

COLORSIFT とアンサンブル学習を用いた単品食事画像の識別

Recognition of Single-Food Images Using Ensemble Learning and COLORSIFT

梶原 祐輔†

Yusuke KAJIWARA

中村 宗広‡

Munehiro NAKAMURA

木村 春彦‡

Haruhiko KIMURA

1. 序論

現在、日本は65歳以上の高齢者が総人口の23.3%を占める超高齢社会である[1][2]。また、この高齢者の割合は年々増加傾向にあり、2050年には3人に1人が高齢者になると推計されている[1]。そのため、今後益々高齢者向け福祉システムの重要性は高まっていくと予想される。

一方、高齢者は高齢になるに従い、健康状態に不安を抱く傾向にあることが報告されている[1]。これは、視覚、聴覚、触覚等の感覚の鈍化や体力の低下が心理状態に影響していることが理由の一つに挙げられる[3]。また、一人暮らしでかつ地域社会との接点が少ない高齢者（以下、独居老人）が罹患した場合、独力で対処することは難しく、また近親者や近所の住人も独居老人の身体の異常に気づき難いため、最悪『住居内で誰にも看取られることなく死亡する』（以下、孤独死）場合もある[4]。独居老人の数と孤独死の数は年々増加しており、対策が急務となっている[1][4]。また、独居老人は罹患した際の事後の対応が困難であるため、日々の健康管理が重要となる。高齢者の健康管理を行う上で重要な項目として栄養管理がある。高齢者は咀嚼力、嚥下力、腸蠕動力、味覚が低下しているため、若年者と比較して栄養が偏りやすく、必要なタンパク質やエネルギーが不足する低栄養に陥ることが多い。この低栄養は老年症候群の一つであり、低栄養が長期に渡ると筋力の低下、体脂肪の低下、免疫力の低下等を引き起こし、寝たきりや要介護に結びつくことも報告されている[5][6]。

そこで本研究では独居老人の健康を維持する上で重要な栄養管理をアシストするシステムを開発する。老人ホームや老人福祉施設等では管理栄養士が栄養を管理するケースが多いが、独居老人の場合、栄養管理を自ら行う必要がある。しかし、独居老人が管理栄養士と同等の知識、経験を持っていることは少ない。また、訪問介護サービスなどを活用することも考えられるがコストが高く、介護職員も慢性的に不足している[7]。

この問題を解決する方法として携帯電話で撮影した食事画像を管理栄養士に送信し、健康を維持するために必要なレシピを推薦してもらうシステムが提案されている[8]。このシステムは訪問介護サービス等の利用と比較して安価だが、大量のユーザが利用した場合、サービスの遅延や質の低下等が起こることが予想される。今後高齢者は年々増加すると予想されているため[2]、栄養管理システムを自動化する必要がある。

本研究ではユーザが食事を撮像すると自動的に食事の種類を識別し、識別した食事の栄養を把握できるシステムを開発する。本稿ではこのシステムの前段階として、食事画像から食事の種類を識別するシステムを開発する。

†立命館大学情報理工学部

‡金沢大学自然科学研究科

食事画像から食事の種類を識別するシステムとして様々な単品の食事が撮像されている単品食事画像から食事の種類を識別する手法[9]-[12]や複数の食事が撮像されている食事画像からそれぞれの食事の種類を識別する手法[13]が提案されているが、複数品目を同時に識別することは疎か単品食事の識別も未だ十分な精度に至っていない。

そこで本稿はまず単品食事画像の識別精度を向上させる。従来研究[12]では複数のカーネルを統合した統合カーネルを用いて識別を行う MKL-SVM を用いて識別を行ない、識別精度の向上を図っている。本稿では複数のカーネルだけでなく、Random Forest や Adaboost 等の識別器も併用し、識別精度の向上を図る。また、色情報を多く含む領域、彩度情報が多く含む領域、明度情報を多く含む領域を重点的に学習し、識別精度の向上を図る。

第2章では食事画像から特徴量を抽出する手法や本稿で用いた学習手法について説明する。第3章でインターネットから取得した 50 種類の食事画像を提案手法で識別し、識別した結果と MKL-SVM 等の識別器を用いた場合の識別結果と比較する。第4章で結論と今後の展望を述べる。

2. システムの概要

ユーザが携帯やカメラ等を用いて食事を撮像し、撮像した食事画像から自動的に食事を識別するシステムを開発する。まず、食事画像の特徴量を抽出し、その後、アンサンブル学習を行い、食事を識別する。特徴量には、従来研究[12][13]で用いられている Scale-Invariant Feature Transform(以下、SIFT)[14]、COLORSIFT[15]、色情報、Histograms of Oriented Gradients (以下、HOG) [16]、ガボール特徴[17]を用いた。また、SIFT、COLORSIFT、色情報はそれぞれ Bag of Features (以下、BoF) [18]によるベクトル量子化を行い、その後空間ピラミッド表現[19]を適用し、位置情報を付加する。また、従来研究[12]では複数のカーネルを統合させ、一つの識別器を構築する MKL-SVM を用いていたが、本研究では複数のカーネルを持つ SVM だけでなく、Random Forest や Adaboost、ニューラルネットワーク等様々な識別器を併用し、識別精度の向上を図る。

2.1 特徴量の抽出

2.1.1 SIFT

SIFT はキーポイントを検出しキーポイント周辺の特徴量を抽出する。以下に SIFT を用いた特徴量の抽出手順を示す。

1. カラー画像のグレースケール化を行う。
2. グレースケール画像に対して異なるスケールのガウスフィルタを適用した画像の差分 Difference of Gaussian (以下、DoG) 画像を用いてキーポイントを検出する。DoG の極値の最大値を求め、濃淡情報を多く含む領域を抽出する。
3. DOG の二次微分であるヘッセ行列 H を用いてエッジ上のキーポイントを削除する。 α は閾値である。

$$\frac{\text{tr}(H)^2}{\det(H)} < \frac{(\alpha+1)^2}{\alpha} \quad (1)$$

4. 位置 (x,y) とスケール (σ) に 2 次関数をフィッティングさせ、サブピクセル推定を行う。DoG 関数 D をテイラー展開し、その後 x,y,σ について偏導関数を求め、式を変形すると (2) 式が導出できる。

$$\begin{bmatrix} \bar{x} \\ \bar{y} \\ \bar{\sigma} \end{bmatrix} = - \begin{bmatrix} \frac{\partial^2 D}{\partial x^2} & \frac{\partial^2 D}{\partial x \partial y} & \frac{\partial^2 D}{\partial x \partial \sigma} \\ \frac{\partial^2 D}{\partial x \partial y} & \frac{\partial^2 D}{\partial y^2} & \frac{\partial^2 D}{\partial y \partial \sigma} \\ \frac{\partial^2 D}{\partial x \partial \sigma} & \frac{\partial^2 D}{\partial y \partial \sigma} & \frac{\partial^2 D}{\partial \sigma^2} \end{bmatrix}^{-1} \begin{bmatrix} \frac{\partial D}{\partial x} \\ \frac{\partial D}{\partial y} \\ \frac{\partial D}{\partial \sigma} \end{bmatrix} \quad (2)$$

$\bar{x}$, $\bar{y}$ はサブピクセル位置と $\bar{\sigma}$ は補正後のスケールである。(2) 式を適用し、キーポイントの位置とスケールの誤差を軽減する。

5. 補正後の DoG 関数 $D(\bar{x}, \bar{y}, \bar{\sigma})$ に対して、閾値 β 以下のキーポイントを削除する。コントラストの低いキーポイントを削除し、ノイズに対して頑強にする。
6. オリエンテーションを計算する。まず、輝度値 f_{xy} からキーポイント毎に勾配強度 $q(x,y)$ と勾配方向 $a(x,y)$ を求める。 x,y は画像上のピクセル位置である。

$$q(i, j) = \sqrt{q_x(x, y)^2 + q_y(x, y)^2} \quad (3)$$

$$a(i, j) = \tan^{-1} \frac{q_y(x, y)}{q_x(x, y)} \quad (4)$$

$$q_x(i, j) = f_{x+1, y} - f_{x-1, y} \quad (5)$$

$$q_y(i, j) = f_{x, y+1} - f_{x, y-1} \quad (6)$$

(4) 式で導出した $a(x,y)$ を 36 方向に量子化し、勾配方向毎に重み付けを行い、重み付き勾配方向ヒストグラム $h(\hat{a})$ を求める。

$$h(\hat{a}) = \sum_x \sum_y G(x, y, \sigma) q(x, y) \delta[\hat{a}, a(x, y)] \quad (7)$$

$G(x, y, \sigma)$ はガウス関数、 δ はクロネッカーのデルタ関数で量子化した $a(x,y)$ と $\hat{a}$ が一致した場合に 1 を出力し、それ以外は 0 となる。この $h(\hat{a})$ の最大値の 80% となる勾配方向をオリエンテーションとする。

7. キーポイントを中心に 3σ のガウス窓を抽出し、 4×4 のブロックに分割する。そして、オリエンテーションを元に 3σ のガウス窓を回転させ、ブロック毎に 8 方向に量子化した勾配ヒストグラムを算出する。
8. 導出した 128 次元の特徴量の総和で正規化を行う。

SIFT 特徴量は、DoG 画像の極値を用いているためスケールの変化に頑強である。また、オリエンテーションを元に回転させるため角度の変化にも強く、8. で 128 次元の特徴量を正規化することで照度の変化にも頑強である。

2.1.2 COLORSIFT

SIFT はグレースケール画像に対してキーポイントの検出、特徴量の抽出を行ったが、COLORSIFT では色相 H 、彩度 S 、明度 V それぞれに対して SIFT 特徴量を求める。RGB から HSV の変換は式 (8) ~ (10) を用いる。 MAX_{rgb} は RGB の最大値、 MIN_{rgb} は RGB の最小値である。また、 H が負の場合、360 を足す。SIFT では濃淡情報を多く含むキーポイントを検出していたが、COLORSIFT では色情報を多く含むキーポイント、彩度情報を多く含むキーポイント、明度情報を多く含むキーポイントを検出し、そのキーポイントの特徴量を抽出する。

$$H = \begin{cases} 60 \left(\frac{G-B}{MAX_{rgb} - MIN_{rgb}} \right) & \text{if } MAX = R \\ 60 \left(\frac{B-R}{MAX_{rgb} - MIN_{rgb}} \right) & \text{if } MAX = G \\ 60 \left(\frac{R-G}{MAX_{rgb} - MIN_{rgb}} \right) & \text{if } MAX = B \end{cases} \quad (8)$$

$$S = \frac{MAX_{rgb} - MIN_{rgb}}{MAX_{rgb}} \quad (9)$$

$$V = MAX \quad (10)$$

2.1.3 色情報, BoF と空間ピラミッド表現

BoF を用いて RGB ヒストグラムを作成する。BoF の手順を以下に示す。

1. カラー画像から画素毎に RGB ベクトルを抽出する。
2. RGB ベクトル群を K-means++ 法 [20] を用いてクラスタリングし、各クラスタ中心 Visual Words を求める。
3. Visual Words と RGB ベクトルを比較し、ユークリッド距離が最も近い Visual words に投票を行い、頻出ヒストグラムを作成する。

次に空間ピラミッド表現を行い、位置情報を付加する。空間ピラミッド表現はカラー画像を 1×1 , 2×2 , 3×3 のブロックに分割し、ブロック毎に頻出ヒストグラムを構築する。また、RGB の他に前述の SIFT や COLORSIFT にもこの BoF、空間ピラミッド表現を適用する。本稿では Visual Words の数を 50 とし、 3×3 までの空間ピラミッド表現を用いる。

2.1.4 HOG

HOG 特徴量は線形的な輝度変化に強く輝度に変化する画像に対しても大まかな形状を抽出できる。HOG はまず、グレースケール画像を $C_w \times C_h$ 個のセルに分割する。次にセル毎に式 (3) ~ (6) を適用し、 0° から 180° までの勾配方向 $a(i,j)$ を 9 方向に量子化し、勾配ヒストグラム $h_{hog}(\hat{a})$ を算出する。

$$h_{hog}(\hat{a}) = \sum_x \sum_y q(x, y) \delta[\hat{a}, a(x, y)] \quad (11)$$

最後に式 (12) を用いて $h_{hog}(\hat{a})$ の正規化を行う。ブロックには、セルが $L_w \times L_h$ 個含まれている。

$$\hat{h}_{hog}(\hat{a}) = \frac{h_{hog}(\hat{a})}{\sqrt{\left(\sum_{k=1}^{L_w \times L_h} h_{hog}(k) \right)^2 + \epsilon}} \quad (12)$$

式 (12) の分母は正規化項、 $\epsilon (=1)$ は補正項である。この正規化で照度の線形的な変化に頑強になる。また、HOG 特徴量の次元数 $\dim$ は式 (13) で求める。

$$\dim = (C_w - L_w + 1)(C_h - L_h + 1)L_w L_h \times \theta \quad (13)$$

本研究では画像中に含まれるセルの数は 8×8 、ブロックに含まれるセルの数は 3×3 である。

2.1.5 ガボール特徴量

ガボール特徴量は空間フィルタの一種で局所的な濃淡情報を抽出できる。ガボール特徴量 [13][17] を式 (14) に示す。

$$Gabor_{r,a}(x, y) = \frac{E^{2r}}{\sigma^2} \exp \left\{ -\frac{E^{2r}(x^2 + y^2)}{2\sigma^2} \right\} \times [\exp \{ jE^r (x \cos \frac{n\pi}{A} + y \sin \frac{n\pi}{A}) \} - \exp \{ -\frac{\sigma^2}{2} \}] \quad (14)$$

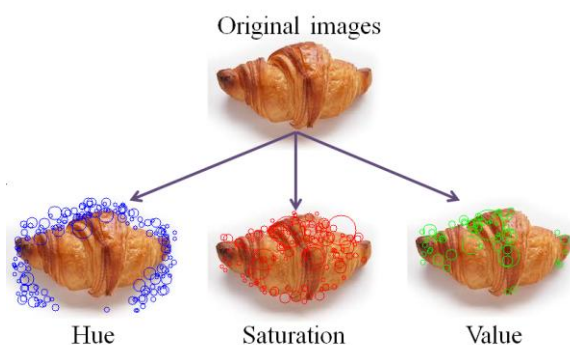


Fig.1 Original images are divided into hue, saturation, and value zone using COLORSIFT.

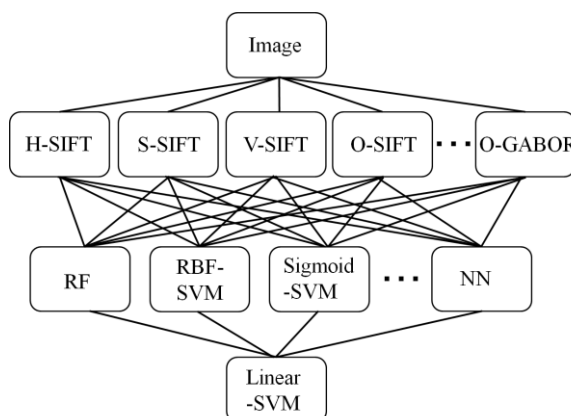


Fig.2 The proposed ensemble learning is constructed to combine support vector machines using multiple kernel and the other machines learning.

r は解像度, n は方向であり, それぞれ $0 \leq r \leq R-1$, $0 \leq n \leq A-1$ である。また, E は拡大率, R は解像度の数, A は方向の数である。本稿では $R=4$, $A=6$ とする。また各フィルタで強度の平均を求め, ヒストグラムで表す。

2.2 アンサンブル学習

アンサンブル学習とは複数の弱識別器に学習させ, 各弱識別器の汎化性能を評価し, 組み合わせることで高い汎化性能を持つ識別器を構成する手法である。一般的なアンサンブル識別器として Adaboost[21], RandomForest[22]等がある。また, 複数のカーネルを線形結合させ, 統合カーネルを構築する Multiple Kernel Learning(以下, MKL)法[23]を用いた MKL-SVM も提案されている。従来研究[12]でも MKL-SVM を用いて識別精度の向上を図っている。本研究ではカーネルが異なる複数の SVM の他に RandomForest やニューラルネットワーク等, n 個の識別器を用いて特徴量から食事を識別し, 識別結果から得られる汎化性能を線形 SVM で評価し, 識別精度の向上を図る。また, COLORSIFT を用いて食事画像に対して色情報を多く含む領域, 彩度情報を多く含む領域, 明度情報を多く含む領域の3つの領域に含まれる特徴量を重点的に学習し, 識別精度の向上を図る。領域の抽出方法を以下に示す

1. COLORSIFTを行い, 色相 H, 彩度 S, 明度 V 毎のキーポイントを検出する。検出例を Fig.1 に示す。色相領域では色情報を多く含む食事の外縁部にキーポイン

Table.1 Classification of 50 food categories

① Broiled eel and rice	② Shrimps with chili sauce	③ Oden
④ Omelet	⑤ Savoury pancake with various ingredients	⑥ Udon
⑦ Pork cutlet on rice	⑧ Curry and rice	⑨ Kimpira gobo
⑩ Gratin	⑪ Croissant	⑫ Corn soup
⑬ Croquette.	⑭ Rice	⑮ Zaru soba
⑯ Sandwiches	⑰ Stew	⑱ Sukiyaki
⑲ Spaghetti	⑳ Rice fried with chicken.	㉑ Fried rice
㉒ Toast	㉓ Hamburger on a bun	㉔ Hamburger
㉕ Pizza	㉖ Bibimbap	㉗ Hot dog
㉘ Potato salad	㉙ Ramen	㉚ Stuffed cabbage
㉛ Sushi	㉜ Yakisoba	㉝ Roast fish
㉞ Rice topped with chicken and eggs	㉟ Sweet and sour pork	㊱ Chawan-mushi
㊲ Fried chicken	㊳ Tempura Udon	㊴ Tianjin rice bowl
㊵ Tendon	㊶ Niku-jyaga	㊷ Natto
㊸ Mabo-tofu	㊹ Miso soup	㊺ Sunny-side up
㊻ Vegetable tempura	㊼ Sauteed vegetables	㊽ Hiyashi chuka
㊾ Cold tofu	㊿ Gyoza	

Table.2 Probability of correct classification by this study and previous study

	AVG	SD
Proposed method A	73.0 %	±11.6
Proposed method B	66.6 %	±13.9
MKL-SVM	63.3 %	±15.8
Random Forest	60.0 %	±12.5
Adaboost	43.4 %	±13.3

トが検出され, 彩度領域, 明度領域は彩度情報と明度情報を多く含む食事内にキーポイントが検出される。

2. これらの領域で検出したキーポイントの中心座標(x, y)の最小値と最大値から $1/10$ を代表値として取得し, 代表値の平均値を矩形の両端とし矩形領域を抽出する。
3. 抽出した矩形領域内で SIFT, COLORSIFT, 色情報, HOG, ガボール特徴量を求める。

本稿で用いるアンサンブル識別器を Fig.2 に示す。また, 学習手順を以下に示す。

1. 原画像群に対して色相領域, 彩度領域, 明度領域を抽出し, 原画像を含むそれぞれの領域に SIFT や COLORSIFT 等を行い, 色相領域から抽出した H-SIFT, 彩度領域から抽出した S-SIFT, 明度領域から抽出した V-SIFT, 原画像から抽出した O-SIFT 等の特徴量を抽出する。抽出した特徴量はトレーニングデータ A とする。

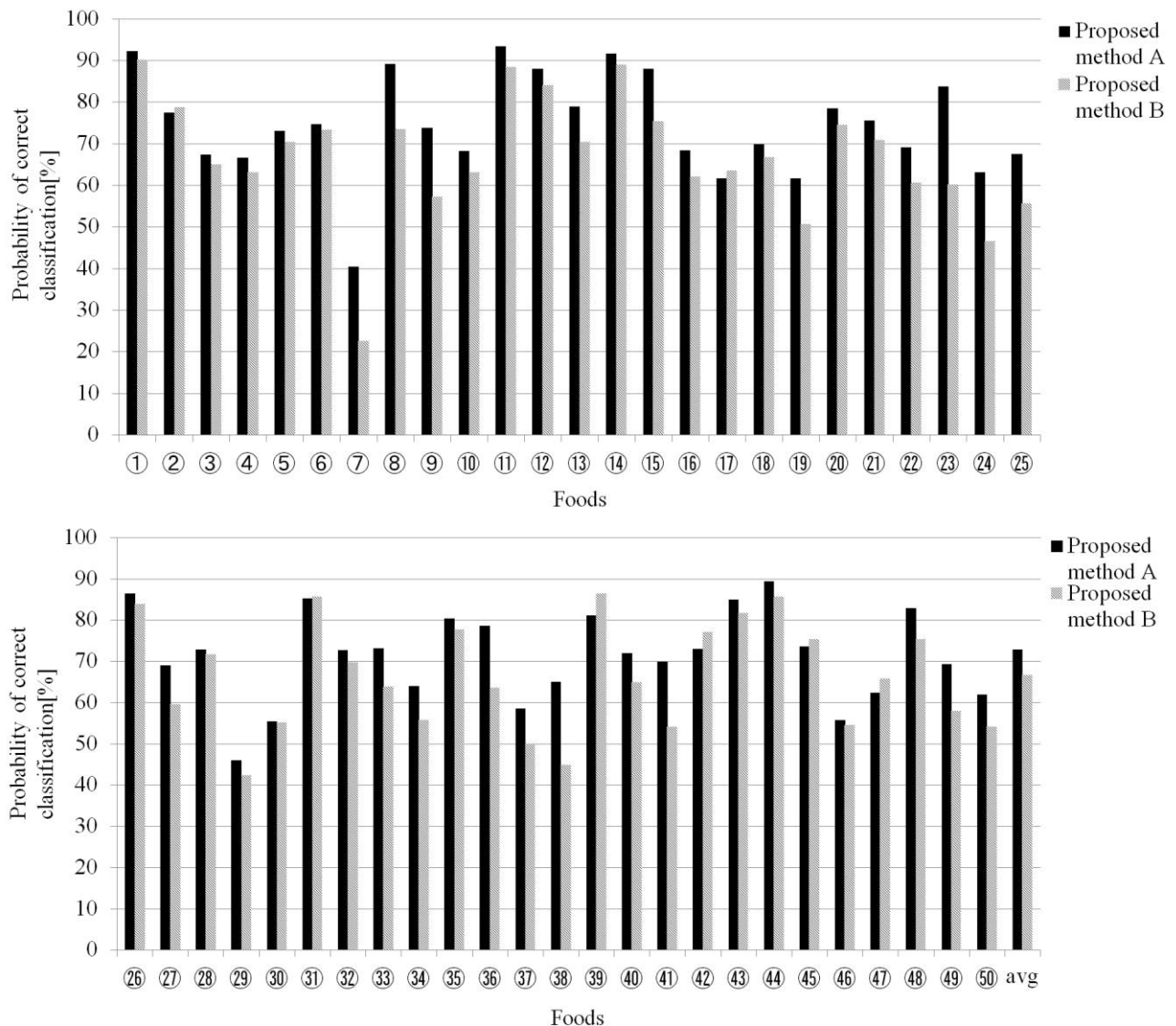


Fig.3 Result of classified 50 food categories by proposed method

2. トレーニングデータ A を 10 分割し, 1 つをテストデータ B, テストデータ B 以外をトレーニングデータ B とする 10-fold-crossvalidation を行い, 各データの識別結果を出力する。これを Random Forest や SVM 等, n 個の識別器で同様に識別する。出力した識別結果は各識別器の汎化性能を示している。
3. 2. で出力した識別結果を線形 SVM で評価する。

3. 評価実験

3.1 実験方法

Google 画像検索で取得した 50 種類の食事画像を用いて食事の種類を識別する。また, 食事画像毎に 40 枚取得した。識別に用いた食事の種類を Table.1 に示す。また, 特徴量の識別に用いた識別器は AdaBoost, Laplacian-LS-SVM, RBF-LS-SVM, ナイーブベイズ, ニューラルネットワーク, Random Forest, Linear SVM, RBF-SVM, Sigmoid-SVM を用いた。評価値として再現率 R と適合率 P の調和平均の F 値を用いる。評価は 10-fold-crossvalidation で行う。ま

た, 従来研究[12]で最も高い識別率を示した MKL-SVM や代表的なアンサンブル学習である Random Forest, Adaboost を単体で用いた場合の識別率と比較する。提案手法の各識別器のパラメータは単体で行った場合と同様のパラメータを用いた。また, 識別器は統計解析フリーソフト R を用いて構築した。統計解析ソフト R は SVM や Random Forest などの機械学習アルゴリズムや検定などの統計的手法をパッケージにして提供しているフリーソフトである。また, 識別に用いた計算機は, CPU が Intel Core i7 870 2.93GHz, メモリが 16GB である。

3.2 実験結果

実験結果を Table.2 に示す。Table.2 より, COLORSIFT を用いて色相領域, 彩度領域, 明度領域を重点的に学習した場合 (以下, 提案手法 A) は MKL-SVM と比較して識別率が 9.7% 向上しており, 他の Random Forest, Adaboost と比較し, それぞれ 12.9%, 29.6% 向上している。また, 色相領域, 彩度領域, 明度領域を学習せず, 複数の異なるカーネルを持つ SVM と他の識別器を組み合わせた場合

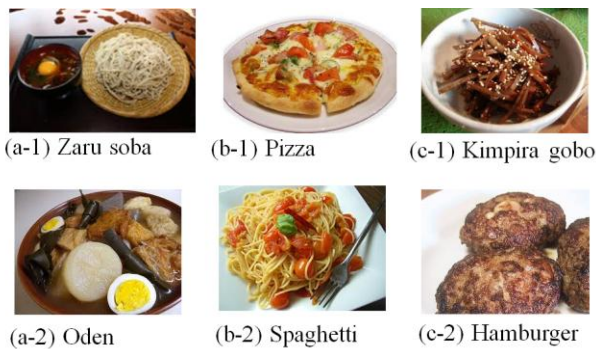


Fig.4 Food images are succeeded in recognition using proposed method A



Fig.5 Food images are missed in recognition using proposed method A

(以下、提案手法 B) も MKL-SVM や Random Forest, Adaboost と比較して識別率がそれぞれ 3.3%, 6.6%, 23.2% 向上している。また、本稿で提案した手法で 50 種類の食事画像を識別した結果を Fig.3 に示す。Fig.3 のグラフより、提案手法 A は提案手法 B と比較してカツ丼、カレーライス、きんぴらごぼう、ざるそば、すき焼き、ハンバーガー、ハンバーグ、ピザ、茶碗蒸し、天ぷらうどん、冷奴の識別率が 10% 以上向上した。

3.2 考察

提案手法 A と提案手法 B で識別率が顕著に改善された画像に注目し、色相領域、彩度領域、明度領域を重点的に学習した場合の影響を考察する。提案手法 B を用いた場合に食事の種類を誤識別し、提案手法 A を用いた場合に食事を正しく識別した食事画像を Fig.4 に示す。Fig.4 より提案手法 B ではざるそばとおでん、ピザとスパゲティ、きんぴらごぼうとハンバーグ等を誤識別しているが、提案手法 A で正しく識別している。提案手法 B において誤識別した理由としては Fig.4 のピザやスパゲティなど形状が異なるが、色が酷似していることが挙げられる。一方、提案手法 A は元画像に加え、それぞれ色、彩度、明度の変化が大きい領域である色相領域、彩度領域、明度領域を重点的に学習することで食事毎の形状の違いを捉えることが可能になったと考えられる。また、提案手法 A を用いた場合と提案手法 B を用いた場合の両方で誤識別した食事画像を Fig.5 に示す。Fig.5 より形状、色が酷似している食事に対しては画像を撮影した状況や匂い、食事量などを考慮する等 2D 静止画像で得られる情報以外の情報が必要と考えられる。

4. 結論・今後の展望

独居老人を対象とした栄養管理システムを開発する前段階として単品の食事画像から食事の種類を識別した。識別した結果、従来研究[12]より 9.7% 向上した。また、色相領域、彩度領域、明度領域を重点的に学習することで色情報が酷似している食事画像も識別可能となった。また、本稿では単品食事画像を対象としたが、従来研究[13]で提案されている複数品目の食事の識別手法と併用して用いることでさらなる精度向上が期待できる。一方、色情報や匂いが酷似している食事画像は誤識別した。これは 2D の静止画像だけでは識別が困難であると考えられる。そのため、今後は動画画像やニオイセンサや触覚センサ等を用いて画像を撮影した状況や匂い、食事量などを取得し、識別率の向上を図る。また、2つのレンズを装着している 3 次元カメラを用いてステレオ法で食事の 3 次元形状把握を行い、食事量を測定し、栄養を表示するシステムを開発する。

参考文献

- [1] 平成24年版高齢社会白書： <http://www8.cao.go.jp/kourei/whitepaper/w-2012/zenbun/>
- [2] 平成23年国民生活基礎調査の概況： <http://www.mhlw.go.jp/toukei/saikin/hw/k-tyosa/k-tyosa11/>
- [3] 芳賀博，七田恵，他：“健康度自己評価と社会・心理・身体的要因”，社会老年学，pp.15-24(1983)
- [4] 東京23区における孤独死の実態： <http://www.fukushihoken.metro.tokyo.jp/kansatsu/kouza.files/19-kodokushinoinjittai.pdf>
- [5] 永井晴美，七田恵子，他：“地域在宅老人の血清アルブミンの加齢変化と生命予後のとの関係”，日老医誌，Vol.21，pp.588-592(1984)
- [6] 湯川晴美：“在宅高齢者の食と健康に関する長期銃弾研究”，日本食生活学会誌，Vol.16，No.2，pp.100-103(2005)
- [7] 周燕飛：“介護施設における介護職員不足問題の経済分析”，公益財団法人医療科学研究 所 医療と社会，Vol. 19，No. 2，pp.151-168 (2009)
- [8] 長谷川聡，吉田友敬，江上いすず，横田正，村上洋子：“ケータイ栄養管理システムによる食育と栄養教育”，コンピュータ & エデュケーション，Vol.21，pp.107-113(2006)
- [9] S.Yang, M.Chen, D.Pomerleau, and R.Sukthankar: "Food recognition using statistics of pairwise local features", Proc.IEEE Computer Vision and Pattern Recognition, pp. 2249 - 2256 (2010)
- [10] Z.Zong, D.T.Nguyen, P.Ogunbona, and W.Li: "On the combination of local texture and global structure for food classification," IEEE International Symposium on Multimedia, pp.204-211(2010)
- [11] M.Chen, K.Dhingra, W.Wu, L.Yang, R.Sukthankar, and J. Yang: "PFID: Pittsburgh fast-food image dataset", 16th IEEE International Conference on Image Processing(ICIP), pp.289-292(2009)
- [12] 上東太一，甫足創，柳井啓司：“Multiple kernel learning による50種類の食事画像の認識”，

電子情報通信学会論文誌, Vol.J93-D, No.8,
pp.1397-1406(2010)

- [13] 松田裕司, 甫足創, 柳井啓司: "候補領域推定に基づく複数品目食事画像認識", 電子情報通信学会論文誌, Vol.J95-D, No.8, pp.1554-1564(2012)
- [14] D.Lowe: "Distinctive Image Features from Scale-Invariant Keypoints", International Journal of Computer Vision, Vol.60, No.2, pp.91-110(2004)
- [15] A.E.Abdel-Hakim and A.A.Farag: "C sift: A sift descriptor with color invariant characteristics", IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol.2, pp.1978-1983(2006)
- [16] N.Dalal, B.Triggs: "Histograms of oriented gradients for human detection", Proc.of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp.886-893(2005)
- [17] J.Jones, L.Palmer: "An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex", J Neurophysiol, Vol.58, No.6, pp.1233-1258(1987)
- [18] G.Csurka, C.R.Dance, L.Fan, and C.Bray: "Visual Categorization with Bags of Keypoints", European Conference on Computer Vision, pp.1-22(2004)
- [19] S.Lazebnik, C.Schmid and J.Ponce: "Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories", Proc.IEEE Computer Vision and Pattern Recognition, pp.2169-2178(2006)
- [20] D.Arthur and S.Vassilvitskii: "k-means++: the advantages of careful seeding", Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms, pp.1027-1035(2007)
- [21] Y.Freund, R.E.Schapir: "A Decision-Theoretic Generalization of on-Line Learning and an Application to Boosting", Journal of Computer and System Sciences, Vol.55, No.1, pp.119-139(1997)
- [22] L.Breiman: "Random Forests", Machine Learning, Vol.45, No.1, pp.5-32(2001)
- [23] Lanckriet, G. R. G., Cristianini, N., Ghaoui, L. E., Bartlett, P., & Jordan, M. I.: "Learning the kernel matrix with semidefinite programming", Journal of Machine Learning Research, Vol.5, pp.27-72(2004)
- [24] P.Domingos, M.Pazzani: "On the Optimality of the Simple Bayesian Classifier under Zero-One Loss", Journal of Machine Learning Research, Vol.29, No.2, pp 103-130(1997)