Detection of urban trees in multiple-source aerial data (optical, infrared, DSM)

L. Pibre\textsuperscript{1,4}, M. Chaumont\textsuperscript{1,2}, G. Subsol\textsuperscript{1}, D. Ienco\textsuperscript{3} and M. Derras\textsuperscript{4}

\textsuperscript{1}LIRMM, CNRS/University of Montpellier, France, \textsuperscript{2}University of Nîmes, France, \textsuperscript{3}IRSTEA, \textsuperscript{4}Berger-Levrault company, Toulouse, France

lionel.pibre@lirmm.fr

Introduc}\textsuperscript{1}ion

- Deep Learning \cite{lecun15} methodology for localization of urban trees in multiple-source aerial data,
- Evaluation of Convolutional Neural Networks (CNNs) on this task,
- Comparison to standard machine learning methods that exploit hand-crafted descriptors.

Vaihingen database\textsuperscript{1}

- Channels Red, Green, Near-infrared and DSM
- 1,600 trees annotated on 19 images
- Use of data augmentation to get about 6,000 images “tree” and 40,000 images “other”

\textsuperscript{1} The Vaihingen data set was provided by the German Society for Photogrammetry, Remote Sensing and Geoinformation (DGPF) \cite{cramer10}.

\textbf{Method}

\textbf{Results}

<table>
<thead>
<tr>
<th>Area Strategy</th>
<th>Recall</th>
<th>Precision</th>
<th>F-Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>AlexNet \cite{alexnet}</td>
<td>41.56%</td>
<td>24.28%</td>
<td>30.41%</td>
</tr>
<tr>
<td>GoogleNet \cite{googlenet}</td>
<td>46.99%</td>
<td>29.24%</td>
<td>35.68%</td>
</tr>
<tr>
<td>HOG+SVM \cite{hogsvm}</td>
<td>26.66%</td>
<td>0.95%</td>
<td>1.83%</td>
</tr>
<tr>
<td>HOG+RF \cite{hogrf}</td>
<td>38.67%</td>
<td>7.77%</td>
<td>10.91%</td>
</tr>
</tbody>
</table>

\textbf{Overlap Strategy}

<table>
<thead>
<tr>
<th>Recall</th>
<th>Precision</th>
<th>F-Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>AlexNet \cite{alexnet}</td>
<td>49.28%</td>
<td>22.57%</td>
</tr>
<tr>
<td>GoogleNet \cite{googlenet}</td>
<td>48.96%</td>
<td>25.71%</td>
</tr>
<tr>
<td>HOG+SVM \cite{hogsvm}</td>
<td>21%</td>
<td>1.83%</td>
</tr>
<tr>
<td>HOG+RF \cite{hogrf}</td>
<td>33.47%</td>
<td>10.91%</td>
</tr>
</tbody>
</table>

Conclusions

- HOG descriptor is not sufficient to this task,
- Deep Learning approach gives better results than the standard approach,
- A simple way to deal with multi-source aerial data,
- Two different strategies used for the fusion of bounding boxes,
- For the future:
  - Integration of the localization/detection step directly in the Deep Learning methods,
  - Have better management of multi-source data.

Bibliography

\cite{lecun15, cramer10, alexnet, googlenet, hogsvm, hogrf, vapihingen, sagard16, kolesnikov15}

\textsuperscript{1} The Vaihingen data set was provided by the German Society for Photogrammetry, Remote Sensing and Geoinformation (DGPF) \cite{cramer10}.

\textsuperscript{2} The Vaihingen data set was provided by the German Society for Photogrammetry, Remote Sensing and Geoinformation (DGPF) \cite{cramer10}.