**Amazon SageMaker の使い方**

**アノテーション用データのアップロード**

1. Amazon S3 に移動し、任意のバケットにフォルダを作成し、画像ファイルをアップロード

図 実際の画面

1. 画像ファイルをアップロードしたフォルダと同じ位置に出力用フォルダを作成



図 ここでは、figuredatasetというバケットにobjectdetectiontestを作成し、  
その中に、入力用にsrc、出力用にdstを作成

**ジョブの詳細の指定**

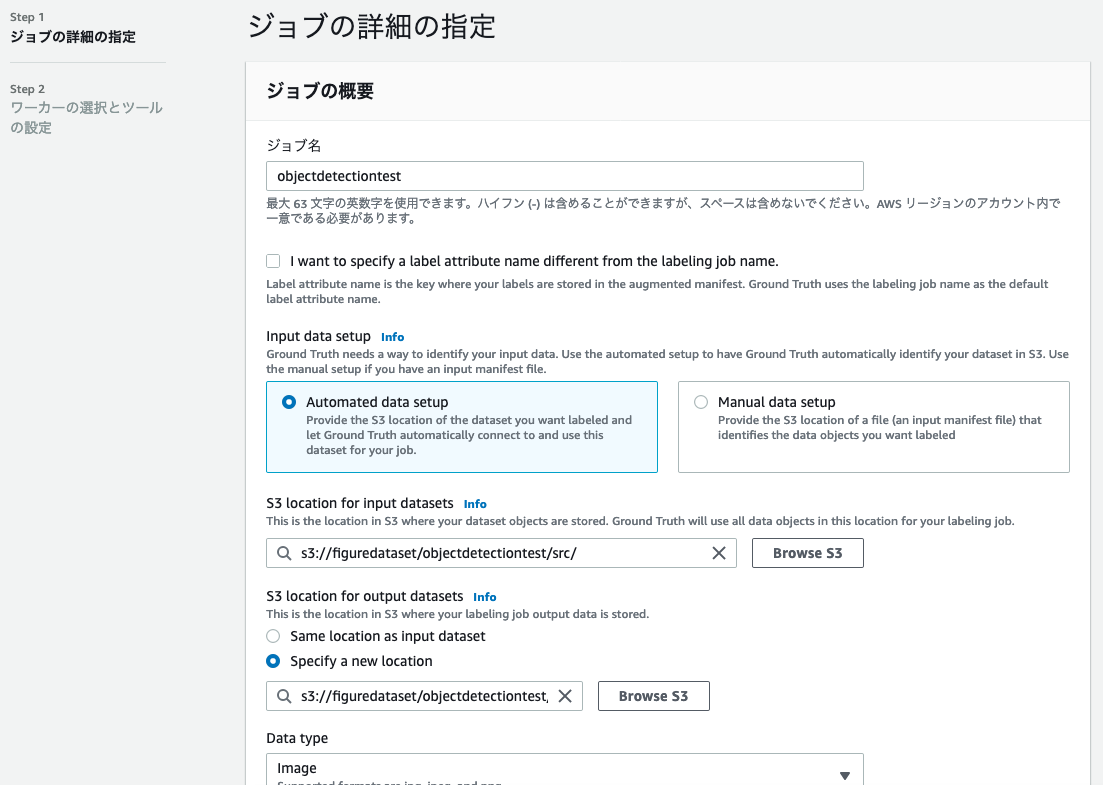
1. Amazon SageMaker > ラベリングジョブ > ラベリングジョブの作成 を選択
2. ジョブの概要を以下のように設定
   1. 任意のジョブ名を設定
   2. Input data setupをAutomated data setupに設定
   3. S3 location for input datasetsに前述の手順でアップロードした入力画像フォルダのパスを入力(この手順書では、s3://figuredataset/objectdetectiontest/src/を入力)
   4. S3 location for output datasetsでSpecify a new locationを選択し、前述の手順で作成した出力画像フォルダのパスを入力（この手順では、出力を指定したが、Same location as input datasetを選択してもよい）
   5. Data typeをImageに設定
   6. IAMロールにAmazonSageMaker-ExecutionRoleがアタッチされたロールを指定（ロールが無い場合は、ここで新規作成することも可能）

図 実際の画面

1. **ジョブの概要の設定が完了したら、Complete data setupボタンをクリック**。問題がなければ、Input data connection successfulのメッセージが出現（この操作により、入力画像フォルダにmanifestファイルが作成される）
2. 続いて、タスクのタイプを設定
   1. Task categoryをImageに設定
   2. タスクの選択で、タスクを設定（ここでは、境界ボックスを設定）

図 実際の画面

1. ここまでできたら、「次へ」をクリック

**ワーカーの選択とツールの選択**

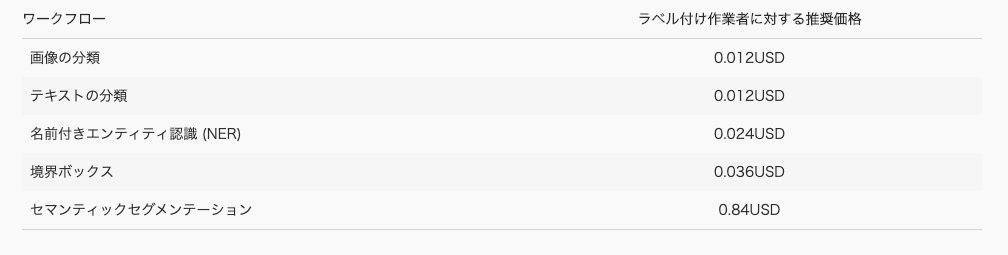
1. ワーカーを以下のように設定
   1. ワーカータイプをパブリックに設定（ここでプライベートに設定してから一旦ジョブを作成し、そのジョブをクローンすることによって、後からパブリックにすることも可能）
   2. タスクのタイムアウト（ワーカーが1つのタスクで作業できる最大時間）を設定
   3. Task expiration time（タスクが期限切れになる前にワーカーが利用可能な時間数）を設定
   4. タスクあたりの料金を設定。以下に推奨価格を示す。

図 <https://aws.amazon.com/jp/sagemaker/groundtruth/pricing/>より転載

* 1. 利用規約（アダルトコンテンツは含まれておりませんなど）にチェックを入れる

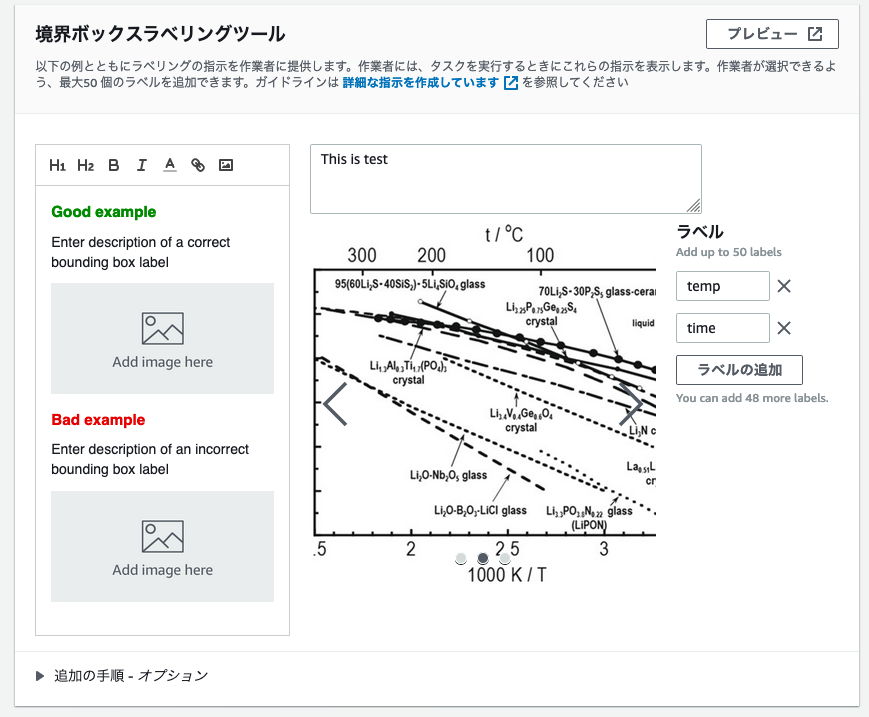
1. 続いて、ツールの設定を行う。ツールの設定は以下の画像のように、ラベルの設定や、詳細な指示の設定が可能なため、タスクに応じて設定。詳細な指示については、以下のAWSが書いた以下の記事が参考になる。  
   <https://aws.amazon.com/jp/blogs/news/create-high-quality-instructions-for-amazon-sagemaker-ground-truth-labeling-jobs/>  
   

図 実際の画面

1. ここまでできたら、「作成」をクリックすると、MechanicalTurkへ依頼がいく
2. 依頼後は、Amazon SageMaker > ラベリングジョブから、進捗状況を確認することが可能。作業完了後は出力フォルダにアノテーションデータが出力される。

図 作業の進捗状況