$$C(k, j) = \gamma_{mask} C_{mask}(\widehat{m}^{(k)}, m^{(j)}) + C_{cls}(\widehat{l}^{(k)}, l^{(j)}),$$