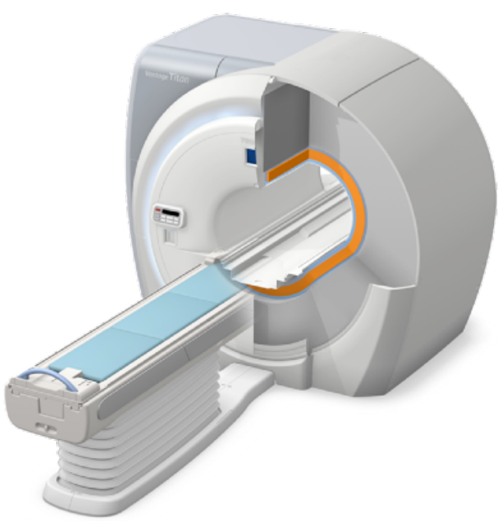
HUMUS-Net

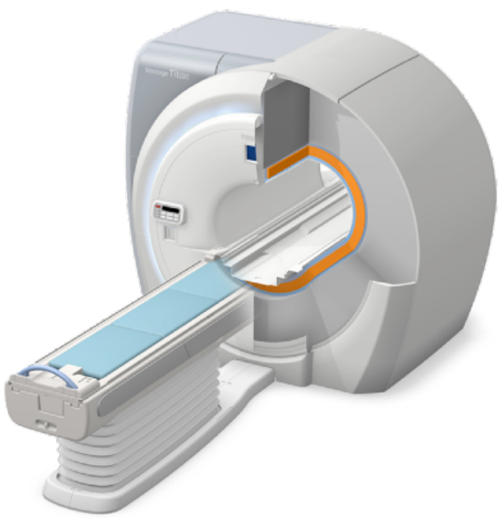
Hybrid Unrolled Multi-scale Network Architecture for Accelerated MRI Reconstruction

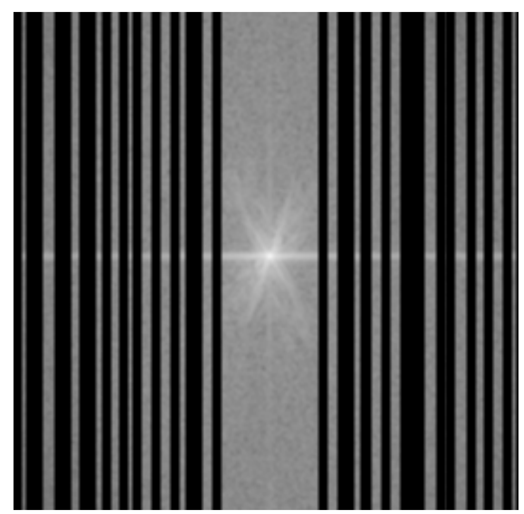
Zalan Fabian, Mahdi Soltanolkotabi

MRI reconstruction

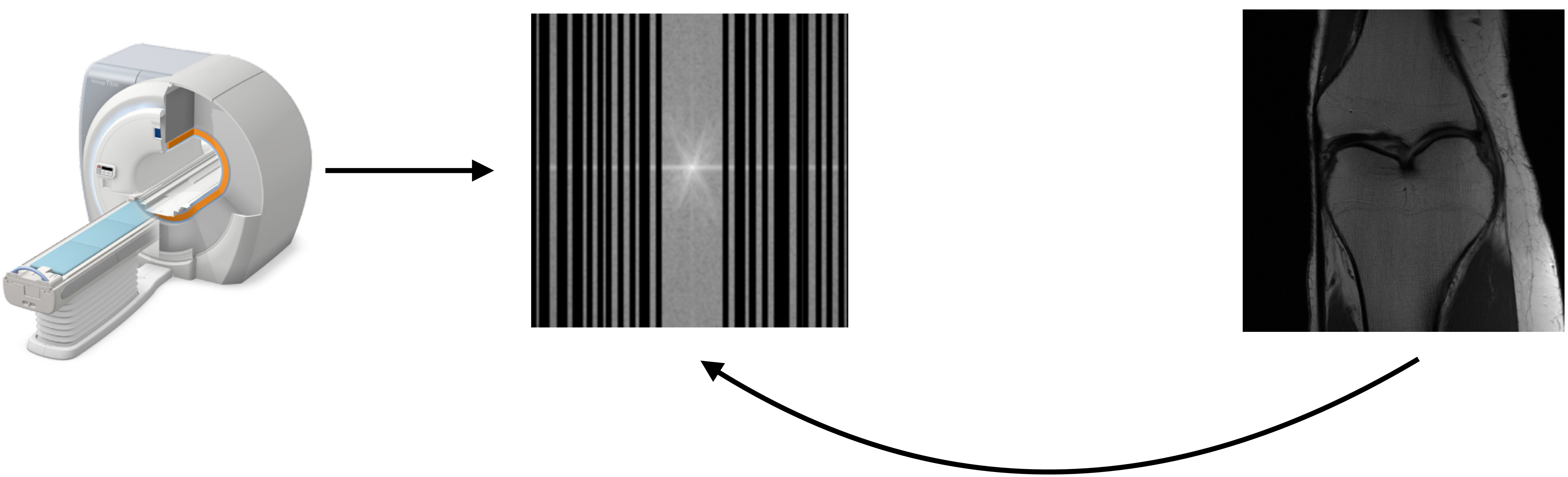


MRI reconstruction

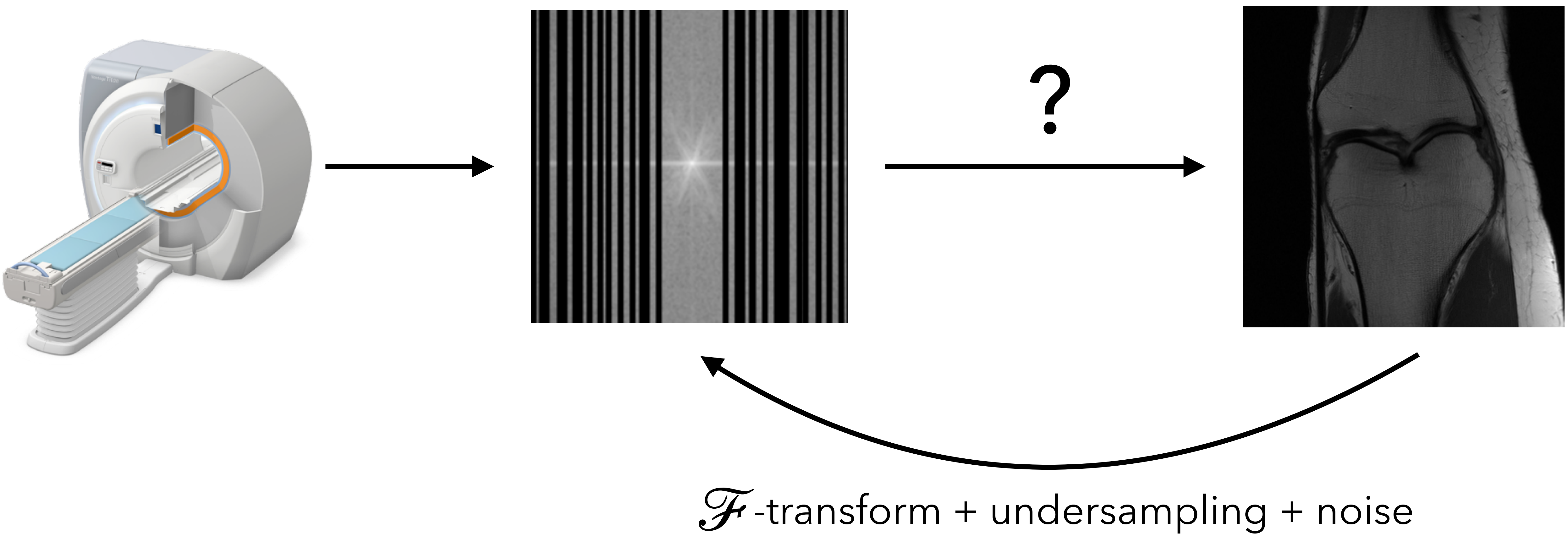




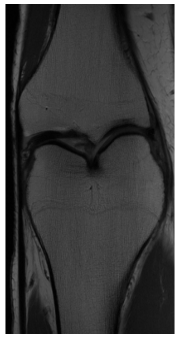
MRI reconstruction



MRI reconstruction

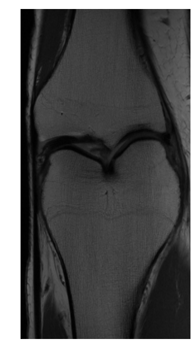


MRI forward model



ground truth
anatomy

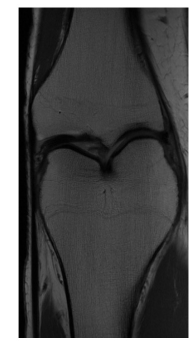
MRI forward model

 $\circ$ 

sensitivity map of ith coil

ground truth anatomy

MRI forward model

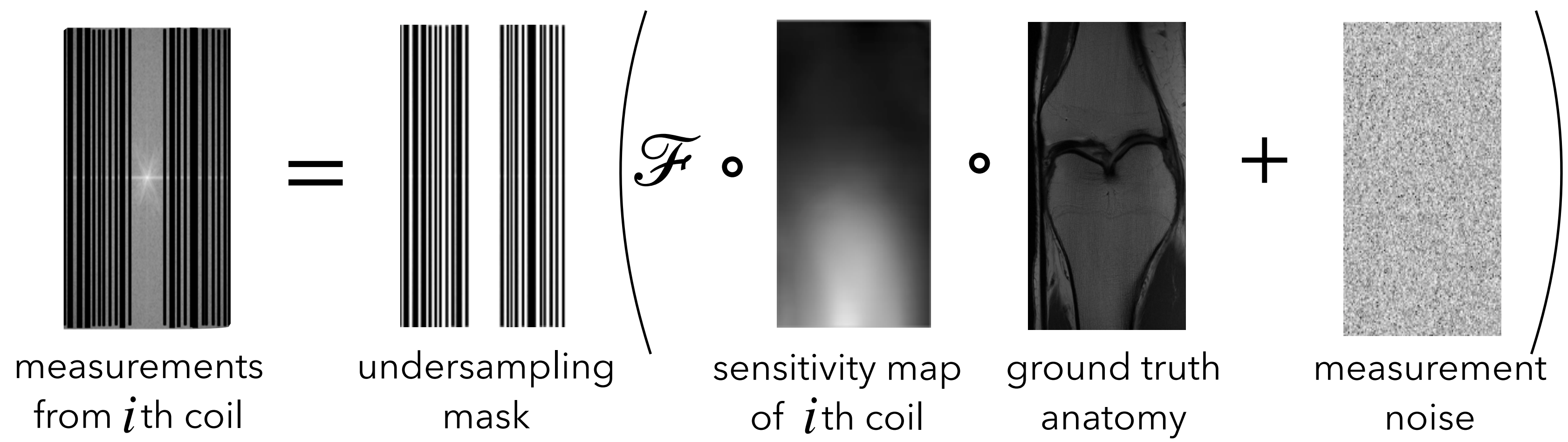
$\mathscr{F} \circ$  $\circ$  $+$ 

sensitivity map of ith coil

ground truth anatomy

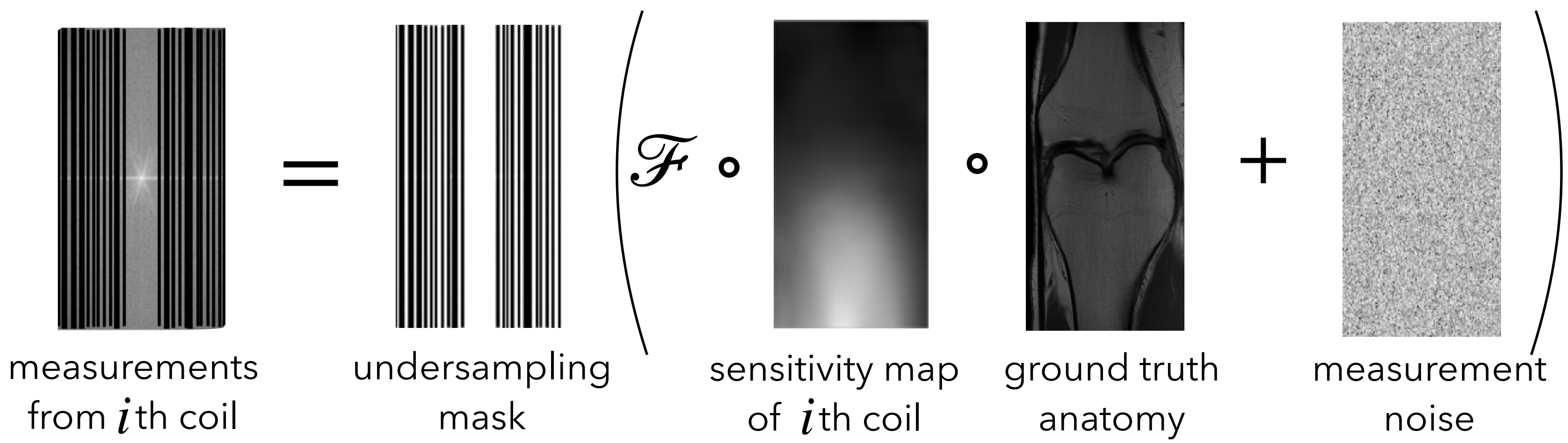
measurement noise

MRI forward model



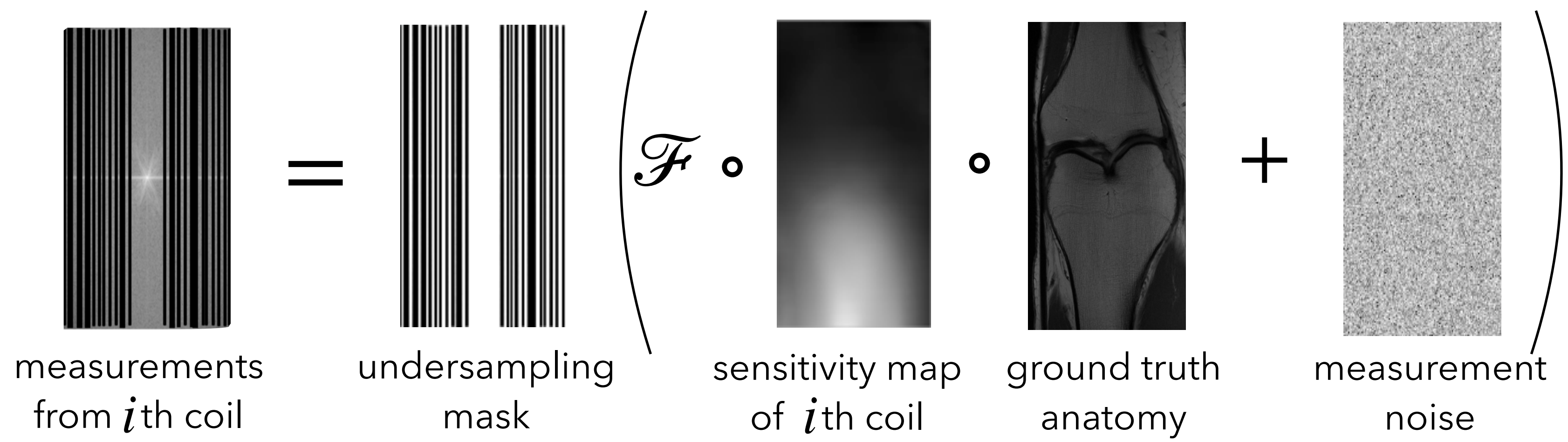
MRI forward model

$$\tilde{k}_i = M(\mathscr{F} S_i x^* + z_i) \quad i = 1,..,N$$



MRI forward model

$$\tilde{k}_i = M(\mathscr{F} S_i x^* + z_i) \quad i = 1,..,N$$



Compressed sensing reconstruction

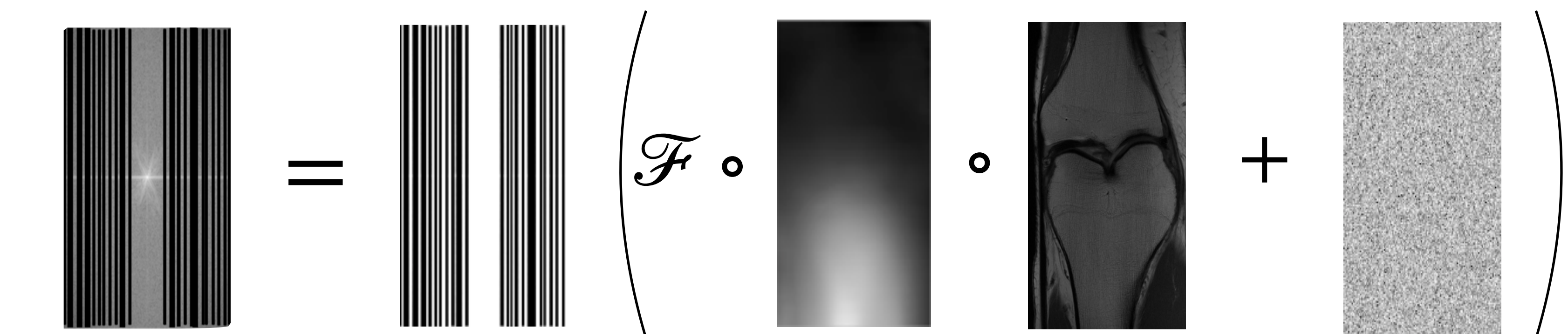
$$\hat{x} = arg \min_x \|\mathscr{A}(x) - \tilde{k}\|^2 + \mathscr{R}(x)$$

data consistency

prior knowledge

MRI forward model

$$\tilde{k}_i = M(\mathscr{F} S_i x^* + z_i) \quad i = 1,..,N$$



measurements from ith coil = undersampling mask $\left(\mathscr{F} \circ \text{sensitivity map of } i\text{th coil} \circ \text{ground truth anatomy} + \text{measurement noise}\right)$

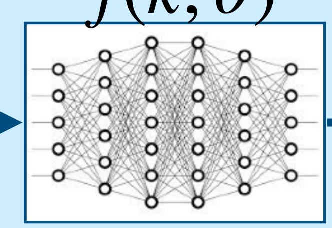
Compressed sensing reconstruction

$$\hat{x} = \underbrace{arg \min_x \|\mathscr{A}(x) - \tilde{k}\|^2}_{\text{data consistency}} + \underbrace{\mathscr{R}(x)}_{\text{prior knowledge}}$$

End-to-end DL reconstruction

$f(\tilde{k};\theta)$

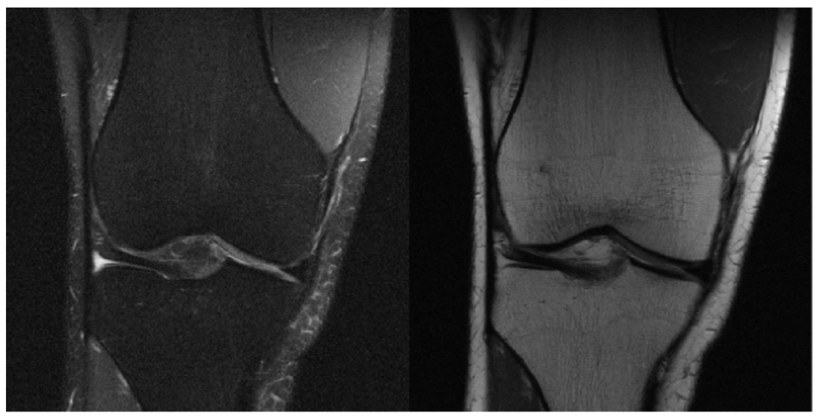
$k \rightarrow$ DL model $\rightarrow \hat{x} \rightarrow \mathscr{L}(\hat{x}, x^*) \leftarrow x^*$



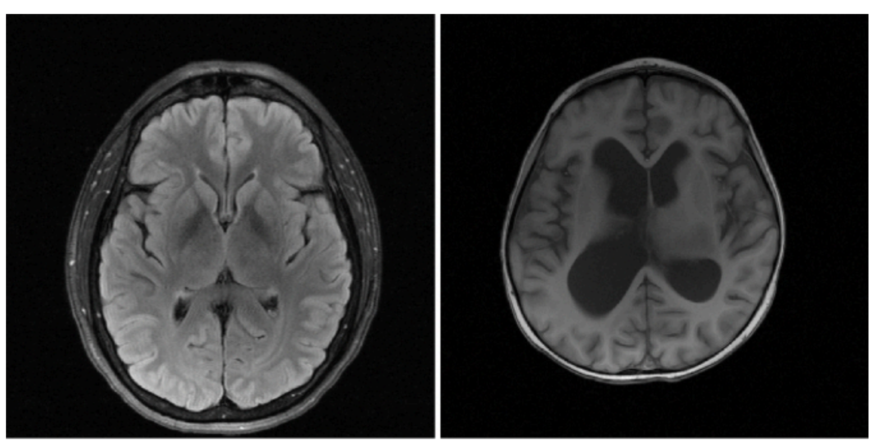
DL model

fastMRI dataset

- Largest public dataset of fully sampled raw MRI measurements



	Volumes		Slices	
	Multi-coil	Single-coil	Multi-coil	Single-coil
training	973	973	34,742	34,742
validation	199	199	7,135	7,135
test	118	108	4,092	3,903
challenge	104	92	3,810	3,305



Field Strength	1.5T	3T
T1	375	407
T1 POST	849	641
T2	1651	2515
FLAIR	126	406
Total	3001	3969

Zbontar, J., Knoll, F., Sriram, A., Murrell, T., Huang, Z., Muckley, M. J., ... & Lui, Y. W. (2018). fastMRI: An open dataset and benchmarks for accelerated MRI. *arXiv preprint arXiv:1811.08839*.

fastMRI Public Leaderboard

Single-Coil Knee | Multi-Coil Knee | Multi-Coil Brain

How well can you reconstruct a single MRI image given a masked k-space and signal from multiple coils? This challenge provides a space for researchers familiar with the physics of MRI and to build solutions compatible with modern MR machinery.

	Acceleration 8x	NMSE	SSIM	PSNR	NYU DATA ONLY
fastMRI Repo End-to-End VarNet 11/11/2020	8x	0.0085	0.8920	37.1	✓
SubtleMR 6/23/2020	8x	0.0085	0.8919	37.1	✓
Deneme4 10/7/2021	8x	0.0085	0.8919	37.1	✓
Wd_UN_I_VN_Sp 10/12/2021	8x	0.0083	0.8917	37.1	✓
Wd_Av_UN_I_with_VN 10/12/2021	8x	0.0083	0.8914	37.1	✓

fastMRI Public Leaderboard

Single-Coil Knee | **Multi-Coil Knee** | Multi-Coil Brain

How well can you reconstruct a single MRI image given a masked k-space and signal from multiple coils? This challenge provides a space for researchers familiar with the physics of MRI and to build solutions compatible with modern MR machinery.

	Acceleration 8x	NMSE	SSIM	PSNR	NYU DATA ONLY
fastMRI Repo End-to-End VarNet 11/11/2020	8x	0.0085	0.8920	37.1	✓
SubtleMR 6/23/2020	8x	0.0085	0.8919	37.1	✓
Deneme4 10/7/2021	8x	0.0085	0.8919	37.1	✓
Wd_UN_I_VN_Sp 10/12/2021	8x	0.0083	0.8917	37.1	✓
Wd_Av_UN_I_with_VN 10/12/2021	8x	0.0083	0.8914	37.1	✓

Top of the leaderboard for almost 2 years!

Unrolled Networks

- Inverse problem formulation

$$\hat{x} = arg \min_x \|\mathscr{A}(x) - y\|^2 + \mathscr{R}(x)$$

- Iterative solution via GD

$$x^{t+1} = x^t - \mu^t \left[\mathscr{A}* \left(\mathscr{A}(x^t) - y\right) + \nabla \mathscr{R}(x^t)\right]$$

Unrolled Networks

- Inverse problem formulation

$$\hat{x} = arg\min_{x} \|\mathscr{A}(x) - y\|^2 + \mathscr{R}(x)$$

- Iterative solution via GD

$$x^{t+1} = x^t - \mu^t \left[\mathscr{A}* \left(\mathscr{A}(x^t) - y\right) + \nabla \mathscr{R}(x^t)\right]$$

What is the best regularizer?

Unrolled Networks

- Inverse problem formulation

$$\hat{x} = arg \min_x \|\mathscr{A}(x) - y\|^2 + \mathscr{R}(x)$$

- Iterative solution via GD

$$x^{t+1} = x^t - \mu^t \left[\mathscr{A}* \left(\mathscr{A}(x^t) - y\right) + \Phi_\theta(x^t)\right]$$

Parameterize regularizer gradient as NN!

Unrolled Networks

- Inverse problem formulation

$$\hat{x} = arg\min_{x} \|\mathscr{A}(x) - y\|^2 + \mathscr{R}(x)$$

- Iterative solution via GD

$$x^{t+1} = x^t - \mu^t \left[\mathscr{A}^* \left(\mathscr{A}(x^t) - y\right) + \Phi_\theta(x^t)\right]$$

Parameterize regularizer gradient as NN!

$$\xrightarrow{x_0}$$

Unrolled Networks

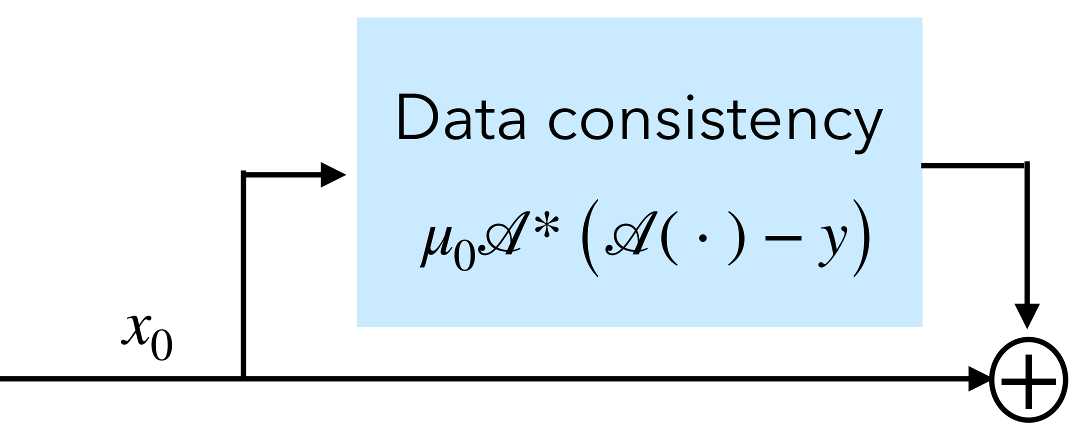
- Inverse problem formulation

$$\hat{x} = arg\min_x \|\mathscr{A}(x) - y\|^2 + \mathscr{R}(x)$$

- Iterative solution via GD

$$x^{t+1} = x^t - \mu^t \left[\mathscr{A}^* \left(\mathscr{A}(x^t) - y\right) + \Phi_\theta(x^t)\right]$$

Parameterize regularizer gradient as NN!



Unrolled Networks

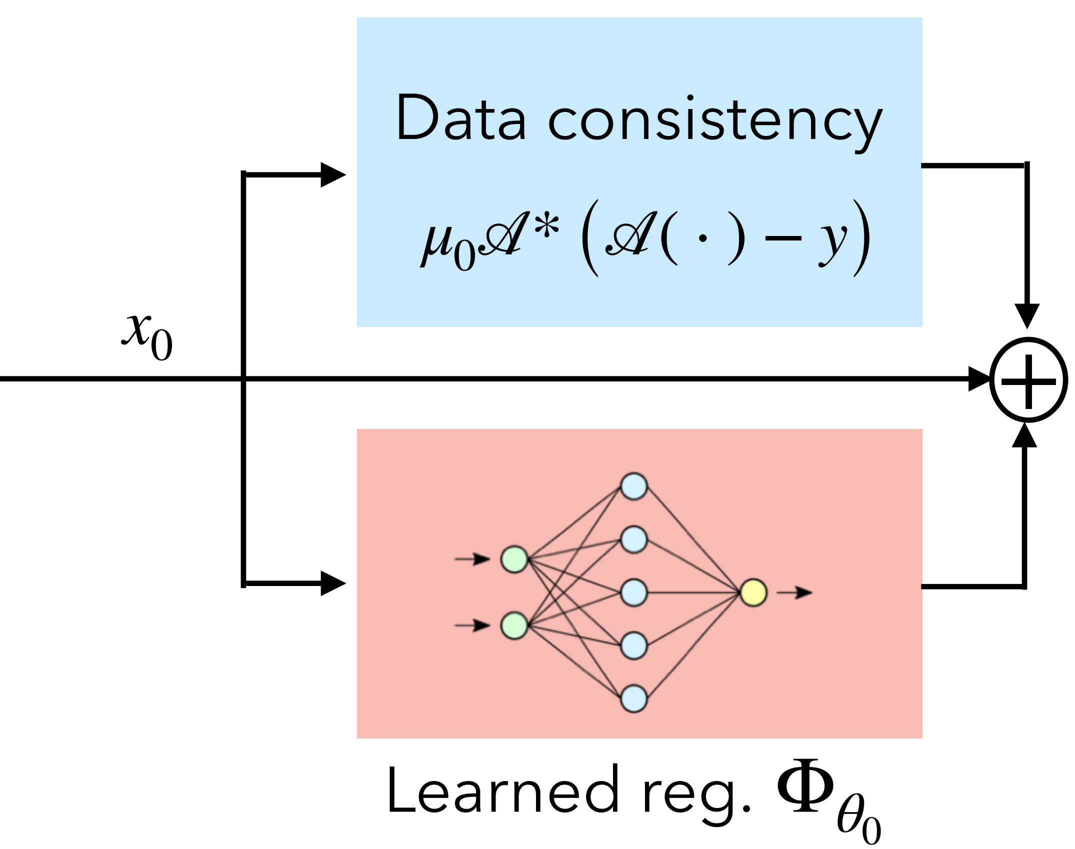
- Inverse problem formulation

$$\hat{x} = arg\min_x \|\mathscr{A}(x) - y\|^2 + \mathscr{R}(x)$$

- Iterative solution via GD

$$x^{t+1} = x^t - \mu^t \left[\mathscr{A}^* \left(\mathscr{A}(x^t) - y\right) + \Phi_\theta(x^t)\right]$$

Parameterize regularizer gradient as NN!



Unrolled Networks

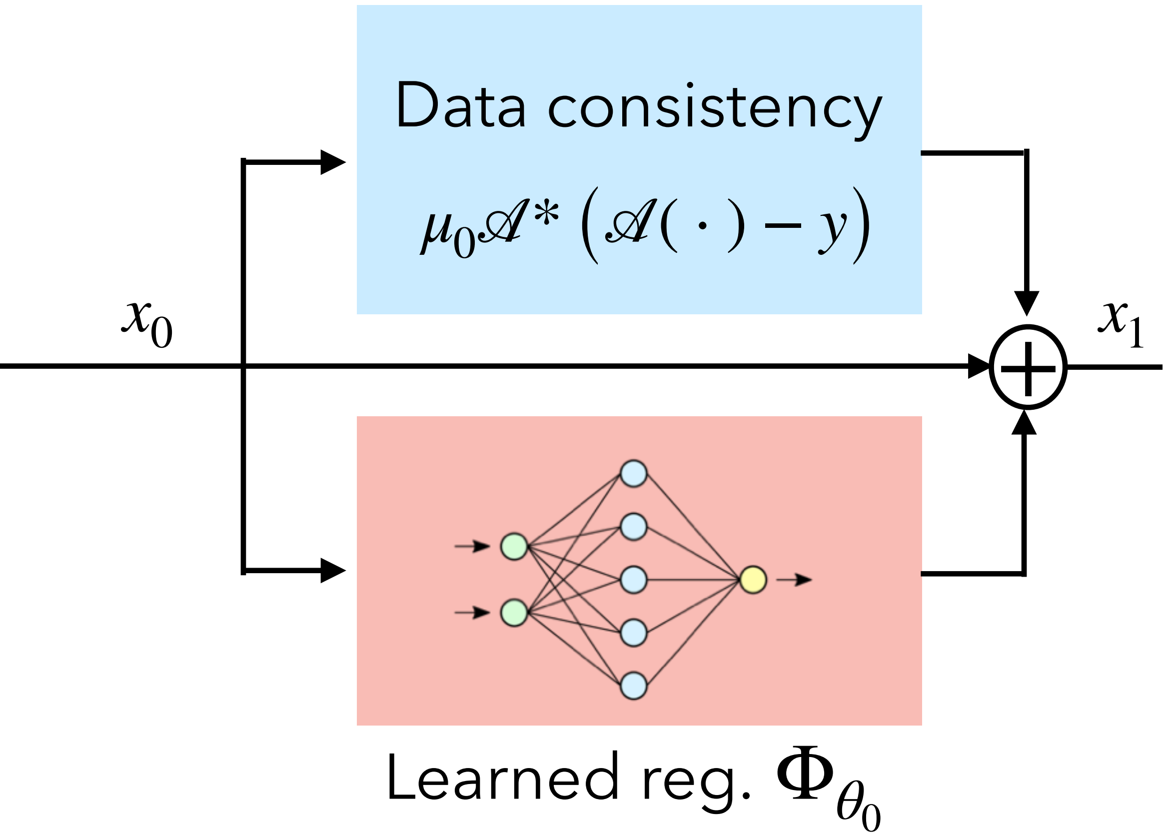
- Inverse problem formulation

$$\hat{x} = arg\min_x \|\mathscr{A}(x) - y\|^2 + \mathscr{R}(x)$$

- Iterative solution via GD

$$x^{t+1} = x^t - \mu^t \left[\mathscr{A}^* \left(\mathscr{A}(x^t) - y\right) + \Phi_\theta(x^t)\right]$$

Parameterize regularizer gradient as NN!



Unrolled Networks

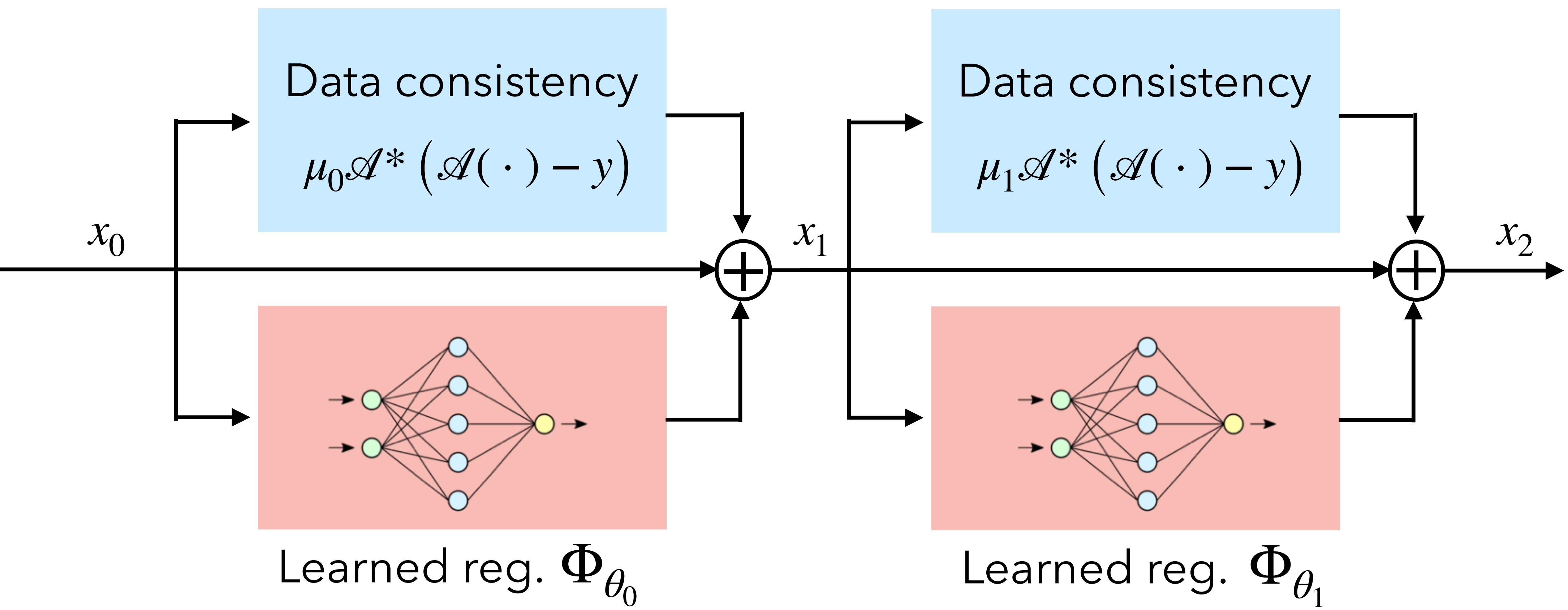
- Inverse problem formulation

$$\hat{x} = arg\min_{x} \|\mathscr{A}(x) - y\|^2 + \mathscr{R}(x)$$

- Iterative solution via GD

$$x^{t+1} = x^t - \mu^t \left[\mathscr{A}^* \left(\mathscr{A}(x^t) - y\right) + \Phi_\theta(x^t)\right]$$

Parameterize regularizer gradient as NN!



Unrolled Networks

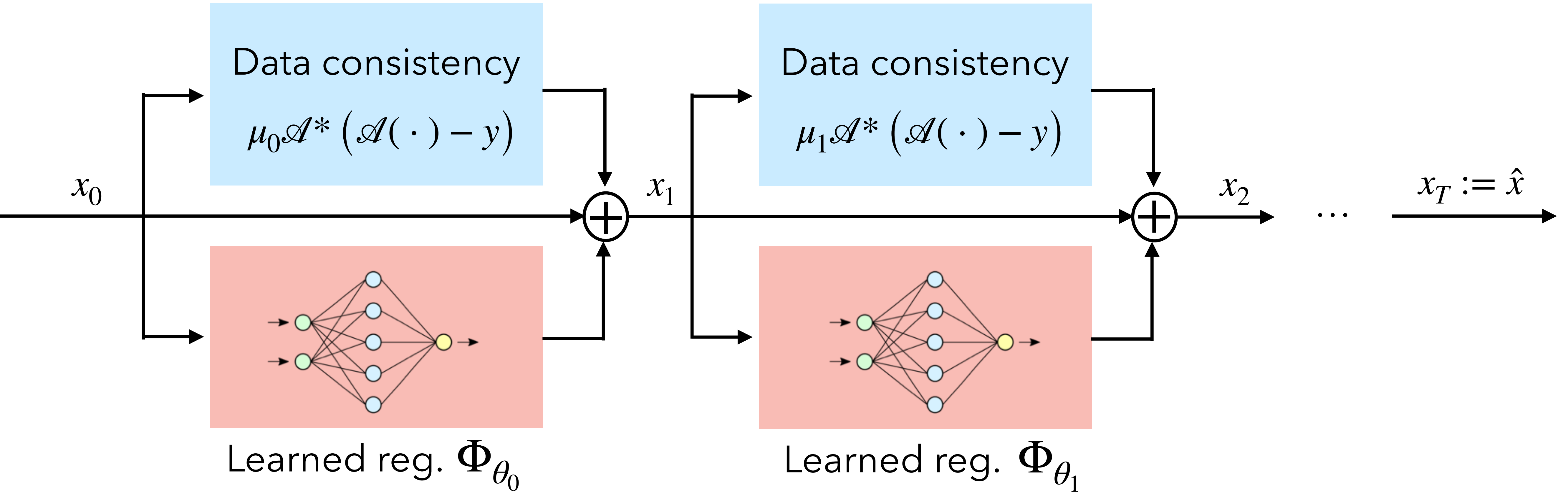
- Inverse problem formulation

$$\hat{x} = arg \min_x \|\mathscr{A}(x) - y\|^2 + \mathscr{R}(x)$$

- Iterative solution via GD

$$x^{t+1} = x^t - \mu^t \left[\mathscr{A}^* \left(\mathscr{A}(x^t) - y\right) + \Phi_\theta(x^t)\right]$$

Parameterize regularizer gradient as NN!



E2E VarNet

- Unroll GD iterations in k-space

Sriram, A., Zbontar, J., Murrell, T., Defazio, A., Zitnick, C. L., Yakubova, N., ... & Johnson, P. (2020). End-to-end variational networks for accelerated MRI reconstruction. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 64-73). Springer, Cham.

E2E VarNet

- Unroll GD iterations in k-space

$$x^{t+1} = x^t - \mu^t \left[\mathscr{A}^* \left(\mathscr{A}(x^t) - y \right) + \Phi_\theta(x^t) \right]$$

Sriram, A., Zbontar, J., Murrell, T., Defazio, A., Zitnick, C. L., Yakubova, N., ... & Johnson, P. (2020). End-to-end variational networks for accelerated MRI reconstruction. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 64-73). Springer, Cham.

E2E VarNet

- Unroll GD iterations in k-space

$$x^{t+1} = x^t - \mu^t \left[\mathscr{A}^* \left(\mathscr{A}(x^t) - y \right) + \Phi_\theta(x^t) \right] \quad \xrightarrow{\mathscr{F}}$$

Sriram, A., Zbontar, J., Murrell, T., Defazio, A., Zitnick, C. L., Yakubova, N., ... & Johnson, P. (2020). End-to-end variational networks for accelerated MRI reconstruction. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 64-73). Springer, Cham.

E2E VarNet

- Unroll GD iterations in k-space

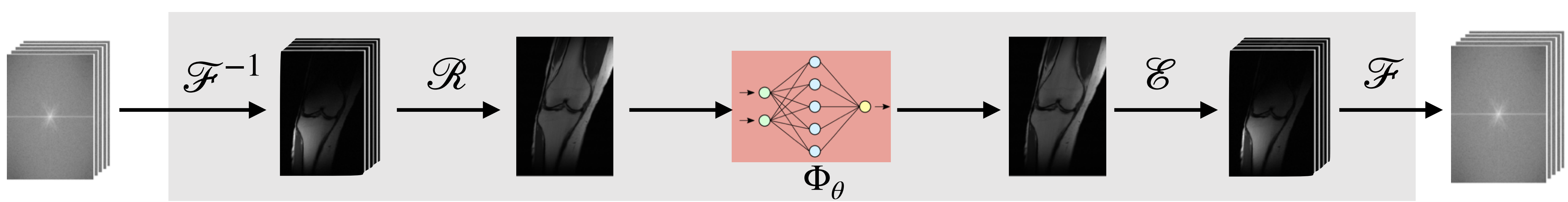
$$x^{t+1} = x^t - \mu^t \left[\mathscr{A}^* \left(\mathscr{A}(x^t) - y\right) + \Phi_\theta(x^t)\right] \quad \xrightarrow{\mathscr{F}} \quad k^{t+1} = k^t - \mu^t M(k^t - \tilde{k}) + \Psi(k^t)$$

Sriram, A., Zbontar, J., Murrell, T., Defazio, A., Zitnick, C. L., Yakubova, N., ... & Johnson, P. (2020). End-to-end variational networks for accelerated MRI reconstruction. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 64-73). Springer, Cham.

E2E VarNet

- Unroll GD iterations in k-space

$$x^{t+1} = x^t - \mu^t \left[\mathscr{A}^* \left(\mathscr{A}(x^t) - y\right) + \Phi_\theta(x^t)\right] \xrightarrow{\mathscr{F}} k^{t+1} = k^t - \mu^t M(k^t - \tilde{k}) + \Psi(k^t)$$

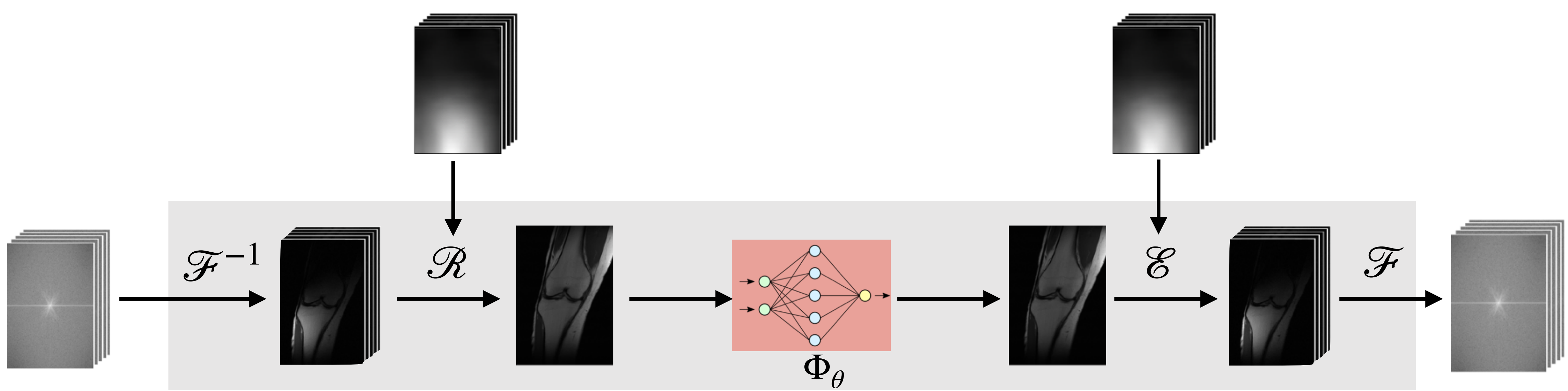


Sriram, A., Zbontar, J., Murrell, T., Defazio, A., Zitnick, C. L., Yakubova, N., ... & Johnson, P. (2020). End-to-end variational networks for accelerated MRI reconstruction. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 64-73). Springer, Cham.

E2E VarNet

- Unroll GD iterations in k-space

$$x^{t+1} = x^t - \mu^t \left[\mathscr{A}^* \left(\mathscr{A}(x^t) - y\right) + \Phi_\theta(x^t)\right] \xrightarrow{\mathscr{F}} k^{t+1} = k^t - \mu^t M(k^t - \tilde{k}) + \Psi(k^t)$$

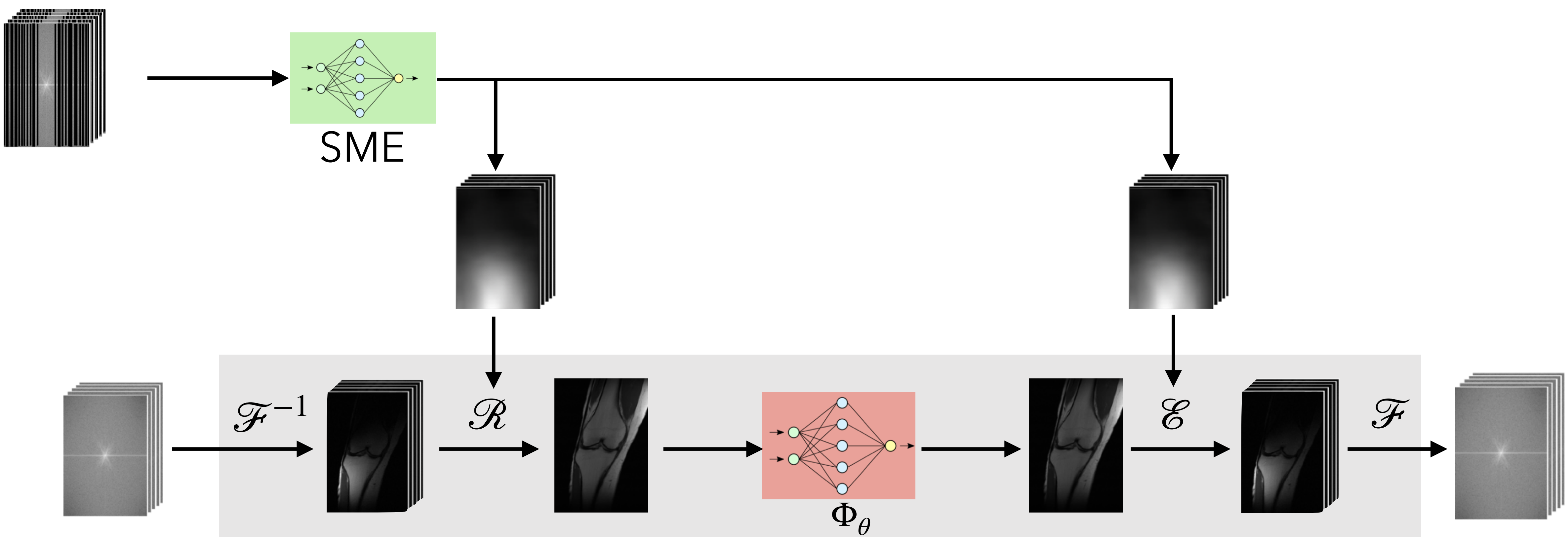


Sriram, A., Zbontar, J., Murrell, T., Defazio, A., Zitnick, C. L., Yakubova, N., ... & Johnson, P. (2020). End-to-end variational networks for accelerated MRI reconstruction. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 64-73). Springer, Cham.

E2E VarNet

- Unroll GD iterations in k-space

$$x^{t+1} = x^t - \mu^t \left[\mathscr{A}^* \left(\mathscr{A}(x^t) - y \right) + \Phi_\theta(x^t) \right] \xrightarrow{\mathscr{F}} k^{t+1} = k^t - \mu^t M(k^t - \tilde{k}) + \Psi(k^t)$$

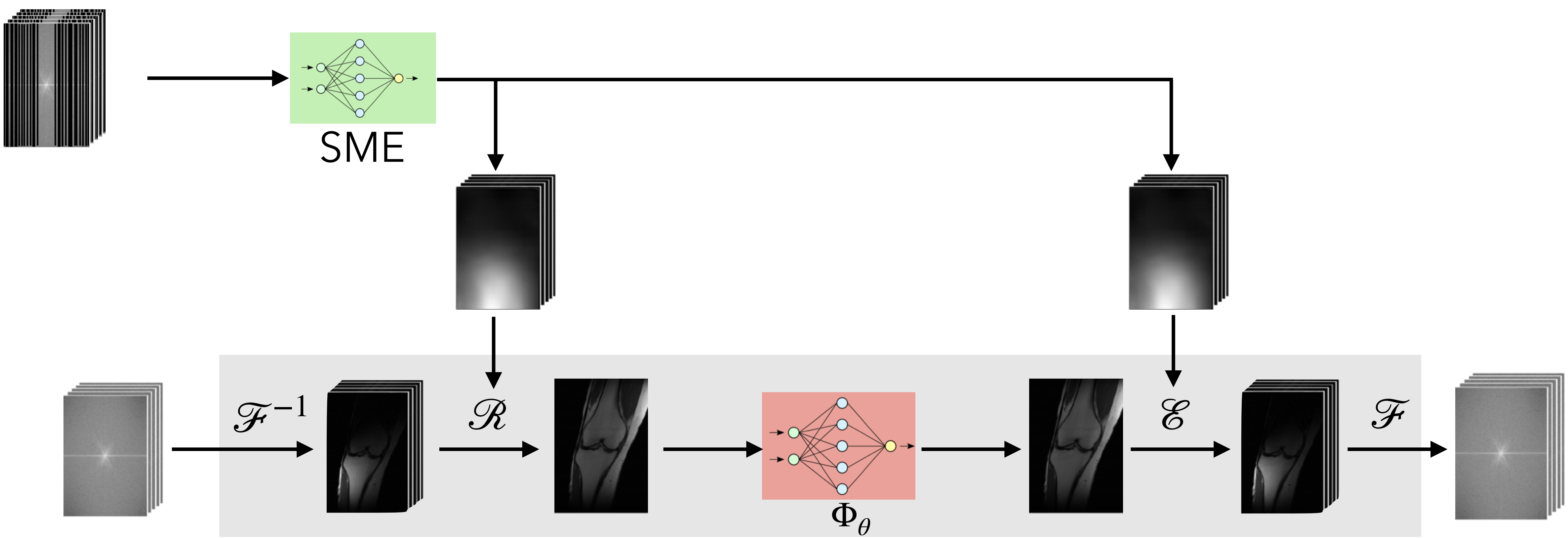


Sriram, A., Zbontar, J., Murrell, T., Defazio, A., Zitnick, C. L., Yakubova, N., ... & Johnson, P. (2020). End-to-end variational networks for accelerated MRI reconstruction. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 64-73). Springer, Cham.

E2E VarNet

- Unroll GD iterations in k-space

$$x^{t+1} = x^t - \mu^t \left[\mathscr{A}^* \left(\mathscr{A}(x^t) - y \right) + \Phi_\theta(x^t) \right] \xrightarrow{\mathscr{F}} k^{t+1} = k^t - \mu^t M(k^t - \tilde{k}) + \Psi(k^t)$$



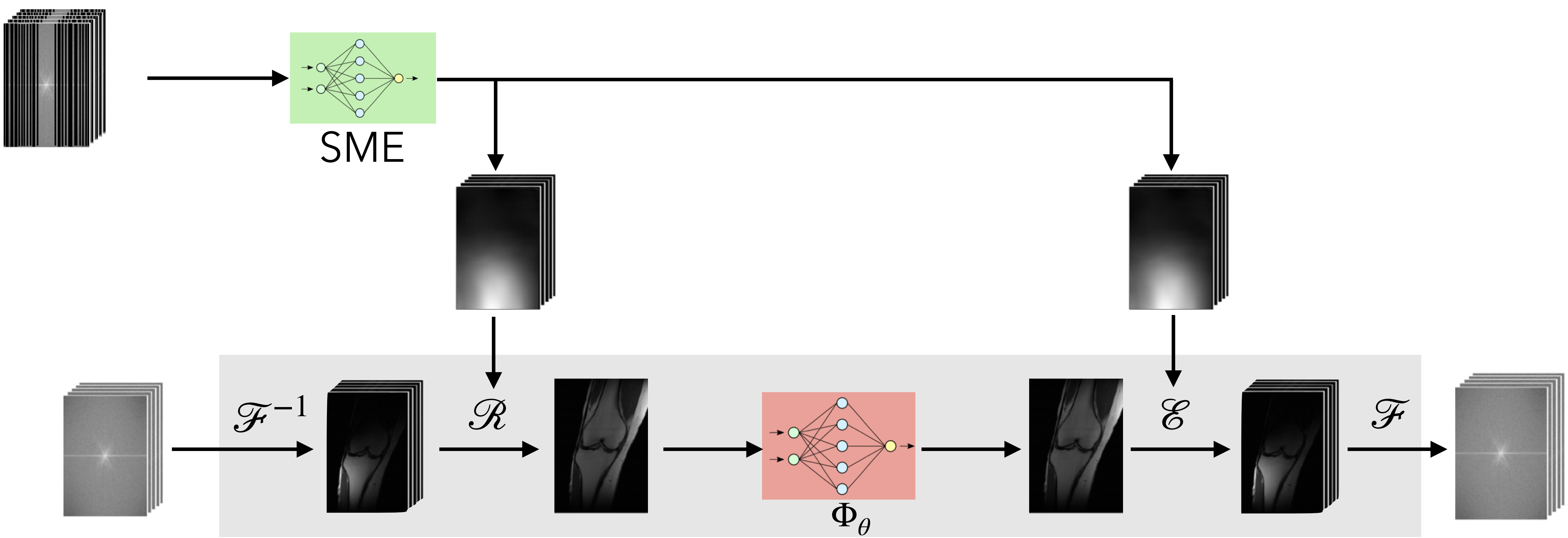
- Denoiser Φ_θ is a U-Net

Sriram, A., Zbontar, J., Murrell, T., Defazio, A., Zitnick, C. L., Yakubova, N., ... & Johnson, P. (2020). End-to-end variational networks for accelerated MRI reconstruction. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 64-73). Springer, Cham.

E2E VarNet

- Unroll GD iterations in k-space

$$x^{t+1} = x^t - \mu^t \left[\mathscr{A}^* \left(\mathscr{A}(x^t) - y\right) + \Phi_\theta(x^t)\right] \xrightarrow{\mathscr{F}} k^{t+1} = k^t - \mu^t M(k^t - \tilde{k}) + \Psi(k^t)$$



- Denoiser Φ_θ is a U-Net

Can we do better with modern architectures?

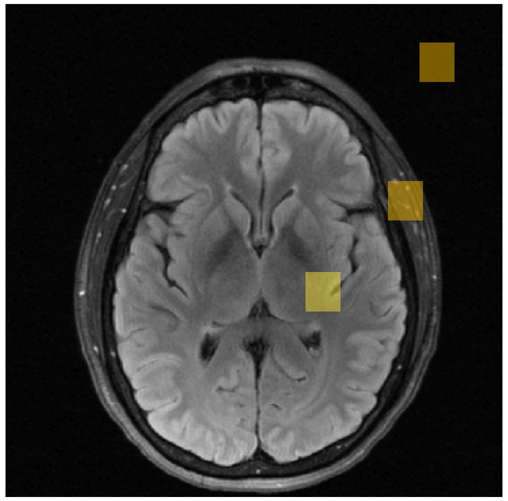
Sriram, A., Zbontar, J., Murrell, T., Defazio, A., Zitnick, C. L., Yakubova, N., ... & Johnson, P. (2020). End-to-end variational networks for accelerated MRI reconstruction. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 64-73). Springer, Cham.

Transformers?

- Benefits of Transformers

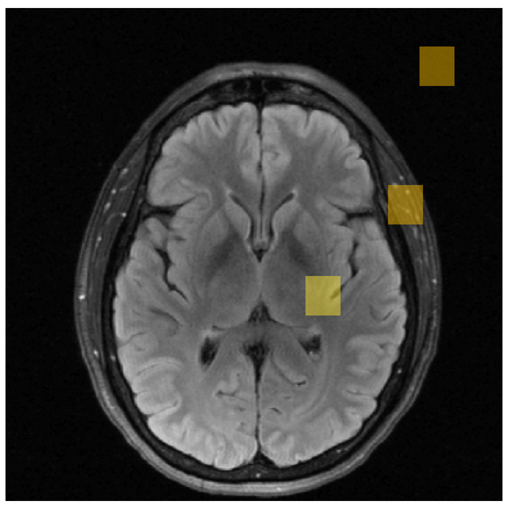
Transformers?

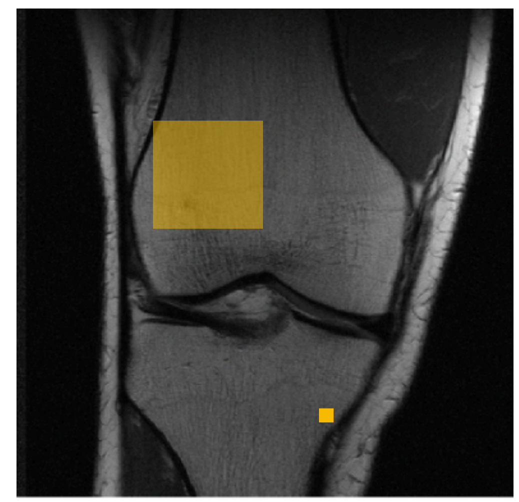
- Benefits of Transformers
  - conv kernels are content-independent



Transformers?

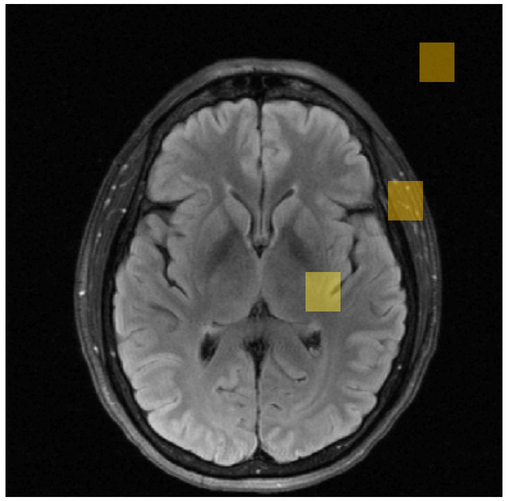
- Benefits of Transformers
  - conv kernels are content-independent
  - conv is not efficient for long-range dependency modelling

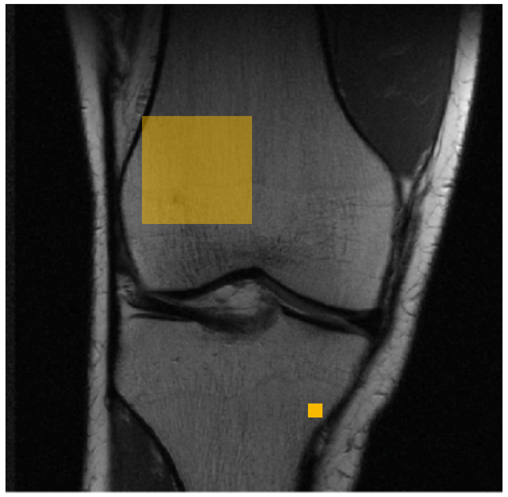




Transformers?

- Benefits of Transformers
  - conv kernels are content-independent
  - conv is not efficient for long-range dependency modelling

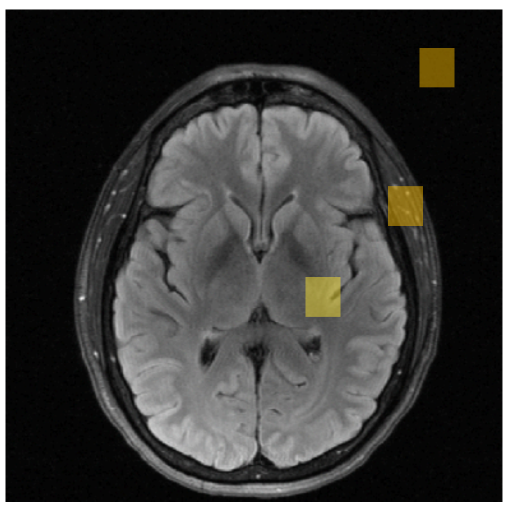


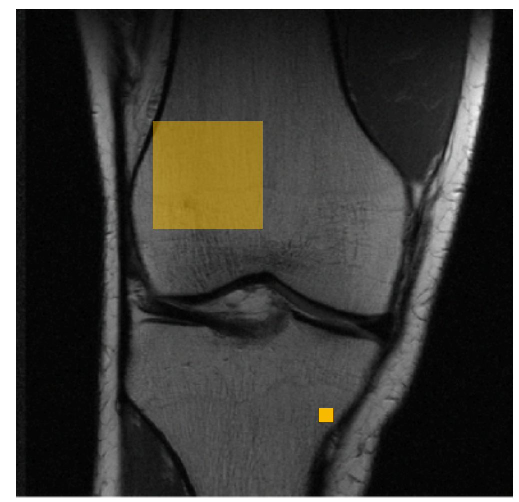


- Self-attention mechanism

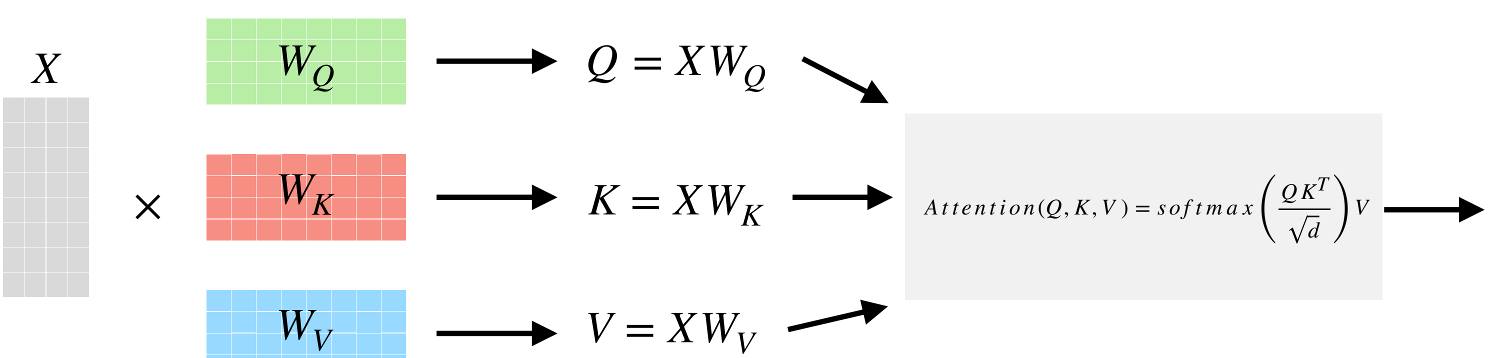
Transformers?

- Benefits of Transformers
  - conv kernels are content-independent
  - conv is not efficient for long-range dependency modelling



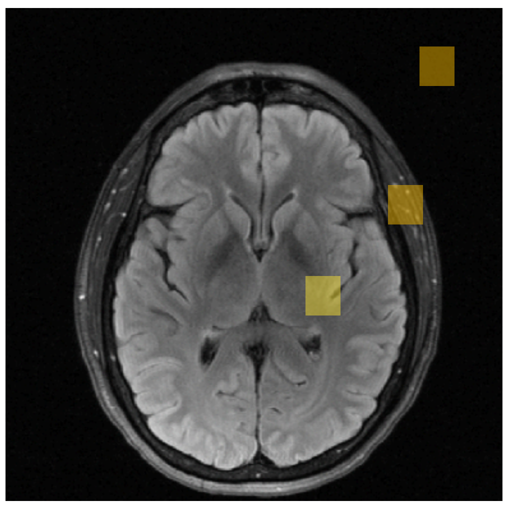


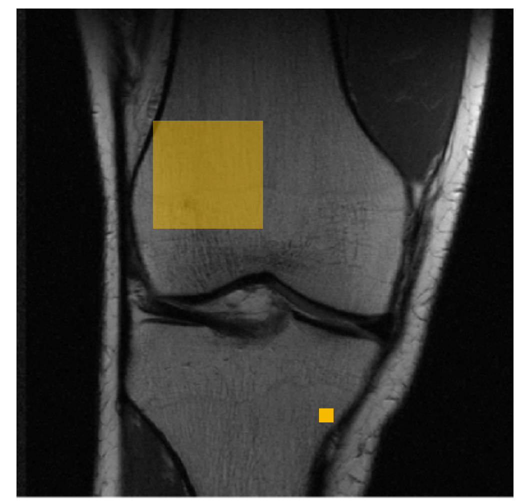
- Self-attention mechanism



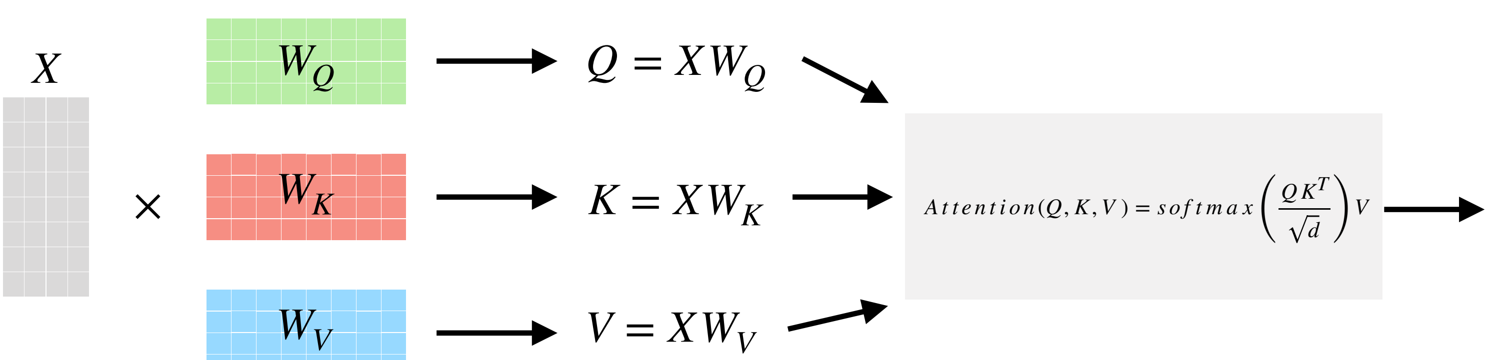
Transformers?

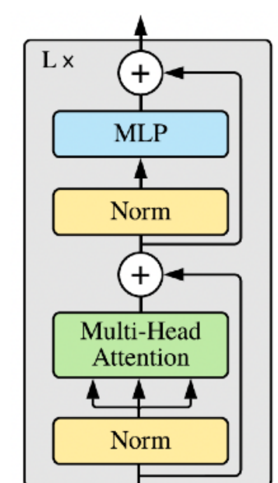
- Benefits of Transformers
  - conv kernels are content-independent
  - conv is not efficient for long-range dependency modelling





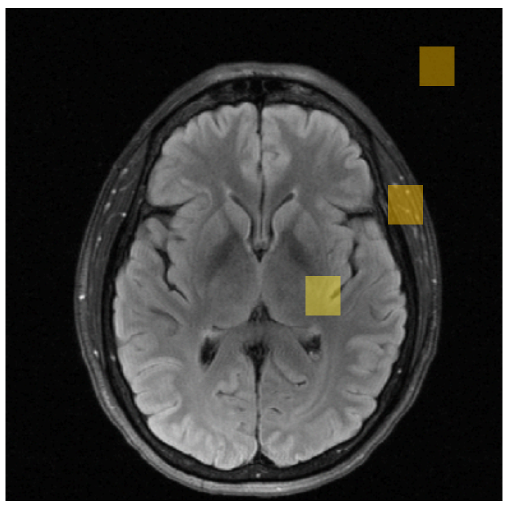
- Self-attention mechanism

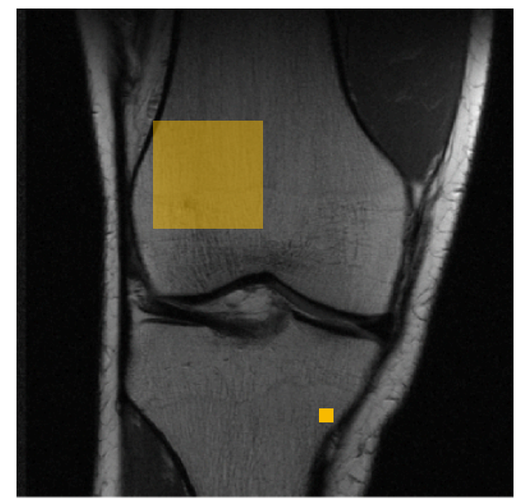




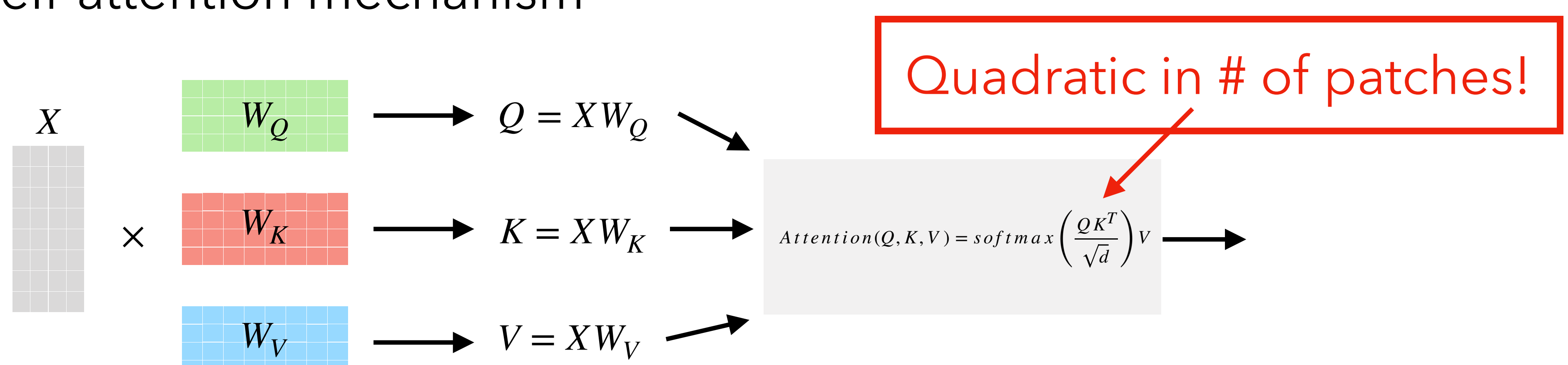
Transformers?

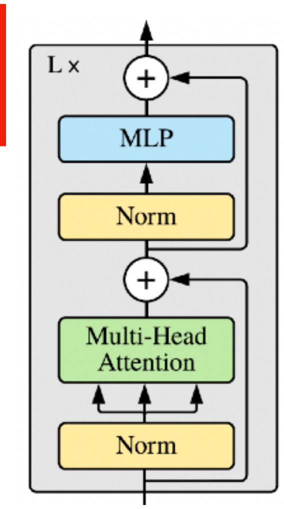
- Benefits of Transformers
  - conv kernels are content-independent
  - conv is not efficient for long-range dependency modelling





- Self-attention mechanism



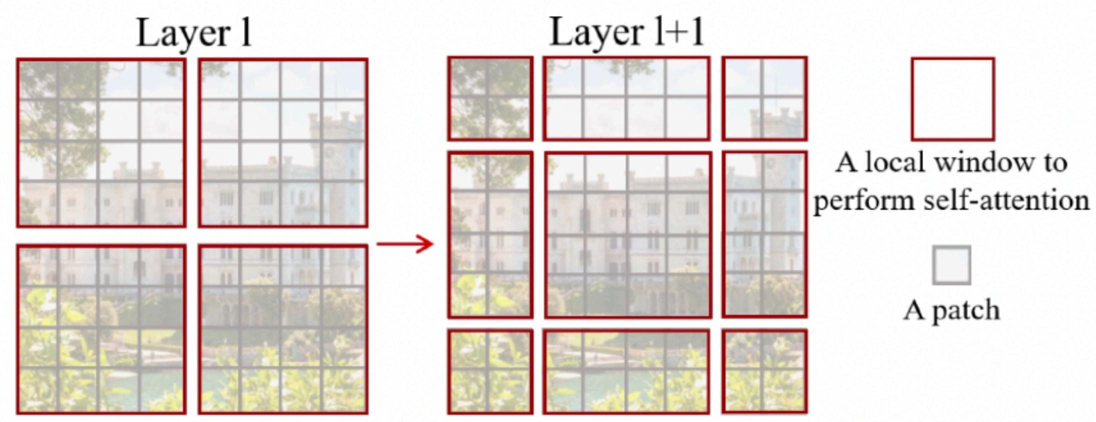


Dealing with quadratic scaling

Liu, Ze, et al. "Swin transformer: Hierarchical vision transformer using shifted windows." *arXiv preprint arXiv:2103.14030* (2021).
Liang, Jingyun, et al. "Swinir: Image restoration using swin transformer." *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2021.

Dealing with quadratic scaling

- Local window attention

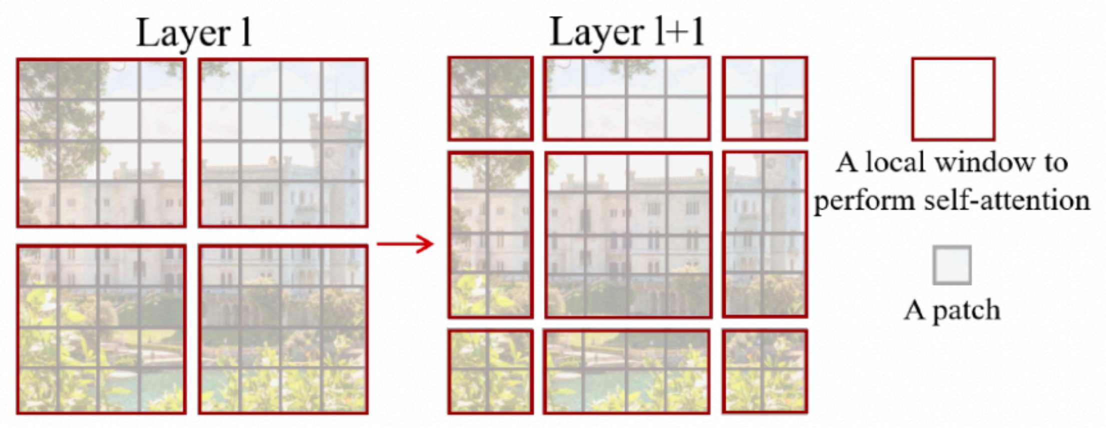


Liu, Ze, et al. "Swin transformer: Hierarchical vision transformer using shifted windows." *arXiv preprint arXiv:2103.14030* (2021).

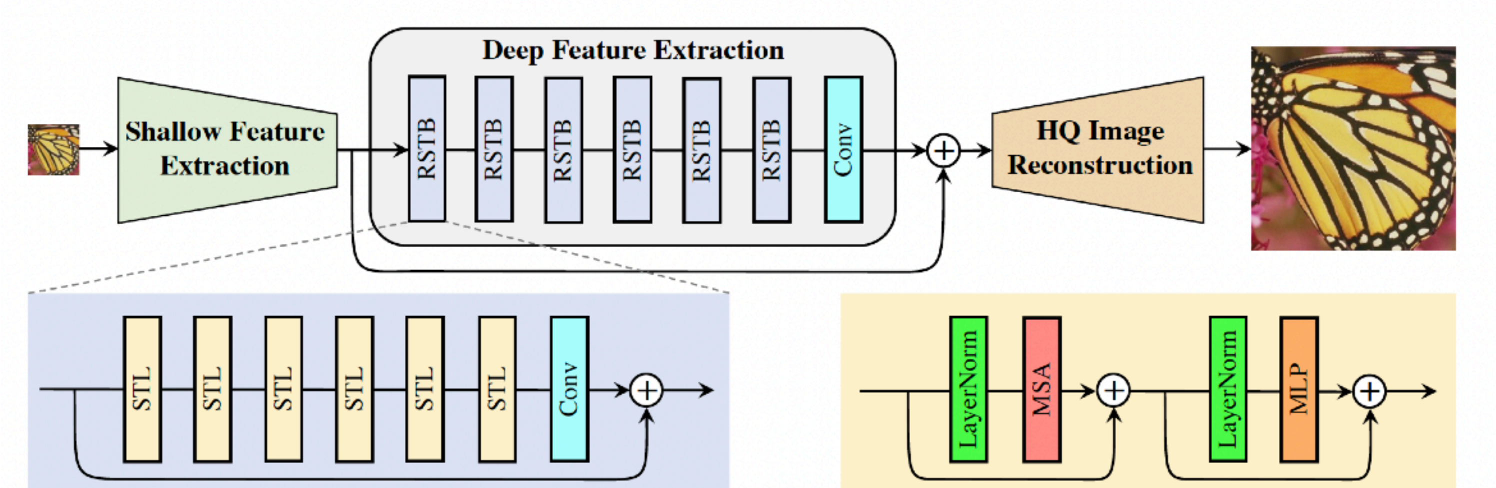
Liang, Jingyun, et al. "Swinir: Image restoration using swin transformer." *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2021.

Dealing with quadratic scaling

- Local window attention



- Swin Transformer for image restoration and super-resolution



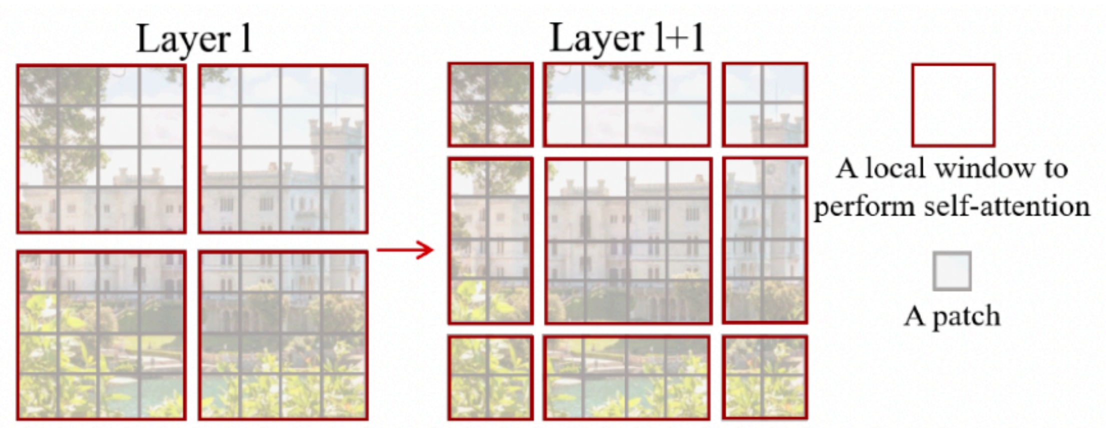
(a) Residual Swin Transformer Block (RSTB)
(b) Swin Transformer Layer (STL)

Liu, Ze, et al. "Swin transformer: Hierarchical vision transformer using shifted windows." *arXiv preprint arXiv:2103.14030* (2021).

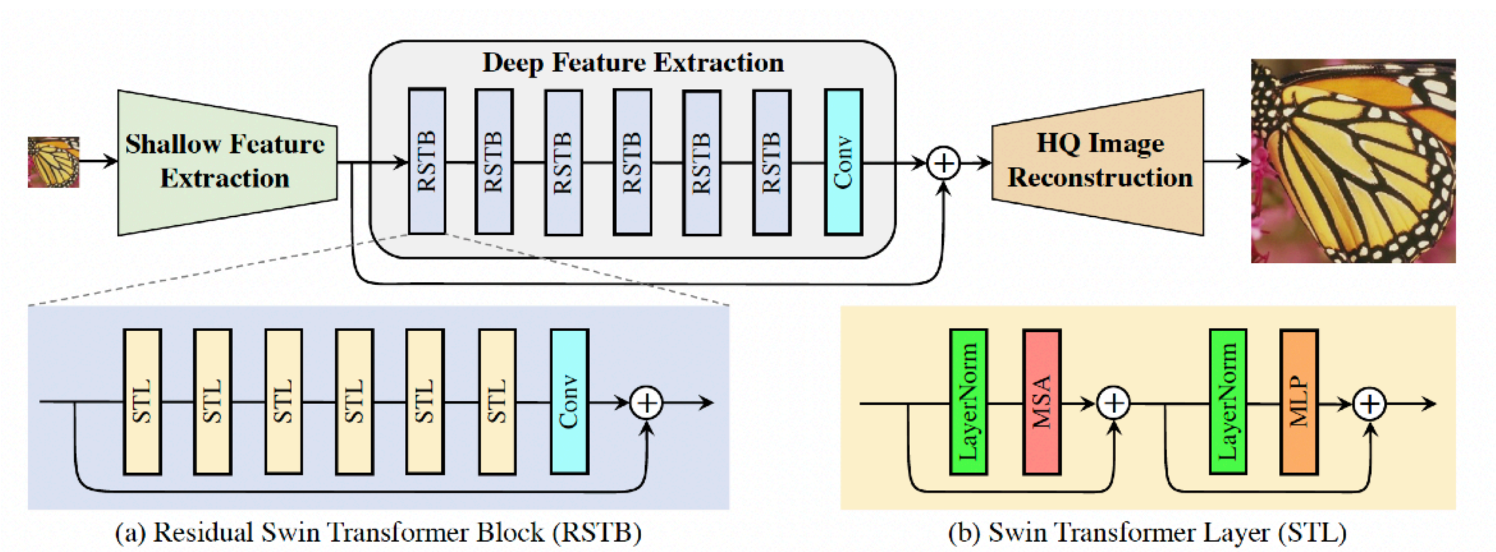
Liang, Jingyun, et al. "Swinir: Image restoration using swin transformer." *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2021.

Dealing with quadratic scaling

- Local window attention



- Swin Transformer for image restoration and super-resolution



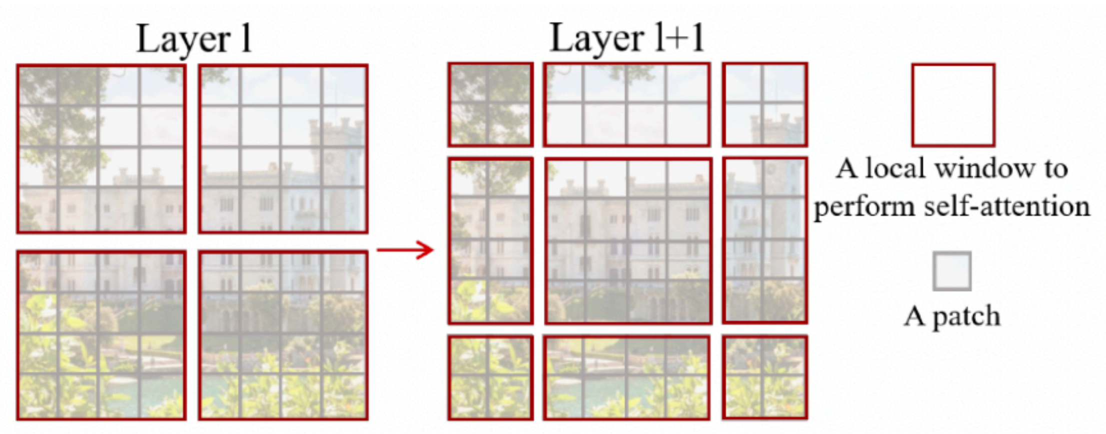
- Transformer-Conv hybrid!

Liu, Ze, et al. "Swin transformer: Hierarchical vision transformer using shifted windows." *arXiv preprint arXiv:2103.14030* (2021).

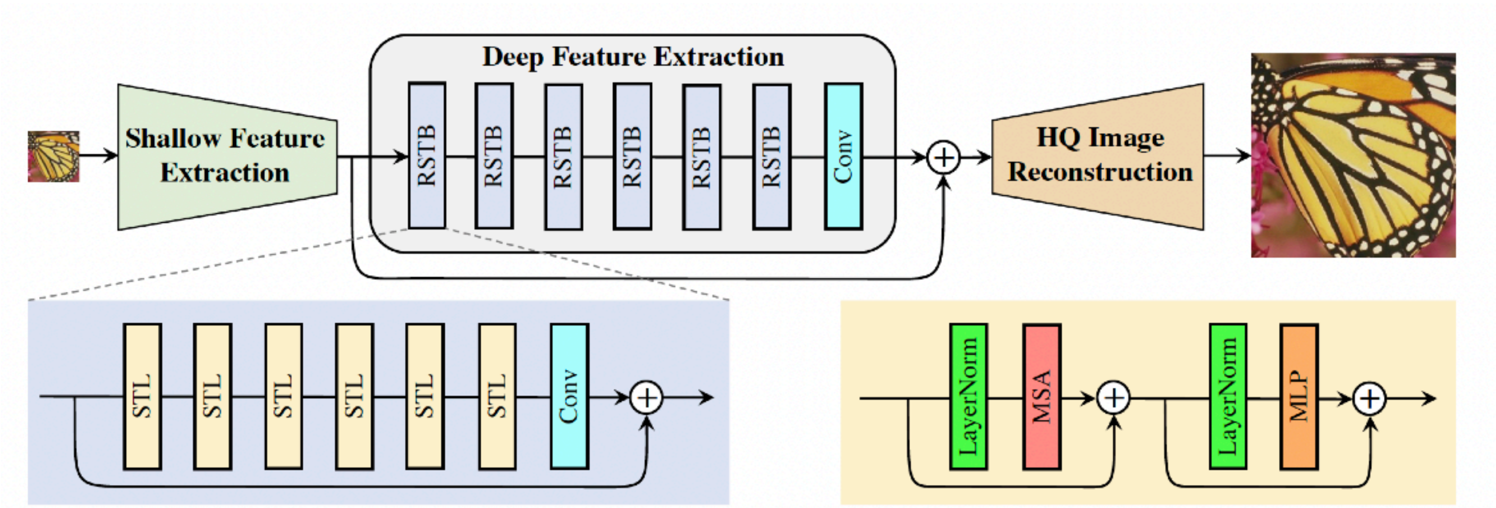
Liang, Jingyun, et al. "Swinir: Image restoration using swin transformer." *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2021.

Dealing with quadratic scaling

- Local window attention



- Swin Transformer for image restoration and super-resolution



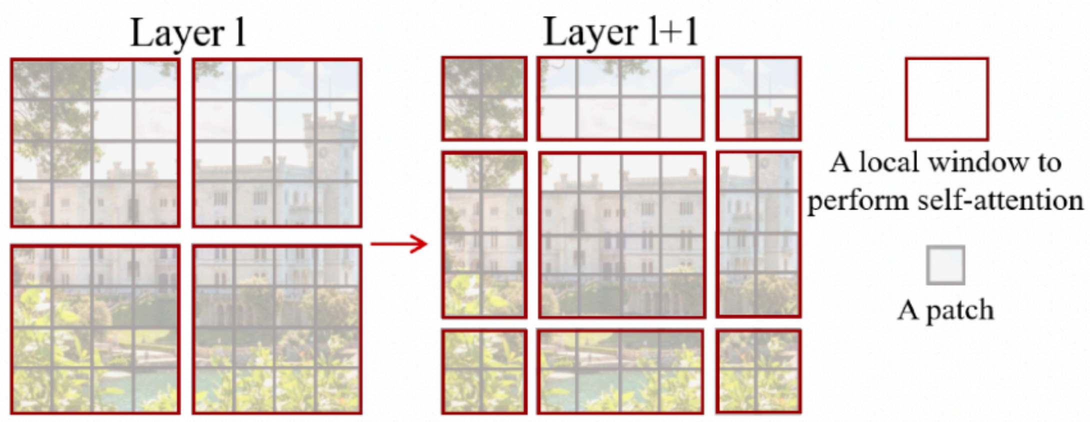
(a) Residual Swin Transformer Block (RSTB)
(b) Swin Transformer Layer (STL)

- Transformer-Conv hybrid!
- Long-range dependencies via SA

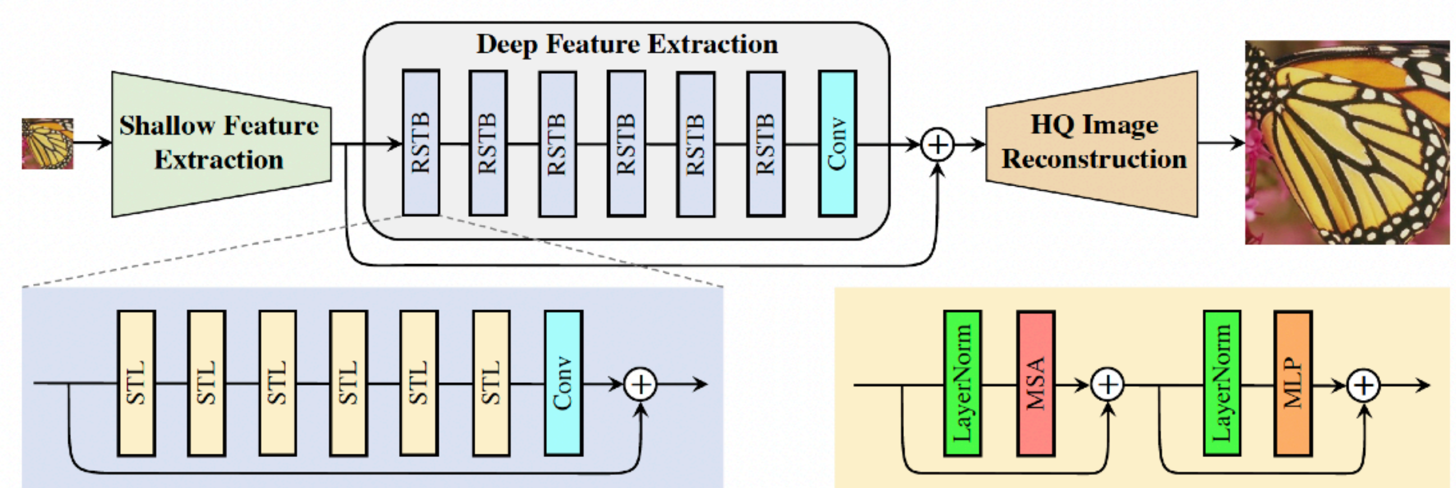
Liu, Ze, et al. "Swin transformer: Hierarchical vision transformer using shifted windows." *arXiv preprint arXiv:2103.14030* (2021).
Liang, Jingyun, et al. "Swinir: Image restoration using swin transformer." *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2021.

Dealing with quadratic scaling

- Local window attention



- Swin Transformer for image restoration and super-resolution



(a) Residual Swin Transformer Block (RSTB)

(b) Swin Transformer Layer (STL)

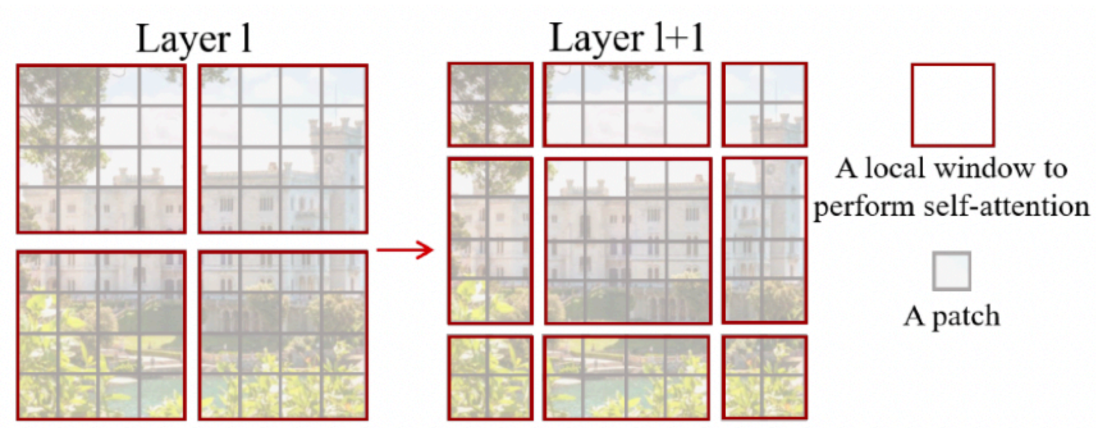
- Transformer-Conv hybrid!
- Long-range dependencies via SA
- Implicit bias via Conv

Liu, Ze, et al. "Swin transformer: Hierarchical vision transformer using shifted windows." *arXiv preprint arXiv:2103.14030* (2021).

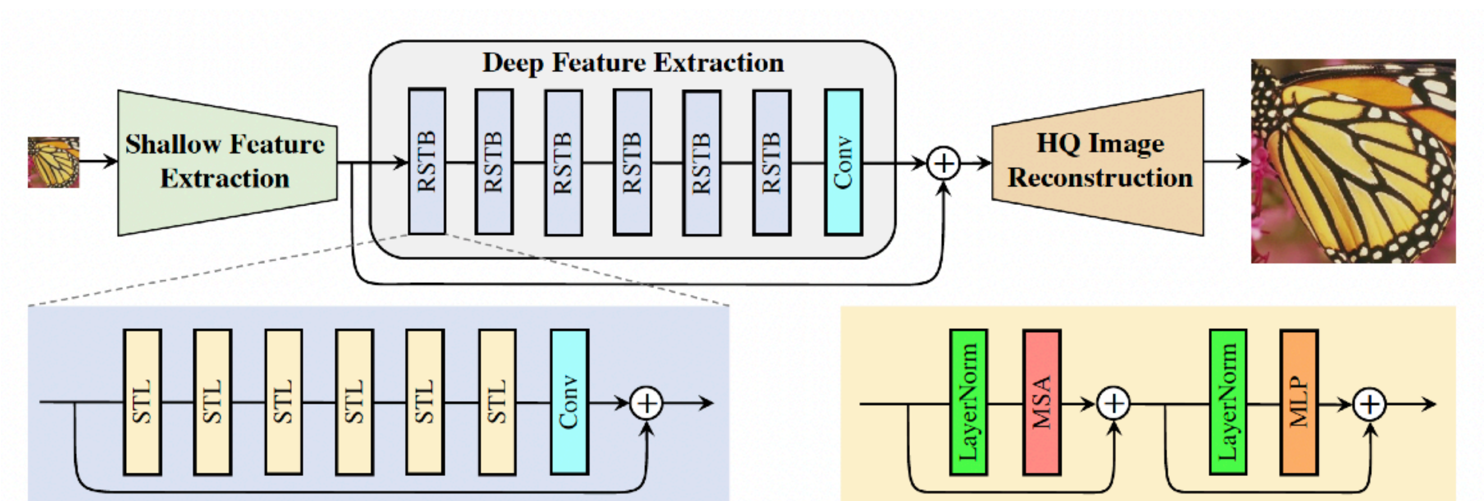
Liang, Jingyun, et al. "Swinir: Image restoration using swin transformer." *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2021.

Dealing with quadratic scaling

- Local window attention



- Swin Transformer for image restoration and super-resolution



(a) Residual Swin Transformer Block (RSTB)
(b) Swin Transformer Layer (STL)

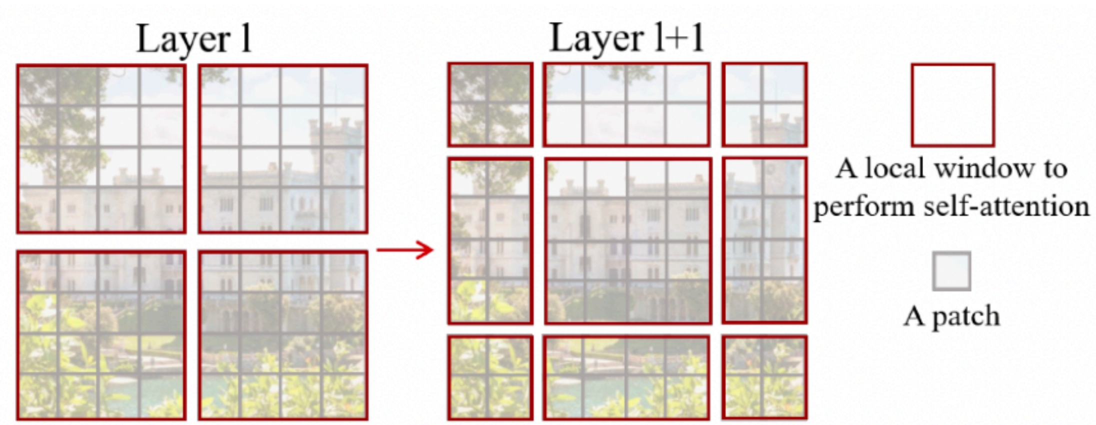
- Transformer-Conv hybrid!
- Long-range dependencies via SA
- Implicit bias via Conv
- Single-scale processing

Model	GPU mem.	mins/epoch	Val. SSIM

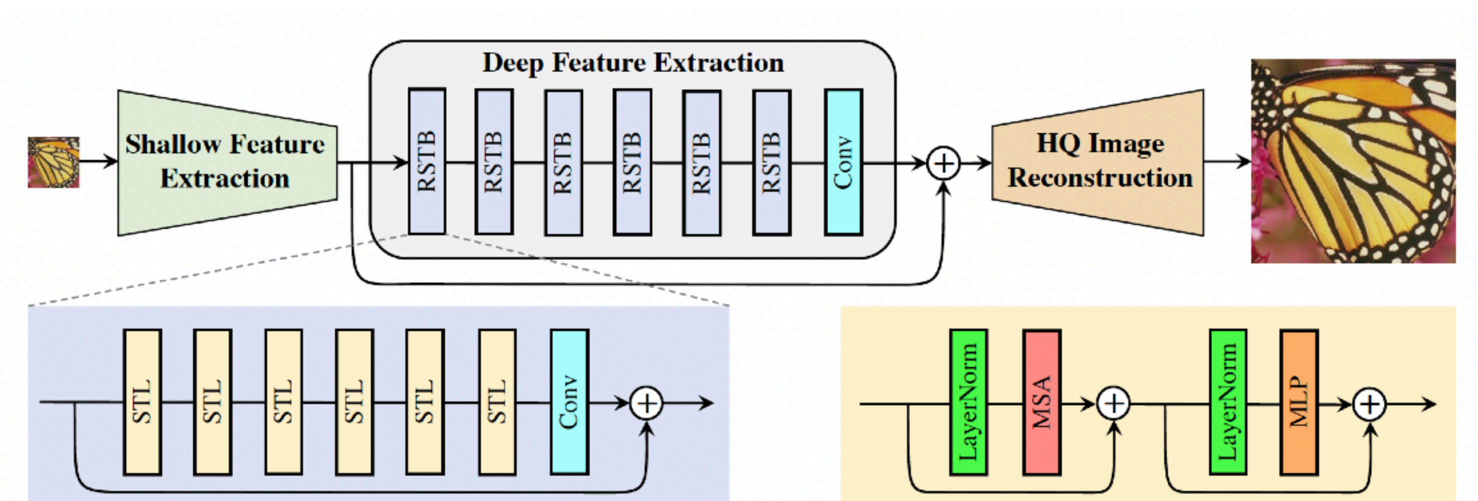
Liu, Ze, et al. "Swin transformer: Hierarchical vision transformer using shifted windows." *arXiv preprint arXiv:2103.14030* (2021).
Liang, Jingyun, et al. "Swinir: Image restoration using swin transformer." *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2021.

Dealing with quadratic scaling

- Local window attention



- Swin Transformer for image restoration and super-resolution



(a) Residual Swin Transformer Block (RSTB)
(b) Swin Transformer Layer (STL)

- Transformer-Conv hybrid!
- Long-range dependencies via SA
- Implicit bias via Conv
- Single-scale processing

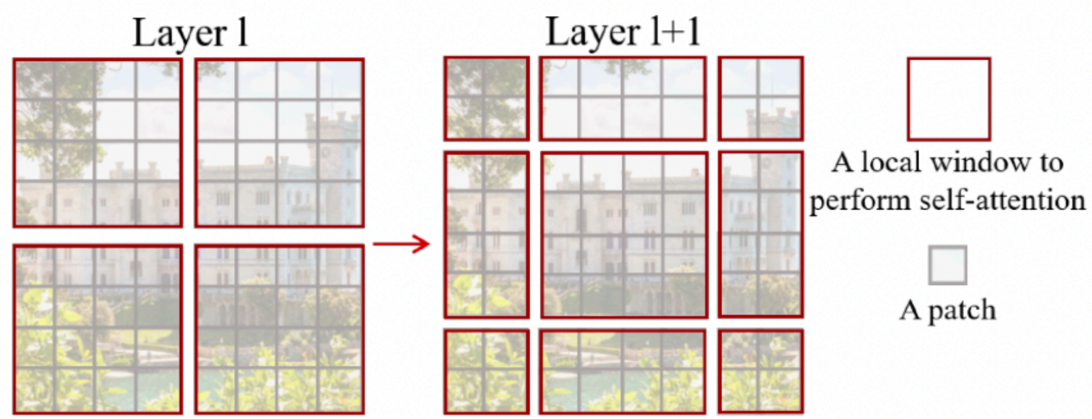
Model	GPU mem.	mins/epoch	Val. SSIM
E2E-VarNet	≈16GB	9	**0.9313**

Liu, Ze, et al. "Swin transformer: Hierarchical vision transformer using shifted windows." *arXiv preprint arXiv:2103.14030* (2021).

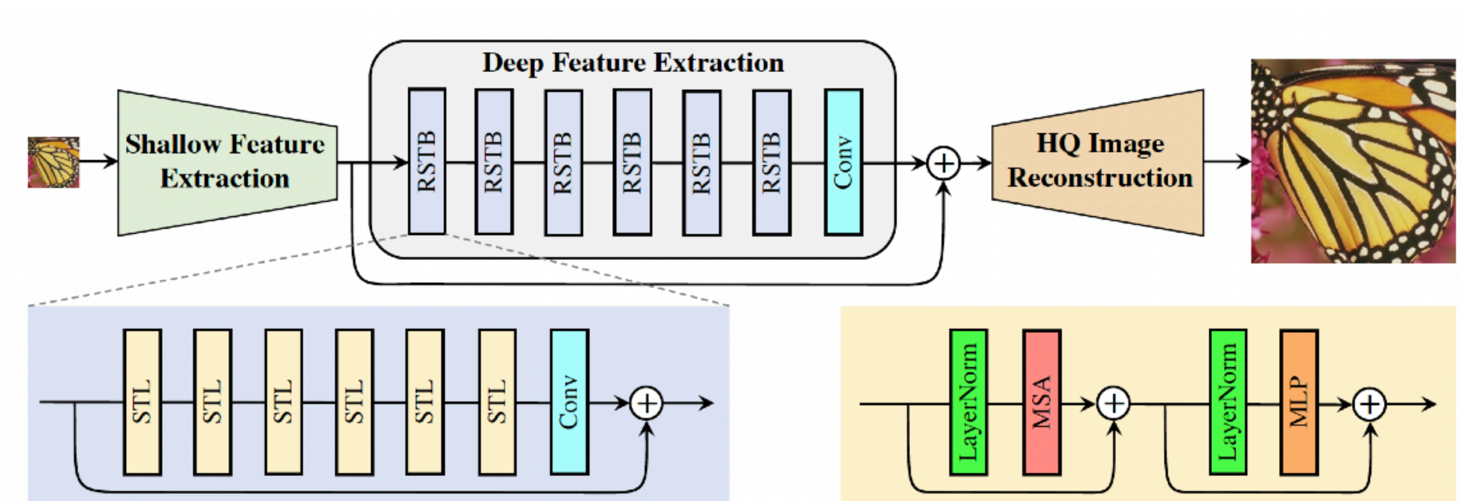
Liang, Jingyun, et al. "Swinir: Image restoration using swin transformer." *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2021.

Dealing with quadratic scaling

- Local window attention



- Swin Transformer for image restoration and super-resolution



(a) Residual Swin Transformer Block (RSTB)
(b) Swin Transformer Layer (STL)

- Transformer-Conv hybrid!
- Long-range dependencies via SA
- Implicit bias via Conv
- Single-scale processing

Model	GPU mem.	mins/epoch	Val. SSIM
E2E-VarNet	≈16GB	9	**0.9313**
Unrolled SwinIR	≈16GB	72	0.9216

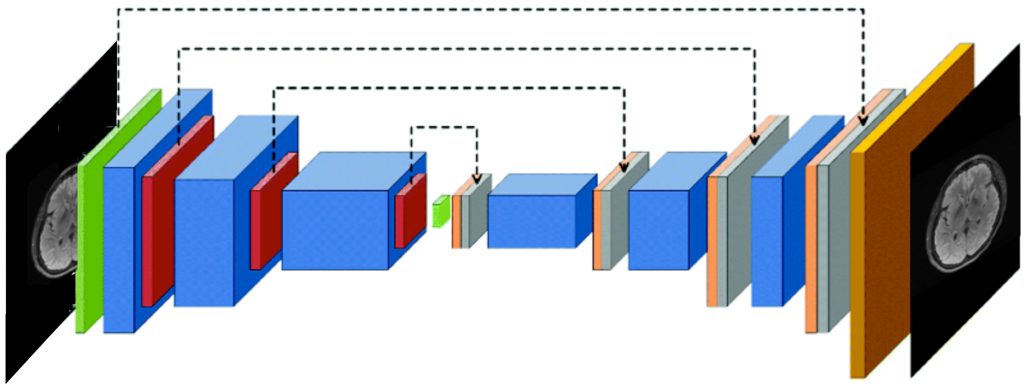
Liu, Ze, et al. "Swin transformer: Hierarchical vision transformer using shifted windows." *arXiv preprint arXiv:2103.14030* (2021).
Liang, Jingyun, et al. "Swinir: Image restoration using swin transformer." *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2021.

Multi-scale SwinIR

- Missing component: hierarchical, multi-scale representations

Multi-scale SwinIR

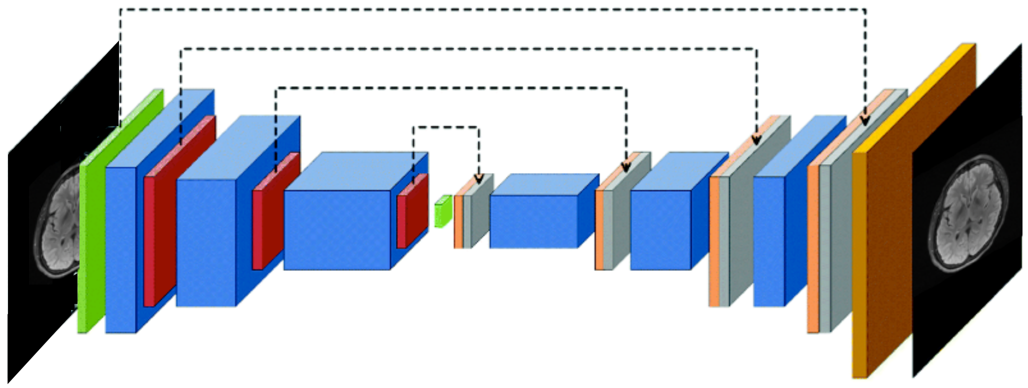
- Missing component: hierarchical, multi-scale representations



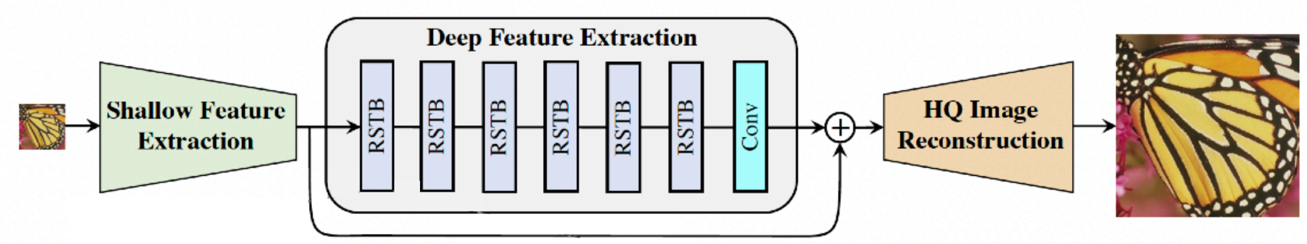
U-Net: multi-scale

Multi-scale SwinIR

- Missing component: hierarchical, multi-scale representations



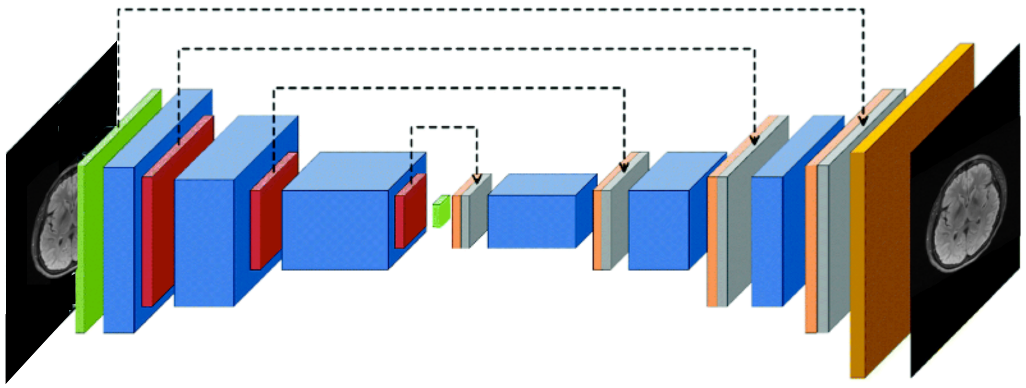
U-Net: multi-scale

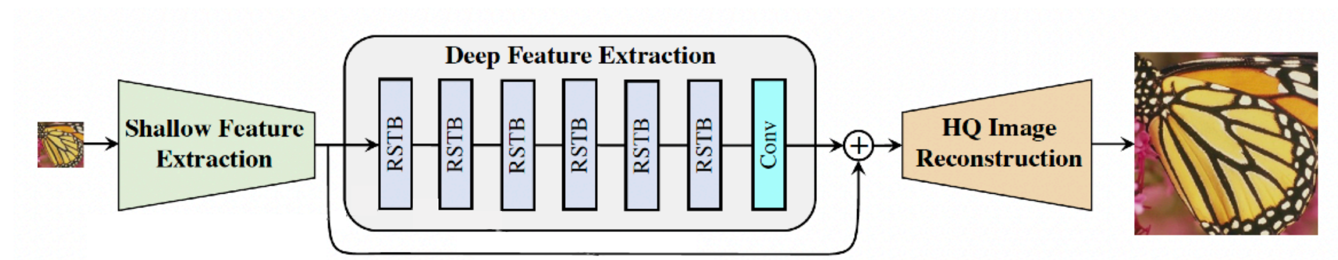


SwinIR: Transformer+Conv

Multi-scale SwinIR

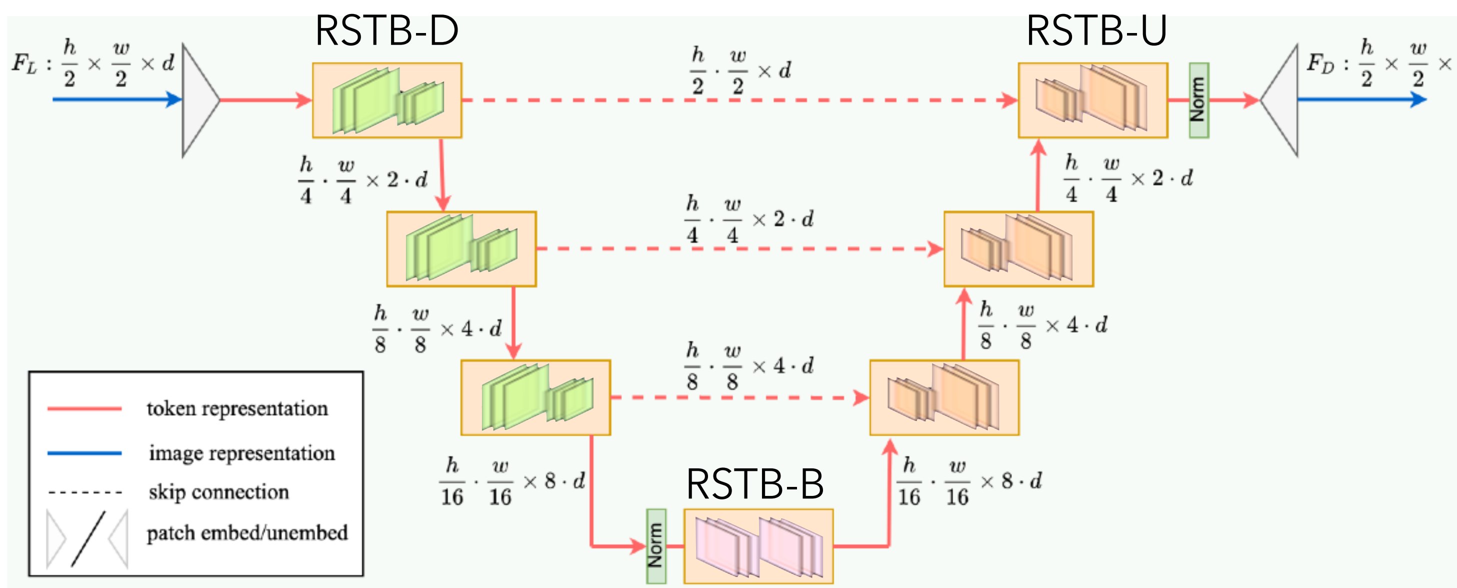
- Missing component: hierarchical, multi-scale representations





U-Net: multi-scale

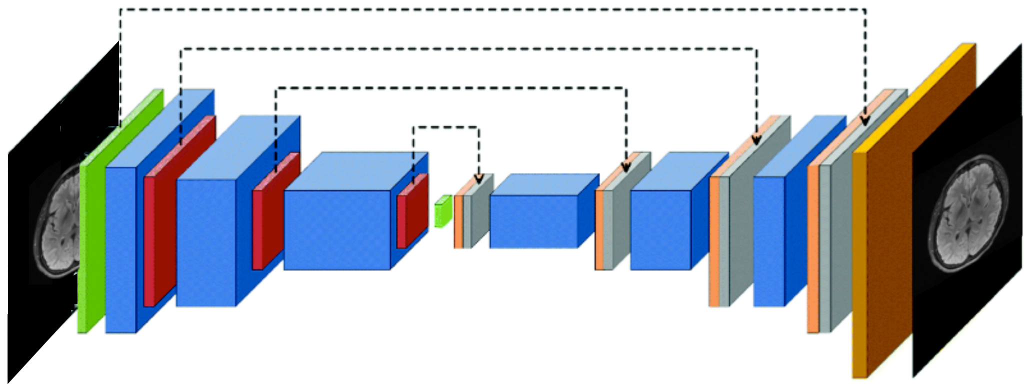
SwinIR: Transformer+Conv

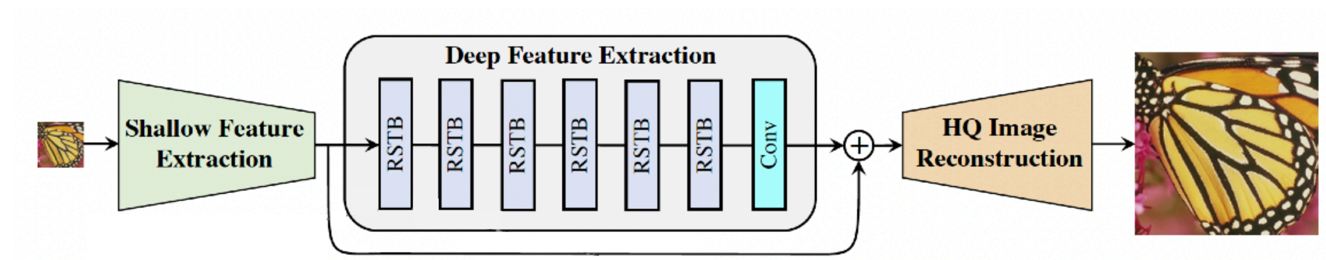


Multi-scale Hybrid Feature Extractor

Multi-scale SwinIR

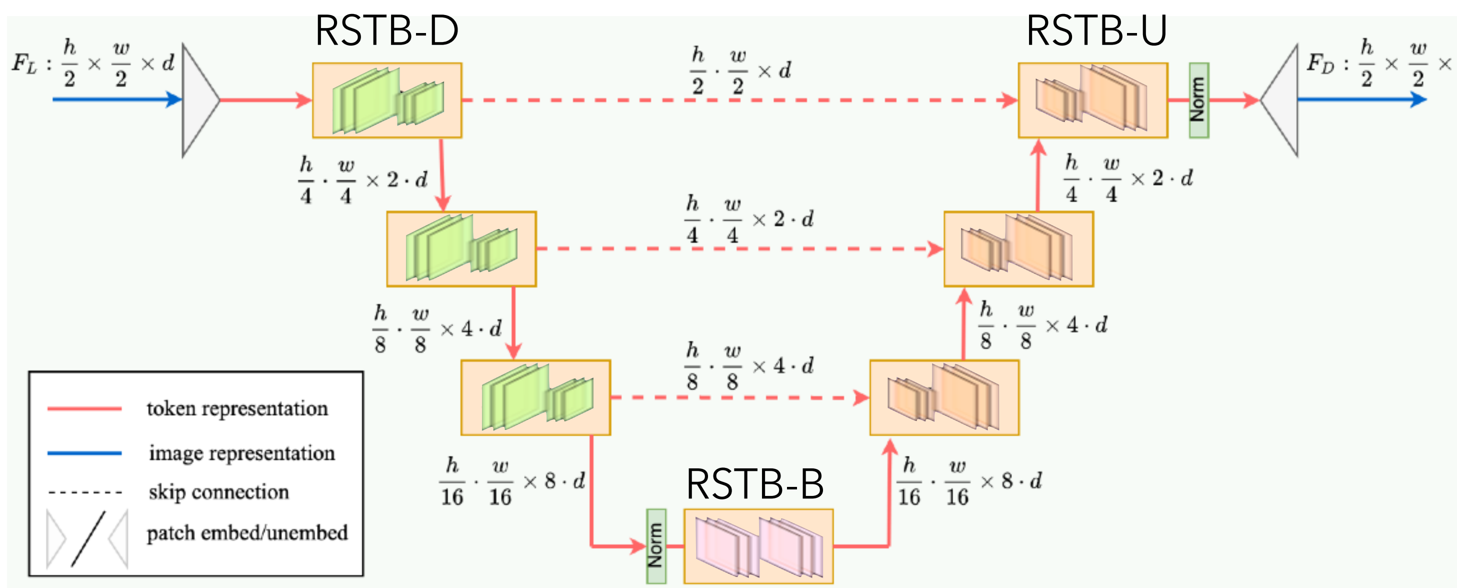
- Missing component: hierarchical, multi-scale representations





U-Net: multi-scale

SwinIR: Transformer+Conv

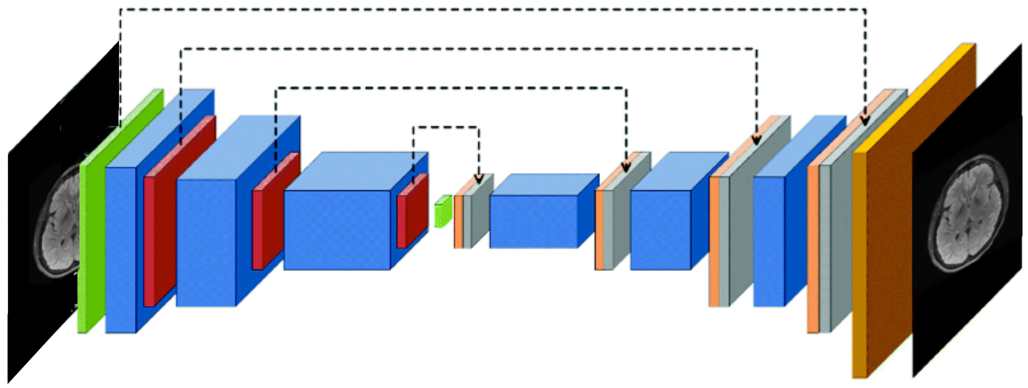


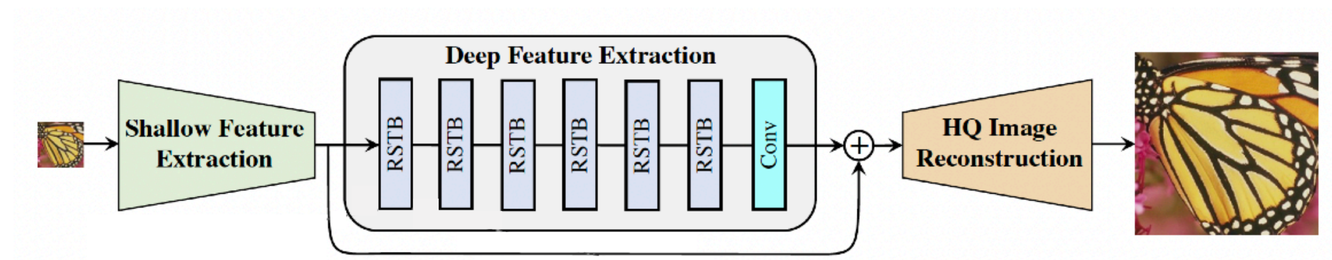
Multi-scale Hybrid Feature Extractor

Model	GPU mem.	mins/epoch	Val. SSIM
E2E-VarNet	≈16GB	9	**0.9313**
Unrolled SwinIR	≈16GB	72	0.9216

Multi-scale SwinIR

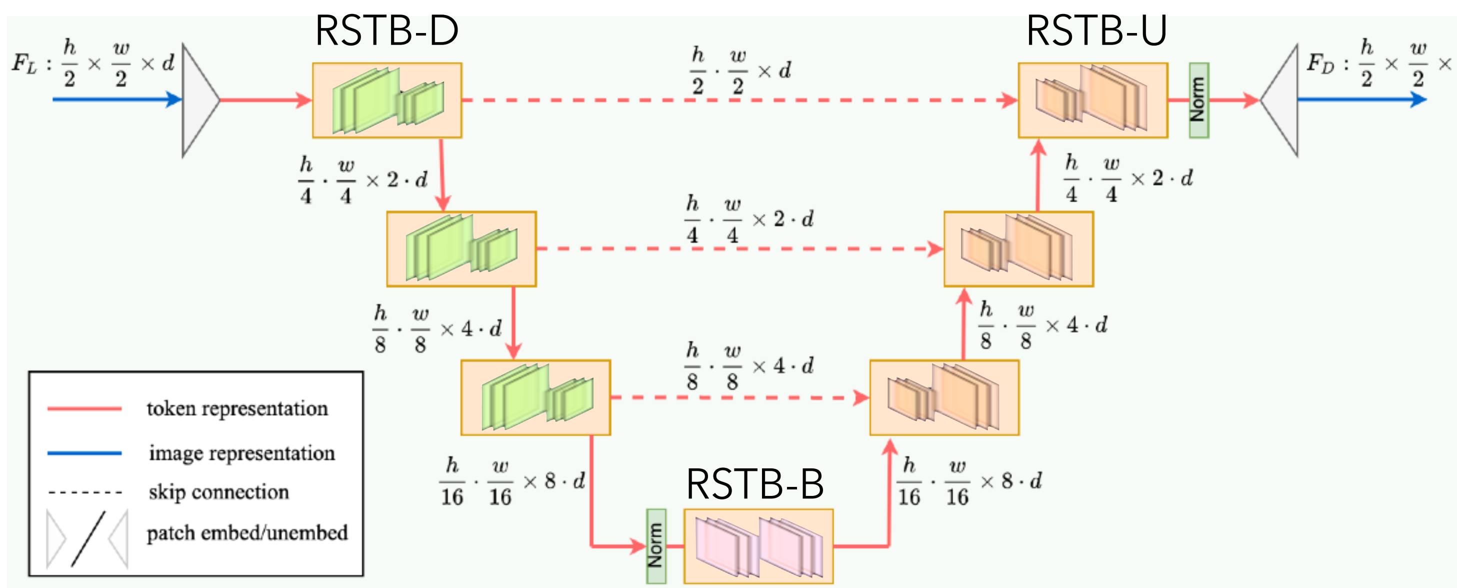
- Missing component: hierarchical, multi-scale representations





U-Net: multi-scale

SwinIR: Transformer+Conv



Multi-scale Hybrid Feature Extractor

Model	GPU mem.	mins/epoch	Val. SSIM
E2E-VarNet	≈16GB	9	**0.9313**
Unrolled SwinIR	≈16GB	72	0.9216
Unrolled, multi-scale SwinIR	≈16GB	66	0.9311

High-resolution challenge

- Key challenge: high-resolution input images **AND** dense prediction task

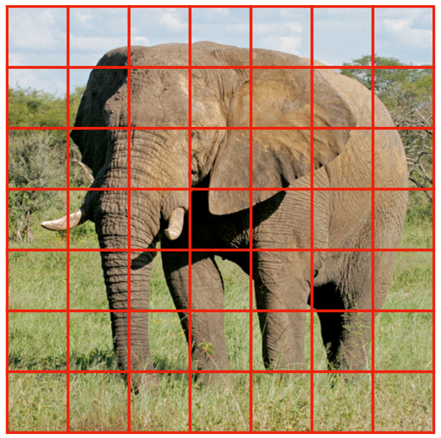
High-resolution challenge

- <u>Key challenge:</u> high-resolution input images **AND** dense prediction task

ImageNet
224×224

High-resolution challenge

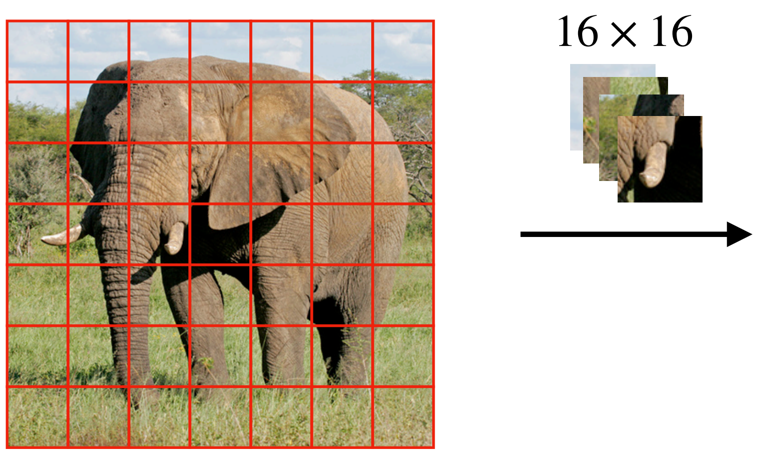
- <u>Key challenge:</u> high-resolution input images **AND** dense prediction task



ImageNet
224×224

High-resolution challenge

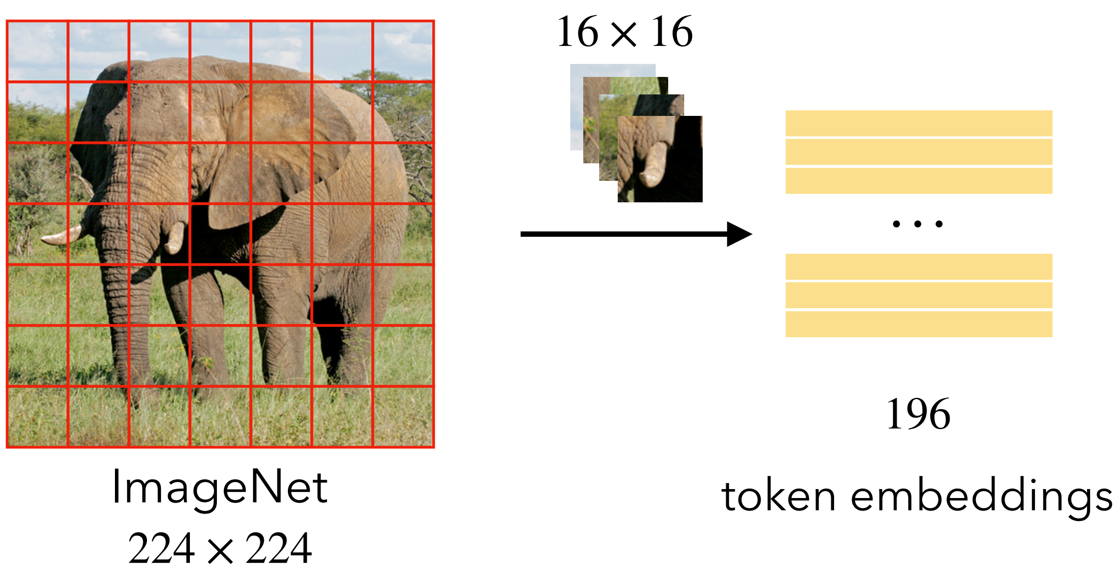
- <u>Key challenge:</u> high-resolution input images **AND** dense prediction task



ImageNet
224×224

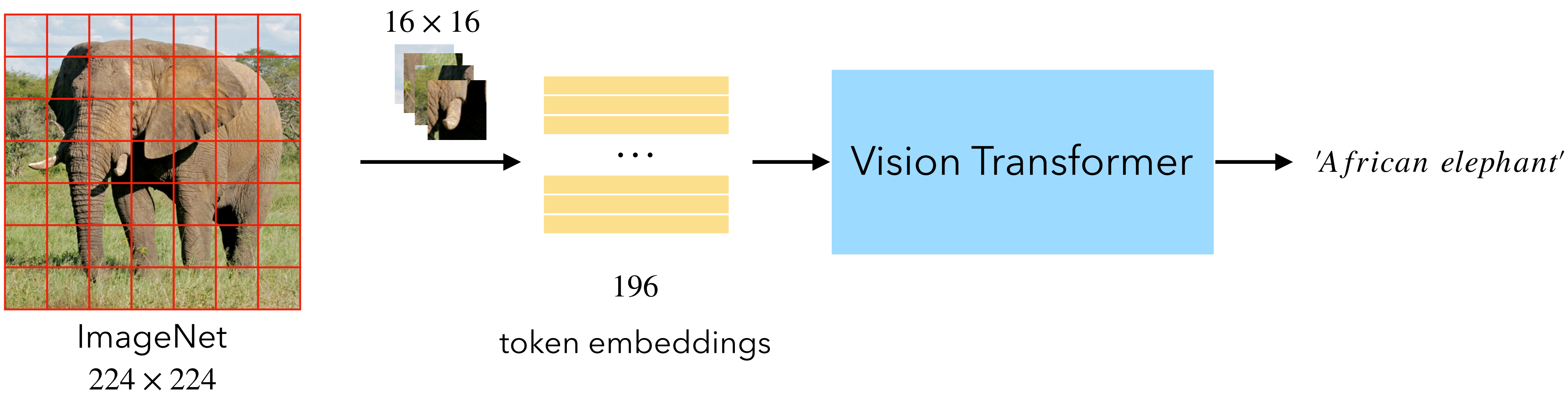
High-resolution challenge

- <u>Key challenge:</u> high-resolution input images **AND** dense prediction task



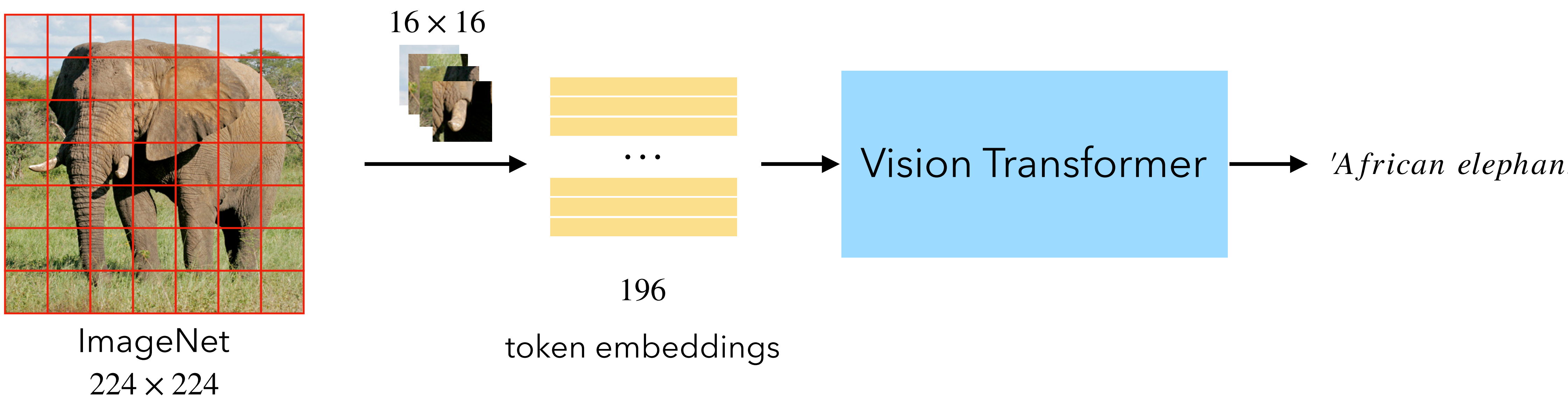
High-resolution challenge

- <u>Key challenge:</u> high-resolution input images **AND** dense prediction task



High-resolution challenge

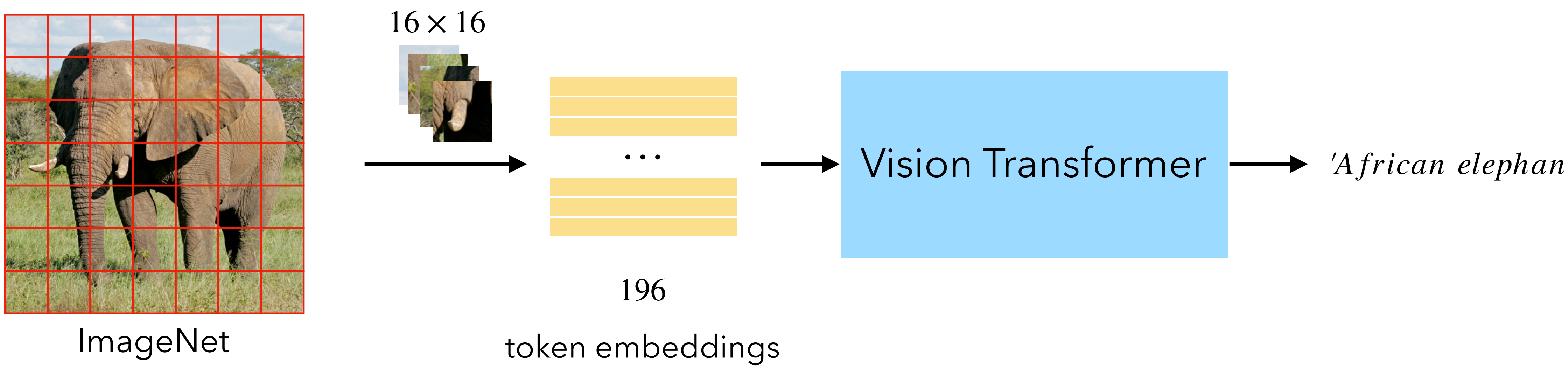
- Key challenge: high-resolution input images **AND** dense prediction task



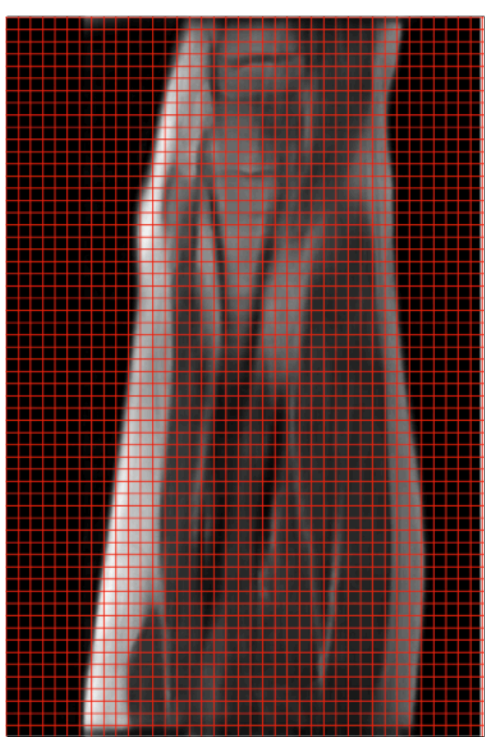
Noisy MR image
640 × 368

High-resolution challenge

- <u>Key challenge:</u> high-resolution input images **AND** dense prediction task



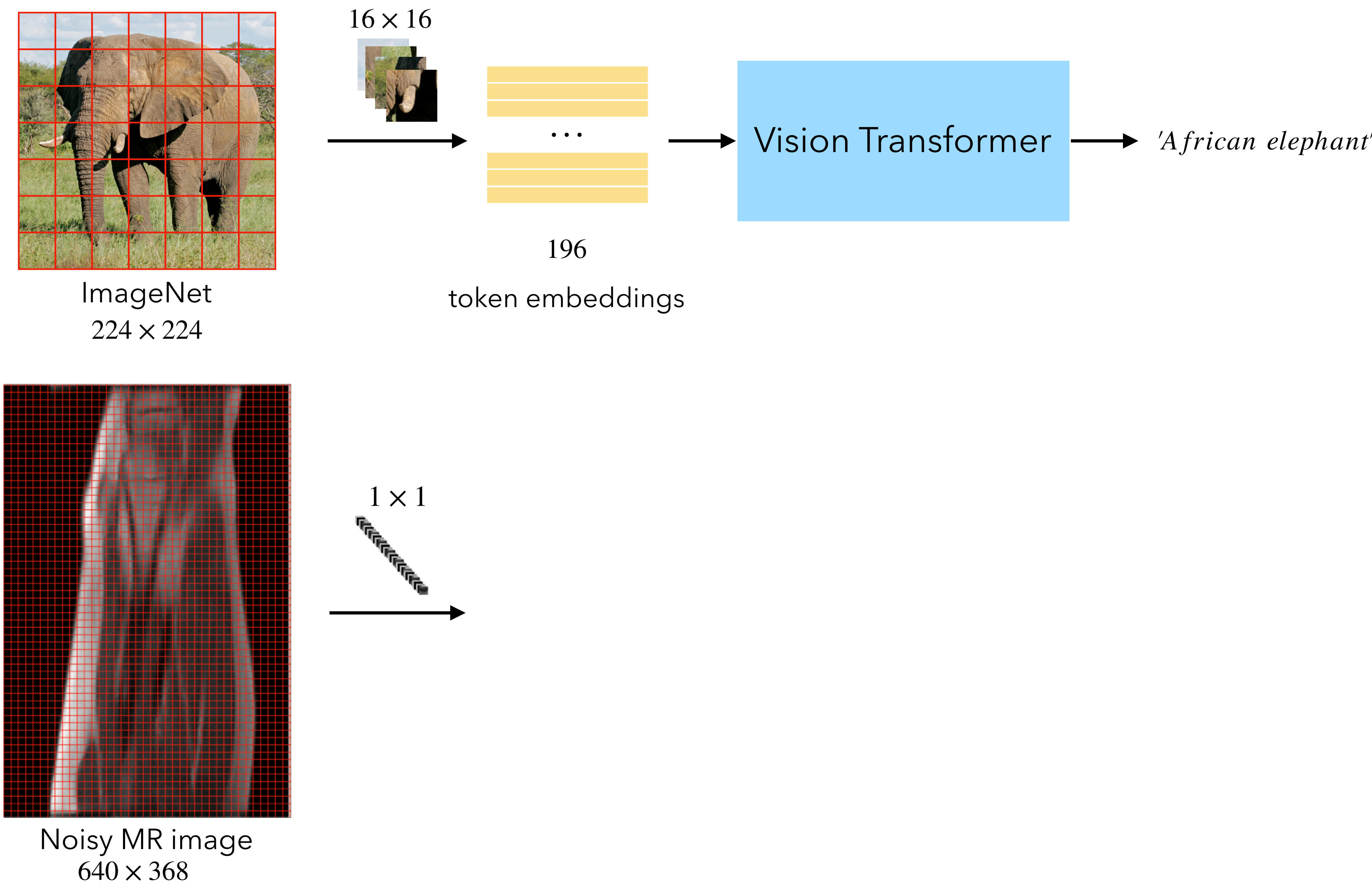
ImageNet
224×224



Noisy MR image
640×368

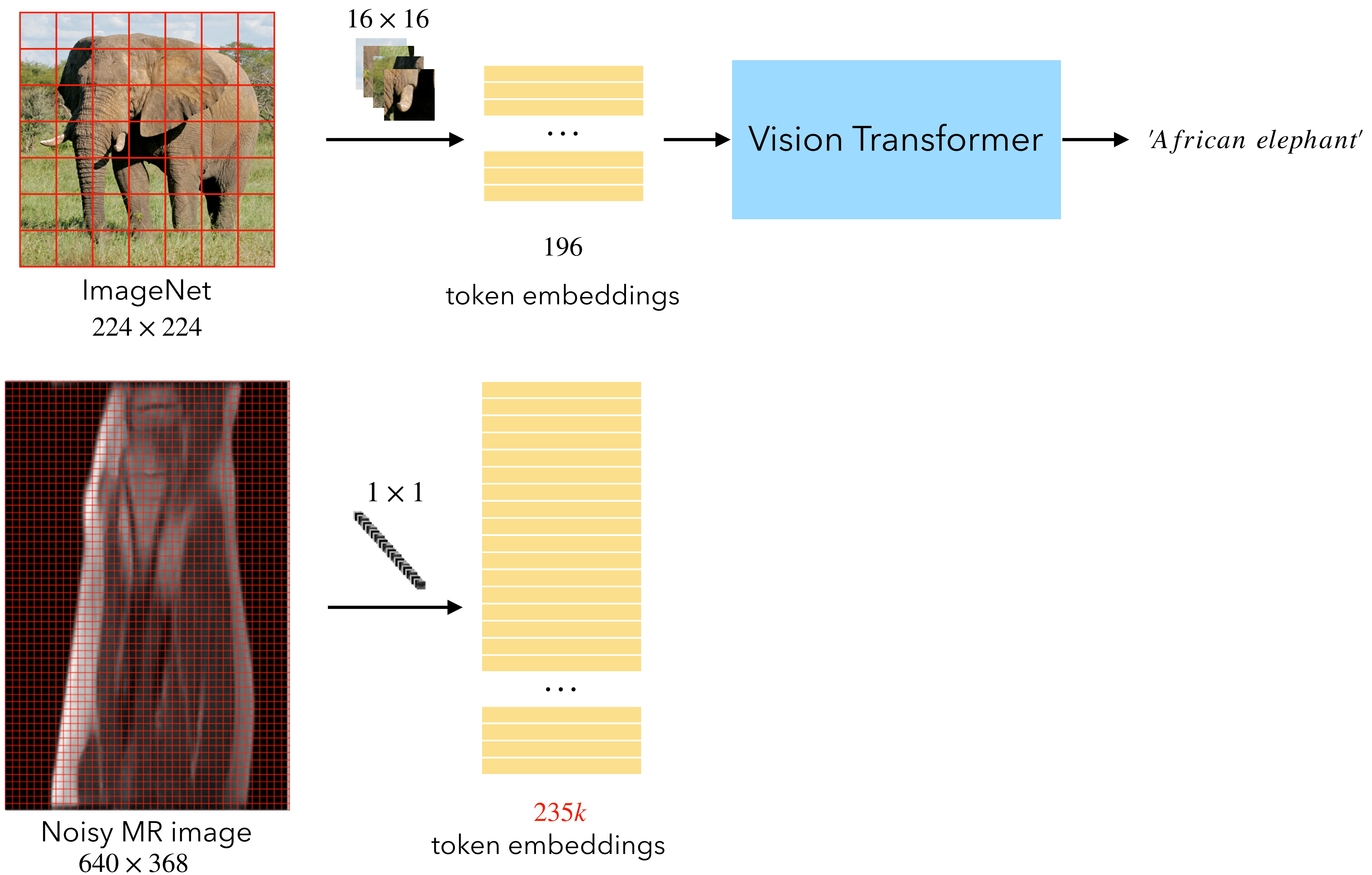
High-resolution challenge

- <u>Key challenge:</u> high-resolution input images **AND** dense prediction task



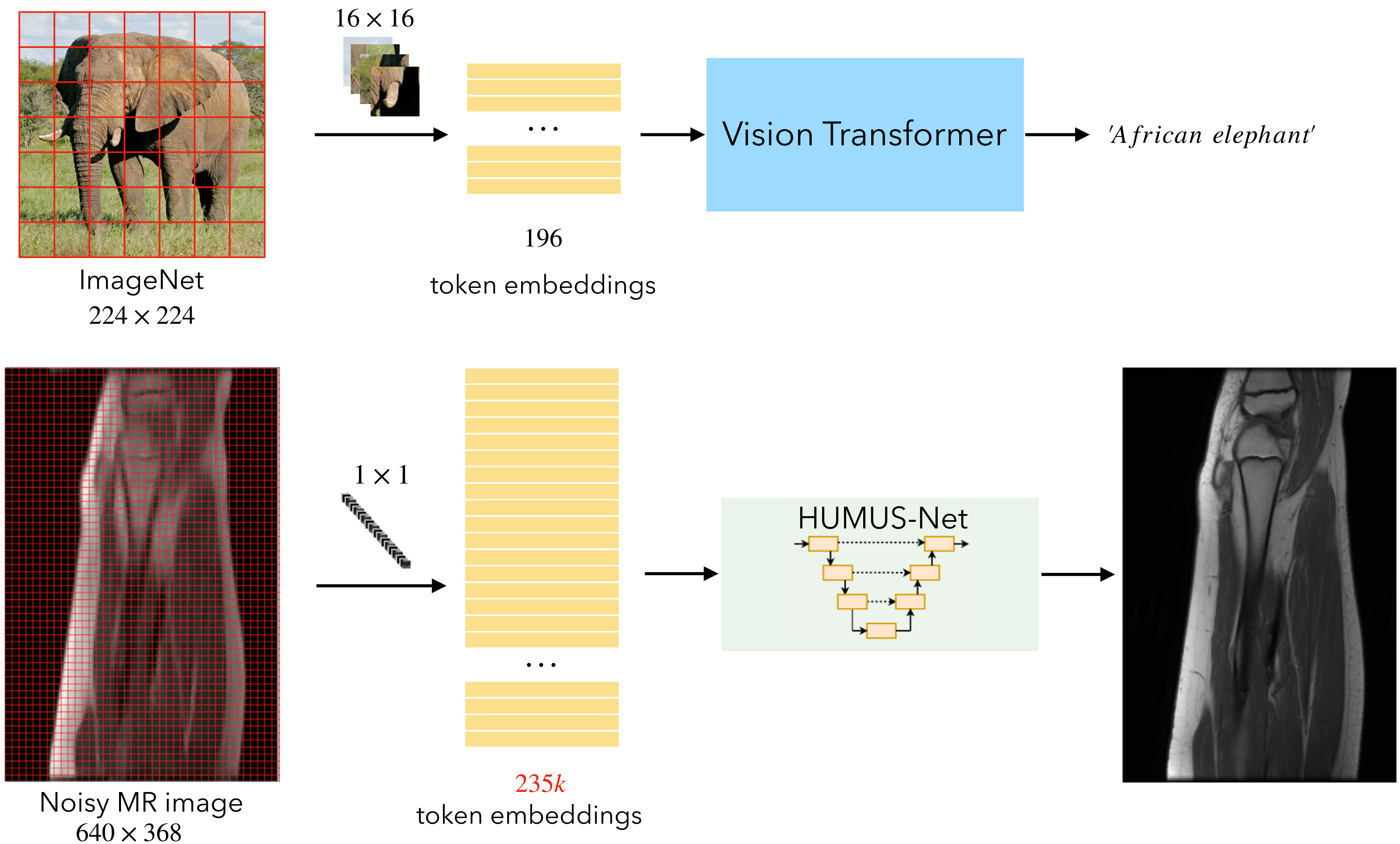
High-resolution challenge

- <u>Key challenge:</u> high-resolution input images **AND** dense prediction task



High-resolution challenge

- <u>Key challenge:</u> high-resolution input images **AND** dense prediction task



Tackling high-resolution

- Larger patch size?

Tackling high-resolution

- Larger patch size?

Model	GPU mem.	mins/epoch	Val. SSIM
E2E-VarNet	$\approx$ 16GB	9	**0.9313**

Tackling high-resolution

- Larger patch size?

Model	GPU mem.	mins/epoch	Val. SSIM
E2E-VarNet	$\approx$ 16GB	9	**0.9313**
Unrolled, M-S SwinIR	$\approx$ 16GB	66	0.9311

Tackling high-resolution

- Larger patch size?

Model	GPU mem.	mins/epoch	Val. SSIM
E2E-VarNet	$\approx$ 16GB	9	**0.9313**
Unrolled, M-S SwinIR	$\approx$ 16GB	66	0.9311
Unrolled, M-S SwinIR, patch size 2	$\approx$ 16GB	30	0.9171

Tackling high-resolution

- Larger patch size?

Model	GPU mem.	mins/epoch	Val. SSIM
E2E-VarNet	$\approx$ 16GB	9	**0.9313**
Unrolled, M-S SwinIR	$\approx$ 16GB	66	0.9311
Unrolled, M-S SwinIR, patch size 2	$\approx$ 16GB	30	0.9171

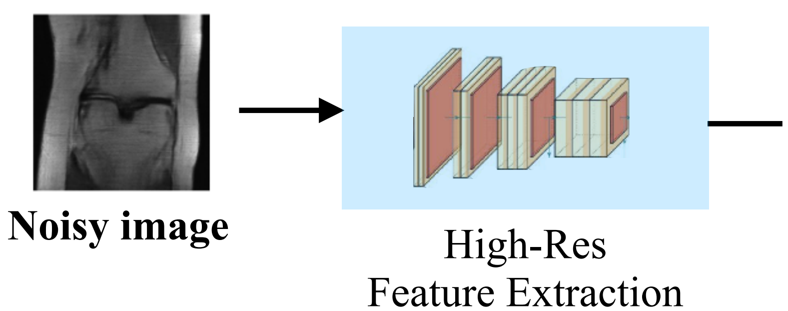
- Our solution:

Tackling high-resolution

- Larger patch size?

Model	GPU mem.	mins/epoch	Val. SSIM
E2E-VarNet	$\approx$ 16GB	9	**0.9313**
Unrolled, M-S SwinIR	$\approx$ 16GB	66	0.9311
Unrolled, M-S SwinIR, patch size 2	$\approx$ 16GB	30	0.9171

- Our solution:
  - extract high-res features via convolutions

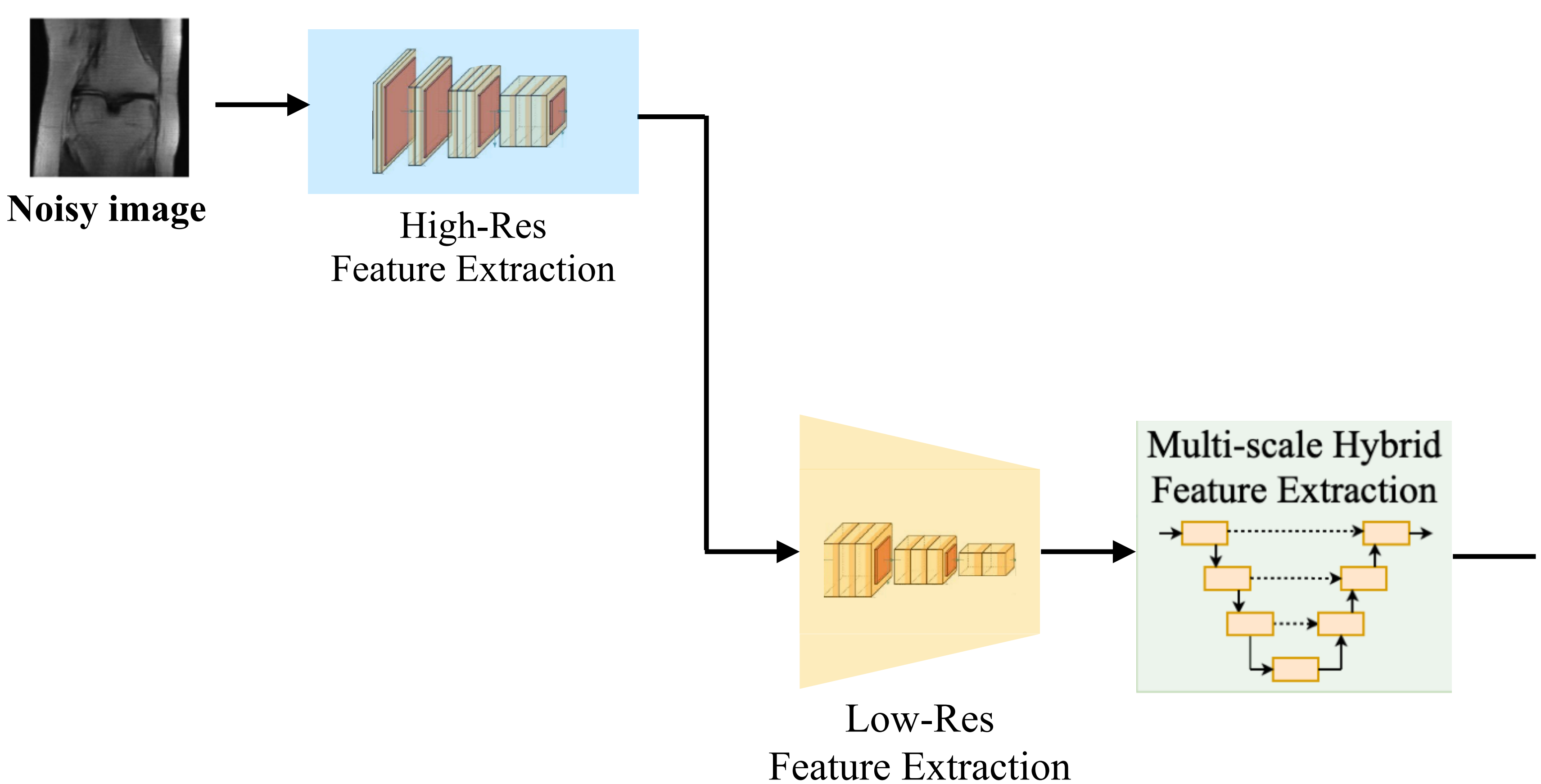


Tackling high-resolution

- Larger patch size?

Model	GPU mem.	mins/epoch	Val. SSIM
E2E-VarNet	≈ 16GB	9	**0.9313**
Unrolled, M-S SwinIR	≈ 16GB	66	0.9311
Unrolled, M-S SwinIR, patch size 2	≈ 16GB	30	0.9171

- Our solution:
    - extract high-res features via convolutions
    - process only lower resolution features via Transformers

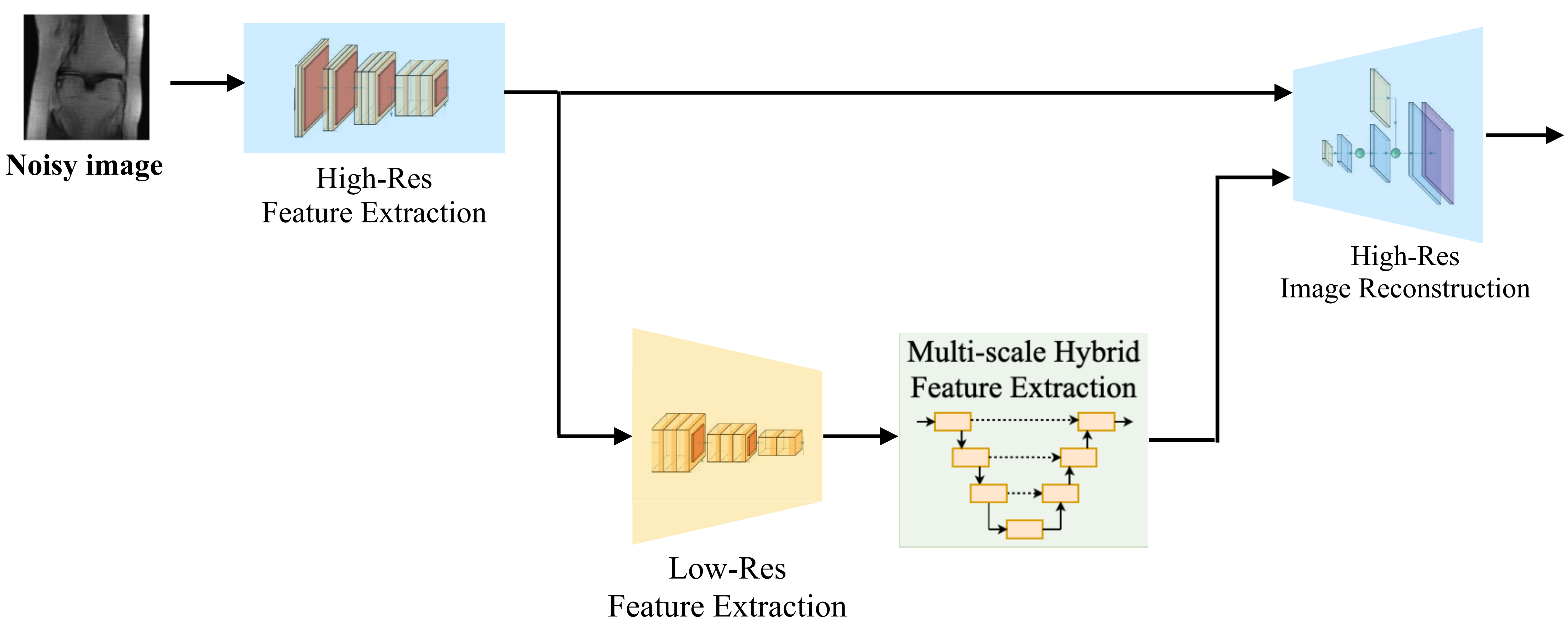


Tackling high-resolution

- Larger patch size?

Model	GPU mem.	mins/epoch	Val. SSIM
E2E-VarNet	≈ 16GB	9	**0.9313**
Unrolled, M-S SwinIR	≈ 16GB	66	0.9311
Unrolled, M-S SwinIR, patch size 2	≈ 16GB	30	0.9171

- Our solution:
  - extract high-res features via convolutions
  - process only lower resolution features via Transformers
  - synthesize high-res and low-res features for reconstruction

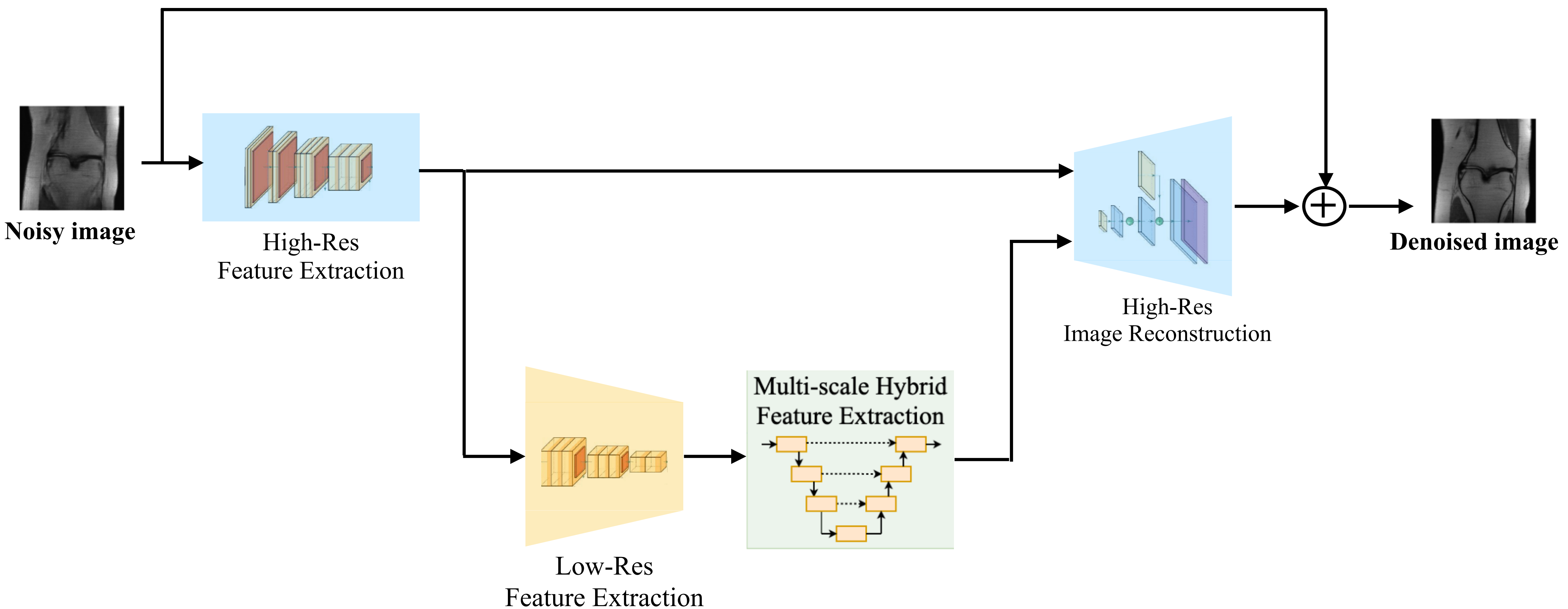


Tackling high-resolution

- Larger patch size?

Model	GPU mem.	mins/epoch	Val. SSIM
E2E-VarNet	≈ 16GB	9	**0.9313**
Unrolled, M-S SwinIR	≈ 16GB	66	0.9311
Unrolled, M-S SwinIR, patch size 2	≈ 16GB	30	0.9171

- Our solution:
  - extract high-res features via convolutions
  - process only lower resolution features via Transformers
  - synthesize high-res and low-res features for reconstruction

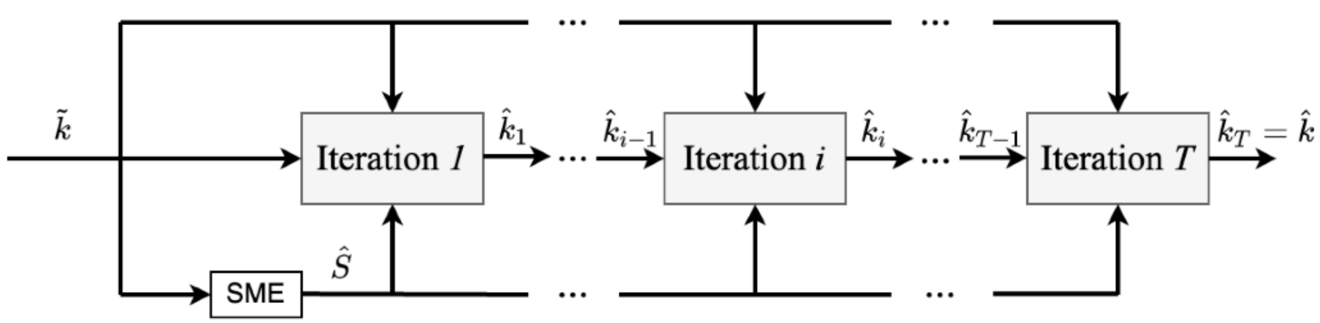


Further improvements

- Adding unrolling

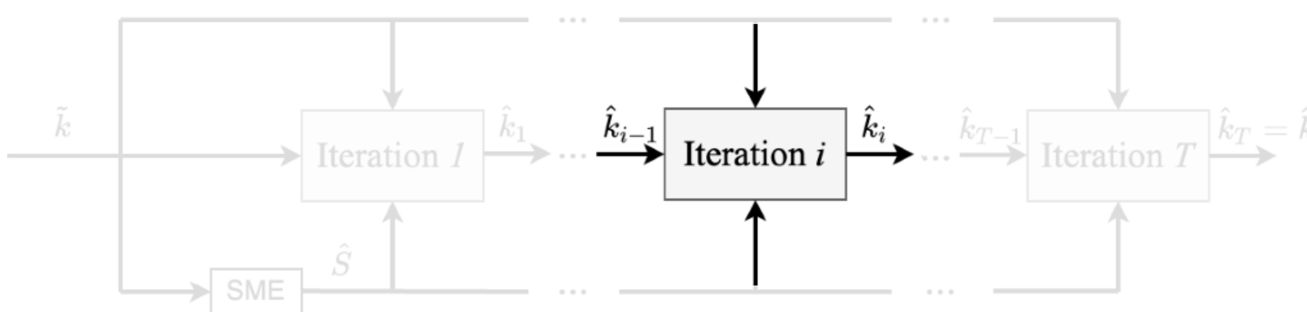
Further improvements

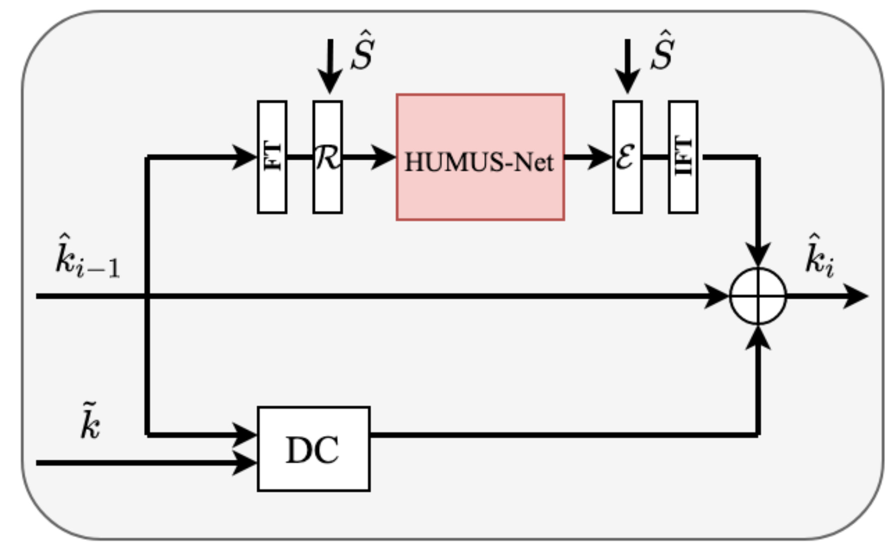
- Adding unrolling



Further improvements

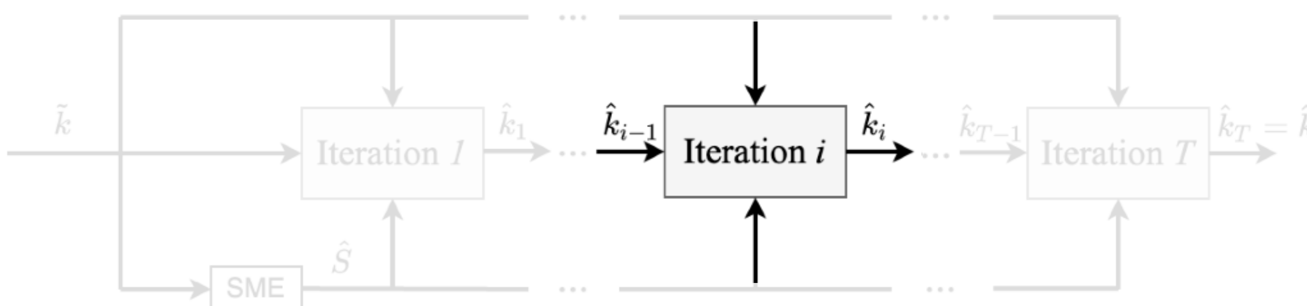
- Adding unrolling

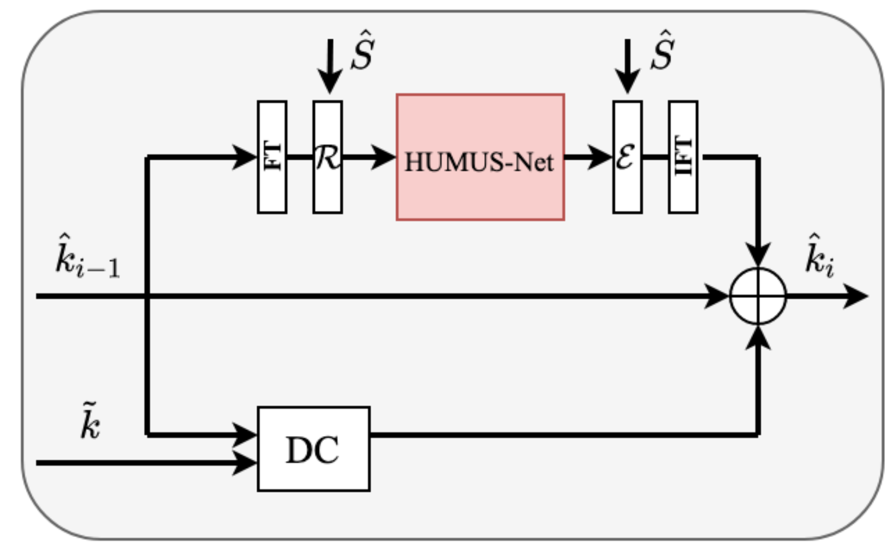




Further improvements

- Adding unrolling

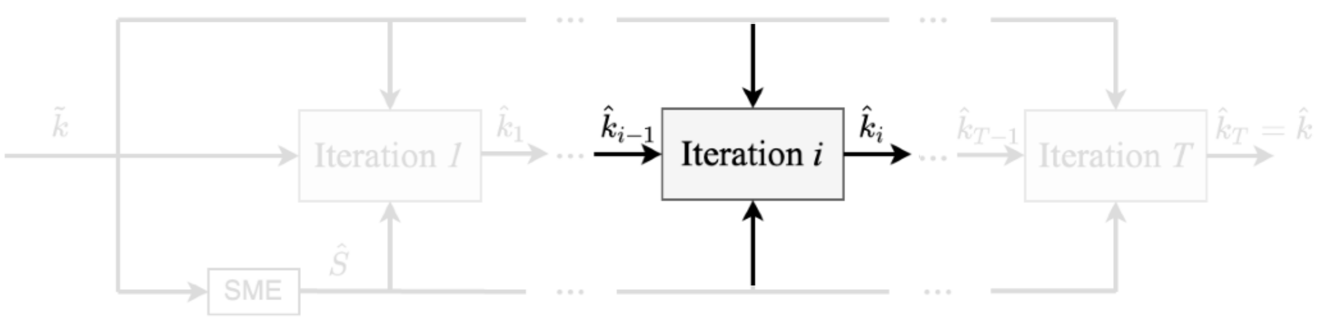


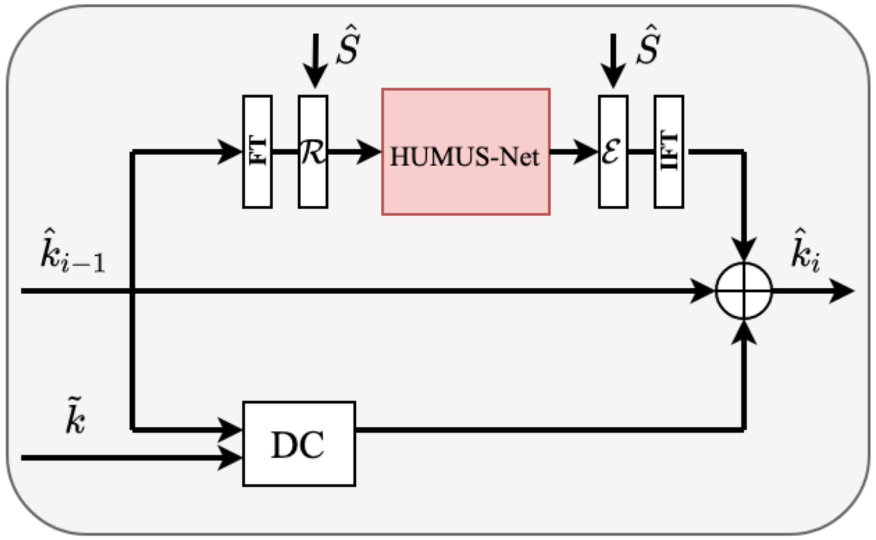


- Adjacent slice reconstruction

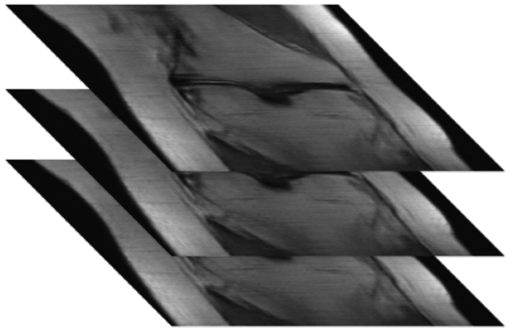
Further improvements

- Adding unrolling



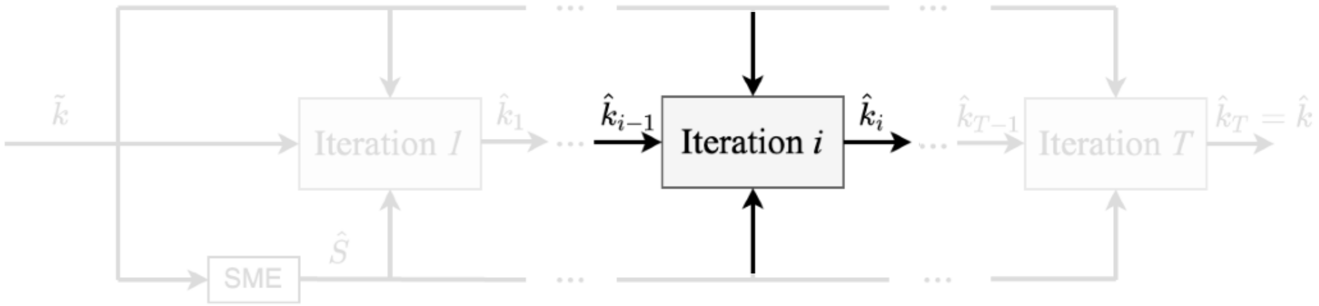


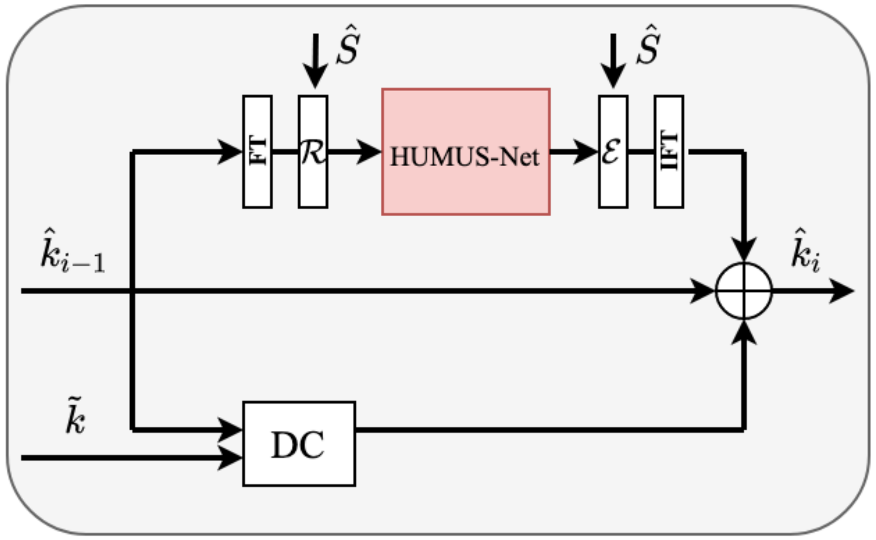
- Adjacent slice reconstruction



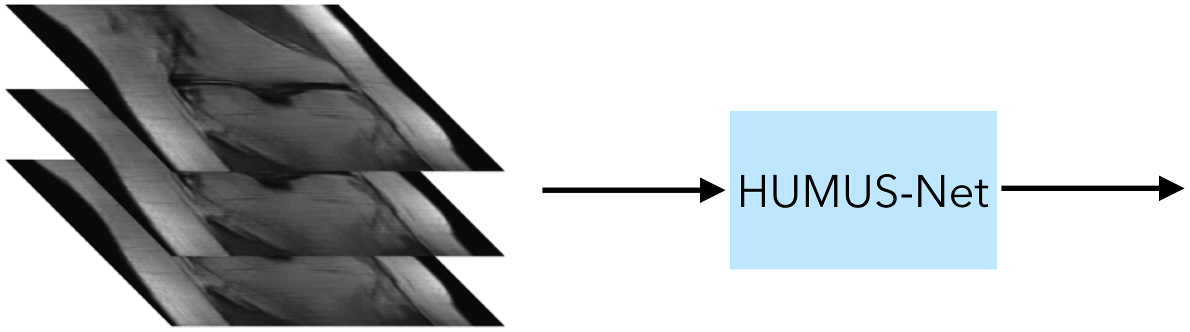
Further improvements

- Adding unrolling



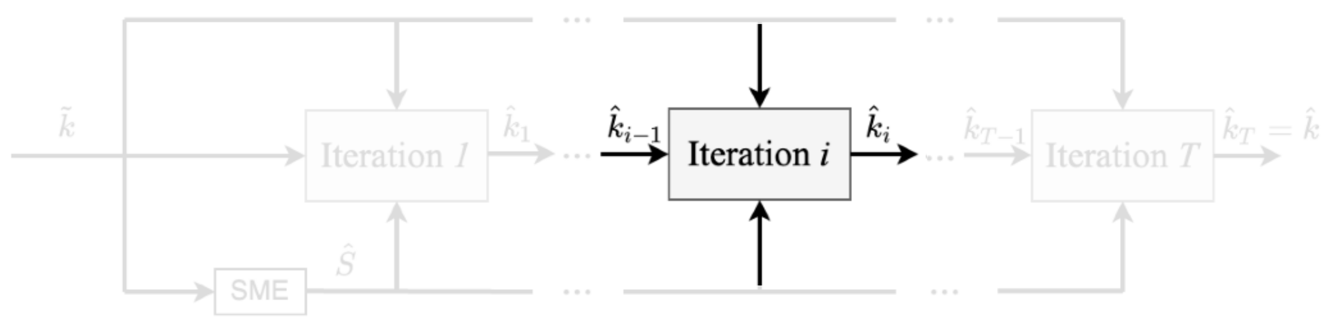


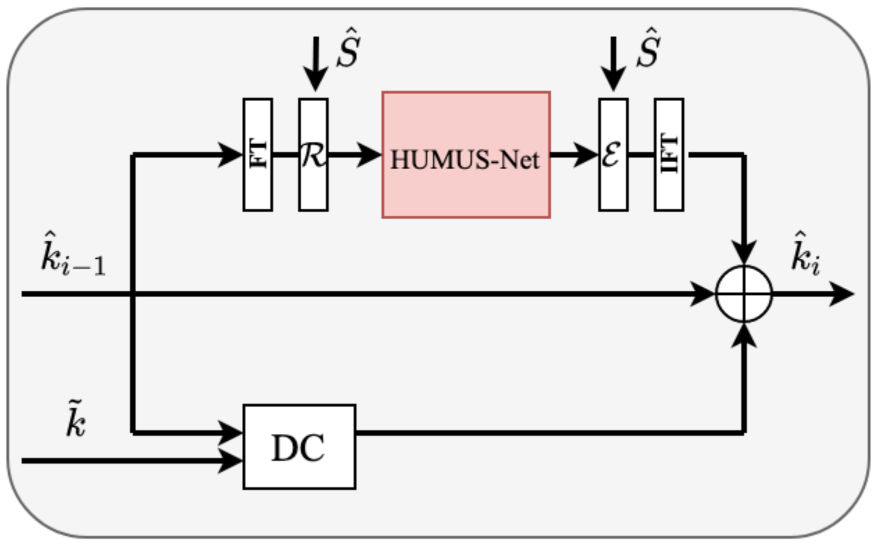
- Adjacent slice reconstruction



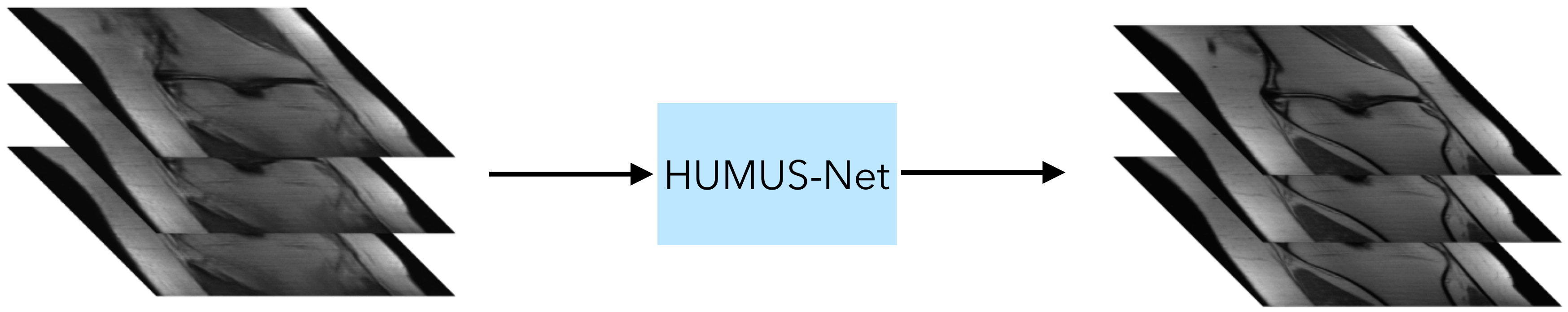
Further improvements

- Adding unrolling



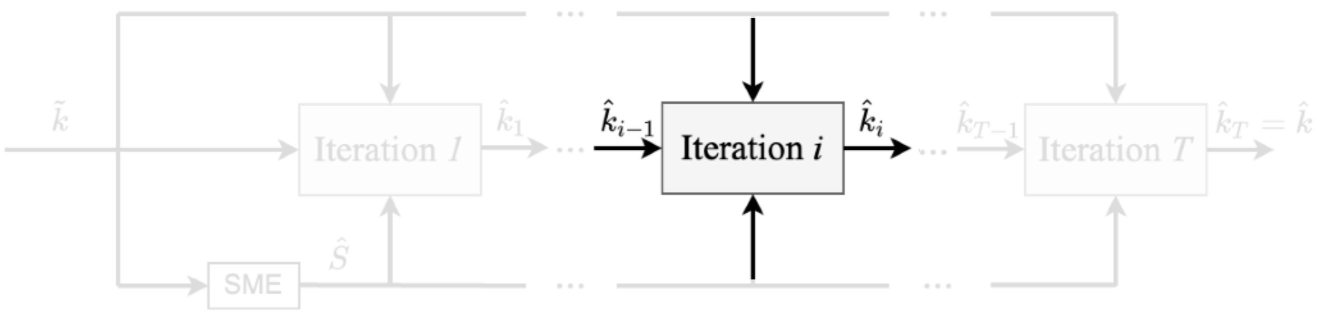


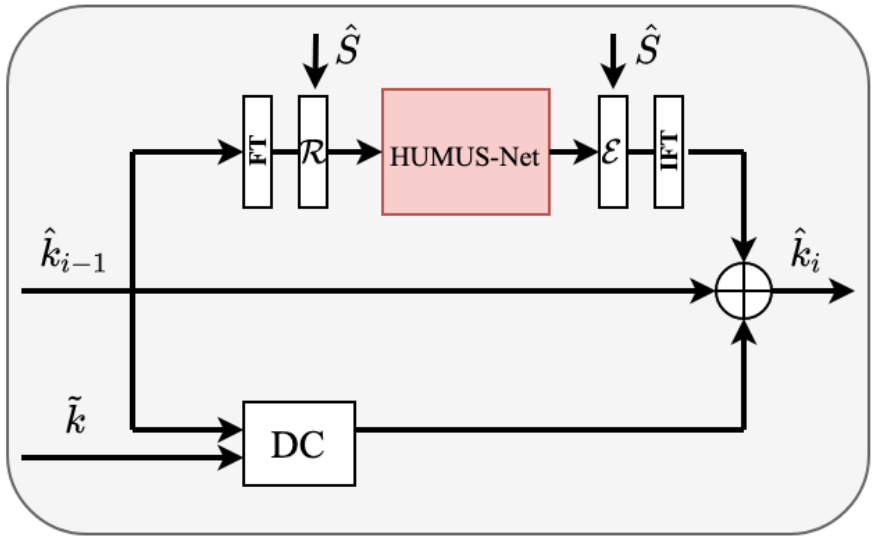
- Adjacent slice reconstruction



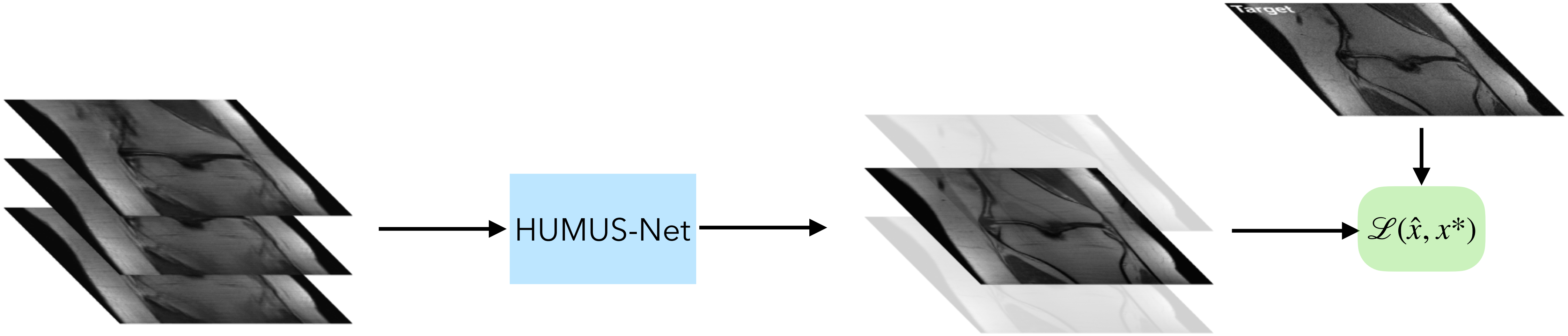
Further improvements

- Adding unrolling





- Adjacent slice reconstruction



HUMUS-Net Experiments

fastMRI multi-coil knee, 8x acceleration

Model	Val. SSIM	Val. PSNR

HUMUS-Net Experiments

fastMRI multi-coil knee, 8x acceleration

Model	Val. SSIM	Val. PSNR
E2E-VarNet	0.8908	36.77

HUMUS-Net Experiments

fastMRI multi-coil knee, 8x acceleration

Model	Val. SSIM	Val. PSNR
E2E-VarNet	0.8908	36.77
HUMUS-Net	**0.8934**	**37.04**

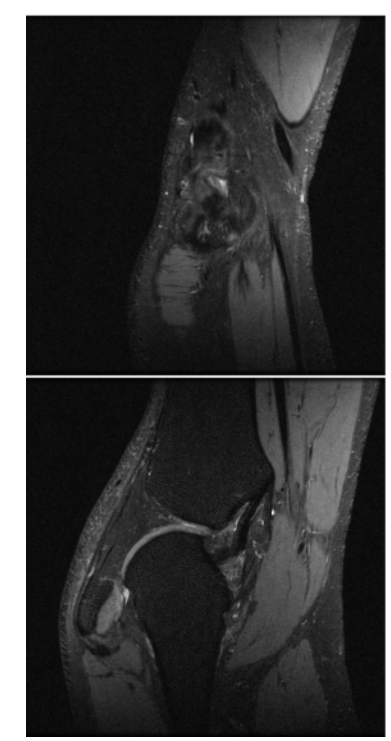
HUMUS-Net Experiments

fastMRI multi-coil knee, 8x acceleration

Model	Val. SSIM	Val. PSNR
E2E-VarNet	0.8908	36.77
HUMUS-Net	**0.8934**	**37.04**

Stanford3D multi-coil knee, 8x acceleration

Model	Val. SSIM	Val. PSNR



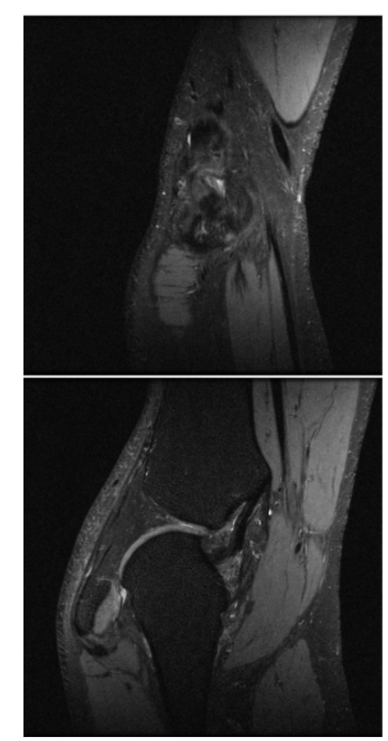
HUMUS-Net Experiments

fastMRI multi-coil knee, 8x acceleration

Model	Val. SSIM	Val. PSNR
E2E-VarNet	0.8908	36.77
HUMUS-Net	**0.8934**	**37.04**

Stanford3D multi-coil knee, 8x acceleration

Model	Val. SSIM	Val. PSNR
E2E-VarNet	0.9432 (±0.0063)	39.99 (± 0.6144)



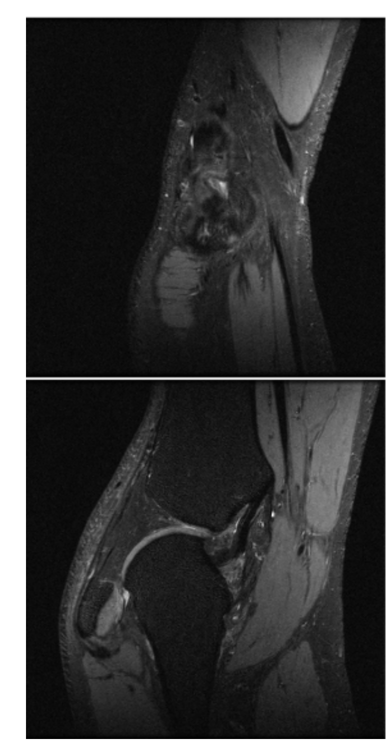
HUMUS-Net Experiments

fastMRI multi-coil knee, 8x acceleration

Model	Val. SSIM	Val. PSNR
E2E-VarNet	0.8908	36.77
HUMUS-Net	**0.8934**	**37.04**

Stanford3D multi-coil knee, 8x acceleration

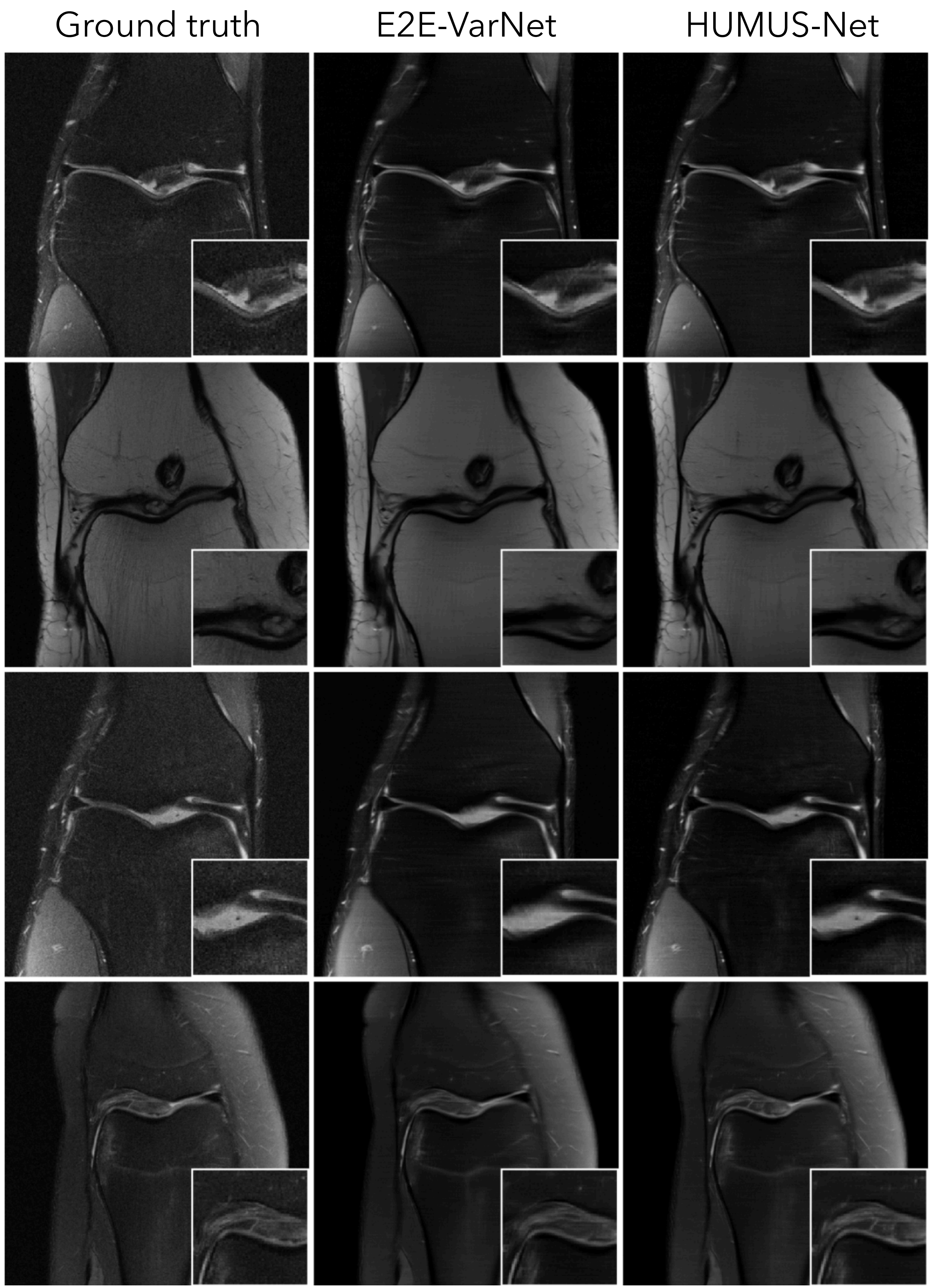
Model	Val. SSIM	Val. PSNR
E2E-VarNet	0.9432 (±0.0063)	39.99 (± 0.6144)
HUMUS-Net	**0.9453** (± 0.0065)	**40.35** (± 0.6460)



HUMUS-Net Experiments

- fastMRI Public Leaderboard

fastMRI knee multi-coil $\times 8$ test data			
Method	**SSIM(↑)**	**PSNR(↑)**	**NMSE(↓)**
E2E-VarNet [Sriram et al., 2020]	0.8900	36.9	0.0089
E2E-VarNet† [Sriram et al., 2020]	0.8920	37.1	0.0085
XPDNet [Ramzi et al., 2020]	0.8893	37.2	0.0083
Σ-Net [Hammernik et al., 2019]	0.8877	36.7	0.0091
i-RIM [Putzky et al., 2019]	0.8875	36.7	0.0091
U-Net [Zbontar et al., 2019]	0.8640	34.7	0.0132
HUMUS-Net (ours)	0.8936	37.0	0.0086
HUMUS-Net (ours)†	**0.8945**	**37.3**	**0.0081**


Ground truth
E2E-VarNet
HUMUS-Net