Based on Mask R-CNN, the PANet propose three features to improve the performance.

First, they create bottom-up path augmentation, to shorten information path and enhance feature pyramid in low-levels, which is really useful for localization.

Second, they develop adaptive feature pooling, to recover broken information path between each proposal and all feature levels. It is a simple component to aggregate features from all feature levels for each proposal, avoiding arbitrarily assigned results.

Finally, they add tiny fully-connected layers for mask prediction, which possess complementary properties to FCN. It is helpful to differentiate instances and recognize separate parts belonging to the same object.

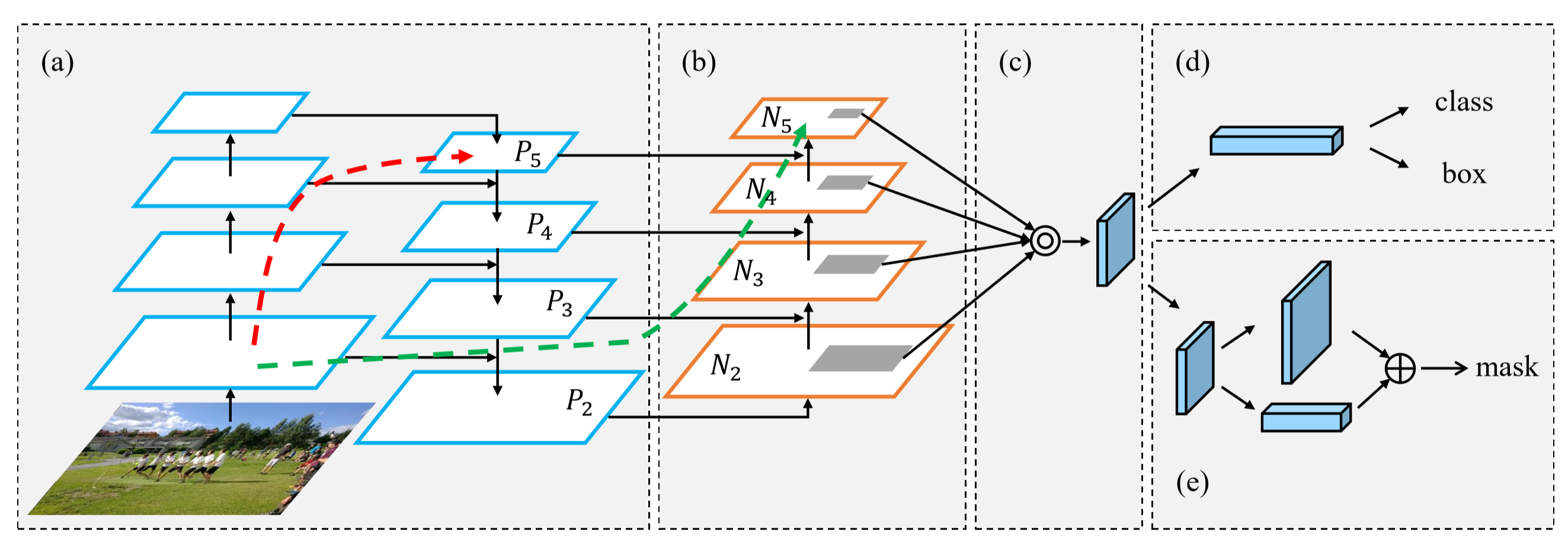


Figure 1. Illustration of our framework. (a) FPN backbone. (b) Bottom-up path augmentation. (c) Adaptive feature pooling.

(d) Box branch. (e) Fully-connected fusion. Note that we omit channel dimension of feature maps in (a) and (b) for brevity.

[1] Liu, Shu, et al. "Path aggregation network for instance segmentation." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2018.