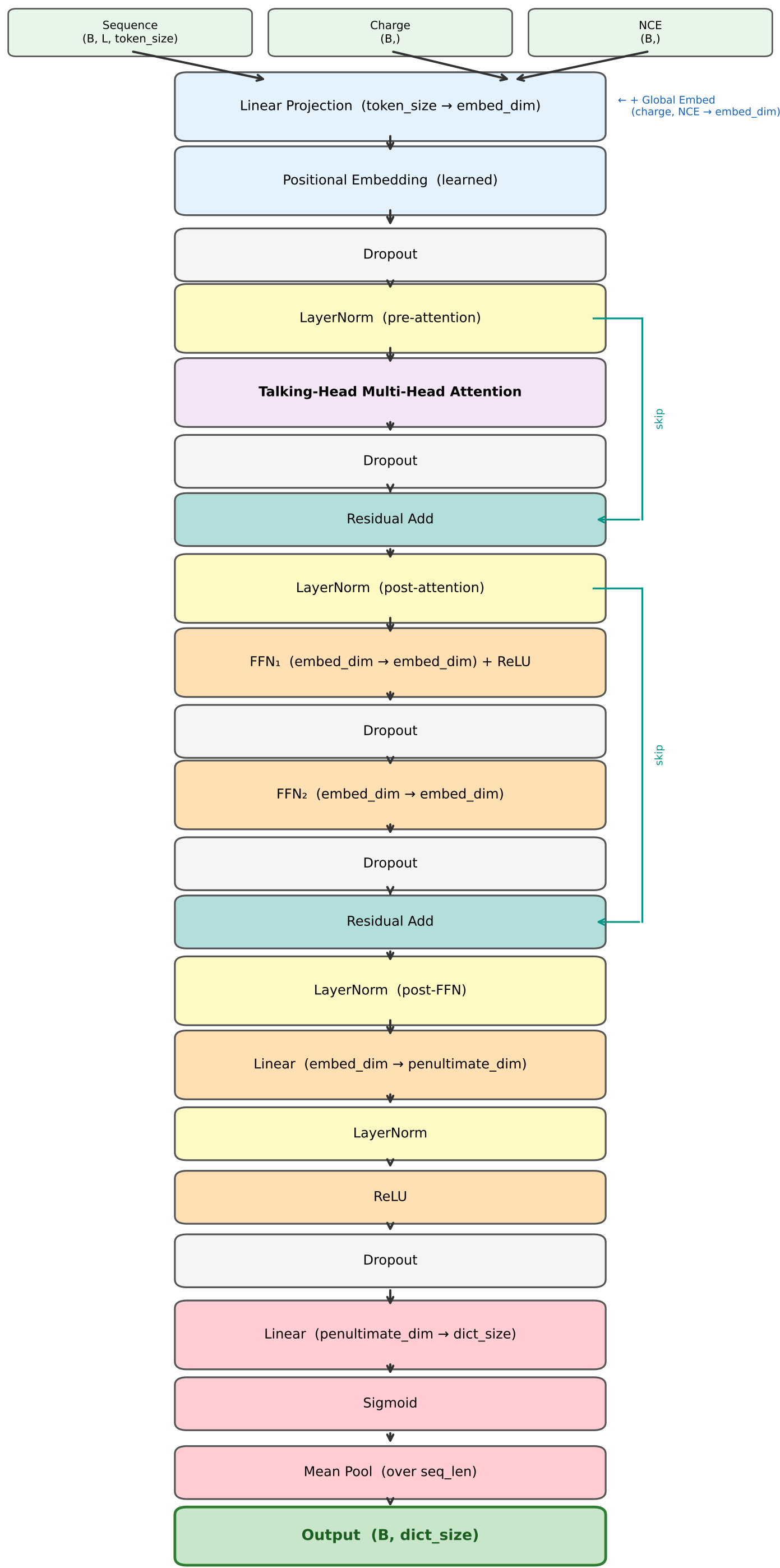


SpecFormer Architecture



- Legend**

Input

LayerNorm

Feedforward

Residual

Embedding / Projection

Attention

Dropout

Output

Default dims: embed_dim=256 | num_heads=64 | head_dim=4 | penultimate_dim=2048 | dict_size=26916