一般社団法人　電子情報通信学会
THE INSTITUTE OF ELECTRONICS,
INFORMATION AND COMMUNICATION ENGINEERS

信学技報
IEICE Technical Report
IMQ2021-68,IE2021-130,MVE2021-97(2022-03)

† † †† †

† 470–0393 101
†† 464–8601
E-mail: †miyazaki.m@md.sist.chukyo-u.ac.jp, {kdoman,y-mekada}@sist.chukyo-u.ac.jp
††ide@i.nagoya-u.ac.jp

あらまし

キーワード

A study on multi-task learning for the attractiveness estimation of food photography

Mitsuaki MIYAZAKI†, Keisuke DOMAN†, Ichiro IDE††, and Yoshito MEKADA†

† School of Engineering, Chukyo University
101 Tokodachi, Kaizu-cho, Toyota, Aichi, 470–0393 Japan
†† Graduate School of Informatics, Nagoya University
Furo-cho, Chikusa-ku, Nagoya, Aichi, 464–8601 Japan
E-mail: †miyazaki.m@md.sist.chukyo-u.ac.jp, {kdoman,y-mekada}@sist.chukyo-u.ac.jp
††ide@i.nagoya-u.ac.jp

Abstract We report the results of a study on multi-task learning for the accuracy improvement of attractiveness estimation of food photography. We have been studying a way of improving the estimation accuracy by effectively using a small-scale image dataset. This paper proposes another approach based on multi-task learning for attractiveness estimation. Multi-task learning is a method that simultaneously learns multiple tasks within one model. Solving a task of interest together with its related tasks can improve the generalization performance of a trained model compared to solving each task independently. Thus, we expect that such a multi-task approach is effective for the attractiveness estimation of food photography. This report proposes a multi-task learning method that simultaneously solves multiple problems including attractiveness estimation as the main task and related estimation/classification problems as subtasks, and also confirm its effectiveness through evaluation experiments.

Key words Food photography, attractiveness estimation, multi-task learning, regression analysis

1. まえがき

1

1a 1b

SNS

Web

This article is a technical report without peer review, and its polished and/or extended version may be published elsewhere.
Copyright ©2022 by IEICE

(a) (b)

1:

Fig. 1 Example of food photos with different attractiveness.

CNN

[1]

Data Augmentation

[2]

[3]

2. 料理写真の魅力度推定手法

2 2

Data Augmentation [2]

4

2.1 マルチタスク学習

3 VGG16 [4]

2:

Fig. 2 Process-flow of the estimation method.

(1)

$$
\begin{aligned}
L &= \alpha L_{\mathrm{main}} + L_{\mathrm{sub}} \\
L_{\mathrm{sub}} &= \sum_{i=1}^{N} \beta_i L_i \\
\alpha &+ \sum_{i=1}^{N} \beta_i = 1.0
\end{aligned}
\tag{1}
$$

L_{main} L_{sub} α β_i N

$\alpha, \beta \geqq 0$

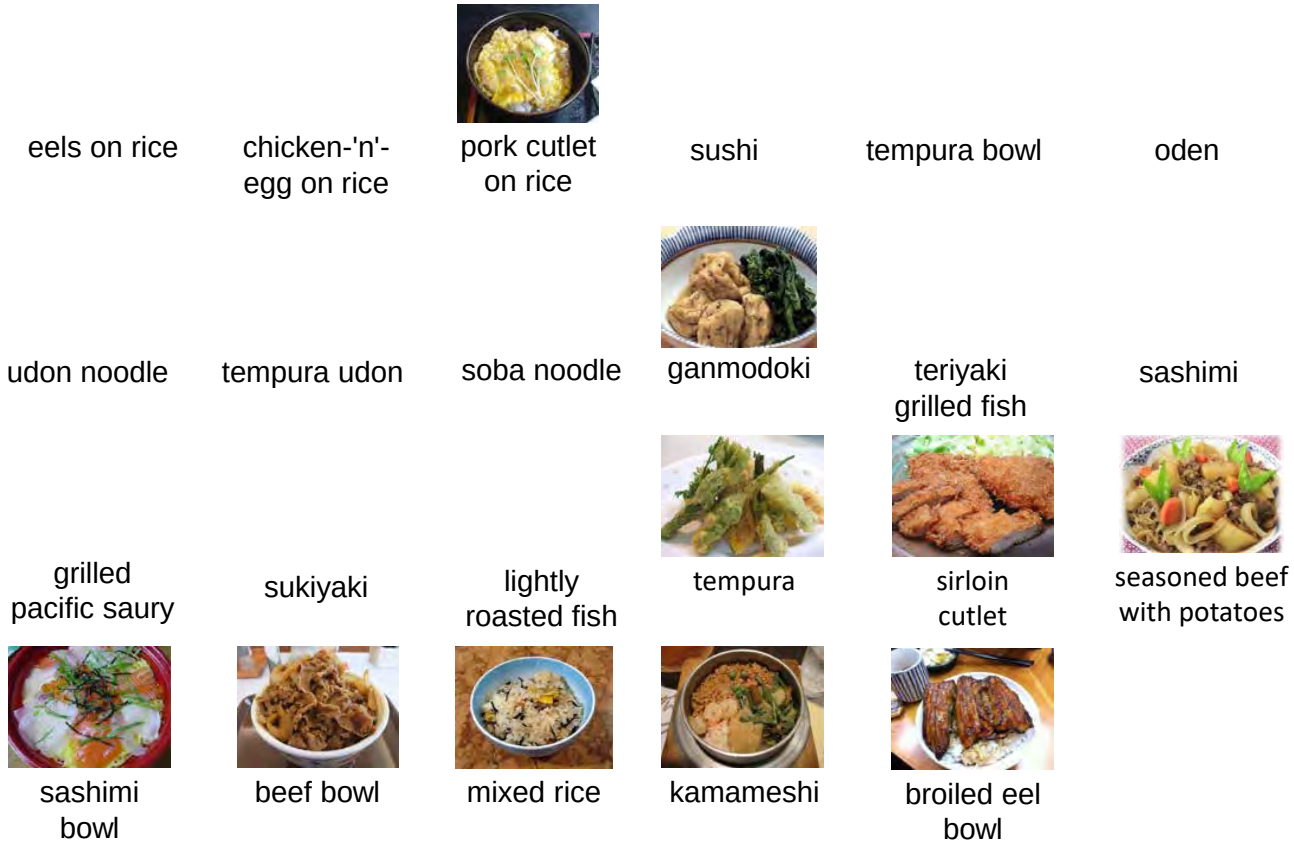
$\alpha > \beta_i$

NU FOOD 360x10 [1] NU FOOD 360x10 10 36 0 1

2.1.1

10

1 NU FOOD 360x10: https://www.cs.is.i.nagoya-u.ac.jp/opensource/nufood/



8: UEC FOOD 256 [8] 23

Fig. 8 Twenty three Japanese food categories selected from the UEC FOOD 256 dataset [8] for attractiveness estimation.

9: UEC FOOD 256 [8] 1 5

Fig. 9 Example of attractiveness values for Japanese food images extracted from the UEC FOOD 256 dataset [8] (1: Low to 5: High).

[6] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," Proc. 2009 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, pp.248–255, July 2009.

[7] R.R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: Visual explanations from deep networks via gradient-based localization," Proc. 16th IEEE Int'l Conf. on Computer Vision, pp.618–626, Aug. 2017.

[8] Y. Kawano and K. Yanai, "Automatic expansion of a food image dataset leveraging existing categories with domain adaptation," Proc. ECCV 2014 Workshop on Transferring and Adapting Source Knowledge in Computer Vision, pp.585–597, Apr. 2014.