Introduction

Background

The pathological information from the two paired views (i.e., medio-lateral oblique and cranio-caudal) are highly relational and complementary, which is crucial for diagnosis in clinical practice.

How to model clinical experience?

**Motivation:** Radiologists take the reasoning procedure explicitly: 1. Extract suspicious regions in the examined view; 2. Search the regions in the other view to make comprehensive decisions.

**Proposed method:** We imitate this process and propose to extract discriminating correlative features, using relation agent to operate two views' complementary information, structuring representations via iterated, message-passing-like modes of processing.

Methodology

**Fig. 2.** The architecture of our CVR-RCNN framework. A paired input image set is fed into the two-branch Faster RCNNs to get the ROIs. The visual and geometry features of ROIs are used by cross-view relation network to learn the effective relation features.

**Two-branch Faster RCNNs:** A two pathway architecture is applied to extract the discriminating correlative features from each representative view. The backbone network adopt two weight-shared Faster-RCNN connected by several relation blocks.

**Cross-view Relation Networks:** The objective of the relation network is to transfer both semantic and geometric information of ROIs from the second (or first) view to the first (or second) one to help detect masses more effectively.

**Channel-wise attention module:** A channel-wise attention based feature aggregation mechanism that supervised by classification loss to reweight the feature maps of different views.

Experiments

**Table 1.** Comparisons with state-of-the-art methods with true positive rate (TPR) versus FPI on the public DDSM dataset.

<table>
<thead>
<tr>
<th>Methods</th>
<th>$F_1$ score</th>
<th>TPR/FPI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Campanini et al. [25]</td>
<td>-</td>
<td>0.80/0.1</td>
</tr>
<tr>
<td>Ellonci et al. [26]</td>
<td>0.92/0.5/4, 0.88/0.2/4, 0.81/0.06</td>
<td></td>
</tr>
<tr>
<td>Sumpor et al. [27]</td>
<td>0.88/0.2/5, 0.85/0.15, 0.80/0.1</td>
<td></td>
</tr>
<tr>
<td>Li et al. [29]</td>
<td>0.90/0.3/4, 0.87/0.10, 0.84/0.10</td>
<td></td>
</tr>
<tr>
<td>Min et al. [3]</td>
<td>0.89/0.2/3, 0.86/0.17, 0.78/0.1</td>
<td></td>
</tr>
<tr>
<td>Yen et al. [1]</td>
<td>0.87/0.2/3, 0.86/0.17, 0.78/0.1</td>
<td></td>
</tr>
<tr>
<td>Faster RCNN [22]</td>
<td>0.52</td>
<td>0.75/0.12</td>
</tr>
<tr>
<td>two-branch Faster RCNNs</td>
<td>0.57</td>
<td>0.75/0.12</td>
</tr>
<tr>
<td>CVR-RCNN</td>
<td>0.73</td>
<td>0.92/0.2/2, 0.88/0.19, 0.85/0.12</td>
</tr>
</tbody>
</table>

**Table 2.** Effect of relation block(s) in the cross-view relation network on the private dataset. $N = 0$ corresponds to the two-branch faster rcnns without relation blocks.

<table>
<thead>
<tr>
<th>Relation Block ($N$)</th>
<th>Precision (%)</th>
<th>Recall (%)</th>
<th>$F_1$ score</th>
<th>FPI</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N = 0$</td>
<td>65.27</td>
<td>71.93</td>
<td>0.69</td>
<td>0.42</td>
</tr>
<tr>
<td>$N = 1$</td>
<td>69.66</td>
<td>71.70</td>
<td>0.71</td>
<td>0.55</td>
</tr>
<tr>
<td>$N = 2$</td>
<td>70.10</td>
<td>72.13</td>
<td>0.71</td>
<td>0.53</td>
</tr>
<tr>
<td>$N = 3$</td>
<td>71.12</td>
<td>75.33</td>
<td>0.73</td>
<td>0.30</td>
</tr>
<tr>
<td>$N = 4$</td>
<td>76.56</td>
<td>70.39</td>
<td>0.73</td>
<td>0.27</td>
</tr>
</tbody>
</table>

**Table 3.** Effect of design loss in the cross-view relation network on the private dataset.

<table>
<thead>
<tr>
<th>COCNet</th>
<th>MSE/SSIM</th>
<th>Dice/SSIM</th>
<th>Precision (%)</th>
<th>Recall (%)</th>
<th>F1-Score</th>
<th>FPI</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>2.1</td>
<td>68.18</td>
<td>72.13</td>
<td>0.49</td>
<td>0.40</td>
<td></td>
</tr>
<tr>
<td>1.2</td>
<td>1.2</td>
<td>68.54</td>
<td>72.86</td>
<td>0.50</td>
<td>0.37</td>
<td></td>
</tr>
<tr>
<td>2.1</td>
<td>1.1</td>
<td>71.72</td>
<td>75.33</td>
<td>0.73</td>
<td>0.30</td>
<td></td>
</tr>
<tr>
<td>2.1</td>
<td>1.1</td>
<td>70.22</td>
<td>74.46</td>
<td>0.52</td>
<td>0.35</td>
<td></td>
</tr>
<tr>
<td>1.2</td>
<td>1.1</td>
<td>71.53</td>
<td>72.71</td>
<td>0.52</td>
<td>0.33</td>
<td></td>
</tr>
<tr>
<td>1.2</td>
<td>1.1</td>
<td>69.37</td>
<td>71.26</td>
<td>0.49</td>
<td>0.35</td>
<td></td>
</tr>
<tr>
<td>1.2</td>
<td>1.2</td>
<td>69.76</td>
<td>71.99</td>
<td>0.50</td>
<td>0.36</td>
<td></td>
</tr>
</tbody>
</table>