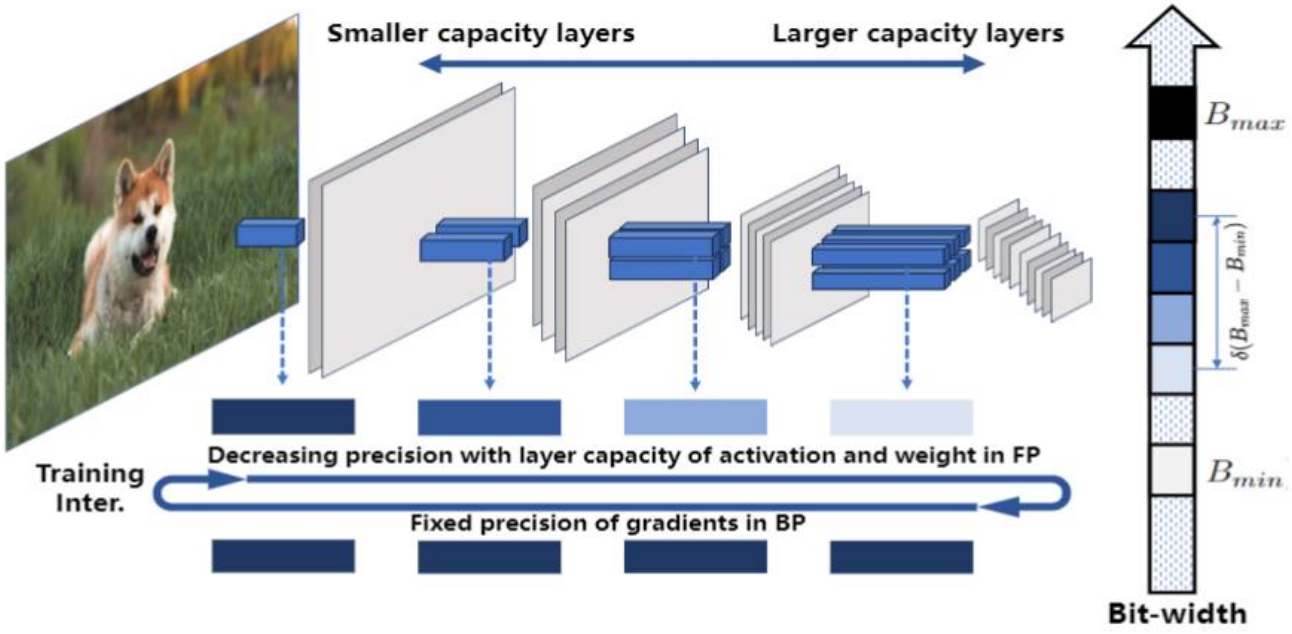
# Efficient Deep Neural Network Training via Decreasing Precision with Layer Capacity

**Ao SHEN, Zhiquan LAI, Tao SUN, Shengwei LI, Keshi GE, Weijie LIU, Dongsheng LI**

Frontiers of Computer Science, DOI: 10.1007/s11704-024-40669-3

# Problems & Ideas

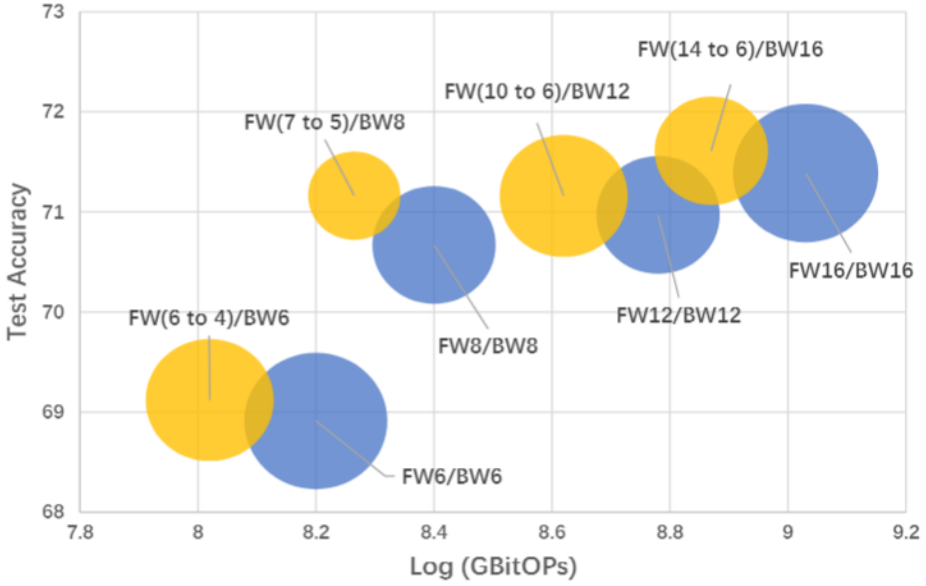
- Traditional DNN Low-Precision Training Methods:
  - Focus primarily on minimizing low-precision noise, requiring additional analysis.
  - Overlook potential benefits of low-precision during training.
- Ideas:
  - Systematic analysis of potential benefits of low-precision training: Regularization effects & Enhanced generalization capability
  - Proposal of Decreasing Precision with layer Capacity (DPC).




- Layer-specific bit-width allocation based on model architecture
- Determination of viable bit-width ranges for quantization
- Direct precision assignment according to model characteristics

# Main Contributions

- Contributions:
  - We analyzed low precision benefits in DNN training: showing enhanced generalization, regularization effects, and correlation with layer capacity;
  - We introduce DPC - a method that assigns bit-widths based on layer capacity without requiring complex analysis or precision criteria;
  - Results show DPC's effectiveness: reduced training costs and improved performance, with consistent results across trials. Additional experiments confirm low precision's training benefits, supporting our theoretical analysis.




Training ResNet on CIFAR while optimizing accuracy, computational efficiency, and variance

| Model | Method | Prec.(spatial/temporal) | Acc(%) | BitOPs-Sav(%) |
|---|---|---|---|---|
| ResNet-18 | Full prec. | 32 / 0 | 70.64 | - |
| | CPT | 32 / 24 | 70.67 (+0.03) | 39.01 |
| | **DPC** | 32 to 8 / 0 | 70.68 (**+0.04**) | 39.51 |
| | SBM | 22 / 0 | 70.53 | 52.53 |
| | **DPC+CPT** | 32 to 8 / -50% | 70.53 (0.00) | 55.08 |
| ResNet-34 | Full prec. | 32 / 0 | 74.13 | - |
| | CPT | 32 / 24 | 74.14 (+0.01) | 39.01 |
| | **DPC** | 32 to 8 / 0 | 74.15 (**+0.02**) | 43.24 |
| | SBM | 22 / 0 | 73.99 | 52.53 |
| | **DPC+CPT** | 32 to 8 / -50% | 74.08 (**+0.09**) | 58.53 |
| ResNet-50 | Full prec. | 32 / 0 | 76.59 | - |
| | **DPC** | 32 to 10 / 0 | 76.63 (**+0.04**) | 44.37 |
| | SBM | 22 / 0 | 76.48 | 52.53 |
| | **DPC+CPT** | 32 to 10 / -50% | 76.57 (**+0.11**) | 59.82 |
| MN-V2 $\alpha = 2$ | Full prec. | 32 / 0 | 70.51 | - |
| | **DPC** | 32 to 12 / 0 | 70.51 (0.00) | 37.68 |
| | SBM | 22 / 0 | 70.33 | 52.53 |
| | **DPC+CPT** | 32 to 12 / -50% | 70.53 (**+0.20**) | 51.49 |

Training on ImageNet achieves a win-win of high accuracy and reduced computation