One Shot Learning via Compositions of Meaningful Patches

Alex Wong Alan L. Yuille
University of California, Los Angeles
http://ccvl.stat.ucla.edu/

CCVL

motivation

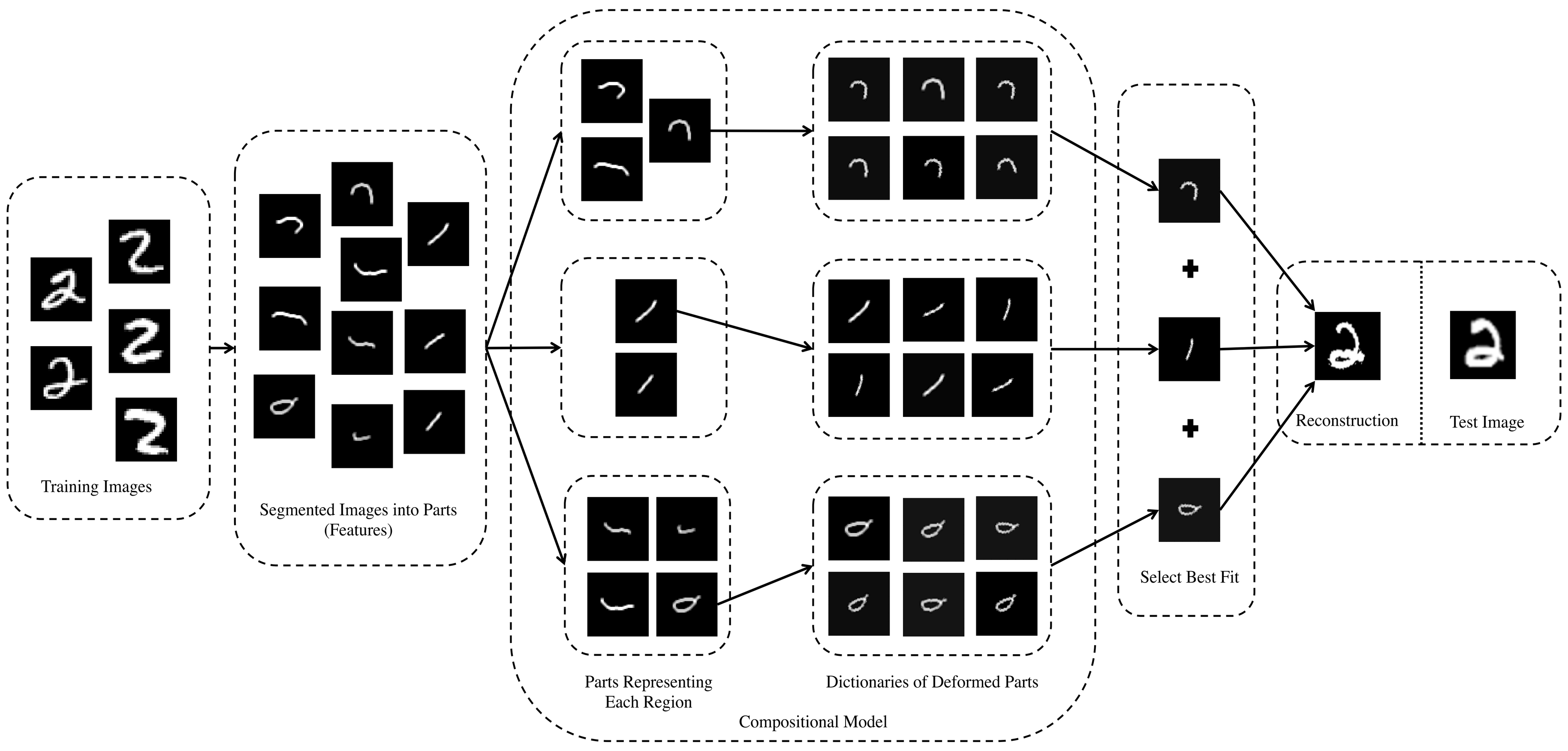
Current state-of-the-art algorithms perform very well on most common datasets **when trained on thousands of examples**
However, humans are able to **learn a concept from very few examples, perhaps even just one**

what is one shot learning?

One shot learning is an object categorization task where very few examples (1-5) are given for training

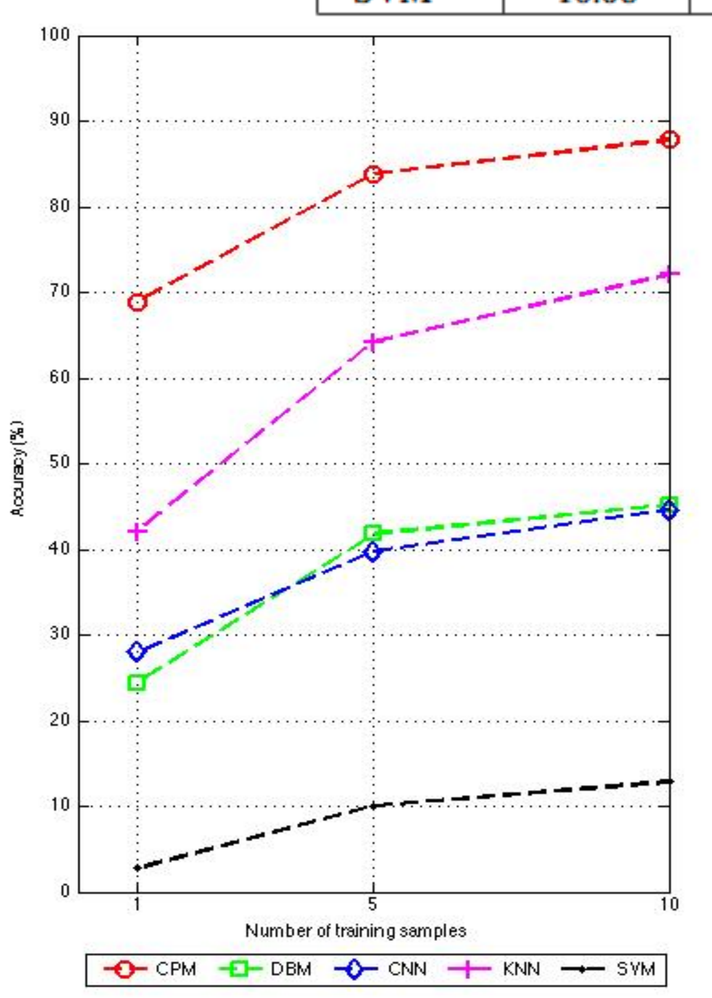
our approach

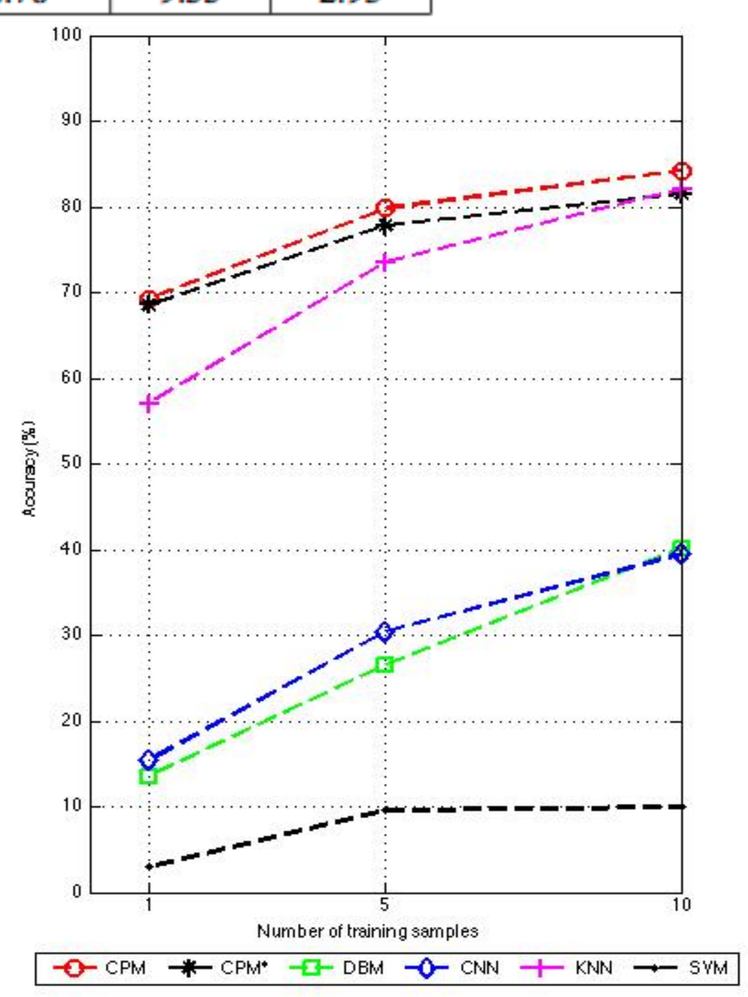
- Learn a meaningful patch-based representation of the underlying structure of an object without human supervision
- Build a compositional model composed of a set of compact dictionaries of meaningful patches
- Reconstruct the target image with deformations of the meaningful patch dictionaries by patch matching
- Select the class of the best proposed reconstruction as label



experimental results

Method	MNIST n=5	MNIST n=1	USPS n=5	USPS n=1
CPM	**83.79**	**68.86**	**79.88**	**69.31**
CPM*	-	-	77.81	68.58
DBM	41.76	24.37	26.60	13.56
CNN	39.80	28.01	30.42	15.37
K-NN	64.26	42.08	73.59	56.98
SVM	10.08	2.78	9.55	2.93





conclusion

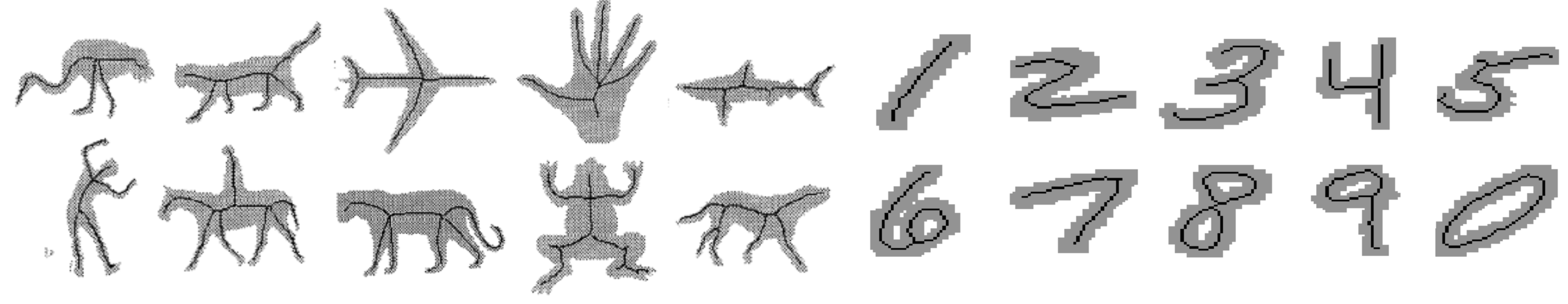
- Our compositional model outperforms popular algorithms on the recognition task under one shot learning
- The extracted features are semantically meaningful
- The model generalizes beyond the training set and demonstrates transferability between separate datasets

acknowledgements

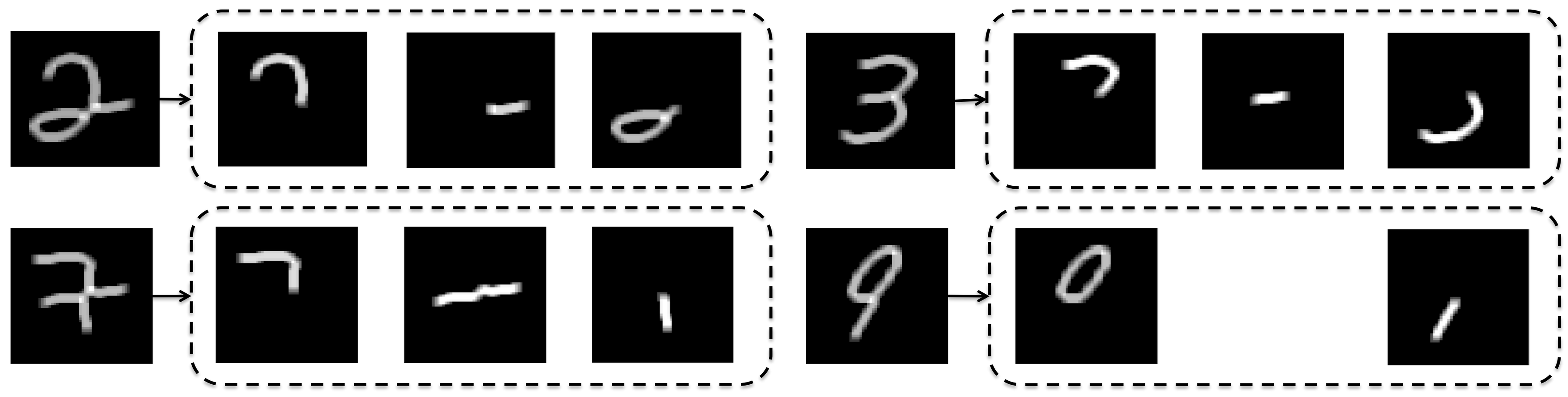
We would like to thank Brian Taylor for performing experiments and editing the manuscript. This work was supported by NSF STC award CCF-1231216 and ONR N00014-12-1-0883.

feature extraction

- Symmetry axis acts as a robust object descriptor



- Branch points separate one meaningful part from another
- Small segments are merged with nearby meaningful parts

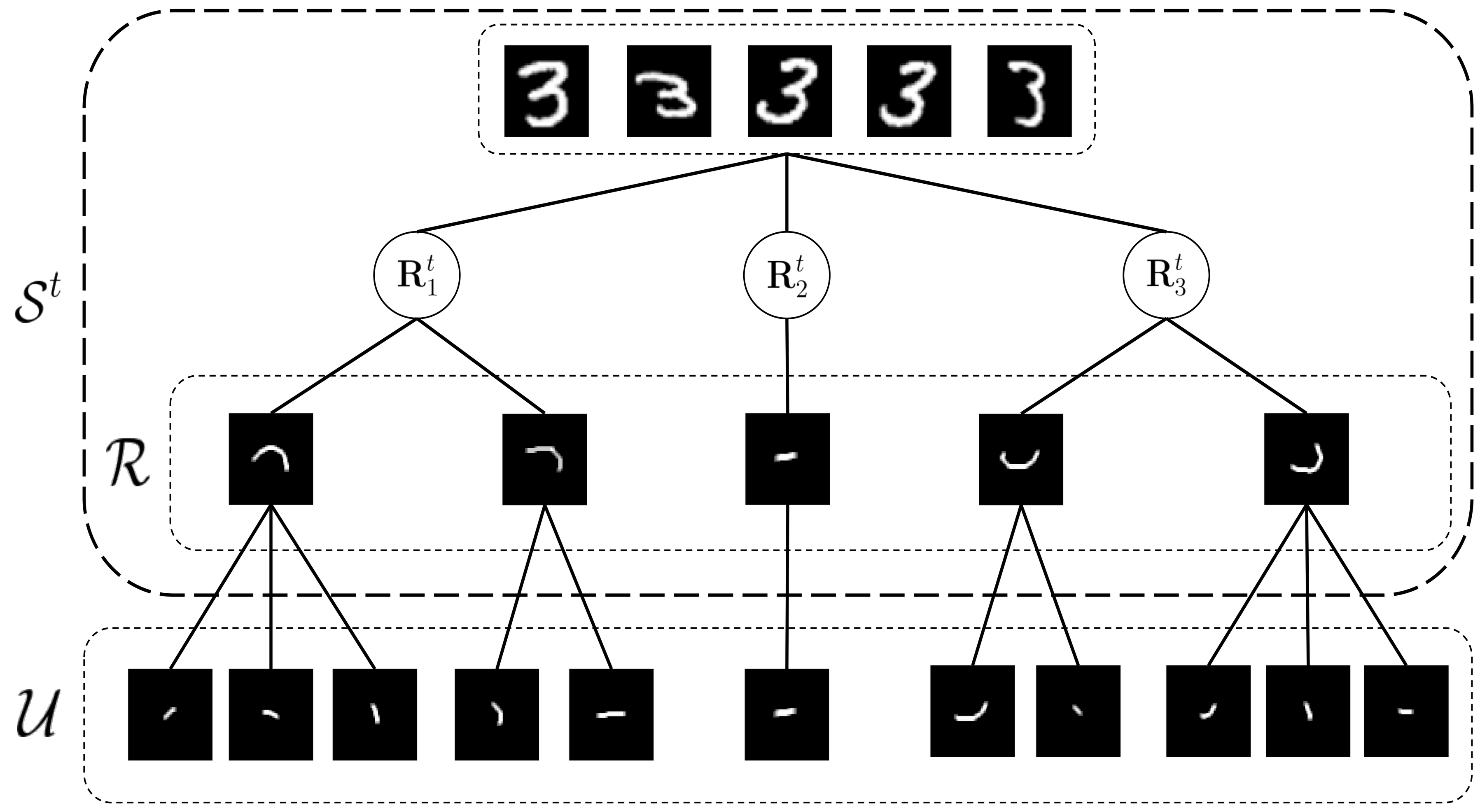


compositional model

- Similar parts, defined by a high match score via Normalized Cross Correlation, are merged to create a compact dictionary
- An AND-OR graph of the part relations is construction for m patches for samples t and u:

$$\mathbf{S}^t = (\mathbf{R}_1^t \vee \mathbf{R}_1^u)\wedge\ldots\wedge\mathbf{R}_k^t\wedge\ldots\wedge(\mathbf{R}_m^t \vee \mathbf{R}_m^u)$$

- Deformations are applied to the meaningful patches



sample reconstructions

*left image denotes test image, right image denotes reconstruction