Paper #1005, ISMRM 2020

Attention-gated convolutional neural networks for off-resonance correction of spiral real-time MRI

Yongwan Lim, Shrikanth S. Narayanan, Krishna S. Nayak

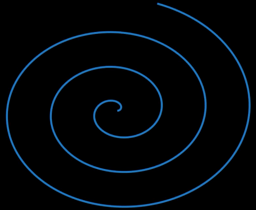
Ming Hsieh Department of Electrical and Computer Engineering, Viterbi School of Engineering,
University of Southern California, Los Angeles, California, USA

USC Viterbi
School of Engineering

Declaration of
Financial Interests or Relationships

Speaker Name: Yongwan Lim

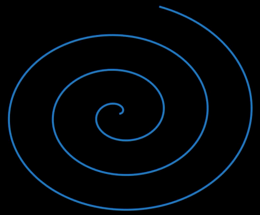
I have no financial interests or relationships to disclose with regard to the subject matter of this presentation.



Spiral Real-time MRI

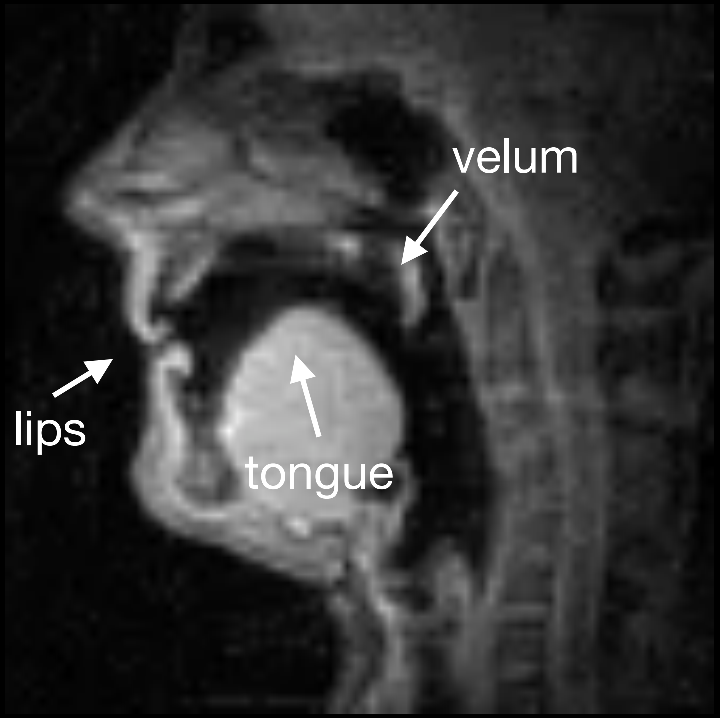
Vocal tract

2.4mm^2, 12ms/frame, R=6.5
@ USC



Spiral Real-time MRI

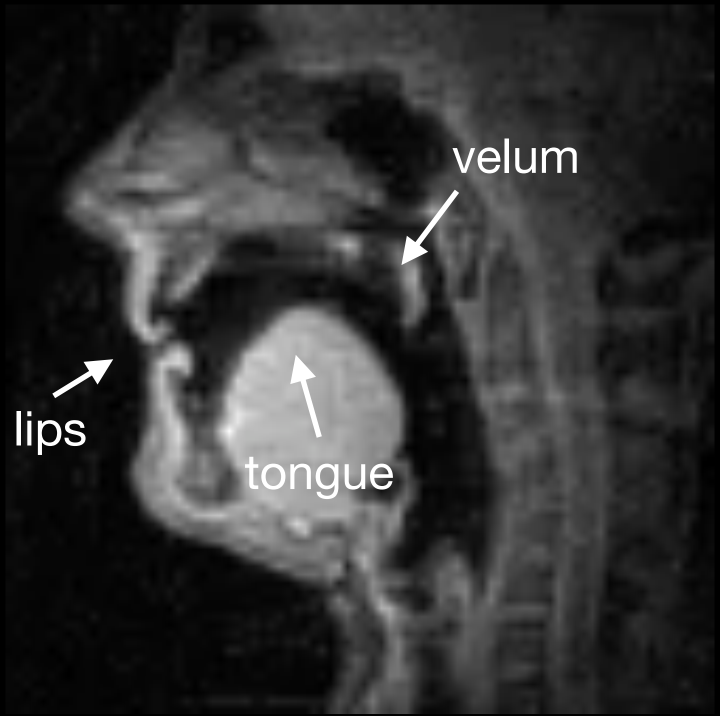
Vocal tract



2.4mm², 12ms/frame, R=6.5
@ USC

Spiral Real-time MRI

Vocal tract

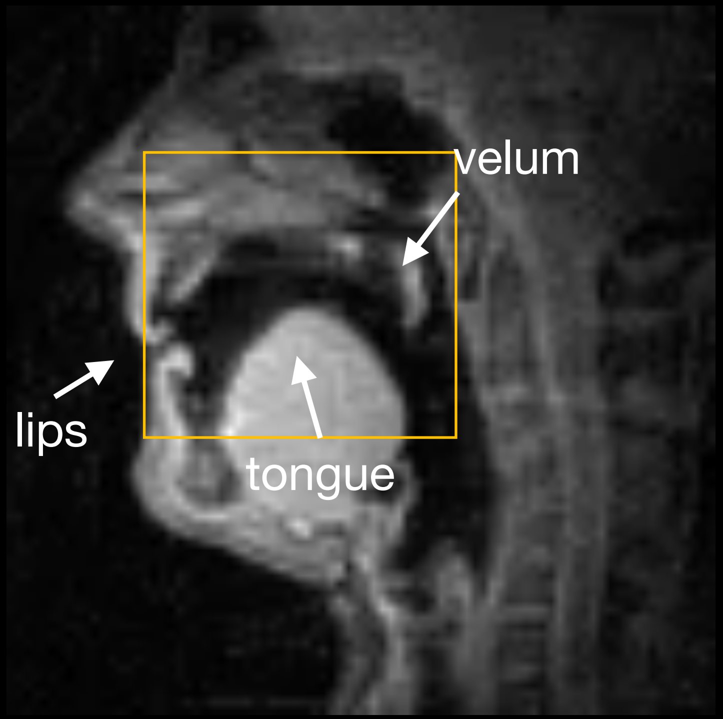


2.4mm², 12ms/frame, R=6.5
@ USC

- Off-resonance artifacts due to local susceptibility difference between air and tissue
  - Spatially and temporally varying

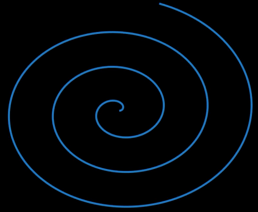
Spiral Real-time MRI

Vocal tract



2.4mm², 12ms/frame, R=6.5
@ USC

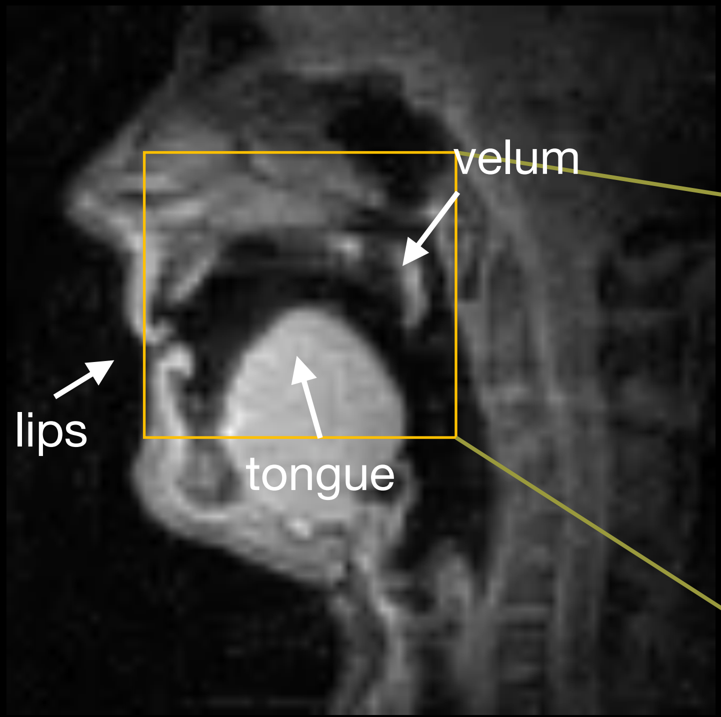
- Off-resonance artifacts due to local susceptibility difference between air and tissue
  - Spatially and temporally varying



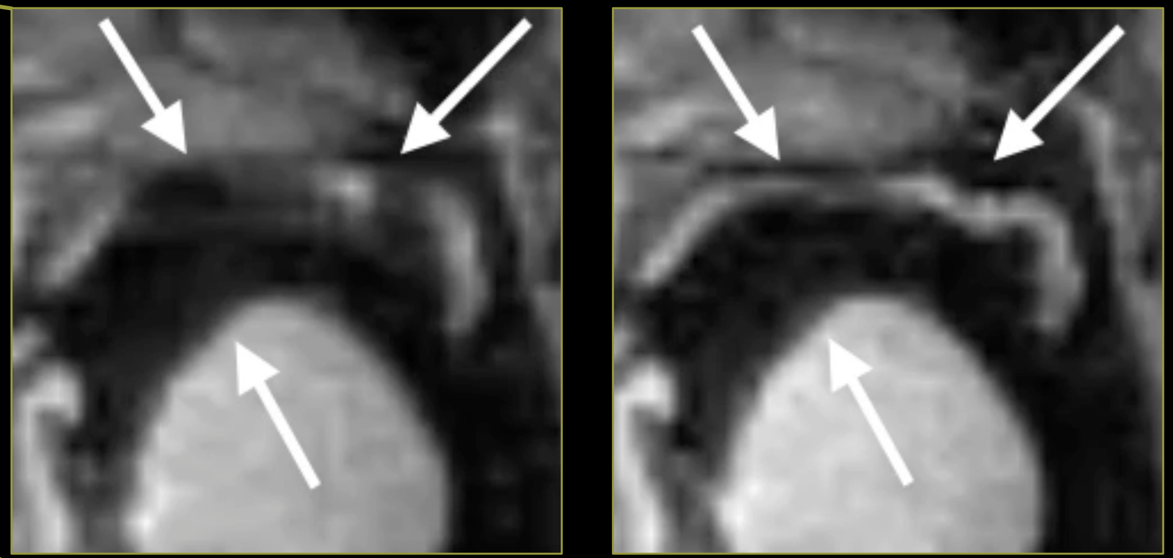
Spiral Real-time MRI

- Off-resonance artifacts due to local susceptibility difference between air and tissue
  - Spatially and temporally varying

Vocal tract



2.4mm^2, 12ms/frame, R=6.5
@ USC



Blurring Artifact **After De-Blurring**

Off-resonance Deblurring

- Standard Approaches[1-4]:

Blurry Image

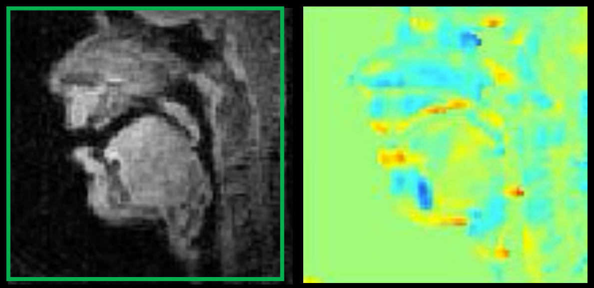
Deblurred Image

[1] KS Nayak et al, MRM. 2001
[2] BP Sutton et al, JMRI. 2010
[3] Y Lim et al. MRM. 2019
[4] DC Noll et al, MRM. 1992

Off-resonance Deblurring

- Standard Approaches[1-4]:

Blurry Image Field Map



Deblurred Image

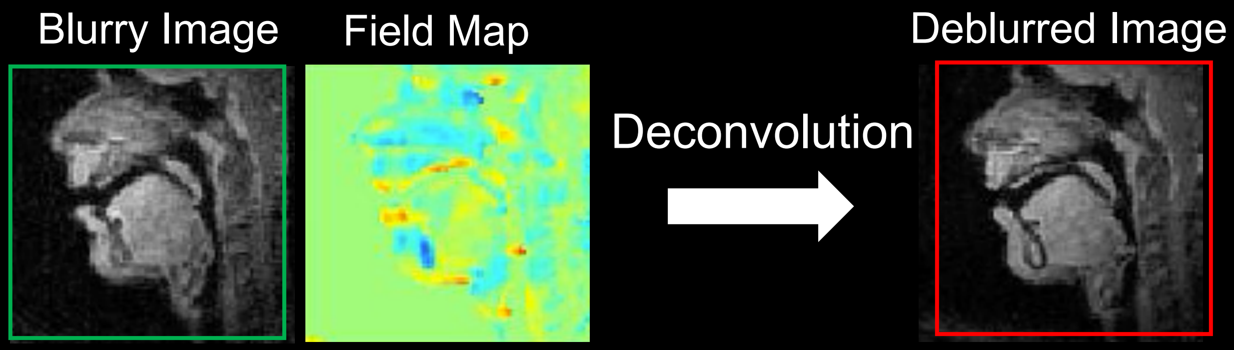
1. Field map acquisition
   - Dual-TE (cf: single-TE or auto-focus)
   - Reduced scan efficiency
2. Spatially-varying deconvolution
   - Non-iterative or iterative methods
   - Computationally slow (~minutes)

USC Viterbi School of Engineering

[1] KS Nayak et al, MRM. 2001
[2] BP Sutton et al, JMRI. 2010
[3] Y Lim et al. MRM. 2019
[4] DC Noll et al, MRM. 1992

Off-resonance Deblurring

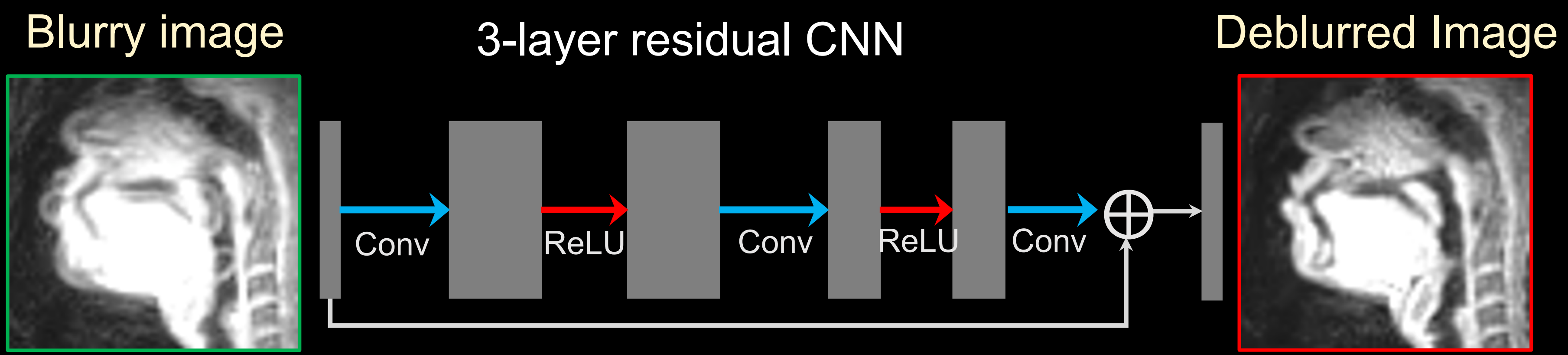
- Standard Approaches[1-4]:



1. Field map acquisition
   - Dual-TE (cf: single-TE or auto-focus)
   - Reduced scan efficiency
2. Spatially-varying deconvolution
   - Non-iterative or iterative methods
   - Computationally slow (~minutes)

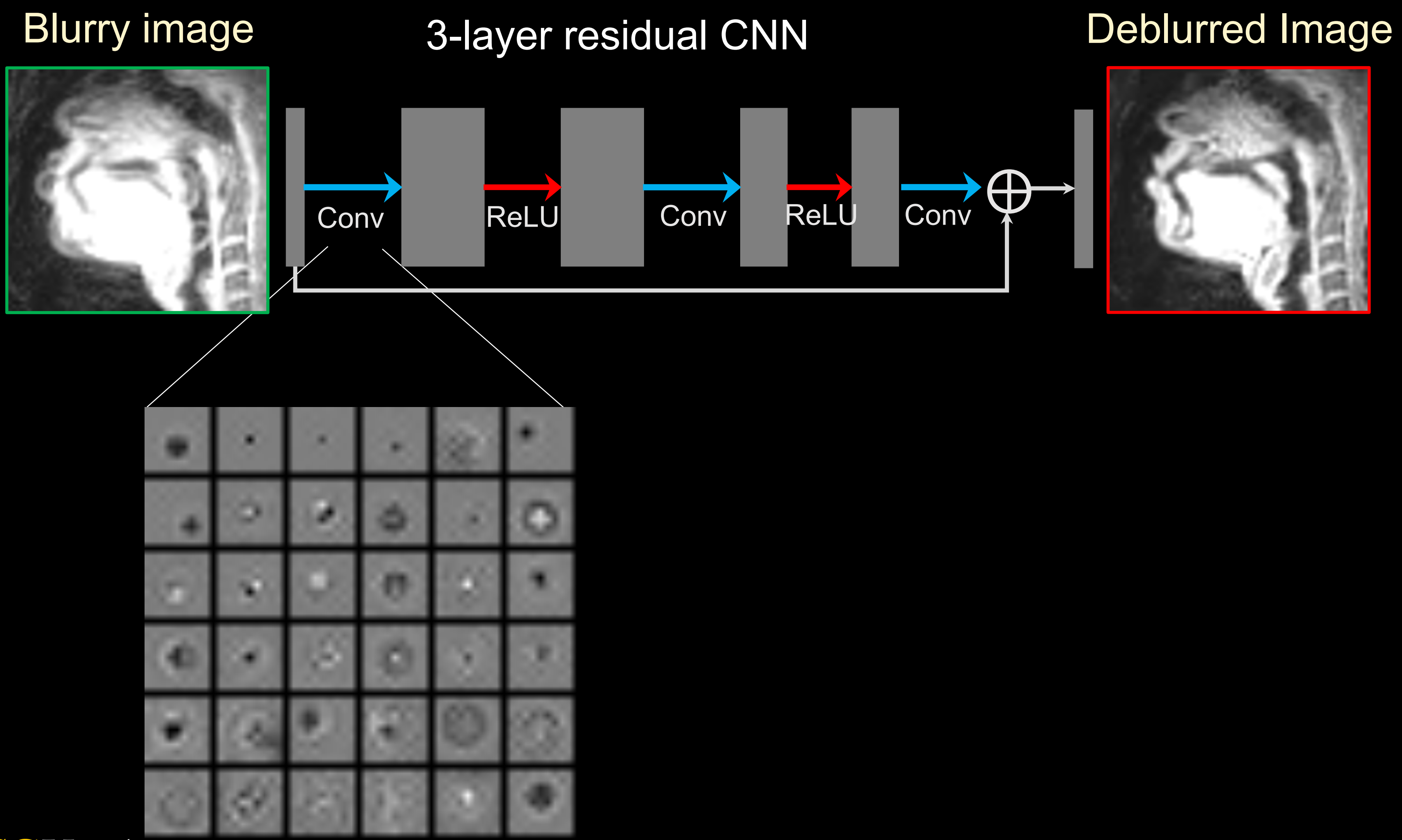
[1] KS Nayak et al, MRM. 2001
[2] BP Sutton et al, JMRI. 2010
[3] Y Lim et al. MRM. 2019
[4] DC Noll et al, MRM. 1992

CNN-based Deblurring[1]



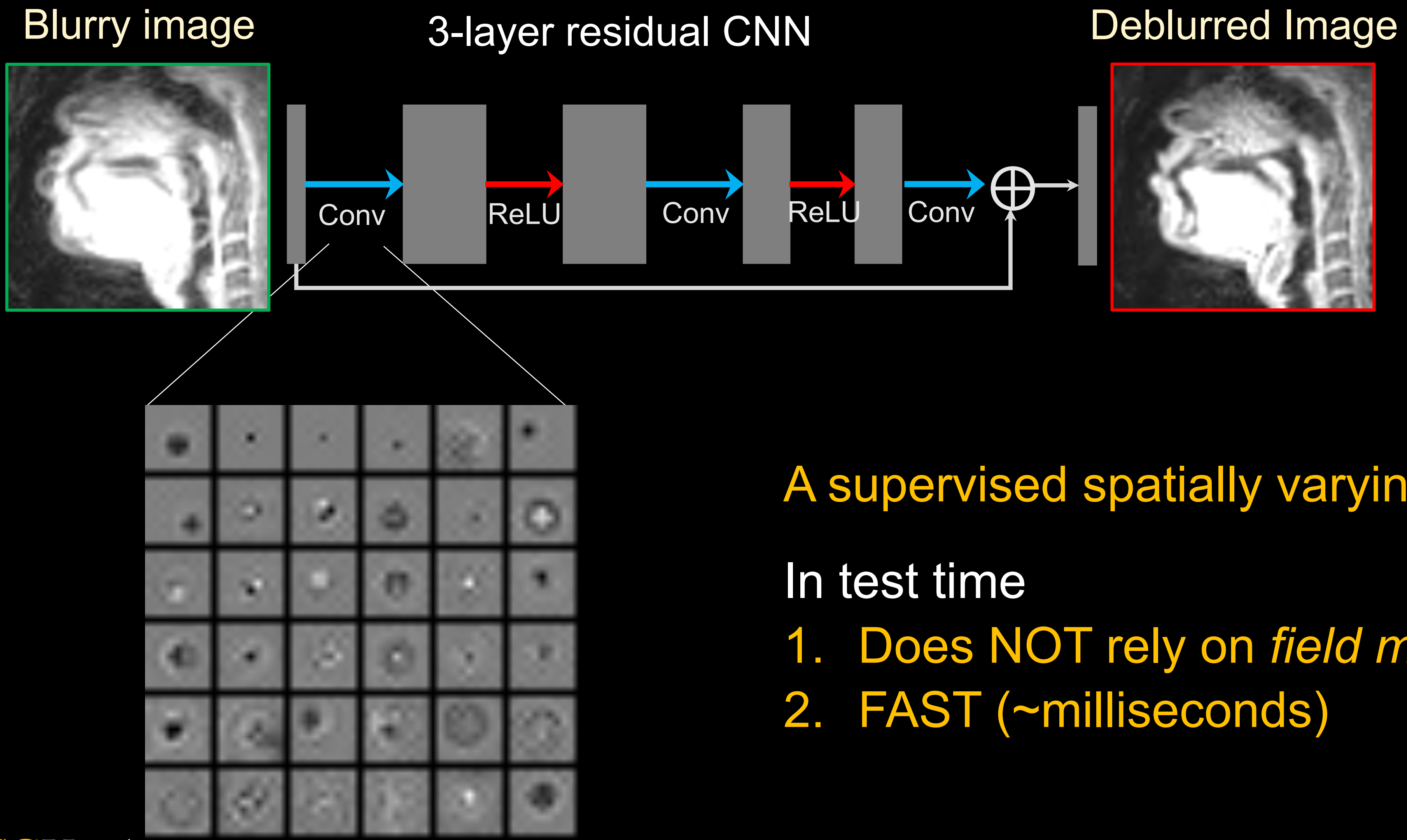
[1] Y Lim, et al, MRM. 2020. 10.1002/mrm.28393

CNN-based Deblurring[1]



[1] Y Lim, et al, MRM. 2020. 10.1002/mrm.28393

CNN-based Deblurring[1]



A supervised spatially varying deconvolution

In test time

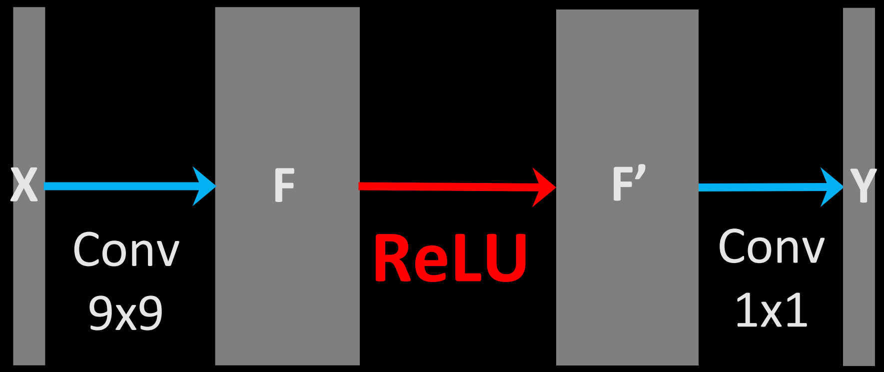
1. Does NOT rely on *field map*
2. FAST (~milliseconds)

[1] Y Lim, et al, MRM. 2020. 10.1002/mrm.28393

Motivation

ReLU nonlinearity

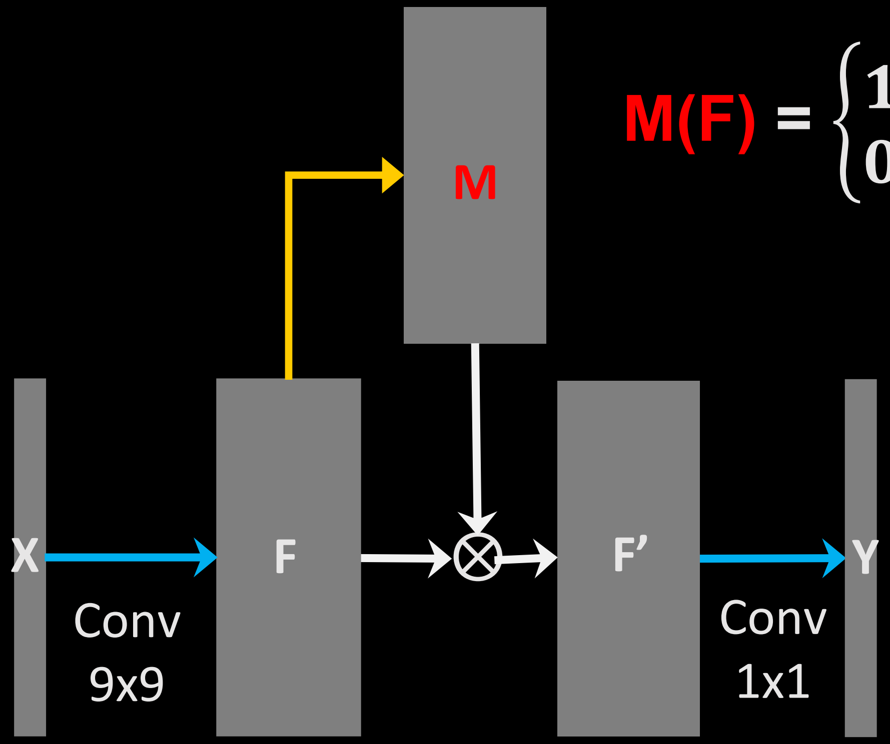
- Provides a spatially-varying binary mask to convolution filters, enabling spatially-varying convolution.



Motivation

ReLU nonlinearity

- Provides a spatially-varying binary mask to convolution filters, enabling spatially-varying convolution.



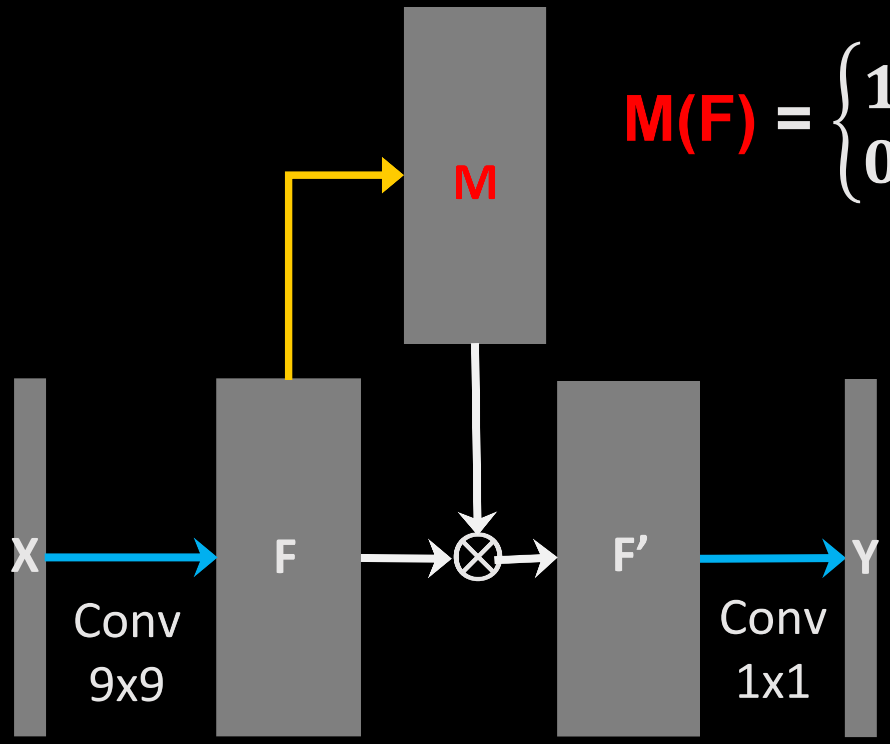
$$\mathbf{M(F)} = \begin{cases} 1 & \text{if } \mathrm{F} > 0 \\ 0 & \text{o.w} \end{cases}$$

[1] LC Man et al, MRM 1997 [2] DC Noll et al, MRM 1992

Motivation

ReLU nonlinearity

- Provides a spatially-varying binary mask to convolution filters, enabling spatially-varying convolution.



$$\mathbf{M(F)} = \begin{cases} 1 & \text{if } F > 0 \\ 0 & \text{o.w} \end{cases}$$

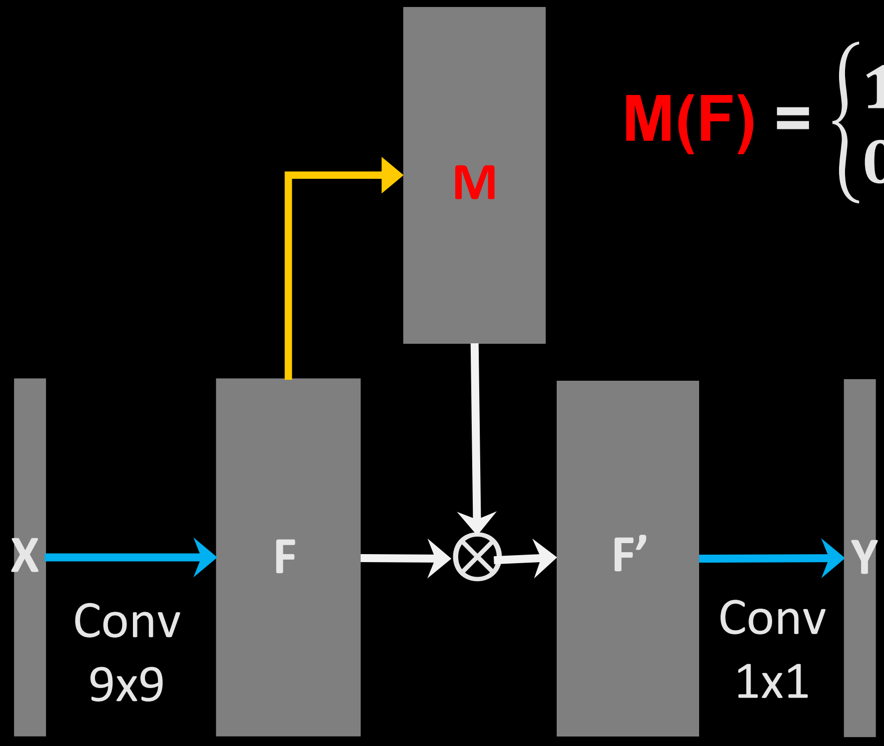
$$F' = F \otimes M(F)$$

[1] LC Man et al, MRM 1997 [2] DC Noll et al, MRM 1992

Motivation

ReLU nonlinearity

- Provides a spatially-varying binary mask to convolution filters, enabling spatially-varying convolution.



$$M(F) = \begin{cases} 1 & \text{if } F > 0 \\ 0 & \text{o.w} \end{cases}$$

$$F' = F \otimes M(F)$$

- The binary mask is computed only based on the sign of pixel value in an element-wise manner.

- It cannot exploit local spatial or channel (filter) dependency, unlike the conventional deblurring methods such as multi-frequency reconstruction[1] or autofocus[2].

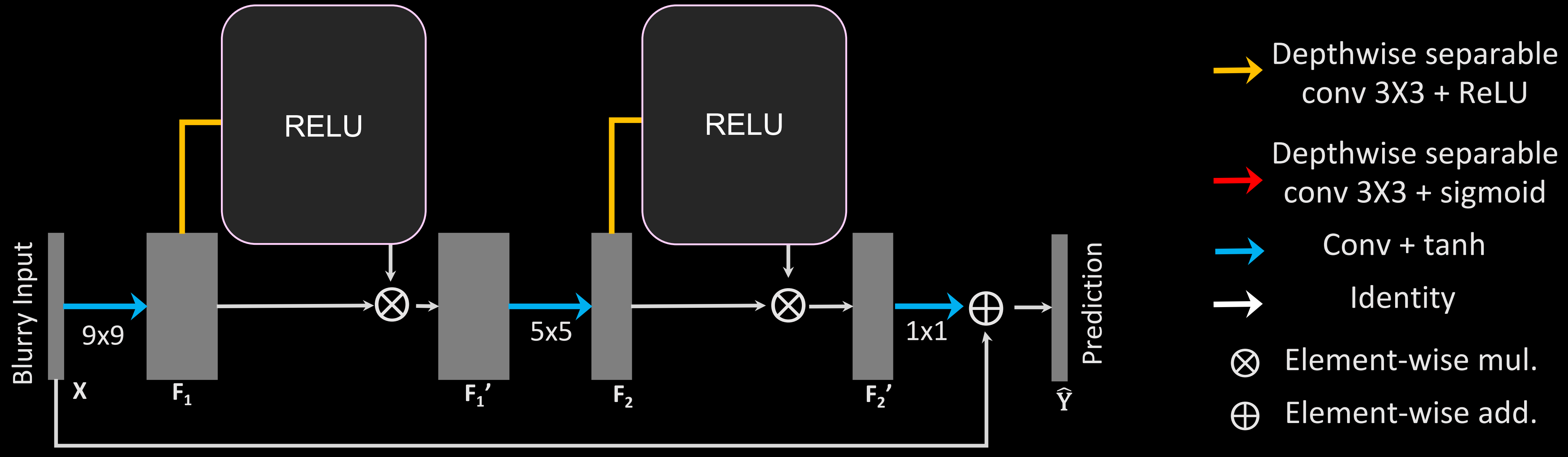
[1] LC Man et al, MRM 1997 [2] DC Noll et al, MRM 1992

Goal of This Work

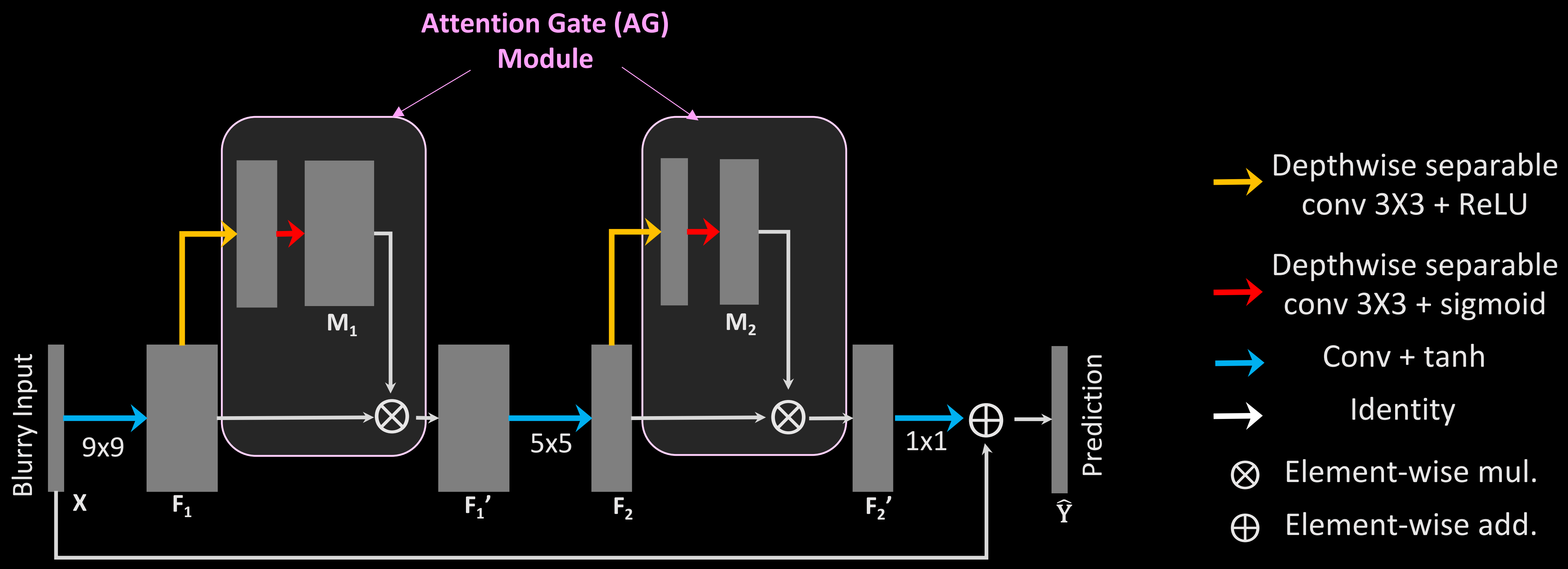
To exploit spatial and channel relationships of filtered outputs to improve the expressiveness of a network

…and enables an efficient off-resonance deblurring in the application of spiral RT-MRI of speech

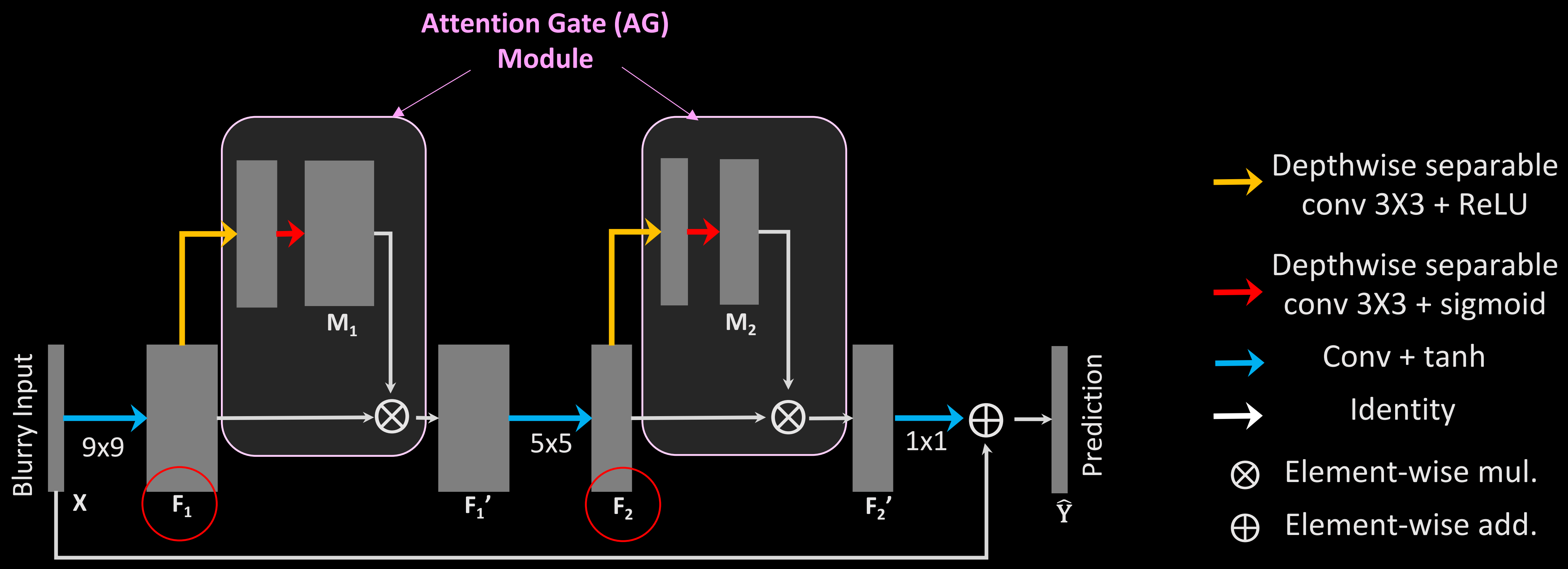
Attention-gate CNN (AG-CNN)



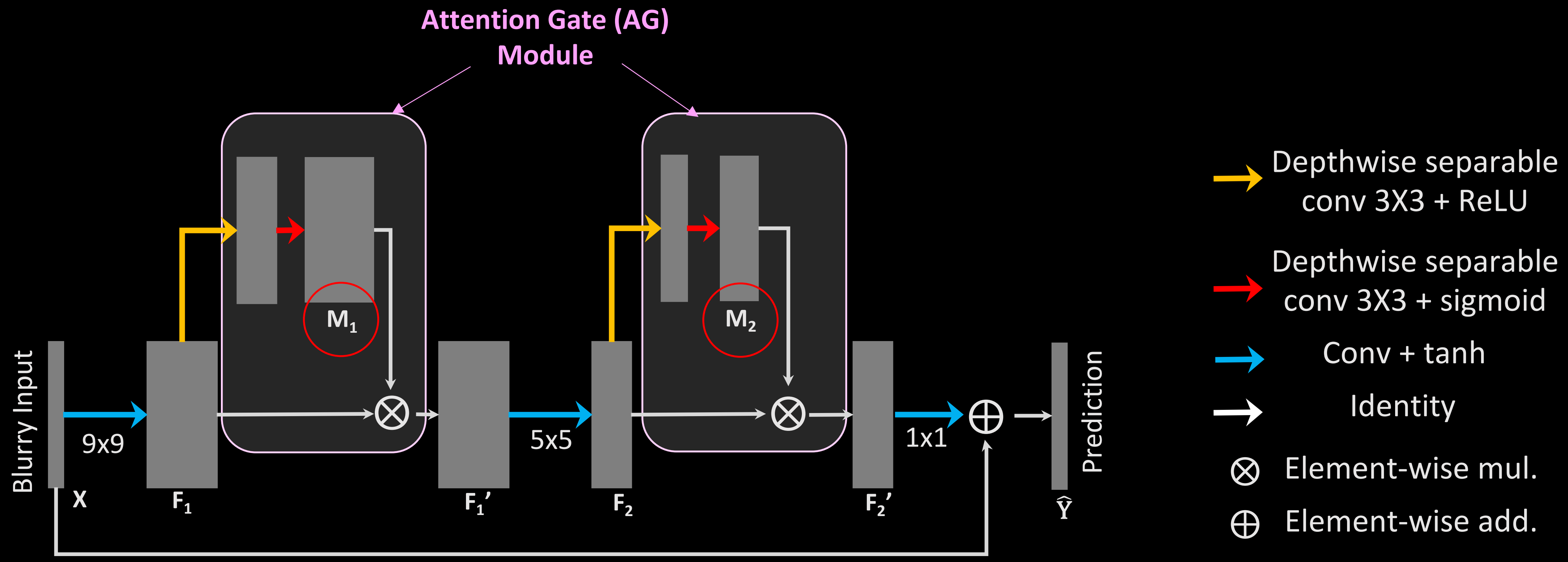
Attention-gate CNN (AG-CNN)



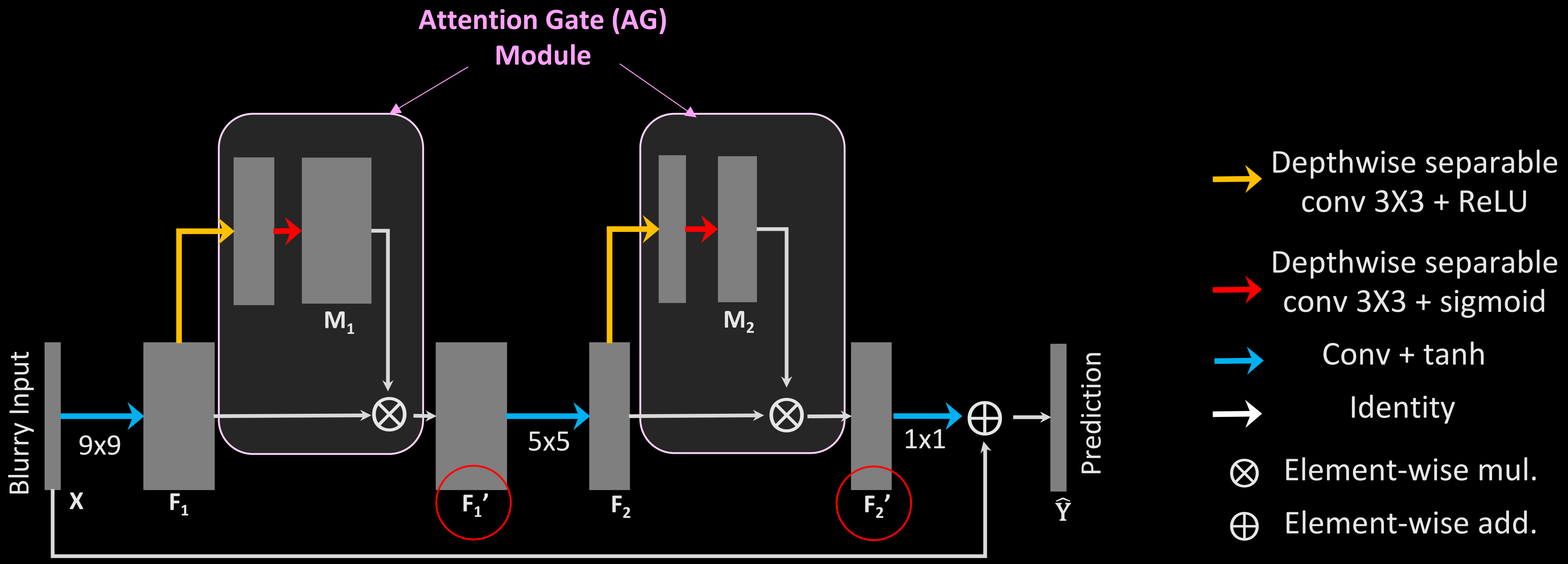
Attention-gate CNN (AG-CNN)



Attention-gate CNN (AG-CNN)



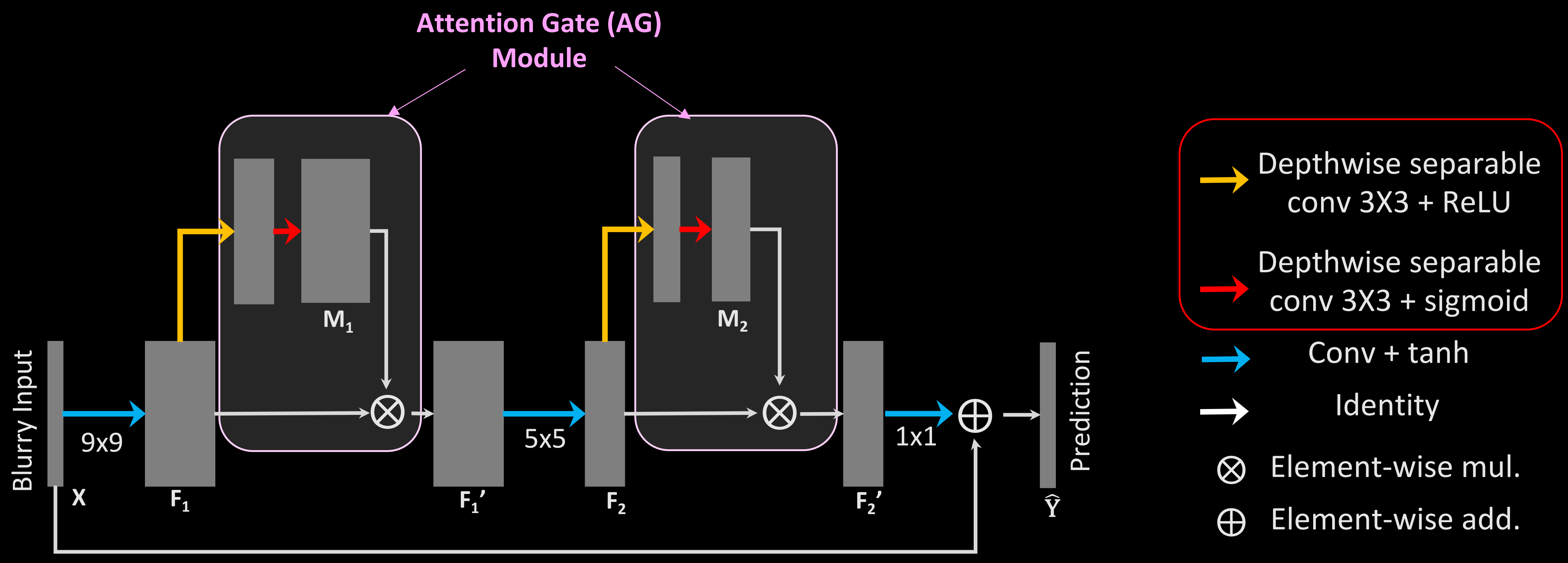
Attention-gate CNN (AG-CNN)



$$F_1' = F_1 \otimes M_1(F_1)$$

$$F_2' = F_2 \otimes M_2(F_2)$$

Attention-gate CNN (AG-CNN)



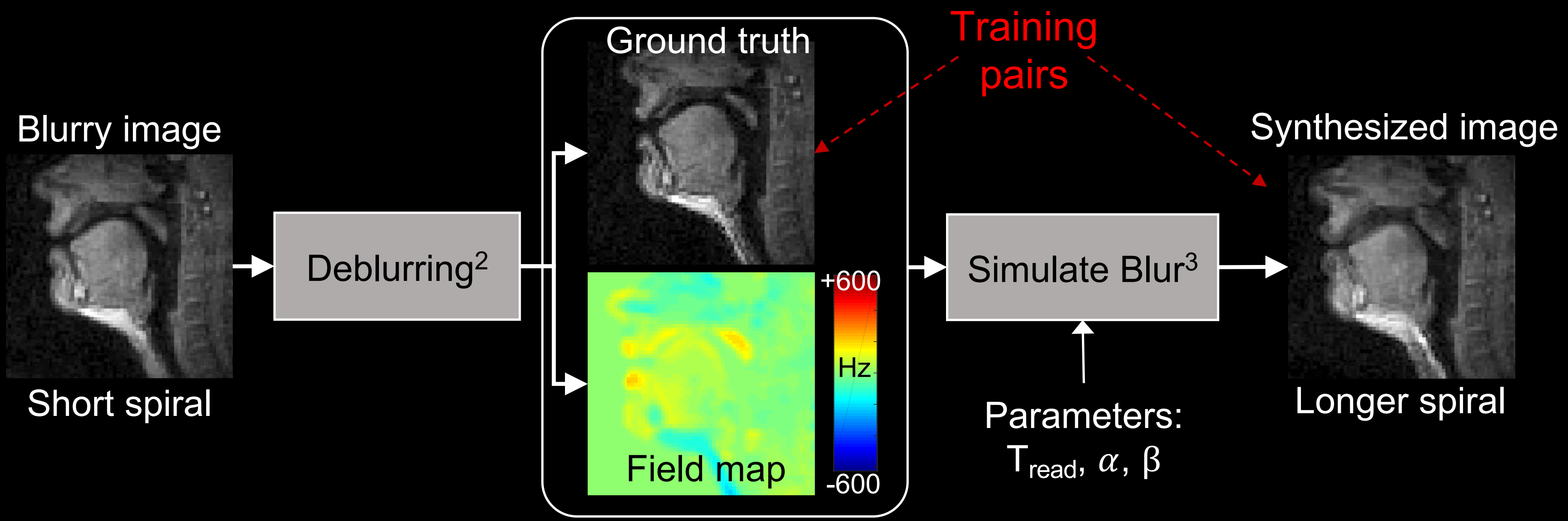
$$F_1' = F_1 \otimes M_1(F_1)$$

$$F_2' = F_2 \otimes M_2(F_2)$$

USC Viterbi School of Engineering

Methods

- **Data**:
  - 2D midsagittal speech spiral RT-MRI scans[1]
  - Training data generation
    - Off-resonance correction[2] and simulation[3]



[1] S Lingala et al, MRM 2017
[2] Y Lim, et al, MRM 2019
[3] Y Lim et al, MRM 2020
[4] M Mathieu M, et al, ICLR 2015
[5] BP Sutton, et al, IEEE TMI 2003

Methods

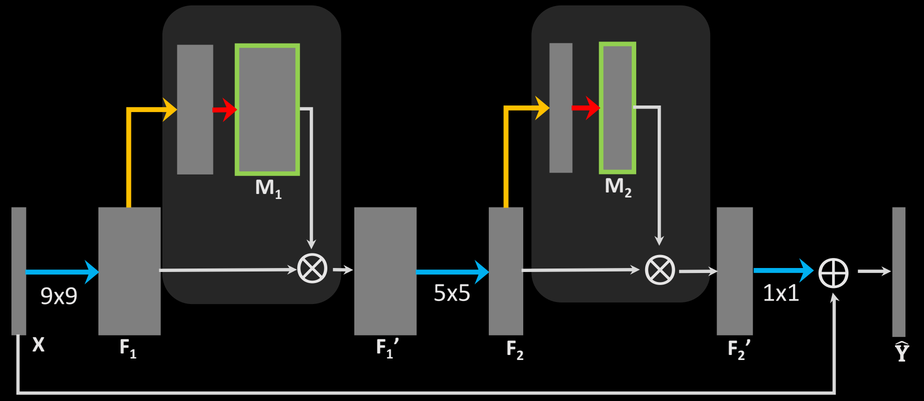
- **Data**:
  - 2D midsagittal speech spiral RT-MRI scans[1]
  - Training data generation
  - Train, validation, and test: 23, 5, and 5 subjects

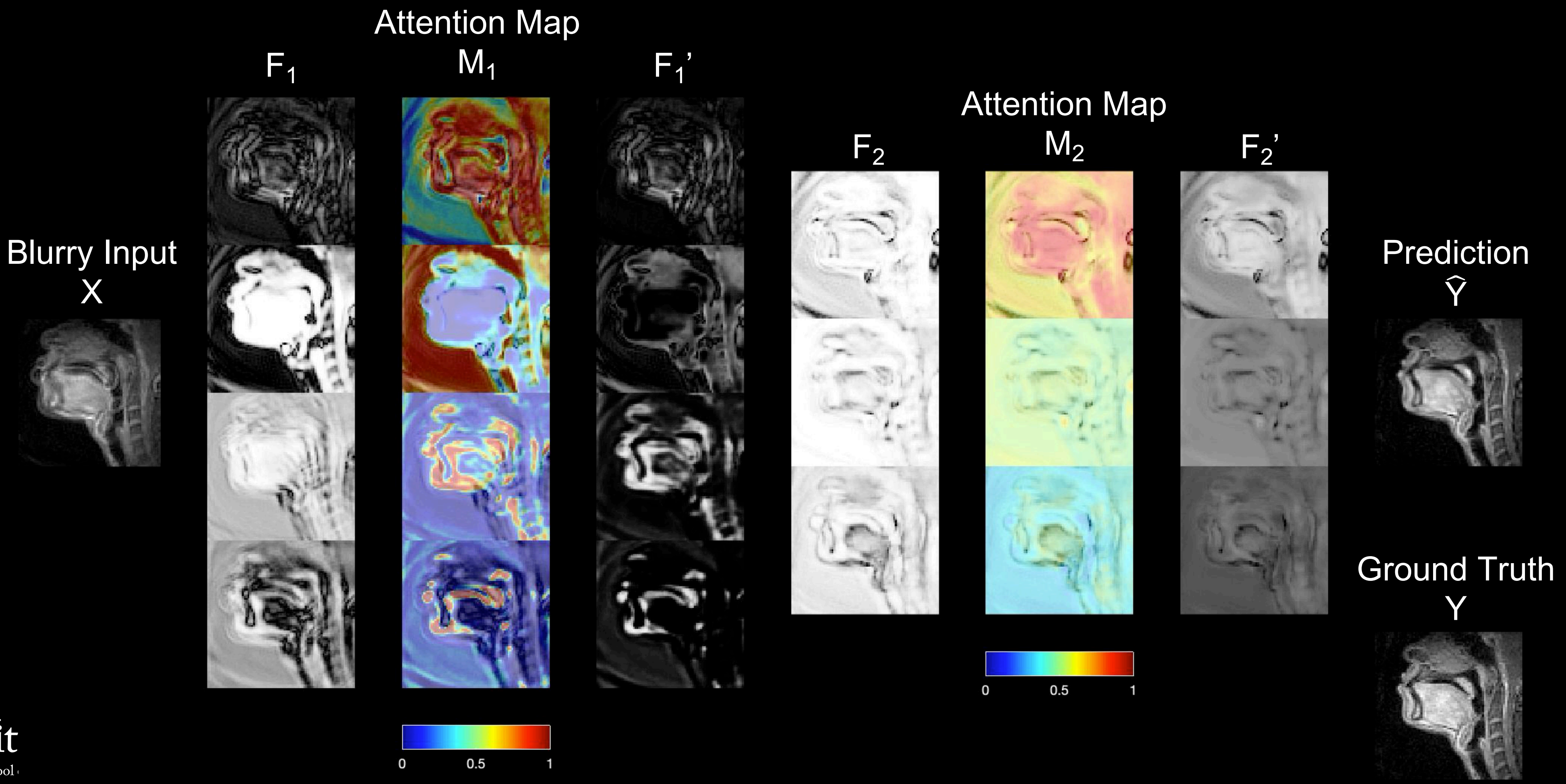
- **Network**:
  - Loss function: $\mathcal{L} = \mathcal{L}_1 + \mathcal{L}_{gdl}$ ($\mathcal{L}_{gdl}$: gradient difference loss[4])
  - Adam optimizer, batch size = 64, epoch = 200

- **Evaluation:**
  - Comparisons: AG-CNN, CNN[3], IR (iterative reconstruction)[5]
  - Quality measures: PSNR, SSIM, HFEN

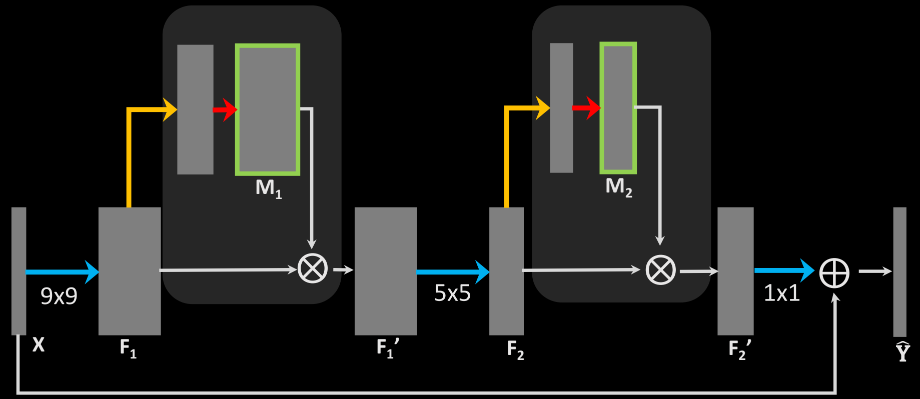
[1] S Lingala et al, MRM 2017
[2] Y Lim, et al, MRM 2019
[3] Y Lim et al, MRM 2020
[4] M Mathieu M, et al, ICLR 2015
[5] BP Sutton, et al, IEEE TMI 2003

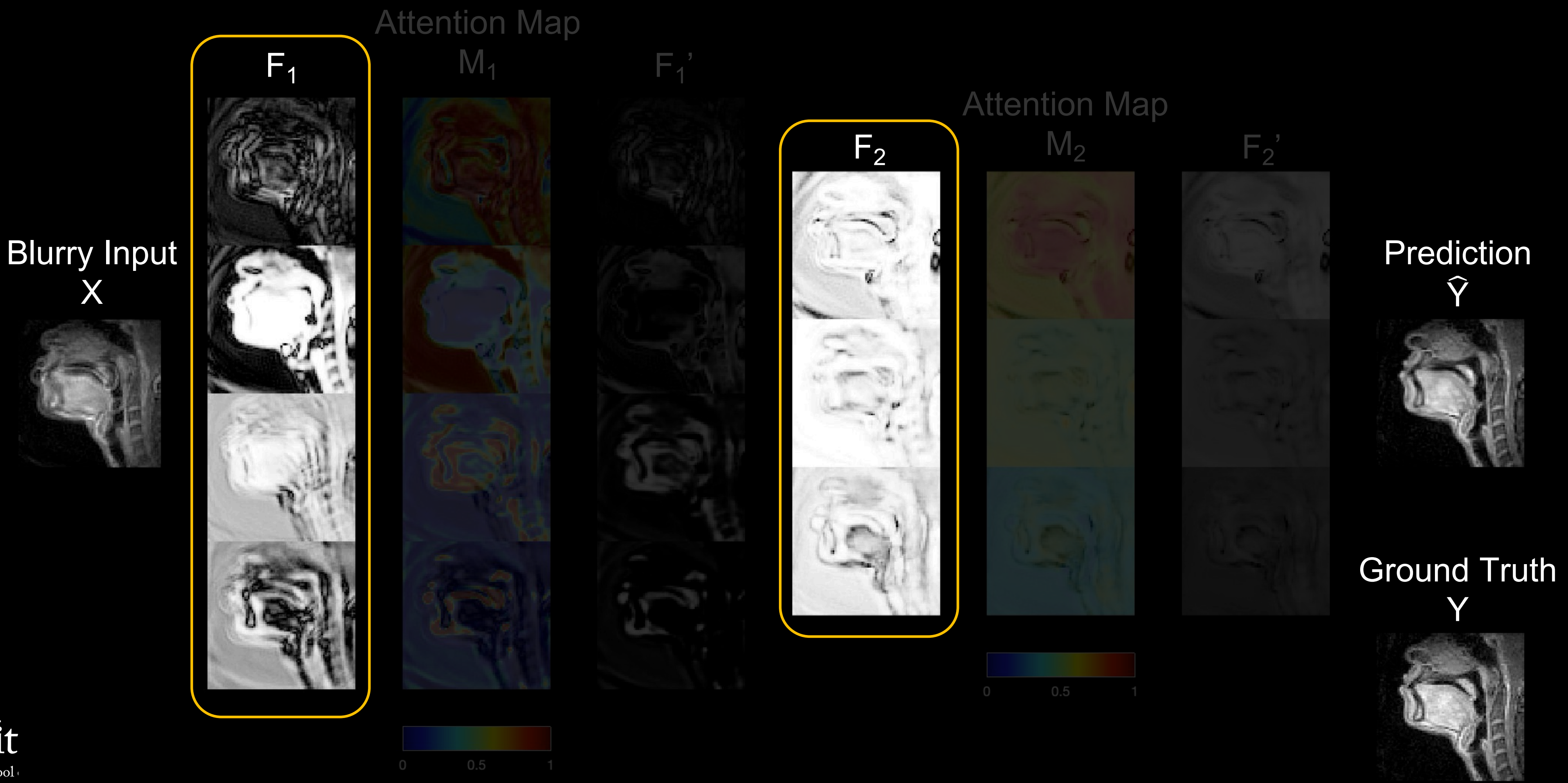
Results: Intermediate Layers





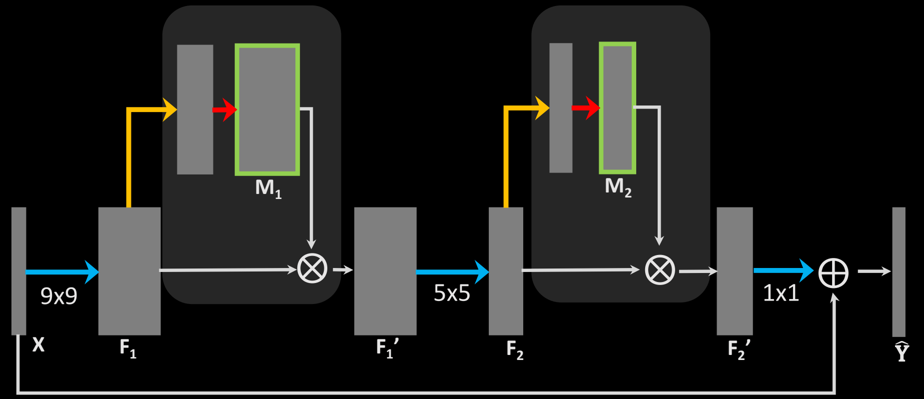
Results: Intermediate Layers

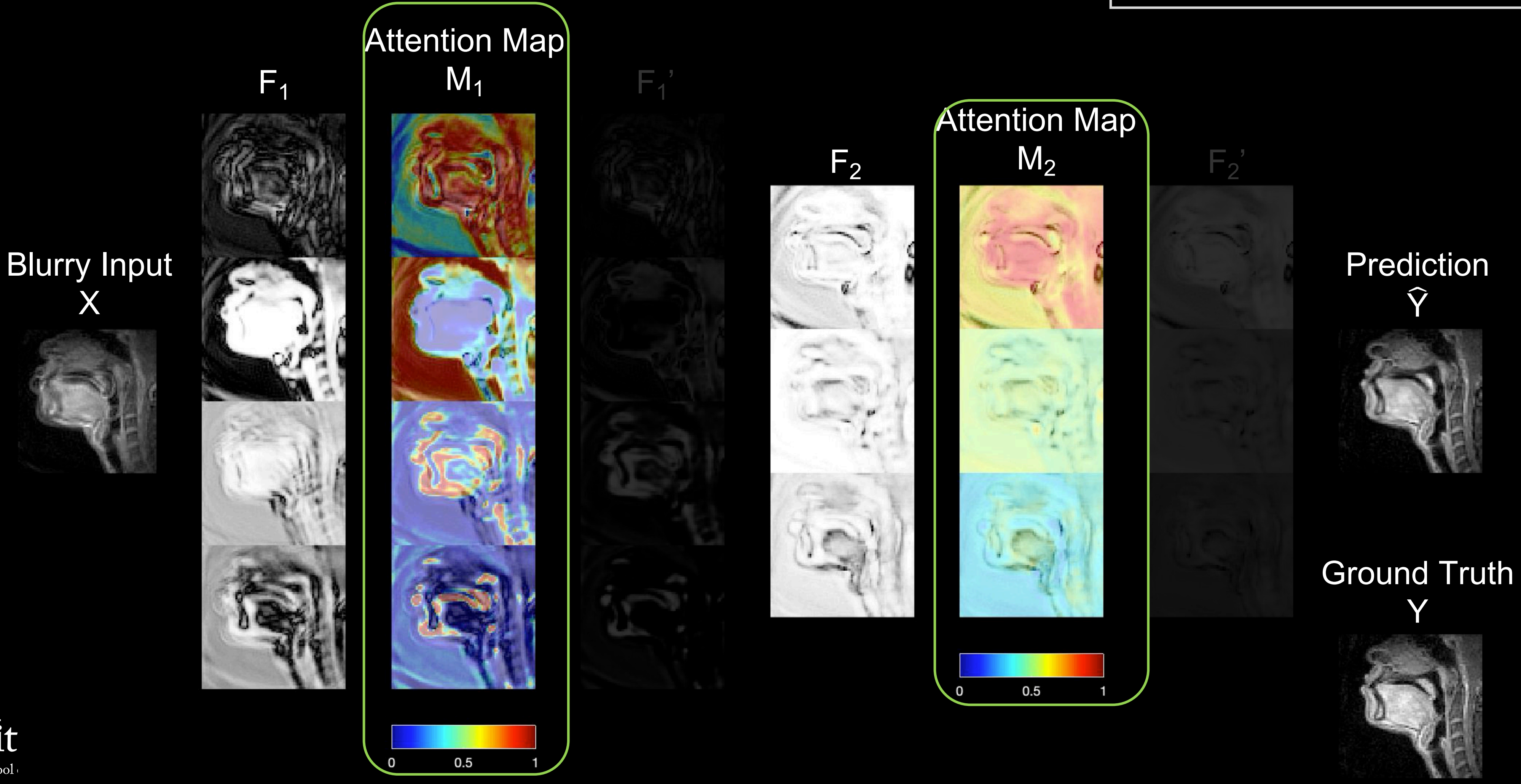




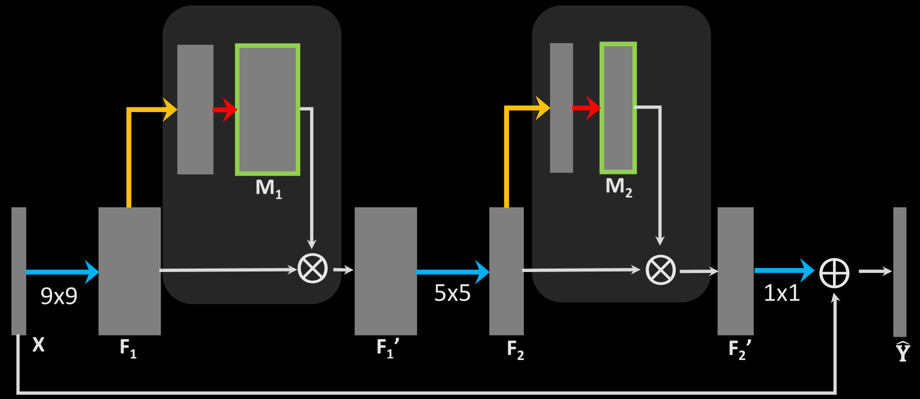
USC Vit
School

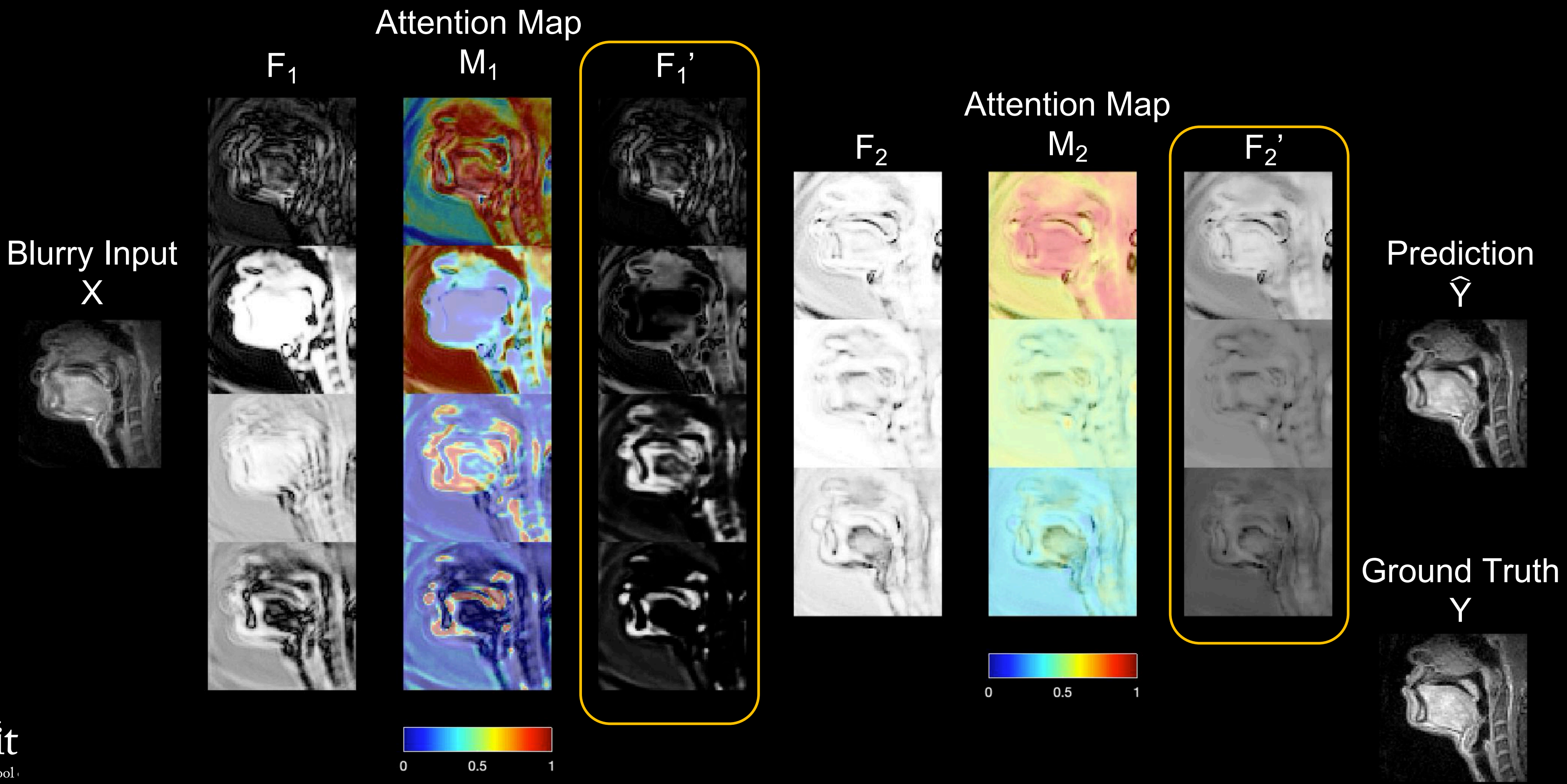
Results: Intermediate Layers





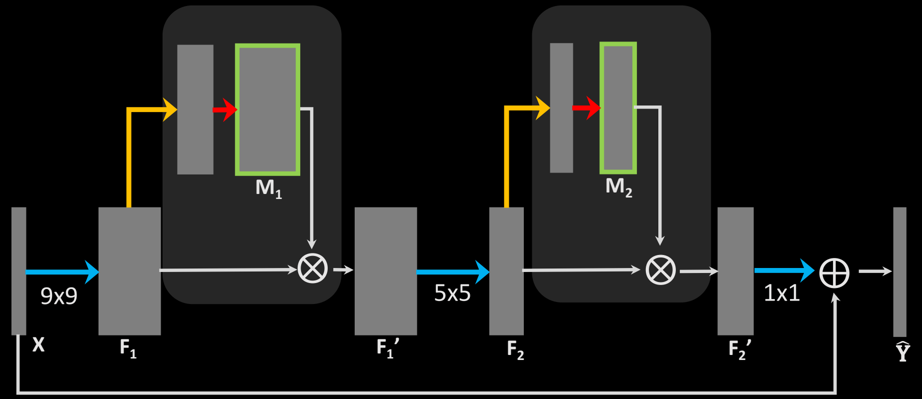
Results: Intermediate Layers

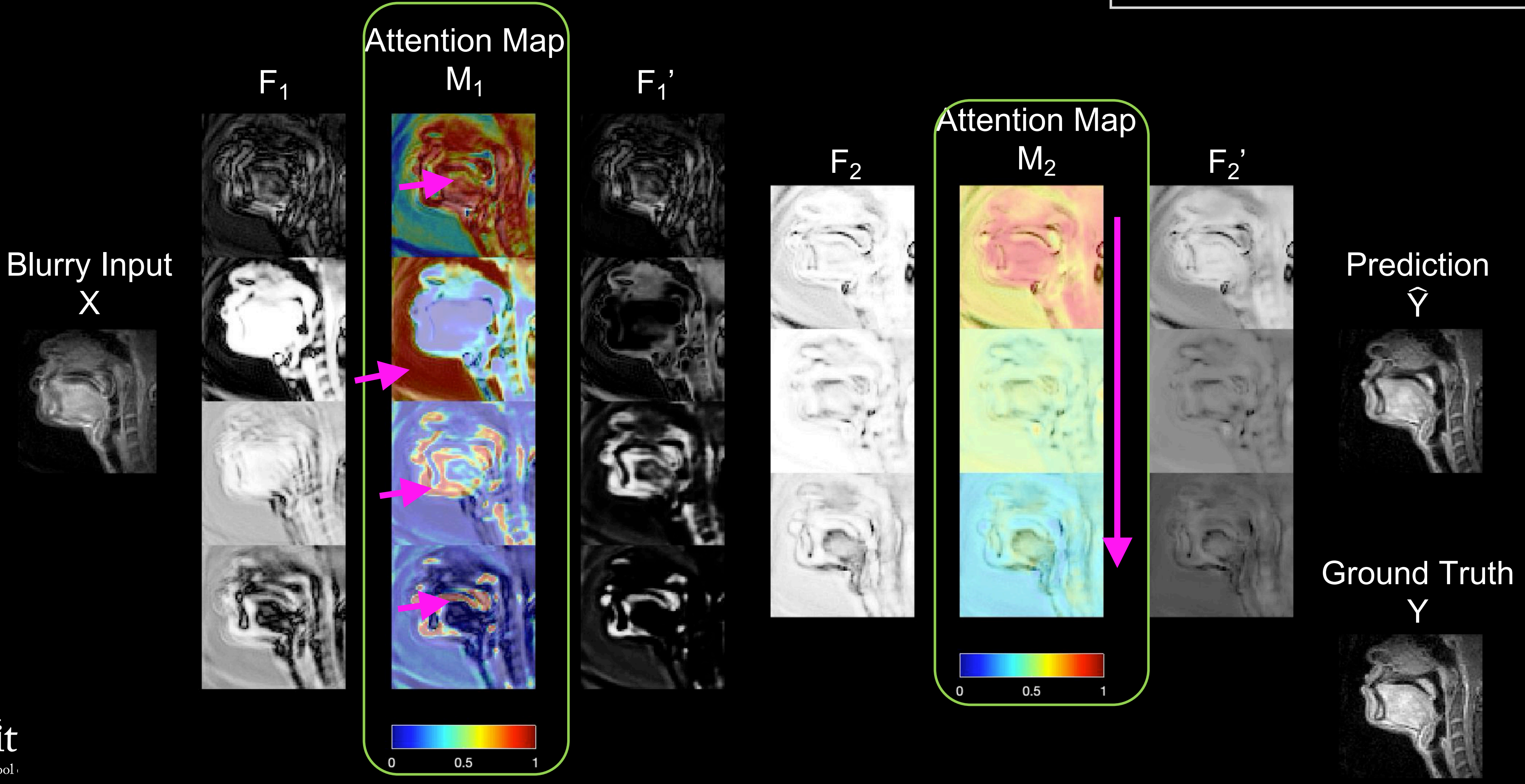




USC Vit
School

Results: Intermediate Layers





Results: Performance vs. Filter size

Architecture	(f_1, f_2)	Params	PSNR	SSIM	HFEN (x100)
CNN (9-5-1)	-	61.7K	29.29	0.944	0.088
+AG	(5,5)	70.7K	30.63	0.959	**0.053**
+AG	(5,3)	70.0K	30.62	**0.959**	0.057
+AG	(5,1)	69.6K	30.61	0.959	0.057
+AG	(3,3)	68.4K	**30.69**	0.958	0.055
+AG	(3,1)	68.1K	30.58	0.958	0.058
(Blurred) Input	-	-	22.16	0.812	0.568

- Improved deblurring performance with less sensitivity to the kernel size but with a slight overhead.
- (f1, f2)=(3, 3) is chosen.

Results: Performance vs. Filter size

filter size in AG module

Architecture	(f_1, f_2)	Params	PSNR	SSIM	HFEN (x100)
CNN (9-5-1)	-	61.7K	29.29	0.944	0.088
+AG	(5,5)	70.7K	30.63	0.959	**0.053**
+AG	(5,3)	70.0K	30.62	**0.959**	0.057
+AG	(5,1)	69.6K	30.61	0.959	0.057
+AG	(3,3)	68.4K	**30.69**	0.958	0.055
+AG	(3,1)	68.1K	30.58	0.958	0.058
(Blurred) Input	-	-	22.16	0.812	0.568

- Improved deblurring performance with less sensitivity to the kernel size but with a slight overhead.
- (f1, f2)=(3, 3) is chosen.

Results: Performance vs. Filter size

filter size in AG module

Architecture	(f_1, f_2)	Params	PSNR	SSIM	HFEN (x100)
CNN (9-5-1)	-	61.7K	29.29	0.944	0.088
+AG	(5,5)	70.7K	30.63	0.959	**0.053**
+AG	(5,3)	70.0K	30.62	**0.959**	0.057
+AG	(5,1)	69.6K	30.61	0.959	0.057
+AG	(3,3)	68.4K	**30.69**	0.958	0.055
+AG	(3,1)	68.1K	30.58	0.958	0.058
(Blurred) Input	-	-	22.16	0.812	0.568

- Improved deblurring performance with less sensitivity to the kernel size but with a slight overhead.
- (f1, f2)=(3, 3) is chosen.

Results: Performance vs. Filter size

filter size in AG module

Architecture	(f_1, f_2)	Params	PSNR	SSIM	HFEN (x100)
CNN (9-5-1)	-	61.7K	29.29	0.944	0.088
+AG	(5,5)	70.7K	30.63	0.959	**0.053**
+AG	(5,3)	70.0K	30.62	**0.959**	0.057
+AG	(5,1)	69.6K	30.61	0.959	0.057
+AG	(3,3)	68.4K	**30.69**	0.958	0.055
+AG	(3,1)	68.1K	30.58	0.958	0.058
(Blurred) Input	-	-	22.16	0.812	0.568

- Improved deblurring performance with less sensitivity to the kernel size but with a slight overhead.
- (f1, f2)=(3, 3) is chosen.

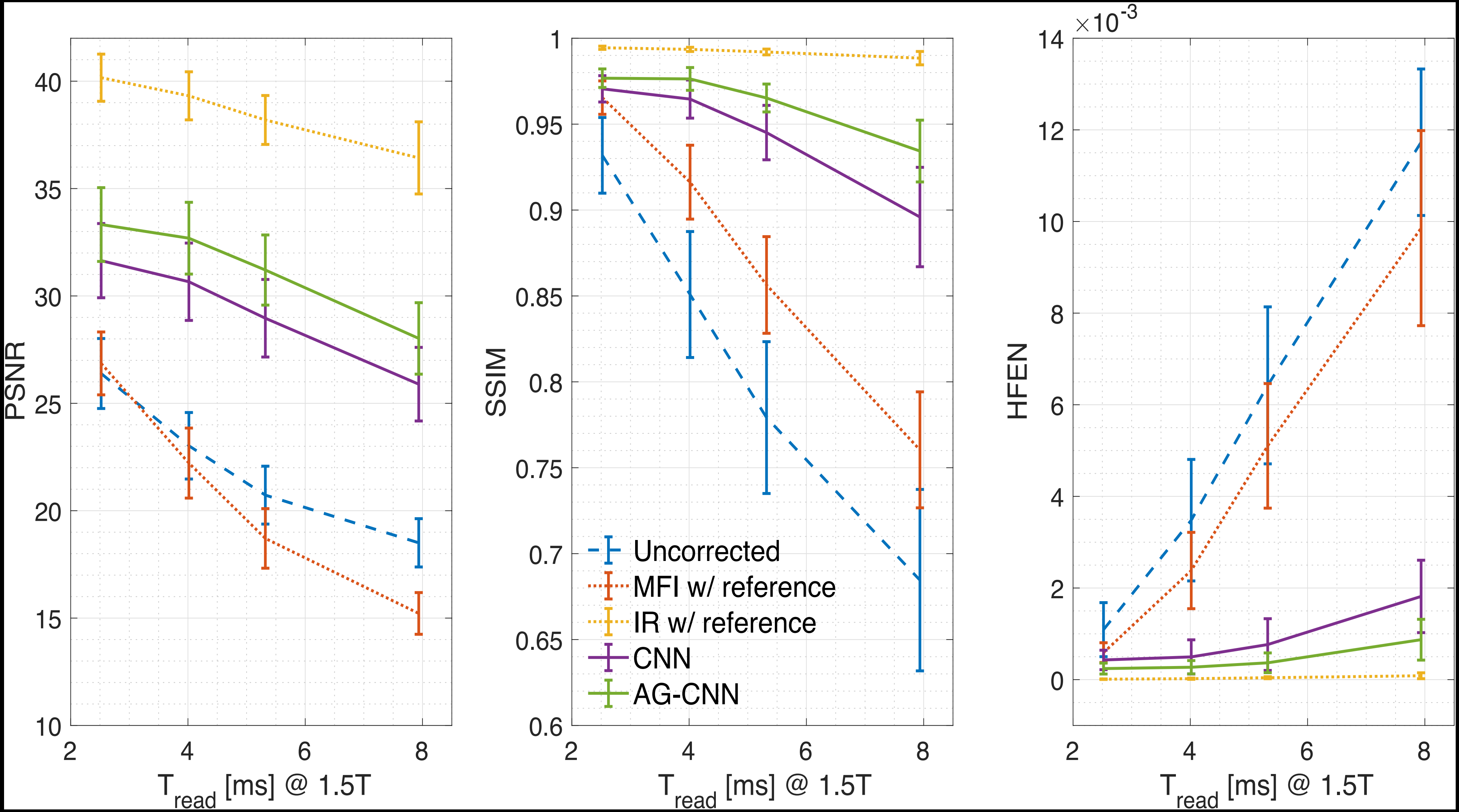
Results: Performance vs. Filter size

filter size in AG module

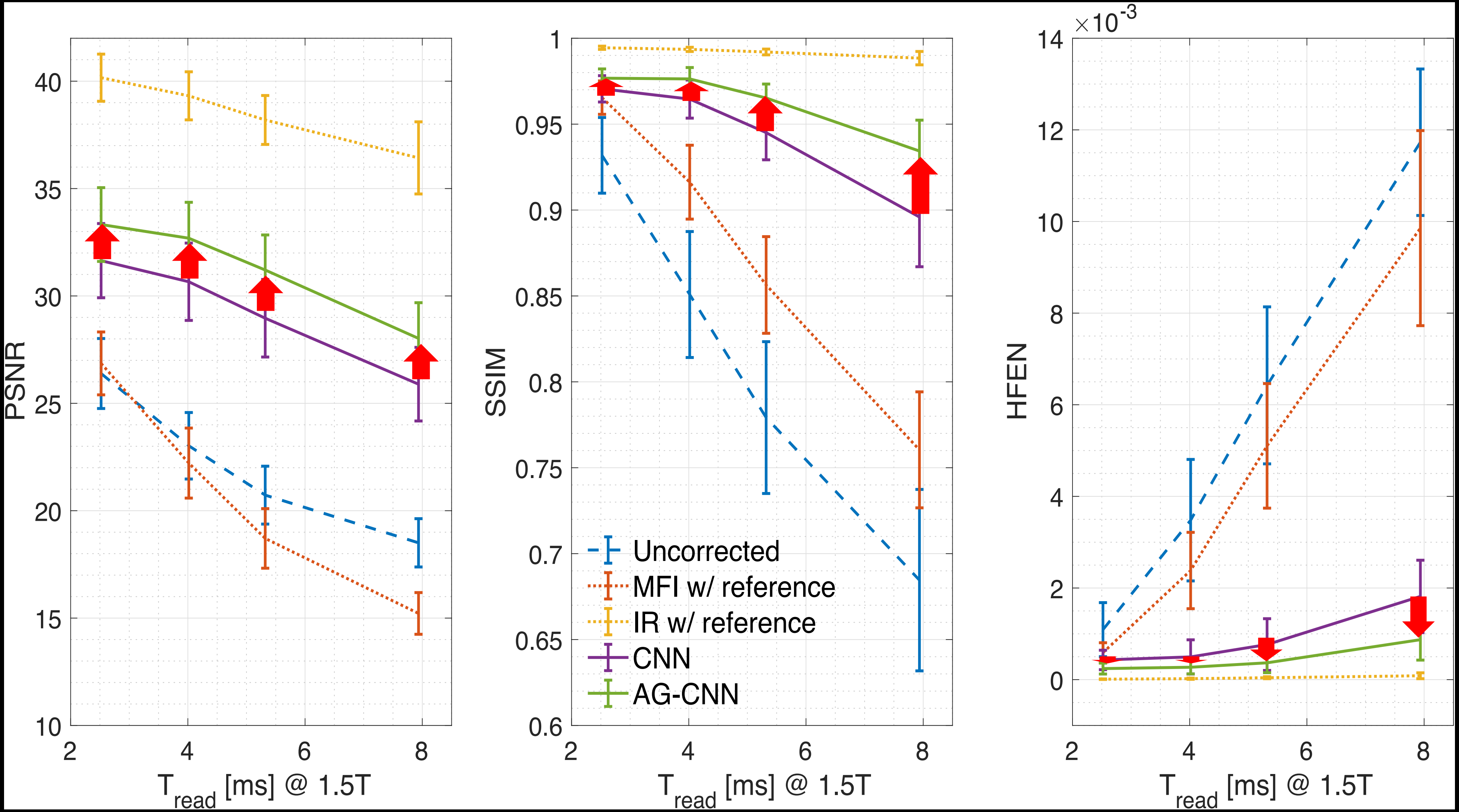
Architecture	(f_1, f_2)	Params	PSNR	SSIM	HFEN (x100)
CNN (9-5-1)	-	61.7K	29.29	0.944	0.088
+AG	(5,5)	70.7K	30.63	0.959	**0.053**
+AG	(5,3)	70.0K	30.62	**0.959**	0.057
+AG	(5,1)	69.6K	30.61	0.959	0.057
+AG	(3,3)	68.4K	**30.69**	0.958	0.055
+AG	(3,1)	68.1K	30.58	0.958	0.058
(Blurred) Input	-	-	22.16	0.812	0.568

- Improved deblurring performance with less sensitivity to the kernel size but with a slight overhead.
- (f1, f2)=(3, 3) is chosen.

Results: Comparisons

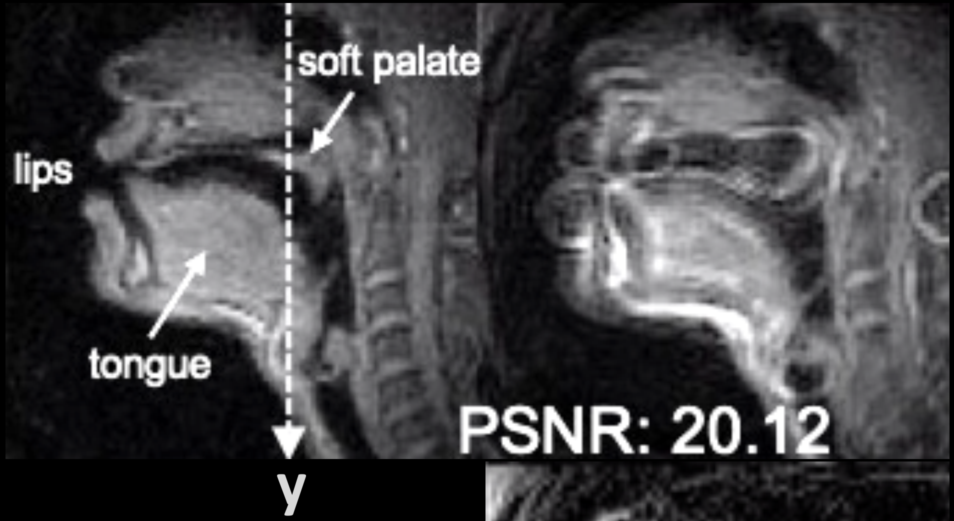


Results: Comparisons



Results: Test Data

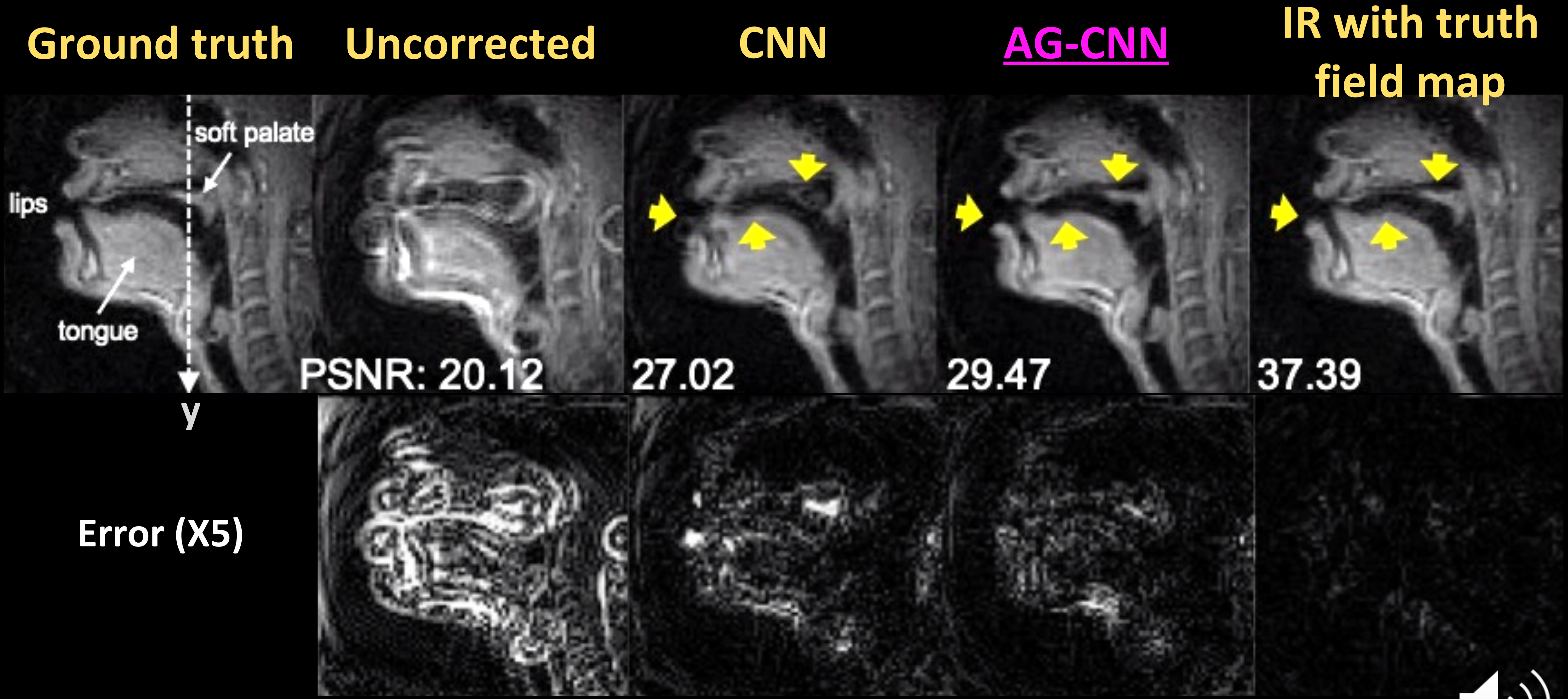
Ground truth **Uncorrected**



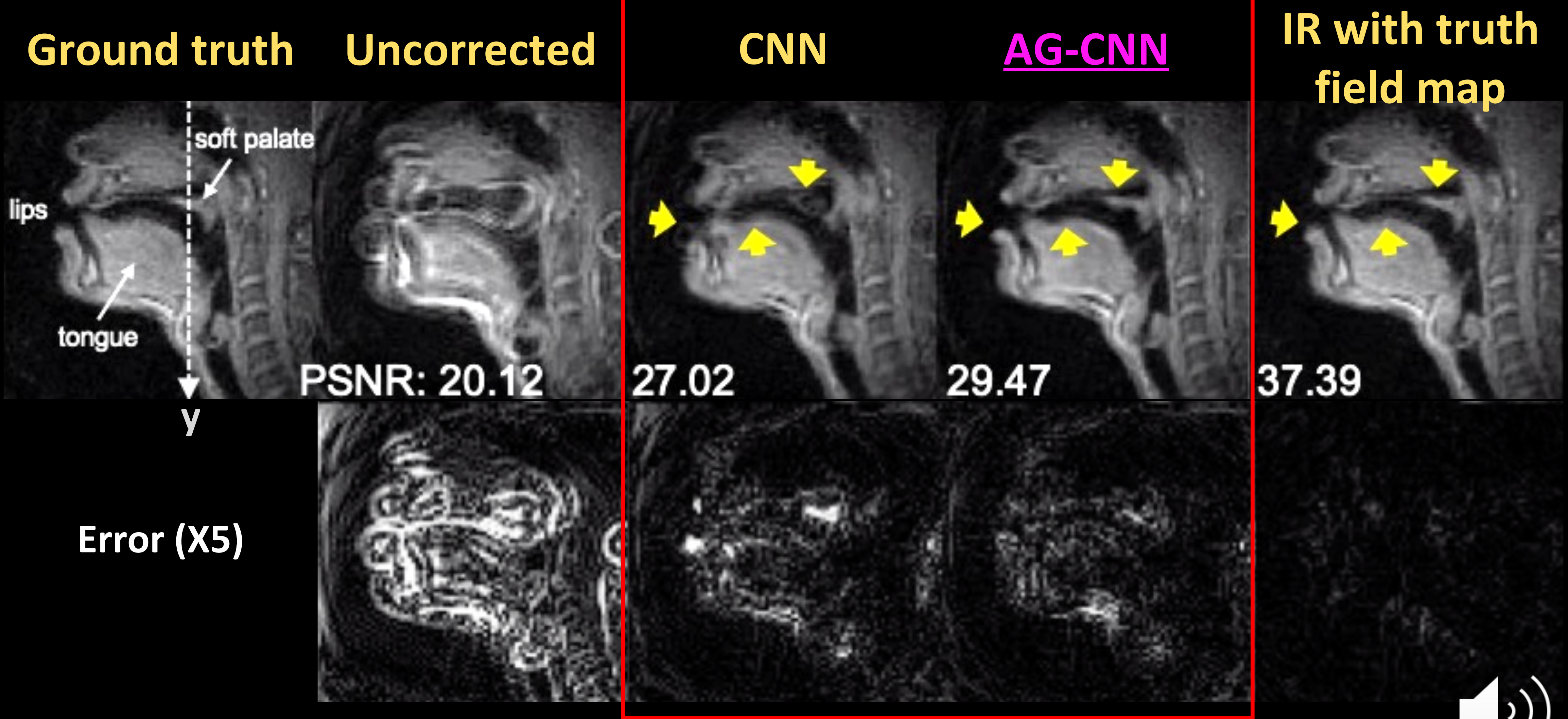
Error (X5)

USC Viterbi
School of Engineering

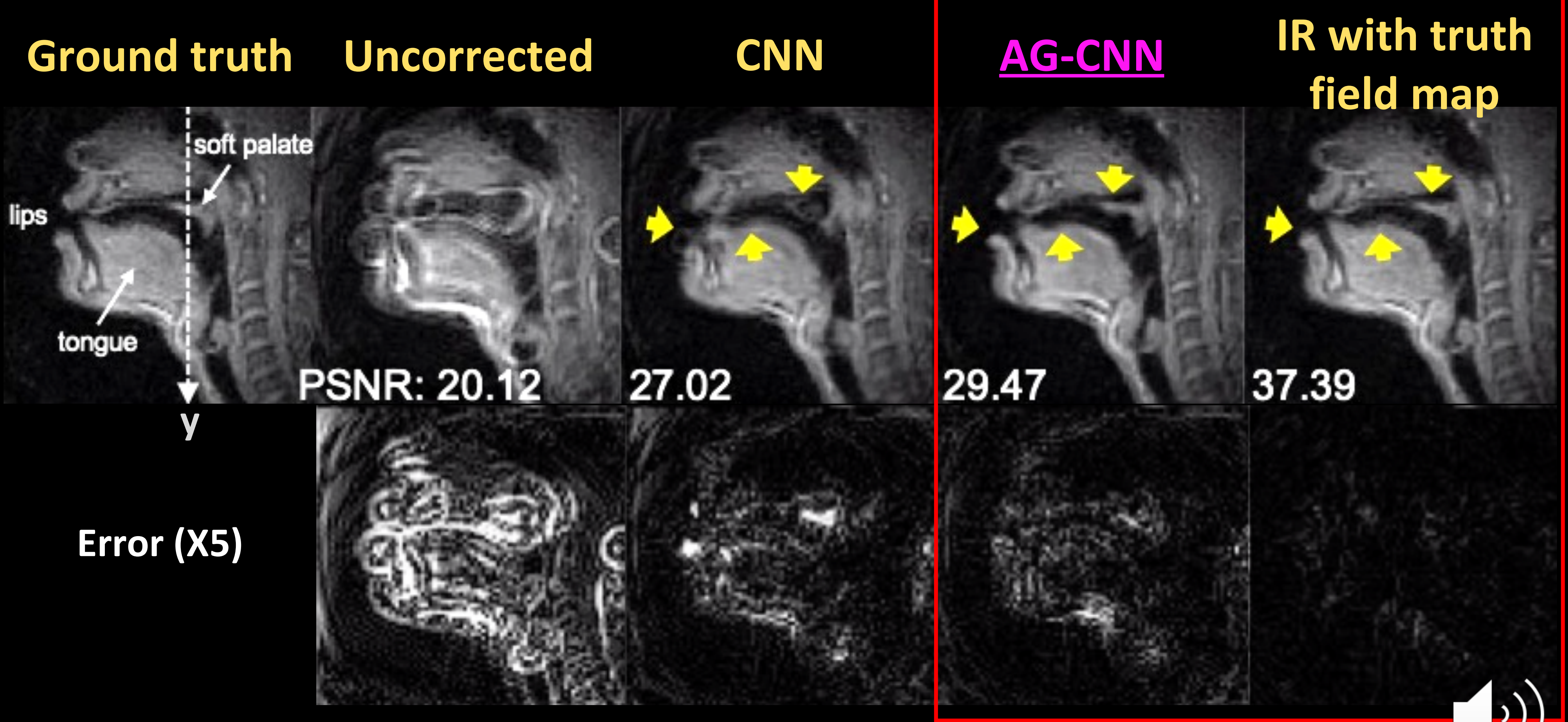
Results: Test Data



Results: Test Data

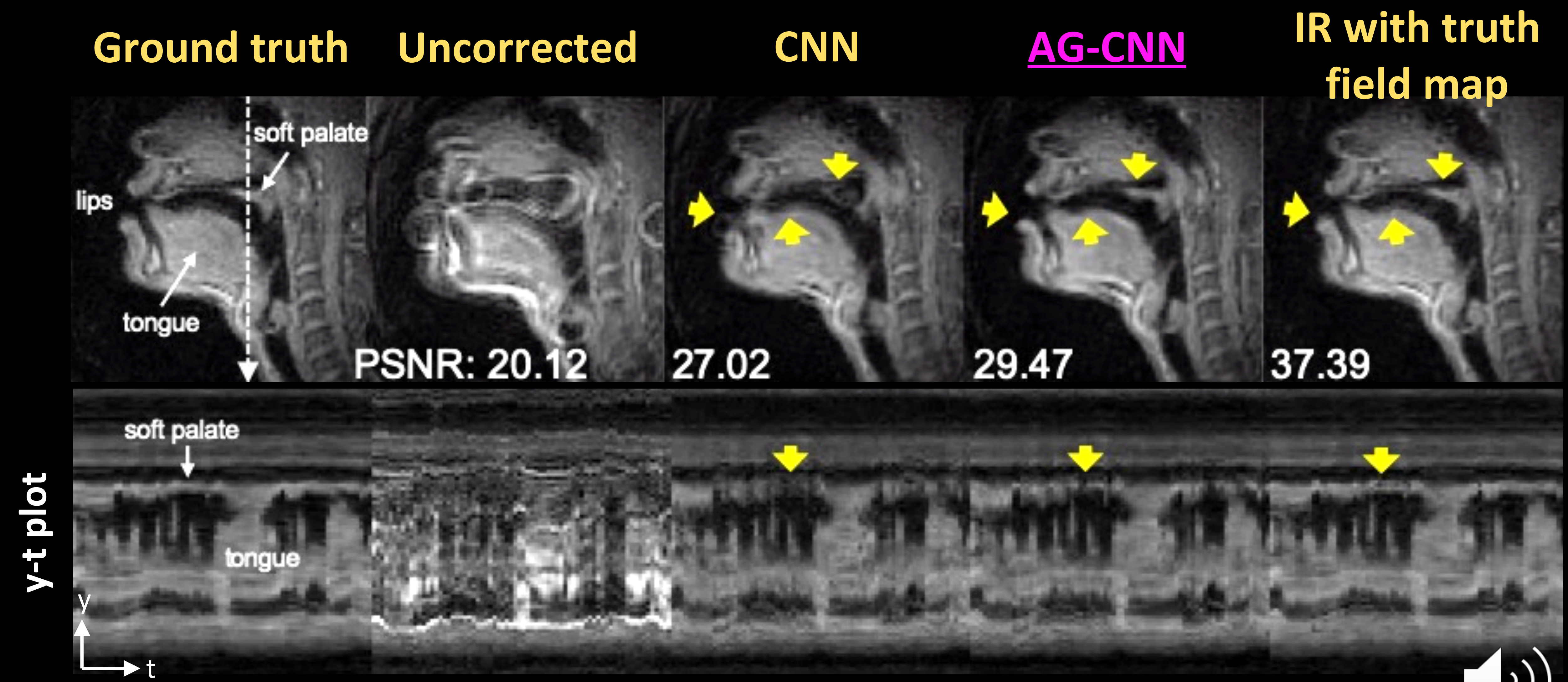


Results: Test Data



USC Viterbi School of Engineering

Results: Test Data



Conclusion

- We develop the AG-CNN-based deblurring method for spiral RT-MRI in speech production.

- AG module could capture spatial and channel relationships of filtered outputs and improves deblurring performance with a slight overhead.

- An extensive comparison with existing attention approaches applicable to this task remains as future work.

Paper #1005, ISMRM 2020

Attention-gated convolutional neural networks for off-resonance correction of spiral real-time MRI

Yongwan Lim, Shrikanth S. Narayanan, Krishna S. Nayak

Thank you for your attention!

If you have any questions, please contact me: YONGWANL@USC.EDU

USC Viterbi
School of Engineering