Evaluation of the Use of Convolutional Neural Networks with Variable Stride for Skin Lesion Classification

Oge Marques, PhD and Luiz Zaniolo, MS
Florida Atlantic University (Boca Raton, FL)
Motivation

- Convolutional Neural Networks (CNNs): extremely popular and effective in image classification tasks.
- Convolution layers of a CNN include a *stride* parameter that dictates how big are the steps for sampling when scanning the input layer to run convolutional operations.
  - The vast majority of CNNs use a *fixed* stride value.
- Previous research (Zaniolo & Marques, MTAP 2019): changing the stride value in CNNs depending on the position of the pixel within the image leads to an increase in processing speed can be achieved without sacrificing accuracy.
Convolution with Fixed Stride

Source: https://ujjwalkarn.me/2016/08/11/intuitive-explanation-convnets/
Variable Stride Mechanism

(Zaniolo & Marques, MTAP 2019)
Early Experiments: good results

Tests on emotions recognition
(Radboud Dataset)

<table>
<thead>
<tr>
<th>Stride</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>94.01 %</td>
</tr>
<tr>
<td>2</td>
<td>91.93 %</td>
</tr>
<tr>
<td>3</td>
<td>86.97 %</td>
</tr>
<tr>
<td>Variable</td>
<td>95.83 %</td>
</tr>
</tbody>
</table>

(Zaniolo & Marques, MTAP 2019)
Early Experiments: not-so-good results

Tests on 8-class scene classification  
(MIT Places Dataset)

<table>
<thead>
<tr>
<th>Stride</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>71.33 %</td>
</tr>
<tr>
<td>2</td>
<td>68.75 %</td>
</tr>
<tr>
<td>3</td>
<td>64.25 %</td>
</tr>
<tr>
<td>Variable</td>
<td>62.50 %</td>
</tr>
</tbody>
</table>

(Zaniolo & Marques, MTAP 2019)
Additional Experiments: good results

Tests on digit recognition (MNIST Dataset)

<table>
<thead>
<tr>
<th>Stride</th>
<th>Accuracy</th>
<th>Training Time</th>
<th>Inference Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>98.8 %</td>
<td>105 min</td>
<td>6.6 ms</td>
</tr>
<tr>
<td>2</td>
<td>97.4 %</td>
<td>41 min</td>
<td>2.6 ms</td>
</tr>
<tr>
<td>3</td>
<td>96.2 %</td>
<td>32 min</td>
<td>1.6 ms</td>
</tr>
<tr>
<td>Variable</td>
<td>98.1 %</td>
<td>41 min</td>
<td>2.6 ms</td>
</tr>
</tbody>
</table>

Tests on object recognition (Fashion MNIST Dataset)

<table>
<thead>
<tr>
<th>Stride</th>
<th>Accuracy</th>
<th>Training Time</th>
<th>Inference Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>91.7 %</td>
<td>335 min</td>
<td>6.6 ms</td>
</tr>
<tr>
<td>2</td>
<td>88.5 %</td>
<td>139 min</td>
<td>2.6 ms</td>
</tr>
<tr>
<td>3</td>
<td>87.2 %</td>
<td>111 min</td>
<td>1.6 ms</td>
</tr>
<tr>
<td>Variable</td>
<td>89.7 %</td>
<td>139 min</td>
<td>2.6 ms</td>
</tr>
</tbody>
</table>

(Zaniolo & Marques, MTAP 2019)
Additional Experiments: not-so-good results

Tests on image classification
(CIFAR-10 Dataset)

<table>
<thead>
<tr>
<th>Stride</th>
<th>Accuracy</th>
<th>Training Time</th>
<th>Inference Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>76.0 %</td>
<td>332 min</td>
<td>59.1 ms</td>
</tr>
<tr>
<td>2</td>
<td>69.9 %</td>
<td>132 min</td>
<td>22.5 ms</td>
</tr>
<tr>
<td>3</td>
<td>62.0 %</td>
<td>93 min</td>
<td>12.9 ms</td>
</tr>
<tr>
<td>Variable</td>
<td>65.5 %</td>
<td>132 min</td>
<td>22.5 ms</td>
</tr>
</tbody>
</table>

(Zaniolo & Marques, MTAP 2019)
Hypothesis

The use of variable stride in skin lesion images whose main contents are in the central portion of the image will lead to improved performance (when compared to the baseline case of comparable computational complexity, i.e., fixed stride = 2).
Methods

- **Baseline CNN**
- **HAM10000 dataset**
  - 7 classes of skin lesions: Melanocytic nevi, Melanoma, Benign keratosis-like lesions, Basal cell carcinoma, Actinic keratoses, Vascular lesions, and Dermatofibroma.
  - 95% of the images were randomly selected for **training** and 5% used for **validation**.
- **MATLAB Deep Learning Toolbox**
- **MacBookPro with 2.7 GHz Intel Core i5 processor, and 8 GB 1867 MHz DDR3 memory.**
Training

The network was trained using 20 epochs and a decaying learning rate, starting with 0.0003 and reducing it by half every 2 epochs.
## Results

<table>
<thead>
<tr>
<th>Stride</th>
<th>Accuracy</th>
<th>Training Time</th>
<th>Inference Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>77.0 %</td>
<td>258 min</td>
<td>31.9 ms</td>
</tr>
<tr>
<td>2</td>
<td>76.4 %</td>
<td>105 min</td>
<td>10.9 ms</td>
</tr>
<tr>
<td>3</td>
<td>75.5 %</td>
<td>70 min</td>
<td>5.8 ms</td>
</tr>
<tr>
<td>Variable</td>
<td>78.1 %</td>
<td>105 min</td>
<td>10.9 ms</td>
</tr>
</tbody>
</table>
Extended work - Brain MRI

253 images (155 tumorous, 98 non-tumorous) from Kaggle’s Brain MRI Images for Brain Tumor Detection dataset.

Tumorous brains

Healthy brains
Extended work - Training

The network was trained using 20 epochs and a fixed learning rate of 0.0003.
## Extended work - Results

<table>
<thead>
<tr>
<th>Stride</th>
<th>Accuracy</th>
<th>Training Time</th>
<th>Inference Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>89.6 %</td>
<td>12.5 min</td>
<td>50.6 ms</td>
</tr>
<tr>
<td>2</td>
<td>84.0 %</td>
<td>4.2 min</td>
<td>15.3 ms</td>
</tr>
<tr>
<td>3</td>
<td>81.6 %</td>
<td>2.4 min</td>
<td>7.4 ms</td>
</tr>
<tr>
<td>Variable</td>
<td>87.2 %</td>
<td>4.2 min</td>
<td>15.3 ms</td>
</tr>
</tbody>
</table>
Conclusion

We evaluated a method for using CNNs with variable stride in the context of skin lesion classification and demonstrated that they can achieve higher accuracy and take less time to train or make predictions than the same networks using fixed stride.

The results of the current study show the potential for wider adoption of variable stride mechanism in a variety of medical image classification applications.