Les images contiennent des données éparses (fragments de textes) avec une distribution spatiale précise (obligatoire à conserver) et des échelles différentes.

Plusieurs méthodes éventuellement possibles sont éliminées, on retient 2 modèles CUTIE-A et -B.

CNNs :