

How much knowledge  
is useful?

How much knowledge  
is changed?

How complex the  
knowledge change is?

How much existing  
knowledge is used?

$$H = m_p \circ \left( \left( \underbrace{W + s \cdot m_o}_{\text{Pre-trained knowledge}} \circ \underbrace{W_B (m_r \circ W_A \circ m_i)}_{\text{Task-specific knowledge}} \right) X + b \right)$$

Freeze

Pre-trained knowledge

Tuning

Task-specific knowledge