Histopathological Image Classification Using Discriminative Feature-Oriented Dictionary Learning

國立中興大學
National ChungHsing University

Institute of Statistics &
Department of Applied Mathematics

$\oint_{\mathrm{NCHU}} f(\mathrm{amath})\, dS_t$

Huynh Huu Dinh
Email: huynhhuudinh@iuh.edu.vn

Histopathological Image Classification Using Discriminative Feature-Oriented Dictionary Learning

5 Author(s) Tiep Huu Vu ; Hojjat Seyed Mousavi ; Vishal Monga ; Ganesh Rao ; U. K. Arvind Rao View All Authors

Abstract:
In histopathological image analysis, feature extraction for classification is a challenging task due to the diversity of histology features suitable for each problem as well as presence of rich geometrical structures. In this paper, we propose an automatic feature discovery framework via learning class-specific dictionaries and present a low-complexity method for classification and disease grading in histopathology. Essentially, our Discriminative Feature-oriented Dictionary Learning (DFDL) method learns class-specific dictionaries such that under a sparsity constraint, the learned dictionaries allow representing a new image sample parsimoniously via the dictionary corresponding to the class identity of the sample. At the same time, the dictionary is designed to be poorly capable of representing samples from other classes. Experiments on three challenging real-world image databases: 1) histopathological images of intraductal breast lesions, 2) mammalian kidney, lung and spleen images provided by the Animal Diagnostics Lab (ADL) at Pennsylvania State University, and 3) brain tumor images from The Cancer Genome Atlas (TCGA) database, reveal the merits of our proposal over state-of-the-art alternatives. Moreover, we demonstrate that DFDL exhibits a more graceful decay in classification accuracy against the number of training images which is highly desirable in practice where generous training is often not available.

Published in: IEEE Transactions on Medical Imaging (Volume: 35 , Issue: 3 , March 2016)

Page(s): 738 - 751

Date of Publication: 26 October 2015

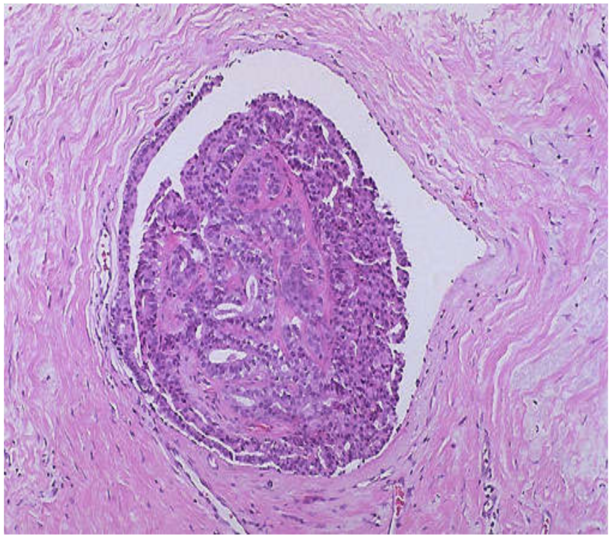
ISSN Information:

INSPEC Accession Number: 15823120

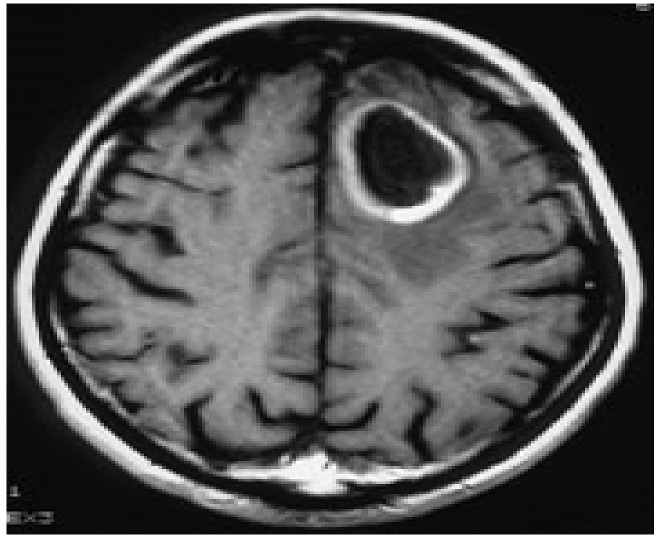
DOI: 10.1109/TMI.2015.2493530

Publisher: IEEE

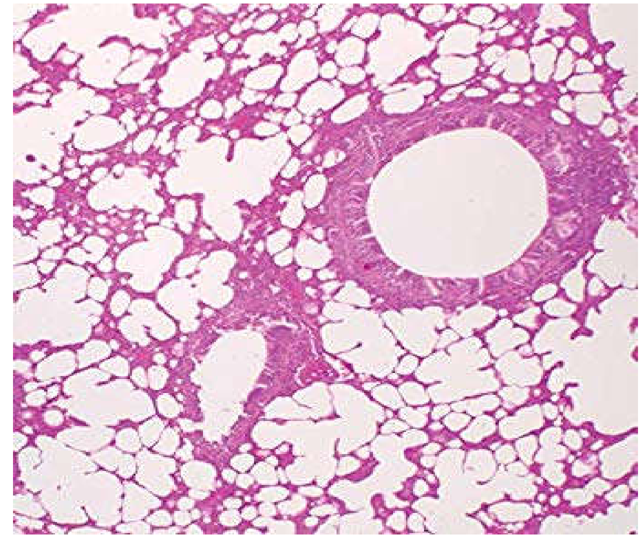
Motivation



Histopathological images of breast lesions (IBL)



Brain tumor images from The Cancer Genome Atlas (TCGA)



Lung images provided by the Animal Diagnostics Lab (ADL) at Pennsylvania State University

Contributions

- A dictionary learning method for automatic feature discovery in histopathological images to mitigate the generally difficult problem of feature extraction in medical images

- Their framework is a discriminative feature-oriented dictionary learning method that emphasizes inter-class differences while keeping intra-class differences small, resulting in enhanced classification performance.

- The proposed method is applied on three diverse histopathological data sets to show the capability of their method in handling a variety of diagnosis and grading problems.

Discriminative Feature - Oriented Dictionary Learning Algorithm

Data: $\mathbf{Y}$, $\bar{\mathbf{Y}}$: collection of all in-class samples and complementary samples.
k: number of learned bases.
ρ: regularization parameter. L: sparsity level
Result: $\mathbf{D}$: dictionary
1. Initializing $\mathbf{D}$ by randomly picking k columns of $\mathbf{Y}$
while *not converged* **do**
  2. Fix $\mathbf{D}$ and update $\mathbf{S}, \bar{\mathbf{S}}$ by solving Problem ($\boxed{2}$);
  3. Fix $\mathbf{S}, \bar{\mathbf{S}}$, calculate:

$$\mathbf{E} = \frac{1}{N}\mathbf{Y}\mathbf{S}^T - \frac{\rho}{\bar{N}}\bar{\mathbf{Y}}\bar{\mathbf{S}}^T; \quad \mathbf{F} = \frac{1}{N}\mathbf{S}\mathbf{S}^T - \frac{\rho}{\bar{N}}\bar{\mathbf{S}}\bar{\mathbf{S}}^T.$$

  if $\mathbf{F}$ *is not PSD* **then**
    $\rho \leftarrow 0.9\rho$; **go to** 3.;
  end
  4. Update $\mathbf{D}$ by solving Problem ($\boxed{4}$);
end

Algorithm 1: DFDL for sparse representation-based classification

Results for IBL Dataset

Class	UDH	DCIS	Method
UDH	91.75	8.25	WND-CHARM
	68.00	32.00	SRC
	93.33	6.67	SHIRC
	84.80	15.20	FDDL
	90.29	9.71	LC-KSVD
	85.71	14.29	Nayak
	96.00	*4.00*	*DFDL*
DCIS	5.77	94.23	WND-CHARM
	44.00	56.00	SRC
	10.00	90.00	SHIRC
	10.00	90.00	FDDL
	14.86	85.14	LC-KSVD
	23.43	76.57	Nayak
	0.50	***99.50***	*DFDL*

Results for ADL Dataset

Class	Not MVP	MVP	Method
Not MVP	76.68 92.92 96.46 92.04 ***94.69***	23.32 7.08 3.54 7.96 *5.31*	WND-CHARM Nayak LC-KSVD FDDL *DFDL*
MVP	21.62 16.22 8.10 18.92 *5.41*	78.38 83.78 91.90 81.08 ***94.59***	WND-CHARM Nayak LC-KSVD FDDL *DFDL*

Results for TCGA Dataset

	Kidney		Lung		Spleen		
Class	Healthy	Inflammatory	Healthy	Inflammatory	Healthy	Inflammatory	Method
Healthy	83.27	16.73	83.20	16.80	87.23	12.77	WND-CHARM$^{(*)}$ [11]
	87.50	12.50	72.50	27.50	70.83	29.17	SRC$^{(*)}$ [24]
	82.50	17.50	75.00	25.00	65.00	35.00	SHIRC [3]
	83.26	16.74	***93.15***	6.85	86.94	13.06	FDDL [31]
	86.84	13.16	85.59	15.41	***89.75***	10.25	LC-KSVD [29]
	73.08	26.92	89.55	10.45	86.44	13.56	Nayak's *et al.* [4]
	88.21	11.79	**96.52**	3.48	**92.88**	7.12	DFDL
Inflammatory	14.22	85.78	14.31	83.69	10.48	89.52	WND-CHARM$^{(*)}$ [11]
	25.00	75.00	24.17	75.83	20.83	79.17	SRC$^{(*)}$ [24]
	16.67	***83.33***	15.00	85.00	11.67	88.33	SHIRC [3]
	19.88	80.12	10.00	***90.00***	8.57	91.43	FDDL [31]
	19.25	81.75	10.89	89.11	8.57	91.43	LC-KSVD [29]
	26.92	73.08	25.90	74.10	6.05	**93.95**	Nayak's *et al.* [4]
	9.92	**90.02**	2.57	**97.43**	7.89	***92.01***	DFDL

$^{(*)}$ Images are classified in whole image level.

Complexity

Method	Complexity	Running time
DFDL	$c^2kN(2d + L^2)$	~ 0.5 hours
LC-KSVD	$c^2kN(2d + 2ck + L^2)$	~ 3 hours
Nayak	$c^2kN(2d + 2qck) + c^2dk^2$	~ 8 hours
FDDL	$c^2kN(2d + 2qck) + c^3dk^2$	> 40 hours