

レシピフローグラフによる調理後食材の抽出を用いた レシピのカロリー推定の検討

匹田 祐平^{†1} 清 雄一^{†1} 田原 康之^{†1} 大須賀 昭彦^{†1}

概要：料理レシピは世界中で生成・利用されている人気のコンテンツであり、近年そのレシピ数は急速に増大している。現在、Web で検索可能なレシピの大半を占めるのは、一般ユーザーが作成・投稿したユーザー投稿型レシピである。一方で、多くのユーザー投稿型レシピサイトでは料理レシピのカロリー記載がなく、レシピから正確なカロリー推定を行う必要性は高まっている。そこで、本研究では各食材が様々な調理方法によってカロリーが変化する点に着目し、レシピからカロリーを推定する方法を提案した。料理レシピの手順部分をフローグラフと呼ばれるグラフに変換し、食材とその食材に対する調理動作を求めることで調理後食材を求め、日本食品標準成分表を参照することで正確なカロリーの推定を試みた。実験では 100 件のレシピに対しカロリー推定を行い、結果として平均 16%程度の誤差が生じることが分かった。

1. はじめに

料理レシピは世界中で生成・利用されている人気のコンテンツであり、近年そのレシピ数は急速に増大している。現在、Web で検索可能なレシピの大半を占めるのは、一般ユーザーが作成・投稿したユーザー投稿型レシピである。米国では Allrecipes や Genius Kitchen などが、日本においてはクックパッドや楽天レシピなどがユーザー投稿型レシピサイトに該当する。特にクックパッドは 500 万件以上のレシピを持つ国内最大級のレシピポータルサイトとなっており、日に日にユーザー数は増加している。一方で、2015 年に内閣府により行われた「食育に関する意識調査」[1]では日頃から健全な食生活の実践を心掛けている人は 4 人に 3 人であると明らかにされ、国民の健康意識が高まっていることが示唆されている。

しかし、通常ユーザー投稿型レシピサイトにおいて、投稿者はレシピの材料名、調理手順などの情報しか掲載することができない場合が多く、各レシピの料理完成後のカロリーまで記載できることは少ない。また、一般的にカロリーを推定する方法として、ユーザーが各食材のカロリーを推定し計算するツールが存在するが、これらのツールは単に食材の種類とその量を指定し単純な和で計算されることが多い。これに対し、一般に調理によるカロリーの変化は ゆでる<網焼き<蒸す<煮る<炒める<揚げる の順で上がりやすくなるということが知られており、正確なカロリー推定を行うにはその食材に行われる調理動作を考慮する必要がある。[2]

2. 関連研究

森ら[3]は先行研究として挙げられる、フローチャートによる一般の手順文書の表現[4]や作業動作とその対象と道具の関係性を表現する木構造[5]といった研究を参考に、一般の手順文章をフローグラフと呼ばれるグラフで表現するこ

とを提案し、レシピを対象としてアノテーションを行った。また、後にレシピテキストからレシピフローグラフを推定する方法を提案した。山肩ら[6]はレシピフローグラフを用いて食材とその食材に対する調理動作の組み合わせを元により詳細にレシピ検索をする手法を提案し、既存手法との優位性を示した。西村ら[7]は写真付きレシピの作成を容易にするために、検索課題として取り込まれてきた手法を文生成モデルに組み込むことで、写真列を入力としてレシピを生成する課題を解決する手法を提案した。また、大野ら[8]はユーザーが作りたい料理のレシピ情報や雰囲気やクエリとして入力すると相性のいい料理レシピを推薦するシステムを提案し、ユーザーが推薦されたレシピ群からレシピを選択することで、そのレシピも含めたクエリで次の料理を推薦し、それを繰り返すことで献立作成の支援を試みた。これらの研究はレシピを生成するものや、レシピデータの中からユーザーに適したレシピを推薦するまでにとどまっており、そのレシピのカロリーについては考慮されていない。よってそれらの生成・推薦されたレシピに対しカロリーを推定するような研究は有用であると考えられる。カロリー推定に関しては加藤ら[9]は既存の料理レシピ提供サイトで提供される料理レシピから、食材や調理方法を抽出し Doc2Vec でレシピのベクトル化を行い、文章ベクトルとカロリーデータを用い、scikit-learn と SVR で自動カロリー判定器を構築し、未知のレシピに対してカロリーを推定する方法を提案した。第 8 章で述べるが、本研究におけるベースラインモデルはこの加藤らの研究手法である。加藤らの研究と本研究との違いは、手法の違いであり、使用するデータや目的は同様であるため今回はこの研究をベースラインモデルとして比較することとした。

3. 研究背景

通常、ユーザー投稿型レシピサイトにおいて、その料理のカロリーが掲載されていることは少ない。投稿者はレシピ

^{†1} 現在、電気通信大学

の食材名、調理手順などの情報しか掲載せず、一般的にカロリーを掲載する項目はない。そこでカロリーを知りたいユーザーが自らカロリーを求める必要がある。そこでカロリーを推定する方法として、レシピの食材からカロリーを自動計算するツールが存在する。しかしこのようなツールは食材の種類とその量のみを考慮し単純な和で計算されていることが多い。これに対し、一般的に調理によるカロリーの変化は「ゆでる<網焼き<蒸す<煮る<炒める<揚げ」の順で上がりやすくなるということが知られており*1、正確なカロリー推定を行うためには調理動作を考慮する必要がある。表1に生の茄子100gに対する調理方法の違いによるカロリーの違いを示す。表1に示した通り、ゆでた茄子と素揚げした茄子を比較すると10倍ほどのカロリーの差が生じていることがわかる。また、なかには調理後食材を扱うことのできるカロリー計算ツールも存在するが、各食材に対してどのような調理が施されているかをユーザーがレシピを参照しながら求める方法は労力を要してしまう。本研究ではこの問題を解決することを研究課題とした。

表1 茄子の調理方法の違いによるカロリーの変化
(出典:[10])

調理方法	エネルギー (kcal)
生	18
ゆで	17
油いため	73
天ぷら	165
素揚げ	177

4. 研究目的

本研究では、ユーザー投稿型レシピサイトなどのネット上のレシピに対して、ある食材に対しどのような調理が施されるかによってカロリーが変化するという点に着目し、調理手順をフローグラフと呼ばれるグラフに変換し、そのフローグラフを用いて正確にレシピのカロリー推定する方法を提案する。

5. アプローチ

この章では5.1章にて本研究で用いられるレシピフローグラフについて説明し、5.2章で本研究の提案手法を説明する。

5.1 レシピフローグラフ[11]

レシピ: ホットドッグ Recipe of "Baked Hot Dog."

Title: ホットドッグ (Baked Hot Dog)
 Ingredients: フランクフルト8つ (8 frankfurters), ホットドッグパン8つ (8 hot dog buns), 豆入りのチリ1缶(14.5オンス) (1 (14.5 ounce) can of chili with beans), カットオニオン1/2カップ (1/2 cup onion (diced)), 千切りチeddarチーズ2カップ (2 cups shredded cheddar cheese), マヨネーズ (mayonnaise), マスタード (mustard), 甘味料 (sweet relish)
 Steps: 1. 各ホットドッグパンの内側にマヨネーズ、マスタード、甘味料を広げる。(Spread the inside of each hot dog bun with mayonnaise, mustard and sweet relish.)
 2. 各ホットドッグにチリ、チーズ、オニオンをふりかける。(Sprinkle each hot dog with chili, cheese, and onion.)
 3. アルミホイルで覆い、オーブンに置く。(Cover with aluminum foil and place into the oven.)
 そして、350度で45分間焼く。(Then bake at 350 degrees for 45 minutes.)

図1 一般的なレシピの構成

レシピフローグラフについて詳しく説明する。まず多くのレシピは図1のような

- タイトル
- 食材リスト
- 調理手順

の形式に従って記述されている。レシピフローグラフはこの内、「調理手順」の意味内容の形式表現で、

- 頂点：食材や調理動作といったレシピ特有の表現 (レシピ用語)
 - 辺：レシピ用語間の関係
- からなる根つき無閉路有向グラフ (rooted DAG) で表される。

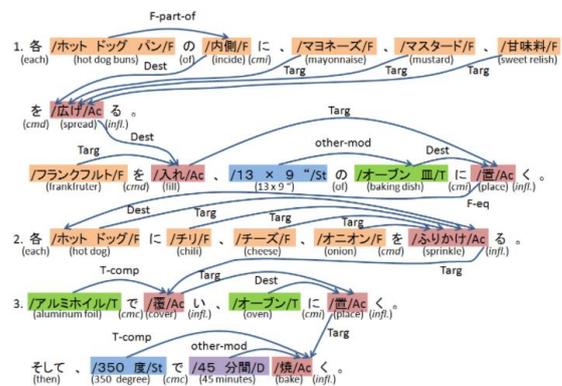


図2 可視化されたフローグラフ

図2は図1のレシピをフローグラフ化し、可視化したものである。無閉路有向グラフの根(図2における「焼/Ac」)が最終的に生成される料理に対応する。フローグラフの頂点は、食材や動作を表す単語列(レシピ用語、r-NE)とタグの組であり、図2では頂点は背景色のある単語によって示されている。また、一般的な固有表現と同様に単語中の用語の境界はないとしており、用語が入れ子関係になることもない。食材や道具といった名詞句だけでなく、調理者の動作のような用言もレシピ用語としており、同一性の簡便な判定などのために、活用語は互換のみをレ

レシピ用語としている。表 2 にレシピ用語の一覧を示す。

表 2 レシピ用語の一覧 (出典:[11])

タグ	意味	備考
F	食材	代名詞・中間・最終生成物を含む
T	道具	調理道具や器など (代名詞を含む)
D	継続時間	概数表現を含む
Q	分量	概数表現を含む
Ac	調理者の動作	語幹のみ
Af	食材の動作	語幹のみ
Sf	食材の状態	味、切り方など
St	道具の状態	温度設定など

フローグラフの辺は、用語間に関係があることを表す。文書中のレシピ固有表現に以下の 3 種類のうちのいずれかの関係が認められる場合、それらのレシピ用語に対応する頂点の間に有向辺を張り、ラベルとしてその関係の種類を付与する。

- 述語項関係がある場合 (ゼロ照応も含む)
 - 同種の二つのレシピ固有表現の間に何らかの参照関係がある場合
 - 二つの動作の間に何らかの関係がある場合
- 辺のラベルの一覧を表 2 に示す。また、頂点間の関係は大きく分けて以下の 4 つがある

(1) **述語項関係 (ゼロ照応も含む)**

調理手順の内容は主に「食材」と「動作」の関係であるため、これらの述語項関係が最も重要とされる。Agent, Targ, Dest, T-comp, F-comp が該当する。

(2) **2 つの食材や道具の参照関係**

既に出現している食材や、食材を加工した物体を指すレシピ用語を関連付ける。F-eq, F-part-of, F-set, T-eq, T-part-of が該当する。

(3) **2 つの動作の関係**

実際には一度しか行われていない動作を繰り返して記述している部分や、動作のタイミングを関連付ける。A-eq, V-tm が該当する。

(4) **その他**

その他の装飾関係を Other-mod としている。
辺は、同一文内の頂点の間だけではなく、異なる文の頂点の間に対しても付与される。したがって、調理手順全体の構造を表現することが可能となる。つまり各食材からの辺を辿ることで、その食材がどの頂点と関係がありどのような調理を施されるかを追うことができる。

表 3 辺のラベルの一覧

ラベル	意味	概説
Agent	主語(カ格)	主に「が」や「は」で表される動作と主語の関係
Targ	対象(ヲ格)	主に「を」で表される動作と対象の関係
Dest	方向(ニ格)	主に「に」で表される動作と方向や場所の関係
T-comp	道具デ	主に「で」で表される動作とその手段の関係
F-comp	食材デ	主に「で」で表される動作とその手段の関係
F-eq	同一の食材	既出の食材とそれに対する参照表現
F-part-of	食材の一部	既出の食材とその一部に対する参照表現
F-set	食材の集合	既出の複数種の食材とその全体に対する参照表現
T-eq	同一の道具	既出の道具とそれに対する参照表現
T-part-of	道具の一部	既出の道具とその一部に対する参照表現
A-eq	同一の動作	既出の動作とそれに対する参照表現
V-tm	動作のタイミング	別の動作を行う条件やタイミングを示す句の動詞
other-mod	その他の修飾語句	

5.2 **カロリー推定手法**

ここで本研究におけるカロリー推定手法について説明する。本研究では、前節で説明したレシピフローグラフの特性である、ある食材 (頂点) に着目するときその頂点に向けられた辺を辿ることでその食材がどのような調理動作を施されるかを特定できるという点に着目する。また文部科学省により公表されている日本標準食品成分表 2020 年版 (八訂) を用いることでカロリー推定を行う。日本標準食品成分表は食品の可食部 100 グラムあたりに含まれる成分値をまとめたデータであり、2,478 種類の食品と 54 項目の成分値が収録されている。また、同じ食材においても調理方法 (生、焼き、煮る、油炒め、など) ごとの食材のカロリーも収録されている。これを用いることで調理後食材の正確なカロリーを推定することができる。以下に本研究で提案する手法の流れを示す。また図 3 に提案手法の概要図を示す。

1. レシピの調理手順部分をレシピフローグラフに変換する。
2. レシピフローグラフから食材とその食材に対する調理手順のペアを調べ、「調理後食材」の情報を抽出する。
3. 「調理後食材」に対応する項目を日本食品標準成分表から参照し、分量ごと調理後食材のカロリーを決定する。
4. レシピの食材リストから食材とその分量を参照し、調理後食材と掛け合わせたあと、すべてのカロリーを足し合わせることでレシピ全体のカロリーを推定する。

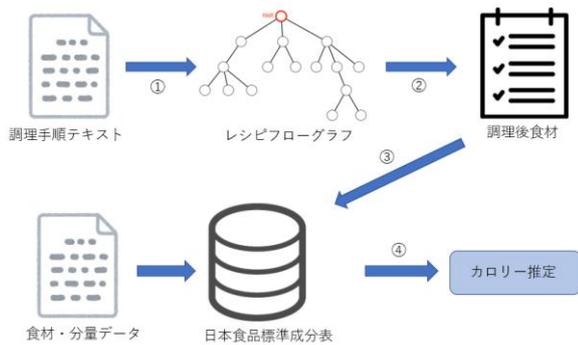


図3 提案手法の概要図

図4～図6で架空のレシピに提案手法を用いた場合の処理のイメージを示す。図4のような調理手順の例に対してフローグラフ化したものが図5である。このフローグラフの食材の頂点のエッジを辿ることで食材とその食材に対して行われる調理の組を得ることができる(図6)。その後カロリーに影響のある調理動作(焼く, ゆでる)のような頂点以外を除去することで、日本食品標準成分表に対応するような食材+調理動作の組を得ることができる。

- 調理手順
1. 豚バラ肉を一口大に切り、にんじん とねぎを細かく切って、フライパンで焼く
 2. ジャガイモとなすを5センチ角に切り、5分茹でる
 3. 1. と2. を混ぜ合わせタレで味付けをする

図4 調理手順の例

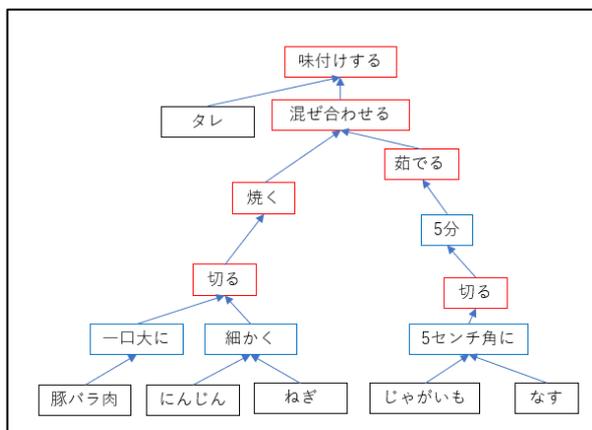


図5 調理手順をフローグラフ化

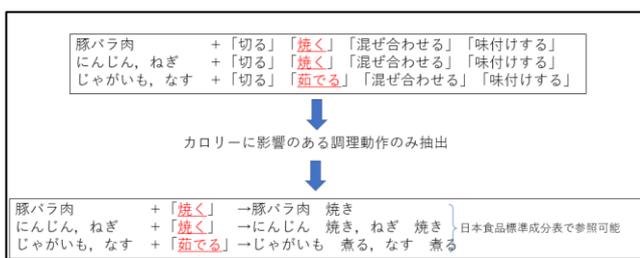


図6 フローグラフを辿り食材と調理の組を抽出

6. 実験方法

6.1 実験手法

アプローチに記載の手法を元の実験を行った。ただし実際には、レシピの手順部分をフローグラフ化するという工程は人手で正確に推定することが難しく、また時間的制約があったためすべてのレシピをフローグラフ化することが困難であったため、レシピを解説し各食材に対してどのような調理手順が施されているかを確認することで調理後食材を得た。その後 Eatreat[12]という日本食品標準成分表2020年版(八訂)のデータを用いてレシピのカロリーを計算するサイトを用いて、調理後食材とその食材の質量を入力しカロリーを推定した。単位は[kcal]である。また正データと比較するために誤差を以下のように割合で定義し各レシピに対して求めた。

$$\text{誤差} = \frac{|\text{推定値} - \text{正データ}|}{\text{正データ}}$$

6.2 使用データ

本実験では味の素が公開している「レシピ大百科」を実験データとして使用した。このレシピサイトは他のユーザー投稿型レシピサイトと同様に、タイトル、食材名および分量、調理手順で構成されている。このレシピサイトの特徴として各レシピに対してカロリーの記載があり、評価基準として用いることができる点が挙げられる。また、ベースラインモデルである加藤ら[9]の実験でも同様のデータを用いているため結果を公平に比較することができる。実際に使用したデータはレシピ大百科のレシピの内、「肉, 魚, 野菜」のタグで検索されたレシピの内ランダム選ばれた100レシピである。

7. 実験結果

6.1で示した手法を用いて100件のレシピに対しカロリーを推定した。100レシピの誤差の平均値, 最小値, 最大値及び中央値を表4に示す。また各レシピの正データと推定値の差を誤差値とし、表5に100レシピの誤差値の平均値, 最小値, 最大値及び中央値を示す。また図7に誤差が降順になるように並べたレシピごとの誤差のグラフを示す。表4に示すように、推定値は正データに対し平均で16.8%の誤差があるという結果が得られた。誤差の分布としては、全レシピの内42レシピが誤差10%以内、62レシピが誤差20%以内に収まる結果となった。また表5に示すように各レシピに対する平均の誤差は62.8kcalとなり、中央値は51kcalとなった。

表4 100 レシピの誤差

誤差(n=100)			
平均値	最小値	最大値	中央値
0.16758	0.00172	0.69436	0.13276

表5 100 レシピの誤差値

誤差値[kcal](n=100)			
平均値	最小値	最大値	中央値
62.76	0.5	240.3	51

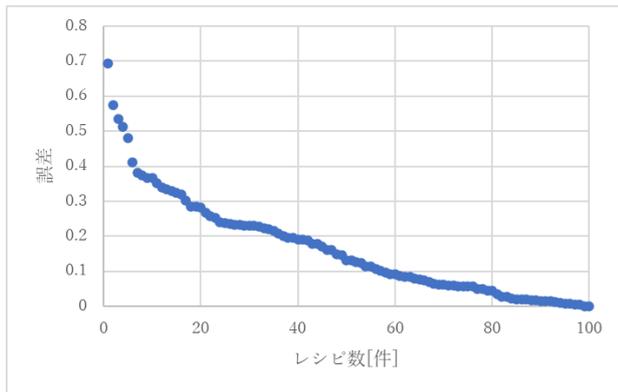


図7 レシピ件数と誤差

8. 評価

ベースラインモデルとの比較として、加藤ら[9]の論文で示されている、誤差が小さい上位7レシピ及び誤差が大きい上位7レシピの結果を本研究の結果と比較する。表6及び表8に本研究における結果のうち、誤差が小さい上位7レシピと誤差が大きい上位7レシピを示す。また、表7及び表9にベースラインモデルにおける結果のうち、誤差が小さい上位7レシピと誤差が大きい上位7レシピを示す。

表6 本研究における誤差の小さい上位7レシピ

料理名	正解カロリー (kcal)	カロリー推定値 (kcal)	誤差 (%)
マグロとチーズで簡単おつまみ	582	583	0.172
山盛りキャベツの絶品豚汁	780	784	0.513
トンテキ	852	857	0.587
豚バラキャベツ鍋	936	943	0.748
豚バラときのこのみぞれ鍋	2256	2273	0.754
豚バラと白菜の重ね鍋	1288	1301	1.009
切り干し大根のごまマヨサラダ	368	373	1.359

表7 ベースラインモデルにおける誤差の小さい上位7レシピ

料理名	正解カロリー (kcal)	カロリー推定値 (kcal)	誤差 (%)
餃子鍋	216	216.09	0.004
チョコレートムース	218	217.92	0.035
味の黒酢鍋照り	223	223.15	0.068
シャキシャキたけのこと卵の塩味炒め	218	217.84	0.073
ブロッコリーの実もサラダ	212	212.19	0.088
きのこごぼうとにらのかき揚げ	226	225.48	0.232
ひじき入り豆腐ハンバーグ風	224	224.65	0.29

表8 本研究における誤差の大きい上位7レシピ

料理名	正解カロリー (kcal)	カロリー推定値 (kcal)	誤差 (%)
白菜と牛肉のすき焼き風	1384	423	69.44
牛薄切り肉の甘辛煮	608	258	57.57
豚こま切れ肉で作る肉豆腐	1048	1610	53.63
すき煮風	622	941	51.29
中華屋さんの肉野菜炒め	444	657	47.97
基本の豚のしょうが焼き	448	632	41.07
レンジローストビーフ	704	436	38.07

表9 ベースラインモデルにおける誤差の大きい7レシピ

料理名	正解カロリー (kcal)	カロリー推定値 (kcal)	誤差 (%)
本格チキンカレー	1128	225.7	79.99
豚の角煮	1040	216.63	79.17
グリル玉ねぎ、ゆでじゃがいも、ソーセージのコンソメハーブクリーム添え	998	223.79	77.58
チキンかつトーストサンド	987	226.49	77.05
サーモンマリネの洋風すし	975	219.77	77.46
ホクホク長いも入りビーフカレー	976	222.96	77.16
ジンジャーブレッド	926	220.97	76.14
シナモン風味の生クリーム添え			

正解のカロリーと誤差の小さい上位7レシピに関して、提案手法ではベースラインモデルを上回る結果を出すことはできなかった。一方で正解のカロリーと誤差の大きい上位7レシピにおいては、提案手法がベースラインを上回り、優位性があると考えられる。ベースラインモデルでは全レシピの平均値が示されていないため、提案手法と比較することはできないが、平均16%程度の誤差はおおむね良好な結果になったと考えられる。

9. 考察

本研究の提案手法を用いた推定値の誤差について考察する。まず誤差が生じてしまった原因として考えられる点として、レシピの食材の表記が曖昧であり、具体的な食材名はユーザーに任せられているという点が挙げられる。例えばあるレシピの食材リストに「牛薄切り肉」という食材名が記載されているとする。今回の提案手法を元にカロリー推定を行う場合、日本食品標準成分表の食材を参照して食材を決定するが、このような食材名であった場合、部位、産地、脂身ありなしなどの条件は指定されていないため、決定すべき食材の候補が複数現れてしまう。また、それらの条件の組み合わせではカロリーの差が2倍近く生まれてしまう場合がある。このような理由から、レシピの食材が日本食品標準成分表のリストと正確に参照できる関係でない場合、その食材の特定はユーザー個人の決め方で変わってしまうため誤差が生じてしまったと考えられる。また本研究の提案手法の軸である、調理後食材を考慮する部分において、本来、どのような食材においてもある調理を施すことでその食材の重量は変化する。例えば生肉100gを加熱した場合、その加熱温度や方法などで多少の変化はあるが、生の状態の肉に対し70~90%程度の質量に減少するこ

とが知られている。本来であればこの質量変化を本提案手法にも取り入れる必要があるが、食材によってその変化率は一定ではなく、また調理手法によっても差異があるため今回は考慮しないで推定を行った。このことも推定カロリーと正解のカロリーの誤差として表れたと考えられる。さらに、正確に食材を決定できた場合においても、レシピ上でその分量が「少々」や「お好みで」等定量的に示されていないようなレシピに関しては筆者が常識的な量を想定し推定を行ったため誤差につながっていると考えられる。

10. まとめ

本研究では、ネット上の料理レシピに対し、レシピの調理手順をフローグラフと呼ばれるグラフに変換し、食材とその食材に対する調理動作を求めることで調理後食材を求め、日本食品標準成分表を参照することで正確なカロリー推定を行う方法を提案した。そして、本研究では味の素提供の”レシピ大百科”上の100レシピに対しカロリー推定の実験を行った。平均で約16%程度の誤差となり、誤差の小ささの項目ではベースラインモデルを上回ることはできなかったが、誤差の大きさの項目ではベースラインモデルに対し優位性を示すことができた。本研究の問題点として、レシピの使用食材を正確に特定することができない点や、調理による質量変化を考慮していないために実際の使用量とは異なる量で計算を行ったために誤差が生じてしまった点などが考えられる。今後はその問題点に対し、使用食材の特定に関しては、そのような問題の起こりやすい肉などの食材を扱っていないレシピに対し、検証を行うといった対策や、考える候補となる食材それぞれに対しカロリーを推定する方法などが考えられる。質量変化の問題に対しては、実験を通して各食材の質量変化率を調査するという方法が考えられる。

謝辞

本研究はJSPS 科研費 JP18H03229, JP18H03340, JP18K19835, JP19H04113, JP19K12107, JP21H03496 の助成を受けたものです。

参考文献

- [1] “食をめぐる意識と実践の現状” .
https://www.maff.go.jp/j/syokuiku/wpaper/pdf/b_1_2_1.pdf, (参照 2022-01-28)
- [2] “調理方法による違い…焼く・煮る・蒸すで料理のカロリーは変わるのか” .
<https://allabout.co.jp/gm/gc/470922/>, (参照 2022-01-28)
- [3] Mori, S., Maeta, H., Yamakata, Y. and Sasada, T.: Flow graph corpus from recipe texts, Proc. LREC14, pp. 2370-2377 (2014)
- [4] Momouchi, Y.: Control structures for actions in procedural texts

- and PT-Chart, Proc. COLING80, pp. 108-114 (1980)
- [5] 荻米志帆乃, 藤井 敦: 料理レシピテキストの構造解析とその応用, 言語処理学会第18回年次大会予稿集, pp. 839-842 (2012)
- [6] Yamakata, Y., Maeta, H., Kadowaki, T., Sasada, T., Imahori, S. and Mori, S.: Cooking recipe search by pairs of ingredient and action - Word sequence v.s. Flow-graph representation, Trans. of JSAL, Vol. 32, No. 1, pp. WII-F 1-9 (2017)
- [7] 西村 太一, 橋本 敦史, 森 信介, 重要語に着目した写真列からのレシピの自動生成, 自然言語処理, 2020年27巻2号, p257-279
- [8] 大野 礼儀, 福原 知宏, 山田 剛一, 増田 英孝, 献立の雰囲気考慮した料理推薦システムの提案, 2015年度人工知能学会全国大会(第29回)
- [9] 加藤 頌大, 上田 真由美, パノット シリアラヤ, 中島 信介, Doc2Vecを用いた料理レシピのベクトル化に基づく自動カロリー計算方式, ファジスティックシンポジウム講演論文集, 34巻, 55-59, 2018-09
- [10] “日本食品標準成分表 2020年度版(八訂)” .
https://www.mext.go.jp/a_menu/syokuhinseibun/mext_01110.html, (参照 2022-01-28)
- [11] “フローグラフコーパス (FG corpus)” .
<http://www.lsta.media.kyoto-u.ac.jp/resource/data/recipe/>, (参照 2022-01-28)
- [12] “Eatreat 栄養価計算/カロリー計算” .
<https://eat-treat.jp/calculation2021>, (参照 2022-01-28)