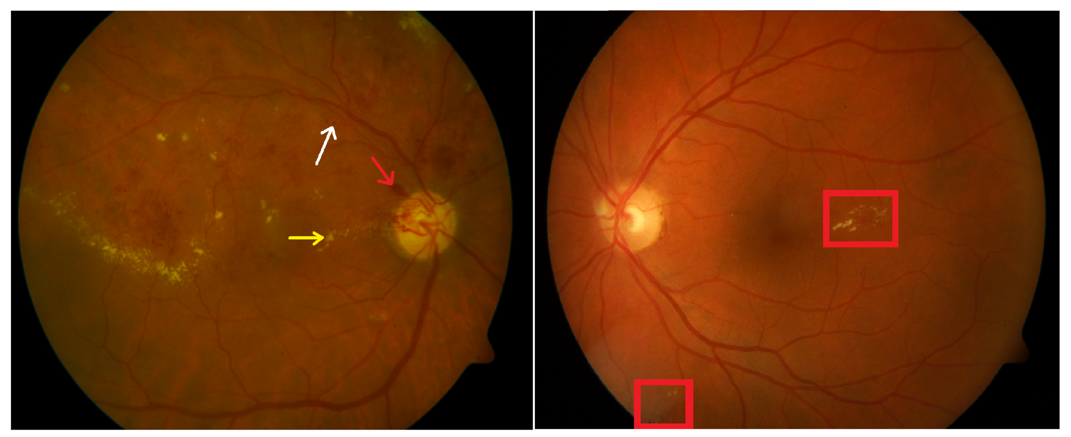
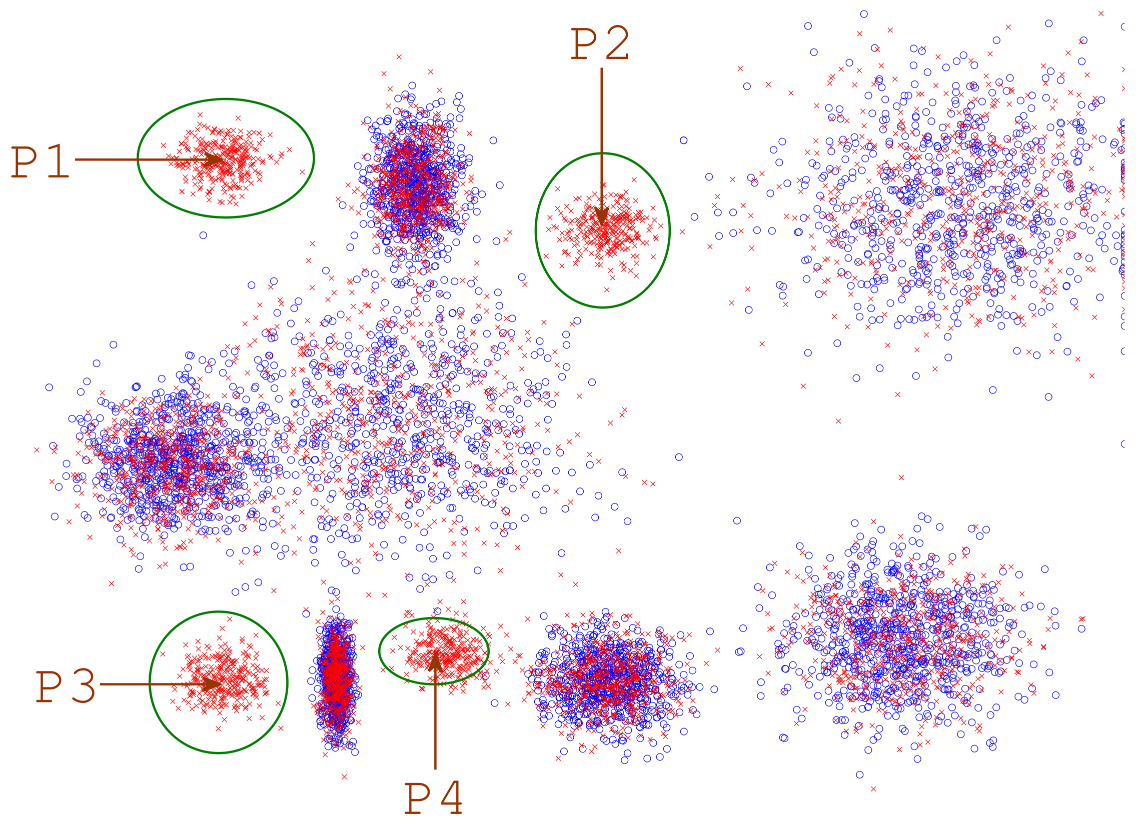
Simple non-parametric methods provide as good or better results to multiple-instance learning.

Ragav Venkatesan, Parag Shridhar Chandakkar, Baoxin Li
Arizona State University.

Multiple Instance Space





$$D = \begin{bmatrix} X^{(1)} & Y^{(1)} \\ X^{(2)} & Y^{(2)} \\ & \vdots \\ X^{(n)} & Y^{(n)} \end{bmatrix}$$

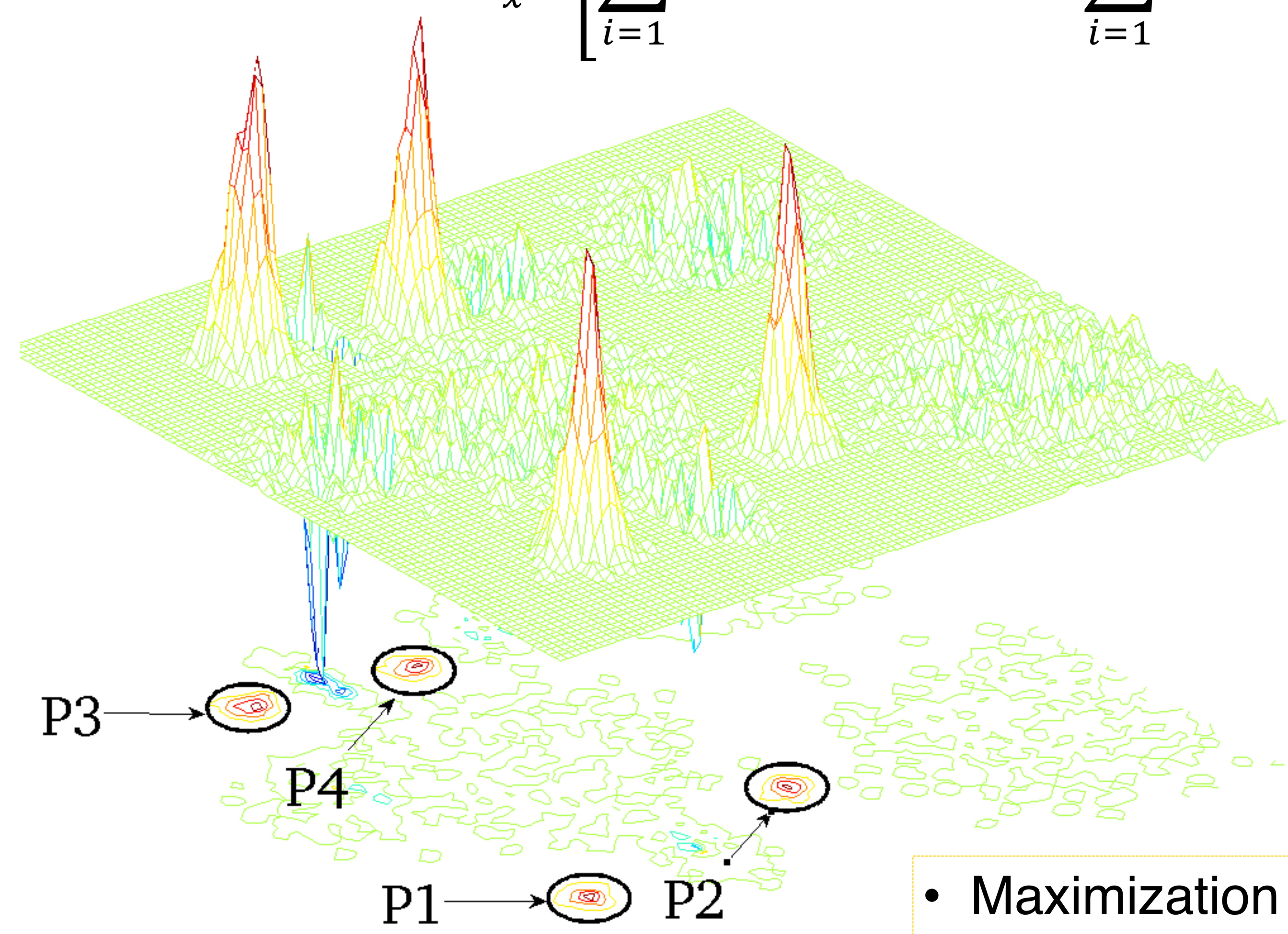
$$X^{(i)} = \{x_1^{(i)}, x_2^{(i)}, \dots x_{m_i}^{(i)}\}$$

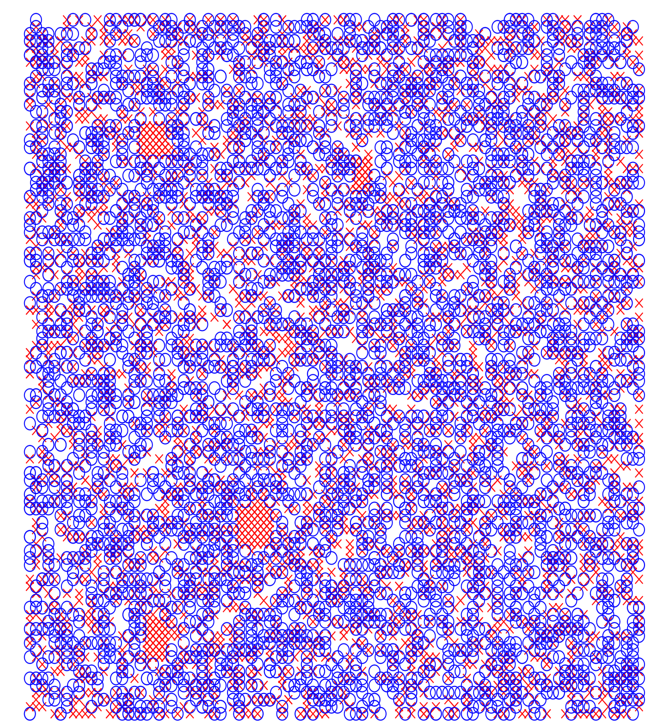
Method

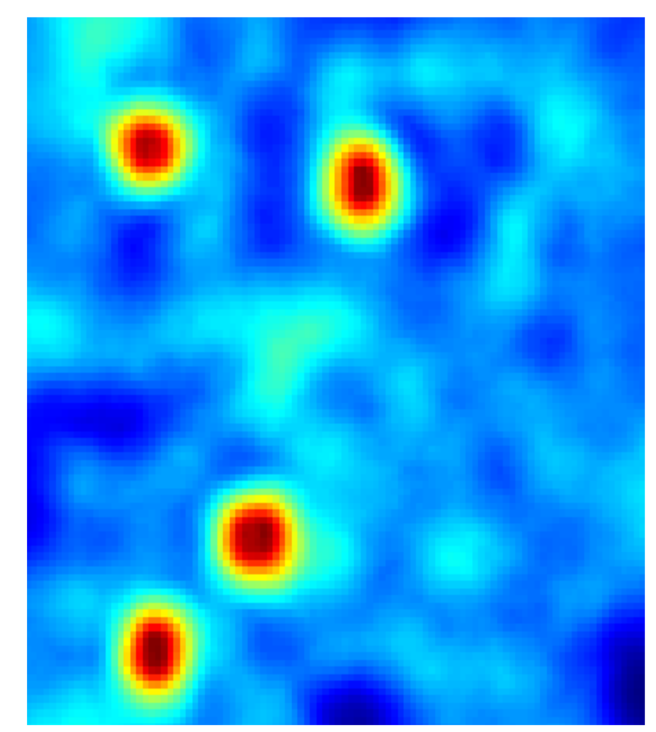
- Bag is a collection of instances.
- Training set has bag-level labels.
- To learn: instance-level labels for unseen and seen bags.

Diverse Density: $$\max_x \prod_i \Pr\left(x = t|B_i^+\right) \prod_i \Pr\left(x = t|B_i^-\right)$$

Proposed Method: $$\max_x \left[\sum_{i=1}^{|k^-|} \Psi\left(\|x - k_i^-\|\right) - \sum_{i=1}^{|k^+|} \Psi\left(\|x - k_i^+\|\right)\right], |k^+| = |k^-| = k$$







- Maximization is performed by from many initial points.
  - All initial points are only instances from positive bags.
  - Y. Chen proposed this methodology of optimization for MIL [3]

Prototyping

- Prototyping is the process of maximizing the density-likes to produce points of interest on the space.
- In a MIL setting a positive prototype is that point on the feature space that maximizes the formulation.
  - This is because the formulation describes the 'positivity condition'. The heat maps show the 'positivity'.
- We can now construct hyper-spherical decision boundaries around these prototypes.

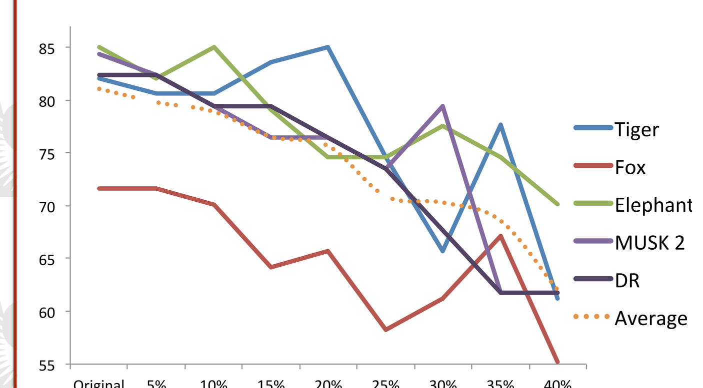
Results

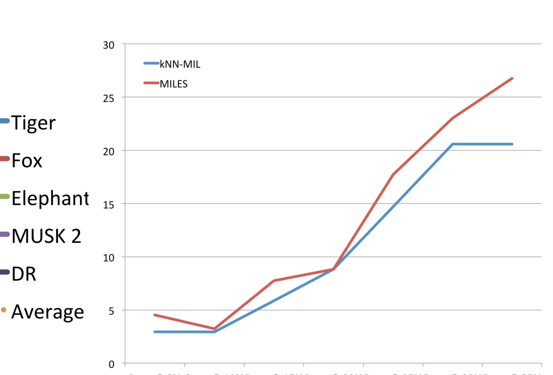
Methods	MUSK 1	MUSK 2
DD	88.9%	82.5%
EM-DD	84.8%	84.9%
citation (*k*)-NN	**92.4%**	86.3%
mi-SVM	87.4%	83.6%
MI-SVM	77.9%	84.3%
DD-SVM	85.8%	**91.3%**
MILES	86.3%	87.7%
MIforest	85%	82%
MILIS	88.6%	91.1%
ISD	85.3%	79.0%
ALP-SVM	87.9%	86.6%
MIC-Bundle	84%	85.2%
Ensemble	89.22%	85.04%
Proposed	**92.4%**	86.4%

Table 1. Performance of various MIL algorithms on the musk dataset.

Methods	Accuracy
DD	61.29%
EM-DD	73.5%
citation *k*-NN	78.7%
mi-SVM	70.32%
MILES	71%
Proposed	**81.3%**

Table 4. Performance of various MIL algorithms on DR dataset.





Related Methods

- Diverse Density
- Expectation-Maximization Diverse Density
- MILES
- ...

All methods in literature are parametric density estimators or parametric embedders

References

[1] O. Maron and T. Lozano-P´erez. A framework for multiple instance learning. NIPS, pages 570–576, 1998.

[2] Q. Zhang and S. Goldman. Em-dd: An improved multiple instance learning technique. Advances in neural information processing systems, 14:1073–1080, 2001.

[3] Y Chen, J. Bi, and J. Z. Wang. Miles: Multiple instance learning via embedded instance selection. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 28(12):1931–1947, 2006.

[4] Q. Wang, L. Si, and D. Zhang. A discriminative data dependent mixture-model approach for multiple instance learning in image classification,. In In Proceedings of the 12th European Conference on Computer Vision (ECCV-12), 2012.

Conclusions

We proposed a simple, yet novel use of MIL featurespace using non-parametric methods that yields as good if not better results than SOTA.

IEEE International Conference on Computer Vision 2015, Santiago, Chile.