

# 深層学習を用いた石造物の検出と分類

小池隆<sup>1</sup>

**概要**：本論文では、石造物研究における深層学習の活用例として、(1) 車載カメラで撮影した動画からの石造物の検出と分類、(2) 画像認識による月待塔オープンデータのデータエンリッチメントについて報告する。

**キーワード**：石造物、深層学習

## Detection and Classification of Stone Stele using Deep Learning

TAKASHI KOIKE<sup>†1</sup>

**Abstract**: In this paper, we report two use cases of deep learning in the study of stone steles. (1) Detection and classification of stone steles from videos taken with car-mounted camera. (2) Data enrichment of Tsukimachi-to open data through image recognition.

**Keywords**: Stone stele, Deep learning

### 1. はじめに

石造物の研究には現地調査で収集されたデータが不可欠である。特に、石造物の分布や地域特性について定量的な分析をするためには、網羅的で詳細なデータが広域にわたって必要になる。

しかし、石造物悉皆調査の結果が報告書として出版されている市町村は少なく、多くの地域では、網羅性に欠けていたり、個々の石造物についての情報が足りなかったり、データが古かったりするのが現状である。さらに、デジタル化して公開された石造物データベースは極めて少ない。

そこで、技術の発展と扱いやすいツールの開発によって様々な分野へと適用範囲が広がりつつある深層学習（ディープラーニング）を用いて、物体検出による石造物データベースの作成と、既存のデータベースに対する情報の付加を試行した。

### 2. 学習モデルの作成

深層学習を石造物の検出に使用するためには、既存の学習済みモデルをそのまま流用することはできない。石造物の検出と分類に特化した学習モデルを新たに作成する必要がある。

#### 2.1 検出対象クラス

学習モデルを作成する場合、検出対象を離散的なクラスへと分けるクラス分類は重要なステップである。ドメイン

固有のニーズに最大限応えられることが求められるが、得られる学習データの量が少ない場合には、過度に詳細な分類をすると精度が低くて使い物にならないモデルになってしまう可能性もある。

石造物の分類方法には様々なものがあり、広く一般的に用いられる分類体系がある訳ではない。『日本石仏事典』では、石造物を(1)像容、(2)信仰、(3)形態の部に分け、各部毎に分類をしている[1]。市町村が発行する石造物調査報告書では主に信仰により分類されることが多い。像容により分類している『しもつまの野仏』では、例えば「観音菩薩」の項に十九夜塔／墓塔／巡拝塔など異なる信仰に基づく塔が含まれている[2]。

本研究では、石造物本体とその部分（パーツ）に分けて検出対象をクラス分類し、検出したクラスの組み合わせによって石造物を任意の観点から分類可能にするというアプローチを採用することとした。

石造物本体としては、石塔・石碑、石祠、五輪塔、石灯籠の4クラスを検出対象とした。石塔・石碑には、角柱型の一般的な石塔のほか、像が浮彫りされた光背型、不定形の自然石型、近代の石碑に多い板状型など多様な形状の石造物を含めた。

部分（パーツ）は像と文字に大別される。像としては、比較的数量が多く像容も特徴的に地蔵菩薩と如意輪観音の2クラスを検出対象とした。地蔵菩薩では坐像も立像も多く見られるが、これらを区別することなく1つのクラスとし、全身を検出範囲とした。

<sup>1</sup> 合同会社ミドリアイティ  
Midori IT, LLC

地藏菩薩は背丈より高い錫杖を持っていることがあるが、全身を収める矩形からはみ出す部分は検出範囲外とした。また、如意輪観音の宝冠は全体を検出範囲に含めた。

如意輪観音に似た像容として、赤子を抱いた子安観音があり、如意輪観音との中間的な形態のものも見られる。赤子を抱いていても、右膝を立て、右手を頬に当てた思惟相であれば如意輪観音に含めた。

像は、背景となる面を持つかどうかで浮彫り(レリーフ)と丸彫りに区別することができる。本研究における学習モデルでは、浮彫りの場合には石塔・石碑の中に像を検出し、丸彫りの場合は像のみを検出する。

文字では「庚申」「道祖」「馬頭」「甲子」「地神」「念佛」「夜」の7クラスを検出対象とした。それぞれ順に庚申塔、道祖神塔、馬頭観音塔、甲子塔、地神塔、念仏供養塔、月待塔(十九夜塔や二十三夜塔など)に多く見られる。ただし、「庚申」と「甲子」は紀年銘の中に「寛政十二庚申年」のように使用されることがあり、「夜」は石灯籠に「常夜灯」として使用されることがある。

文字のクラス分類においては、各クラスともなるべく少ない文字数となるようにした。例えば馬頭観音塔の場合、「馬頭観世音」や「馬頭尊」「馬頭大士」などと刻まれるが、これらのすべてを1つの「馬頭」クラスとして検出するためである。また、文字には崩し字や篆書体など多様な書体を使用され、陰刻や陽刻、縁彫(枠彫)などの彫り方があるが、特にそれらを区別はしなかった。

以上の13の分類クラスと『日本石仏事典』の各部との対応を表1に示す。

表1 検出対象の分類

『日本石仏事典』の部	分類クラス(ラベル)
像容	地藏菩薩(jizo) 如意輪観音(nyoirin)
信仰	庚申(koshin) 道祖(douso) 馬頭(batou) 甲子(kasshi) 地神(jijin) 念佛(nenbutsu) 夜(night)
形態	石塔・石碑(stone) 石祠(sekishi) 五輪塔(gorin) 石灯籠(toro)

## 2.2 YOLO

YOLO (You Only Look Once) は物体検出の手法である。R-CNN などの従来の手法では、物体が含まれている可能性のある領域の検出とそこに含まれる物体の分類の2段階に分けて実行する。それに対して YOLO では、領域に対する分類の信頼度という回帰問題に帰着させることにより、検出と分類を1つのニューラルネットワークで実行する。そ

のため処理速度が非常に高速である[3]。

YOLO の実装には、いくつかの種類がある。本研究は Darknet YOLOv4[4]を用いて開始した後に、Ultralytics が開発した高速で精度の高い YOLOv5[5]に移行した。

## 2.3 学習データ

学習に使用する画像ファイルは主にインターネットで収集した。以下の学習サイクルを繰り返しながら検出対象クラスを少しずつ増やすとともに、検出精度を向上させるために学習データの件数を増加させた。

- 手作業による画像ファイル収集とラベル付け
- YOLO を用いて学習モデルを作成
- Web クローラーを使用して画像ファイルを収集し、YOLO を用いて自動でラベル付け
- YOLO を用いて学習モデルを再作成

学習サイクルの初期においては、手作業あるいは精度の低い学習モデルによるラベル付けを行うため、ラベル付けが適切に行われていないこともあった。そこで、ラベルデータと YOLO による検出結果の差分を抽出するプログラムを作成し、ラベルがないのに検出されたり、逆にラベルがあるのに検出されなかったりした場合には、目視による確認をおこない、適宜ラベル付けを修正をした。

最初は 500 枚の画像を用い、2 クラスの検出を学習させた。上記の作業を繰り返し、現時点では 17,296 枚の画像を用いて 13 クラスの検出を学習させた。画像ファイルのうち 3,953 枚はラベル付けの対象外であり、これは False Positive を防ぐための学習データである。ラベル付けされた件数は分類クラスによって大きく異なる。「石塔・石碑」は他の多くの分類クラスに共通して現れ、1 つのファイルで複数回ラベル付けされることもあるため最も多い。比較的最近になって分類クラスに追加した「甲子」や「地神」はまだ少ない。詳細な内訳は表2の通りである。

表2 分類クラス毎のラベル付け件数

分類クラス(ラベル)	件数
石塔・石碑(stone)	16,419
夜(night)	3,253
庚申(koshin)	2,688
地藏菩薩(jizo)	2,295
道祖(douso)	1,589
石灯籠(toro)	1,373
如意輪観音(nyoirin)	1,332
馬頭(batou)	903
石祠(sekishi)	515
五輪塔(gorin)	464
念佛(nenbutsu)	308
甲子(kasshi)	147
地神(jijin)	128

### 3. 車載カメラ動画からの石造物の検出

作成した石造物の学習モデルを利用して、車載カメラで撮影した動画から石造物の検出を実施した。

#### 3.1 動画の撮影

ジンバル付きの小型ハンドヘルドカメラを車内に設置し、車両進行方向に対して左側方の路傍風景を撮影した。右側方の場合、石造物までの距離が遠すぎたり、対向車線の車両によってカメラ視界が塞がれる可能性があるためである。解像度はフル HD (1920px×1080px)、フレームレートは 60fps とした。使用したカメラは 4K (3820px×2160px) での撮影も可能だが、より長時間録画するために解像度は低く設定した。

#### 3.2 検出

石造物の置かれた周囲の環境を含めて記録するため、物体検出した前後 15 フレームを含む動画を保存するように、YOLOv5 に含まれる検出プログラム detect.py を改造して使用した。また、動画に写った石造物の目視による確認が容易になるよう、実際には 60fps の動画を 10fps として保存し、スローモーションで再生されるようにした。

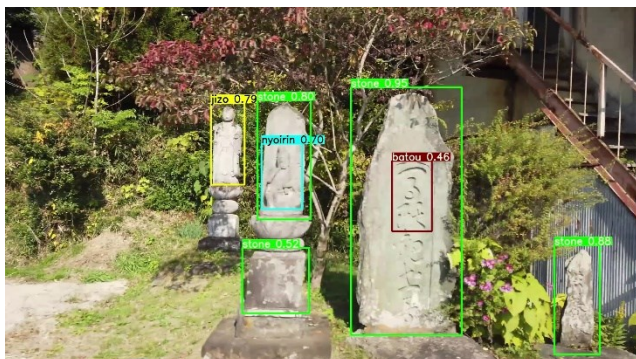


図 1 検出例

図 1 は、保存された検出結果の動画の 1 フレームである。4 つの石塔・石碑、地藏菩薩、如意輪観音、「馬頭」の文字を検出している。

#### 3.3 マッピングと公開

本研究において使用したカメラでは、保存した動画ファイルに位置情報は記録されない。そのため、別途 GPS ログを取得しておき、動画と GPS ログを時刻で突き合わせることで石造物の位置を特定した。

石造物検出結果の動画は、位置情報を含む GeoJSON 形式のメタ情報ファイルを用いて地図上にマッピングし、Web で公開した [6]。マーカーをクリックするとポップアップウィンドウで動画を再生することができる。メタ情報と動画ファイルのライセンスは CC BY とした。

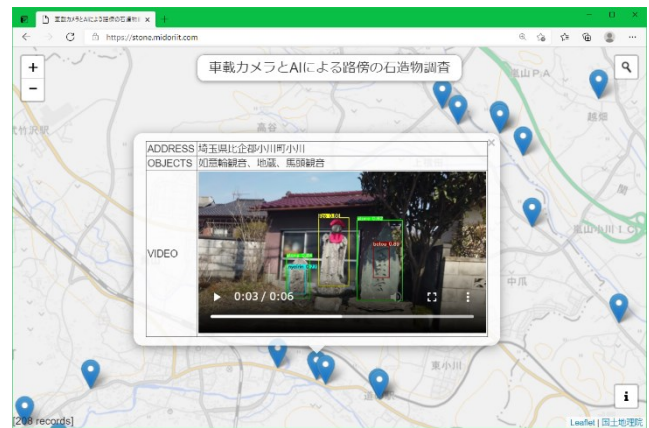


図 2 検出結果動画のマップ表示

### 4. データエンリッチメント

元データに情報を付加してより有用なものに拡張するデータエンリッチメントを、月待塔オープンデータに対して適用した。

#### 4.1 月待塔オープンデータ

月待塔オープンデータは、市民参加型の石造物調査プロジェクトである「月待ビンゴプロジェクト」[7]によって作成されたデータであり、CSV、GeoJSON および KML 形式のファイルを GitHub 上で CC0 で公開している[8]。

データは、参加者が Twitter に投稿したツイートから自動的に生成され、ツイートとツイートに添付された写真の URL、月待塔の種類、位置情報、ツイートから抽出した造立年、POI (point of interest) 等の情報が含まれる。

月待ビンゴの参加者は、「#月待ビンゴ」のハッシュタグを付け、緯度経度を記載して月待塔の写真を添付してツイートする。月待塔には「十九夜」「二十三夜」のように月待行事を行う「アタリ日」が刻まれており、これらを「#19」「#23」のようにツイートに含めることによって月待塔の種類を示すとともに、ビンゴゲームとしてのゲーミフィケーションを実現している。

2021 年 3 月 31 日末時点で、延べ 20 名を超える参加者のツイートから、21 都県に分布する 3,923 件の月待塔データが生成された。全ツイートを地図上にマッピングして表示した「月待ビンゴ 攻略マップ」[9]のほか、月待塔の種類毎の地域分布密度を簡易的に可視化した「月待塔ヒートマップ」[10]、造立年による月待塔の広がり可視化する「月待塔 時系列マップ」[11]などを提供している。

#### 4.2 物体検出によるツイートへのタグ付け

2021 年 3 月 31 日までのデータ 3,923 件に対して、ツイートに添付された写真から物体検出し、人手により精査されていない AI による「タグ付け」という意味で ai\_tags 列に格納した。タグの対象クラスは、「庚申」「念佛」の文字、

地藏菩薩と如意輪観音の像に限定した。

「庚申」と「念佛」については、おなじ画像の中に「夜」も検出し、さらにそれらの中心点が同じ石塔・石碑のバウンディングボックス内であることを条件とした。これは、月待塔とは別に庚申塔や念仏供養塔などがある場合にタグ付けを行わないようにするためである。

地藏菩薩と如意輪観音の像については、「夜」を検出した場合には、像と「夜」の中心 X 座標が 1 つの石塔・石碑の X 座標範囲内に含まれることを条件とした。Y 座標を判定に用いない理由は、丸彫りの場合には像と石塔・石碑が縦に並ぶためである。光背型の刻像塔では文字が小さく検出が困難なこともあるため、「夜」を検出しない場合には像のみの検出でタグ付けをした。

タグ毎のツイート件数を表 3 に示す。1 つのツイートに複数のタグが付けられる場合もあるため、何らかのタグが付けられたツイートの数は、タグ毎のツイート数の合計よりも少ない 1,302 件となった。

表 3 タグ付けされたツイート件数

タグ	ツイート件数
庚申	20
念佛	248
地藏菩薩	73
如意輪観音	1,176

### 4.3 併刻塔の発見

図 3 は、「庚申」タグの付けられたツイートの画像のうち、二十三夜の月待と庚申塔という異なる信仰に基づいて造立された、併刻塔と呼ばれる珍しいタイプのものである。

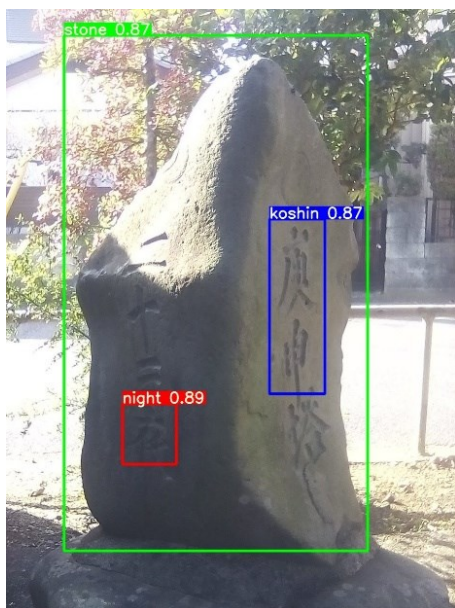


図 3 二十三夜と庚申の併刻塔

このような併刻塔の発見を期待して「庚申」のタグ付けを実施したが、実際には、紀年銘の中にある「庚申」を検出する場合も多く、図 4 のように、「庚申」の中心が「夜」の刻まれた石塔・石碑のバウンディングボックス内にぎりぎり含まれることによって期待と異なるタグ付けされたケースもあった。

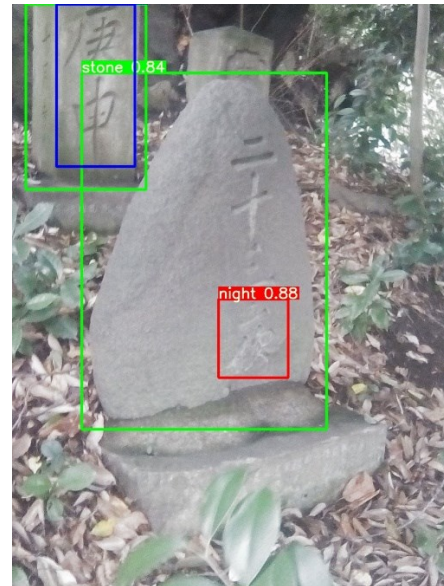


図 4 「庚申」タグが付けられたツイートの画像例

「庚申」タグが付けられた 20 件のツイートのうち、併刻塔は 6 件、紀年銘に「庚申」が含まれていた例は 5 件、図 4 と同様の例が 7 件、その他の誤検出によるものが 2 件であった。

### 4.4 月待塔と「念佛」

「念佛」のタグが付けられた 248 件のツイートを月待塔の種類（異なるアタリ日との併刻塔は除く）毎に集計した結果を表 4 に示す。ツイート数では圧倒的に十九夜塔が多く、十九夜塔の 15.8% に「念佛」が含まれていた。比率では十六夜塔の 34.2% が最も高かった。

表 4 「念佛」を含む月待塔の種類別集計

月待塔の種類	「念佛」あり	ツイート総数
十九夜塔	223	1,408
十六夜塔	13	38
二十一夜塔	4	17
十五夜塔	3	20
十四夜塔	1	3
十七夜塔	1	56
十八夜塔	1	61
二十夜塔	1	56
二十二夜塔	1	304



二十一夜塔と二十二夜塔は、(1)女人講による造立が多く、(2)刻像塔の場合には如意輪観音が彫られることが多いという、十九夜塔との共通点を持っている。それにもかかわらず、二十二夜塔ではほとんど「念佛」が検出されなかったということは注目に値する。

月待塔オープンデータでは二十三夜塔が 1,767 件と最も多いが、「念佛」を含む二十三夜塔は 1 つもなかった。

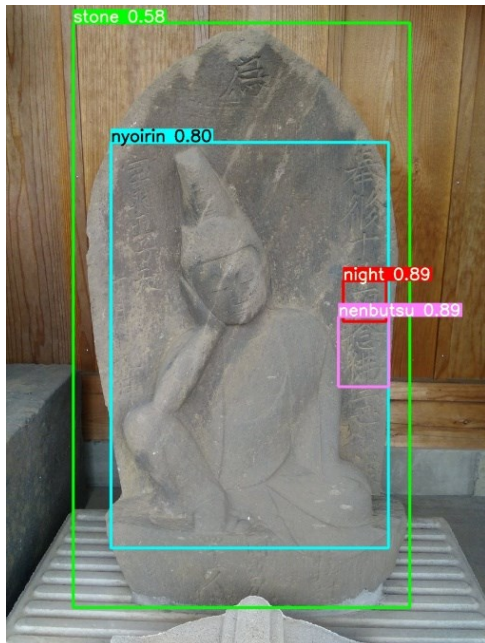


図 5 十九夜塔における「念佛」の検出例

「念佛」を含む十九夜塔を都道府県毎に集計した結果を表 5 に示す。

表 5 「念佛」を含む十九夜塔の県別集計

県	「念佛」あり	ツイート総数	「念佛」率
千葉県	103	645	16.0%
栃木県	43	254	16.9%
茨城県	34	164	20.7%
群馬県	25	90	27.8%
埼玉県	15	81	18.5%
福島県	2	136	1.5%
長野県	1	34	2.9%

関東地方の 5 県と比べて、福島県と長野県において十九夜塔に「念佛」が含まれる割合が著しく低い。十九夜塔の主な造立者である十九夜講について、『日本石仏事典』では「十九夜の月待講といわれるが婦人たちの念仏講でもある」[12]と記されているが、関東地方と福島・長野県とでは状況が異なる可能性もある。

今回のデータ分析は、あくまでも月待ビンゴの参加者によって集められたデータに基づく分析であり、調査地域に

偏りがあることは否めない。

また、必ずしも石造物に信仰の実態が反映されているとは限らない。自然石を利用した文字塔の多い地域では、石造物から得られる文字情報は少なくなっている可能性がある。

データ分析から得られた結果の解釈においては、民俗調査の記録も合わせて検討しなければならない。しかし、既に多くの地域において月待信仰は失われてしまっているため、石造物データの定量的な分析は今後さらに重要な研究手段となる。

## 5. おわりに

深層学習を用いて路傍の石造物を発見し、データベース化できることが示された。さらに、石造物のデジタル画像を含むデータベースにおいては、深層学習を用いたデータエンリッチメントによってさらに多くの知見が得られる可能性が示された。

今後は、データに基づく定量的な石造物研究の推進に向けて、さらに有用なデータを蓄積できるよう、多様な石造物の検出と分類を可能にしたい。

YOLOv5 はカメラからの映像入力をリアルタイムに処理することもできるが、現時点で車載可能な PC の処理能力の制約もあり、本研究においては録画したファイルからの検出を実施した。しかし、コンピュータの処理能力向上は目覚ましい。PC やスマホが現在とは比べ物にならないほど圧倒的なパワーを持った未来を見据えて、本研究を継続的に発展させていきたい。

## 参考文献

- [1] 懇親懇話会. 日本石仏事典. 雄山閣出版, 1980
- [2] 下妻市石仏調査委員会. しもつまの野仏: 人々と石仏との語らい. 下妻市教育委員会, 1991
- [3] Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi. You Only Look Once: Unified, Real-Time Object Detection. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2016, pp. 779-788
- [4] “AlexeyAB/darknet”. <https://github.com/AlexeyAB/darknet>
- [5] “ultralytics/yolov5”. <https://github.com/ultralytics/yolov5>
- [6] “車載カメラと AI による路傍の石造物調査”. <https://stone.midoriit.com>
- [7] “月待ビンゴプロジェクト”. <https://moon.midoriit.com>
- [8] “月待塔オープンデータ by 月待ビンゴ”. <https://github.com/midoriit/tsukimachito>
- [9] “月待ビンゴ攻略マップ”. <https://moon.midoriit.com/map/>
- [10] “月待塔ヒートマップ”. <https://moon.midoriit.com/map/heatmap.html>
- [11] “月待塔時系列マップ”. <https://moon.midoriit.com/map/timeseries.html>
- [12] 懇親懇話会. 日本石仏事典. 雄山閣出版, 1980, p.171.