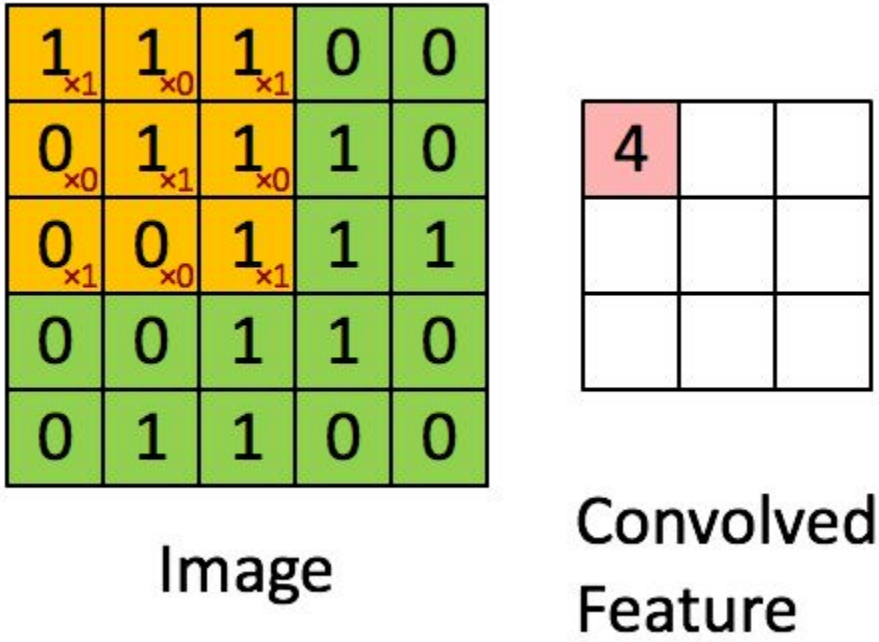
Evaluation of the Use of Convolutional Neural Networks with Variable Stride for Skin Lesion Classification

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

Motivation

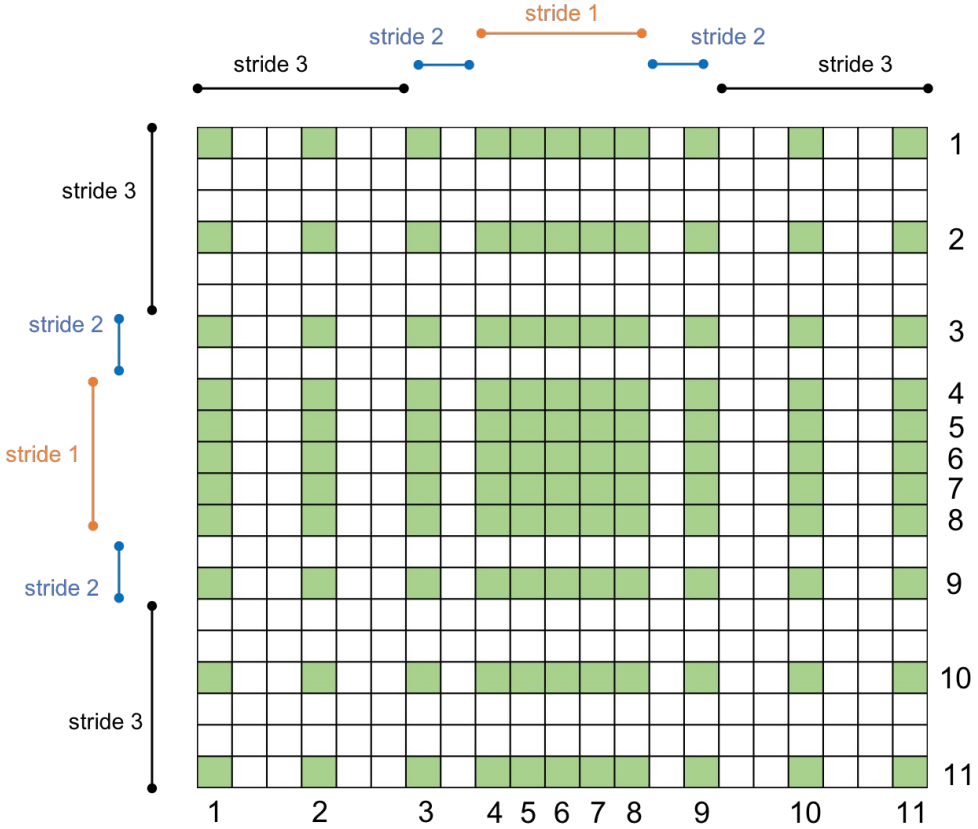
- Convolutional Neural Networks (CNNs): extremely popular and effective in image classification tasks.
- Convolution layers of a CNN include a *stride* parameter that dictates how big are the steps for sampling when scanning the input layer to run convolutional operations.
  - The vast majority of CNNs 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

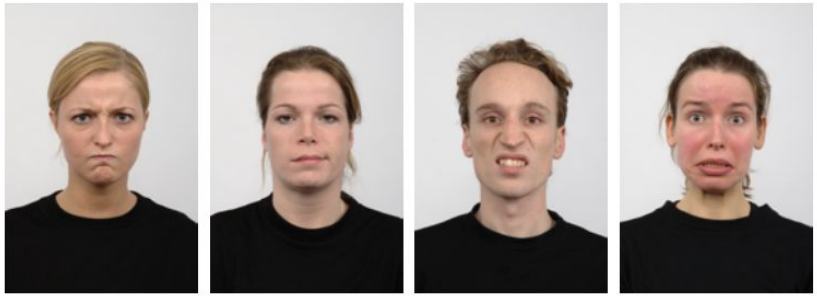


(Zaniolo & Marques, MTAP 2019)

Early Experiments: good results

Tests on emotions recognition
(Radboud Dataset)

Stride	Accuracy
1	94.01 %
2	91.93 %
3	86.97 %
Variable	**95.83** %



(a) Angry (b) Contemptuous (c) Disgusted (d) Fearful

(a) Happy (b) Neutral (c) Sad (d) Surprised

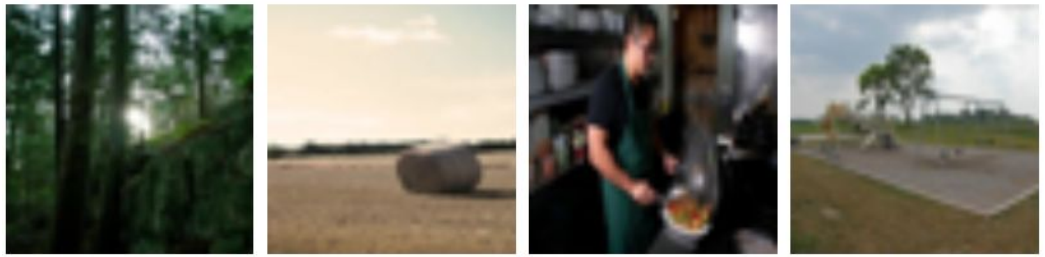
(Zaniolo & Marques, MTAP 2019)

Early Experiments: not-so-good results

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

Stride	Accuracy
1	71.33 %
2	68.75 %
3	64.25 %
Variable	**62.50 %**

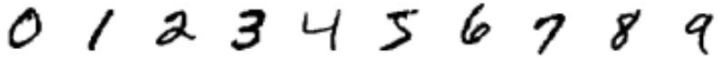
(a) Attic (b) Bathroom (c) Beach (d) Bedroom



(a) Forest (b) Hayfield (c) Kitchen (d) Playground

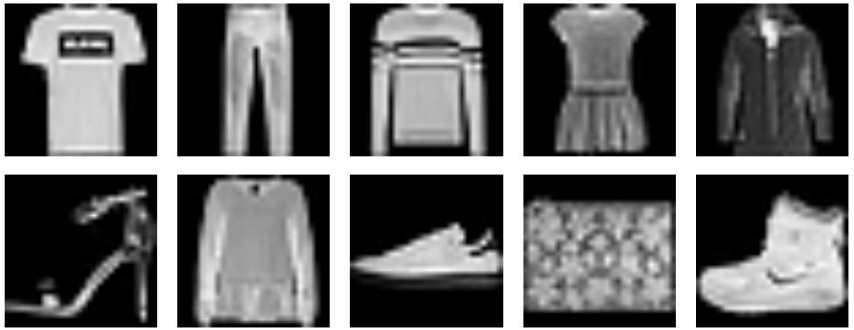
(Zaniolo & Marques, MTAP 2019)

Additional Experiments: good results



Tests on digit recognition
(MNIST Dataset)

Stride	Accuracy	Training Time	Inference Time
1	98.8 %	105 min	6.6 ms
2	97.4 %	41 min	2.6 ms
3	96.2 %	32 min	1.6 ms
Variable	**98.1 %**	**41 min**	**2.6 ms**



Tests on object recognition
(Fashion MNIST Dataset)

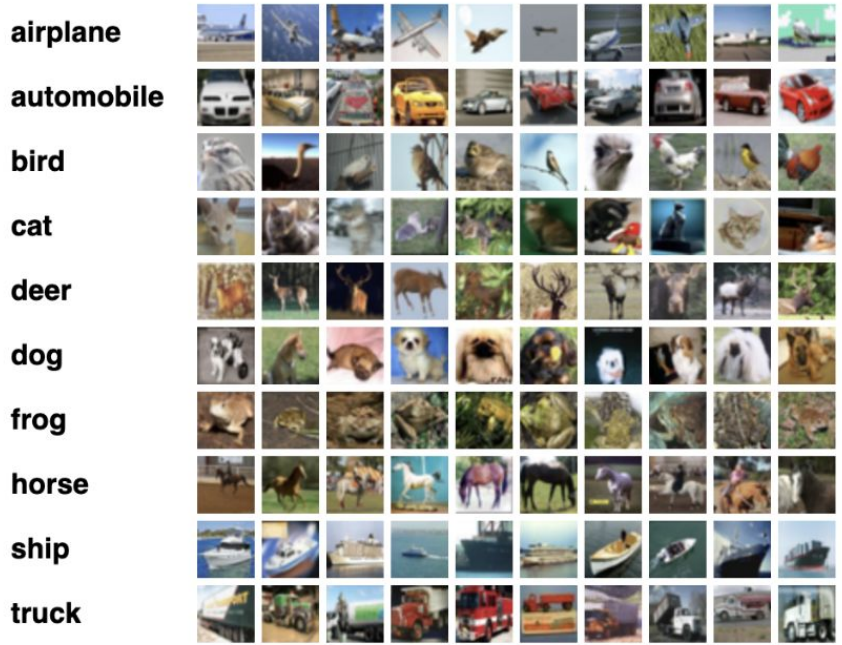
Stride	Accuracy	Training Time	Inference Time
1	91.7 %	335 min	6.6 ms
2	88.5 %	139 min	2.6 ms
3	87.2 %	111 min	1.6 ms
Variable	**89.7 %**	**139 min**	**2.6 ms**

(Zaniolo & Marques, MTAP 2019)

Additional Experiments: not-so-good results

Tests on image classification
(CIFAR-10 Dataset)

Stride	Accuracy	Training Time	Inference Time
1	76.0 %	332 min	59.1 ms
2	69.9 %	132 min	22.5 ms
3	62.0 %	93 min	12.9 ms
Variable	**65.5 %**	**132 min**	**22.5 ms**



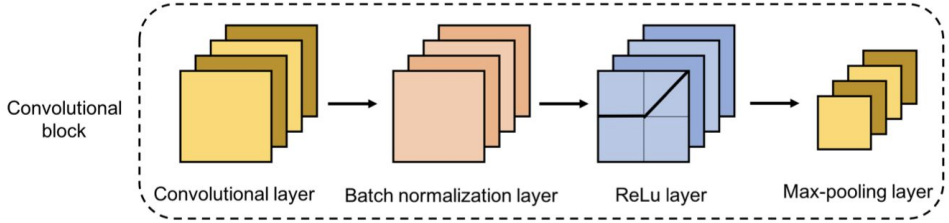
(Zaniolo & Marques, MTAP 2019)

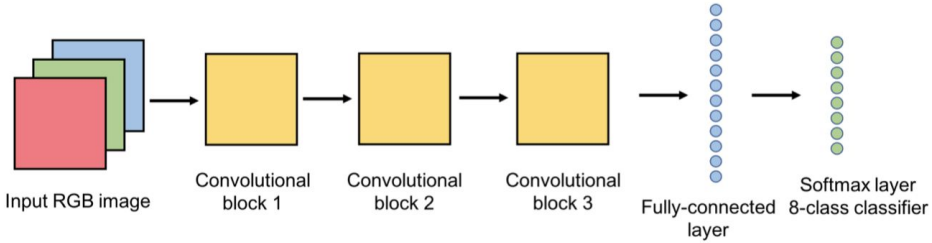
Hypothesis

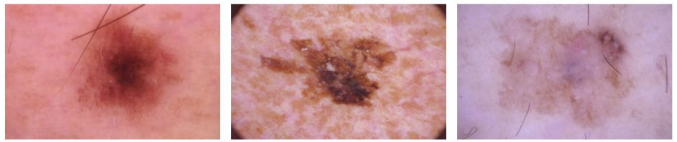
The use of **variable stride** in **skin lesion images** whose main contents are in the central portion of the image will lead to improved performance (when compared to the baseline case of comparable computational complexity, i.e., fixed stride = 2).

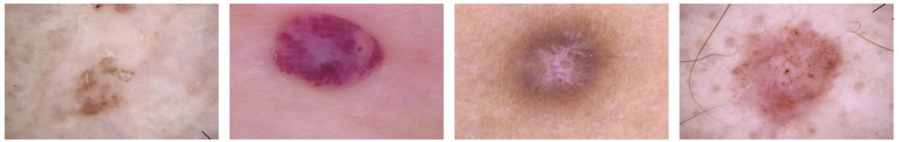
Methods

- **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.



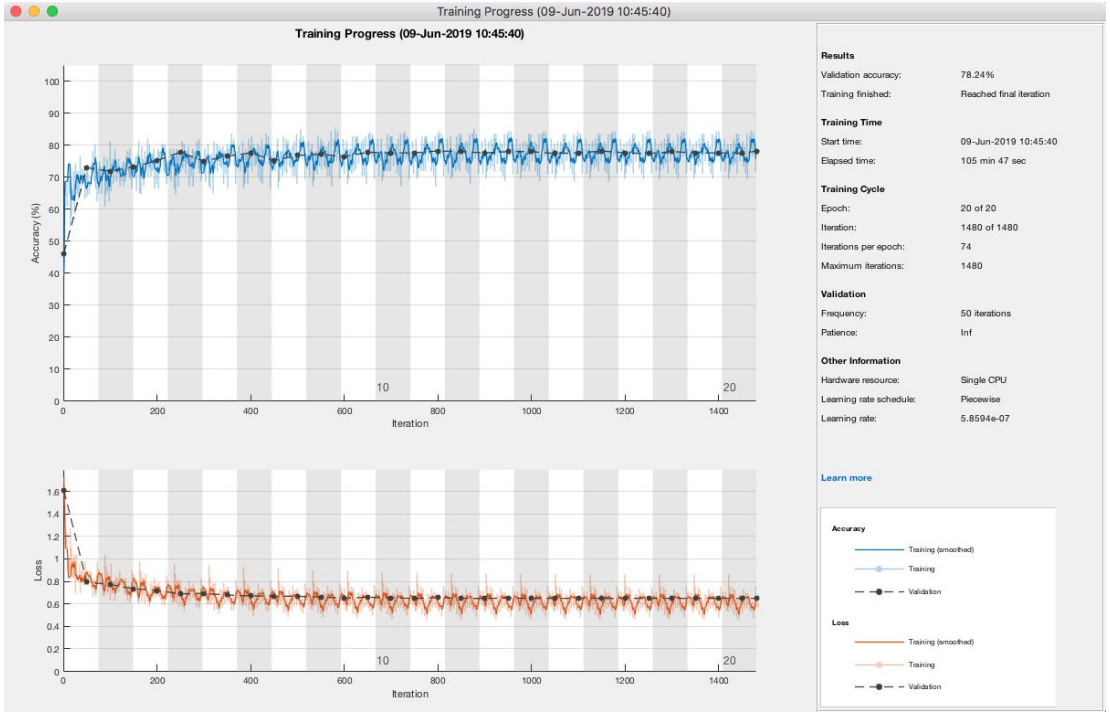






Training

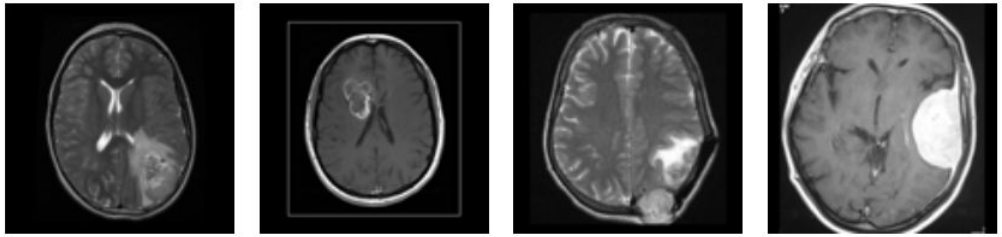
The network was trained using 20 epochs and a decaying learning rate, starting with 0.0003 and reducing it by half every 2 epochs.



Results

Stride	Accuracy	Training Time	Inference Time
1	77.0 %	258 min	31.9 ms
2	76.4 %	105 min	10.9 ms
3	75.5 %	70 min	5.8 ms
Variable	**78.1 %**	**105 min**	**10.9 ms**

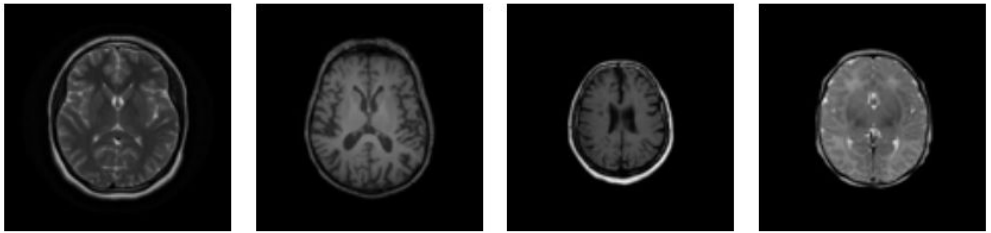
Extended work - Brain MRI



Tumorous brains

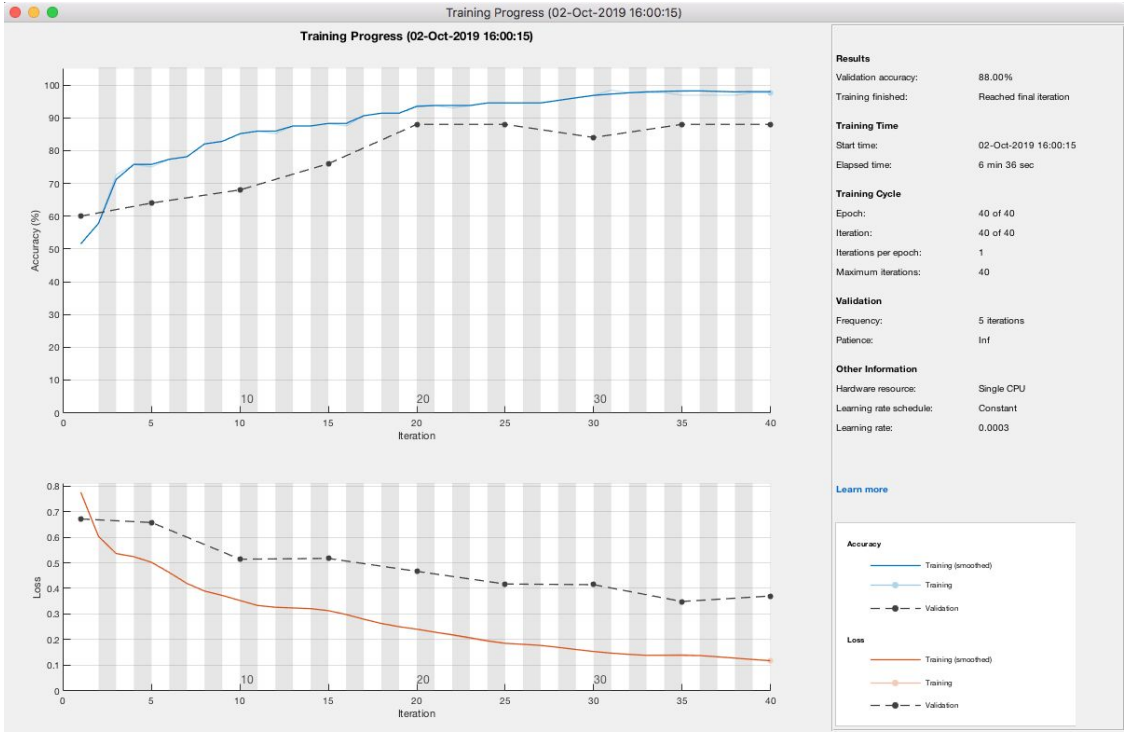
253 images (155 tumorous, 98 non-tumorous) from Kaggle's *Brain MRI Images for Brain Tumor Detection* dataset.

Healthy brains



Extended work - Training

The network was trained using 20 epochs and a fixed learning rate of 0.0003.



Extended work - Results

Stride	Accuracy	Training Time	Inference Time
1	89.6 %	12.5 min	50.6 ms
2	84.0 %	4.2 min	15.3 ms
3	81.6 %	2.4 min	7.4 ms
Variable	**87.2 %**	**4.2 min**	**15.3 ms**

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.