4. Convolutional Neural Networks

Convolutional Neural Networks (C-NNs) are hierarchical multi-layered neural networks, which are inspired by biological research results on the visual cortex of mammals. They were introduced in [3] and successfully applied in several computer vision tasks, like optical character recognition [4] or face detection [1, 2].

Kernel based feature extraction is a widely used approach in a lot of computer vision applications. In most cases it is an ad hoc decision which kernel is used to calculate descriptive features. Instead of applying pre-defined task specific feature operators before the classification process, C-NN include the feature extraction into the optimization problem.

During the training stage the error is not only backpropagated within the net itself, but even into the kernel. Therefore, no fixed prior choice for a specific kernel function is made, but the net choose those kernels, which are most suitable for the specific task.

The architecture of a C-NN consists of several convolutional and subsequent sub-sampling layers. The subsampling reduces the computational effort during training and application.

The one and only extension this work makes to standard C-NN is to use complex-valued weights and inputs in all layers. Therefore, complex domain specific learning has to be used.

3. Complex-Valued MLP

Fully-polarimetric SAR measures the amplitude and phase of the backscattered signal in four different combinations of transmitted and received polarisation. Therefore, it forms a complex-valued signal.

Most approaches for object classification are designed for real-valued data and require a projection of complex-valued PoSAR data to the real domain. Those projections make the classification result dependent on the used method. Complex-Valued Multi-Layer Perceptrons (CV-MLPs) can be applied to complex-valued PoSAR data directly and circumvent those dependencies.

In particular the lack of complex functions, which are bounded as well as analytical (Laplace's Theorem), make the extension of standard MLPs to CV-MLPs non-trivial. Nevertheless, several approaches were proposed, for example the usage of spli-complex activation functions as the approach was proposed, for example the approach introduced in [3] and successfully applied in several computer vision tasks, like optical character recognition [4] or face detection [1, 2].

As expected the test error usually decreases with more hidden neurons. A more complex topology lead to a better performance as well. However, the largest decrease of error can be achieved by using more features, meaning more neurons within the convolutional layer. The right side of the 2nd row of Figure 1 shows a visual representation of the classification result.

Figure 1. The images above show details of the classification process. The used data shown at the left side of the first row is a fully-polarimetric SAR image (E-SAR, L-Band, Oberpfaffenhofen). On the right side the reference data (no label (black), forest (dark green), low vegetation (light green), field (yellow), settlement (red), large building/industrial area (pink), street/railway (blue)) is shown. The obtained classification result can be seen on the right side of the second row. All other images are visualisations of the activation of a few exemplary convolutional neurons.

Table 1

<table>
<thead>
<tr>
<th>Topology</th>
<th>Error Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>10-10-6</td>
<td>31% ± 5%</td>
</tr>
<tr>
<td>10-10-6</td>
<td>26% ± 1%</td>
</tr>
<tr>
<td>10-10-6</td>
<td>24% ± 2%</td>
</tr>
<tr>
<td>10-10-6</td>
<td>25% ± 2%</td>
</tr>
<tr>
<td>10-10-6</td>
<td>24% ± 1%</td>
</tr>
<tr>
<td>10-10-6</td>
<td>24% ± 2%</td>
</tr>
</tbody>
</table>

References


7. Conclusion

The presented preliminary results show that CC-NN can be successfully utilized to classify PoSAR data.

Although the net topology with only one convolutional layer is quite simple, the classification performance increased significantly if compared to CV-MLPs.

Indeed, the achieved error rates are comparable to those acquired by applying a speckle reduction technique before classifying with a CV-MLP.

5. Learned Kernels

Figure 1 visualizes the activation of several exemplary neurons of the convolutional layer. As can be seen the activation differs greatly for different neurons. While some of them seem to specialize on local line features for different orientations, others try to capture intensity variations or even polarimetric information.

6. Results

The aim of this paper is to investigate the general applicability of this learning scheme to classification of PoSAR data. Therefore, the number of convolutional layers was restricted to one in this study.

Different numbers of neurons within the convolutional layer, as well as different numbers of hidden layers and neurons per hidden layer were investigated.

Each row of Table 1 denotes different topologies ranging from a single layer net (with six output neurons due to the six-class problem) to a net with two hidden layers each with fifty neurons. The columns mean different numbers of neurons in the single convolutional layer.

As expected the test error usually decreases with more hidden neurons. A more complex topology lead to a better performance as well. However, the largest decrease of error can be achieved by using more features, meaning more neurons within the convolutional layer. The right side of the 2nd row of Figure 1 shows a visual representation of the classification result.

A CV-MLP (two hidden layers with 50 neurons per layer) evaluated on the same data set with identical performance measurements achieved an error rate of 34%.

The introduction of contextual knowledge as well as the usage of learned convolutional features is able to significantly improve the performance.