ID: PS1-18

AR技術とモバイル深層学習を活用した食事カロリー量推定システム

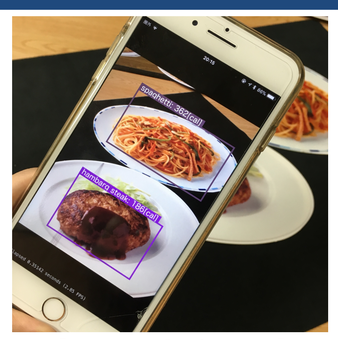
丹野良介, 會下 拓実, Jaehyeong Cho, 柳井 啓司 (電気通信大学)

背景

近年, 食事管理アプリでの**カロリー量**記録が可能に, しかし....
ユーザーによる料理情報の手入力が必要.
コストがかかる人手によるサービス.

with AR (拡張現実)

目的 : 食事画像からの全自動カロリー量推定


[DeepCalorieCam]


[AR DeepCalorieCamV2]

+ Segmentation


[AR DeepCalorieCamV3]

手法 : 補助情報を用いたマルチタスク推定

関連研究①

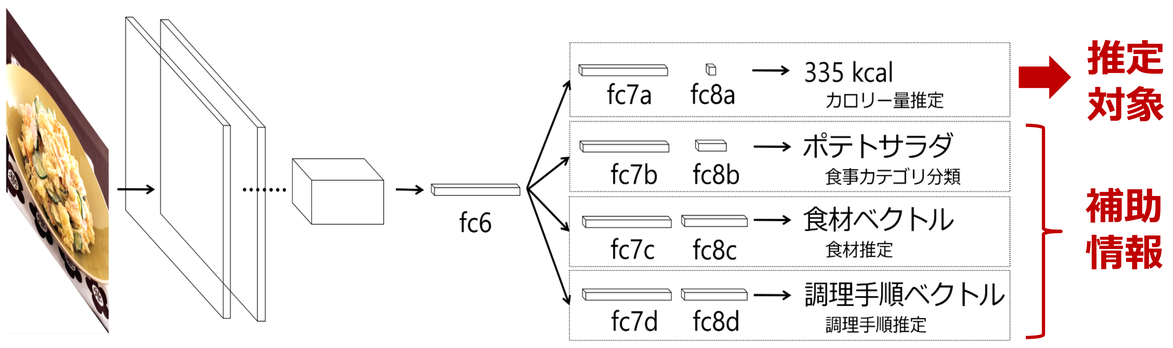
2011年, 宮崎ら[1]による食事画像からのカロリー量**直接推定**
SURF特徴量や**カラーヒストグラム**などの従来の特徴量の使用
高精度のカロリー量推定は実現できなかった

➡ 本研究では**CNN**による食事画像からカロリー量を**直接推定**

関連研究②

2016年, Chenら[2]による画像からの食材&食事カテゴリ同時推定
Multi-task CNN[3]による食材&食事カテゴリのマルチタスク学習
マルチタスク学習により**両タスクの精度が向上**した

➡ 本研究では**カロリー**, **食事カテゴリ**, **食材**, **調理手順**の**マルチタスク学習**



➡ **相乗効果によりカロリー量推定の精度向上が期待できる**

全体損失関数 L :

(λ は ロスの重み)

$$L = \frac{1}{N}\sum_{n=0}^{N}\left(\underline{L_{cal}} + \lambda_{cat}\underline{L_{cat}} + \lambda_{ing}\underline{L_{ing}} + \lambda_{dir}\underline{L_{dir}}\right)$$

L_{cal} : カロリー量推定ロス, L_{cat} : カテゴリ分類ロス, L_{ing} : 食材推定ロス, L_{dir} : 調理手順推定ロス

1. カロリー量推定ロス L_{cal} :

$$L_{cal} = \lambda_{re}L_{re} + \lambda_{ab}L_{ab}$$

$$L_{re} = \frac{|y-g|}{g} \quad L_{ab} = |y-g|$$

(L_{re} : 相対誤差, L_{ab} : 絶対誤差) (y : 推定カロリー量, g : 正解カロリー量)

2. カテゴリ分類ロス L_{cat} :

$$L_{cat} = -\sum_{k=1}^{n} g_k \log y_k \quad (\ Cross\ entropy\ )$$

3, 4. 食材＆調理手順推定ロス L_{ing} & L_{dir} :

$$L_{ing,dir} = \frac{1}{2}\sum_{k=1}^{n}(y_k - g_k)^2 \quad (\ Mean\ squared\ error\ )$$

[クックパッドデータセットの調理手順文章を用いたWord2Vecの学習]

Word2Vec[5]により食材 / 調理手順情報を分散表現に変換し, 正解値として使用.

食材情報

材料(4人分)	
生ざけ	4切れ (260g)
じゃがいも・小	3個
ブロッコリー	1/4個
レモン・輪切り	4枚
パセリ・みじん切り	適量
小麦粉	適量
A 「AJINOMOTO 肝芽の恵みコーン油」	大さじ1

調理手順情報

さけのムニエルの作り方

(1)さけは「コンソメ」をふって両面になじませ、小麦粉をまぶす。

(2)フライパンにAを熱し、(1)のさけの両面を中火で色よく焼き、弱火にしてフタをし、約3分蒸し焼きにする。

(3)じゃがいもは皮をむいて3等分に し、水に10分ほどさらして水気をきる。鍋に入れ、ヒタヒタの水を加えて火にかけ、煮立ったら弱火にし、フタをしてやわらかくなるまで約1

➢ 学習データ : 約 8,710,000 文 (調理手順文章)

➢ Word2Vec辞書 ： 46,304 単語

例）ムニエル



生ざけ		[0.053423 -0.087293 ... 0.007234]
じゃがいも	Word2Vec →	[0.037983 -0.009285 ... 0.034096]
ブロッコリー		[0.002938 -0.048793 ... 0.001240]
・・・		・・・

→ [0.078389 -0.015354 ... 0.002023] **食材情報ベクトル**

さけ		[0.008439 -0.046298 ... 0.023428]
コンソメ	Word2Vec →	[0.023982 -0.004298 ... 0.073249]
ふって		[0.003428 -0.092392 ... 0.006345]
両面		・・・
・・・		

→ [0.005998 -0.035861 ... 0.019575] **調理手順ベクトル**

検索結果例①

クエリ : 豚バラ肉	
豚バラ	0.9232
豚ばら肉	0.8743
豚肉	0.8636
バラ肉	0.7855

検索結果例②

クエリ : にんじん	
人参	0.8059
だいこん	0.6358
たまねぎ	0.6132
ニンジン	0.6097

実験 : シングルタスクとマルチタスクの比較

カロリー量付き食事画像セットの構築

Web上のレシピサイト

1. AJINOMOTO.Park
2. E・レシピ
3. kikkoman
4. レタスクラブニュース
5. みんなのきょうの料理 NHKエデュケーショナル
6. オレンジページnet

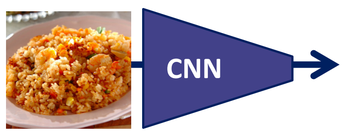


レシピデータとして収集
- 食事画像
- カロリー量
- 食材情報
- 調理手順情報

食事 **15** カテゴリ
合計 **4877** 枚

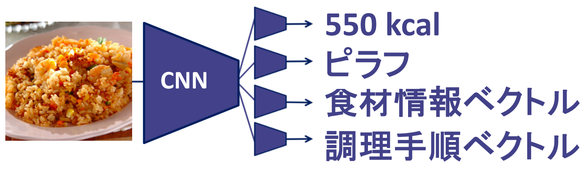


[Single-task CNN (baseline)]



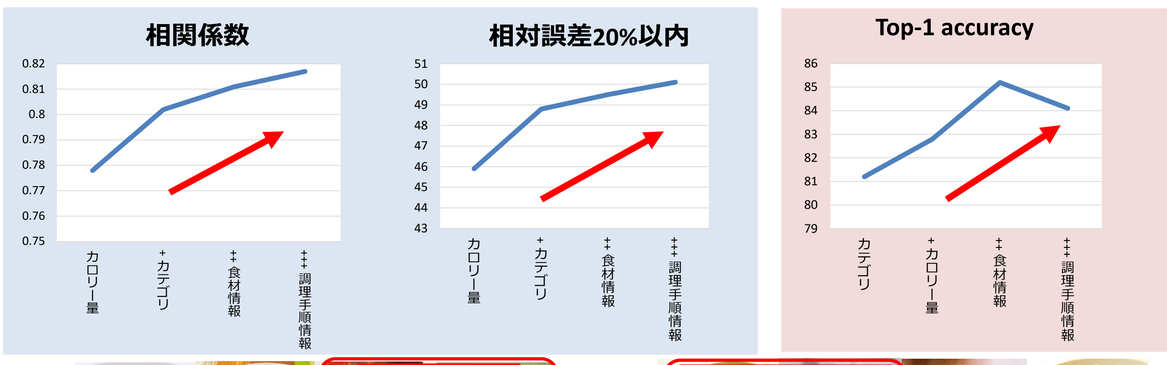
相関係数 (カロリー推定)	0.78
相対誤差20%以内 (カロリー推定)	46 %
Top-1 accuracy (カテゴリ分類)	81 %

[Multi-task CNN (提案手法)]



相関係数 (カロリー推定)	0.82 **+0.04**
相対誤差20%以内 (カロリー推定)	50 % **+4 %**
Top-1 accuracy (カテゴリ分類)	84 % **+3 %**

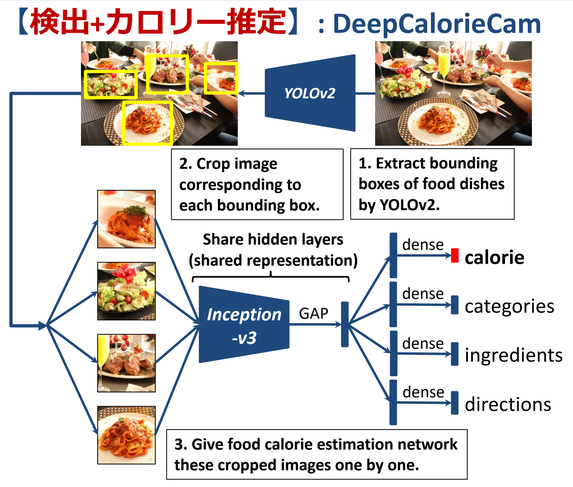
➡ **相乗効果によりカロリー量・カテゴリ推定の精度が向上**





アプリケーション : カロリー量推定システム

【検出+カロリー推定】 : DeepCalorieCam



デモ 2

【ARによる食事面積自動計量 + カロリー推定】
領域分割を利用した食事領域自動算出
: AR DeepCalorieCamV3
領域分割: Mobilenet V2 + U-Net

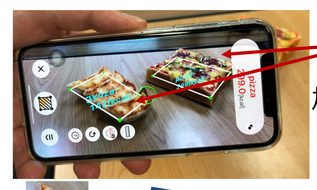


学習画像: UEC-Food100 (GrabCutでマスク生成)

デモ 3

【ARによる食事面積計量 食事の実領域を算出 +カロリー推定】 : AR DeepCalorieCamV2

カメラ入力と加速度センサを併用した
モバイルＡＲ技術
- **キャリブレーションカードなし**で, **実寸値**を測定
- **実際の面積に基づくカロリー量推定**を実現



加速度センサーやカメラから物体の距離を算出 (ARKit framework利用)



Calorie =142 [kcal] (gyoza)

Calorie = 880 [kcal] (pizza)

[カロリー, 食事領域面積]



食品領域の面積 : F_a , ピクセル数 : F_p カロリー量 : C , 係数 : a, b, c
基準物体領域の面積 : B_a ピクセル数 : B_p

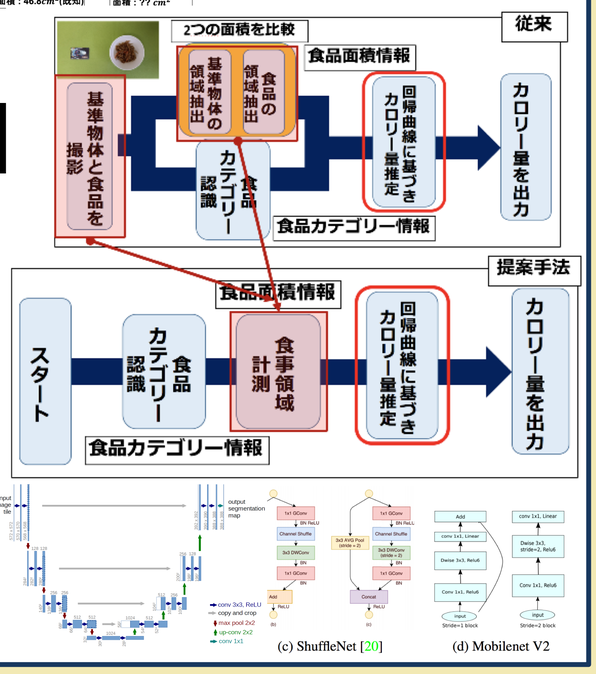
$$F_a = B_a * \frac{F_p}{B_p}, \quad C = a * F_a^2 + b * F_a + c$$

（参考）従来：面積ベースのカロリー量推定


基準物体 ピクセル数 : 43149 面積 : 46.8cm^2(既知)


きんぴらごぼう ピクセル数 : 68066 面積 : ?? cm^2

食品領域と基準物体領域のピクセル数から比率を求め、比率を基準物体の面積に掛け合わせる



まとめと今後の課題

- CNNによる食事画像からのカロリー量直接推定
- マルチタスク学習により各タスクの精度が向上
- カロリー量情報付き食事画像データセットの構築
- 領域分割などによる食品の量を考慮したカロリー量推定
- 大規模データセットの構築

[1] T. Miyazaki, G. Chaminda, D. Silva, and K. Aizawa. Image - based calorie content estimation for dietary assessment. In Proc. of IEEE ISM Workshop on Multimedia for Cooking and Eating Activities, 2011.
[2] J. J. Chen and C. W. Ngo. Deep-based ingredient recognition for cooking recipe retrieval. In Proc. of ACM International Conference Multimedia, 2016.
[3] H. A. Abrar, W. Gang, L. Jiwen, and J. Kui. Multi-task CNN model for attribute prediction. IEEE Transactions on Multimedia, Vol. 17, No. 11, pp. 1949–1959, 2015.

ECCV Workshop

Home · Call for Papers · Call for Artworks · Fashion Challenge · Program · Organizers · Contact

食事変換プロジェクトHP

First Workshop on

Computer Vision for Fashion, Art and Design

丹野良介, 堀田大地, 成冨志優, 柳井啓司

（宣伝）Artworks部門でAcceptされました

食事変換例