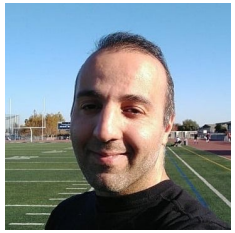
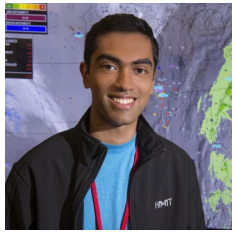
Mind the Pad – CNNs can Develop Blind Spots



Bilal Alsallakh

Narine Kokhlikyan



Vivek Miglani

Jun Yuan

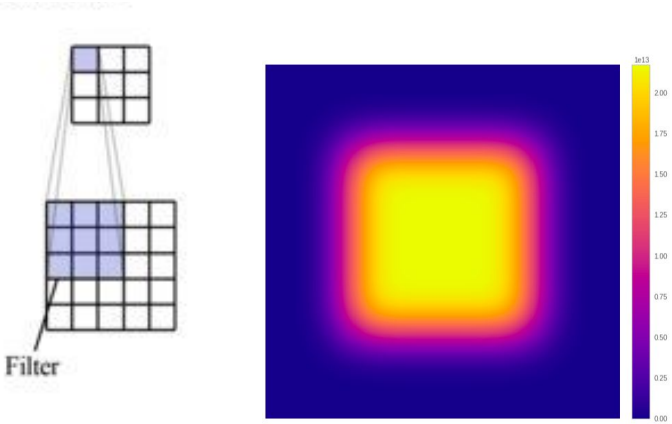
Orion Reblitz-Richardson

FACEBOOK

NYU

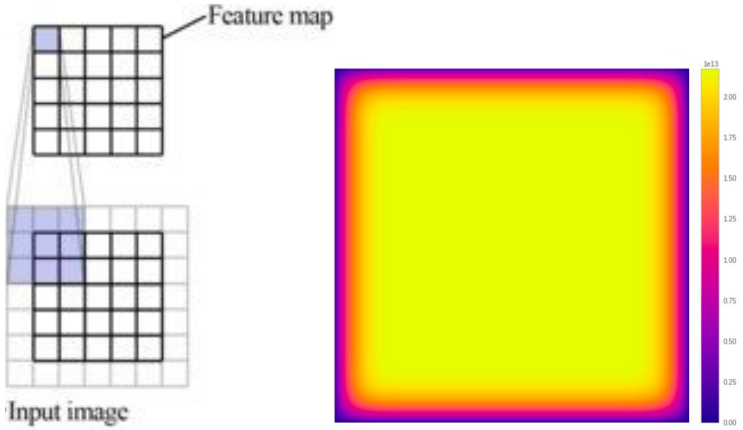
Why Padding?

No Padding



pixel space

SAME 0-Padding



pixel space

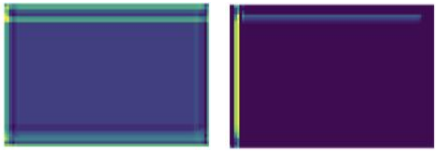
Color represents the number of times an input pixel is utilized by VGG-19

Problems with Zero Padding

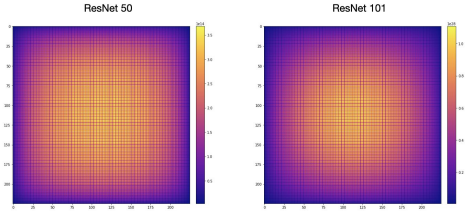
0	0	0	0	0	0
0	35	19	25	6	0
0	13	22	16	53	0
0	4	3	7	10	0
0	9	8	1	3	0
0	0	0	0	0	0

It can **skew the learnt filters** if applied unevenly.

It can **induce line artifacts** in feature maps.



It can **incur mild foveation** in the pixel space.

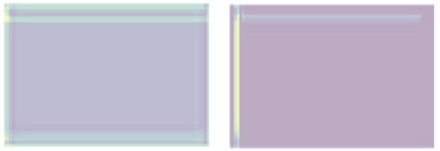


Problems with Zero Padding

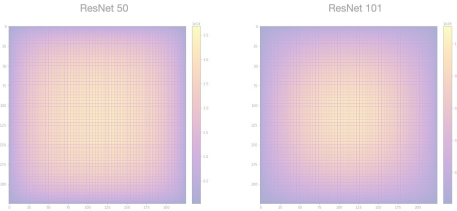
0	0	0	0	0	0
0	35	19	25	6	0
0	13	22	16	53	0
0	4	3	7	10	0
0	9	8	1	3	0
0	0	0	0	0	0

It can **skew the learnt filters** if applied unevenly.

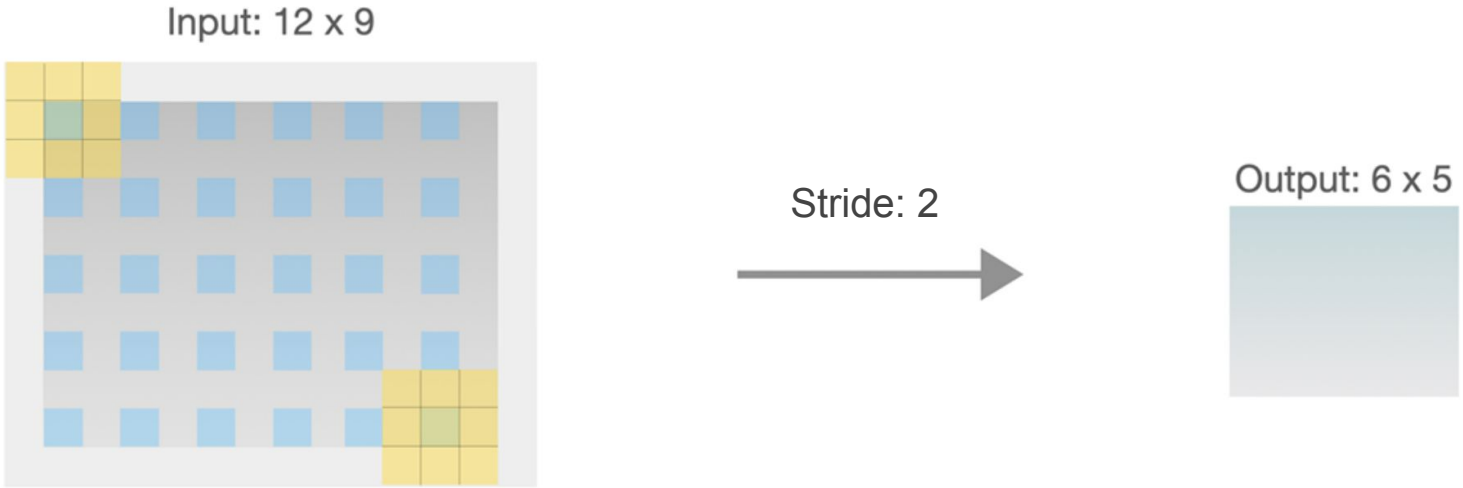
It can **induce line artifacts** in feature maps.



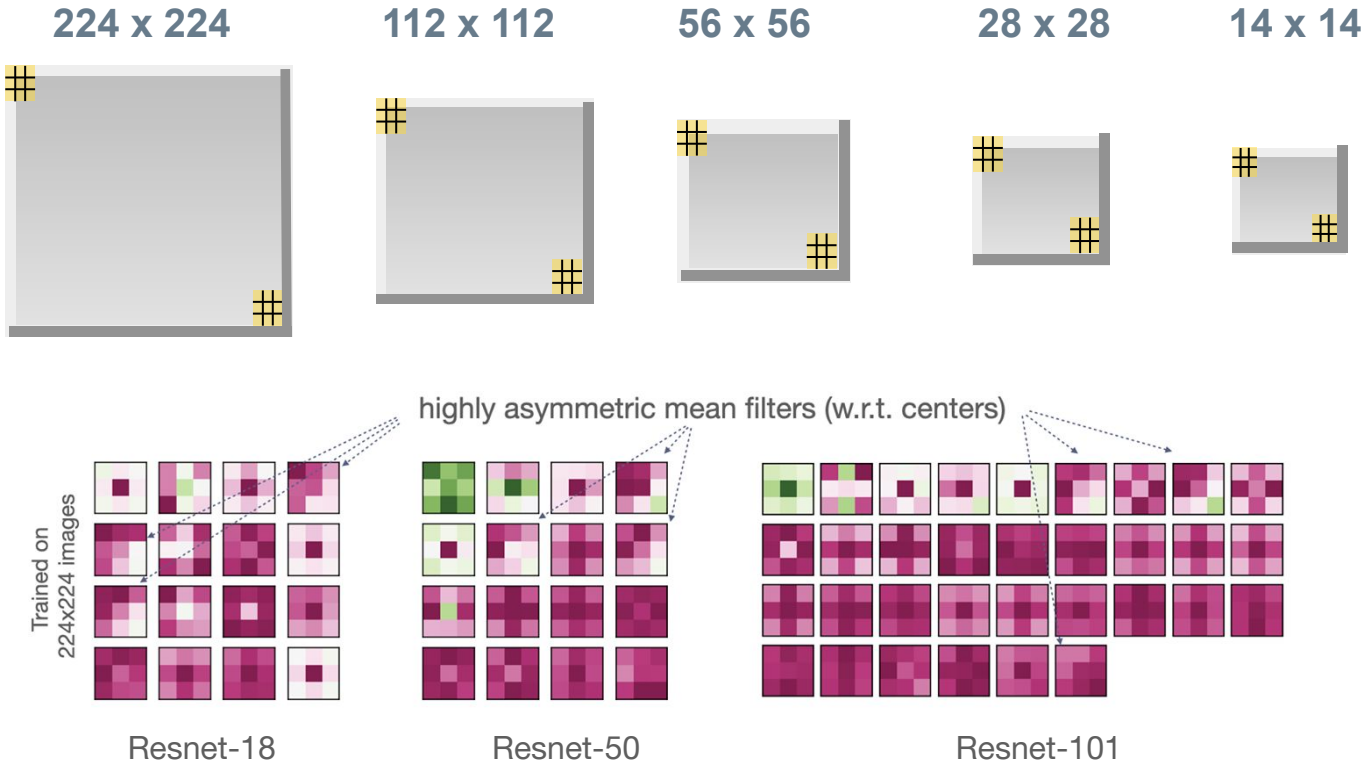
It can **incur mild foveation** in the pixel space.



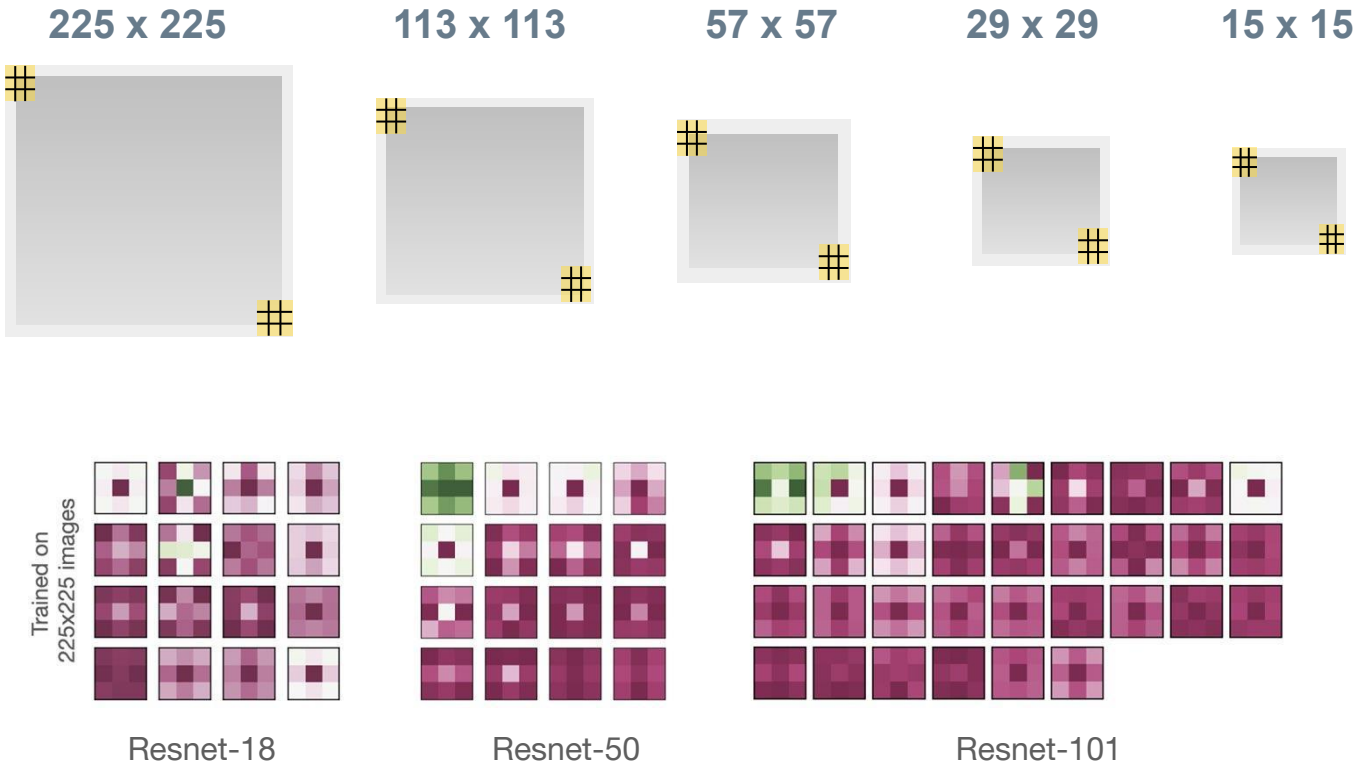
Strided convolution might consume 0-padding unevenly:



Uneven padding ⟹ skewed filters:



Even padding ⟹ symmetric mean filters:



Symmetric filters ⟹ higher accuracy (and shift invariance)

Top-1 (and Top-5) Accuracy on ImageNet

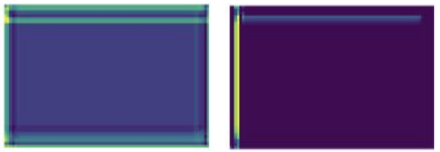
Input Size	MobileNet	ResNet-18	ResNet-34	ResNet-50	ResNet-101
224×224	68.19 (88.44)	69.93 (89.22)	73.30 (91.42)	75.65 (92.47)	77.37 (93.56)
225×225	68.80 (88.78)	70.27 (89.52)	73.72 (91.58)	76.01 (92.90)	77.67 (93.81)

Problems with Zero Padding

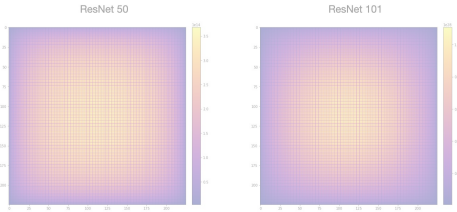
0	0	0	0	0	0
0	35	19	25	6	0
0	13	22	16	53	0
0	4	3	7	10	0
0	9	8	1	3	0
0	0	0	0	0	0

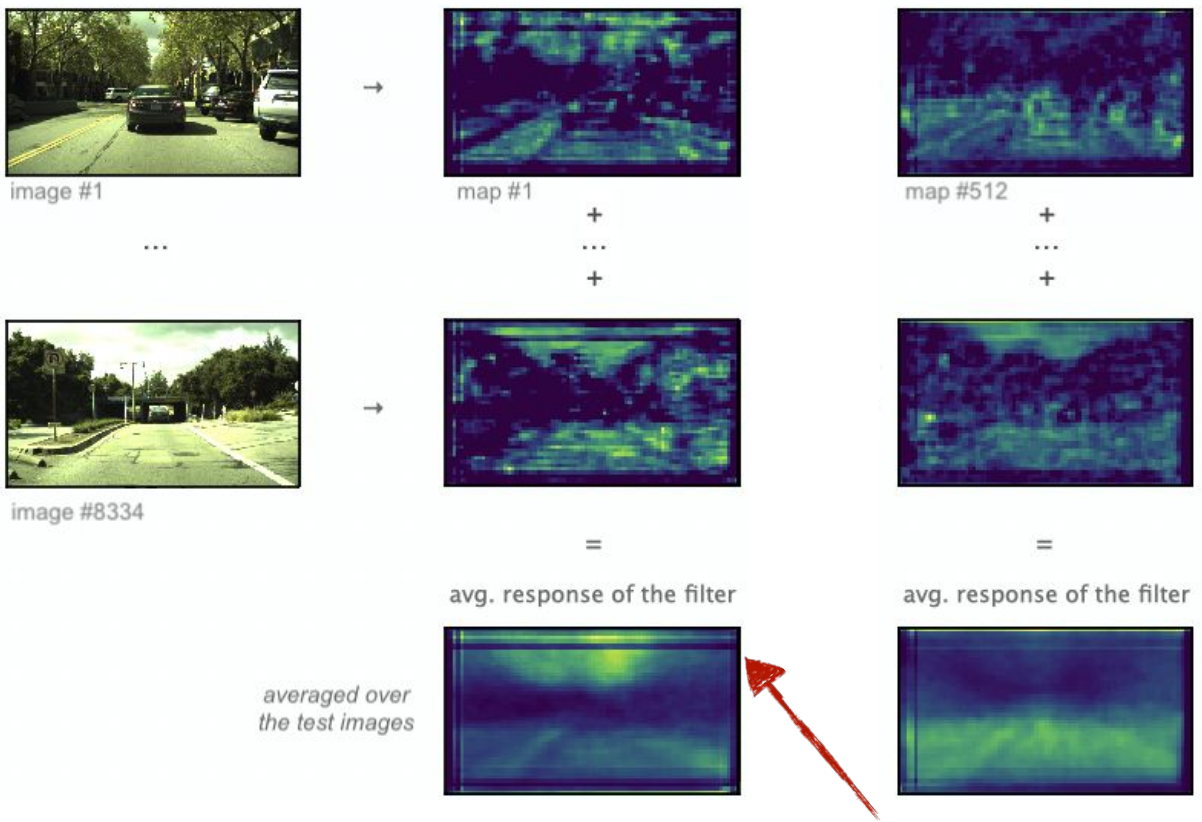
It can **skew the learnt filters** if applied unevenly.

It can **induce line artifacts** in feature maps.



It can **incur mild foveation** in the pixel space.

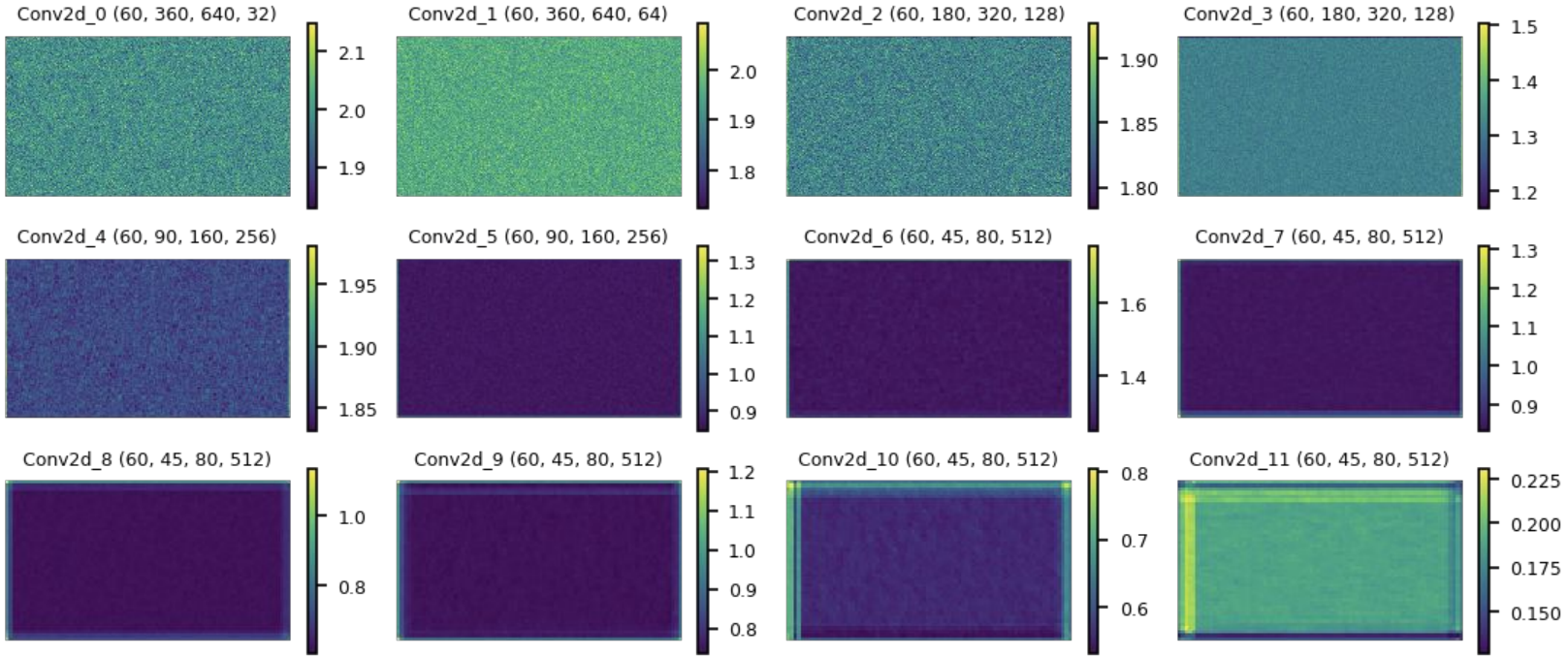




Artificially suppressed activation at certain locations

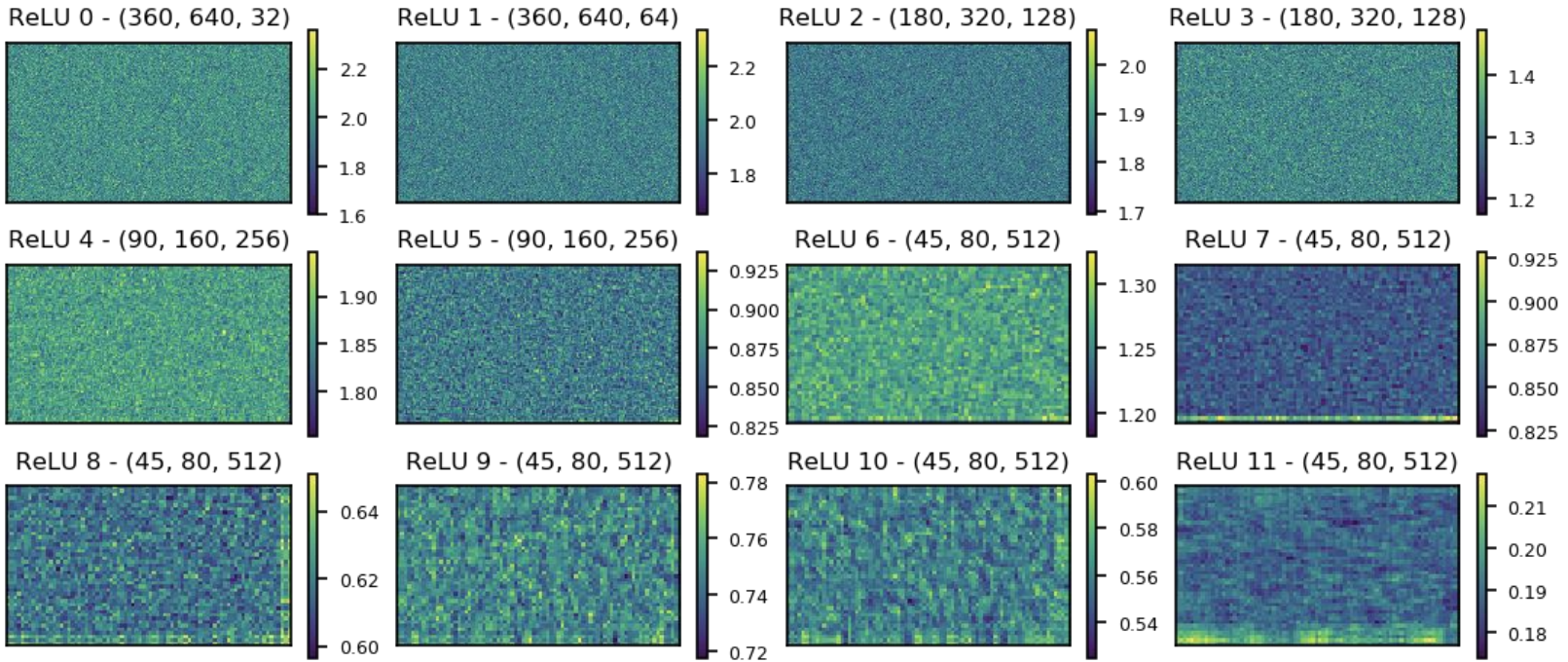
Feature maps under 0-padding

Mean filter response per layer (with random inputs)

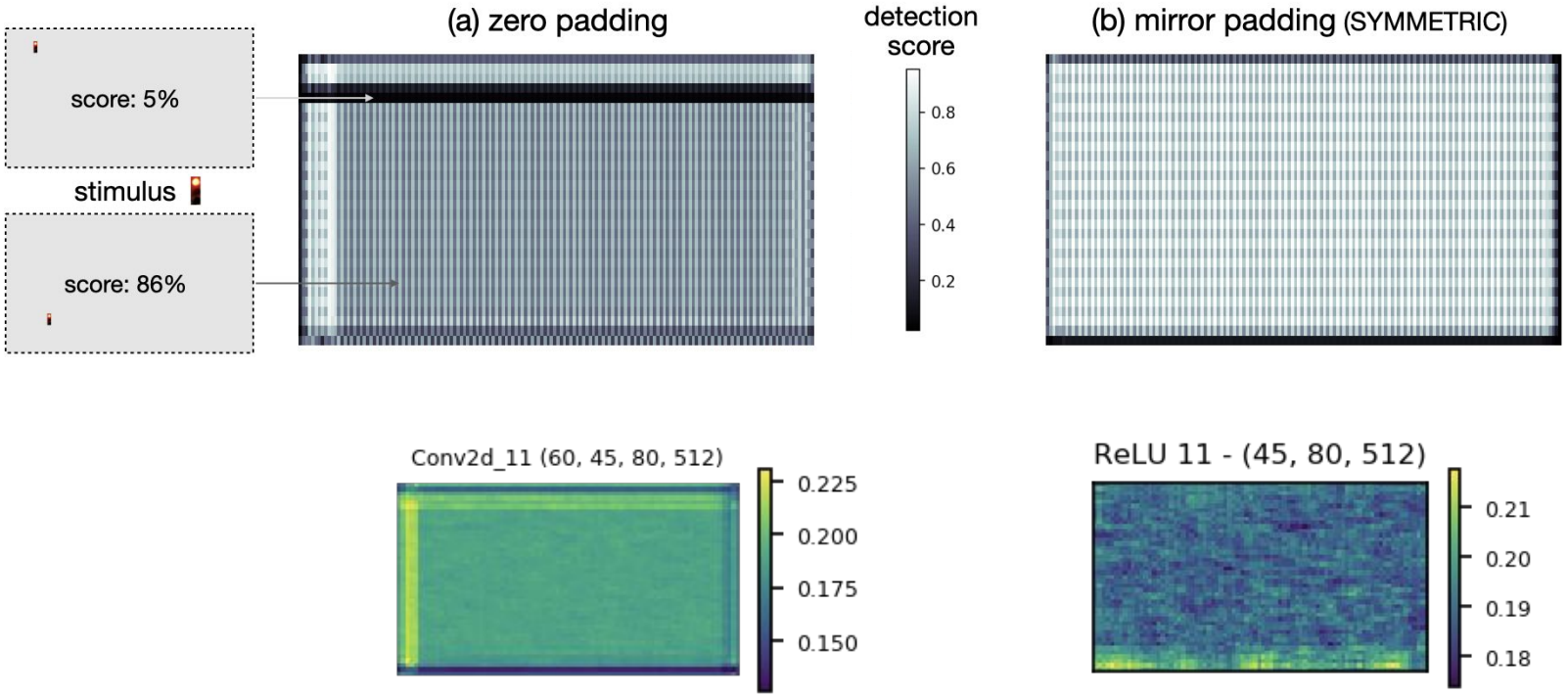


Feature maps under mirror padding

Mean filter response per layer (with random inputs)



Stimulus Detectability at Different Locations



Mitigated artifacts ⟹ higher accuracy (and shift invariance)

44% (shifted upwards)

7% (baseline)

82% (shifted downwards)

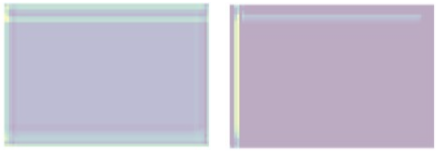
Average Precision (AP)	AP@.20IOU	AP@.50IOU	AP@.75IOU
Zero Padding	80.24%	49.58%	3.7%
Mirror Padding	83.20%	57%	8.44%

Problems with Zero Padding

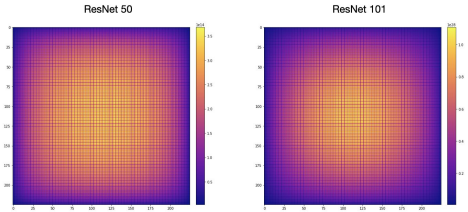
0	0	0	0	0	0
0	35	19	25	6	0
0	13	22	16	53	0
0	4	3	7	10	0
0	9	8	1	3	0
0	0	0	0	0	0

It can **skew the learnt filters** if applied unevenly.

It can **induce line artifacts** in feature maps.

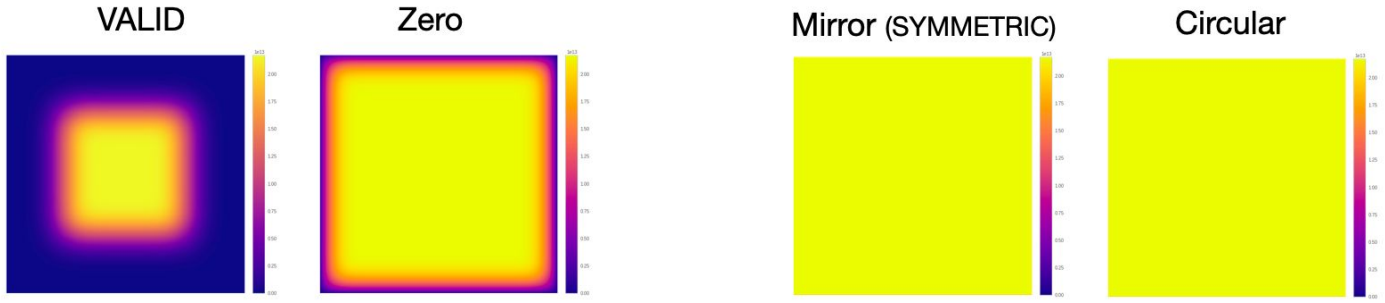


It can **incur mild foveation** in the pixel space.

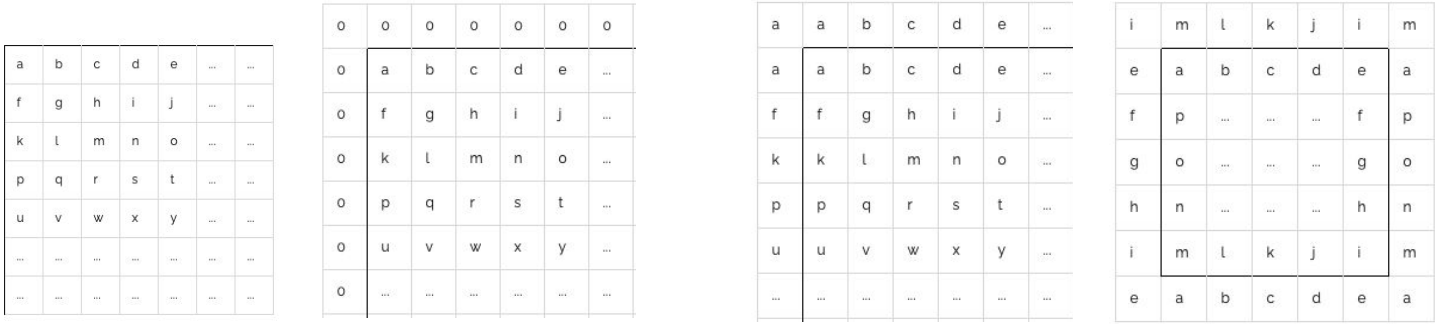


SAME 0-padding mildly marginalizes the periphery:

Input Space (512x512) of VGG-19



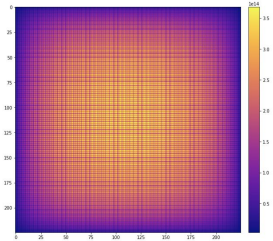
Illustrations on small inputs and one layer



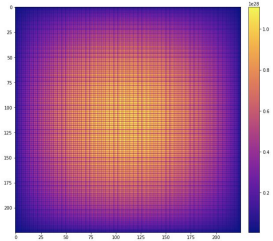
SAME 0-padding mildly marginalizes the periphery:

Deeper CNN => more “foveation”

ResNet-50

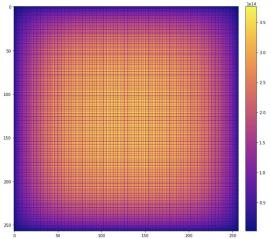


ResNet-101

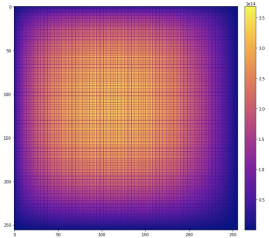


Uneven padding => skewed effects

225x225 input



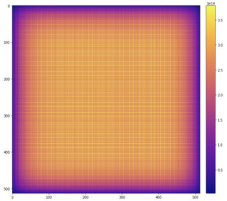
224x224 input



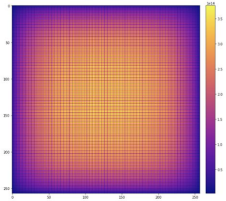
Boundary marginalization is more extensive in smaller inputs (relative to input size):

computed for ResNet-50

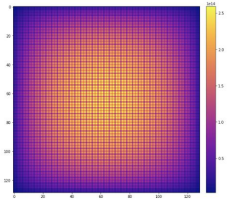
513 x 513



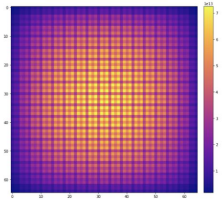
257 x 257



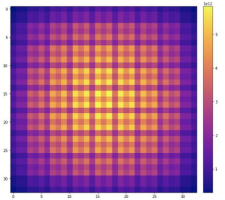
129 x 129



65 x 65



33 x 33



Summary

Zero padding can:

- **Skew the learned filters** during downsampling.
  - Consider eliminating uneven padding.
  - Consider rigid kernels (maxpool, avgpool).
- **Induce line artifacts** in the feature maps**.**
- **Marginalize the periphery** in the pixel space**.**

mind-the-pad.github.io

Circular or symmetric mirror padding:

- Mitigate these issues.

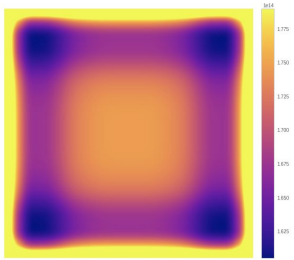
  Might fit your task better - worth trying.

Backup Slides

Other Padding Schemes Can Also Incur “Foveation” Effect

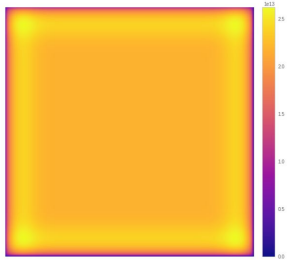
Under Different Padding Schemes

Replicate (5x5 kernel)



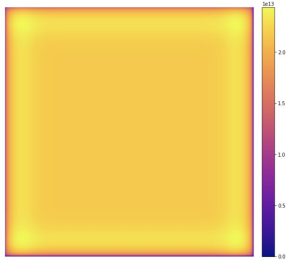
a	a	a	b	c	d	e
a	a	a	b	c	d	e
a	a	a	b	c	d	e
f	f	f	g	h	i	j
k	k	k	l	m	n	o
p	p	p	q	r	s	t
u	u	u	v	w	x	y

Mirror (REFLECT)



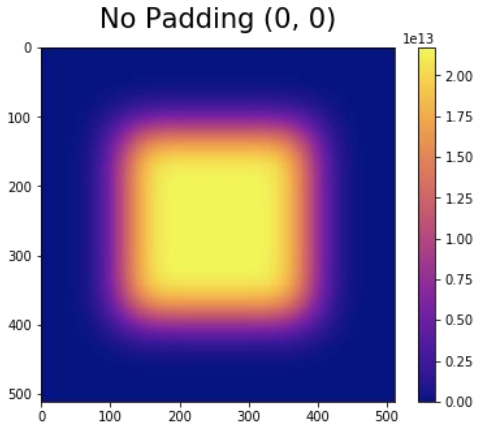
g	f	g	h	i	j
b	a	b	c	d	e
g	f	g	h	i	j
l	k	l	m	n	o
q	p	q	r	s	t
v	u	v	w	x	y

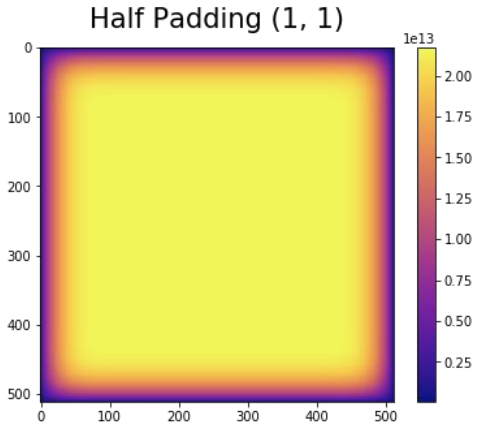
PartialConv

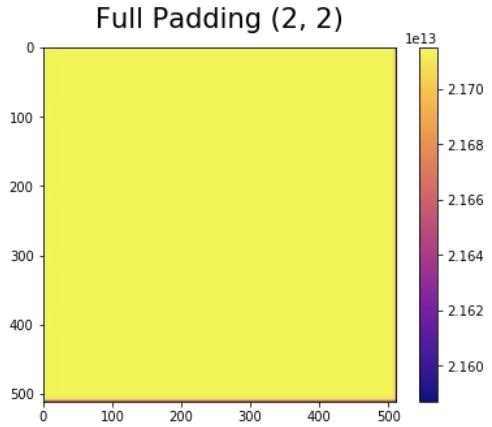


	a	b	c	d	e
	f	g	h	i	j
	k	l	m	n	o
	p	q	r	s	t
	u	v	w	x	y

Full Padding

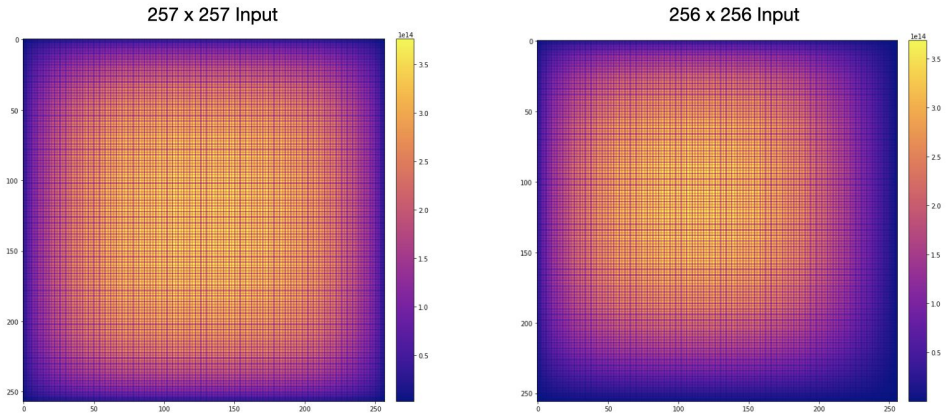






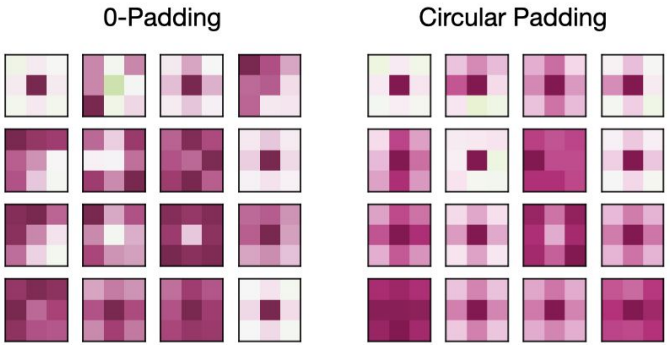
Uneven Application of Padding

Color represents the number of time an input pixel is utilized in ResNet-50

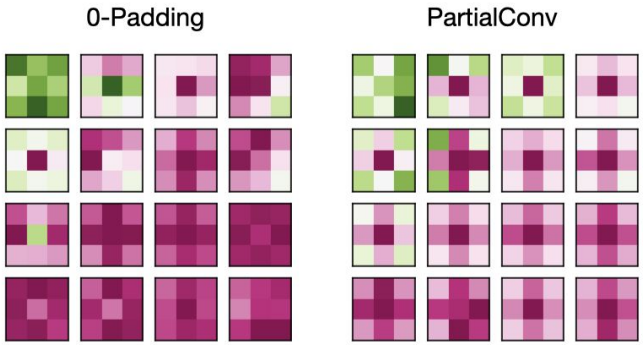


Padding Mode and Symmetry of Average Filter

Input size is 224x224, leading to uneven padding at every downsampling layer



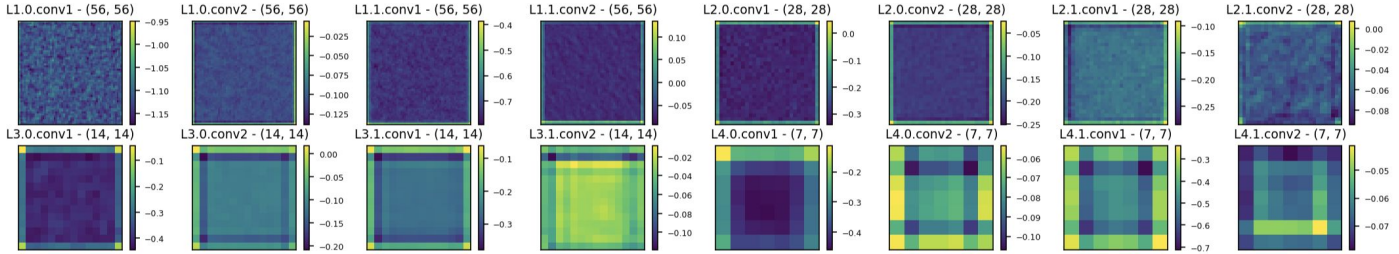
ResNet-18



ResNet-50

Feature-Maps in ResNet-18

0-Padding



Mirror (SYMMETRIC)