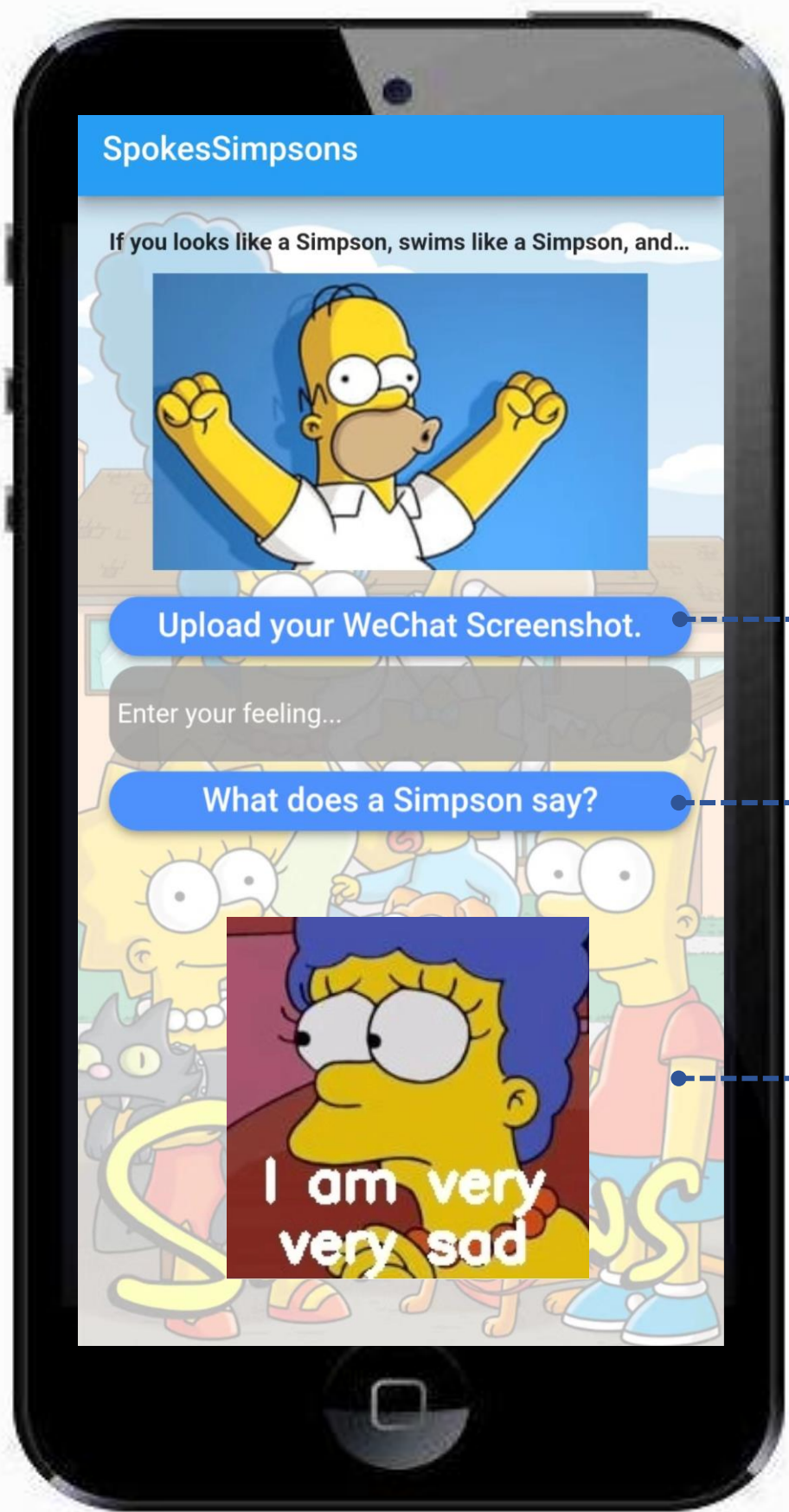


SpokesSimpsons

- Generate Stickers of the Simpsons via Wechat Screenshots

Team members: Fang Liu, Zeyu Yang, Yongqi Zheng



Upload your WeChat screenshots (English only)

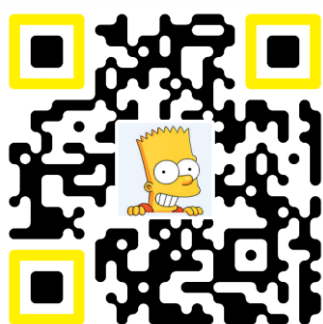
or

Submit what you want to say (English only)

Your stickers of the Simpsons will be here!

Overview

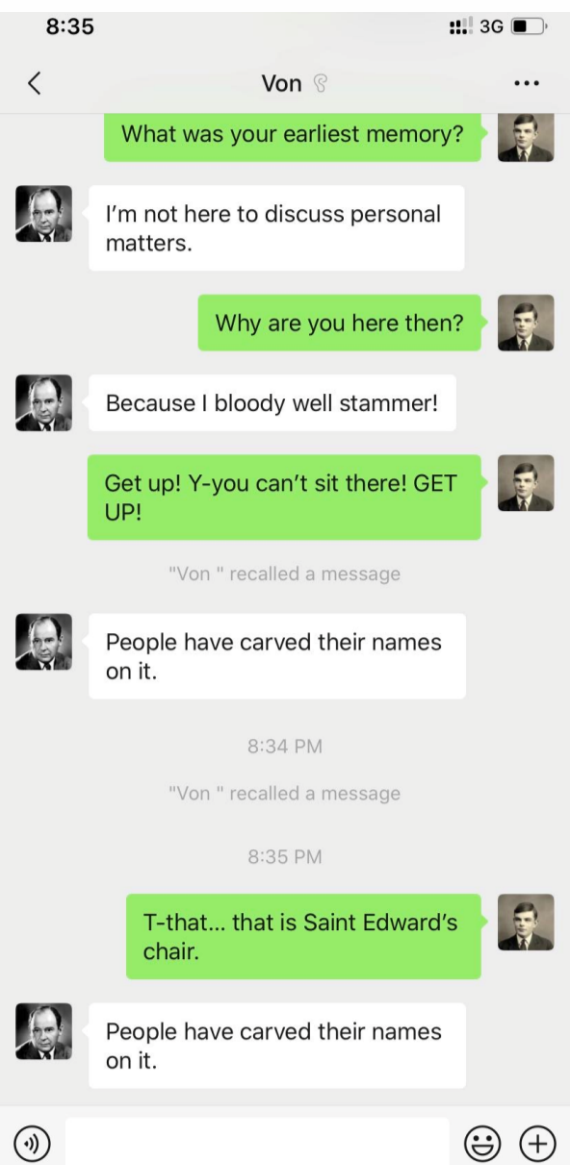
Want to use memes but can't find the right one? Don't know what to say when chatting? Here comes SpokesSimpsons, a meme generator!



You can choose to submit a sentence or upload a WeChat screenshot. We will find the dialogues and images in the dataset that most closely match the mood of the sentence you submitted (or the sentence we predict from historical dialogue), and post the dialogues on the images.

In particular, for submitted sentences or predicted sentences, we use Sentence BERT to get the mood embeddings, then find 3 top dialogues which have nearest embeddings (we have prepared their embeddings in advance). For these dialogues, find nearest images, then put each pair of sentences and pictures together.

OCR & Predict

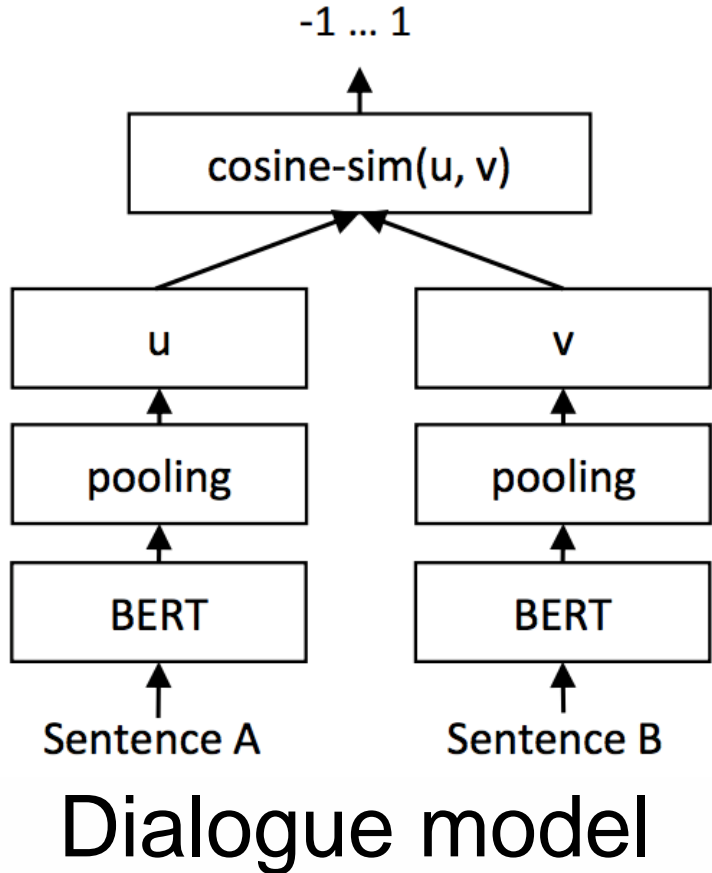


- OCR
Extract historical dialogues from WeChat screenshots. Our OCR module is consist of two parts: detection and recognition. We use a connectionist text proposal network based model to conduct detection. Recognition is handle by Tesseract.
- Predict
From historical dialogues predict the next sentence. Use the Pytorch port of OpenNMT, an open-source (MIT) neural machine translation system.

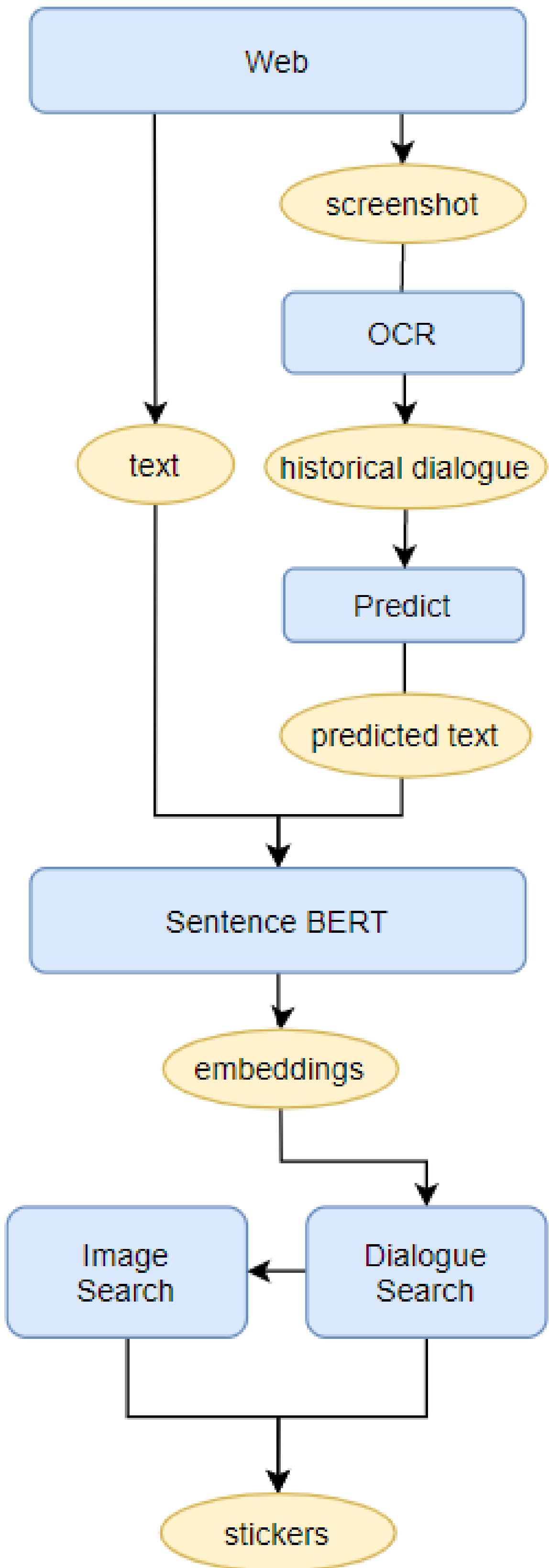
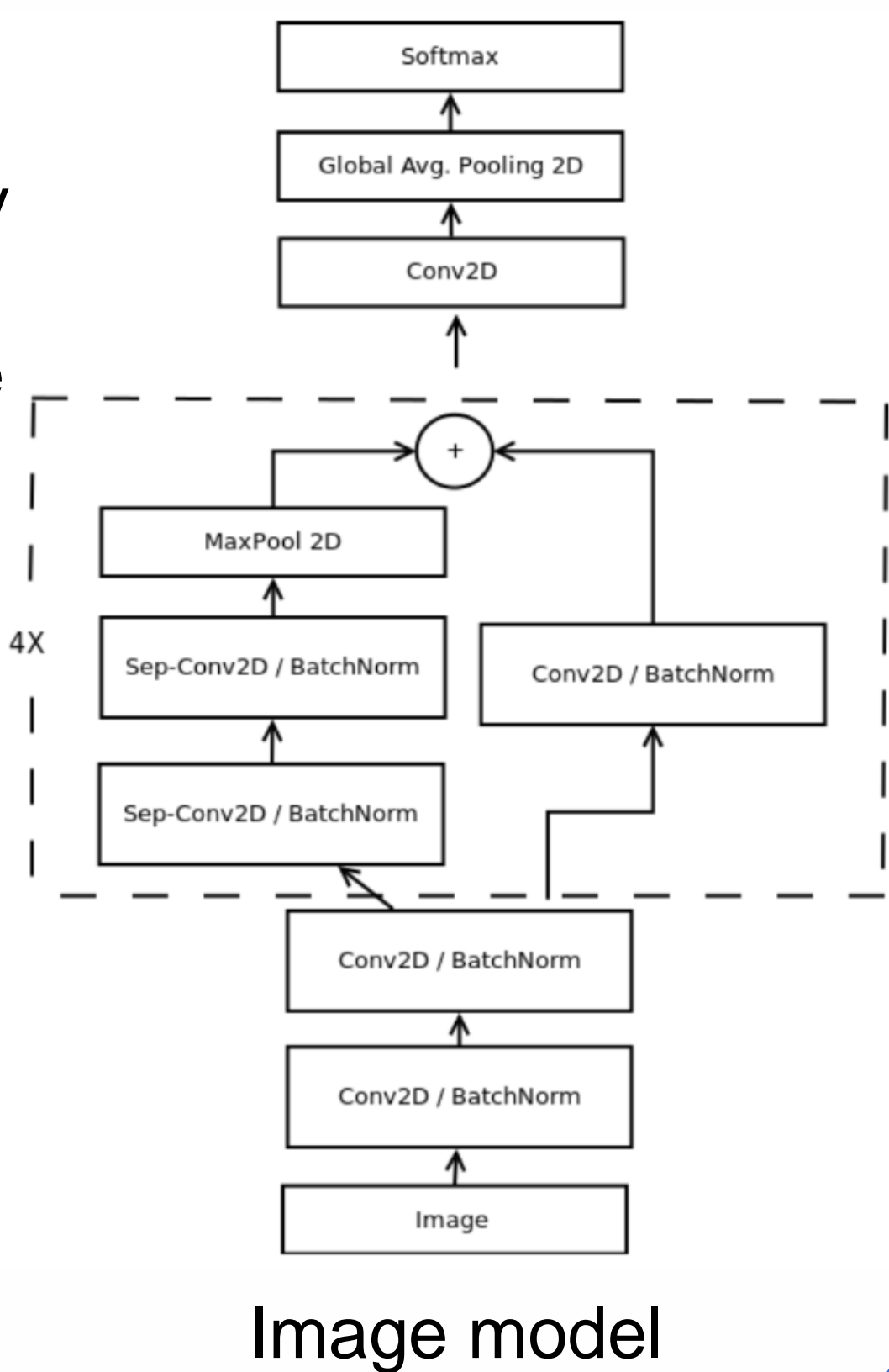
Image & dialogue search

We map both images and texts into a 7-dimension sentiment vector, reflecting respective degrees of *angry*, *disgust*, *scared*, *happy*, *sad*, *surprised*, and *neutral*. In this sentiment space, we provide each text the nearest pictures by L2 distance.

- Dialogue search
Given an input text either by user typed in or chatbot generated, we use pretrained sentence embeddings by Sentence-BERT to retrieve most semantically similar sentence from Simpson corpus as candidates. Then we map embeddings of each candidate text into sentiment vectors by calculate cosine similarity of respective sentiment word.



- Image search
We apply a pretrained emotion classification model on human faces and fine-tune it with some Simpson faces we've picked out from the dataset which have typical expressions.



Learn more at: <https://github.com/yongzhengqi/spokessimpsons>

投票区域

