# CROSS-LAYER FEATURES IN CONVOLUTIONAL NEURAL NETWORKS FOR GENERIC CLASSIFICATION TASKS

Cornell University

Kuan-Chuan Peng and Tsuhan Chen {kp388, tsuhan}@cornell.edu

## Inspiration

1. Features extracted from the top-most layer of convolutional neural networks (CNN) have the best discriminative ability. [1, 8]
2. Is there any useful features captured by "shallower" networks but not captured by "deeper" networks?
3. What if we combine features learned from CNNs with different depths?

## Contributions

1. The first paper explicitly utilizing the features extracted from multiple CNNs with different depths.
2. Propose cross-layer features in CNN for generic classification tasks and validate the efficacy of cross-layer features by showing their superior performance on three classification tasks.

## Example Images of the Tasks

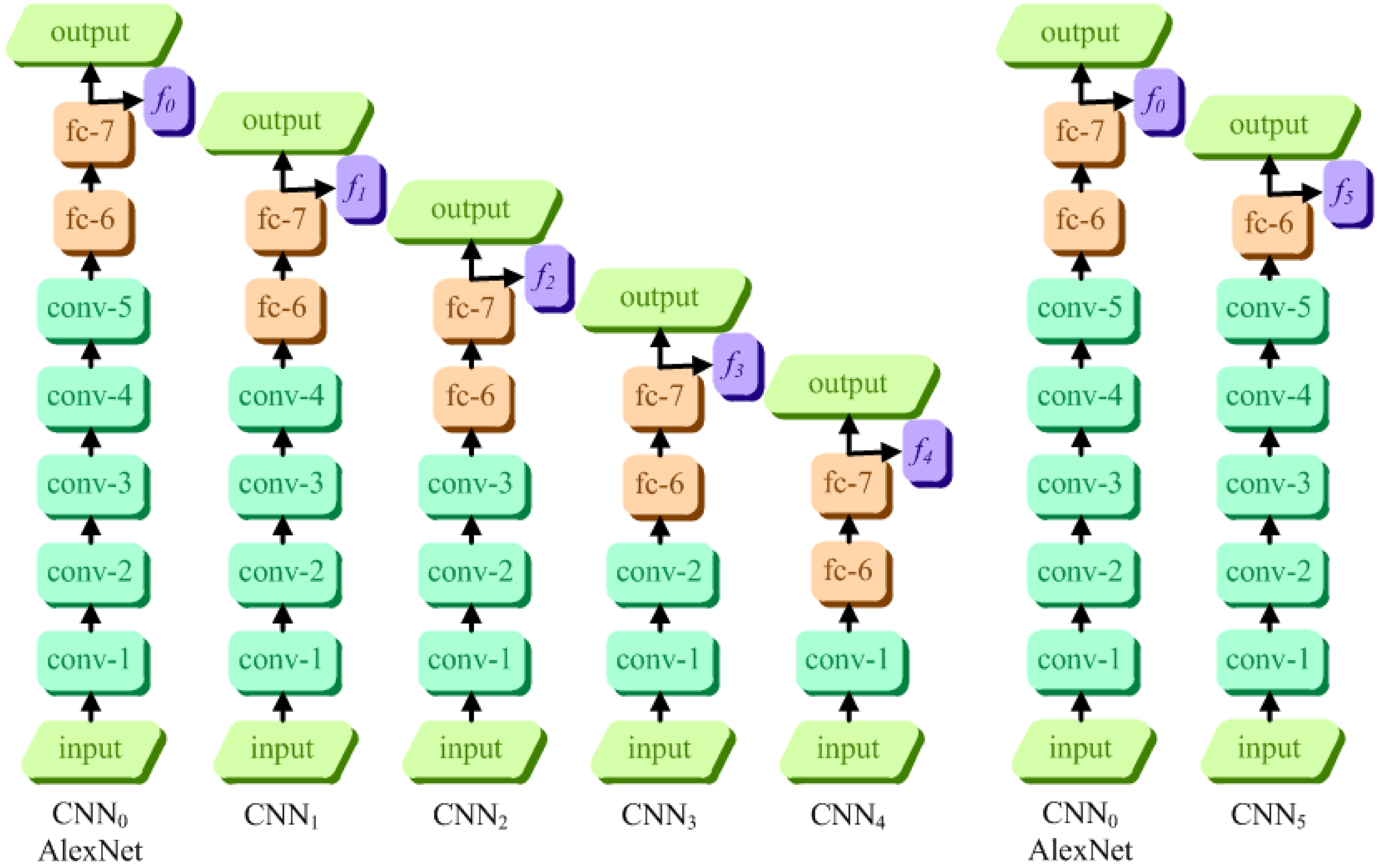
artist style & artistic style classification


architectural style classification


## Specification of Tasks and Datasets

| dataset | Painting-91 [3] | Painting-91 [3] | arcDataset [7] |
|---|---|---|---|
| task | artist style classification | artistic style classification | architectural style classification |
| number of classes | 91 | 13 | 10 / 25 |
| number of images | 4266 | 2338 | 2043 / 4786 |
| image type | painting | painting | architecture |
| examples of class labels | Rubens, Picasso. | Baroque, Cubbism. | Georgian, Gothic. |
| number of training images | 2275 | 1250 | 300 / 750 |
| number of testing images | 1991 | 1088 | 1743 / 4036 |
| training/testing split | specified [3] | specified [3] | random |
| number of fold(s) | 1 | 1 | 10 |
| evaluation metric | accuracy | accuracy | accuracy |
| reference of the above setting | [3] | [3] | [7] |

## Formation of Cross-layer Features




| feature ID | cross-layer features | dimension |
|---|---|---|
| $F_0$ (baseline) | $f_0$ | $k$ |
| $F_1$ | $f_0 + f_1$ | $2k$ |
| $F_2$ | $f_0 + f_1 + f_2$ | $3k$ |
| $F_3$ | $f_0 + f_1 + f_2 + f_3$ | $4k$ |
| $F_4$ | $f_0 + f_1 + f_2 + f_3 + f_4$ | $5k$ |
| $F_5$ | $f_0 + f_5$ | $2k$ |

1. We pre-train each CNN$_i$ with ImageNet [2] and fine-tune with the task of interest.
2. For each task, we train a linear SVM classifier for each $F_i$.

## Experimental Results

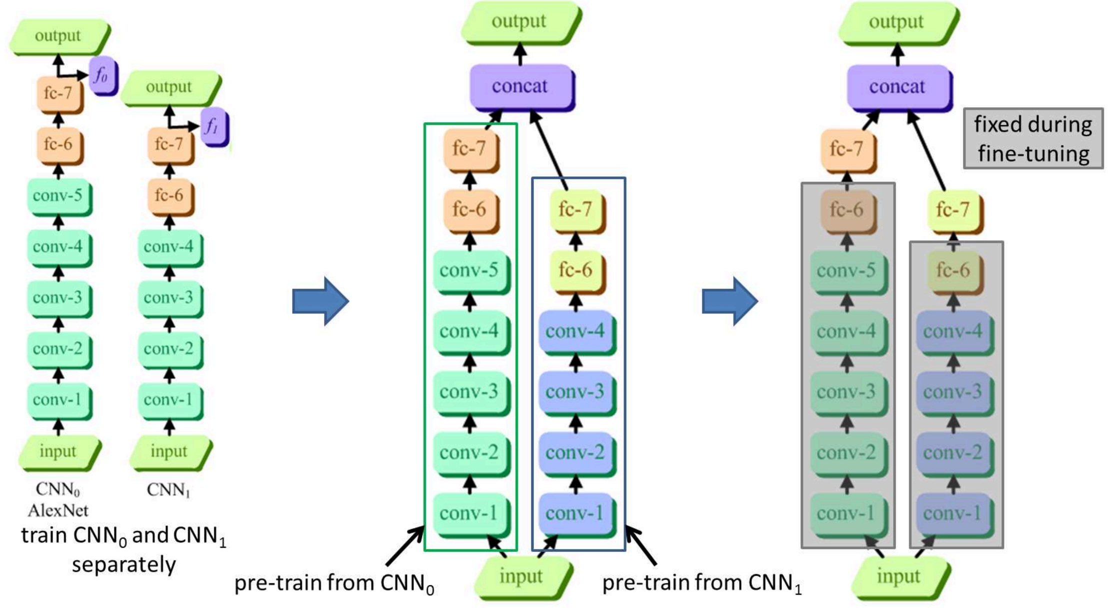
| task | artist style classification | artistic style classification | architectural style classification |
|---|---|---|---|
| prior work | 53.10 | 62.20 | 69.17 / 46.21 |
| reference | [3] | [3] | [7] |
| $F_0$ (baseline) | 55.15 | 67.37 | 70.64 / 54.84 |
| $F_1$ | 55.40 | 68.20 | **71.34** / **55.57** |
| $F_2$ | **56.40** | 68.57 | 70.73 / 55.35 |
| $F_3$ | 56.35 | **69.21** | 70.68 / 55.32 |
| $F_4$ | 56.35 | **69.21** | 70.68 / 55.31 |
| $F_5$ | 56.25 | 68.29 | 70.94 / 55.44 |

## Conclusion

Extensible to generic classification tasks, cross-layer features outperform the state-of-the-art performance and AlexNet [4] baseline on artist style, artistic style, and architectural style classification.

## Ongoing Work

Jointly training on cross-layer features:




## References

[1] P. Agrawal, R. Girshick, and J. Malik. Analyzing the performance of multilayer neural networks for object recognition. In *ECCV '14.*

[2] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and F.-F. Li. ImageNet: A large-scale hierarchical image database. In *CVPR '09.*

[3] F. S. Khan, S. Beigpour, J. V. D. Weijer, and M. Felsberg. Painting-91: a large scale database for computational painting categorization. In *Machine Vision and Applications '14.*

[4] A. Krizhevsky, I. Sutskever, and G. E. Hinton. ImageNet classification with deep convolutional neural networks. In *NIPS '12.*

[5] K.-C. Peng, and T. Chen. A framework of extracting multi-scale features using multiple convolutional neural networks. In *ICME '15.*

[6] K.-C. Peng, and T. Chen. Toward correlating and solving abstract tasks using convolutional neural networks. Submitted to *WACV '16.*

[7] Z. Xu, D. Tao, Y. Zhang, J. Wu, and A. C. Tsoi. Architectural style classification using multinomial latent logistic regression. In *ECCV '14.*

[8] M. D. Zeiler and R. Fergus. Visualizing and understanding convolutional networks. In *ECCV '14.*