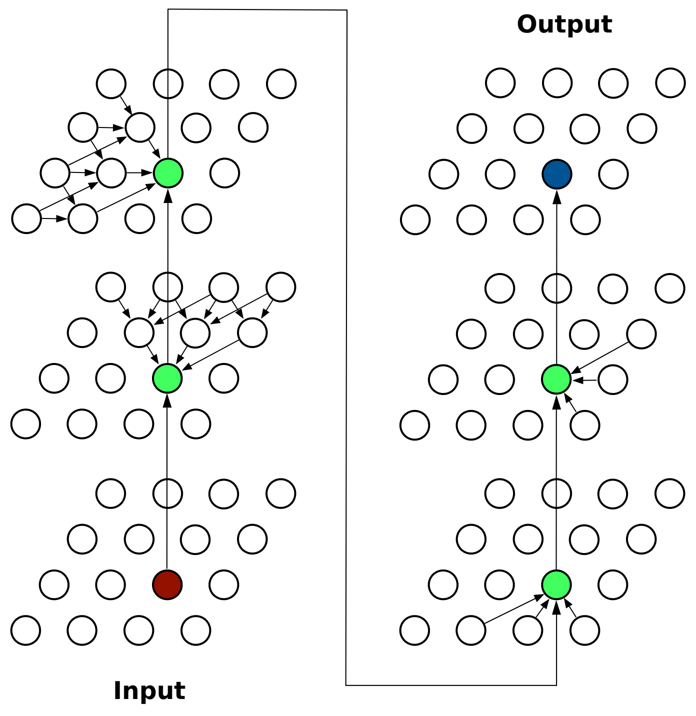


Spatial Dependency Networks

Neural Layers for Improved Generative Image Modeling

Đorđe Miladinović, Aleksandar Stanić, Stefan Bauer, Jürgen Schmidhuber & Joachim M. Buhmann

GitHub @djordjemila

@djordjemila

ETH Zürich

USI/SUPSI IDSIA Istituto Dalle Molle di studi sull'intelligenza artificiale

USI Università della Svizzera italiana

Max Planck Institute for Intelligent Systems

ICLR 2021

Motivation: Image Generation

(Miladinović et al. 2021)

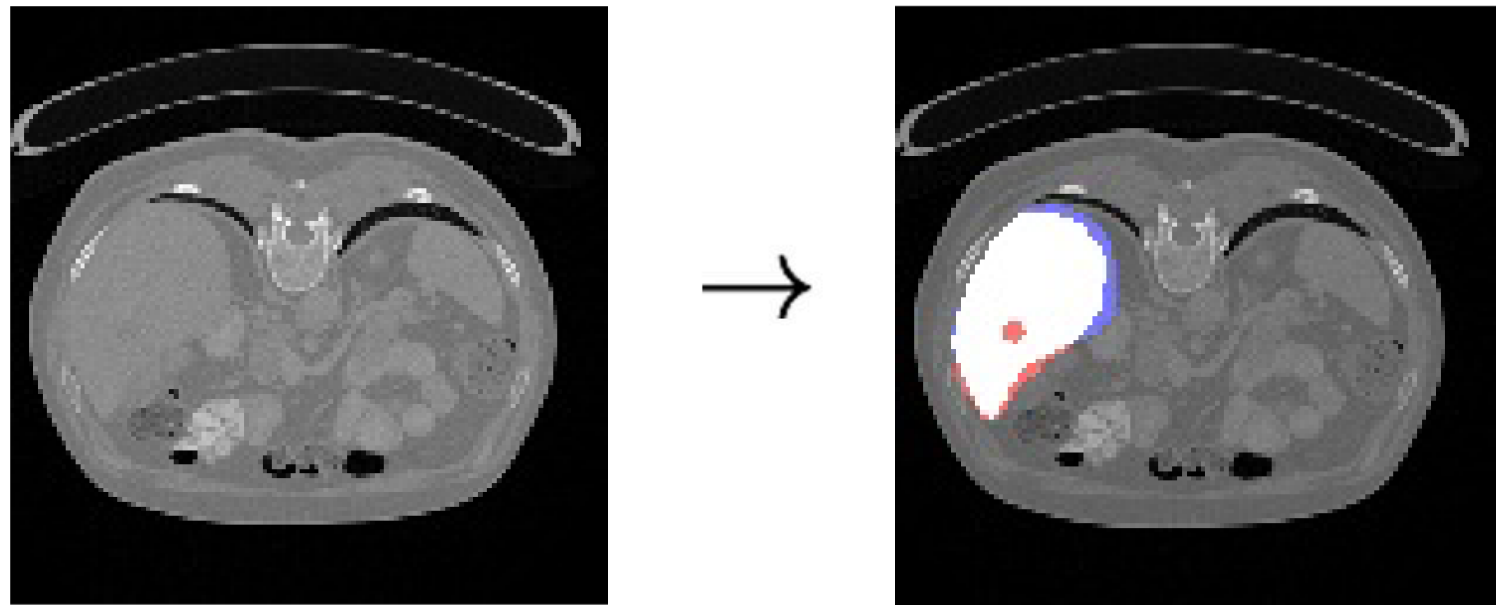
Motivation: Image Generation

realistic image synthesis **[this work]**

(Miladinović et al. 2021)

Motivation: Image Generation

realistic image synthesis **[this work]**

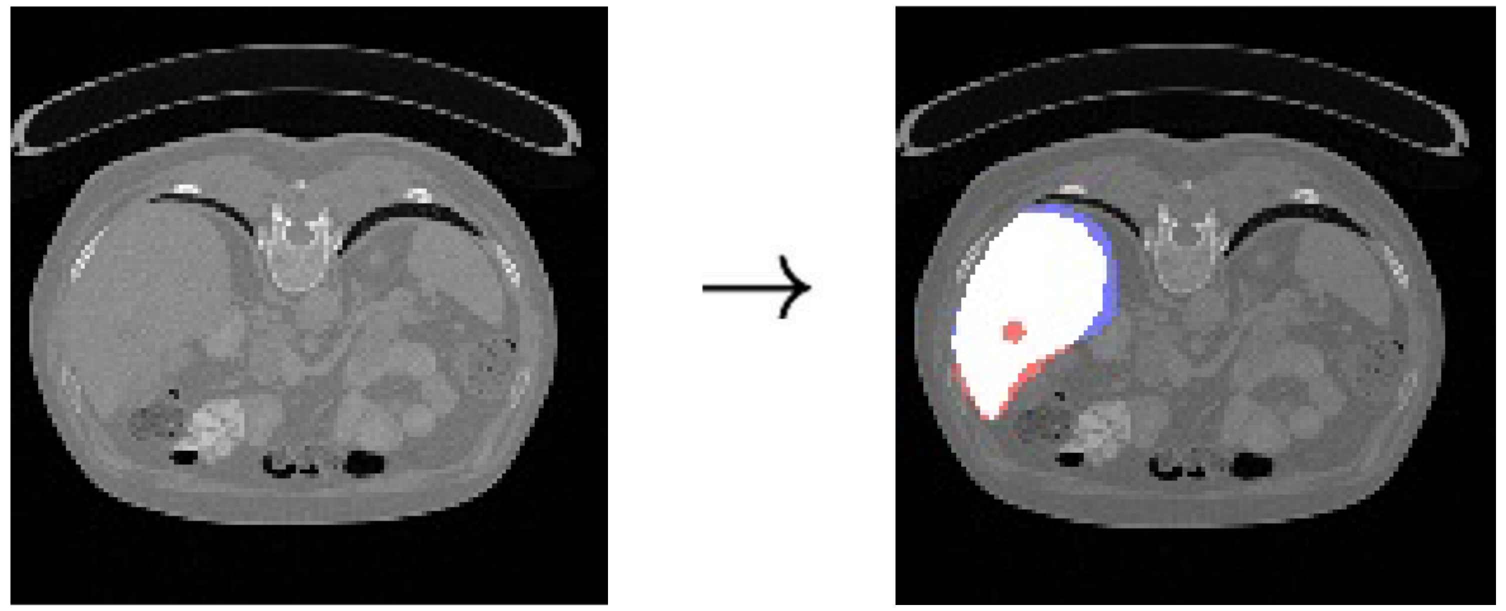


segmentation [Carvalho et al. 2021] **[follow-up]**

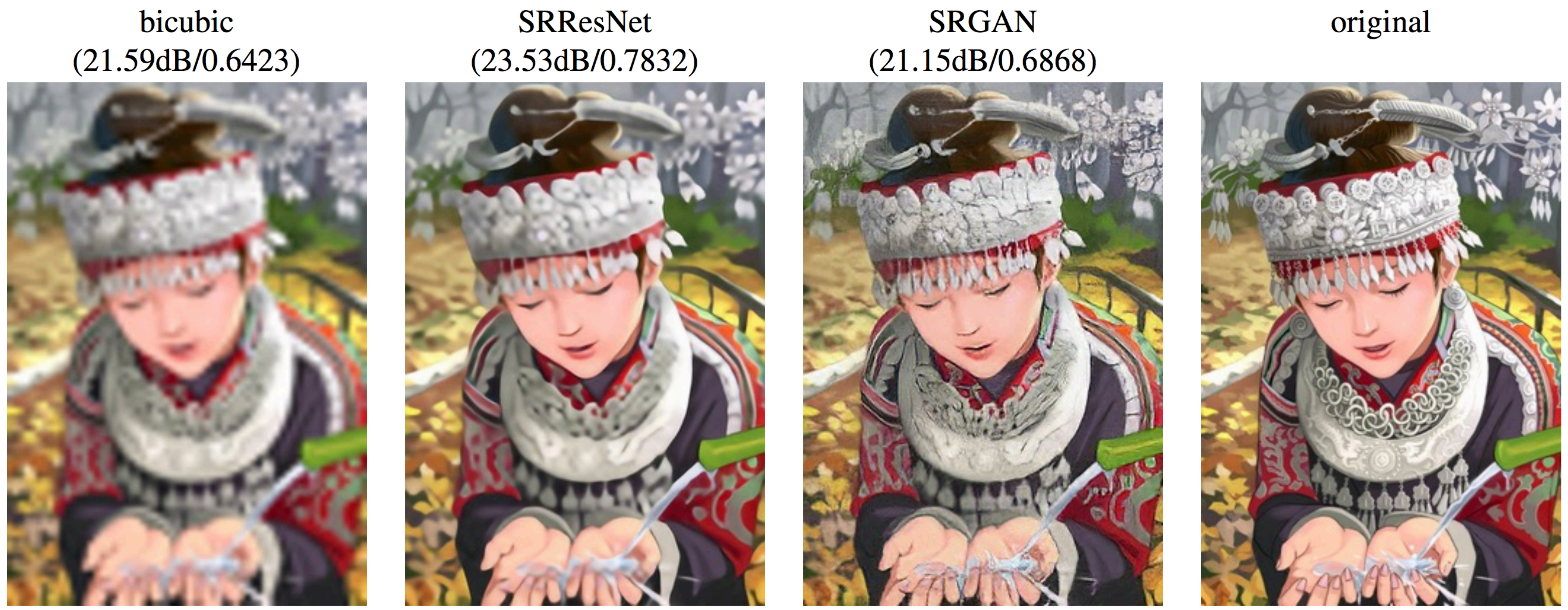
(Miladinović et al. 2021)

Motivation: Image Generation

realistic image synthesis **[this work]**



segmentation [Carvalho et al. 2021] **[follow-up]**

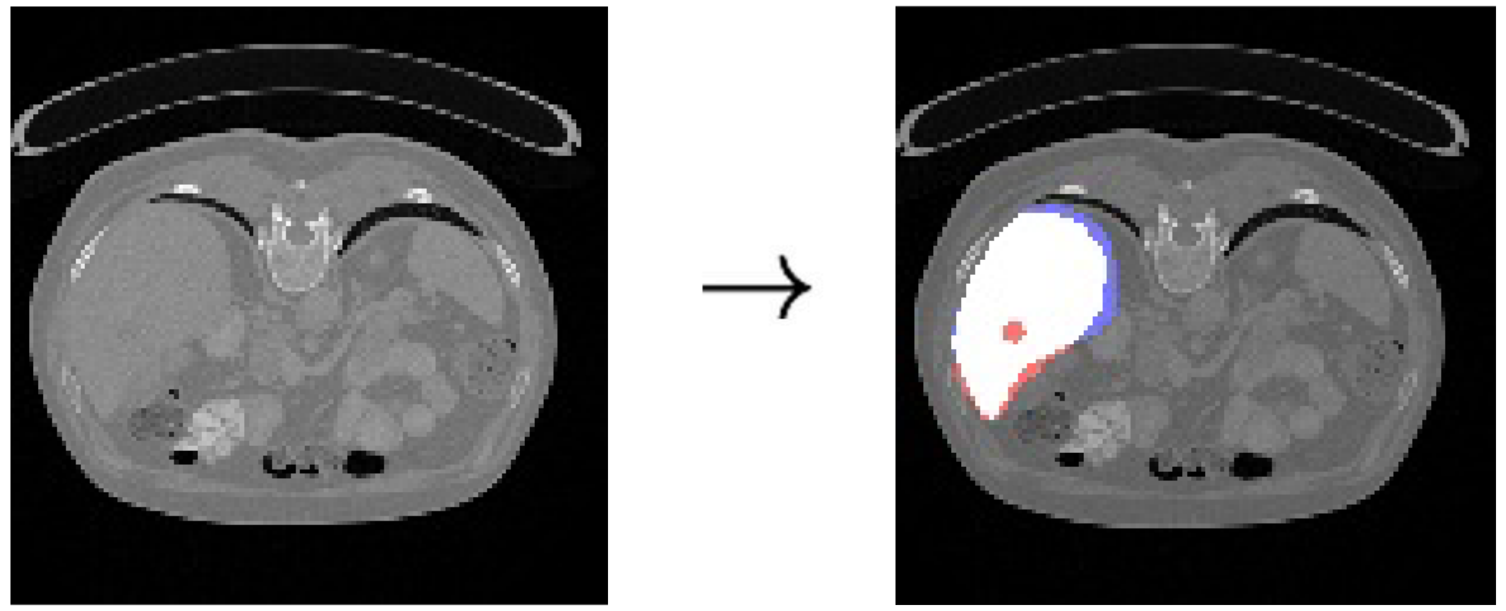


super-resolution [Ledig et al. 2017]

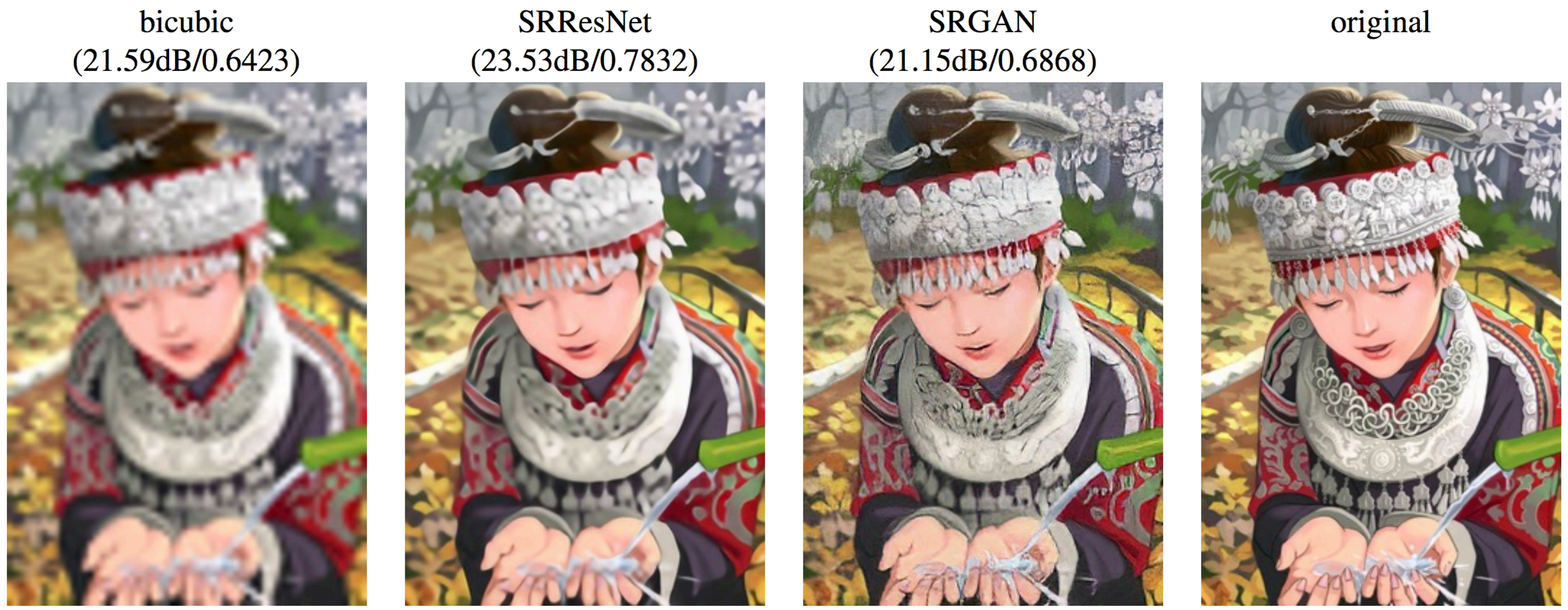
(Miladinović et al. 2021)

Motivation: Image Generation

realistic image synthesis **[this work]**



segmentation [Carvalho et al. 2021] **[follow-up]**



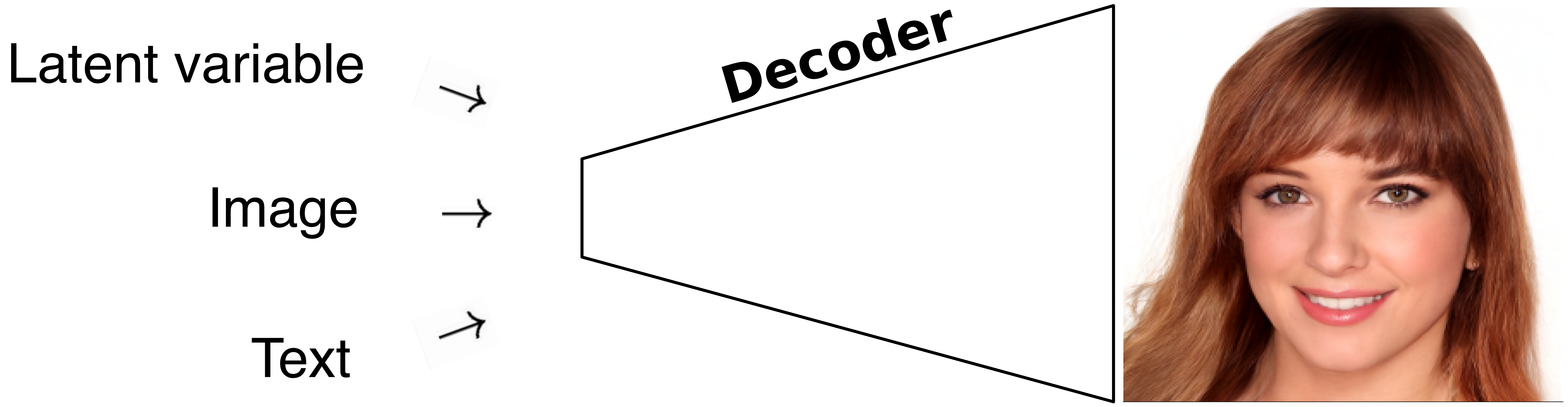
super-resolution [Ledig et al. 2017]

image-to-image translation

[Zhu et al. 2017]

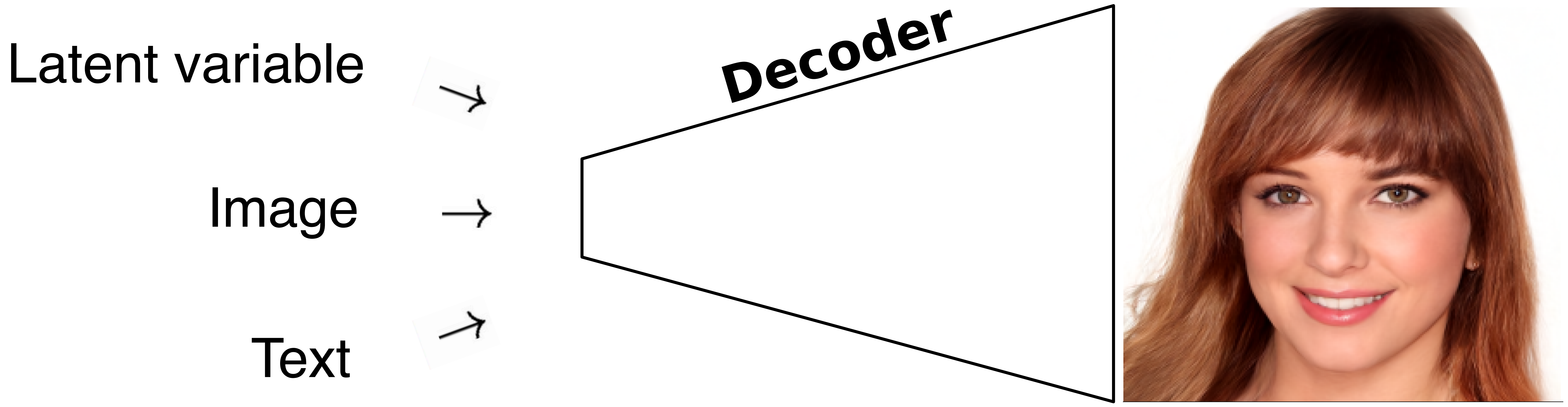
(Miladinović et al. 2021)

Image Generation requires an Image Decoder/Generator



(Miladinović et al. 2021)

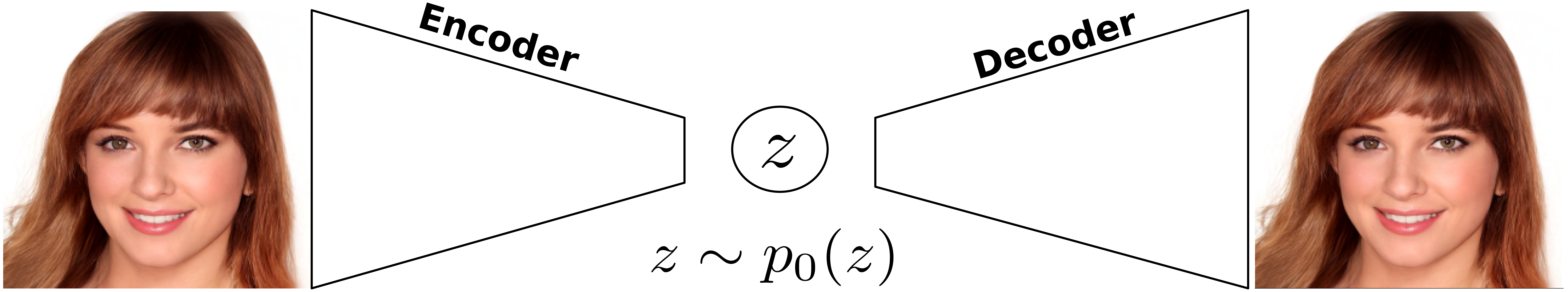
Image Generation requires an Image Decoder/Generator



Focus of this paper: **Variational Autoencoders** [Kingma et al. 2013].

(Miladinović et al. 2021)

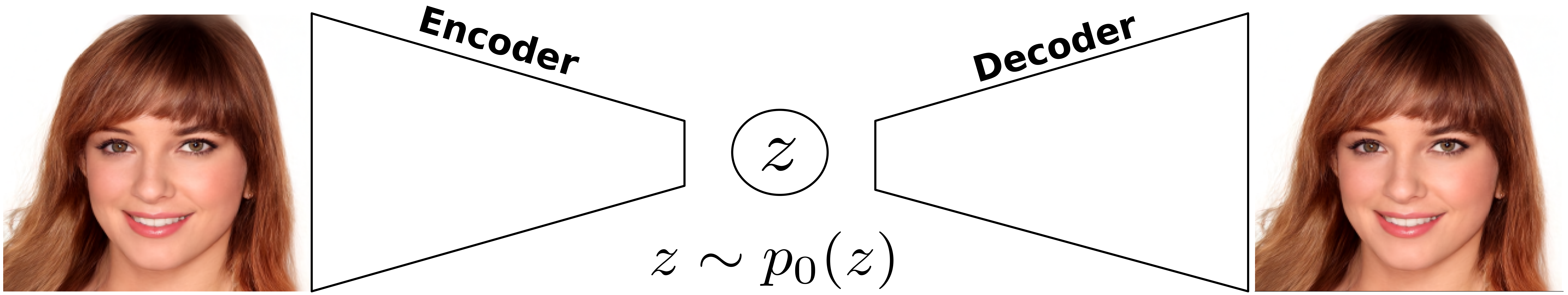
Image Generation requires an Image Decoder/Generator



Focus of this paper: **Variational Autoencoders** [Kingma et al. 2013].

(Miladinović et al. 2021)

Image Generation requires an Image Decoder/Generator



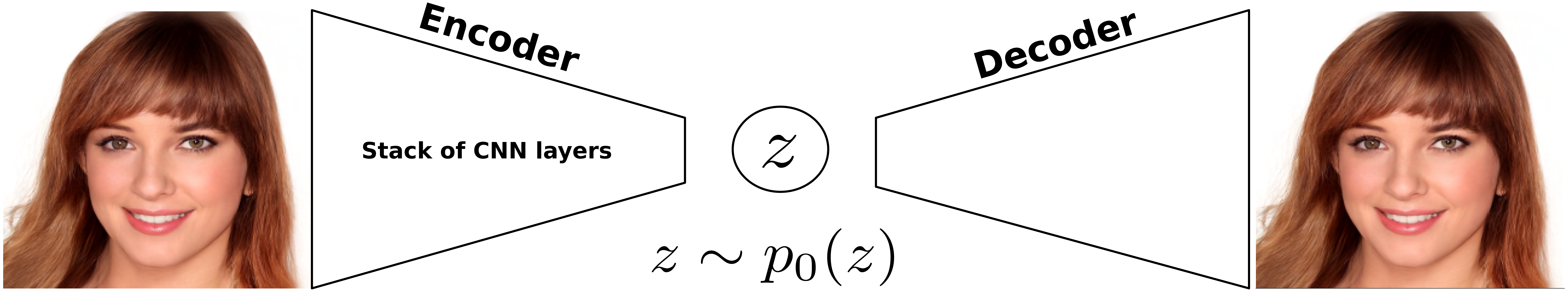
Focus of this paper: **Variational Autoencoders** [Kingma et al. 2013].

Our hypothesis:

- Convolutional network (CNN) — good in extracting patterns from images — encoding.
- **Spatial dependency network (SDN)** — better in synthesizing images — decoding.

(Miladinović et al. 2021)

Image Generation requires an Image Decoder/Generator



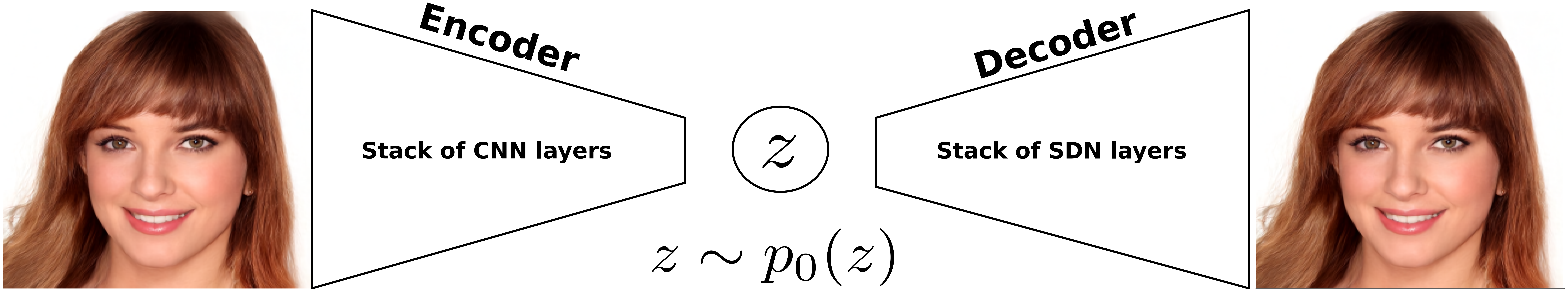
Focus of this paper: **Variational Autoencoders** [Kingma et al. 2013].

Our hypothesis:

- Convolutional network (CNN) — good in extracting patterns from images — encoding.
- **Spatial dependency network (SDN)** — better in synthesizing images — decoding.

(Miladinović et al. 2021)

Image Generation requires an Image Decoder/Generator



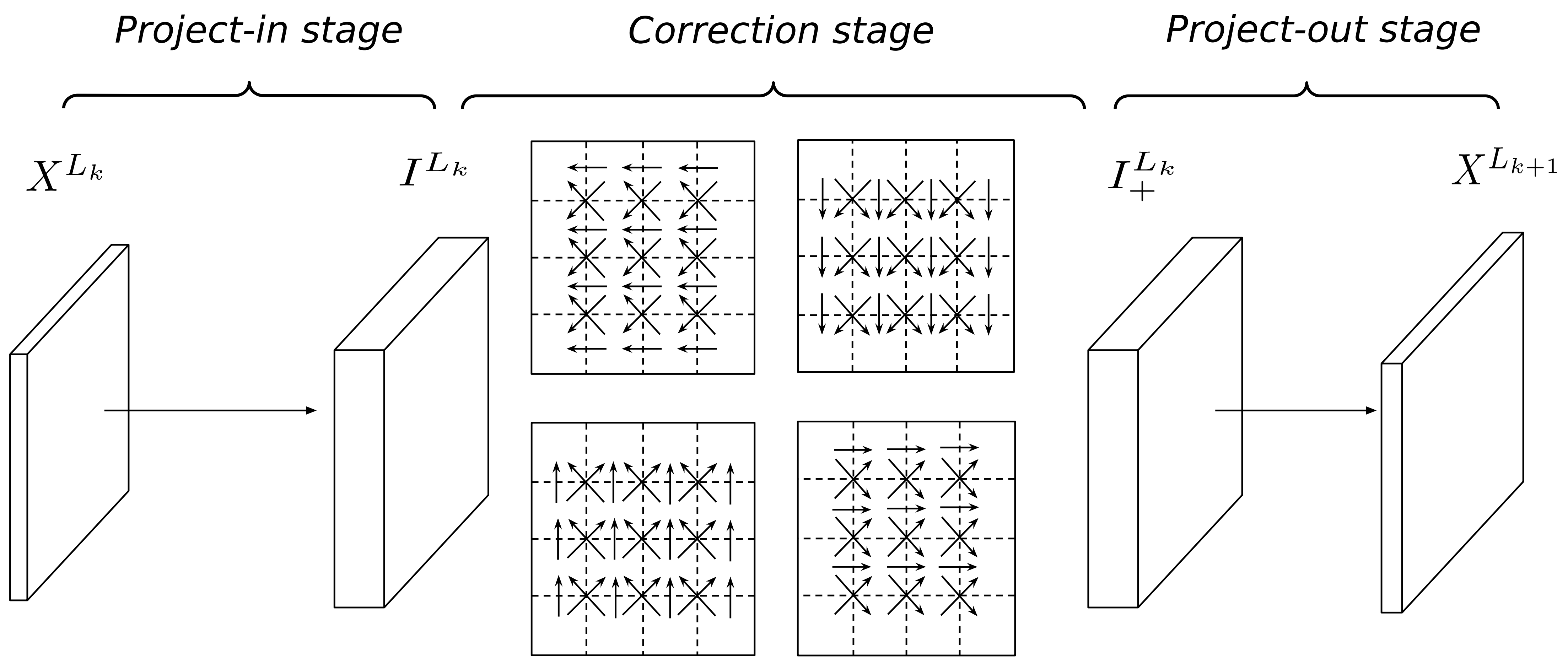
Focus of this paper: **Variational Autoencoders** [Kingma et al. 2013].

Our hypothesis:

- Convolutional network (CNN) — good in extracting patterns from images — encoding.
- **Spatial dependency network (SDN)** — better in synthesizing images — decoding.

(Miladinović et al. 2021)

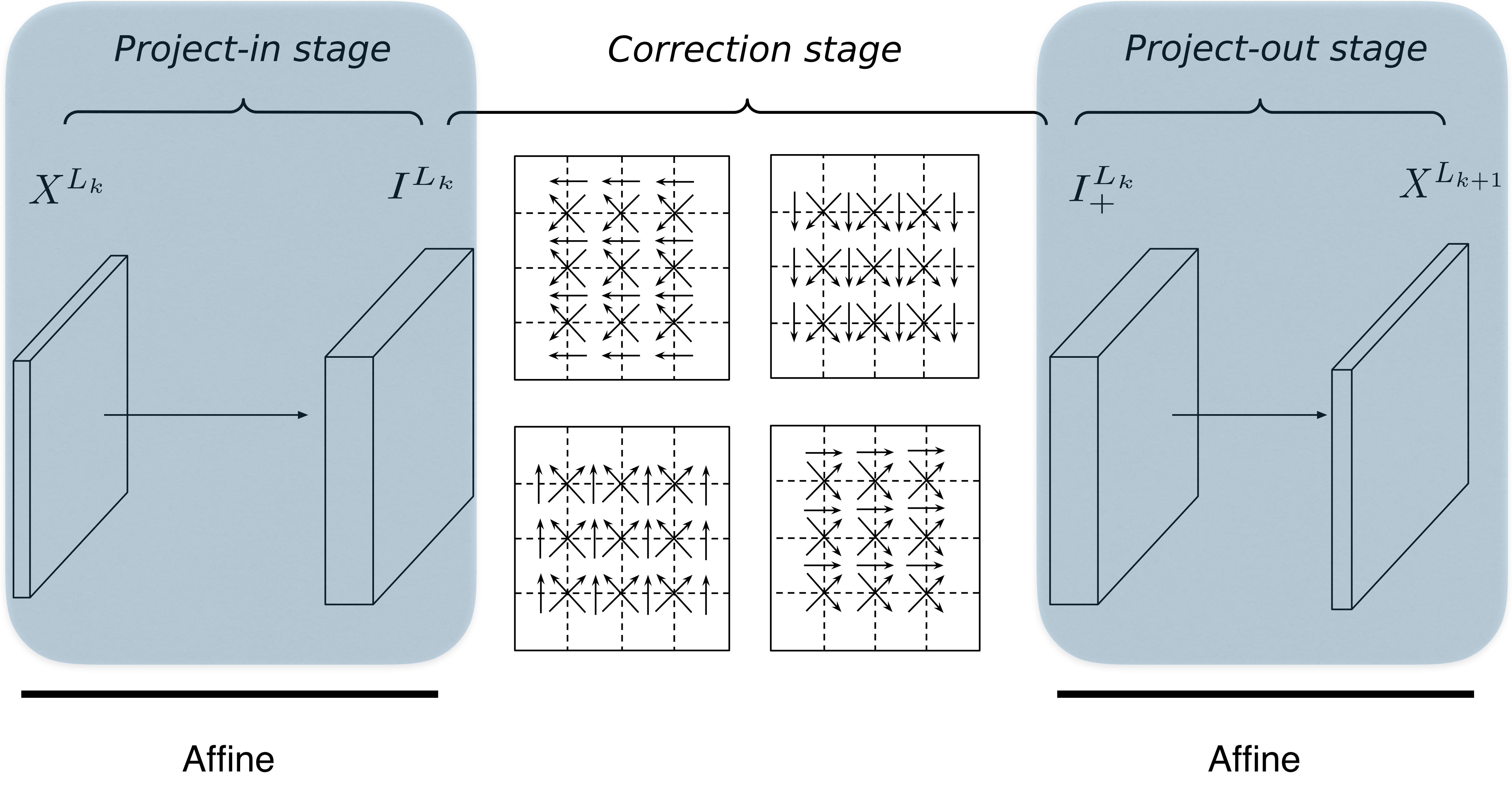
Spatial Dependency Network (SDN)



Computational flow of an SDN layer.

(Miladinović et al. 2021)

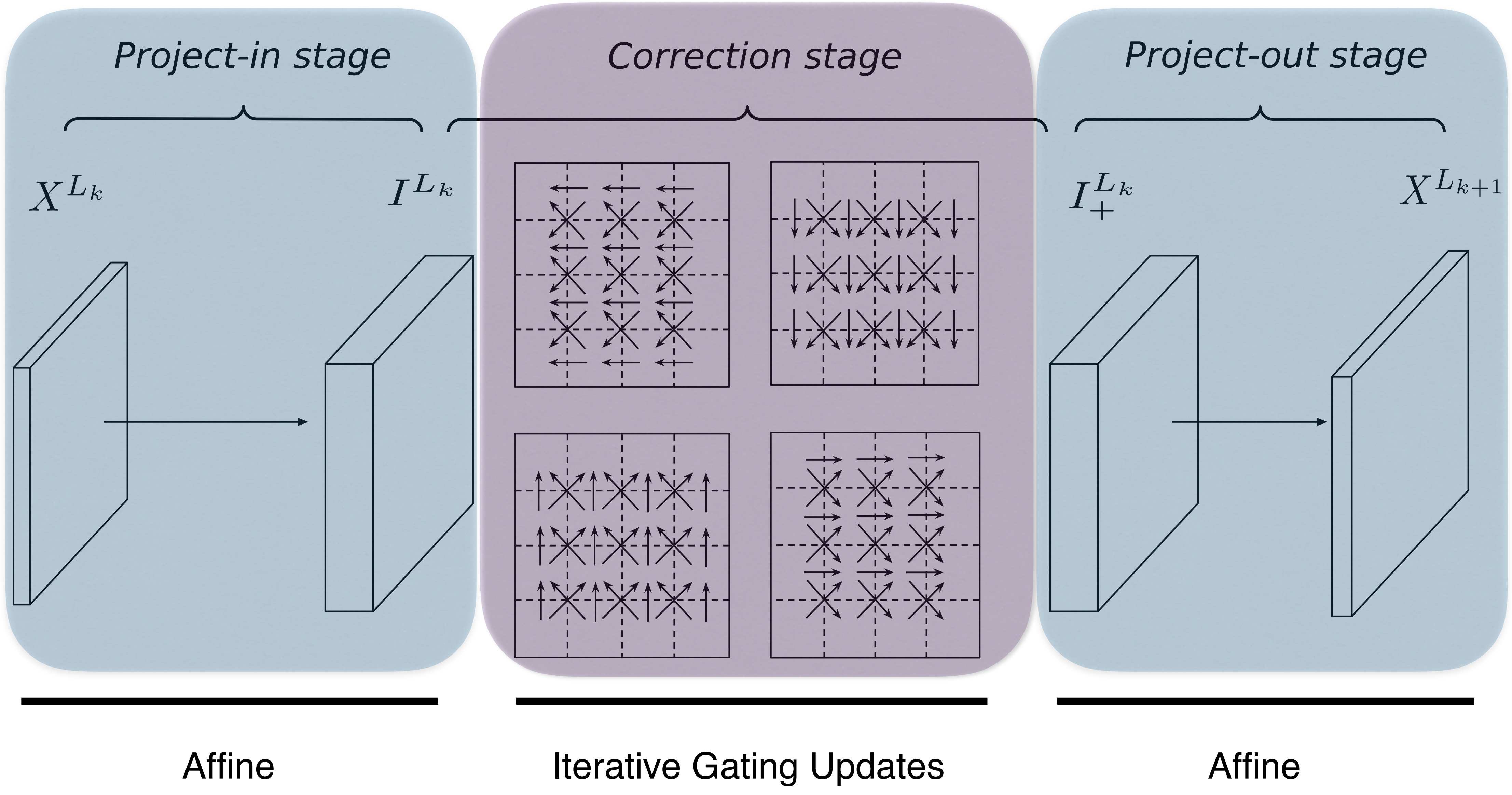
Spatial Dependency Network (SDN)



Computational flow of an SDN layer.

(Miladinović et al. 2021)

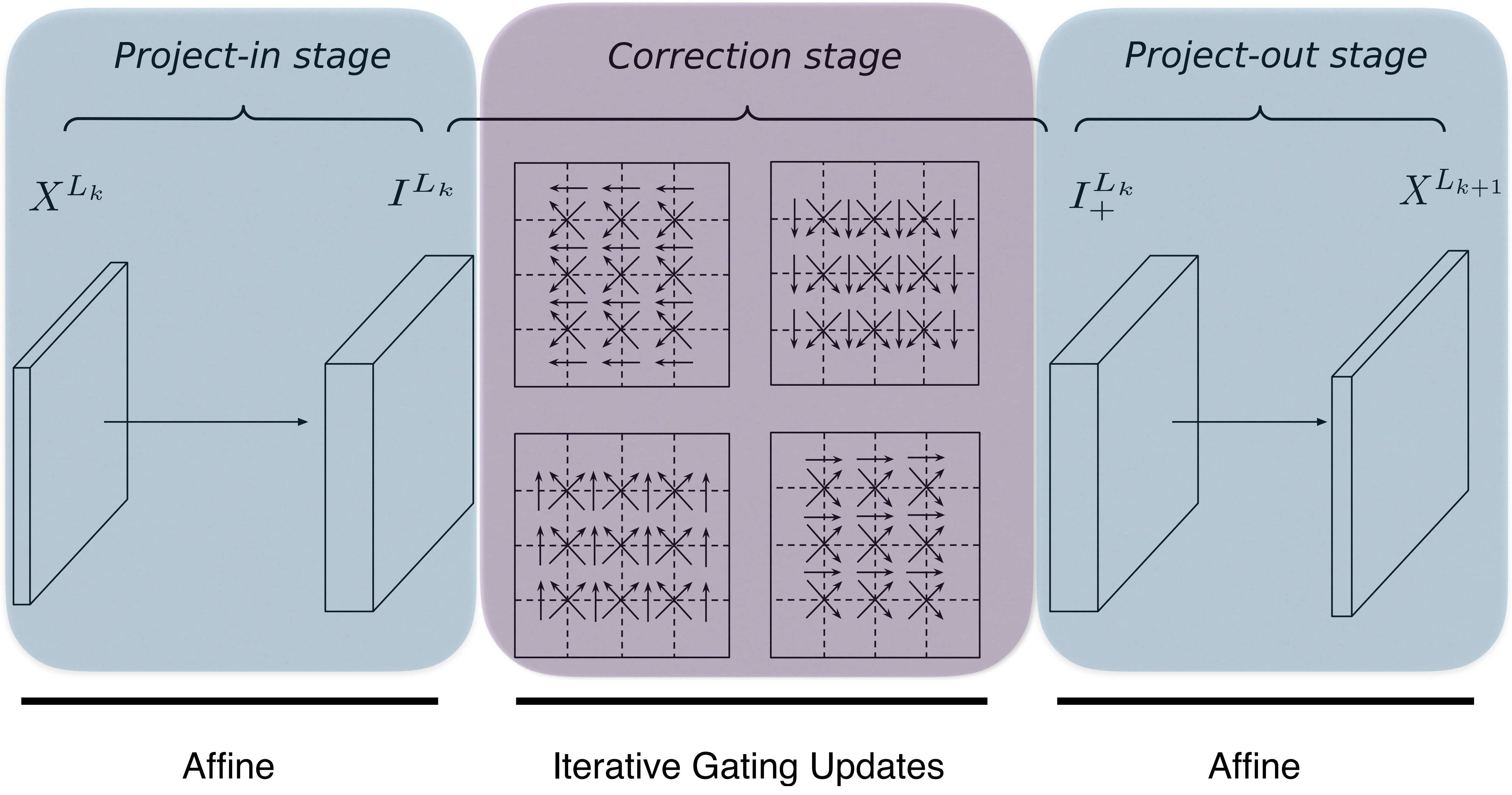
Spatial Dependency Network (SDN)



Computational flow of an SDN layer.

(Miladinović et al. 2021)

Spatial Dependency Network (SDN)

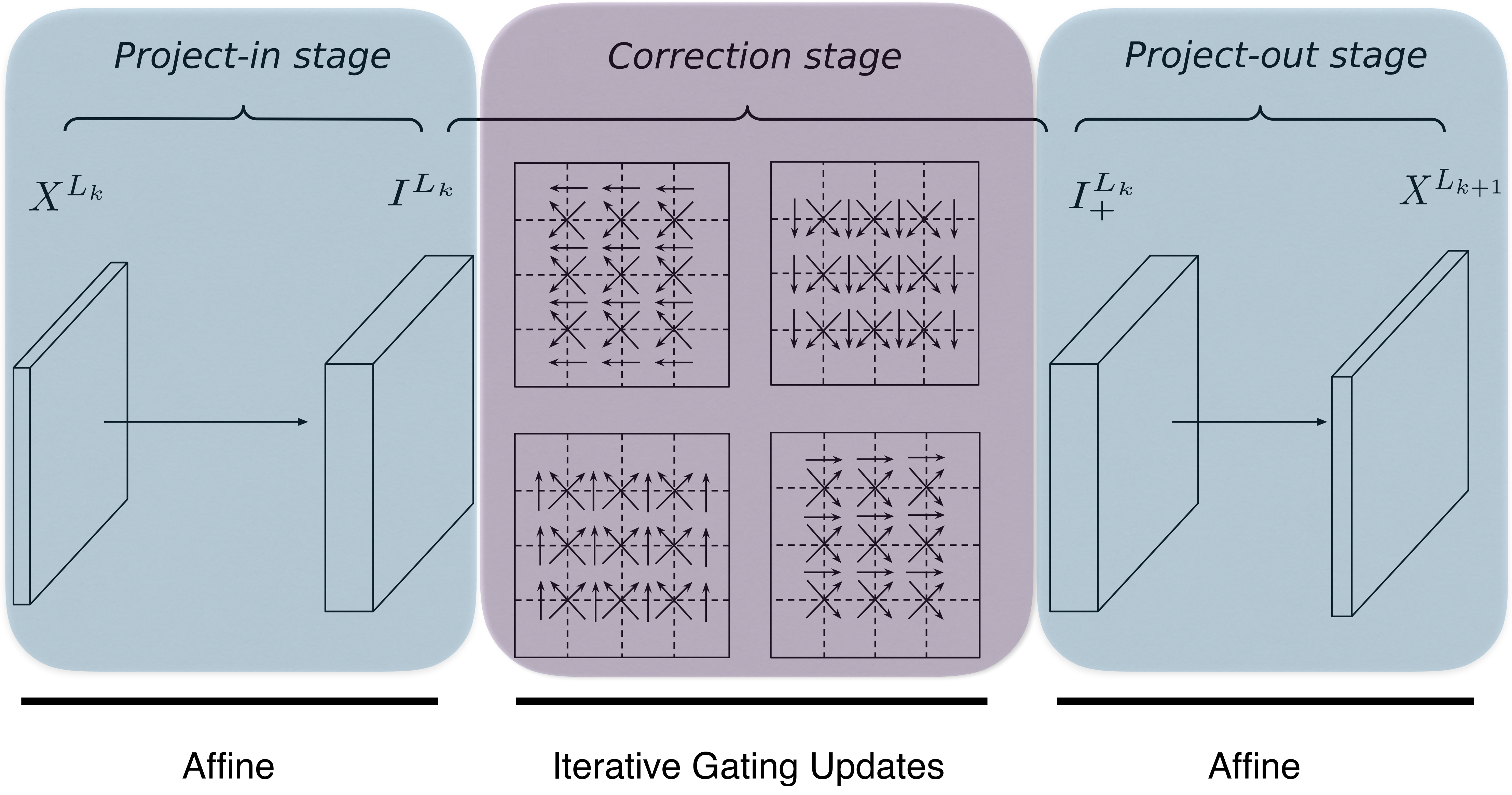


Computational flow of an SDN layer.

SDN layers:

(Miladinović et al. 2021)

Spatial Dependency Network (SDN)



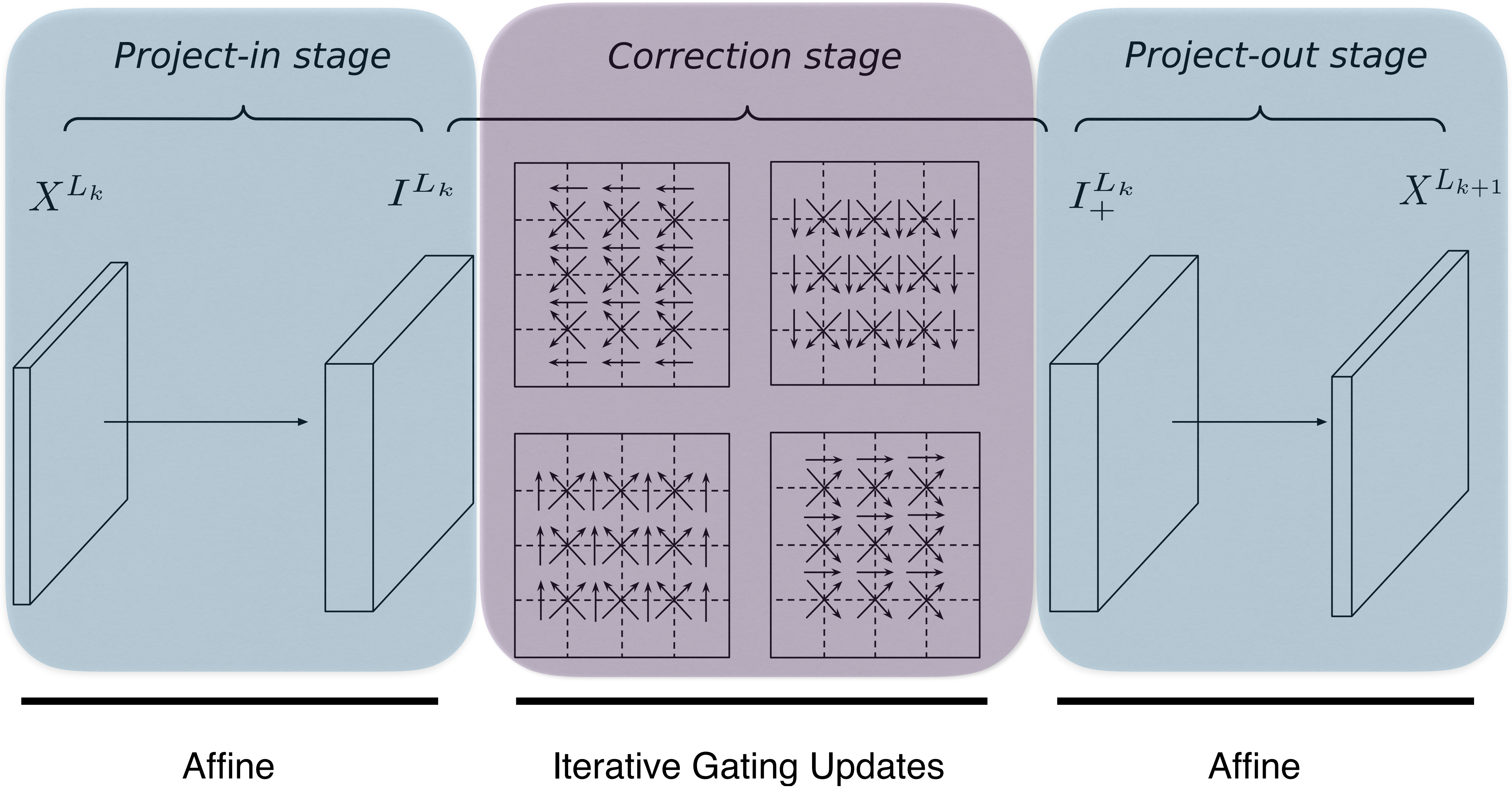
Computational flow of an SDN layer.

SDN layers:

1. **enforce spatial coherence** — the similarity of neighboring feature vectors.

(Miladinović et al. 2021)

Spatial Dependency Network (SDN)



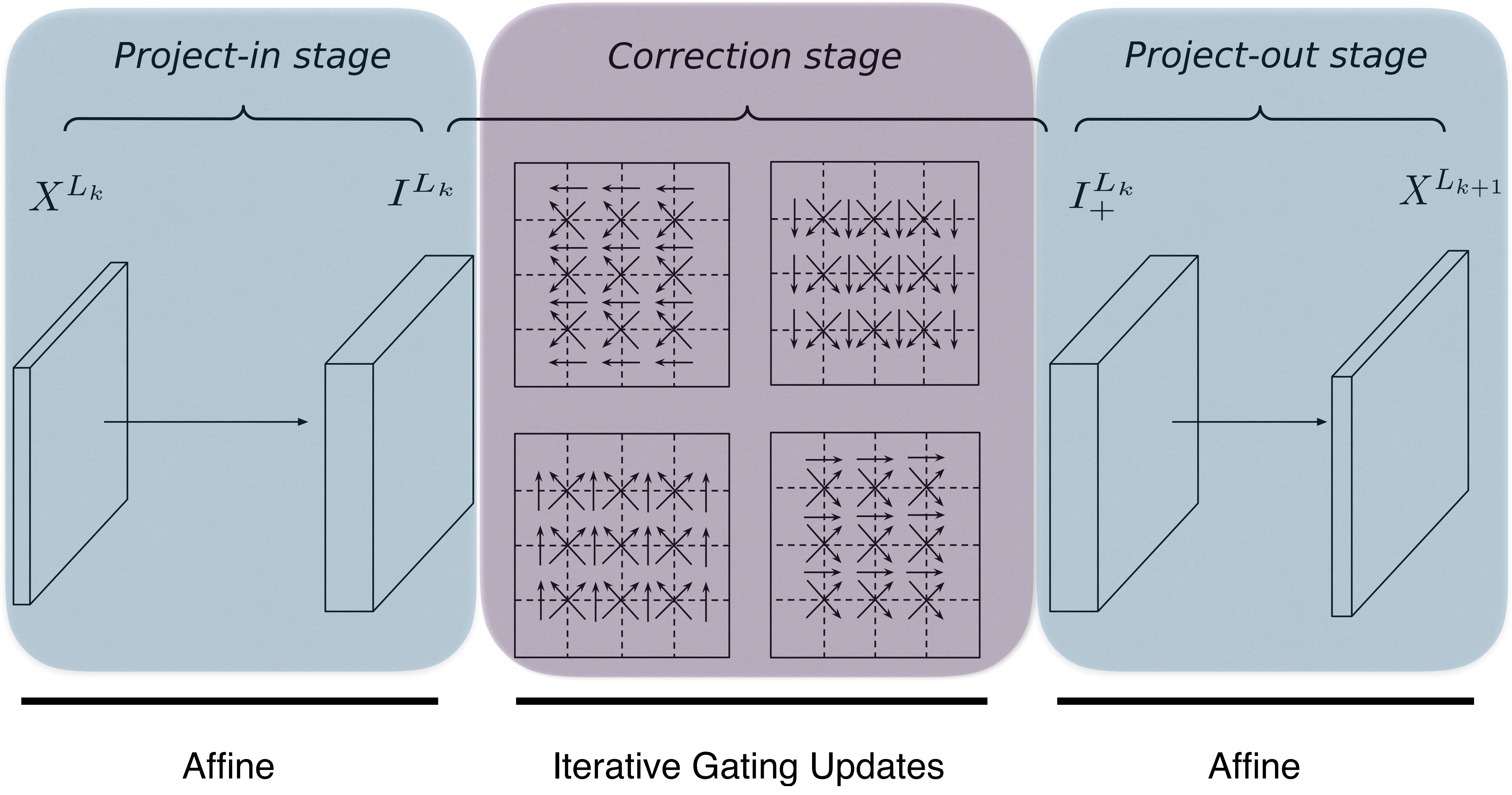
Computational flow of an SDN layer.

SDN layers:

1. **enforce spatial coherence** — the similarity of neighboring feature vectors.
2. **model long-range spatial dependencies** at each level of a NN.

(Miladinović et al. 2021)

Spatial Dependency Network (SDN)



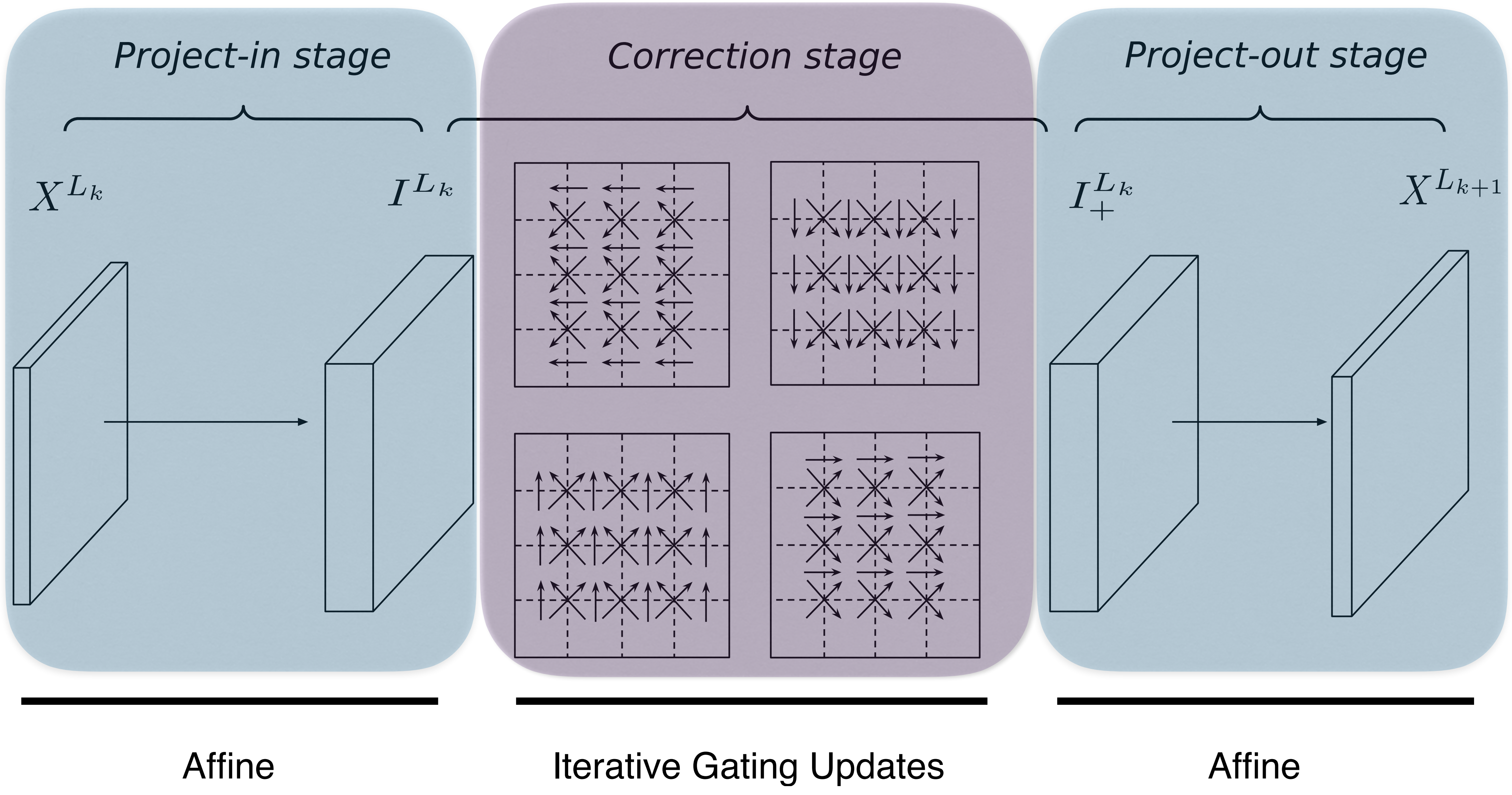
Computational flow of an SDN layer.

SDN layers:

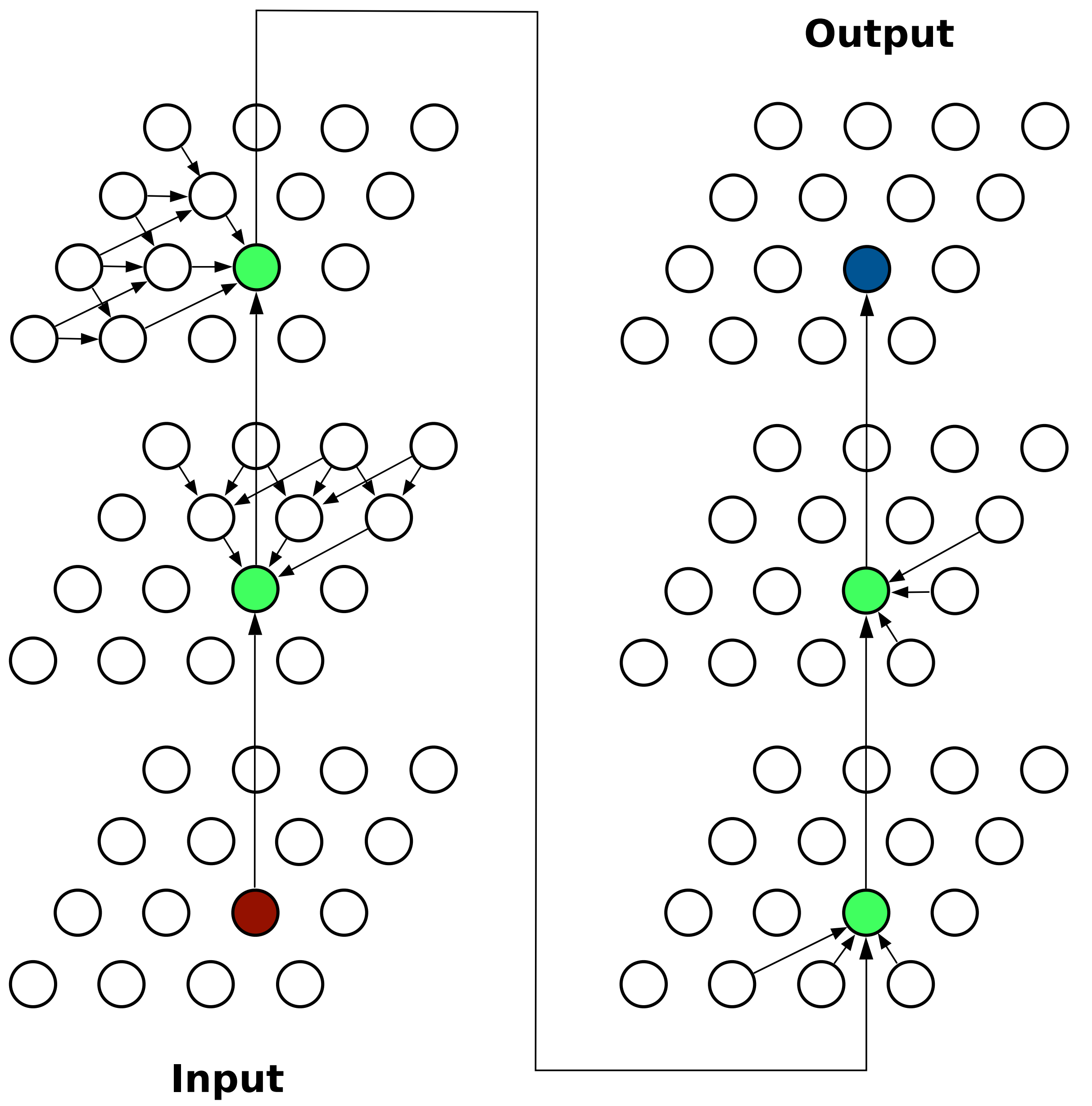
1. **enforce spatial coherence** — the similarity of neighboring feature vectors.
2. **model long-range spatial dependencies** at each level of a NN.
3. **achive equivariance to translation** through parameter sharing.

(Miladinović et al. 2021)

Spatial Dependency Network (SDN)



Computational flow of an SDN layer.



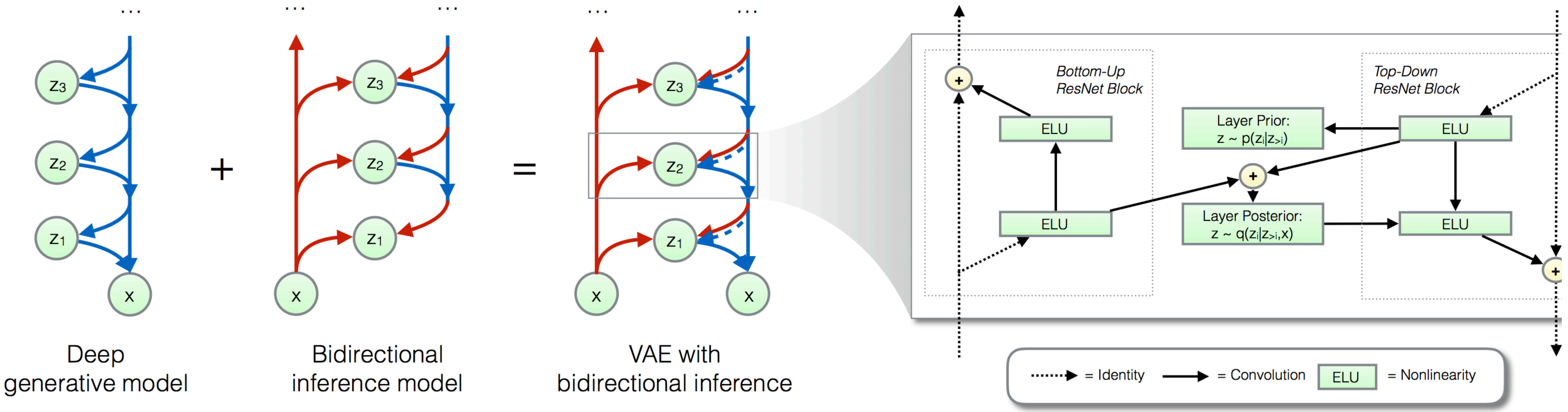
Graphical model of an SDN layer.

SDN layers:

1. **enforce spatial coherence** — the similarity of neighboring feature vectors.
2. **model long-range spatial dependencies** at each level of a NN.
3. **achive equivariance to translation** through parameter sharing.

(Miladinović et al. 2021)

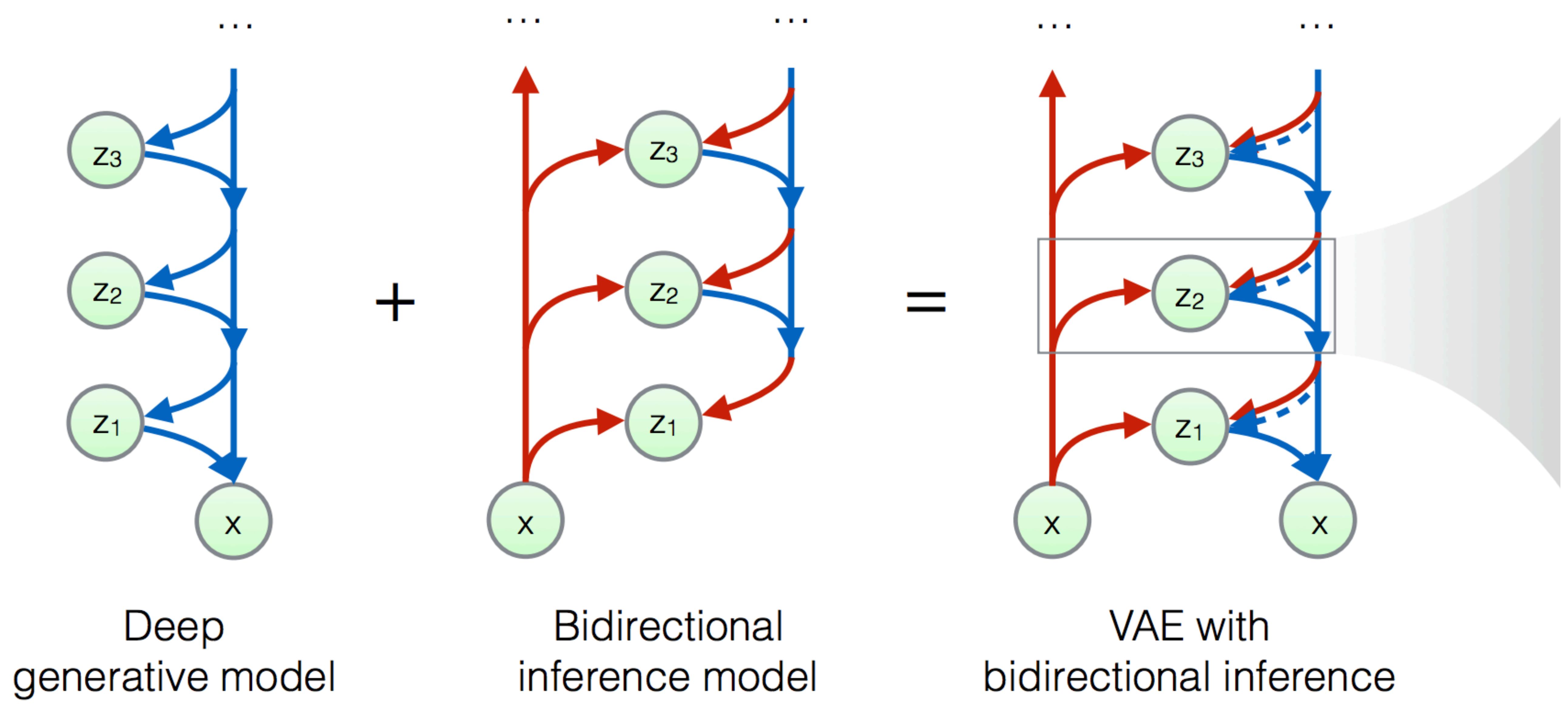
SDN-VAE: IAF-VAE enhanced with SDN layers

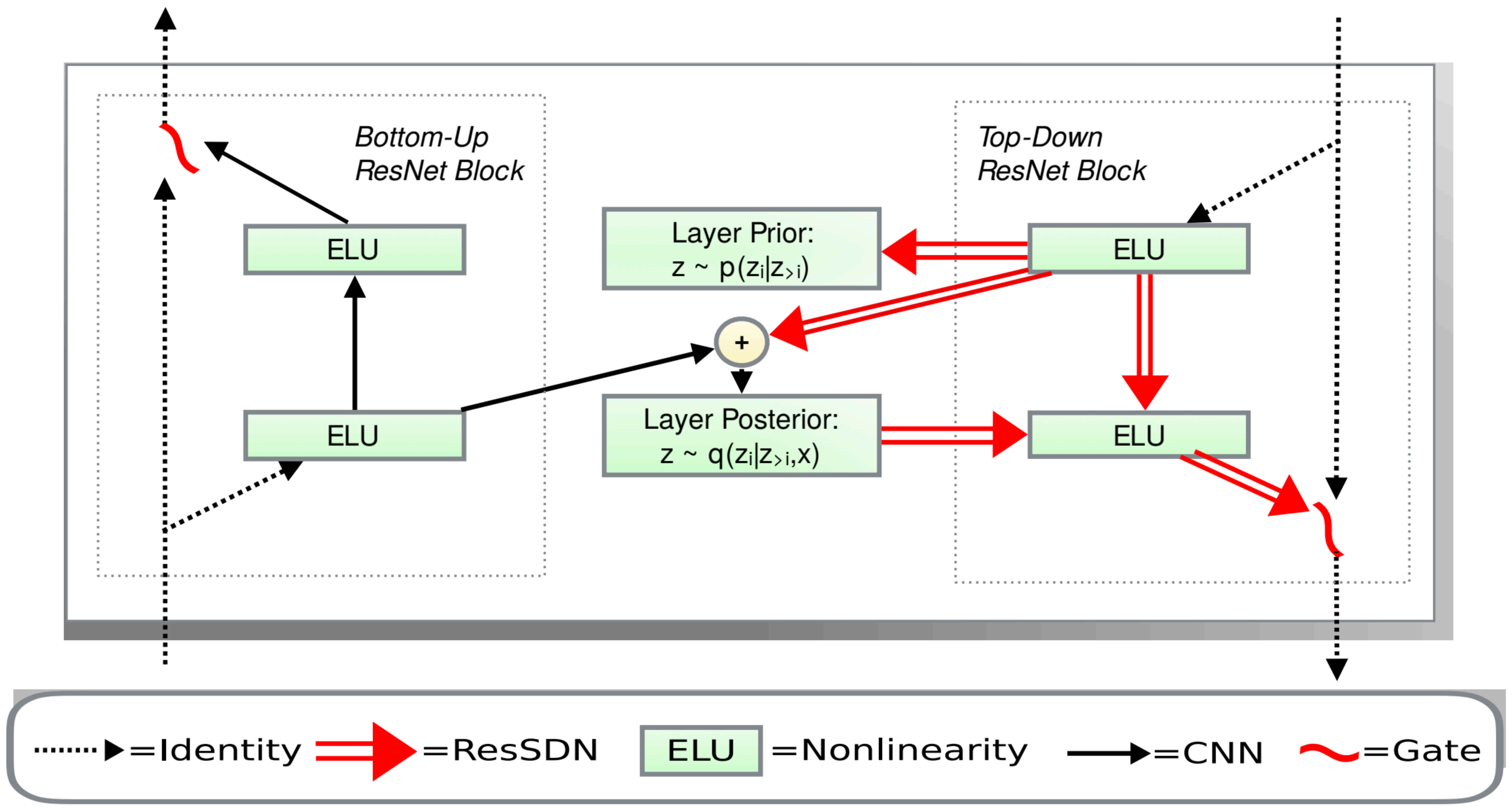


IAF-VAE: Hierarchical (Ladder) VAE [Sønderby et al. 2016] with ResNet blocks [Kingma et al. 2017]

(Miladinović et al. 2021)

SDN-VAE: IAF-VAE enhanced with SDN layers

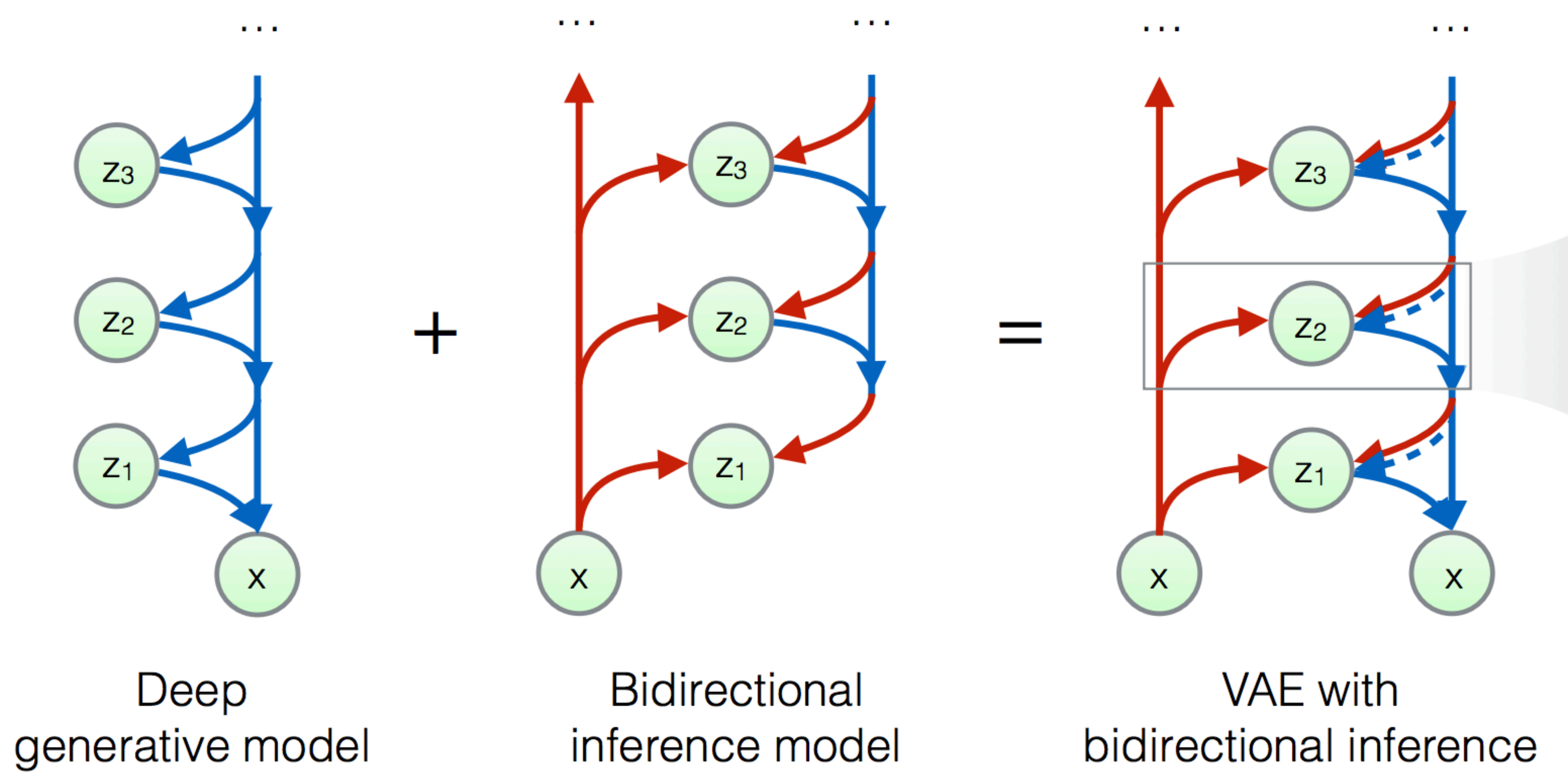


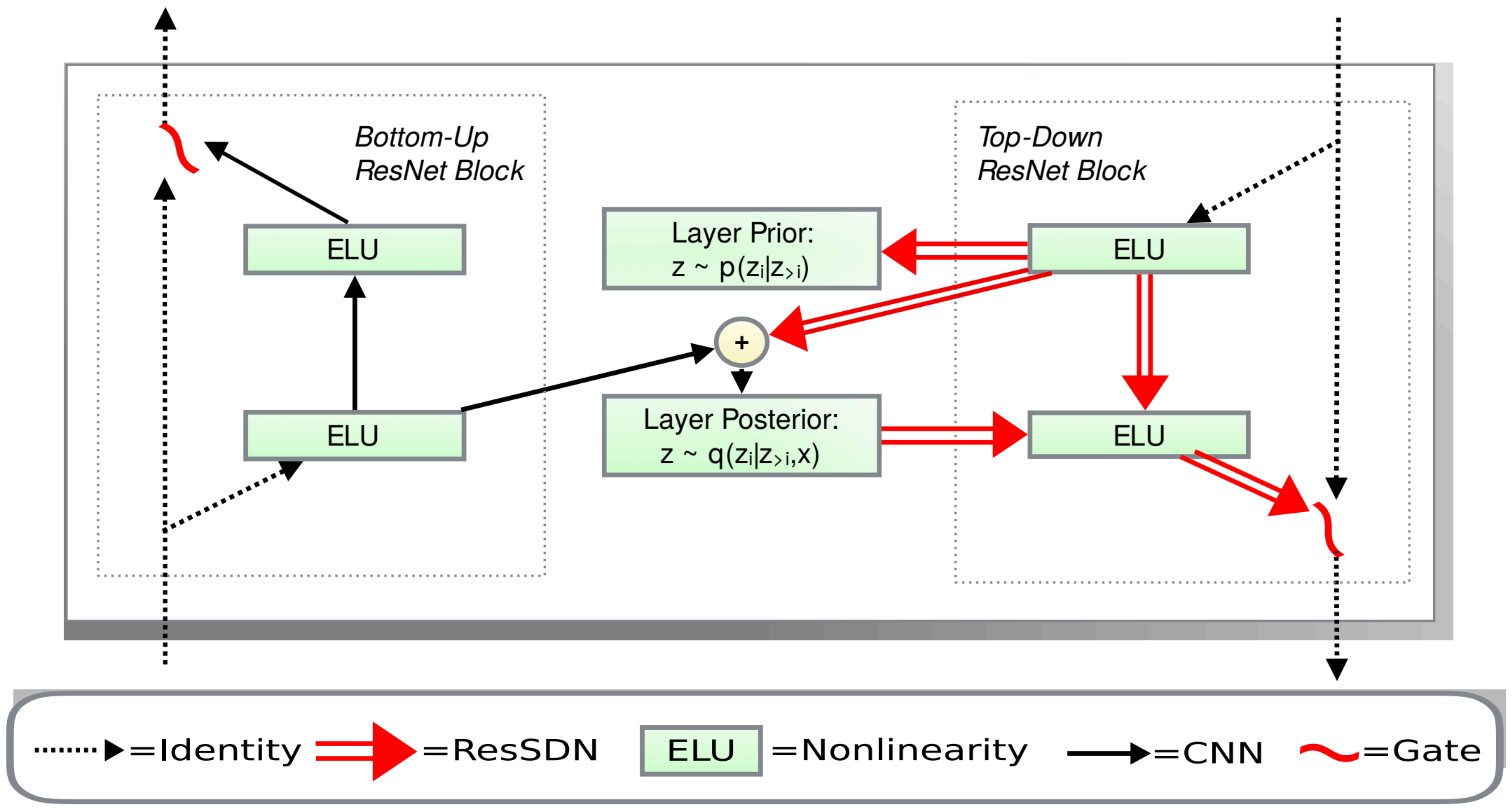


IAF-VAE: Hierarchical (Ladder) VAE [Sønderby et al. 2016] with ResNet blocks [Kingma et al. 2017]

(Miladinović et al. 2021)

SDN-VAE: IAF-VAE enhanced with SDN layers





IAF-VAE: Hierarchical (Ladder) VAE [Sønderby et al. 2016] with ResNet blocks [Kingma et al. 2017]

with additional modifications:

- Mixed-precision [Micikevicius et al. 2017]
- Gated "highway" residual [Srivastava et al. 2015]
- Discretized Mixture of Logistics [Salimans et al. 2017]

(Miladinović et al. 2021)

SDN-VAE: Density Estimation

Type	Method	CIFAR-10	ImageNet32	CelebAHQ256
VAE-based	**SDN-VAE (ours)**	2.87	3.85	0.70
	IAF-VAE+ (ours)	3.05	4.00	0.71
	IAF-VAE (Kingma et al., 2016)	3.11	X	X
	BIVA (Maaløe et al., 2019)	3.08	X	X
	NVAE (Vahdat & Kautz, 2020)	2.91	3.92	0.70
Flow-based	GLOW (Kingma & Dhariwal, 2018)	3.35	4.09	1.03
	FLOW++ (Ho et al., 2019)	3.08	3.86	X
	ANF (Huang et al., 2020)	3.05	3.92	0.72
	SurVAE (Nielsen et al., 2020)	3.08	4.00	X
Autoregressive*	PixelRNN (Van Oord et al., 2016)	3.00	3.86	X
	PixelCNN (Van den Oord et al., 2016)	3.03	3.83	X
	PixelCNN++ (Salimans et al., 2017)	2.92	X	X
	PixelSNAIL (Chen et al., 2018)	**2.85**	3.80	X
	SPN (Menick & Kalchbrenner, 2018)	X	3.85	**0.61**
	IT (Parmar et al., 2018)	2.90	**3.77**	X

Negative test log-likelihood.

(Miladinović et al. 2021)

SDN-VAE: Density Estimation

Type	Method	CIFAR-10	ImageNet32	CelebAHQ256
VAE-based	**SDN-VAE (ours)**	2.87	3.85	0.70
	IAF-VAE+ (ours)	3.05	4.00	0.71
	IAF-VAE (Kingma et al., 2016)	3.11	X	X
	BIVA (Maaløe et al., 2019)	3.08	X	X
	NVAE (Vahdat & Kautz, 2020)	2.91	3.92	0.70
Flow-based	GLOW (Kingma & Dhariwal, 2018)	3.35	4.09	1.03
	FLOW++ (Ho et al., 2019)	3.08	3.86	X
	ANF (Huang et al., 2020)	3.05	3.92	0.72
	SurVAE (Nielsen et al., 2020)	3.08	4.00	X
Autoregressive*	PixelRNN (Van Oord et al., 2016)	3.00	3.86	X
	PixelCNN (Van den Oord et al., 2016)	3.03	3.83	X
	PixelCNN++ (Salimans et al., 2017)	2.92	X	X
	PixelSNAIL (Chen et al., 2018)	**2.85**	3.80	X
	SPN (Menick & Kalchbrenner, 2018)	X	3.85	**0.61**
	IT (Parmar et al., 2018)	2.90	**3.77**	X

CNN vs SDN

Negative test log-likelihood.

(Miladinović et al. 2021)

SDN-VAE: Density Estimation

SotA !!

CNN vs SDN

Type	Method	CIFAR-10	ImageNet32	CelebAHQ256
VAE-based	**SDN-VAE (ours)**	2.87	3.85	0.70
	IAF-VAE+ (ours)	3.05	4.00	0.71
	IAF-VAE (Kingma et al., 2016)	3.11	X	X
	BIVA (Maaløe et al., 2019)	3.08	X	X
	NVAE (Vahdat & Kautz, 2020)	2.91	3.92	0.70
Flow-based	GLOW (Kingma & Dhariwal, 2018)	3.35	4.09	1.03
	FLOW++ (Ho et al., 2019)	3.08	3.86	X
	ANF (Huang et al., 2020)	3.05	3.92	0.72
	SurVAE (Nielsen et al., 2020)	3.08	4.00	X
Autoregressive*	PixelRNN (Van Oord et al., 2016)	3.00	3.86	X
	PixelCNN (Van den Oord et al., 2016)	3.03	3.83	X
	PixelCNN++ (Salimans et al., 2017)	2.92	X	X
	PixelSNAIL (Chen et al., 2018)	**2.85**	3.80	X
	SPN (Menick & Kalchbrenner, 2018)	X	3.85	**0.61**
	IT (Parmar et al., 2018)	2.90	**3.77**	X

Negative test log-likelihood.

(Miladinović et al. 2021)

SDN-VAE: Image Synthesis

temperature

temperature

(Miladinović et al. 2021)

SDN-VAE: Image Manipulation

Neighborhood exploration

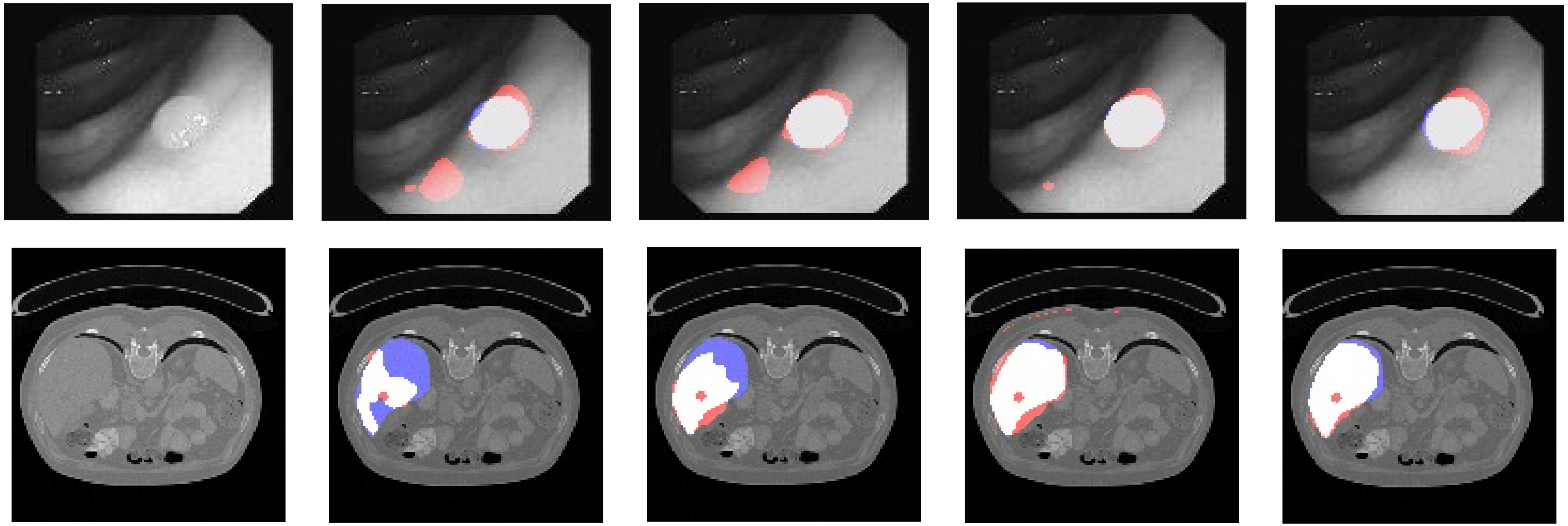
temperature

Interpolation

(Miladinović et al. 2021)

SDU-Nets for Medical Image Segmentation

SDN + U-Net [Carvalho et al. 2021]



Original **U-Net** **U-Net++** **SDN + U-Net** **SDN + U-Net++**

False Positives

False Negatives

Correct

(see the paper for quantitative results)

(Miladinović et al. 2021)

Implementation & Other Information

PyTorch

PyTorch Lightning

GitHub https://github.com/djordjemila/sdn

(PDT)

(Miladinović et al. 2021)

Implementation & Other Information

PyTorch

PyTorch Lightning

GitHub

https://github.com/djordjemila/sdn

```
from torch.nn import Conv2d
img_decoder = nn.Sequential(
    ...
    Conv2d(in_ch, out_ch, 4, 2, 1),
    nn.ReLU(True),
    ...
)
```

From CNN to SDN layer
in just one line of code !

```
from sdnlib.nn import SDNLayer
img_decoder = nn.Sequential(
    ...
    SDNLayer(in_ch, out_ch, num_features, 4, 2, 1),
    nn.ReLU(True),
    ...
)
```

(PDT)

(Miladinović et al. 2021)

Implementation & Other Information

PyTorch PyTorch Lightning GitHub https://github.com/djordjemila/sdn

```
from torch.nn import Conv2d
img_decoder = nn.Sequential(
    ...
    Conv2d(in_ch, out_ch, 4, 2, 1),
    nn.ReLU(True),
    ...
)
```

From CNN to SDN layer
in just one line of code !

```
from sdnlib.nn import SDNLayer
img_decoder = nn.Sequential(
    ...
    SDNLayer(in_ch, out_ch, num_features, 4, 2, 1),
    nn.ReLU(True),
    ...
)
```

arXiv link for the SDN paper: https://arxiv.org/abs/2103.08877

arXiv link for the follow-up, SDU-Net paper: https://arxiv.org/abs/2103.11713

Make sure you are up to date on Twitter ! Follow: https://twitter.com/djordjemila

Poster Session 4	May 4, 2021, 1 a.m.	May 4, 2021, 3 a.m.

(PDT)

(Miladinović et al. 2021)