

# 実践 2 : データ収集と画像の切り出し

土井 伸洋

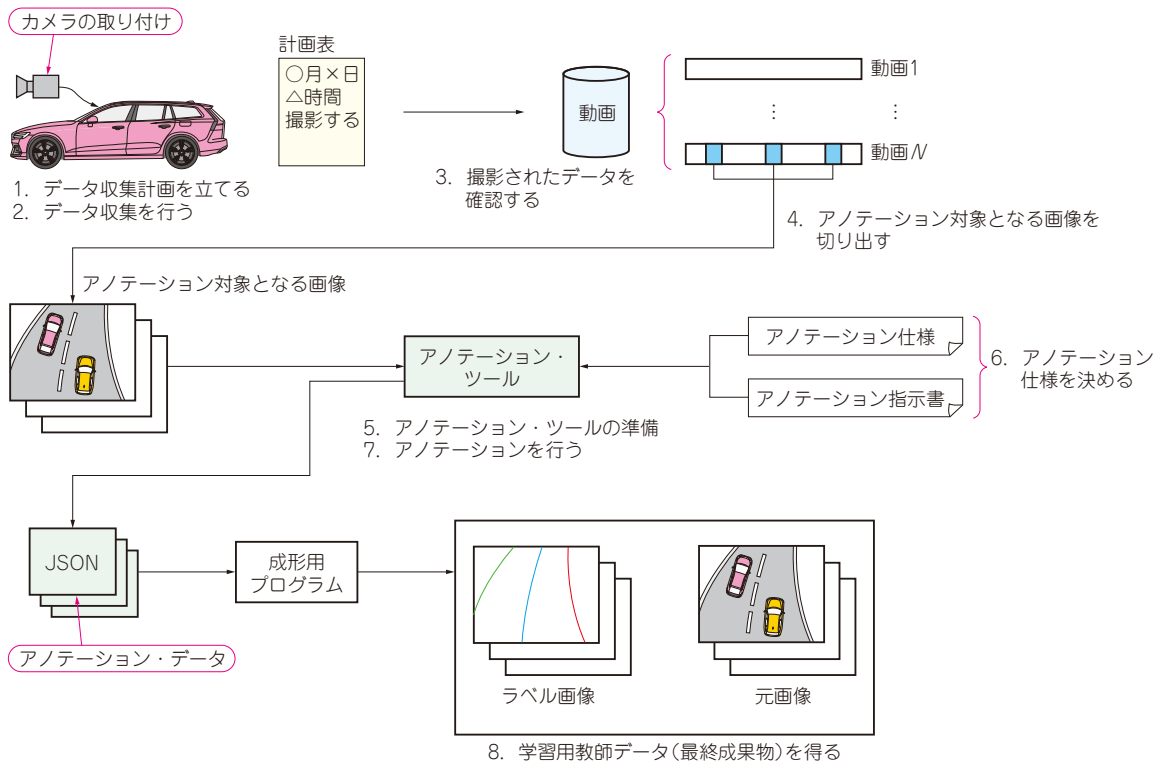


図1 データ収集からアノテーションまでの手順

問題設定ができたら、データ収集、そして教師データの作成に取り掛かります。全行程の中で最も泥臭く、根気の要る作業になります。

このフェーズでやるべきことは複数のステップに分けられます(図1)。

1. データ収集計画を立てる
2. データ収集を行う
3. 撮影されたデータを確認する
4. アノテーション対象となる画像を切り出す
5. アノテーション仕様を決める
6. アノテーション・ツールを準備する
7. アノテーションを行う
8. 後処理を行い学習用の教師データを得る

← 次章

以降、順に説明します。上記ステップの4以降では、動画像データの成形にプログラムを記述する必要があります。今回は、機械学習分野で利用者が多いPythonと、画像処理ライブラリの定番であるOpenCVを使って処理を行うことにしました

### ● 動画像データの成形に必要なプログラム環境

以下に示す性能のPCとソフトウェアを使いました。

- CPU : Core i7-6800K CPU @3.40GHz
- GPU : RTX 2080Ti (11Gバイト, エヌビディア)
- RAM : 64Gバイト
- ディスク : 1TバイトSSD