, Convolutional layer consists of filters initialised by random weights which need to be learned by the network. Further, a down-sampling layer Max Pooling is incorporated which reduces the input dimensionality. Finally, after feature extraction Fully-Connected layer classifies high level features. Moreover, Dropout is used as regularization to avoid overfitting on training data.
4. RESULTS ANALYSIS AND DISCUSSION

The performance of VGG19_v2 is evaluated based on Accuracy, Sensitivity, Specificity, F1-Score and Mathew’s Correlation Coefficient (MCC) as depicted in Table 1. It has been observed that VGG19_v2 outperforms the baseline methods P4M and Pi+ and achieved an ROC-AUC score of 99.02%. Moreover, among all experimentations for TDLH the results of VGG19_v2 is comparable to other models ResNet, NASNet and VGG19.

Table 1: Performance metrics of the proposed CNN-oriented DL-based TDLH methods along with baseline methods statistics on PCam dataset.

<table>
<thead>
<tr>
<th>Research Contribution</th>
<th>Method</th>
<th>Accuracy (in %)</th>
<th>Sensitivity (in %)</th>
<th>Specificity (in %)</th>
<th>F1 Score (in %)</th>
<th>MCC (in %)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Veeling et al., 2018)</td>
<td>P4M</td>
<td>89.80</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>(Teh and Taylor, 2019)</td>
<td>Pi+</td>
<td>90.36 ± 0.41</td>
<td>96.49</td>
<td>91.28</td>
<td>97.84</td>
<td>91.23</td>
</tr>
<tr>
<td>Proposed (TDLH)</td>
<td>Resnet</td>
<td>93.58</td>
<td>96.49</td>
<td>91.28</td>
<td>94.04</td>
<td>87.88</td>
</tr>
<tr>
<td>Proposed (TDLH)</td>
<td>NASNet</td>
<td>94.47</td>
<td>93.55</td>
<td>95.78</td>
<td>94.61</td>
<td>89.36</td>
</tr>
<tr>
<td>Proposed (TDLH)</td>
<td>VGG19</td>
<td>95.60</td>
<td>96.88</td>
<td>94.32</td>
<td>95.65</td>
<td>91.23</td>
</tr>
<tr>
<td>Proposed (TDLH)</td>
<td>VGG19_v2</td>
<td>95.94</td>
<td>96.63</td>
<td>95.25</td>
<td>95.97</td>
<td>91.90</td>
</tr>
</tbody>
</table>

References


