

# Stable Diffusion を用いた画像生成による河川巡視 AI のデータ増強

八千代エンジニアリング 正会員 ○高橋 悠太  
正会員 藤井 純一郎  
正会員 天方 匡純

## 1. 背景

我が国は多くの河川を有し、効率的な巡視のため、ドローン(UAV)と AI を用いた河川巡視の高度化が検討されている<sup>1)</sup>。河川巡視は、多様な背景の中から、多様な対象を検出する必要があるため、多様で大量の学習データを必要とするが、ドローンによる空撮に関する制限は未だ多いため、データを増やすことは容易でない。本研究では Stable Diffusion<sup>2)</sup>で生成した画像を学習に加えることでデータ増強可能か検証する。既往研究<sup>2,3)</sup>で得られた学習セットとモデルに対し、生成画像を追加して学習し、推論結果の変化を評価する。

## 2. 実験概要

学習データセットの配分は表-1 に示す。既往研究で得られたモデルを Benchmark、生成画像を加えて学習したモデルを Proposal とする。Benchmark は空撮画像が 280 枚、生成画像は 0 枚、計 280 枚で学習されている。これに対し、Proposal は空撮画像 250 枚、生成画像 443 枚、計 693 枚で学習した。Benchmark と比較し、空撮画像が少なくなっているが、これは Proposal に対し不利な条件である点に留意する。学習条件は既往の研究<sup>1)</sup>と同様の条件を用いた。

生成画像は Stable Diffusion の txt2img 機能を用い、プロンプトに入力して生成した。GUI として、ロー

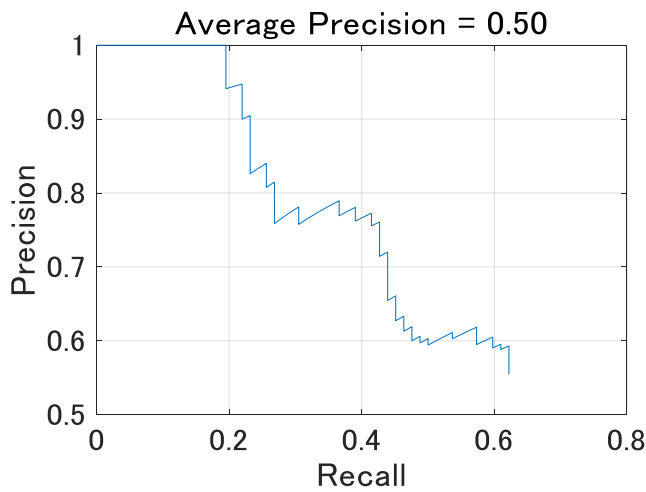
カル環境で使用可能な日本語化ツール<sup>4)</sup>を使用した。生成する画像は不法投棄として多い、ペットボトル(plastic bottle)、ゴミ袋(plastic bag)、三角コーン(safe cone)、段ボール箱(cardboard box)の四種類について、プロンプトに入力した。空撮されたような画像を得たいため、空撮画像(aerial image)を入力した。空撮画像であるため、対象は小さく写る必要があることから、Small を入力し、定冠詞 a をつけて対象が大量に生成されることを防止した。加えて Negative ワードに many, garbage を追加することで、同じく大量のゴミが生成され、実際の環境とも異なる上、検知結果の検証が困難な画像が生成されることを防止した。背景について、通常河川敷の状況は各河川で異なるため、今回草原(grass)、岩場(rock)、粘土(clay)、土(dirt)の四種類を入れ替えて生成した。また河川敷であることが強調されるよう、river side をポジティブに入れた。すなわちペットボトルが岩場にある画像を生成したい場合、例えばプロンプトは a small plastic bottle, on rock field, in river side, aerial image.となる。画像サイズは 768\*768 pixel, サンプリング回数は 30, CFG スケールは 14 とした。その他、画像生成に関する条件は GUI のデフォルトを使用した。一部三角コーンの生成画像に崩壊が見られ、追加生成したが、各背景・対象毎 25 枚を最低限学習できるよう生成した。

表-1 学習データセットの配分

	Benchmark	Proposal
空撮	280	250
生成	0	443
合計	280	693



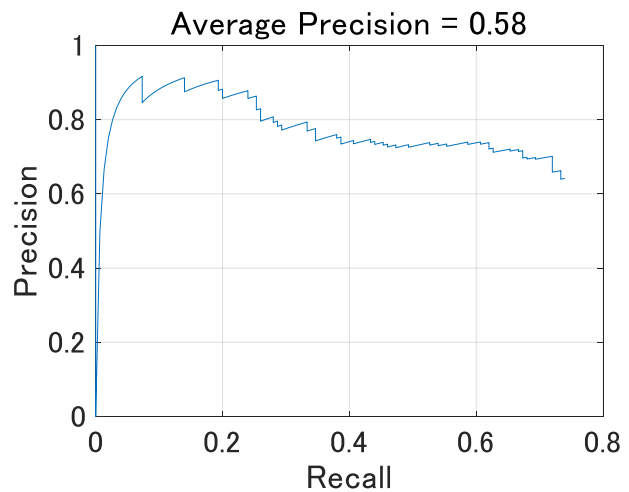
図-1 生成画像：岩場の上のペットボトル



(a) PR 曲線 : Benchmark



(c) 推論結果の一例 : Benchmark



(b) PR 曲線 : Proposal



(d) 推論結果の一例 : Proposal

図-2 Benchmark と Proposal モデルの推論結果 : PR 曲線(a),(b)と推論結果の一例(c),(d)

### 3. 結果と考察

Benchmark(既往研究の結果)と, Proposal の PR 曲線と Average precision(AP)および Benchmark と同一画像に対する推論結果を図-2 に示す. 図-2(a)において Benchmark の AP が 0.50 に対し, 図-2(b)が示す通り Proposal では 0.08 上昇していることが確認できる. 次に同じ画像に推論を行った場合の結果を比較する図-2(c)に示す通り Benchmark ではほとんど検知していないことが分かる. この時, 左上の物体は教師として設定していないため, 今回検知できていなくとも問題ない. 一方, 図-2(d)が示す Proposal の推論結果ではビニール袋やペットボトル等が検知されている. これらの結果から, 生成画像を加えることでデータセットを増強し, 学習を改善可能なことが分かる.

### 4. まとめ

ドローン河川巡視 AI に必要となる学習データを増強するため, Stable Diffusion で生成した画像を学習データセットに追加し, 学習が改善可能か検証を行った. 空撮画像のみで学習した Benchmark と生成

画像を追加した Proposal モデルで未学習画像に対し水路を行った. PR 曲線から得られた Average Precision と同一画像への推論結果の定性的比較から, 生成画像追加による増強効果を確認した.

今後, より多くの空撮画像を学習したモデルに足りない特徴量を, 生成画像によって補完可能か検証する.

### 参考文献

- 1) 高橋悠太, 藤井純一郎, 天方匡純, 山下隆義: UAV と画像認識 AI による河川巡視を補う地上画像の特徴量とその利用法検討, AI・データサイエンス論文集, 第1巻, J1号, pp.580-587, 2020.
- 2) R. Rombach, A. Blattmann, D. Lorenz, P. Esser, B. Ommer: High-Resolution Image Synthesis with Latent Diffusion Models, arXiv:2112.10752, 2021.
- 3) <https://huggingface.co/stabilityai/stable-diffusion-2-1>
- 4) <https://github.com/AUTOMATIC1111/stable-diffusion-webui>