Attention-based Deep Multiple Instance Learning

Maximilian Ilse
University of Amsterdam

Jakub M.Tomczak
Qualcomm Research

Max Welling
University of Amsterdam

Presented by **Qianwei Wang**
2018.10.15

About Author

Maximilian Ilse

关注

PhD candidate in Deep Learning, University of Amsterdam
在 uva.nl 的电子邮件经过验证

medical imaging   deep learning   machine learning

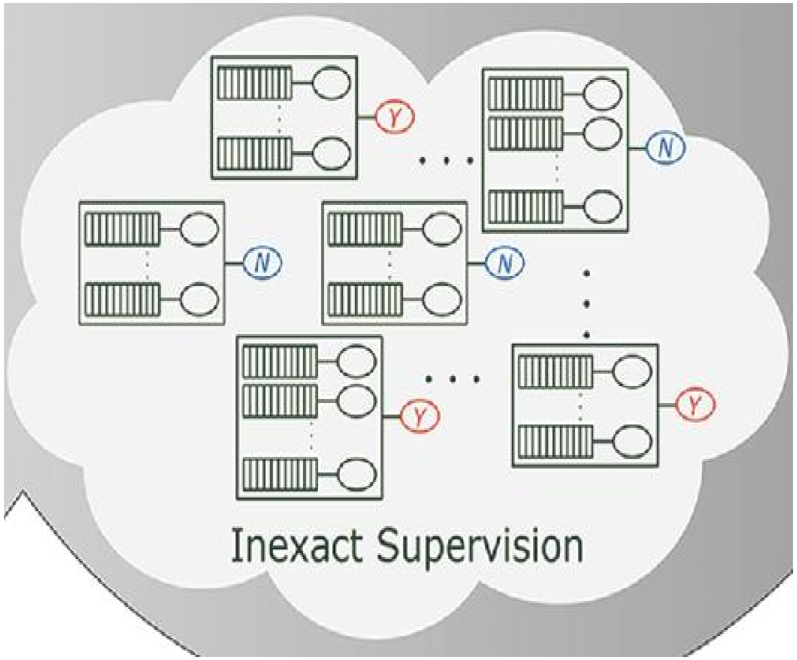
标题	引用次数	年份
Attention-based deep multiple instance learning M Ilse, JM Tomczak, M Welling arXiv preprint arXiv:1802.04712	7	2018
Deep Learning with Permutation-invariant Operator for Multi-instance Histopathology Classification JM Tomczak, M Ilse, M Welling arXiv preprint arXiv:1712.00310	3	2017
Histopathological classification of precursor lesions of esophageal adenocarcinoma: A Deep Multiple Instance Learning Approach JM Tomczak, M Ilse, M Welling, M Jansen, HG Coleman, M Lucas, ...		2018

Outline

- Background
  - Multiple Instance Learning (MIL)
- Proposed Method
  - A General Three-step Approach of MIL
  - MIL with Neural Networks
  - Attention-based MIL pooling
- Experiments
- Conclusion

Background

- Multiple Instance Learning (MIL)
  - The training set is composed of labeled *bags* each consists of many unlabeled instances, and the goal is to predict unseen bags.



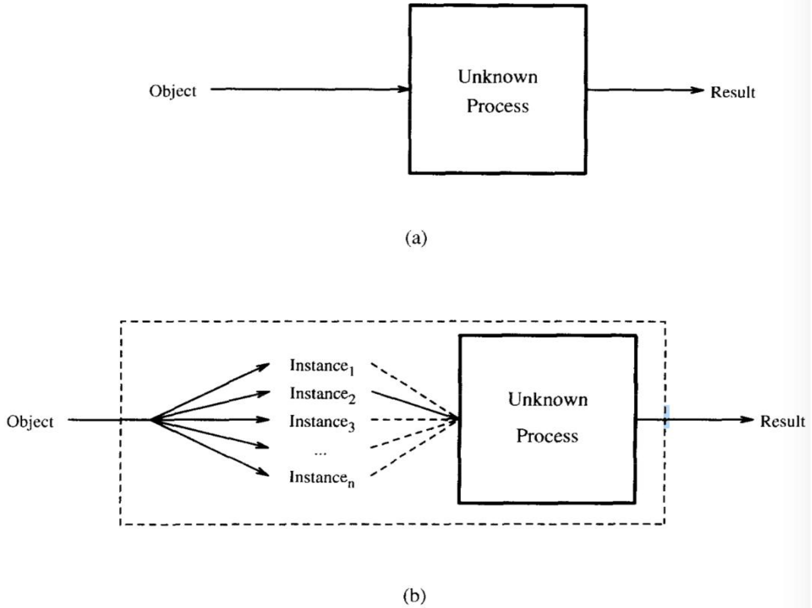


Fig. 1. Supervised learning: (a) usual situation and (b) multiple instance situation.

Proposed Method

- Multi-instance Learning In this Paper:
  - Binary classification;
  - Neither dependency nor ordering among instances.

Three-step Approach of MIL

Neither dependency nor ordering among instances

A **permutation-invariant** scoring function S(X)∈[0, 1]

Theorem 1. *A scoring function for a set of instances X, $S(X) \in \mathbb{R}$, is a symmetric function (i.e., permutation-invariant to the elements in X), if and only if it can be decomposed in the following form:*

$$S(X) = g\Big(\sum_{\mathbf{x} \in X} f(\mathbf{x})\Big), \tag{3}$$

where f and g are suitable transformations.

Three step approach: (i) transformation of instances by function f , (ii) combination of transformed instances by premutation-invariant pooling function σ , (iii) transformation of combined instances by g.

Three-step Approach of MIL

Three step approach: (i) transformation of instances by function f, (ii) combination of transformed instances by premutation-invariant pooling function σ, (iii) transformation of combined instances by g.

$$f + \sigma + g = \text{MIL algorithm}$$

instance-level classifier	+	max-pooling	+	identity function	=	instance-level MIL algorithm
low-dimensional embedding	+	mean-pooling	+	bag-level classifier	=	embedding-level MIL algorithm

MIL with Neural Networks

Learning And Mining from DatA

$$f \quad + \quad \sigma \quad + \quad g \quad = \quad \text{MIL algorithm}$$

Neural Networks   Neural Networks

Advantages: (i) flexible , (ii) can be trained end-to-end

Attention-based MIL Pooling

- Why Attention?
  - Previous pooling: pre-defined, non-trainable;
  - Seek for a flexible and ***interpretable*** pooling method.
- Attention Mechanism

$$\mathbf{z} = \sum_{k=1}^{K} a_k \mathbf{h}_k, \tag{7}$$

where:

$$a_k = \frac{\exp\{\mathbf{w}^\top \tanh\left(\mathbf{V}\mathbf{h}_k^\top\right)\}}{\sum_{j=1}^{K} \exp\{\mathbf{w}^\top \tanh\left(\mathbf{V}\mathbf{h}_j^\top\right)\}}, \tag{8}$$

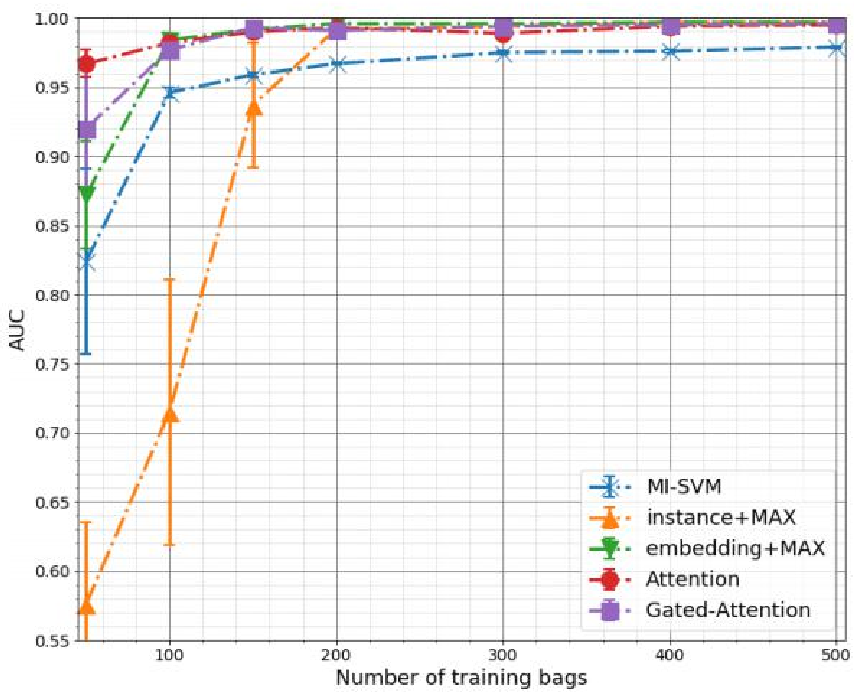
Experiments

- Classical MIL Datasets

	Method	Musk1	Musk2	Fox	Tiger	Elephant
Classical MIL method	mi-SVM [1]	0.874±N/A	0.836±N/A	0.582±N/A	0.784±N/A	0.822±N/A
	MI-SVM [1]	0.779±N/A	0.843±N/A	0.578±N/A	0.840±N/A	0.843±N/A
	MI-Kernel [2]	**0.880**±0.031	**0.893**±0.015	**0.603**±0.028	0.842±0.010	0.843±0.016
	EM-DD [3]	0.849±0.044	**0.869**±0.048	**0.609**±0.045	0.730±0.043	0.771±0.043
	mi-Graph [4]	**0.889**±0.033	**0.903**±0.039	**0.620**±0.044	**0.860**±0.037	**0.869**±0.035
	miVLAD [5]	**0.871**±0.043	**0.872**±0.042	**0.620**±0.044	0.811±0.039	**0.850**±0.036
	miFV [5]	**0.909**±0.040	**0.884**±0.042	**0.621**±0.049	0.813±0.037	**0.852**±0.036
Deep MIL in (Wang et al.)	mi-Net [6]	**0.889**±0.039	**0.858**±0.049	**0.613**±0.035	0.824±0.034	**0.858**±0.037
	MI-Net [6]	**0.887**±0.041	**0.859**±0.046	**0.622**±0.038	**0.830**±0.032	**0.862**±0.034
	MI-Net with DS [6]	**0.894**±0.042	**0.874**±0.043	**0.630**±0.037	**0.845**±0.039	**0.872**±0.032
	MI-Net with RC [6]	**0.898**±0.043	**0.873**±0.044	**0.619**±0.047	**0.836**±0.037	**0.857**±0.040
Proposed method	Attention	**0.892**±0.040	**0.858**±0.048	**0.615**±0.043	**0.839**±0.022	**0.868**±0.022
	Gated-Attention	**0.900**±0.050	**0.863**±0.042	**0.603**±0.029	**0.845**±0.018	**0.857**±0.027

Experiments

- MNIST-bags
  - Constructed by sample images from MNIST as instances
  - A bag is positive if it contains one or more '9'



Experiments

- MNIST-bags

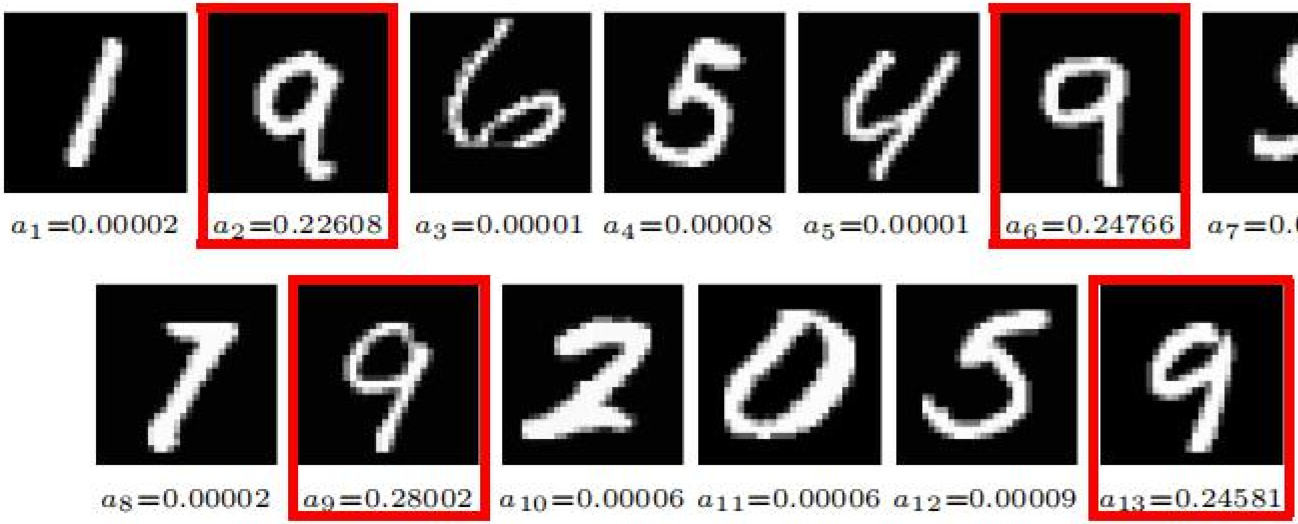


Figure 4. Example of attention weights for a positive bag.

Experiments

- Real-world Dataset
  - Divide every image into 32 x 32 patches.

Table 2. Results on BREAST CANCER. Experiments were run 5 times and an average ($\pm$ a standard error of the mean) is reported.

METHOD	ACCURACY	PRECISION	RECALL	F-SCORE	AUC
Instance+max	0.614±0.020	0.585±0.03	0.477±0.087	0.506±0.054	0.612±0.026
Instance+mean	0.672±0.026	0.672±0.034	0.515±0.056	0.577±0.049	0.719±0.019
Embedding+max	0.607±0.015	0.558±0.013	0.546±0.070	0.543±0.042	0.650±0.013
Embedding+mean	**0.741**±0.023	**0.741**±0.023	0.654±0.054	0.689±0.034	**0.796**±0.012
Attention	**0.745**±0.018	0.718±0.021	**0.715**±0.046	**0.712**±0.025	0.775±0.016
Gated-Attention	**0.755**±0.016	**0.728**±0.016	**0.731**±0.042	**0.725**±0.023	**0.799**±0.020

Conclusion

- General three-step approach for MIL
- Flexible and interpretable MIL model using neural networks and attention-based pooling

Q & A

Thanks！