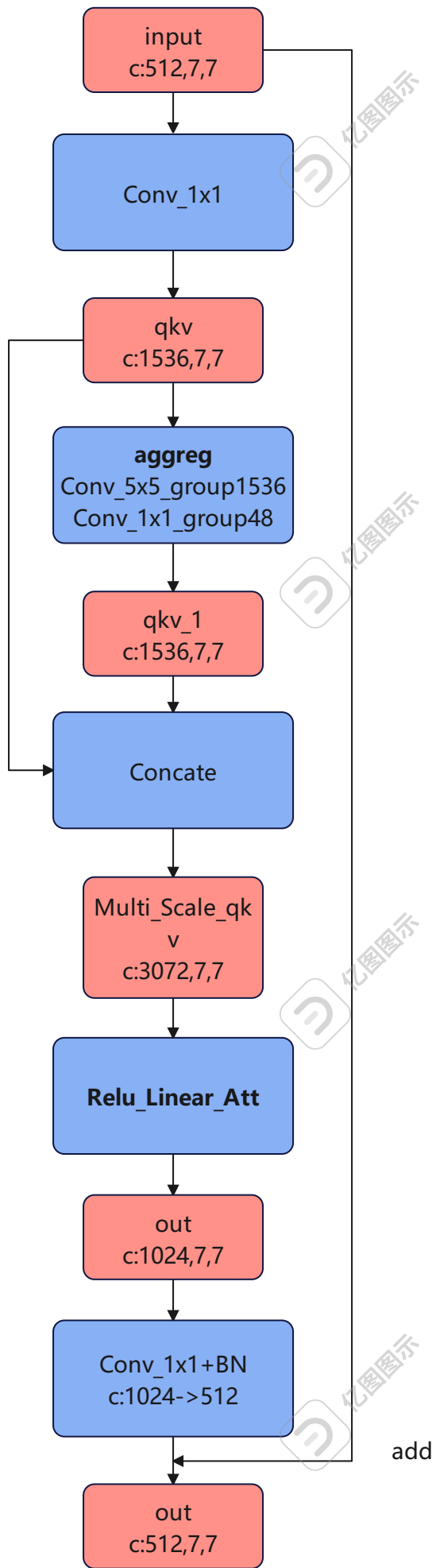
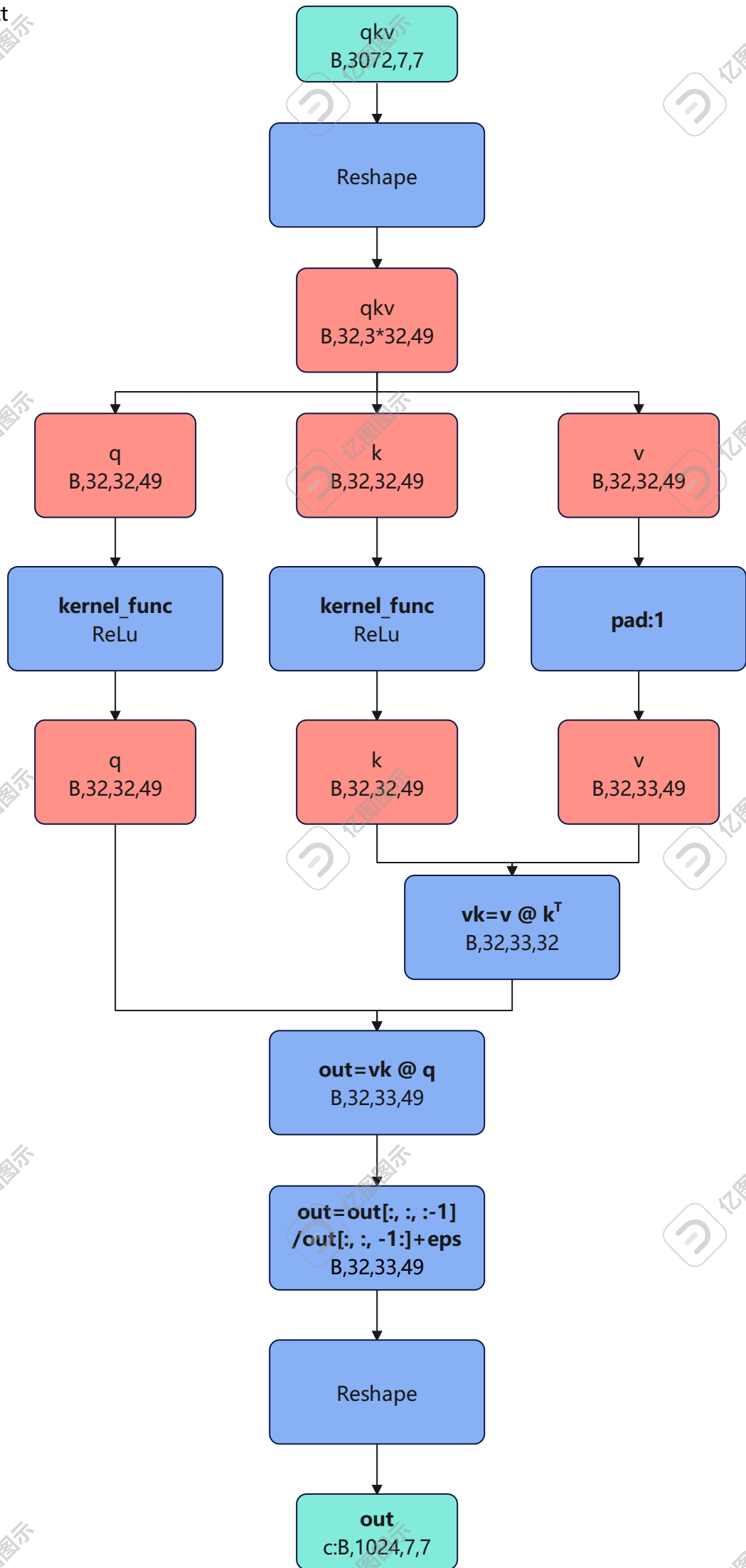




MLA



Relu_Linear_
Att



$$Sim(Q, K) = \exp(\frac{QK^T}{\sqrt{d}})$$

$$Sim(Q, K) = \text{ReLU}(Q)\text{ReLU}(K)^T.$$

$$O_i = \sum_{j=1}^N \frac{\text{ReLU}(Q_i)\text{ReLU}(K_j)^T}{\sum_{j=1}^N \text{ReLU}(Q_i)\text{ReLU}(K_j)^T} V_j$$
$$= \frac{\sum_{j=1}^N (\text{ReLU}(Q_i)\text{ReLU}(K_j)^T) V_j}{\text{ReLU}(Q_i) \sum_{j=1}^N \text{ReLU}(K_j)^T}.$$

$$O_i = \frac{\sum_{j=1}^N [\text{ReLU}(Q_i)\text{ReLU}(K_j)^T] V_j}{\text{ReLU}(Q_i) \sum_{j=1}^N \text{ReLU}(K_j)^T}$$
$$= \frac{\sum_{j=1}^N \text{ReLU}(Q_i) [(\text{ReLU}(K_j)^T V_j)]}{\text{ReLU}(Q_i) \sum_{j=1}^N \text{ReLU}(K_j)^T}$$
$$= \frac{\text{ReLU}(Q_i) (\sum_{j=1}^N \text{ReLU}(K_j)^T V_j)}{\text{ReLU}(Q_i) (\sum_{j=1}^N \text{ReLU}(K_j)^T)}.$$