食事画像領域分割 データセットの作成とその活用

岡本　開夢　Cho Jaehyeong　會下拓実　柳井啓司

電気通信大学大学院　情報理工学研究科　情報学専攻

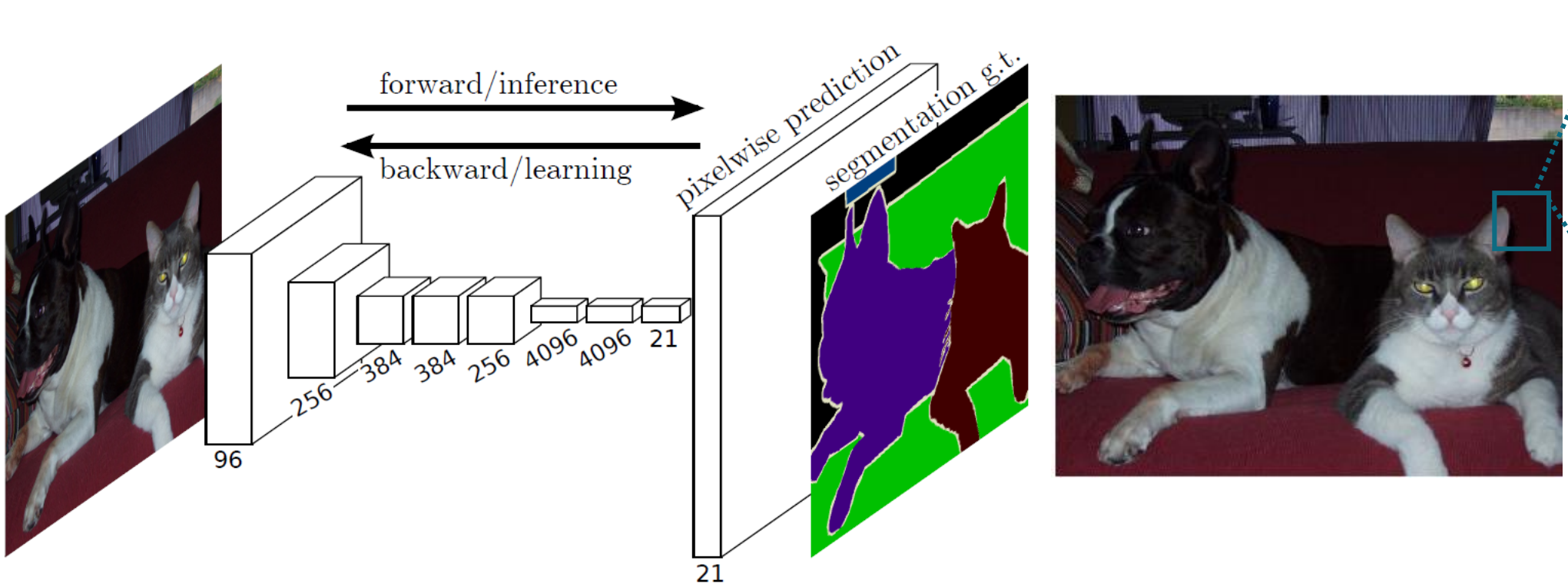
本スライドの流れ

1. 背景
2. 本研究の目的
3. データセットの作成
4. 実験
5. まとめ

1.背景：CNNによる領域分割

深層学習の発達によりピクセルレベルにおけるカテゴリ推定が可能

⇒画像内における各カテゴリごとの領域分割が可能

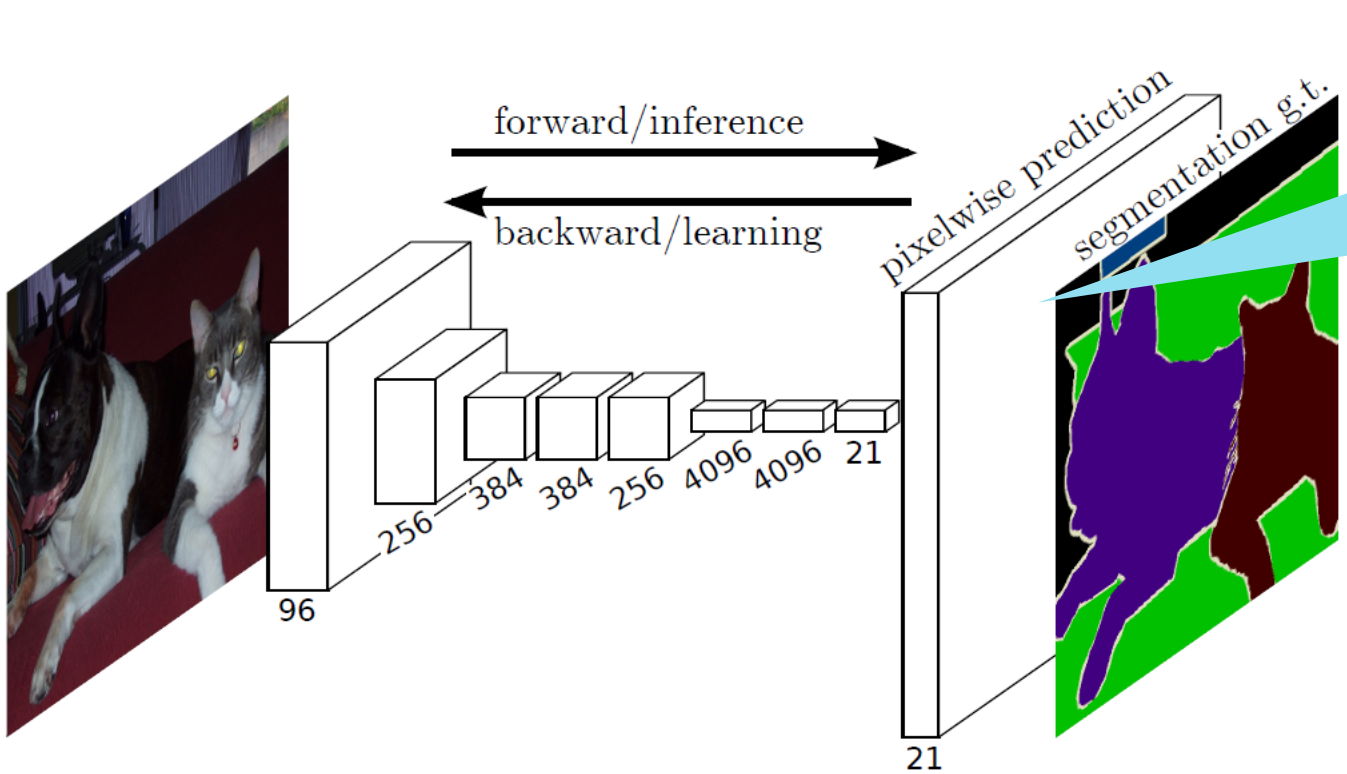


Long, Jonathan, Evan Shelhamer, and Trevor Darrell. "Fully convolutional networks for semantic segmentation." In Proc. of the IEEE conference on computer vision and pattern recognition. 2015.

1.背景：CNNによる領域分割

深層学習の発達によりピクセルレベルにおけるカテゴリ推定が可能

⇒画像内における各カテゴリごとの領域分割が可能



学習する際に
実画像とマスク画像が必要

Long, Jonathan, Evan Shelhamer, and Trevor Darrell. "Fully convolutional networks for semantic segmentation." In Proc. of the IEEE conference on computer vision and pattern recognition. 2015.

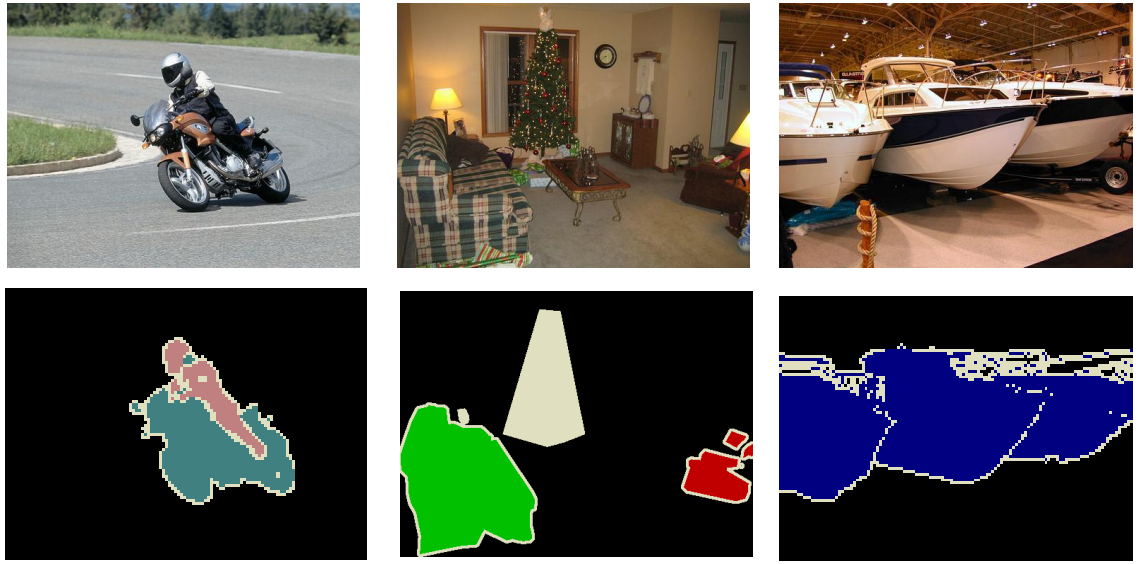
1.背景：CNNによる領域分割

領域分割に用いられる主な学習用のデータセット[1][2]

22カテゴリ
約1万枚

Microsoft
Common Objects in Context

80カテゴリ
約33万枚



1.背景：CNNによる領域分割

領域分割に用いられる主な学習用のデータセット[1][2]

22カテゴリ
約1万枚

Microsoft
Common Objects in Context

80カテゴリ
約33万枚

そのうちの食事カテゴリは
わずか10カテゴリ…
(バナナ、サンドウィッチ、ピザなど)

1.背景：CNNによる領域分割

領域分割に用いられる主な学習用のデータセット[1][2]

22カテゴリ
約1万枚

80カテゴリ
約33万枚

そのうちの食事カテゴリは
わずか10カテゴリ…
(バナナ、サンドウィッチ、ピザなど)

食事領域分割モデルの
学習には不十分

1.背景：CNNによる領域分割

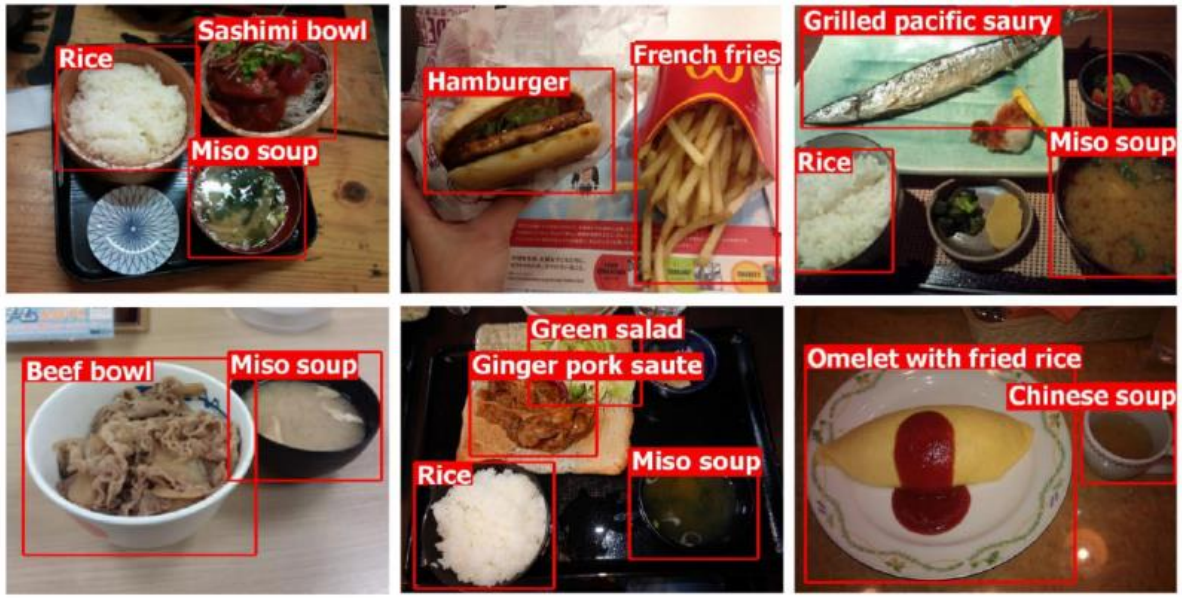
食事画像データセット[10][3]は多数存在

Food101[10]

101カテゴリ
約10万枚

UECFood-100[3]

100カテゴリ
約1.3万枚



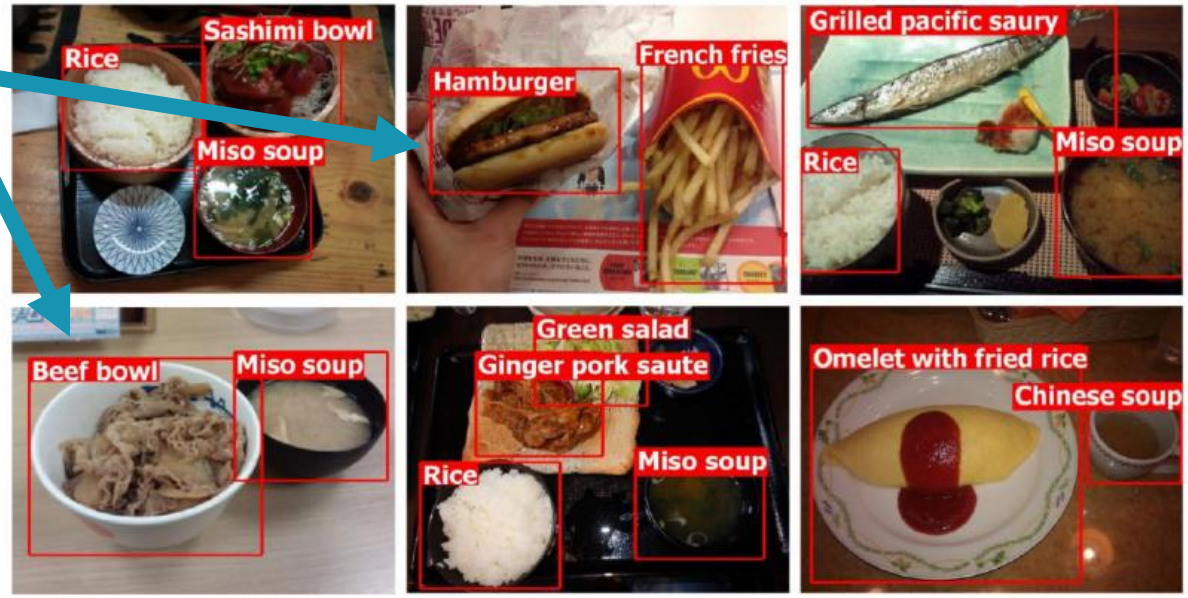
1.背景：CNNによる領域分割

食事画像データセット[10][3]は多数存在

そのうち<u>バウンディングボックス</u>がついているのは少数

UECFood-100[3]

100カテゴリ
約1.3万枚



1.背景：CNNによる領域分割

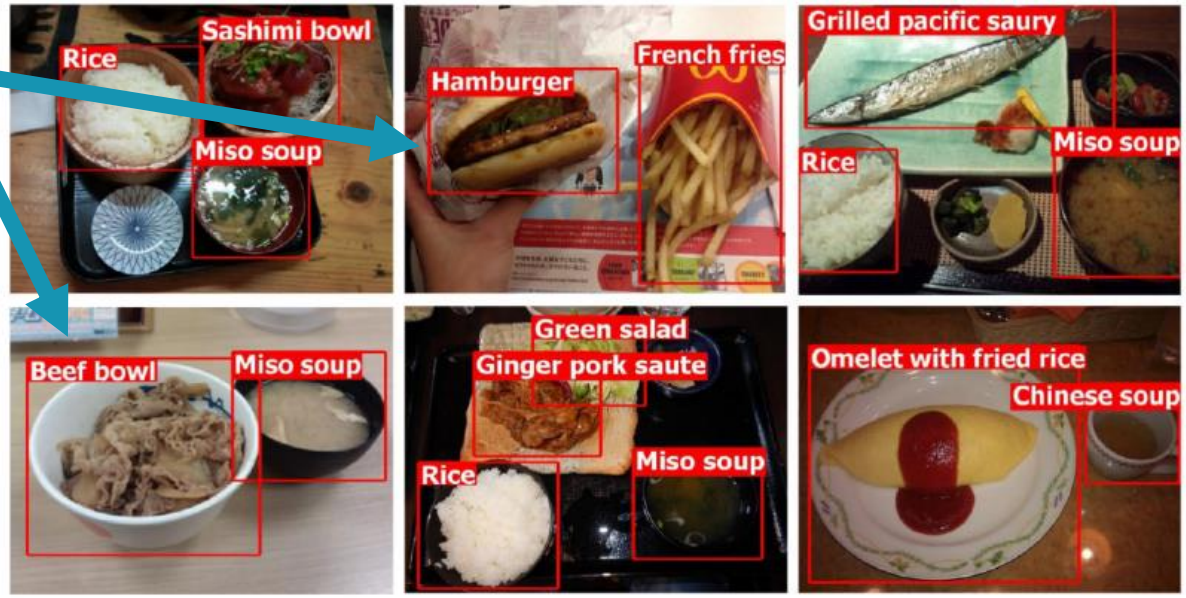
食事画像データセット[10][3]は多数存在

そのうち**バウンディングボックス**
がついているのは少数[3][11]

食事領域分割に使うことができる
ような大規模データセットは
存在していない

UECFood-100[3]

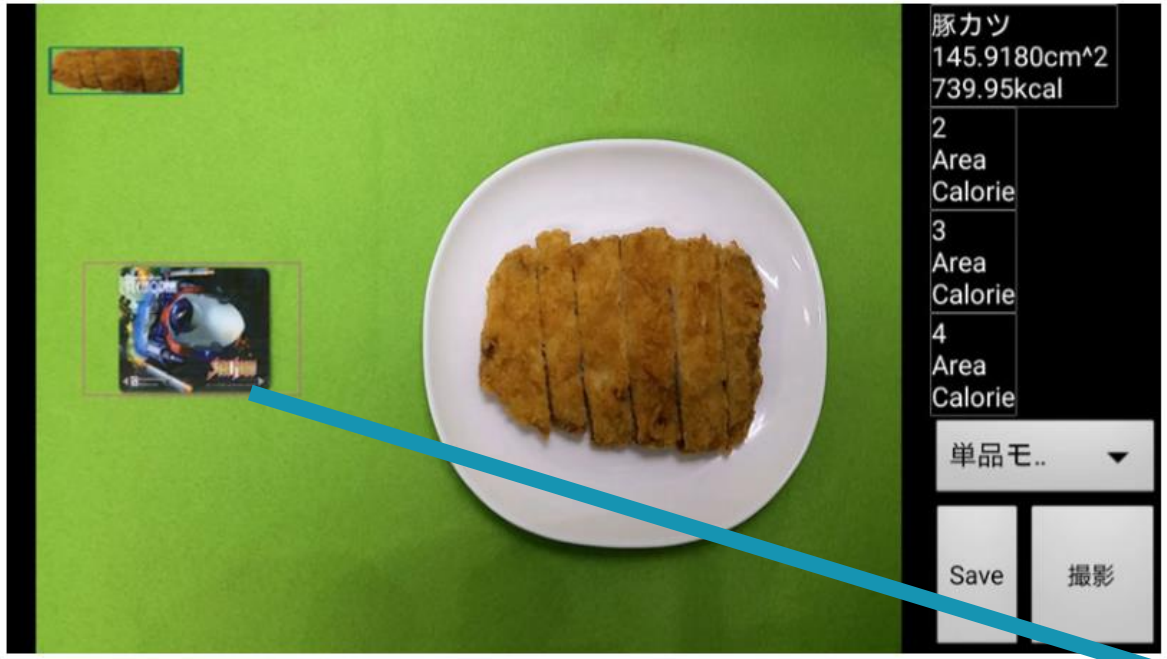
100カテゴリ
約1.3万枚



1.背景：食事画像を用いた研究

岡元らのカロリー量推定アプリケーション[6]

⇒基準物体と食事領域の抽出を行い回帰式をもとにカロリー量推定



Grab-cutによる食事領域抽出[5]

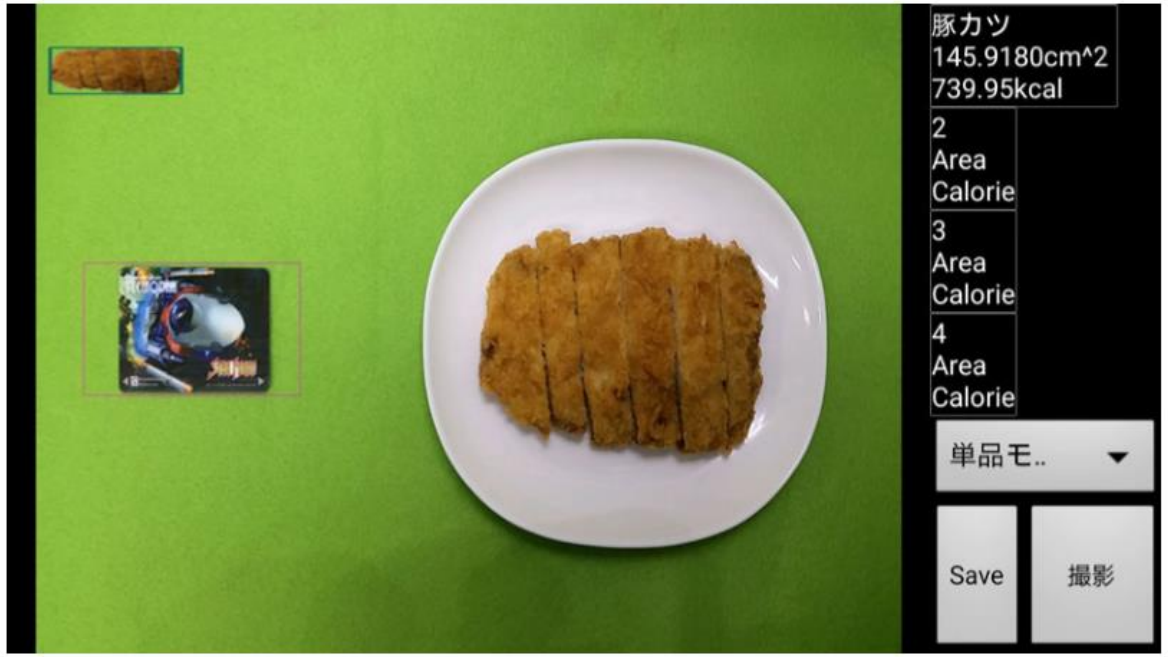
基準物体と食事領域のピクセルの対比関係により実面積を求める。

K. Okamoto and K. Yanai. An Automatic Calorie Estimation System of Food Images on a Smartphone, MADiMa, 2016

1.背景：食事画像を用いた研究

岡元らのカロリー量推定アプリケーション[6]

⇒基準面積と食事領域の抽出を行い回帰式をもとにカロリー量推定



K. Okamoto and K. Yanai. An Automatic Calorie Estimation System of Food Images on a Smartphone, MADiMa, 2016

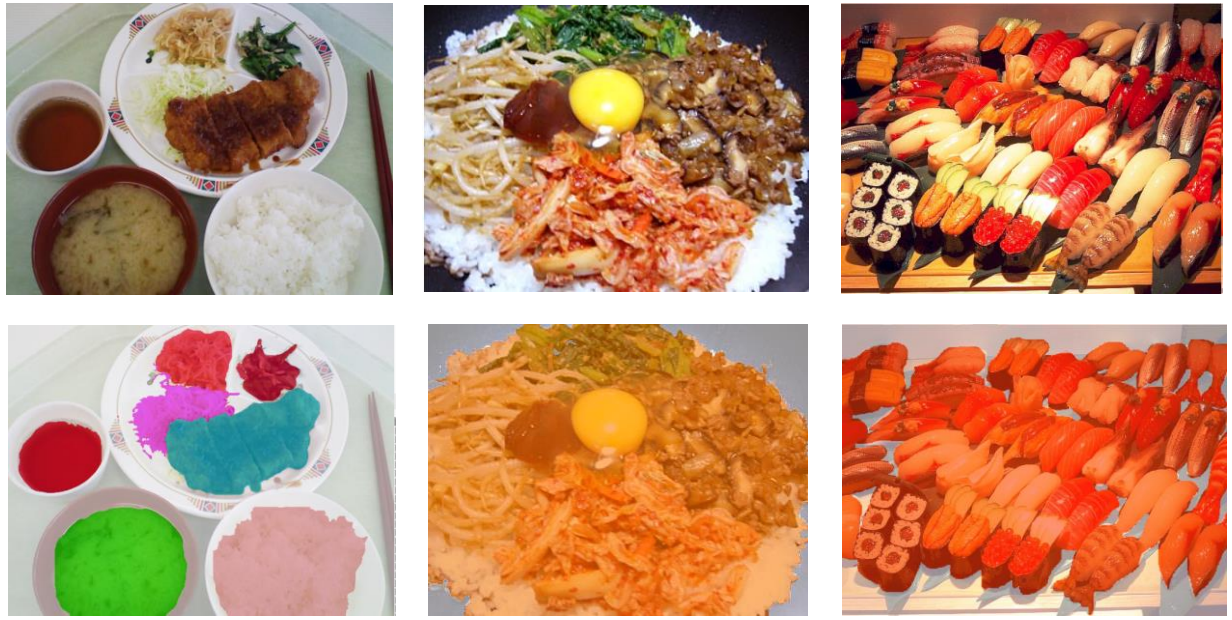
Grab-cutによる食事領域抽出[5]

カロリー量推定には
食事領域分割は不可欠

食事領域分割用の学習用データセットの作成

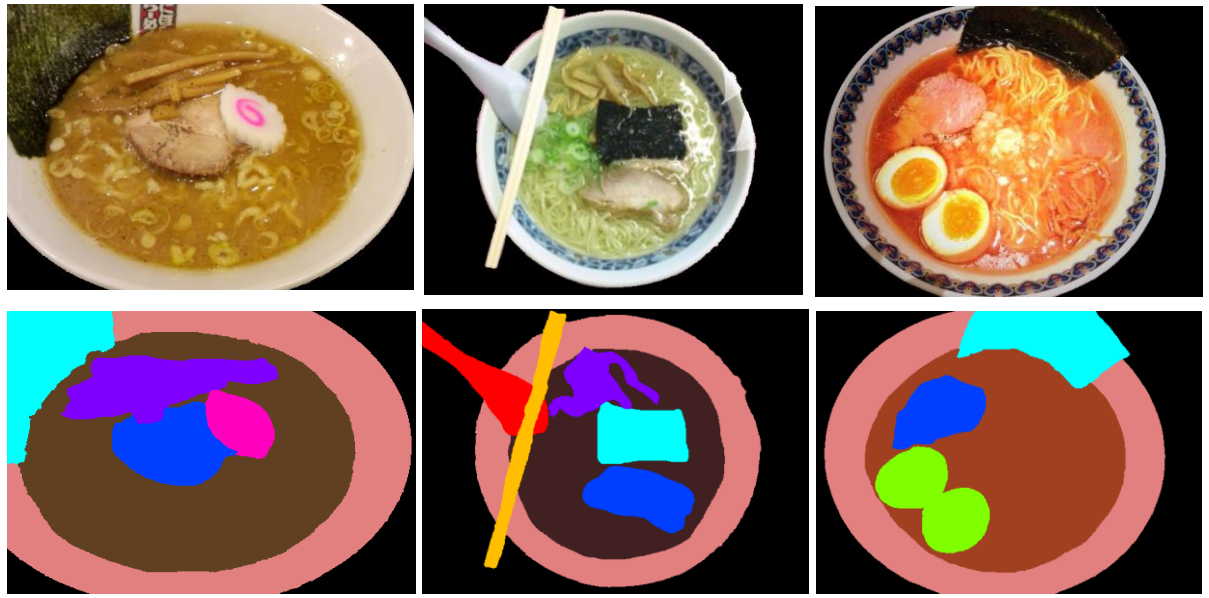
UECFoodPIX

102カテゴリ　1万枚



UEC-Ramen555

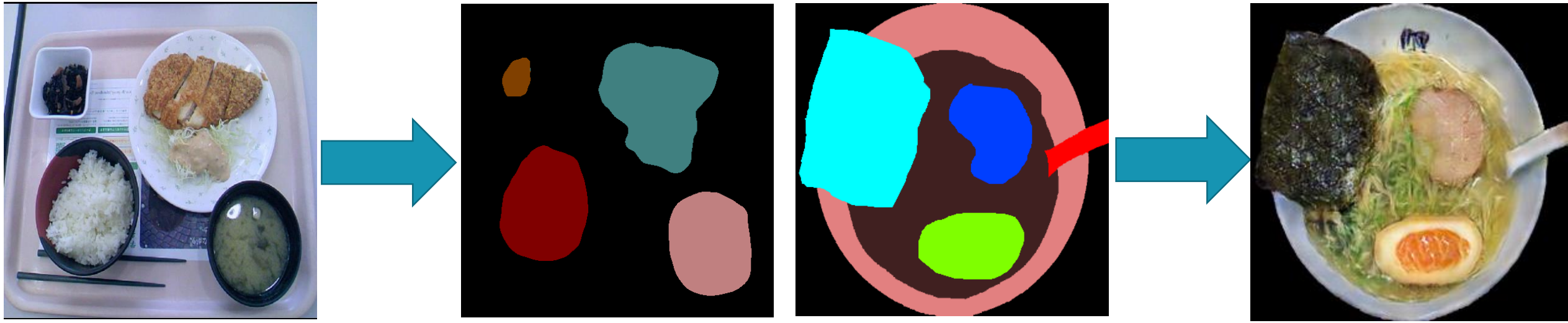
15カテゴリ　555枚



作成したデータセットを用いた活用

食事領域分割

食事画像生成

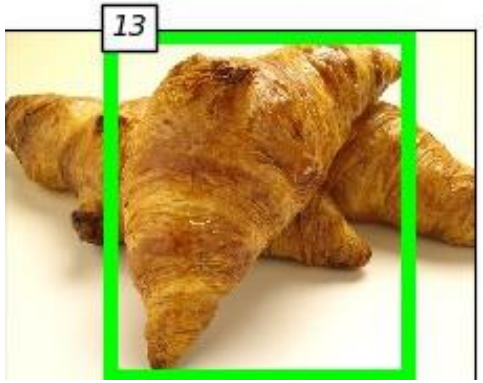


従来のUECFOOD100の問題点

- インスタンスや100カテゴリ以外の食事領域を考慮していない

※学習する際にバウンディングボックスがない領域が不明瞭

従来のバウンディングボックス

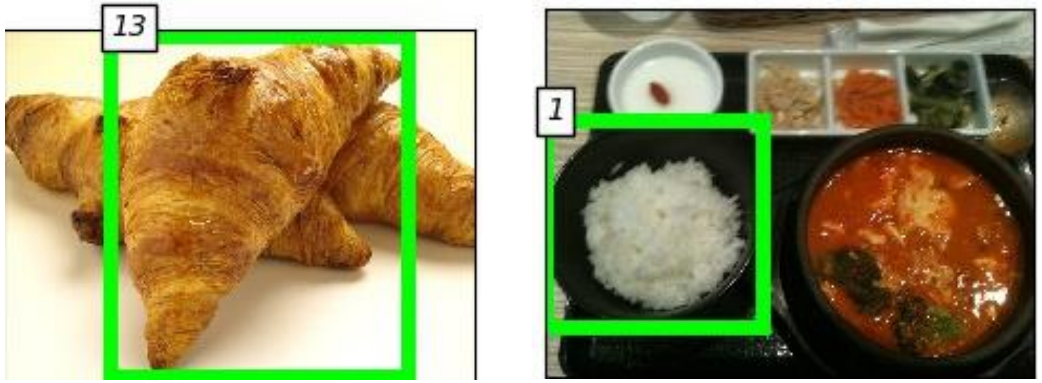


従来のUECFOOD100の問題点

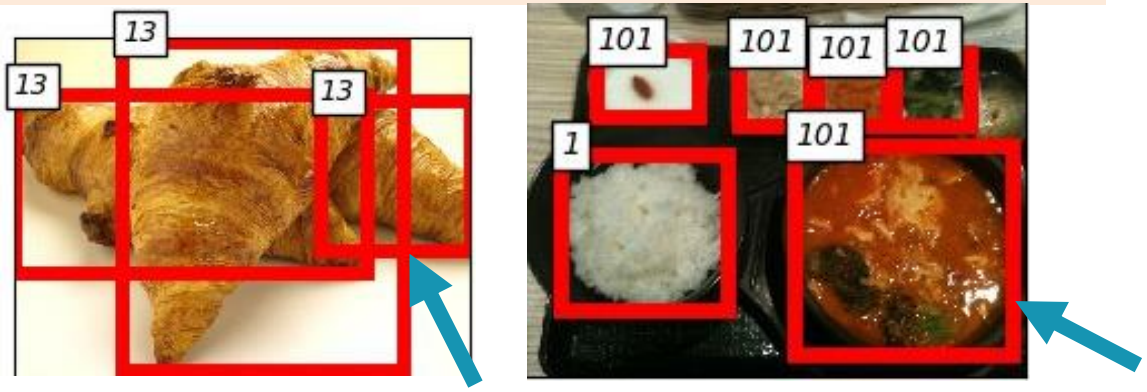
- インスタンスや100カテゴリ以外の食事領
  ※バウンディングボックスがない部分

インスタンスと100カテゴリ以外の食事領域を考慮
手作業で作成

従来のバウンディングボックス



新たなバウンディングボックス



3.データセットの作成　UECFoodPIX

バウンディングボックスのカテゴリラベル付与

- バウンディングボックスが従来の2倍、アノテーションコストが増大

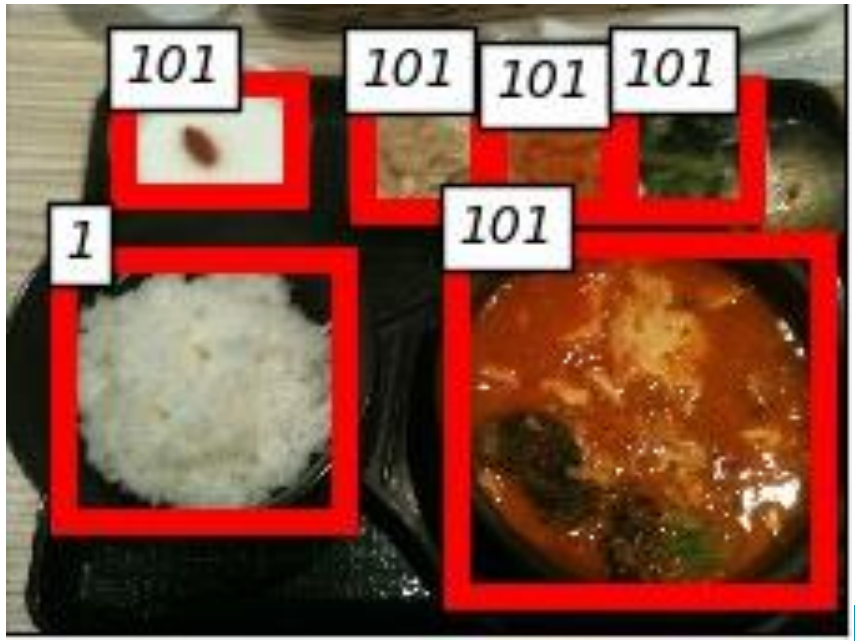
　⇒食事カテゴリをa_o(重なる領域の比)による自動的に付与

$$a_o = \frac{area(B_{uecfoodseg} \cap B_{uecfood100})}{\min(B_{uecfoodseg}, B_{uecfood100})}$$

$a_o \geq 0.5$：元々のカテゴリ

$a_o < 0.5$：’その他の食事’カテゴリ(101)

別に’飲み物’カテゴリも付与(102)



バウンディングボックスのカテゴリラベル付与

- バウンディングボックスが従来の2倍、アノテーションコストが増大

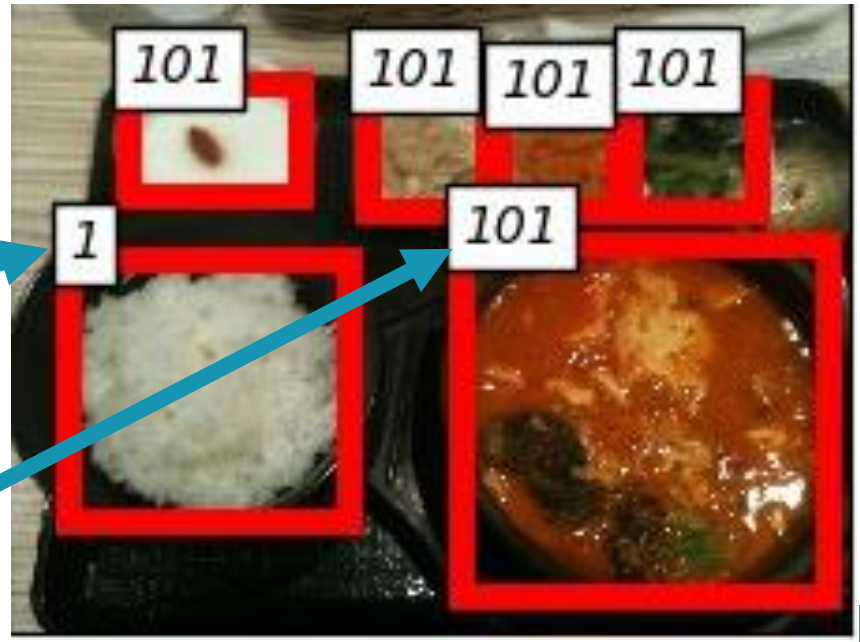
　⇒食事カテゴリをa_o(重なる領域の比)による自動的に付与

$$a_o = \frac{area(B_{uecfoodseg} \cap B_{uecfood100})}{\min(B_{uecfoodseg}, B_{uecfood100})}$$

$a_o \geq 0.5$：元々のカテゴリ

$a_o < 0.5$：’その他の食事’カテゴリ(101)

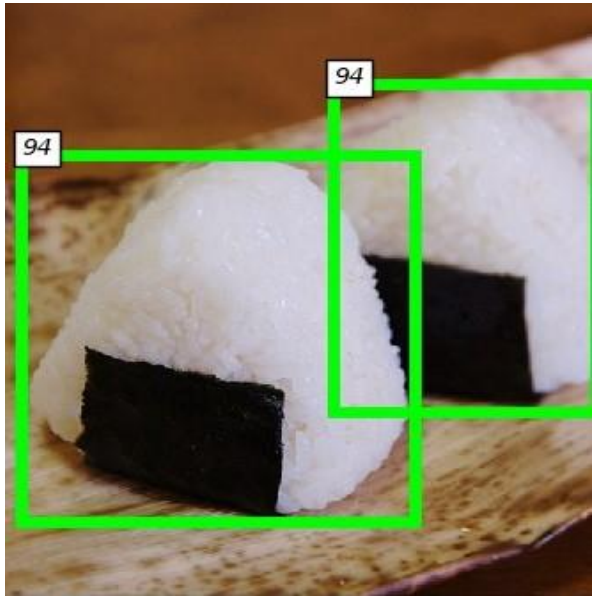
別に’飲み物’カテゴリも付与(102)



マスク画像の作成

⇒新たなバウンディングボックスの領域に対してGrab-cut[5]を適用

バウンディングボックス

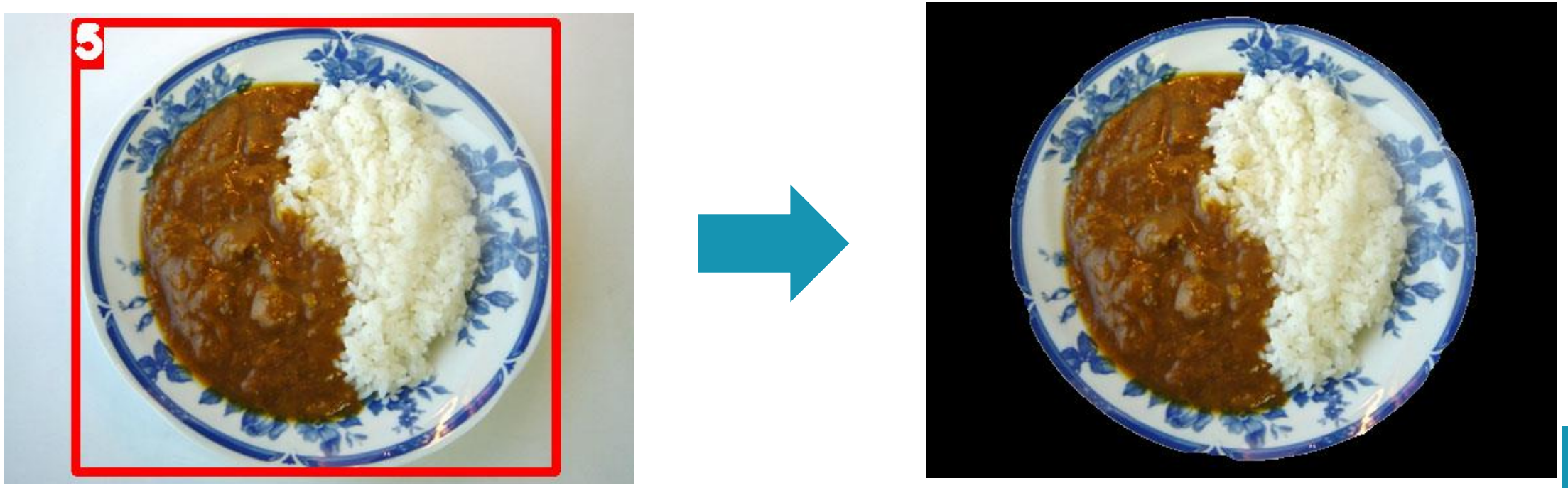


4.実験：食事領域分割 UECFoodPIX

Grab-cut[5]：グラフカットを繰り返し用いて前景と背景を抽出する手法

⇒短径領域を与えるだけで領域分割することが可能

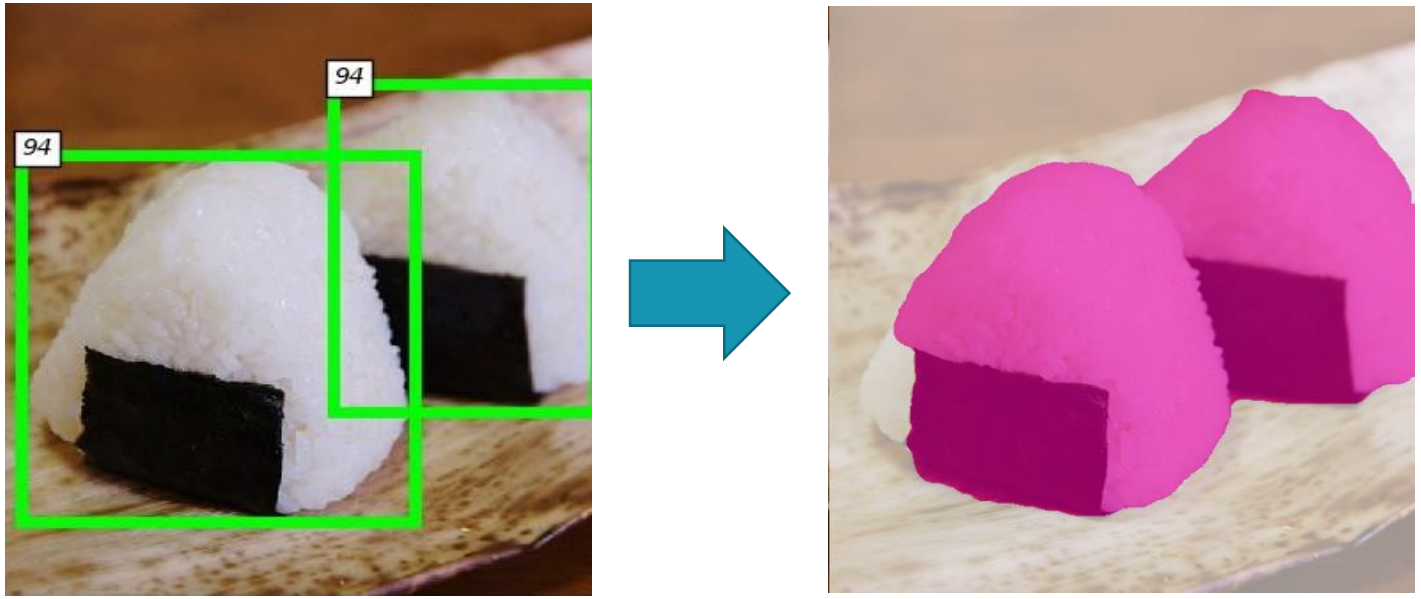
UECFood[3]のバウンディングボックスに対してGrab-cutを適用した例



マスク画像の作成

⇒新たなバウンディングボックスの領域に対してGrab-cut[5]を適用

バウンディングボックス　　Grab-cutによる自動生成



マスク画像の作成

⇒新たなバウンディングボックスの領域に対してGrab-cut[5]を適用

⇒人手で精査することでマスク画像を作成

バウンディングボックス　Grab-cutによる自動生成　マスク画像

食事カテゴリの問題点

- 同じ食事カテゴリでも異なる形や色をしているものが存在

※同じラーメンの画像でもスープや具材によって全く違く見える

同一のラーメンカテゴリ

食事カテゴリの問題点

- 同じ食事カテゴリでも異なる形や色をしているものが存在

※同じラーメンの画像でもスープや具材によって全く違く見える

同一のラーメンカテゴリ

材料の違いを考慮した
ラーメンデータセットの作成

3.データセットの作成　UEC-Ramen555

ラーメンマスク画像

⇒SNS上で収集した555枚のラーメン画像を人手でアノテーション

⇒スープの種類や具材を考慮した15クラスで構成

実画像

マスク画像

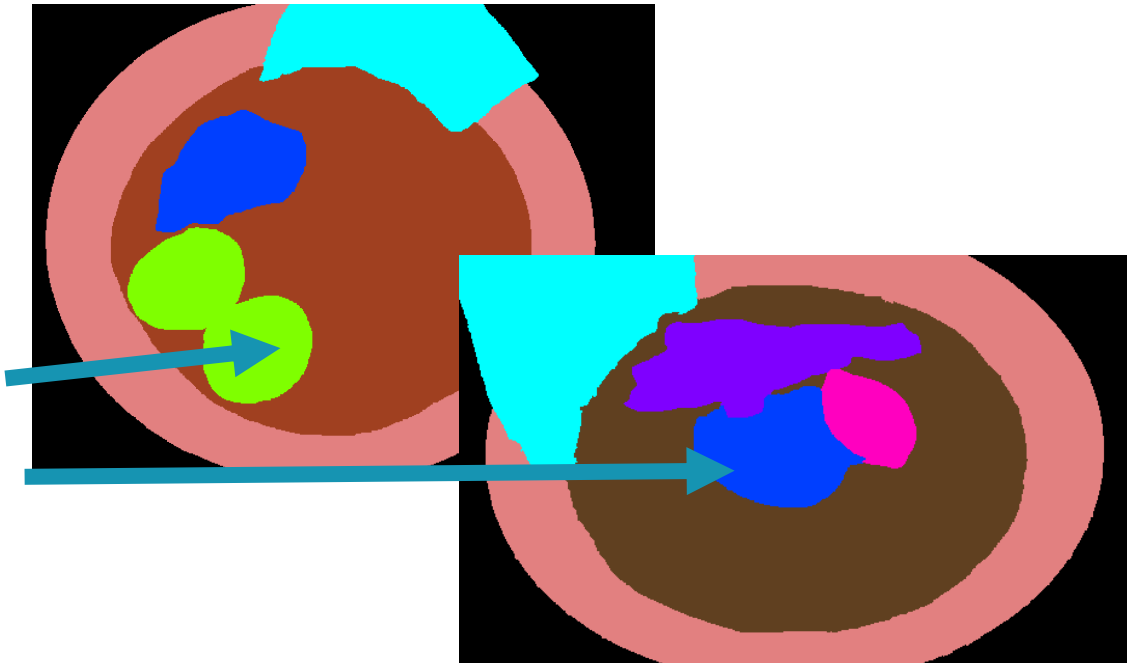
ラーメンマスク画像

⇒SNS上で収集した555枚のラーメン画像を人手でアノテーション

⇒スープの種類や具材を考慮した15クラスで構成

実画像

マスク画像



スープや具材を考慮した
ラーメンマスク画像

4.実験：食事領域分割 UECFoodPIX

UECFoodPIXを用いた食事領域分割を行う

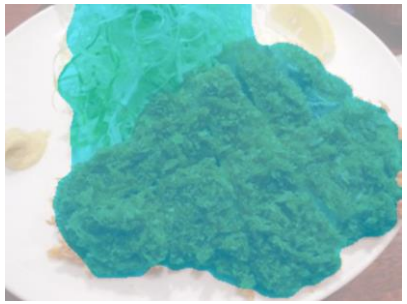
⇒DeeplabV3+[13]による食事領域分割

学習画像：1)自動生成[5]9000枚　2)自動生成[5]7000枚＋マスク画像2000枚

評価画像：マスク画像1000枚

各クラスのAccuracyの平均とmIoUによる評価

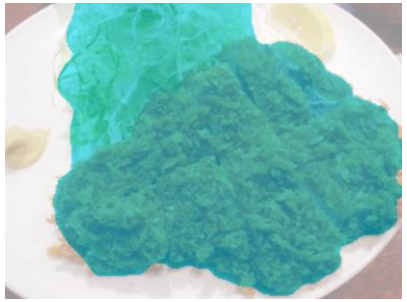
1)の学習画像



自動生成

自動生成

2)の学習画像



自動生成

マスク画像

実験結果

1)と2)の領域分割の各数値

- 2)のほうがどちらとも高い数値

学習画像	Acc	mIoU
1)自動生成9000枚	0.560	0.416
2)自動生成7000枚＋マスク画像2000枚	0.597	0.436

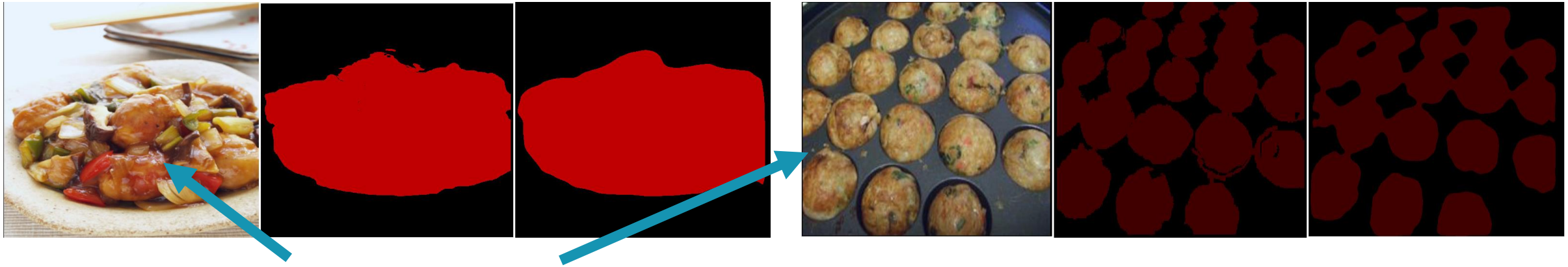
4.実験：食事領域分割 UECFoodPIX

実験結果

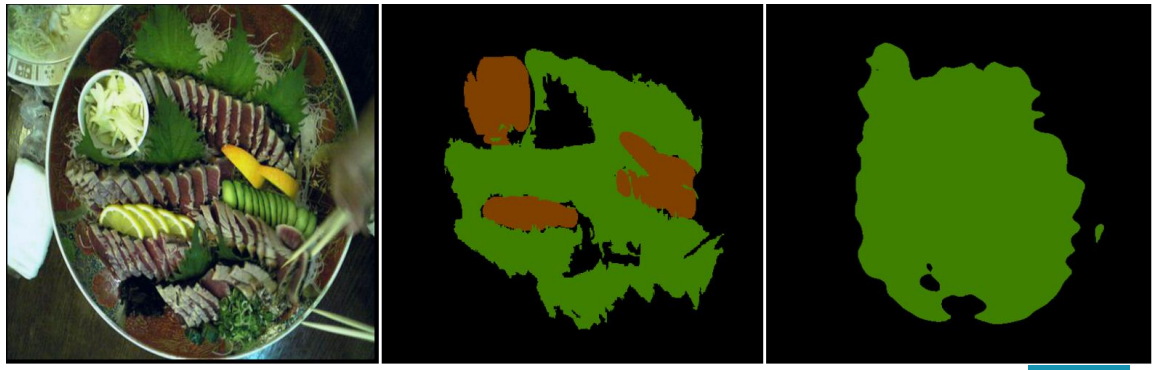
2)による領域分割結果(左：入力画像、中：正解画像、右：推定画像)

実験結果

2)による領域分割結果(左：入力画像、中：正解画像、右：推定画像)



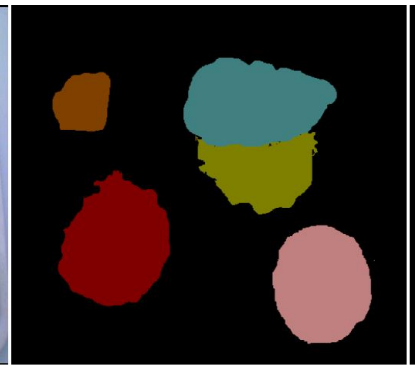
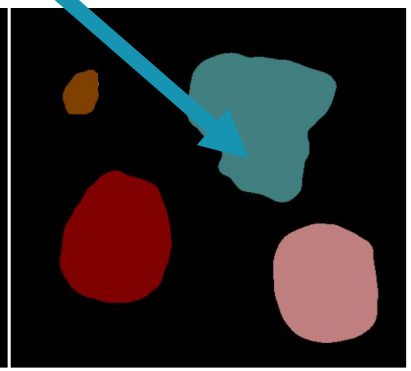
単品品目の食事領域分割はおおよそ可能

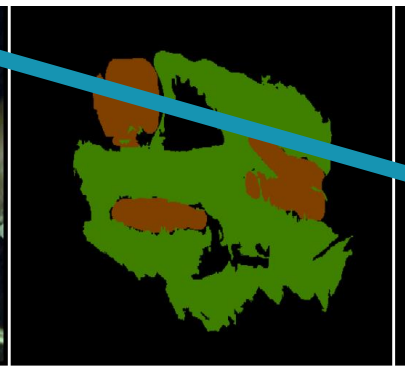
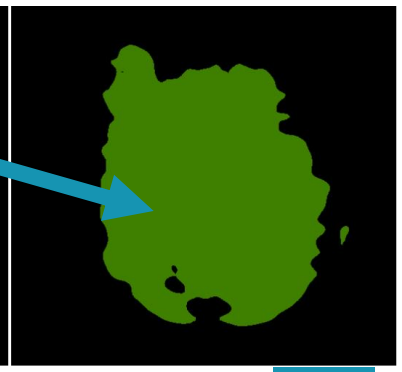


実験結果

2)による領域分割結果(左：入力画像、中：正解画像、右：推定画像)

複数品目の分割において
誤ったカテゴリ分類
(キャベツやレモン)

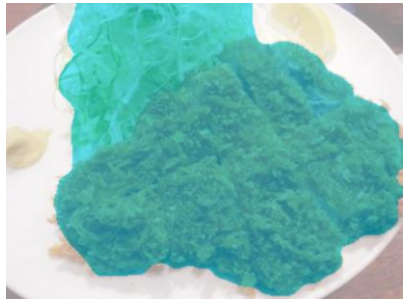



実験結果

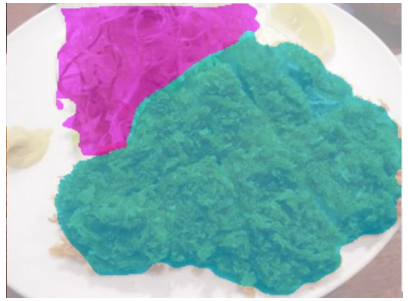
2)による領域分割結果(左：入力画像、中：正解画像、右：推定画像)

複数品目の分割において
誤ったカテゴリ分類
(キャベツやレモン)

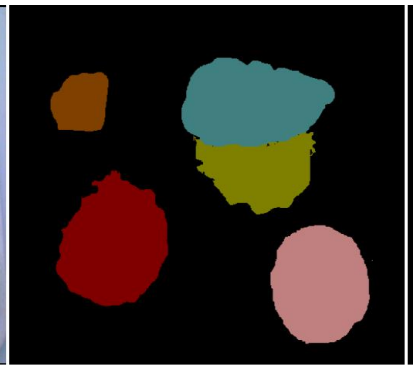
原因：自動生成とマスク画像の違い

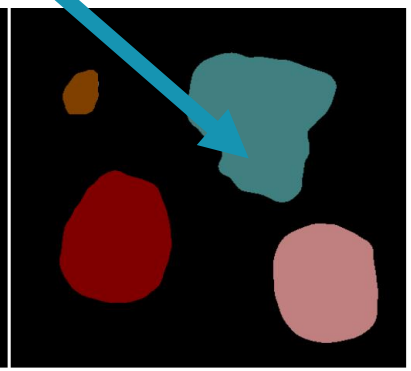


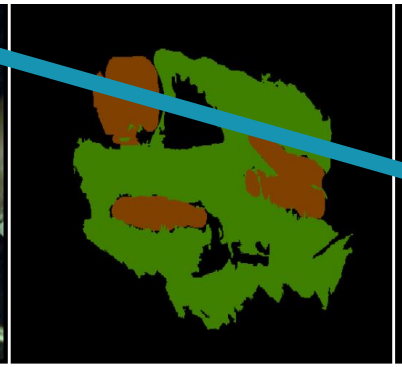
自動生成

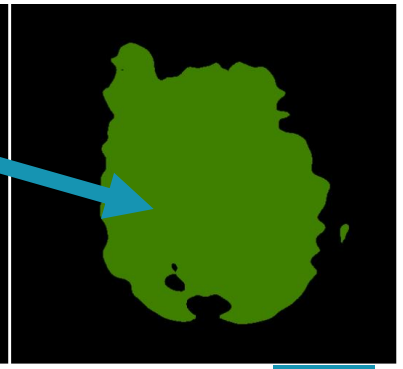


マスク画像







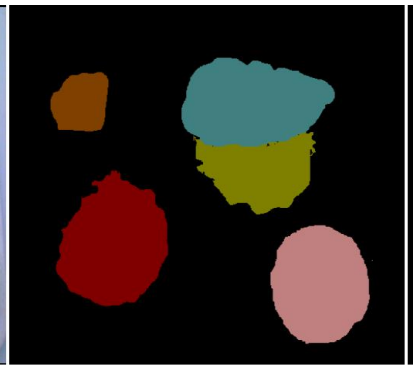
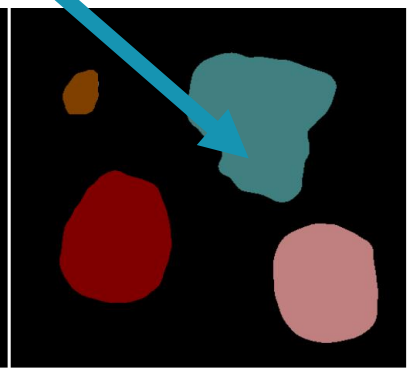


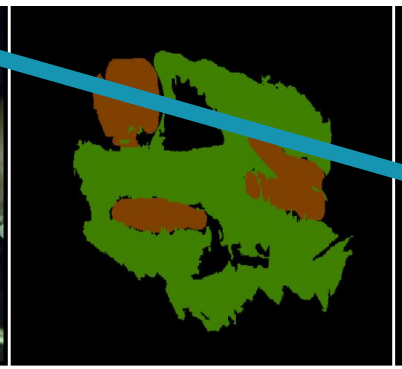
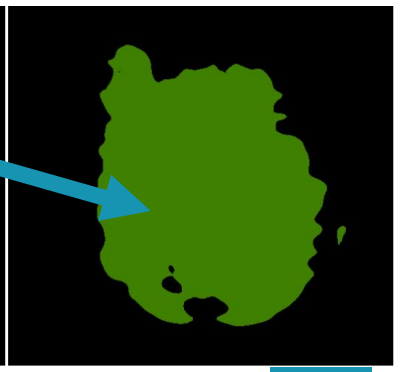
実験結果

2)による領域分割結果(左：入力画像、中：正解画像、右：推定画像)

複数品目の推定において
誤ったカテゴリ分類
(キャベツやレモン)

マスク画像（人手で精査）の
枚数を増やすことで精度の
向上が期待

4.実験：画像生成 UEC-Ramen555

UEC-Ramen555を用いた画像生成を行う

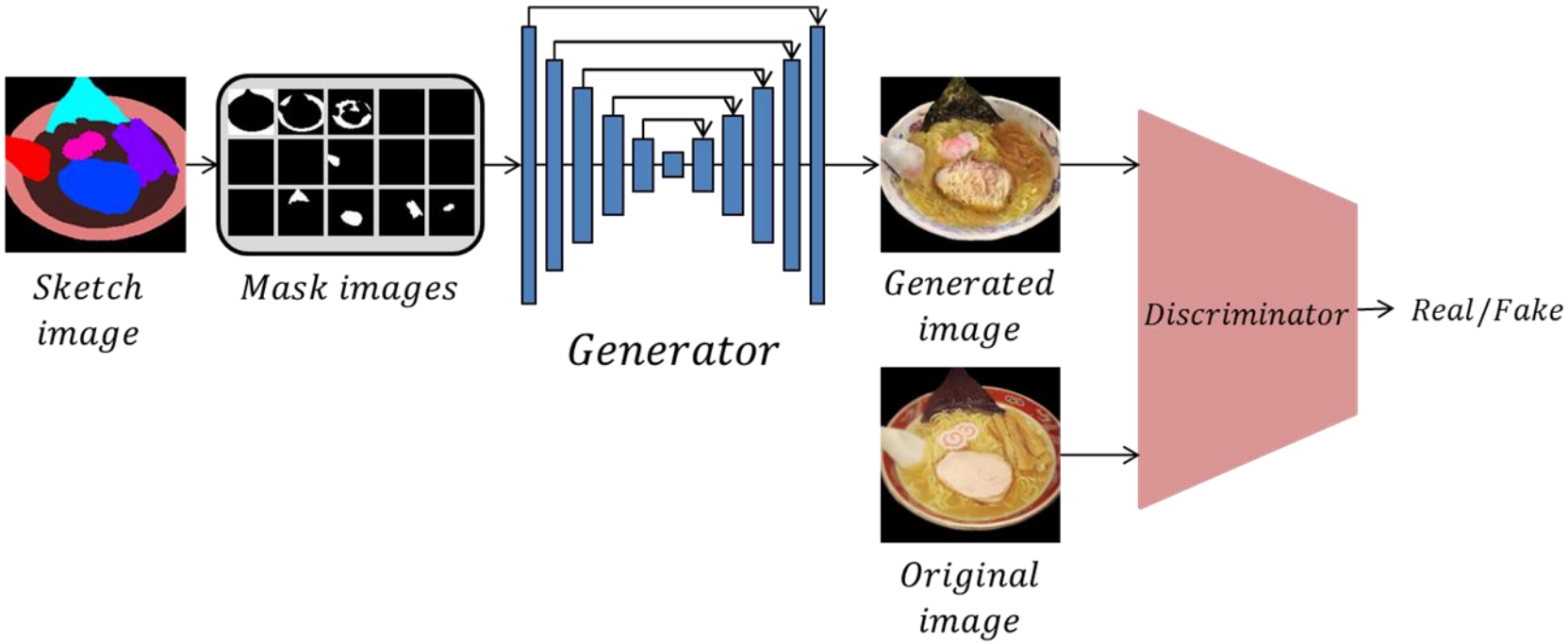
⇒ DeeplabV3+[13]による食事領域分割

4.実験：画像生成 UEC-Ramen555

UEC-Ramen555を用いた画像生成を行う

⇒ DeeplabV3+[13]による食事領域分割

⇒領域抽出したマスク画像を用いてpix2pix[14]による画像生成



画像生成に用いたネットワーク図

4.実験：画像生成 UEC-Ramen555

UEC-Ramen555を用いた画像生成を行う

⇒ DeeplabV3+[13]による食事領域分割

⇒領域抽出したマスク画像を用いてpix2pix[14]による画像生成

学習画像：500枚　評価画像：55枚

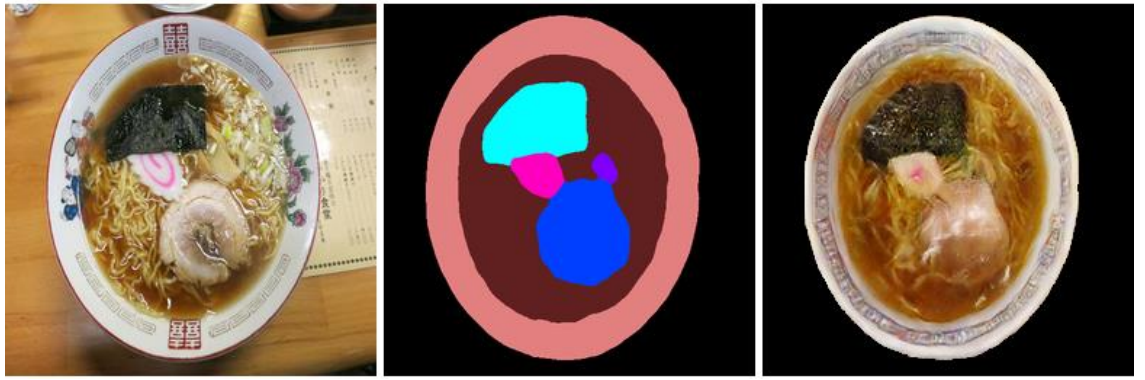
学習したモデルを用いてユーザーが画像を修正できる

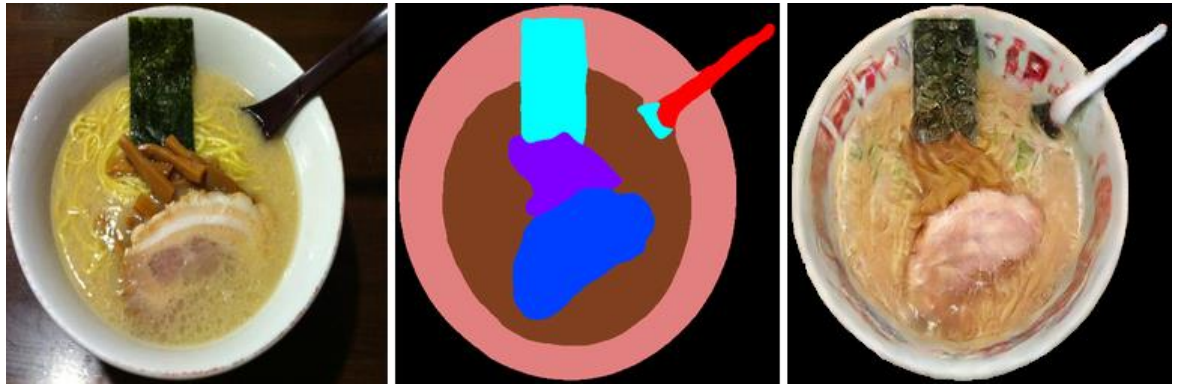
webベースシステムである“RamenAsYouLike” [17]を実装

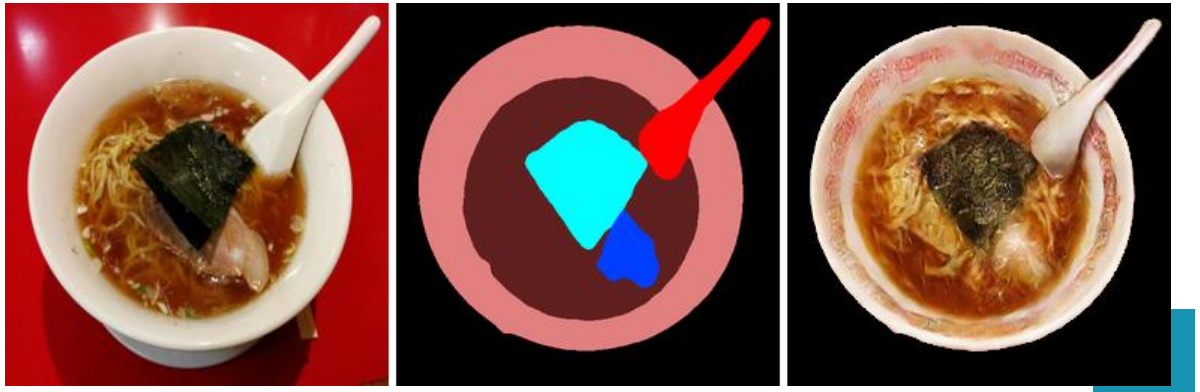
4.実験：画像生成 UEC-Ramen555

実験結果

画像生成結果(左：入力画像、中：領域分割結果、右：生成画像)



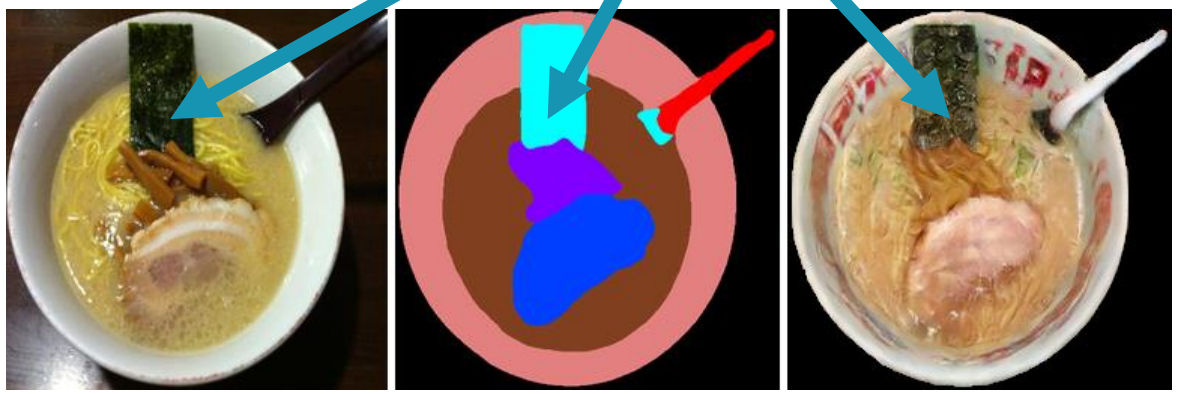


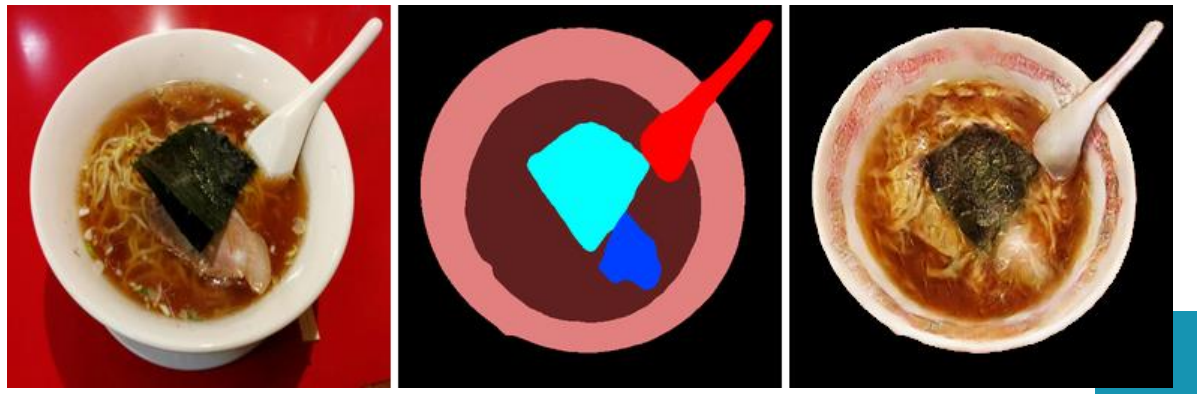


実験結果

画像生成結果(左：入力画像、中：領域分割結果、右：生成画像)

抽出した領域と同じ位置にスープや具材が生成



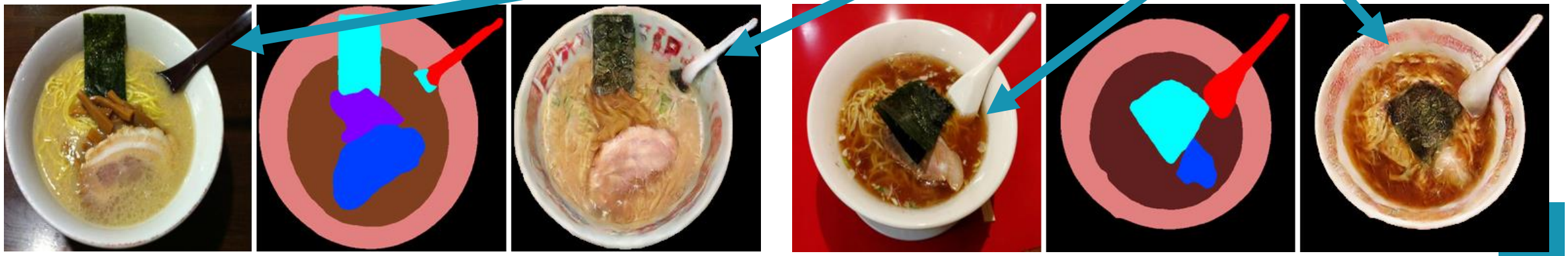


実験結果

画像生成結果(左：入力画像、中：領域分割結果、右：生成画像)

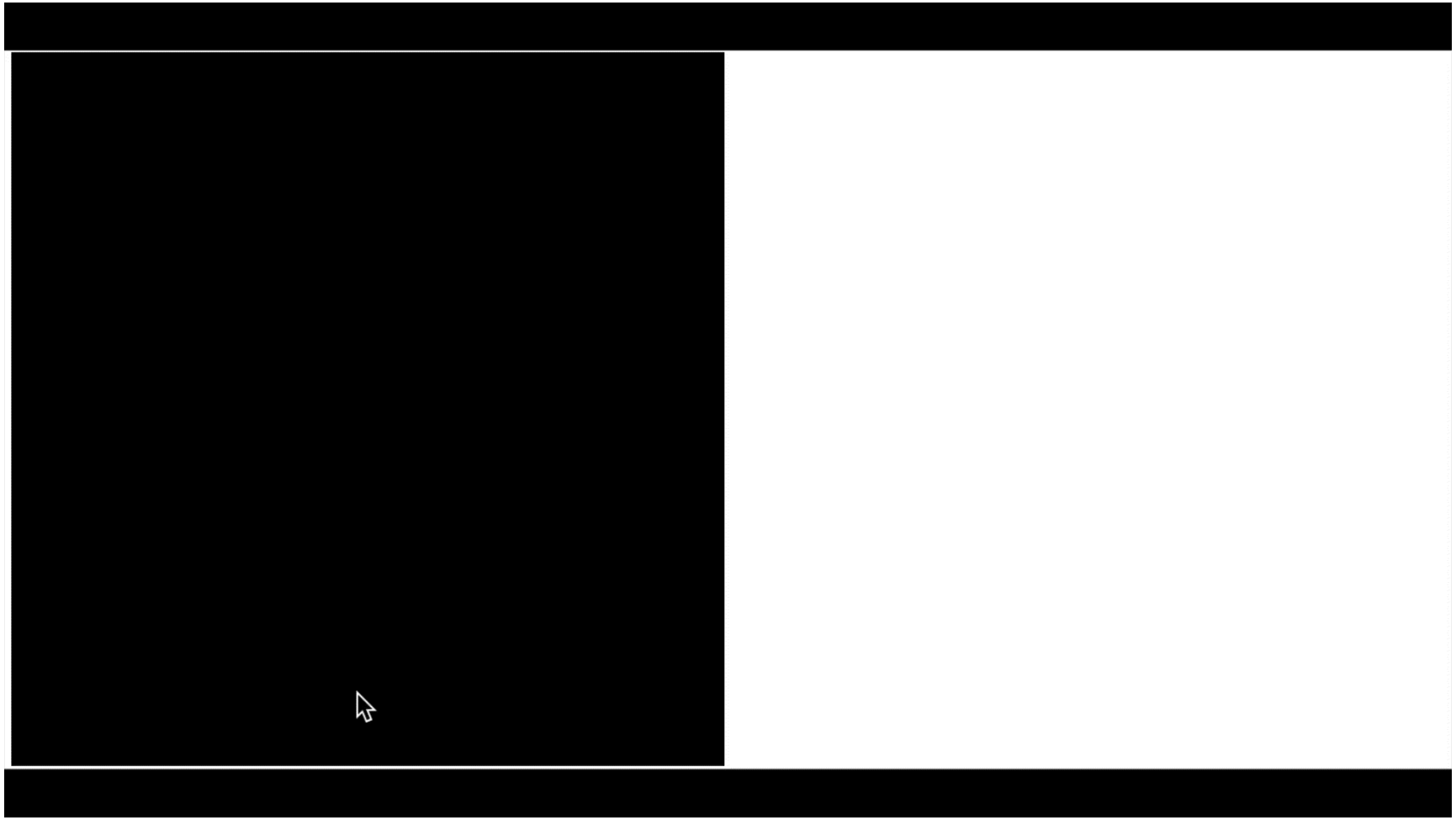
抽出した領域と同じ位置にスープや具材が生成

どんぶりやレンゲのスタイルが保持されていない



4.実験：画像生成 UEC-Ramen555

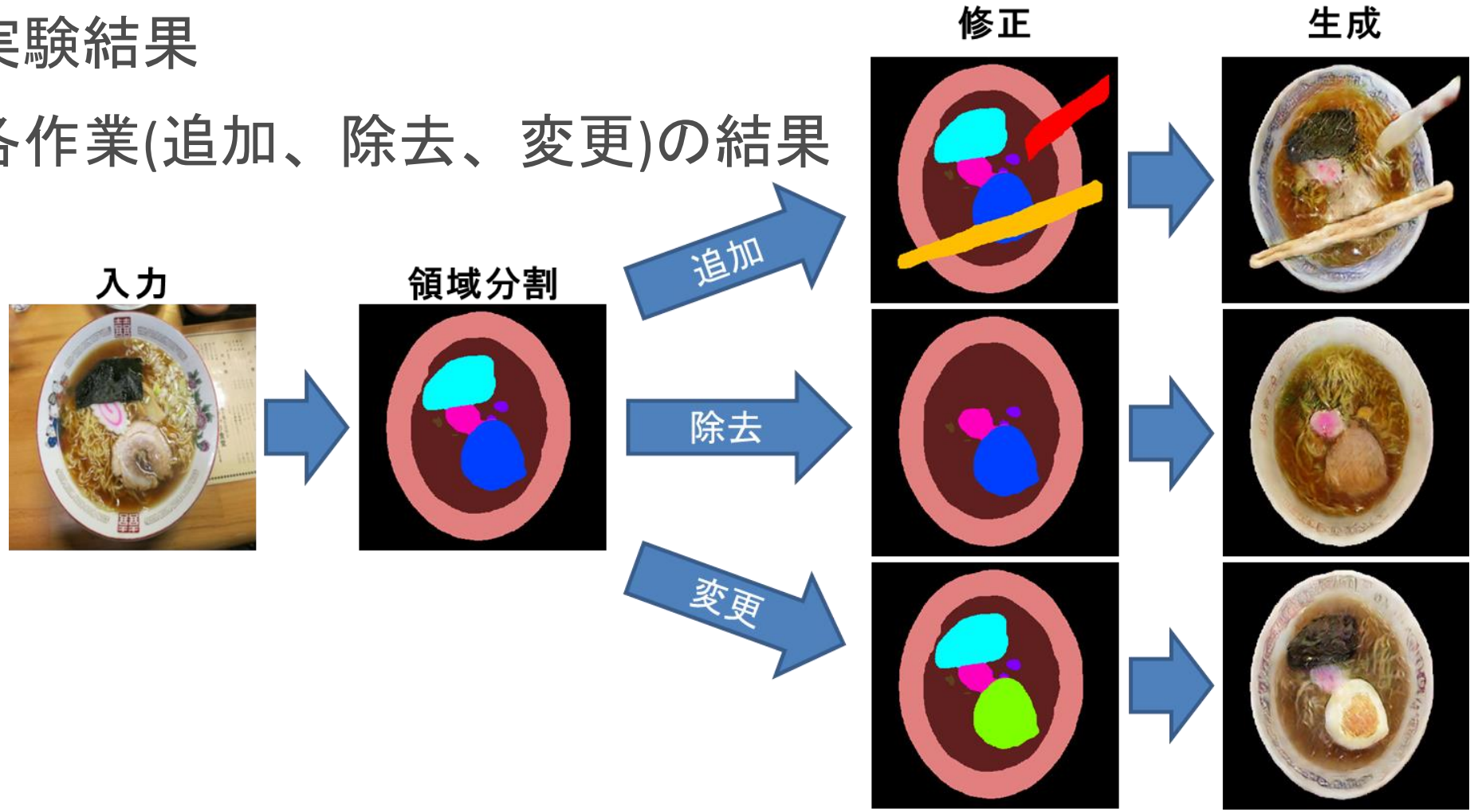
ユーザーが画像を修正できるwebベースシステム“RamenAsYouLike” [17]



ユーザーが画像を修正できるwebベースシステム“RamenAsYouLike” [17]

実験結果

各作業(追加、除去、変更)の結果



5.まとめ

本研究の貢献

- 食事領域分割用データセットUECFoodPIXとUEC-Ramen555の作成
  http://foodcam.mobi/dataset/にて公開予定
- 活用例として領域分割と画像生成を行った

今後の課題

- 食事量やカロリー量推定のベンチマークとなる「量」を考慮したデータセットの整備
- 画像生成においてはスタイルを維持するようなスタイル抽出とスタイルを考慮した画像生成