

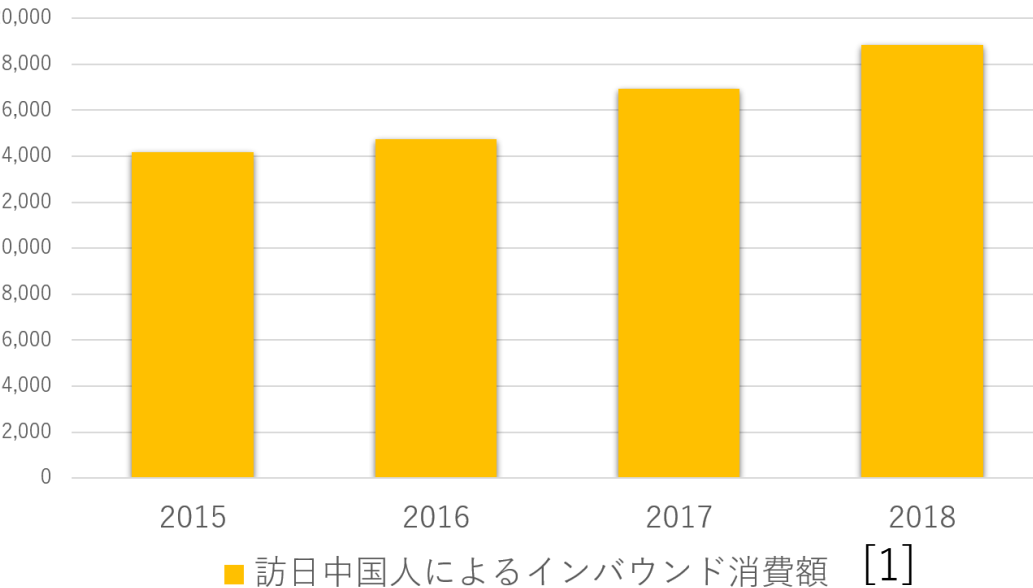


中国のマイクロブログ(Weibo) データ解析による中国人観光客向けの 注文サポートアプリ開発

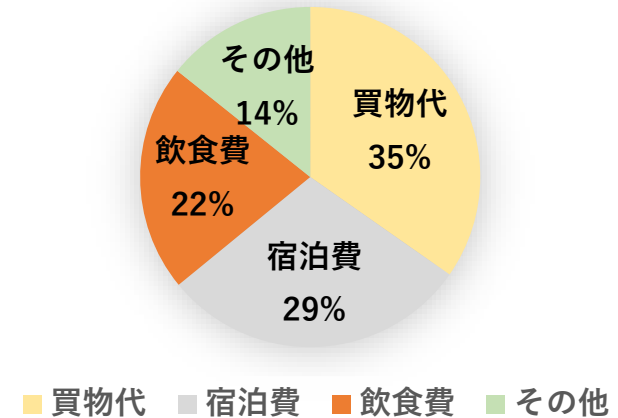
2015311050 孫イク
情報システムコース
(指導教員：小栗宏次)

研究背景

[1]日本政府観光局 (JNTO) 発表統計よりJTB総合研究所作成2018
[2]観光庁『訪日外国人の消費動向』



2018年訪日中国人旅行支出内訳



「食」がインバウンドの経済効果へ大きい影響を与える

■ 「食」についての課題

観光庁[2]：

観光客が最も困るを感じる場所：飲食店

最も困ること：注文のし辛さ (65.8%)

注文における課題

- 料理の内容が不明
- 料理が自分の好みに合うかどうか

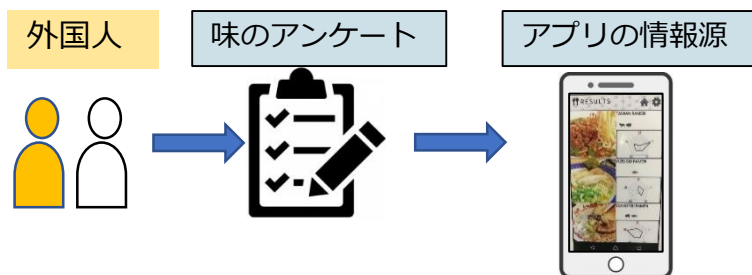


先行研究

[3]森田ら, “外国人向け日本食フードサーチアプリ”
「NEXT COMMUNICATION AWARDS 2016」

森田ら^[4]

外国人向け日本フードサーチアプリ



豚骨ラーメン



味のレーダーチャートの提供より対象店の料理注文サポート

■ 課題

- ・ 情報収集の局限性
- ・ 食文化の差より好みの差異が存在

農林水産省[4]: **好みの傾向性**

中国: 冷たい料理苦手

ヨーロッパ: 馬肉, 鯨肉, 活作り好まない

オーストラリア: 甘いもの好き

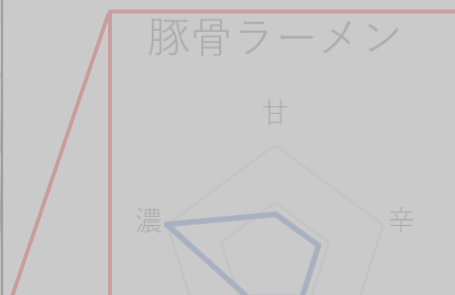
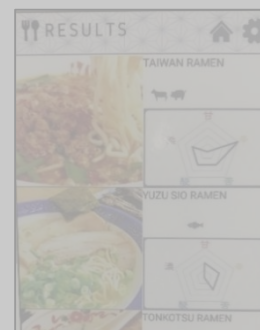
本研究

中国人の好みより中国人観光客向けの注文サポート

OK LAB.

Aichi Prefectural University

森田ら^[4]
外国人向け日本フードサーチアプリ



研究目的：中国のマイクロブログ「Weibo」の解析による中国人の好み情報を提供する注文サポートアプリの提案，開発

- それぞれ国の食文化により好みの差異が存在

農林水産省^[4]：好みの傾向

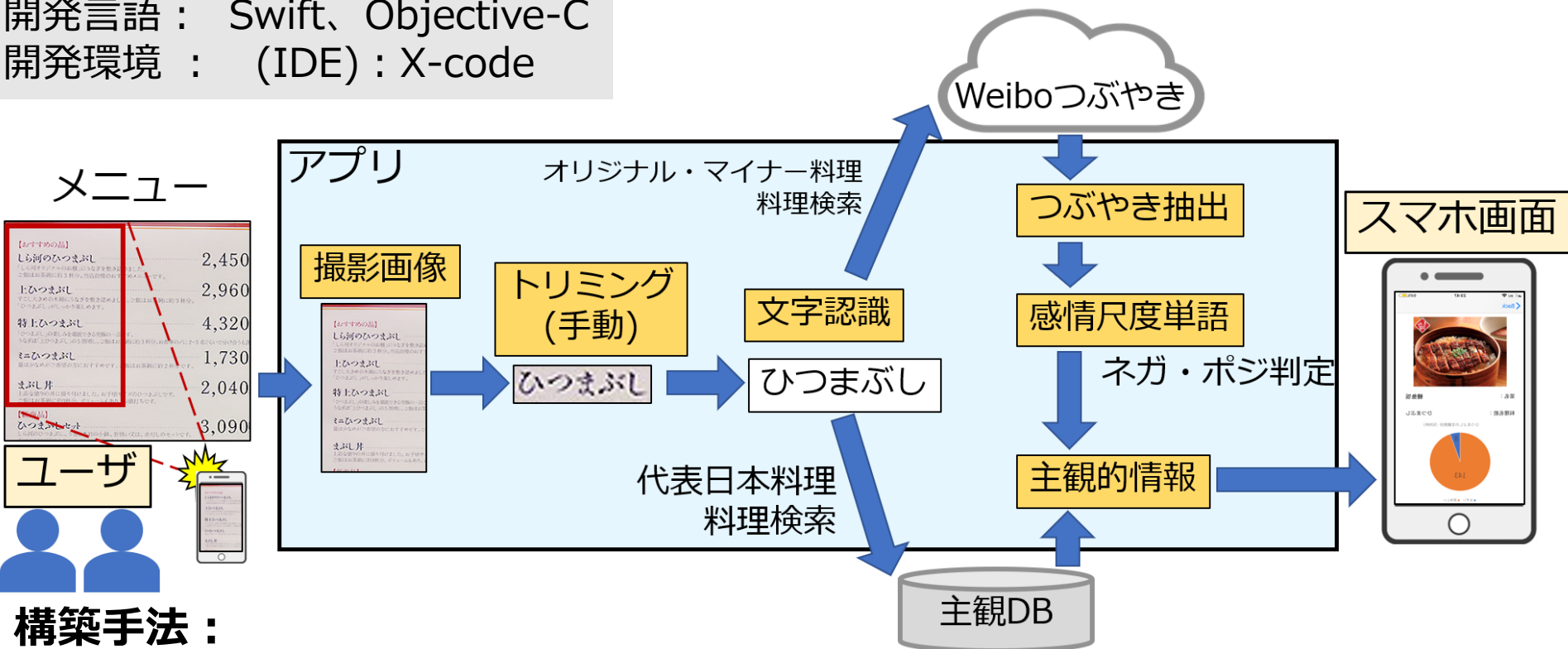
中国人：冷たい料理苦手

ヨーロッパ：馬肉，鯨肉，活作り好まない

オーストラリア：甘いものの好き

アプリ概要

開発言語 : Swift、Objective-C
開発環境 : (IDE) : X-code



構築手法 :

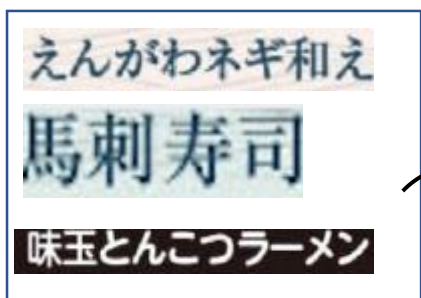


メニュー文字認識

■OCRライブラリ「SwiftyTesseract」の精度検証と実装

2種類日本語学習データ

- tessdata_best(正確性重視)
- tessdata_fast(速度重視)



100件料理名の写真

認識

Tessdata_best : 味 玉 と ん こ つ ラ ー メ ン
Tessdata_fast : 味玉とんこつラーメン

出力

STEP2

学習データ	特徴	正解率	平均計算時間
tessdata best	正確性重視	78%	0.278s
Tessdata fast	速度重視	89%	0.108s

アプリ実装 :

tessdata_fast(速度重視)



「Weibo」 つぶやき情報抽出

マイクロブログ特徴：

ユーザの趣味や行動を反映する

中国のマイクロブログ「Weibo」

認識された料理名

料理のつぶやきを抽出

■ 代表料理

ひつまぶし, 納豆, 天ぷらなど

特徴：料理名の認知度が高い

「Weibo」情報量が豊富

スクレイピングツールより大量収集

料理名	ひつまぶし	納豆	馬刺身	白子
収集件数	520	1033	673	616
時間(分)	70	120	108	80

鰻魚飯真的好好吃, 满满足足
ひつまぶし美味しい大満足だ

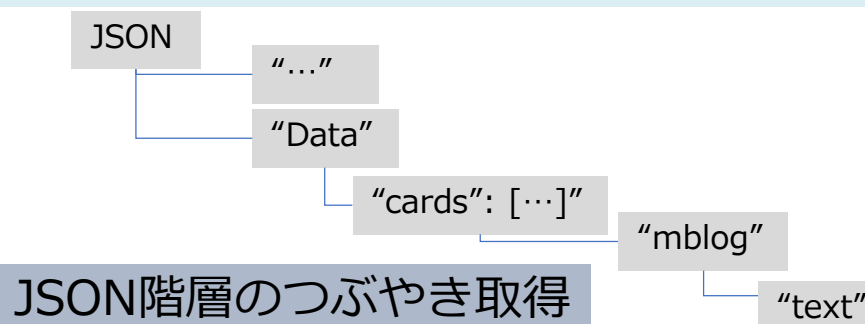
■ オリジナル料理

辛味チキン, 極上ヒレカツ

特徴：店のオリジナル料理名

メニュー数が膨大

「Weibo」JSONデータ解析より収集



中国人の好みポジ・ネガ度抽出

「好み」尺度：

つぶやき中の単語「美味しい」と「まずい」の割合

ポジティブ	ネガティブ
好吃（美味しい）	难吃（まずい）

■ 代表料理

抽出つぶやき

■ オリジナル料理

抽出つぶやき

今天的鰻鱼饭真的太好吃了！中午吃的鰻鱼饭很好吃。随手拍的鰻鱼饭。鰻鱼饭也太腥了 难吃

今天的鰻鱼饭真的太~~好吃~~了！中午吃的鰻鱼饭很~~好吃~~。随手拍的鰻鱼饭。鰻鱼饭也太腥了 难吃

好吃 照合

[今天的鰻鱼饭真的太], [了！中午吃的鰻鱼饭很], [随手拍的鰻鱼饭。鰻鱼饭也太腥了 难吃]

Array.length = 3, Count(好吃) = 3 - 1 = 2回

料理名	抽出件数	ポジ/ネガ度
ひつまぶし	520	143/8
納豆	1033	89/23
馬刺身	673	123/14
白子	616	12/8
天ぷら	652	185/4

料理名	抽出件数	ポジ/ネガ度
辛味チキン	10	2/0
エスカルゴ	10	1/1
マルゲリータピサ	10	1/0

実験検証

実験目的：

本研究で開発されたアプリの有効性検証

実験対象：

初回訪日中国人観光客3名

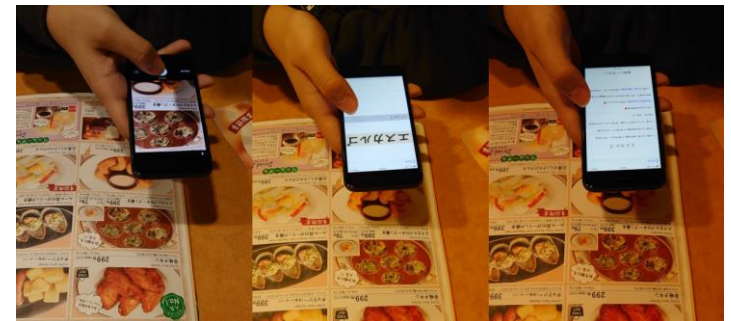
実験場所： くら寿司（代表料理）
 サイゼリヤ（オリジナル料理）

実験方法： 二つの店で本アプリを通じ、注文サポートの有効性をアンケートで回答してもらう。

実験風景：



代表日本料理



オリジナル料理



実験結果

提案の有効性：高い有効性がある

- 写真がない場合
- 多数の料理対応
- 好みの傾向性があることを認める

現段階アプリの有効性：有効性がある

代表料理	評価結果 (5>4>3>2>1)
全体の実用性	5,5,5
認識精度	4,4,4
好み情報の参考性	4,4,4

- イラストと翻訳名が見れるよかった
- データ料理が多いため、信頼性がより高い
- 白子料理さすがに中国人は無理

オリジナル料理	評価結果 (5>4>3>2>1)
全体の実用性	5,5,5
認識精度	4,4,3
好み情報の参考性	4,4,3

- オリジナル料理の情報量が少ない
- オリジナル料理のつぶやき内容が確認でき
- エスカルゴを避けてよかった



まとめ

研究目的

「Weibo」から中国人の好み抽出手法を用いて
中国人向けの注文サポートアプリの提案，開発

アプローチ

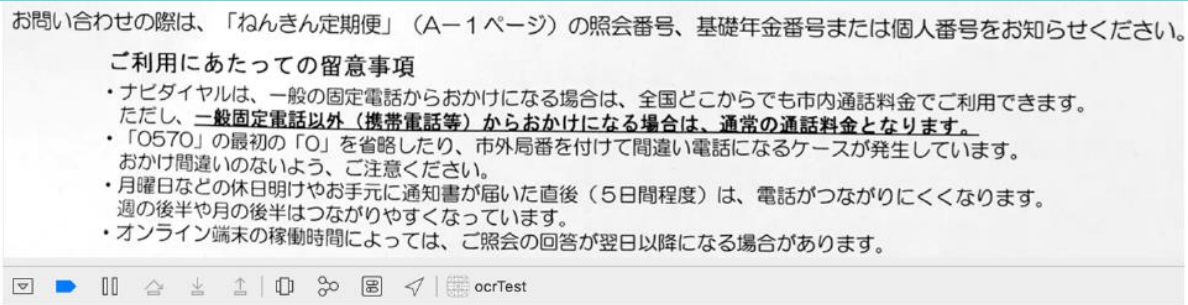
1. メニュー認識機能
2. 「Weibo」つぶやき抽出
3. つぶやきポジ・ネガ度抽出

結果

1. 提案の有効性が高い評価
2. 開発されたアプリ有効性を確認



なぜスピード重視のほうが精度が高い？



誤字率：4/262

お問い合わせの際は、「ねんきん定期便」（A-1 ページ）の照会番号、基礎年金番号または個人番号をお知らせください。

ご利用にあたっての留意事項

- ・ナビダイヤルは、一般の固定電話からおかけになる場合は、半国としからでも市内通話料金でご利用できます。

ただし、

- ・「0570」の最初の「0」を省略したり、市外局番を付けて間違い電話になるケースが発生しています。

おかけ間違いのないよう、ご注意ください。

- ・月曜日などの休日8時めお手元に通知書が届いた直後（5日間程度）は、電話がつながりにくくなります。

週の後半や月の後半はつながりやすくなっています。

- ・オンライン端末の稼働時間によっては、ご照会の回答が翌日以降になる場合があります。

1.2855770587921143

誤字率：1/262

図 3.16: tessdata.fast (LSTM 正確性重視)

お問い合わせの際は、「ねんきん定期便」（A-1 ページ）の照会番号、基礎年金番号または個人番号をお知らせください。

ご利用にあたっての留意事項

- ・ナビダイヤルは、一般の固定電話からおかけになる場合は、全国どこからでも市内通話料金でご利用できます。

ただし、

- ・「0570」の最初の「0」を省略したり、市外局番を付けて間違い電話になるケースが発生しています。

おかけ間違いのないよう、ご注意ください。

- ・月曜日などの休日明けやお手元に通知書が届いた直後（5日間程度）は、電話がつながりにくくなります。

週の後半や月の後半はつながりやすくなっています。

- ・オンライン端末の稼働時間によっては、ご照会の回答が翌日以降になる場合があります。

4.8244229555130005

図 3.17: tessdata.best (LSTM 正確性重視)

本研究：

- ・メニュー画像対応
- ・キーワードの取得
(スペースなし)
- ・処理スピード

代表料理とオリジナル料理

代表料理：

- ・ グーグル中国語訳名で500件以上スクレイピング取得できる
- ・ 中国語訳名が多くの人に理解されている
- ・ 店を限定することができない

オリジナル料理：

- ・ グーグルの中国語訳名が中国人に理解されていない
- ・ 料理を提供する店に限定できる



代表料理

- 「Weibo」情報量が多い（処理時間，「Weibo」収集量の制限）
- 知名度が高い（検索される確率が高い）
- 観光客の目的性が高い（高い信頼性）

オリジナル料理

- 「Weibo」情報量が少ない
- メニューのオリジナル性が高い



アプリの操作

