LOCO-Reg

Locality-Promoting Representation Learning

Johannes Schneider
Assistant Professor in Data Science
Weights in Convolutional Networks

... are not of the same magnitude
On average weights near the center are larger

3x3 Filter

|  S(mall) | M(edium) | S  |
|---------|---------|----|
|  M      | Large   | M  |
|  S      | M       | S  |

Distribution in log-scale of absolute weights of 3x3 filters at center, left middle and left upper
Let us think about that...

Filter = common pattern in feature maps

Red or violet? Red preferred because they have a large center

⇒ More robust (to noise, variation)
Model from Physics

- We want maximal feature cohesion
- Activation, weight = Mass
- Cohesion = Gravitational Force

3x3 spatial filter

|   |   |   |
|---|---|---|
| 1 | 2 | 2 |
| 2 | 6 | 3 |
| 1 | 3 | 2 |

\[
d(m_{0,0}, m_{0,1}) = 1 \\
d(m_{0,0}, m_{0,2}) = 2 \\
d(m_{0,0}, m_{1,1}) = 2.5 \\
d(m_{0,0}, m_{1,2}) = 5.5 \\
d(m_{0,0}, m_{2,2}) = 8.5
\]

Theorem 1. For any feature strength distribution \( m' \leq m_c, m_{co}, m_n < (1 + \epsilon)m' \) with \( \epsilon \in [0, 0.675] \), the cohesion \( F_{tot} \) of the feature is increased most by increasing \( m_c \), and more by increasing any \( m_n \in M_n \) than any \( m_{co} \in M_{co} \) for arbitrary \( m' \), center \( m_c = m_{1,1} \), direct neighbors \( M_n := \{m_{1,0}, m_{0,1}, m_{2,1}, m_{1,2}\} \) and corners \( M_{co} := \{m_{0,0}, m_{2,0}, m_{2,2}, m_{0,2}\} \) (Figure 3).
Implementation: LOCO-Reg

Standard L2-regularization pushes all weights to be equal
⇒ This reduces central weights too much
LOCO-Reg: Regularize outer weights more than more central weights

Base L2
Regularization Constant

|       | LOCO- Regularization weights |
|-------|-------------------------------|
| γ > η |                   | η > 1 |                   | γ > η |
| η > 1 |                   | 1     |                   | η > 1 |

|       |       |       |       |       |
|-------|-------|-------|-------|-------|
| λ     | γ > η | η > 1 | γ > η |       |

| Dataset | Architecture | (η, γ) | Avg. Accuracy for different λ | Best Acc. |
|---------|--------------|--------|-------------------------------|-----------|
| cifar10 | MobileNet    | (1,1)  | .8611 .8686 .8688 .8647 .8688 | .8688     |
| cifar10 | MobileNet    | (1.4, 1.56) | .8618 .8701* .8714 .8657 .8714 |           |
| cifar10 | MobileNet    | (1.8, 2.13) | .8619* .8692* .8721*γ .8668*γ .8721*γ |           |
| cifar10 | ResNet      | (1,1)  | .9191 .9227 .9236 .9222 .9236 | .9236     |
| cifar10 | ResNet      | (1.4, 1.56) | .921 .9253* .9242 .9224 .9253* | .9253*    |
| cifar10 | ResNet      | (1.8, 2.13) | .9186 .9244* .9237 .9236 .9244* | .9244*    |
| cifar10 | VGG        | (1,1)  | .8754 .8761 .882 .8858 .8858 | .8858     |
| cifar10 | VGG        | (1.4, 1.56) | .8722 .884** .8858** .8869 .8869 |           |
| cifar10 | VGG        | (1.8, 2.13) | .8808** .8816* .8875*** .8884* .8884* |           |
| cifar100| MobileNet  | (1,1)  | .5926 .6116 .6182 .6155 .6182 | .6182     |
| cifar100| MobileNet  | (1.4, 1.56) | .5941 .6124 .6182 .6149 .6182 | .6182     |
| cifar100| MobileNet  | (1.8, 2.13) | .5935 .6144 .6199 .6184* .6199 | .6199     |
| cifar100| ResNet     | (1,1)  | .702 .71 .7156 .7124 .7156 | .7156     |
| cifar100| ResNet     | (1.4, 1.56) | .702 .7129* .7163 .7146 .7163 | .7163     |
| cifar100| ResNet     | (1.8, 2.13) | .7022 .7116 .7198*** .7142 .7198*** |           |
| cifar100| VGG       | (1,1)  | .6415 .6551 .6597 .6599 .6599 | .6599     |
| cifar100| VGG       | (1.4, 1.56) | .6432 .6583* .6665*** .6645* .6665*** |           |
| cifar100| VGG       | (1.8, 2.13) | .6449* .6629*** .6653** .6671*** .6671*** |           |
| fashion | MobileNet  | (1,1)  | .9403 .9402 .939 .9369 .9403 | .9403     |
| fashion | MobileNet  | (1.4, 1.56) | .9398 .9406 .9385 .9372 .9406 | .9406     |
| fashion | MobileNet  | (1.8, 2.13) | .9402 .9408 .9398 .9371 .9408 | .9408     |
| fashion | ResNet    | (1,1)  | .9501 .9504 .9494 .9492 .9504 | .9504     |
| fashion | ResNet    | (1.4, 1.56) | .9496 .951 .9506** .9489 .951 | .951      |
| fashion | ResNet    | (1.8, 2.13) | .9509* .9505 .9515 .9494 .9504 | .9504     |
| fashion | VGG       | (1,1)  | .9404 .942 .9417 .9426 .9426 | .9426     |
| fashion | VGG       | (1.4, 1.56) | .941 .9414 .9419 .9436* .9436* | .9436*    |
| fashion | VGG       | (1.8, 2.13) | .9423 .9417 .9436* .9437* .9437* | .9437*    |
THANKS