Transfer Learning by Borrowing Examples for Multiclass Object Detection

Joseph J. Lim (MIT), Ruslan Salakhutdinov (UToronto), Antonio Torralba (MIT)
lim@csail.mit.edu, rsalakhu@utstat.toronto.edu, torralba@csail.mit.edu

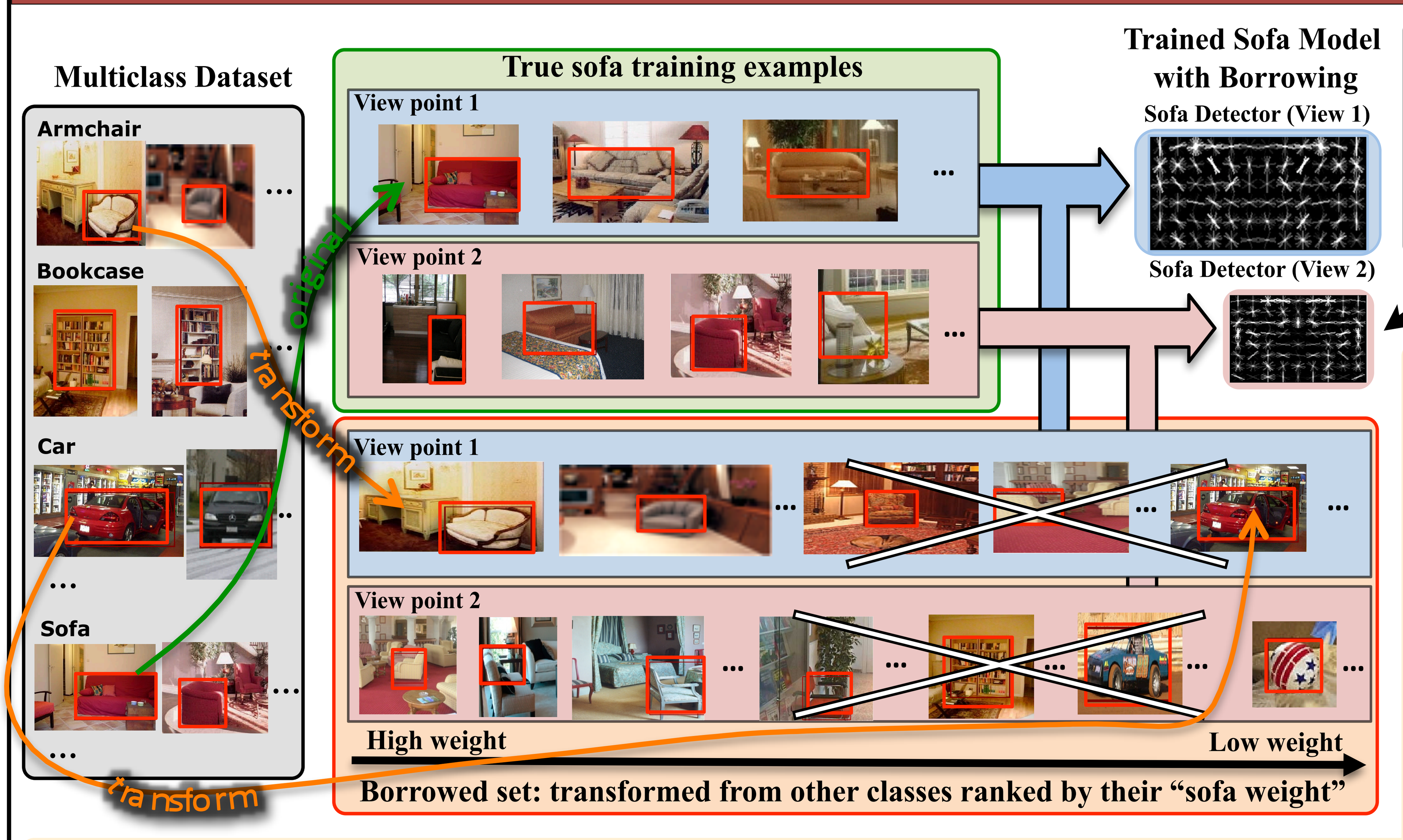
Problem Formulation

The Challenge: there are many classes with few training examples for object detection, despite the trend of increasingly large datasets

Our Goal: learning which examples from other classes (or dataset) to train together with target class examples

Our Contribution:
(1) borrowing examples from other classes rather than sharing at the parameter space
(2) learning the transformation to apply during borrowing so that borrowed examples look more alike to target class

Approach



We are borrowing examples for a sofa detector. Front-view armchair examples are elongated so that they will look more like sofa. All examples from other classes will learn weights and then some will be selected to train together with original examples.

Standard Training
In a standard training for object detector for class C, the following form is used:

$$\min_{\beta^c}\left(\sum_{i=1}^{n_c+b}\text{Loss}(\beta^c\cdot \mathbf{x}_i, \text{sign}(y_i)) + \lambda R(\beta^c)\right)$$

Here, β^c is the object detector model.

Positive examples: all examples from C
Negative examples: any non-positive example

There is no borrowing / sharing information between classes

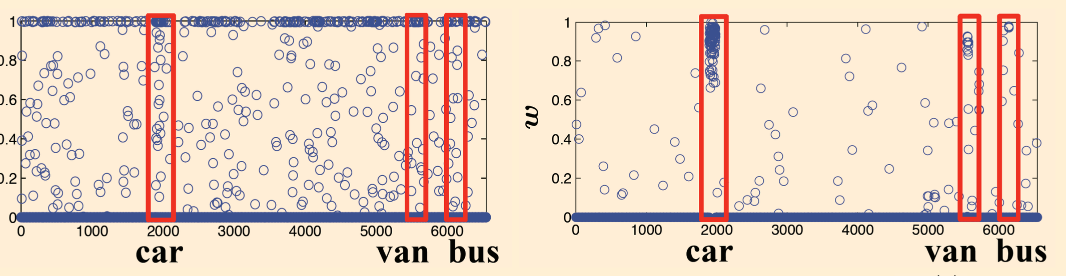
Our Approach
In addition to train β^c, another goal is to learn w_i^c, which indicates how much we borrow x_i from class y_i for training β^c.

$$\sum_{c\in\mathcal{C}}\min_{\beta^c}\min_{\mathbf{w}^{*,c}}\left(\sum_{i=1}^{n+b}(1-w_i^{*,c})\text{Loss}(\beta^c\cdot x_i, \text{sign}(y_i)) + \lambda R(\beta^c) + \Omega_{\lambda_1,\lambda_2}(\mathbf{w}^{*,c})\right)$$ where $\Omega_{\lambda_1,\lambda_2}(\mathbf{w}^*) = \lambda_1\sum_{l\in\mathcal{C}}\sqrt{n_l}\|\mathbf{w}^*_{(l)}\|_2 + \lambda_2\|\mathbf{w}^*\|_1$ = sparse group lasso

Positive examples: {all examples from C} and {other similar examples according to w_i^c}
Negative examples: any examples with w_i^c = 0

The sparse group lasso criteria controls how much examples are borrowed at the group level as well as keeps the sparsity of weights at the individual level.

Solving this optimization problem is non-convex. Hence, we use an iterative algorithm
(1) solving for β^c given w_i^c and (2) solving for w_i^c given β^c.
We initialize w_i^c by setting 1 for all examples from C, and 0 for all others. The 1st iteration is equivalent to solving the standard approach without borrowing.

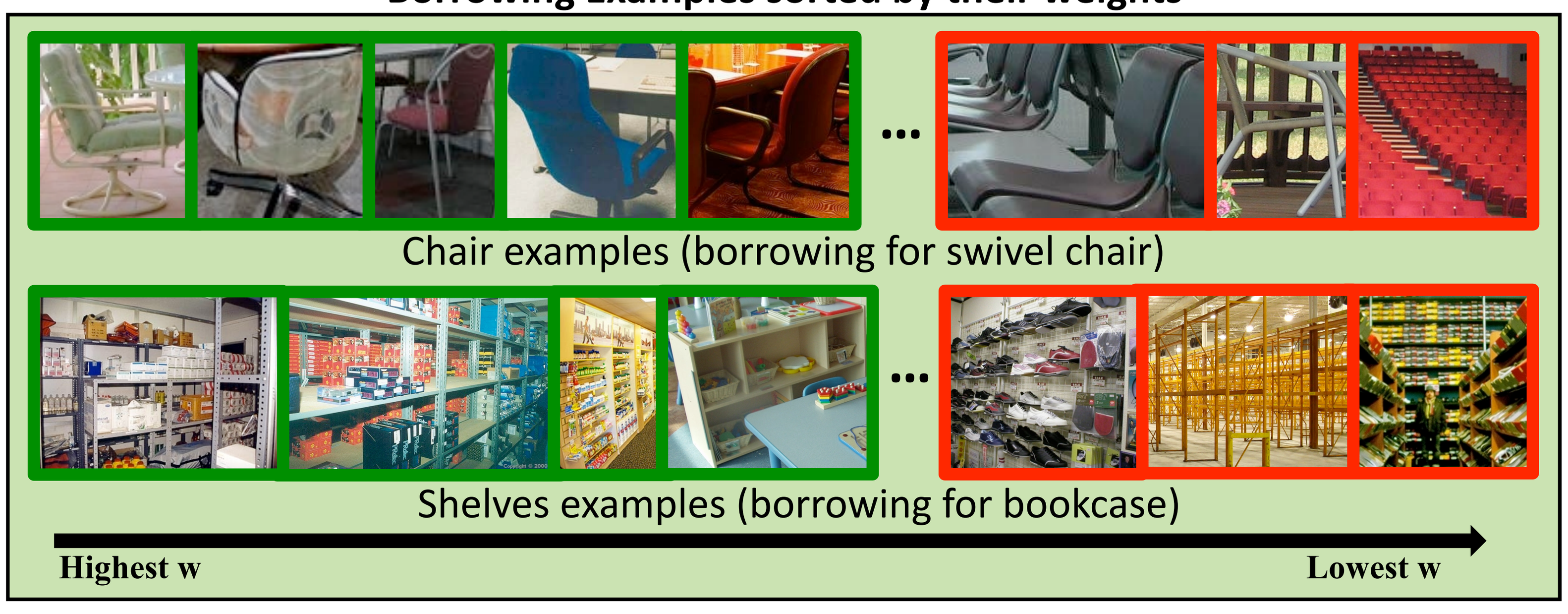


(a) only with L1 norm (b) learned with $\Omega(\cdot)$

Borrowing weights for truck
of all examples from other examples

Experiment 1 - Borrowing from *other classes* within the same dataset

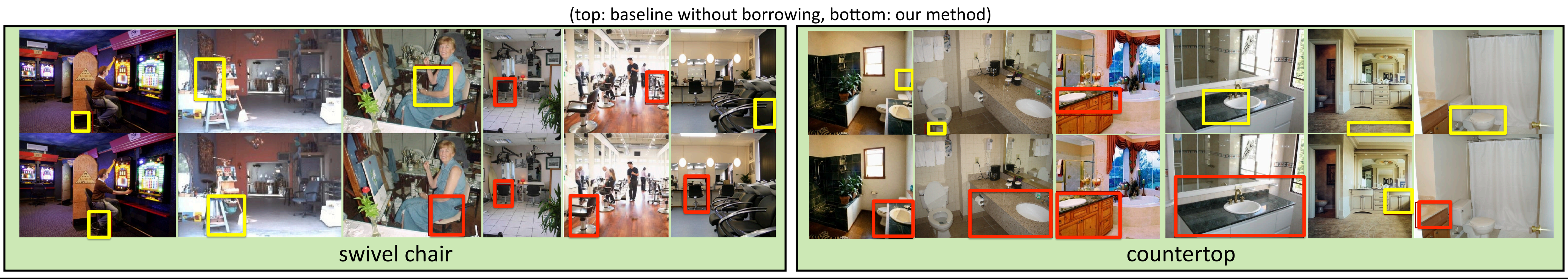
Borrowing Examples sorted by their weights



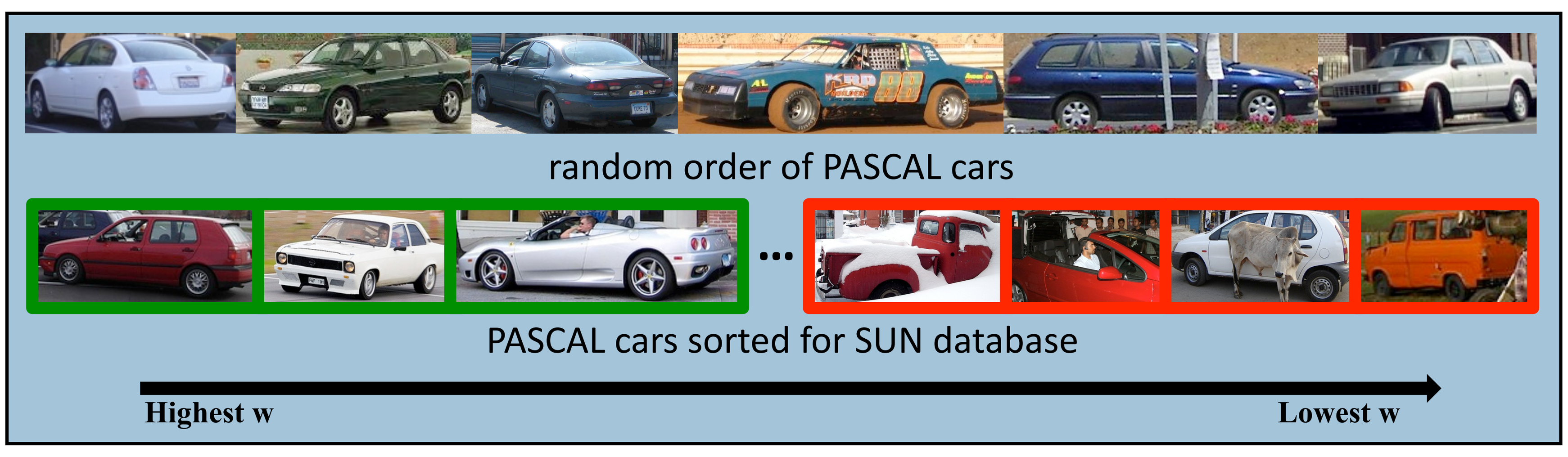
- The experiment here is borrowing examples from *other classes*.
- **Our method** improves **1.36** over the baseline while borrowing all examples from similar classes **without any selection** improves only 0.30.

Methods	Borrow without transformation	Borrow all examples from the same classes without selection	Borrow with transformation
AP without borrowing	14.99	16.59	16.59
AP improvements	+1.00	+0.30	**+1.36**

Detection Results
(top: baseline without borrowing, bottom: our method)



Experiment 2 - Borrowing from *another dataset*



- The experiment here is borrowing examples from an*other dataset*.
- When we naively merge two datasets *without any selection*, the performances improves (a) 0.15 and (b) 1.93 on SUN09 and PASCAL, respectively. (some performances dropped)
- **Our method** *with selection* by learning weights improves the performances by (a) **1.98** and (b) **2.86** on SUN09 and PASCAL, respectively.

	SUN09 only	PASCAL only	SUN09 +PASCAL	SUN09 +borrow PASCAL
car	43.31	39.47	43.64	**45.88**
person	45.46	28.78	46.46	**46.90**
sofa	12.96	11.97	12.86	**15.25**
chair	18.82	13.84	18.18	**20.45**
mean	30.14	23.51	30.29	**32.12**

(a) Testing on SUN09

	PASCAL only	SUN09 only	PASCAL +SUN09	PASCAL +borrow SUN09
car	49.58	40.81	49.91	**51.00**
person	23.58	22.31	26.05	**27.05**
sofa	19.91	13.99	20.01	**22.17**
chair	14.23	14.20	**19.06**	18.55
mean	26.83	22.83	28.76	**29.69**

(b) Testing on PASCAL

Conclusion

- We proposed an effective method for transfer learning across object categories, based on the sparse group Lasso framework.
- We showed that our method finds useful examples to borrow and improves the state-of-the-art detector performance with/without transformation.

Object detection transfer learning dataset is available at: http://csail.mit.edu/~lim/lst_nips2011/