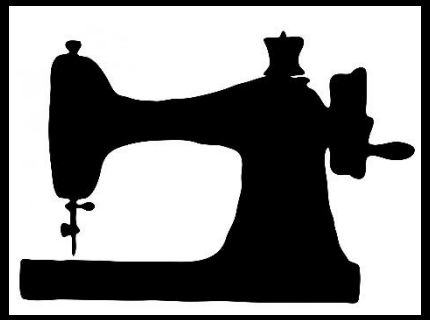
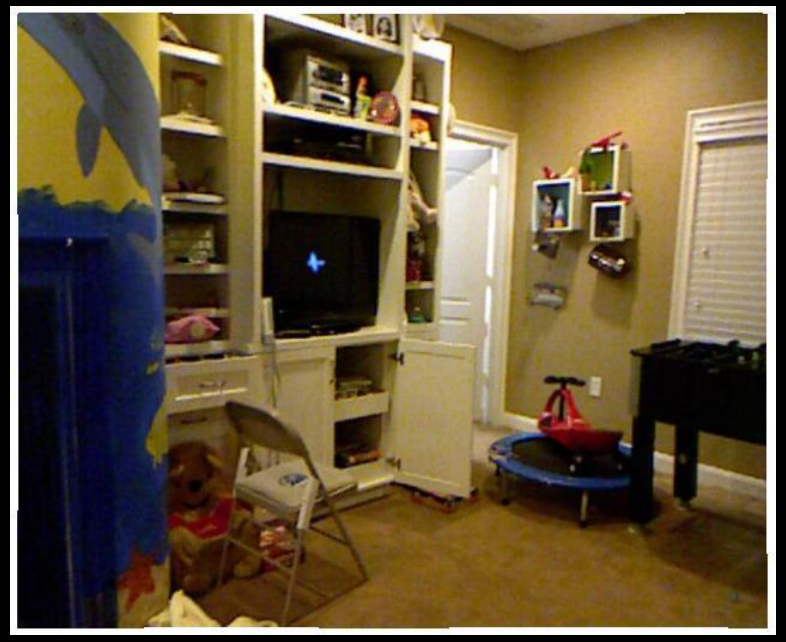


Cross-stitch Networks for Multi-task Learning

Ishan Misra*, Abhinav Shrivastava*, Abhinav Gupta, Martial Hebert

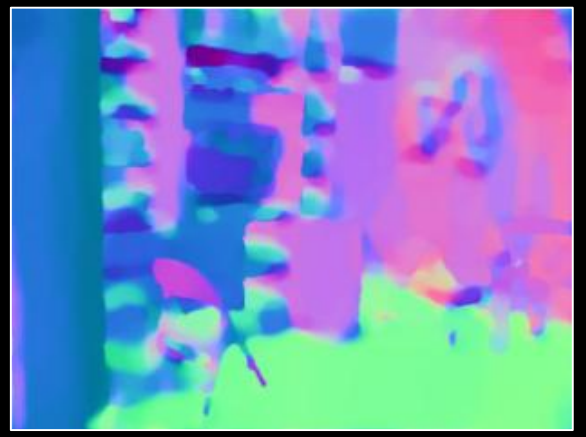
Multiple supervision

Input



Labels

Surface Normals



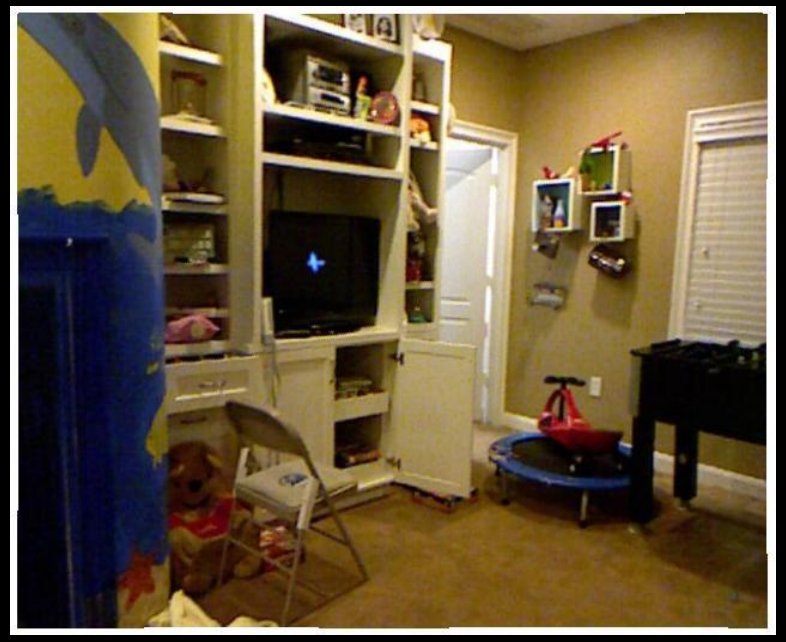
Surface orientation

Semantic Segmentation

Semantic labels

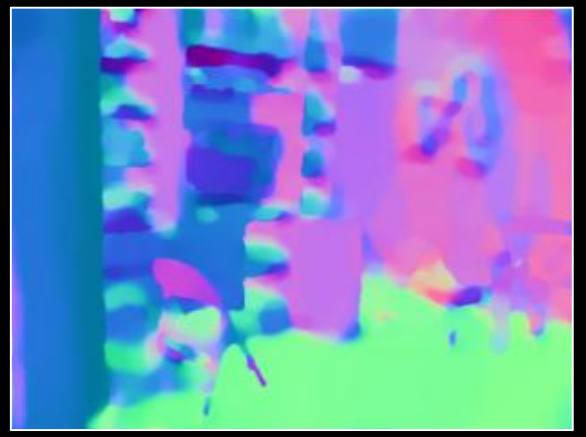
“Free” supervision

Input



Labels

Surface Normals



Surface orientation

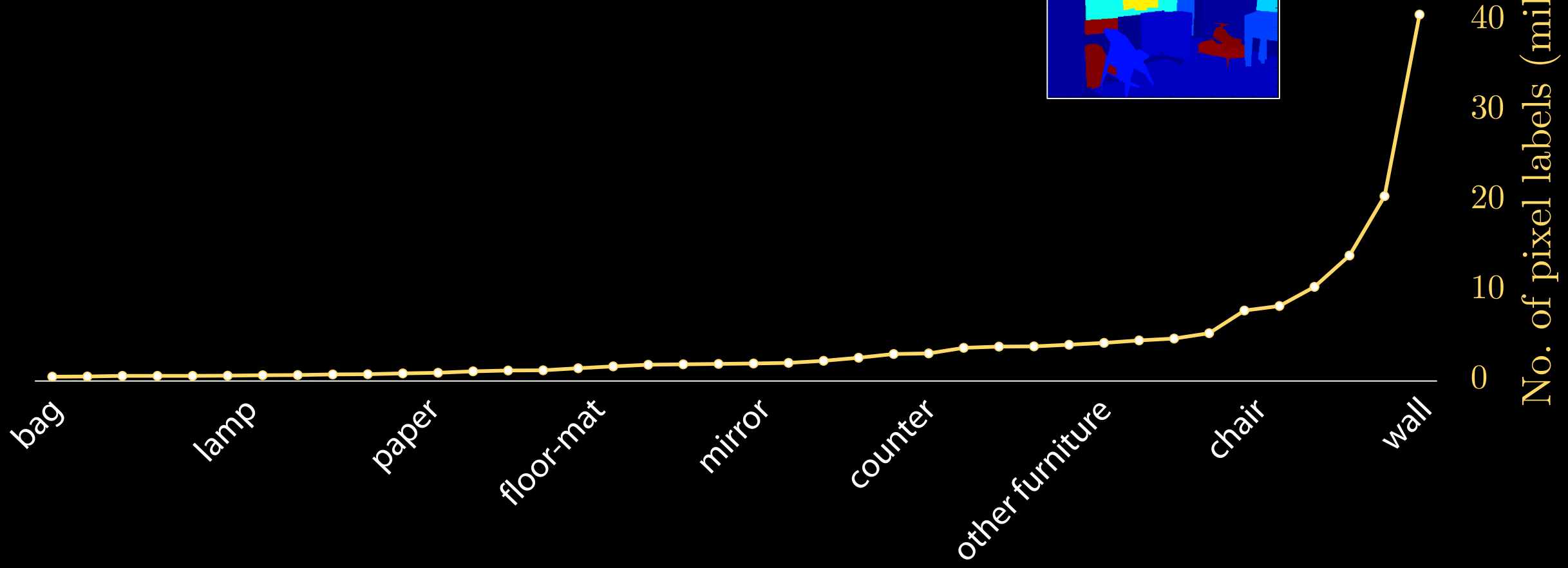
Using Kinect:
Free

Semantic Segmentation

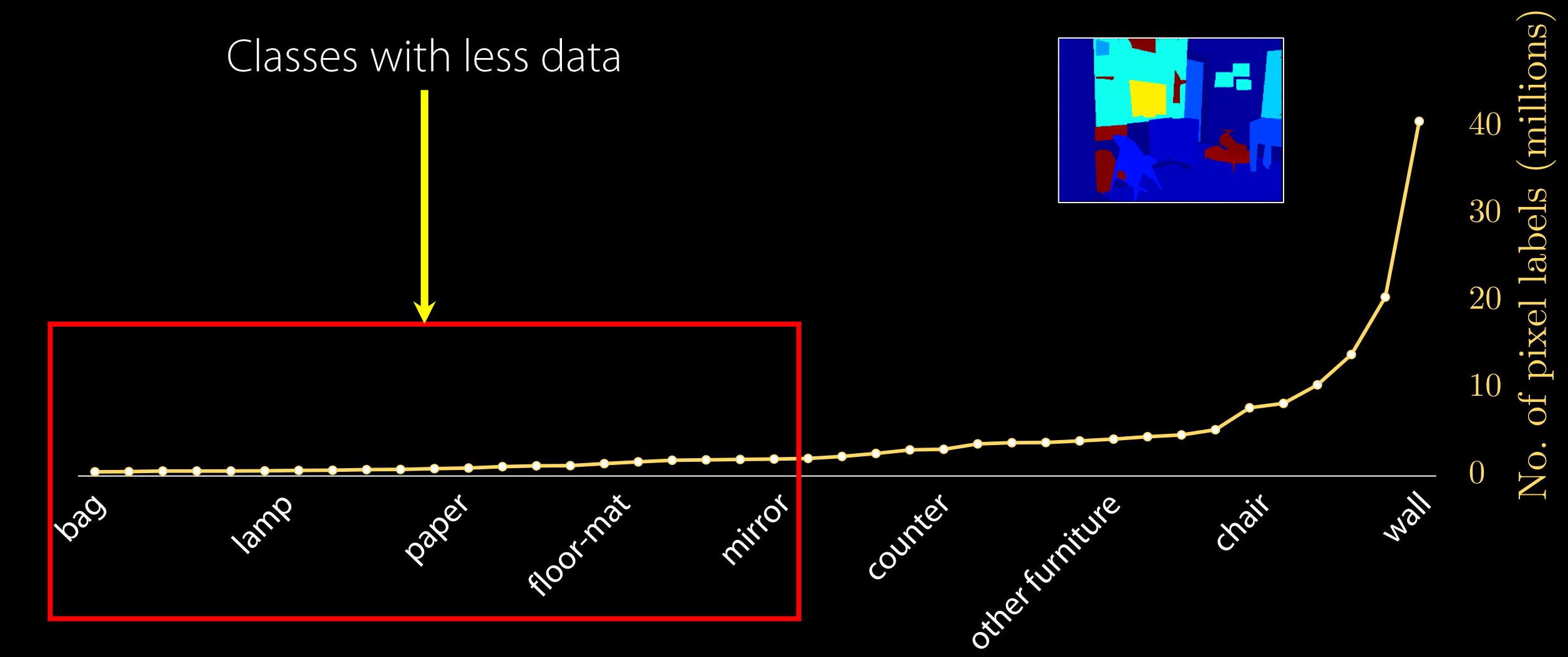
Semantic labels

Manually marked:
Expensive

No. of pixels per segmentation class

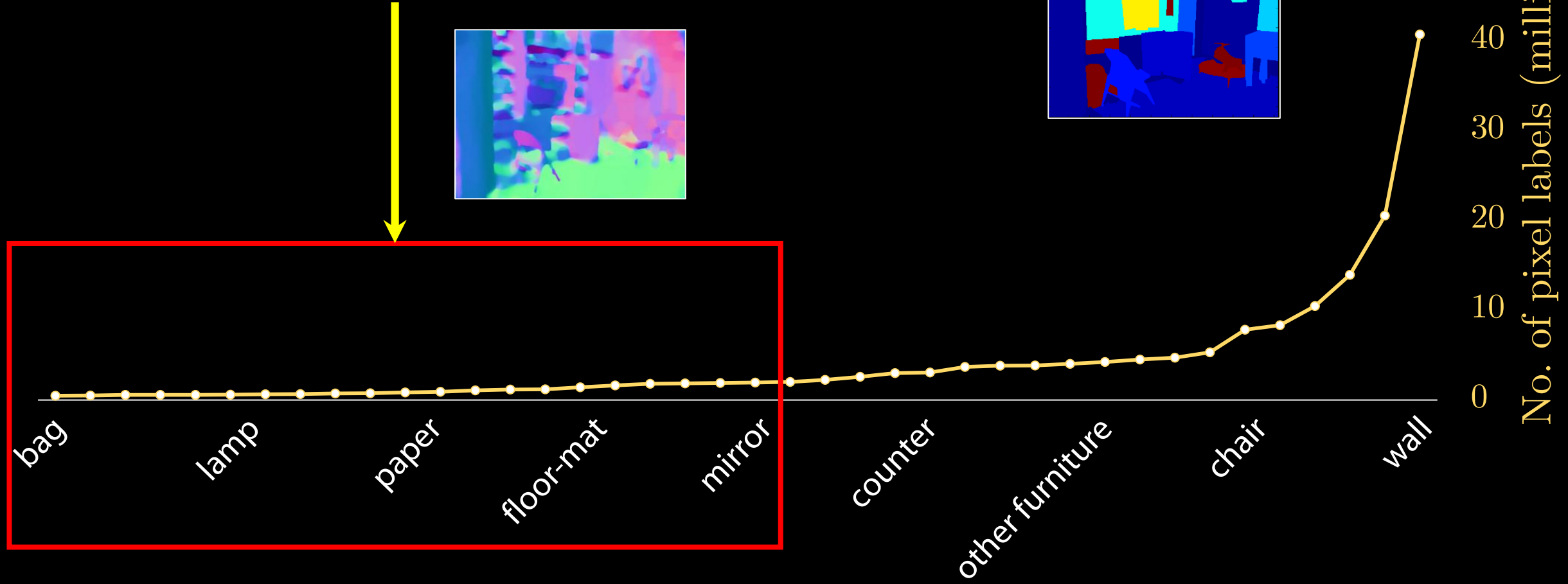


No. of pixels per segmentation class



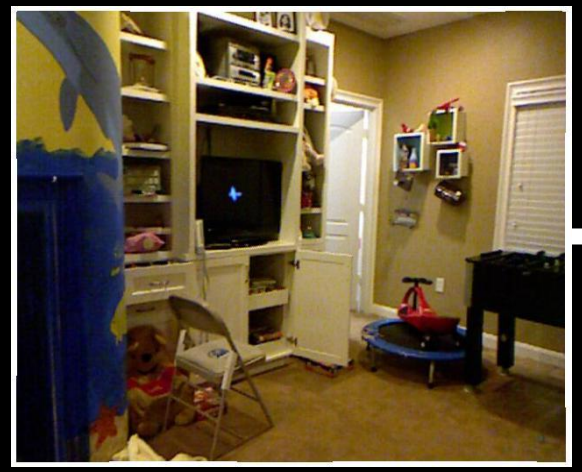
No. of pixels per segmentation class

Classes with less data

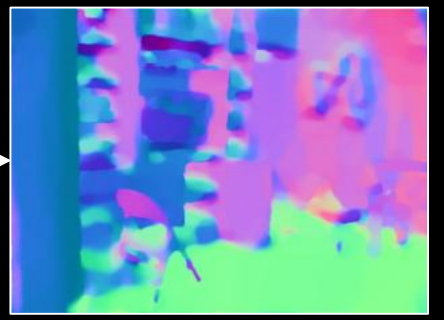


Multi-task: Shared network

Input



Task A



Task B

Shared Layers

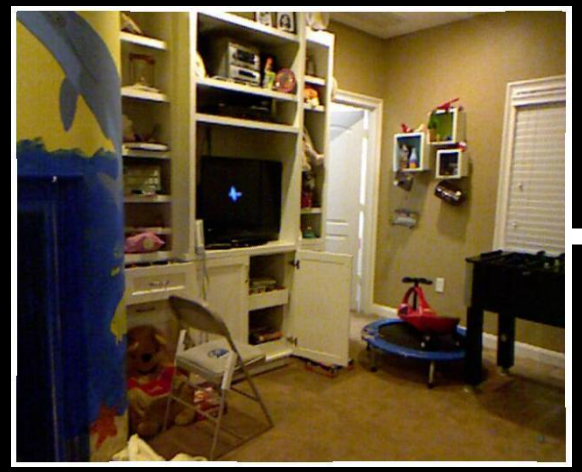
Task A layers

Task B layers

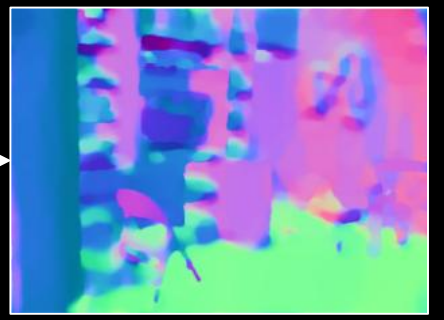
Complete sharing of parameters between tasks

Split at a lower layer

Input



Task A



Task B

Split here?

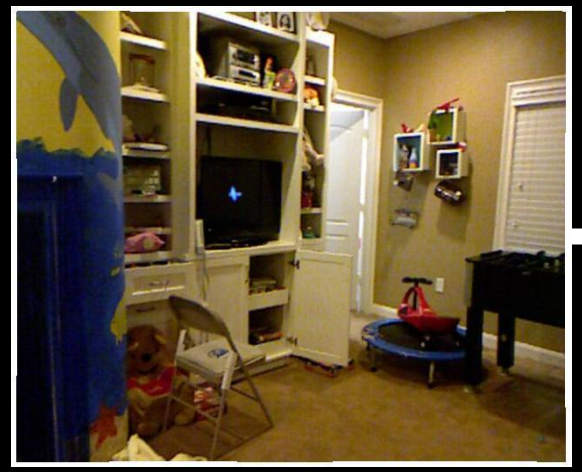
Shared Layers

Task A layers

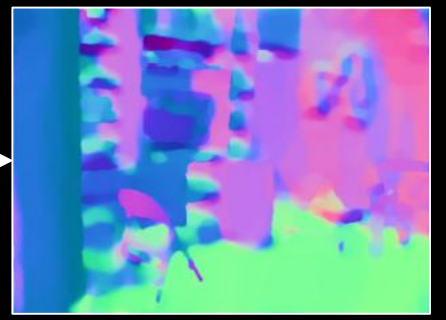
Task B layers

Split at a lower layer

Input



Task A



Task B

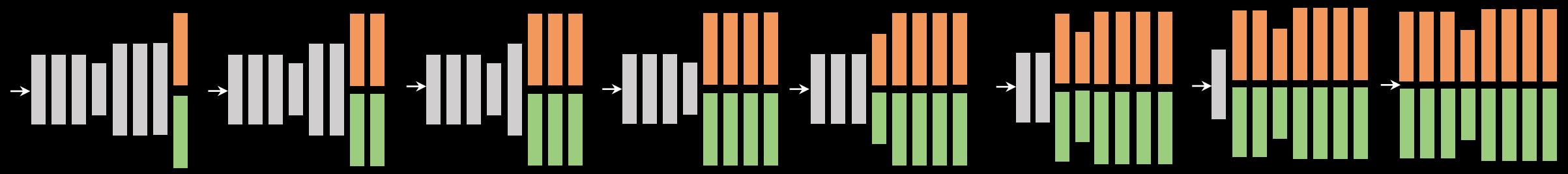
Shared Layers

Task A layers

Task B layers

Spectrum of sharing: Standard Multi-task

9 8 7 6 5 4 3 2,1

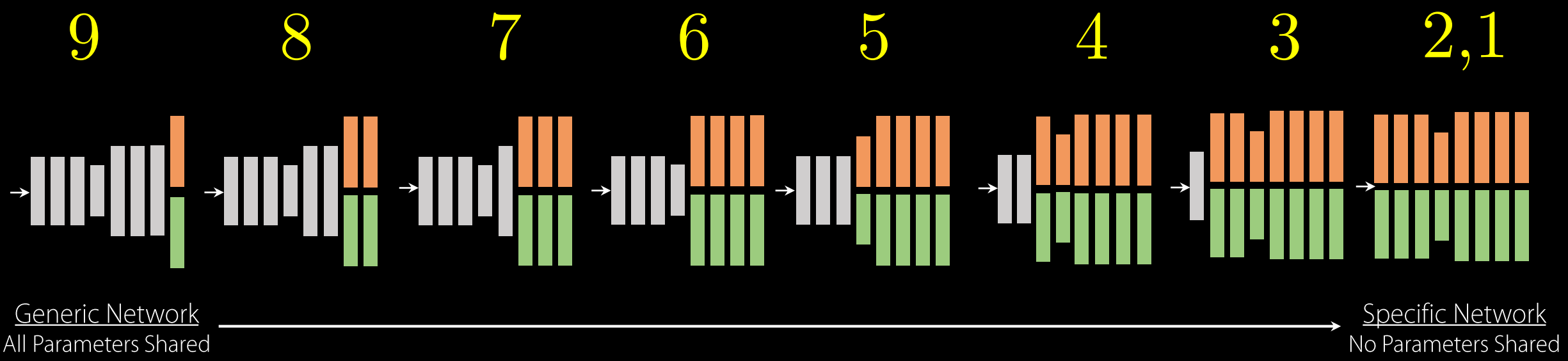


Generic Network
All Parameters Shared

Specific Network
No Parameters Shared

Shared Layers
Task A layers
Task B layers

Standard Multi-task Learning



Train and Test all of them

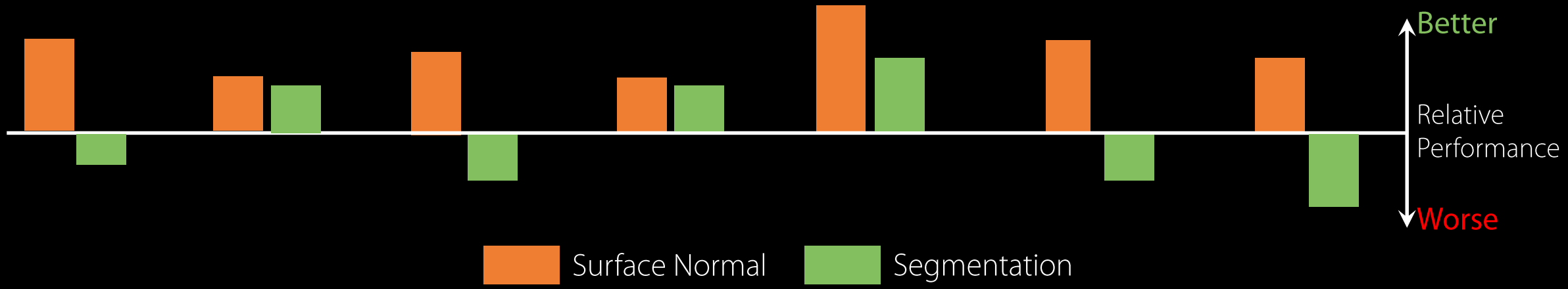
Standard Multi-task Learning

9 8 7 6 5 4 3 2,1

Generic Network
All Parameters Shared

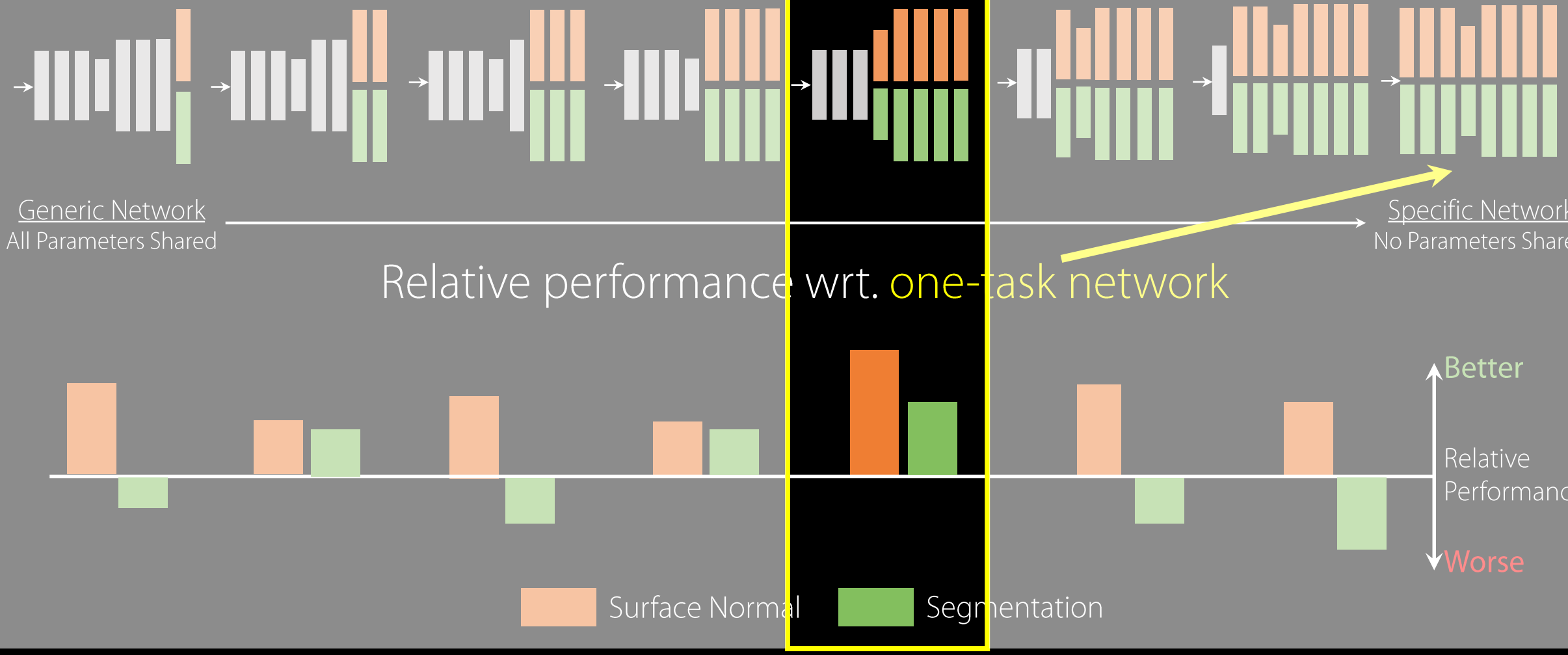
Specific Network
No Parameters Shared

Relative performance wrt. one-task network



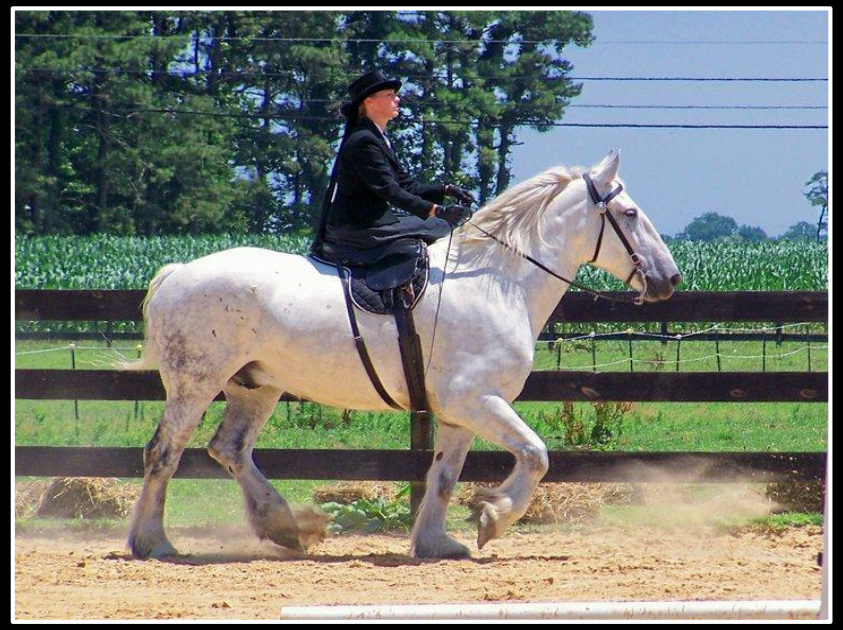
Pick the best

9 8 7 6 5 4 3 2,1



New Tasks?

Input



Labels

Attributes

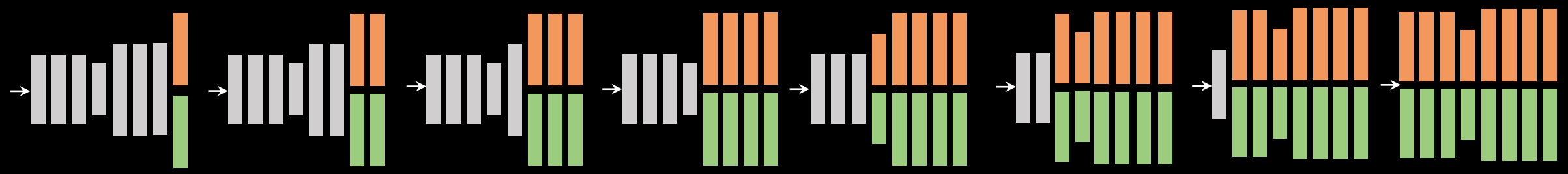
Has saddle Four legs

Detection

Object location

Training and measuring performance

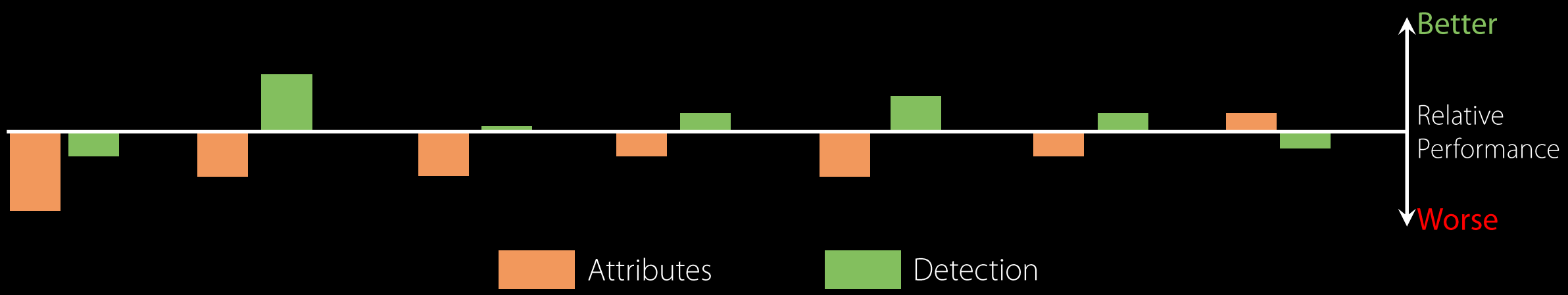
9 8 7 6 5 4 3 2,1



Generic Network
All Parameters Shared

Specific Network
No Parameters Shared

Relative performance wrt. one-task network

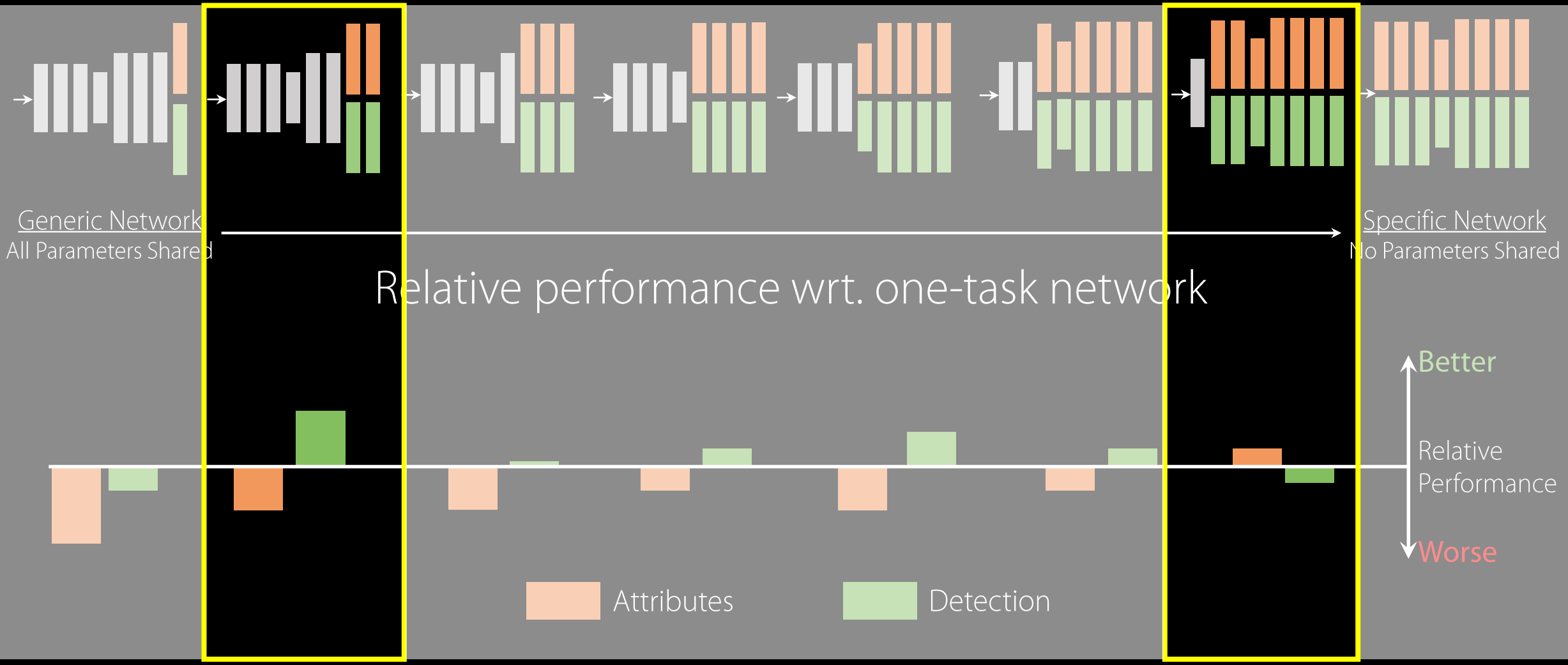


No single best network

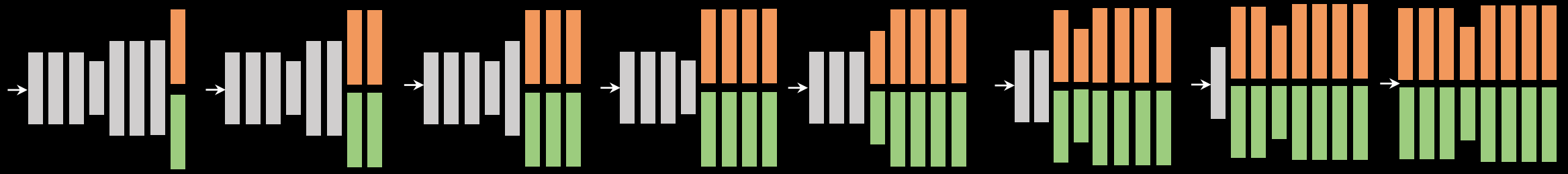
9 8 7 6 5 4 3 2,1

Best for Detection

Best for Attributes



Problems with standard approach



Generic Network
All Parameters Shared

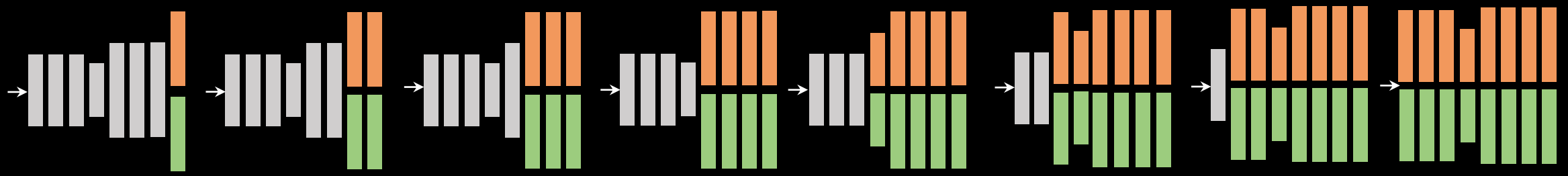
Specific Network
No Parameters Shared

Problems

Given a new set of tasks:

- Enumerate many ConvNet architectures

No principled way of exploring



Generic Network
All Parameters Shared

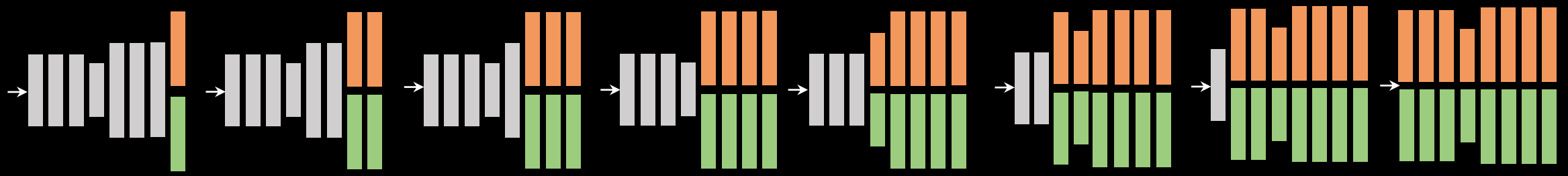
Specific Network
No Parameters Shared

Problems

Given a new set of tasks:

- Enumerate many ConvNet architectures
- Train all of them

Practically Expensive



Generic Network
All Parameters Shared

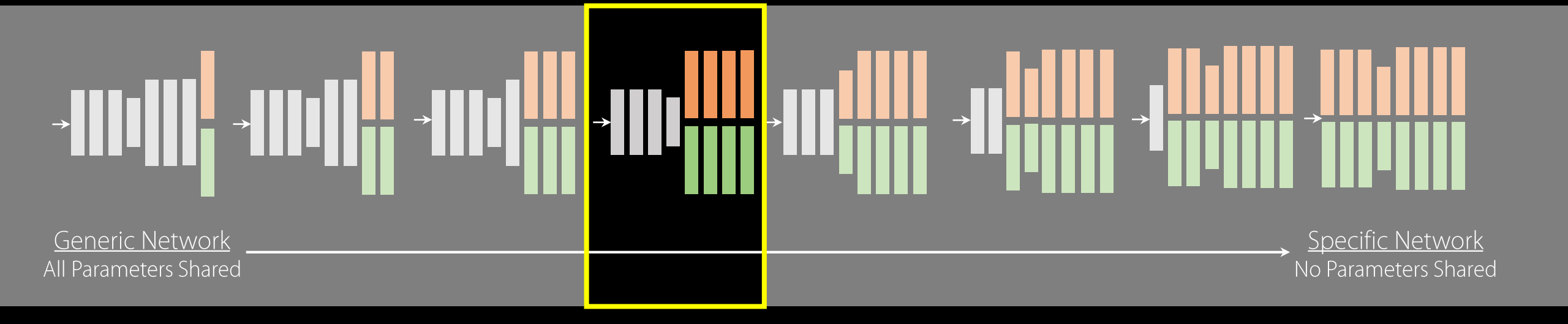
Specific Network
No Parameters Shared

Problems

Given a new set of tasks:

- Enumerate many ConvNet architectures
- Train all of them
- Pick the best one for your tasks

Best architecture depends on tasks

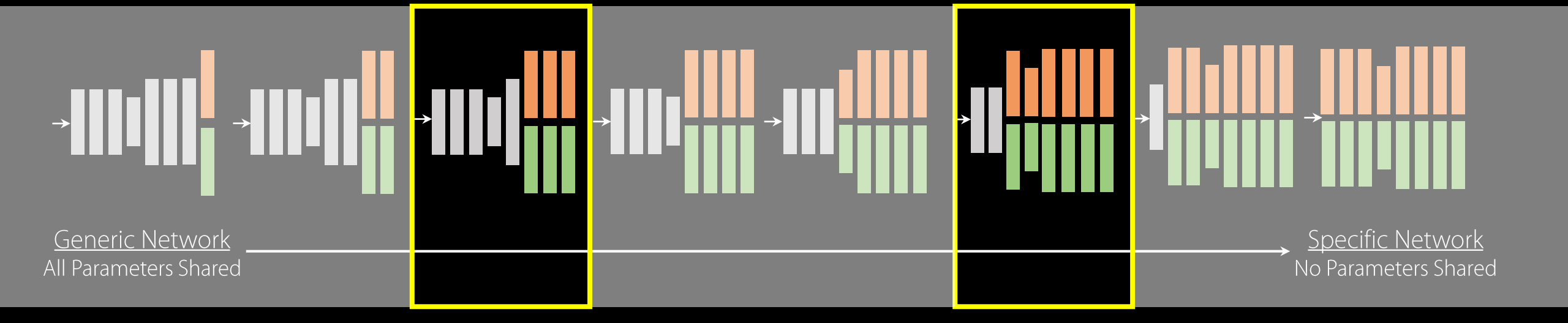


Problems

Given a new set of tasks:

- Enumerate many ConvNet architectures
- Train all of them
- Pick the best one for your tasks

One architecture may not perform well on all tasks



Our contribution

- Principled exploration of architectures
  - **Without** training all of them

Our contribution

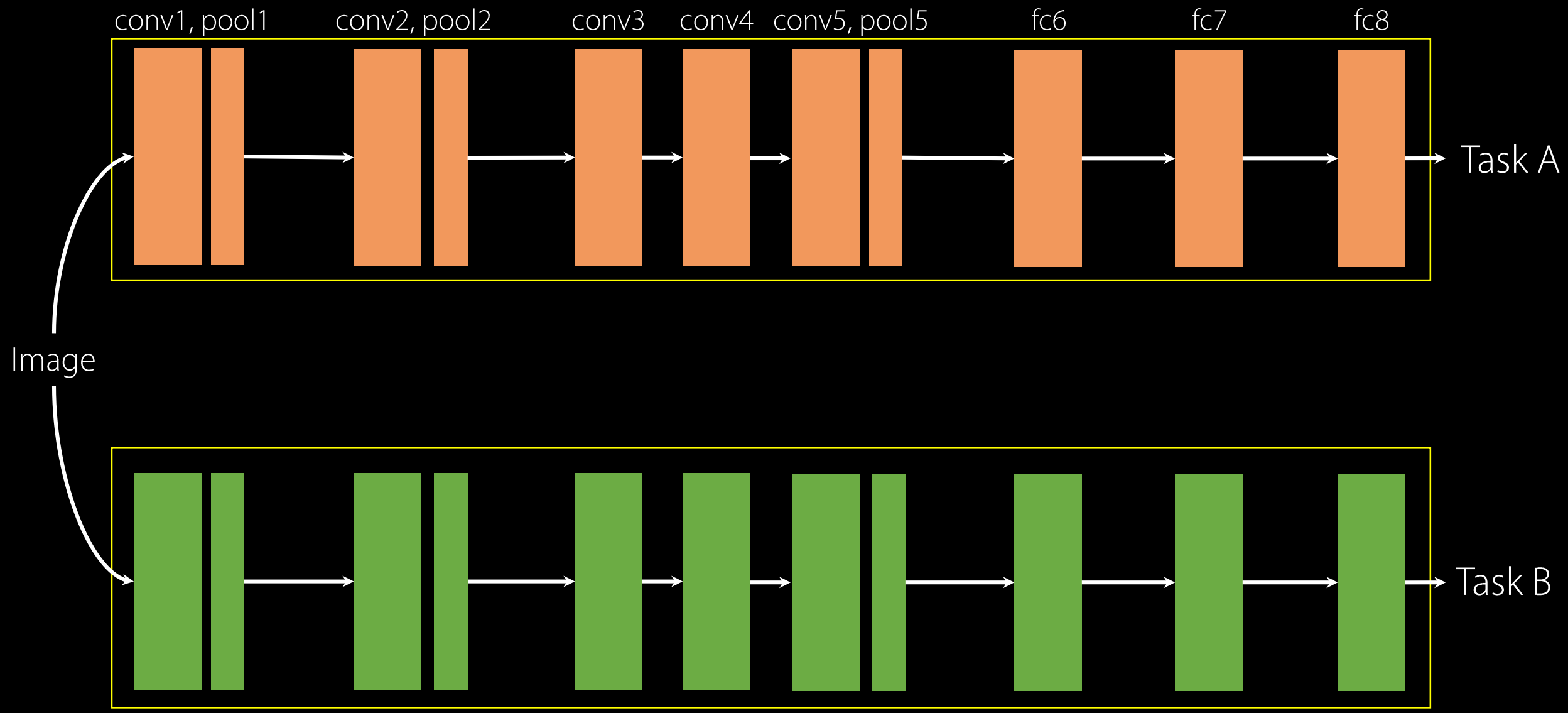
- Principled exploration of architectures
  - **Without** training all of them
- Finds network architecture automatically
  - Generalizes across tasks

Our contribution

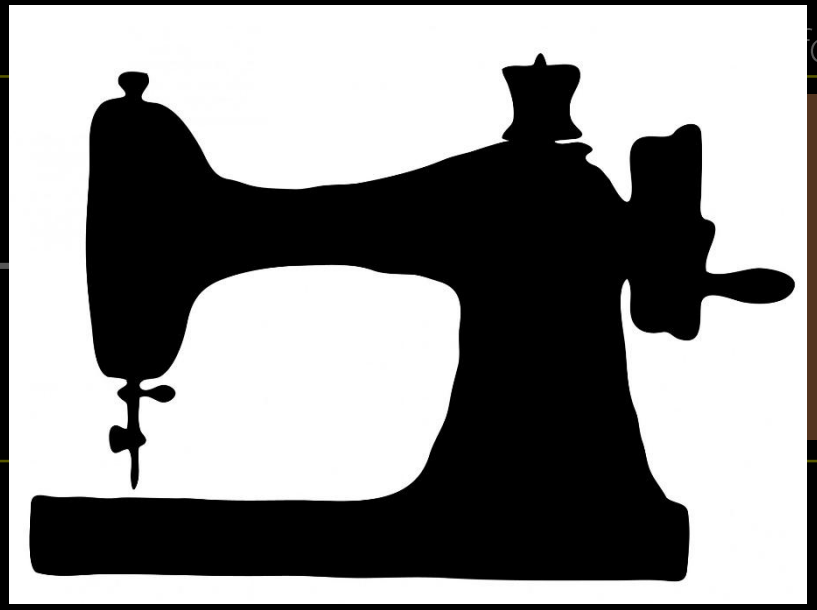
- Principled exploration of architectures
  - **Without** training all of them
- Finds network architecture automatically
  - Generalizes across tasks

- Implicitly explores **more architectures** than brute-force
- **Performs better** than architectures found via brute-force

Two one-task networks



Idiom: A stitch in time saves nine



conv1, pool1 conv2, pool2 fc6 fc7 fc8

Task A

Image

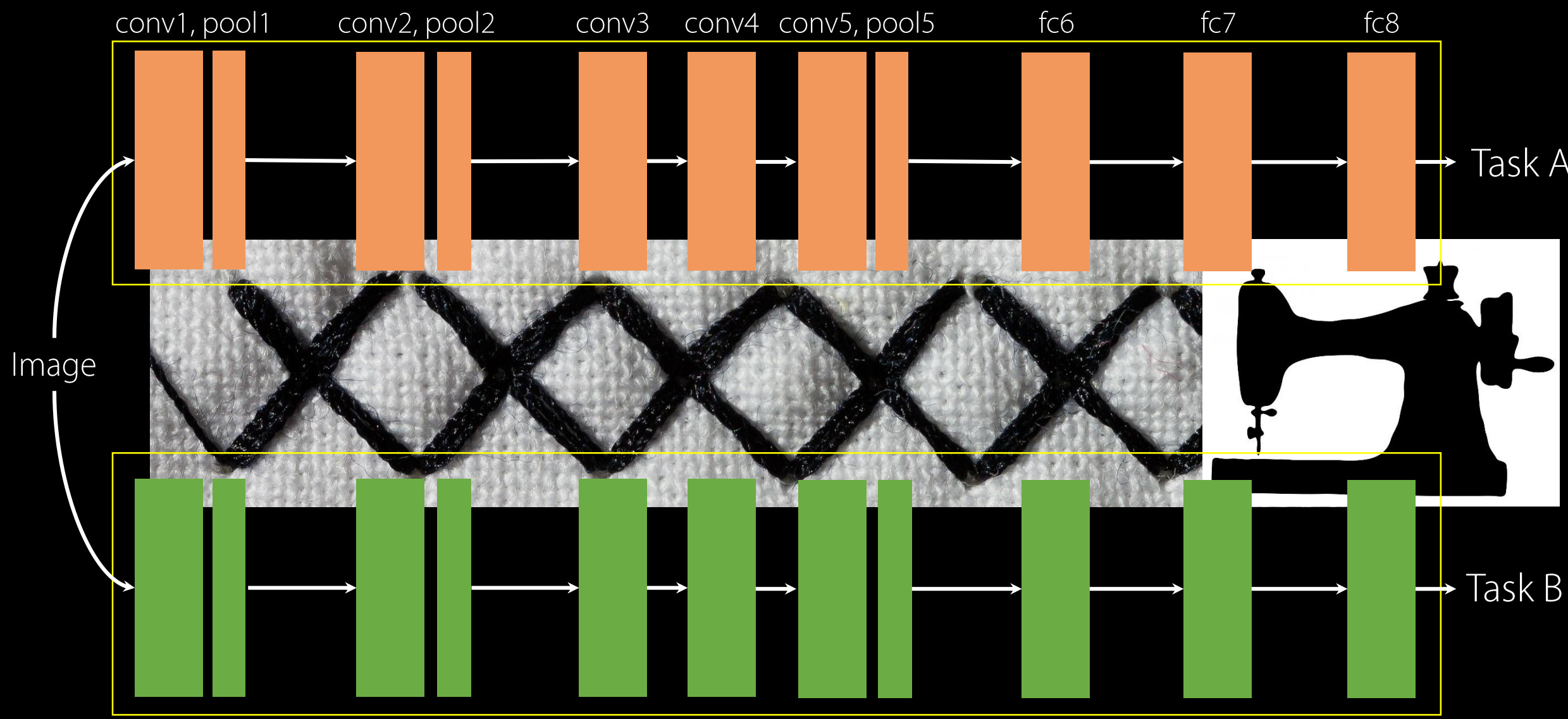
9 8 7 6 5 4 3 2,1

Task B

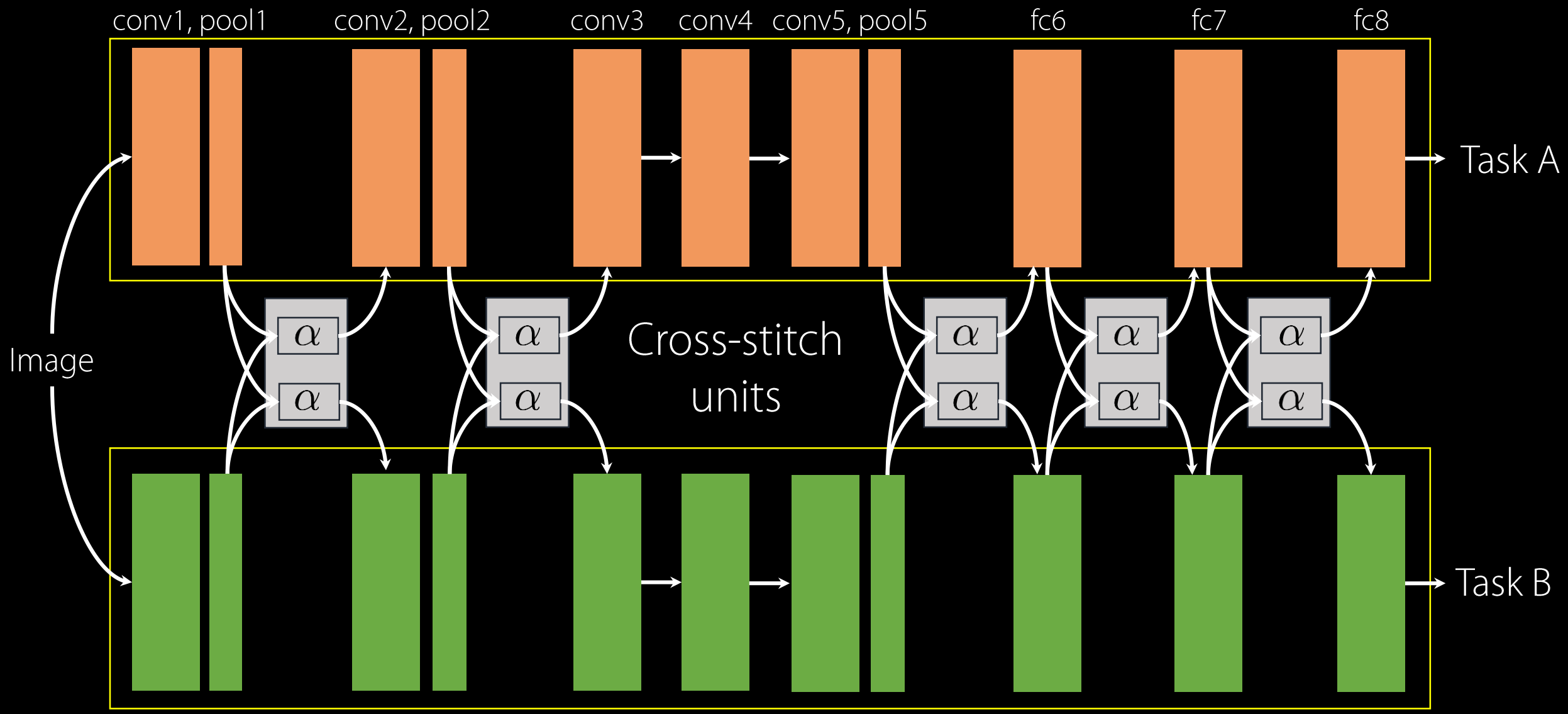
Generic Network
All Parameters Shared

Specific Network
No Parameters Shared

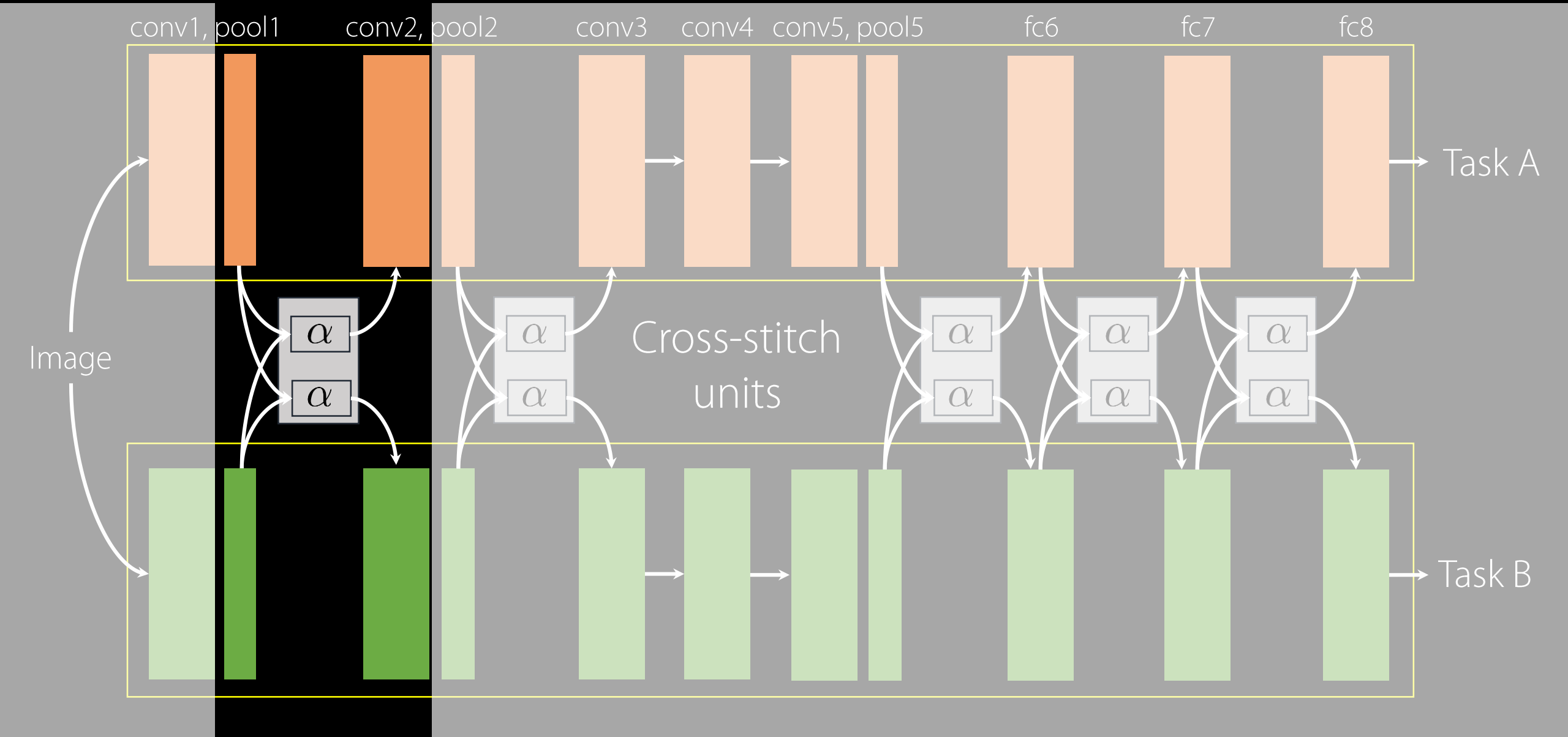
Art of cross-stitching

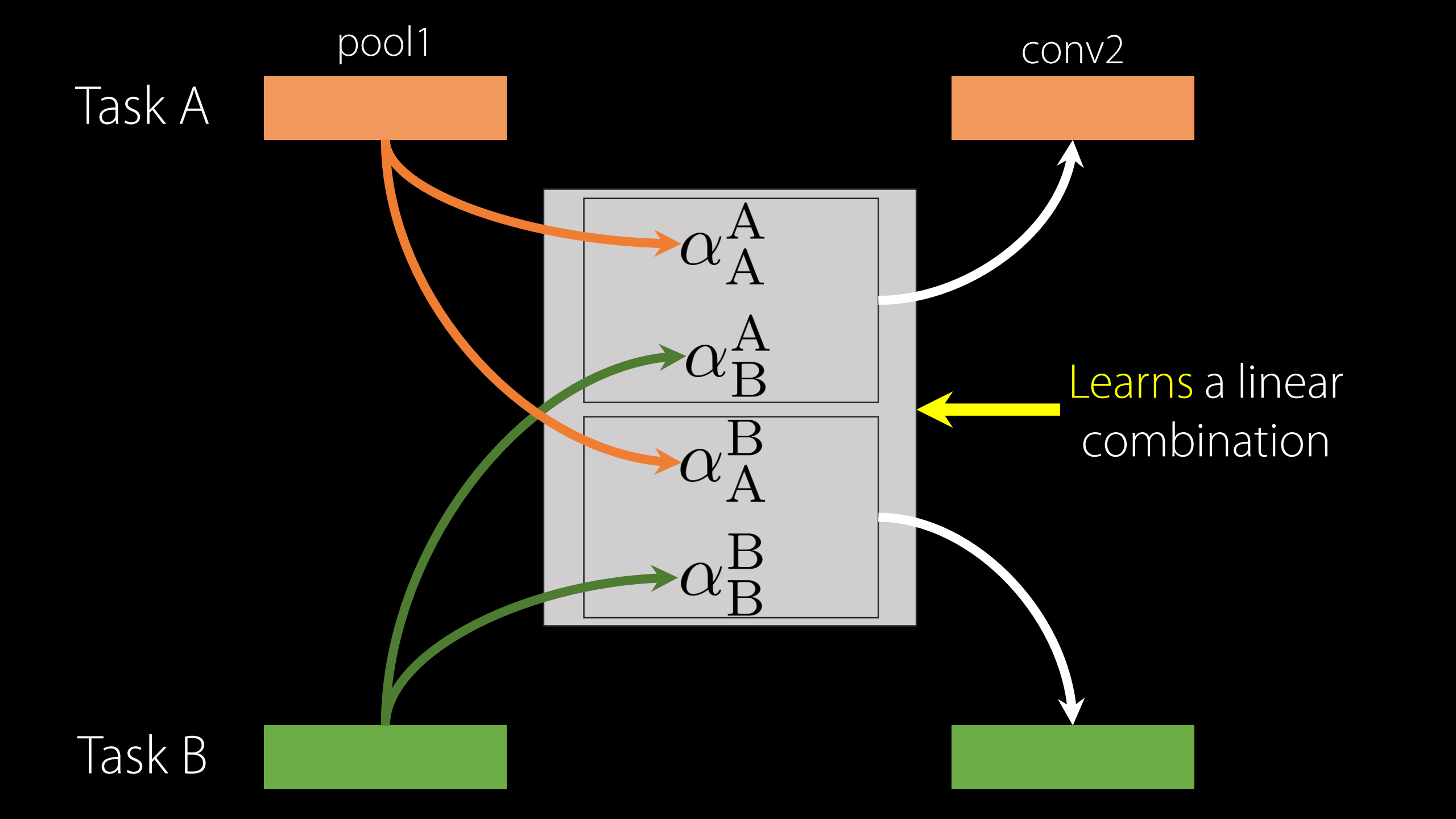


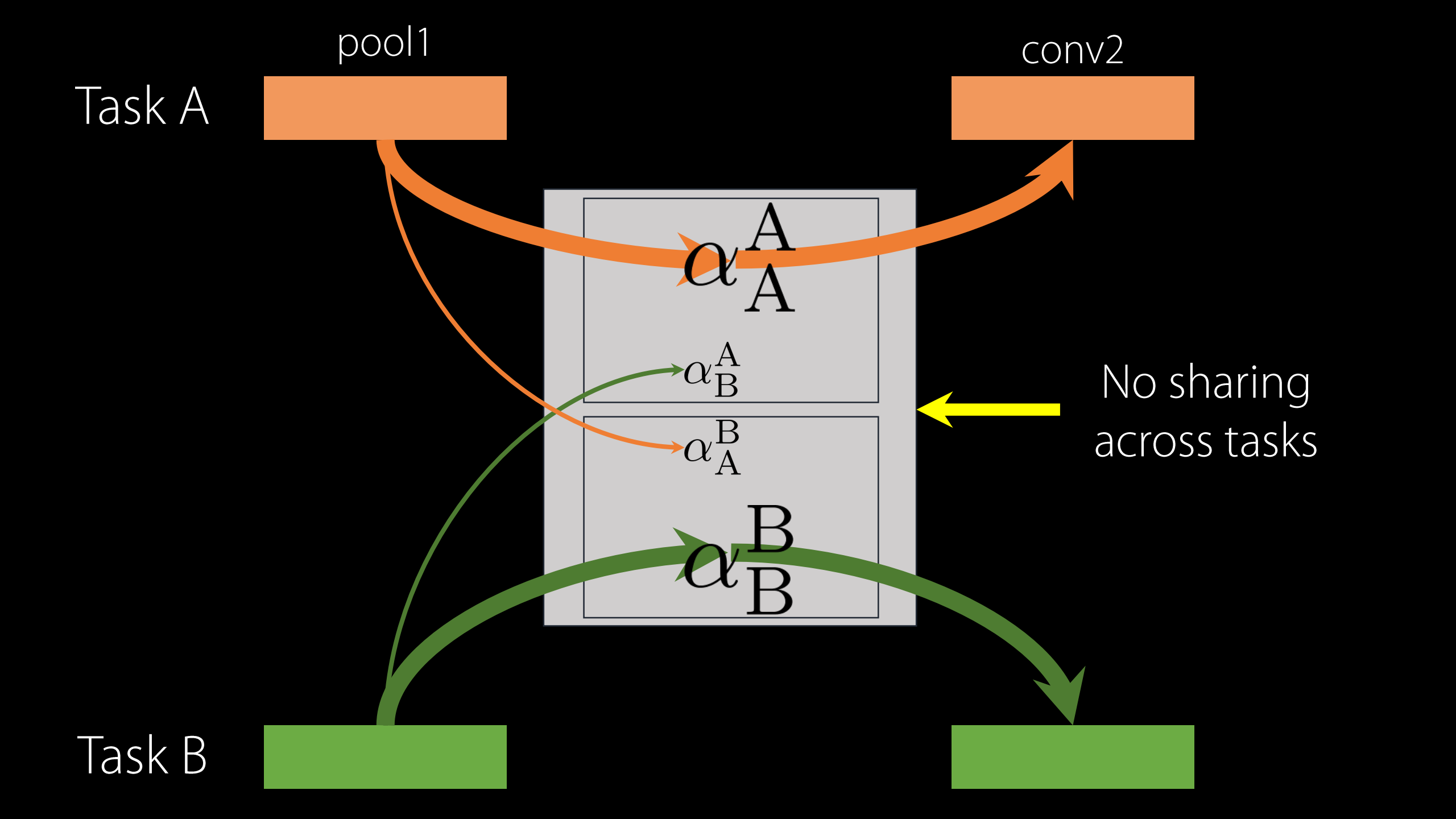
Adding cross-stitch

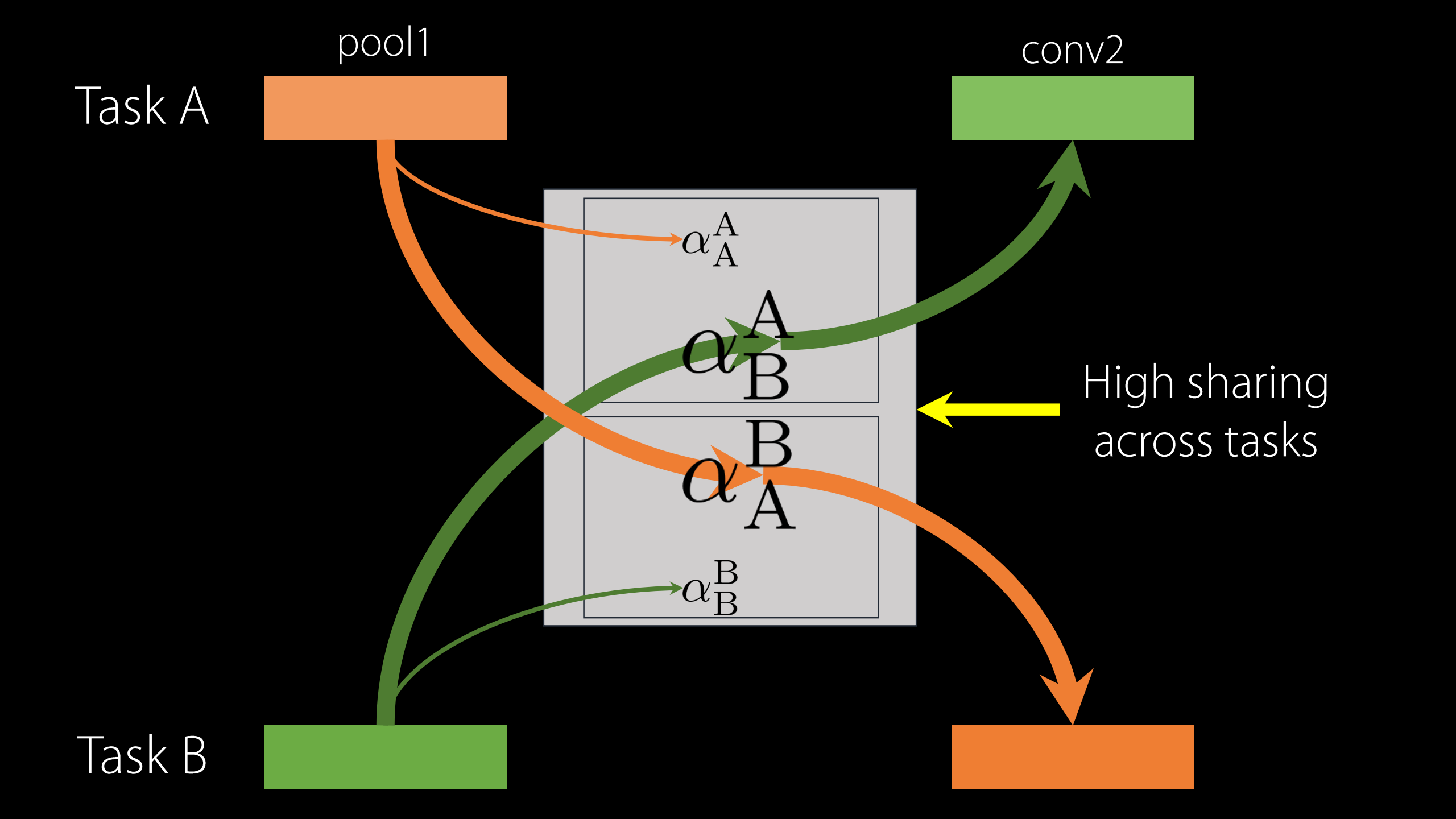


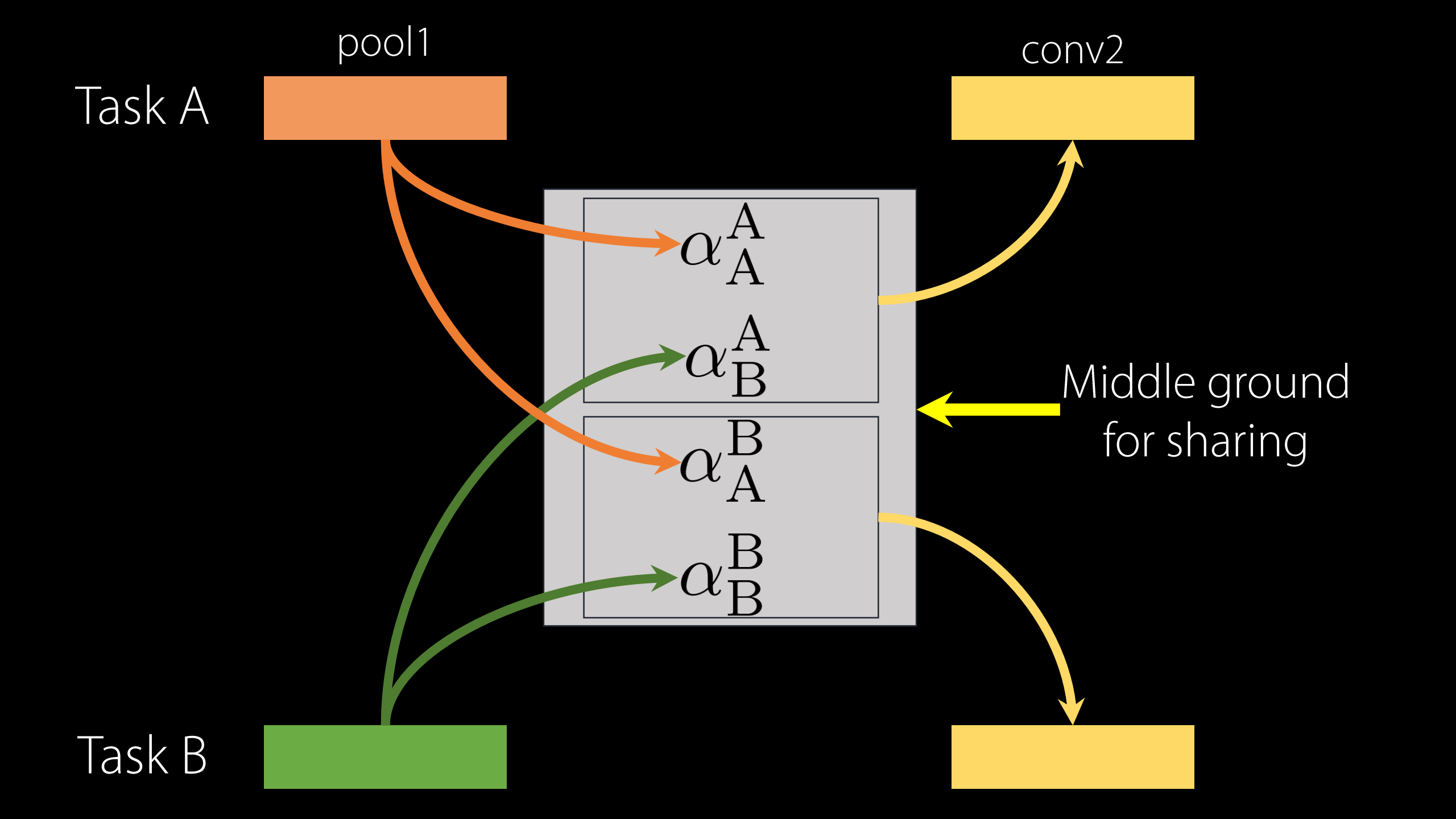
Our proposed unit: Cross-stitch



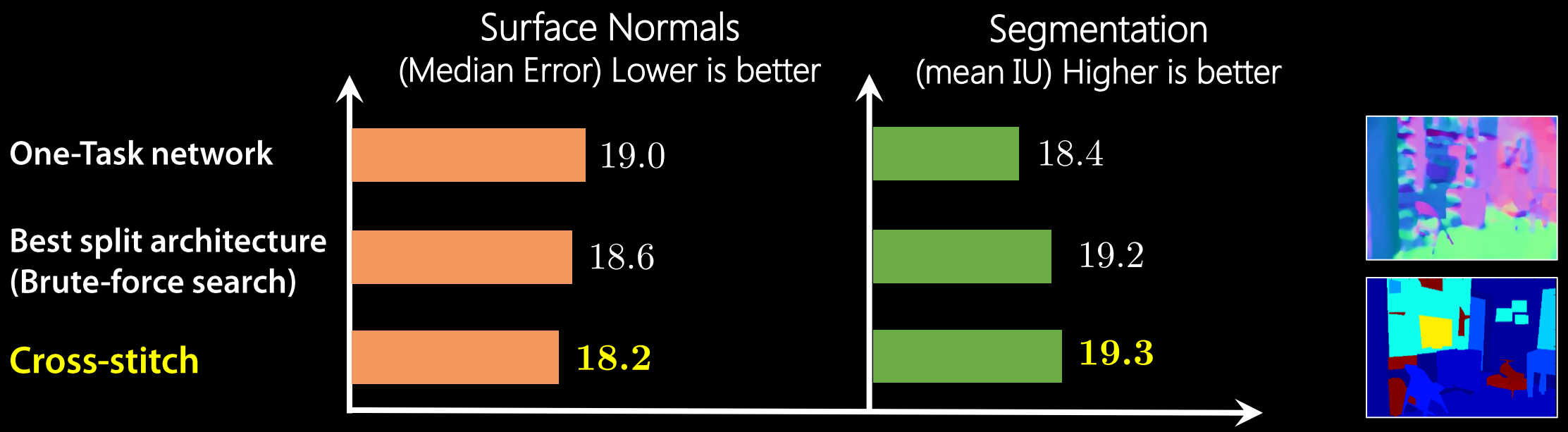

pool1
conv2
Task A
α_A^A
α_B^A
α_A^B
α_B^B
Learns a linear combination
Task B

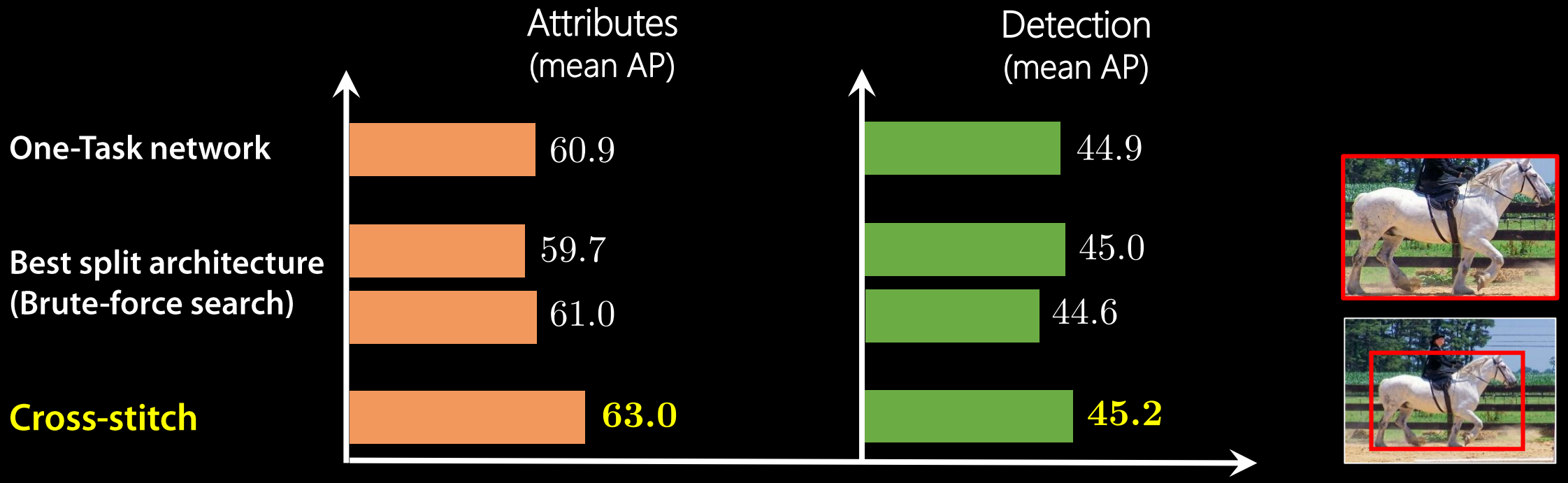

pool1
conv2
Task A
α_A^A
α_B^A
α_A^B
α_B^B
No sharing
across tasks
Task B


pool1
conv2
Task A
α_A^A
α_B^A
α_A^B
α_B^B
High sharing
across tasks
Task B

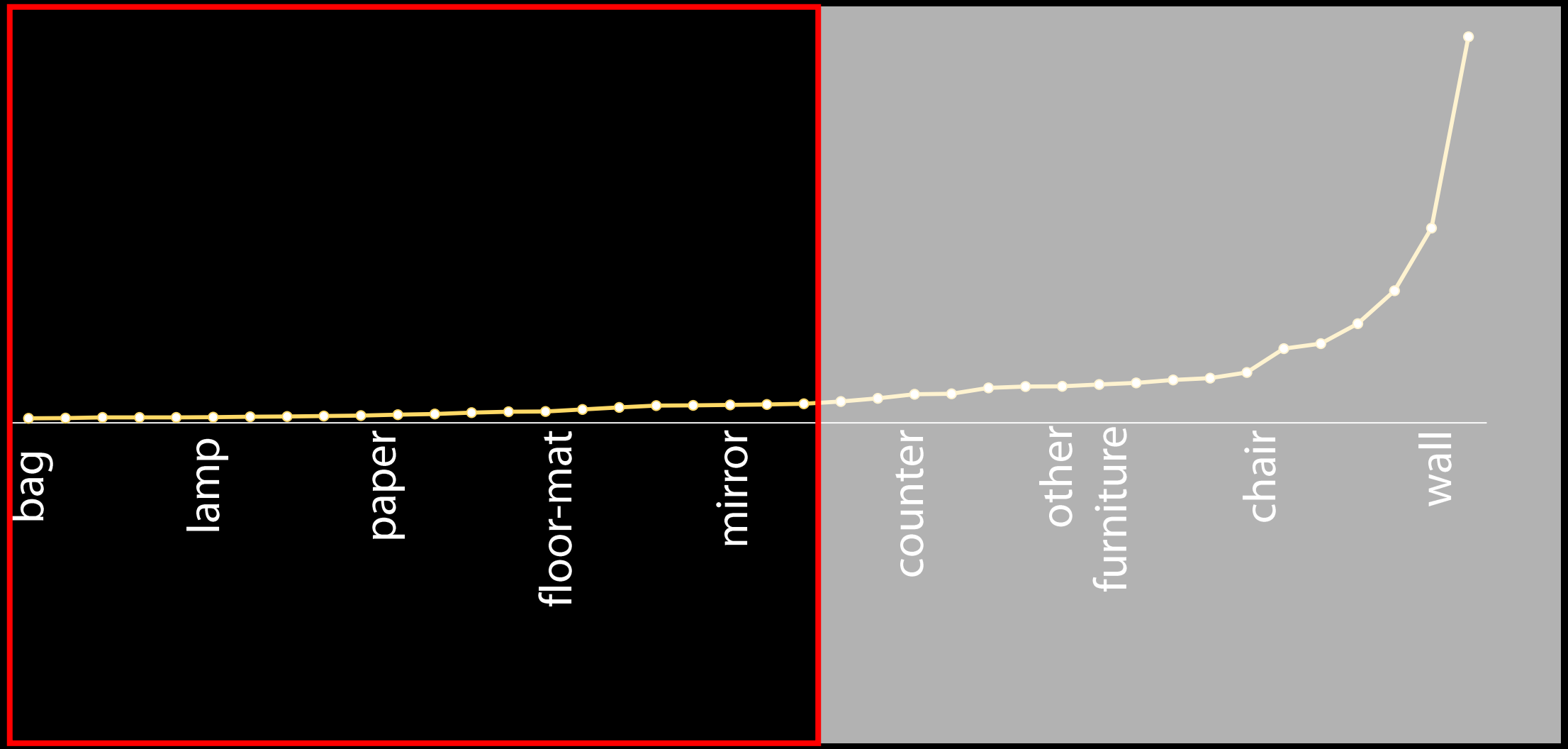

pool1
conv2
Task A
α_A^A
α_B^A
α_A^B
α_B^B
Middle ground
for sharing
Task B

Results

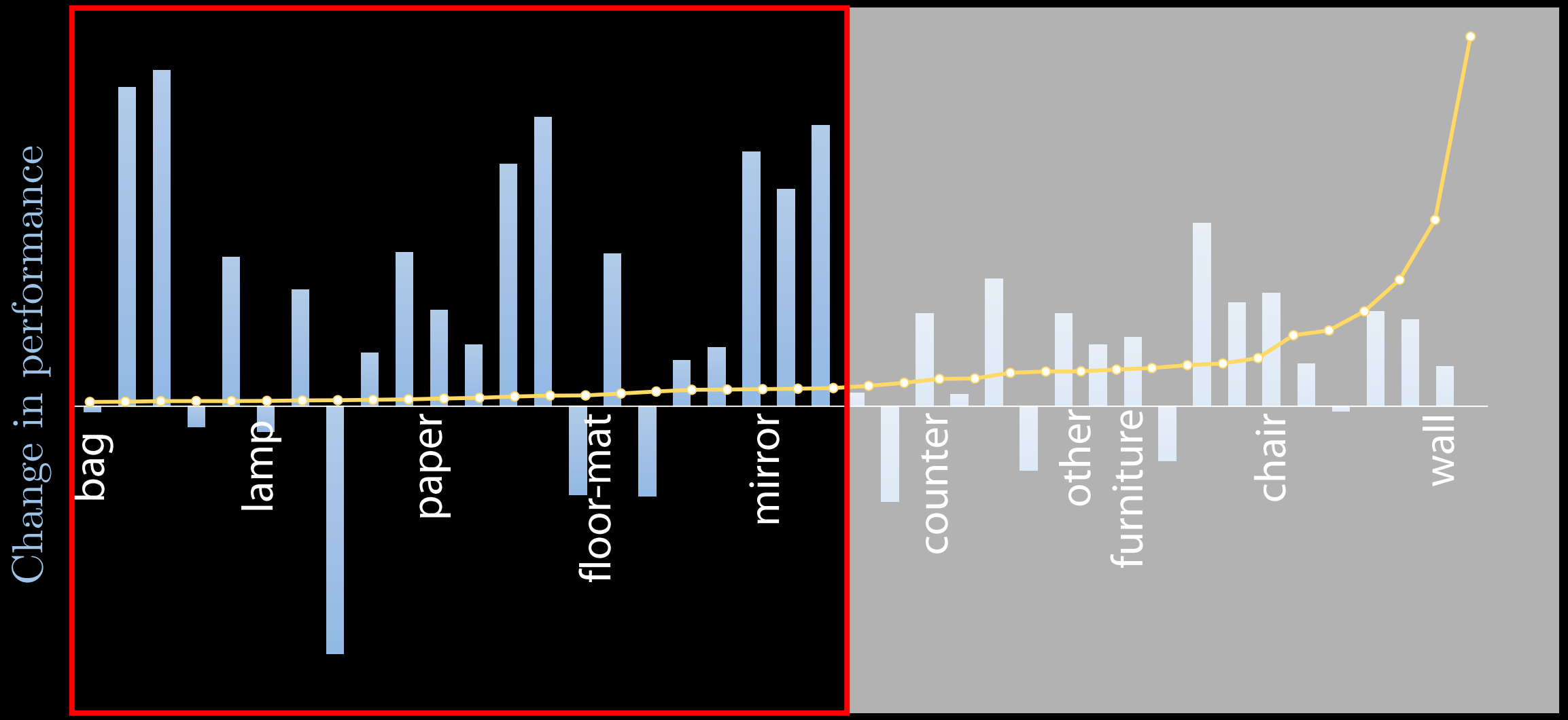

Surface Normals
(Median Error) Lower is better
Segmentation
(mean IU) Higher is better
One-Task network
Best split architecture
(Brute-force search)
Cross-stitch
19.0
18.6
18.2
18.4
19.2
19.3


Attributes
(mean AP)
Detection
(mean AP)
One-Task network
Best split architecture
(Brute-force search)
Cross-stitch
60.9
59.7
61.0
63.0
44.9
45.0
44.6
45.2

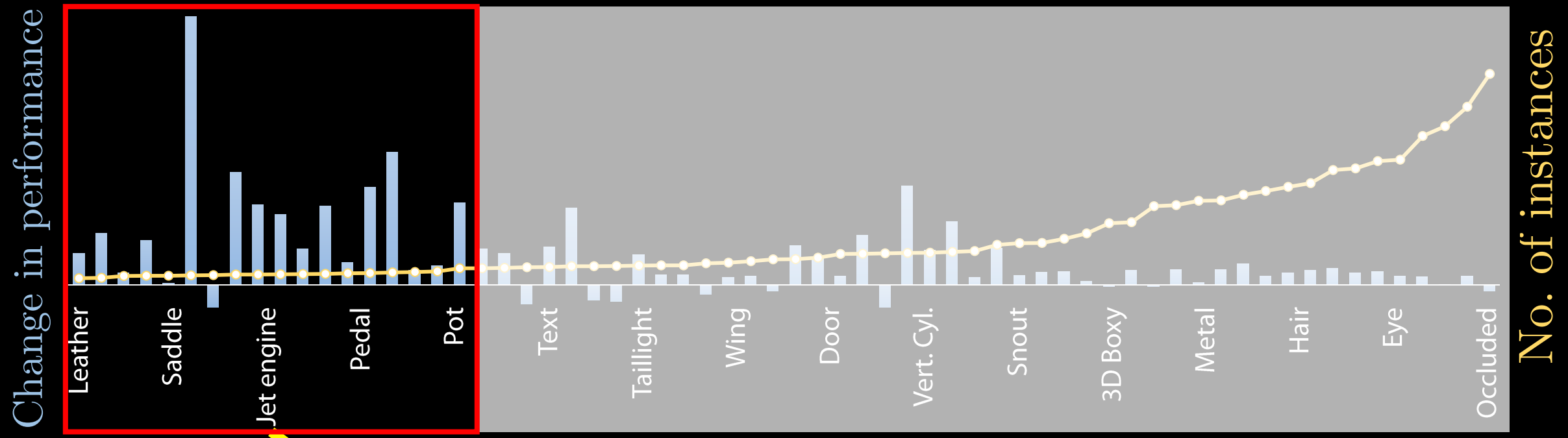
Data starved segmentation



Data starved segmentation



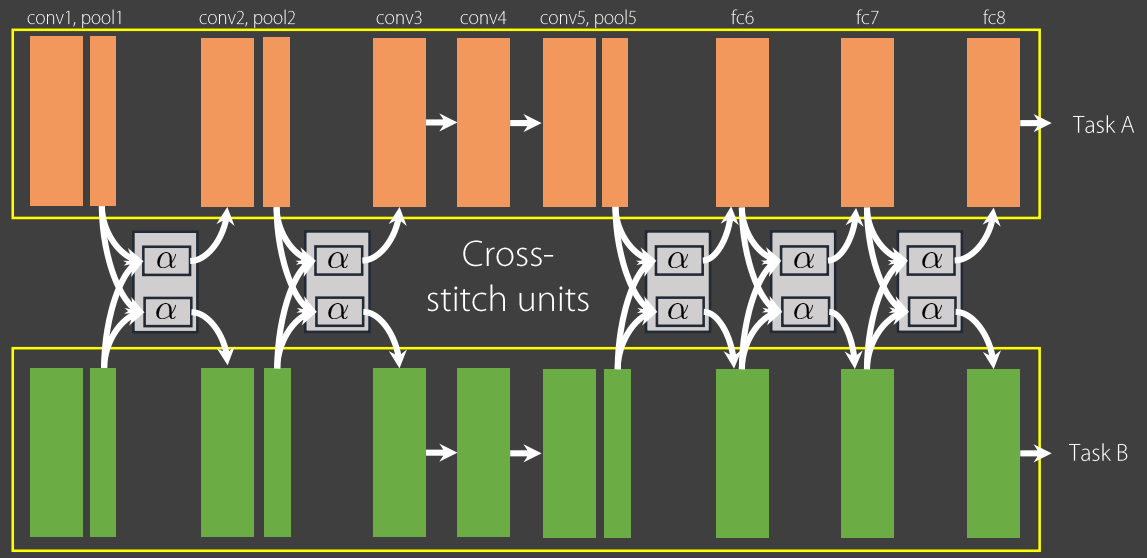
Data starved Attributes

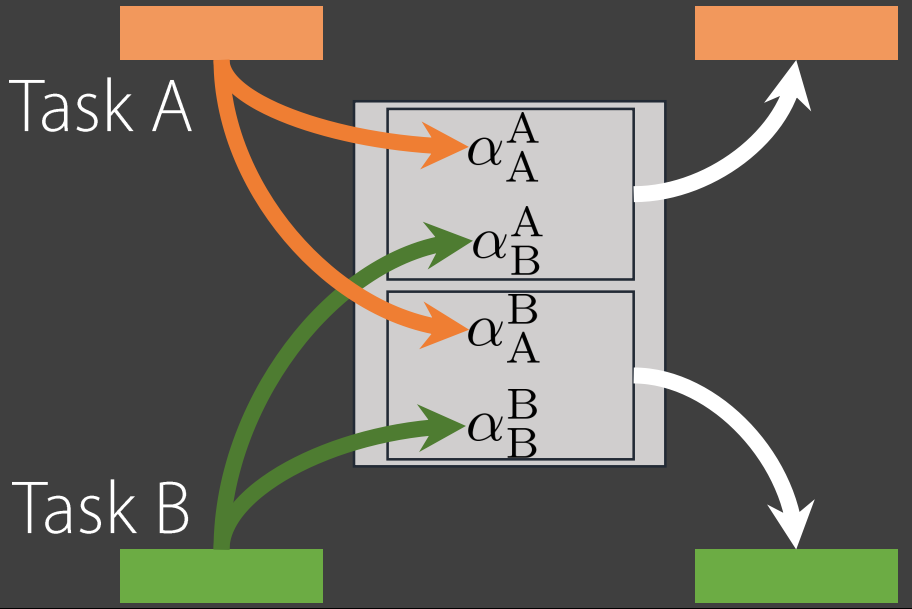


+4.6% mAP for bottom 10 attributes

+4.3% mAP for bottom 20 attributes

Cross-stitch Networks





Poster # 22

Session 3-2 (4:45pm-6:45pm)