

食事画像認識を用いた モバイルレシピ推薦システム

電気通信大学

丸山拓馬 秋山瑞樹 柳井啓司

❖ 背景



画像認識技術の進歩



スマートフォンの普及
処理能力が向上



モバイルデバイス上でも
画像認識が実行可能に



❖ 提案内容

モバイル



レシピ検索



物体認識



買い物時のレシピ検索
直感的に操作可能

モバイル上で
食材認識によるレシピ検索は
本研究が初めての試み

❖ 目的

□ 食材画像認識でレシピ検索



一般物体認識

スマートフォン内で認識

動画で認識

30種類の食材が対象

❖ 目的

食事作成のどの部分
で利用するか？

買い物中の
レシピ検索を支援する

1. 食事を作ることを決定



2. レシピを決定する



3. 食材を調達に出かける



4. レシピを確認・変更する



5. 食材を購入する

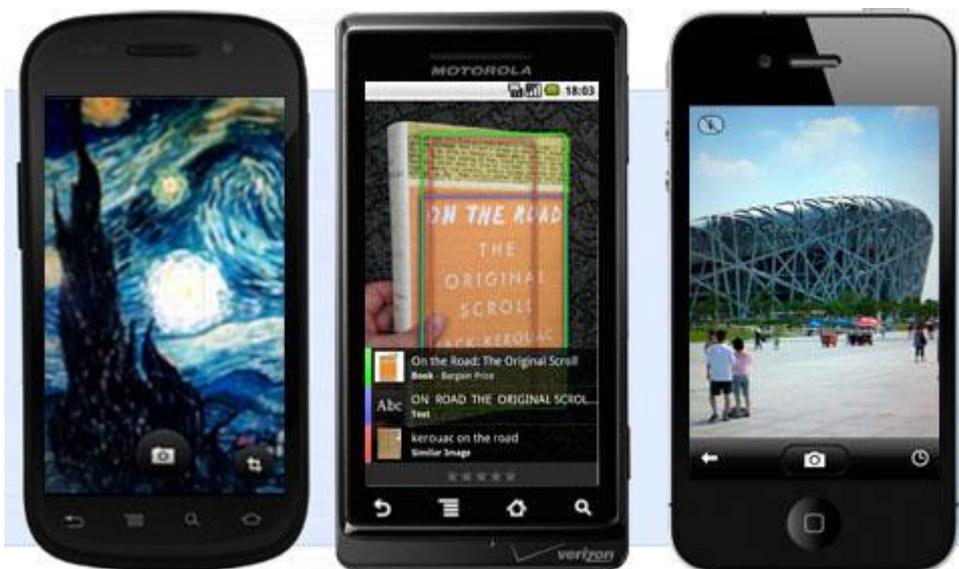


6. 調理をする



7. 料理を食べる

□ Google Goggles



- ランドマーク
- 本
- テキスト
- ロゴ

- 特定物体の認識
- 類似画像の検索

一般物体が対象では正しい認識結果が得られない

❖ 提案システムの流れ

1. スマートフォンを食材にかざす

2. 食材を認識

3. 食材のレシピデータを取得

4. レシピリストを表示

5. レシピの選択

6. レシピの表示



食材候補



1位



6位



❖ 提案システムの流れ

3. 食材 レシピリスト 6. レシピの 食材候補 トを表示

 <p>トマトとツナの簡単パスタ！ by LoveHeart</p>	
 <p>簡単～あと一品☆いんげんとトマトのサラダ by りすのお母さん</p>	 <p>肉</p>
 <p>鶏肉と野菜のトマト煮込み by sindy0922</p>	 <p>魚 tomato</p>
 <p>サバ入りトマトパスタ by にこちゃん111</p>	 <p>野菜 0.0[sec]</p>
 <p>超簡単トマトサラダ by nayuママ</p>	 <p>果物</p>

手動選択

取得

映

ク

❖ レシピデータの取得



COOKPAD



ユーザ投稿型レシピサイト

108万以上のレシピ

日本最大のレシピサイト

現状では新着レシピ順に
取得データをそのまま表示

改善の余地あり

❖ 画像認識手法

特徴について

Bag-of-Color Histogram, Bag-of-SURF
(KD-treeで高速化、1000次元コードブック)

マルチフレームで特徴統合(精度向上)

分類手法

線形SVM(非線形に比べ大幅に速度向上)

モバイルでの高速化に重点、1枚0.15秒で処理可能

❖ 認識にかかる時間

Galaxy S2

- 2011/6/3発売
- 1GHz Dual Core
- 1GB RAM



HTC Desire HD

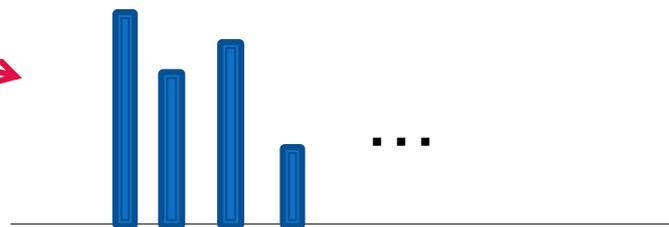
- 2010/11/12発売
- 1GHz
- 756MB RAM



機種名							平均
GAXALY S2	216	176	172	150	125	167.8	
HTC DesireHD	390	536	368	350	326	394.0	

単位はmsec

❖ カラーヒストグラム



各色次元を4分割

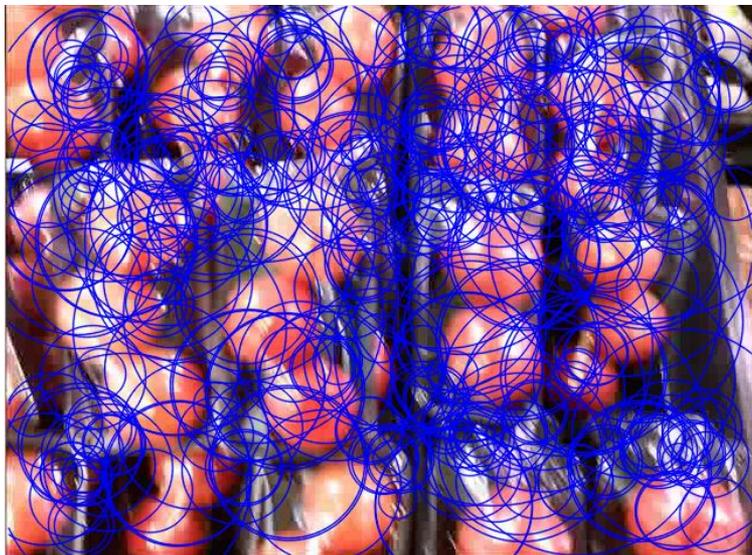
$4 \times 4 \times 4 = 64$ 次元特徴

12 x 12分割



1枚から144個の特徴抽出

❖ SURF

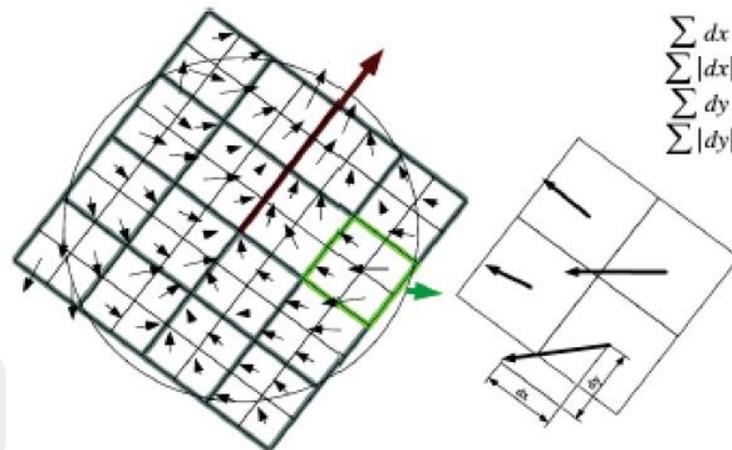


2通りの検出方法

- 通常のキーポイント検出
- 12x12, 24x24, 48x48, 96x96のマルチスケールグリッド検出

4x4の区間から4要素を抽出=64次元

$$\sum dx, \sum dy, \sum |dx|, \sum |dy|$$

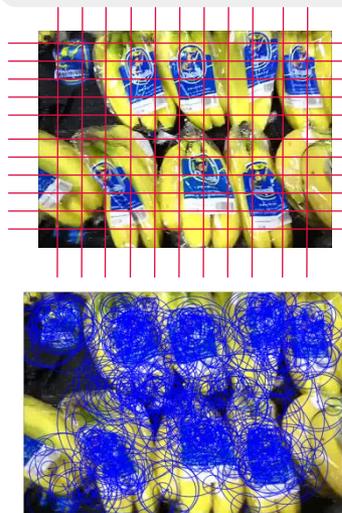


❖ Bag-of-Features手法

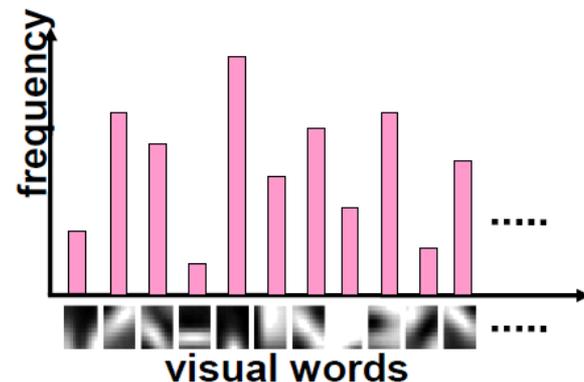
入力画像



特徴抽出



特徴ベクトル



モバイルシステムでのコードブックへの投票

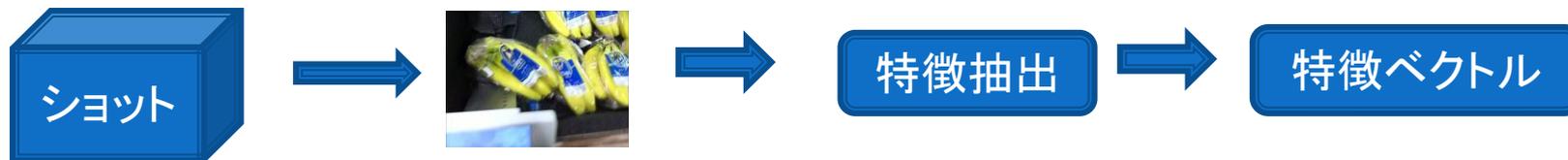
1000次元のコードブック作成(N=1000)

KD-treeを作成.O(logN)回の比較で対象のVisual Wordを検索

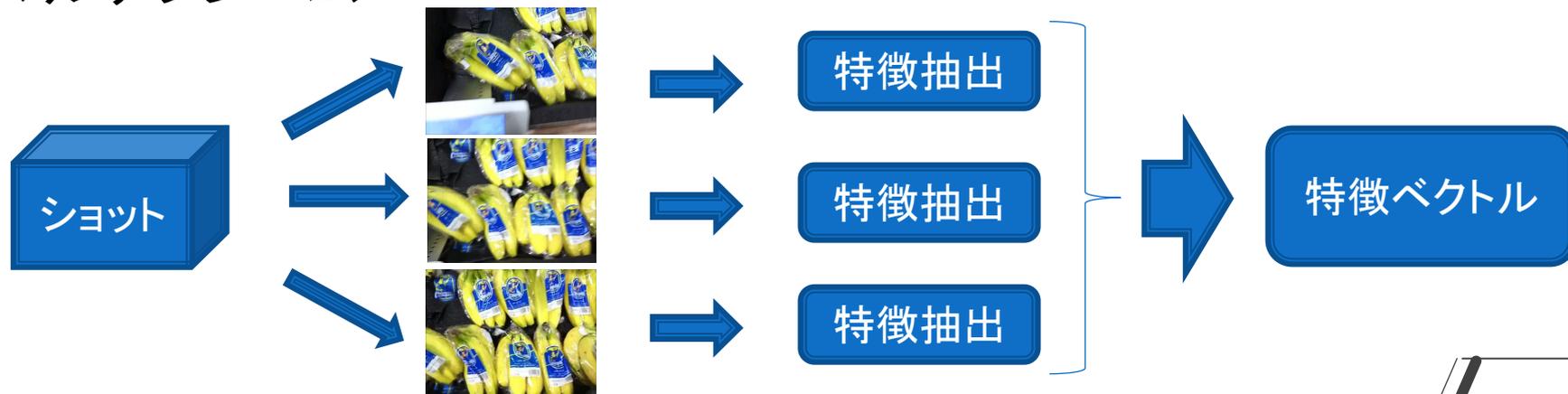
❖ マルチフレーム認識

複数フレームの特徴を統合

○単一フレーム



○マルチフレーム



❖ 高速化の工夫：線形SVM

線形SVMの判別式

$$y(x) = \sum_{i=1}^N w_i x_i x + b$$

$O(DN)$

$$= x \sum_{i=1}^N w_i x_i + b$$

$O(D)$

サポートベクトルとその重みは
学習時に関数値を計算して保存

識別時に高速

カーネル関数

$$K(x_i, x)$$

y : 判別値

x : テストベクトル

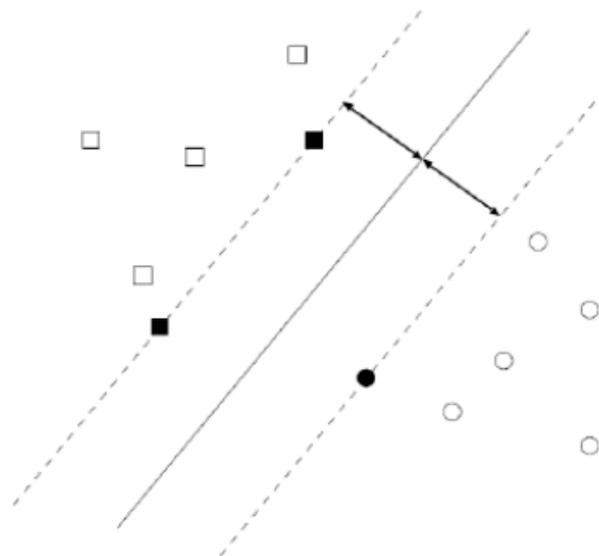
x_i : サポートベクトル

w_i : サポートベクトルの重み

N : サポートベクトルの数

D : サポートベクトルの次元

❖ 高速化の工夫：線形SVM



省メモリ、高速

Androidの1アプリは16MByteまで

モバイルに実装する際には

コードブックサイズ(1000次元)

$$1000 \times 64 \times 4\text{byte} \cong 256\text{Kbyte}$$

サポートベクトル(30食材分)

$$1000 \times 30 \times 4\text{byte} \cong 30\text{Kbyte}$$

❖ 実験

実例

スーパーで撮影した食材

画像認識精度

認識精度(手法ごと、食材ごと、上位k位)

ユーザ評価

食材3種類を対象に目的のレシピを検索、時間

システムの評価、ユーザのコメント

❖ 対象食材

肉類

鶏肉

牛肉

豚肉

ハム

ウインナー

ひき肉

魚類

まぐろ

イカ

タコ

エビ

しゃけ

野菜

しめじ

じゃがいも

ナス

にんじん

レタス

大根

トマト

きゅうり

キャベツ

ネギ

白菜

しいたけ

玉ねぎ

野菜

りんご

いちご

パイナップル

みかん

バナナ

グレープフルーツ

❖ 実例

スーパーで使うと認識されやすい

グレープフルーツ



1位で自動表示



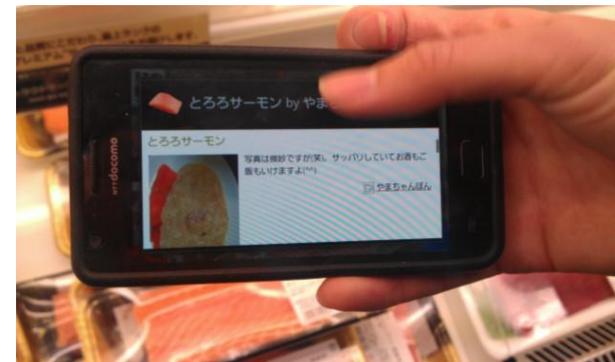
レシピ



鮭

3位の候補を選択

レシピ



❖ 学習データ

30食材が対象

10以上のスーパーで動画撮影

1秒以上、6秒未満の動画を各食材で10本ずつ

動画単位の10fold-crossvalidationで評価

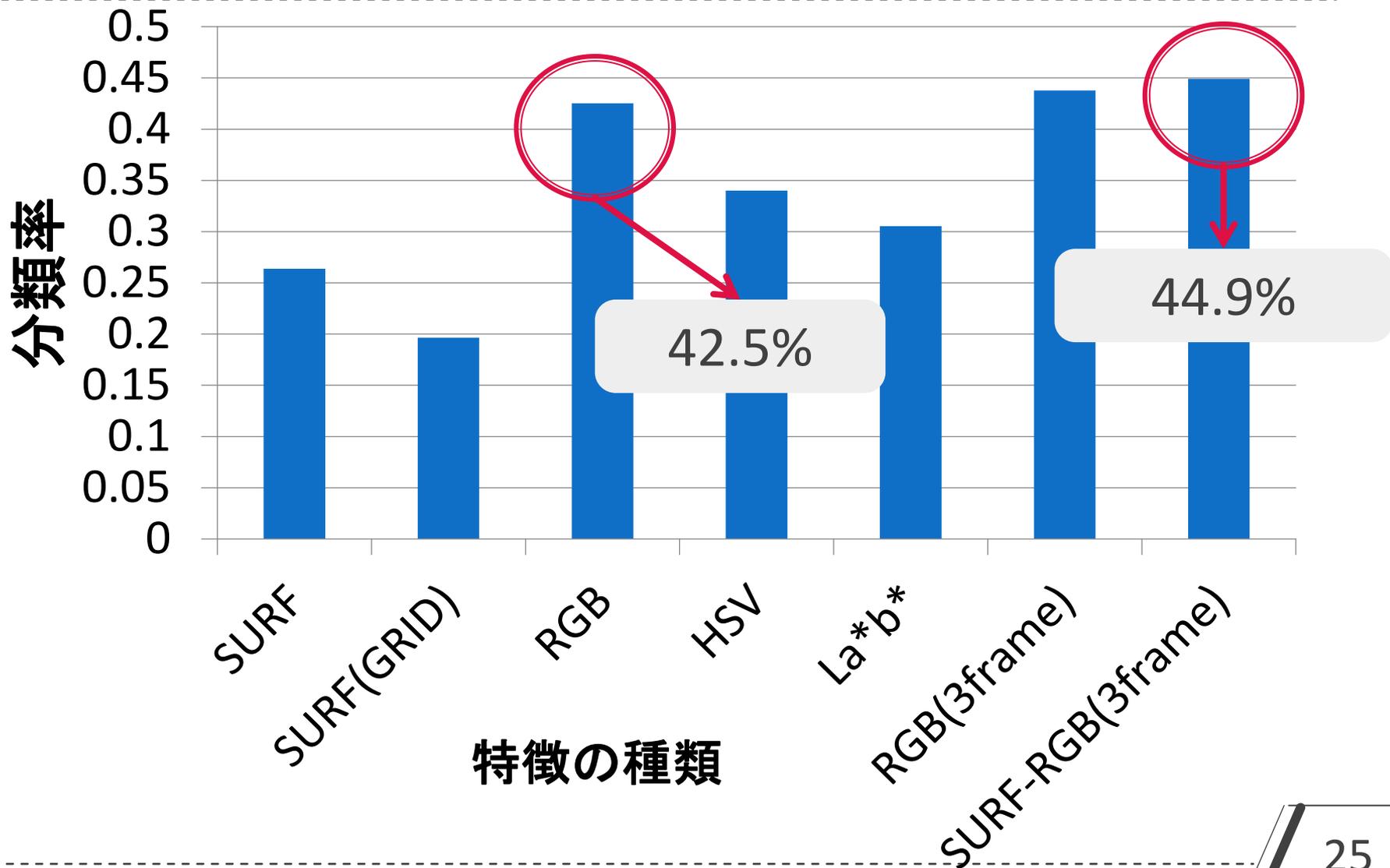
袋詰めされている

光反射が大きい

難しい対象



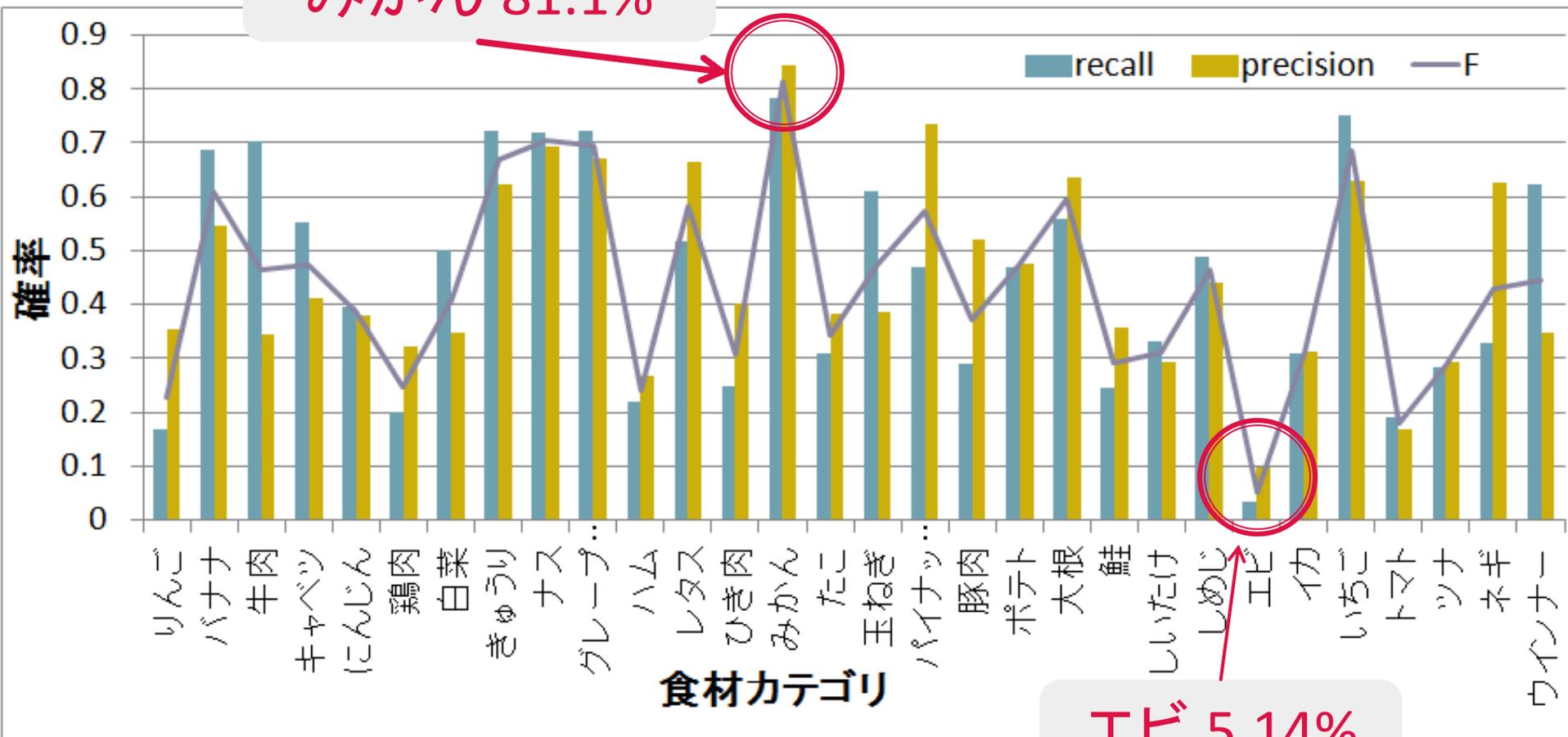
❖ 各手法の認識精度比較



❖ 食材ごとの結果

30食材での結果

みかん 81.1%



エビ 5.14%

❖ 成功した食材、失敗した食材

みかん：色が単純



エビ：色が様々



❖ 間違いやすい食材

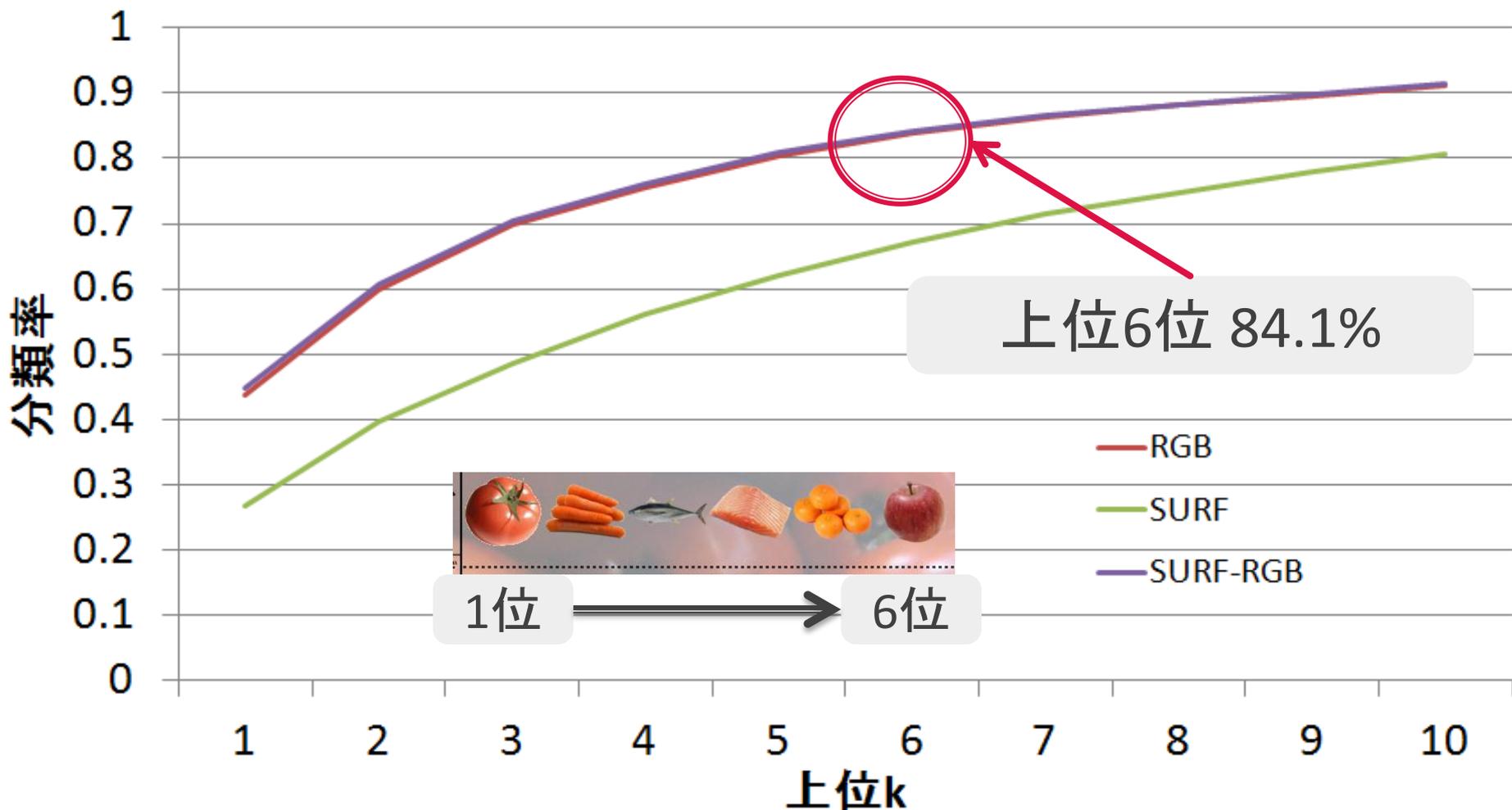
りんご：赤りんごと青りんご



間違えやすい食材：トマト、グレープフルーツ



上位k位まで考慮した結果



❖ ユーザ評価

方法

3食材x3レシピ=9レシピを用意

手動と認識で探す場合の時間計測

認識精度について (5段階)

使いやすさ(5段階)

画像と手動どちらが使いやすいか(5段階)

コメント

被験者数

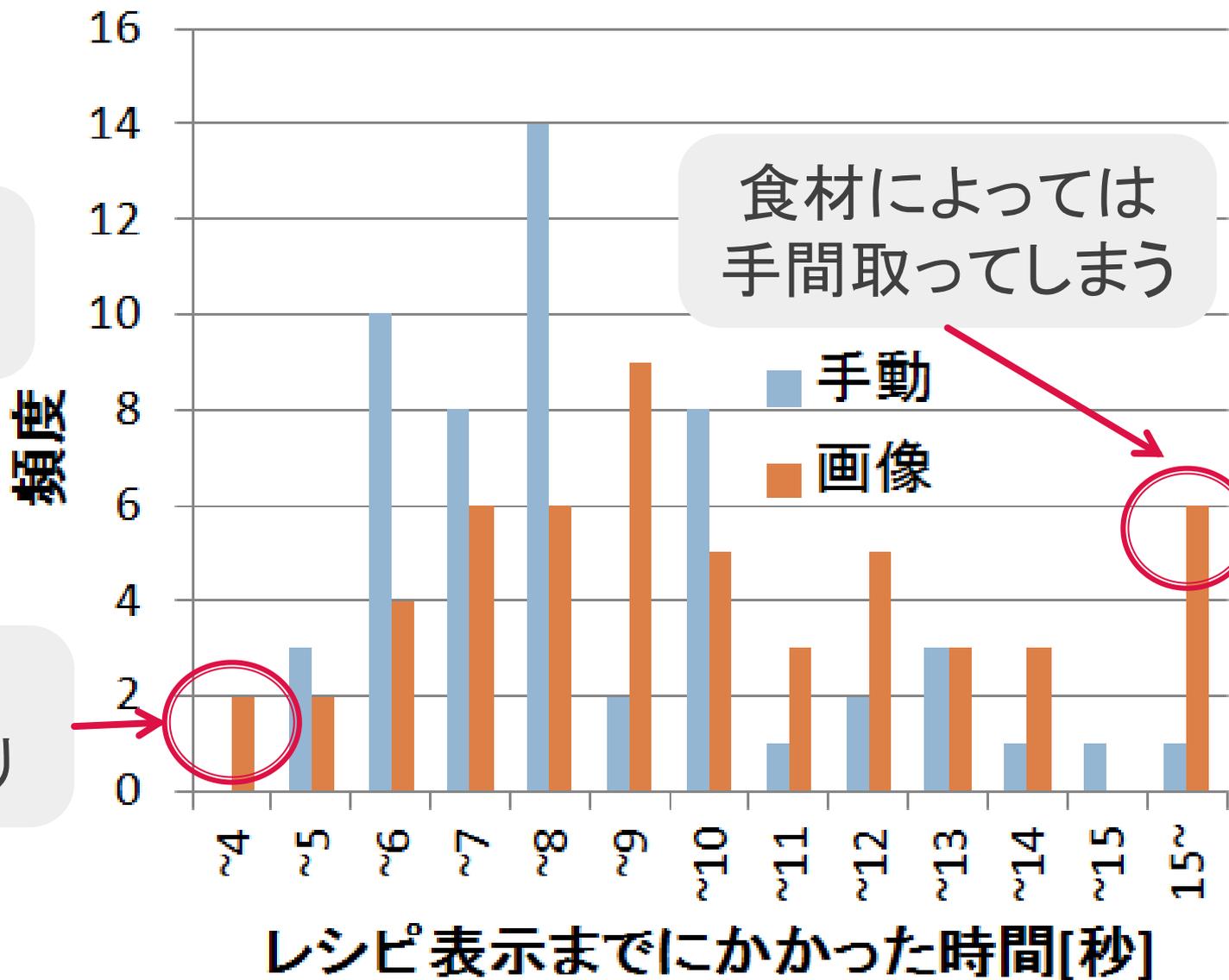
5人

❖ ユーザ評価: レシピ選択までの時間

中央値

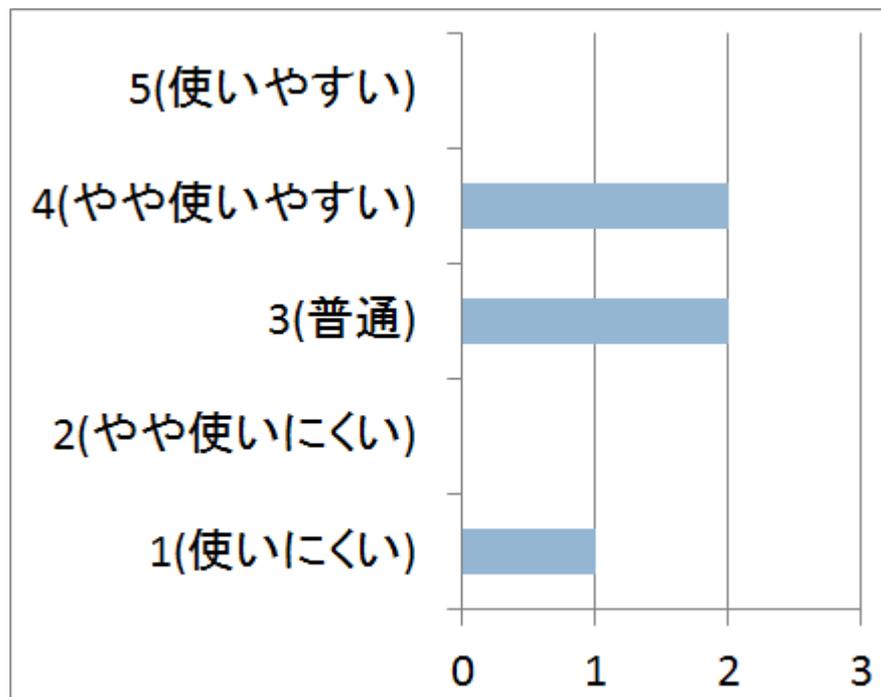
画像: 8.51秒
手動: 7.30秒

手動より
早い場合あり



❖ ユーザ評価

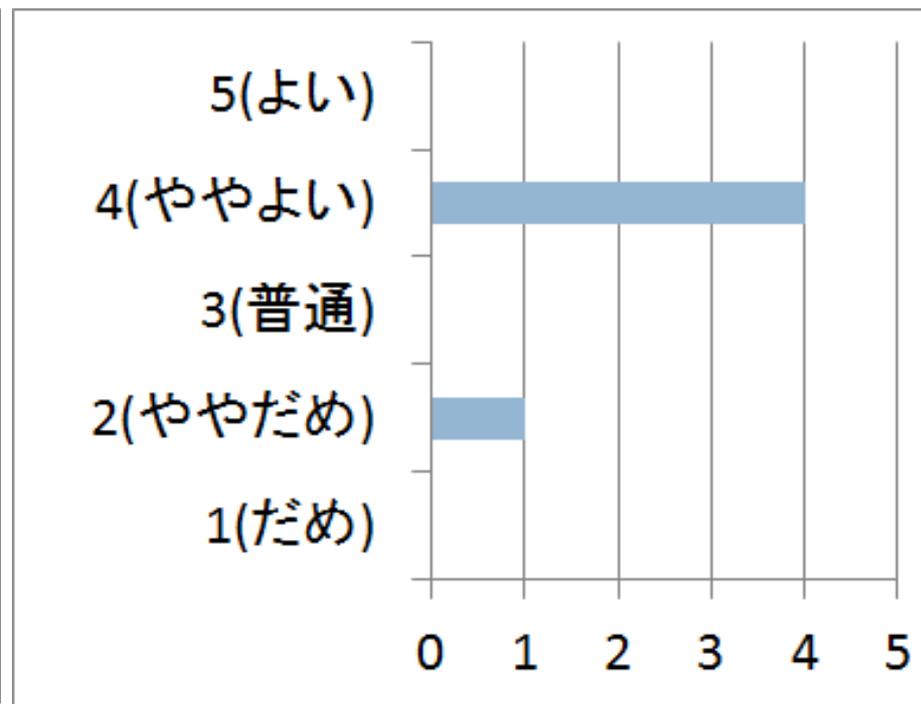
システムの使いやすさ



コメント

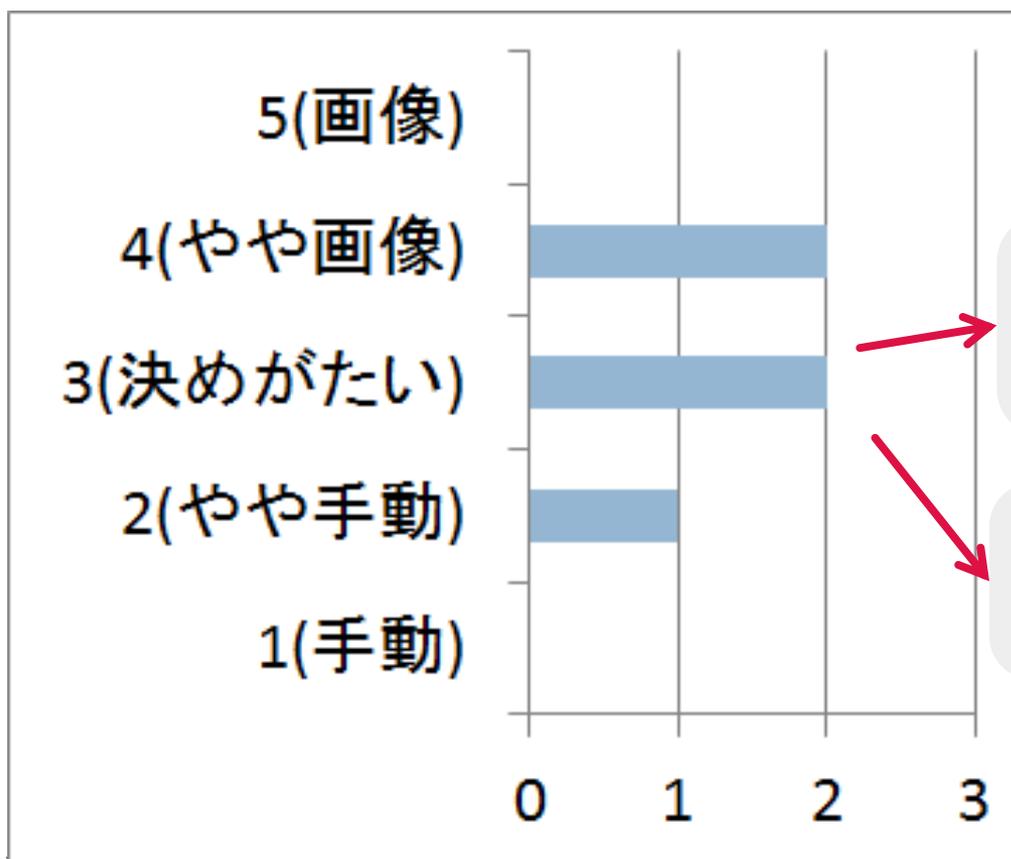
食材候補の更新が激しくて使いづらい

画像認識の精度



❖ ユーザ評価

画像認識と手動、
どちらが使いやすいか？



コメント

「よく認識できる食材は
画像認識のほうが良かった」

「精度が上がれば現状でも
かなりいいのでは」

❖ ユーザ評価

コメント

スーパーでその場で安い食材のレシピが分かるのでいい

精度が上がれば現状でも使いやすいかも

食材によって認識にばらつきがある

価格情報を踏まえてレシピ推薦してくれるとうれしい

❖ まとめ

提案システム

スマートフォンをかざすことで、食材を認識してレシピを提示するシステムを構築した

- ✓ 一般物体認識を0.15[sec]で高速に行う。
- ✓ 30種類の食材が対象

ユーザ実験

必ずしもいい結果ではなかった。
システムの改善が必要。

画像認識

1位 44.9%
6位まで84.1%

❖ 今後の展望

システム

より実用的な機能を実現

複数品食材の条件付き検索

価格情報を加味したレシピ^o推薦

冷蔵庫で事前撮影した食材を組み合わせて検索

画像認識

1位70%, 上位6位90%を目指す

カメラのブレに対応可能な色以外の特徴が必要

❖ ありがとうございました

アプリは公開中です。
是非ダウンロードしてお試してください。

ダウンロード



<http://miru.yni.jp/>

プロジェクトページ

<http://mm.cs.uec.ac.jp/maruyama/recipe/>