

SONY



Vision and Sensing Application SDK CVAT画像アノテーション 機能仕様書

Copyright 2023 Sony Semiconductor Solutions Corporation

Version 0.2.0

2023 - 1 - 30

AITRIOS™、およびそのロゴは、ソニーグループ株式会社またはその関連会社の登録商標または商標です。

目次

1. 更新履歴	1
2. 用語・略語	2
3. 参照資料	3
4. 想定ユースケース	4
5. 機能概要、アルゴリズム	5
6. 操作性仕様、画面仕様	9
7. 目標性能	18
8. 制限事項	19
9. その他特記事項	20
10. 未決定事項	21

1.更新履歷

Date	What/Why
2023/01/30	初版作成

2. 用語・略語

Terms/Abbreviations	Meaning
tag/タグ	正解ラベル
annotation/アノテーション	ラベル情報、ラベリング
アノテーション情報	画像に対するメタデータ
データセット	学習やモデルの評価を行うための画像とアノテーションのデータの集合体

3. 参照資料

- Reference/Related documents (関連資料)
 - Codespaces port forwarding
 - <https://docs.github.com/en/codespaces/codespaces-reference/security-in-codespaces#port-forwarding>
 - CVAT
 - <https://github.com/opencv/cvat>

4. 想定ユースケース

- データセット作成
 - Image Classificationタスク
 - Dev Container上でCVATからデータセットを出力しシームレスにImage ClassificationのAIモデルを作成したい
 - Object Detectionタスク
 - ユーザーが持つ自前の画像データに対し、アノテーションを行いたい
 - データセットにデータを追加したりタグを編集するなど、カスタマイズを行いたい

5. 機能概要、アルゴリズム

Functional Overview

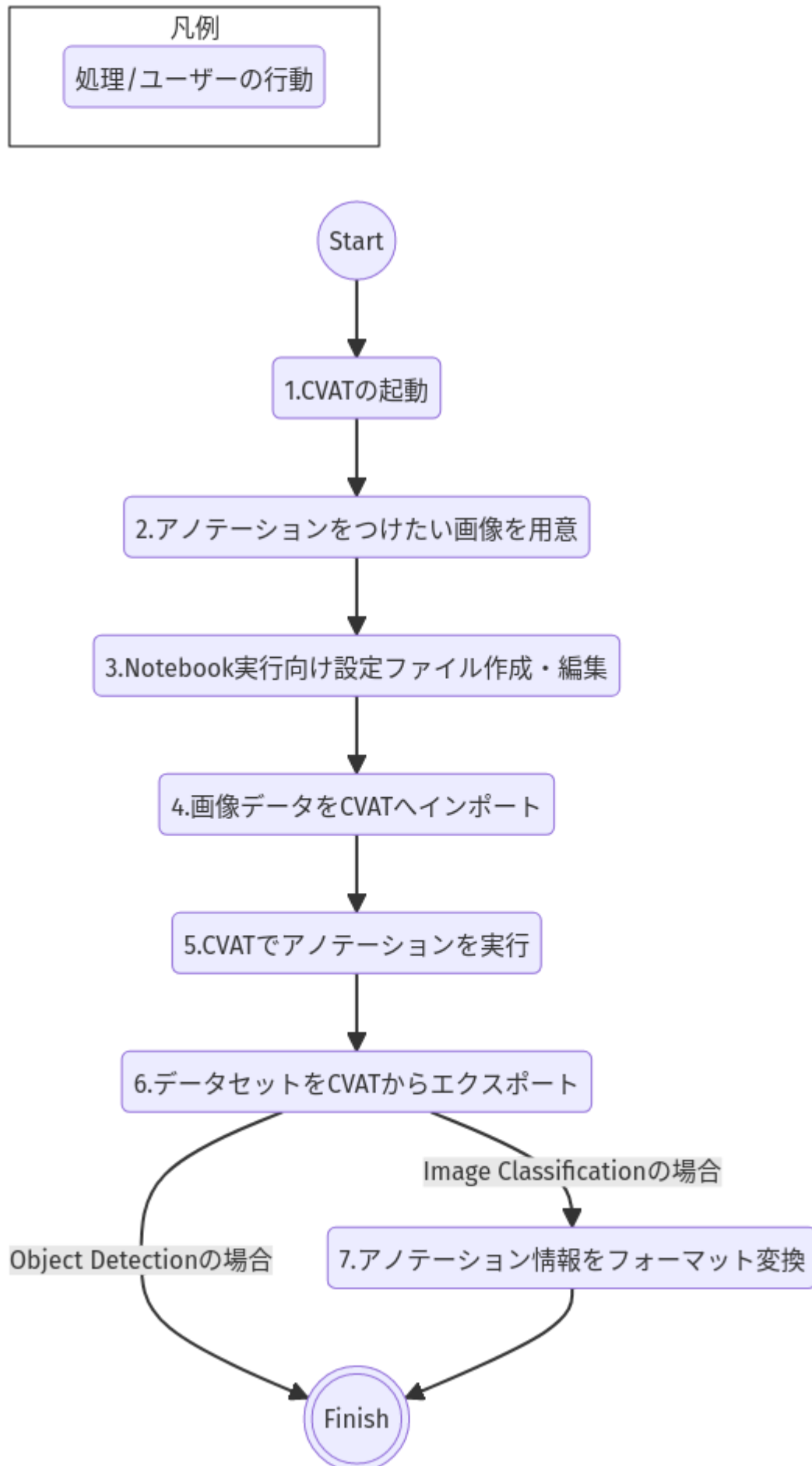
- SDKのDev Container(Local PCまたはCodespaces)において、CVATを起動する
- CVATの機能でデータのアノテーション追加や編集、マニュアルでのアノテーションを実現できる
- Local file systemやCodespacesからCVATへの画像データのインポートができる
 - インポートが可能な画像の格納先は下記表の通り
 - SDKでは「Local PC」又は「Codespaces」をサポートする
 - Codespaces上の画像をインポートする場合はCVATのAPIをNotebookで利用して行う
- CVATからLocal file systemやCodespacesへのアノテーション情報のエクスポートができる
 - エクスポートフォーマットはCVATで任意の形式を選択可能
 - Image Classificationの場合、SDKでは「ImageNet 1.0」のみサポートする
 - Object Detectionの場合、アノテーションの情報を持つ形式のみサポートする

CVATがサポートするエクスポートフォーマット	SDKにおけるサポート
CamVid 1.0	Yes
Cityscapes 1.0	Yes
COCO 1.0	Yes
COCO Keypoints 1.0	Yes
CVAT for images 1.1	Yes
CVAT for video 1.1	Yes
Datumaro 1.0	Yes
ICDAR Localization 1.0	Yes
ICDAR Recognition 1.0	No
ICDAR Segmentation 1.0	No
ImageNet 1.0	Yes

CVATがサポートするエクスポートフォーマット	SDKにおけるサポート
KITTI 1.0	Yes
LabelMe 3.0	Yes
LFW 1.0	No
Market-1501 1.0	No
MOT 1.1	Yes
MOTS PNG 1.0	No
Open Images V6 1.0	Yes
PASCAL VOC 1.1	Yes
Segmentation mask 1.1	Yes
TFRecord 1.0	Yes
VGGFace2 1.0	Yes
WiderFace 1.0	Yes
YOLO 1.1	Yes

- Image Classificationの場合、CVATからエクスポートしたアノテーション情報を、SDKでAI学習や量子化に使用するためのフォーマットに変換できる
- SDKにてサポートする画像フォーマットはJPEGとする

- フロー概要



- フロー詳細

1. CVATの起動

- READMEの手順に従いCVATをセットアップ

2. アノテーションをつけたい画像を用意

- アノテーションを付ける画像を用意する

3. Notebook実行向け設定ファイル作成・編集

- 設定ファイル[configuration.json](#)を作成・編集してNotebook実行時の設定を行う



Notebookを実行する場合のみ

4. 画像をCVATへインポート

- 画像のインポートを行うNotebook又はCVATのWeb UIを使って画像をインポートする

5. CVATでアノテーションを実行

- CVATのWeb UIでインポートした画像にアノテーションをつける

6. データセットをCVATからエクスポート

- データセットのエクスポートを行うNotebook又はCVATのWeb UIを使ってデータセットをエクスポートする

7. アノテーション情報をフォーマット変換(Image Classificationの場合のみ)

- CVATからエクスポートしたアノテーション情報をSDKでAI学習や量子化に使用するためのフォーマットに変換する

6. 操作性仕様、画面仕様

How to start each function

1. SDK環境Topの **README.md** に含まれるハイパーリンクから、**tutorials** ディレクトリの **README.md** にジャンプする
2. **tutorials** ディレクトリの **README.md** に含まれるハイパーリンクから、**2_prepare_dataset** ディレクトリにジャンプする
3. **2_prepare_dataset** の **README.md** に含まれるハイパーリンクから、**annotate_images** の **README.md** にジャンプする
4. **annotate_images** の **README.md** に含まれるハイパーリンクから、**image_classification** または **object_detection** の **README.md** を開く
5. Set up CVAT を実行し起動ログが止まるまで待機する
6. VS CodeのPort Forwardingタブで8080ポートをWebブラウザで開く
 - 起動完了しCVATログイン画面が表示されるまで待つ
 - (初回のみ) VS CodeのTerminalタブで、CVATのsuperuser権限のアカウントを作成するコマンドを実行する (コマンドは **image_classification** または **object_detection** の **README.md** に記載)
 - WebブラウザのCVATログイン画面にて、CVATのsuperuser権限のアカウント情報を入力する
 - 認証に成功するとCVATの初期画面に遷移する

アノテーションをつけたい画像を用意

1. **image_classification** または **object_detection** フォルダ配下に **images** のフォルダを作成し、CVATにインポートしアノテーションしたい画像を格納する



フォルダは任意の構成とする(子フォルダがある場合、子フォルダ内の画像もインポート対象となる)

Notebook実行向け設定ファイル作成・編集

1. 「Dev Containerのローカルストレージから画像をインポートする場合」又は「Dev Containerのローカルストレージへアノテーション情報をエクスポートする場合」又は「アノテーション情報をフォーマット変換する場合」に、実行ディレクトリへ設定ファイル (**configuration.json**)を作成し、編集する



特別な記載がある場合を除き、原則として省略は不可。



特別な記載がある場合を除き、原則として大文字小文字を区別する。



原則としてシンボリックリンクのフォルダパス、ファイルパスは使用不可。

Configuration	Meaning	Range	Remarks
cvat_username	CVATにログインするユーザー名		インポートorエクスポートする際に指定する
cvat_password	CVATにログインするユーザーのパスワード		インポートorエクスポートする際に指定する
cvat_project_id	CVATに画像をインポートorデータセットをエクスポートするプロジェクトID		インポートorエクスポートする際に指定する
import_dir	CVATにインポートしアノテーションを行う画像を格納するパス	絶対パスまたは Notebook(*.ipynb)からの相対パス	インポートする際に指定する
import_image_extension	CVATにインポートしアノテーションを行う画像の拡張子		インポートする際に指定する
import_task_name	CVATにインポートする際に作成されるタスク名		インポートする際に指定する

Configuration	Meaning	Range	Remarks
export_format	CVATからアノテーション情報をエクスポートする際のフォーマット形式		エクスポートする際に指定する
export_dir	CVATからアノテーション情報をエクスポートする保存先のパス	絶対パスまたは Notebook(*.ipynb)からの相対パス	エクスポートorフォーマット変換する際に指定する
dataset_conversion_base_file	フォーマット変換するファイルパス	絶対パスまたは Notebook(*.ipynb)からの相対パス	フォーマット変換する際に指定する(Image Classificationのみ)
dataset_conversion_dir	CVATからエクスポートしたアノテーション情報をSDKのAIモデル学習や量子化に使用するためにフォーマット変換する保存先のパス	絶対パスまたは Notebook(*.ipynb)からの相対パス	フォーマット変換する際に指定する(Image Classificationのみ)。フォルダ内に既存のデータセットが含まれている場合、エラーメッセージを表示し処理を中断する。
dataset_conversion_validation_split	データセットをフォーマット変換する際に、データセット内の画像のうち、Trainingに使用せずValidation用に使用する画像の割合	0.0より大 - 1.0より小	フォーマット変換する際に指定する(Image Classificationのみ)
dataset_conversion_seed	データセットをフォーマット変換する際に、データセット内の画像をシャッフルする際のランダムseed値	0 - 4294967295	フォーマット変換する際に指定する(Image Classificationのみ)

画像をCVATへインポート

- Dev Containerのローカルストレージから画像をインポートする場合
 1. (プロジェクトを未作成の場合のみ)CVATのWeb UIにてメニュー **[Project]** の **[+]** から **[Create a new project]** を選択し、プロジェクトを作成する
 2. 作成したプロジェクトから **[Constructor]** の **[Add label]** を選択しラベルを追加する
 3. **image_classification** または **object_detection** の **import_api.ipynb** を実行して、import_dir内の画像をインポートする(このときimport_task_nameで指定した名称でタスクが作成されプロジェクトに関連付けられる。なお同じタスク名を指定して複数回インポートを実行した場合、別タスクIDで同じタスク名のタスクが作成される。)
 - その後下記の動作をする
 - 実行ディレクトリの **設定ファイル** の存在をチェックする
 - エラー発生時はその内容を表示し、中断する
 - セルを実行中にセルの停止ボタンを押下すると処理を中断する
 - **設定ファイル** の各パラメータの存在をチェックする
 - エラー発生時はその内容を表示し、中断する
 - セルを実行中にセルの停止ボタンを押下すると処理を中断する
 - **設定ファイル** の各パラメータの値を読み取り、APIクライアント認証に必要な情報を用意する
 - エラー発生時はその内容を表示し、中断する
 - セルを実行中にセルの停止ボタンを押下すると処理を中断する
 - **設定ファイル** の各パラメータの値を読み取り、画像を読み込む
 - エラー発生時はその内容を表示し、中断する
 - セルを実行中にセルの停止ボタンを押下すると処理を中断する
 - 認証に成功し、プロジェクトに画像を表示する
 - エラー発生時はその内容を表示し、中断する
 - セルを実行中にセルの停止ボタンを押下すると処理を中断する
 - CVATのWeb UIにてプロジェクトのタスク内に画像がインポートされていることを確認できる

- Webブラウザを起動しているローカル環境から画像をインポートする場合
 1. (プロジェクトを未作成の場合のみ)CVATのWeb UIにてメニュー [Project] の [+] から [Create a new project] を選択し、プロジェクトを作成する
 2. 作成したプロジェクトの下部にある [+] から [Create a new task] を選択しタスクを作成する
 3. タスクの項目の [Select files] にある [My computer] タブの [Click or drag files to this area] を開き、画像ファイルを選択する
 4. [Submit & Open] ボタンを押下してインポートする



インポート手順は [Documentation](#) を参照

CVATでアノテーションを実行

1. 必要に応じてCVATのプロジェクトにて [Constructor] の [Add label] を選択しラベルを追加する
2. プロジェクト内のタスクにて [Jobs] を選択しジョブの画面に遷移する
3. 関連付けたいタグを [Setup tag] から選択してクリックし、画像にアノテーション付けを行う
4. 次の画像に遷移するために、画像の上部にある [>] ボタンをクリックし、次の画像に対して上記と同様にキーを押下してタグを関連づける
5. 最後の画像までアノテーションを行った後、[≡(menu)] ボタンよりメニューを表示させて [Finish the job] をクリックして完了する



アノテーション手順は [Documentation](#) を参照

データセットをCVATからエクスポート

- Dev Containerのローカルストレージへデータセットをエクスポートする場合
 1. `image_classification` または `object_detection` の `export_api.ipynb` を実行して、`cvat_project_id`で指定したプロジェクトからデータセットのエクスポートを実行する
 - その後下記の動作をする
 - 実行ディレクトリの **設定ファイル** の存在をチェックする
 - エラー発生時はその内容を表示し、中断する
 - セルを実行中にセルの停止ボタンを押下すると処理を中断する
 - **設定ファイル** の各パラメータの存在をチェックする
 - エラー発生時はその内容を表示し、中断する
 - セルを実行中にセルの停止ボタンを押下すると処理を中断する
 - **設定ファイル** の各パラメータの値を読み取り、APIクライアント認証に必要な情報を用意する
 - エラー発生時はその内容を表示し、中断する
 - セルを実行中にセルの停止ボタンを押下すると処理を中断する
 - 認証に成功し、`export_dir` に、データセットのzipファイルをダウンロードする
 - エラー発生時はその内容を表示し、中断する
 - セルを実行中にセルの停止ボタンを押下すると処理を中断する
 - `export_dir` が存在しないパスの場合は、フォルダを生成する
- Webブラウザを起動しているローカル環境へデータセットをエクスポートする場合
 1. CVATのWeb UIにてプロジェクトの [i] をクリックして、表示されたメニューより **[Export dataset]** をクリックする
 2. **[Export project ~ as a dataset]** ダイアログの **[Export format]** の中から **[ImageNet 1.0]** を選択してクリックする
 3. **[Custom name]** にダウンロードするファイルの名前を記入する
 4. **[Save images]** をチェックして画像ファイルをエクスポートファイルに含める
 5. ブラウザのダウンロード機能にてダウンロード先を指定しzipファイルをダウンロードする

- Image Classificationの場合、エクスポートしたzipファイル内のフォルダ構成は、下記のようにアノテーションの名称でフォルダがあり、各フォルダ内にアノテーションで関連付けした画像ファイルが入る構成となる

Object Detectionの場合、フォルダ構成は各フォーマットによって異なる

エクスポートしたzipファイル

```
└ タグA/  
  │  └ 画像ファイル  
  │  └ 画像ファイル  
  │  └ ・・・  
└ タグB/  
  │  └ 画像ファイル  
  │  └ 画像ファイル  
  │  └ ・・・  
└ ・・・
```

アノテーション情報をフォーマット変換(Image Classificationの場合のみ)

1. dataset_conversion_base_file にデータセットのzipファイルが存在する状態で、
image_classification の **convert_dataset.ipynb** を実行して、フォーマット変換を実行する
 - dataset_conversion_dir が tutorials/_common/dataset の場合、
tutorials/_common/datasetフォルダ内に下記のようにアノテーション情報が格納される

```
tutorials/
├── 2_prepare_dataset/
│   └── L_annotate_images/
│       └── L_image_classification/
│           ├── configuration.json
│           └── L_images/
│               ├── 画像ファイル
│               ├── 画像ファイル
│               └── . . . .
└── L_common
    └── L_dataset
        ├── *.zip (1)
        ├── cvat_exported/ (2)
        │   ├── 画像の分類名/
        │   │   └── L_画像ファイル
        │   ├── 画像の分類名/
        │   │   └── L_画像ファイル
        │   └── . . . .
        ├── labels.json (3)
        ├── training/ (4)
        │   ├── 画像の分類名/
        │   │   └── L_画像ファイル
        │   ├── 画像の分類名/
        │   │   └── L_画像ファイル
        │   └── . . . .
        └── L_validation/ (5)
            ├── 画像の分類名/
            │   └── L_画像ファイル
            ├── 画像の分類名/
            │   └── L_画像ファイル
            └── . . . .
```

- (1) 変換元データ。CVATからエクスポートしたzipファイル
- (2) 変換時中間出力データ。CVATからエクスポートしたzipファイルの中身がこのフォルダ配下に解凍される
- (3) 変換時中間出力データ。cvat_exportedフォルダから作成したlabel情報ファイル
- (4) 変換出力データ。cvat_exportedフォルダから、training用に抽出したもの

(5) 変換出力データ。cvat_exportedフォルダから、validation用に抽出したもの

- label情報ファイルのフォーマットは下記のようにlabel名とそのid値が記載されたjsonファイルとする

```
{"daisy": 0, "dandelion": 1, "roses": 2, "sunflowers": 3, "tulips": 4}
```

- dataset_conversion_dir が存在しないパスの場合は、フォルダを生成する

7. 目標性能

- ユーザビリティ
 - SDKの環境構築完了後、追加のインストール手順なしにCVATを使用できること
 - UIの応答時間が1.2秒以内であること

8. 制限事項

- CodespacesのMachine Typeが最小構成(2-core)だとCVATが起動失敗する場合があります、4-core以上のMachine Typeを選択することを推奨する
- インポートやエクスポート処理を途中でキャンセルして再開する場合、途中からの再開ではなく各処理を最初から実行する

9. その他特記事項

- SDK内で定義するエラーコード、メッセージはなし
- パスワードをドキュメントに記載することについて
 - [port forward](#)がデフォルトでprivate設定になっており、Codespacesの作成者以外がそのポートにアクセスできないようになっているため、セキュリティ上の問題はなし
- CVATのバージョン確認方法について
 - CVATのWeb UIでログイン後、ユーザー名をクリックして[**About**] をクリックすると表示されるダイアログにバージョン番号が記載されている

10. 未決定事項

なし