

分散表現を用いた複数料理レシピからの典型手順の抽出

重田 識博[†] 難波 英嗣[†] 竹澤 寿幸[†]

[†] 広島市立大学大学院 情報科学研究科 〒731-3194 広島県広島市安佐南区大塚東 3-4-1

E-mail: †{shigeta, nanba, takezawa}@ls.info.hiroshima-cu.ac.jp

あらまし 本研究では、ある料理に関する複数のレシピを比較して、その料理の典型的な手順を出力するシステムを構築する。ある料理の典型的な手順は、複数のレシピの手順を比較し、その共通部分を同定することで実現できる。ここで、料理レシピでは、同じ材料や調理器具であっても、レシピによって表記が異なるという問題がある。これらの問題を解決するため、分散表現を用いて手順の同定を行う。本研究では、典型的な手順の出力を複数テキスト要約ととらえ、テキスト要約の評価手法に基づいて、提案手法を評価する。

キーワード 料理レシピ, 複数テキスト要約, 手順要約, 分散表現

1. はじめに

近年、Cookpad や楽天レシピなどのユーザ投稿型レシピサイトの利用が盛んになり、数多くのレシピを検索することができるようになってきている。このようなサイトでは、ある料理に対してユーザ毎に材料の表記や手順の順序が異なるレシピが大量に存在するため、料理スキルの低いユーザにとって、その料理の典型的な作成手順を知るのが必ずしも容易ではないという問題がある。そこで、本研究では、ある料理に関する複数のレシピを入力とし、多くのレシピに共通する手順を典型的な手順として出力するシステムを構築する。また、その典型的な手順を集めて典型的なレシピ出力するシステムを構築する。このようなシステムが実現できれば、本システムで出力する典型的な手順と個々のレシピの手順を比較し、その差異を抽出することで、個々のレシピの特徴を検出することが可能になる。

入力された複数のテキスト間で内容が類似する箇所を認識し、それらを重要箇所として出力するというタスクは、一般に複数テキスト要約と呼ばれており、本研究もその一種であると捉えることができる。複数テキスト要約では、複数のテキストを入力として、テキスト間の類似点と相違点を検出できる。本研究では、複数レシピを入力として、前者は典型的な手順、後者は個々のレシピの特徴として捉え、複数テキスト要約でレシピの要約を試みる。その際、食材や手順の表記が異なる表記揺れという問題がある。内容が類似する箇所を認識する手法は、これまで様々な手法が提案されてきたが、本研究では、Mikolov ら [1] が提案する Paragraph Vector と呼ばれる文の分散表現を用いる。この手法が従来手法よりも優れていることを実験により検証する。

2. 関連研究

手順テキストを対象に要約する研究 [2, 3] やフローチャートや類似度グラフを用いて手順や文章を用いて

要約する研究 [4, 5]、オントロジーや定型表現を用いた従来の代表的な手法 [6] として存在しているが、近年では、分散表現を用いた研究が盛んである。Mikolov ら [7] は自然言語処理の手法の一つであり、従来のアルゴリズムよりも少ない学習時間で飛躍的な精度向上を可能にした Word2Vec を開発している。

Mikolov ら [1] は、さらにこの技術を文に適応して文レベルでの多次元空間のベクトルを表現している。テキストに関して、最も有名な手法として bag-of-words や bag-of-n-grams を挙げられるが、bag-of-words は言葉の順序や単語の意味を考慮しておらず、また bag-of-n-grams は短い文では文脈を考慮するが希少性の高い文や長い文では考慮できない。これらの問題点に対応できる手法として、Paragraph Vector を提案している。Paragraph Vector は、ベクトル表現から次の単語を予測するように学習される 2 つのモデルがある。単語ベクトルとパラグラフ ID を入力として、次の単語を予測するモデルを Distributed Memory Model of Paragraph Vectors (PV-DM) と、パラグラフ ID を入力として、周辺単語を予測するモデルを Distributed Bag of Words version of Paragraph Vector (PV-DBOW) がある。PV-DM では、Word2Vec と同じように単語ベクトルに加えて、文脈や意味情報を記憶したパラグラフ ID も入力することで文をベクトル表現している。また、PV-DBOW では、文脈や意味情報から周辺単語を予測するように学習して文をベクトルで表現している。本研究ではこの手法を適用しており、表層的な情報だけでなく、文の意味を考慮した類似度で手順の対応付けを行っている。

3. 複数料理レシピ要約

関連研究でも述べたように、料理レシピの要約では、ユーザごとに使用する食材や調理器具の表記揺れや、述語や目的語の省略、同じ料理名のレシピでも代用品を用いたレシピが存在するため、同義語辞書や述語項

構造解析、事前知識として料理レシピに使用される食材や調理器具、調理動作などを体系化したオントロジーを用いて、手順の同定が行われてきた。本研究では、典型的なレシピを生成するため、まず、複数のレシピに共通する典型的な手順の同定を行う。典型的な手順の同定では、含意関係認識を拡張した意味的類似度認識のタスクに着目し、料理レシピのテキストを対象に、食材や調理器具の表記揺れや代用品、手順の表記揺れを認識して意味的類似度を算出し、手順の対応付けを行う。本研究では、このタスクを類似手順の検索(タスク 1)として定義する。また、ある料理に関する複数のレシピの類似手順の集合からその料理の典型的な手順を出力して1つの典型的なレシピを作成する。本研究では、このタスクを典型的なレシピの生成(タスク 2)と定義して典型的なレシピが生成できたか評価する。本研究の類似手順の検索(タスク 1)と典型的なレシピの生成(タスク 2)は、図 1 のような位置付けとなる。

3.1 節では、2つのタスクの説明や問題提起について述べ、3.2 節では、類似手順を出力する際に用いる手法について説明する。3.3 節では、3.2 節で述べた結果から、典型的なレシピを生成する手順について述べる。

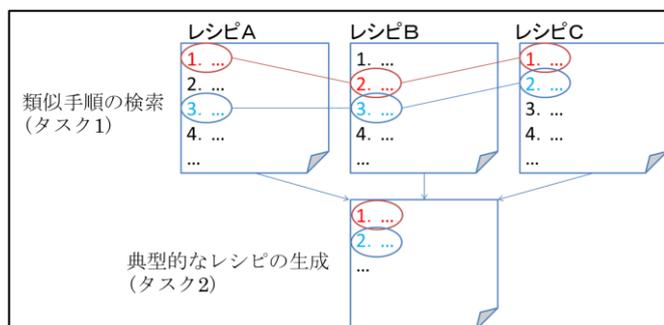


図 1: 典型的な手順の抽出

3.1. 問題提起

図 2 では、「親子丼」に関する 3 つのレシピから、内容が類似していると判断できる手順を列挙したものである。

- ・ 鶏肉は小さめの乱切り、玉ねぎは 3 mm 幅の薄切りにする。
- ・ 鶏肉・たまねぎを適当な大きさに切ります。
- ・ 鶏肉は一口サイズ、タマネギはスライスしておく。

図 2: 3 つの「親子丼」レシピに記載された手順の例

1 節で述べた「テキスト間で内容が類似する箇所を認識する」とは、図 2 に示す 3 文の内容が類似してい

ることを自動認識することを示す。図 2 の例の場合、「鶏肉」や「玉ねぎ」という単語は共通して出現するものの、例えば「切る」という動詞に関して、「小さめの乱切り」、「3mm 幅の薄切り」、「適当な大きさに切る」、「一口サイズに切る」、「スライスする」のように表記が異なる。単語頻度を利用した従来手法で図 2 の 3 文の対応付けを行うと図 3 のように形態素解析の結果から名詞や動詞、形容詞など特定の品詞に絞り、表記が同じ単語の出現回数によって文間の類似度を算出する。2 つの手順間で数多くの同じ単語が共通して出現すれば、それらの手順は内容が類似していると判断する。図 2 では、たまねぎのように読み方は同じでも表記の異なる単語は別の単語とみなしてしまう。よって、従来手法では、図 2 の 3 手順をあまり類似していないと判断する。本研究では、このような問題に対処するため、分散表現を用いて解決を試みる。

- ・ 鶏肉/小さめ/乱切り/玉ねぎ/3/mm/幅/薄切り/する
- ・ 鶏肉/たまねぎ/適当/大きさ/切り
- ・ 鶏肉/一口/サイズ/タマネギ/スライス/し/おく

図 3: 単語頻度を利用した従来手法の対応付け例

3.2. 類似手順の検索(タスク 1)

本研究では、3.1 節の図 2 のようにある料理に関する複数のレシピから手順を 1 つに入力し、他の複数のレシピに存在する入力手順と類似した手順を出力する。その際、3.1 節で述べた問題を考慮する必要がある。タスク 1 において、従来手法をレシピに適用する手法と分散表現を用いる提案手法について述べる。

従来手法

3.1 節で述べたように、数多くの単語の表層情報が共通して出現する数によって、手順は内容が類似していると判断する。代表的な手法として TF-IDF がある。ある文書に含まれる単語の出現頻度 TF とある単語が出現する文書頻度の逆数 IDF によって構成される文書中に含まれる単語の重要度を評価する手法の 1 つである。この手法の欠点として、文長を考慮していないため、文長が長く重要単語を多く含む文が選ばれやすくなる。料理レシピでは、食材や単純な手順の省略が多くあり、1 手順の長さにはばらつきがあるため、この手法は適していない。そこで本研究では、Okapi BM25 を従来手法として用いる。この手法は、TF-IDF の欠点を補うため、平均的な文の長さの値を計算して、その文書がおおよそどれくらいの長さであるかを比率的に計算した値を TF-IDF の式に入力して単語の重要度を評価する。一般的に、TF-IDF よりも良い結果が得

られると言われており、比較手法のベースラインとして用いられている。Okapi BM25 を式(1)に示す。

$$\text{score}(D,Q) = \sum_{q_i \in Q} \text{IDF}(q_i) \left(\frac{\text{TF}(q_i,D) * (k_1 + 1)}{\text{TF}(q_i,D) + k_q * (1 - b + b * \frac{\text{DL}}{\text{avgDL}})} \right) \quad (1)$$

TF:文中の出現頻度

IDF:多くの文に出現しない頻度

DL:文中の総単語数

avgDL:平均的な文の長さ

提案手法

分散表現とは、単語や文章を高次元の実数ベクトルで表現する技術である。近い意味の単語や文章を近いベクトルに対応させて単語の意味を補完する。分散表現には、関連研究で述べたように単語の分散表現を獲得する Word2Vec[7] と文の分散表現を獲得する Paragraph Vector [1] の 2 種類ある。Word2Vec は「類似する文脈で利用される単語は、似た意味を持つ」という仮説に基づき、「単語」の特徴をベクトルで表現する技術である。単語の表層情報だけでは認識できなかった表記揺れに対応するため、Word2Vec を用いた研究が多くある。Word2Vec では図 4 のような、類似した文脈に出現する単語の表記揺れには効果を発揮することができない。図 4 は、人がどれも類似した工程と判断した手順文章であるが、「焼く」以外のすべての手順に共通した単語はなく、「両面」と「片面焼いたら、ひっくり返して焼く」といったフレーズレベルでの対応付けが必要と考えられる。そのため、単語ベクトルだけでは表記揺れに対応できない。本研究では、単語の分散表現では対応できないフレーズレベルに表記揺れにも対応するため、2.2 節で述べた Mikolov ら [1] が提案した文の分散表現を獲得する Paragraph Vector の手法を料理レシピに適用する。具体的には、Mikolov らが提案した 2 つのモデルのうち、PV-DBOW モデルを用いて料理レシピの手順テキストの分散表現を獲得する。Paragraph Vector では、入力長に関係なく実行でき、語の意味的な近さを考慮した類似度計算が可能になる。また、ベクトルを作成する際、単語情報に加えて、語の順番や意味を保持した段落情報も入力して文間の類似度を測ることができる。

- 両面こんがり焼いて完成。
- 両面に焼き目がついたら、蓋をして、火が通るまで焼く。
- フライパン（弱めの中火）で片面に焼き色をつけたら、ひっくり返して弱火にして蓋をして 3 分焼きます。
- 油を少量ひいたフライパンで両面をこんがり焼いたら出来あがりです。

図 4:単語ベクトルで表記揺れが解消できない例

3.3. 典型的なレシピの生成(タスク 2)

類似手順のそれぞれのクラスタから、典型的な手順を 1 つずつ抽出して典型的なレシピの生成を行う。手順のクラスタリングには、汎用的に利用できるデータクラスタリングツール bayon¹を用いて行う。bayon では、Repeated Bisection 法 と K-means 法によってクラスタリングを行うことができるが、本研究では、一般的に広く知られている K-means 法でクラスタリングを行う。あらかじめ料理名ごとに与えられた要約長と同じ数だけ典型的な手順を抽出して典型的なレシピの生成を行う。本研究では、典型的なレシピで作成する際、手順順序を考慮していない。関連研究で述べたように、手順順序は手順テキストを扱う上で重要になるが、本研究では、典型的な手順を抽出して典型的なレシピを構築しており、冗長性なく抽出することに重きをおいている。3.2 節で説明した従来手法と提案手法を構成する値を用いて典型的なレシピの生成を試みる。

従来手法

Okapi BM25 は、入力した文の単語ごとにベクトル値を算出する技術である。各単語に付与されたベクトル値を要素として、典型的な手順をクラスタリングする。クラスタ結果から手順を抽出して典型的なレシピを生成する。

提案手法

Paragraph Vector によって文の分散を取得し、その分散の各次元の値を要素として典型的な手順をクラスタリングする。クラスタ結果から手順を抽出して典型的なレシピを生成する。

4. 実験

・実験データ

本研究では、テキスト自動要約の評価型プロジェクト TSC² の第 3 回(TSC-3)で行われた複数文書要約のタ

¹ <https://github.com/fujimizu/bayon/wiki>

² <http://lr-www.pi.titech.ac.jp/tsc/>

スクに従い、データを作成した。TSC-3 では、1 トピックあたり平均約 10 記事、計 30 トピックについて新聞記事データを用いて、評価用データを構築している。投稿型レシピサイトの 1 つである楽天レシピ(約 80 万レシピ、約 240 万文)の 1 つの料理あたり平均約 10 レシピ、計 35 の料理のレシピデータを抽出し、このうち 5 つの料理をパラメータ調整用データ、残り 30 個の料理名を評価用に用いた。料理名の選定は、幅広く料理文化をカバーできるように考慮しており、図 5 にトピックの一覧を示す。トピックの選定は、和食/洋食/中華、あるいはメインディッシュ/サイドディッシュ/デザートなど、35 トピックが特定のカテゴリの料理に偏らないよう配慮した。

001 酢豚	019 梅酒
002 角煮	020 パンパンジー
003 ミートソースパスタ	021 高野豆腐
004 豚肉のしょうが焼き	022 エビフライ
005 パンプディング	023 大学いも
006 イチゴジャム	024 麻婆豆腐
007 パパロア	025 だし巻き卵
008 みたらし団子	026 おはぎ
009 おからハンバーグ	027 ゴーヤチャンプルー
010 パンプキンスープ	028 水餃子
011 カレーピラフ	029 スパニッシュオムレツ
012 カステラ	030 カスタードプリン
013 あんぱん	031 唐揚げ
014 かぼちゃコロッケ	032 カルボナーラ
015 サバの味噌煮	033 親子丼
016 カキフライ	034 茶碗蒸し
017 切干大根	035 エビマヨ
018 いなりずし	

図 5:複数レシピ要約に用いたトピック一覧

本研究では、3 章で述べた 2 つのタスクについてそれぞれ実験を行った。分散表現には、Python の深層学習ライブラリ gensim³ で Paragraph Vector の手法を実装した Doc2Vec というツールを用いた。

タスク 1:類似手順の検索

このタスクでは、ある料理に関する複数のレシピが 3.1 節の図 2 のように与えられたとき、これらの手順の 1 つを入力として残りの手順を検索する実験を行い、人手によって典型的な手順の対応付けを行った結果と比較評価した。

評価尺度

技術研究を目的に開催されているワークショップ TREC⁴ (Text REtrieval Conference)で用いられている評価ツールを用いて precision と recall で評価した。このツールは、上位 N 件における 11 点補完平均適合率を出力する。本研究では、N=30 に設定した。この N は、今回正解データで用いる全手順が各料理で用いられる平均手順数の約半数を参考に人手で設定した。

³ <https://radimrehurek.com/gensim/models/doc2vec.html>

比較手法

提案手法 (Doc2Vec):分散表現を利用する。

ベースライン手法 (Okapi BM25):文の表層情報を利用する。

タスク 2:典型的なレシピの生成

このタスクでは、3.3 節でクラスタリングした結果から各料理に与えられる要約長に応じて、ランダムに典型的な手順を抽出して典型的なレシピを作成する実験を行い、レシピとして成立しているか評価した。本実験では、ランダムにレシピの生成を行う過程を 100 回実行して、評価した。

評価尺度

TSC-3 で定義されている coverage と accuracy を用いて評価した [8]。coverage は、抽出すべき文がどれだけ抽出されたかを評価する値である。また accuracy は、システムが抽出した文がどれだけ正解かを評価する値である。

比較手法

提案手法 (Doc2Vec):分散表現でクラスタリングされた手順集合を利用した。

ベースライン (Okapi BM25):単語の表層情報でクラスタリングされた手順集合を利用した。

・実験結果

タスク 1:類似手順の検索

類似手順の検索の実験結果を図 6 に示す。また、提案手法を使うことでベースライン手法の抽出精度を改善できたかを調査するため、各手順の抽出精度を比較評価した結果を表 1 に示す。

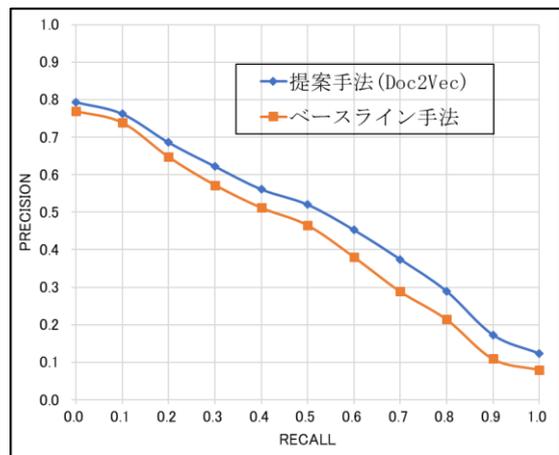


図 6:類似手順の検索の結果

⁴ <http://trec.nist.gov/>

表 1:手法の抽出比較

	改善	同等	改悪	計
ベースライン手法に対する提案手法の抽出比較	665	3	414	1082

タスク 2:典型的なレシピの生成

各料理の典型的なレシピの生成を行い、レシピとして成立しているかを評価したベースライン手法の結果を図 7、提案手法の結果を図 8 に示す。また、表 2 に各手法の平均値を示す。

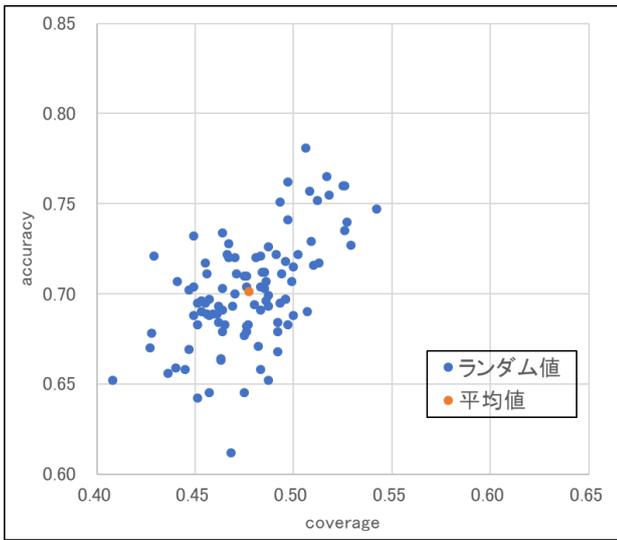


図 7:ベースライン手法で作成したレシピの評価のばらつき

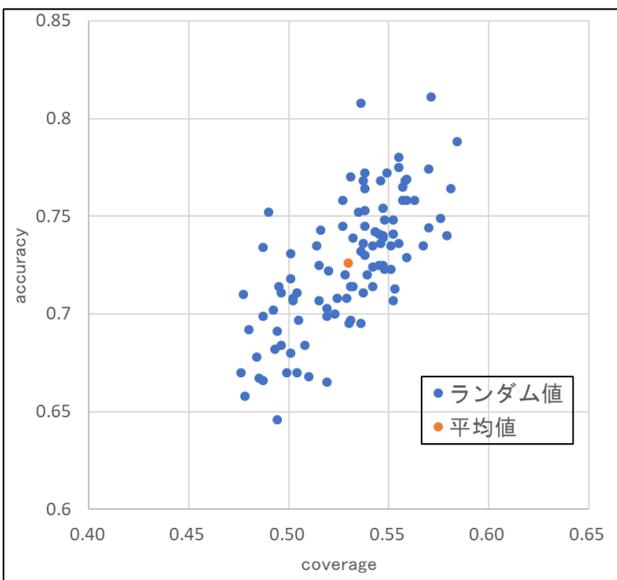


図 8:提案手法で作成したレシピの評価のばらつき

表 2:典型的なレシピ作成の結果

	coverage	accuracy
ベースライン手法	0.477	0.701
提案手法	0.530	0.726

・考察

タスク 1:類似手順の検索の対応付け

提案手法がベースライン手法より良い結果を示した例を図 9 に示す。図 9 の手順が対応していることを人間は判断できるが、ベースライン手法では、このような文の対応付けが困難であった。提案手法では、入力と出力の単語がほとんど一致していない手順でも対応付いている例が多くあった。

- ・ 鍋に鶏肉、玉ねぎ、出汁、調味料を全部入れ、10分位煮込む
- ・ 温まった麺つゆに↑の具材を入れて味がしみ込むまで弱火で煮詰めていきます

図 9:抽出結果の一例

また、対応付けが失敗した例として図 10 では、句点で 1 文に分割しても 1 文中に複数の手順が含まれている手順では、どちらの手法でも対応付けが困難であった。本研究のように手順の内容をできるだけ一致させるために、句点に区切った文の比較を行っているが、その文中に異なる手順の内容を記載して、1 文とした手順を入力し、他の手順を抽出することには、対応できなかったと推測できる。

- ・ 豚肉は 1.5cm 角切りにし下味をつけ卵、水を加えてよくもみ込み片栗粉で絡めてねっとりなるまで混ぜる
- ・ 豚こま肉は 2cm 角くらいにきって、醤油、酒各大 2 つつまぜしておく。

図 10:対応付けが困難な一例

タスク 2:典型的なレシピの生成

提案手法で典型的なレシピを生成した結果からベースライン手法より提案手法が上手く典型的なレシピを作成した例と提案手法が典型的な手順を抽出できなかったレシピの例を用いて考察を行った。

◆ ベースライン手法より提案手法が上手く典型的なレシピを作成した事例

図 11、図 12 に典型的なレシピの生成結果例を示す。本実験では、抽出した手順の順番を考慮していない。手順集合内の下線の手順が提案手法で抽出されており、「*」の食材を元のレシピから参照すれば、調理

可能であることがわかった。また、ベースライン手法と比較したところ、ベースライン手法で「片栗粉をつけて鶏肉を揚げる」手順が抽出できているが「長芋」を混ぜた手順は典型的な手順でないため、不正解となっている。手順内容に冗長性があるものが選ばれていない点で提案手法の方が、典型的なレシピとして成立している。

- フライパンなどに油を入れて、低温から肉を揚げる。
- 鶏肉は一口大に切って*印の調味料に 30 分以上漬けておく。

図 11:提案手法で抽出された典型的なレシピ

- 鶏肉を好きな大きさに切り、ビニール袋又はボールに入れ、★の調味料を入れてよく揉む。
- × **長芋**に片栗粉を入れてまぜ、180℃の油で揚げる。

図 12:ベースライン手法で抽出された典型的なレシピ

◆ 提案手法が典型的な手順を抽出できなかったレシピ事例

図 13 に提案手法で「カキフライ」の典型的なレシピの生成の結果、間違った抽出をした手順を示す。

- × 小さい牡蠣の場合は 2 個を一緒に胡椒がふつてない面をくっつけて合わせる。

図 13:「カキフライ」の典型的なレシピの結果

しかし、この出力した手順が属している手順集合の他の手順の内訳をみると、図 14 であった。今回、抽出してしまった手順以外、すべて正解手順と判断された手順がクラスタリングされている。従って、分散によってクラスタリングされた手順からランダム抽出によって偶然間違った手順を抽出したことが分かり、分散によるクラスタリングは有効であることが分かります。

- × 小さい牡蠣の場合は 2 個を一緒に胡椒がふつてない面をくっつけて合わせる。
- 牡蠣は洗って水気をしっかりふき取る
- 1 を取り出しキッチンペーパーでよく水気をきる
- カキは水でふり洗いして、キッチンペーパーで丁寧に水分のふきとりし、軽く胡椒をふる
- ペーパータオルで水気をしっかりふきとり、片面にコショウをふる

図 14:分散表現で分類された手順集合

5. おわりに

本研究では、ある料理に関する複数のレシピを入力とし、多くのレシピに共通する手順を典型的な手順として出力するシステムを構築し、類似手順の検索を行

った。また、その結果から典型的な手順を集めて典型的なレシピ出力するシステムを構築し、どちらの実験でも従来手法より提案手法が優れている結果が得られた。類似手順の検索では、従来手法で抽出できない手順を抽出できていることが確認できた。また、提案手法で出力された典型的なレシピでは、分散によるクラスタリングが良い結果を示した。

6. 今後の課題

レシピ生成でのクラスタ数や別の抽出手法を検討することで、精度向上が期待できる。また、手順内容をさらに一致されるには、レシピに用いる食材や調理器具の統一したレシピを事前に判別する組み込む必要がある。入れる食材や使う調理器具が異なれば、手順表現が異なり、手順数に違いが生じる。本研究では、手順位置を考慮せず、実験や評価を行っているため、手順位置を考慮することで実際に調理可能な典型的なレシピの生成を目指す。

参 考 文 献

- [1] T. Mikolov, and Q. Le, “Distributed Representations of Sentences and Documents”, In Proceedings of the The 31st International Conference on Machine Learning, ICML 2014, pp.1188-1196, 2014.
- [2] 瀧本 洋喜, 笹野 遼平, 高村 大也, 奥村 学, “施設配置問題に基づく同一料理のレシピ集合からの基本手順の抽出”, 言語処理学会第 21 回年次大会, 2015.
- [3] 高木 優, 藤井 敦, “手順テキストを対象とした比較対象要約”, 言語処理学会第 21 回年次大会発表論文集, pp.573-576, 2015.
- [4] 山肩 洋子, 今堀 慎治, 森 信介, 田中 克己, “ワークフロー表現を用いたレシピの典型性評価と典型的なレシピの生成”, 電子通信学会論文誌 J99-D(4), pp.378-391, 2016.
- [5] 北島 理沙, 小林 一郎, “潜在トピックを考慮したグラフ表現に基づく複数文書要約”, 日本知能情報フジイ学会論文誌, Vol.25, No.6, pp.914-923, 2013.
- [6] H. Nanba, Y. Doi, M. Tsujita, T. Takezawa, and K. Sumiya, “Construction of a cooking ontology from cooking recipes and patents”, In Proceedings of Workshop on Smart Technology for Cooking and Eating Activities, CEA2014, 2014.
- [7] T. Mikolov, I. Sutskever, K. Chen, G. Corrado, and J. Dean, “Distributed Representations of Words and Phrases and their Compositionality”, In Proceedings of the 26th Neural Information Processing Systems, NIPS 2013, pp.3111-3119, 2013.
- [8] T. Hirao, M. Okumura, T. Fukushima, and H. Nanba, “Text Summarization Challenge 3 - Text Summarization Evaluation at NTCIR Workshop 4 -”, In Proceedings of the 4th NTCIR Workshop, 2004.