

Gaussian-Based Pooling for Convolutional Neural Networks

Contributions: New Pooling Module

- VWe propose Parametric Probabilistic Pooling Module based on inverse-softplus Gaussian model.
- It naturally combines two approaches of

[Zeiler+13] M. Zeiler&R. Fergus, "Stochastic Pooling", ICLR2013

Previous Local Pooling



✓ Our Local Pooling





Takumi Kobayashi

National Institute of AIST, Japan

Parametric Probabilistic Model

Probabilistic model is formulated based on pooling functionality.



github.com/tk1980/GaussianPooling

Experimental Results

Ablation study by 13-layer Net (2 local pooling) on Cifar100.

Probabilistic Model

| Method | Error (%) |
|---|---|
| Gaussian Half-Gauss (fixed) Half-Gauss iSP-Gauss | $\begin{array}{c} 24.51 {\pm} 0.36 \\ 24.25 {\pm} 0.25 \\ 23.48 {\pm} 0.22 \\ 23.52 {\pm} 0.37 \end{array}$ |
| Average iSP-Gauss ($\sigma_0 = 0$) | $\begin{array}{c} 24.78 {\pm} 0.18 \\ 24.12 {\pm} 0.17 \end{array}$ |

Stochastic Method

| Method | Error (%) |
|--|--|
| Stochastic [33] Mixed [15] Half-Gauss iSP-Gauss | $\begin{array}{c} 24.52{\pm}0.18\\ 24.33{\pm}0.23\\ 23.48{\pm}0.22\\ 23.52{\pm}0.37\end{array}$ |
| S^{3} [34] + Stochastic [33] S^{3} [34] + Mixed [15] S^{3} [34] + Half-Gauss S^{3} [34] + iSP-Gauss | $\begin{array}{c} 24.01{\pm}0.20\\ 23.31{\pm}0.12\\ 23.12{\pm}0.17\\ 22.98{\pm}0.02 \end{array}$ |

Parametric Model

| Method | Error |
|--------------------------------------|-------------------------|
| NiN [17] ResNiN [7, 17] SE [9] | 24.49 24.33 23.99 |
| iSP-Gauss | 23.52 |

Stochasticity

| Method | Full $(\epsilon_{\boldsymbol{q}}^{c})$ | Р |
|-------------------------|---|--------|
| Half-Gauss iSP-Gauss | $\begin{array}{c} 23.48 {\pm} 0.22 \\ 23.52 {\pm} 0.37 \end{array}$ | 2 2 |

Global Pooling?

| Method | Error (%) |
|--|---|
| GAP GAP + DropOut [16] Half-Gauss iSP-Gauss | $\begin{array}{c} 24.78 {\pm} 0.18 \\ 24.58 {\pm} 0.27 \\ 24.54 {\pm} 0.14 \\ 23.83 {\pm} 0.18 \end{array}$ |

✓ Performance comparison by local pooling.

| Cifar 100 (a) 13-layer Net | | | IM AGENET | | | | | | | |
|-----------------------------------|------------------------------|-------------------|-----------|-------------------|------------|-------|---------------------|------------|-------|-------|
| | | (b) MobileNet [8] | | (c) ResNet-50 [7] | | | (d) ResNeXt-50 [30] | | | |
| Method | Error (%) | Method | Top-1 | Top-5 | Method | Top-1 | Top-5 | Method | Top-1 | Top-5 |
| skip | $24.83 {\pm} 0.15$ | skip | 29.84 | 10.35 | skip | 23.53 | 7.00 | skip | 22.69 | 6.65 |
| avg | $24.78{\pm}0.18$ | avg | 28.94 | 10.00 | avg | 22.61 | 6.52 | avg | 22.14 | 6.35 |
| max | $24.74{\pm}0.08$ | max | 29.23 | 10.02 | max | 22.99 | 6.71 | max | 22.20 | 6.24 |
| Stochastic [33 | $3] 24.52 \pm 0.18$ | Stochastic | 30.26 | 10.64 | stochastic | 25.47 | 7.87 | stochastic | 25.02 | 7.73 |
| Mixed [15] | $24.33{\pm}0.23$ | Mixed | 29.49 | 10.14 | Mixed | 22.81 | 6.53 | Mixed | 21.83 | 6.09 |
| DPP [22] | $24.59{\scriptstyle\pm0.15}$ | DPP | 28.92 | 9.92 | DPP | 22.52 | 6.63 | DPP | 21.84 | 5.98 |
| Gated [15] | $24.42{\pm}0.45$ | Gated | 28.62 | 9.86 | Gated | 22.27 | 6.33 | Gated | 21.63 | 5.99 |
| GFGP [1] | $24.41 {\pm} 0.22$ | GFGP | 27.68 | 9.27 | GFGP | 21.79 | 5.95 | GFGP | 21.35 | 5.74 |
| Half-Gauss | 23.48 ± 0.22 | Half-Gauss | 5 27.96 | 9.38 | Half-Gauss | 21.66 | 5.88 | Half-Gauss | 20.89 | 5.72 |
| iSP-Gauss | 23.52 ± 0.37 | iSP-Gauss | 27.33 | 9.00 | iSP-Gauss | 21.37 | 5.68 | iSP-Gauss | 20.66 | 5.60 |

✓ Estimated parameters during training.





- (%)

Partial (ϵ^c 23.60 ± 0.07 23.68 ± 0.06

| 3- | layer | Ne |
|----|-------|----|
| | conv | |
| | conv | |
| | conv | |
| | pool | |
| | conv | |
| | conv | |
| | conv | |
| | pool | |
| | conv | |
| | conv | |
| | conv | |
| | GAP | |
| | FC | |
| | | |

