Contributions
- state of the art on SUN397 benchmark[1]
- 49.5% (SIFT & LCS + Fisher Vector + MTL)
- consistent improvement over standard one-vs-all single task learning (STL)
  - w/ and w/o color cues
- 5.50 training examples per class
- top-K accuracy for all K
- scalability to Fisher Vector features
  - 2,000,000 dimensions, dense

Scene Classification
- previous state of the art[2] (47.2%)
- Fisher Vector on SIFT & LCS (color feature)
- independent one-vs-all SVMs
- SUN397 challenges
  - groups of related (ambiguous) classes
  - ≤50 training examples per class
  - combined with high dimensional FV features ⇒ overfitting
- existing relations between classes could be exploited

Algorithm
1. start with an image representation \( x \) (e.g., Fisher Vector, but could be any)
2. train one-vs-all SVMs on \( x \) (first layer, initialization for MTL)
3. stack learned predictors into \( U_k \)
4. iterate (multitask learning)
  - train SVMs on \( U_k \)
  - update \( U \)
- prediction cost is effectively the same as STL since additional product is low dim.
5. final prediction:
  - \( \arg\max \{\alpha_k(i, U_k x)\} \) for \( i = 1, \ldots, K \)

Implementation Details (code on GitHub!)
- adapt SDCA solver[3]
  - (Stochastic Dual Coordinate Ascent)
- no primal variables, all in dual
- learning \( U \) via SDCA-variant
- both subproblems via SDCA (hence MTL-SDCA)
- use precomputed kernels (dual optimization: \( n=20\times K = 260K \))
- closed-form updates, also for \( \alpha_k \)

Runtime Comparison*
- SDCA training
  - STL vs MTL
  - Overhead
  - STL training
  - x11
- \( \times1.4 \)
- \( \times1.8 \)
- \( \times1.07 \)
- Evaluation on SUN397
  - STL vs MTL
  - SDCA, SIFT (46.9)
  - STL vs MTL
  - SDCA, SIFT (45.1)
  - STL vs MTL
  - Sanchez et al., SIFT (43.3)
  - Donahue et al. (40.9)
  - Su and Jurie (35.6)
  - Xiao et al. (38.0)

References
  Large-scale scene recognition from Kitchen to Zoo. CVPR'10.
  for regularized loss minimization. JMLR'13.

Conclusion
- effective MTL regularization consistently improves over STL
- achieves state of the art results
- scales to dense high dimensional image representation (Fisher Vector)