

複数のコグニティブAPIを活用した宅内コンテキスト認識

～ 多数決アプローチ ～

陳 思楠[†] 佐伯 幸郎[†] 中村 匡秀^{†,††}

[†] 神戸大学 〒657-8501 神戸市灘区六甲台町 1-1

^{††} 理化学研究所・革新知能統合研究センター 〒103-0027 東京都中央区日本橋 1-4-1

E-mail: [†]chensinan@ws.cs.kobe-u.ac.jp, ^{††}sachio@carp.kobe-u.ac.jp, ^{†††}masa-n@cs.kobe-u.ac.jp

あらまし 宅内の居住者や環境に関する状況情報（総じて宅内コンテキストと呼ぶ）をいかに精度よく認識するかは、従来より重要な研究課題とされ、特にユビキタス・コンピューティングの分野で長年研究されてきた。我々は、コグニティブAPIが提供する汎用的な画像認識と、軽量の機械学習を組み合わせることで、一般家庭にも導入可能なコンテキスト認識手法を実現すべく研究を進めている。先行研究では、複数のAPIの認識性能を比較・調査し、単一APIによる特徴量を用いた宅内コンテキスト認識手法を提案した。しかしながら、評価実験において、食事や遊び、全体ミーティング等、多人数が参加するコンテキストの区別が難しいという問題点を残している。本研究の目的は、一般家庭への導入容易性を維持しながら、従来区別が難しい宅内コンテキストの認識精度をより一層向上させることである。提案手法では、複数のコグニティブAPIを利用し、各APIによる特徴量を用いた認識モデルを構築する。これらモデルの認識結果を多数決することで、コンテキストの分類精度を向上させることを狙う。深層学習を用いるアプローチに比べ、提案手法は強力なマシンパワーや大量のデータラベリングを必要しない。これにより、宅内の様々な状況情報を高精度にかつ効率よく認識することが期待できる。

キーワード 宅内コンテキスト, 画像認識, コグニティブAPI, 機械学習, 多数決

Recognizing Fine-Grained Contexts at Home Using Multiple Cognitive APIs

～ Majority Voting Approach ～

Sinan CHEN[†], Sachio SAIKI[†], and Masahide NAKAMURA^{†,††}

[†] Kobe University Rokkodai-cho 1-1, Nada-ku, Kobe, Hyogo 657-8501 Japan

^{††} Riken AIP 1-4-1 Nihon-bashi, Chuo-ku, Tokyo 103-0027 Japan

E-mail: [†]chensinan@ws.cs.kobe-u.ac.jp, ^{††}sachio@carp.kobe-u.ac.jp, ^{†††}masa-n@cs.kobe-u.ac.jp

Abstract Recognizing fine-grained contexts in individual houses and residents has been an important but challenging research topic. It has been studied for many years in the field of ubiquitous computing. We are studying fine-grained context recognition affordable for general households, by integrating general-purpose image-based cognitive APIs and light-weight machine learning. In our previous study, we first evaluated the recognition performance of commercial APIs, and then developed a recognition method that uses the recognition results of the single API as feature values. However, we found that the method could not distinguish different contexts with multiple people such as “Eating” and “Playing games” and “General meeting”. The goal of this research is to improve the recognition accuracy of such difficult contexts, with preserving the affordability to general households. In the proposed method, we use multiple image-based cognitive APIs. For each API, we construct an independent recognition model using feature values of the API. Then, the context is determined by majority voting among results of the independent models. Compared to the popular approach with deep learning, the proposed method does not require a huge amount of labeled data or vast computing resources. As a result, it can recognize home contexts more accurately and economically.

Key words Home context, Image recognition, Cognitive API, Machine learning, Majority voting

1. はじめに

住居や建物内において、居住者の日常生活行動や宅内の環境状態（総じて宅内コンテキストと呼ぶ）を認識する研究は、これまで特にユビキタスコンピューティングの分野で数多く行われてきた。伝統的には、ウェアラブルセンサや宅内測位システム、環境センサ等を利用して宅内の様々なコンテキストを認識し、状況に応じた情報提供や機器操作を行う付加価値サービスの実現に役立てられる。また近年では、深層学習（Deep Learning）を用いたマルチメディアデータの認識技術の進歩が著しい。一般的に音声や画像等のマルチメディアデータは、従来のセンサデータより情報量が多いため、より細粒度の宅内コンテキストを高精度に認識できる可能性がある。

しかしながら、各種センサを用いた手法では身体への装着や建物への取り付け、それらの保守等が必要となり、一般家庭に導入して日常的に運用するに未だ敷居が高い。また、単純に深層学習を適用するアプローチでは、世帯ごとのレイアウトや環境状態の違いもあり、世帯固有の認識モデルの構築に膨大な学習データと計算資源が必要になる。これらの理由から、既存手法の多くは一般家庭に広く普及するには至っていない。

本研究の最終目的は、世帯ごとの個性や要求に対応でき、かつ、一般家庭で実施可能な高精度の宅内コンテキスト認識技術を実現することである。我々はこれまで、一般家庭に導入が容易なカメラデバイスに着目し、コグニティブ API による画像認識と軽量の機械学習を組み合わせることで、宅内コンテキストを高精度かつ少ない資源で認識する手法の研究を進めている。

先行研究 [1] [2] では、研究室の日常活動画像をコグニティブ API の画像認識にかけ、タグ集合を抽出、文書特徴量と文書間類似度によって、画像がコンテキストに応じたクラスに分割できるかを評価した。大手 3 社のコグニティブ API を試したが、これらの API が提供する汎用的な画像認識のみでは、研究室に特化したコンテキストをうまく認識できないことが分かった。[3] では、あるコグニティブ API から得られたタグ集合から文書ベクトルを生成し、これを特徴量として教師あり学習にかけ、研究室の 7 種類のコンテキストを識別する分類器を構築した。その結果、平均精度 0.92 以上でコンテキストを認識するモデルが構築できた。しかしながら、食事やゲーム、ミーティングといった多人数が参加するコンテキストは、区別が難しいということが分かった。

そこで本稿では、これらの区別が難しかったコンテキストの認識精度をより一層向上させることを目的とする。そのためのアプローチとして、我々は複数のコグニティブ API を用いてこれらをアンサンブル学習 [4] することを考える。具体的には、宅内の同一画像に対し、異なる API によって特徴量を抽出して、複数の独立した認識モデルを構築する。これらの認識結果を多数決することで、認識精度を向上させる（図 1）。異なる API によって同一の画像を多角的に観察し特徴を抽出できるようになるため、宅内の多様なコンテキストをより高精度で認識できる可能性がある。

提案手法の有効性を評価するため、我々は研究室の 7 つのコ

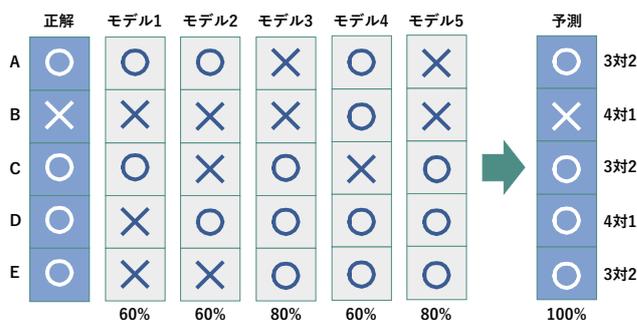


図 1 アンサンブル学習による多数決の例

ンテキストを認識する実験を行った。認識するコンテキストは、食事 (Dining together), 全体ミーティング (General meeting), 無人 (Nobody), 個別ミーティング (One-to-one meeting), 研究 (Personal study), 遊び (Play games), 掃除 (Room cleaning) である。各コンテキストについて代表的な画像を 100 枚、異なる日に撮影したものから選別し、ラベル付けた。次に、選別した合計 700 枚の画像データを、Microsoft Azure Computer Vision API [5], IBM Watson Visual Recognition API [6], Clarifai API [7], Imagga REST API [8], Paralleldots API [9] という 5 種類の API に送信し、それぞれ API の認識結果からタグ集合を抽出した。そして、抽出したタグ集合を文書と見なし、TF-IDF (Term Frequency Inverse Document Frequency) [10] 法を用いてベクトル化した。続いて、ベクトル化したタグ集合と対応するコンテキストのラベルを Microsoft Azure Machine Learning Studio [11] に導入し、Multiclass Neural Network によって 5 つの分類器を構築した。最後に、それらの分類器の結果を多数決し、認識結果を出力した。

評価実験の結果に関し、単一 API による特徴量を用いたモデルの認識精度では、Azure API の全体精度 (Overall accuracy) は 0.85, Watson API は 0.80, Clarifai API は 0.91, Imagga API は 0.94, Paralleldots API は 0.77 という結果になった。それに対し、提案する複数 API による多数決での全体精度は 0.98 であった。さらに、混同行列による評価では、全体ミーティング、無人、個別ミーティング、遊びの認識精度は 1.00, 食事、研究、掃除は 0.96 となった。以上の結果により、7 種類のコンテキストに対し、アンサンブル学習による多数決によって、コンテキストの認識精度が顕著に向上したことが分かった。

2. 準備

2.1 宅内のコンテキスト認識

宅内のコンテキストとは、スマートホームにおける居住者の日常生活行動や宅内の環境状態に関するあらゆる状況情報を意味とする。宅内の時々刻々の状況変化をリアルタイムに把握することは、提供すべきサービスの内容やタイミングに対して極めて重要な意味を持つ。そのため、宅内のコンテキストに対する高精度の認識や、居住者の要求に応じたタイムリーなサービスの提供に関しては従来の重要な研究課題とされ、特にユビキタス・コンピューティングの分野に長年研究されてきた。

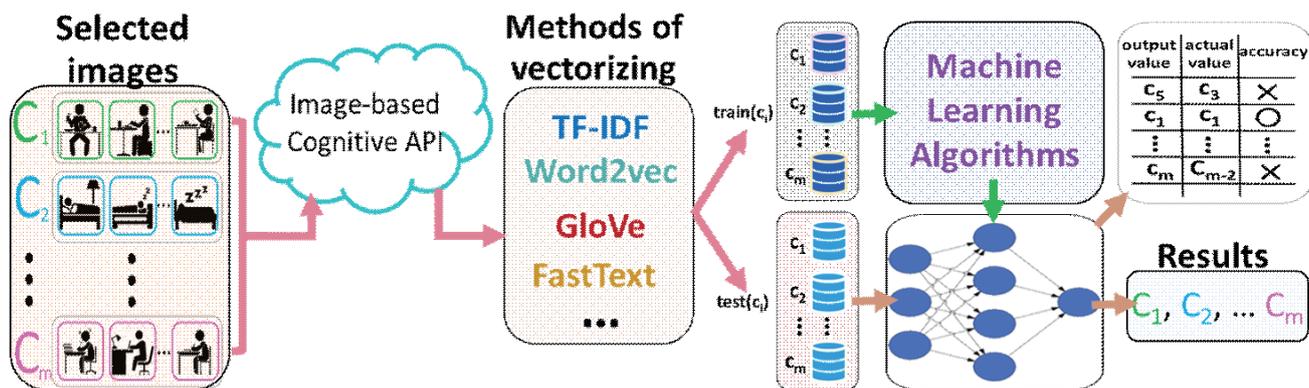


図2 先行研究における提案の流れ

宅内のコンテキスト認識に関するこれまでの研究では、ウェアラブルセンサや環境センサ、スマートフォン等から得られる数値データを収集・分析し、それらのデータから特徴量を抽出・活用したコンテキストを認識する研究が主流であった。例えば、スマートハウス内センシングを活用した生活行動推薦システム [12] や、スマートホームにおける複数のセンシングデータに基づいた生活行動データ抽出システム [13]、宅内の環境変化と声掛けに基づく在宅高齢者の日常生活行動センシングシステム [14] 等が提案されている。

しかしながら、専用のセンサによる居住者の生活への侵襲性や、世帯ごとの個性による運用の複雑さから、一般家庭に広く普及するには至っていない。昨今に実用化が著しいスマートホームでは、宅内のコンテキスト認識の多様な応用が期待される。従来の代表的な応用例として、独居高齢者の見守りや、居住者の生活リズムの維持向上等が有している。

2.2 画像に基づくコグニティブ API

コグニティブサービスは、画像や音声、テキストといったマルチメディアデータを認識するクラウドサービスであり、一般的にクラウドの豊富な計算資源を生かして構築された学習済み機械学習モデルによって実装されている。大規模で複雑な認識処理をアプリケーションに容易に組み込めるため、コグニティブ API といったコグニティブサービスを外部アプリケーションから呼び出して利用するための API が近年に登場している。

画像に基づくコグニティブ API は、コンピュータビジョン分野技術を基に、与えられた画像を認識し、画像内に含まれた様々な情報を自動的に抽出して JSON (JavaScript Object Notation) 形式で認識結果を返す。代表的なコグニティブ API としては、Microsoft Azure Computer Vision, IBM Watson, Google Cloud Vision, Clarifai 等が存在している。

コグニティブ API の画像認識結果に関し、サービスの提供するクラウド会社によって、API の機能種類や各 API の画像認識に用いる内部アルゴリズム、認識結果に含まれる要素項目が一部に異なっている。一般に画像から人間の顔や、物体、場所、テキストなどを主な検出・認識の対象として、年齢や、顔感情、カテゴリ、タグ、オブジェクト座標などの特徴値が返される。本研究では、異なる API の認識結果中の共通的な要素

項目であるタグの集合を画像の特徴量と見なす。

2.3 先行研究 [1] [2] [3]

先行研究では、我々は宅内における時系列の画像特徴を抽出・認識する高い負荷処理をクラウド側に委譲したいと考え、クラウドサービスの 1 種である画像に基づくコグニティブ API を活用している。

最初に、我々は汎用的なコグニティブ API の宅内での適用可能性を探求することを目的とした。まず、研究室の日常生活画像を収集・分類し、Microsoft Azure Computer Vision, IBM Watson Visual Recognition, Google Cloud Vision の 3 種類の API よりそれぞれ出力した認識結果を取得し、特徴量とするタグ集合を抽出した。次に、タグ集合がオリジナルのコンテキストを反映しているかをチェックするため、TF-IDF 法を用いてそれらのタグ集合をベクトル化し、同じコンテキストの凝集、異なるコンテキストの分離が可能かを、文書間類似度を表すためのコサイン類似度によって評価した。その結果、API が出力するタグ情報を、そのまま研究室のコンテキスト認識に適用しても、十分な性能が得られないことが分かった [1] [2]。

その次に、我々は高精度のコンテキスト認識を達成するためにコグニティブ API による特徴量をトレーニングしたいと考え、その前に利用した 3 種類の API 中から最も出力タグが多い Microsoft Azure Computer Vision を用い、研究室の新しく収集・分類した活動画像を認識された。API の認識結果よりタグ集合を抽出し、TF-IDF 法によるベクトル化をした。さらに、それらのベクトルデータを学習データとして、クラウド機械学習プラットフォームである Microsoft Azure Machine Learning Studio に導入し、軽量の教師あり機械学習アルゴリズムである多値分類の Multiclass Neural Network に向け、認識精度 0.92 以上のモデルが構築できた (図 2) [3]。

2.4 アンサンブル学習

アンサンブル学習とは、複数のモデルを融合させて 1 つの学習モデルを生成することで、モデルの精度が大幅に向上することである (図 1) [4]。基本的なアンサンブル学習には、バギング、ブースティング、スタッキングの 3 種類がある。本稿では、アンサンブル学習のスタッキングと似る手法を利用し、モデルを積み上げて最終的に多数決による認識結果を生成する。

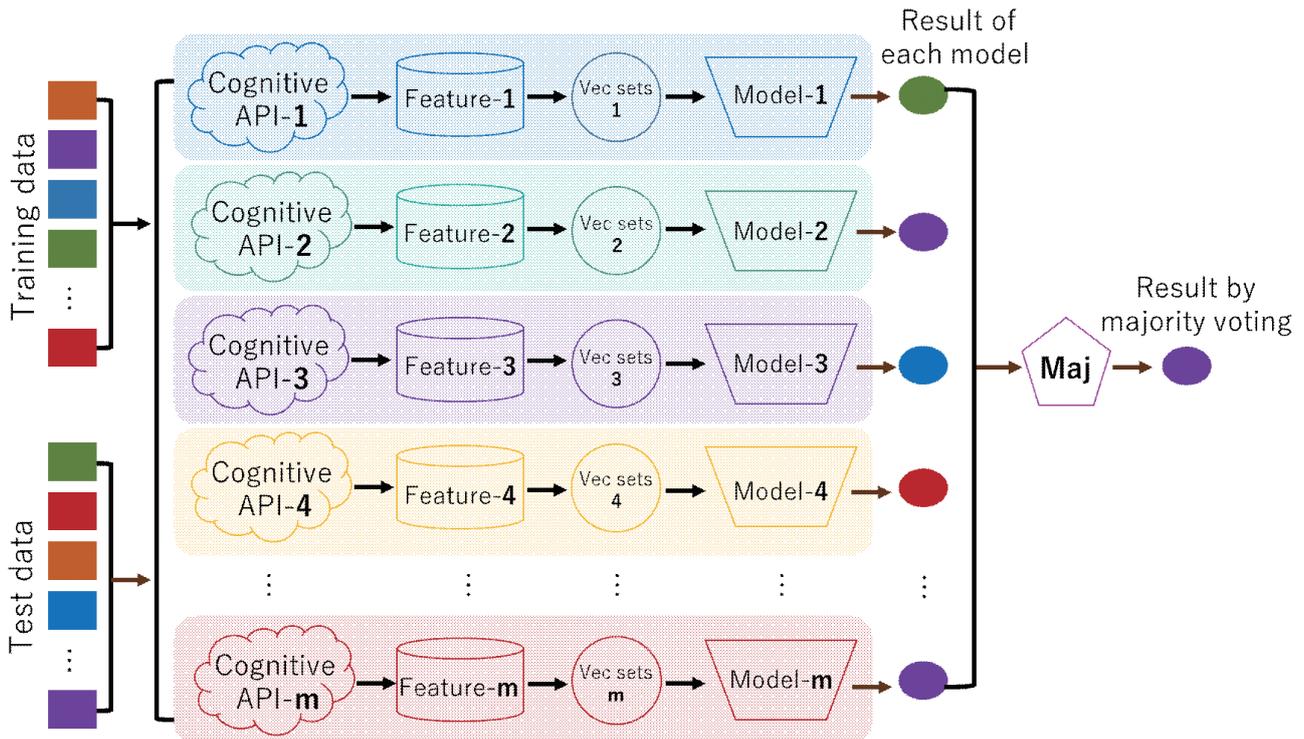


図3 提案手法の流れ

3. 提案手法

3.1 概要

同一の画像に対し、サービスの提供するクラウド会社によって、画像に基づくコグニティブ API による特徴量（タグ）の数量も内容も異なる。また、世帯ごとの個性や要求によって、API が認識する画像内容も異なる。こういった様々な違いが存在し、モデルを構築する際にいかに高精度な認識結果を維持できるかは、我々が検討している重要な課題である。

本稿では、複数のコグニティブ API を利用するアプローチを考える。各 API による特徴量をそれぞれ活用し、認識モデルを構築する。さらに、各モデルの認識結果を併用し、アンサンブル学習を用いて最終的に多数決によるモデルを生成する。これによって、モデルの認識精度を向上することが期待できる。

3.2 提案手法の流れ

提案する手法は次の 5 ステップからなる（図 3）：

STEP1：データセットの作成

空間内で認識したい l 個のコンテキスト $C = \{c_1, c_2, \dots, c_l\}$ を定義し、各コンテキスト $c_i \in C$ を特徴的に反映している n 個の画像データ $IMG(c_i)_j = \{img(c_i)_1, img(c_i)_2, \dots, img(c_i)_n\} (0 \leq i \leq l) (0 \leq j \leq n)$ を取得し、 $l \times n$ のデータセットを作成する。作成したデータセットを訓練用とテスト用に分割する。具体的には、各 c_i の n 個の画像データを α 個と $n - \alpha$ 個の 2 つの集合 $train(c_i)_j$ と $test(c_i)_j$ に分割する。

STEP2：各コグニティブ API による特徴量の抽出

$train(c_i)_j$ と $test(c_i)_j$ それぞれの中の各コンテキスト c_i の特徴的に反映している画像データ $IMG(c_i)_j$ を m 種のコグニティブ API ($API = \{api_1, api_2, \dots, api_m\}$) にそれぞれ投入し、

$api_q \in API (0 \leq q \leq m)$ が $IMG(c_i)_j$ 内に認識するタグ集合 $Tag((IMG(c_i)(j_q)) = \{t_1, t_2, t_3, \dots\}$ を得る。本ステップでは、訓練データとして $l \times \alpha \times m$ 個のタグ集合が得られ、テストデータとして $l \times (n - \alpha) \times m$ 個のタグ集合が得られる。

STEP3：特徴量のベクトル化

$train(c_i)_j$ と $test(c_i)_j$ のそれぞれに含まれた全てのタグ集合 $Tag((IMG(c_i)(j_q))$ をベクトル化し、各タグ集合よりベクトル集合 $Vec(Tag((IMG(c_i)(j_q)) = \{v_1, v_2, v_3, \dots\}$ を得る。

STEP4：各特徴量を用いたモデルの構築

STEP2 で分割した各 api_q による特徴量（ベクトル集合）の α 個を訓練データとして利用し、通常の教師あり機械学習のアルゴリズムを実行し、認識モデル $Model = \{model_1, model_2, \dots, model_m\} (0 \leq q \leq m)$ を構築する。ここで、 $model_q \in Model$ は入力ベクトルを c_1, c_2, \dots, c_l に分類する多値分類器である。

STEP5：各モデル認識結果を用いたアンサンブル学習

STEP2 で分割した各 api_q による特徴量（ベクトル集合）の残りの $n - \alpha$ をテストデータとして $model_q$ に入力し、各画像 $IMG(c_i)_j$ に対応する各 api_q の出力結果 $C_q = \{c_1, c_2, c_3, \dots, c_m\}$ をアンサンブル学習によって、最終の多数決による認識結果 $Maj\{c_1, c_2, c_3, \dots, c_m\}$ を評価する。

4. 評価実験

3. で提案した手法に従い、実際の環境下での撮影した画像データを基に評価実験を行う。

4.1 実験環境

先行研究を行って以来、我々は研究室内の固定スペースで定点カメラを設置し、昨年 7 月から現時点までの画像データを 5 秒間隔で撮影・蓄積している。我々はまず認識したいコン

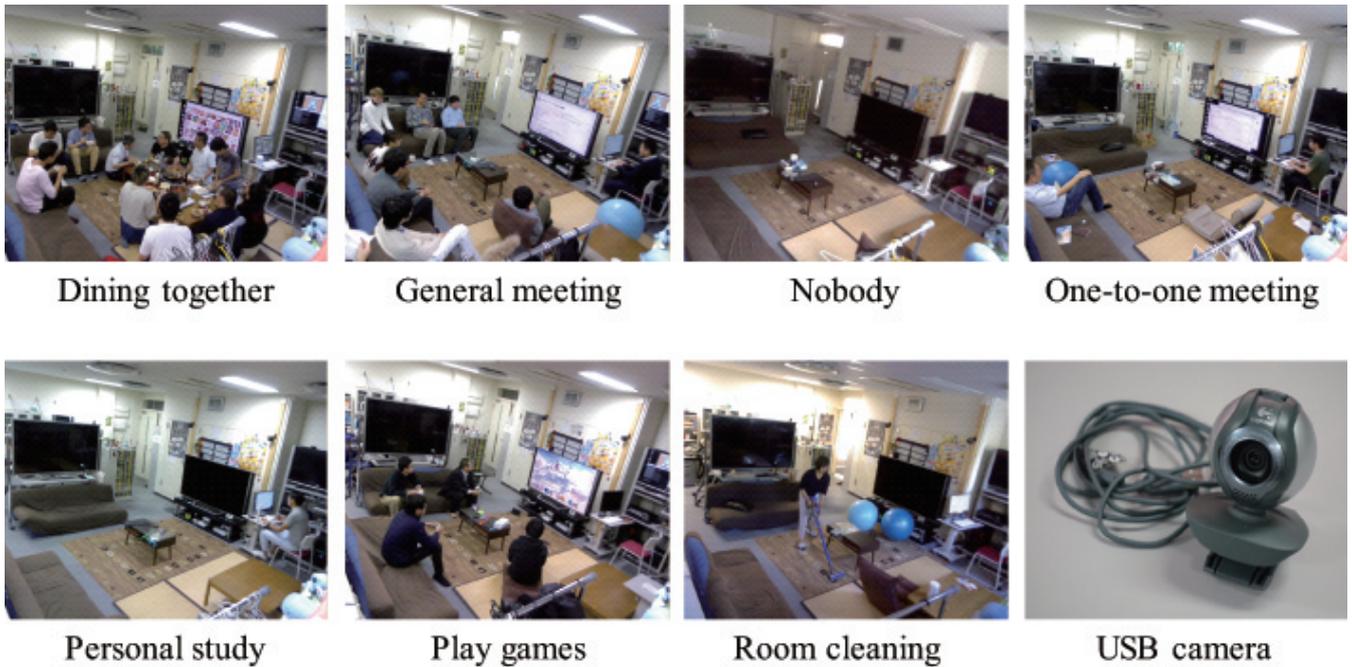


図 4 各コンテキストにおける代表的な画像と USB カメラ

テキストとして、食事 (Dining together), 全体ミーティング (General meeting), 無人 (Nobody), 個別ミーティング (One-to-one meeting), 研究 (Personal study), 遊び (Play games), 掃除 (Room cleaning) の 7 つを定義し、蓄積した全ての画像データの中から、各コンテキストに代表する異なる日の一定間隔で撮影した画像を 100 枚に選別してラベル付けした。図 4 に各コンテキストにおける代表的な画像と USB カメラを示す。次に、選別した合計 700 枚の画像データをランダム化し、半分ずつ訓練データとテストデータに分割した。そして、訓練データとテストデータをそれぞれに *Microsoft Azure Computer Vision API* [5], *IBM Watson Visual Recognition API* [6], *Clarifai API* [7], *Imagga REST API* [8], *Paralleldots API* [9] という 5 種類の API に送信し、それぞれ API の認識結果からタグ集合を抽出した。図 5 に同一の画像の各 API による特徴量 (タグ集合) の例を示す。

4.2 各 API による特徴量のベクトル化

訓練データとテストデータのそれぞれから、各 API により抽出したタグ集合を文書と見なし、*TF-IDF (Term Frequency Inverse Document Frequency)* [10] 法を用いてベクトル化した。

4.3 単一 API による特徴量を用いたモデルの構築

訓練データから各 API による特徴量のベクトル集合と対応するコンテキストのラベルを *Microsoft Azure Machine Learning Studio* [11] に導入し、Multiclass Neural Network によってそれぞれのモデルを構築した。テストデータから各 API による特徴量のベクトル集合をそれぞれモデルの入力として、モデルの出力ラベルとあらかじめ付与されたラベルを比較して、それぞれモデルの精度を評価した。

4.4 複数 API による特徴量を用いたモデル出力の多数決構築したそれぞれモデルの出力ラベル結果を併用し、アン

Using API name	Tags of the image about room cleaning
Microsoft Azure Computer Vision API	indoor,living,room,table,television,furniture,sitting,messy,cluttered,area,computer,fireplace,fire,bedroom,large,flat,view,screen,desk,video,woman,young,playing,bed,game,man,standing,dog,people
IBM Watson Visual Recognition API	living,room,indoors,classroom,basement,support,supporting structure
Clarifai API	room,furniture,indoors,table,desk,seat,chair,trading,floor,interior,design,home,hospital,medicine,technology>window,mirror,business,computer,people,production
Imagga REST API	room,interior,furniture,table,home,house,modern,floor,decor,chair,sofa,design,wood>window,luxury,living,lamp,apartment,indoors,light,home,theater,building,office,architecture,comfortable,wall,residential,couch,inside,theater,carpet,desk,fireplace,kitchen,living,room,pillow,3d,structure,relax,seat,lighting,decoration,estate,glass,furnishings,style,bedroom,indoor,empty,domestic,real,decorate,relaxation,cozy,chairs,residence,family,rest,area,space,contemporary,comfort,equipment,monitor,leather,render,television,classroom,hardwood,nobody,vase,hotel,bed,business,elegance,clean,upscale,lifestyle,computer,studio,apartment,rug,new,plant,elegant,furnishing,stylish,guest,spacious,cabinet,ceiling,armchair,device,marble,restaurant,work,mirror,dining,ottoman,shelf,fixtures,wooden,suburbs,suite,suburban,dwelling,lounge,tile,display,fashion,place,book
Paralleldots API	Room,Interior,design,Property,Vehicle,Building,Home,Sport,venue,Screenshot,Furniture

図 5 同一の画像の各 API による特徴量 (タグ集合) の例

サンプル学習で最終的に多数決による認識結果を生成した。Python で実行する各モデル認識結果の多数決に関するコードは `find_majority = data.mode().iloc[0]` である。

4.5 実験結果

本実験の構築した各モデルの認識精度と多数決による認識

Model name	Overall accuracy	Dining together	General meeting	Nobody	One-to-one meeting	Personal study	Play games	Room cleaning
Azure API - model	0.85	0.96	0.89	1.00	0.66	0.92	0.84	0.77
Watson API - model	0.80	0.89	0.67	0.82	0.80	0.94	0.80	0.71
Clarifai API - model	0.91	0.91	0.98	0.91	0.84	0.92	0.92	0.92
Imagga API - model	0.94	0.96	0.93	1.00	0.89	0.96	0.92	0.96
Paralldots API - model	0.77	0.80	0.89	0.93	0.45	0.88	0.67	0.80
Majority voting	0.98	0.96	1.00	1.00	1.00	0.96	1.00	0.96

図6 実験結果

精度は図6に示す。構築した各モデルの全体精度（Overall accuracy）に関し、Azure APIは0.85、Watson APIは0.80、Clarifai APIは0.91、Imagga APIは0.94、Paralldots APIは0.77という結果になった。各コンテキストのラベルに対し、各モデルの最も低い認識精度に関し、Paralldots APIモデルの個別ミーティングに対する認識精度は0.45であり、最も高い認識精度に関し、Azure APIモデルとImagga APIモデルの無人に対する認識精度は1.00であった。

それに対し、各モデルの認識結果を用いた多数決による認識精度に関し、全体精度は0.98であった。各コンテキストのラベルに対し、全体ミーティング、無人、個別ミーティング、遊びの認識精度は1.00、食事、研究、掃除は0.96となった。

4.6 考察

本実験で各APIによる特徴量を用いて構築したモデルの評価結果において、一対一の個別ミーティングという画像に対する認識精度がより低く、APIごとの出力タグの内容範囲や数量が限られる要因に影響があると推察される。多数決による認識精度の結果において、多人数が参加する食事や一人での研究、掃除というコンテキストのラベル認識精度がより低く、誤認識の個別データが集中する可能性があると考えている。また、宅内活動の差異により、人間が刻々に無関係な情報（据え置きインテリアや電気製品）を若干に遮ることで、APIよりそれらが認識できずタグを省略する場合がある。総じて言えば、APIの出力特性や訓練データの選択範囲を把握することがモデル構築の精度向上に重要なのではないであろう。

5. おわりに

本稿では、世帯ごとの個性や要求に対応できる一般家庭で実施可能な高精度の宅内コンテキスト認識を実現することを最終目的とし、我々は複数のAPIを用いた宅内のコンテキスト認識モデルの認識結果の多数決手法を提案した。提案手法では、宅内の同一画像に対し、異なるAPIによる特徴量で構築したそれぞれのモデルの認識結果を併用し、アンサンブル学習を用いて最終的に多数決する。これにより、宅内の多様なコンテキストへの認識精度が向上できる。また、評価実験では、5種類のAPIで抽出したタグ集合をTF-IDF法でベクトル化し、それぞれに研究室の7種類のコンテキストを認識する実験を行った。その結果、Azure APIの方の全体精度（Overall accuracy）は0.85、Watson APIは0.80、Clarifai APIは0.91、Imagga APIは0.94、Paralldots APIは0.77という結果になった。各モデルの認識結果を併用し、アンサンブル学習で最終的に多数決による認識結果に関し、全体精度は0.98であった。

今後の課題として、各APIによる特徴量を結合し、協力して構築したモデルの認識結果を評価したいと考える。また、各APIの出力特性や訓練データの選択範囲に関する検討もしたい。最後に、構築したモデルの研究室での配備・運用を行って、その性能と有効性を評価したい。

謝辞 本研究の一部はJSPS科研費JP19H01138, JP17H00731, JP18H03242, JP18H03342, JP19K02973の助成を受けている。

文 献

- [1] 陳 思楠, 佐伯幸郎, 中村匡秀, “画像に基づくコグニティブ api の宅内センシングへの適用可能性,” 電子情報通信学会技術研究報告 = IEICE technical report : 信学技報, vol.118, no.195, SC2018-19, pp.31-36, Aug. 2018.
- [2] S. Chen, S. Saiki, and M. Nakamura, “Evaluating feasibility of image-based cognitive apis for home context sensing,” ICSPIS2018, Nov. 2018.
- [3] 陳 思楠, 佐伯幸郎, 中村匡秀, “コグニティブ api による特徴量を用いた宅内コンテキスト認識手法の提案,” 電子情報通信学会技術研究報告 = IEICE technical report : 信学技報, vol.118, no.511, SC2018-38, pp.7-12, Aug. 2019.
- [4] codExa, “アンサンブル学習の仕組みと3つの種類について解説します,” <https://www.codexa.net/what-is-ensemble-learning/>. visited on 2019-02-01.
- [5] M. Azure, “Visual recognition,” <https://cloud.ibm.com/apidocs/visual-recognition>. visited on 2019-04-15.
- [6] I. Watson, “Computer vision,” <https://azure.microsoft.com/ja-jp/services/cognitive-services/computer-vision/>. visited on 2019-02-01.
- [7] Clarifai, “Transforming enterprises with computer vision ai,” <https://clarifai.com/>. visited on 2019-04-15.
- [8] Imagga, “Imagga api,” <https://docs.imagga.com/>. visited on 2019-04-15.
- [9] ParallelDots, “Image recognition,” <https://www.paralleldots.com/object-recognizer>. visited on 2019-04-15.
- [10] “Tf-idf で文書をベクトル化,” <http://ailaby.com/tfidf/>, Aug. 2016. visited on 2019-02-01.
- [11] M. Azure, “Azure machine learning studio,” <https://azure.microsoft.com/ja-jp/services/machine-learning-studio/>. visited on 2019-02-01.
- [12] 中村笙子, 廣森聡仁, 山口弘純, 東野輝夫, 山口容平, 下田吉之, “スマートハウス内センシングを活用した生活行動推薦システム,” マルチメディア、分散協調とモバイルシンポジウム 2014 論文集, vol.2014, pp.1557-1566, July 2014.
- [13] 上田健揮, 玉井森彦, 安本慶一, “スマートホームにおける複数のセンシングデータに基づいた生活行動データ抽出システムの提案,” マルチメディア、分散協調とモバイルシンポジウム 2014 論文集, vol.2014, pp.1884-1891, July 2014.
- [14] 玉水一柔, 榊原誠司, 佐伯幸郎, 中村匡秀, 安田 清, “宅内の環境変化と声掛けに基づく在宅高齢者の日常生活行動センシングシステムの検討 (情報通信マネジメント),” 電子情報通信学会技術研究報告 = IEICE technical report : 信学技報, vol.116, no.404, pp.7-12, Jan. 2017.