Architecture Configurations	
MR architecture	CT architecture
Conv 1: 64×64 input	Conv 1: 128×128 input
Conv 1: 3×3 kernel, $64 \times 64 \times 32$ map Conv 2: 3×3 kernel, $64 \times 64 \times 32$ map	Conv 1: 3×3 kernel, $128 \times 128 \times 32$ map Conv 2: 3×3 kernel, $128 \times 128 \times 32$ map
-	2×2 Max-pooling layer 1, $64 \times 64 \times 32$ map
Conv 3: 3×3 kernel, $64 \times 64 \times 64$ map Conv 4: 3×3 kernel, $64 \times 64 \times 64$ map	Conv 3: 3×3 kernel, $64 \times 64 \times 64$ map Conv 4: 3×3 kernel, $64 \times 64 \times 64$ map
2×2 Max-pooling layer 2, $32 \times 32 \times 64$ map	
Conv 5: 3×3 kernel, $32 \times 32 \times 96$ map Conv 6: 3×3 kernel, 32×32 map	Conv 5: 3×3 kernel, $32 \times 32 \times 96$ map Conv 6: 3×3 kernel, $32 \times 32 \times 96$ map
2×2 Max-pooling layer 3, $16 \times 16 \times 96$ map	
Fused Conv 1: 3×3 kernel, $16 \times 16 \times 192$ map 2×2 Max-pooling layer 4, $8 \times 8 \times 96$ map Fused Conv 2: 3×3 kernel, $8 \times 8 \times 192$ map	
Fused Deconv 1: 3×3 kernel, $16 \times 16 \times 128$ map	
Deconv 2: 3×3 kernel, $32 \times 32 \times 92$ map Deconv 3: 3×3 kernel, $64 \times 64 \times 64$ map -	Deconv 2: 3×3 kernel, $32 \times 32 \times 92$ map Deconv 3: 3×3 kernel, $64 \times 64 \times 64$ map Deconv 4: 3×3 kernel, $128 \times 128 \times 32$ map
Connect 1: 1×1 kernel, $16 \times 16 \times 96$ map, pool 3$\rightarrow$ deconv 2 Connect 2: 1×1 kernel, $32 \times 32 \times 64$ map, pool 2$\rightarrow$ deconv 3 Connect 3: 1×1 kernel, $64 \times 64 \times 32$ map, pool 1$\rightarrow$ deconv 4	

Table 1: Fusion CT&MR segmentation neural network architecture