

クラウドソーシングによる食事画像 データセットの自動構築

電気通信大学 総合情報学専攻

河野 憲之, 柳井 啓司

背景

- データセットの大規模化
 - 数百種類や数千種類など
 - クラウドソーシングの利用
 - ノイズデータの除去



- クラウドソーシングの問題点
 - コストが少なからずかかる
 - 性能は完璧ではない

➡ データのノイズ除去を事前に自動で行うことによって、コスト削減と精度向上を目指す。

目的

- 既存の画像データセットに新カテゴリを追加
 - 「食事画像データ」を題材とする.
 - Webから画像収集し, ノイズ除去
 - 画像認識による自動ノイズ除去
 - Amazon Mechanical Turk (AMT)を用いて, ノイズ除去.
- 画像に付与する情報
 - クラスラベル
 - 位置情報 (バウンディングボックス)

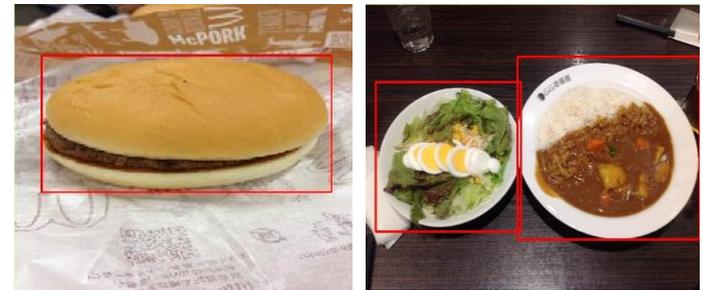


関連研究

- 食事画像データセット

- UEC Food100 [ICME2013]

- 本研究では, これを拡張



- クラウドソーシングによる画像知識ベース構築

- ImageNet (<http://image-net.org>) [CVPR2009]

- Web画像検索 + AMT

IMAGENET

- 画像認識 + active learning + クラウドソーシング

- 物体検出器とAMTを組み合わせ [CVPR2012]

データセット自動構築の流れ

1. Webから食事画像を収集

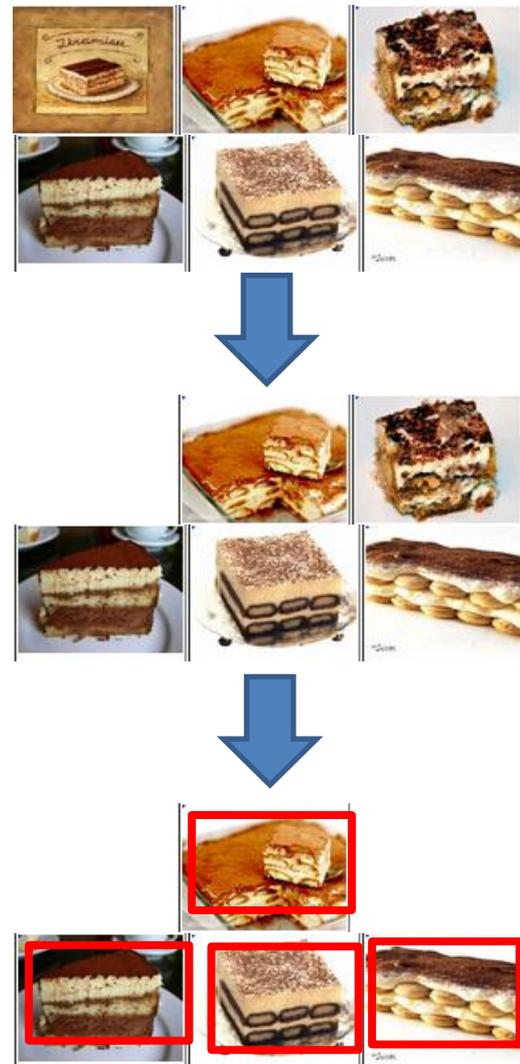
- キーワードによるWeb画像収集

2. 画像認識による自動ノイズ除去

- 食事画像判別と個別食事識別

3. AMTによる手動ノイズ除去

- サンプル選択, 画像選択,
バウンディングボックス描画



1. Webから食事画像を収集

- 新規カテゴリのキーワードを準備
 - 例)「マンゴープリン」「ティラミス」

- 収集   

- WebAPI (Flickr, Twitter, Bing)で
キーワードによる画像検索
- 日本語キーワードと、英訳キーワードの
両方を用いて、画像収集
 - 中華, タイ料理, ベトナム料理などは,
現地語キーワード+英語キーワード

2. 画像認識によるノイズ除去

- 2段階フィルタリングによる自動ノイズ除去
 1. 食事画像判別器
 - 食事・非食事画像分類
 - 既存食事データセットを学習画像
 2. 個別食事識別器
 - 新規カテゴリ・非新規カテゴリ分類
 - 食事画像判別器の評価値の高い画像を学習画像

2.1. 食事画像判別器

- 目的
 - 食事画像 と 非食事画像を判別
- 画像認識器による判別
 - 画像特徴量+ liner SVM
 - 正例: UEC-FOODの100カテゴリを13カテゴリに集約,
 - 負例: 食事名で検索した時に間違えやすい画像.
 - レストラン、出店、飲食店、夕食会
 - 13の識別器を学習し, 最大スコアを「食事度(foodness)」とする.

2.1 作成されたグループと属する食事例

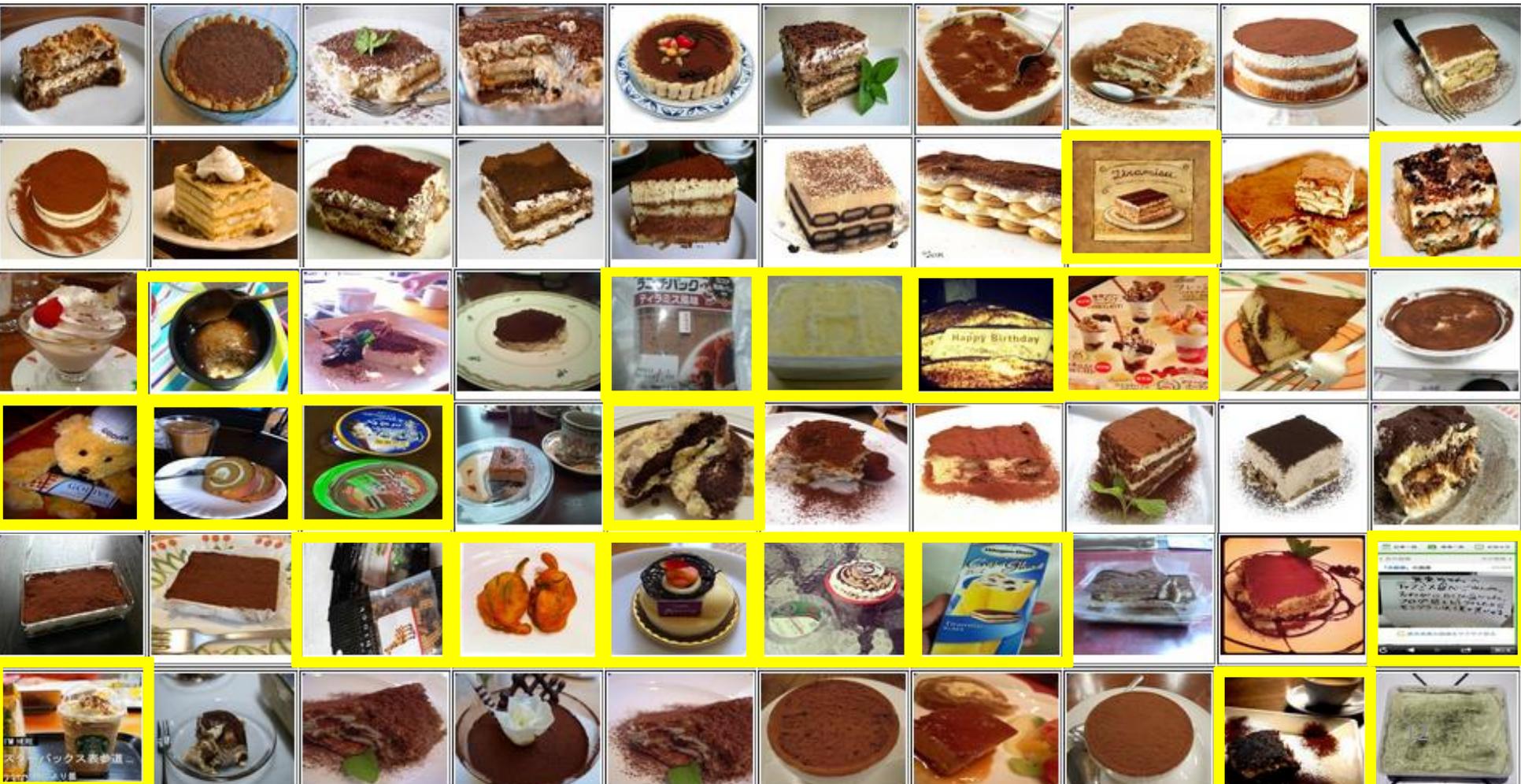
noodles	udon noodles, dipping noodles, ramen
yellow color	omlet, potage, steamed egg hotchpotch
soup	miso soup, pork miso soup, japanese tofu and vegetable chowder
fried	takoyaki, japanese-style pancake, fried noodle
deep fried	croquette, sirloin cutlet, fried chicken
salad	green salad, sauteed vegetables, vegetable tempura
bread	sandwiches, raisin bread, roll bread
seafood	sashimi, sashimi bowl, sushi
rice	rice, pilaf, fried rice
fish	grilled salmon, grilled pacific saury, dried fish
boiled and seasoned	seasoned beef with potatoes simmered ganmodoki seasoned beef with potatoes
sauteed	sauteed vegetables, goya chanpuru, kinpira-style sauteed burdock
sauce	stew, curry, stir-fried shrimp in chili sauce

2.2. 個別食事識別器

- 目的
 - 目的のカテゴリの食事画像かどうか識別
 - [問題点]新規カテゴリのため学習データが存在しない.
 - 転移学習を利用して, 既存カテゴリの知識を利用
- 画像認識器による判別 (A-SVM以外は2.1と同じ)
 - 画像特徴量 + **adaptive SVM**
 - [擬似正例] 食事画像判別器の評価値上位画像
 - [負例] 食事画像判別器の評価値下位画像
 - 最も類似している食事グループ画像をソースドメインデータとする.

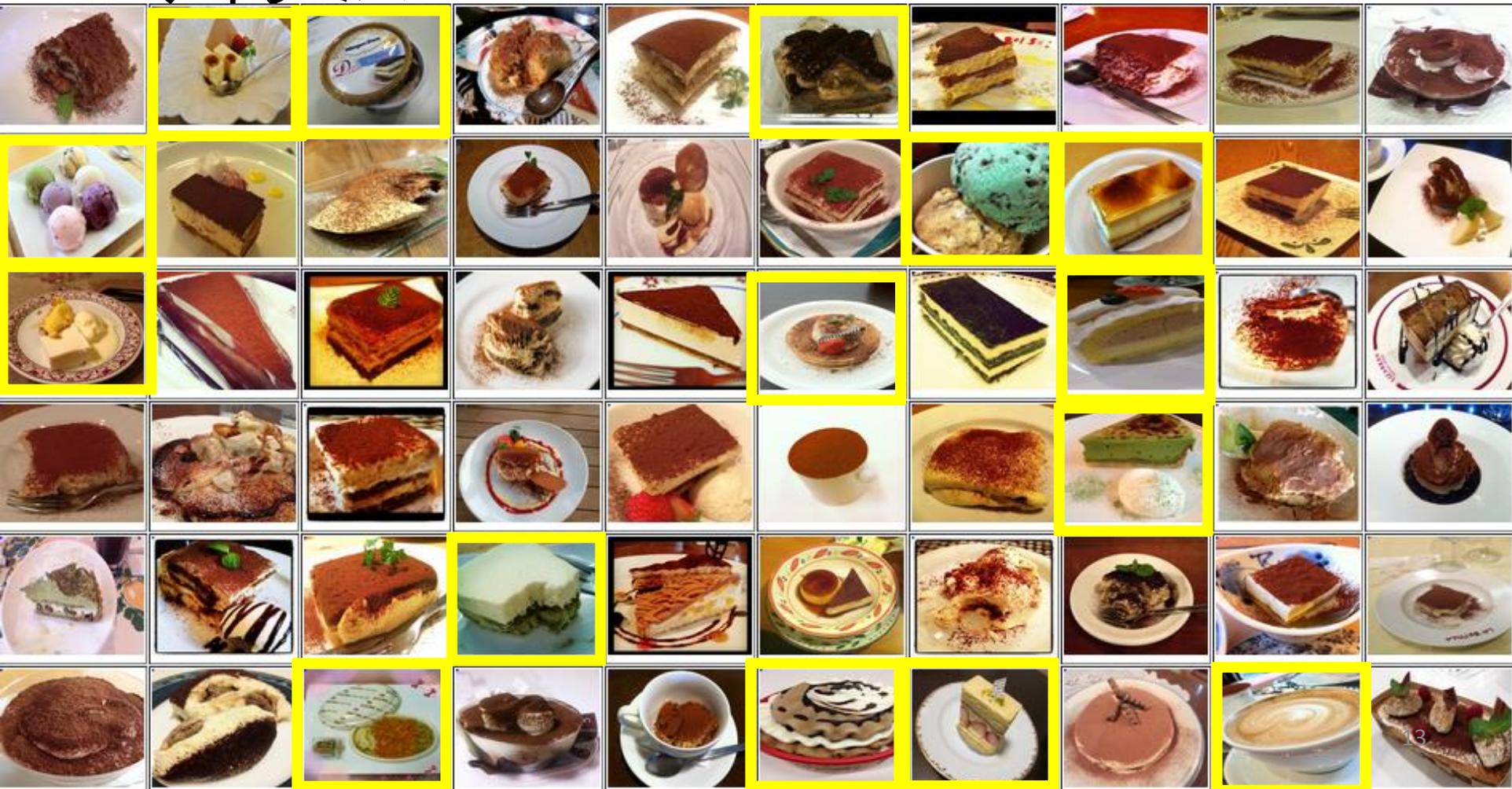
Webから収集

- ティラミス



食事画像判別適用後

- ティラミス



個別識別器適用後

- ティラミス



転移学習で、ソースドメインとなった画像例

ターゲット

最も類似した食事グループ



沖縄そば
プリン
ローネ
ナチヨス
マンゴー
ミネスト

3. AMTによる手動ノイズ除去

- 人間の力を借りてノイズ除去
最後に, バウンディングボックス(BB)付与
- 3段階のクラウドソーシング処理
 - サンプル画像選択タスク
 - ノイズ除去タスク
 - BB付与タスク



3.1. クラウドソーシングでサンプル画像を取得

- 問題点

- ワーカーは対象の料理について知らない
- ノイズを除去しても、少なからず含まれる



- 目的

- 後のタスク(ノイズ除去, BB付与)でワーカーに提示するための一般的な少数で高信頼のサンプル画像の取得



- AMTで処理する画像

- 評価値上位を画像類似度で並び替えして、使用

3.1. サンプル選択タスク

- タスクの特徴

- 30枚を与えて、最大7枚、
一般的なサンプルを選択してもらう

- リンクを与えるので勉強してもらう

- \$0.06、4人

- 少し高め、外部リンクで勉強させる、また重要であるため

- 質の向上

- サンプル選択をしないとヒットを提出できない

- 設置したリンクをたどらないとヒットを提出できない

3.1. タスク例

チゲのサンプル画像選択タスク



3.1. 実際のページ例

Select normal food: **Vietnamese Popiah** in images.

Description

Overview

select normal Vietnamese Popiah in images. (max: 10 images)

You can select images by click on it.

The selected image is displayed as a red circle.

And if you won't this selected state you can remove it by clicking on this image again.

The red circle will disappear.

Note

1. About image selecting

You must select Vietnamese Popiah images. Do not select none Vietnamese Popiah images.

(There's a possibility that may be show none Vietnamese Popiah images.)

You should select normal Vietnamese Popiah not abnormal one, for example the food is used many topping.

Next, you select photograph well images as possible as you can.

Finally, **you should visit some image search site. if you do not visit, you can not submit this hit.**

(I provide you with link to site where is center in this page.)

2. About image loading

You have 30 images in this HIT.

If the images are not loading, please click "Click here to Reload" tab and wait a moment.

But if reloading still doesn't solve the problem, please skip this HIT.

[Click Here to Reload](#)

3. About "cooking-in-progress"

"cooking-in-progress" image means "no Vietnamese Popiah".

Please select only on the image which is "ready-to-eat" only.

"Toy foods" also should not be marked.

Examples

Finally, I will show the good examples.

GOOD example: tiramisu



GOOD example: pancake



What is Vietnamese Popiah?

About Vietnamese Popiah information.

Please you visit on web site [Google Image Search](#), [bing](#) or [Wikipedia](#) for more detail.

Please start this HIT now.

TARGET IS Vietnamese Popiah



Please tell us your nationality (country name)

Fill this form everytime when you submit.

Thank you for your contribution.

Please provide any comment or suggestion your opinions to make this HIT more friendly to workers.

We appreciate your feedback!

(this form is optional)

Your state:

You have not visit site yet.

You have not select image yet.

Have you read "Note" section above? **YOU WILL BE REJECTED** if you are regarded you don't follow the rules in "Note" section. Before submitting, make sure that you worked correctly.

3.1. サンプル選択結果

- 5~7枚取得し、対象の食事である
 - 100カテゴリ100%
- ワーカーからの評価
 - 他のタスクで有用性を調査

	有用	普通	有用でない
ノイズ除去	89.59	7.90	2.52
BB付与	91.68	7.02	1.31

サンプル有無による性能評価は未調査

3.1. 收集結果例



3.2. クラウドソーシングでノイズ画像の除去

- タスクの目的

- 人手による対象でない食事画像の除去



- タスクの特徴

- 1タスク \$0.03, 5人
- 結果の結合は多数決
- 未判定が5つ以上あると、提出できない



3.2. タスク例

取得したサンプル画像



ノイズ除去タスク



その他、タスクの説明や、よい例、悪い例などを提示

3.2. 実際のページ例

Mark **panggang** and mark non panggang in images.

Description

Overview

You mark panggang images by left-clicking and non panggang images by right-clicking. A target food image you mark as panggang is displayed as a red circle, and a non target image is displayed as a red cross. Now, target (panggang) is shown below.



You can remark by left-clicking or right-clicking on a image again. Previous one will be ignored or disappeared.

If there are **no panggang** in a image, the target is **invisible more than half of it and blurred**, the target is **not a real food item** such as illustration, animation and package or the target is **not "ready-to-eat"**,

check the image as "no panggang" by **right clicking** on it. In images checked as "**no panggang**", a red cross is drawn as shown below.



You can redo everything except submitting in this HIT. Don't be so nervous.

Note

- When you are not sure
 - When you are not sure if you should mark as a panggang image or "no panggang" image, please you mark as "no panggang".
- About image loading
 - You have 25 images in this HIT.
 - When there are images that are not loaded, please reload this page after a while.
 - If reloading doesn't solve the problem, please skip this HIT.
 - [Click Here to Reload](#)
- About "cooking-in-progress"
 - "cooking-in-progress" image means "no panggang".
 - Please you mark as panggang images only on a image which is "ready-to-eat".
 - "Toy foods" also should not be marked.

Examples

Finally, I will show the good examples.

GOOD example1: paella



GOOD example2: waffle



What is panggang?

About panggang information.

Please you visit on web site [Google Image Search](#), [bing](#) or [Wikipedia](#) for more detail.

I show the target (panggang) images again.



Please start this HIT now.

TARGET IS panggang



Please tell us your nationality (country name)

Fill this form everytime when you submit.

Thank you for your contribution.

sample images of panggang are useful or useless for this task ? useful so so useless

: nationality

Please provide your opinions to make this HIT more friendly to workers.

We appreciate your feedback!

(this form is optional)

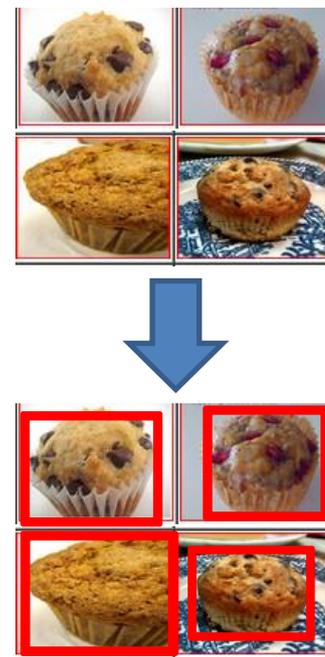
Your state:

You have not visit site yet.
 mark as panggang images: 0 image.
 make as non panggang images: 0 image.
 unmark: 25 images.

Have you read "Note" section above? YOU WILL BE REJECTED if you are regarded you don't follow the rules in "Note" section. Before submiting, make sure that you worked correctly.

3.3. クラウドソーシングでBB付き食事画像を取得

- タスクの目的
 - バウンディングボックスの付与
 - 人手による対象でない食事画像の除去
- タスクの特徴
 - 1タスク \$0.05、4人
 - 質の向上
 - 現在の状況の提示
 - 入力時にノイズのチェック



3.3. タスク例

取得したサンプル画像



バウンディングボックス付与タスク



3.3. 実際のページ例

Draw bounding box around **rice and tea** in image.

Description

Overview

Draw Bounding-Box around rice and tea in images.
You can draw Bounding-Box by Drag-and-Drop.
A Bounding-Box you draw is displayed as a red rectangle.
If target is curry, an example is shown below.



Now, target (rice and tea) is shown below(automatically selected).



You can redraw Bounding-Box just by left-clicking on a image again.
Previous one will disappear.

If there are **no rice and tea** in a image,
the target is **invisible more than half of it and blurred**,
the target is **not a real food item** such as illustration, animation and package or
the target is **not "ready-to-eat"**,
check the image as "no rice and tea" by **right clicking** on it.
In images checked as "no rice and tea", a red cross is drawn as shown below.



You can uncheck a image just by right-clicking on a image again.
You can redo everything except submitting in this HIT.
Don't be so nervous.

Note

1. Images

We have already filtered no rice and tea images by mturk, but images maybe include some no rice and tea images.

2. When you are not sure

When you are not sure if you should draw Bounding-Box or check as "no rice and tea", please check as "no rice and tea".

3. About image loading

You have 10 images in this HIT.
When there are images that are not loaded, please reload this page after a while.
If reloading doesn't solve the problem, please skip this HIT.

[Click Here to Reload](#)

4. About "cooking-in-progress"

"cooking-in-progress" image means "no rice and tea".
Please draw Bounding-Box only on a image which is "ready-to-eat".
"Toy foods" also should not be marked.

5. About Dish

If rice and tea is on a dish in a image, please draw Bounding-Box **not including the dish**.
So, when there are multiple dishes which have rice and tea on, you must choose whichever to draw Bounding-Box around too.
Remember, **"One Bounding-Box, one food"**.
Don't let a Bounding-Box include a dish.

Examples

Last, I show the good example and bad example for curry.

GOOD example



BAD example



What is rice and tea?

About rice and tea information.

Please you visit on web site [Google Image Search](#), [bing](#) or [Wikipedia](#) for more detail.

I show the target (rice and tea) images again.



Now please start this HIT

TARGET IS NOW **rice and tea**



Please answer one question and tell us your nationality (country name)

Fill this form everytime you submit.

Thank you for your contribution.

sample images of rice and tea are useful or useless for this task ? useful so so useless

: nationality

Please provide your opinions to make this HIT more friendly to workers.

We appreciate your feedback!

(this form is optional)

: nationality

Please provide your opinions to make this HIT more friendly to workers.

We appreciate your feedback!

(this form is optional)

Your state:

You have not visit site yet.

draw bounding box: 0 image.

checked noObject: 0 image.

unchecked: 10 images.

Have you read "Note" section above?

YOU WILL BE REJECTED if you are regarded you don't follow the rules in "Note" section.

Before submitting, make sure that you worked correctly.

3.3. 収集結果例

アップルパイ



実験設定

- 目的

- 100種類各100枚以上のラベル付きデータの取得
- クラウドソーシングに用いる画像と性能の評価
- タスク分散による性能の評価

- 選別方法

- 食事画像判別器: FC
- 個別食事カテゴリ識別器: A-SVM
- AMTによるノイズ除去タスク: AMT

実験内容

- 実験1
 - コスト (100枚をデータセットに追加する賃金)
 - 回収率 (投げた画像の内、パスした割合)
- 実験2
 - 構築されたデータセットの適合率

実験結果

コスト

	ノイズ除去		BB付与		総量
	回収率[%]	コスト[\$]	回収率[%]	コスト[\$]	コスト[\$]
FC	-	-	64.2	3.11	3.11
FC + A-SVM	-	-	74.7	2.68	2.68
FC + A-SVM + AMT	80.9	0.74	86.7	2.31	3.16

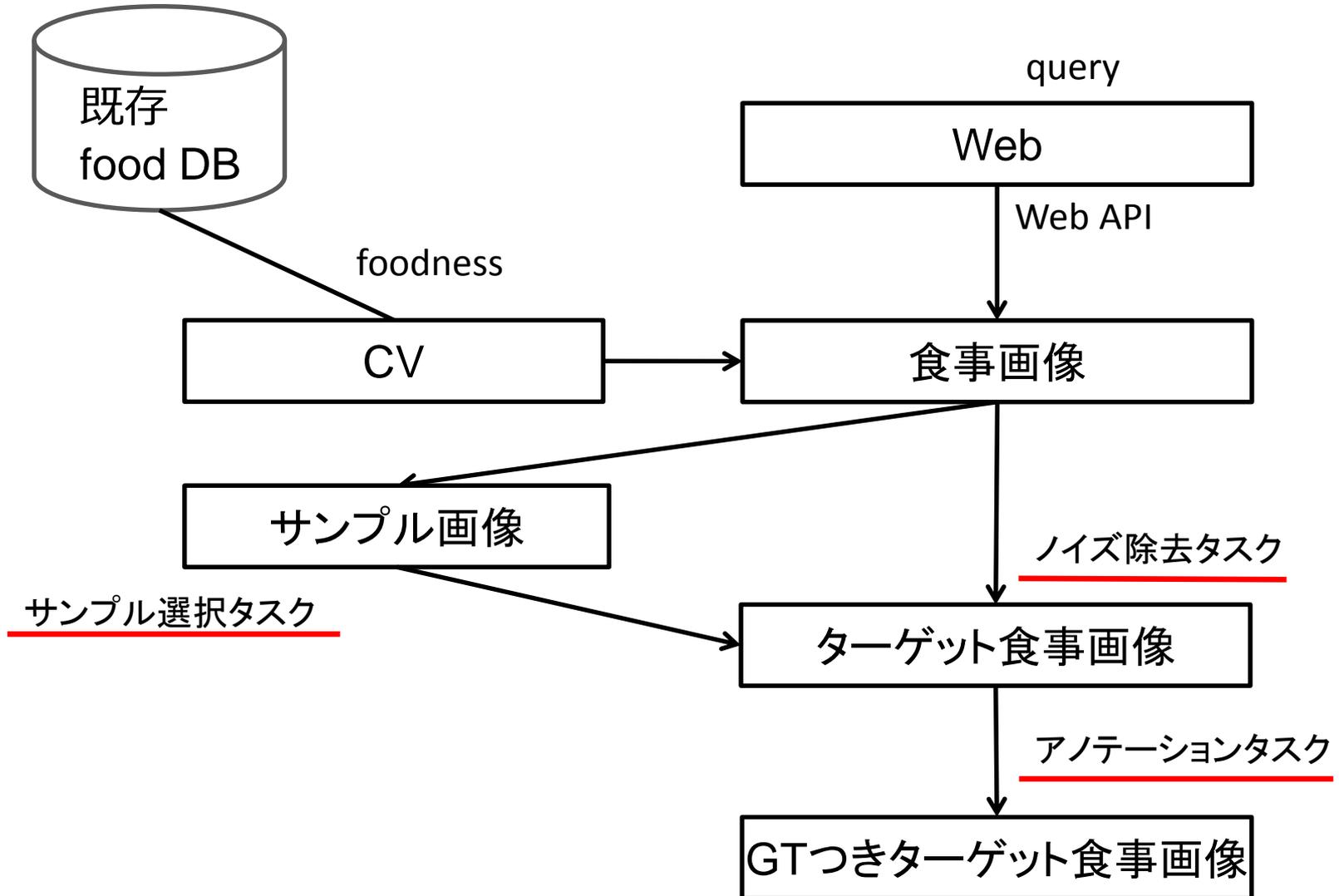
適合率

	適合率[%]	上昇[%]
FC	91.10	-
FC + A-SVM	94.19	+ 3.09
FC + A-SVM + AMT	97.83	+ 3.64

まとめと課題

- クラウドソーシングを用い、食事画像データセットの自動拡張に取り組んだ
- データのノイズ除去を事前に行うことによって、コスト削減と性能が向上した
- 今後の課題
 - 食事画像以外のデータセットの構築
 - 個別のワーカーの信頼性評価の導入

システムの流れ

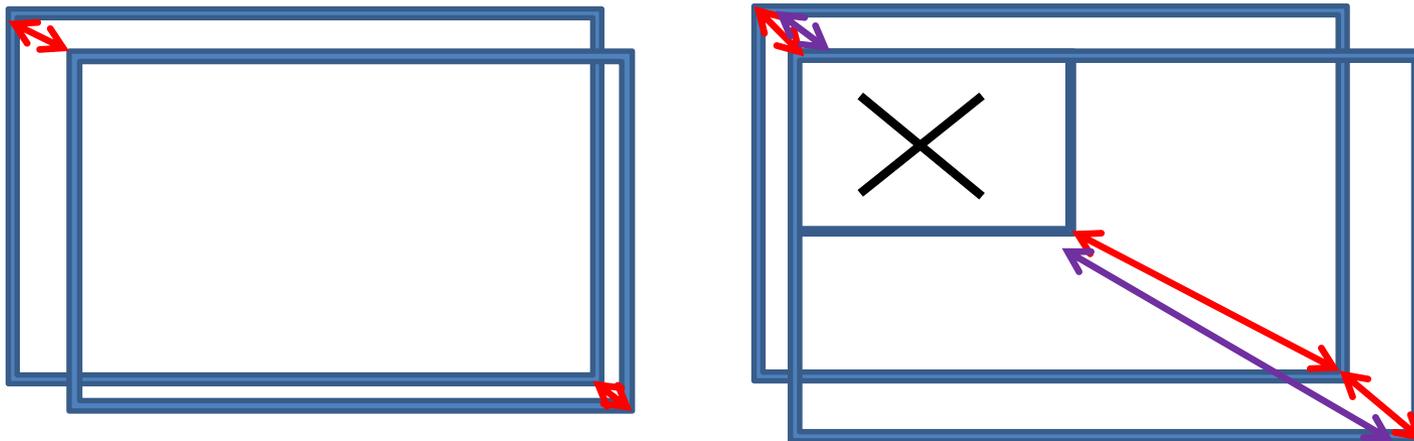


2.2. 未知カテゴリごとの識別器

- Adaptive SVMの学習データ
 - ソースドメイン
 - 食事グループ
 - ターゲットドメイン
 - 未知カテゴリ

結果の結合

- 複数のワーカーによる結果の結合
 - 複数のバウンディングボックスが付近x%内に存在 (始点、終点)
 - そのバウンディングボックスの平均
 - データベースに加える



ワーカーの国別人口

- 自由回答によるワーカーの国別人口
– インド人が圧倒的に多い

india	usa	Canada	French	uk	Asia
1668	394	20	3	3	4

- 今回対象にした食事は、日本を含む5か国
– インド人は未知の食事も多い

悪意のあるワーカー

- 特定の一人の悪意のあるワーカーの結果



3.3. 收集結果例

杏仁豆腐

