

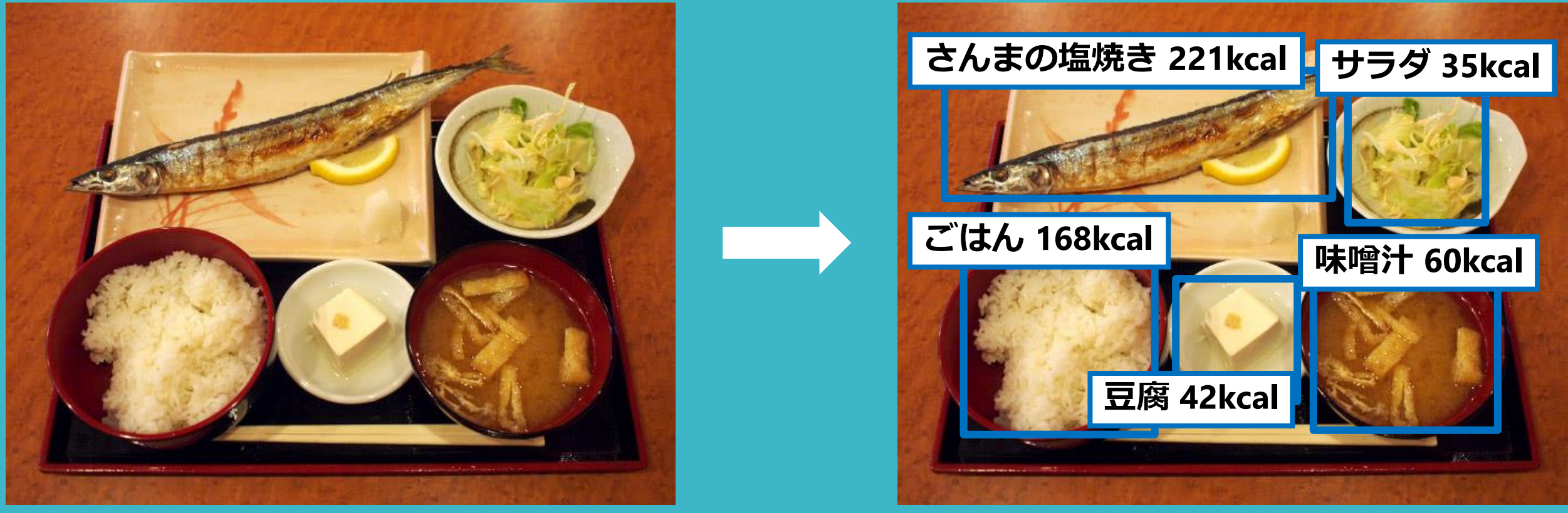
會下 拓実, 柳井 啓司 (電気通信大学)

背景と目的

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...

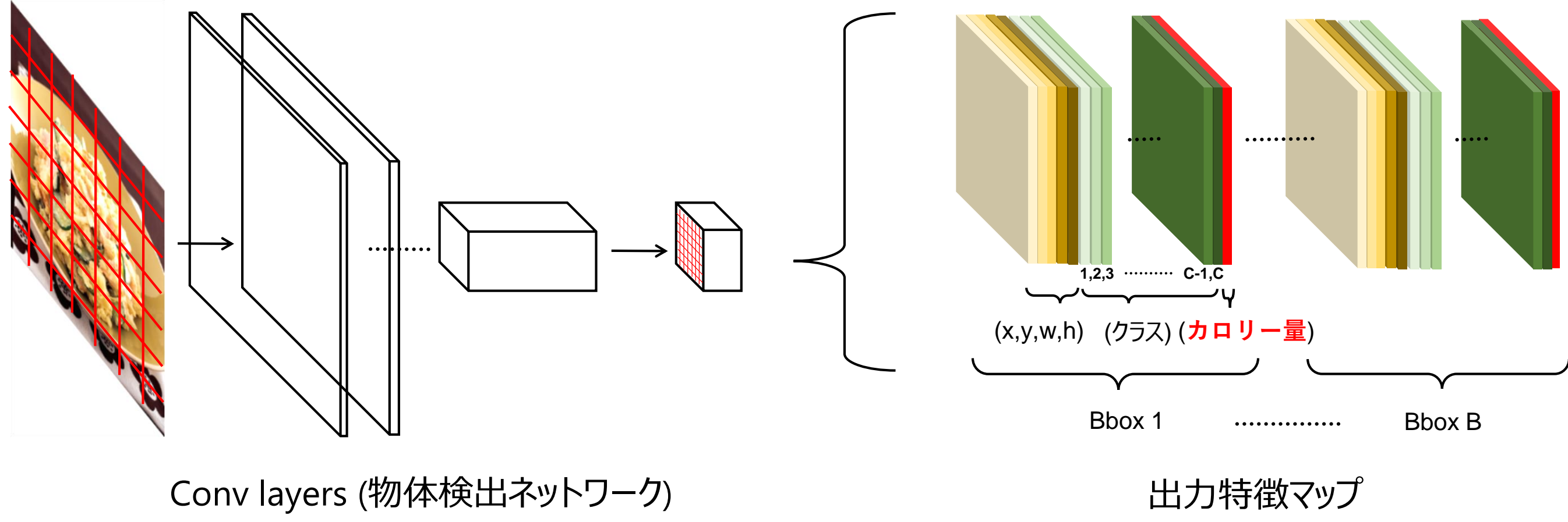
- 料理名や量などの情報のユーザによる手入力が必要
- 1品ずつカロリー量を計算
- 人手によるサービスにより料金が発生

複数料理画像からの全自動カロリー量推定



手法 : 料理検出とカロリー量推定

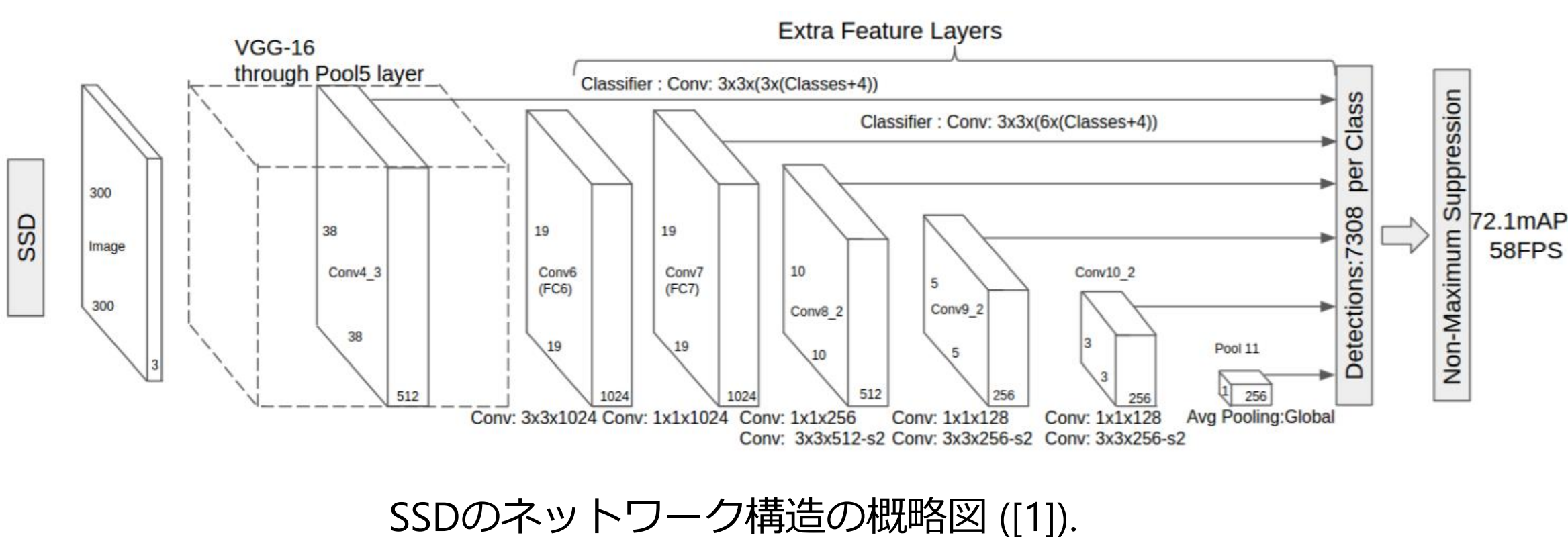
- 単一CNNによる料理検出とカロリー量推定のマルチタスク学習。
- 複数料理画像から単品料理を検出し, そのカロリー量を同時に推定。
- 物体検出モデルの出力層にカロリー量を出力する**チャンネル**を追加。



単一ネットワークでの同時推定による高速化と省メモリ化が期待できる。

1. SSD : Wei Liu et. al. [1] 2016 (物体検出)

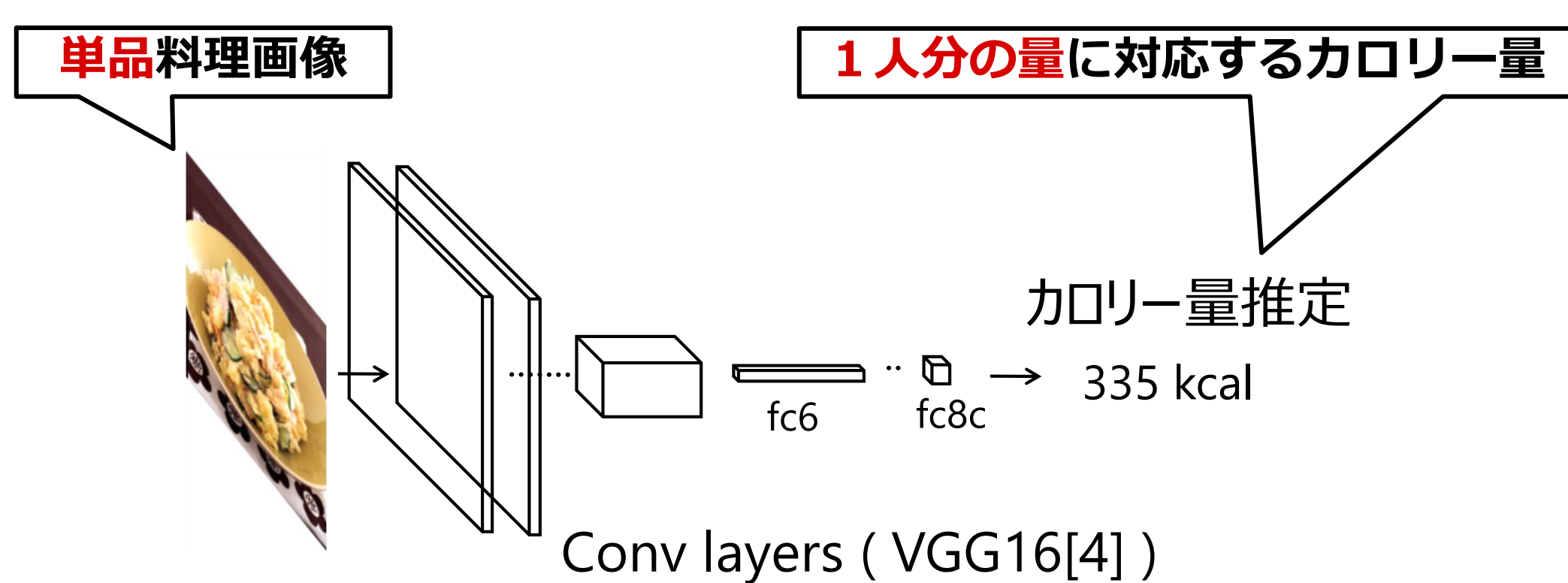
- CNNを用いた高速かつ高精度な物体検出ネットワーク。
- ネットワーク全体のEnd-to-end学習が可能。



本研究ではSSDを用いて複数料理画像からの単品料理検出を行う。

2. Single-task CNN : Ege and Yanai [2] 2017 (カロリー量推定)

- CNNを用いた回帰による料理画像からのカロリー量直接推定。
- カロリー量と料理カテゴリのマルチタスク学習により精度向上を達成。



- 絶対誤差を L_{ab} , 相対誤差を L_{re} とすると, カロリー量推定タスクの損失関数 L_{cal} は下のように定義される。

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

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

(y : 推定されたカロリー量, g : 正解カロリー量)

本研究ではこのネットワークを用いて料理画像からのカロリー量推定を行う。

実験 : シングルタスク学習との比較

- ベースライン (シングルタスク学習)
 - 料理検出のシングルタスク学習 (UECFood-100 dataset [3]を使用) (a)
 - カロリー量推定のシングルタスク学習 (Calorie-50 datasetを使用) (b)
 - 二段階推定による複数料理画像からのカロリー量推定手法 (a) により検出された単品料理画像を(b)に与えることでカロリー量推定を行う。
- 本手法 (マルチタスク学習)
 - 単一CNNによる料理検出とカロリー量推定のマルチタスク学習
 - UECFood-100 dataset に対しては物体検出に関する損失関数を使用。
 - Calorie-50 dataset に対してはカロリー量推定に関する損失関数を使用。

[Calorie50 dataset] [UEC Food-100 dataset [3]] [擬似Bboxアノテーション例]

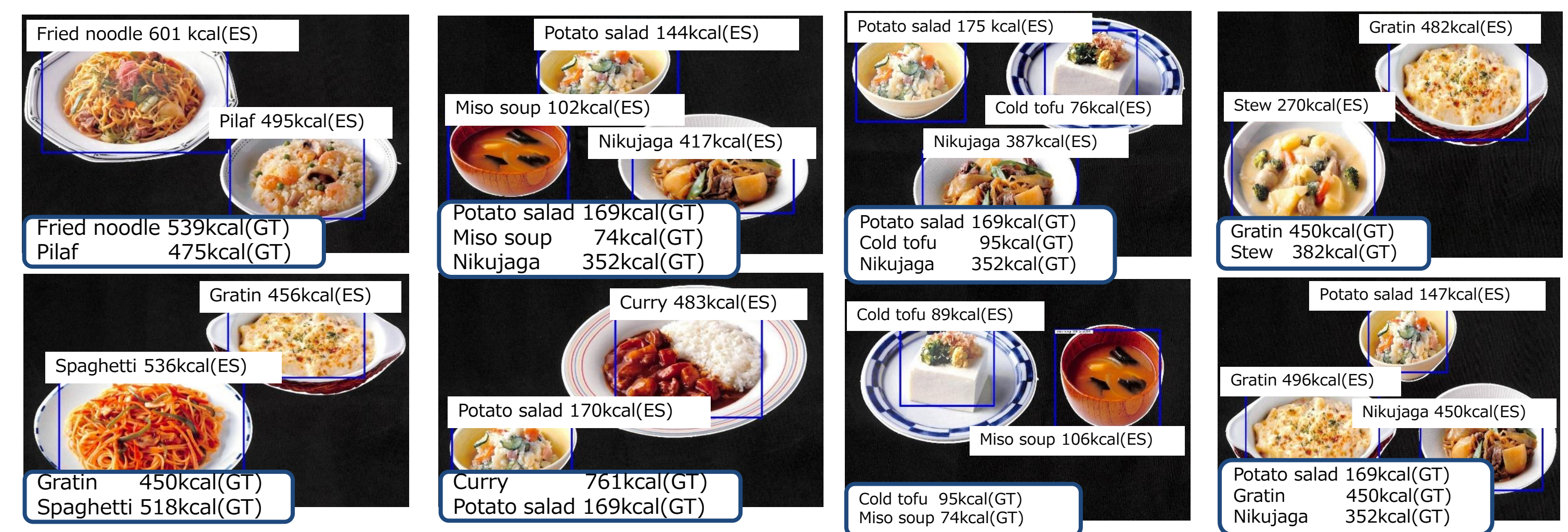
○ カロリー量 × カロリー量
 × バウンディングボックス ○ バウンディングボックス

Calorie50 datasetにBboxを付与

実験結果

手法	mAP (%)	相対誤差 (%)	絶対誤差 (kcal)	相関係数	相対誤差 20%以内 (%)
料理検出 (uecfood100)	34.1	---	---	---	---
料理検出 (uecfood100+calorie50) (a)	37.8	---	---	---	---
カロリー量推定 (calorie50) (b)	---	27.1	91.8	80.7	50.9
二段階推定 (a)→(b)	---	27.3	92.5	80.5	50.6
料理検出+カロリー量推定 (本手法)	37.7	26.6	89.4	81.0	50.7

- 本手法はシングルタスク学習と比較して、精度の低下なしに複数料理画像からのカロリー量推定を実現。
- 単一ネットワークでの同時推定により**高速化と省メモリ化**を達成。



複数料理画像からの料理検出とカロリー量推定の例 (物体検出ネットワークとしてYOLOv2 [Redmon et al. CVPR 2017] を使用.)

まとめと今後の課題

- 複数料理画像からの同時カロリー量推定を行い, 単一CNNでの同時推定による高速化と省メモリ化を実現。
- マルチタスク学習によりカロリー量推定精度が向上。
- 領域分割などを用いた食品の量を考慮したカロリー量推定。
- 大規模料理画像データセットの構築。

[1] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, and S. E. Reed. SSD: single shot multibox detector. CoRR, abs/1512.02325, 2015.

[2] T. Ege and K. Yanai. Simultaneous estimation of food categories and calories with multi-task cnn. In Proc. of IAPR International Conference on Machine Vision Applications(MVA), 2017.

[3] Y. Matsuda, H. Hajime, and K. Yanai. Recognition of multiple-food images by detecting candidate regions. In Proc. of IEEE International Conference on Multimedia and Expo, 2012.

[4] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. In arXiv preprint arXiv:1409.1556, 2014.