Bibliography


BIBLIOGRAPHY


[131] Oh Song, H., Xiang, Y., Jegelka, S., Savarese, S.: Deep metric learning via lifted structured


[150] Iscen, A., Tolias, G., Avrithis, Y., Furon, T., Cham, O.: Efficient diffusion on region mani-


BIBLIOGRAPHY


[347] Lu, J., Xiong, C., Parikh, D., Socher, R.: Knowing when to look: Adaptive attention via


# List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Name / Short Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>DCNNs</td>
<td>Deep Convolutional Neural Networks / A regularized version of multilayer perceptrons based on convolution kernels</td>
</tr>
<tr>
<td>CBIR</td>
<td>Content-based Image Retrieval / An image search task according to the content contained in images</td>
</tr>
<tr>
<td>MAC</td>
<td>Maximum Activations of Convolutions / Maximum value over a convolutional feature map</td>
</tr>
<tr>
<td>R-MAC</td>
<td>Regional Maximum Activations of Convolutions / Maximum value over a region on a convolutional feature map</td>
</tr>
<tr>
<td>CroW</td>
<td>Cross-dimensional Weighting / Weighting the activations over different feature maps</td>
</tr>
<tr>
<td>SPoC</td>
<td>Sum-Pooled Convolutional / Sum pooling over different feature maps</td>
</tr>
<tr>
<td>ReLU</td>
<td>Rectified Linear Unit / An activation function returns 0 if it receives any negative input</td>
</tr>
<tr>
<td>SPM</td>
<td>Spatial Pyramid Modeling / An method to model feature in a pyramid way</td>
</tr>
<tr>
<td>t-SNE</td>
<td>t-Distributed Stochastic Neighbor Embedding / A method to visualize high-dimensional data</td>
</tr>
<tr>
<td>RPNs</td>
<td>Region Proposal Networks / A network to obtain proposal for a region or an object</td>
</tr>
<tr>
<td>FC</td>
<td>Fully-Connected (layer)</td>
</tr>
<tr>
<td>KNN</td>
<td>K-Nearest Neighbors</td>
</tr>
<tr>
<td>BoW</td>
<td>Bag-of-Words / Method to embed features according to the number of feature occurrences</td>
</tr>
<tr>
<td>VLAD</td>
<td>Vector of Locally Aggregated Descriptors / Method to embed features based on their residuals w.r.t. each visual word</td>
</tr>
<tr>
<td>FV</td>
<td>Fisher Vector / Method to embed features by using Gaussian mixture model</td>
</tr>
<tr>
<td>GeM</td>
<td>Generalized Mean / A pooling method to apply over each feature map</td>
</tr>
<tr>
<td>CAM</td>
<td>Class Activation Maps / A feature weighting method based on an activated class output</td>
</tr>
<tr>
<td>PCA</td>
<td>Principal Component Analysis</td>
</tr>
<tr>
<td>MMD</td>
<td>Maximum Mean Discrepancy</td>
</tr>
<tr>
<td>FGIR</td>
<td>Fine-Grained Image Retrieval</td>
</tr>
<tr>
<td>RKHS</td>
<td>Reproducing Kernel Hilbert Space</td>
</tr>
<tr>
<td>DKD</td>
<td>Dual Knowledge Distillation / Knowledge distillation based on two teacher models</td>
</tr>
<tr>
<td>GCNs</td>
<td>Graph Convolutional Networks</td>
</tr>
<tr>
<td>VQA</td>
<td>Visual Question Answering / A computer vision task</td>
</tr>
<tr>
<td>KL-divergence</td>
<td>Kullback–Leibler divergence / A metric to measure the distance between two distributions</td>
</tr>
</tbody>
</table>