DFDL: DISCRIMINATIVE FEATURE-ORIENTED DICTIONARY LEARNING FOR HISTOPATHOLOGICAL IMAGE CLASSIFICATION

Tiep H. Vu†, Hojjat S. Mousavi†, Vishal Monga†, UK Arvind Rao*, Ganesh Rao*

†Department of Electrical Engineering, Pennsylvania State University, University Park, PA, USA *UT MD Anderson Cancer Center, Houston, TX, USA

Histopathological Image Classification Problems

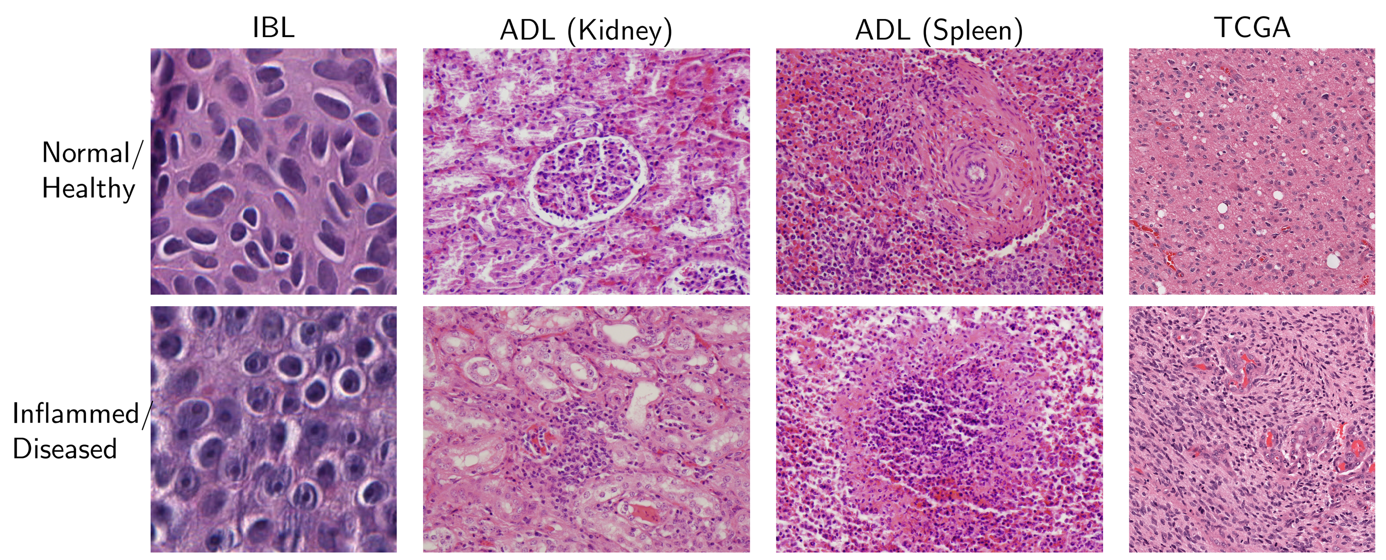


Figure: Samples from three datasets

- Difficulties:
  - Diversity of histology features suitable for each problem
  - Presence of rich geometrical structures
- Pathologists' approaches:
  - Look for problem-specific visual cues:
    - Shape, color, size of cells.
    - Distribution of cells.
    - The presence of a specific set of cells.
- Challenging question: **How to (automatically) extract features?**

Our main contributions

- A **discriminative** dictionary learning method for **automatic feature discovery**.
- Two **low-complexity** procedures for classification and detection problems.
- Extensive experimental results on **three different datasets:**.
  - **IBL** - Intraductal Breast Lesions.
  - **ADL** - Animal Diagnostic Laboratory: Kidney, Lung and Spleen.
  - **TCGA** - The Cancer Genome Atlas: Glioblastoma Multiforme.

Motivation and Problem Formulation

1. **Motivation**
   - A dictionary **D** **sparsely represents** in-class samples (**Y**): $\min_{\|\mathbf{s}_i\|_0 \le L} \|\mathbf{Y} - \mathbf{DS}\|_F^2$ small.
   - But it is incapable of expressing complementary samples ($\bar{\mathbf{Y}}$) with small number of bases:

$$\min_{\|\bar{\mathbf{s}}_i\|_0 \le L} \|\bar{\mathbf{Y}} - \mathbf{D}\bar{\mathbf{S}}\|_F^2 \text{ large}$$

$$V_{L,\varepsilon}(\mathbf{D}) = \{\mathbf{y} : \min_{\|\mathbf{s}\|_0 \le L} \|\mathbf{y} - \mathbf{Ds}\|_2^2 \le \varepsilon\}$$

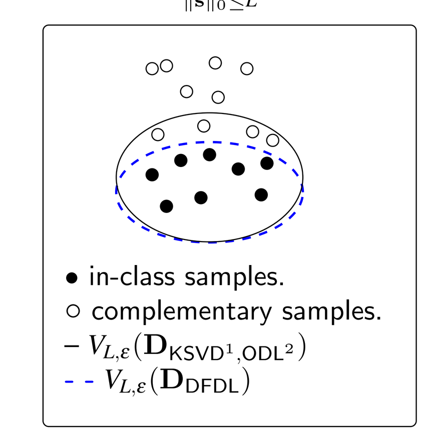


Figure: Main idea

2. **Problem Formulation**

Optimization problem

$$\mathbf{D}^* = \arg\min_{\mathbf{D}} \left(\frac{1}{N} \min_{\|\mathbf{s}_i\|_0 \le L} \|\mathbf{Y} - \mathbf{DS}\|_F^2 - \frac{\rho}{\bar{N}} \min_{\|\bar{\mathbf{s}}_i\|_0 \le L} \|\bar{\mathbf{Y}} - \mathbf{D}\bar{\mathbf{S}}\|_F^2 \right)$$

subject to: $\|\mathbf{d}_j\|_2^2 = 1$

ρ is a regularization parameter.

- Initial **D** could be obtained from:

$$(\mathbf{D}^*, \mathbf{S}^*) = \arg\min_{\mathbf{D},\mathbf{S}} \{\|\mathbf{Y} - \mathbf{DS}\|_F^2 + \lambda\|\mathbf{S}\|_1\} \text{ then choose } L \approx \frac{1}{N}\|\mathbf{S}^*\|_0$$

- $\mathbf{D}^*$ can be found by alternately solving two problems:
  1. fix **D**, find $\mathbf{S}, \bar{\mathbf{S}}$
  2. fix $\mathbf{S}, \bar{\mathbf{S}}$, find **D**

[1] M. Elad et. al., IEEE Trans. on Signal Processing, 2006

[2] J. Mairal et. al., The Journal of Machine Learning Research, 2010

Discriminative Feature-oriented Dictionary Learning (DFDL) Algorithm

Algorithm 1 Learning DFDL Dictionary for each class

INPUT: $\mathbf{Y}, \bar{\mathbf{Y}}, k, \rho$.

1. Choose L and initial **D** by ODL on **Y**

while not converged **do**

2. Fix **D** and update $\mathbf{S}, \bar{\mathbf{S}}$ by solving an OMP problem;
3. Fix $\mathbf{S}, \bar{\mathbf{S}}$, calculate: $\mathbf{E} = \frac{1}{N}\mathbf{YS}^T - \frac{\rho}{\bar{N}}\bar{\mathbf{Y}}\bar{\mathbf{S}}^T$; $\quad \mathbf{F} = \frac{1}{N}\mathbf{SS}^T - \frac{\rho}{\bar{N}}\bar{\mathbf{S}}\bar{\mathbf{S}}^T$.
4. Update **D** from: $\mathbf{D}^* = \arg\min_{\mathbf{D}} \{ -2\mathrm{trace}(\mathbf{ED}^T) + \mathrm{trace}(\mathbf{D}(\mathbf{F} - \lambda_{\min}(\mathbf{F})\mathbf{I})\mathbf{D}^T)\}$

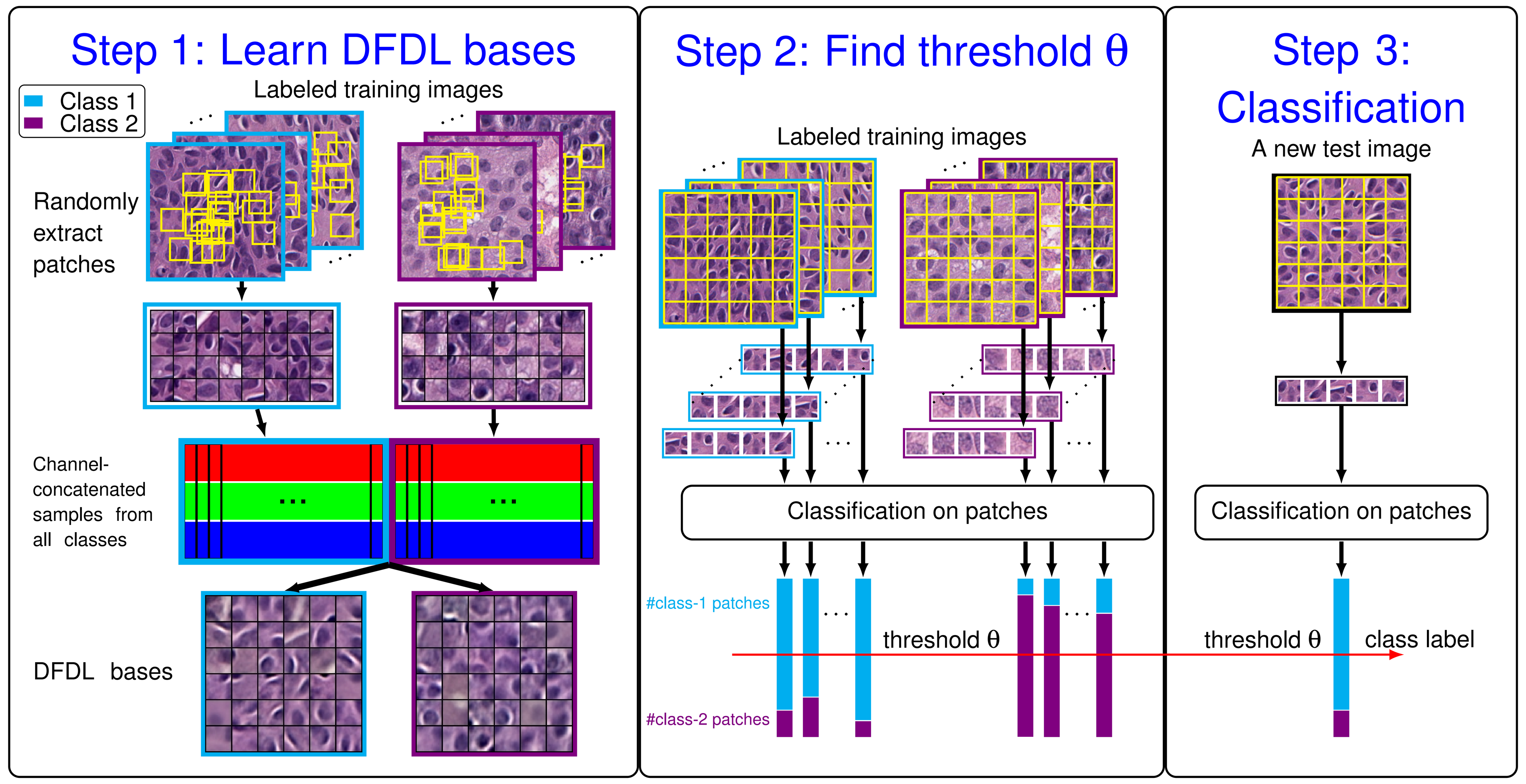
end while

RETURN: D

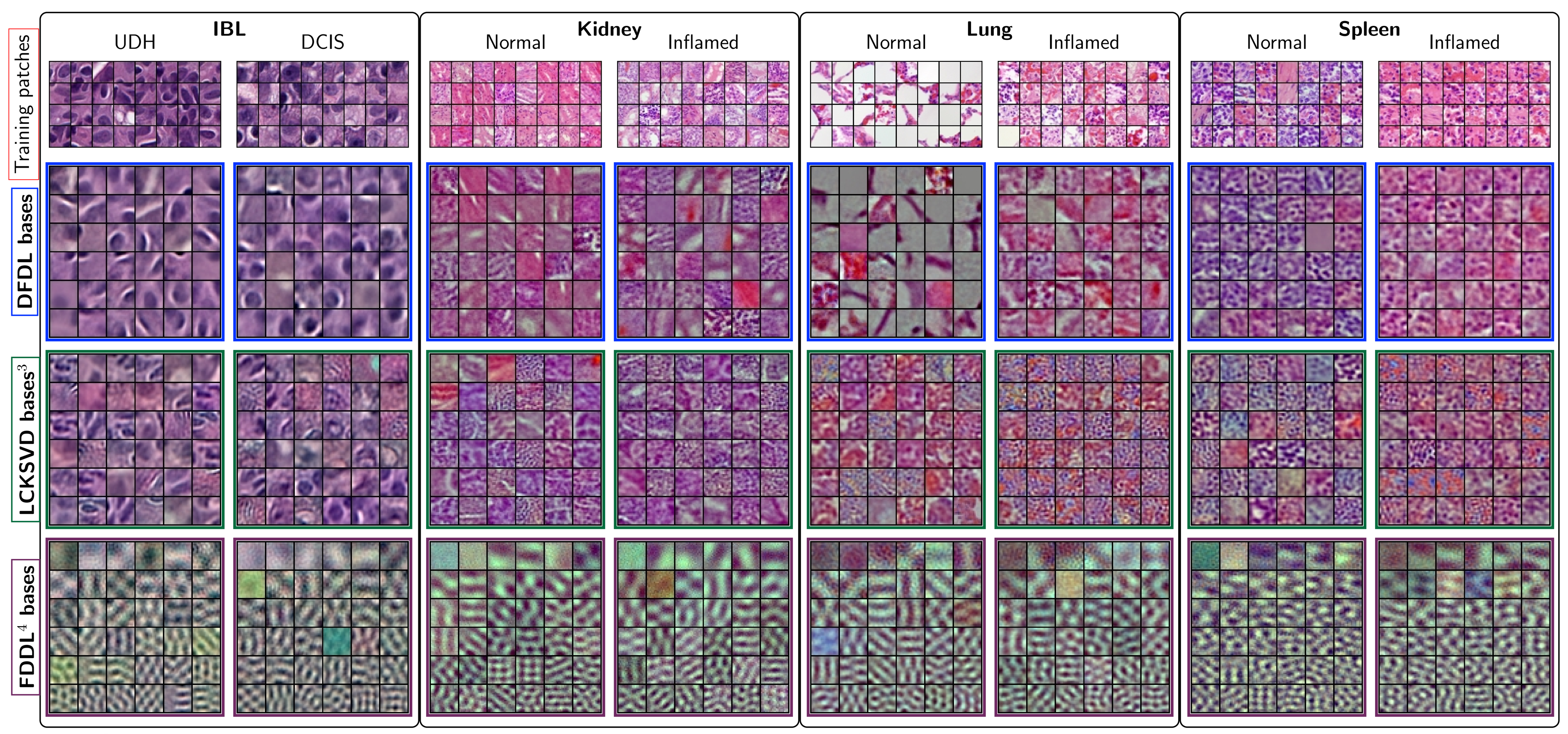
Classification step: 1. Find sparse code: $\hat{\mathbf{s}} = \arg\min_{\mathbf{s}}\{\|\mathbf{y} - [\mathbf{D}_1, \ldots \mathbf{D}_c]\mathbf{s}\|_2^2 + \lambda\|\mathbf{s}\|_1\}$

2. identity($\mathbf{y}$) $= \arg\min_i \|\mathbf{y} - \mathbf{D}_i\delta_i(\hat{\mathbf{s}})\|_2$ where $\delta_i(\hat{\mathbf{s}})$ is the part of $\hat{\mathbf{s}}$ associated with class i.

IBL/ADL classification procedure (proposed for RGB images)



Example bases learned from different Dictionary Learning methods



[3] Z. Jiang et. al., IEEE Trans. on Pattern Analysis and Machine Int.. (TPAMI), 2013; [4] M. Yang et. al., Proc. IEEE Inter. Conf. on Computer Vision (ICCV), 2011

IBL and ADL: Overall accuracies and ROC curves

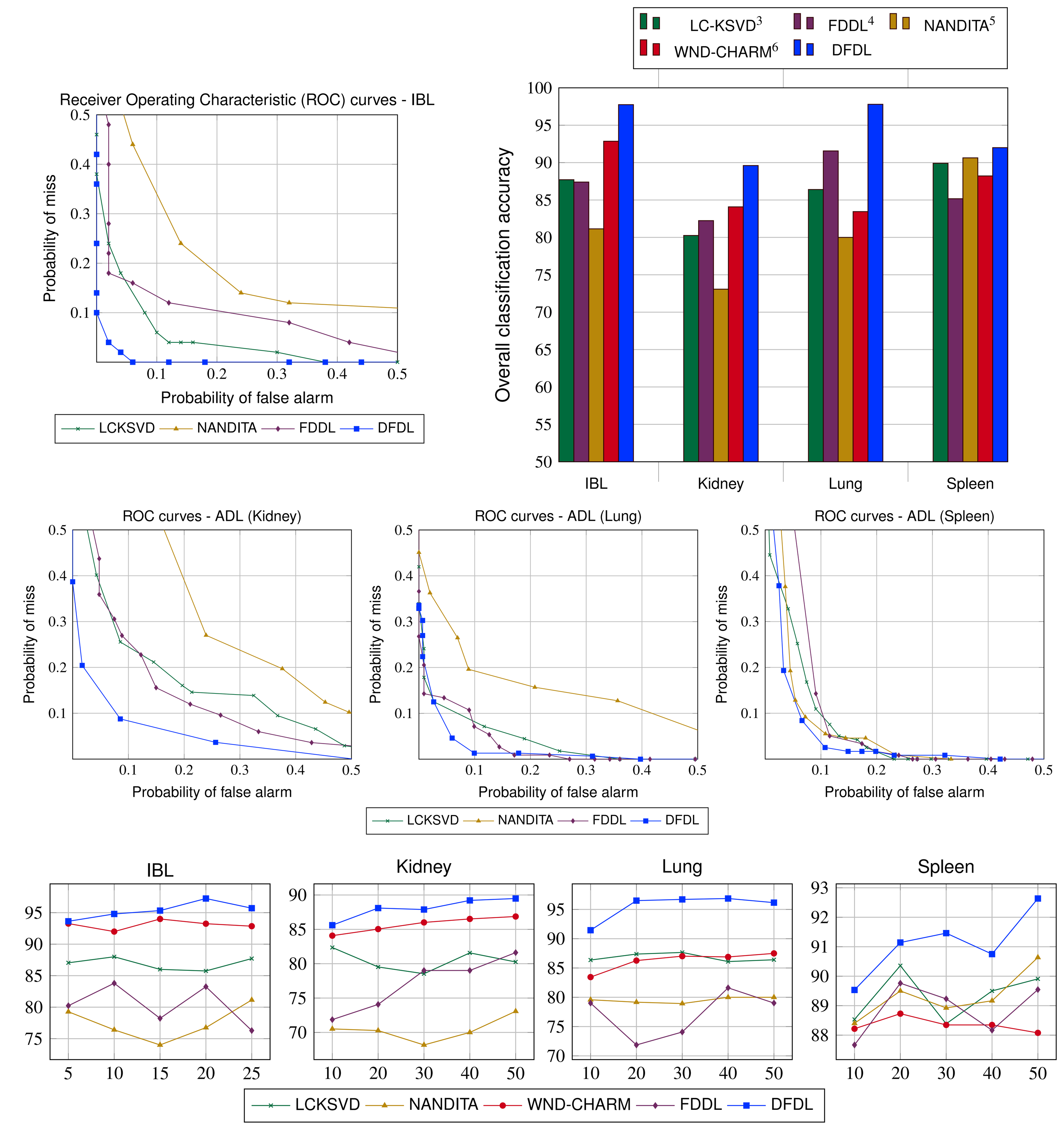


Figure: Overall accuracies over number of training images

[3] Z. Jiang et. al., IEEE Trans. on Pattern Analysis and Machine Int., 2013

[4] M. Yang et. al., Proc. IEEE Conf. on Computer Vision, 2011

[5] N. Nayak et. al., IEEE Int. Symp. Biomed. Imag., 2013

[6] L. Shamir et. al., Source Code Biol. Med., 2008

TCGA classification procedure/results

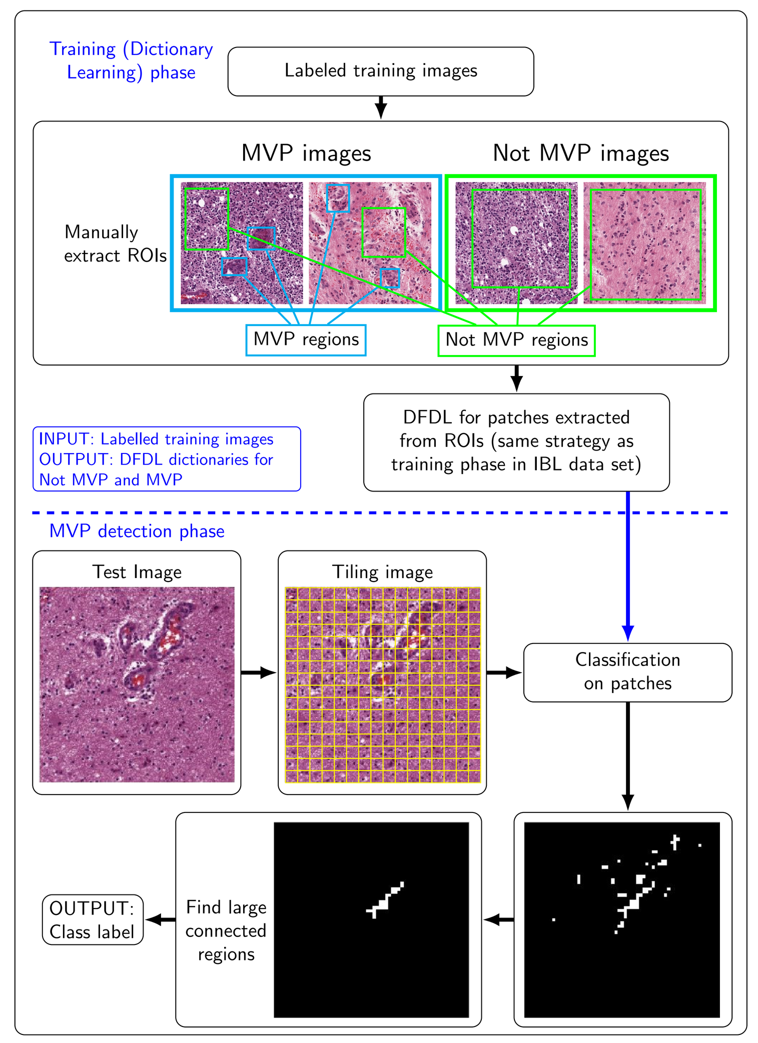


Table 1: CONFUSION MATRIX: TCGA

Class	Not MVP	MVP	Method
Not VMP	76.68	23.32	WND-CHARM
	96.46	3.54	LC-KSVD
	92.92	7.08	NANDITA
	95.57	4.43	DFDL
MVP	21.62	78.38	WND-CHARM
	8.10	**91.90**	LC-KSVD
	16.22	83.78	NANDITA
	10.81	***89.19***	DFDL

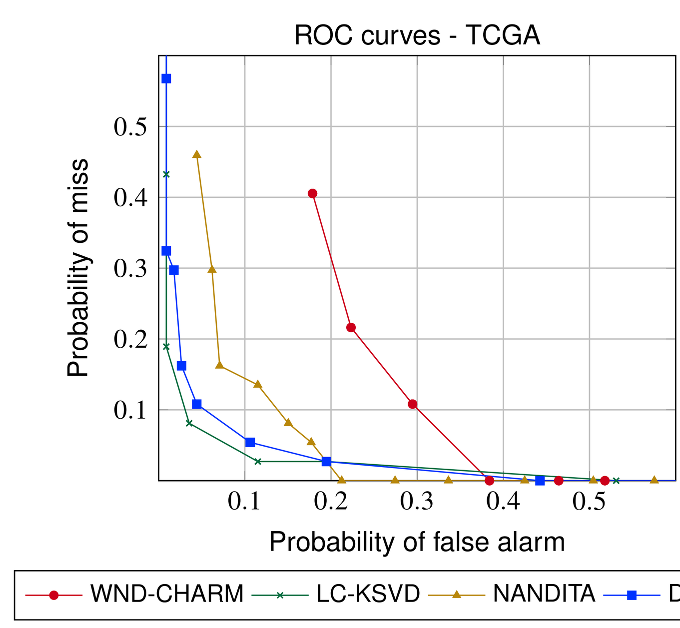


Figure: Microvascular proliferation (MVP) detection procedure and results