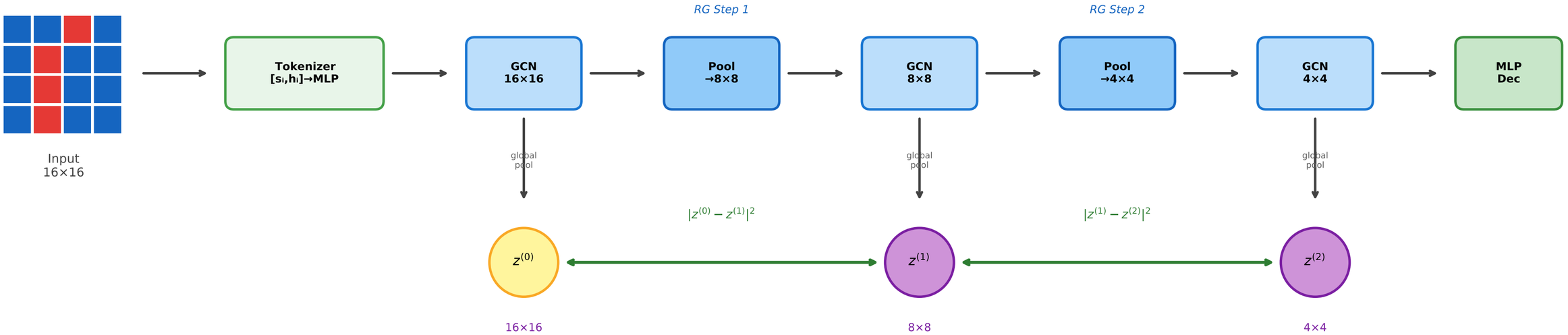
GNN-VAE with Multi-Scale RG Encoder

Each RG level outputs its own latent z



Latent Computation at Each Level:

Level 0 (16×16):

$h^{(0)} = GCN(x_{16\times16})$

$\mu^{(0)}, \sigma^{(0)} = Linear(GlobalPool(h^{(0)}))$

$z^{(0)} = \mu^{(0)} + \sigma^{(0)} \cdot \varepsilon$

Level 1 (8×8):

$h^{(1)} = GCN(Pool(h^{(0)}))$

$\mu^{(1)}, \sigma^{(1)} = Linear(GlobalPool(h^{(1)}))$

$z^{(1)} = \mu^{(1)} + \sigma^{(1)} \cdot \varepsilon$

Level 2 (4×4):

$h^{(2)} = GCN(Pool(h^{(1)}))$

$\mu^{(2)}, \sigma^{(2)} = Linear(GlobalPool(h^{(2)}))$

$z^{(2)} = \mu^{(2)} + \sigma^{(2)} \cdot \varepsilon$

Total Loss: $\mathcal{L} = \mathcal{L}_{recon} + \beta \cdot \mathcal{L}_{KL} + \lambda_{RG} \cdot \mathcal{L}_{RG}$

$$\mathcal{L}_{RG} = \frac{1}{2}\left(|z^{(0)} - z^{(1)}|^2 + |z^{(1)} - z^{(2)}|^2\right)$$

GCN Layer · Pooling (RG) · $z^{(0)}$ (Fine) · $z^{(i)}$ (Coarse) · Decoder