

USING ARTIFICIAL INTELLIGENCE FOR AUTOMATING PAVEMENT CONDITION ASSESSMENT

O. D. Aslan¹, E. Gultepe², I. J. Ramaji^{3*}, S. Kermanshachi⁴

¹Department of Civil Engineering, College of Computing, Engineering and Construction, University of North Florida, Jacksonville, FL, USA

²Department of Computer Science, College of Computing, Engineering and Construction, University of North Florida, Jacksonville, FL, USA

³Department of Construction Management, School of Engineering, Computing and Construction Management, Roger Williams University, Bristol, RI, USA

⁴Department of Civil Engineering, College of Engineering, University of Texas at Arlington, Arlington, TX, USA

* Corresponding author

ABSTRACT The financial burden due to pavement damage on road networks is a major handicap to the economic development of a country. According to an ASCE report, this issue may cost as much as \$67 billion per year. Regularly planned condition assessments and repairs of pavement can mitigate any derived costs and increase traffic safety. However, due to the large extent of civil infrastructure networks, required periodic inspections and assessments can be expensive and time-consuming. Further compounding the issue is that the majority of damage assessment mechanisms rely on human visual analysis, which can be prone to potential user bias and errors. In this study, we present a framework to automate roadway assessment by implementing a Convolutional Neural Network (CNN) that classifies various types of cracks in pavements. CNNs are a special type of deep artificial neural networks that demonstrate high accuracy and efficiency in image-based machine learning tasks. One of the main advantages of CNNs is that they can automatically learn the salient features of an image dataset without any prior knowledge or pre-processing by the user. Thus, the need for feature engineering is obviated and thereby eases the deployment of our assessment framework. Our framework was developed and tested on a balanced dataset containing 400 color images and consisting of four types of pavement damage: (1) longitudinal, (2) transverse, (3) alligator, and (4) pothole cracks. We apply image augmentation using a bundle of transformations to improve the crack classification accuracy of our CNN. The classification accuracy of the four types of cracks was found to be 76.2%. Demonstrating that the proposed CNN model can predict crack types without any user intervention at a good level of accuracy. To improve the robustness and accuracy of our assessment framework, we will analyze more types of cracks, using a larger dataset size in future studies.

1. Introduction

Asphalt roads are developed with delicate materials that exhibit distress due to loading, aging and environmental conditions. In most circumstances, these distresses appear as surface cracks. In the United States, the road network includes more than four million miles of public roadways, in which 32% of the major roads are in poor or mediocre condition (Herrmann, A.W, 2013). Thus, as a preventative measure and to reduce the cost of maintenance, early crack detection and classification are critical aspects of asphalt road assessment.

In addition, automotive safety is significantly affected by pavement conditions due to destructive effects on vehicles. During 2005 to 2009, of the 189,645 deaths on U.S. highways, one-third of the fatalities were due to poor roadway conditions (Nelli et al., 2014). Current pavement condition assessment schemes can be very time consuming, biased, laborious and costly. Furthermore, these approaches can pose safety risks to personnel involved in the process. In many autonomous pavement assessment systems, image-processing methods are employed for crack detection. Nevertheless, environmental conditions such as light and shadow condition, different

asphalt textures, and non-crack patterns can compromise the assessment outcome. Performance evaluation of various commercially available systems shows that they may have problems with non-crack patterns, which may result in false crack detection and classification (Yashon and Michael, 2016). Recent improvements in the field of artificial neural networks, especially in deep learning, have paved a new way of applying computer vision methods to pavement crack detection in pavement images.

Deep learning algorithms have been shown to successfully handle existing limitations in image processing for crack detection applications (Cha et al., 2016). One of the most established algorithms among various deep learning models for image analysis and classification is Convolutional Neural Networks (CNNs). In recent years, there have been several efforts to advance crack detection and classification using deep learning techniques. Baoxian et al. (2018) proposed a novel CNN model for pavement cracking classification based on 3D pavement images. Tong et al. (2018) proposed a two-stage model for crack detection, which first selects images that may contain cracks using clustering analysis and then follows up with a CNN classification. Zhang et al. (2016) utilized a CNN

to classify whether individual pixels belonged to a crack using localized image patches. Nevertheless, the method disregards the spatial relations between pixels and overestimates crack width (Zhang et al., 2016). Furthermore, their method is still dependent on non-automatic designed feature extractors for pre-processing and the CNN is only used as a classifier. Besides, its network architecture is strictly related to the input image size, which disallows the generalization of the method (Allen et al., 2017).

1.1 Convolutional Neural Network

CNNs were introduced in 1980 (LeCun et al., 1989; Fukushima, 1980) and over the past two decades have become the machine learning technique of choice for image classification. Since the early 2000s, CNNs have been applied with great success to classification, segmentation, and recognition of objects. In CNNs, the convolutional feature learning layers take advantage of the innate properties of images, as opposed to autoencoder neural networks, which are ambivalent to image data. CNNs take advantage of the local coherence of pixels in images to significantly reduce the number of operations needed to fit a model. By processing images on successive patches of pixels using matrix convolution, a convolution feature map is learned and used to filter the images. Pooling layers are used to reduce the size of the feature maps by extracting the most salient features. Feature pooling also introduces spatial invariance to the mapping. This has the effect of making the CNN model more robust to rotated, sheared, and flipped images.

The depth of the network (i.e., the number of convolutional and pooling layers) plays an important role in the ability of the convolutional networks to give accurate results (Krizhevsky et al., 2012). However, it is not always necessary to have very deep CNNs to achieve high accuracy. Nonetheless, using deep layered networks do tend to improve the accuracy of classification. The main advantage of CNN is that it automatically learns the important features, without any human supervision, directly from the raw data. The convolutional and pooling layers are followed by fully connected and probabilistic sigmoidal layers to provide the final classification labels (Christian et al., 2015; Krizhevsky et al., 2012).

The architecture of a typical CNN, structured with convolutional, pooling and fully connected layers is shown in Figure 1. The convolution layer is always first, which takes images as input and are represented as a matrix of pixel values. Then the model selects small image patches in order to learn the feature filters (or neurons) (Figure 3). After, the learned filters are used by applying matrix convolution to the entire input images. The filter's task is to multiply its values by the original pixel values. Convolved images are followed by a nonlinear activation function, such as rectified linear units (ReLU) and maximum pooling layers. When the image passes through one convolution layer, the output of the first layer becomes the input to the second convolutional layer. The nonlinear activation layer provides the CNN model the ability to capture complex pixel relationships images with respect to

the response variable (as a class label) (Gopalakrishnan et al., 2017; Schmidhuber et al., 2015). The pooling layer is applied to the width and height of an image and basically performs a downsampling operation by selecting the maximum activated pixel value (Long et al., 2015).

1.2 Research Aim

This study aims to develop a deep neural network-based model to detect and classify various cracks on pavement surfaces. The developed crack detection and classification model relies on a CNN. An image dataset consisting of longitudinal, transverse, alligator, and pothole cracks was used to develop the model. The network was created in Python with the Keras deep learning library using the TensorFlow backend. Thus, we demonstrate the capabilities of deep learning in pavement crack classification.

2. Materials and Methods

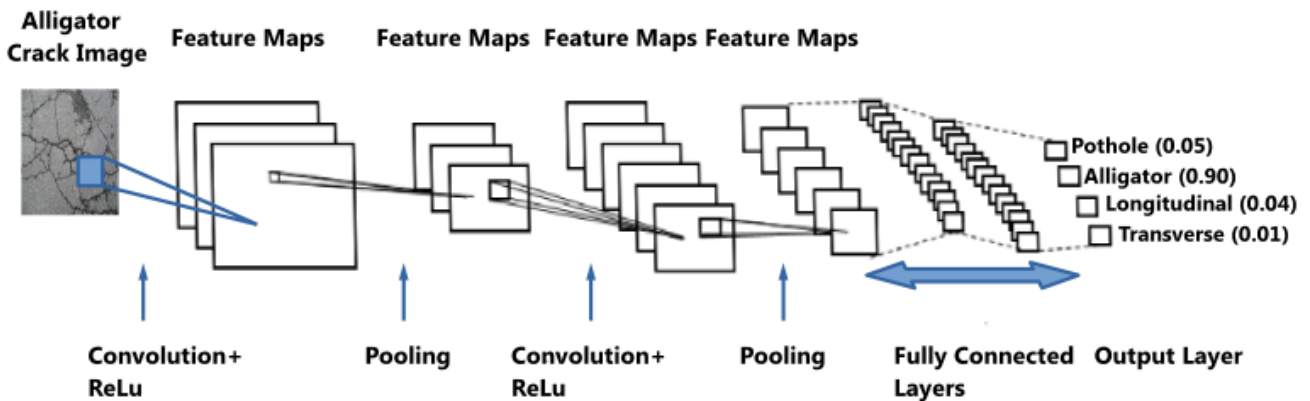
This section reviews the processing framework of the proposed crack detection and classification method. The processing framework in this study consists of three main parts: (1) creation of a labeled image dataset, (2) establish the deep learning model using a CNN, and (3) parameter selection using training and testing dataset.

2.1 Data Preparation

The dataset used in this study was collected from Google, Bing, and Yandex search engines. The dataset is comprised of 400 color images, which have been interpolated and scaled to have dimensions of 32×32 pixels. The goal was to collect images that provide good coverage of the diversity of cracks that are typically found in pavements. Consequently, the images in the dataset were manually organized and labeled into four crack categories: longitudinal, transverse, alligator, and pothole (Figure 2). To classify the images using the CNN, the data was divided into a training set consisting of 80% of the images and a test set containing the remaining 20% of the data. The images were randomly divided into the train and test sets, such that the class distribution of the images was balanced. Meaning that the ratio of example images from each type of crack is the same in both training and test sets. By ensuring that the underlying distribution of the data is correctly divided, it can be safely assumed that the overall model accuracy is an unbiased measure of the CNN model performance.

Many deep learning models require large amounts of data. Usually the larger the dataset size, the better the classification accuracy. The dataset in our study contains a relatively small amount of data compared to typical deep learning studies. To mitigate the small size of the dataset and improve classification accuracy, images were augmented using a predetermined set of image transformations. These transformations consisted of feature-wise pixel centering, normalization using the mean and standard deviation, and whitening. Therefore, these transformations help to capture the variation in the training data.

Figure 1 General and basic principles of a CNN.



2.2 Deep learning model architecture and setup

The CNN architecture consisted of two convolutional units applied before a 16 neuron fully-connected layer and a softmax classification layer with 4 output neurons (Figure 4). Each convolutional unit is composed of: (1) two successive convolutions with 32 filters each and a receptive field size of 3×3 pixels, (2) a ReLu activation (Ramachandran et al., 2017), and (3) a dropout regularization of 0.3 applied to the model weights activation (Wu et al., 2015). The model was trained until convergence using: (1) a batch size of 32 images, (2) learning rate of 0.01 with the AdaGrad optimization algorithm (Kingma et al., 2014), and (3) the categorical cross entropy loss function (Farsad et al. 2017). The CNN performance was assessed using the accuracy metric, which is defined as all true positive and negative labels divided by the total number of images in the dataset.

Figure 2 Examples of asphalt pavement crack.



For our deep learning CNN model, we implemented the model in Python using the Keras deep learning and neural network API. The Keras library was used with the TensorFlow GPU backend. The hardware setup consisted of an NVIDIA GeForce GTX 1070 GPU, 32GB of RAM memory, and a n Intel Core i7 7800X 3.5 GHz CPU. The operating system used was Ubuntu 16.04.

3. Results and Discussion

Determination and evaluation of pavement cracks have been generally accomplished using regularly performed manual inspections. This procedure typically involves a pavement condition survey form, written out by a trained practitioner of a highway agency, who travels along the road and collects both visual and quantitative data from the pavement surface by visual inspection. Although traditional manual inspection methods are the common form of practice, they can have various problems. From a reliability point of view, the manual inspection methods include a natural bias, as the major conclusions are taken based on the practitioner's point of view, which is generally dependent on the educational and expertise level of those people. From a practical point of view, this is not only monotonous, but also inefficient as it takes too much time to process the pavement data and make judgments about the pavement condition. Finally, these methods usually require efficient management of trained experts, which may bring additional constraints. Since these people are under the continuous risk of exposing themselves to traffic, there are numerous worker safety regulations. Given these onerous constraints of manual crack assessment, it is obvious why automated classification is preferred.

In this study, the focus was to develop a robust pavement crack classification system. After training the CNN model, the testing accuracy of the dataset consisting of the four crack types was found to be 76.2%. This performance is comparable to Zhang et al. (2016) who had achieved an F-score accuracy of 89.7% for a binary classification problem, which only looked at finding a crack or a non-crack in their images. Furthermore, they had used many more images in their dataset.

Our classification problem was more difficult considering that we attempted to classify four different types of cracks.

Regardless of the test case and utilized model, the performance of our dataset with fewer samples is comparable to other studies with a high number of samples. The size of training dataset tends to positively influence accuracy. When the dataset size is small, there is a risk of overfitting, unless the data is not augmented, as we have performed in our CNN.

The deep learning framework developed in this study has several strengths: it is data-driven and effective at resolving complex patterns in pavement images. Also, the deep learning system has some weakness. First, it can be data-hungry and not suitable when data size is small unless correct processing steps are not taken into consideration. Second, the computational cost of learning is high unless the user does not how to utilize the correct model parameters. Third, since the CNN is a black box model, it is not good at providing inferences for decision making.

4. Future Work

Our future work will focus on three aspects: (1) the database can be extended up to 1000 images in every class, to give a more precise and broad classification model, (2) increase the number of CNN classification groups to enrich the result benchmarks, and (3) implement general adversarial networks in order to understand the feature representation of the input images.

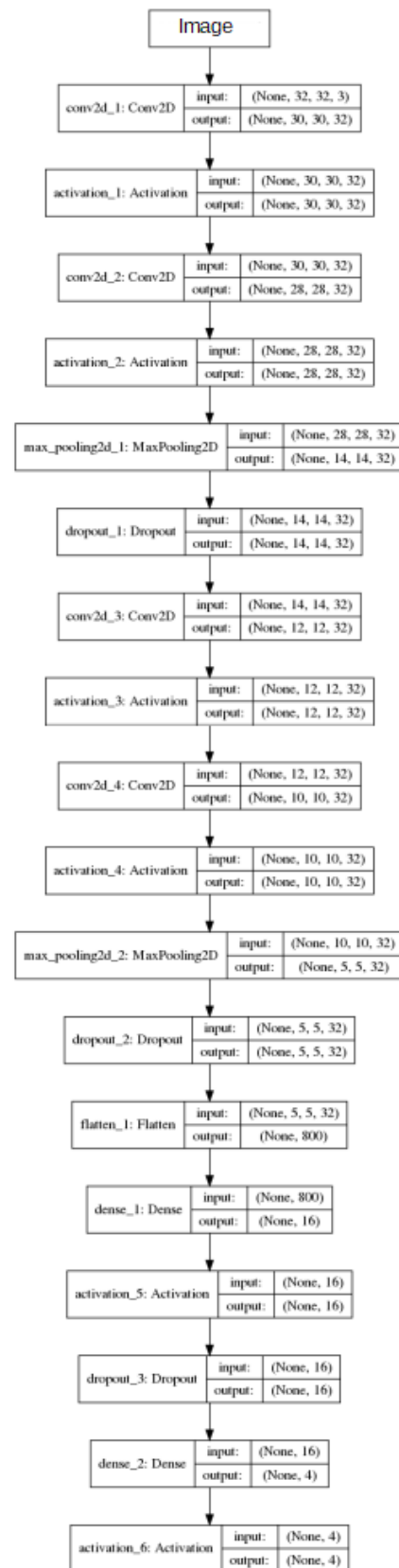
Figure 3 First layer convolutional feature maps obtained from the CNN model.



5. Conclusions

This paper proposed an automated crack detection and classification method based on deep learning (convolutional neural network) to detect and classify cracks on a small set of images taken from search engines. The proposed CNN is used to classify pavement cracks into four categories: transverse, longitudinal, pothole, and alligator cracks. The overall accuracy of the proposed CNN is 76.2%. This high-performance accuracy demonstrates the reliability of CNNs as powerful tools for classifying pavement crack images.

Figure 4 Implemented CNN architecture.



6. References

- Herrmann AW (2013). ASCE Report Card for America's Infrastructure. In: International Association for Bridge and Structural Engineering (IABSE) Symposium Report, vol. 99, pp. 9–10.
- Nelli D, Hansen M, and O'Neil M (2014). *State data system crash data report: 2000–2009*. (Report No. DOT HS 812 052). Washington, DC: National Highway Traffic Safety Administration.
- Yashon OO and Michael H (2016). 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, no. 3, pp. 481–499.
- Cha YJ, You K and Choi W (2016). Vision-based detection of loosened bolts using the Hough transform and support vector machines, *Automation in Construction*, 71(2), 181–88
- Tong Z, Gao J, Han Z, Wang Z (2018). Recognition of asphalt pavement crack length using deep convolutional neural networks. *Road Mater. Pavement Des.* 2018;19:1334–1349. doi: 10.1080/14680629.2017.1308265.
- Zhang L, Yang F, Zhang YD, Zhu YJ (2016). Road crack detection using deep convolutional neural network; *Proceedings of the IEEE International Conference on Image Processing (ICIP)*; Phoenix, AZ, USA. 25–28 September 2016; pp. 3708–3712
- Allen. Z, Kelvin. CW, Baoxian. L, Enhui. Y, Xianxing. D, Yi P, Yue F, Yang L, Joshua QL, and Cheng C (2017). Automated pixel-level pavement crack detection on 3D asphalt surfaces using a deep-learning network, *Computer-Aided Civil and Infrastructure Engineering*, vol. 32, no. 10, pp. 805–819.
- Fukushima, Kunihiko. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological Cybernetics*, 36(4):193–202, 1980. *Biological Cybernetics*, 36 (4) (1980), pp. 193-202
- Fukushima, Kunihiko. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological Cybernetics*, 36(4):193–202, 1980. *Biological Cybernetics*, 36 (4) (1980), pp. 193-202
- Christian K, Georgieva, K, Kasireddy V, Akinci B and Fieguth, P (2015). A review on computer vision based defect detection and condition assessment of concrete and asphalt civil infrastructure, *Adv. Eng. Informatics*, 29(2), 196-210.
- Krizhevsky A, Sutskever I, and Geoffrey EH (2012). Imagenet classification with deep convolutional neural networks, *Advances in neural information processing systems* ;25:1097-1105.
- Schmidhuber J, Deep learning in neural networks: An overview. *Neural Netw.* 61, 85–117 (2015). 10.1016/j.neunet.2014.09.003pmid:25462637
- Gopalakrishnan K, Khaitan SK, Choudhary A, Agrawal A (2017). Deep Convolutional Neural Networks with transfer learning for computer vision-based datadriven pavement distress detection, *Construction and Building Materials* 157: 322–330.
- Long J, Shelhamer E, Darrell T (2015). Fully convolutional networks for semantic segmentation. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3431–3440
- Ramachandran, P., Zoph, B., & Le, Q. V. (2017). Searching for activation functions. arXiv:1710.05941 [cs.NE].
- Wu, H., Gu, X.: Towards dropout training for convolutional neural networks. *Neural Netw.* 71, 1–10 (2015)
- Kingma, Diederik P. and Ba, Jimmy. Adam: A Method for Stochastic Optimization. arXiv:1412.6980 [cs.LG], December 2014.
- N. Farsad, A. Goldsmith, "Detection algorithms for communication systems using deep learning", [online] Available: <http://arxiv.org/abs/1705.08044>.
- Krizhevsky, A., Sutskever, I., & Hinton, G. (2012). ImageNet classification with deep convolutional neural networks. In NIPS.
- Li L, Sun L, Ning G and Tan S (2014). Automatic Pavement Crack Recognition Based on BP Neural Network. *Swarm Intelligence in Transportation Engineering*, PROMET - Traffic and Transportation 26(1),11-22