Input Image

RoI pooled features

Node identification and Categorization (Text and Non-Text)

Node Indexing and structured graph creation

Iterative node merging to reduce text bias

Final structured graph to perform distillation

Text1 Text2 Text3 NT1 NT2 Text4 Text5 Text6 NT3

Text[1,2,3] NT1 NT2 Text4 Text5 Text6 NT3

Text[1,2,3] NT1 NT2 Text[4,5,6] NT3