A New Convex Loss Function For Multiple Instance Support Vector Machines

Sang-Baeg Kim
Department of Industrial Engineering
Sungkynkwan University
Suwon, South Korea
Email: kimsb01@skku.edu

Jung-Man Bae
AI Vision Analytics Devision
KAI Networks Co. LTD
Suwon, South Korea
Email: jm.bae@kai-networks.com

Contents

- Introduction
- Mathematical Models
- Experiments
- Conclusions

MIL; Multiple Instance Learning

- Weakly Supervised Learning
  - Training instances are arranged in sets, called *bags*
  - Labels are provided for entire bags, not for instances
  - Task: Find a bag classifier to predict the labels of unseen bags
- Applications of MIL
  - Drug Activity Prediction Problem: the first MIL Model (1997)
  - Content Base Image Retrieval
  - Computer Aided Diagnosis (from images)
  - Semantic Image Segmentation
  - Anomaly Detection in Videos
  - Video Classification

Formulations of MIL

✓ SVM Formulations
- mi-SVM, MI-SVM
- $\propto$SVM
- RMI-SVM

✓ WR-SVM
- A New SVM based on the *Witness Rate(WR)* of a positive bag
- Maximizing the minimum WR among positive bags
- Estimation of WR of a positive bag using *tanh*(·) for unknown labels

✓ Contributions of WR-SVM
- Proposing a new convex loss function for MIL
- Providing a very simple neural network framework for MIL

WR-SVM(1)

✓ Binary MIL Model: Training dataset: $\{(X_i, Y_i)\}_{i=1}^{N}$

- $X_i = \{x_1^i, x_2^i, \dots, x_{M_i}^i\}$: bag i
- $x_j^i \in R^d$: instances of bag i
- $Y_i \in \mathcal{Y} = \{-1,1\}$ is the known label of the bag X_i.
- The label y_j^i of an instance x_j^i is unknown, $y_j^i \in \{-1,1\}$

✓ Standard MIL Assumptions

- If $Y_i = 1$, then $y_j^i = 1$ for at least one $j \in \{1, \cdots, M_i\}$.
- If $Y_i = -1$, then $y_j^i = -1$ for all $j \in \{1, \cdots, M_i\}$.

WR-SVM(2)

WR-SVM

- The Witness Rate (WR) ρ_i of the i-th positive bag is defined by

$$\rho_i = \frac{1}{M_i} \sum_{j=1}^{M_i} \mathbb{1}_{\{y_j^i = 1\}}$$

- WR-SVM maximizes $\min_{i:Y_i=1} \{\rho_i\}$

$$\min_{y_j^i, w, b, \xi_j^i} \frac{\lambda}{2} \|w\|^2 + \frac{1}{N} \sum_{i:Y_i=-1} \xi_j^i + \frac{1}{N} \frac{1}{\min_{i:Y_i=1} \{\rho_i\}}$$

subject to

$$-w^T x_j^i - b \geq 1 - \xi_j^i, \qquad \forall i: Y_i = -1$$

$$\sum_j \frac{y_j^i + 1}{2} \geq 1, \forall i: Y_i = 1$$

$$y_j^i \in \{-1, 1\}, \forall j, i: Y_i = 1$$

$$\xi_j^i \geq 0, \forall i: Y_i = -1$$

WR-SVM(3)

✓ Relax the integer variable y_j^i to be a continuous variable

- Approximate the label y_j^i of an instance x_j^i in positive bags with a continuous variable $z_j^i = \tanh(w^T x_j^i + b) \in (-1,1)$

- Using this relaxation, WR can be approximated as follows:

$$\hat{\rho}_i = \frac{1}{M_i} \sum_{j=1}^{M_i} \mathbb{1}_{\{z_j^i \geq z_0\}} z_j^i$$

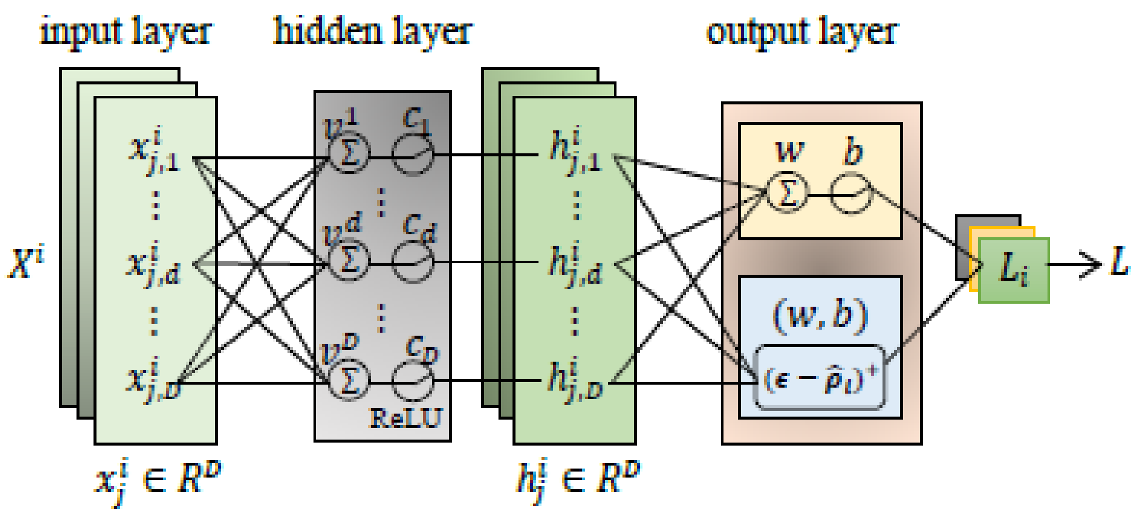
- Loss function L of WR-SVM:

$$L = \frac{\lambda}{2} \|w\|^2 + \frac{1}{N} \sum_{i:Y_i=-1} \sum_{j=1}^{M_i} \left(1 + w^T x_j^i + b\right)_+ + \frac{1}{N} \sum_{i:Y_i=1} (\epsilon - \hat{\rho}_i)_+$$

Deep WR-SVM

✓ DNN architecture of WR-SVM

- The loss function L is convex.
- MIL pooling function for WR-SVM is $\hat{\rho}_i > 0$.
- Deep WR-SVM need not the MIL Pooling Layer.
- The first Deep MIL without MIL Pooling Layer.



Performance of Deep WR-SVM

Video Datasets (30 classes)

- WIDER bags: sampled WIDER images from 30 classes (class 0class 29) to make artificial video bags
- CCV + WIDER bags
- HMDB51
- UCF-101

Classifier	Accuracy(%)			
	WIDER	CCV+	HMDB51	UCF-101
mi-SVM	25.42	23.24	21.33	19.41
MI-SVM	27.73	28.45	25.46	23.72
alter $\propto$SVM	35.33	31.35	29.37	33.30
Single-granular $\propto$SVM	37.45	34.85	31.65	28.75
RMI-SVM	37.10	38.15	35.78	34.26
Ensemble of CNNs	68.32	58.42	64.75	66.37
AWR-SVM	71.65	69.53	68.71	65.66

Conclusions

✓ Contributions of Our Works

- We introduce a new convex formulation, WR-SVM, of the MIL problem based on the WRs of positive bags.
- Our NN framework of WR-SVM is one of the simplest NN models for MIL.

✓ Further Research

- Test WR-SVM for larger classes and develop efficient bag generators and
- Optimal DNN architectures (i.e., depths and widths) for WR-SVM.

Thank you for attention.

ICPR 2020