Classification of Usual Interstitial Pneumonia Pattern on Computed Tomography using Spin Image Texture Descriptors

Stephen Humphries MS¹, Kunihiro Yagihashi MD², Joyce Schroeder MD, Rohit Sood MD, PhD³ and David Lynch MB² for the IPFNet Investigators
¹Dept. of Bioengineering, UC Denver ²Dept. of Radiology, National Jewish Health, Denver, CO ³PAREXEL Informatics, Billerica, MA

Rationale

• Visual identification of a usual interstitial pneumonia (UIP) pattern on CT is subject to high inter-observer variation [1]
• Effective computer methods for analysis of UIP on CT would be valuable
• Quantifying UIP pattern (Fig. 1) requires measurement of patterns in both pixel intensities and their spatial arrangement, i.e. image texture
• Our goal is to develop and test a texture-based algorithm capable of quantifying extent of UIP on CT

Methods - Training

• Volumetric CT of 284 subjects enrolled in the IPFNet ACE [3] were used for training
• Regions of interest (ROIs) demonstrating normal lung and key features of a UIP pattern were outlined by an experienced radiologist and reviewed by a second
• Points of interest (POIs) were sampled within ROIs and Spin Image features (Fig. 2) computed for each
• A Random Forest (RF) classifier [2] was trained using spin image features from labeled POIs
• Leave-one-out cross validation (CV) was used to validate RF training (Fig. 3)

Results - Training

• A total of 900 ROIs were collected for training
• CV results show the RF classifier correctly distinguished fibrotic from non-fibrotic lung in 97.4% of POIs (Table 1)
• Qualitative testing (Fig. 4) shows good agreement

Table 1 - Cross validation results

<table>
<thead>
<tr>
<th>Radiologist label</th>
<th>Fibrotic</th>
<th>Non-fibrotic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity</td>
<td>98.4%</td>
<td>98.2%</td>
</tr>
<tr>
<td>Specificity</td>
<td>93.7%</td>
<td>94.4%</td>
</tr>
</tbody>
</table>

Results - Volumetric Testing

• Radiologist visual scores (extent of reticular abnormality + honeycombing) showed relatively poor inter-observer agreement (weighted $\kappa = 0.3$, $p < 0.001$)
• Volumetric testing showed moderate correlation between % of lung classified as fibrotic, mean visual score and physiologic measures (Table 2)

Table 2 - Pearson Correlations (p<0.001)

<table>
<thead>
<tr>
<th>Measure</th>
<th>Visual Score</th>
<th>Algorithm Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Score</td>
<td>(95% Conf. Interval)</td>
<td>(95% Conf. Interval)</td>
</tr>
<tr>
<td>Visual Score</td>
<td>0.38 (0.28, 0.48)</td>
<td>-</td>
</tr>
<tr>
<td>Algorithm Score</td>
<td>-0.48 (-0.56, -0.38)</td>
<td>-0.55 (-0.63, -0.46)</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>-0.54 (-0.62, -0.46)</td>
<td>-0.61 (-0.67, -0.53)</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>-0.21 (-0.31, -0.09)</td>
<td>-0.54 (-0.62, -0.46)</td>
</tr>
</tbody>
</table>

Conclusions

• Observer agreement on visual scoring of extent of fibrosis is relatively poor
• Automatic texture-based classification of a UIP pattern on CT can quantify extent of fibrosis in subjects enrolled in a multicenter trial, scanned with a variety of CT scanners
• This technique should permit objective evaluation of disease progression or response to treatment
• Further validation, including evaluation of sequential scans, is ongoing

Acknowledgments:
This work was supported in part by the Colorado Bioscience Discovery Evaluation Grant Program (10BGF-21) and the National Heart, Lung and Blood Institute (IPF Network)

References: