Using humans to build mid-level features

Genevieve Patterson[1] Tsung-Yi Lin[2] James Hays[1]

[1]Brown University [2]University of California, San Diego

Original Discriminative Patch Discovery Method

[Singh, Gupta, Efros, ***Unsupervised Discovery of Mid-Level Discriminative Patches***, ECCV 2012]

Algorithm 1 Discover Top n Discriminative Patches

Require: Discovery set $\mathcal{D}$, Natural World set $\mathcal{N}$

1: $\mathcal{D} \Rightarrow \{D_1, D_2\}$; $\mathcal{N} \Rightarrow \{N_1, N_2\}$ ▷ Divide $\mathcal{D}, \mathcal{N}$ into equal sized disjoint sets
2: $S \Leftarrow rand\_sample(D_1)$ ▷ Sample random patches from D_1
3: $K \Leftarrow kmeans(S)$ ▷ Cluster patches using KMeans
4: **while not** $converged()$ **do**
5: **for all** i such that $size(K[i]) \geq 3$ **do** ▷ Prune out small ones
6: $C_{new}[i] \Leftarrow svm\_train(K[i], N_1)$ ▷ Train classifier for each cluster
7: $K_{new}[i] \Leftarrow detect\_top(C[i], D_2, m)$ ▷ Find top m new members in other set
8: **end for**
9: $K \Leftarrow K_{new}$; $C \Leftarrow C_{new}$
10: $swap(D_1, D_2)$; $swap(N_1, N_2)$ ▷ Swap the two sets
11: **end while**
12: $A[i] \Leftarrow purity(K[i]) + \lambda \times discriminativeness(K[i])\ \forall\ i$ ▷ Compute scores
13: **return** $select\_top(C, A, n)$ ▷ Sort according to scores and select top n patches

Human-in-the-loop Discriminative Patch Discovery:

New Algorithm

Require: Discovery set $\mathcal{D}$, Natural World set $\mathcal{N}$

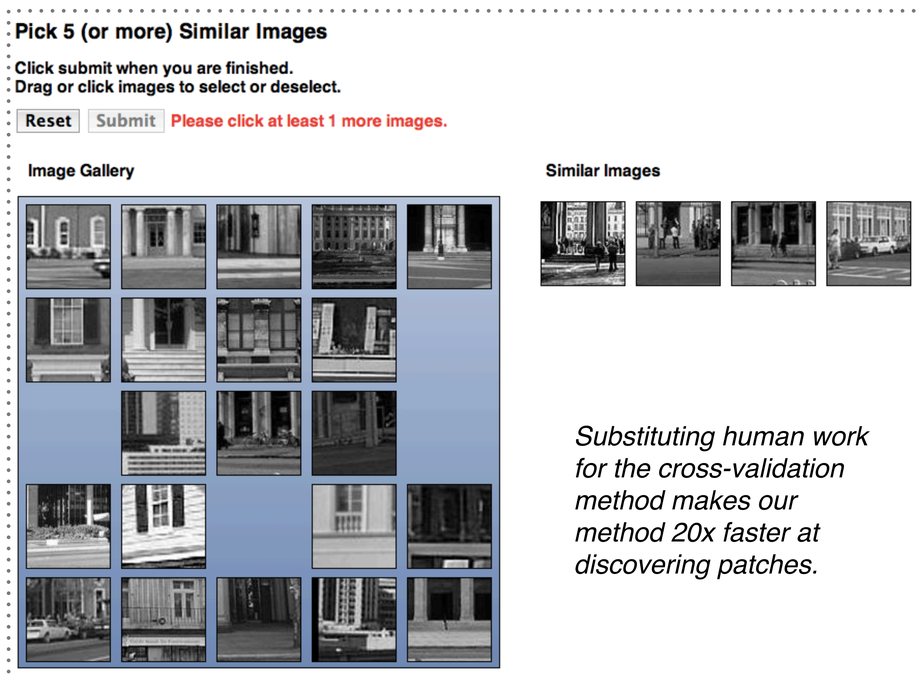
1: $\mathcal{D} \Rightarrow \{D_1, D_2\}$; $\mathcal{N} \Rightarrow \{N_1, N_2\}$ ▷ Divide $\mathcal{D}, \mathcal{N}$ into equal sized disjoint sets
2: $S \Leftarrow rand\_sample(D_1)$ ▷ Sample random patches from D_1
3: $K \Leftarrow kmeans(S)$ ▷ Cluster patches using KMeans

Ask Humans to refine initial NN clusters using MTurk

13: **return** $select\_top(C, A, n)$ ▷ Sort according to scores and select top n patches

Mechanical Turk User Interface:

Users view a group of 25 nearest neighbors.
They select the 5 patches they believe are the most similar.

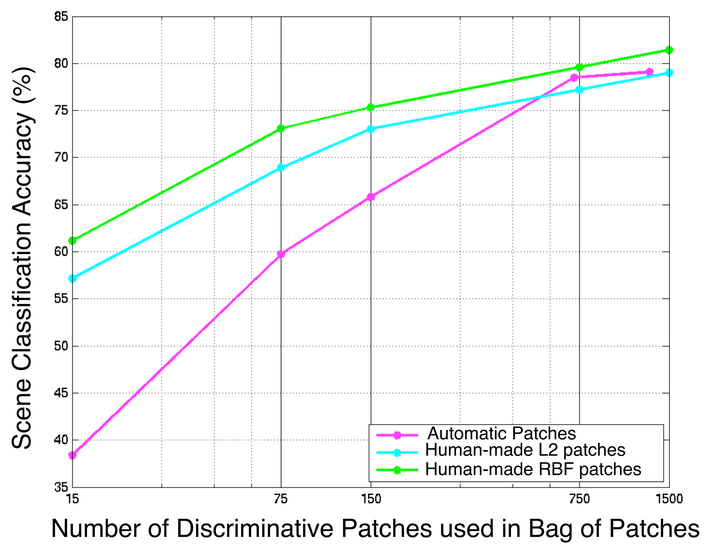


Substituting human work for the cross-validation method makes our method 20x faster at discovering patches.

Test Images and their Top 5 Most Confidently Detected Human-made Patches

Patches discovered with scene category labels as weak supervision

Performance on 15 Scene Datatset



Scene Category Confusions for Both Types of Patches