

MATHEMATICAL MODEL FOR THE PROBLEM OF CLASSIFICATION OF PAVEMENT DEFECTS

Pham Thi Lien^{1*}, Tran Thi Tuyet¹, Nguyen Quang Hiep¹, Nguyen Thi Dung¹, Kieu Tuan Dung²,
 Nguyen Thi Phuong Dung²

¹TNU - University of Information and Communication Technology, ²Thuy Loi University

ARTICLE INFO		ABSTRACT
Received:	19/3/2022	The road surface defect detection and classification system based on machine learning algorithms is already very advanced and is increasingly proving its outstanding advantages. In this paper, some image segmentation algorithms used in practice are presented, compared and evaluated. In this study, we present the convolution neural network—VGG16 structure to classify pavement defects, with a graph-based method to optimize the image segmentation on the pavement defect image. This proposed method is intended to overcome limitations caused by objective factors, such as high sensitivity to data of certain types of light and noise dependence, such as defect data of the road surface. Three different datasets were collected from the Center for Telecommunication and Multimedia, INESC TEC - Portugal (1200 images), Irkutsk city - Russian Federation (800 images) and Thai Nguyen city - Vietnam (550 images). The classification results based on the VGG-16 methods of the datasets in turn are good because the curves have a state closer to 1 than 0.5.
Revised:	24/6/2022	
Published:	04/7/2022	

KEYWORDS

Machine learning
 Deep learning
 Pavement defect classification
 VGG16
 Features extraction
 Convolutional Neural Network

MÔ HÌNH TOÁN HỌC CHO BÀI TOÁN PHÂN LOẠI CÁC KHUYẾT TẬT MẶT ĐƯỜNG

Phạm Thị Liên^{1*}, Trần Thị Tuyết¹, Nguyễn Quang Hiệp¹, Nguyễn Thị Dung¹, Kiều Tuấn Dũng²,
 Nguyễn Thị Phương Dung²

¹Trường Đại học Công nghệ thông tin và Truyền thông – ĐH Thái Nguyên, ²Trường Đại học Thủy lợi

THÔNG TIN BÀI BÁO		TÓM TẮT
Ngày nhận bài:	19/3/2022	Hệ thống phân loại và phát hiện khuyết tật mặt đường dựa trên thuật toán máy học vốn đã rất tiên tiến và ngày càng chứng tỏ những ưu điểm vượt trội của mình. Trong bài báo này, một số thuật toán phân đoạn hình ảnh được sử dụng trong thực tế được trình bày, so sánh và đánh giá. Trong nghiên cứu này, chúng tôi trình bày mạng nơ-ron tích chập - cấu trúc VGG16 để phân loại các khuyết tật mặt đường, với phương pháp dựa trên đồ thị để tối ưu hóa việc phân đoạn hình ảnh trên hình ảnh khuyết tật mặt đường. Phương pháp đề xuất này nhằm khắc phục những hạn chế do các yếu tố khách quan gây ra, chẳng hạn như độ nhạy cao đối với dữ liệu của một số loại phụ thuộc vào ánh sáng và tiếng ồn, chẳng hạn như dữ liệu khuyết tật của mặt đường. Ba bộ dữ liệu khác nhau được thu thập từ Trung tâm Viễn thông và Đa phương tiện, INESC TEC - Bồ Đào Nha (1200 hình ảnh), thành phố Irkutsk - Liên bang Nga (800 hình ảnh) và thành phố Thái Nguyên - Việt Nam (550 hình ảnh). Kết quả phân lớp dựa trên các phương pháp VGG-16 của lần lượt các bộ dữ liệu cho thấy kết quả phân loại là tốt vì các đường cong có trạng thái gần giá trị 1 hơn giá trị 0,5.
Ngày hoàn thiện:	24/6/2022	
Ngày đăng:	04/7/2022	

TỪ KHÓA

Học máy
 Học sâu
 Phân loại khuyết tật mặt đường
 VGG16
 Trích chọn đặc trưng
 Mạng nơ-ron tích chập

DOI: <https://doi.org/10.34238/tnu-jst.5722>

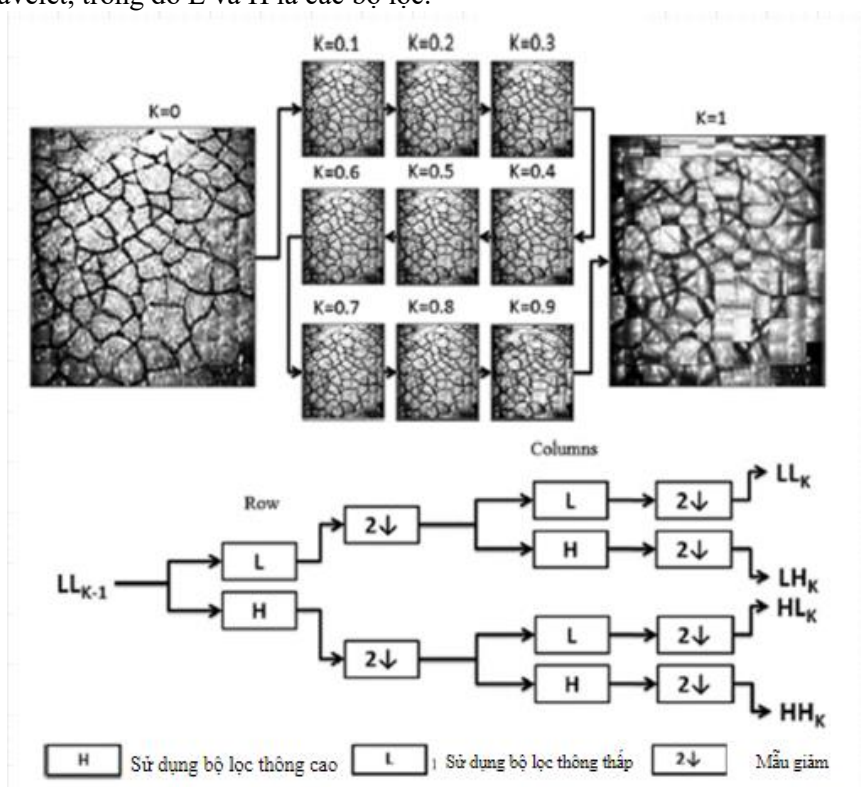
* Corresponding author. Email: ptlien@ictu.edu.vn

1. Giới thiệu

Trong việc phát triển hệ thống phân loại các khuyết tật mặt đường, các hình ảnh thu được từ những nơi được quan sát sẽ được xử lý để tạo thành dữ liệu đầu vào là các đặc trưng của đối tượng cho hệ thống phân tích và xử lý. Tuy nhiên, các hệ thống vẫn gặp phải một số tồn tại như hiệu quả của việc quan sát luôn phụ thuộc vào điều kiện môi trường quan sát (ánh sáng, tính chất đường giao thông...), kiểu khuyết tật mặt đường (các kiểu khuyết tật nứt, ổ gà, lún...), phương pháp tiếp cận để giải quyết vấn đề về lựa chọn thuật toán, lựa chọn các đặc trưng, kết hợp các phương pháp xử lý. Vì vậy, các hệ thống này vẫn còn đang được nhiều nhà khoa học, trung tâm nghiên cứu trên thế giới quan tâm phát triển.

Các khuyết tật mặt đường được coi là các đối tượng trong ảnh. Trích chọn đặc trưng của đối tượng là bước quan trọng trong bài toán phân tích hình ảnh, video, công việc này khái quát lại đó là việc sử dụng các kỹ thuật xử lý ảnh để phân tích, biến đổi các thông tin trong ảnh thành các con số có ý nghĩa phục vụ cho việc phát hiện và phân loại các khuyết tật. Các phương pháp sử dụng biến đổi wavelet [1], phương pháp lấy mẫu [2], phương pháp đo đẳng vị (anisotropy measure) [3], phương pháp lý thuyết Fuzzy [4], phương pháp trích chọn đặc trưng dựa vào kết cấu [5] hoạt động hiệu quả giải quyết các vấn đề trong trích chọn các đặc trưng khuyết tật mặt đường trong ảnh và video.

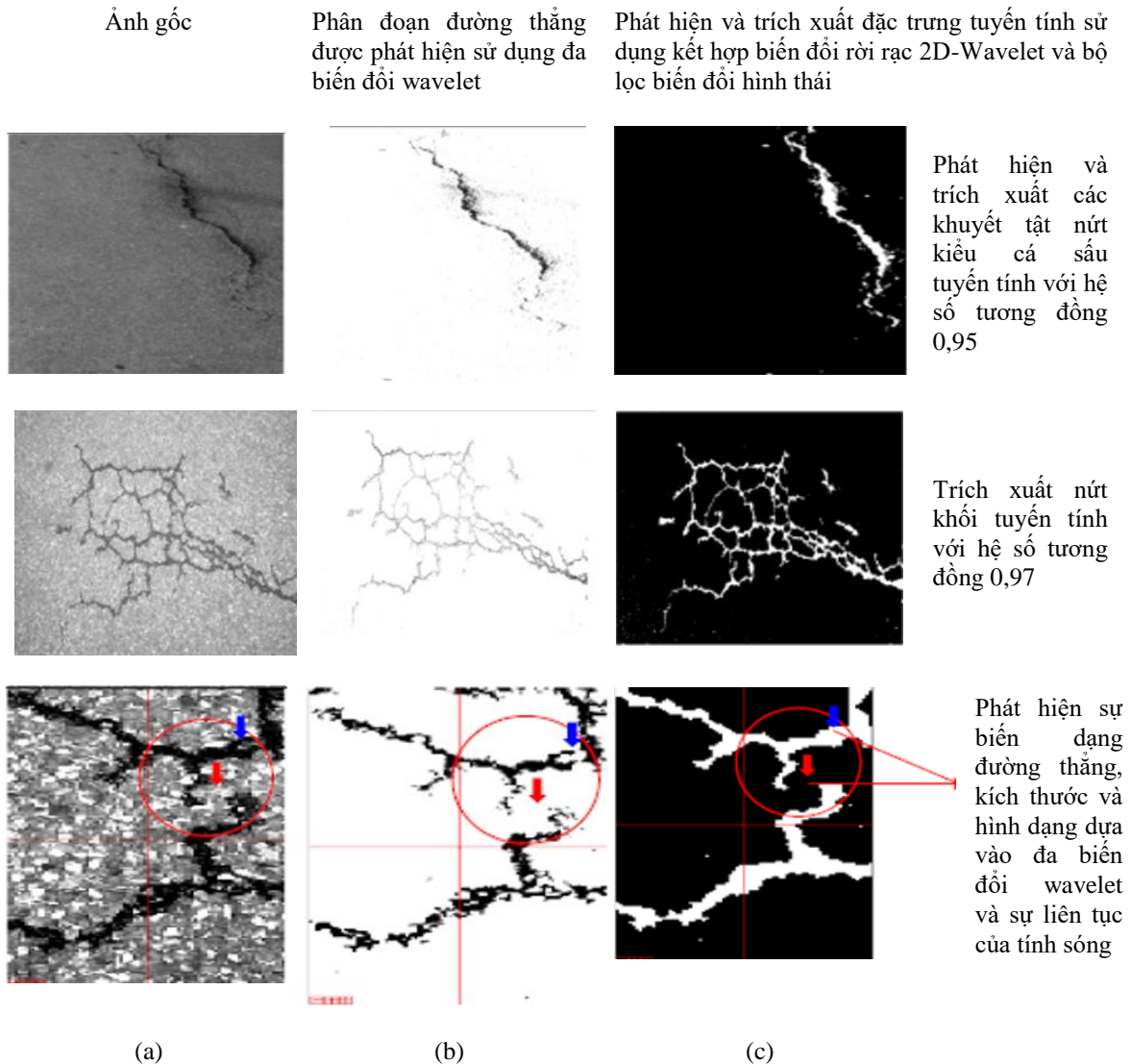
Trong [6] đã trình bày phương pháp trích chọn đặc trưng sử dụng biến đổi Wavelet. Sử dụng biến đổi Wavelet cung cấp một công cụ tốt cho quá trình phân tích đa phân giải (Multi-resolution) đối với các khuyết tật nứt trên mặt đường. Hình 1 mô tả quá trình phân rã ảnh dựa vào biến đổi Wavelet, trong đó L và H là các bộ lọc.



Hình 1. Mô tả phương pháp trích chọn đặc trưng khuyết tật mặt đường bằng Wavelet – Random transform [6]

Phân tích hình dạng rất quan trọng trong việc phân biệt các điểm ảnh khuyết tật với điểm ảnh lân cận. Phương pháp trích chọn đặc trưng tuyến tính được sử dụng trong nghiên cứu [7]. Hình 2 mô tả việc sử dụng kết hợp biến đổi rời rạc 2D-Wavelet (discrete wavelet transform - DWT) và

bộ lọc biến đổi hình thái (successive morphologic transformation filtering – SMF). Sử dụng bộ lọc thích ứng cho trích chọn đặc trưng hình thái và sự liên tục của khuyết tật nứt trên mặt đường.



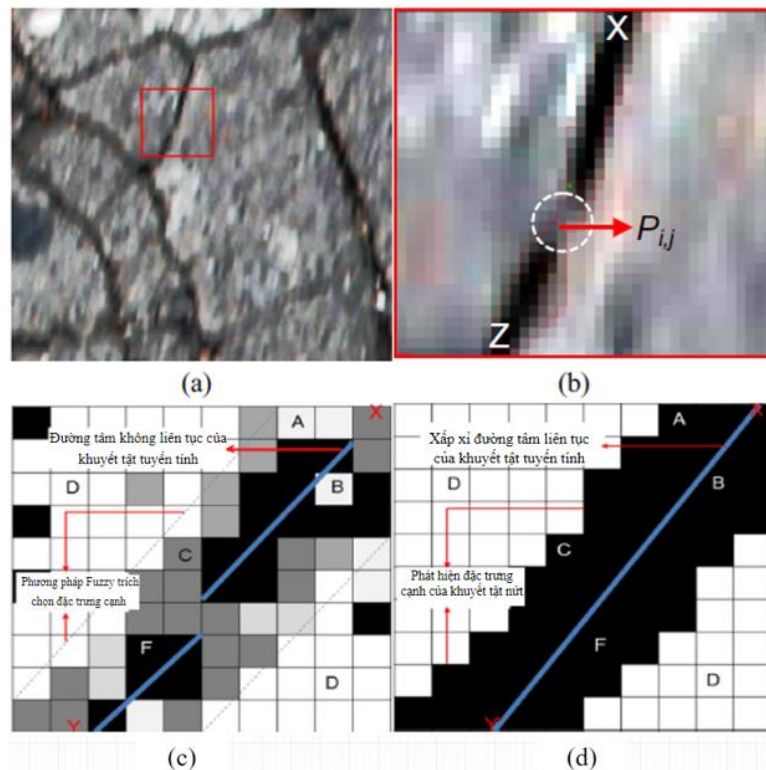
Hình 2. Minh họa phương pháp trích chọn đặc trưng dựa vào DWT-SMF [7]

Phương pháp fuzzy được áp dụng trong [8] để trích chọn các đặc trưng khuyết tật nứt trên mặt đường trong ảnh và video. Hình (3a) biểu diễn hình ảnh ban đầu, hình (3b) biểu diễn kết quả phân đoạn ảnh vùng ảnh khuyết tật, hình (3c, d) minh họa kết quả trích chọn đặc trưng cạnh cho khuyết tật bằng phương pháp Fuzzy.

Các kỹ thuật xử lý hình ảnh để loại bỏ nền ảnh, những kết quả thu được vẫn tồn tại những thành phần không phải là các khuyết tật mặt đường. Các phương pháp máy học thường được dùng để tách các khuyết tật mặt đường ra khỏi các thành phần không phải là khuyết tật. Các tính năng khác nhau đã được sử dụng trong các nghiên cứu. Các tính năng về độ dày bề mặt, diện tích và độ dẻo là các tính năng được sử dụng rộng rãi để xác định các vùng ảnh quan tâm [9] – [12].

Tuy nhiên, trong thực tế việc lựa chọn các đặc trưng để xử lý đều trực quan và dựa trên quan sát của con người. Để khắc phục vấn đề này thì cần phải kết hợp sử dụng các đặc trưng để mô tả

được tối ưu các khuyết tật mặt đường. Các thông tin có thể được suy ra từ mỗi đặc trưng và tầm quan trọng của các giá trị đặc trưng để đánh giá các khuyết tật mặt đường.



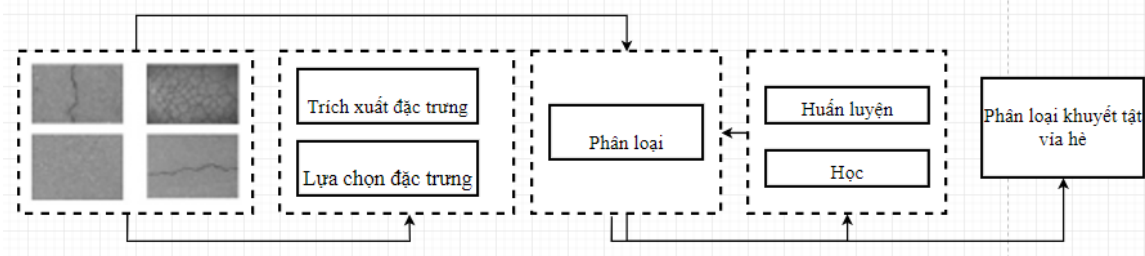
Hình 3. Hình ảnh minh họa kết quả áp dụng phương pháp fuzzy trích chọn đặc trưng diện tích khuyết tật mặt đường [8]

2. Các phương pháp máy học và học sâu được ứng dụng để phát hiện và phân lớp các khuyết tật mặt đường

Các phương pháp máy học, học sâu được ứng dụng và phát triển mạnh mẽ để giải quyết nhiệm vụ phát hiện và phân loại các khuyết tật mặt đường trên ảnh tĩnh và video. Các phương pháp máy học, học sâu được đánh giá dựa trên các tiêu chí: hiệu suất dự đoán, thời gian thực hiện, tính ổn định của hệ thống đối với dữ liệu mất cân đối (dữ liệu ảnh không có khuyết tật và ảnh có khuyết tật mặt đường), tính ổn định của hệ thống đối với các yếu tố bên ngoài (nhiều, ánh sáng...), khả năng diễn giải các kết quả và tính rõ ràng của thủ tục.

Nhiều nhà nghiên cứu đã chứng minh những lợi thế của học sâu để xây dựng các ứng dụng. Một ví dụ về điều đó bao gồm việc chuyển đổi các cảm biến giao thông trên đường thành một tác nhân thông minh tự động phát hiện tai nạn và dự đoán các điều kiện của xe trong tương lai [13]. Ngoài ra, có nhiều phương pháp Deep Learning được sử dụng trong vận chuyển như ANN. ANN có thể được sử dụng để lập kế hoạch đường bộ [14], giao thông công cộng [15], phát hiện sự cố giao thông [16] và dự đoán các điều kiện của xe [17]. Nó được phân loại thành các phương pháp học tập có giám sát và không giám sát. Các phương pháp được giám sát bao gồm Máy Vector hỗ trợ (SVM), Mạng nơ-ron xác suất (PNN), Mạng cơ sở xuyên tâm (RBN), Hàng xóm K-Gần nhất và Cây quyết định,... trong khi các mạng nơ-ron không được giám sát bao gồm phân tích theo tầng và phân tầng. Để phát hiện các khuyết tật mặt đường thường dựa trên các phương pháp, chẳng hạn như ngưỡng cường độ, phát hiện cạnh, lý thuyết đồ thị, phân tích kết cấu, thuật toán học máy và phương pháp dựa trên mạng thần kinh. Các thuật toán ngưỡng được dựa trên giả định rằng các vết nứt được biểu thị bằng cực tiểu cường độ cục bộ. Các tác giả ở [18] đã khám phá và đánh giá phương pháp học chuyển đổi sâu, viz., sử dụng các mô hình học sâu được đào tạo trên

dữ liệu hình ảnh dữ liệu lớn của Image (ImageNet) và chuyển giao khả năng học hỏi của họ để phát hiện vết nứt mặt đường tự động từ bề mặt số hóa hình ảnh thu được từ cơ sở dữ liệu FHWA/LTPP. Bản chất công việc trong [19] dựa trên sự khác biệt ở mức xám giữa các pixel nứt và pixel nền, có giới hạn trực tiếp cho các vết nứt đó mà không thay đổi cường độ rõ ràng với các khu vực xung quanh. Thông tin sâu sắc như một nhân vật bẩm sinh khác liên quan đến các vết nứt sẽ thuận lợi trong việc phát hiện khi các đặc điểm mức độ màu xám không đủ nổi bật. Các kỹ thuật phát hiện cạnh bao gồm việc sử dụng các bộ lọc Canny, bộ lọc phát hiện cạnh Sobel và các bộ lọc hình thái khác [20], [21]. Với sự phát triển của trí tuệ nhân tạo trong những năm gần đây, một số hệ thống phân loại được xây dựng dựa trên các thuật toán học máy được đưa vào lĩnh vực phát hiện khuyết tật mặt đường tự động như máy vector hỗ trợ [22], Rừng ngẫu nhiên [23]. Với sự ra đời của công nghệ học tập sâu, Mạng nơ-ron tích chập đã bắt đầu thống trị lĩnh vực phát hiện và nhận dạng đối tượng.



Hình 4. Cấu trúc chung của một hệ thống phát hiện và phân loại các khuyết tật mặt đường

Một số phương pháp, thuật toán máy học đã được ứng dụng và phát triển thành công trong tự động phát hiện và phân loại các khuyết tật mặt đường và có cấu trúc chung được minh họa ở hình 4. Các thuật toán đó được thống kê ở bảng 1.

Bảng 1. Phân tích các thuật toán máy học (****: Tốt nhất, *: Tồi nhất, trên cùng bộ dữ liệu test)

	Decision Trees	Neural Networks	Naïve Bayes	SVM
Loại máy học	Unsupervised learning	Supervised learning	Supervised learning	Supervised learning
Độ chính xác	**	***	*	****
Tốc độ học đối với số lượng thuộc tính	***	*	****	*
Tốc độ phân lớp	****	****	****	****
Dung sai với tiếng ồn	**	**	***	**
Xử lý các thuộc tính: rời rạc/ nhị phân/ liên tục	****	(không rời rạc)	(không liên tục)	(không rời rạc)
Xử lý vấn đề tràn bộ nhớ	**	*	***	**
Cải thiện quá trình học	**	***	****	**
Xử lý tham số mô hình	***	*	****	*

Trong bài báo [24], tác giả đã xây dựng hệ thống gồm 3 phần chính: tiền xử lý và nâng cao chất lượng ảnh số, trích chọn các đặc trưng hình học trong từng vùng ảnh, phát hiện và nhận dạng các khuyết tật dựa vào biến đổi wavelet cho ảnh có độ phân giải thấp, sử dụng bộ lọc trung vị để lấy ngưỡng, sử dụng bộ lọc hình thái đối với đặc trưng hình dạng, sử dụng biến đổi Random để phân loại các kiểu khuyết tật trên đường.

Nghiên cứu [25] sử dụng phương pháp máy học để phát hiện và phân loại các khuyết tật ổ gà trong video. Ảnh được phân đoạn thành các vùng, trích chọn các đặc trưng hình dạng và kết cấu trong phát hiện khuyết tật ổ gà. Các đặc trưng hình học được ước lượng thông qua sự phức hồi hình thái và hồi quy logistic.

Nejad và Zakeri [26] đã trình bày về một hệ thống tự động phát hiện và phân loại khuyết tật nứt trên mặt đường dựa trên sự kết hợp biến đổi Wavelet và mạng thần kinh động. Hệ thống này gồm 3 phần chính: phần thứ nhất là xử lý và lưu trữ ảnh đầu vào, phần thứ 2 là trích chọn đặc trưng dựa vào biến đổi Fourier và cân bằng Histogram, phân loại các khuyết tật mặt đường sử dụng mạng thần kinh, phần thứ 3 là hiển thị kết quả phân lớp.

Các chương trình tự động phân loại khuyết tật mặt đường đã được phát triển trong 20 năm qua. Ban đầu các chương trình làm việc trên dữ liệu ảnh được thu thập và lưu trữ [27], [28]. Nhưng các chương trình này làm việc hiệu quả thấp, độ phân giải dữ liệu không cao, khó khăn trong làm việc máy tính. Từ đó, dẫn đến tăng chi phí và tính phức tạp hệ thống. Trong những năm gần đây, các hệ thống kỹ thuật số đã xuất hiện, trở thành phương pháp thông dụng. Hầu hết họ sử dụng kỹ thuật xử lý hình ảnh trên video. Chất lượng video tốt và đủ ánh sáng để đáp ứng khả năng sử dụng các thuật toán máy học.

Có nhiều hệ thống tự động phát hiện các vết nứt dựa trên các đặc điểm giống nhau sử dụng máy quét dòng. Hệ thống này phát hiện các vết nứt có độ rộng 1 mm. Việc phân loại này chủ yếu dựa vào độ rộng của các vết khuyết tật. Vào năm 1999, Australian Commonwealth Scientific and Industrial Research Organization (CSIRO) là tổ chức đầu tiên phát triển hệ thống tự động phát hiện các khuyết tật nứt – RoadCrack. Hệ thống này có thể phát hiện các vết nứt trên đường cao tốc với độ rộng 1 mm. Các thiết bị thu nhận dữ liệu và xử lý hình ảnh, phân tích được gắn trên một chiếc xe. Nền tảng ARAN [29] được áp dụng rộng rãi ở Mỹ để tự động phân tích đường. Hệ thống này sử dụng máy quét với cấu hình gồm 2 hệ thống. Độ phân giải 2 mm với ánh sáng không ổn định và độ phân giải 1 mm với ánh sáng hồng ngoại để thu thập dữ liệu. Những hình ảnh này được đánh giá thông qua phần mềm tự động phát hiện vết nứt – WiseCrax, bên cạnh đó chương trình tự động phát hiện các khuyết tật mặt đường WayLink's Digital Highway Data Vehicle (DHDV) cũng rất thông dụng. Tại các nước châu Âu, hệ thống PAVUE [30] được triển khai ở các nước Hà Lan và Phần Lan. Hệ thống được trang bị 2 máy quay video để tiến hành thu thập dữ liệu và cấu hình máy tính đã tích hợp các thuật toán máy học để làm việc.

Hình 5 mô tả thiết bị tổ hợp công nghệ với tốc độ cao để thu thập và lưu trữ dữ liệu. Thiết bị bao gồm máy ảnh kỹ thuật số, công nghệ siêu âm và laser để chụp đường ngang của đường, thiết bị laser với tần số cao để thu thập kết cấu mặt đường.



Hình 5. Thiết bị thu nhận hình ảnh mặt đường [30]



Hình 6. Hệ thống LRIS [16]

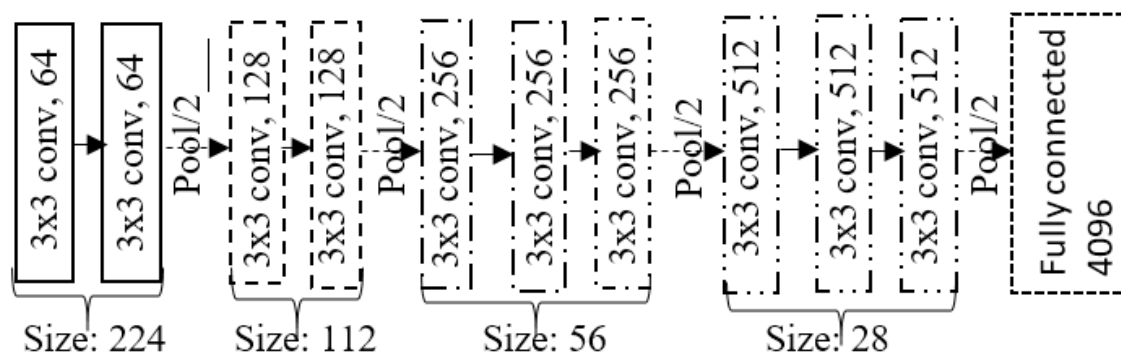
Hình 6 mô tả hệ thống LRIS được tạo bởi các hàng camera song song có độ phân giải cao và nguồn laser cực lớn. Các camera này được gắn vào đằng sau các ô tô chuyên dụng để thu thập dữ liệu và phân loại các khuyết tật mặt đường được thực hiện phân tích offline bởi một máy phát hiện nứt tự động [31]. Ưu điểm của hệ thống thu được hình ảnh nhanh chóng, an toàn và chất lượng hơn.

Hệ thống GIE LaserVISION [32] là một ví dụ điển hình sử dụng công nghệ laser để tự động phát hiện và phân loại các khuyết tật mặt đường. Hệ thống này sử dụng 4 cảm biến laser và cung cấp khả năng tính toán 3D để cải thiện việc tính toán các khuyết tật trong không gian 3D. Tuy nhiên, độ phân giải của hệ thống thấp và chỉ làm việc với vết nứt ngang. Vấn đề chính của việc xử lý dữ liệu video chính là ánh sáng, góc của thiết bị, vật xung quanh... ảnh hưởng kết quả phát hiện và phân loại các khuyết tật. Để giảm bớt sự ảnh hưởng của các yếu tố này hệ thống StereoVision [33], [34] đã sử dụng công nghệ 3D để nâng cao hiệu quả làm việc của hệ thống.

3. Mô hình kiến trúc CNN - VGG16 được đề xuất trong hệ thống

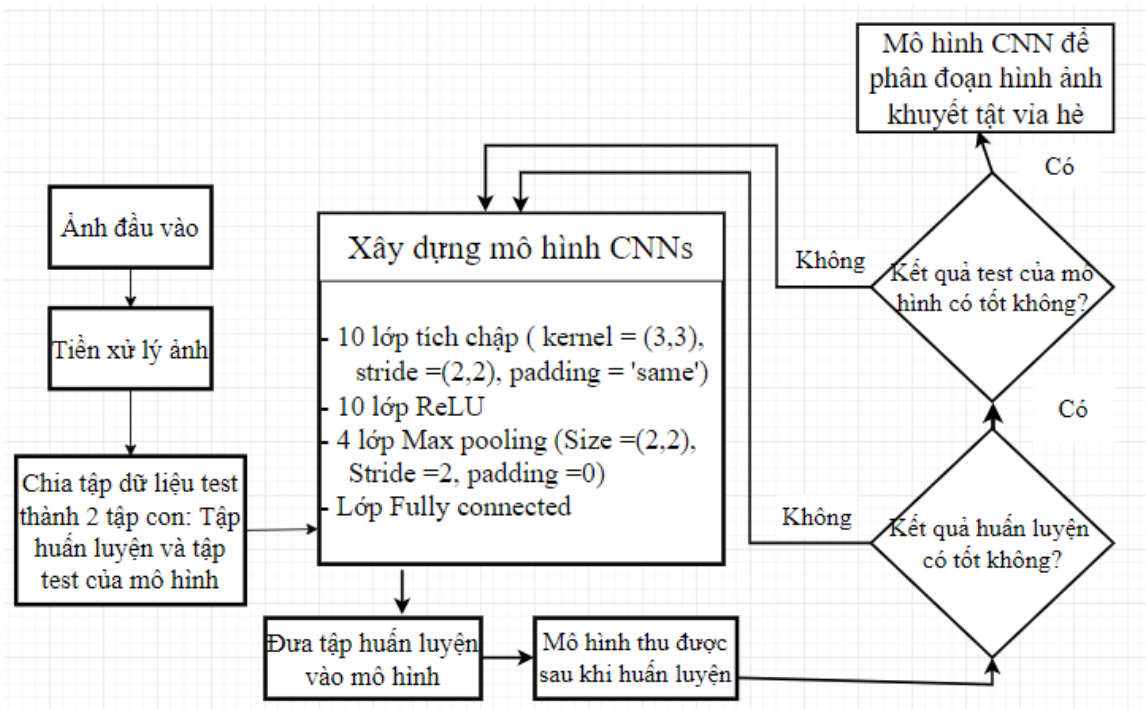
Với ưu thế được thiết kế một cách đặc biệt để phân tích hình ảnh trực quan, VGG 16 (hình 7) không đòi hỏi quá trình xử lý một cách tối ưu như các thuật toán máy học, thị giác máy tính. Kiến trúc của VGG 16 gồm 3 lớp chính: convolutional layers, pooling layers, và fully connected layer, trong đó lớp chập chính là lớp chính. Trong số đó, các hoạt động tích chập và gộp có thể được lặp đi lặp lại nhiều lần để tối ưu hóa hiệu ứng lọc.

CNN thường sử dụng các bộ lọc một cách song song từ 32 đến 512 bộ lọc cho một dữ liệu đầu vào. Một yêu cầu bắt buộc của bộ lọc trong CNN đó là nó phải cùng kênh (tức là cùng độ sâu) với đầu vào. Và sẽ tốt hơn nếu như ta cung cấp cho CNN các giá trị cụ thể cho từng kênh đầu vào. Cuối cùng, các kênh này sẽ được tổng hợp lại để tạo thành một kênh duy nhất (mỗi bộ lọc đều tạo ra kết quả cuối cùng là 1 một bản đồ tính năng 2D).



Hình 7. Kiến trúc của VGG 16 được đề xuất trong hệ thống (kết quả nghiên cứu của đề tài)

Mỗi lớp tích chập trong CNN tạo ra các bản đồ đối tượng từ hình ảnh đầu vào, biểu thị cường độ của từng đối tượng. Nghiên cứu của chúng tôi nhằm phát hiện và phân loại đối tượng bằng cách so sánh bản đồ đặc trưng của nó với bản đồ đặc trưng của ảnh đầu vào trước đó. Để đạt được điều này, các vectơ đặc trưng sẽ được trích xuất từ các bản đồ đặc trưng được tạo từ mỗi lớp tích chập. Mô hình CNN (hình 8) được đào tạo với tỉ lệ học là 0,01 bằng cách sử dụng chức năng kích hoạt softmax trong lớp tổng hợp. Hàm kích hoạt ReLU được sử dụng tất cả các lớp ẩn của lớp chập, số lần học nhiều nhất của mạng là 100 lần. Kiến trúc mạng cơ bản bao gồm: Lớp đầu vào [Lớp Convolution - Lớp MaxPooling - Lớp Kích hoạt] => Lớp đầu ra. Khi huấn luyện mạng, tham số Batch size = 32 được chọn để chọn kích thước mẫu cần học và điều chỉnh các thông số trong các lớp mạng để có độ chính xác tốt nhất. Số lần huấn luyện mạng tối đa là 100 lần. Mô hình có kết quả phân loại tốt nhất trên tập huấn luyện sẽ được lựa chọn và thử nghiệm trên các tập kiểm tra để đánh giá. Nó lưu trữ các bản đồ đặc trưng của các hình ảnh được CNN phân loại các hình ảnh mới nhập bằng cách so sánh chúng với các bản đồ đặc trưng được lưu trong bộ nhớ cache. Bởi vì mô hình VGG-16 được đào tạo trước được tinh chỉnh, hình ảnh dữ liệu đầu vào đã được thay đổi kích thước thành 224x224 pixel thông qua nội suy song tuyến. Kích thước bản đồ đối tượng thu được từ lớp tích tụ đầu tiên là 224x224x6. Bản đồ đối tượng đầu ra ($OF_{k,l,n}$) cho một tích chập chuẩn, dựa trên sai (stride) bằng 1 và khoảng đệm (padding), thường thu được bằng bản đồ của đối tượng đầu vào ($IF_{k+1,j,m}$) và nhân chập ($K_{i,j,m,n}$).



Hình 8. Quá trình huấn luyện và kiểm tra của mô hình CNN trong hệ thống phân loại các khuyết tật mặt đường (kết quả nghiên cứu của đề tài)

$$OF_{k,l,n} = \sum_{i,j,m} K_{i,j,m,n} IF_{k+i-1,l+j-1,m} \quad (1)$$

Dựa trên phương trình (1) chi phí tính toán phụ thuộc vào số kênh đầu ra, kích thước bản đồ tính năng đầu vào, kích thước hạt nhân và số lượng kênh đầu vào. Như đã thấy trong phương trình 2 (x, y là đầu vào và đầu ra), mọi lớp tích chập được kết nối với một lớp đơn vị tuyến tính chỉnh lưu (ReLU). ReLU cũng có thể loại bỏ vấn đề gradient biến mất lan truyền ngược trong quá trình đào tạo và nó có thể giảm thời gian đào tạo.

$$y = \max(0, x) \quad (2)$$

Mô hình VGG 16 bao gồm 10 lớp phức hợp, 4 lớp gộp và một lớp được kết nối đầy đủ. Sau lớp tích chập thứ hai, thứ tư, thứ bảy, thứ 10 với ReLU, một lớp tổng hợp tối đa tiếp theo. Lớp tổng hợp tối đa sử dụng giá trị lớn nhất trong bộ lọc với kích thước được chỉ định và tiến hành lấy mẫu con. Sau 10 lớp phức hợp, 10 lớp ReLU và 4 lớp tổng hợp tối đa, kích thước bản đồ tính năng cuối cùng $28 \times 28 \times 512$.

4. Kết quả thực nghiệm

Kết quả phân loại các khuyết tật mặt đường dựa trên kiến trúc VGG-16 đối với 3 bộ dữ liệu (1), (2) và (3). Các bảng 2, 3, 4 mô tả các ma trận confusion. Bảng 5 mô tả các kết quả phân lớp các khuyết tật mặt đường dựa trên các phương pháp VGG-16, RF, SVM của lần lượt của bộ dữ liệu 1, 2, 3. Đây được gọi là ma trận nhập nhằng dùng để dự đoán và điều chỉnh cho hệ thống ở lớp Pooling.

Ví dụ: Trong bảng 2 giống hàng BC với cột BC là 173 có nghĩa là hệ thống phân loại chính xác 173 bức ảnh là BC, Phân loại sai thành LC là 7 ảnh, và C là 0 ảnh. Từ đó tính là độ chính xác là 96,11%.

Bảng 2. Kết quả phân loại các khuyết tật mặt đường của bộ dữ liệu 1

True class 97,7 (%)	Assigned class			Độ chính xác (%)
	BC	LC	P	
BC	173	7	0	96,11
LC	3	193	4	96,50
P	1	2	167	98,24

Bảng 3. Kết quả phân loại các khuyết tật mặt đường của bộ dữ liệu 2

True class 97,62 (%)	Assigned class			Độ chính xác (%)
	BC	LC	P	
BC	195	2	3	97,50
LC	3	262	5	97,03
P	5	1	325	98,18

Bảng 4. Kết quả phân loại các khuyết tật mặt đường của bộ dữ liệu 3

True class 98,50 (%)	Assigned class			Độ chính xác (%)
	BC	LC	P	
BC	445	2	3	98,89
LC	2	317	1	99,09
P	7	3	420	97,67

Bảng 5. Kết quả phân lớp dựa trên các phương pháp VGG-16, RF, SVM của lần lượt của bộ dữ liệu 1, 2, 3

Các tiêu chí	Bộ dữ liệu	RF	SVM	VGG-16
Độ chính xác (%)	1	91,77	88,52	98,50
	2	91,05	86,76	97,62
	3	90,89	86,35	97,07
Thời gian Phân lớp (Giây)	1	1775	1202	1800
	2	1287	815	708
	3	1153	870	513
MSE	1	0,305	0,670	0,200
	2	0,310	0,566	0,311
	3	0,230	0,412	0,296

5. Kết luận

Phân loại khuyết tật đường tự động là một ứng dụng thực tế của phương pháp học máy, nơi Deep Learning thể hiện thế mạnh của mình trên dữ liệu không yêu cầu xử lý trước hình ảnh quá nhiều. Trong bài báo này, chúng tôi đề xuất cách tiếp cận Học sâu dựa trên việc tạo dữ liệu đào tạo tổng hợp để phân loại khuyết tật đường trong ảnh đường. Thuật toán tạo dữ liệu tổng hợp cho phép chúng tôi dễ dàng có được tập dữ liệu huấn luyện về các khuyết tật trên đường. Kết quả thử nghiệm trên 6 bộ dữ liệu khác nhau với số lượng lớn đối tượng cho thấy hiệu quả và độ chính xác cao. Điều này xác nhận rằng CNN có thể được sử dụng trong học máy để giải quyết vấn đề phát hiện đối tượng trong các hình ảnh có độ nhiễu cao và nhạy cảm với ánh sáng. Đóng góp chính của chúng tôi là đặt bộ nhớ cache trong lớp phức hợp của CNN để đảm bảo rằng nó hoạt động tốt và cải thiện tốc độ phân đoạn hình ảnh trong hệ thống phân loại đối tượng. Phương pháp này được sử dụng để tạo bản đồ đối tượng. Trong hình ảnh thu được trực tiếp, các yếu tố khách quan (ánh sáng, nhiễu,...) có ảnh hưởng mạnh đến việc trích xuất đặc điểm và dẫn đến nhầm lẫn pixel cung cấp thông tin không chính xác. Điều này là do thực tế là các thuật toán thị giác máy tính hoặc kỹ thuật xử lý hình ảnh cũng phụ thuộc vào dữ liệu. Sử dụng tích chập trong VGG16 để phân đoạn các khuyết tật trên đường dựa trên việc tăng độ sâu của mạng nơ-ron tích chập (số lớp ẩn) để tăng cường tiếp nhận có thể dẫn đến cải thiện hiệu suất mô hình.

Lời cảm ơn

Bài báo là sản phẩm của đề tài NCKH có mã số ĐH2020-TN07-01, được tài trợ bởi kinh phí của trường Đại học Công nghệ Thông tin và Truyền thông – ĐH Thái Nguyên. Nhóm tác giả xin chân thành cảm ơn sự tài trợ của quý Nhà trường.

TÀI LIỆU THAM KHẢO/ REFERENCES

- [1] S. Chambon, “Introduction of a wavelet transform based on 2D matched filter in a Markov random field for fine structure extraction: Application on road crack detection,” *Proceedings Image Processing: Machine Vision Applications II, Proc. SPIE*, vol. 7251, 2009, doi: 10.1117/12.805437.
- [2] C. Ma, C. Zaho, and Y. Hou, “Pavement distress detection based on nonsubsampling contourlet transform,” *Proc. IEEE Int. CSSE*, 2008, pp. 28-31.
- [3] T. Nguyen, M. Avila, and B. Stephane, “Automatic detection and classification of defects on road pavement using anisotropy measure,” *Proc. 17th EUSIPCO*, 2009, pp. 617-621.
- [4] H. D. Cheng, “Novel approach to pavement cracking detection based on fuzzy set theory,” *Journal of Computing in Civil Engineering*, vol. 13, no. 4, pp. 270–280, 1999.
- [5] F. Roli, “Measure of texture anisotropy for crack detection on textured surfaces,” *Electronics Letters*, vol. 32, pp. 1274–1279, 1996.
- [6] J. Zhou, P. S. Huang, and F. P. Chiang, “Wavelet-based pavement distress detection and evaluation,” *Optical Engineering*, vol. 45, 2006, doi: 10.1117/1.2172917.
- [7] O. Yashon and M. Hahn, “Wavelet-morphology based detection of incipient linear cracks in asphalt pavements from RGB camera imagery and classification using circular Radon transform,” *Advanced Engineering Informatics*, vol. 30, pp. 481–499, 2016.
- [8] A. Downey, H. Koutsopoulos, and I. El Sanhoury, “Analysis of Segmentation Algorithms for Pavement Distress Images,” *Journal of Transportation Engineering*, vol. 119, pp. 869 – 888, 1993.
- [9] Y. Hu and C. Zhao, “A local binary pattern based method for pavement crack detection,” *Journal of Pattern Recognition Research*, vol. 1, pp. 140–147, 2013.
- [10] M. S. Kaseko and S. G. Ritchie, “A neural network-based methodology for pavement crack detection and classification,” *Transportation Research Part C: Emerging Technologies*, vol. 1, no. 4, pp. 275–291, 1993.
- [11] K. R. Kirschke and S. A. Velinsky, “Histogram based approach for automated pavement crack sensing,” *Journal of Transportation Engineering*, vol. 118, no. 5, pp. 700–710, 1992.
- [12] J. Zhou, P. S. Huang, and F. P. Chiang, “Wavelet-aided pavement distress image processing,” *Optical Science and Technology, SPIE’s 48th Annual Meeting*, 2003, pp. 728-739.
- [13] F. Klugl, A. Bazzan, and S. Ossowski, “Agents in traffic and transportation,” *Transp. Res. Part C: Emerg. Technol.*, vol. 18, pp. 69-70, 2010, doi: 10.1016/j.trc.2009.08.002.
- [14] E. Dogan and A. Akgng, “Forecasting highway casualties under the effect of railway development policy in Turkey using artificial neural networks,” *Neural Comput. Appl.*, vol. 22, pp. 869-877, 2013.
- [15] S. Budalakoti, A. Srivastava, and M. Ote, “Anomaly detection and diagnosis algorithms for discrete symbol sequences with applications to airline safety,” *IEEE Trans. Syst. Man Cybern. Part C: Appl.*, vol. 39, pp. 101-113, 2009.
- [16] R. Wang, S. Fan, and D. Work, “Efficient multiple model particle filtering for joint traffic state estimation and incident detection,” *Transp. Res. Part C: Emerg. Technol.*, vol. 71, pp. 521-537, 2016.
- [17] Y. Lv, Y. Duan, W. Kang, and Z. Li, “Traffic flow prediction with big data: A deep learning approach,” *IEEE Trans. Intell. Transp. Syst.*, vol. 16, pp. 865-873, 2014.
- [18] K. Gopalakrishnan, S. Khaitan, and A. Choudhary, “Deep convolutional neural networks with transfer learning for computer vision-based data-driven pavement distress detection,” *Constr. Build. Mater.*, vol. 157, pp. 322-330, 2017.
- [19] D. Zhang, Q. Li, Y. Chen, and M. Cao, “An efficient and reliable coarse-to-fine approach for asphalt pavement crack detection,” *Image Vis. Comput.*, vol. 57, pp. 130-146, 2017.
- [20] G. Wu, X. Sun, L. Zhou, and H. Zhang, “Research on morphological wavelet operator for crack detection of asphalt pavement,” *Proceedings of the 2016 IEEE International Conference on Information and Automation*, Ningbo, China, 2016, pp. 1573-1577.

- [21] H. Oliveira, J. Caeiro, and P. Correia, "Accelerated unsupervised filtering for the smoothing of road pavement surface imagery," *14 Proceedings of the 2014 22nd European Signal Processing Conference (EUSIPCO)*, Lisbon, Portugal, 2016, pp. 2456-2469.
- [22] M. R. Schlotjes, M. Burrow, H. Evdorides, and T. Henning, "Using support vector machines to predict the probability of pavement failure," *Proc. Inst. Civ. Eng. Transp.*, vol. 168, pp. 212-222, 2015.
- [23] T. Nguyen, T. L. Nguyen, and A. I. Greglea, "Machine learning algorithms application to road defects," *Intelligent Decision Technologies*, vol. 12, pp. 59-66, 2018.
- [24] Y. Boykov and V. Kolmogorov, "An experimental comparison of min-cut/max- flow algorithms for energy minimization in vision," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 9, pp. 1124-1137, 2004.
- [25] C. Koch and I. Brilakis, "Improving Pothole Recognition through Vision Tracking for Automated Pavement Assessment," *Conference or Workshop Item*, 2011, pp. 1-8.
- [26] F. M. Nejad and H. Zakeri, "An optimum feature extraction method based on wavelet-Radon transform and dynamic neural network for pavement distress classification," *Expert Syst. Appl.*, vol. 8, pp. 9442-9460, 2011.
- [27] H. N. Koutsopoulos and A. B. Downey, "Primitive-based classification of pavement cracking images," *Journal of Transportation Engineering*, vol. 119, pp. 402-418, 1993.
- [28] L. Gang, H. Yu-yao, and Z. Yan, "Automatic Recognition Algorithm of Pavement Defect Image Based on OTSU and Maximizing Mutual Information," *Microelectronics Computer*, vol. 7, pp. 241-247, 2009.
- [29] Fugro, "Automatic Road Analyzer," 2011. [Online]. Available: <http://www.roadware.com/products/9000>. [Accessed October 10, 2021]
- [30] Dhdv, "WayLink Digital Highway Data Vehicle," 2011. [Online]. Available: <http://www.waylink.com/DHDV.htm>. [Accessed October 10, 2021]
- [31] K. A. Abaza, S. A. Ashur, and I. A. Al-Khatib, "Integrated Pavement Management System with a Markovian Prediction Model," *Journal of Transp. Eng.*, vol. 130, no. 1, pp. 24-33, 2004.
- [32] Gie, "Technologies Laservision," 2011. [Online]. Available: http://www.gieinc.ca/main_en.html. [Accessed October 10, 2021]
- [33] Q. Li, M. Yao, and X. Yao, "A real-time 3D Scanning System for Pavement Distortion Inspection," *Measurement Science and Technology*, vol. 21, pp. 1-9, 2010.
- [34] A. Makhmalbaf, M. W. Park, and J. Yang, "2D Vision Tracking Methods Performance Comparison for 3D Tracking of Construction Resources," *Proc. of Construction Research Congress.*, 2010, pp. 459-469.