A Multi-Task Self-Normalizing 3D-CNN to Infer Tuberculosis Radiological Manifestations

Pedro M. Gordaliza1
Juan José Vaquero1
Sally Sharpe2
Fergus Gleeson3
Arrate Muñoz-Barrutia1

1 Dpto.Bioingeniería e Ing. Aeroespacial, Universidad Carlos III de Madrid, Leganés, Spain
2 Microbiology Services Division, Public Health England, Porton Down, England
3 Oncology Department, The Churchill Hospital, Headington, Oxford, UK

Editors: Under Review for MIDL 2019

Abstract

We propose a learning method well-suited to infer the presence of Tuberculosis (TB) manifestations on Computer Tomography (CT) scans mimicking the radiologist reports. Latent features are extracted from the CT volumes employing the V-Net encoder and those are the input to a Feed-Forward Neural Network (FNN) for multi-class classification. To overtake the issues (e.g., exploding/vanishing gradients, lack of sensibility) that normally appear when training deep 3D models with datasets of limited size and composed of large volumes, our proposal employs: 1) At the network architecture level, the scaled exponential linear unit (SELU) activation which allows the network self-normalization, and 2) at the learning phase, multi-task learning with a loss function weighted by the task homoscedastic uncertainty. The results achieve $F_1$-scores close to or above 0.9 for the detection of TB lesions and a Root Mean Square Error of 1.16 for the number of nodules.

Keywords: Self-Normalizing Neural Networks, multi-task, homoscedastic uncertainty

1. Introduction

Tuberculosis (TB) is an infectious disease which generally affects the lungs and has a high incidence and mortality (World Health Organization and others, 2018). Due to the severity of the pandemic, the World Health Organization (WHO) has launched an ambitious plan to eradicate TB by 2030, for which, the extraction of sensitive radiological biomarkers (Nachiappan et al., 2017) is a clear need. Traditionally, radiologists through visual inspection of x-ray Computed Tomography (CT) volumes generate reports summarizing the presence of TB-related manifestations. However, using this approach for the extraction of robust TB radiological biomarkers, it is unfeasible and automation required. In recent years, the introduction of deep learning techniques has drastically contributed to this task (Wang et al., 2017; Litjens et al., 2017; Hinton, 2018). For deep learning, knowledge is usually injected into the model in the form of manually segmented masks of the lung lesions. Instead, our proposal directly employs the expertise acquired by the radiologist through years of clinical practice as synthesized in tabular reports.

2. Material and Methods

Chest CT scans (56) acquired from 14 male Cynomolgus macaques at 3, 7, 11 and 16 weeks after TB aerosol exposure were employed. First, the CT volumes were cropped (Gordaliza et al., 2018) and resampled to 1mm × 1mm × 2mm. During the training, data augmentation is performed (elastic transformation, addition of Gaussian noise). As labelled data, we employ the tabular reports elaborated by a radiologist (20 years experienced) that contain the number of detected nodules (0 – 15) and binary annotations indicating the presence per lung lobe of the most common TB manifestations (e.g., cavitations, conglomerations, consolidations and trees in bud) (Nachiappan et al., 2017).

2.1. Model Architecture

Our implementation (Figure 1) extracts latent features from the volumes employing the V-Net (Milletari et al., 2016) encoder. The extracted features are used as input of the Feed-Forward Neural Networks (FNNs) for multi-class classification. The encoder generates 1,376.256 features, which feed FNN \(_1\) (task-shared). The outputs of FNN \(_1\) are employed as the input of two independent FNNs, FNN \(_R\) and FNN \(_B\), corresponding with regression (nodule counting) and binary tasks. Dropout or Batch Normalization (BN) (Ioffe and Szegedy, 2015) layers are included where is needed. When employing SNN, BN is unnecessary.

![Figure 1: 3D-CNN + three FNNs (Feed Forward neural networks): FNN\(_1\) (tasks-shared parameters) and, FNN\(_R\) and FNN\(_B\) for prediction of regression and binary tasks. BN is not present with SELU.](image)

2.2. Self-Normalizing Neural Networks

For regularization and normalization purposes, we apply the Self-Normalizing Neural Networks (SNNs) strategy to our model (Klambauer et al., 2017) that preserves the activation magnitude close to zero mean and unit variance. As activation function, we use the Scaled Exponential Linear Unit (SELU) and as dropout, the \(\alpha\)-dropout (Klambauer et al., 2017).

2.3. Learning Principle: Uncertainty Weighted Multi-task Loss

When working in multi-task classification, the decision on the influence of each task in the final loss is non-trivial. Traditionally, the loss is computed as \(\mathcal{L} = \sum_i w_i \mathcal{L}_i\), where the
weights \(w_i\) are selected either manually or after an exhaustive grid search. This approach is highly influenced by the units and the scale of each task and extremely time-consuming. Recently, (Kendall et al., 2018) et al. proposed a method to compute the weights guiding each specific task loss by homoscedastic uncertainty of the predictions. In this work, we adapt this approach to our multi-label classification problem and derive a loss function for our model (Figure 1).

3. Experiments and Results

For the evaluation, we employed 5-fold Cross-Validation (CV): Validation set (4 CT volumes of one subject); training set (remaining 13 subjects). We compared the proposed model (SELU) and a modified version which employs a Parametric Rectified Linear Unit (PReLU), Batch Normalization and standard dropout (\(BN+PReLU\)). The models are trained with 10,000 iterations of the ADAM optimizer (Kingma and Ba, 2014) (learning rate= \(10^{-5}\), mini-batch size= 15). For \(BN+PReLU\), standard parameters were employed for the ADAM optimizer and a dropout rate of 0.5. For SELU: \(\beta_2 = 0.9\), \(\epsilon = 0.01\) and alpha-dropout= 0.1 (Klambauer et al., 2017). Figure 2 shows for the validation data, the loss at each fold for the \(BN+PReLU\) (in red) and the proposed model (in blue). It can be observed that SELU presents improved convergence. The inference error is estimated by the Root Mean Square Error (RMSE) for the nodules count tasks and the \(F_1\)-score for the twenty binary tasks. Table 1 presents the results per fold. No significant statistical differences were found for paired t-test (\(p \geq 0.05\)). Nevertheless, the proposed model presents better \(RMSE\) and \(F_1\)-score results.

<table>
<thead>
<tr>
<th>Manifestation/Field</th>
<th>Nodules [RMSE]</th>
<th>Cavitations ([\beta]_1)</th>
<th>Conglomeration ([\beta]_2)</th>
<th>Consolidation ([\beta]_3)</th>
<th>Tree in bud ([\beta]_4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BN+PReLU SELU</td>
<td>BN+PReLU SELU</td>
<td>BN+PReLU SELU</td>
<td>BN+PReLU SELU</td>
<td>BN+PReLU SELU</td>
<td>BN+PReLU SELU</td>
</tr>
<tr>
<td>1</td>
<td>0.73 ± 0.04</td>
<td>0.85 ± 0.15</td>
<td>0.89 ± 0.11</td>
<td>0.86 ± 0.12</td>
<td>0.90 ± 0.13</td>
</tr>
<tr>
<td>2</td>
<td>1.15 ± 0.09</td>
<td>1.09 ± 0.33</td>
<td>0.86 ± 0.23</td>
<td>0.89 ± 0.22</td>
<td>0.94 ± 0.17</td>
</tr>
<tr>
<td>3</td>
<td>0.41 ± 0.34</td>
<td>0.22 ± 0.39</td>
<td>0.85 ± 0.12</td>
<td>0.87 ± 0.11</td>
<td>0.90 ± 0.19</td>
</tr>
<tr>
<td>4</td>
<td>1.22 ± 0.16</td>
<td>0.76 ± 0.24</td>
<td>0.94 ± 0.15</td>
<td>0.91 ± 0.19</td>
<td>0.94 ± 0.11</td>
</tr>
<tr>
<td>5</td>
<td>0.41 ± 0.28</td>
<td>0.47 ± 0.18</td>
<td>0.93 ± 0.17</td>
<td>0.94 ± 0.17</td>
<td>0.96 ± 0.12</td>
</tr>
<tr>
<td>Total</td>
<td>0.78 ± 0.69</td>
<td>0.68 ± 0.62</td>
<td>0.89 ± 0.16</td>
<td>0.90 ± 0.17</td>
<td>0.94 ± 0.16</td>
</tr>
</tbody>
</table>

Table 1: Performance measure results the for BN+PReLU and the proposed model (SELU).

Figure 2: Evolution of the validation loss for each fold of the Cross-Validation.

4. Conclusions

Although further validation on large datasets is needed, the work presents a promising inference of the radiologist reports. This is achieved with a reduced computational complexity by avoiding normalization layers and hyperparameter tuning of the loss weights.
Acknowledgments

The research leading to these results received funding from the Innovative Medicines Initiative (www.imi.europa.eu) Joint Undertaking under grant agreement no. 115337, whose resources comprise funding from EU FP7/2007-2013 and EFPIA companies in kind contribution. This work was partially funded by projects RTC-2015-3772-1, TEC2015-73064-EXP and TEC2016-78052-R from the Spanish Ministry of Economy, Industry and Competitiveness (MEIC), TOPUS S2013/MIT-3024 project from the regional government of Madrid and by the Department of Health, UK.

This material is based upon work supported by Google Cloud Platform.

References


