

ARを用いて身近な物の英単語を表示する英語学習支援システムの開発

産業能率大学 情報マネジメント学部 川野邊研究室 4年 佐藤 幸輝 指導教員:川野邊 誠

研究概要

研究背景

背景①

既存の中高の学習スタイルは言語学的な正しさや受験でよく使われる表現項目を重視

身近でない物や文化を題材にしているため日常的な言語として扱えていない人が多い

普段の生活における身近な単語を優先的に学ぶことで日常的な言語感覚に沿った学習が可能

背景②

第二言語を使用する際、第一言語から第二言語への変換を段階的に行っている

幼児の言語習得では単語の理解に視覚や触覚など様々な情報を複合的に扱い判断

長谷川らの研究^[1]や画像優位性効果、二重符号化理論からもビジュアルを用いた学習効果が示されている

普段使っている、目に入る物の英単語を優先的に学ぶことで日常的な言語感覚と人の自然な言語習得プロセスに沿った英語学習が可能なのではないか

研究目的

日常的に使用する物、目に入る物の英単語をARを用いて優先的に学習する自然な言語習得プロセスに沿った英語学習支援システムの開発

期待効果

日常的に使う英単語の習得から英語への親近感を喚起することで英語の習得を促進

画像と物体を掛け合わせた名詞ベースの言語習得により日本語から英語への変換に要する思考時間の短縮

研究手法

撮影した画像に映り込む物体を識別可能な画像処理ライブラリを使用

カメラの映像をリアルタイムに認識し画像内に映る物の上に対応する英単語を表示

直感的に学習を促すARシステムを実装

日常的な使用を目指しシステムのスマートフォン対応

自然な形で言語習得が可能か評価実験を実施

研究成果

使用環境

- ・Unity6.2 (6000.2.7f2)
C#を基本とするゲーム開発プラットフォーム
- ・AR Foundation
Unity上でマルチプラットフォームに対応したARアプリの作成が可能なフレームワーク
- ・Unity AI (Inference Engine)
生成AIを利用した開発支援ツール。推論エンジンをYOLO上で動作するために使用
- ・YOLO(v8,11)
物体検出アルゴリズム。畳み込みニューラルネットワークを活用し、高速検出が可能
Ultralytics社が開発したv8,11のバージョンを使用

システム構造

①インプットモード

自室などの様子を撮影し
検出した物体の画像データを保存

②学習モード

ARを用いて、各物体上にマーカーを表示
マーカーをタッチすると英単語が表示

③テストモード

「出題：画像、回答：単語」
「出題：単語、回答：画像」の2種類用意

システム開発

ベースプログラム: unity/inference-engine-yolo^[2]

Unity上でYOLOを用いて、動画の物体検出を行うライブラリをベースとして使用

開発内容

- ・検出対象メディアの変更
“動画”から“Webカメラを用いたリアルタイム映像”に検出対象メディアを変更
- ・学習済みモデルの用意
データセット学習済みのYOLOv8,11を本環境で利用可能なonnx形式に変換
- ・使用するモデル用のリストを作成
物体検出時に参照する物体のリストを使用するモデルごとに用意
- ・開発を容易にするための各種UI開発
各種パラメータの設定を容易に変更可能なUIを実装

使用モデルの選定

表1. 各モデルの比較

使用モデル	リリース	検出数	mAP* (精度)	特徴
YOLOv8	2023	80	50.2	広く利用され、参考情報が多く成熟したエコシステムを持つ
YOLOv8 (Open Images v7)	2023	600	33.6	データセットをCOCOからOpen Images v7に変更したモデル
YOLO11	2024	80	51.5	v8以上の精度だがサードパーティやコミュニティが未成熟

*mのモデルサイズでの値

動作検証

- ・各モデルと5つのモデルサイズを使用
v8,v8(Open Images v7),11×n,s,n,l,x
- ・信頼度と重なり度のしきい値を調整

検証結果

- ・v8,11は高い検出精度だがモデルサイズやバージョンによる差異は表出せず
- ・v8(Open Images v7)はCOCOのモデルよりも検出がシビアな傾向で、同じしきい値では検出物体数が減少
信頼度のしきい値を大幅に下げると大量の物体を検出したため、細かな調整が必要



図1. v8での物体検出

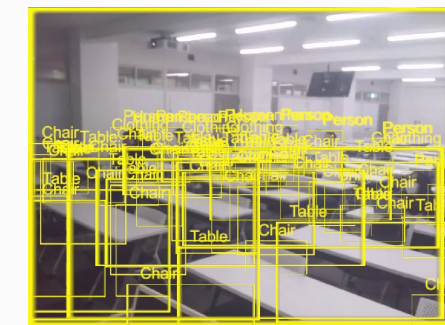


図2. 信頼度0.05以下

インプットモード、暗記モードの実装

手順1:撮影と保存機能を実装（インプットモード）

- ・画面内UIでボタンを実装。ボタンを押すとその場面の撮影と保存を実行
- ・同様にUIでテキストを表示し、保存した物体数を元にステータスを表示

手順2:AR環境への移行（暗記モード）

- ・動作検証用ではWebカメラの映像をそのまま表示する仕様であるためARとして扱えるよう、ARカメラの映像をYOLOに入力する形に変更

手順3:ARを用いた学習機能の実装（暗記モード）

- ・物体を囲う枠の代わりにタッチ可能なボタンを物体上に表示
- ・ボタンを押すと2秒間だけその物体の英単語を表示

今後の予定

- ・テスト機へのインストール
- ・問題管理、出題システム設定
- ・動作テスト
- ・卒業論文の執筆

参考文献

- [1] 長谷川和則, 金子敬一, 都田青子. 異なる方式に基づく英単語学習用システムの開発と評価. 情報処理学会研究報告 = IPSJ SIG technical reports 2006, 74, pp.25-32, 2006.
- [2] Hugging Face. unity/inference-engine-yolo
<https://huggingface.co/unity/inference-engine-yolo>, 2025.