

F1 scores per overlap threshold on testset

