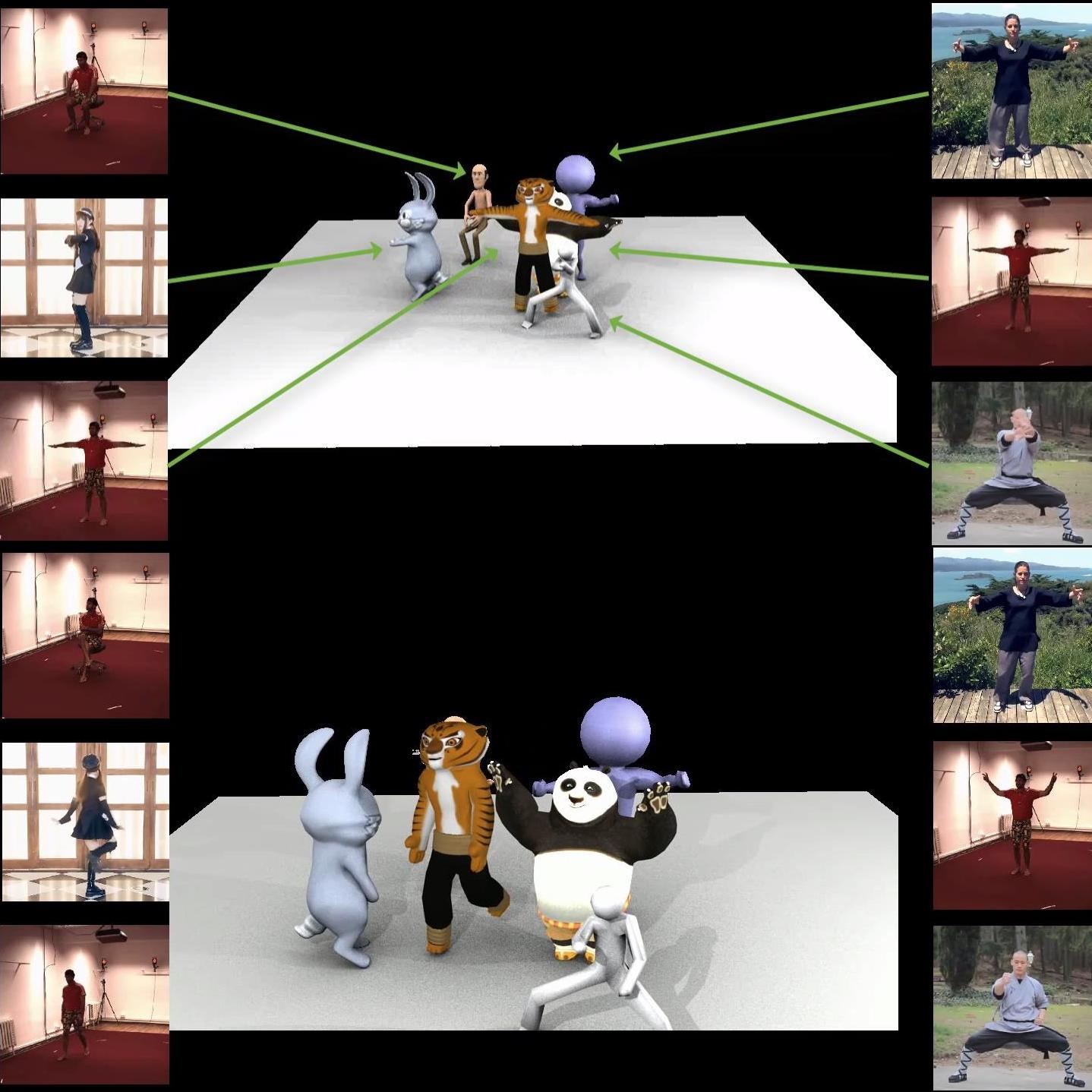




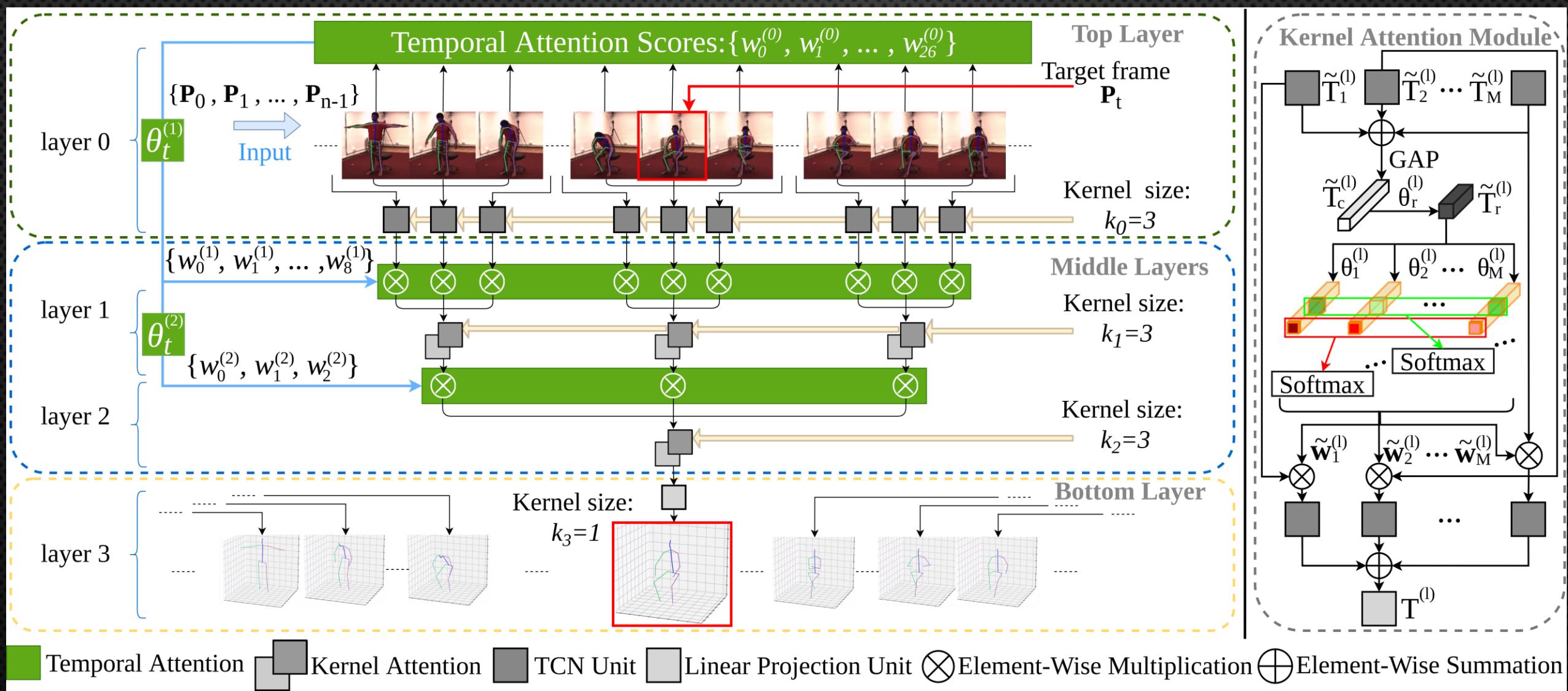
Attention Mechanism Exploits Temporal Contexts: Real-time 3D Human Pose Reconstruction

Ruixu Liu, Ju Shen, He Wang
Chen Chen, Sen-ching Cheung, Vijayan Asari

Code is available at: (<https://github.com/lrxjason/Attention3DHumanPose>)



Attentional Mechanism



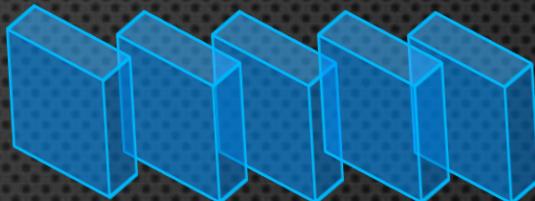
Attention



Temporal Attention

(weights on tensors)

$(\omega_1^t, \omega_2^t, \omega_3^t, \omega_4^t, \omega_5^t \dots)$

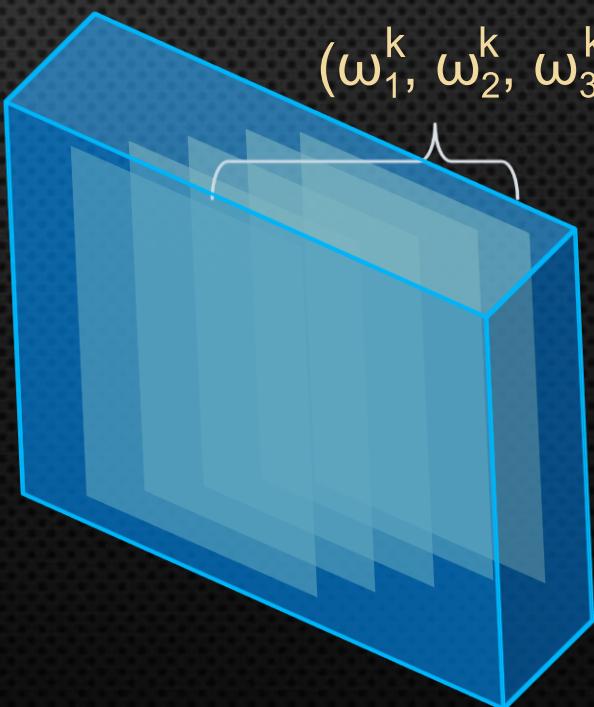


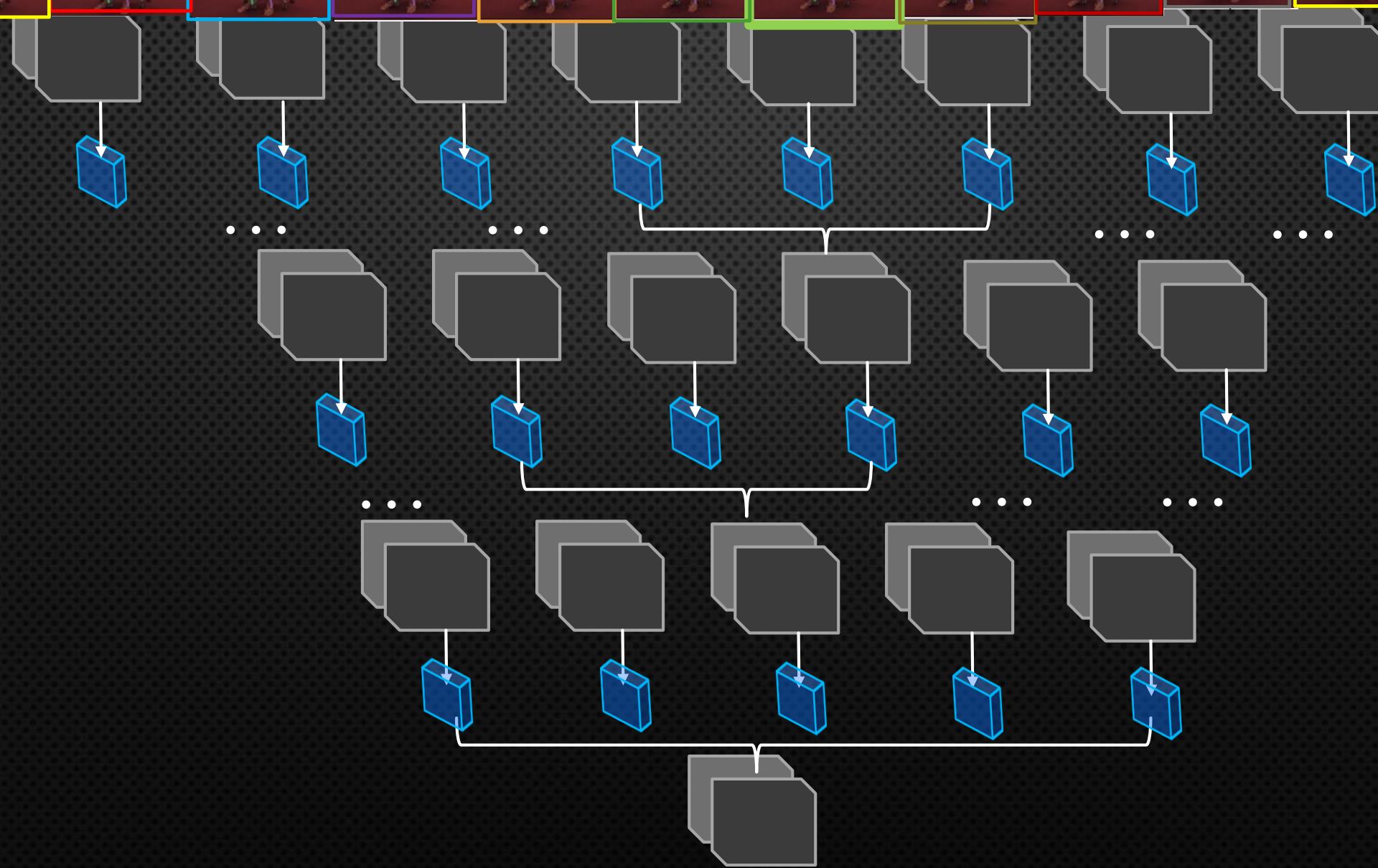
...

Kernel Attention

(weights on channels)

$(\omega_1^k, \omega_2^k, \omega_3^k, \omega_4^k, \omega_5^k \dots)$



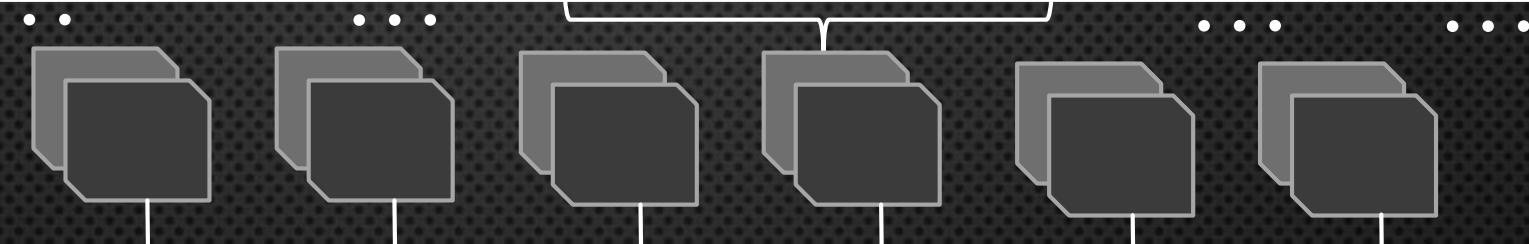


Attention Layer 0

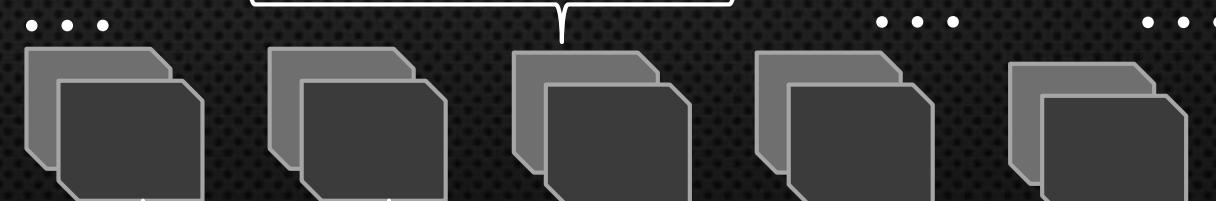


$W^{(1)}$

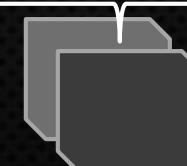
Attention Layer 1



Attention Layer 2



Attention Layer 3

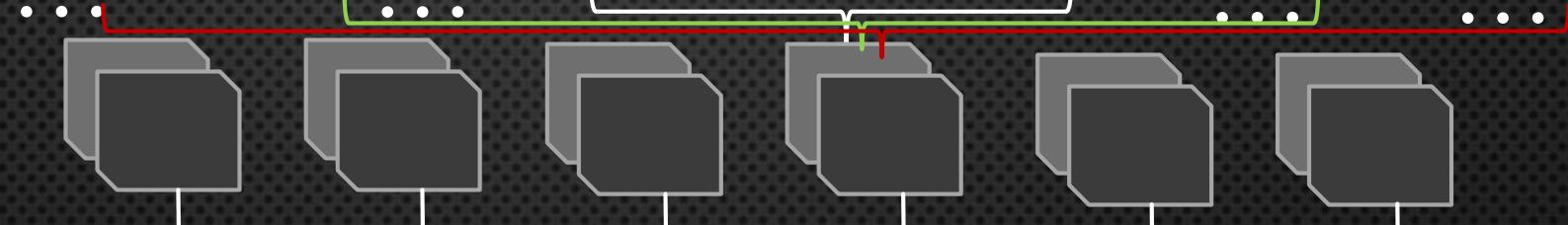


Attention Layer 0



Attention Layer 1

$W^{(1)}$



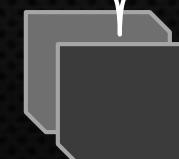
Attention Layer 2

$W^{(2)}$



Attention Layer 3

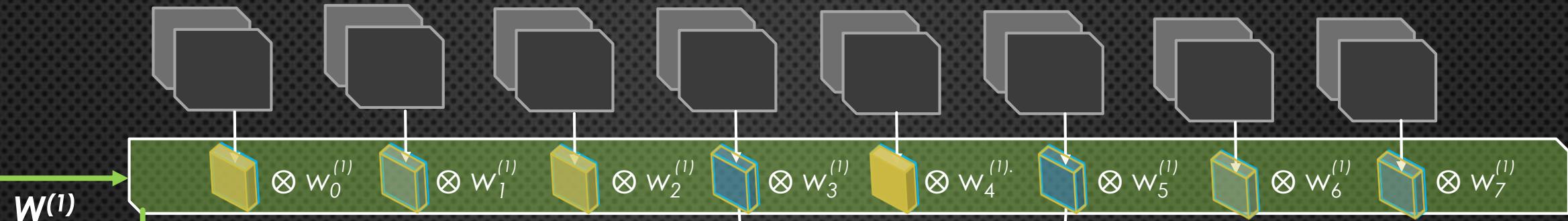
$W^{(3)}$



$$W^{(0)} \leftarrow NCC(f_0, f_1, \dots, f_{n-1})$$

Attention Layer 0

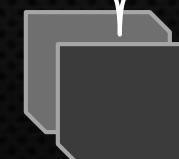
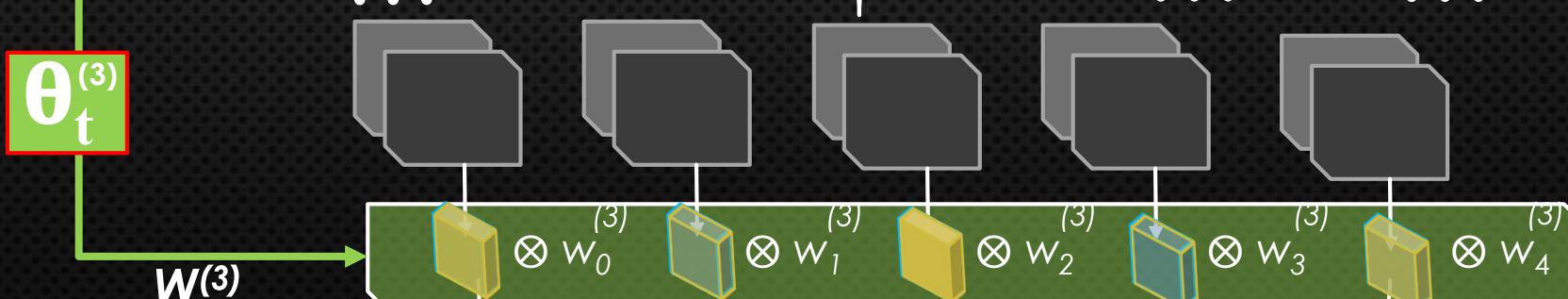
$\Theta_t^{(1)}$



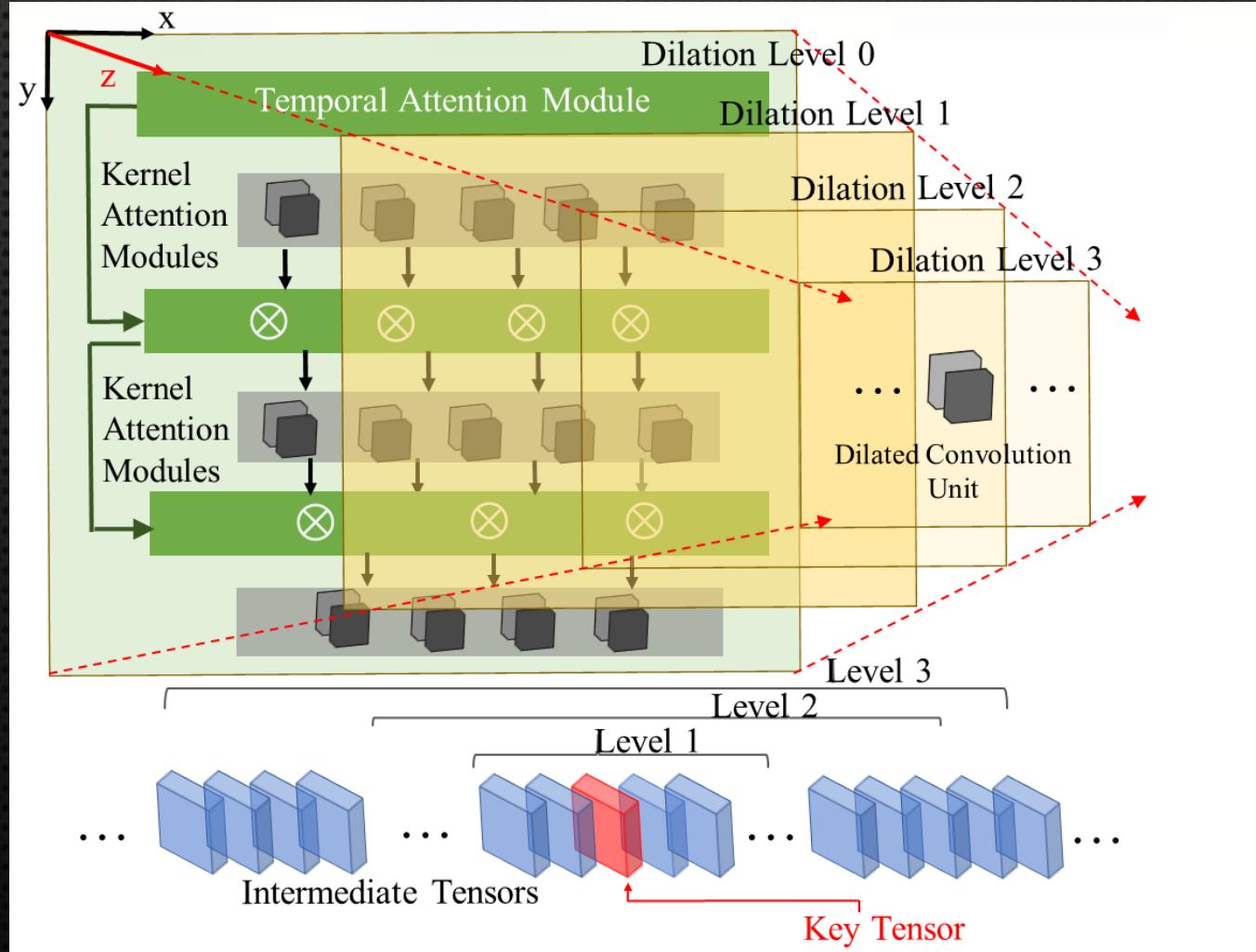
$\Theta_t^{(2)}$



$\Theta_t^{(3)}$



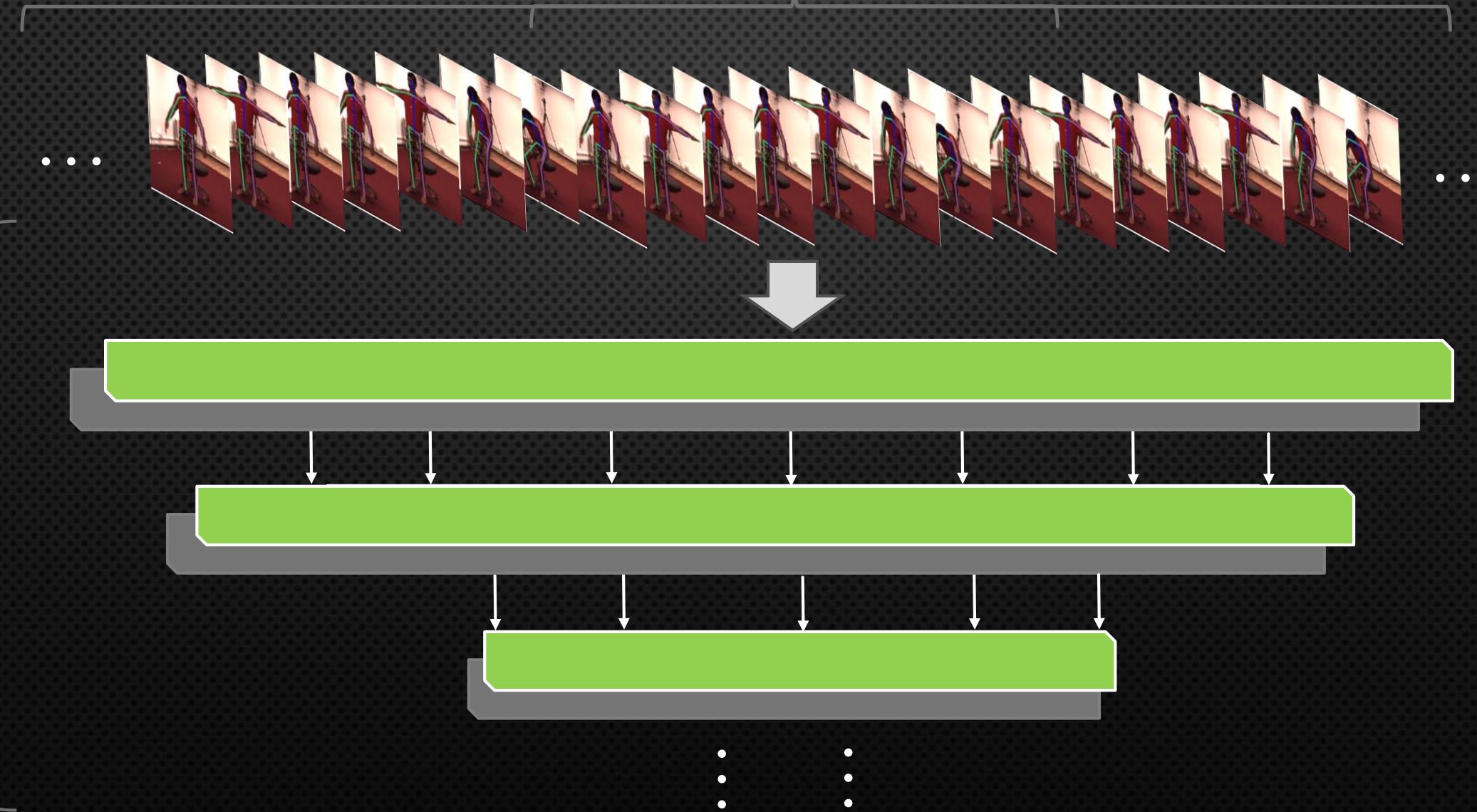
The Multi-scale Dilated Convolution Structure



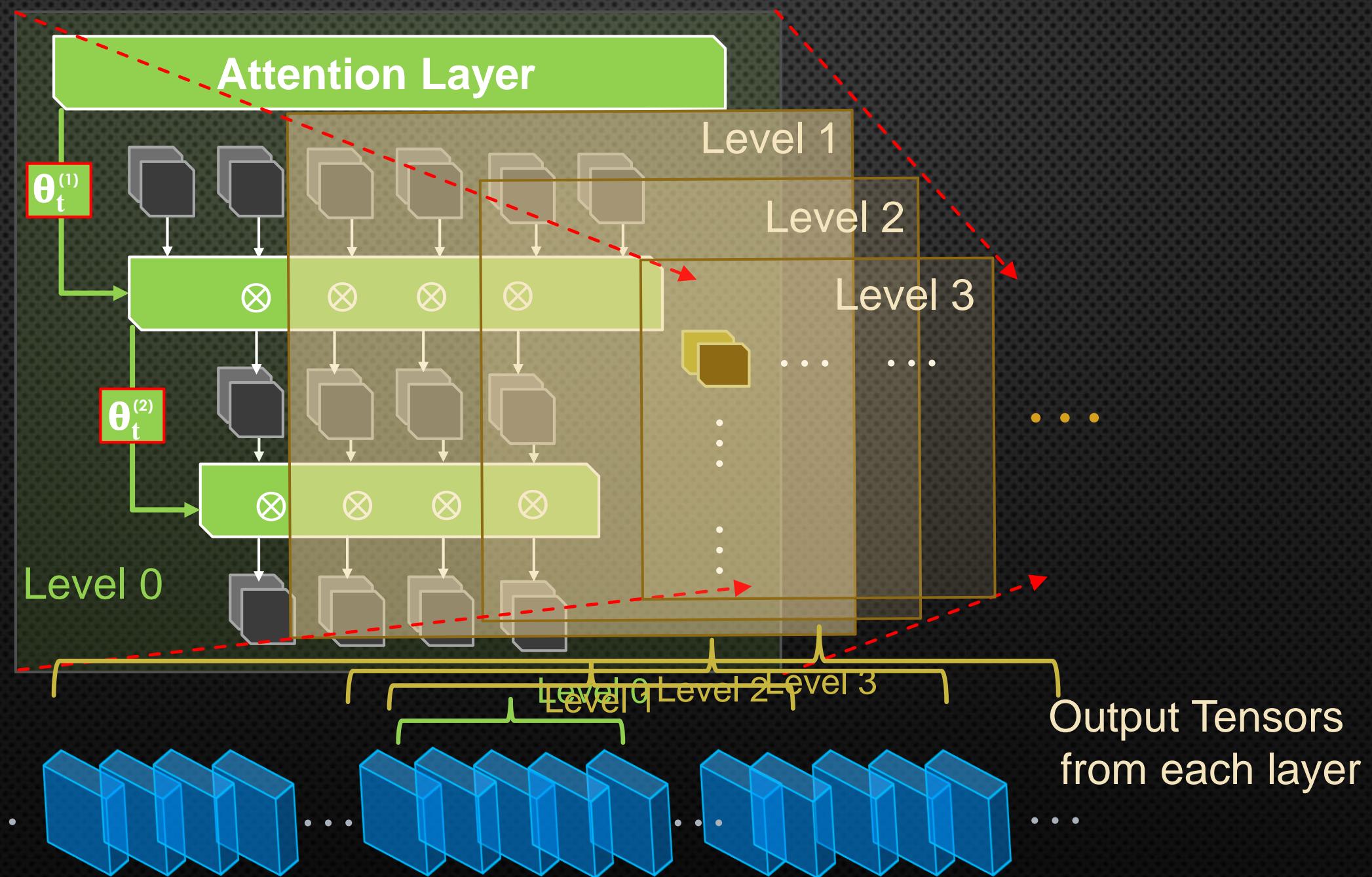
To increase receptive field

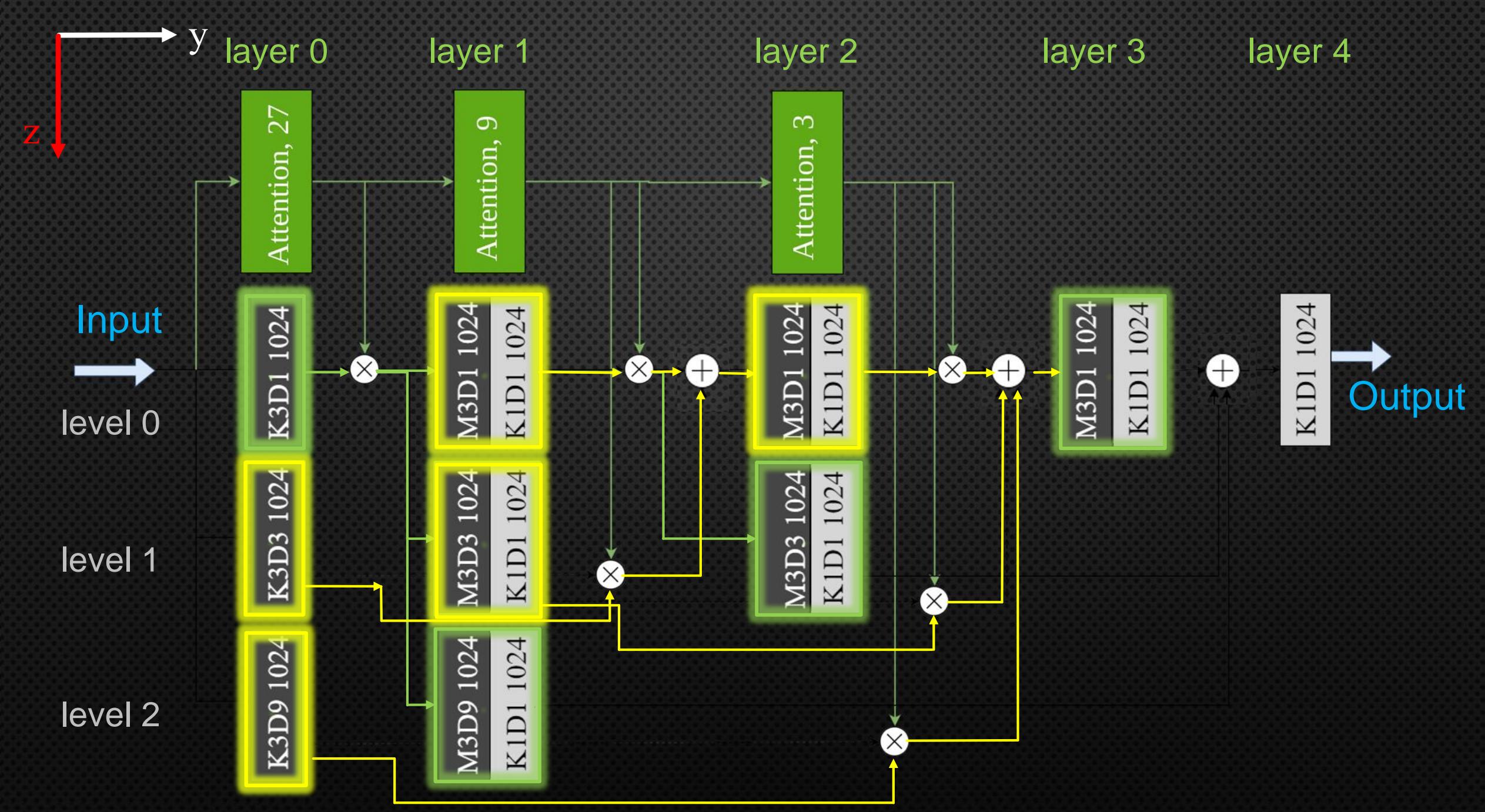
Input

More layers



x
y
z





Quantitative Evaluation:

1. Side-by-side comparison with state-of-the-art

S9 Direction

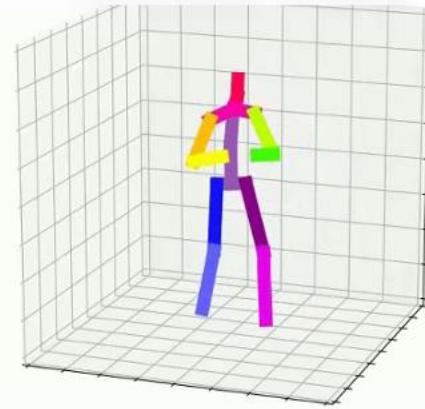
Input



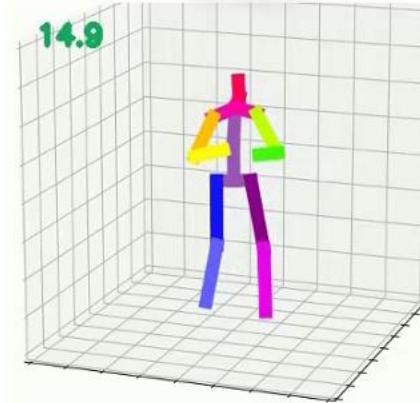
jitter



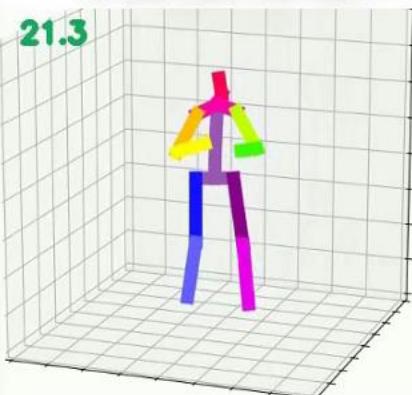
Ground Truth



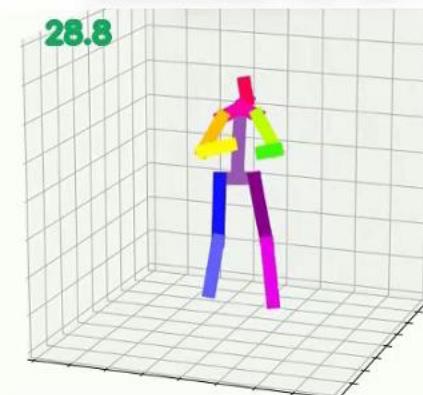
Ours



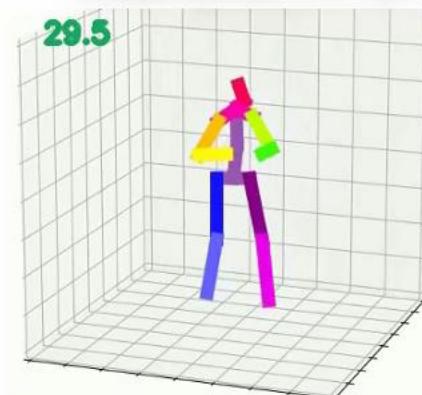
Pavllo et al.



Hossain et al.



Martinez et al.



Quantitative Evaluation:

Motion Retargeting Views

S9 Walking

Ground Truth

Ours



Input

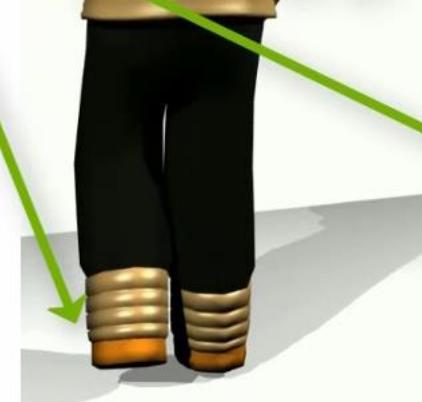


Inaccuracy

Pavllo et al.



Hossain et al.



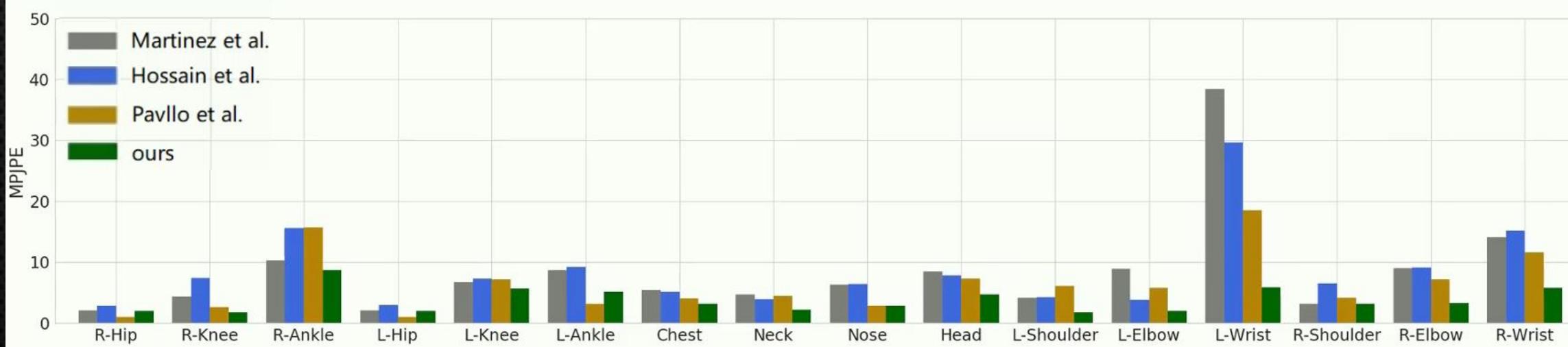
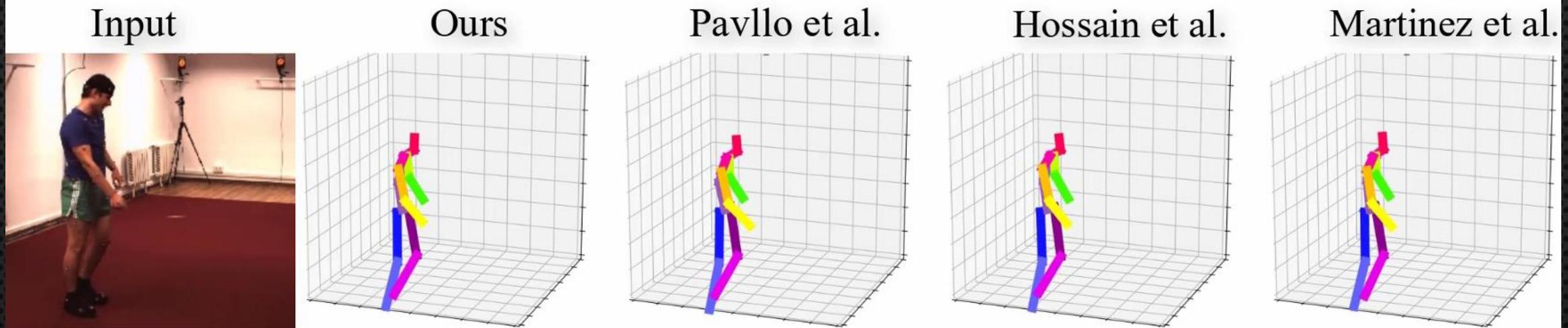
Martinez et al



Quantitative Evaluation:

2. Joint-wise MPJPE: comparison with
state-of-the-art

S11 WalkDog

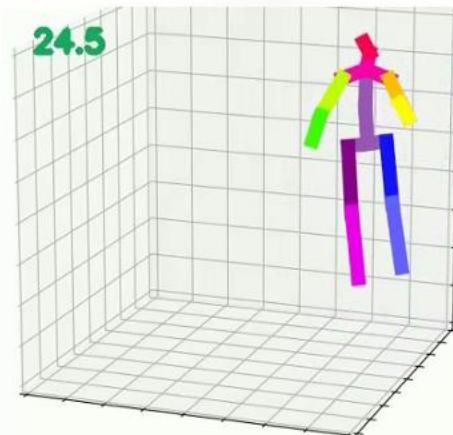


Quantitative Evaluation:

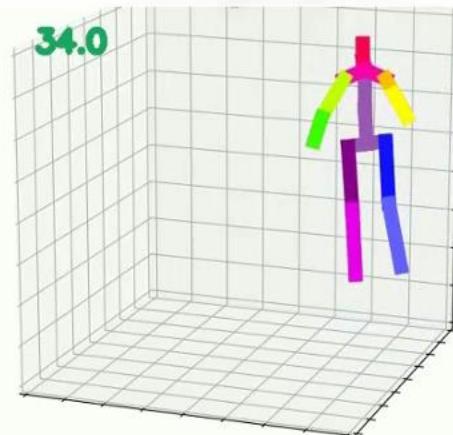
3. Frame-wise MPJPE: comparison with
state-of-the-art

S9 Smoking

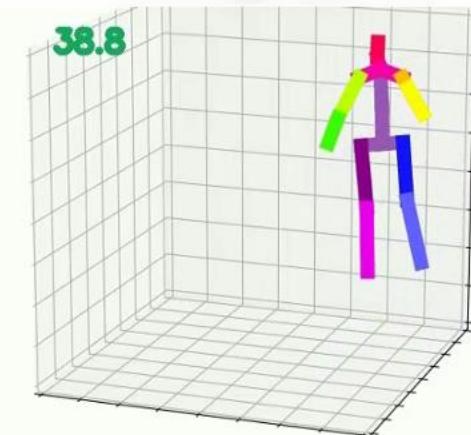
Ours



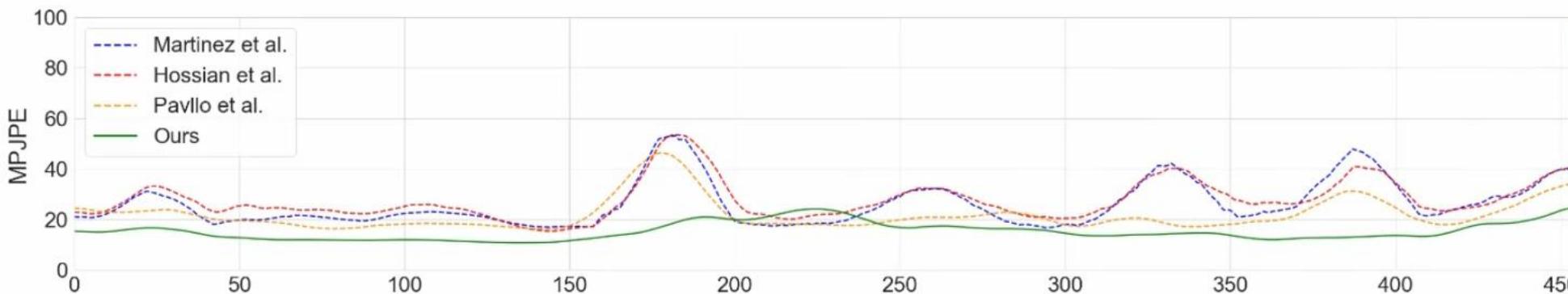
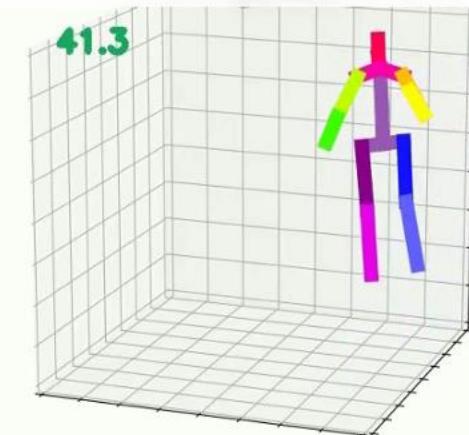
Pavllo et al.



Hossain et al.

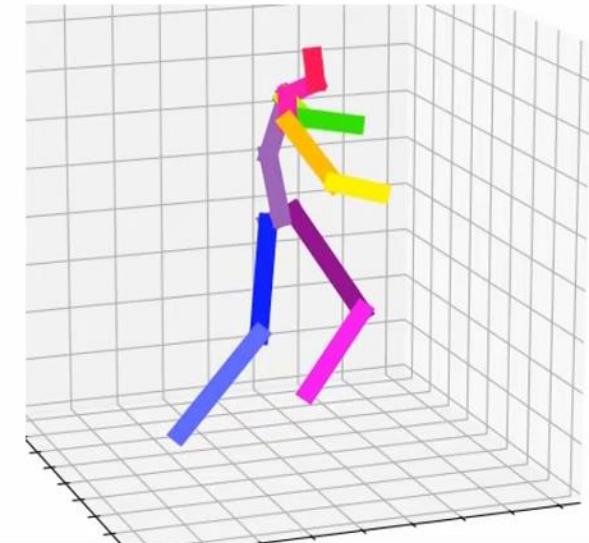
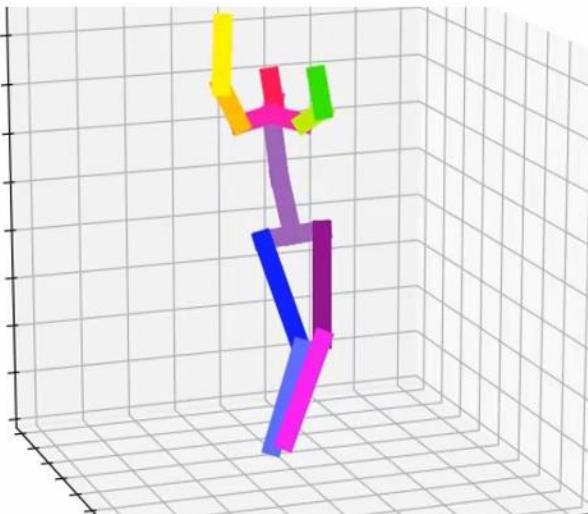
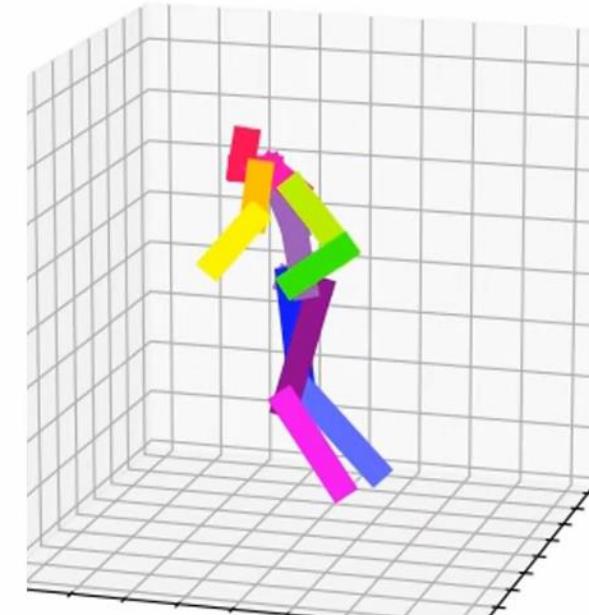
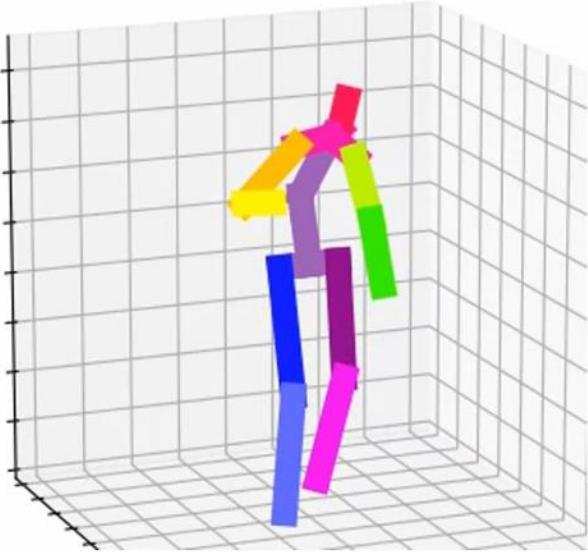


Martinez et al.



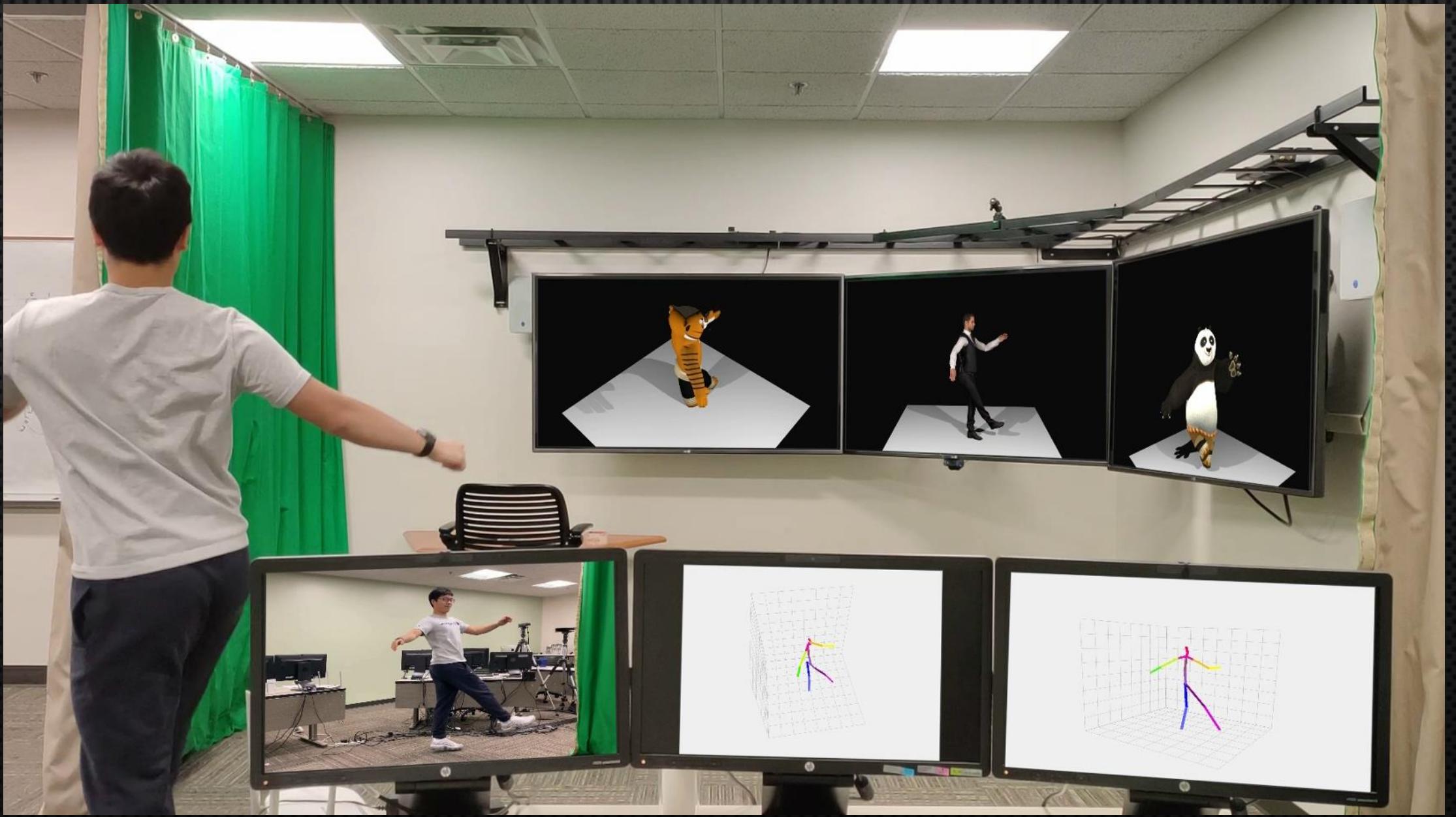
Qualitative Evaluation:

4. Results on wild videos



Qualitative Evaluation:

5. Real-time performance using the causal model



Thank you for watching



Code is available at: (<https://github.com/lrxjason/Attention3DHumanPose>)