Task-Driven Feature Learning for Tumor Histology

Heather D. Couture (heather@cs.unc.edu), J.S. Marron, Nancy Thomas, David Eberhard, Charles Perou, Marc Niethammer – University of North Carolina

Introduction

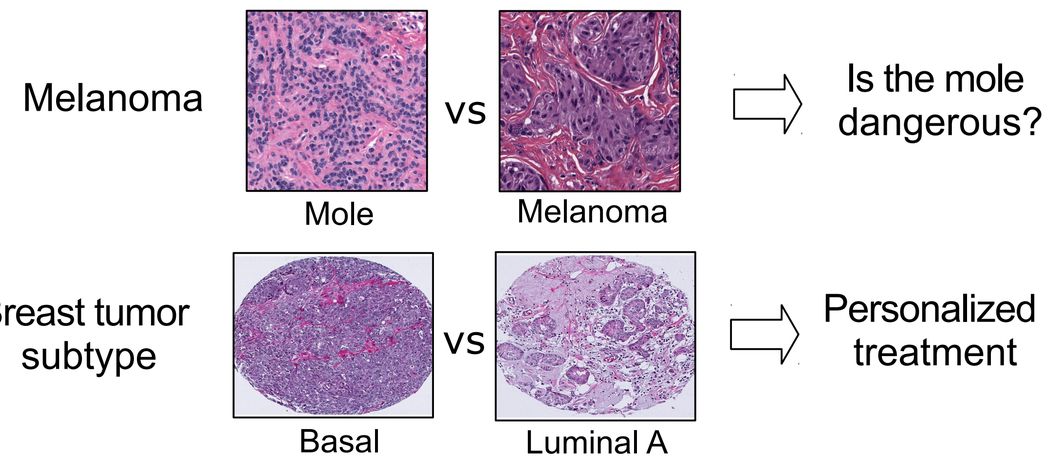
Through learning a dictionary of images features, we can capture the local structure of tumor tissue from histology images. Our work strives to improve the discriminatory power of the dictionary by incorporating class information into the sparse coding formulation. This technique is shown to improve tumor diagnosis and subtyping models and further our goal of interpretable features.

Applications:
- Diagnosis – differentiate between benign and malignant lesions
- Prognosis – identify dangerous tumors
- Subtyping – complement recent methods that target treatment based on molecular analysis

Goal: improve predictions by using automated image analysis
- Faster
- More repeatable
- Capture properties that pathologists cannot

Problem Definition

Image classification:

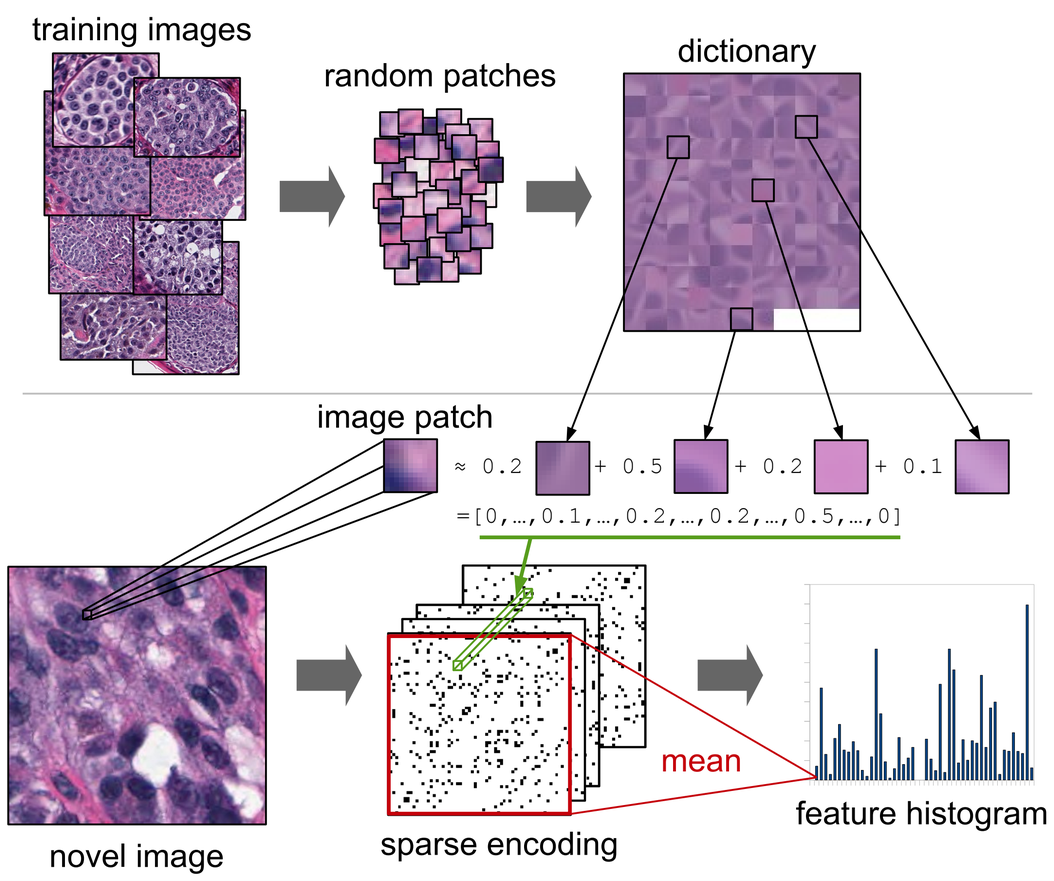


Method: learn dictionary of patch-based features by sparse coding to describe tumor tissue and predict class

Contribution: improve discriminatory power of dictionary by incorporating image labels into dictionary learning process

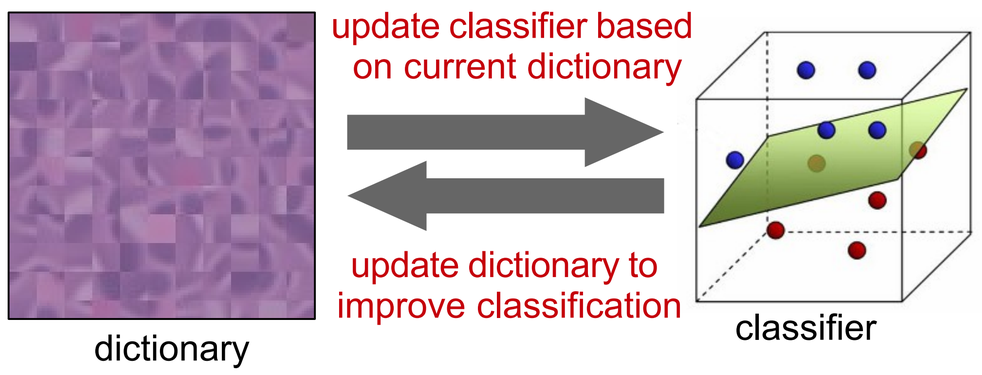
Unsupervised Dictionary Learning

Model image patches as sparse linear combinations of dictionary elements



Task-Driven Dictionary Learning

Tune dictionary to classification task

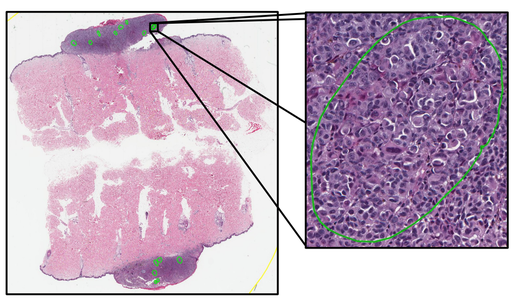


Approach: minimize logistic loss by updating classifier and dictionary using stochastic gradient descent

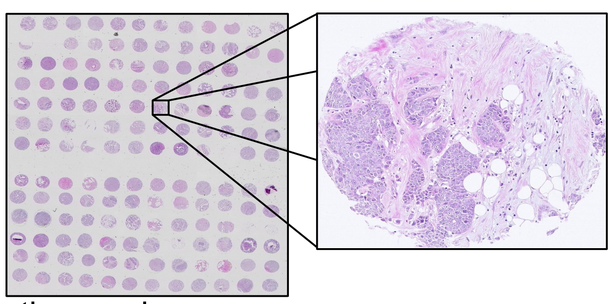
Algorithm:
Initialize dictionary with unsupervised dictionary learning
Initialize classifier with logistic regression on set of patches
Repeat until convergence:
- Select random image region
- Compute sparse encoding and histogram
- Update classifier with gradient descent step
- Update dictionary with gradient descent step

Results

Melanoma vs. mole



Breast tumor subtype



Comparison of cross-validated accuracy for classification applied to individual patches and whole images

	Melanoma vs. mole		Breast tumor subtype Basal vs. Luminal A	
	Unsupervised dictionary	Task-driven dictionary	Unsupervised dictionary	Task-driven dictionary
Patch-level	60%	66%	52%	54%
Image-level	96.2%	TBD	72.9%	TBD

Interpreting classification models

Apply classifier to patches within image to determine which regions are contributing to the predicted class

