

Model	Loss	Update	DP	eval	
				SI	SD
HDNN- $H_{128}L_{10}$	\mathcal{L}^{CE}	All	2	29.4	28.8
HDNN- $H_{128}L_{10}$	\mathcal{L}^{CE}	Gates	2	29.4	28.7
HDNN- $H_{128}L_{10}$ -KD	\mathcal{L}^{KD}	All	1	28.4	27.5
HDNN- $H_{128}L_{10}$ -KD	\mathcal{L}^{KD}	Gates	1	28.4	27.8
HDNN- $H_{128}L_{10}$ -KD	\mathcal{L}^{CE}	All	2	28.4	27.7
HDNN- $H_{128}L_{10}$ -KD	\mathcal{L}^{CE}	Gates	2	28.4	27.1