CLASSIFICATION OF CORALS IN REFLECTANCE AND FLUORESCENCE IMAGES USING CONVOLUTIONAL NEURAL NETWORK REPRESENTATIONS

Lian Xu, Mohammed Bennamoun, Senjian An, Ferdous Sohel, Farid Boussaid

Lian.xu@research.uwa.edu.au

Introduction

Recent works have shown that CNN features from fully connected layers (FC features) of pre-trained deep networks have been transferred to coral classification, achieving superior performance compared to those well-designed handcrafted features [1]. However,

From the perspective of CNN features,

- CNN features from higher layers are more specific to their original task, while features from intermediate or lower layers are more general to other applications [2].
- FC features capture global spatial layout information while deep convolutional features (CONV features) contain rich local information.

From the perspective of coral images,

- Coral datasets have large discrepancy with ImageNet (in which most categories of corals are unseen).
- Some local features (e.g. texture and edge) are the primary representations for corals.

Contribution

- We proposed a method based on CNN and vector of locally aggregated descriptors (VLAD) encoding for coral image classification.
- We investigated the combined strength of two types of deep features (i.e. FC features and CONV features) in coral image classification.
- We evaluated the transferability of deep CNN features to coral classification for two image modalities (reflectance and fluorescence).

Challenges

- Physical properties of the water medium (e.g. absorption and scattering), cause underwater images to suffer from color degradation which is not present in the ground images.
- Water turbidity and floating particles result in underwater images exhibiting low contrast and limited visibility.
- In case of coral species, they have large variations in morphologies, size, color, shape, and texture across classes, whose boundaries are often ambiguous.
- Class imbalance (i.e. non-coral species often predominate in the whole set) results in misclassified minority coral classes.

Figure 1. Cropped coral patches from the EFC dataset. Five randomly selected patches from four classes are cropped out from the annotated point locations.

Approach

First, we use the VGG-F pre-trained model developed by [3] to extract deep CNN features. Then, VLAD encoding is applied to compress the dense CONV features into a compact representation. Finally, a linear SVM is used for classification.

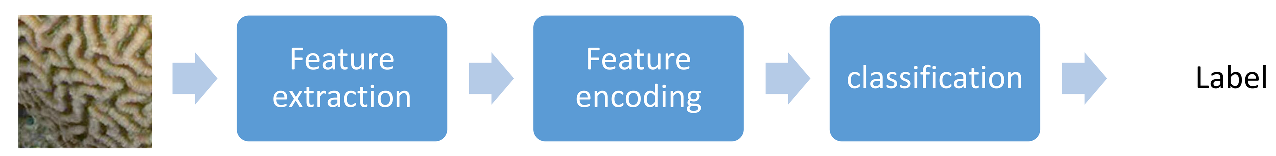

Figure 2. The classification pipeline.

A. Feature extraction

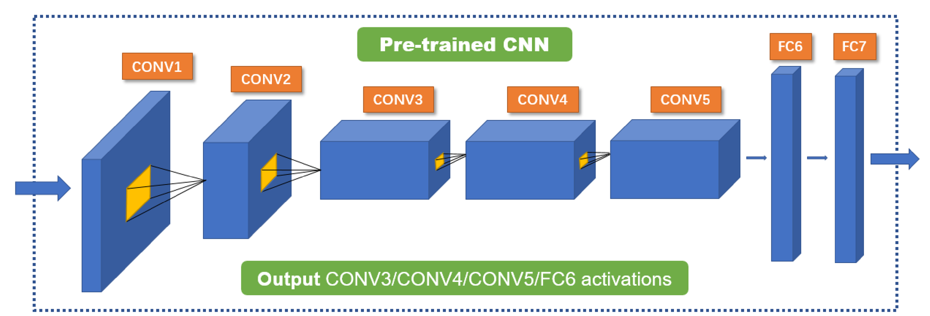

Figure 3. The structure of VGG-F.

Layer	Output size (N×N ×D)
CONV3	13×13×256
CONV4	13×13×256
CONV5	13×13×256
FC6	1×1×4096

Table1. The sizes of CONV and FC feature of VGG-F.

B. Feature encoding

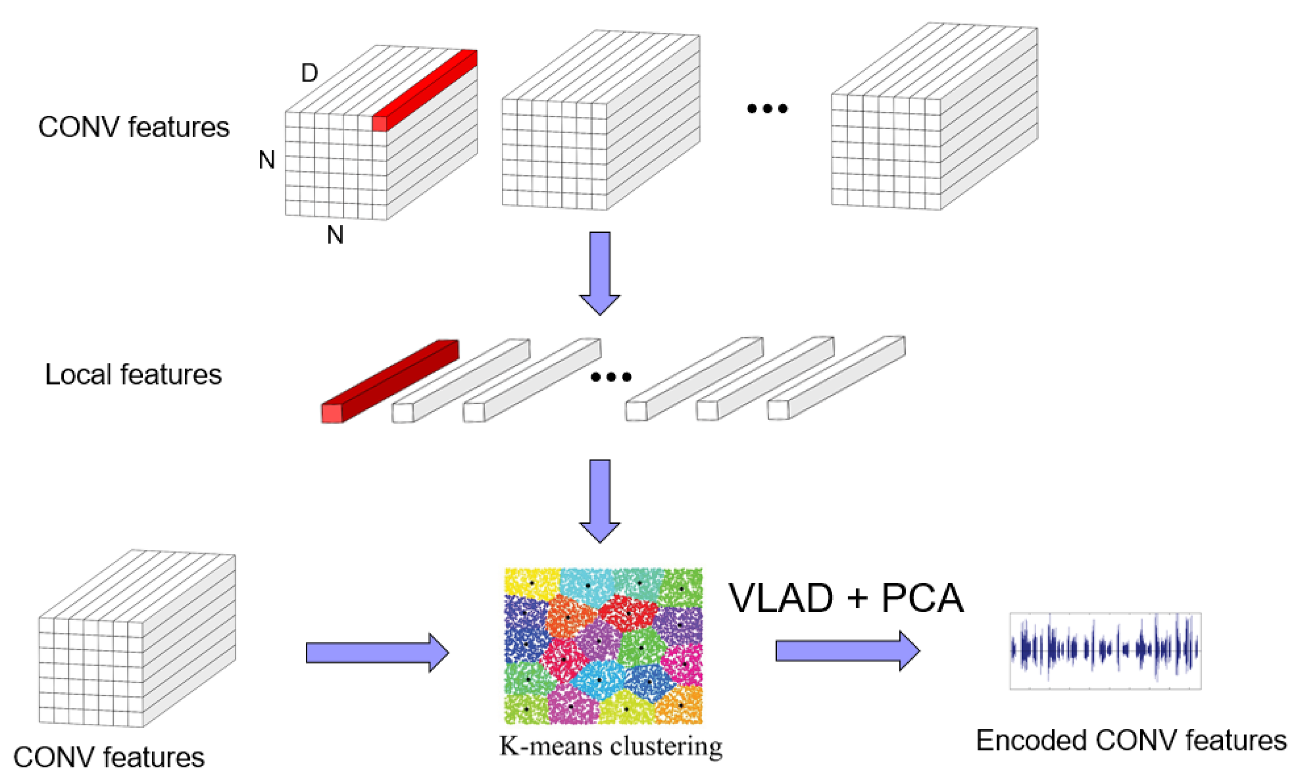

Figure 4. The framework of VLAD encoding [4] for CONV features.

The VLAD descriptor is constructed by accumulating the differences between each local feature and their corresponding nearest centers.

$$v_k^l = \sum_{c_k^l \in rNN(f_s^l)} f_s^l - c_k^l$$

C. Classification via SVM

Experiments

Experiments were performed on the EFC coral dataset [5], which consists of registered reflectance and fluorescence image pairs with annotated points.

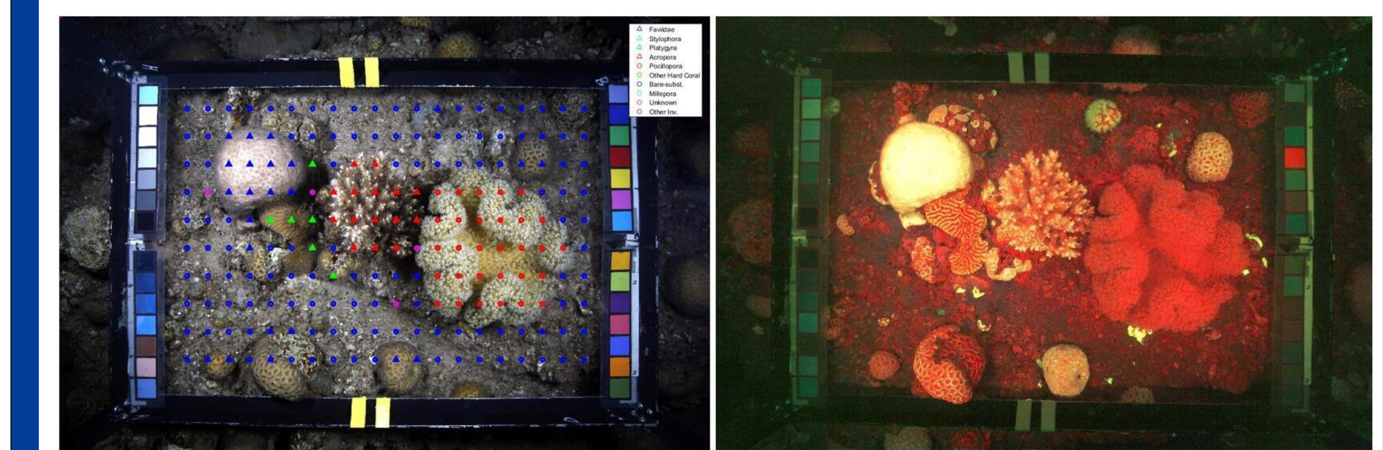

Figure 5. A pair of images from the EFC dataset.(Left)Reflectance image, (Right) Fluorescence image

A. Results on Reflectance images

Features	Feature dimension	Accuracy
CONV3	512	89.2±0.8%
FC6	4096	88.3±0.8%
CONV3+FC6	4608	**89.9**±0.8%
[5]	—	87.8±1.1%

B. Results on Fluorescence images

Features	Feature dimension	Accuracy
CONV3	512	86.5±1.0%
FC6	4096	85.4±1.0%
CONV3+FC6	4608	**86.7**±1.0%
[5]	—	85.5±1.2%

C. Results on fused images

Features	Feature dimension	Accuracy
CONV3	1024	90.7±0.8%
FC6	8192	90.4±0.8%
CONV3+FC6	9216	**91.4**±0.8%
[5]	—	90.5±0.8%

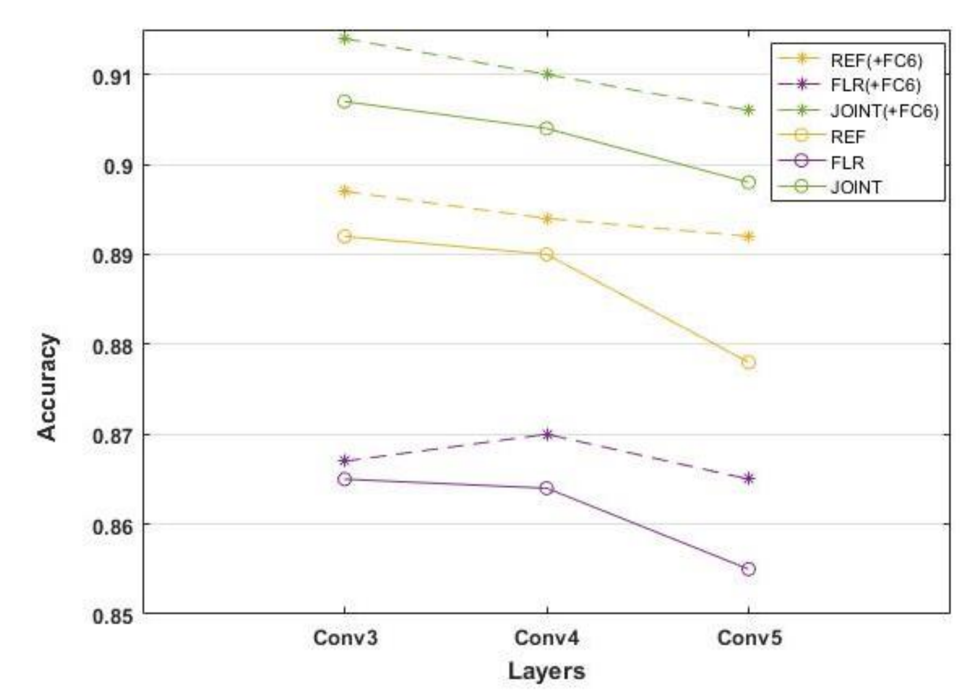

Figure 6. Performance of convolutional features from different layers: solid and dash lines correspond to individual convolutional features and those combined with FC6 features.

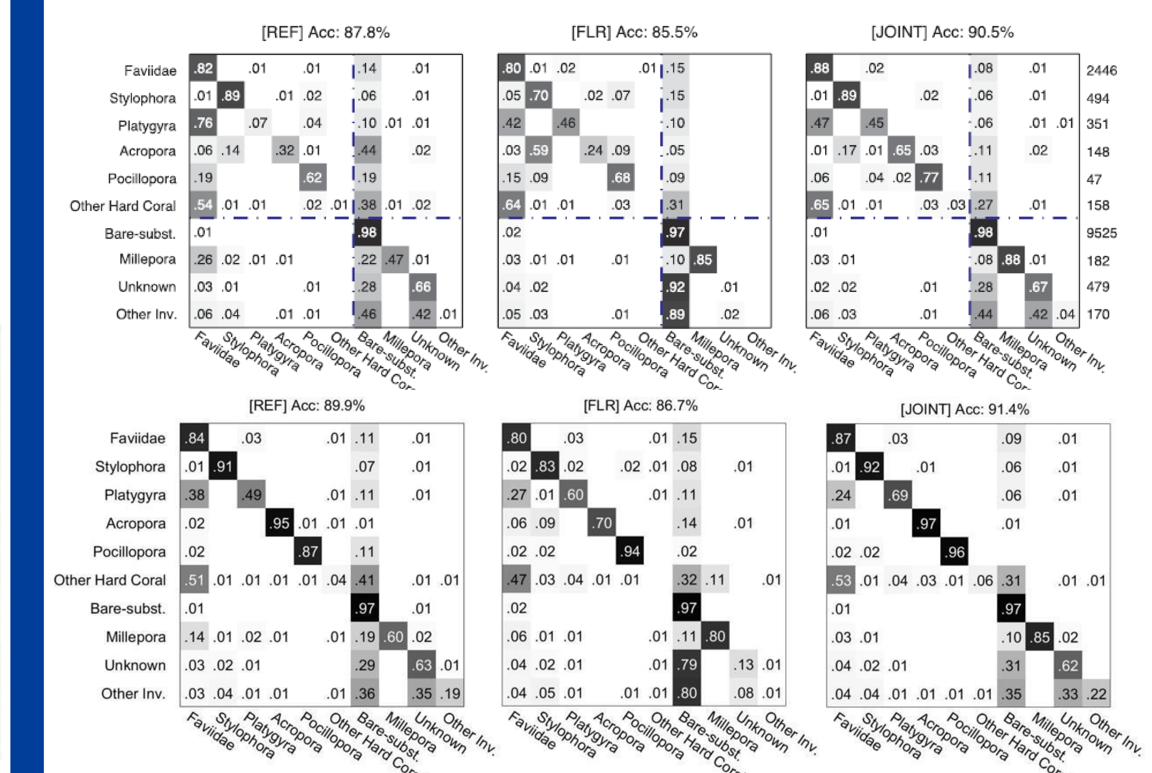

Figure 7. Confusion matrices. Results from (Up) [5] and (bottom) our methods based on combined FC6 and CONV3 features from (left) reflectance images, (middle) fluorescence images and (right) both images.

Conclusion

- In the case of lacking sufficient training data, CNN off-the-shelf features outperform training a small network from scratch.
- Low dimensional compact encoded convolutional features achieve comparable and better results against fully connected features on reflectance and fluorescence coral images.
- We suggest, in visual tasks, deep convolutional features should be the first choice for a new dataset, which is very different from the original dataset for training deep networks.
- Combining features from convolutional and fully connected layers can further improve the overall accuracy.

References

[1] A. Mahmood, M. Bennamoun, S. An, F. Sohel, F. Boussaid, R. Hovey, G.Kendrick, and R. Fisher, "Coral classification with hybrid feature representations," in Image Processing (ICIP), 2016 IEEE International Conference on, pp. 519–523, IEEE, 2016.

[2] J. Yosinski, J. Clune, Y. Bengio, and H. Lipson, "How transferable are features in deep neural networks?," in Advances in neural information processing systems, pp. 3320–3328, 2014.

[3] K. Chatfield, K. Simonyan, A. Vedaldi, and A. Zisserman,"Return of the devil in the details: Delving deep into convolutional nets," arXiv preprint arXiv:1405.3531, 2014.

[4] H. J´egou, M. Douze, C. Schmid, and P. P´erez, "Aggregating local descriptors into a compact image representation," in Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on, pp. 3304–3311, IEEE, 2010.

[5] O. Beijbom, T. Treibitz, D. I. Kline, G. Eyal, A. Khen, B. Neal, Y. Loya, B. G. Mitchell, and D. Kriegman, "Improving automated annotation of benthic survey images using wide-band fluorescence," Scientific reports, vol. 6, p. 23166, 2016.

Acknowledgements

This research is partially supported by China Scholarship Council funds (CSC, 201607565016) and Australian Research Council Grants (DP150104251 and DE120102960). The authors acknowledge NVIDIA for providing a Titan X GPU for the experiments involved in this research.