❒ 1204

Pothole recognition using convolution neural networks and transfer learning

Chayadevi Senigalakuruba, Suraj Pabba
[1]Department of Information Technology Vasavi College of Engineering, Telangana, India

Article Info	ABSTRACT
	Potholes have been and still are a huge problem for every walk of life. There are many deaths and accidents reported daily due to that very problem. For that reason, pothole recognition comes into the picture. To maintain and preserve a road, it is vital to detect potholes. It also helps in the prevention of accidents. Roads play an important part in day-to-day transportation for every person around the world. But the quality of roads decreases drastically due to the way of usage and aging. The existing methods take much time and manpower to repair the damaged areas. The entire process is slowing down just because an expert team is checking whether there is a pothole at the reported location or not. So, if we automate the process of detection of potholes from a set of images reported from a particular location and appropriately alerting the authorities with the amount of damage, the process speeds up exponentially. We must solve the major problem of pothole recognition by using machine learning algorithms. This paper will discuss a convolution neural network-based and a transfer learning-based solution for pothole recognition.

Corresponding Author:

Chayadevi Senigalakuruba
Department of Information Technology, Vasavi College of Engineering
Telangana, India
Email: skchayadevi@staff.vce.ac.in

1. INTRODUCTION

Traffic congestion, as well as brain strain, are both on the rise, increasing the risk of accidents [1]. Our current transportation model relies heavily on road and highway access. Individuals are transported by personal vehicles to and from their destinations. It is more than just a public nuisance when poor road conditions cause discomfort for drivers and passengers. It also damages vehicles and can result in tragic accidents. Road traffic accidents are affected by many factors [2]. Potholes on roads become a very big problem when the vehicles are running at higher speeds [3]. Public Works Departments (PWDS) and private contractors work together to assess and repair city roads, a complicated process. Traffic congestion is becoming a worldwide problem as the number of motor vehicles on roads and population density are increasing [4]. India lost nearly 3% of its gross domestic product (GDP) to road traffic accidents or 58,000 million USD in absolute terms, according to a UNESCAP study [5]. Injuries due to road accidents cost around 518 billion USD every year globally [6]. A person's social contacts are also affected by an accident, in addition to his or her own injuries [7]. Every day, six Indians were killed by potholes in 2016. Besides the fact that it may involve people's lives, it can also have a detrimental effect on socio-economic development [8]. An accident on the road can cause many types of damage, including indirect damage to infrastructure, direct damage to human beings, and many other kinds of administrative damage [9]. Potholes are linked to dozens of deaths that go unreported, so the actual number may be higher. Based on data shared with the Centre by states, Andhra Pradesh, Uttar

Pradesh, Gujarat, West Bengal, Maharashtra, Kerala, and Odisha are among the Top 10 states with the highest death rate due to potholes in India. The number of roads paved in India was 63.24% as of March 31st, 2017. In India, between 2015 and 2017, over 9,300 deaths have been attributed to potholes since 2015. Three thousand five hundred and seventy-nine people were killed by potholes on Indian roads in 2017, while twenty-five thousand were injured. Hence, roads must be constructed, maintained, and constructed with vehicle safety in mind There are already existing methods to solve this problem like image segmentation-based edge detection methods, AlexNet based methods, you only look once (YOLO) based methods as well. But the methods discussed are of better version and level when compared to them.

Potholes could be identified more quickly and more efficiently if the process of identifying them was automated. As a result, more time and money could be spent on fixing the roads instead of identifying them. As discussed in Goldberg [10] automation emphasizes more on quality. An overview of convolutional neural networks (CNNs) and their applications in image recognition is presented in the first part of this paper. CNN's can be applied in many different ways possible; they can be pre trained as discussed in Abdullah and Hasan [11] or deep CNN's can also be used as discussed in Yang and Li [12] for image processing and recognition. We will then move on to the use of transfer learning with Inception V3 and Imagenet in the second part of this paper. As suggested in Degadwala et al. [13] the inception V3 along with transfer learning can be used for image captioning.

CNN's are a category of neural networks which are apt to work with images and predict objects in images [14]. Neural networks are neuron-based systems that use neurons to create outputs based on inputs. It is possible for "neurons" to be called nodes, and each node is a weighted input taken from the presiding nodes or from inputs to determine output. To make a good decision, each weight within "neurons" must be trained. A CNN is a type of deep-feedforward neural network that makes use of spatial relationships of data to recognize images as they use deep learning on large datasets. For a detailed explanation of CNN's refer Albawi et al. [15]. Machine learning method called transfer learning takes a model developed for one task and uses it to start a new model on a second one. Due to the resources required to develop neural network models and the amazing leaps in intelligence that these machines provide, the use of pre-trained neural networks for computer vision and natural language processing has become increasingly popular in deep learning.

We went through multiple papers where people used different methods to approach the problem. Few papers used CNN with Alexnet by Srinidhi and SM [16], YOLO by Yik et al. [17] to solve the issue at hand, while others used non-machine learning methods such as edge detection to detect anomalies in the road. While the methodology used is either not better, or the results that they have obtained are not very high.

Based on the fundamental properties of potholes, the paper by Nienaber et al. [18] discusses an algorithmic approach to detecting potholes without requiring any machine learning. As the approach is visual, it is clear that the solution will depend on lighting, angle or point of view, and other factors that obstruct the view of potholes. Pothole detection begins with grayscale conversion of the road section image. In order to remove noise from this grayscale image, a Gaussian filter was applied. The road surface is then extracted using differentiation-based Canny edge detection. Danti et al. [19] have created a model which detects and recognizes potholes, black clusters were used to extract potential regions hence adding pothole detection to the recognition part. Douanghphanah and Oneyama [20] discussed road roughness by using fast fourier transform (FTT) in frequency space, it analyzes the effect of speed devices and has classified them by the roughness index. The papers have given us a brief knowledge of what all algorithms can be applied to obtain the expected results. They have helped with what to do and what not to. There are some algorithms in Kaggle as well where different models are created to output whether there is a pothole or not.

The models mentioned in this paper are better than the existing models in the literature as we are using a bigger dataset and have got higher accuracies when compared. The models discussed in this paper are faster both in training and predicting the potholes. We are also using a new approach known as transfer learning which is gives better results than the existing models in the literature.

2. METHOD
2.1. Dataset
Data set taken from internet. The title of the dataset is "Pothole Detection" and is located at [21]. The total number of images in data set are 1,440, a sample image is as referred to in the Figure 1. While training the model data set is divided in to desired 3 sets like training, testing and validation. Depends on occurrence and nonoccurrence of pot hole in the image all the 1,440 have been divided. Size of the images resized to 64×64, for enhancing accuracy, size of the image is considered by emphatical method.

2.2. Convolutional neural network
CNN is better at identifying a face, objects, and traffic signboards. It is better at the same because it extends deep learning algorithms [22], [23]. The dataset contains images of 64×64 size, we have resized the

images to 256×256 for achieving better results. We have made sure that the size of the images remains the same throughout so that the process of classifications becomes easy. This brought uniformity to the images. Throughout this research, several CNN models have been created and experimented with. A majority of these models utilize convolutional, ReLu, average pooling and layer-by-layer connections. Only notable models will be discussed here.

Figure 1. Pothole image

Architecture, the main model as referred to in the Figure 2. created with the Keras library in Python. To include dataset into designed network, we have used the ImageDataGenerator function present in Keras. This operation creates a "usable copy" of the image for the model to use for training and testing. Numerous CNN 2D layers were used in the model, have 10 filters with (SAF) Sigmoid activation function. The steps have changed (1,1), (3,3), (5,5). Soft max function used at the output layer". the metrics used is "Accuracy". The model consists of a total of 7 CNN layers, Model run for about 50 times and got a 94% testing accuracy. We found some anomalies in the detection of images as well, at times images that have potholes are not detected and at times images without potholes are detected as images with potholes. Initially, we have checked on the algorithm whether it has any faults in it but later found out that it is the images that are creating the problems, mainly it is the disturbances in those images. Disturbances such as splashes of water which partially cover the potholes, blurry images with oil marks on the roads which are generally circular in shape are being noticed unwantedly. Such images were deteriorating the accuracy but we have removed such images with disturbances so that we achieve better results.
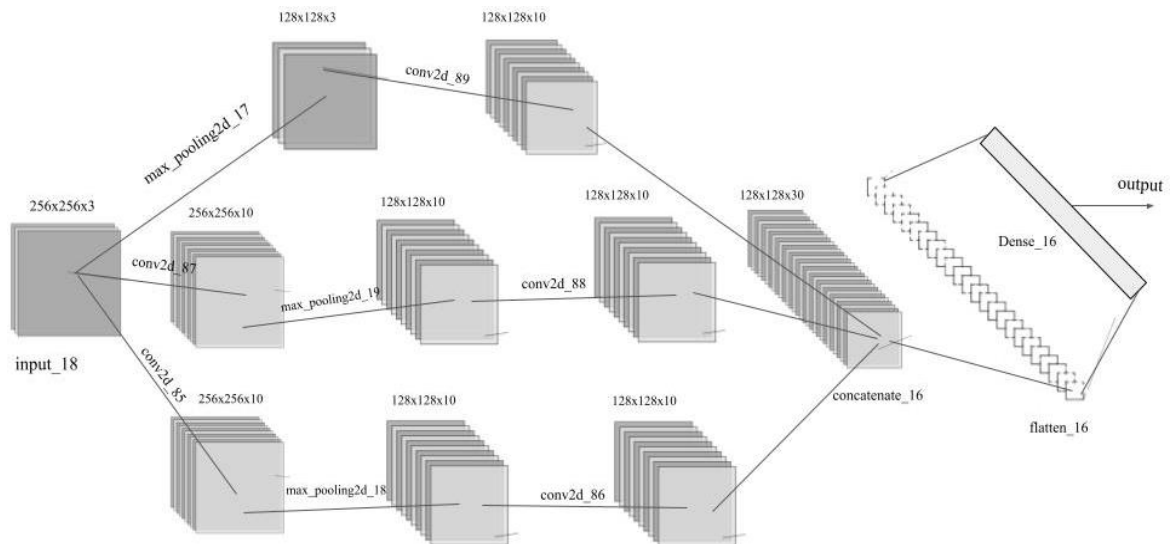

Figure 2. CNN model architecture

2.3. Inception V3

Data Preparation As the pre-trained model only takes images in (224,224,3) format we have augmented all the images into the said size from 64×64 original size as these images is going to run based on a pre-trained model the augmentation of the images made a huge difference in the final results obtained. Data

is pre-processed using the preprocess input () function present in Kera's. Hence, this function when imported from Keras automatically pre-processes all the images present in the dataset to the sizes as required by the inception-v3 model.

Architecture Inception-V3 is a variation of CNN, consists of 48 inception layers deep CNN as referred to in the Figure 3. Each inception layer contains of a 1×1, a 3×3, and a 5×5 layer where all their outputs are combined into a single input vector Training model uses Transfer learning. During training architecture on Google's image-net [24], [25] weights were changed. The entire architecture initialized with the weights of image-net, except output and input layers. The format of the input is 224(width):224(height):3 (RGB) and the format of the output is a vector of 2 rows. If pot hole is present first row is '1', if no pot hole second row vector is '1'. The model consists a total of 4098 trainable params. Model is initialized with weights from image -net data set. The knowledge found in categorizing image-net dataset was used to order pothole images. last layers are retained, have a higher implication in finding object classes with a very sluggish learning rate to get best accuracy. Retraining method has helped in achieving better results, which will be further discussed. One reason behind the increase in the accuracy might be freezing and retraining of the model, the other reason might be the adaptability of the model towards the new images which are specific to pothole detection. The improvement might be very small, but in such models, even small changes make a huge difference. People are still working on such algorithms so that they, even more, improve the accuracy of the algorithms which use these pre-trained models. The image depicts the entire model of inception V3.
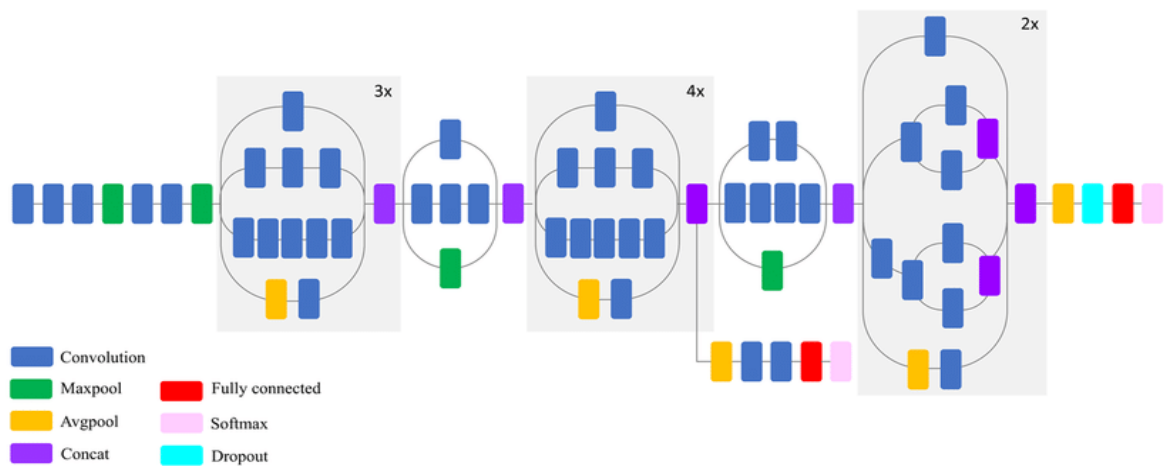

Figure 3. InceptionV3 model architecture

3. RESULTS AND DISCUSSION

Throughout the algorithm development, Kaggle data set and other images from internet were used. As referred to in Table 1 results of both the models were listed. CNN model has given better results when compared to transfer learning but, as the Inception V3 model was pre-trained with many images from image-net the results obtained from that model are better than the CNN model results. The inception v3 algorithm is trained on around 2,330 images, 1,000 being generic images from ImageNet and the rest being related to pothole detection, whereas the CNN model is trained on 1,440 images which are very specific to the sole cause of detecting potholes. The results would even improve if the size of the datasets is increased or the versatility of images is increased like from grayscale to RGB, small-sized images to large.

As already mentioned, we have found some anomalies as referred to in Figure 4, we found that such images are becoming an anomaly as there is no proper view of the pothole in it. Figure 5 is an image, which is perfectly detected, we have made sure that the image is not taken from the dataset and is not of the same size as the dataset images, because such images are the ones that we generally find or are needed to test in our day-to-day life.

Table 1. Comparison of accuracies

Network	Training Accuracy	Training Loss	Testing Accuracy	Testing Loss
CNN	0.96	0.02	0.94	0.032
Inception V3 BEFORE FT (fine tune)	0.96	0.16	0.912	0.2
Inception V3 AFTER FT (fine tune)	0.95	0.16	0.92	0.2

Figure 4. Anomalous image from the dataset   Figure 5. Image recognized as a pothole

4.   CONCLUSION

The proposed convolutional neural network model and inception V3 model have performed well and are ready to predict the potholes from the images. The models can be used by the government authorities also known as public work Departments. They can reduce the time taken to recognize them and the cost put in as well. The models have been worked on so perfectly so that we get as accurate to recognize potholes as we can. The important part of this entire process is entirely dependent on the dataset and the images in it, mainly the sizes of the images are important. There should be uniformity among all the images used so that they are perfectly trained and validated. We have found some anomalies in the testing phase where small-sized potholes are at times ignored. From the results obtained, we can confirm that the Inception V3 based Image net model has performed better as it has been trained on a larger and flexible dataset when compared to the CNN model. The CNN model also has produced comparatively better results from existing models.

5.   FUTURE SCOPE

The performance of the models can be improved by using huge, sophisticated datasets with uniform sizes containing different sizes of potholes. The performance can also be improved by using the datasets from various governments from different countries and continents, either old or new datasets will surely improve the accuracy. The results are better enough to be integrated into a project wherein potholes are to be recognized or detected, say we can use a unmanned aerial vehicle (UAV) to detect the potholes. With project integration, we can even increase datasets by capturing the real data in and around us. There are many more avenues to approach the problem and solve it in an even much better way as in we can apply all the CNN models. There are also many more areas into which we can integrate the pothole recognition and detection solution.

REFERENCES

[1]   H. T. Manjunatha and A. Danti, "Detection and classification of potholes in Indian roads using wavelet based energy modules," *2019 International Conference on Data Science and Communication, IconDSC 2019*, 2019, doi: 10.1109/IconDSC.2019.8816878.
[2]   A. Ditcharoen, B. Chhour, T. Traikunwaranon, N. Aphivongpanya, K. Maneerat, and V. Ammarapala, "Road traffic accidents severity factors: A review paper," *Proceedings of 2018 5th International Conference on Business and Industrial Research: Smart Technology for Next Generation of Information, Engineering, Business and Social Science, ICBIR 2018*, pp. 339–343, 2018, doi: 10.1109/ICBIR.2018.8391218.
[3]   Y. Li, C. Papachristou, and D. Weyer, "Road pothole detection system based on stereo vision," *Proceedings of the IEEE National Aerospace Electronics Conference, NAECON*, vol. 2018-July, pp. 292–297, 2018, doi: 10.1109/NAECON.2018.8556809.
[4]   S. S. Rode, S. Vijay, P. Goyal, P. Kulkarni, and K. Arya, "Pothole detection and warning system: Infrastructure support and system design," *Proceedings - 2009 International Conference on Electronic Computer Technology, ICECT 2009*, pp. 286–290, 2009, doi: 10.1109/ICECT.2009.152.
[5]   S. Sonal and S. Suman, "A framework for analysis of road accidents," *2018 International Conference on Emerging Trends and Innovations In Engineering And Technological Research, ICETIETR 2018*, 2018, doi: 10.1109/ICETIETR.2018.8529088.
[6]   J. Patil, M. Prabhu, D. Walavalkar, and V. B. Lobo, "Road accident analysis using machine learning," *2020 IEEE Pune Section International Conference, PuneCon 2020*, pp. 108–112, 2020, doi: 10.1109/PuneCon50868.2020.9362403.
[7]   J. R. Asor, G. M. B. Catedrilla, and J. E. Estrada, "A study on the road accidents using data investigation and visualization in Los Baños, Laguna, Philippines," *2018 International Conference on Information and Communications Technology, ICOIACT 2018*,

vol. 2018-January, pp. 96–101, 2018, doi: 10.1109/ICOIACT.2018.8350662.

[8] E. Johnson, J. M. Abraham, S. Sulaiman, L. Padma Suresh, and S. Deepa Rajan, "Study on road accidents using data mining technology," *Proc. IEEE Conference on Emerging Devices and Smart Systems, ICEDSS 2018*, pp. 250–252, 2018, doi: 10.1109/ICEDSS.2018.8544370.

[9] E. J. Reddy, P. N. Reddy, G. Maithreyi, M. B. C. Balaji, S. K. Dash, and K. A. Kumari, "Development and analysis of pothole detection and alert based on nodeMCU," *International Conference on Emerging Trends in Information Technology and Engineering, ic-ETITE 2020*, 2020, doi: 10.1109/ic-ETITE47903.2020.347.

[10] K. Goldberg, "What is automation?," *IEEE Transactions on Automation Science and Engineering*, vol. 9, no. 1, pp. 1–2, 2012, doi: 10.1109/TASE.2011.2178910.

[11] Abdullah and M. S. Hasan, "An application of pre-trained CNN for image classification," *20th International Conference of Computer and Information Technology, ICCIT 2017*, vol. 2018-January, pp. 1–6, 2018, doi: 10.1109/ICCITECHN.2017.8281779.

[12] J. Yang and J. Li, "Application of deep convolution neural network," *2016 13th International Computer Conference on Wavelet Active Media Technology and Information Processing, ICCWAMTIP 2017*, vol. 2018-February, pp. 229–232, 2017, doi: 10.1109/ICCWAMTIP.2017.8301485.

[13] S. Degadwala, D. Vyas, H. Biswas, U. Chakraborty, and S. Saha, "Image captioning using inception V3 transfer learning model," *Proceedings of the 6th International Conference on Communication and Electronics Systems, ICCES 2021*, pp. 1103–1108, 2021, doi: 10.1109/ICCES51350.2021.9489111.

[14] S. K. Roy, G. Krishna, S. R. Dubey, and B. B. Chaudhuri, "HybridSN: Exploring 3-D-2-D CNN feature hierarchy for hyperspectral image classification," *IEEE Geoscience and Remote Sensing Letters*, vol. 17, no. 2, pp. 277–281, 2020, doi: 10.1109/LGRS.2019.2918719.

[15] S. Albawi, T. A. Mohammed, and S. Al-Zawi, "Understanding of a convolutional neural network," *Proceedings of 2017 International Conference on Engineering and Technology, ICET 2017*, vol. 2018-January, pp. 1–6, 2018, doi: 10.1109/ICEngTechnol.2017.8308186.

[16] G. Srinidhi and R. D. SM, "Pothole detection using CNN and AlexNet," *SSRN Electronic Journal*, 2020, doi: 10.2139/ssrn.3648822.

[17] Y. K. Yik, N. E. Alias, Y. Yusof, and S. Isaak, "A real-time pothole detection based on deep learning approach," *Journal of Physics: Conference Series*, vol. 1828, no. 1, 2021, doi: 10.1088/1742-6596/1828/1/012001.

[18] S. Nienaber, M. J. (Thinus) Booysen, and R. Kroon, "Detecting potholes using simple image processing techniques and real-world footage," *Proceedings of the 34th Southern African Transport Conference (SATC 2015)*, pp. 153–164, 2015.

[19] A. Danti, J. Y. Kulkarni, and P. S. Hiremath, "An image processing approach to detect lanes, pot holes and recognize road signs in Indian roads," *International Journal of Modeling and Optimization*, pp. 658–662, 2012, doi: 10.7763/ijmo.2012.v2.204.

[20] V. Douangphachanh and H. Oneyama, "A study on the use of smartphones for road roughness condition estimation," *The Journal of Eastern Asian Society for Transportation Studies*, vol. 30, no. 5, pp. 1551–1564, 2016, doi: 10.11175/easts.10.1551.

[21] A. Kumar, "Pothole detection dataset | Kaggle," *Kaggle Datasets*, 2020, [Online]. Available: https://www.kaggle.com/atulyakumar98/pothole-detection-dataset.

[22] R. He, X. Wu, Z. Sun, and T. Tan, "Wasserstein CNN: Learning invariant features for NIR-VIS face recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 41, no. 7, pp. 1761–1773, 2019, doi: 10.1109/TPAMI.2018.2842770.

[23] C. Szegedy *et al.*, "Going deeper with convolutions," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 07-12-June-2015, pp. 1–9, 2015, doi: 10.1109/CVPR.2015.7298594.

[24] G. Pan, L. Fu, R. Yu, and M. Muresan, "Evaluation of alternative pre-trained convolutional neural networks for winter road surface condition monitoring," *ICTIS 2019 - 5th International Conference on Transportation Information and Safety*, pp. 614–620, 2019, doi: 10.1109/ICTIS.2019.8883540.

[25] G. Pan, M. Muresan, R. Yu, and L. Fu, "Real-time winter road surface condition monitoring using an improved residual cnn," *Canadian Journal of Civil Engineering*, vol. 48, no. 9, pp. 1215–1222, 2021, doi: 10.1139/cjce-2019-0367.

BIOGRAPHIES OF AUTHORS

Dr. Senigalakuruba Chayadevi [ID] [G] [SC] [O] an associate professor at Vasavi College of Engineering with an experience of 20 years. I received my Ph.D from JNTU in the year 2020 under Electronics And Communication Engineering specialization. Completed my M.tech from JNTU in the year of 2009 under Digital systems and Computer Electronics as specialization and my B.Tech from NIT Warangal in the year 1992 from ECE Department. I have published more than 15 journals as of date and hope to keep publishing. She can be contacted at email: Skchayadevi@staff.vce.ac.in,

Mr. Suraj Pabba [ID] [G] [SC] [O] a student from Vasavi College of Engineering from the Information Technology Department completed B.Tech in the year 2021. I have completed my intermediate in the year 2017 from Narayan junior college and schooling in the year 2015 from St. Patrick's High School. I have high interest in the areas like Machine learning, Artificial intelligence and robotics. I am currently working as a system engineer at TCS hyderabad. He can be contacted at email: surajpabba1947@gmail.com.