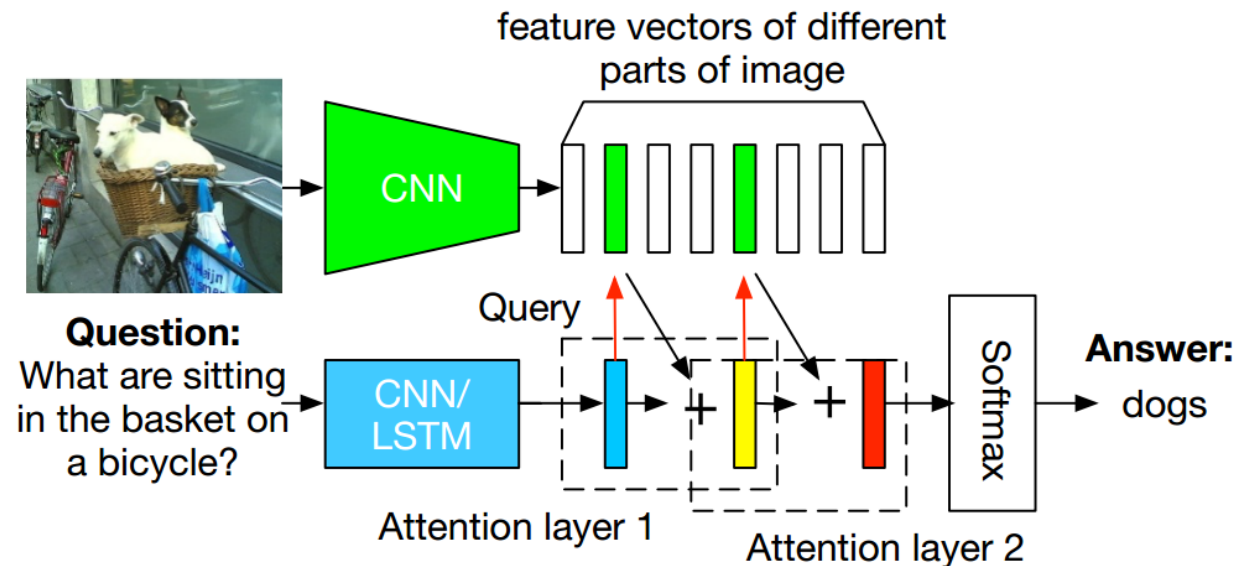




ICDAR 2019 Robust Reading Challenge Scene Text Visual Question Answering

IFLYTEK Research & USTC
frwang@iflytek.com

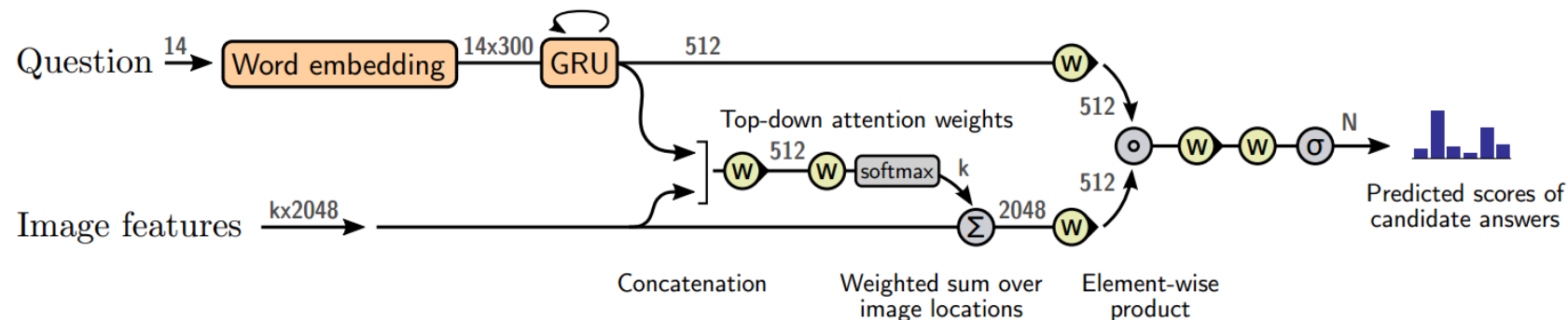
For VQA Problem:



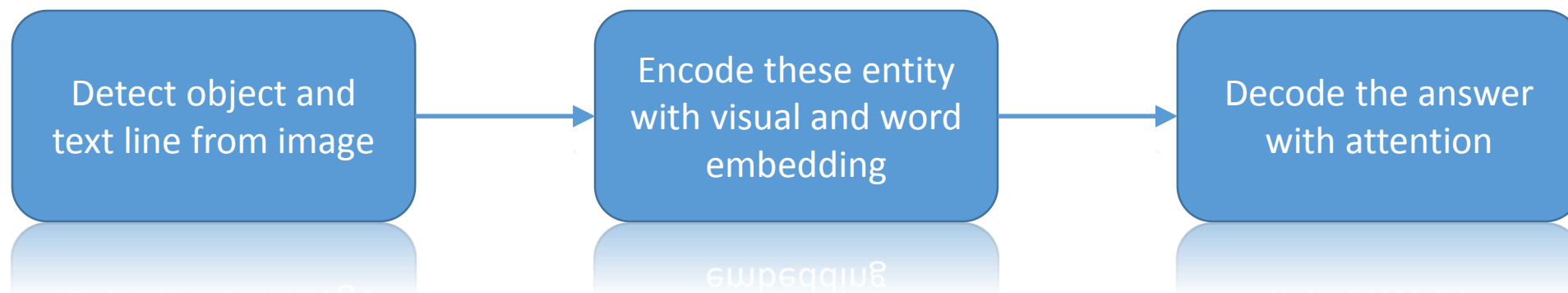
Attention Mechanism

For ST-VQA Problem:

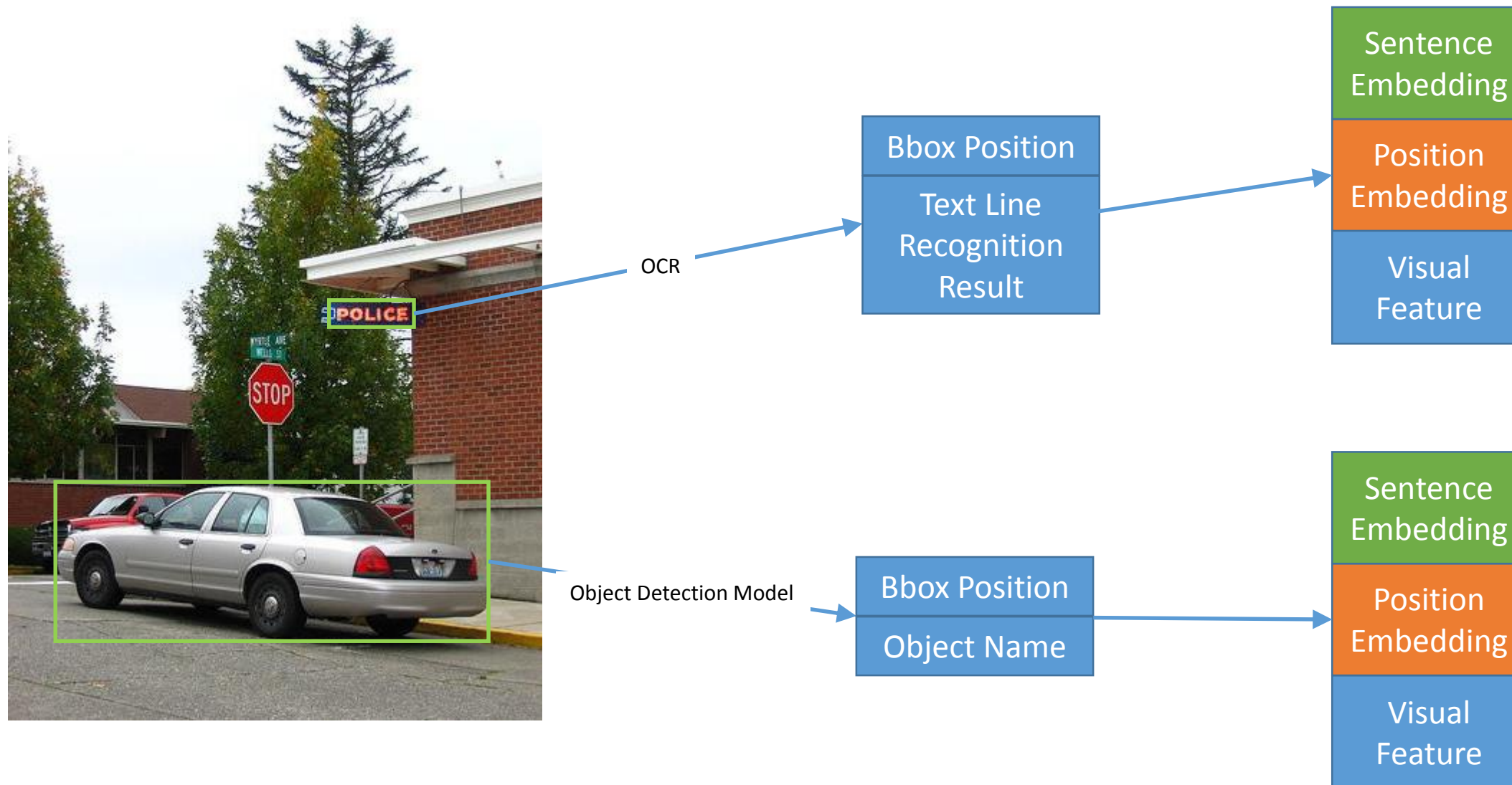
Pay more attention to text information!

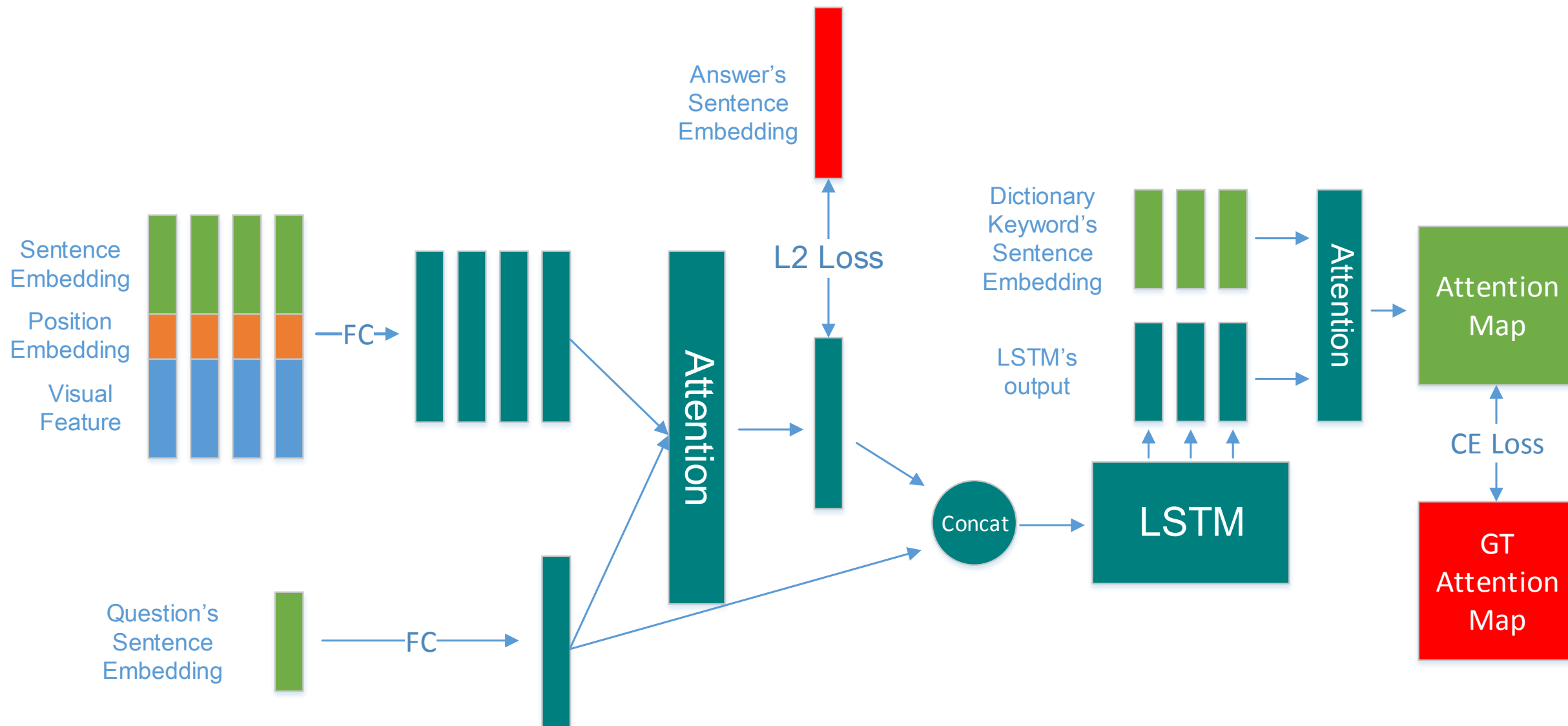


Top-Down Strategy

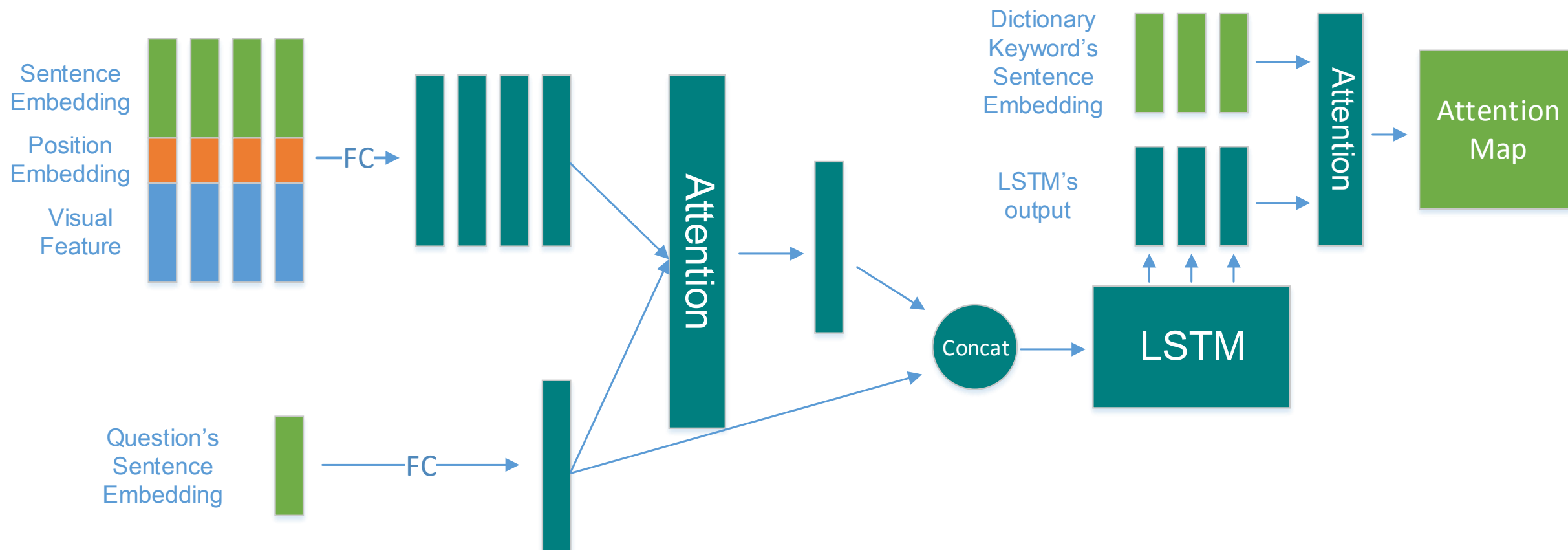


Visual Feature Extract





Task1 Inference



Task2 & Task3 Inference

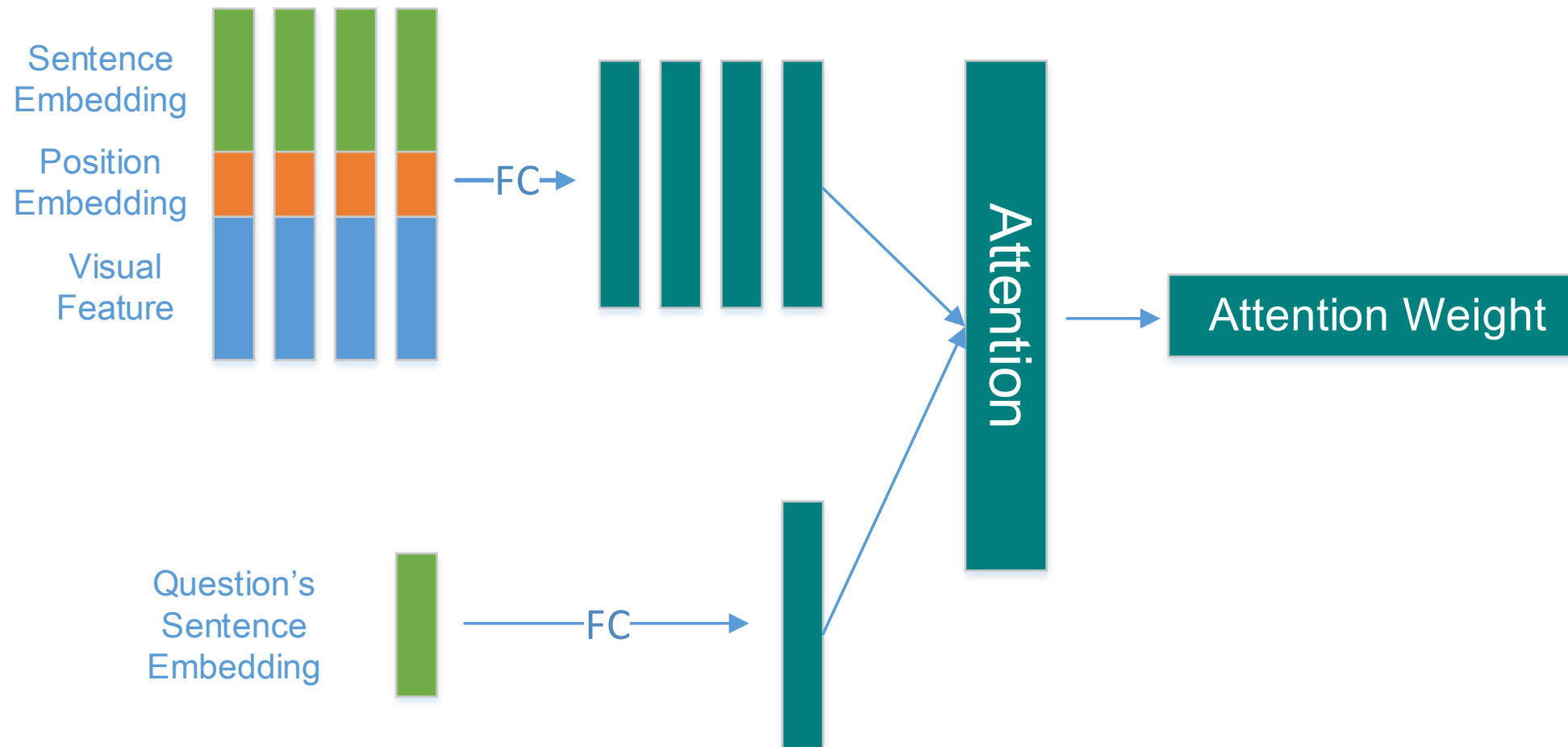


Table 1

	CE Loss	L2 Loss	Valid Score (ANLS)
Task1	√		0.4939
	√	√	0.4963
Task3	√		0.2096
	√	√	0.3002

Table 2

	Bert Type	Valid Score (ANLS)
Task1	Bert Base	0.4963
	Bert Large	0.4982
Task3	Bert Base	0.3002
	Bert Large	0.3257



Question: what color is this duck?
Pred Answer: orange



Question: name of this taxi company?
Pred Answer: comcab

W N A N S

