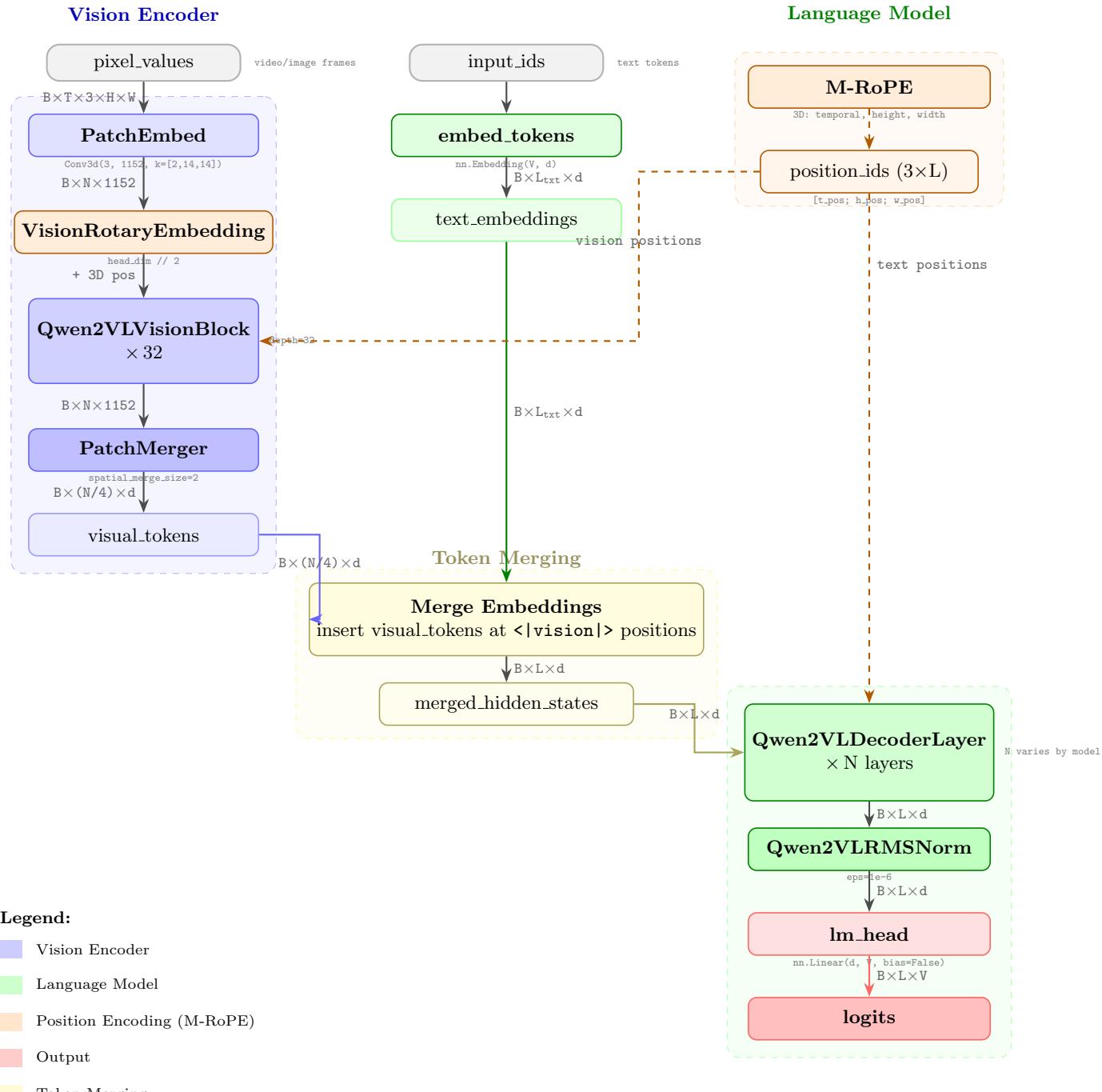


Qwen2VLForConditionalGeneration

Overall Architecture – Forward Pass



Model Variants:

- 2B: $d=1536, N=28, V=151936$
- 7B: $d=3584, N=28, V=152064$
- 72B: $d=8192, N=80, V=152064$
- Vision: $\text{embed_dim}=1152, \text{depth}=32$