Cutting-Edge Techniques for Depth Map Super-Resolution

Ryan Peterson and Josiah Smith

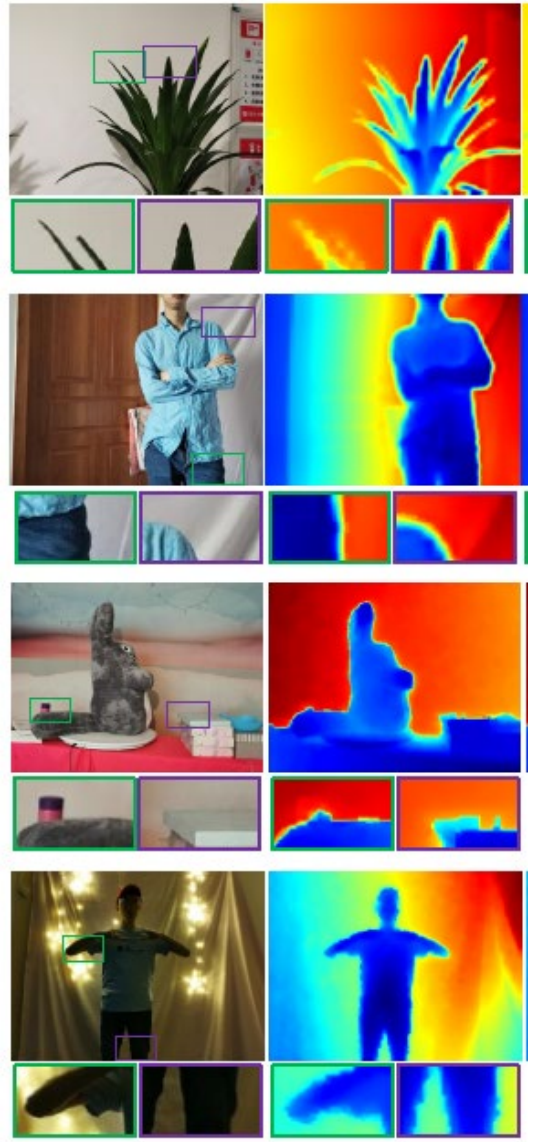
Outline

1. Depth Map Super-Resolution

2. Methods
   - **S**hifted **Win**dow (Swin) Transformer, **N**onlinear **A**ctivation **F**ree (NAF) Networks, **D**eformable **K**ernel **N**etworks (DKN)

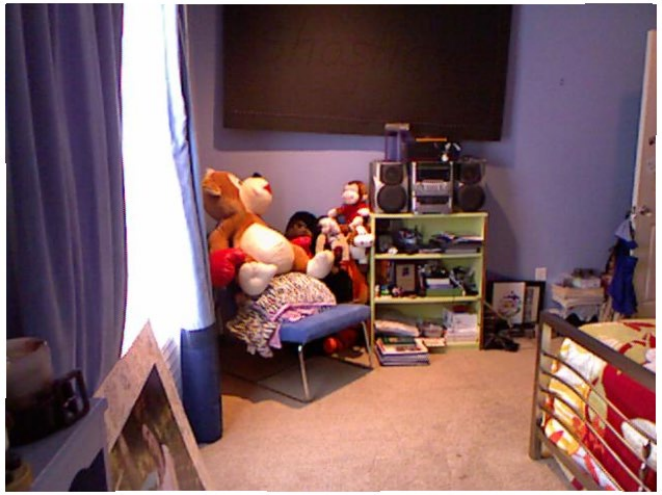
3. Proposed Architectures
   - Swin Transformer- and NAF-based **D**epth **M**ap **S**uper **R**esolution (DMSR)

4. Results and Conclusions



Depth Map Super-Resolution

- Input:
  - High-resolution RGB image
  - Low-resolution depth image



H, W, C

H/8, W/8, 1

Network

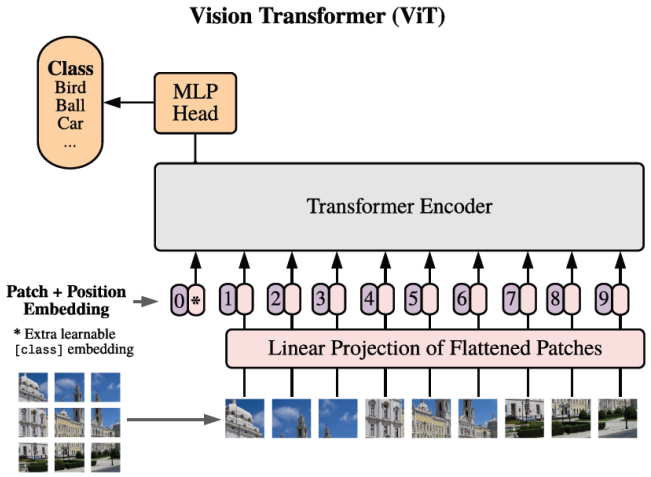
- Output:
  - High-resolution depth image

H, W, 1

Swin Transformer Architecture

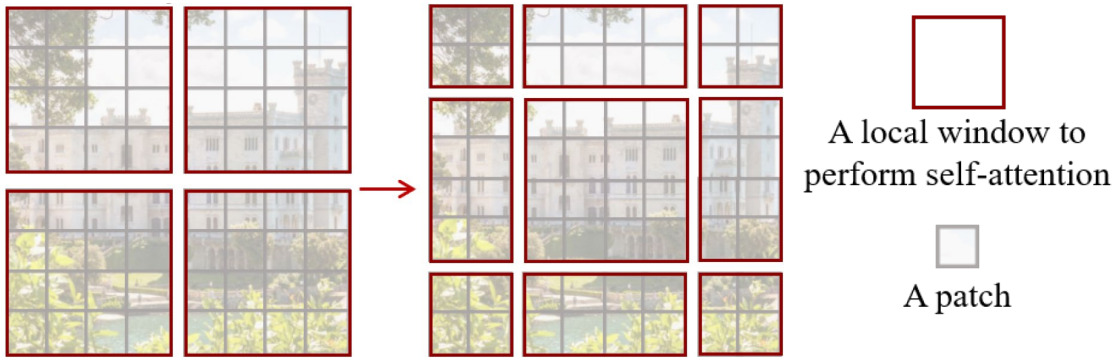
Vision Transformer (ViT)

- Split image into patches
- Use MSA encoder-decoder
- Global context, translationally variant



Swin Transformer

- Hierarchical (different patch sizes)
- Local self-attention
- Shifting windows → translationally equivariant (CNN), local context

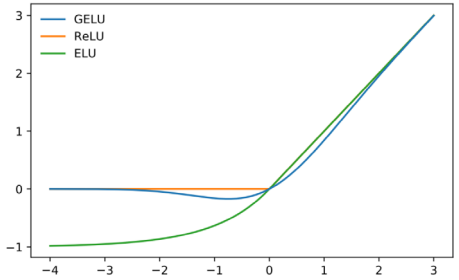


Nonlinear Activation Free (NAF) Networks

Nonlinear Activation Functions

- GELU

$$\mathrm{GELU}(\boldsymbol{x}) = \boldsymbol{x} \odot \phi(\boldsymbol{x})$$

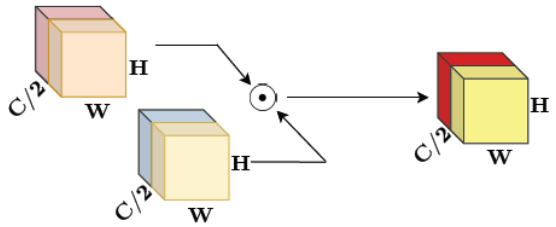


- Gated Linear Unit (GLU)

$$\mathrm{GLU}(\boldsymbol{x}) = f(\boldsymbol{x}) \odot \sigma\big(g(\boldsymbol{x})\big)$$

- SimpleGate (SG)

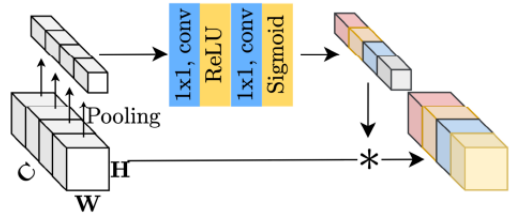
$$\mathrm{SG}(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{x} \odot \boldsymbol{y}$$



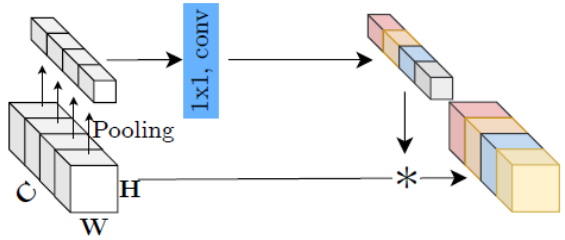
(c)Simple Gate

Simplified Channel Attention

- Channel attention



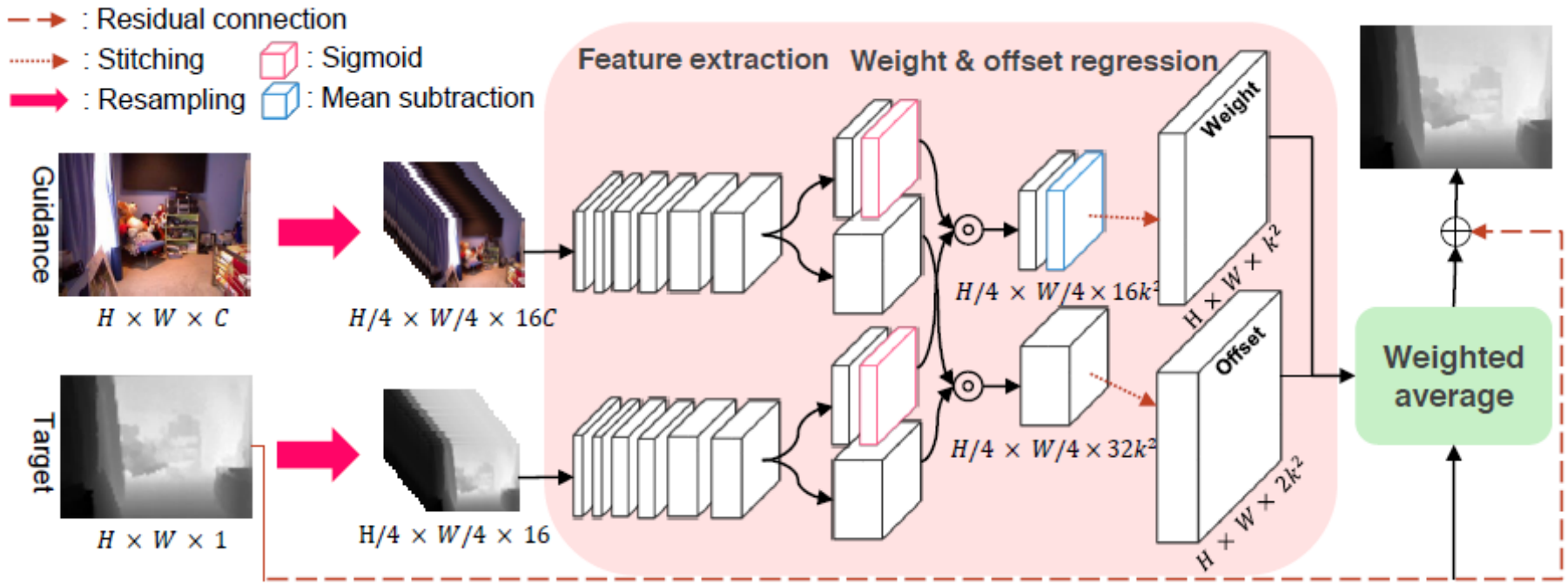
- Simplified Channel Attention



Deformable Kernel Networks (DKN)

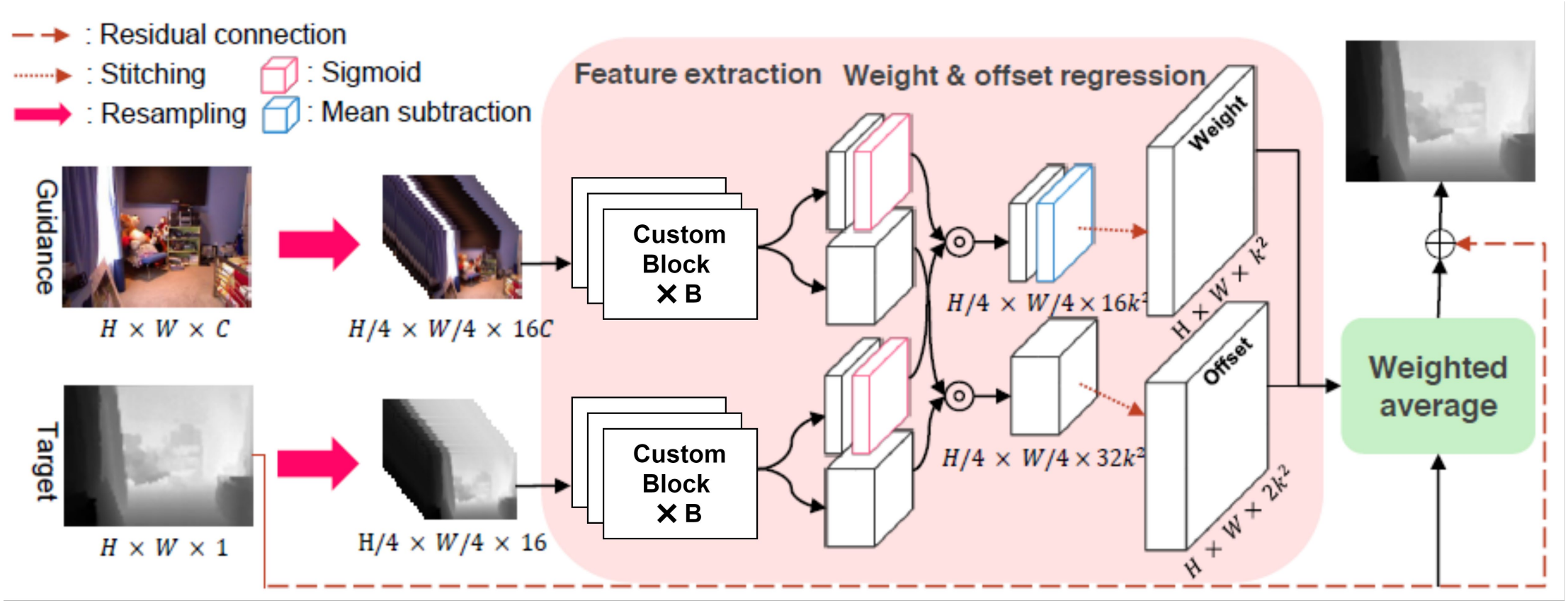
- Images are resampled
- Features from each image are extracted

- Combined
- Stitched
- Averaged with original depth



Our Networks

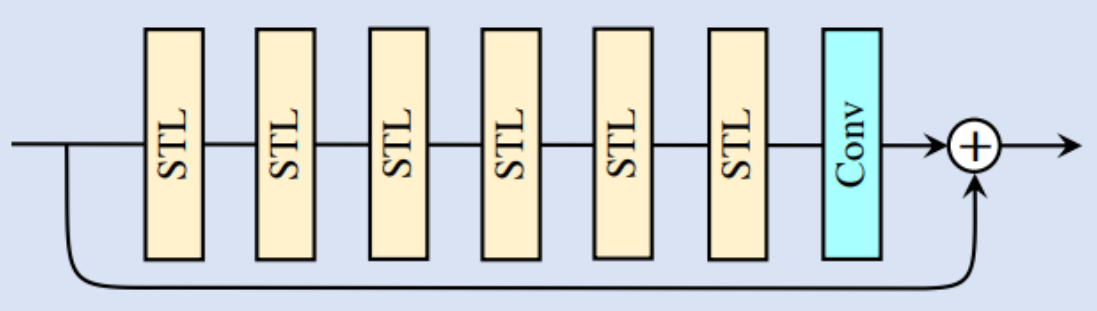
1. Swin transformer-based blocks
2. NAF-based blocks



Custom Network Architecture

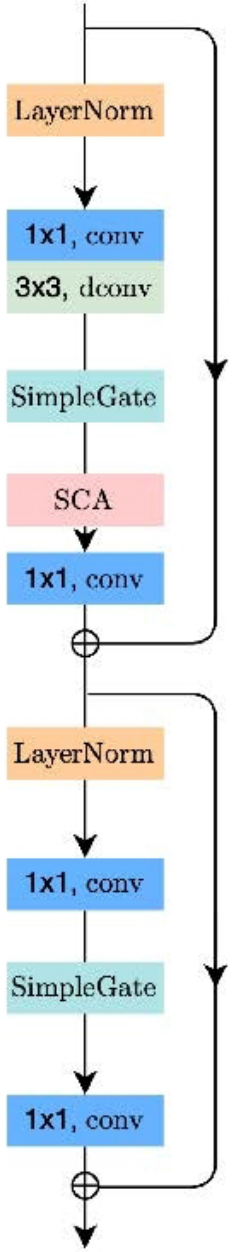
Residual Swin Transformer Block (RSTB)

- Sequential Swin layers
- Residual connection



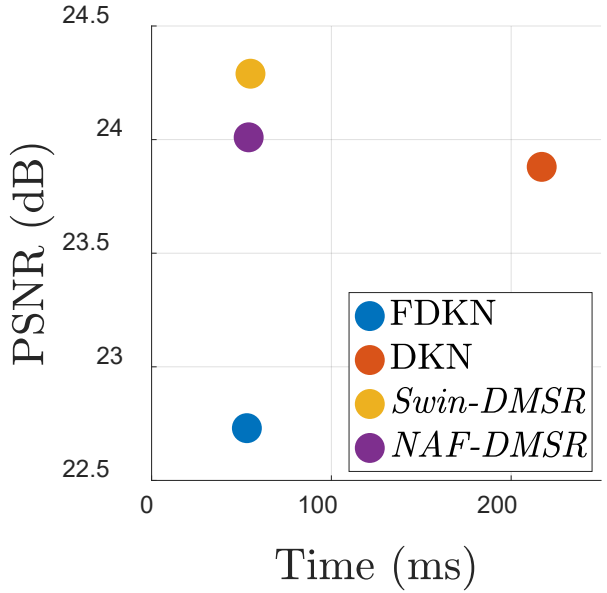
NAF Block

- Simplified in comparison to state-of-the-art blocks
- Uses convolution and attention



Results

Low Resolution | GT | FDKN | DKN | *Swin-DMSR* | *NAF-DMSR*



	FDKN	DKN	Swin-DMSR	NAF-DMSR
PNSR (dB)	22.73	23.88	**24.29**	24.01
Time (ms)	**53**	217	55	54


Questions?
Thank You!