Mind the Pool: Convolutional Neural Networks can Overfit Input size

Bilal Alsallakh
Voxel AI

David Yan
Meta AI

Narine Kokhlikyan
Meta AI

Vivek Miglani
Meta AI

Orion Reblitz-Richardson
Meta AI

Pamela Bhattacharya
Meta

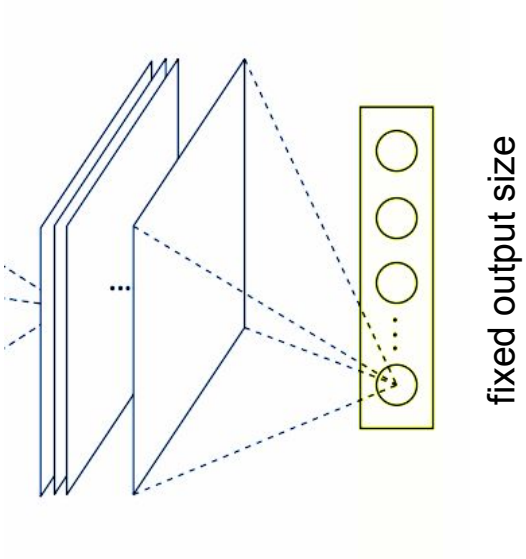
Motivation

CNNs can handle arbitrary input size, thanks to:

- Fully-convolutional backbone =>

  Feature map size can vary with input size.
- Global pooling (e.g. global average pooling) =>

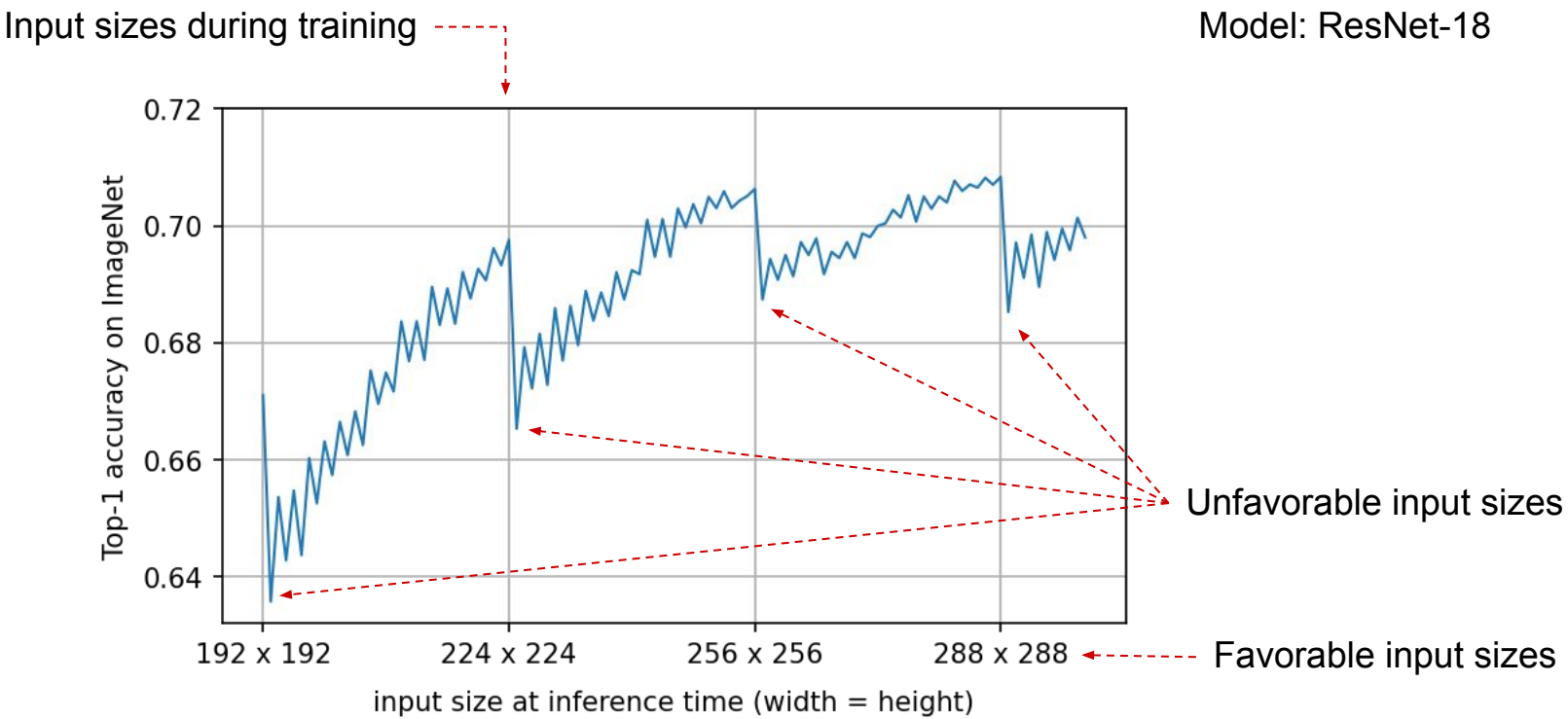
  Fixed output size.



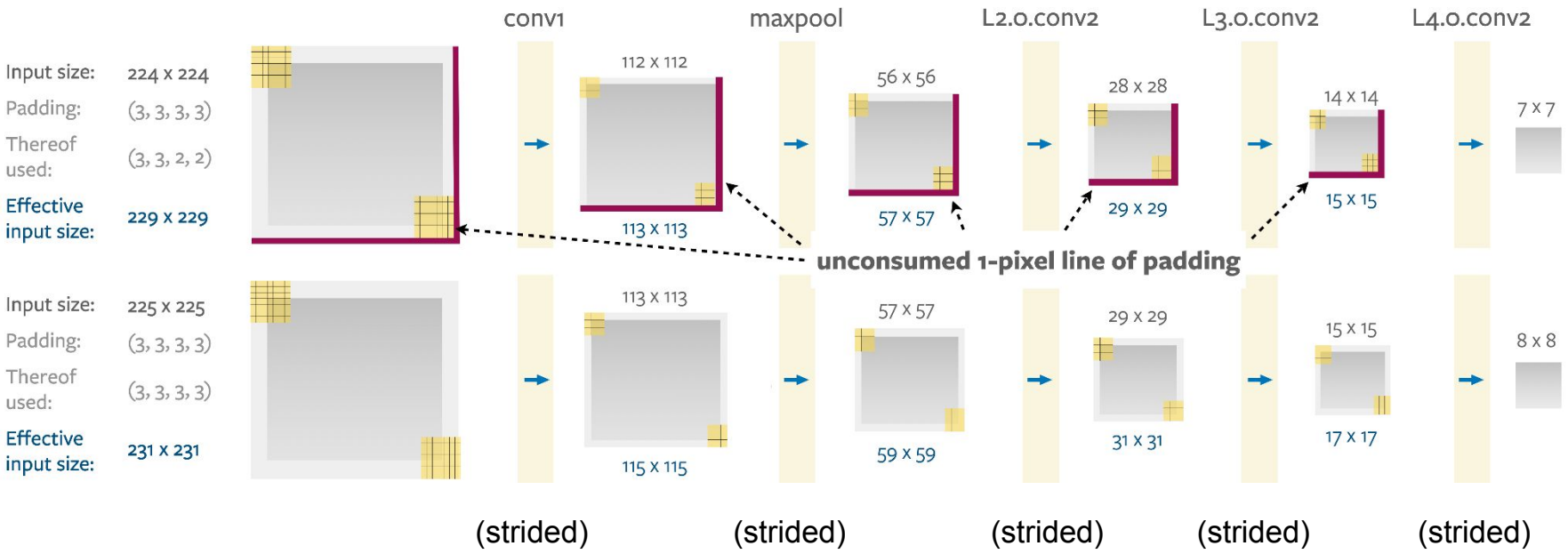
Question

Are CNNs sensitive to small changes in the input size?

Standard CNNs Exhibit Periodic Sensitivity

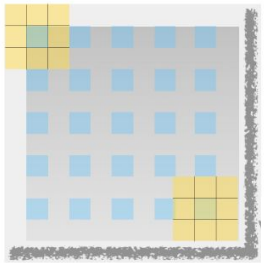


The Culprit: Pooling Layers Can Leave Input Unconsumed



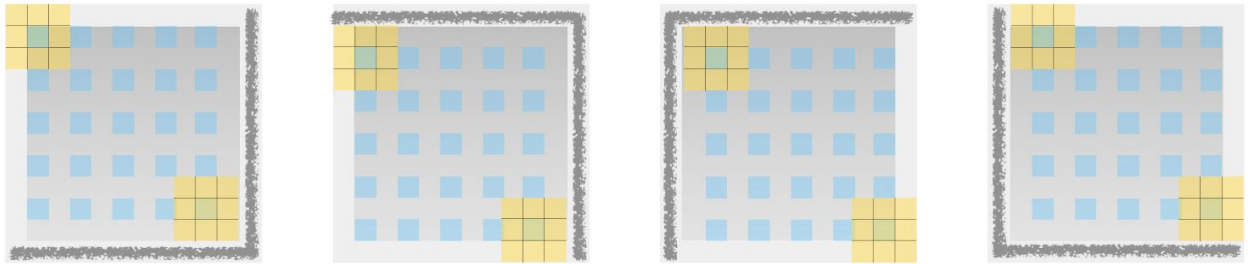
Solution: Spatially-Balanced Pooling (SBPool)

Standard Pooling

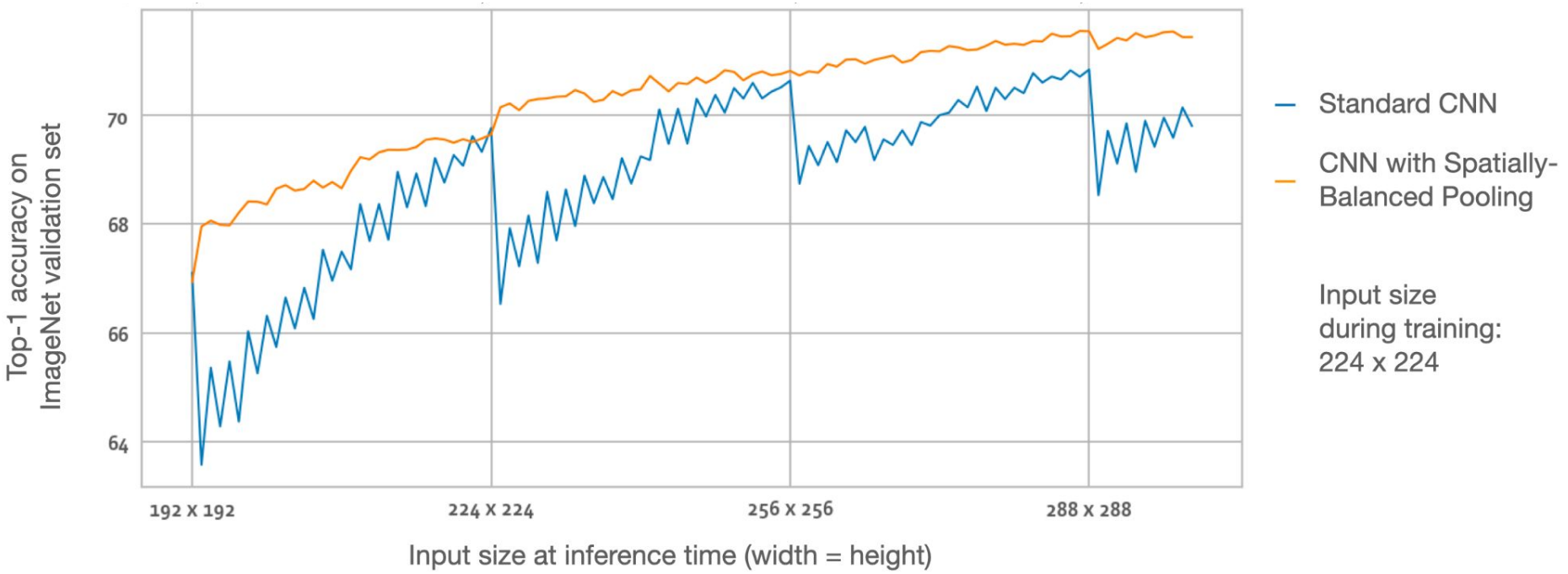


unconsumed part of the input

SBPool: Randomly select from the following variants during training



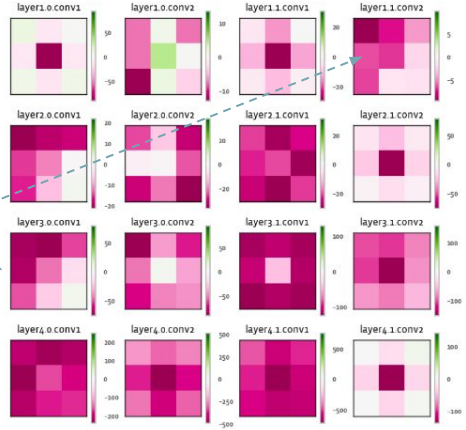
Results: ImageNet Classification with ResNet-18



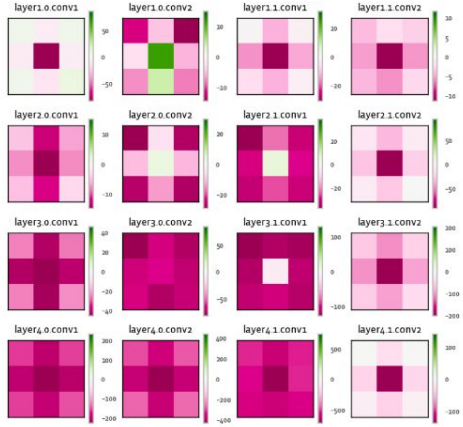
SBPool Mitigates Potential Skewness in Learned Filters

Mean weight kernel per layer

Skewness



ResNet-18 trained under standard downsampling



ResNet-18 trained under SBPool

Skewness mitigated

Takeaways

CNNs can process inputs of arbitrary size, leveraging their full resolution.

However, standard pooling arithmetic in CNNs can lead to:

- Overfitting boundary conditions dictated by the training input size.
- Skewing the learned weights.

SBPool mitigates the overfitting and skewness:

- This improves robustness to changes in input size and to translational shifts.
- This can improve the model accuracy even when fixing the input size.