



Pavement Crack Detection Using a Lightweight Convolutional Neural Network

Alaa Sheta, Hamza Turabieh, Sultan Aljahdali, and Abdulaziz Alangari

¹ Computer Science Department, Southern Connecticut State University, New Haven, CT, USA
ShetaA1@southernct.edu

² Department of Information Technology, Taif University, Taif, Saudi Arabia
h.turabieh@tu.edu.sa

³ Department of Computer Science, Taif University, Taif, Saudi Arabia
aljahdali@tu.edu.sa

⁴ Taif University, Taif, Saudi Arabia
alankary@tu.edu.sa

Abstract

Automating the process of detecting pavement cracks became a challenge mission. In the last few decades, many methods were proposed to solve this problem. The reason is that maintaining a stable condition of roads is essential for the safety of people and public properties. It was reported that maintaining one mile of roads in New York City in the USA might cost from four to ten thousand dollars. In this paper, we explore our initial idea of developing a lightweight Convolutional Neural Network (CNN or ConvNet) model that can be used to detect pavement cracks. The proposed CNN was trained using the AigleRN data set, which contains 400 images of road cracks of 480×320 resolution. The proposed lightweight CNN architecture performed a better fitting to the image data set due to the reduction in the number of parameters. The proposed CNN was capable of detecting cracks with a various number of sample images. We simulated the CNN architecture over different sizes of training/testing (i.e., 90/10, 80/20, and 70/30) data sets for 11 runs. The obtained results show that 90/10 data division for training and testing is outperformed other categories with an average accuracy of 97.27%.

1 Introduction

Roads are the heart of modern cities due to their importance either for public or private transportation. Keeping these roads in excellent condition is needed for safe driving. Roads should provide acceptable characteristics (i.e., shape, surface, and friction) to keep vehicles while traveling in safe conditions at the desired speed. Automatic pavement crack detection is a challenging task in road maintenance since the cracks have different intensity inhomogeneity and complex background structure [10]. The World Bank reports that roads and transportation infrastructures are the most valuable assets in any country, and it is essential to keep them in good working condition [30]. As a result, countries try to maintain their roads regularly and keep

them in good condition, which reduces the maintenance cost in the long run. In general, roads should have a predetermined schedule to maintain to keep them safe for the public [26].

Pavement inspections are a process for both highways and local airport roads implemented by various authorities such as transportation agencies. Pavement engineers are responsible for inspecting all types of cracks, evenness, and distress by routinely collecting road condition data. A frequent pavement distress survey usually executed according to a pre-defined plan to monitor the pavement state and regulate the appropriate maintenance and rehabilitation actions. Various types of distress can be inspected, such as longitudinal cracking, crosswise cracking, chunk cracking and patching [17]. The pavement inspection process has to be implemented in various weather and traffic conditions; thus, it is time-consuming and involves human errors. Moreover, efficient pavement monitoring strategies can enhance the pavement life-cycle maintenance costs [5].

Road data collection is the first phase of the pavement crack detection process. Cracking is the most popular type of pavement distress [4]. A myriad of things could cause distress in the pavement, such as inadequate structural support for the given loading, poor construction, weather conditions like sundry, and deteriorate it [18]. Drought, soil erosion, and extended periods of rain can cause the ground beneath the pavement to shift. Poor road conditions will reduce the safety level for drivers and will increase the vehicle maintenance cost. Several municipalities prefer to construct new roads or rehabilitate dilapidated roads rather than have regular maintenance. In general, developing new pavement crack detection systems attracts many academic researchers and pavement engineers to find such systems that reduce time and cost, more efficiency, and safer [29].

Several methods have been proposed to detect pavement cracks based on image processing concepts such as adaptive thresholding and morphological processing. Sy *et al.* [22] employed bilevel thresholding, morphological operation, and projection to detect pavement cracks. The authors tested their proposed approach on three kinds of images: laboratory images, static images, and AMAC reg images. Li *et al.* [12] proposed a pavement crack thresholding method based on neighboring differential histogram statistics. Oliveira and Correia [15] proposed a local thresholding method for pavement crack images that are non-overlapping blocks. Li *et al.* [11] proposed new unsupervised multi-scale fusion crack detection (MFCFD) algorithm to detect pavement cracks. The proposed approach consists of three phases: extract the candidate cracks phase, determine the crack correspondences across different scales phase, and evaluating phase based on a multivariate statistical hypothesis test.

Image processing technique called Histograms of oriented gradients is used [6] for crack recognition. The images are collected using a camera placed orthogonally to the road pavement axis. In [27], the authors developed a supervised classification method. Such that each crack image goes across gray-scale transformation, median filtering, k-means segmentation, and finally, classification using the neighbor method. Three layers of self-studying back propagation neural network for pavement crack recognition were presented in [24]. A pavement segmentation method is developed based on the Genetic Algorithm and entropy theory to determine the threshold value accurately.

Lately, CNN provided promising results. It was used for crack detection as a supervised method based on deep learning for crack detection [30]. The features are automatically learned from manually annotated image patches acquired by a low-cost sensor, i.e., smartphone. CNN showed superior crack detection performance. Most recent work includes grid-based pavement crack detection using deep learning [23], to perform pixel-wise classification of surface cracks on road and pavement images [3], deep crack end-to-end trainable deep CNN for automatic crack detection is proposed [31] by learning high-level features of pavement crack representation.

Automating the inspection processes using robotics has also been investigated [7, 13]. The development of an autonomous robot capable of automating the pavement inspection is the primary goal of many research works. The autonomous mobile robot can be programmed to behave as an intelligent agent (IA). IA can observe the surrounding environment and react with a suitable decision. ARS can produce high productivity in inspecting the road and complete the mission in a much better, faster, and safer way. This process will reduce costs and human mistakes.

In this paper, the standard CNN architecture is presented in Section 2. The pavement crack data set used in this paper is presented in Section 3. We provide our initial idea of developing a Convolution Neural Networks (CNN) to classify the structure of the cracks from raw images collected from [19] and the system simulation and experimental results are provided in Section 4. Finally, the conclusion and future works are presented in Section 5.

2 Deep Neural Networks

Over the past few years, the CNN paradigm has shown excellent performance in tackling numerous problems in computer vision, speech recognition, natural language processing, machine translation in real-time, and drug design [16]. CNN is based on an artificial neural network (ANN) but with deep structure. The word “deep” in “deep learning” denotes the increased number of layers in the network.

The use of deep learning is superior compared to other Machine Learning algorithms the most crucial reason the model capabilities to capture features from the data. Deep-learning networks can perform automatic feature extraction without human intervention, unlike most traditional machine-learning algorithms [14]. Each layer of nodes trains on a distinct set of features based on the previous layer’s output. The deeper we go into the network, the more intricate features can be recognized. This feature hierarchy makes deep learning networks capable of handling large, high-dimensional data sets with billions of parameters that pass through nonlinear functions.

With the growing size of training data, traditional machine learning techniques like support vector machine (SVM) and logistic regression do not perform well [2]. They tend to plateau after a certain point. In contrast, deep neural network performance increases with the increased data fed to the model. Data, computation time, and algorithms drive a typical in-depth learning process.

2.1 Convolutional Neural Network

CNN is a particular type of ANN that can learn several patterns inside images by extracting features. CNN shows an excellent performance in several domains such as traffic sign recognition, image classifications, medical image segmentation [21]. Moreover, CNN enhances several machine learning methods such as transfer learning based on CNNs for object detection problems [20]. The first architecture of CNN (i.e., LetNet-5) appears in 1989, where the Back-Propagation algorithm [8] optimizes the connection weights. Later on, several architectures for CNN have been proposed, such as ResNet, LeNet, AlexNet, VGG, and GoogLeNet [25]. In general, CNN is a Feedforward Neural Networks that consists of two parts: feature extractor and classifier. The feature extractor part has convolutional and max-pooling layers, while the classifier part has a fully connected layer. The convolutional layer is the primary building block of CNN. Convolution is an operation that merges two sets of information, which has weights and biases to various aspects/objects in the image and be able to differentiate one from the other. The

main aim of the convolutional layer is to learn feature representations of the input images. The convolution is applied using a convolution filter (i.e., kernel) to produce a feature map. Each neuron located in the same feature map is employed to extract several local characteristics from different locations in the former layer. Low-level features are extracted from the first convolutional layer, such as edges, color, gradient orientation, and many others. To extract more new features, the first convolutional layer mapped with a learned kernel, where the results are moved to a nonlinear activation function. Applying different kernels will generate several new features.

A ConvNet can extract the spatial and temporal dependencies from the input image using a set of relevant filters [28]. The leading architecture of ConvNet manipulates a better fitting to the input data set (i.e., crack image) by reducing the number of parameters employed in ConvNet and the re-usability of weights. ConvNet can be trained to learn and examine the sophistication of the input image in an excellent way. A pictorial demonstration of CNN architecture is shown in Figure 1. The input image is a pavement image that either has a crack or not. The feature extractor has three convolutional and pooling layers. While the classifier has one fully connected layer, the output can be a Softmax layer indicating whether there is a crack or not.

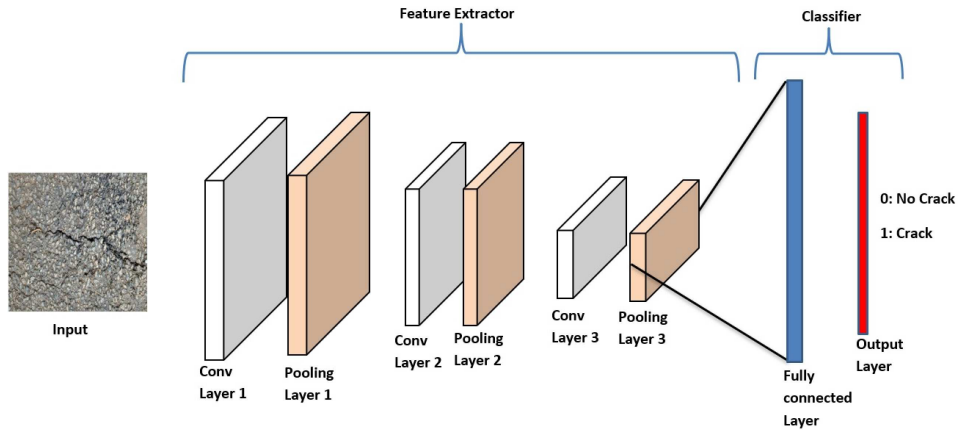


Figure 1: The CNN architecture.

3 Pavement Crack Data set

We used the AigleRN [1] data set that contains 400 images of road cracks where the image size is 480×320 with full-color [19]. The dimensions of the input images used to train the CNN is 230×230 pixels. The total number of examples is 400, where 90% was used for training, and the rest 10% for testing, as shown in Table 1.

Table 1: Training and Testing Data set

Image size	Positive examples	Negative examples
Images 320×320	200	200



Figure 2: Examples of images in AigleRN data set. Sample size is 230×230 pixel. This patch shows negative samples, that do not contain crack pixels.

4 CNN Architecture

Convolution neural networks are the state of the art technology in computer vision and other artificial intelligence fields. Their strengths lie within its ability to extract features automatically [9]. It is implemented by applying a sequence of filters (i.e., convolutions) to the input images. The proposed CNN architecture has defined three convolution layers, each with 16 filters followed by max pool layers. The classical CNN contains three layers: convolutional layers, max-pooling layers, and fully-connection (FC) layers. In this work, CNN architecture has 3 convolutional layers with 3 max-pooling layers and 2 FC layers. Figure 4 demonstrates the CNN architecture for the pavement crack problem. The size of input pavement crack image is $230 \times 230 \times 3$.

4.1 Model Evaluation

Several methods were employed to evaluate the pavement crack detection systems such as Correctly Classified Instances (CCI) and Incorrectly Classified Instances (ICI). In this paper, we used CCI and ICI to examine the overall performance of the CNN model. The equations



Figure 3: Examples of images in AigleRN data set. Sample size is 230×230 pixel. This patch shows positive samples, that contain crack pixels.

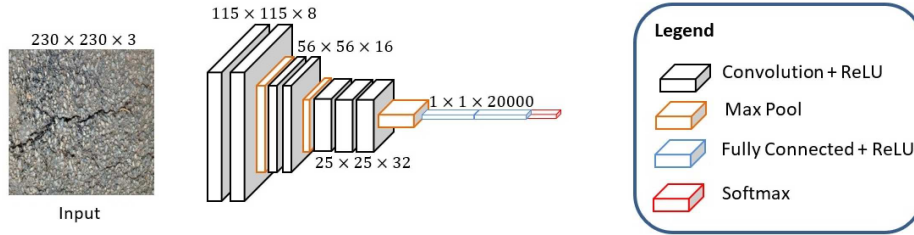


Figure 4: ConvNet architecture for pavement crack detection.

which described the CCI and ICI are presented in Equation 1 and Equation 2.

$$CCI = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

$$ICI = \frac{FP + FN}{TP + TN + FP + FN} \tag{2}$$

Where TP refers to the true-positive prediction for positive instances, TN refers to the true-negative prediction for negative instances, FP refers to false-positive prediction for positive instances, and FN refers to the false-negative prediction for negative instances. Table 2 demonstrates the Confusion Matrix (CM) to examine the pavement crack detection system.

Table 2: The confusion matrix (CM).

Actual Class	Predicted Class		
		Class = Yes (Crack exists)	Class = No (Crack not exists)
	Class = Yes	True Positive (TP)	False Negative (FN)
Class = No	False Positive (FP)	True Negative (TN)	

4.2 CNN Simulation

In this paper, we divided the data set using the holdout method (i.e., Training/Testing) into three groups: 90/10, 80/20, and 70/30. To examine the overall performance of the proposed CNN, we executed each category 11 times independently for 60 epochs for each run until we reached convergence and obtained the best accuracy. Table 3 demonstrates the obtained results based on Average, maximum, minimum, and standard deviation for each category. It is clear that based on the average accuracy results for testing data set, 90/10 gain accuracy value equals 97.27, which is outperforming other categories. Figure 5, we show the accuracy and loss of the network during the training process for the 90/10 category, respectively. The increase in accuracy and decrease in loss often represent the network's ability to capturing common patterns. Figure 6 shows the performance of CNN based on the box-plot diagram for both training and testing data sets. It is evident that, based on the average accuracy values, the 90/10 category has a stable performance over the testing data set. Finally, we can notice that the performance of CNN depends on the size of the input data. It is crucial to select a good ratio between training and testing data sets.

Table 3: CNN Simulation Results for 11 independent runs.

Training/Testing %	Training data set.				Testing data set.			
	Avg.	Max.	Min.	Std.	Avg.	Max.	Min.	Std.
70/30	99.12	100.00	97.14	0.87	96.52	98.33	95.00	0.97
80/20	97.84	100.00	85.94	4.09	95.91	98.75	80.00	5.36
90/10	98.69	100.00	94.17	1.62	97.27	100.00	92.50	2.08

5 Conclusion and Future Work

In this paper, we employed a CNN architecture to detect pavement cracks from a benchmark data set obtained from AigleRN. A set of experiments were performed using different sizes of training/testing data set. The achieved results show that 90/10 data division can produce a stable and robust automated pavement cracks system. CNN can produce a high classification accuracy in an acceptable real-time execution. Our future work includes the development of an autonomous robot system (ARS), that consists of the lightweight CNN, WiFi camera, and GPS sensor compents that can be to be executed online to detect pavement cracks. We also plan to build a GUI to analyze the detection and analysis of pavement in a better way.

References

- [1] Sylvie Chambon and Jean-Marc Moliard. Automatic road pavement assessment with image processing: Review and comparison. *International Journal of Geophysics*, 2011, 03 2011.

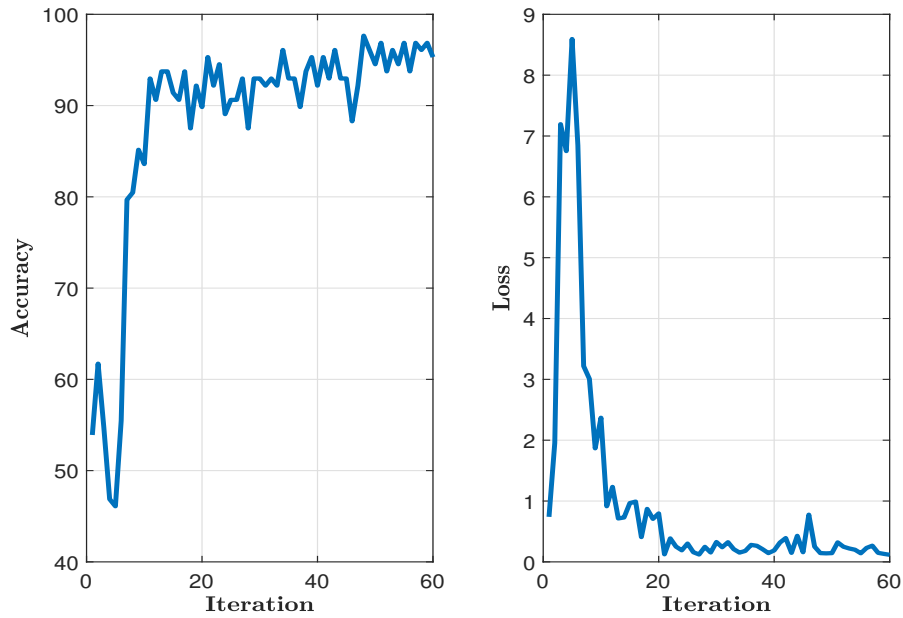


Figure 5: CNN Accuracy and Loss conversion curves

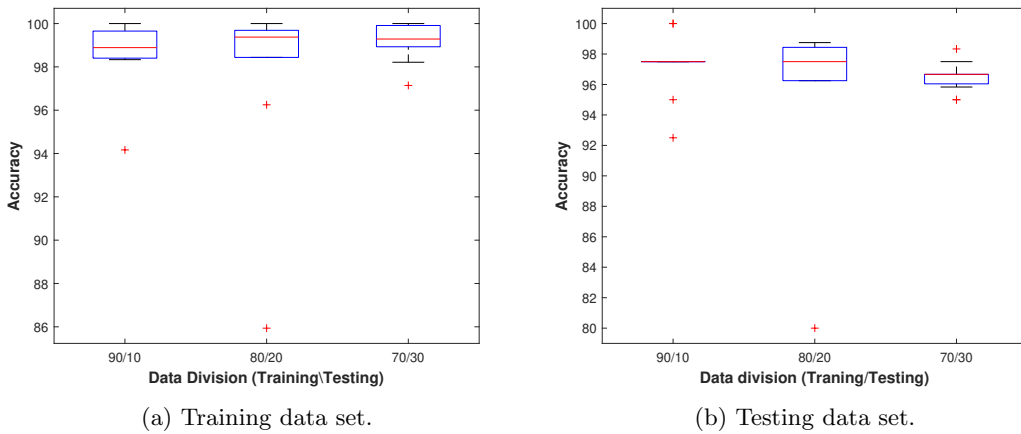


Figure 6: Boxplot figures for training and testing data set with different data division.

[2] Shaveta Dargan, Munish Kumar, Maruthi Rohit Ayyagari, and Gulshan Kumar. A survey of deep learning and its applications: A new paradigm to machine learning. *Archives of Computational Methods in Engineering*, Jun 2019.

[3] M. David Jenkins, T. A. Carr, M. I. Iglesias, T. Buggy, and G. Morison. A deep convolutional neural network for semantic pixel-wise segmentation of road and pavement surface cracks. In *2018 26th European Signal Processing Conference (EUSIPCO)*, pages 2120–2124, Sep. 2018.

[4] Zhun Fan, Yuming Wu, Jiewei Lu, and Wenji Li. Automatic pavement crack detection based on

- structured prediction with the convolutional neural network. *ArXiv*, abs/1802.02208, 2018.
- [5] Kasthurirangan Gopalakrishnan, Siddhartha K. Khaitan, Alok Choudhary, and Ankit Agrawal. Deep convolutional neural networks with transfer learning for computer vision-based data-driven pavement distress detection. *Construction and Building Materials*, 157:322 – 330, 2017.
 - [6] R. Kapela, P. Śniatała, A. Turkot, A. Rybarczyk, A. Pożarycki, P. Rydzewski, M. Wyczalek, and A. Błoch. Asphalt surfaced pavement cracks detection based on histograms of oriented gradients. In *2015 22nd International Conference Mixed Design of Integrated Circuits Systems (MIXDES)*, pages 579–584, June 2015.
 - [7] T. Le, S. Gibb, N. Pham, H. M. La, L. Falk, and T. Berendsen. Autonomous robotic system using non-destructive evaluation methods for bridge deck inspection. In *2017 IEEE International Conference on Robotics and Automation (ICRA)*, pages 3672–3677, May 2017.
 - [8] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel. Backpropagation applied to handwritten zip code recognition. *Neural Computation*, 1(4):541–551, Dec 1989.
 - [9] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *nature*, 521(7553):436, 2015.
 - [10] Giovanni Leonardi, Vincenzo Barrile, Rocco Palamara, Federica Suraci, and Gabriele Candela. Road degradation survey through images by drone. In Francesco Calabrò, Lucia Della Spina, and Carmelina Bevilacqua, editors, *New Metropolitan Perspectives*, pages 222–228, Cham, 2019. Springer International Publishing.
 - [11] H. Li, D. Song, Y. Liu, and B. Li. Automatic pavement crack detection by multi-scale image fusion. *IEEE Transactions on Intelligent Transportation Systems*, 20(6):2025–2036, June 2019.
 - [12] Q. Li and X. Liu. Novel approach to pavement image segmentation based on neighboring difference histogram method. In *2008 Congress on Image and Signal Processing*, volume 2, pages 792–796, May 2008.
 - [13] Christopher Maynard, Robert L. Williams, Paul Bosscher, and Lindsey Bryson. Autonomous robot for pavement construction in challenging environments. *Earth & Space*, 2006, 03 2006.
 - [14] Heba Mohsen, El-Sayed A. El-Dahshan, El-Sayed M. El-Horbaty, and Abdel-Badeeh M. Salem. Classification using deep learning neural networks for brain tumors. *Future Computing and Informatics Journal*, 3(1):68 – 71, 2018.
 - [15] H. Oliveira and P. L. Correia. Automatic road crack segmentation using entropy and image dynamic thresholding. In *2009 17th European Signal Processing Conference*, pages 622–626, Aug 2009.
 - [16] Samira Pouyanfar, Saad Sadiq, Yilin Yan, Haiman Tian, Yudong Tao, Maria Presa Reyes, Meiling Shyu, Shu-Ching Chen, and S. S. Iyengar. A survey on deep learning: Algorithms, techniques, and applications. *ACM Comput. Surv.*, 51(5):92:1–92:36, September 2018.
 - [17] S. Saat, A. R. M. Kamil, M. Z. M. Tumari, and A. S. R. A. Subki. Development of an autonomous robot for inspection system. In *2018 IEEE 14th International Colloquium on Signal Processing Its Applications (CSPA)*, pages 272–276, March 2018.
 - [18] Anton K. Schindler and B. Frank McCullough. Importance of concrete temperature control during concrete pavement construction in hot weather conditions. *Transportation Research Record*, 1813(1):3–10, 2002.
 - [19] Yong Shi, Limeng Cui, Zhiquan Qi, Fan Meng, and Zhensong Chen. Automatic road crack detection using random structured forests. *IEEE Transactions on Intelligent Transportation Systems*, 17(12):3434–3445, 2016.
 - [20] H. Shin, H. R. Roth, M. Gao, L. Lu, Z. Xu, I. Nogues, J. Yao, D. Mollura, and R. M. Summers. Deep convolutional neural networks for computer-aided detection: Cnn architectures, dataset characteristics and transfer learning. *IEEE Transactions on Medical Imaging*, 35(5):1285–1298, May 2016.
 - [21] Y. Sun, B. Xue, M. Zhang, and G. G. Yen. Evolving deep convolutional neural networks for image classification. *IEEE Transactions on Evolutionary Computation*, pages 1–1, 2019.

- [22] N. T. Sy, M. Avila, S. Begot, and J. C. Bardet. Detection of defects in road surface by a vision system. In *MELECON 2008 - The 14th IEEE Mediterranean Electrotechnical Conference*, pages 847–851, May 2008.
- [23] X. Wang and Z. Hu. Grid-based pavement crack analysis using deep learning. In *2017 4th International Conference on Transportation Information and Safety (ICTIS)*, pages 917–924, Aug 2017.
- [24] G. Xu, J. Ma, F. Liu, and X. Niu. Automatic recognition of pavement surface crack based on bp neural network. In *2008 International Conference on Computer and Electrical Engineering*, pages 19–22, Dec 2008.
- [25] Akira Yamada, Kazuki Oyama, Sachie Fujita, Eriko Yoshizawa, Fumihito Ichinohe, Daisuke Komatsu, and Yasunari Fujinaga. Dynamic contrast-enhanced computed tomography diagnosis of primary liver cancers using transfer learning of pretrained convolutional neural networks: Is registration of multiphasic images necessary? *International Journal of Computer Assisted Radiology and Surgery*, 14(8):1295–1301, Aug 2019.
- [26] F. Yang, L. Zhang, S. Yu, D. Prokhorov, X. Mei, and H. Ling. Feature pyramid and hierarchical boosting network for pavement crack detection. *IEEE Transactions on Intelligent Transportation Systems*, pages 1–11, 2019.
- [27] G. Younes, C. Hadda, N. Attia, and Z. Djelloul. Supervised learning and automatic recognition of asphalt pavement deteriorations. In *2009 IEEE/ACS International Conference on Computer Systems and Applications*, pages 205–210, May 2009.
- [28] H. Yue, J. Liu, J. Yang, X. Sun, T. Nguyen, and F. Wu. Ienet: Internal and external patch matching convnet for web image guided denoising. *IEEE Transactions on Circuits and Systems for Video Technology*, pages 1–1, 2019.
- [29] Dejin Zhang, Qingquan Li, Ying Chen, Min Cao, Li He, and Bailing Zhang. An efficient and reliable coarse-to-fine approach for asphalt pavement crack detection. *Image and Vision Computing*, 57:130–146, 2017.
- [30] L. Zhang, F. Yang, Y. Daniel Zhang, and Y. J. Zhu. Road crack detection using deep convolutional neural network. In *2016 IEEE International Conference on Image Processing (ICIP)*, pages 3708–3712, Sep. 2016.
- [31] Q. Zou, Z. Zhang, Q. Li, X. Qi, Q. Wang, and S. Wang. Deepcrack: Learning hierarchical convolutional features for crack detection. *IEEE Transactions on Image Processing*, 28(3):1498–1512, March 2019.