

# 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における撮影手法の改善

東京情報大学

大木 隼矢、永井 保夫、マッキン ケネスジェームス

## 研究背景

- ▶ 近年スマートフォンの発展、普及によりSNSなどで評価を得ることができる魅力的な写真を撮りたいという需要が増えている
- ▶ 写真を魅力的に撮るうえで構図や色などを考慮する必要がある
- ▶ 写真を撮る上で色は写真編集アプリなどで簡単に手直しができる
- ▶ しかし非専門家が構図を考えて写真を撮るのはなかなか難しい
- ▶ 構図を考慮した魅力的な写真を撮る補助をするシステムを作れないか？

## 前回の提案

- ▶ 深層学習を使いユーザが構図を決定する上での配置や撮る視点を考慮する必要がなく撮影が行えるのではないか
- ▶ SNSなどに上げる機会が多い料理写真を魅力的な（おいしそうに見える）視点で撮る
- ▶ 料理写真がおいしそうに見える度合いを魅力度とする
- ▶ 深層学習を使った魅力度推定器を利用し、魅力的な視点で写真を撮る補助をするシステムの開発を行った
  - 畳み込みニューラルネットワーク（CNN：Convolution Neural Network）のVGG16とVGG19に転移学習を行い魅力度推定器を作成を行った
  - スマートフォンなどから様々な方向で撮影した料理の動画から最適な視点の写真を出力を行った

# 研究結果

- ▶ VGG16の代わりとしてVGG19でも転移学習を行い魅力度推定器として使用することができる。
- ▶ 実際に動画から出力を行うと、撮影方法によりデータセットに近い（魅力度の高い）画像を選び出すことができた
  - ディープラーニングを使用することでユーザが配置や撮る視点を考慮する必要がなく撮影が行える
- 余計なものなどが映り込んだ画像を選んでしまう場合もあった
- 写真撮影補助システムとして、見切れている画像などの画像を選ばないように改良する必要がある。

# 目的

- ▶ 入力の際の見切れている画像を連写で撮ることで安定して撮影が行える
- ▶ 連写で撮ることで入力の際に見切れている画像を減らす
- ▶ 見切れている画像を少なくすることで魅力度の向上を行う

# 提案

前回の提案：入力に使用する動画を画像化し入力する

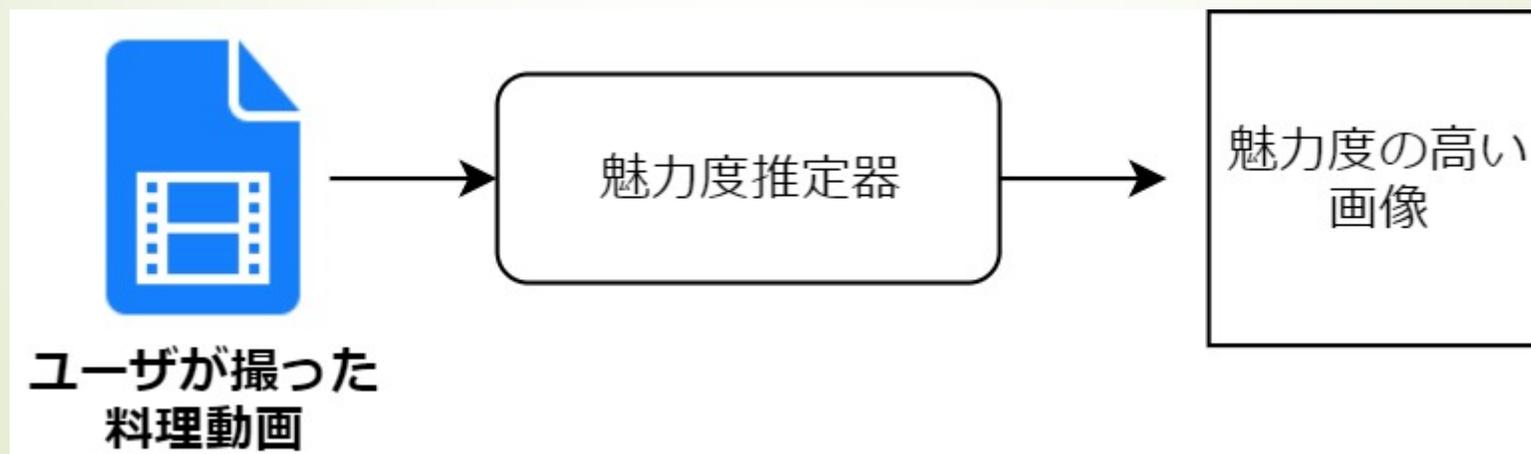


本提案：カメラの連写機能を使い入力する

- 見切れている画像を少なくすることにより魅力度の向上を行い、システムの精度を改善する

# システムの概要

- ▶ CNNに転移学習を行い、料理写真の魅力度推定器を作成する
- ▶ 動画で料理を様々な方向から撮影を行い、魅力度推定器を使い動画の中から最も魅力的な写真を抽出する



# 魅力度推定器の作成

- ▶ 魅力度推定器の作成には先行研究で評価の高い、学習済みのVGG16を使用した
- ▶ ネットワークの学習には、高橋ら[5]のデータセット「NUFOOD 360×10」の料理写真を244×244にリサイズを行い入力した
- ▶ 画像に対する魅力度を教師データとして入力した
- ▶ 畳み込み層の重みを固定し、全結合層のみ学習を行う

# データセット:NUFOOD 360×10

- 物体を様々な方向から撮影できるシステムを使い、36の異なる視点から10種類食品サンプル撮影
- 被験者に異なる視点の画像の組み合わせを見せ、魅力度が高いものを選ばせ、結果を0～1の値からなる魅力度を写真に設定している



魅力度 : 1

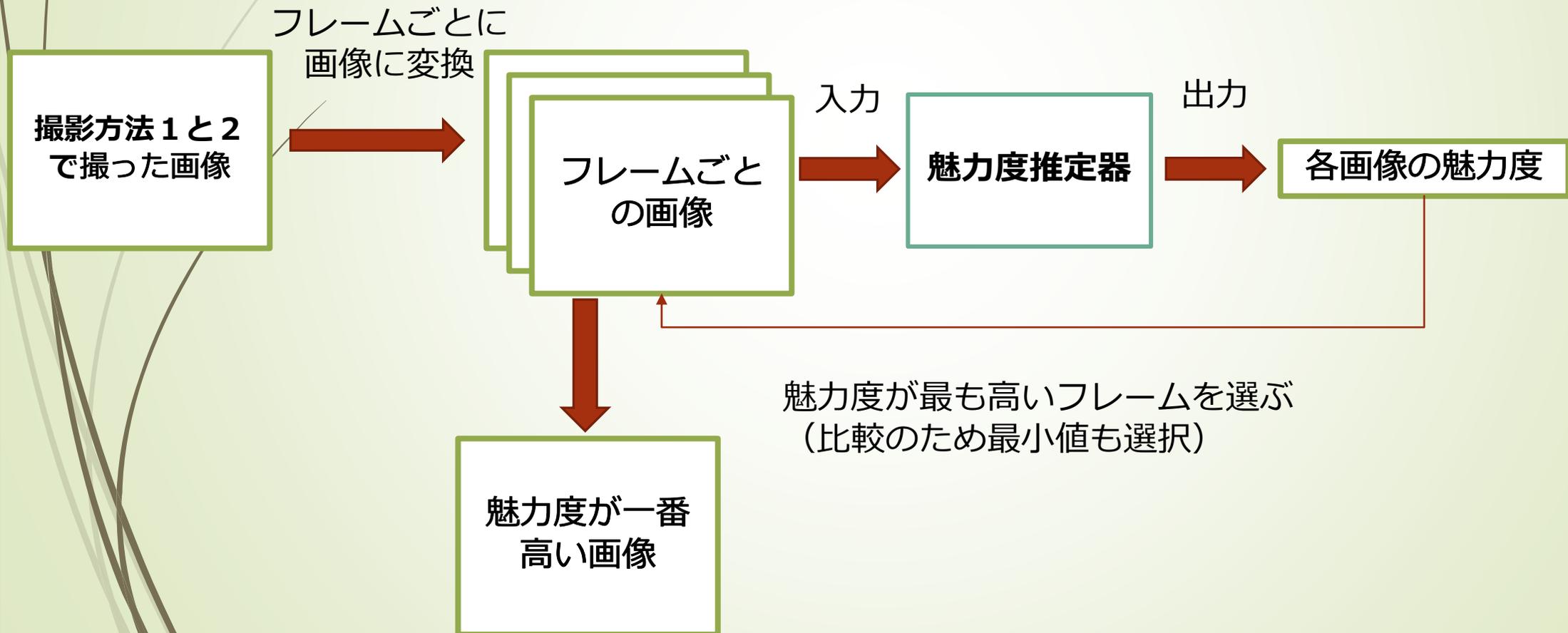


魅力度 : 0.595024437



魅力度 : 0

# 実際の動き



# ネットワークのパラメータ

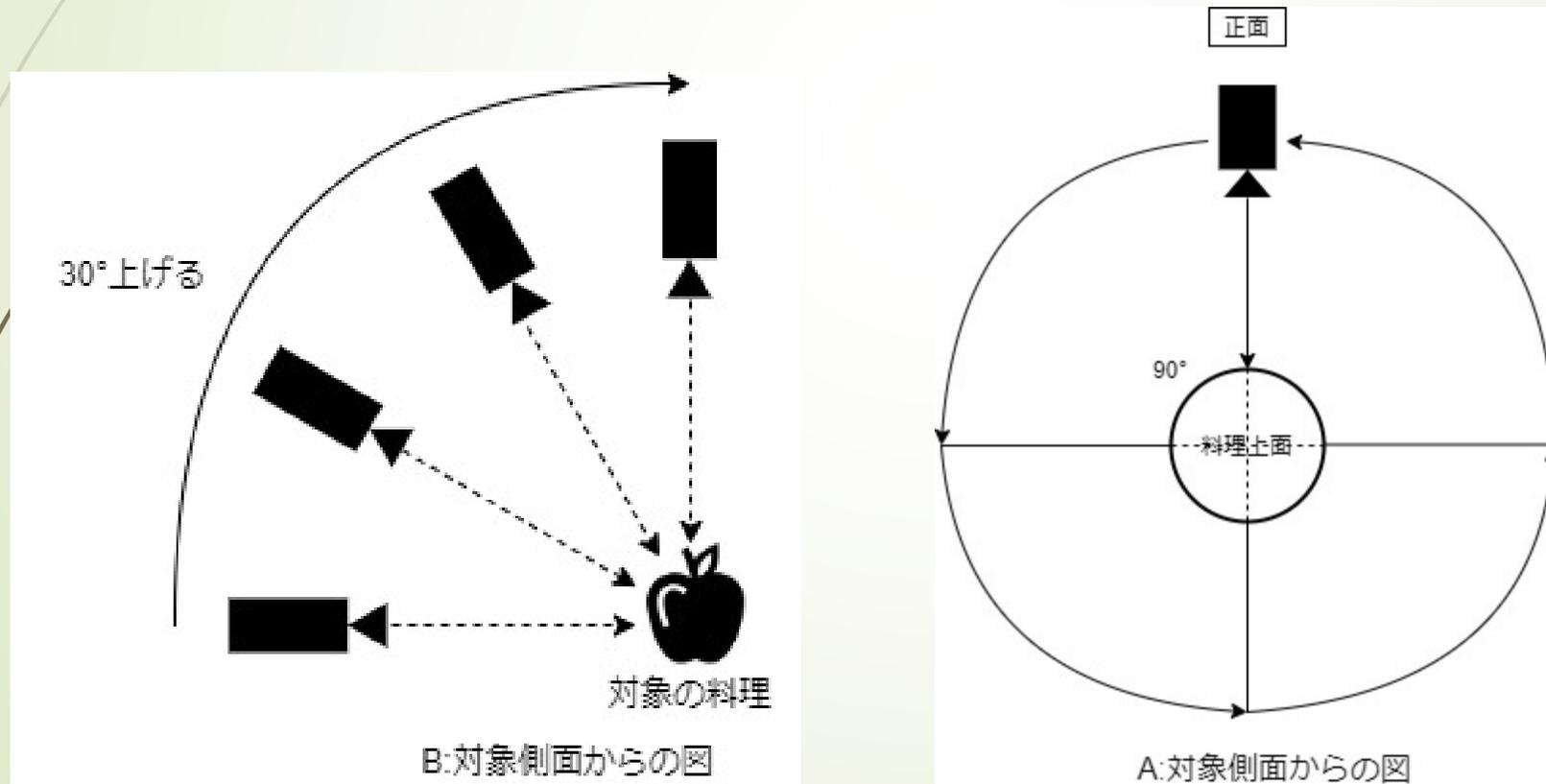
- ▶ 全結合層のユニット数 :  $256 \times 2$
- ▶ 全結合層の活性化関数 : ReLU ( $f(u) = \max(u, 0)$ )
- ▶ 出力層のユニット数 : 1
- ▶ 出力層の活性化関数 : シグモイド関数 ( $f(u) = \frac{1}{1+e^{-u}}$ )
- ▶ 最適化手法 : Adam (Adaptive Moment Estimation)
- ▶ 評価指数 : 出力と教師信号との平均絶対誤差  
(MAE : Mean Absolute Error )
- ▶ 学習回数 : 100回

# 画像の撮影

- ▶ 様々な角度から最適な構図を抽出するため対象の料理を様々な方向から撮影を行った
- ▶ 撮影にはスマートフォンのカメラで連写機能を使い画像を撮影した
- ▶ 撮影方法として2つの手法と比較のために、動画でも同じ手法で撮影をした

# 入力画像の撮影方法 1

- 対象の料理を俯角0度で360度を回し、その後おおよそ俯角を30度ずつ上げ撮影し、俯角90度まで撮影する



# 入力動画の撮影方法 1

## メリット

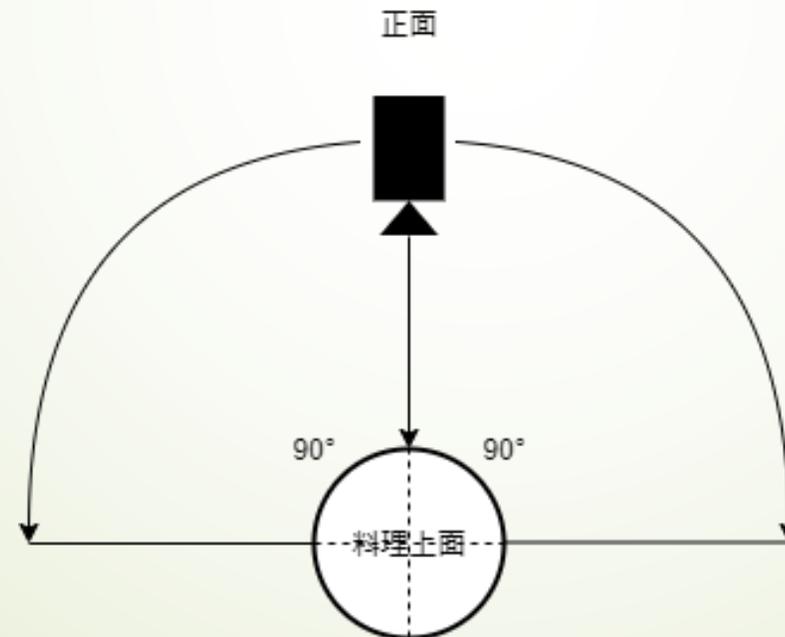
- ▶ 正面を考慮せずに撮影できる
- ▶ データーセットの撮影方法に近い手法で撮影できる

## デメリット

- ▶ 対象が見切れているフレームや写っていないフレームが出やすい
- ▶ ユーザーが意図しない方向で撮影される場合がある

## 入力動画の撮影方法2

- 対象の料理に対し正面から俯角0度で左右90度に動かし、俯角をおおよそ30度ずつ撮影方法1と同じように上げ、俯角90度まで撮影する



## 入力動画の撮影方法2

### メリット

- ▶ 対象が見切れているフレームや写っていないフレームが出にくい
- ▶ ユーザーが意図しない方向で撮影されにくい

### デメリット

- ▶ ユーザーが料理の正面を考慮しなければならない
- ▶ 入力に使えるデータが少なくなる

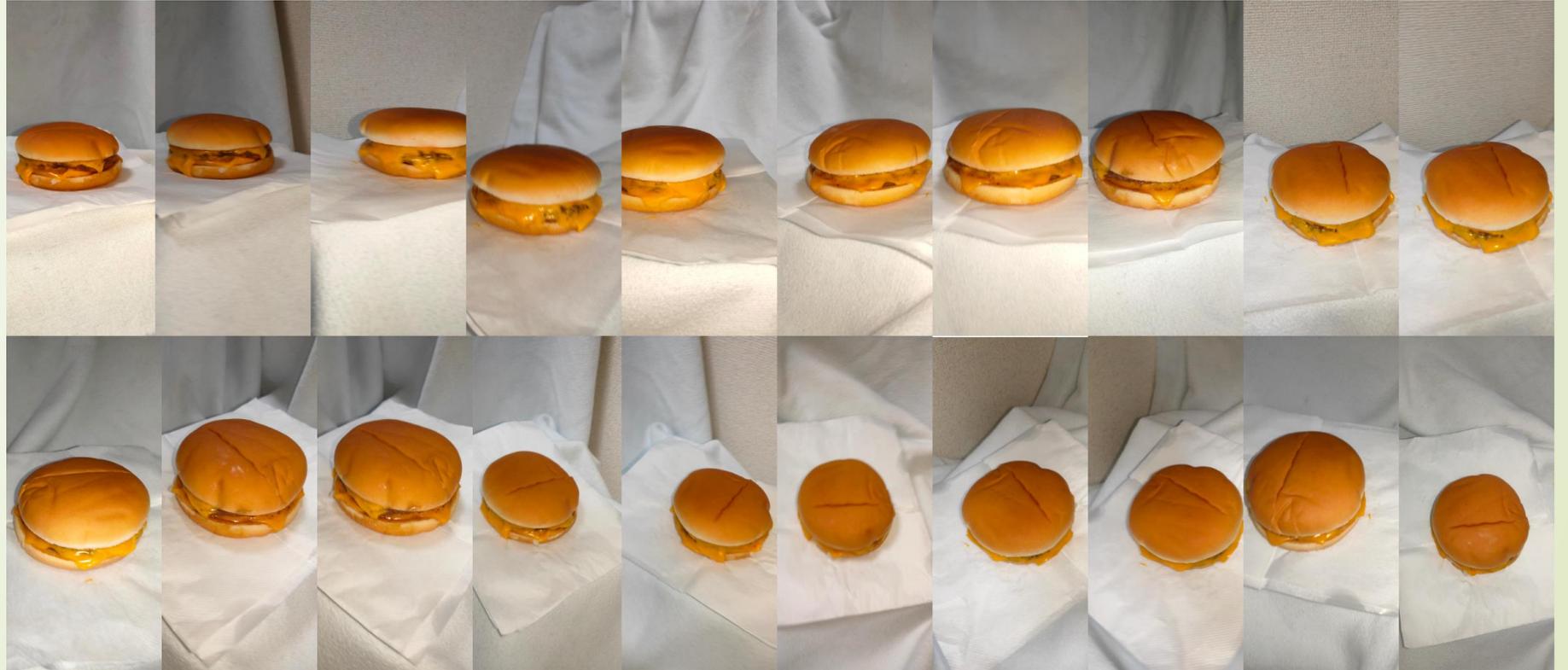
# 評価方法

- ▶ 撮影方法の手法1と手法2で撮った動画を魅力度推定器に入力を行う
- ▶ 連写機能で撮影方法の手法1と手法2で撮った画像を魅力度推定器に入力を行う
- ▶ 動画と連写機能のそれぞれの見切れている画像（フレーム）数と魅力度が最大の画像を出力する。

# 手法 1 で実際に撮った動画



# 手法 1 で実際に撮った連続写真



# 結果 1

- ▶ 手法1で撮影した動画と連続写真を魅力度推定器に入力した結果



データセット[5]の魅力  
度が最大値の画像



A: 動画での撮影



B: 連写での撮影

## 結果 2

- ▶ 手法2で撮影した動画と連続写真を魅力度推定器に入力した結果



A: 動画での撮影



B: 連写での撮影

## 出力された魅力度と見切れている画像の割合

	魅力度最大	入力枚数	見切れている画像	割合
結果1-動画での出力	0.8086604	1570	496	0.315
結果1-連写での出力	0.8642167	763	219	0.287
結果2-動画での出力	0.5976593	264	30	0.113
結果2-連写での出力	0.820611	164	35	0.213

# 考察

- ▶ 動画で撮影するよりも写真で撮影することにより、入力データの見切れている画像の割合を少なくすることができた
- ▶ 魅力度は動画で撮影するより、連写で撮影することで魅力度を高くすることができた
- ▶ 実験1の連写での出力画像は真上からの視点で、データセットでの魅力度が最も高いチーズバーガーの視点とは違うものになってしまった
- ▶ 違う料理の魅力度が高い画像に近い視点の画像を抽出してしまったためであると考えられる

# 結論

- 本研究では、動画を画像化し入力する代わりに、カメラの連写機能を使うことで手振れや見切れている画像を少なくする撮影手法の改善策を提案し、より良い構図の写真を出力できるか実験を行った
- 結果、連写で撮影を行うことで見切れている画像を減らすことができた
- 出力される魅力度を前回報告より魅力度を高く出力することができた
- しかし、出力された画像は対象の料理のデータセットとは違う視点になった画像が出ていた

# 今後の展望

- ▶ 撮影している対象を判断する分類器を作成する
  - ▶ 違う料理の魅力度が高い画像に近い画像を抽出するのを防止する
- ▶ VGG16、VGG19以外のCNNのネットワークでのテスト
  - ▶ ネットワークの精度を向上させる
- ▶ 料理写真以外での抽出を行えるようにする

## 参考文献

- [1]角野杏早比、太田祐揮、志津野之也、原大輔、濱川礼：写真撮影における構図決定支援システム ～E-cose～
- [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
- [3]田中直人、高島健太郎：機械学習技術を応用した写真撮影における個性発揮支援システムの提案、情報処理学会研究報告.HCI、研究報告ヒューマンコンピュータインタラクション
- [4]Keras / Tensorflowで転移学習を行う  
<https://qiita.com/yampy/items/6f1f48fee16db7888f07>
- [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