

PROBLEM DEFINITION



Various studies of psychophysics imply that material perception in human vision is interconnected with perception of object category. An observation is that human can perceive some material properties and categories of objects only after correct recognition of the object categories. The problem is how to model such dependencies, and utilize them to accurately perform a task of interest (material recognition in our case).

OUR CONTRIBUTIONS

In this paper, we propose a feature selection method to select and combine deep features learned using a transfer learning setting. Our contributions are:

- We propose a method for material recognition by selecting and integrating multiple features of different CNN models. They are pre-trained on different datasets/tasks and, if possible and necessary, they are further fine-tuned on the target task/dataset in advance.
- We introduce an extended version of the benchmark material dataset (namely, FMD [1]), called EFMD which is ten times larger than the FMD dataset. The images of EFMD are selected according to surface properties of objects observed in the images that are similar to that of FMD. By the employment of EFMD for transfer learning, we achieve $84.0\% \pm 1.8\%$ accuracy on FMD, which is close to human performance (84.9%).

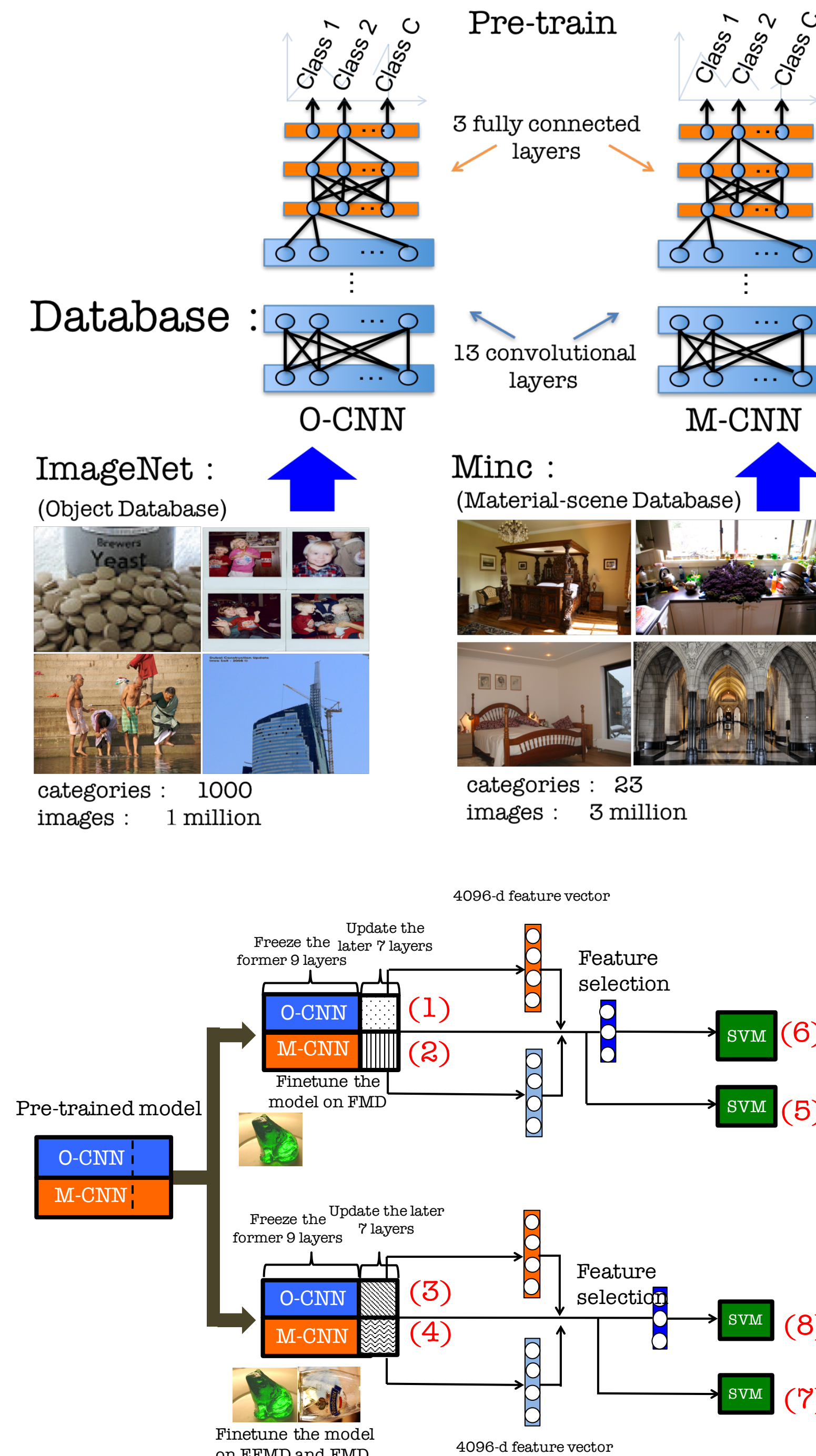
REFERENCES

- [1] L. Sharan, R. Rosenholtz and E. Adelson, "Material perception: What can you see in a brief glance?" *JOV2014*.
- [2] B. Sean, U. Paul, S. Noah and B. Kavita, "Material recognition in the wild with the Materials in Context Database," *CVPR*, 2015.

ACKNOWLEDGEMENT

This work was partly supported by CREST, JST and by JSPS KAKENHI Grant Number 15H05919 (Grant-in-Aid for Scientific Research on Innovative Areas Innovative SHITSUKSAN Science and Technology).

METHODOLOGY



FEATURE SELECTION

Input:

- $\{\Phi_n\}_{n=1}^N$: A set of representations each of which is learned using a CNN obtained using a dataset \mathcal{D}_n .
- $\mathcal{D}_b = \{(I^{(i)}, Y^{(i)})\}_{i=1}^{M_b}$: A dataset of images that will be used for inference of representations.
- T : The number of integrated features.
- K : The number of samples that have maximal activation values for each feature.

Output:

- \mathcal{S} : A set of integrated features.

Initialization:

- For each $(I, Y) \in \mathcal{D}_b$, extract features using a representation $\Phi_n(\cdot)$ such that $\mathbf{x}_n = \Phi_n(I)$, $\forall n = 1, 2, \dots, N$.
- Concatenate $\mathbf{x}_n, \forall n$ and construct $\mathbf{x}_c = [\mathbf{x}_n]_{n=1}^N$.
- Define $\mathcal{C} = \{(\mathbf{x}_c^i, Y^i)\}_{i=1}^{M_b}$, and set equal weight $w_i = 1$ to each sample belonging to \mathcal{C} .
- For each individual feature of \mathbf{x}_c , define the sets $\mathcal{F} = \{x_{c,j} \in \mathbf{x}_c\}_{j=1}^{|\mathbf{x}_c|}$, and $\mathcal{S} = \emptyset$.

for $j \leftarrow 1$ to $|\mathcal{F}|$ do

Construct K samples that have maximal feature values on $x_{c,j}$.

Compute class entropy on the set of top K samples.

end

for $t \leftarrow 1$ to T do

Normalize the sample weights.

Select the feature that minimizes the weighted class entropy.

Penalize the samples the top K samples of the integrated feature.

end

EXPERIMENTAL ANALYSES AND RESULTS

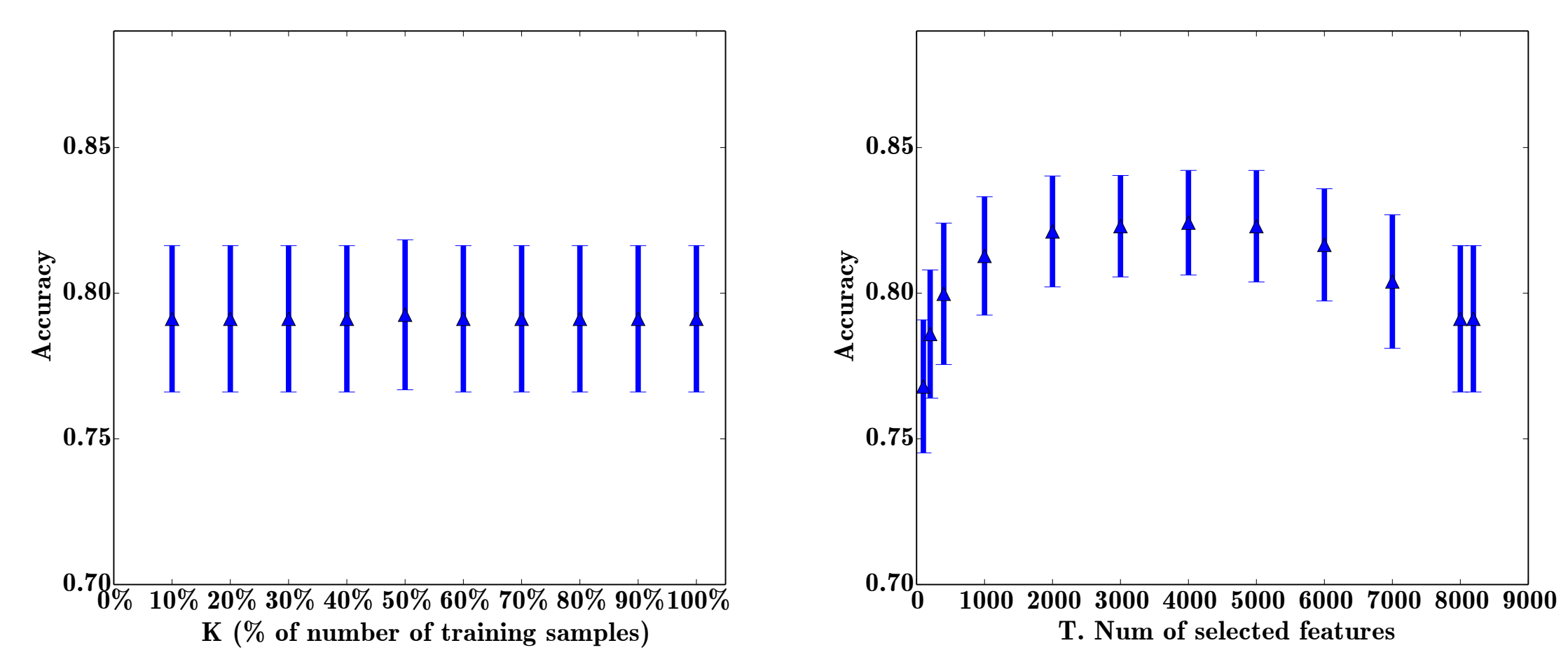
- Performance (accuracy) comparison for different tasks. M: Material features learned using MINC. O: Object features learned using ILSVRC2012. MO: Concatenated material and object features ($\mathbf{x}_c \in \mathcal{F}$). SMO: Features integrated using the proposed method ($\mathbf{x}_c \in \mathcal{S}$).

Task	M (%)	O (%)	MO (%)	SMO (%)
FMD	80.4 ± 1.9	79.6 ± 2.1	79.1 ± 2.5	82.3 ± 1.7
FMD-2	82.5 ± 2.0	82.9 ± 1.6	83.9 ± 1.8	84.0 ± 1.8
EFMD	88.7 ± 0.2	88.8 ± 0.3	89.7 ± 0.13	89.7 ± 0.16
MINC-val	82.45 [2]	68.17	83.48	83.93
MINC-test	82.19 [2]	68.04	83.12	83.60

- Average entropy values of distributions of detections of concatenated and integrated features (MO and SMO). In addition, we provide a diversity analysis of decisions of classifiers employed on individual feature sets (O and M), where \downarrow (\uparrow) indicates that the smaller (larger) the measurement, the larger the diversity.

Diversity Measures	FMD	FMD-2	EFMD
$MO(\hat{H}^{\mathcal{F}})$	2.40	2.29	2.21
$SMO(\hat{H}^{\mathcal{S}})$	2.19	2.08	1.99
$\kappa(\downarrow)$	0.6471	0.7103	0.7116
Q Statistics (\downarrow)	0.9860	0.9944	0.9996
Kohavi-Wolpert Variance (\uparrow)	0.0070	0.0043	0.0002
Disagreement (\uparrow)	0.0279	0.0172	0.0008
Generalized Diversity (\uparrow)	0.3383	0.2892	0.2795

- Analysis of classification performance for (left) different number of Top- K samples, and (right) different number (T) of integrated features on FMD.



- Comparison of number of object and material features belonging to the set of selected features. We show the number of selected object and material features in FMD, FMD2, EFMD, and MINC (from left to right).

