

Long-Tailed Instance Segmentation using Extreme Value Distribution

Konstantinos Alexandridis, Shan Luo, Anh Nguyen and Jiankang Deng

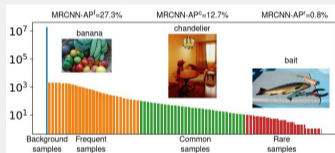


SUMMARY

1. Long-tailed Instance segmentation using Gumbel Distribution.
2. Gumbel activation function for classification layer.
3. Gumbel improves performance on LVIS [1] long-tailed instance segmentation benchmark.

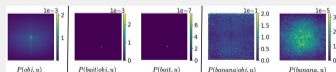
Contact-konsa15@liverpool.ac.uk

MOTIVATION



Instance segmentation models like MaskRCNN [2] do not detect rare category objects having low average precision (AP) in long-tailed benchmarks like LVIS.

IMBALANCED OBJECT DISTRIBUTION



Let obj be object occurrence and u the location in the normalized grid. The long-tailed object distribution has low expected values for both *frequent* and *rare* objects. This is due to **class** and **location** imbalance problem.

HYPOTHESIS

Sigmoid/Softmax activation functions cannot effectively model long-tailed object distribution due to imbalance problem. Gumbel distribution is a better choice for modeling the extreme values of long-tailed object distribution.

GUMBEL ACTIVATION

We develop Gumbel activation η_γ as:

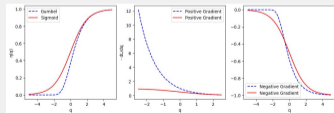
$$\eta_\gamma(q_i) = \exp(-\exp(-q_i)) \quad (1)$$

The loss using Gumbel Cross Entropy is:

$$L(\eta_\gamma(q_i), y_i) = \begin{cases} -\log(\eta_\gamma(q_i)), & \text{if } y_i = 1 \\ -\log(1 - \eta_\gamma(q_i)), & \text{if } y_i = 0 \end{cases} \quad (2)$$

The gradient of Eq. 2 is:

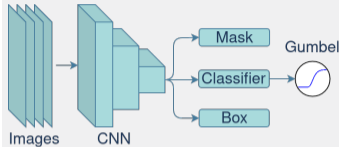
$$\frac{dL(\eta_\gamma(q_i), y_i)}{dq_i} = \begin{cases} -\exp(-q_i), & \text{if } y_i = 1 \\ \frac{\exp(-q_i)}{\exp(\exp(-q_i)) - 1}, & \text{if } y_i = 0 \end{cases} \quad (3)$$



REFERENCES

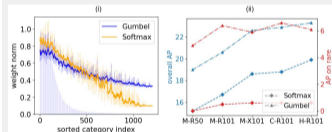
- [1] Gupta et al.: *LVIS: A dataset for large vocabulary instance segmentation*, CVPR (2019)
- [2] He et al.: *Mask r-cnn*, CVPR (2017)

GUMBEL INTEGRATION



Gumbel is integrated in the classification layer.

RESULTS



(i) Gumbel produces balanced weight norms in the classification layer, reducing classification bias. (ii) Gumbel has better performance than Softmax on (M)askRCNN-(R)esNet-50,101, ResNe(X)t-101, (C)ascade-MaskRCNN and (H)ybrid-Task-Cascade using LVIS dataset.

