

(44) Lightweight Road Manager : 深層学習を用いた 路面損傷自動判定スマートフォンアプリ

前田 紘弥¹・関本 義秀²・瀬戸 寿一³

¹ 非会員 東京大学学生 工学系研究科社会基盤学専攻 (〒153-8505 東京都目黒区駒場 4-6-1)

E-mail:maedahi@iis.u-tokyo.ac.jp

² 正会員 東京大学准教授 生産技術研究所 (〒153-8505 東京都目黒区駒場 4-6-1)

E-mail:sekimoto@iis.u-tokyo.ac.jp

³ 非会員 東京大学特任講師 空間情報科学研究センター (〒277-8568 千葉県柏市柏の葉 5-1-5)

近年、情報通信技術 (ICT) の普及によって、例えば FixMyStreet のように、市民が地域の課題 (道路の損傷など) を発見すると、スマートフォンを用いて位置情報等を付与した形で、行政にレポートするという仕組みが世界中の先進的な自治体で始まっている。昨今の地方自治体のインフラ維持管理を取り巻く厳しい状況を鑑みるに、そのような仕組みで大量に蓄積される地域の課題データをインフラ維持管理に役立てられないかという期待感がある。そこで本研究では、サーバー側に日々蓄積される市民の投稿データを、サーバー側で学習しつつ、スマホアプリ側で起動時にサーバーから最新の判定器パラメータを更新して、瞬時に判断できる仕組みを作った。その結果、一般的なスマートフォンで 1.1 秒で 80% を超える精度で判定することができた。

Key Words: Smartphone application, deep learning, road damage status, citizen data

1. はじめに

日本のインフラ維持管理を巡る状況が厳しいことは周知の通りである。このような状態は、財政や専門家が不足している市区町村などで特に深刻であり、そもそも十分なデータ・工学的知識・技術力を有していない地方自治体も多い¹⁾。

一方で、近年世界中の先進的な自治体では ICT を最大限活用し、市民協働で地域の課題を解決してく仕組み運用され始めている²⁾。イギリスの FixMyStreet³⁾、アメリカのシカゴ 311⁴⁾、日本国内では本研究の研究対象である千葉県千葉市のちばレポ⁵⁾をはじめとして、年々数が増えつつある。このような仕組みでは、市民が見つけた道路の損傷等の状況を示す写真をスマートフォンで撮影し、位置情報付きで専用の Web サイトに投稿する仕組みが基本である。このようにして、専門家ではない市民目線ではあるが、損傷箇所の画像と位置情報、損傷状況の説明という貴重なデータ (市民データ) が蓄積されていくことになる。

財源・専門家の不足などのインフラを取り巻く厳しい状況の中で、市民参加型の仕組みで蓄積される市民データを行政の維持管理業務に役立てたいとともに、こうし

たまちづくりの活動に市民を巻き込み理解を深めて欲しいという期待感がある。

しかしながらこのような仕組みで蓄積される市民データをインフラの維持管理に実務レベルで役立て、運用を継続していくためには解決せねばならない課題がある。その最たるものは、道路の損傷程度の判定がそもそも素人には難しく、ゆえに軽微な損傷 (行政にとっては修繕不要と考えられる箇所) が多数レポートされてしまうことである。現状、こう言った軽微な損傷に関するレポートを行政担当職員が一つ一つ目視点検をしており、行政にとってコストとなりうる。

そこで、本研究では素人 (専門ではない行政職員も含む) が道路の損傷程度を判断することを補助するスマートフォンアプリケーション (Lightweight Road Manager : LRM) を開発した。これによって、行政の負担になりうる軽微なレポートが減ることが期待される。技術的には、千葉市役所管轄内の土木事務所職員 (道路管理者) によって「損傷なし」「損傷あり、修繕不要」「損傷あり、修繕必要」という三段階でラベルづけされた画像を深層学習を用いて学習させ、LRM に学習済みモデルを組み込んでいる。さらに、本研究で開発された仕組みでは、日々大量に蓄積される市民の声と行政の点検結果をサー

バー側で学習し、アプリ起動時に新たに学習されたパラメータを更新するようになっている。いうなれば、日をおうごとに、データが増えるごとに、アプリは賢くなるのである。最終的に5章で述べるように、一般的なスマートフォンで81.4%の精度で、1.08秒という処理速度で損傷程度を判定することができた。

LRMを用いれば専門知識を持たない行政職員も点検作業に貢献できるようになるので、専門家不足に直面している地方自治体にとっても活用の可能性が大いにあることはいうまでもない。さらに、ちばレポのような市民協働システムの継続的な運用に大きな貢献をするだろう。

2. 関連研究

(1) ICT を用いた自治体の市民協働システム

ICT を用いた自治体の市民協働協働システムの例を表-1に示す。表-1はOffenhuberら²⁾を参考に筆者が加筆したものである。このような仕組みが世界中で運用されており、ちばレポの効率的な運用や有効活用法の確立は横展開の可能性を秘めたものであると言える。

(2) 道路路面の点検

近年では道路路面の点検をより効率的に点検を行うために、画像処理技術を用いて、車載カメラから路面性状を把握しようとする試みも多く存在する⁶⁾⁷⁾⁸⁾⁹⁾。しかし、このような研究では比較的性能の高い専用カメラとサーバーを用いていることがほとんどである¹⁰⁾。本研究では、あくまでも広く普及している器具のみを用いて、専門的な知識を有していない市民が、道路の点検を行うことに主眼を置いているため、一般的なスマートフォンのみを用いている。

表-1 世界中に広がる市民協働システムの例

System	Since	Initiated by	Area
Chiba Report (Chiba-repo)	August 2014	Public sector	Chiba City
Cambridge iReport	December 2011	Public sector	Cambridge, MA
City sourced	September 2009	Start-up	Location independent
Standard open311 Specification	June 2009	Consortium/Public	Location independent
Ushahidi	January 2008	Start-up/NPO	Location independent
FixMyStreet	February 2007	NPO	UK, Japan
NYC311 online	March 2003	Public sector	New York City

(3) 深層学習を用いた画像認識とモデルの軽量化

近年、畳込ニューラルネットワーク(CNN)を用いた画像認識と画像分類における研究が数多くなされている。CNNの良い点は特徴量の重みだけでなく、特徴量そのものも学習することができる点である。CNNはここ最近で最も高い判定精度を実現している。深層学習を用いた画像処理などについては、Srinivasら¹¹⁾に詳しい。ただし、CNNは、学習させるパラメータ数が多く、学習済みのモデルが数百MBを超えることも珍しくない。また、パラメータ数の分だけ、学習にも判定にも時間がかかってしまう。よって、学習済みモデルをスマートフォン上で実行させることを考えると、大きな容量・パラメータ数がボトルネックとなり、処理が遅くなってしまふ。このような背景でCNNの構造をできるだけ軽量化しようという試みが多くなされている¹²⁾。本研究では、CNNの基本的な構造となっている、AlexNet¹³⁾について、畳込層のフィルタ数と全結合層のニューロンの数のみを変更して、学習済みモデル容量・パラメータ数を調整し、処理速度や精度がどのように変化するかを考察した。

3. Light-Weight Road Manager

図-1と図-2に本研究で提案する仕組みの全体像とアプリそのものの画面キャプチャを示す。この仕組みは四つのステップから成り立つ。

Step 1. 行政側サーバーに日々蓄積される市民からのレポートと行政の点検結果を学習

Step 2. 市民が最新のモデルをスマートフォンにダウンロード

Step 3. スマートフォンアプリケーション LRM を用い道路路面の損傷程度を判定

Step 4. 行政側サーバーにレポート

という、上記四つのステップを繰り返すことで、LRMはより賢く・効果的になるのである。



図-1 本研究で提案するシステムの全体像

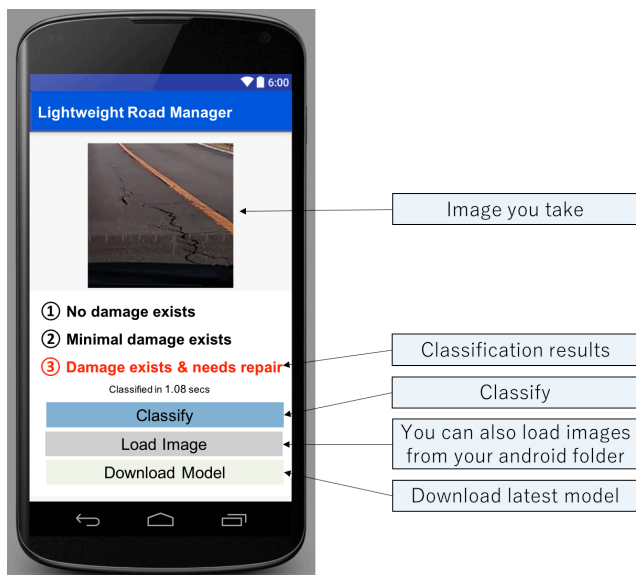


図-2 LRM の画面イメージ

学習に用いる教師データを作成するにあたり、まずはレボを通して蓄積された画像を用いることを考えた。しかしながら、ちばレボは地域の課題をレポートするという仕組み上、損傷なしの画像を収集することが困難であった。そこで筆者らは、教師データを作成するために、千葉市内約 500km の路線をスマートフォンで路面を動画撮影しながら走行した。撮影された動画から路面損傷と思われる画像を筆者らが目視で抽出した。次に、目視で抽出した画像に対して、損傷程度の判定を、千葉市内 4 つの土木事務所の方々に管轄区ごとに依頼した。三つの区分：①損傷なし、②損傷あり修繕不要、③損傷あり修繕必要は千葉市役所内土木事務所が通常業務で使用している判断基準である。その結果、図-3 のような合計 1003 枚の教師データを作成することができた。ただし、4 つの異なる土木事務所にそれぞれ教師データ作成を依頼しているので、各土木事務所間で微妙な差異があり、十分な教師データを作成できていない恐れがある。また、本研究では data augmentation を行い、各クラスが 500 枚ずつになるようにして学習を進めた。一般に data augmentation は教師データが少ない場合に、過学習を防ぐ効果があり、より汎用性の高いモデルを学習できるとされる。特にサンプルのばらつきが予想できる場合に有効である¹⁴⁾。道路をスマートフォンで撮影する場合、日射の有無・手の角度によって大きな影響を受けると考えられる。そこで、本研究では、data augmentation として日射の有無に対応するために画像の色彩や明るさを変更し、手の角度の変化に対応するためにアフィン変換を施している。

本研究では一般的なスマートフォンでどの程度の性能が出るかの目安を作るために、ニューラルネットワークの基本的な構造は固定して、カーネル数のみを変化させ

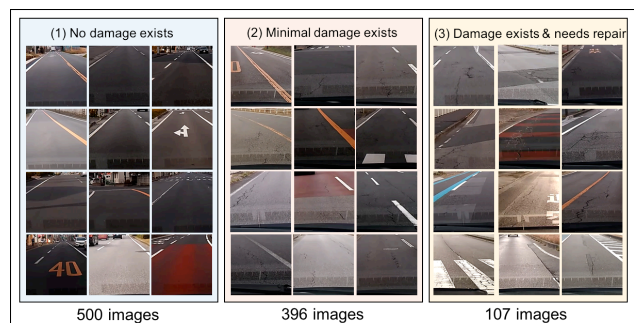


図-3 道路管理者による教師データサンプル

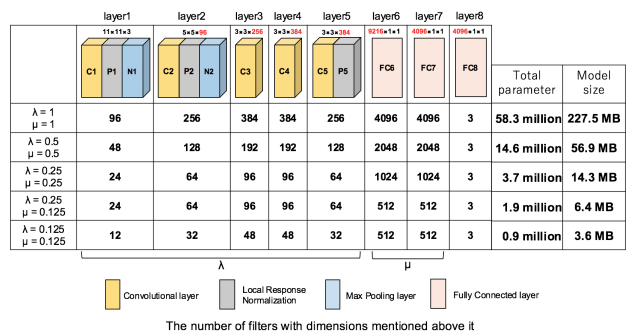


図-4 AlexNet の構造とカーネル数変化

る簡単な実験を行った。今回は、CNN のベースモデルとなっている AlexNet を基本的な構造として、そのカーネル数のみを変化させた。図-4 は、AlexNet の構造と、カーネル数を変化させた時の構造を示したものである。ここでは、便宜的に畳み込み層のフィルタ数を λ (96, 254, 384, 384, 256), 全結合層のニューロン数を μ (4046, 4096) で表している。また各モデルのパラメータ数合計と、本研究で用意している 1500 images を用いて学習を行ったときの 1000Epoch におけるモデルサイズを示している。

4. 実験と結果

本章で述べる結果は、第 3 章で説明した教師データ 1500 枚を学習データ 1200 枚（各クラス 400 枚ずつ）とテストデータ 300 枚（各クラス 100 枚ずつ）に分けて K=5 のクロスバリデーションを行った結果を示している。図-5 はカーネル数を変化させた時、GPU サーバー（本研究では Amazon EC2 g2.2xlarge を用いている）を用いて学習をさせる場合の処理時間を示している。 λ と μ の値が小さくなると、処理時間も同様に短くなるが見て取れる。しかし、 $\lambda = 0.25$, $\mu = 0.25$ と $\lambda = 0.25$, $\mu = 0.125$ の時を比較するとそれほど処理時間は変化しないことがわかる。これは、全結合層のパラメータ数を変化させても、行列のサイズが変化するだけで、行列計算の数は一回のまま変化しないためだと思われる。

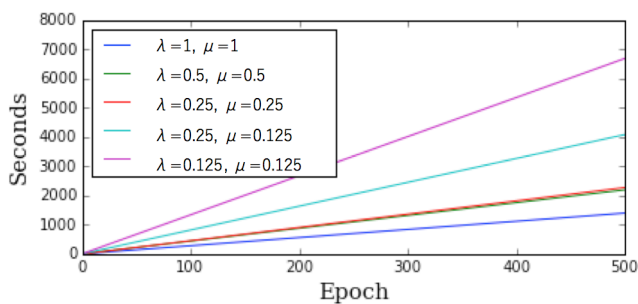


図-5 学習にかかる時間

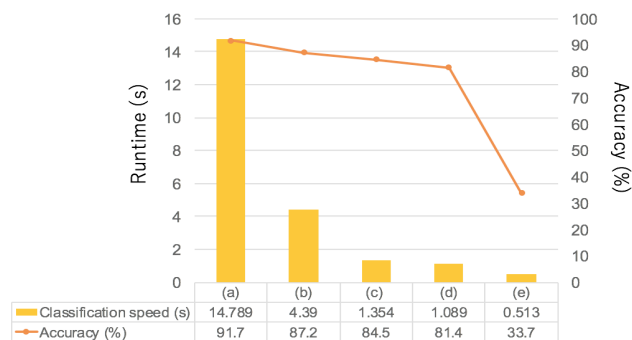


図-6 スマートフォン上での処理速度と判定精度

$\lambda=0.25, \mu=0.125$			Actual		Precision	
			Smooth	Damaged		
		No Need Repair		Need Repair		
Prediction	Smooth		392	4	1	98.7%
	Damaged	No Need Repair	95	385	55	72.0%
		Need Repair	13	111	444	78.2%
Recall			78.4%	77.0%	88.8%	Overall Accuracy 81.4%

図-7 $\lambda=0.25, \mu=0.125$ の時の混同行列

続いて、図-6 は学習済みモデルをスマートフォン（本研究では Motorola Moto G を用いている）上で操作した場合の処理速度と判定精度が示されている。 $\lambda=0.25, \mu=0.125$ を境にして、精度が大きく低下していることがわかる。また $\lambda=0.25, \mu=0.125$ の時の、混同行列が図-7 である。損傷ありの中で、修繕が必要か不要かという判断における誤判定の数が多くなっていることがわかる。

5. 考察

本研究の大きな成果は、サーバー側で深層学習させたモデルを Android 上から呼び出し、道路損傷を 80%以上の精度を担保しつつ、1.08 秒という高速で判定できるようにしたことである。そのために、本研究では深層学習を用いた画像処理でモデルを学習させた。本研究で開発したアプリケーションにより、専門的な知識を持たない

市民や行政職員が、道路路面の状況について判定することができる。ちばレポなど、市民協働システムに寄せられるレポートの自動分類や、専門家不足に直面している地方自治体などで大いに役立つだろう。

参考文献

- 1) 富山 和也, 川村 彰, 藤田 旬, 石田 樹 : 地方自治体の舗装維持管理実態を考慮した市街地道路の効果的な路面点検手法の開発, 土木学会論文集 F3 (土木情報学), Vol.69, No.2, I_54-I_62, 2013.
- 2) Offenhuber, Dietmar. : Infrastructure legibility a comparative analysis of open311 based citizen feedback systems, *Cambridge Journal of Regions*, 2014.
- 3) King, Stephen F., and Paul Brown : Fix my street or else: using the internet to voice local public service concerns, *Proceedings of the 1st international conference on Theory and practice of electronic governance*, 2007.
- 4) 311 City Services, the City of Chicago's Official Site, <http://www.cityofchicago.org/city/en/depts/311/supp_info/311hist.html>, (最終閲覧日 2016 年 6 月 30 日).
- 5) ちばレポ, <<http://chibarepo.force.com>> (最終閲覧日 2016 年 6 月 30 日).
- 6) Youngtae, Jo, and Seungki Ryu : Pothole detection system using a black-box camera, *Sensors* 15.11, pp.29316-29331, 2015.
- 7) Zou, Qin, et al. : CrackTree: Automatic crack detection from pavement images, *Pattern Recognition Letters* 33.3, pp.227-238, 2012.
- 8) 井原 正, 加藤 晃, 浅野 耕司 : デジカメ診断ソフトによる舗装路面のひび割れ測定 (性能評価特集号), 舗装 43. 7, pp.31-34, 2008.
- 9) 全 邦 釘, 橋本 和明, 片岡 望, 蔵本 直弥, 大賀 水田生 : ナイーブベイズ法によるアスファルト舗装撮影画像からのひび割れ自動検出手法, 土木学会論文集 E1 (舗装工学), Vol.70, No.3, I_1-I_8, 2014.
- 10) Yu, X., & Salari, E. : Pavement pothole detection and severity measurement using laser imaging, *Electro/Information Technology (EIT), IEEE International Conference*, pp.1-5, 2011.
- 11) Srinivas, S., Sarvadevabhatla, R. K., Mopuri, K. R., Prabhu, N., Kruthiventi, S. S., & Babu, R. V. : A taxonomy of deep convolutional neural nets for computer vision, *arXiv preprint arXiv:1601.06615*, 2016.
- 12) Iandola, F. N., Moskewicz, M. W., Ashraf, K., Han, S., Dally, W. J., & Keutzer, K. : SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and < 1MB model size, *arXiv preprint arXiv:1602.0736*, 2016.
- 13) Krizhevsky, A., Sutskever, I., & Hinton, G. E. : Imagenet classification with deep convolutional neural networks, *Advances in Neural Information Processing Systems*, pp.1097-1105, 2012.
- 14) Chatfield, K., Simonyan, K., Vedaldi, A., & Zisserman, A. : Return of the devil in the details: delving deep into convolutional nets, 2014.