

部分グラフに基づくレシピフローグラフ分散表現の比較評価

Comparative evaluation on subgraph-based distributed representations of recipe flow graphs

二宮 あかり¹ 尾崎 知伸^{1,2} *
Akari Ninomiya¹ Tomonobu Ozaki^{1,2}

¹ 日本大学 大学院総合基礎科学研究科

¹ Graduate School of Integrated Basic Sciences, Nihon University

² 日本大学 文理学部

² College of Humanities and Sciences, Nihon University

Abstract: Recipe flow graphs are graph-based representation of cooking procedures. Since recipe flow graphs hold detailed information on cooking recipes, distributed representations which reflect various aspects on recipe flow graphs precisely are expected to be useful for a wide range of applications on cooking activities. In this paper, to obtain useful insight for obtaining good distributed representation of recipe flow graphs, we compare and evaluate experimentally several distributed representations derived in various settings.

1 はじめに

近年、健康志向の高まりと SNS の発展に伴い、クックパッド¹ や楽天レシピ² に代表されるような料理メディアに関するコミュニティーサイトが日常的に利用されるようになってきている。これらのサイトにおいて、レシピ推薦や類似レシピ検索、代替食材の提案など、利用者が持つ種々の高度な要求に精度良く応えるためには、蓄積された調理レシピを分析するとともに、データ全体を分類・構造化することが必要となる。

調理レシピは通常、タイトル、食材リスト、調理手順の3つで構成される。このうち本研究では、調理手順に対するグラフ表現であるレシピフローグラフ [1, 2, 3] を対象とする。詳しくは後述するが、レシピフローグラフは調理手順の意味構造を詳細に保持することが可能であり、高度なサービスにおいて必要不可欠な役割を果たすことが期待できる。これまでに文献 [4] において、レシピ中のパターンを用いたレシピフローグラフに対する分散表現の獲得が議論されている。本研究は [4] を出発点とし、レシピフローグラフに対するより適した分散表現の獲得を目的に、(1) 利用パターン、(2) 基礎となる分散表現獲得技術、(3) ベクトルの次元数に関して多様な設定を行い、それぞれから獲得される分散表現を実験的に比較・評価する。

本論文の構成は以下の通りである。2章で関連研究について言及する。3章では、レシピフローグラフについて導入を行った後、文献 [4] に基づき、レシピフローグラフに対する分散表現獲得手法を説明する。4章において、種々の設定で獲得される分散表現を比較・評価し、最後に5章でまとめを行う。

2 関連研究

これまでに、レシピデータを対象とした分散表現技術の適用に関して種々の研究が行われている。文献 [5] では、代替食材発見タスクに、代表的な分散表現獲得技術である word2vec [6] を適用している。同様に文献 [7] では、代替食材発見タスクに関し、word2vec を用いて得られる食材ベクトルと doc2vec [8] を用いて得られるレシピテキストベクトルを併用することを提案している。さらに文献 [9] では、レシピの郷土料理スタイルの変更に分散表現技術を適用している。一方文献 [10] では、レシピ検索精度の向上を目的に、調理画像とレシピテキストを同一空間に埋め込むマルチモーダルニューラルモデルを提案している。

異なる面からの関連研究として、分散表現獲得技術の構造データへの適用に関しては、部分構造に着目した手法がいくつか提案されている。文献 [11] では、ギャップ制約を伴う系列パターンを用い、系列データに対する分散表現を獲得する手法が提案されている。同様に、文献 [12] や [4] では、頻出部分グラフパターンを用い

*連絡先：日本大学文理学部情報科学科
〒156-8550 東京都世田谷区桜上水 3-25-40
E-mail: tozaki@chs.nihon-u.ac.jp

¹<http://cookpad.com/>

²<http://recipe.rakuten.co.jp/>

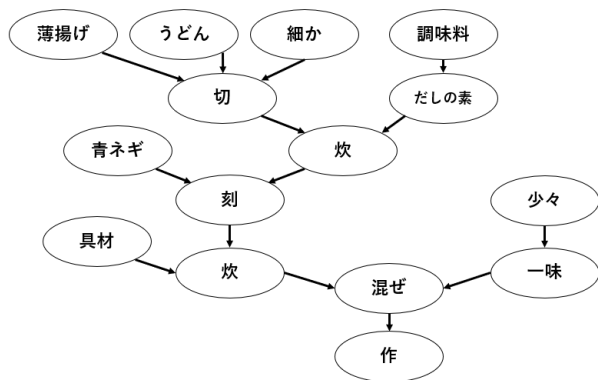


図 1: レシピフローグラフの例

たグラフトランザクションのベクトル化に関して議論している。

3 レシピフローグラフの分散表現

3.1 レシピフローグラフ

レシピは通常、タイトル、食材リスト、調理手順の3つで構成される。レシピフローグラフ [1, 2, 3] は、有向非巡回グラフを用いた調理手順に対する意味内容の構造化表現であり、食材や道具、継続時間などのレシピ用語を保持する頂点と、レシピ用語間の関係を表す有向辺から構成される。8種の頂点と13種の辺が準備され、頂点間に(1)述語項関係がある場合、(2)同種の2つの固有表現の間に何らかの参照関係がある場合、(3)2つの動作の間に何らかの関係がある場合のいずれかの関係が認められる場合に辺が張られる。図1にフローグラフの例を示す。

3.2 頻出部分グラフパターン

ラベルの全体集合を L としたとき、レシピフローグラフ $r = \langle V_r, E_r, l_r \rangle$ とは、頂点集合 V_r と辺集合 $E_r \subseteq V_r \times V_r$ 、ラベル関数 $l_r : V_r \cup E_r \rightarrow L$ の3項組で表現される有向グラフである。2つのグラフ $p = \langle V_p, E_p, l_p \rangle$ と $q = \langle V_q, E_q, l_q \rangle$ に関し、下記の条件を満たす関数 $f : V_p \rightarrow V_q$ が存在するとき、 p を q の部分グラフと呼び、 $p \subset q$ と表記する。

1. $\forall u \in V_p [l_p(u) = l_q(f(u))]$
2. $\forall (u, v) \in E_p \exists (f(u), f(v)) \in E_q$ s.t.
 $l_p(u, v) = l_q(f(u), f(v))$

複数のレシピフローグラフに共通して含まれる部分グラフをパターンと呼ぶ。また、 n 個のレシピフローグラフで構成されるデータベース $R = \{r_1, \dots, r_n\}$ に対し、パターン p の支持度を

$$\text{sup}(p, R) = |\{r \in R \mid p \subset r\}| / |R|$$

と定義する。パラメタとして最小支持度 σ ($0 < \sigma \leq 1$) が与えられたとき、条件 $\text{sup}(p, R) \geq \sigma$ を満たすパターンを頻出パターンと呼ぶ。また、データベース R に対する頻出パターンの集合を F と表記する。支持度に関しては逆単調性 ($p \subset q \rightarrow \text{sup}(p, R) > \text{sup}(q, R)$) が成り立つため、 F には類似する頻出パターンが多数含まれる可能性がある。類似パターンを排除するために、本研究では、支持度に着目した圧縮表現を利用する。パラメタ $\delta \geq 0$ に対し、条件

$$\forall q \in F (q \subset p \rightarrow \text{sup}(q, R) > (1 + \delta) \times \text{sup}(p, R))$$

を満たす頻出パターン p を、頻出 δ フリーパターンと呼ぶ。頻出 δ フリーパターンは、同程度の支持度を持つパターンを同値類と見做した場合の極小元に相当する。一方で、同程度の支持度を持つパターンを同値類と見做した場合の極大元に相当するパターンとして、下記の条件を満たす頻出パターン p を、頻出 δ 飽和パターンと呼ぶ。

$$\forall q \in F (q \supset p \rightarrow \text{sup}(q, R) < (1 - \delta) \times \text{sup}(p, R))$$

頻出 δ フリーおよび飽和パターンの集合をそれぞれ G , C と表記する。このとき $G, C \subseteq F$ が成り立つ。

3.3 パターンを用いた分散表現の獲得

文献 [4] と同様、本論文では、レシピフローグラフが含むパターンを利用して各レシピフローグラフの分散表現を獲得する。

レシピフローグラフデータベース R とパターン集合 $FP \in \{F, G, C\}$ に対し、各レシピフローグラフが含むパターンの集合をトランザクションとする新たなデータベース

$$D_{R,FP} = \{ \langle r, \{p \in FP \mid p \subset r\} \rangle \mid r \in R \}$$

を考える。データベース $D_{R,FP}$ に対し、word2vec[6] や GloVe[13], StarSpace[14] などの既存の分散表現獲得手法 M を適用することで得られるパターン $p \in FP$ に対する n 次元実数ベクトル (分散表現) を \overline{p}_n^M と表記する。また、レシピフローグラフ $r \in R$ に対する分散表現 $\overline{r}_n^M = \sum_{p \in FP, p \subset r} \overline{p}_n^M$ は、 r が含むパターンに対するベクトルの総和として獲得する。

以降では、パターン集合 $FP \in \{F, G, C\}$ に対して手法 M を適用することで得られるレシピフローグラフの n 次元ベクトル表現の集合を

$$FP_n^M = \{\vec{r}_n^M \mid r \in R\}$$

と表記する。

4 実験と考察

4.1 データセット

本研究では、クックパッド株式会社が国立情報学研究所と協力して研究者に提供しているクックパッドデータセット³ と、それらを基に山肩らが構築したレシピフローグラフコーパス⁴ を対象に、レシピタイトルがそれぞれ「オムライス」、「コロッケ」、「ハンバーグ」、「チャーハン」で終わるレシピを利用した。データセットの基本情報を表1に示す。

頻出パターン集合 F の獲得には、代表的な頻出部分グラフマイナーである $gSpan[15]$ を用いた。なお、最小支持度は $\sigma = 10/33,678$ を採用している。また、 F に対する後処理として、 $\delta = 0.05$ とする頻出 δ フリーパターン集合 G と頻出 δ 飽和パターン集合 C を獲得した。また、抽出できた頻出部分グラフは、全体で 21,502 件、頻出 δ フリーパターン集合は 21,089 件、頻出 δ 飽和パターン集合は 21,100 件である。

分散表現の獲得には、次元数を $n \in \{100, 200, 300\}$ とし、モデルとして StarSpace (以降 SS) と GloVe (以降 GV) を用いた。 SS は、単語や文書に加え、グラフや画像に対しても統一的な枠組みでベクトル表現を獲得することのできる高い汎用性を特徴とする。一方 GV は、文書全体の単語の共起に着目したモデルであり、単語の意味と構文的な規則性を明示的に捉えることを特徴とする。

表 1: データセットの概要

料理種	レシピ数	平均 食材数	平均 手順数	平均 ノード数
オムライス	4,013	9.96	6.34	52.20
コロッケ	7,734	8.85	6.01	49.43
ハンバーグ	12,934	10.63	6.03	51.13
チャーハン	8,997	8.30	4.84	38.71
合計	33,678	9.52	5.76	47.55

³<https://www.nii.ac.jp/dsc/idr/cookpad/cookpad.html>

⁴<http://www.ar.media.kyoto-u.ac.jp/data/recipe/>

4.2 多次元尺度法による視覚化

各分散表現の分布を確認するため、コサイン距離に基づく多次元尺度法を用いて視覚化を行った。結果を図2に示す。図中において、“O”はオムライス、“C”はコロッケ、“T”はチャーハン、“H”はハンバーグに関するレシピフローグラフを表す。また視認性向上のため、各料理種からランダムに選択した100件のみを表示している。

同一モデルを用いた場合、頻出パターンの種類や次元数による分布の差異は確認できない。その一方で、モデルによって分布に大きな差異があることが分かる。 SS では、オムライスとチャーハン、コロッケとハンバーグが重なり合うように分布しているのがわかる。これに対し GV では、料理種の重なり具合が判別しにくく、一点に集中するような密集した分布に対して大きな外れ値を含んでいることが伺える。

4.3 Jaccard 係数を用いた比較

獲得された各分散表現を比較するために、2つの分散表現間で同一レシピに対するコサイン類似度上位100レシピ集合をそれぞれ算出し、それらの Jaccard 係数を算出した。得られた Jaccard 係数の平均値を表2にまとめる。

SS 同士での Jaccard 係数平均の最大値は $F_{300}^{SS} - G_{300}^{SS}$ 間の 0.290、 GV 同士での最大値は $F_{300}^{GV} - G_{300}^{GV}$ 間の 0.842 であり、最大値に関しては SS 同士の方が低いことが分かる。また、 SS 間の Jaccard 係数平均は全体としても GV 間のそれよりも値が低い。これらの結果より、 SS を用いた場合、次元数、パターン種によって得られる分散表現が大きく異なることが分かる。それに対し GV を用いた場合は、次元数、パターン種の影響は大きくないことも伺える。一方、 $SS - GV$ 間での Jaccard 係数平均は 0.05 を下回っており、利用するモデルが獲得ベクトルに対して大きな影響を持つことが確認できる。

4.4 料理種の分類

分散表現を説明変数、料理種をクラスとする分類問題を通じ、獲得された分散表現の比較を行う。分類器として SVM (Support Vector Machine) を採用し、約 34,000 件のデータのうち約 27,000 件をトレーニングデータ、残りの約 7,000 件をテストデータとした。各手法に対する分類結果 (適合率, 再現率, F 値) を表3に示す。

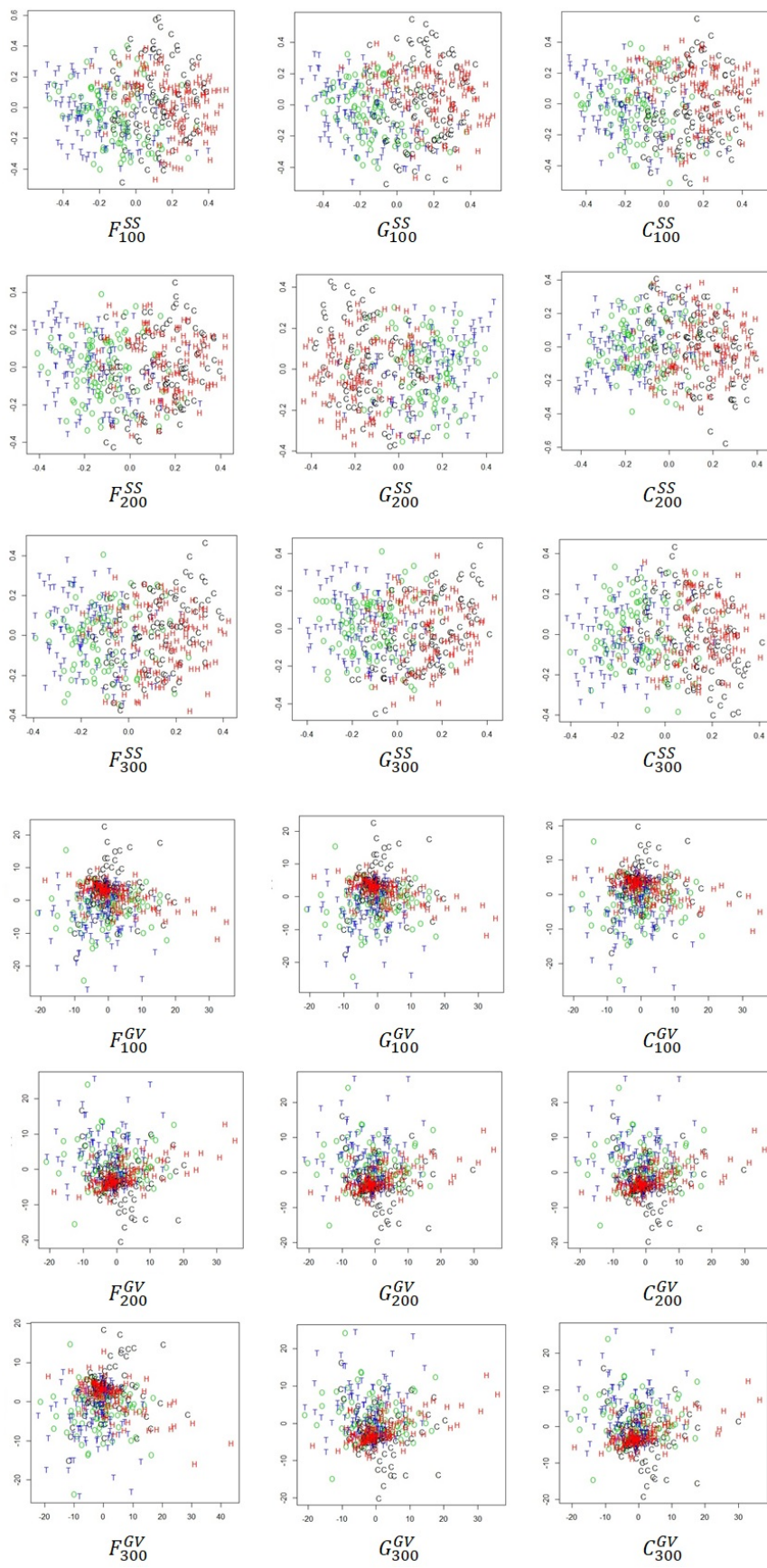


図 2: 多次元尺度法による視覚化

表 2: Jaccard 係数の平均

	F_{100}^{SS}	F_{200}^{SS}	F_{300}^{SS}	G_{100}^{SS}	G_{200}^{SS}	G_{300}^{SS}	C_{100}^{SS}	C_{200}^{SS}	C_{300}^{SS}	F_{100}^{GV}	F_{200}^{GV}	F_{300}^{GV}	G_{100}^{GV}	G_{200}^{GV}	G_{300}^{GV}	C_{100}^{GV}	C_{200}^{GV}	C_{300}^{GV}
F_{100}^{SS}		0.175	0.187	0.144	0.174	0.186	0.145	0.174	0.186	0.036	0.034	0.033	0.036	0.034	0.032	0.036	0.033	0.032
F_{200}^{SS}	0.175		0.257	0.174	0.231	0.256	0.174	0.231	0.255	0.041	0.038	0.036	0.041	0.038	0.036	0.041	0.037	0.036
F_{300}^{SS}	0.187	0.257		0.255	0.255	0.289	0.186	0.256	0.289	0.043	0.040	0.038	0.043	0.040	0.038	0.043	0.040	0.038
G_{100}^{SS}	0.144	0.174	0.255		0.174	0.187	0.145	0.175	0.186	0.035	0.033	0.031	0.035	0.033	0.031	0.035	0.032	0.031
G_{200}^{SS}	0.174	0.231	0.255	0.174		0.257	0.175	0.232	0.256	0.040	0.038	0.036	0.040	0.038	0.036	0.040	0.037	0.035
G_{300}^{SS}	0.186	0.256	0.289	0.187	0.257		0.186	0.257	0.290	0.042	0.039	0.038	0.042	0.039	0.038	0.042	0.039	0.037
C_{100}^{SS}	0.145	0.174	0.186	0.145	0.175	0.186		0.175	0.187	0.035	0.033	0.031	0.035	0.033	0.031	0.035	0.032	0.031
C_{200}^{SS}	0.174	0.231	0.256	0.175	0.232	0.257	0.175		0.257	0.040	0.037	0.036	0.040	0.037	0.035	0.040	0.037	0.035
C_{300}^{SS}	0.186	0.255	0.289	0.186	0.256	0.290	0.187	0.257		0.043	0.040	0.038	0.043	0.039	0.037	0.042	0.039	0.037
F_{100}^{GV}	0.036	0.041	0.043	0.035	0.040	0.042	0.035	0.040	0.043		0.749	0.679	0.588	0.591	0.569	0.583	0.606	0.591
F_{200}^{GV}	0.034	0.038	0.040	0.033	0.038	0.039	0.033	0.037	0.040	0.749		0.842	0.594	0.640	0.637	0.591	0.664	0.669
F_{300}^{GV}	0.033	0.036	0.038	0.031	0.036	0.038	0.031	0.036	0.038	0.679	0.842		0.571	0.636	0.650	0.569	0.662	0.688
G_{100}^{GV}	0.036	0.041	0.043	0.035	0.040	0.042	0.035	0.040	0.043	0.588	0.594	0.571		0.745	0.670	0.610	0.635	0.616
G_{200}^{GV}	0.034	0.038	0.040	0.033	0.038	0.039	0.033	0.037	0.039	0.591	0.640	0.636	0.745		0.839	0.610	0.699	0.705
G_{300}^{GV}	0.032	0.036	0.038	0.031	0.036	0.038	0.031	0.035	0.037	0.569	0.637	0.650	0.670	0.839		0.585	0.697	0.729
C_{100}^{GV}	0.036	0.041	0.043	0.035	0.040	0.042	0.035	0.040	0.042	0.583	0.591	0.569	0.610	0.610	0.585		0.723	0.658
C_{200}^{GV}	0.033	0.037	0.040	0.032	0.037	0.039	0.032	0.037	0.039	0.606	0.664	0.662	0.635	0.699	0.697	0.723		0.842
C_{300}^{GV}	0.032	0.036	0.038	0.031	0.035	0.037	0.031	0.035	0.037	0.591	0.669	0.688	0.616	0.705	0.729	0.658	0.842	

GVを用いた場合の精度はSSと比較し大きく低下している。これは、多次元尺度法による視覚化でも確認された一点に集中するような分散表現の値や大きな外れ値に起因すると考えられる。SSを用いた場合の精度は全体的に高い傾向にあるが、オムライスを対象とした場合は例外であり、その適合率は0.6程度、再現率が0.4程度となった。また、チャーハンと判別されるケースが多数確認された。これは、調理内容においてオムライスのご飯を卵で包む際のご飯がチャーハンの調理とほぼ同様であることが要因であると考えられる。同様にコロケとハンバーグでも溶き卵を似たような方法で利用するためか、コロケとハンバーグ間において判別ミスが多かった。

利用パターンの違いによる精度の差は確認できないが、次元数に関しては、 $n = 100$ に比べて $n = 200, 300$ の方が精度が高い傾向が伺える。

5 まとめと今後の課題

本論文では、レシピフローグラフを適切に表現する分散表現の獲得を目的に、利用パターンと次元数、基礎となる手法に関する種々の設定を用いて分散表現を導出し、それらを実験的に比較・評価した。

今後の課題としては、代替食材発見などの異なるタスクを用いた評価と、種々のグラフ分散表現獲得手法[16]や非ユークリッド空間への埋め込み手法[17]との比較があげられる。

謝辞： レシピフローグラフをご提供頂きました東京大学の山肩洋子氏、京都大学の森信介教授に感謝いたします。また本研究では、クックパッド株式会社と国立情報学研究所が提供する「クックパッドデータ」を利用しました。本研究の一部はJSPS科研費17K00315の助成を受けたものです。

参考文献

- [1] S. Mori, H. Maeta, Y. Yamakata, and T. Sasada : Flow graph corpus from recipe texts, *Proc. the 9th International Conference on Language Resources and Evaluation*, pp.2370–2377 (2014)
- [2] S. Mori, H. Maeta, T. Sasada, K. Yoshino, A. Hashimoto, T. Funatomi, and Y. Yamakata : FlowGraph2Text: Automatic sentence skeleton compilation for procedural text generation, *Proc. of the 8th International Natural Language Generation Conference*, pp.118–122 (2014)
- [3] Y. Yamakata, H. Maeta, T. Kadowaki, T. Sasada, S. Imahori and S. Mori : Cooking recipe search by pairs of ingredient and action, *人工知能学会論文誌*, Vol.32, No.1, pp.1–9 (2017)
- [4] A. Ninomiya and T. Ozaki : Learning Distributed Representation of Recipe Flow Graphs via Frequent Subgraphs, *Proc. of the 11th Workshop on Multimedia for Cooking and Eating Activities*, pp.25–28 (2019)
- [5] 野沢 健人, 中岡 義貴, 山本 修平 : word2vecを用いた代替食材の発見手法の提案, *信学技報*, Vol.114, No.204, DE2014-30, pp.41–46 (2014)
- [6] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado and J. Dean : Distributed Representations of Words and Phrases and their Compositionality, *Advances in Neural Information Processing Systems 26*, pp.3111–3119 (2013)
- [7] 梅本 晴弥, 豊田 哲也, 大原 剛三 : 料理レシピの分散表現を用いた代替食材の発見手法, *行動変容と社会システム*, Vol.03, 2018-03-01 (2018)
- [8] Q. Le and T. Mikolov : Distributed representations of sentences and documents, *arXiv preprint*, arXiv:1405.4053v2 (2014)
- [9] M. Kazama, M. Sugimoto, C. Hosokawa, K. Matsushima, L. R. Varshney, and Y. Ishikawa : A neural network system for transformation of regional cuisine style, *Frontiers in ICT*, Vol.5, No.14 (2018)
- [10] A. Salvador, N. Hynes, Y. Aytar, J. Marin, F. Ofli, I. Weber, and A. Torralba : Learning cross-modal embeddings for cooking recipes and food images, *Proc. of the 2017 IEEE Conference on Computer Vision and Pattern Recognition*, pp.3068–3076 (2017)
- [11] D. Nguyen, W. Luo, T. D. Nguyen, S. Venkatesh, and D. Phung : Sqn2Vec: Learning sequence representation via sequential patterns with a gap constraint, *Proc. of the 2018 Joint European Conference on Machine Learning and Knowledge Discovery in Databases, Part 2*, pp.569–584 (2018)
- [12] D. Nguyen, W. Luo, T. D. Nguyen, S. Venkatesh, and D. Phung : Learning graph representation via frequent subgraphs, *Proc. of the 2018*

表 3: SVM の結果

n	料理種	適合率	再現率	F 値	適合率	再現率	F 値	適合率	再現率	F 値
		F^{SS}			G^{SS}			C^{SS}		
100	オムライス	0.60	0.39	0.47	0.56	0.33	0.42	0.58	0.35	0.44
	コロッケ	0.85	0.81	0.83	0.84	0.80	0.82	0.85	0.79	0.82
	チャーハン	0.78	0.88	0.82	0.76	0.87	0.81	0.76	0.87	0.81
	ハンバーグ	0.85	0.89	0.87	0.84	0.89	0.86	0.84	0.89	0.87
200	オムライス	0.61	0.41	0.49	0.62	0.41	0.50	0.60	0.41	0.49
	コロッケ	0.86	0.81	0.84	0.86	0.79	0.82	0.84	0.80	0.82
	チャーハン	0.78	0.88	0.83	0.78	0.88	0.83	0.79	0.87	0.83
	ハンバーグ	0.85	0.89	0.87	0.85	0.90	0.87	0.85	0.93	0.89
300	オムライス	0.61	0.41	0.49	0.63	0.43	0.51	0.60	0.40	0.48
	コロッケ	0.86	0.80	0.83	0.92	0.82	0.87	0.86	0.80	0.83
	チャーハン	0.78	0.88	0.83	0.78	0.87	0.82	0.78	0.87	0.83
	ハンバーグ	0.85	0.90	0.87	0.86	0.90	0.88	0.88	0.90	0.89
		F^{GV}			G^{GV}			C^{GV}		
100	オムライス	0.24	0.01	0.02	0.20	0.01	0.01	0.24	0.01	0.02
	コロッケ	0.81	0.42	0.56	0.83	0.41	0.54	0.83	0.40	0.54
	チャーハン	0.61	0.72	0.66	0.61	0.73	0.66	0.61	0.72	0.66
	ハンバーグ	0.59	0.87	0.70	0.59	0.88	0.70	0.58	0.87	0.70
200	オムライス	0.25	0.01	0.02	0.23	0.01	0.02	0.38	0.01	0.02
	コロッケ	0.81	0.43	0.56	0.84	0.56	0.67	0.83	0.40	0.54
	チャーハン	0.61	0.71	0.66	0.61	0.72	0.66	0.61	0.72	0.66
	ハンバーグ	0.59	0.87	0.70	0.66	0.87	0.75	0.58	0.88	0.70
300	オムライス	0.23	0.01	0.02	0.24	0.01	0.02	0.30	0.01	0.02
	コロッケ	0.81	0.43	0.57	0.82	0.41	0.55	0.82	0.41	0.55
	チャーハン	0.62	0.71	0.66	0.55	0.72	0.62	0.61	0.72	0.66
	ハンバーグ	0.59	0.88	0.70	0.58	0.88	0.70	0.58	0.88	0.70

SIAM International Conference on Data Mining, pp.306–314 (2018)

- [13] J. Pennington, R. Socher and C. D. Manning : GloVe: Global Vectors for Word Representation, *Proc. of the 2014 Conference on Empirical Methods in Natural Language Processing*, pp.1532–1543 (2014)

- [14] L. Wu, A. Fisch, S. Chopra, K. Adams, A. Bordes, and J. Weston : StarSpace:Embed all the things!, *Proc. of 32nd AAAI Conference on Artificial Intelligence*, pp.5569–5577 (2018)

- [15] X. Yan and J. Han : gSpan : Graph-based substructure pattern mining, *Proc. of the 2002 IEEE International Conference on Data Mining* pp.721–724 (2002)

- [16] H. Cai, V. W. Zheng, and K. C.-C. Chang : A Comprehensive Survey of Graph Embedding: Problems, Techniques and Applications. arXiv:1709.07604 (2017)

- [17] M. Nickel and D. Kiehl, Poincaré Embeddings for Learning Hierarchical Representations, arXiv:1705.08039 (2017)