

## 料理写真撮影における基本構図決定システムの検討 Basic Composition Decision System for Food Photography

大木 慧矢<sup>†</sup> マッキンケネスジェームス<sup>†</sup> 永井保夫<sup>†</sup>  
Seiya Ohki Kenneth J. Mackin Yasuo Nagai

### 1. はじめに

近年スマートフォンの発展・普及により写真撮影が手軽となり、SNSなどへの公開が容易になってきた。また、撮影上での色合いなどは最近のHDR (High Dynamic Range) 機能や写真の編集アプリでの調節が可能となるとともに、簡単に編集可能となってきている。そのため、写真に対する一層のクオリティーの高さが求められている。しかしながら、露出や色合いは簡単に編集可能であるのに対して、最適な構図を決めて上手く撮影することは専門家でないとおこなえていない。

我々はこの問題を解決するために、深層学習を利用し、魅力的な構図で撮影をする補助システムの開発を行う。

### 2. 基本構図決定システム

我々は、魅力的な構図で撮影をする補助システムの開発を行うため、柿森ら[3]が使用した料理写真の構図の要素に着目した。ここでは、構成要素の中で「写真内における料理の数、それぞれの位置、向き、占有面積(写真にどの程度大きく写っているか)、撮影時のカメラの角度」の5つの内、「向き」と「撮影時のカメラの角度」を取り上げる。提案する構図決定システムでは、佐藤らの魅力度推定器[2]を使用し、様々な角度から動画で撮影を行い、動画から魅力的な「向き」と「撮影時のカメラの角度」の画像を出力する。

さらに、撮影時のカメラの角度のなかで、「向き」と「撮影時のカメラの角度」以外の、写真内における位置と占有面積の2つの要素を考慮し、撮影することができるシステムの検討をおこなった。そこでの判定器[4]では、Twitter上の画像と画像に対する「いいね」の値を学習データとし深層学習を使い、YOLOV3で料理の切り抜きを行うことで、位置や占有面積を考慮し構図を評価する手法を取り入れていた。

しかし、この手法では、撮影の補助を行うシステムを作成する上でユーザが判定器の評価を元に試行錯誤的に構図を評価しなければならない。そのため、撮影補助システムとして、簡単に使用することができないという問題があった。

そこで、本提案では、ユーザが対象を撮影する際、対象を真上から一度撮影し、システムに入力を行う。これにより、提案システムでは、柿森らが提案した基本構図の決定手法で求められた基本構図を、ユーザに提示することで、ユーザが簡単に構図を決定できるようになる。

### 3. 提案手法

本提案では、ユーザに基本構図を提示するシステムを開発するため、深層学習を適用し、お皿の数や料理の面積、お皿の形から、料理写真の基本構図を決定する。ここでの基

本構図は、丸形のお皿が1枚の場合の5つの分類と、お皿が2枚以上であった場合の5つの分類の合計10種類の分類により決定される。前者のお皿が1枚の場合は、メイン料理の皿の面積がメイン料理の料理部分の面積の2倍以上(以降、小料理と呼ぶ)(図1A)、丸形のお皿が1枚でメイン料理の料理部分の面積の2倍未満(以降、大料理と呼ぶ)(図1B)、正方形のお皿1枚で小料理(図1C)、長方形のお皿1枚で小料理(図1D)、正方形のお皿1枚で大料理(図1E)の5つの分類が考えられる。お皿が2枚以上であった場合は、1枚の場合の5種類の分類が考えられる。本研究では、これらの10種類の分類の前段階としてお皿が1枚のみでの想定で5種類の分類器を作成した。

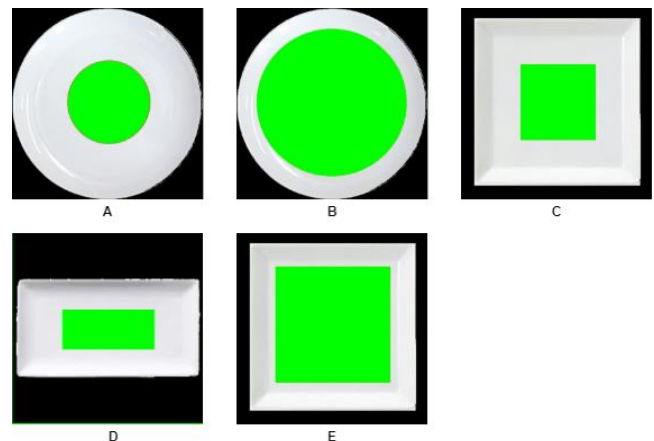


図1 お皿と料理面積のイメージ(緑部分が料理)

#### 3.1 データセットの作成

データセット作成には3種類のお皿(丸型、正方形、長方形)と10種類の料理のデータ[5]を使用した。図2は皿の面積が料理の料理画像の面積の2倍以上(小料理)と、面積が料理の料理画像の面積の2倍未満(大料理)になるように重ね合わせた画像を示している。本研究では、合計で50枚の画像(1分類10枚)を作成した。



図2 お皿と料理写真を重ね合わせた例

次に、水平移動、垂直移動、回転、拡大、縮小、アフィン変換、上下反転、左右反転の8種類に加工を行うことで、合計450枚(1分類90枚)の画像を加工し、グレースケール化を行った。

<sup>†</sup> 東京情報大学 Tokyo University of Information Sciences

### 3.2 分類器の作成

分類器は、VGG16に転移学習を行うことで実現した。転移学習にあたって作成したデータセットは  $244 \times 244$  にリサイズし入力した。画像に対する教師データを入力し全結合層のみ学習を行う。重みの学習を行う全結合層の活性化関数にはReLUを使用し、出力の全結合層にはソフトマックス関数を用いた。最適化手法にはAdam、評価関数にはMAE (Mean Absolute Error) を用い、学習回数は200回とした。

### 4. 評価方法

図3は、3.1で作成したデータセットの80%を訓練データ、残りの20%をテストデータとして5分類で教師あり学習を行った場合の正解率を示している。図4は教師データと予測値との絶対平均誤差を示している。

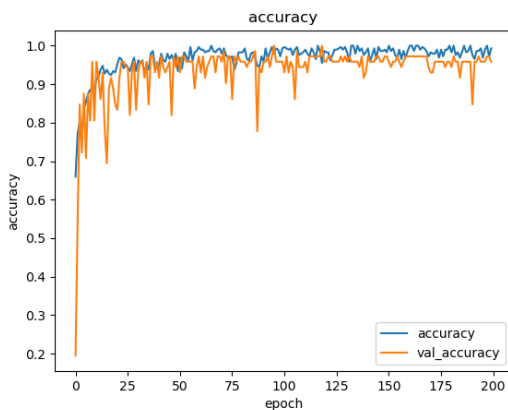


図3 データセットを学習した正解率

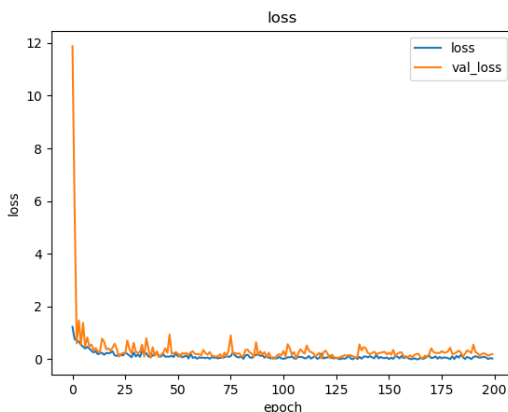


図4 データセットとの絶対平均誤差

データセットを学習した結果、正解率と平均誤差ともに早い段階で収束することができ、訓練データの学習でも、多少の揺らぎが見られたが収束することができた。

さらに、図5のような、実際の料理写真を分類器に入力した結果、5件中4件正しく分類することができた。



図5 実際の料理写真の入力例  
(左: 長方形・小料理, 右: 丸形・小料理)

### 5. 結論

我々が以前取り上げた撮影補助システムでは、ユーザが判定器の評価を元に構図を試行錯誤して評価しなければ簡単に撮影できないことが明らかになった。これに対して、本研究では、深層学習を適用することで、ユーザが撮影対象を真上から一度撮影し、お皿の数や料理の面積、お皿の形から料理写真の基本構図を決定し、ユーザに提示するシステムを開発した。

その結果、作成された分類器では、お皿が1枚のみである場合、元の画像が1分類あたり10枚と少ない画像では正解率と絶対平均誤差ともに高い精度を出すことができた。実際のデータでも、ユーザが撮影対象を真上から料理を撮影することで、分類した結果から最適な構図を提示することができた。

しかし、実際の入力結果では、長方形のお皿と正方形のお皿の判別が困難であり、本提案ではお皿が2枚以上ある場合に分類が十分に行えていない。今後は、データセットの画像を2枚以上ある分類とし、長方形と正方形のお皿の画像のパターンを増やすことを考えている。

最終的には、「向き」と「撮影時のカメラの角度」の要素を使ったシステムと組み合わせることで5つの料理写真の構図の要素すべてを考慮した撮影補助システムの実現を検討していく。

### 参考文献

- [1] 大木 慧矢, ケネス マッキン, 永井保夫: 料理写真の魅力度推定器を使った写真撮影補助システムの検討, 第82回情報処理学会 全国大会, 2020年.
- [2] 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, vol. 117, no.252, MVE2017-32, pp. 107-111, 2017年.
- [3] 柿森 隆生, 岡部 誠, 尾内 理紀夫: おいしそうな料理写真撮影を支援するシステムの検討, 第56回プログラミング・シンポジウム予稿集, pp. 131-141, 2015年.
- [4] 大木 慧矢, ケネス マッキン, 永井保夫: 料理写真の構図決定補助システムにおける構図評価手法の検討, 第83回情報処理学会 全国大会, 2021年.
- [5] Kazuma Takahashi, Keisuke Doman, Yasutomo Kawanishi, Takatsugu Hirayama, Ichiro Ide, Daisuke Deguchi, Hiroshi Murase: "Estimation of the Attractiveness of Food Photography Focusing on Main Ingredients", Proc. of 9th Workshop on Multimedia for Cooking and Eating Activities (CEA), pp. 1-6, Aug. 2017.