Analysis Dictionary Learning: An Efficient and Discriminative Solution

Wen Tang, Ashkan Panahi, Hamid Krim, Liyi Dai
Department of Electrical and Computer Engineering
North Carolina State University, Raleigh, NC, USA
1Raytheon Integrated Defense Systems, Tewksbury, MA, USA

Abstract

Discriminative Dictionary Learning (DL) methods have been widely advocated for image classification problems. To further sharpen their discriminative capabilities, most state-of-the-art DL methods have additional constraints included in the learning stages. These various constraints, however, lead to additional computational complexity. We hence propose an efficient Discriminative Convolutional Analysis Dictionary Learning (DCADL) method, as a lower cost Alternative DL method, to preserve the discriminative nature of the image structures and refine the interclass structure representations.

Introduction

In the past decade, sparse representation has been widely used and successfully applied to address a variety of image processing and computer vision problems. One well known approach is the Synthesis Dictionary Learning (SDL), which recovers the signal by learning a dictionary with corresponding coefficients. To overcome the shortcomings of classical patch-based sparse representation and better translation invariance, convolutional filters were also introduced in SDL for signiﬁcant signal processing applications. Due to this success in image processing, SDL has also been explored in image inference problems, such as image classiﬁcation, by augmenting with some supervised learning constraints. Besides SDL, Analysis Dictionary Learning (ADL) has recently been of interest on account of its fast encoding and stability attributes. Inspired by the SDL methodology in image classiﬁcation, ADL has also been adapted to the supervised learning problems by promoting discriminative sparse representations. In all above methods, both the structure of images and the structure between different classes play important roles in the classiﬁcation task. Such structure increases the accuracy, but they also require a substantial amount of computation and time for training and testing. It is hence desirable to forego this potentially costly structure-promoting regularization and to instead embed the discriminating characteristics of ADL methods in the dictionary formulation itself.

This intuitive DCADL framework is deﬁned as follows

\[
\begin{align*}
\textbf{arg min} & \sum_k \sum_{x \in D_k} \left( \| x - H_k \|_2^2 + \lambda \| H_k \|_F^2 \right) \\
\text{s.t.} & \| H_k \|_1 \leq 1; \forall k = 1, \ldots, m, \\
& x \in X_k, \quad \text{where } x \in \mathbb{R}^n \\
& \text{and } \phi_k(x) = \sum_{j=1}^{J} \alpha_j \phi_j(x), \\
& \text{where } \alpha_j \in \mathbb{R}^J \\
& \text{and } \phi_j(x) = \sum_{l=1}^{L} \phi_{jl}(x), \\
& \text{where } \phi_{jl}(x) \text{ is the } j \text{th image, and } \phi_j(x) \in \mathbb{R}^p \\
& \text{is the } j \text{th response map of the } j \text{th image corresponding to the } j \text{th atom}.
\end{align*}
\]

Segment an image \( x \) into \( p \) patches \( \{ x_1, \ldots, x_p \} \) with \( s \times s \) pixels, being of the same size as the atom, and let \( \hat{X} = \{ x_1, \ldots, x_p \} \) be the set of all patches. Let \( H = \left[ \begin{array}{cccc} H_1 & \cdots & H_p \end{array} \right] \in \mathbb{R}^{n \times mp} \). The problem in Eq. (3) can then be rewritten in the same form as in Eq. (2):

\[
\begin{align*}
\textbf{arg min} & \sum_k \sum_{x \in D_k} \left( \| x - H_k \|_2^2 + \lambda \| H_k \|_F^2 \right) \\
\text{s.t.} & \| H_k \|_1 \leq 1; \forall k = 1, \ldots, m, \\
& x \in X_k, \quad \text{where } x \in \mathbb{R}^n \\
& \text{and } \phi_k(x) = \sum_{j=1}^{J} \alpha_j \phi_j(x), \\
& \text{where } \alpha_j \in \mathbb{R}^J \\
& \text{and } \phi_j(x) = \sum_{l=1}^{L} \phi_{jl}(x), \\
& \text{where } \phi_{jl}(x) \text{ is the } j \text{th image, and } \phi_j(x) \in \mathbb{R}^p \\
& \text{is the } j \text{th response map of the } j \text{th image corresponding to the } j \text{th atom}.
\end{align*}
\]

For Yealed: Our proposed DCADL method achieves the highest classification accuracy with the shortest training time and an extremely fast testing time, while securing an at least 1% greater accuracy relative to others’. In the second part of Table 1, a 2-layer hierarchical dictionary learning approach [8] and a 3-layer convolutional network[7] are also included for comparison. Both of these two methods also worked on raw pixels of images.

Experiments and Results

Four widely used visual classiﬁcation datasets have been applied to evaluate our proposed DCADL.

Table 1: Classification Results on Extended YaleB Dataset

<table>
<thead>
<tr>
<th>Method/Number</th>
<th>Accuracy/%</th>
<th>Training Time/s</th>
<th>Testing Time/s</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADL-SvdM[8][1]</td>
<td>83.81</td>
<td>1671</td>
<td>6.78 x 10^3</td>
</tr>
<tr>
<td>2-layer ADL-SvdM[8][2]</td>
<td>81.67</td>
<td>1006</td>
<td>6.29 x 10^3</td>
</tr>
<tr>
<td>3-layer ADL-SvdM[8][3]</td>
<td>80.92</td>
<td>528</td>
<td>6.17 x 10^3</td>
</tr>
<tr>
<td>DCADL[5][4]</td>
<td>90.97</td>
<td>5.09</td>
<td>3.82</td>
</tr>
<tr>
<td>DADL[3][5]</td>
<td>88.91</td>
<td>1.33</td>
<td>1.93 x 10^3</td>
</tr>
</tbody>
</table>

For Yealed: Our proposed DCADL method achieves the highest performance again in our experiments, achieving the fastest training and testing time. Though its accuracy is slightly lower than the reported one in DADL, DCADL is at least 1.5 times faster than DADL in training and testing time.

For Scene 15: Our accuracy is barely lower than LC-KSVD [2] but is still higher than all other methods and the reported performance of LC-KSVD. In addition, compared with all other methods, DCADL still registers a much greater training and testing time gain.

Conclusion

We proposed an efficient discriminative convolutional ADL method for classiﬁcation tasks. Our DCADL consists of learning a convolutional ADL together with a universal linear classiﬁer. We further improved the optimization framework of DCADL to a more efﬁcient discriminative DL framework by eliminating structural constraint costs, while preserving the discriminative power. Our extensive numerical studies show the DCADL exhibits its highly competitive and highly accurate classiﬁcation performance with signiﬁcantly efﬁcient computation.

Acknowledgements

This research was generously supported in part by the U.S. Army Research Ofﬁce under agreement W911NF-16-2-0005.

Reference