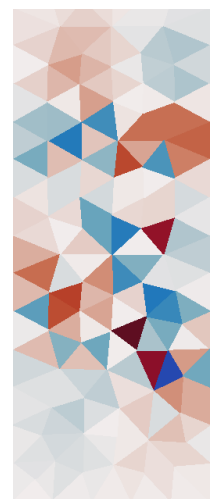


(a)

Input



Features



Output



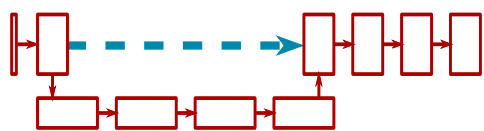
Projection

P

P^\dagger

Linear

U-Net



(b)

Conv 7x7
depth-wise

Layer
Norm
+

Conv 1x1
+ Gelu

Conv 1x1

+

