

料理の種類に応じた盛り付けに関する典型度評価

中村 真務[†] 川西 康友[†] 道満 恵介^{†,†} 平山 高嗣^{††} 井手 一郎^{†††,†}
出口 大輔[†] 村瀬 洋[†]

[†] 名古屋大学 大学院情報学研究科 〒464-8601 愛知県名古屋市千種区不老町

^{††} 中京大学 工学部 〒470-0393 愛知県豊田市貝津町床立 101

^{†††} 名古屋大学 未来社会創造機構 〒464-8601 愛知県名古屋市千種区不老町

^{††††} 名古屋大学 数理・データ科学教育研究センター 〒464-8601 愛知県名古屋市千種区不老町

E-mail: [†] nakamuram@murase.is.i.nagoya-u.ac.jp, {kawanishi,ide,murase}@i.nagoya-u.ac.jp

[†] ddeguchi@nagoya-u.jp, ^{††} kdoman@sist.chukyo-u.ac.jp, ^{†††} takatsugu.hirayama@nagoya-u.jp

あらまし 本報告では、料理レシピ検索支援を目的として、料理の種類に応じて盛り付けの典型度を評価する手法を提案する。典型的な盛り付けとは、その料理の種類において頻出する特徴を備えている盛り付けである。そこで本研究では、典型的な盛り付けは出現頻度が高いことと、同一の盛り付け同士では似た画像特徴をもつことを考慮して定量的に典型度を推定する。提案手法ではまず、料理写真の背景領域を除去し、色特徴とテクスチャ特徴を抽出する。次に、特徴量を正規化し、カーネル密度推定によって密度分布を生成する。そして、ある料理写真が与えられたとき、その料理写真の特徴量と密度分布から典型度を計算する。評価実験ではカーネル密度推定を用いた典型度評価手法により、色特徴の有効性を確認した。

キーワード 典型度評価, 料理写真, カーネル密度推定

Typicality evaluation of food-type specific presentation

Masamu NAKAMURA[†], Yasutomo KAWANISHI[†], Keisuke DOMAN^{††,†}, Takatsugu

HIRAYAMA^{†††}, Ichiro IDE^{††††,†}, Daisuke DEGUCHI[†], and Hiroshi MURASE[†]

[†] Graduate School of Informatics, Nagoya University

Furo-cho, Chikusa-ku, Nagoya-shi, Aichi, 464-8601 Japan

^{††} School of Engineering, Chukyo University

101 Tokodachi, Kaizu-cho, Toyota-shi, Aichi, 470-0393 Japan

^{†††} Institutes of Innovation for Future Society, Nagoya University

^{††††} Mathematical and Data Science Center, Nagoya University

Furo-cho, Chikusa-ku, Nagoya-shi, Aichi, 464-8601 Japan

E-mail: [†] nakamuram@murase.is.i.nagoya-u.ac.jp, {kawanishi,ide,murase}@i.nagoya-u.ac.jp

[†] ddeguchi@nagoya-u.jp, ^{††} kdoman@sist.chukyo-u.ac.jp, ^{†††} takatsugu.hirayama@nagoya-u.jp

Abstract We propose a method to evaluate the typicality of a food presentation within its food-type for supporting cooking recipe search. A typical food presentation is a food presentation which shares common features of a food-type. Therefore, in this research, we evaluate the typicality of food presentations by considering that a typical food presentation has a high frequency of appearance and that similar food presentations have similar image features. The proposed method first removes the background region and extracts image features from food photos. Next, kernel density estimation generates the density distribution on the normalized features. Finally, given a food photo, the typicality of its food presentation is calculated by the features of the food photo and the density distribution. Evaluation results showed the effectiveness of using color features for the typicality evaluation.

Key words Typicality evaluation, food photo, kernel density estimation

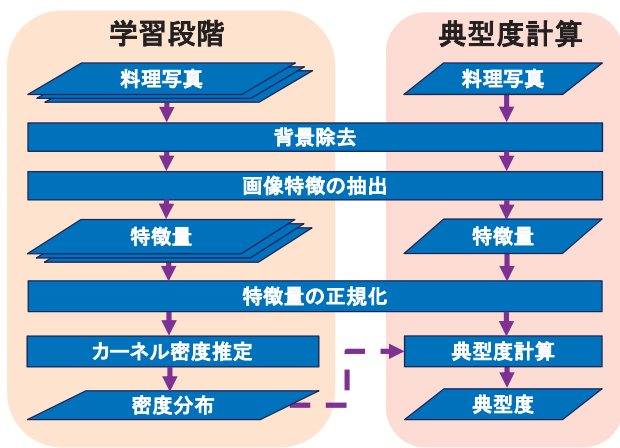


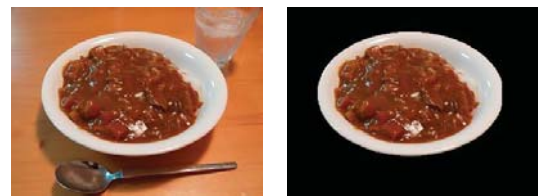
図1 提案手法の処理手順.

1. はじめに

近年、インターネット上の Consumer Generated Media の普及により投稿型料理レシピサイトが増加している。例えば、2020年2月現在、クックパッド^(注1)では約320万件、楽天レシピ^(注2)では約180万件の料理レシピが投稿されており、多くの利用者がいる。利用者は、料理の作り方を調べることや、参考にすることを目的として料理レシピサイトを利用する。その際に料理レシピ検索では、料理の完成図である料理写真のような直感的にイメージを捉えやすい情報を参考にすることが多い。また、料理の完成図そのものを参考にするために利用することも多い。しかし、膨大な検索候補から直感に合った料理レシピを探すのは困難であるため、検索候補の絞り込みが必要である。

料理レシピ検索に関しては、様々な研究が行われている[1]。横井らは、食材の組み合わせに関する料理レシピの典型度を分析し、適切な典型度の料理レシピを推薦する手法を提案した[2]。この手法では、料理レシピの食材リストを数値化した特徴ベクトルを用いて、料理の種類で分類された料理カテゴリごとに部分空間を構築することで典型度を計算する。この部分空間はその料理カテゴリの典型的な食材の組み合わせを表現することができる。この研究では画像特徴を用いていないが、上記のように料理の見えを考慮するためには、画像特徴を用いた手法について検討する必要がある。

一方、料理の見た目を定量的に評価する研究も行われている。高橋らは、与えられた料理の向きと主食材の見えを考慮した画像特徴を用いて、料理画像の魅力度を推定する手法を提案している[3]。また、佐藤らは、あらかじめ大規模データセットで学習されたCNNモデルに対して料理画像を用いて転移学習することで、End-to-Endに与えられた料理画像の撮影角度に関する魅力度を推定する手法を提案している[4]。これらのように、料理に対する撮影角度を考慮した料理写真の魅力度を評価する手法がいくつか提案されてきたが、そもそも料理の魅力的な盛り付けに関しては考慮していない。



(a) 処理前 (b) 処理後

図2 GrabCutによる前景領域抽出例.

以上の背景から、本報告では、料理の盛り付けとして典型的なものと新奇性があるものが魅力的であると考え、料理の種類ごとに盛り付けの典型度を料理画像から定量的に評価する手法を提案する。提案手法では、料理の種類ごとに画像特徴を抽出し、その料理の盛り付けを分析することで典型度を評価する。具体的には、典型的な盛り付けは出現頻度が高いことと、同一の盛り付け同士では似た画像特徴をもつことに着目して定量的に典型度を推定する。

2. 料理写真に基づく盛り付けの典型度評価手法

本節では、料理写真の画像特徴を用いて、料理の種類に応じた盛り付けの典型度を計算する。図1に本論文の提案手法の手順を示す。以下、各処理について説明する。

2.1 背景除去

料理の盛り付けの典型度を評価するため、盛り付けと直接関係しない料理と皿以外の領域である背景領域を除去する。そこで、入力された料理写真から GrabCut [5] により人手で前景領域を抽出する。この手法は、前景として抽出したい領域を大まかに囲む矩形を与えることにより、矩形の中央に存在する前景とその周辺の背景を分離する領域分割手法である。図2に背景除去の例を示す。本研究では、料理と皿の領域を「料理領域」と定義し、料理領域から画像特徴を抽出する。また、画像特徴を抽出する際には、GrabCutで指定した矩形の大きさに切り抜いた画像を用いる。

2.2 画像特徴の抽出

料理写真から色特徴とテクスチャ特徴を抽出する。以下に各画像特徴について紹介する。

HSV色ヒストグラム

HSV色空間がRGB色空間に比べて人間の感覚に近い色空間であることから、色特徴の1つとしてHSV色ヒストグラムを用いる。HSV色ヒストグラムは、HSV色空間で表現された画像の各画素値である色相(Hue)と彩度(Saturation)と明度(Value)のうち、色相と彩度の出現頻度を表したものである。

HSV色ヒストグラムの抽出方法は以下の通りである。まず、入力された料理写真をHSV色空間の画像に変換する。次に、式1に示すように、明度(Value)を除いて色相 $h_{i,j}$ と彩度 $s_{i,j}$ で表した画素値 $c_{i,j}$ に変換する。

$$c_{i,j} = (h_i, s_j) \quad (0 \leq i \leq m, 0 \leq j \leq n) \quad (1)$$

また、次元削減のために画素値を量子化する。最後に、ヒスト

(注1): <http://cookpad.com/>

(注2): <http://recipe.rakuten.co.jp/>

グラムの各値を料理領域の画素数で除算することで正規化する．以上から得られたヒストグラムを HS 色ヒストグラムとする．

HS 色コリログラム

HS 色ヒストグラムと同様な理由で色特徴の 1 つとして HS 色コリログラムを用いる．色コリログラムは, Jing らによって提案された色同士の空間的な相関を表す画像特徴量である [6]．

HS 色コリログラムの抽出方法は以下の通りである．まず入力された料理写真を $L \times L$ 画素に拡張した HSV 色空間の画像に変換する．次に, HS 色ヒストグラムの場合と同様に式 (1) に示すように, 色相 $h_{i,j}$ と彩度 $s_{i,j}$ で表した画素値 $c_{i,j}$ に変換する．そして量子化された各画素値 $c_{i,j}$ の画素 p_1 に対して, 無限大ノルムで表された距離 k にある画素 p_2 の画素値が $c_{i',j'}$ である確率を HS 色コリログラムとする．

SIFT + Bag-of-Features

テキストチャ特徴の 1 つとして SIFT 特徴量を用いた Bag-of-Features (BoF) 表現を用いる．SIFT (Scale-Invariant Features Transform) は, Lowe らが提案したアルゴリズムで, 画像の局所特徴量を抽出する手法である [7]．また, BoF は, Csuska らによって提案された画像の局所特徴量をヒストグラムで表現した画像特徴量である [8]．SIFT 特徴量は照明変化や回転, 拡大縮小に頑健であるため, 撮影環境や撮影角度の影響が小さいと考えられるため, 本研究では, BoF の局所特徴量に SIFT 特徴量を用いる．

BoF を抽出する際は, コードブックの作成とヒストグラムの作成の 2 段階の工程がある．コードブックの作成では, まず学習段階で入力された全ての料理写真から局所特徴量を抽出する．次に, k -means 法を用いて局所特徴量のクラスタリングをし, 各クラスタの中心であるプロトタイプをコードブックに記述する．ヒストグラムの作成では, 特徴量を抽出する料理写真の全ての局所特徴量を, プロトタイプに基づいて最近傍法で分類する．そして各プロトタイプに分類された局所特徴量の数からヒストグラムを作成する．

Histograms of Oriented Gradients

画像空間全体の大まかな勾配特徴を抽出することができるため, テキスチャ特徴の 1 つとして HOG 特徴量を用いる．HOG (Histograms of Oriented Gradients) は, Dalal らが提案した画像特徴量で, 輝度勾配をヒストグラムで表したものである [9]．

HOG 特徴量の抽出方法は以下の通りである．まず, 式 2 と式 3 から, 勾配強度 $m(x, y)$ と勾配方向 $\theta(x, y)$ を計算する．ただし, 勾配方向 $\theta(x, y)$ はヒストグラムのビンとして用いるために量子化する． $L(x, y)$ は料理写真を輝度画像に変換したときの各画素値である．

$$m(x, y) = \sqrt{f_x(x, y)^2 + f_y(x, y)^2} \quad (2)$$

$$\theta(x, y) = \begin{cases} \frac{\pi}{2} & (f_y(x, y) = 0) \\ \arctan \frac{f_x(x, y)}{f_y(x, y)} & (\text{otherwise}) \end{cases} \quad (3)$$

$$f_x(x, y) = L(x, y) - L(x + 1, y) \quad (4)$$

$$f_y(x, y) = L(x, y) - L(x, y + 1) \quad (5)$$

次に, 画像を等間隔に分割した部分領域であるセルごとに勾配方向のヒストグラムを作成する．このヒストグラムを作成する際の投票には, 勾配強度で重み付けする．そして, 注目したセルの周辺の $n \times n$ 個のセルの集まりをブロックとして, ブロック内のセルのヒストグラムを連結し, ブロック内で正規化する．ブロック内の k 番目のセルのヒストグラムを $h_c(k)$, それらを連結させたヒストグラムを h_b とおくと, 正規化の計算は式 6 と式 7 となる．ただし, $\epsilon (= 1)$ はゼロ除算を防ぐための定数である．

$$h'_b = \frac{h_b}{H} \quad (6)$$

$$H = \sqrt{\sum_{k=1}^{n \times n} \|h_c(k)\|^2 + \epsilon} \quad (7)$$

最後に, ブロックごとのヒストグラムを全て連結したものを HOG 特徴量とする．

2.3 特徴量の正規化

抽出した特徴量の正規化について説明する．まず, 料理の種類ごとに抽出した N 個の特徴量 $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}$ から分散共分散行列 Σ を式 8 により計算する．次に, 分散共分散行列 Σ の固有値問題として式 9 から固有値 λ_i とそれに対応した固有ベクトル \mathbf{e}_i ($i = 1, 2, \dots, N$) を計算する．ただし, μ は \mathbf{X} の平均ベクトルである．

$$\Sigma = \frac{1}{N} \sum_{k=1}^N (\mathbf{x}_k - \mu)(\mathbf{x}_k - \mu)^t \quad (8)$$

$$\Sigma \mathbf{e} = \lambda \mathbf{e} \quad (9)$$

このようにして得られる固有ベクトルのうち, 対応する固有値が大きい固有ベクトルは, その料理の種類において見えの違いをよく表現している成分である．つまり, 対応する固有値が大きい固有ベクトルで構築された部分空間上で, 類似した特徴をもつ料理同士は, 似た盛り付けであると考えられる．よって, 対応する固有値が大きい固有ベクトルを基底ベクトルとすることで, その料理の種類における盛り付けの違いをよく表現した部分空間を構築することができる．最後に, 部分空間に対して白色化の処理し, 各主成分の分散を均一にする．

2.4 カーネル密度推定

カーネル密度推定による密度分布の生成方法について説明する．まず, 学習データ内の N 枚の料理写真から特徴量 $\mathbf{X}^* = \{\mathbf{x}_1^*, \mathbf{x}_2^*, \dots, \mathbf{x}_N^*\}$ を抽出する．次に, \mathbf{X} を構築した部分空間に射影し, 射影 $\mathbf{X}^* = \{\mathbf{x}_1^*, \mathbf{x}_2^*, \dots, \mathbf{x}_N^*\}$ を得る．そして, カーネル密度推定 (式 10) を用いて, ある入力の射影 \mathbf{x}^* に対する密度分布 $\sigma(\mathbf{x}^*)$ を計算する．ただし, $K(\mathbf{x})$ をカーネル関数, h をバンド幅とする．

$$\sigma(\mathbf{x}^*) = \frac{1}{N} \sum_{k=1}^N K\left(\frac{\mathbf{x}^* - \mathbf{x}_k^*}{h}\right) \quad (10)$$

2.5 典型度の計算

典型度の計算方法について説明する．まず, 入力された料理

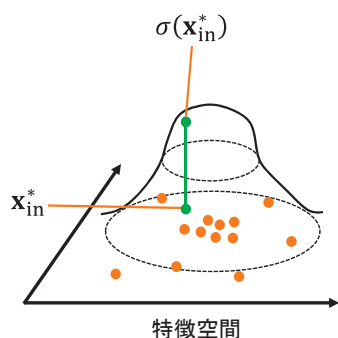


図 3 カーネル密度推定による典型度の計算.

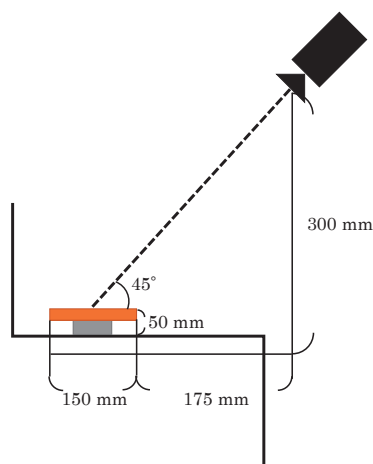


図 4 撮影環境の略図.

写真から特徴量 x_{in} を抽出する．次に， x_{in} を構築した部分空間に射影し，射影 x_{in}^* を得る．そして，図 3 に示すように， x_{in}^* を密度分布 $\sigma(x^*)$ に入力することで，典型度である $\sigma(x_{in}^*)$ を計算する．

3. 評価実験

本節では，提案手法の有効性を確認するための評価実験について報告する．はじめに，評価実験のために作成したデータセットについて述べる．次に，5種類の盛りつけパターンのカレーライスについて提案手法が有効であるかを評価する．

3.1 評価用データセットの作成

評価実験に用いるデータセットを作成するために，5種類の盛りつけパターンのカレーライスを撮影した．照明条件を一定にして撮影角度を俯角 45° に固定し，カレーライスを 12° 刻み回転させて撮影した．図 4 に撮影環境の略図を示す．また，図 5 に実際に撮影した 5種類の盛りつけパターンの例を示す．

次に，撮影したカレーライスについて，典型度に基づいて盛りつけパターンを順位付けするために被験者実験を行った．被験者実験では，図 5 に示す 5種類のカレーライスを提示して，「より一般的であると思う順に高い順位を付ける」ように指示した．10代 1名，20代 26名の男女のべ 27名の被験者実験から回答を集めた．その集計結果を表 1 に示す．図 5 の 5種類の

表 1 各カレーライスの盛りつけパターンの順位毎の得票数．

盛りつけパターン	1位	2位	3位	4位	5位
パターン 1	26	1	0	0	0
パターン 2	0	5	16	3	3
パターン 3	0	0	7	18	2
パターン 4	1	21	2	2	1
パターン 5	0	0	2	4	21

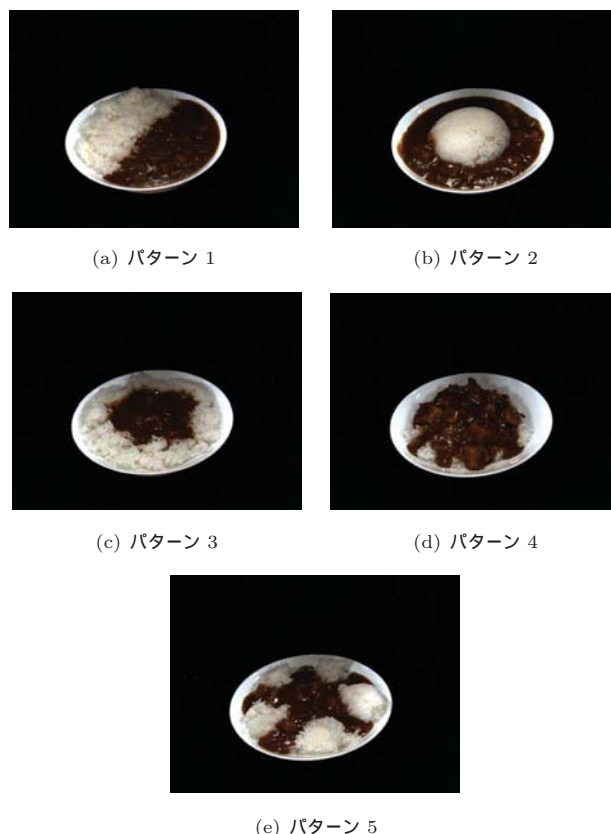


図 5 撮影したカレーライスの盛りつけパターン．

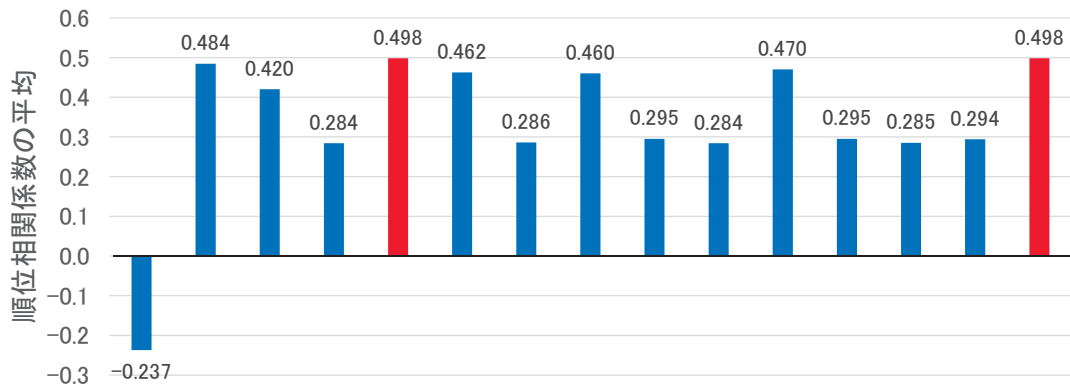
盛りつけパターンは，1，4，2，3，5の順に一般的な盛りつけであるという結果になった．本研究では，この被験者実験で得られた順位を盛りつけの典型度の順位として，以降の実験で評価に用いる

3.2 実験方法

本項では，撮影方向の影響に対する頑健性と盛りつけの違いに対する典型度の推定について，提案手法の有効性を確認する．学習データには料理画像公開データセットである UEC-FOOD256 [10] のカテゴリ「beef curry」内の 252 枚の画像を用いた．そして，評価用データセット中の各盛りつけパターンから抽出した 5 枚の写真の典型度と，被験者実験の結果との順位相関を全ての組合せである 24,300,000 通り ($= 30^5$) で計算してその平均を計算した．順位相関を表す評価指標には Spearman の順位相関係数を用いた．

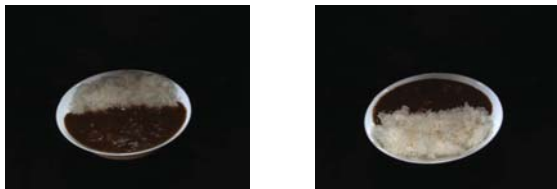
3.3 実験結果

図 6 に，特徴の組合せ毎の推定結果による Spearman の順位相関係数の平均を示す．HS 色ヒストグラムと HS 色コリログラムを組み合わせた特徴と，全特徴を組み合わせた特徴の



ヒストグラム	●				●	●	●				●	●	●		●
コリログラム		●			●			●	●		●	●		●	●
SIFT+BoF			●			●		●		●	●		●	●	●
HOG				●			●		●	●		●	●	●	●

図 6 使用特徴の組合せ毎の Spearman 順位相関係数の平均。



(a) ルーの領域が大きい撮影方向 (b) ライスの領域が大きい撮影方向

図 7 撮影方向による見え方の違い。

Spearman の順位相関係数の平均が最も高い結果になった。以上から、色特徴の組合せが有効であることが確認された。

3.4 考察

本実験の結果では、回転不変であるにも関わらず、HS 色ヒストグラムによる推定結果は負の相関になった。これは、被写体の回転によってカレーライスが見え方が変化したためであると考えられる。具体的には、図 7 に示すように、回転によりライスとルーの位置関係が変化し、ライスが見える領域が大きい撮影方向とルーが見える領域が大きい撮影方向があるからである。以上のことから、撮影角度や撮影方向の影響を受けにくい特徴を用いたり、影響を軽減するための手法を考える必要がある。

4. まとめ

本報告では、典型的な盛り付けは出現頻度が高いことと、同一の盛り付け同士では似た画像特徴をもつことに着目して定量的に典型度を推定する手法を提案した。具体的には、料理写真の背景除去処理をした後、色特徴とテキスト特徴を抽出し、その特徴の正規化をしてカーネル密度推定によって密度分布を生成した。そして、ある料理写真が与えられた時に、密度分布に入力することで典型度を算出した。評価実験では、色特徴の組み合わせによる有効性を確認した。今後はより複雑な盛り付

けの料理への適用や学習用データセットの拡張を検討していく必要がある。

謝 辞

本研究の一部は、科学研究費補助金による。

文 献

- [1] 高橋哲朗, 井手一郎, “レシピ・献立検索 (特集: 食とコンピューティング),” 情報処理, Vol. 52, No. 11, pp. 1376–1381, Oct. 2011.
- [2] S. Yokoi, K. Doman, T. Hirayama, I. Ide, D. Deguchi, and H. Murase, “Typicality analysis of the combination of ingredients in a cooking recipe for assisting the arrangement of ingredients,” Proc. 2015 IEEE Int. Conf. on Multimedia and Expo, No. h75, pp. 1–6, July 2015.
- [3] K. Takahashi, T. Hattori, K. Doman, Y. Kawanishi, T. Hirayama, I. Ide, D. Deguchi, and H. Murase, “Estimation of the attractiveness of food photography based on image features,” IEICE Trans. Info & Sys., vol. E102-D, no. 8, pp. 1590–1593, Aug. 2019.
- [4] 佐藤陽昇, 道満恵介, 平山高嗣, 井手一郎, 川西康友, 出口大輔, 村瀬 洋, “畳み込みニューラルネットワークを用いた料理写真の魅力度推定,” 信学技報, MVE2017-32, Oct. 2017.
- [5] C. Rother, V. Kolmogorov and A. Blake, “GrabCut: Interactive foreground extraction using iterated graph cuts,” ACM Trans. Graph., vol.23, no.3, pp. 309–314, Aug. 2004.
- [6] J. Huang, S Ravi Kumar, M. Mitra, W. Jing, and Z. Zabih, “Image indexing using color correlogram,” Proc. 1997 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, pp.762–768, June 1997.
- [7] D.G. Lowe, “Object recognition from local scale-invariant features”, Proc. 7th IEEE Int. Conf. on Computer Vision, pp.1150–1157, Sept. 1999.
- [8] G. Csurka, C. Bray, C. Dance and L. Fan, “Visual categorization with bags of keypoints,” Proc. ECCV2004 Workshop on Statistical Learning in Computer Vision, pp.59–74, May 2004.
- [9] N. Dalal and W. Triggs, “Histograms of oriented gradients

for human detection,” Proc. 2005 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, pp.886–893, June 2005.

- [10] Y. Kawano and K. Yanai, “Automatic expansion of a food image dataset leveraging existing categories with domain adaptation,” Proc. ECCV2014 Workshop on Transferring and Adapting Source Knowledge in Computer Vision, pp. 3–17, Sept. 2014.