Initialization Using Perlin Noise for Training Networks with a Limited Amount of Data

Nakamasa Inoue\textsuperscript{1*}, Eisuke Yamagata\textsuperscript{1*}, Hirokatsu Kataoka\textsuperscript{2}

\textsuperscript{1}Tokyo Institute of Technology, \textsuperscript{2}AIST

* equal contribution

ICPR 2020
Overview

We propose a network initialization method using Perlin noise.

Gaussian Initialization

Proposed Initialization
Key Idea

Initialize network parameters by solving an artificial Perlin noise classification problem.
Perlin Noise

Gradient noise proposed by Ken Perlin in 1983

(a) $2^n \times 2^m$ Grid

(b) Random Vectors

(c) Gradients

(d) Perlin Noise Sample
Perlin Noise Categories

Fine-to-cause categories on Perlin noise

$2^n \times 2^m$ grid

$1 \leq n \leq N, 1 \leq m \leq M$

$NM$ categories

$y, 2^n \times 2^m$ : Noise Category and Grid
Instances

Intra-category variation of noise samples

1, $2^1 \times 2^1$

13, $2^2 \times 2^3$

43, $2^5 \times 2^3$
Experiments

Comparison with other methods on four datasets.

**TABLE I**

**PERFORMANCE COMPARISON ON FOUR DATASETS. CLASSIFICATION ACCURACIES (%) FOR EACH DATASET WITH TWO TYPES OF NETWORK ARE SHOWN.**

<table>
<thead>
<tr>
<th>Method</th>
<th>Cifar-10</th>
<th>Cifar-100</th>
<th>Omniglot</th>
<th>DTD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal initialization</td>
<td>92.62</td>
<td>93.47</td>
<td>75.16</td>
<td>75.59</td>
</tr>
<tr>
<td>Xavier initialization</td>
<td>92.30</td>
<td>93.58</td>
<td>73.85</td>
<td>75.14</td>
</tr>
<tr>
<td>He initialization</td>
<td>93.50</td>
<td>93.43</td>
<td>74.17</td>
<td>75.73</td>
</tr>
<tr>
<td>Proposed method</td>
<td><strong>93.76</strong></td>
<td><strong>94.27</strong></td>
<td><strong>77.42</strong></td>
<td><strong>78.21</strong></td>
</tr>
</tbody>
</table>

Fig. 5. Visualization of filters of the first convolutional (conv1) layer. Filters before and after training on Cifar-10 are shown. (a) He initialization and (b) proposed method.

Come to our poster!