

# Menu Expert

Chung Yu Wang<sup>1</sup>, Yu-Sheng Chen<sup>2</sup>, Zi Yu Lang<sup>2</sup>

<sup>1</sup>Department of Computer Science, Stanford University,

<sup>2</sup>Institute for Computational & Mathematical Engineering, Stanford University



## Motivation

When visiting foreign restaurants, people are often unfamiliar with the names on the menu. This is not only due to language barrier, but more importantly due to cultural differences. In this project, we would like to help people lower such hurdle by presenting a translated name of the dish along with the image of the dish on the side of the original dish name on the menu. Via our proposed pipeline, people may further extend the application to showing the information on a wearable VR device.

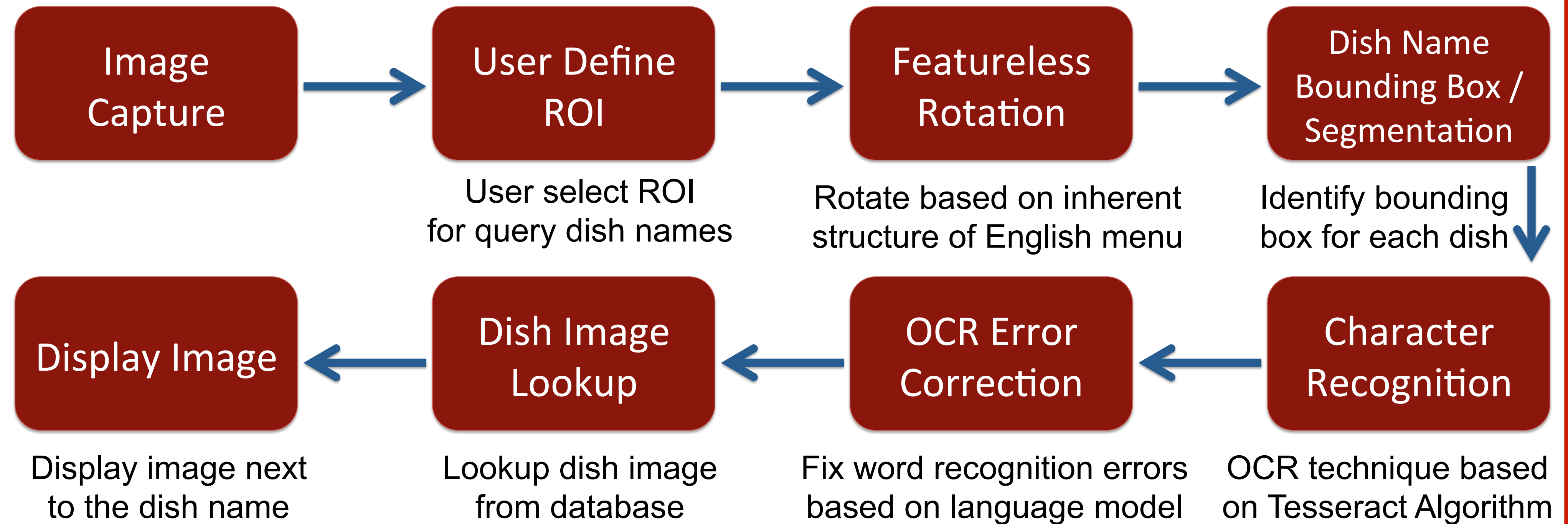
## Related Work

After studying current research and key techniques related to this topic, we find that the applications aiming at translating text content in images have common limitations:

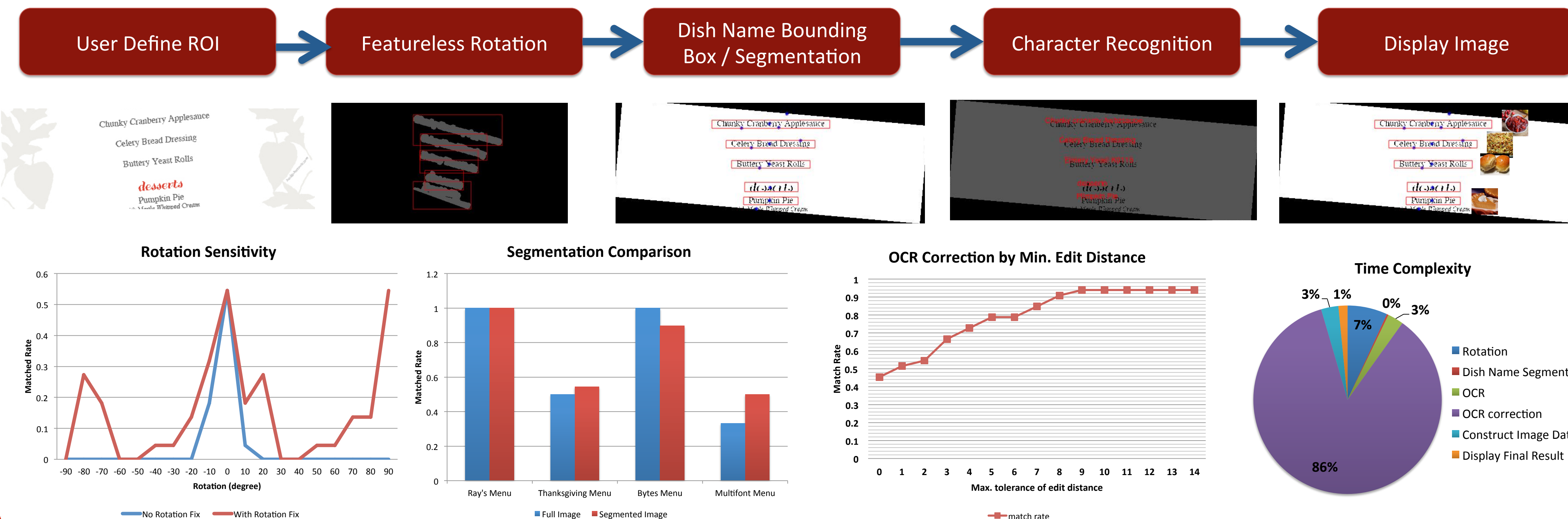
- Feature matching methods are not applicable to different font
- Lacking flexibility to incorporate meaning of individual characters
- Cannot always achieve satisfying results without preprocessing techniques

So we set our goal as building a text to image translation system for restaurant menus with emphasis on enhancing the robustness by implementing certain pre and post processing techniques.

## Menu to Dish Image Implementation Technique



## Experimental Results



## Conclusion & Future Work

OCR has been an active research in the past few years, ranging from SIFT based feature matching to the more modern day CNN techniques. However, these techniques often produce suboptimal results without any preprocessing techniques.

To build a robust menu recognition engine, we focused on building a digital image preprocessing implementation, by providing a way to automatically correct the input image's rotation, identified the interested ROI to reduce search space and noise. Since text recognition is inherently also a language model based problem, we further provided a post-processing edit-distance error correction to improve accuracy from the vanilla OCR engine. With our pipeline, the increase in accuracy varies depending on the cases, with some cases reaching over 50% increase in correct recognition rate.

Some remaining works include making the segmentation and rotation angle's underlying dilation pipeline more robust by building using a pyramid of different scales of the dilation structure.