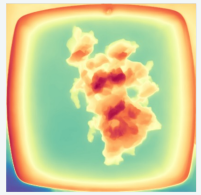
(42) 大規模視覚言語モデルを用いた食事画像からのカロリー量推定

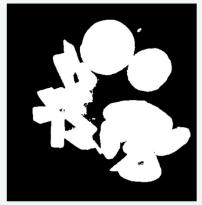
田邊光[1], 柳井啓司[1], 1電気通信大学

- **LLM の知識や推論性能を活用して** 食事画像からカロリー量を推定
- 手法１：大規模視覚言語モデルをファインチューニング
  - 食事カロリー量推定について複数の評価指標でスコアが改善
- 手法２：食品体積推定モデルと大規模視覚言語モデルを使用
  - ゼロショット食事カロリー量推定についてスコアが改善

体積推定値：445.3 mL
カロリー量正解値：149 kcal
カロリー量推定値：144 kcal



体積推定値：600.3 mL
カロリー量正解値：156 kcal
カロリー量推定値：222 kcal



手法２のカロリー量推定結果（下段：体積推定過程の領域分割・単眼深度推定の結果）