



Title: Estimating Anime Sacred Sites Based on Geolocalization by Hybrid Model of Classification and Regression (分類と回帰のハイブリッドモデルによる地理位置情報に基づくアニメ聖地の推定)

Authors: Weixiang Wang, Kaito Honda and Takayasu Fushimi

(王 威翔(研究室の学生), 本多 海斗(研究室の学生), 伏見 卓恭(東京工科大学・CS 学部・講師))

Journal: Proceedings of the 23rd IEEE/WIC International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT2024)

掲載年月: 2025 年 6 月

研究概要: 本研究では、アニメ画像からその撮影場所（聖地）を高精度で推定する新たなハイブリッドモデルを提案しています。従来の地理推定は大まかな地域や都市単位の分類にとどまることが多いですが、本研究では分類と回帰の手法を組み合わせることで、より詳細な位置情報を特定できるように工夫しています。訓練には、実際の写真にアニメスタイルを適用したスタイル変換画像を利用し、多角的なアプローチで学習を行うことで、誤差の少ない位置推定を可能にしました。これにより、アニメの舞台となった場所を正確に特定し、観光や文化振興に寄与できる技術の基盤を築いています。

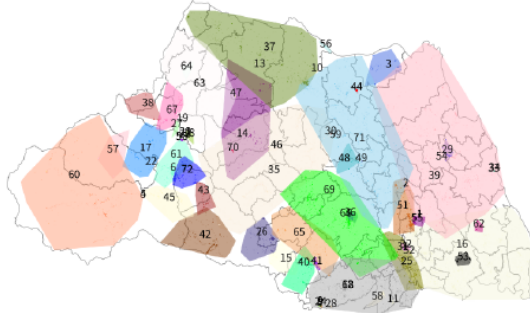
研究背景: 近年、アニメの海外人気とともに、アニメに登場する聖地巡礼が大きな社会現象となっています。これに伴い、地域経済への貢献や地域振興の切り札としての役割も期待されています。一方、従来の地理推定技術は主に写真から大まかな地域や都市レベルを推定できるものの、具体的な場所の特定には限界があります。特に、アニメのシーンから詳細な場所を特定するには、従来手法の精度向上が不可欠です。こうした課題を解決するために、本研究では分類と回帰を融合したハイブリッドモデルの必要性和、その可能性を模索しています。

研究成果: 本研究の重要な成果は、従来よりも正確にアニメの聖地を推定できるハイブリッドモデルの構築とその有効性の確認です。特に、スタイル変換を施した画像をトレーニングデータとして活用することで、多様なスタイルや環境条件に対応し、誤差を大きく削減しました。実験結果では、位置推定の誤差が従来手法よりも大幅に減少し、特に制約の少ない正確な場所の特定を実現しています。これにより、アニメに登場する場所の詳細特定や、ファンによる訪問支援、さらには地域振興や観光促進へ向けた新たなアプローチの可能性が示されました。

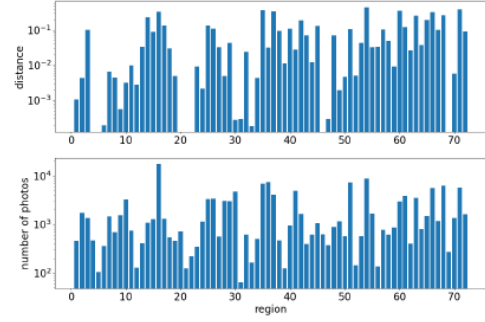
社会への影響: 本研究は、アニメファンや観光業界、地域振興に大きな貢献が期待されます。具体的には、アニメの舞台となる場所を高精度で特定できるため、ファンが気軽に聖地巡礼を楽しめる環境を創出します。また、少ない情報でも場所を特定できるため、情報不足や事前知識の少ない観光客の誘導に役立ちます。さらに、これにより地域の観光資源としてのアニメ聖地の価値が向上し、経済効果や地域振興、文化継承に寄与します。加えて、観光客の安心・安全な巡礼を促進し、地域の持続可能な発展や多文化交流の促進にもつながると期待されます。

専門用語:

- ハイブリッドモデル: 複数の手法（分類と回帰）を組み合わせ、より高精度な位置推定を実現する手法。
- スタイル変換 (Style Transfer): 実写画像をアニメ風に変換する人工知能技術で、データ拡張や特徴抽出に利用される。
- SMOTE (Synthetic Minority Over-sampling Technique): 不均衡なデータセットに対して少数クラスのデータを増やす手法で、分類の学習精度向上に利用される。
- 位置推定 (Geolocalization): 画像や映像から撮影場所の緯度経度を特定する技術や作業。



(a) Visualization result



(b) Size of area and amount of photos

図：類似写真投稿領域と各領域のサイズと写真投稿件数（埼玉県）

表：各種法の予測精度

TABLE I: Accuracy of Label Prediction for Segmented Regions

(a) Training: Real Photos / Test: Real Photos					(b) Training: Real Photos / Test: Anime Images				
	Acc	Pre	Rec	F1		Acc	Pre	Rec	F1
AlexNet	0.427	0.263	0.428	0.316	AlexNet	0.081	0.031	0.007	0.011
VGGNet	0.391	0.231	0.392	0.276	VGGNet	0.270	0.077	0.037	0.050
GoogLeNet	0.027	0.015	0.027	0.006	GoogLeNet	0.000	0.000	0.000	0.000
ResNet	0.076	0.019	0.077	0.015	ResNet	0.000	0.000	0.000	0.000

(c) Training: Style-Transferred Images / Test: Style-Transferred Images					(d) Training: Style-Transferred Images / Test: Anime Images				
	Acc	Pre	Rec	F1		Acc	Pre	Rec	F1
AlexNet	0.372	0.221	0.372	0.262	AlexNet	0.216	0.041	0.020	0.027
VGGNet	0.376	0.228	0.376	0.267	VGGNet	0.216	0.061	0.025	0.035
GoogLeNet	0.032	0.029	0.031	0.012	GoogLeNet	0.000	0.000	0.000	0.000
ResNet	0.062	0.009	0.062	0.014	ResNet	0.000	0.000	0.000	0.000

TABLE II: Accuracy of Label Prediction for Administrative Divisions

(a) Training: Real Photos / Test: Real Photos					(b) Training: Real Photos / Test: Anime Images				
	Acc	Pre	Rec	F1		Acc	Pre	Rec	F1
AlexNet	0.166	0.084	0.167	0.090	AlexNet	0.000	0.000	0.000	0.000
VGGNet	0.168	0.069	0.168	0.089	VGGNet	0.189	0.071	0.024	0.035
GoogLeNet	0.028	0.024	0.029	0.013	GoogLeNet	0.000	0.000	0.000	0.000
ResNet	0.023	0.001	0.023	0.002	ResNet	0.000	0.000	0.000	0.000

(c) Training: Style-Transferred Images / Test: Style-Transferred Images					(d) Training: Style-Transferred Images / Test: Anime Images				
	Acc	Pre	Rec	F1		Acc	Pre	Rec	F1
AlexNet	0.155	0.058	0.154	0.077	AlexNet	0.000	0.000	0.000	0.000
VGGNet	0.160	0.059	0.159	0.081	VGGNet	0.081	0.054	0.008	0.014
GoogLeNet	0.028	0.021	0.028	0.013	GoogLeNet	0.000	0.000	0.000	0.000
ResNet	0.017	0.004	0.017	0.002	ResNet	0.000	0.000	0.000	0.000