データの説明

データ番号 （コモンズセンターで記入します）	DPPSCdbp_2025-07
タイトル	人工衛星画像による常磐自動車道におけるトラック交通量把握のための教師データおよびそれによるトラック交通量の推定″
作成者	内藤久裕 Meiyi Zhuang
問い合わせ先	naito@dpipe.tsukuba.ac.jp
概要（データの内容や作成方法）	1. **2_sentinel2_dataset**：茨城県における道路 AOI でマスクされた Sentinel-2 画像が含まれています。 2. **3_annotated_images**：YOLO キーポイント検出モデルの学習に使用したデータセットとアノテーションファイルが含まれています。 3. **4_prediction_images**：ステップ 3 で学習した YOLO モデルによる予測画像および CSV ファイルが含まれています。 4. **5_traffic_data**：構築されたデータセット（テーブル）および対応する可視化図が含まれています。
	1_ground_truth_video フォルダの詳細情報 • 本フォルダには、茨城県つくば市において取得したグラウンドトゥルース動画が含まれています。 • グラウンドトゥルース動画は 2 地点で収集しました。 • 1 つは Kise 地域（4 つの橋候補あり）で、GPS 座標は（緯度 = 36.0843909、経度 = 140.1554314）です。 • もう 1 つは Heights 橋で、GPS 座標は（緯度 = 36.1059684、経度 = 140.1764893）です。 • Kise 地域には 4 つの橋候補があり、それぞれ 1、2、3、4 とラベル付けされています。動画ファイル名は、橋 ID（1～4）の後に「kise」が付けられています。

- 動画ファイル名において、「south」はカメラが東京方向を向いていることを示し、「north」は水戸方向を向いていることを示します。

2_sentinel2_dataset フォルダの詳細情報

- 本フォルダには、日本の高速道路区間を対象とした Sentinel-2 画像が、AOI（Area of Interest）単位で整理されて格納されています。

- まず、日本の高速道路ネットワークを抽出し、1 km × 1 km の AOI に分割します（各 AOI には 50 メートルのバッファを付加）。各 AOI には固有の ID が割り当てられます。

- 次に、openEO API（openEO は、エンドポイント https://openeo.dataspace.copernicus.eu/openeo/1.2/を通じてアクセスした）を用いて、茨城県の道路バッファ付き Sentinel-2 画像をダウンロードします。

- 以下の Python コードを用いてデータをダウンロードした。 datacube = (connection .load_collection("SENTINEL2_L2A", temporal_extent=TIME_RANGE, bands=BANDS, max_cloud_cover=CLOUD_MAX) .filter_spatial(aoi_geojson)) datacube.download(NC_PATH)

- 茨城県内には合計 388 の AOI があり、2016 年から 2024 年までの期間をカバーしています。

3_annotated_images フォルダの詳細情報

- 本フォルダには、同一内容で異なる形式のアノテーションデータセットが 2 種類含まれています。

1. **keypoint_sentinel2.v2i.coco**：COCO 形式（macOS および Windows の両方に対応）

- データ作成手順

1. ダウンロードした画像の中から、グラウンドトゥルース動画が取得されている高架橋地域に対応する AOI を選択しました。

2. アノテーションには Roboflow を使用し、2024 年の通年データについてアノテーションを実施しました。
3. アノテーション済みデータセットは、train、test、valid の各フォルダに整理されています。各フォルダには、画像ファイルと、それに対応するアノテーションファイル（COCO JSON 形式）が含まれています。

- 再現方法

1. Roboflow にログインし、新しいプロジェクトを作成します。モデルタイプとして Keypoint Detection を選択します。
2. クラスとして「rainbow」を 1 つ定義し、キーポイントとして「red」「green」「blue」の 3 点を設定します。
3. COCO 形式のデータセットフォルダ（解凍済み）をアップロードします。画像ファイルとアノテーションファイルが同時にアップロードされていることを確認し、「Save and Continue」をクリックします。
4. 画面右上の「Train Model」をクリックし、「Custom Training」を選択した後、モデルとして YOLOv11（Fast）を選択し、学習を開始します。
5. モデルの学習には通常約 15 分を要します。

4_prediction_images フォルダの詳細情報

- 茨城県南部および北部の AOI を選択し、2016 年から 2024 年までの期間について予測を実施しました。

- 予測には、ステップ 3 で学習したパラメータと重みを用いた YOLO キーポイント検出モデルを使用しました。

このフォルダには以下の 3 つの CSV ファイルが含まれています：

1. **per_day_box_vs_kpts_filtered**：各 AOI 内のすべての画像に対する時系列情報を含みます。
2. **pred_keypoint_long_all**：すべての予測結果（各画像におけるキーポイント位置および信頼度スコア）を含みます。
3. **predictions_all_batches**：各 AOI についてのバッチ単位の情報（検出された「rainbow」の数など）を記録しています。

また、このフォルダには複数のサブフォルダがあります。各サブフォルダは AOI ID で命名されています。

- 各 AOI サブフォルダ内には、さらに {AOI_ID}_b01、{AOI_ID}_b02 などのサブサブフォルダがあります。「b」は「batch」を意味します。画像数が

	多いため、予測はバッチ単位（1 バッチあたり 128 画像）で処理されています。01、02、03 はそれぞれ第 1、第 2、第 3 バッチを示します。 • 各バッチフォルダ内には labels フォルダがあり、そのバッチ内の各画像における「rainbow」の検出位置が記録されています。 • 各 AOI サブフォルダには、フォルダ全体の CSV ファイルに対応する 2 つの CSV ファイルも含まれています。違いは、それらが対応するバッチに特化した情報を含んでいる点です。
	5_traffic_data フォルダの詳細情報 • このフォルダには 4 つの CSV ファイルが含まれています： 1. **per_day_aoi**：AOI−日付レベルのデータセット 2. **per_day**：すべての AOI における検出数を日付単位で集計したデータセット • また、このフォルダには 4 つの図が含まれています： 1. **by_month**：日付レベルデータを基にした月次集計 2. **by_year**：日付レベルデータを基にした年次集計
更新履歴（版・ファイル名・年月日）	2026 年4月25日
データの形式	ビデオファイルは MP4.Sentinel 2 ファイルは tif ファイル and png ファイル。 Annotated image は coco ファイル。Prediction は CSV と JPG. Traffic data は CSV と png ファイル
データのサイズ	1.9 ギガ(zip に圧縮した状態で)
利用上の注意（メタデータの利用など）	
関連報告書・論文等	
備考	