

Method	Training			Inference	
	FLOPs	$ \Theta $	$ D $	FLOPs	$ \Theta $
<i>New Architectures</i>					
Transformer	1x	1x	1x	1x	1x
Dynamic Conv	1x	1x	1x	1x	1x
<i>Semi-Supervised</i>					
NMT+BERT	60x	3x	>25x	3x	3x
Back-Translation	2x	1x	>50x	1x	1x
<i>Evolution</i>					
(So et al., 2019)	>15000x	1x	1x	1x	1x
<i>Our Data Diversification</i>					
Default Setup	7x	1x	1x	1x	1x