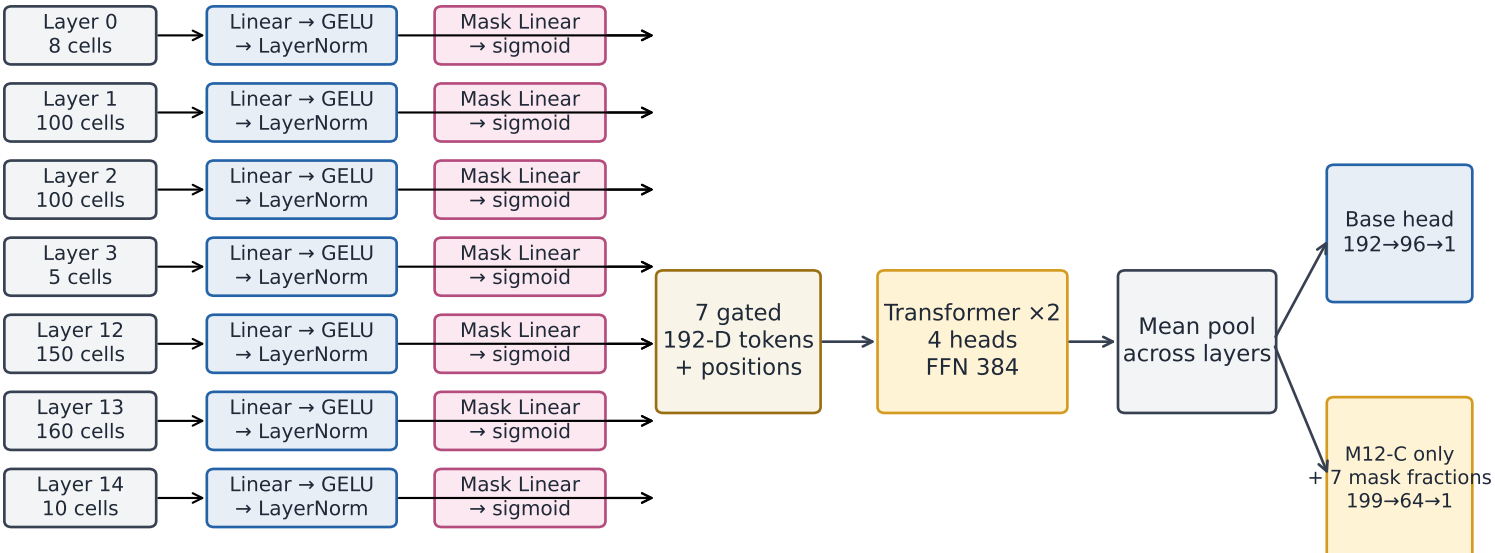
Layer-gated Transformer architecture

M10 and M11 share the base architecture; M12-C adds the detector-state residual branch.