

Thing2Vec: 人の行動理解に向けたモノの使用特性のベクトル埋め込み方法の提案

田仲百音^{†1}
大阪大学

岸野泰恵^{†2}
日本電信電話株式会社, 大阪大学

下西英之^{†3}
大阪大学

1. はじめに

近年、センシング技術の向上により、サーマルカメラの利用が拡大している。それに伴い、これまでデータ化できなかった情報を、スマートハウスなどで活用できる可能性が高まってきている過去の研究では、人が触れた箇所をサーマルカメラを用いて検出することで、感染症対策に活用する取り組みが行われている [1]。この技術と物体検知を組み合わせることで、日常的に触れた物体の使用履歴を記録できるようになる。そこで我々は、接触した物体の履歴を蓄積することで、モノの使われ方を分析できるのではないかと考えた。

モノの使われ方を分析するための手法としては、モノの分散表現を作成するアプローチを提案する。分散表現は、非構造化データをベクトルで表現する手法であり、ベクトル埋め込みとも呼ばれる。埋め込まれたデータ点のうち、近しい性質を示すものは、多次元空間上において互いに近い場所に位置するという特徴がある。最近では、文章、商品などさまざまなものを対象としたベクトル埋め込みの手法が提案され、分析や推薦などの応用に活かす研究が行われている。そのため、モノの使われ方をベクトルで表現することによって、異常検知や生活パターンの特定、人の行動理解など幅広い応用が期待できる。そこで我々は、モノの使われ方がモノを特徴づける情報として重要であるという仮説のもと、モノの使用情報を用いてモノの分散表現を作成する Thing2Vec を提案する。Thing2Vec では、Word2Vec のモデルをもとに、入力として物体 ID を、出力にはモノの使用情報を与えて学習を行うことで、中間層の潜在変数を分散表現として得ることができる。

実験では、可視光カメラとサーマルカメラを使用してモノの使用情報を取得し、得られた情報をもとに分散表現を作成した。生成された分散表現をクラスタリングし分析を行った結果、生成された分散表現は、モノの使用情報に基づ

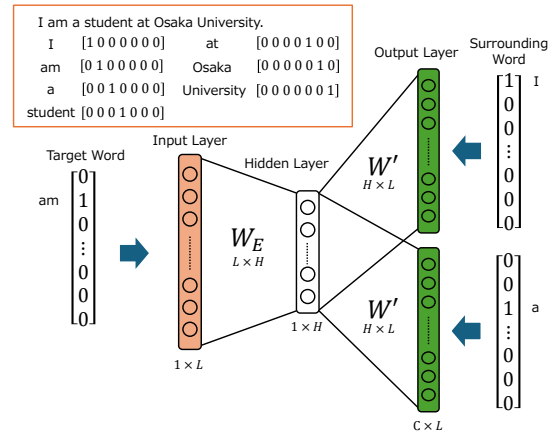


図1 Word2Vec のアーキテクチャ

Figure 1 The Architecture of Word2Vec

いたクラスタを形成することが分かった。今後、より大規模なデータで学習することでモノの使われ方に関する分析や、人の行動理解に活用できる可能性がある。

2. 関連研究

これまでに様々な非構造化データを分散表現として扱う手法が提案されてきた [2, 3, 4, 5]。

Word2Vec [6, 7] は、単語をベクトル埋め込みする手法の一つである。文章内の注目単語について、その周囲の単語を予測するタスクを解くことで単語の分散表現を作成する。Word2Vec には、bag-of-words と Skip-gram の 2 つのモデルがあり、Skip-gram の方が高精度である。Skip-gram モデル (図 1) は入力層と 1 つの隠れ層、出力層からなる。ここで、図中の L はベクトルに埋め込む単語数、 H は分散表現の次元数、 C は出力層に与えられる周辺単語数、すなわちウィンドウサイズである。簡単のため、図中の出力層に与えられている周辺語は前後 1 単語であるが、実際の周辺語数 C は任意に設定することができる。Word2Vec では、以下の 3 つのステップによって学習が行われる。

1. 入力層には単語の対応する位置にフラグが設定された one-hot ベクトルを入力し、隠れ層で重み行列 W_E と

Thing2Vec: Proposal of Object Usage Pattern Embedding Toward Human Activity Understanding.

^{†1} MONE TANAKA, Osaka University

^{†2} YASUE KISHINO, Nippon Telegraph and Telephone Corporation, Osaka University

^{†3} HIDEYUKI SHIMONISHI, Osaka University

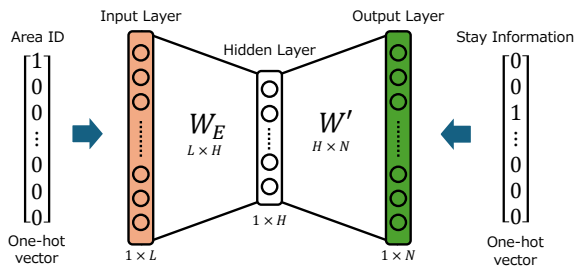


図2 Area2Vec のアーキテクチャ
Figure2 The Architecture of Area2Vec

- の行列積に変換する。
2. 隠れ層の出力を、重み行列 W' との行列積に変換した後、出力層の softmax 関数によって各単語の確率分布として解釈する。
 3. 得られた確率分布と、出力層に与えられる周辺語の one-hot ベクトル間の交差エントロピー損失を計算し、重みを更新する。

ここで、入力層に与えられる入力 は ont-hot ベクトルであるため、入力層から隠れ層への重み行列の各行を、その単語の分散表現として扱うことができる。したがって、学習された重み行列から分散表現が得られる。実験の結果、大量のデータで学習して得られたベクトルは、単語間の関係を意味することが分かった。

Barkan らは、Word2Vec をレコメンデーションシステムに応用する手法として Item2Vec [8] を提案した。協調フィルタリングにおいて、アイテムをユーザが選択した集合として扱うが、ユーザとアイテムのセット関係に関する情報が提供されない場合がある。そこで Item2Vec では、アイテムベース協調フィルタリングのため、Word2Vec における単語をアイテムに、単語列をアイテム集合と置き換えて学習が行われる。Skip-gram を改良した Skip-gram with Negative Sampling (SGNS) を用いて学習することで、ユーザの情報が未開示であるシナリオに対し、アイテム間の関連性を推測することを可能にした。実験の結果、Item2Vec は SVD (特異値分解) モデルと比較して、高精度なジャンル分類を行えることがわかった。

Area2Vec [9] は注目エリアについて、曜日、到着時刻、滞在時間からなる滞在情報に基づいた分散表現を作成する。Skip-gram モデルをもとに考案された、Area2Vec のアーキテクチャを図2に示す。Area2Vec では、50メートル四方に切り取られた区画ごとにエリア ID が割り振られ、エリア ID に対応する one-hot ベクトルが入力層に入力される。隠れ層の出力は、Word2Vec と同様に各エリアの滞在情報の確率分布として解釈される。この構成で学習することで、出力層から得られる確率分布と、離散化された滞在情

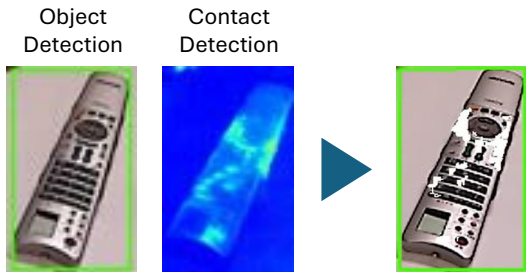


図3 物体検知・接触検知の例
Figure3 The Example of Object and Contact Detection.

報の one-hot ベクトルとの交差エントロピー損失が計算され、重み行列が更新される。実験では、学習した重み行列から得られた分散表現を k-means 法でクラスタリングすることで、繁華街や住宅エリア、オフィスエリアなどを視覚的に区分し、人にとって意味のある分析結果が得られることが分かった。

3. Thing2Vec

近年、IoT 技術やユビキタスコンピューティングの発展により、センシング技術が向上している。個々のモノに関するセンシング技術も進歩しており、使用時刻など、これまで収集されていなかったデータを収集できるようになった。しかし、日常的に使用するモノに対し、それぞれのデータを組み合わせて分析を行う研究は少ない。一方で、実生活において、生活に密着したモノは多種多様である。そこで我々は、さまざまなモノの使われ方を同時に分析することで生活パターンの特定や人の行動理解に活用できるのではないかと考えた。

本研究では、モノの使用情報からモノの分散表現を作成する Thing2Vec を提案する。例えば、仕事のある平日に使うモノと、休日にプライベートで使うモノは異なる。通勤している人の場合、鍵や財布などは平日の朝夕の決まった時間に取りられることが多いだろう。このような情報から分散表現を作成することで、生活パターンの特定や異常検知への応用が期待できる。

我々の提案手法では、可視光カメラとサーマルカメラを用いてモノの使用情報を取得し、Thing2Vec によって使用情報に基づいたモノの分散表現を作成する。本章では、それぞれの手順について詳細に説明する。

3.1. 物体の検出

物体の検出においては、一般的な画像処理アルゴリズムを用いて可視光カメラから物体を検出する。また、人が触れて実際に使った物体であるかを、サーマルカメラによる

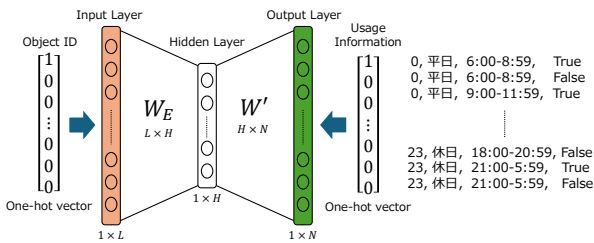


図 4 Thing2Vec のアーキテクチャ
Figure 4 The Architecture of Thing2Vec

熱痕跡を用いて判断する。人は恒温動物であり、通常、環境表面より体温が高いため、人がモノに触れると物体表面に熱が残る。熱痕跡と呼ばれるこの熱が、検出した物体上にあれば、人が触れたものとする (図 3)。人の日常物の操作を分析するためには、物体が単にカメラ画像中に撮影されたか、実際に触れたかの区別が重要であるため、サーマルカメラを併用する方法を取る。

検知された物体はログに記録した後、同一物体であれば一連の物体履歴として統合する。以上の処理により、物体履歴として、物体の出現時間、消失時間、接触の有無を記録していく。

3.2. モデル

提案する Thing2Vec は、Area2Vec モデルを参考に考案したものである。Thing2Vec は 3 層のネットワークから構成され、入力層には各物体の ID に対応した one-hot ベクトルを、出力層には、離散化されたモノの使用情報に対応する one-hot ベクトルを与える。入力層と隠れ層の間の重み行列は $L \times H$ の行列 W_E であり、隠れ層と出力層の間の重み行列は $H \times N$ の行列 W' である。ここで、 L はベクトルに埋め込むモノの数、 N は出力層のノード数 (すなわちモノの使用情報の組み合わせ総数)、 H は分散表現の次元数を示す。

3.3. 定義

我々は、モノの使用情報として「物体 ID」、「曜日」、「出現時刻」、検出された時間内に人が触れたかどうか、すなわちモノが使用されたかを意味する「使用状態」の 4 つを使用した。人がモノに触れて置いたのか、長時間置かれていた物の上にあったものが移動されて、そのモノが撮影されるようになったのかによって、人の物の使い方に対する認識は異なると考えられる。そこでモノが使用される曜日や時刻が、モノを特徴づける上で重要だと考え、使用状態を導入した。サーマルカメラを用いて使用状態を取得し、Thing2Vec に入力することで、より人の認識に近い分散表現になる可能性がある。我々は、表 1 に示すようにモノの使用情報を定

表 1 モノの使用情報の定義
Table 1 Definition of Object Usage Information

モノの使用情報	値
物体 ID	0-23
曜日	平日, 休日 (土曜日と日曜日)
出現時刻	6:00-8:59, 9:00-11:59, 12:00-14:59, 15:00-17:59, 18:00-20:59, 21:00-5:59
使用状態	True, False

義し、それぞれを離散的な値に割り当てた。

3.4. 埋め込み

前節の定義によって離散化結果は、物体 ID、曜日、出現時刻、使用状態の全組合せである $24 \times 2 \times 6 \times 2 = 576$ のカテゴリに振り分けられる。入力層には物体 ID に対応した one-hot ベクトルが、出力層には、576 カテゴリに振り分けられたモノの使用情報に対応する one-hot ベクトルが与えられる。Thing2Vec モデルは、以下の 3 つのステップによって学習を行う。

1. 入力層には物体 ID の対応する位置にフラグが設定された one-hot ベクトルを入力し、隠れ層で重み行列 W_E との行列積に変換する。
2. 隠れ層の出力を、重み行列 W' との行列積に変換した後、出力層の softmax 関数によって各使用情報の確率分布として解釈する。
3. 得られた確率分布と、出力層に与えられる周辺語の one-hot ベクトル間の交差エントロピー損失を計算し、重みを更新する。

学習後、学習された重み行列から分散表現が得られる。

4. 実験

Thing2Vec がモノの特徴を捉えているかを確認するため、小規模な実験を行った。本章では、実験設定および実験の結果について述べる。

4.1. 物体検出の実装

実装上、物体の検出には、我々が過去に提案した物体検知・記録システム [10] を使用する。提案手法は、本来個別の物体を追跡することを想定しているが、個々の日常的な物体を認識して追跡するアルゴリズムを即座に実装することは困難であったため、既存の物体検知アルゴリズムである YOLO-World [11] を使用した。従来の YOLO では事前に定義された語彙の物体しか検知することができなかったが、YOLO-

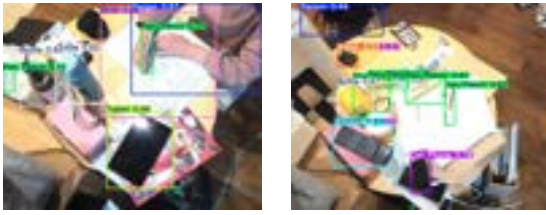


図5 実験画像の例.

Figure 5 The example of experiment images.

World は柔軟な語彙を指定することができる. 今回の実験では, 日常的に使用するモノとして, YOLO-World に以下の 24 語彙を与えた. Person, Glasses, Bottle, Cup, Handbag, Book, Umbrella, Watch, Pen/Pencil, Cell Phone, Clock, Head Phone, Remote, Scissors, Folder, Earphone, Mask, Tissues, Wallet, Tablet, Key, Stapler, Eraser, Lipstick.

実験では, YOLO-World の出力する YOLO ラベルを「物体 ID」として扱った. 検知された物体はログに記録された後, バウンディングボックスの IoU と HSV 色空間の Hue 成分によって同一物体かどうか判定され, 最終的に一連の物体履歴へ統合される.

4.2. データセット

我々は著者の 1 名と 3 名の家族が住む自宅で日常生活の一部を撮影し, 10 日間で 24 シーンを撮影した. 撮影には, 解像度 480×640 の可視光カメラ (GAZO GZ300), 解像度 480×512 のサーマルカメラ (FLIR Boson+ AS50deg) を用いた. 撮影された画像の例を図 5 に示す. ここで, 前節で説明したような画像処理および物体履歴の生成を行った結果, 計 173 の物体履歴が得られた.

4.3. 結果

モデルは 200 エポックで学習され, 分散表現の次元は参考^{*1}に示された以下の式に基づいて 5 に設定された.

$$\text{embedding_dimensions} = \text{number_of_categories}^{0.25} \quad (1)$$

カテゴリ数は $24 \times 2 \times 6 \times 2 = 576$ であるため, $576^{0.25} \approx 4.9$ より 5 次元とした. 得られた分散表現は, クラスタリングによって解釈可能な形に変換される. 本実験では, 代表的なクラスタリング手法である k-means と DBSCAN [12] を用いた. DBSCAN では, ある一定範囲内に 4 点以上集まっているものをクラスタと定義し, クラスタに含まれない点はノイズとした.

得られた分散表現を DBSCAN 法でクラスタリングした結果, 17 クラスタとノイズに分類された. 5 次元の分散表

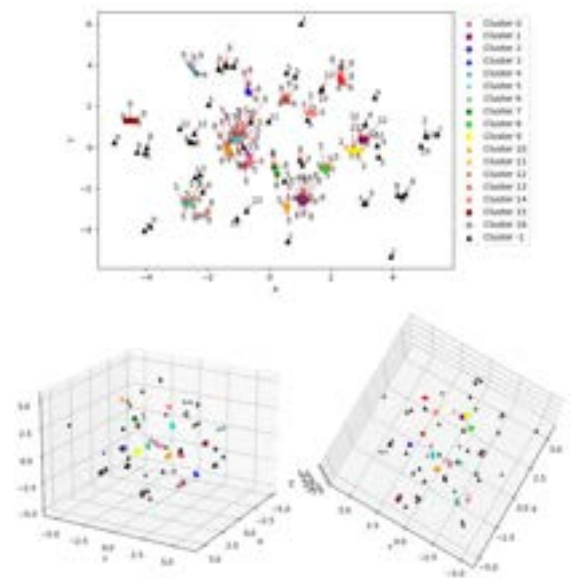


図6 DBSCAN クラスタリングを行った結果. 図中の数字は, 「物体 ID」すなわち, 検知された YOLO ラベルを示している.

Figure 6 2D and 3D visualization of DBSCAN clustering results. The numbers in the figure indicate the "object ID", i.e., the detected YOLO labels.

現を DBSCAN クラスタリングした結果について, 主成分分析を行って 2 次元と 3 次元のグラフに視覚化したものを図 6 に示す. ここで, 図中の数字は「物体 ID」すなわち, YOLO-World の出力した YOLO ラベルを示している. 以降では, 物体 ID とラベルを同義とする.

クラスタ 0, 5, 6, 8, 11, 14 はラベル 5: Book の集合である. これらのクラスタ間の違いは, 出現時刻, 曜日, 使用状態による. クラスタ 0 は平日の 18-21 時の間に使用された本, クラスタ 5 は平日の 9-12 時の間に使用された本, クラスタ 6 は平日の 9-12 時の間に出現した本, クラスタ 8 は平日の 15-18 時の間に使用された本, クラスタ 11 は休日の 15-18 時の間に使用された本, クラスタ 14 は休日の 15-18 時の間に出現した本を意味する. ここでは, クラスタ 5 とクラスタ 6 は冬休み中に朝から小学生が本を使って遊んでいるクラスタであり, 他は授業期間中に宿題を夕方に行うクラスタであった.

クラスタ 2, 3, 7, 10, 12, 16 はラベル 8: Pen/Pencil の集合である. クラスタ 2 は平日の 18-21 時の間に出現したペン, クラスタ 3 は平日の 18-21 時の間に使用されたペン, クラスタ 7 は平日の 9-12 時の間に出現したペン, クラスタ 10 は平日の 15-18 時の間に出現したペン, クラスタ 12 は休日の 15-18 時の間に出現したペン, クラスタ 16 は平日の 9-12 時の間に使用されたペンである. ここでは, 宿題をし

^{*1} <https://developers.googleblog.com/en/introducing-tensorflow-feature-columns/>

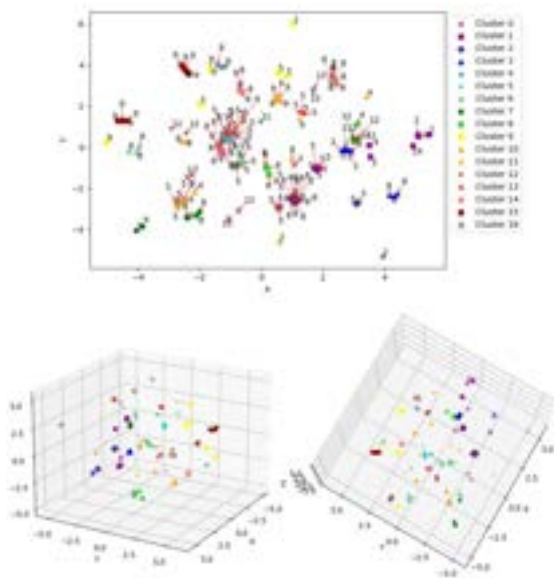


図 7 k-means クラスタリングした結果. 図中の数字は, 「物体 ID」すなわち, 検知された YOLO ラベルを示している.

Figure 7 2D and 3D visualization of k-means clustering results. The numbers in the figure indicate the "object ID", i.e., the detected YOLO labels.

ている途中のペン (クラス 3) と机に置かれたままのペン (クラス 2) が別のクラスに分類された.

クラス 4, 13, 15 はラベル 9: Cell Phone の集合である. クラス 4 は平日の 18-21 時の間に使用された携帯電話, クラス 13 は休日の 15-18 時の間に出現した携帯電話, クラス 15 は平日の 18-21 時の間に出現した携帯電話である. カメラ撮影範囲内に携帯電話をよく充電する場所が含まれており, 平日においては 18 時以降の帰宅後に充電するが, 休日は 15 時から出現しており, 携帯電話の使い方, 場所の違いがクラスに表れていることがわかる.

クラス 1 はラベル 12: Remote のクラス, クラス 9 はラベル 2: Bottle のクラスである. クラス-1 (ノイズ) には, 出現回数の少ないデータが含まれる.

図 7 に k-means クラスタリングの結果を示す. ここでは, DBSCAN クラスタリングと比較するために $k = 17$ とした. k-means は DBSCAN と比較して, 異なるラベルを同じクラスに分類することが多く, DBSCAN の方が本研究の目的に沿ったクラスタリングが可能であることを確認した.

次に, 図 8, 9 に離散化したデータを分散表現に変換せず, 直接クラスタリングした結果を示す. つまり, 表 1 の定義に従って値を割り振り, 物体 ID, 曜日, 出現時刻, 使用状態に関する 4 次元配列をクラスタリングした. k-means ク

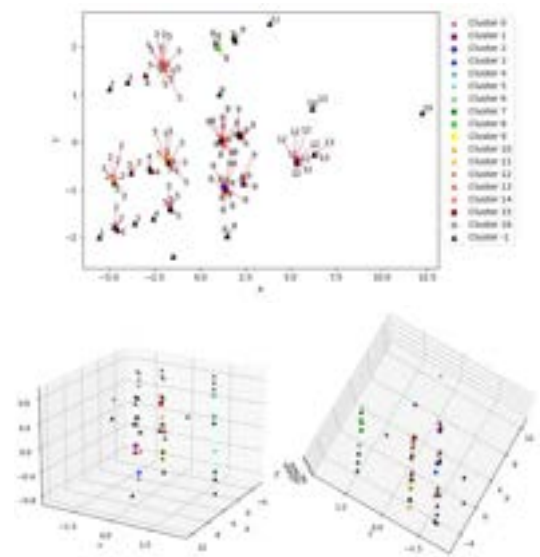


図 8 分散表現に変換せず, データを直接 DBSCAN クラスタリングした結果. 図中の数字は, 「物体 ID」すなわち, 検知された YOLO ラベルを示している.

Figure 8 2D and 3D visualization of DBSCAN clustering results without Thing2Vec. The numbers in the figure indicate the "object ID", i.e., the detected YOLO labels.

ラスタリングでは, 時間帯や曜日は分類できているものの, 2 と 3, 12 と 13 といったように, 隣接する物体ラベルが同じクラスに分布しているケースが多かった. これは, 離散化されたデータを直接プロットすると, 隣接するラベルが空間上で近い位置に配置されるためである. しかし, 曜日や時刻とは異なり, 隣接したラベルに意味はない. また, DBSCAN は隣接したラベルを同じクラスとすることなくクラスタリングできたが, データが増えるにつれ, クラスの収束に時間がかかる. したがって, 得られたデータを直接クラスタリングする方法では不十分であり, データを分散表現に変換することが重要である.

5. 議論

実験の結果, Thing2Vec によって得られた分散表現は, 物体 ID, 曜日, 出現時刻, 使用状態という使用情報に基づいたクラスを形成することが確認できた. また, 宿題の最中のペンの置かれ方の状況によって, 別のクラスに分類されることが分かった. したがって, たとえば宿題中のクラスの遷移回数から, 宿題の難易度 (ペンを置いて考えた頻度), 宿題に対する集中度 (途中の離席回数) を分析できる可能性がある. 携帯電話については, 平日と休日で充電される時間帯が異なることが別クラスに分かれ, 何時に在宅しているかを判別できることが明らかになった. こ

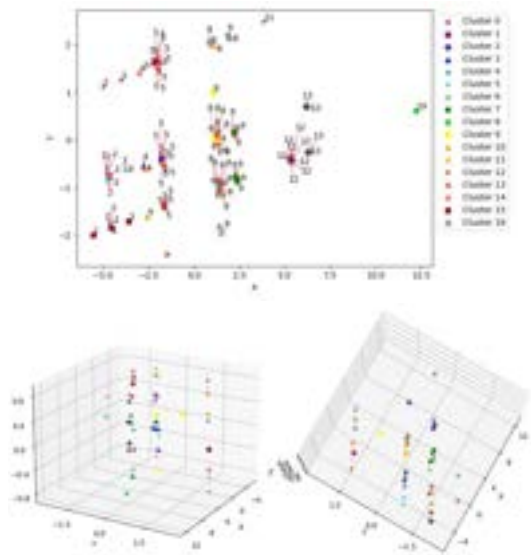


図9 分散表現に変換せず、データを直接 k-means クラスタリングした結果。図中の数字は、「物体 ID」すなわち、検知された YOLO ラベルを示している。

Figure 9 2D and 3D visualization of k-means clustering results without Thing2Vec. The numbers in the figure indicate the "object ID", i.e., the detected YOLO labels.

のような人の無意識な行動の変化を捉えることができれば、さまざまな分析に活かすことができるだろう。さらに冬休み中の例外的な行動が別クラスに分類されたことから、普段と違う行動の変化や異常を検出することもできるのではないかと考えている。ノイズに分類されたデータについては、データセットに出現時間や物体ラベルに関する偏りがあるため、データ量を増やせば新たなクラスを形成する可能性がある。k-means クラスタリングの結果に関しては、少ないデータの特徴を表現するのに出力層の 576 という次元が大きすぎたため、データが粗になる部分が多く、クラスタリングが意図しない結果になってしまったと考えられる。データを増強することで、さらに次元を増やし、月日やモノの位置など本実験で導入しなかった使用情報も埋め込める可能性がある。また、本実験では実装の都合上、YOLO ラベルと物体 ID を同義として扱ったが、個々の物体を追跡できれば、それぞれのモノについてより詳細な分析が可能となるだろう。

分散表現に変換せずに k-means を実行すると、互いに隣り合う異なるラベル (8 と 9 など) が同じクラスに分類されてしまうことがあった。これは、離散化されたデータを直接プロットすると、隣接するラベルが空間的に近接して配置されるためである。DBSCAN は離散化されたデータでも分類できたが、データ数と次元数が増えると収束に時間がかかってしまう。したがって、モノの使用情報を提案

する Thing2Vec を用いて分散表現に変換することで、物体履歴を人の理解に近い形でクラスタリングできるようになり、スケーラビリティも高まる。

Thing2Vec によって得られた分散表現は、異常検知や生活パターンの識別、人の行動理解など幅広い応用が期待できる。今回は少ないデータでの検証であったためクラスタリングが形成されることの確認にとどまったが、より大規模なデータセットを用意することでモノの使い方に関する新たな知見が得られるのではないかと期待している。また、個人ごとのデータを収集することで、生活習慣のクラスタリングや、家族構成によるクラスタリング、家庭状況の変化の検出などが可能になると考えられる。

本実験では固定カメラによるデータ収集を行い、固定された場所でのモノの使い方の情報を収集したが、ウェアラブルカメラを活用したデータ収集方法も考えられる。ウェアラブルカメラを用いると、個人の視点に近いデータを収集できるようになり、個人が使用するモノに関する分析ができるようになる。本手法においては、データの収集方法がボトルネックであるが、室内を常時撮影するような監視やライフログのようなアプリケーションと併用すれば、収集は容易になると考えている。

6. おわりに

本論文では、モノの使用情報に基づいて、モノの使用特性をベクトル埋め込みする手法 Thing2Vec を提案した。Thing2Vec は、Area2Vec を参考にしたモノの分散表現を作成方法であり、「物体 ID」、「曜日」、「出現時刻」、「使用状態」を用いてモノの使用情報をベクトルに埋め込む。モノの使用情報は、固定した可視光カメラとサーマルカメラで撮影した画像に対し、物体検知および接触検知を行うことで収集する。

実験では、実験者の日常生活の一部を撮影し、Thing2Vec によってモノの使用特性を捉えることができるか検証を行った。実験結果、Thing2Vec によって獲得された分散表現は、使用情報に基づいたクラスを形成することが確認できた。今後の取り組みとしては、データをさらに増強し生活パターンの特定や異常検知、人の行動理解に活用したいと考えている。

参考文献

- [1] Kishino, Y., Shirai, Y., Yanagisawa, Y., Ohara, K., Mizutani, S. and Suyama, T.: Identifying human contact points on environmental surfaces using heat traces to support disinfect activities: poster abstract, *Proc. of the Conference on Embedded Networked Sensor Systems* (2020).

- [2] Le, Q. and Mikolov, T.: Distributed representations of sentences and documents, *International conference on machine learning*, PMLR, pp. 1188–1196 (2014).
- [3] Grover, A. and Leskovec, J.: node2vec: Scalable feature learning for networks, *Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 855–864 (2016).
- [4] Narayanan, A., Chandramohan, M., Venkatesan, R., Chen, L., Liu, Y. and Jaiswal, S.: graph2vec: Learning distributed representations of graphs, *arXiv preprint arXiv:1707.05005* (2017).
- [5] Reddy, P., Gharbi, M., Lukac, M. and Mitra, N. J.: Im2vec: Synthesizing vector graphics without vector supervision, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 7342–7351 (2021).
- [6] Mikolov, T.: Efficient estimation of word representations in vector space, *arXiv preprint arXiv:1301.3781*, Vol. 3781 (2013).
- [7] Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S. and Dean, J.: Distributed representations of words and phrases and their compositionality, *Advances in neural information processing systems*, Vol. 26 (2013).
- [8] Barkan, O. and Koenigstein, N.: Item2vec: neural item embedding for collaborative filtering, *2016 IEEE 26th International Workshop on Machine Learning for Signal Processing (MLSP)*, IEEE, pp. 1–6 (2016).
- [9] 庄子之和, 青木俊介, 米澤拓郎, 河口信夫ほか: 大規模ユーザの滞在情報に基づくエリアの特徴付けと COVID-19 による影響分析, *情報処理学会論文誌*, Vol. 62, No. 10, pp. 1644–1657 (2021).
- [10] Tanaka, M., Kishino, Y. and Shimonishi, H.: Efficient Object Retrieval System Using Contact History from a Thermal Camera, *IEEE 43rd International Conference on Consumer Electronics ICCE2025* (2025).
- [11] Cheng, T., Song, L., Ge, Y., Liu, W., Wang, X. and Shan, Y.: YOLO-World: Real-Time Open-Vocabulary Object Detection, *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)* (2024).
- [12] Ester, M., Kriegel, H.-P., Sander, J., Xu, X. et al.: A density-based algorithm for discovering clusters in large spatial databases with noise, *kdd*, Vol. 96, No. 34, pp. 226–231 (1996).