Adding Unlabeled Samples to Categories by Learned Attributes

Jonghyun Choi  Mohammad Rastegari  Ali Farhadi  Larry S. Davis

Problem

To obtain a better classifier
Better classification model
More training data

How to Add Samples

• Active learning
  • Require human in the loop
• Semi-supervised learning
  • Unlabeled samples are assumed to follow the same distribution of labeled samples

Our Solution

By learning data driven attributes based on [2]
• No human required
• No underlying distribution assumed

Approach Overview

Adding by Two Kinds of Attributes

Formulation

Joint optimization for discovering discriminative attributes and unlabeled samples

\[
\min_{\beta,\gamma} \sum_{c \in \mathcal{C}} \left( \alpha J_c^\alpha(I_c, w_c^\alpha) + \beta J_c^\beta(I_c, w_c^\beta) \right) + M(I)
\]
subject to

\[
J_c^\alpha(I_c, w_c^\alpha) = \|w_c^\alpha\|^2 + \lambda_\alpha \sum_{i=1}^n \xi_{c,i}, \quad \xi_{c,i} \geq 0, \quad \forall i \in \{1, \ldots, n\}
\]

\[
I_{c,i} \cdot y_{c,i}(w_c^\alpha x_i) \geq 1 - \xi_{c,i}, \quad \forall i \in \{1, \ldots, n\}
\]

\[
J_c^\beta(I_c, w_c^\beta) = \|w_c^\beta\|^2 + \lambda_\beta \sum_{k=1}^{d_c} \left( \sum_{j=1}^n I_{c,j} \phi(x_{c,j}) \right) + \gamma - \sum_{k=1}^{d_c} \gamma k
\]

\[
I_{c,j} \cdot y_{c,j}(w_c^\beta \phi(x_{c,j})) \geq 1 - \xi_{c,j}, \quad \forall j \in \{1, \ldots, n\}
\]

\[
\sum_{k=1}^{d_c} I_{c,k} \leq \gamma, \quad I_{c,k} = 1, \quad \forall k \in \{1, \ldots, d_c\}
\]

\[
M(I) = \sum_{c \in \mathcal{C}, \neq 2} I_{c,1} \cdot I_{c,2}
\]

Experimental Results

Comparison with other methods

Varying number of added images

By learning data driven attributes based on [2]
• No human required
• No underlying distribution assumed

By learning data driven attributes based on [2]
• No human required
• No underlying distribution assumed

Adding to Training Samples

Adding by Categorical Attributes

Adding by Exemplar Attributes

Formulation

Joint optimization for discovering discriminative attributes and unlabeled samples

\[
\min_{\beta,\gamma} \sum_{c \in \mathcal{C}} \left( \alpha J_c^\alpha(I_c, w_c^\alpha) + \beta J_c^\beta(I_c, w_c^\beta) \right) + M(I)
\]
subject to

\[
J_c^\alpha(I_c, w_c^\alpha) = \|w_c^\alpha\|^2 + \lambda_\alpha \sum_{i=1}^n \xi_{c,i}, \quad \xi_{c,i} \geq 0, \quad \forall i \in \{1, \ldots, n\}
\]

\[
I_{c,i} \cdot y_{c,i}(w_c^\alpha x_i) \geq 1 - \xi_{c,i}, \quad \forall i \in \{1, \ldots, n\}
\]

\[
J_c^\beta(I_c, w_c^\beta) = \|w_c^\beta\|^2 + \lambda_\beta \sum_{k=1}^{d_c} \left( \sum_{j=1}^n I_{c,j} \phi(x_{c,j}) \right) + \gamma - \sum_{k=1}^{d_c} \gamma k
\]

\[
I_{c,j} \cdot y_{c,j}(w_c^\beta \phi(x_{c,j})) \geq 1 - \xi_{c,j}, \quad \forall j \in \{1, \ldots, n\}
\]

\[
\sum_{k=1}^{d_c} I_{c,k} \leq \gamma, \quad I_{c,k} = 1, \quad \forall k \in \{1, \ldots, d_c\}
\]

\[
M(I) = \sum_{c \in \mathcal{C}, \neq 2} I_{c,1} \cdot I_{c,2}
\]

By learning data driven attributes based on [2]
• No human required
• No underlying distribution assumed

Adding to Training Samples

Adding by Categorical Attributes

Adding by Exemplar Attributes

Formulation

Joint optimization for discovering discriminative attributes and unlabeled samples

\[
\min_{\beta,\gamma} \sum_{c \in \mathcal{C}} \left( \alpha J_c^\alpha(I_c, w_c^\alpha) + \beta J_c^\beta(I_c, w_c^\beta) \right) + M(I)
\]
subject to

\[
J_c^\alpha(I_c, w_c^\alpha) = \|w_c^\alpha\|^2 + \lambda_\alpha \sum_{i=1}^n \xi_{c,i}, \quad \xi_{c,i} \geq 0, \quad \forall i \in \{1, \ldots, n\}
\]

\[
I_{c,i} \cdot y_{c,i}(w_c^\alpha x_i) \geq 1 - \xi_{c,i}, \quad \forall i \in \{1, \ldots, n\}
\]

\[
J_c^\beta(I_c, w_c^\beta) = \|w_c^\beta\|^2 + \lambda_\beta \sum_{k=1}^{d_c} \left( \sum_{j=1}^n I_{c,j} \phi(x_{c,j}) \right) + \gamma - \sum_{k=1}^{d_c} \gamma k
\]

\[
I_{c,j} \cdot y_{c,j}(w_c^\beta \phi(x_{c,j})) \geq 1 - \xi_{c,j}, \quad \forall j \in \{1, \ldots, n\}
\]

\[
\sum_{k=1}^{d_c} I_{c,k} \leq \gamma, \quad I_{c,k} = 1, \quad \forall k \in \{1, \ldots, d_c\}
\]

\[
M(I) = \sum_{c \in \mathcal{C}, \neq 2} I_{c,1} \cdot I_{c,2}
\]

By learning data driven attributes based on [2]
• No human required
• No underlying distribution assumed

Adding to Training Samples

Adding by Categorical Attributes

Adding by Exemplar Attributes

Formulation

Joint optimization for discovering discriminative attributes and unlabeled samples

\[
\min_{\beta,\gamma} \sum_{c \in \mathcal{C}} \left( \alpha J_c^\alpha(I_c, w_c^\alpha) + \beta J_c^\beta(I_c, w_c^\beta) \right) + M(I)
\]
subject to

\[
J_c^\alpha(I_c, w_c^\alpha) = \|w_c^\alpha\|^2 + \lambda_\alpha \sum_{i=1}^n \xi_{c,i}, \quad \xi_{c,i} \geq 0, \quad \forall i \in \{1, \ldots, n\}
\]

\[
I_{c,i} \cdot y_{c,i}(w_c^\alpha x_i) \geq 1 - \xi_{c,i}, \quad \forall i \in \{1, \ldots, n\}
\]

\[
J_c^\beta(I_c, w_c^\beta) = \|w_c^\beta\|^2 + \lambda_\beta \sum_{k=1}^{d_c} \left( \sum_{j=1}^n I_{c,j} \phi(x_{c,j}) \right) + \gamma - \sum_{k=1}^{d_c} \gamma k
\]

\[
I_{c,j} \cdot y_{c,j}(w_c^\beta \phi(x_{c,j})) \geq 1 - \xi_{c,j}, \quad \forall j \in \{1, \ldots, n\}
\]

\[
\sum_{k=1}^{d_c} I_{c,k} \leq \gamma, \quad I_{c,k} = 1, \quad \forall k \in \{1, \ldots, d_c\}
\]

\[
M(I) = \sum_{c \in \mathcal{C}, \neq 2} I_{c,1} \cdot I_{c,2}
\]