Enhanced Feature Pyramid Network for Semantic Segmentation

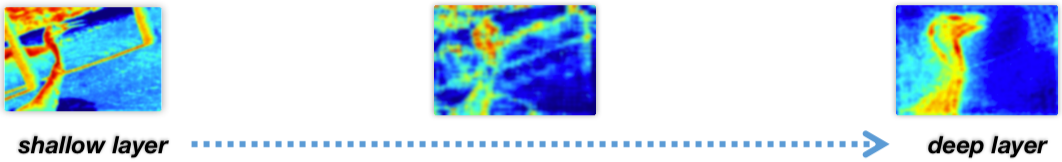
Mucong Ye, Jingpeng Ouyang, Ge Chen, Jing Zhang*, Xiaogang Yu
College of Software, Beihang University, Beijing, China

ICPR2020

Observation

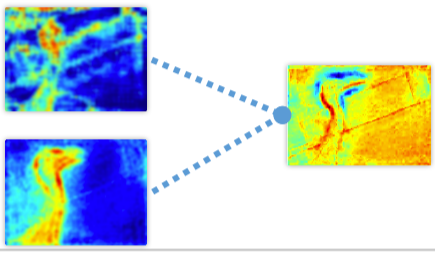
1. Semantic gap between shallow and deep features

- Shallow feature layers contain low-level appearance information (e.g., edges, lines, and corners)
- Deep layers contain more semantic information to distinguish different classes



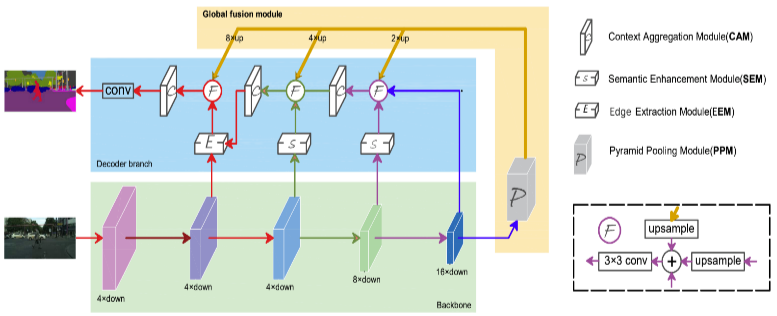
2. In the encoder-decoder architecture, a skip connection is frequently used to aggregate multi-scale context information

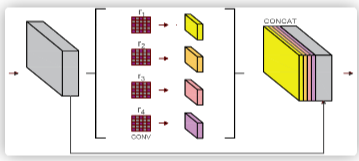
- The semantic difference between the shallow features and the deep features hinders the effective fusion of the different features
- Simply combining the shallow feature with the deep feature will bring some background "noise", which will affect the robustness of the feature



Networks

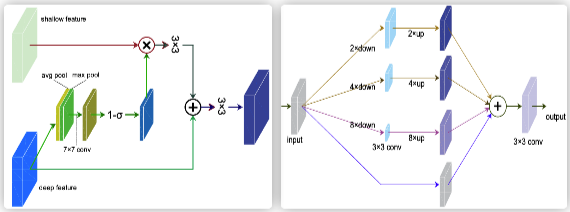
- **Enhanced Feature Pyramid Network (EFPN)** : To bridge the semantic gap and realize the effective fusion of multi-layer features
- **Global fusion model (GFM)** : Remedy the drawback of U-shape networks that top-down signals are gradually diluted





- **Semantic Enhancement Module (SEM)** : Enhance the shallow features

- **Edge Extraction Module (EEM)** : Based on attention
- **Context Aggregation Module (CAM)** : Better aggregation



Experiment & Conclusion

- **First, it is verified that the direct use of jump connections to fuse shallow features and deep features will affect the robustness of features**
- **Second, the effectiveness of global fusion module in encoder branch is further evaluated**
- **Finally, we prove the effectiveness of EFPN in bridging the semantic gap through experiments**