A Theoretical Framework Towards Building a Lightweight Model for Pothole Detection using Knowledge Distillation Approach

Aminu Musa[1*], *Mohamed Hamada*[2**], and *Mohammed Hassan*[3**]

[1]Department of Computer Science, Federal University Dutse, Nigeria
[2] University of Aizu, Japan

[3]Department of Software Engineering, Bayero University, Kano. Nigeria

Abstract. Despite recent advances in deep learning, the rise of edge devices, and the exponential growth of Internet of Things (IoT) connected devices undermine the performance of deep learning models. It is clear that the future of computing is moving to edge devices. Autonomous vehicles and self-driving cars have leveraged the power of computer vision, especially object detection to navigate through traffic safely. Nevertheless, to be able to drive on all types of roads, these new vehicles have to be equipped with a road anomaly detection system, which strikes the need for small deep learning models of detecting road anomalies that can be deployed on these vehicles for safer driving experience. However, the current deep learning models are not practical on embedded devices due to the heavy resource requirements of the models, as such cannot be deployed on embedded devices. This paper proposes a theoretical approach to building a lightweight model from a cumbersome pothole detection model that is suitable on edge devices using knowledge distillation. It presents the theoretical approach of knowledge distillation, why it is a better technique of model compression compared to the rest. It shows that a cumbersome model can be made lightweight without sacrificing accuracy and with a reduced time complexity and faster training time.

Keywords: Knowledge distillation, Deep learning, Model compression, Pothole detection.

1 Introduction

Pothole is a natural cave or a hollow on the road surface formed as a result of erosion or aging of asphalt. Potholes pose a lot of dangers for road transport users in many developing countries, especially in Nigeria. The task of maintaining roads and removing these road anomalies is an expensive and tedious one, due to the nature of landmass and climate conditions in Nigeria. It is reported that pothole is the second largest cause of accidents in Nigeria apart from over speeding and reckless driving, with annual reported accidents surpassing 45% [1]. The problem of potholes in Nigeria cannot be eradicated completely by government but rather how to manage it and drive safely.

This calls the attention of some researchers to investigate the possible ways of detecting potholes and ways to avoid them for safe driving. [2] Proposed a real-time pothole detection system based on mobile sensing. The work was based on the vehicle accelerometer whereby the accelerometer data is normalized by Euler angle computation and is adopted in the pothole detection algorithm to obtain the pothole information by a moving vehicle. The solution was based on the technology that can be considered obsolete nowadays, with advent of recent emergent technologies, it is possible to build a pothole detection model based on these technologies. [3] Proposed a laser-based approach of detecting potholes by a moving vehicle to improve driver's safety. Other several approaches exist such as approaches based on IoT and kinetic sensors as in the case of [4]. However, all the above-mentioned proposed models of pothole detection are subject to an improvement due to technological advancement, meanwhile, there are better ways to do it with greater accuracy and efficiency. The works mentioned above are very tedious and expensive to implement, some even require a driver to be in the pothole before it can detect it, in the case of vibration-based models.

Researchers in computer vision proposed several facile models of pothole detection using convolutional neural network. These solutions achieve a great detection accuracy ranging up to 97%. But one of the major problems of CNN is, it is resource expensive as such the solutions cannot be developed and deployed on embedded devices. Embedded device environment is characterized by limited resources.

* Corresponding author: musa.aminu@fud.edu.ng
** email: hamada@u-aizu.ac.jp
***e-mail: mhassan.se@buk.edu.ng

CNNs have proven to be an excellent algorithm in object detection. However, the architecture of CNN is growing deeper and deeper, with the number of parameters increasing from millions to billions, which limits the use of these models to just high computational devices. With increasing growth of embedded devices like smartphones, tablets, car kits and various other devices, the demand for lighter models with a smaller number of parameters is like never before. This presents a kind of trade-off between the existing models and the environment within which the model will be used.

To eliminate the trade-off, the models have to be made lighter, to enable the models to fit in embedded environments. This is why a novel lightweight model for pothole detection is proposed based on knowledge distillation technique. With an aim of producing a lighter model made from a shallow neural network, capable of detecting potholes as accurately as the deep models with heavy resource requirements.

The model is based on teacher-student learning, where a cumbersome convolutional neural network (CNN) called the teacher model will be built with several hidden layers and enormous parameters to achieve higher accuracy. The model will be trained on the pothole dataset, a soft label will be generated which will serve as distilled knowledge. On the other hand, a shallow NN model with just one hidden layer called the student model will be built which takes the distilled knowledge from the teacher model and images from dataset as input to be trained on. By utilizing the distilled knowledge, the model can be trained faster, with fewer parameters, and still be able to generalize on dataset just like the dense teacher model. The student model will utilize fewer computing resources, and still provide a remarkable performance in terms of accuracy, the model can be deployed easily on embedded as it is very light compared to the teacher model, it can be run locally without internet connectivity and can be trained without GPU intervention, as such eliminating the heavy resources requirement of the existing models.

The rest of the paper is organized as follows: Section 2 discussed the current literature in the field. Section 3 described the conceptual framework of the proposed model. Section 4 present discussions, and assumptions about expected results of the study, as the research is still in progress. Section 5 concludes the paper with a conclusion and recommendation of the paper.

2 Related work

In an effort to produce a low-cost solution that can be deployed in personal cars, researchers in the field of machine learning (ML) specifically computer vision began to explore the possibilities of building a machine learning model to aid in pothole detection. The prevalent nature of machine learning allows it to have application in almost every aspect of life, such as health, communication, commerce, and so on, for instances, [21], [22] used ML techniques in health sector to detect different type of cancer, the idea of whether ML can be used to solve problems in the transportation sector was clearly a viable one.

Recently, researchers in areas of Artificial Intelligent have had a breakthrough by creating autonomous vehicles and self-driving cars, these vehicles are capable of driving themselves around, from one destination to another with minimal or no user intervention, to be able to produce such vehicles so many factors have to be put into consideration. For example, the vehicle must be equipped with a lot of sensors for it to navigate safely, each sensor carrying out a specific task, object detection is at the core of these vehicles, to be able to differentiate between road surface and pavement, to distinguish several road signs and act accordingly, to detect other vehicles and the distance between them. All the aforementioned factors have to be considered for an autonomous vehicle to function normally.

Furthermore, deep learning has made it possible for a convolutional neural network (CNN) to detect and classify different images with higher accuracy. From this view authors of [23] proposed an intelligent disease detection model, using the plant leaves images. This calls the attention of some researchers to investigate the possibility of using CNN in detecting potholes for safer driving.

Jo *et al* [5], proposed ML based approach to pothole detection. In which the locations and size of potholes were detected using black-box camera. The pothole information was transmitted to the central pothole database for road maintenance. Prior to their research, the process of detection and updating database were all done manually, the process requires an automatic pothole detection system, that can collect pothole information at low cost and over a wide area.

Additionally, Silvister et al. [6] used a DCNN to detect potholes in real-time. The system uses a wireless portable camera and GPS for the location tagging, the system used vision object detection and image ZMQ library to stream the frames and process them in the processor PC. The capturing and streaming activity performed very well from the mini portable computer camera which is attached to the vehicle.

Moreover, Rasyid et al. [7] used deep learning technology and smartphone's camera mounted on a moving vehicle to scan road surface and classify whether it is road surface or a pothole, on detecting pothole the system alerts the driver in real-time. The work extends the previous works by providing severity information of the pothole. The system achieved a significant accuracy of 96.7%, but the system requires internet access before it can classify images because the model was built on top Google cloud environment with several high-performance computing resources. As a result, the model can only be used remotely, cannot be installed in a resource-constrained environment or embedded device efficiently.

Almost all the models of pothole detection based CNN found in literature were resources expensive, they make use of several hidden layers and too many parameters to achieve better performance at the cost of high computation resources [8].

Others tried to minimize the huge computational requirements of the model by trying to reduce the model size and suit the embedded environment. Camilleri and Gatt [12] proposed a model of pothole detection based on YOLO algorithm, which has 106 fully convolutional layers stacked together for better performance. The author performed a number of experiments with different models to obtain an optimal model that fits in an embedded device like smartphones. The second model proposed by the author was based on yolo-v3 tiny which has fewer layers than the original YOLO, and model 3 was based on ssd mobile net. The experimental results show that model 3 has the best resource requirement for embedded devices, but the accuracy is not as good as model 2 or 1.

The issue of model compression comes in handy, in an effort to come up with a light model that can fit the environment. Many techniques were proposed in the past to make models lighter, but some techniques achieved that at the expense of accuracy. This is why the selection of techniques to use in designing a light model is still researchable.

In line with this, Zhigiang and Yuyang [10] described several methods of compressing a CNN model, in their paper, they emphasized techniques such as model quantization, model pruning, and knowledge distillation, they carried an experiment in which all three techniques were used and their performance was compared against each other, the results show that knowledge distillation achieves high accuracy and has faster inference time than the remaining techniques.

Knowledge distillation (KD) was first described by Hinton et al. [19], where they describe that large CNN models in an ensemble or single architecture can be trained

With a very strong regularizer. After the training of the teacher model, a different kind of training, called "distillation" allows the transfer of knowledge from the cumbersome model to a small model (student) that is more suitable for deployment.

Abbasi et al. [9] described in their paper that knowledge distillation can be used to optimize CNN model not just for compression purposes. They mentioned that the technique can also be used in a situation where the environment is resource-constrained, in a situation where the dataset is not enough, or in a situation where a network accuracy wants to be enhanced but without adding any network complexity. Which makes it an optimal technique for an embedded environment.

Moreover, Allen-Zhu [18] presented a thorough survey on KD, and extends it to show that KD can be ensemble. That is the student model can have more than one teacher model in order to improve the performance and help the student learns faster. In the end, they proposed a new technique called self-distillation, which can be viewed as implicitly combining ensemble and knowledge distillation to improve model accuracy.

Similarly, [15, 16, 17] all used knowledge distillation to compress a cumbersome model in their papers, with the intent of obtaining a lightweight version that has similar performance. Their work indicates that KD can truly be used to optimize CNN model or compress it, to suit a resource-constrained environment without sacrificing accuracy as in other compression techniques.

However, for almost all the aforementioned model compression techniques found in the literature, knowledge distillation technique has never been used on pothole detection with intention of producing a lightweight, reduced parameter model that is suitable on embedded devices. Therefore, a lightweight model of pothole detection based on knowledge distillation technique is proposed.

3 Methodology

In this section, the methodology proposed in this research will be introduced, the section will describe the proposed model and its components. The section is divided into two, the first part deals with datasets and data preprocessing and section two discussed how the model is built.

3.1 Dataset Description

The dataset to be used in this research is the images of road surfaces which contain both good surfaces and potholes, the dataset is divided into only two classes that are, road surface and pothole. The images were preprocessed in order to make them uniform in terms of dimensions, they are all resized to 256 by 256 pixels. The datasets contain 789 images of good road surfaces and 670 images of potholes. The dataset to be used in this research was not found publically available on the internet therefore the authors gathered the dataset solely for the purpose of this research and made it publically available on kaggle.com for other relevant researches. The images were captured using Samsung galaxy A71 HD camera, on different highways of Nigeria (both state and federal). Fig. 1. Shows the sample of images from the dataset.

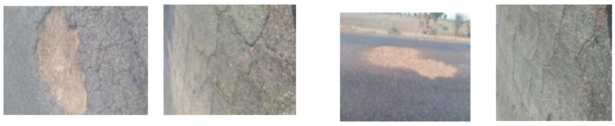

Fig.1 sample images from dataset

3.2 Proposed Model

The proposed model framework is well depicted in Fig.1. .A cumbersome CNN model named TM model is built up with two (2) convolutional layers, pooling layers and a fully connected layer with four (4) hidden layers, and an output layer. The model will be trained on the pothole dataset.

On other hand, a shallow CNN with only one hidden layer is define and is referred to as SM model. The model will take input images from dataset as inputs. The training of SM is divided into two phases. Distillation loss will be obtained during SM training phase 1. Where SM is trained based on distilled knowledge from TM . In SM training phase two SM will be trained on actual images and the two losses will be summed together and take the average loss as general loss.

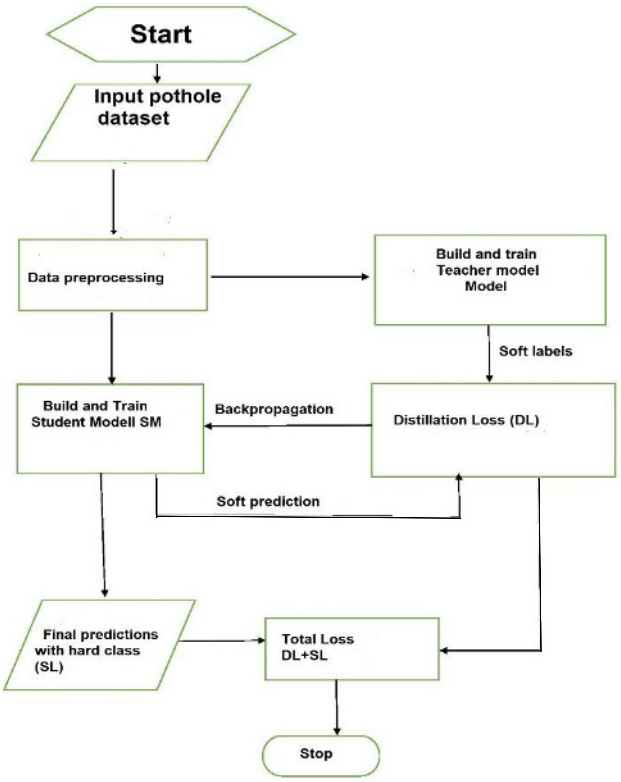

Fig.2 Proposed model framework

The main idea behind KD is to give SM model a place to start, instead of starting from scratch as the teacher does. SM will be fed with distilled knowledge, soft labels from TM, which has high entropy and provides more information than original images, as such allowing the SM to learn faster, reducing the computations needed to perform the task.

Softmax activation function with the temperature greater than 1, will be applied to the logits of TM model to obtain the soft labels, and the same function will be applied to the logits of SM model in training phase 1, in phase two, a temperature of 1 will be used in softmax function to obtained to hard labels and compare it with the ground truth The process of the teacher-student learning is presented in Fig. 2 below.

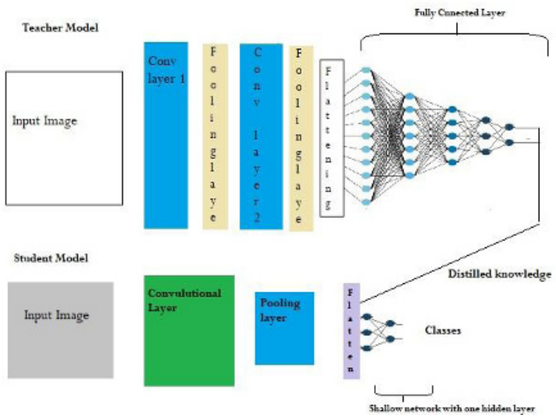

Fig. 3 Teacher-student learning

From Fig.2 The teacher model (TM) is a CNN model with an input layer, three hidden layers, and output layer with softmax activation function. The TM model will be trained on pothole dataset which was described above, to obtain the logits and soft. TM model can be formulated mathematically below:
Given an image I, with filter K of dimension k1, k2 and color channel c, a CNN will take the image and perform convolution operations by applying the filters on the images and extracting meaningful information from the image, an operation that can be regarded as automatic feature extraction.

$$(X)^{i,j} = \sum_{m=0}^{k1-1}\sum_{n=0}^{k2-1}\sum_{c=1}^{c}(K(m,n,c).I(i+m,j+n,c)+b) \quad (1)$$

If **Z**: is the output vector at last layer.

Then

$$\boldsymbol{Z} = f(x)^{i,j} \quad (2)$$

Where: f is the Softmax activation function applied on **z**.

To obtain a class probabilities, softmax activation function with a temperature greater than one (1) will be applied on **z**. the probabilities is regarded as Soft labels or the distilled knowledge, which can be obtained as follows:

$$TM = yi(x\backslash t) = \frac{e^{\frac{z^i(x)}{t}}}{\sum_j e^{\frac{z^j(x)}{t}}} \quad (3)$$

where t is parameter called temperature and **Z**i: is the logit of the input *x*.

When TM is trained with a temperature of 1, it will compute normal softmax function and therefore produces hard labels.
The soft labels from TM model is the distilled knowledge that will be integrated into the student model (SM) model to serve as additional information during the training of the SM model. The advantage of using soft labels is that it has high entropy than the hard labels which provide more information about training set than the hard labels, they also have less variance than the hard targets between training instances. The distillation loss function for the SM model can be computed as:

$$Disti\_l\,oss = \frac{1}{m}\sum_{j=0}^{m} 2t^2\,\alpha D_{KL}(\boldsymbol{p^j},\boldsymbol{q^j}) - (1-\alpha)\sum_{i=1}^{c}(yi)^j\; l\,og(1-(yhat)^j) \quad (4)$$

Such that: *m* is the batch size
P: is the soft labels from teacher network
q: is the softmax scores of student model
Dkl: is the Kullback- Liebler (KL) divergence between p and q
α: is the relative importance of the teacher's control.

The distillation loss function presented in equation (4) uses KL Divergence of the teacher and student's class probabilities. In essence the distillation loss function is the difference between soft student predictions and soft teacher labels. While student loss function is the difference between SM prediction and ground truth using softmax function with temperature of 1.

The cross-entropy loss acts as the first part of the total loss for SM training with weight (α), the other batch, is the one where the SM model is trained to minimize the cross-entropy loss between SM prediction output (with Temperature t=1) and hard labels (real labels), the loss serves as the other part of total loss of SM model with Weight (1- α).

The SM loss function can be computed in equation (5) below:

$$SM \;=\; yi(x\backslash t) = e\hat{}((Z\hat{}i\ (x))/t)/(\textstyle\sum\_j\ (e\hat{}(z\hat{}j\ )\ (x))/t) \quad (5)$$

With t=1.

Therefore the general loss function can be computed by performing a forward propagation on both TM and SM during training to calculate the general of the model.

General loss= $\alpha(SMloss) + (1-\alpha)Dist\_loss$ (6)

The entire learning process of TM and SM models can be summarized in Fig.3: below, which displays proposed model algorithm.

Algorithm 1: KD PSEUDOCODE

- Input: Deep CNN model TM
- Output: Shallow Student model SM
- Declare TM, SM, Temperature T, Train_Set X and Softmax F
- Initialize TM model

1. Input train_set images X
2. Do
   1. train TM:
   2. Forward pass
   3. Backpropagation
3. While CE <=0
4. Stop
5. Apply Softmax with temperature (T) on logits y to soften the labels
6. Obtain soft labels: Z← *F*(*y*)
7. Initialize SM Model:
8. Feed soft labels Z, train_set X
9. Do
   1. Train SM with Z :
   2. Train SM with X
   3. Obtain General loss
10. While CE<=0
11. Stop
12. Obtained Train Student model

Fig 3: Algorithm 1.

4 Discussion

As the paper title implied, theoretical framework, there are no concrete results to be presented yet, but rather proofs of author's assumptions and expected results as the research is still ongoing.

Table I: Model evaluation metrics

Metric	TM	SM
Parameters	Many	Less
Hidden layers	4	1
Running time	O(2n^2)	O(n)

Table I, presented some expected differences in the two models, which have a key impact on model performance. Logically it can be assumed that the model with fewer trainable parameters and few layers will have less complexity and is going to need low computational resources. Whereas deep models like TM with a high number of hidden layers and trainable parameters are typically complex and computationally expensive.

From the methodology, it can be inferred that knowledge distillation can be used to compress large cumbersome models that are resource expensive into a shallow network, by guiding a smaller model called student model with different architecture but similar performance.

The advantages of the SM over the TM are enormous, it is faster to train, requires less computational resources as such can be deployed in a resource-constrained environment.

The main aim of this paper is to come up with a scalable model of pothole detection which can be deployed in an embedded device easily. Therefore the paper has described a theoretical roadmap and cost-effective way of achieving the aim via knowledge distillation technique.

Conspicuously, the architecture of TM and SM differs, features and parameters of cumbersome TM are different from the SM model. The network architecture is entirely different. Unlike in transfer learning where both models' architecture has to be the same. This shows that knowledge distillation is a better option than transfer learning as it can minimize the network architecture. This also indicates that parameters will be reduced significantly, since the number of hidden layers is reduced, which prove that the two models will have different time complexity. This in turn reduced the resource cost of the model. TM is DCNN while SM is a shallow network.

The time complexity of the two algorithms is another interesting metric to consider, the TM is DCNN which has worst running time of O(2n^2) compared to SM model with O(n). The complexity depends on the input size, number of layers, number of neurons at each layer, and trainable parameters.

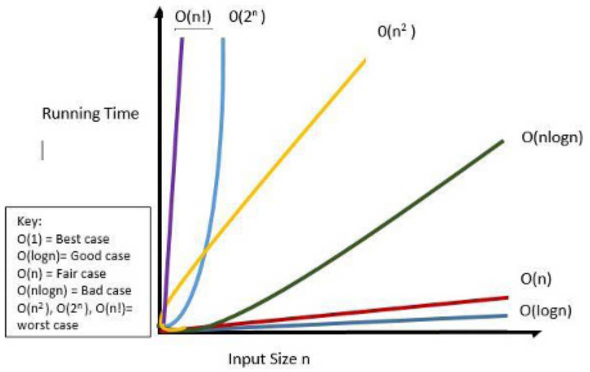

Fig. 4: Time complexity graph

Fig.4 shows the graph of time complexities of algorithm using big oh notation, it describes how running time of an algorithm grows with increasing input size. It can be seen that the best running time is O(1), followed by logarithmic function O(logn) as good, then a linear function O(n) as fair, then O(nlogn) as bad, and followed by factorial, quadratic and polynomials as the worst cases.

When calculated it reveals that TM model has a running time of O(2n^2) while SM has a running time of O(n). This shows that SM is a better model in terms of time complexity or Tm has the worst time complexity.

Obviously, SM will train faster than TM due to the difference in their architecture and fewer parameters, and still be able generalized. This shows that SM is best suited for edge devices. It takes into cognizance the resource constrains of the environment, and provide a model that can fit just right.

5 Conclusion and future work

With the rise of autonomous vehicles, self-driving cars, the need for a pothole detection model is necessary, not just any model, rather a scalable lightweight model that can be deployed on edge devices. According to Statista [20], the total number of installed embedded connected devices is projected to be 21.5 billion by 2025. With this large amount of edge devices, computation is coming down to edge devices eventually, the need for algorithms and models to accommodate these devices is like never before. Cumbersome, resource expensive algorithms like DCNN are clearly not practical on edge devices. Knowledge distillation allows us to compress a model without shortening the performance of the model. The technique was applied in this paper, which demonstrates how it was used to compress a cumbersome pothole detection model into a shallow neural network model. The model can be easily trained and can be deployed on any embedded device decently.

The main goal of the paper is to have a small and compact model of pothole detection to mimic the performance of the cumbersome model. In future, it is recommended that the model be developed and evaluated on standard performance metrics such as accuracy, precision, recall, and F1 score. To give more insight into the performance of the model. The work can also be extended by developing and deploying this model on the embedded device for practical use.

References

[1]P. Nduka, “The Causes of Potholes On Roads, Their Effects And Control Methods,” Engineering All: Engineering (Science & Technology), Technical Posts (Diy), And Business Investments In Nigeria, Aug. 20, 2021. https://www.engineeringall.com/the-causes-of-potholes-on-roads-their-effects-and-control-methods/ (accessed Aug. 27, 2021).

[2]H.-W. Wang, C.-H. Chen, D.-Y. Cheng, C.-H. Lin, and C.-C. Lo, “A Real-Time Pothole Detection Approach for Intelligent Transportation System,” Mathematical Problems in Engineering, vol. 2015, pp. 1–7, 2015, doi: 10.1155/2015/869627.

[3]S. Yamuna, R. Pavithra, and M. Ranjitha, “Automatic Pothole Detection System using Laser,” International Journal for Research in Applied Science and Engineering Technology, vol. V, no. III, pp. 1226–1231, Mar. 2017, doi: 10.22214/ijraset.2017.3226.

[4]C. C. Aggarwal, Neural networks and deep learning: a textbook. New York: Springer, 2019.

[5] Jo, Y., Ryu, S.-K., & Kim, Y.-R. (2016). Pothole Detection Based on the Features of Intensity and Motion. Transportation Research Record, 2595(1), 18–28. https://doi.org/10.3141/2595-03

[6]S. Silvister, D. Komandur, S. Kokate, and A. Joshi, “Deep Learning Approach to Detect Potholes in Real-Time using Smartphone,” IEEE Pune Section International Conference (PuneCon), 2019.

[7]1Alfandino Rasyid et al., “Pothole Visual Detection using Machine Learning Method integrated with Internet of Thing Video Streaming Platform *,” 2019 International Electronics Symposium (IES), 2019.

[8]R. Sadiq, M. Rodriguez, and H. Mian, “Empirical Models to Predict Disinfection By-Products (DBPs) in Drinking Water: An Updated Review,” www.sciencedirect.com, vol. 2, pp. 324–338, Jan. 2019, doi: 10.1016/B978-0-12-409548-9.11193-5.

[9]S. Abbasi, M. Hajabdollahi, N. Karimi, and S. Samavi, “Modeling Teacher-Student Techniques in Deep Neural Networks for Knowledge Distillation,” Jul. 2019. Accessed: Oct. 27, 2021. [Online].

[10]C. Xi, X. Zhiqiang, and C. Yuyang, “Introduction to Model Compression Knowledge Distillation,” *2021 6th International Conference on Intelligent Computing and Signal Processing (ICSP)*, Apr. 2021, doi: 10.1109/icsp51882.2021.9408881.

[11]C.-W. Kuan, W.-H. Chen, and Y.-C. Lin, “Pothole Detection and Avoidance via Deep Learning on Edge Devices,” Jun. 2021, Accessed: Nov. 03, 2021. [Online].

[12]N. Camilleri and T. Gatt, “Detecting road potholes using computer vision techniques,” May 2020, Accessed: Nov. 03, 2021. [Online].

[13]Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015, doi: 10.1038/nature14539.

[14]Y. Cheng, D. Wang, P. Zhou, and T. Zhang, “Model Compression and Acceleration for Deep Neural Networks: The Principles, Progress, and Challenges,” *IEEE Signal Processing Magazine*, vol. 35, no. 1, pp. 126–136, Jan. 2018, doi: 10.1109/msp.2017.2765695.

[15]M.-C. Wu and C.-T. Chiu, “Multi-teacher knowledge distillation for compressed video action

recognition based on deep learning," *Journal of Systems Architecture*, vol. 103, p. 101695, Feb. 2020, doi: 10.1016/j.sysarc.2019.101695.

[16]H. Ni, J. Shen, and C. Yuan, "Enhanced Knowledge Distillation for Face Recognition," *2019 IEEE Intl Conf on Parallel & Distributed Processing with Applications, Big Data & Cloud Computing, Sustainable Computing & Communications, Social Computing & Networking (ISPA/BDCloud/SocialCom/SustainCom)*, Dec. 2019, doi: 10.1109/ispa-bdcloud-sustaincom-socialcom48970.2019.00207.

[17]J. Yim, D. Joo, J. Bae, and J. Kim, "A Gift from Knowledge Distillation: Fast Optimization, Network Minimization and Transfer Learning," 2017, Accessed: Nov. 04, 2021. [Online].

[18]Z. Allen-Zhu, M. Research, and R. Li, "Towards Understanding Ensemble, Knowledge Distillation and Self-Distillation in Deep Learning," 2020. Accessed: Nov. 04, 2021. [Online].

[19]G. Hinton, O. Vinyals, and J. Dean, "Distilling the Knowledge in a Neural Network," 2015. Accessed: Nov. 15, 2021. [Online].

[20]V. Gutierraz Cruz, "Global number of connected IoT devices 2015-2025," Statista, Mar. 2021. https://www.statista.com/statistics/1101442/iot-number-of-connected-devices-worldwide/.

[21]E. Onyinyechi, M. Hamada, S. Yusuf, and M. Hassan, "The Role of Linear Discriminant Analysis for Accurate Prediction of Breast Cancer," Dec. 2021, doi: 10.1109/MCSoC51149.2021.00057.

[22] M. Hamada, H. Kakudi, J. Tanimu, P. Robert, and M. Hassan, "Evaluation of Recursive Feature Elimination and LASSO Regularization-based optimized feature selection approaches for cervical cancer prediction," presented at the 14th IEEE International Symposium on Multi core many core System on chips, Singapore, 2021, doi: 10.1109/MCSoC51149.2021.00056.

[23] A. Musa, M. Hamada, F. Aliyu, and M. Hassan, "An Intelligent Plant Dissease Detection System for Smart Hydroponic using Convolutional Neural Network," presented at the 14th IEEE International Symposium on Multi core many core System on chips, Singapore, 2021, doi:10.1109/MCSoC51149.2021.00058.