	BLEU	TF	BOOK	VIS	SCENE	CNN (training)	CNN (inference)	CRF (trainin
Per movie-book pair	6h	10 min	3 min	2h	1h	3 min	0.2 min	5h