Convolutional Neural Network for Roadside Barriers Detection: Transfer Learning versus Non-Transfer Learning

Mahdi Rezapour * and Khaled Ksaibati

Wyoming Technology Transfer Center, 1000 E University Ave, Dept. 3295, Laramie, WY 82071, USA; khaled@uwyo.edu
* Correspondence: mrezapou@uwyo.edu; Tel.: +1-(307)-766-6230; Fax: +1-(307)-766-6784

Abstract: Increasingly more governmental organizations in the U.S. have started to implement artificial intelligence to enhance the asset management process with an objective of controlling the costs of data collection. To help the Wyoming Department of Transportation (WYDOT) to automate the data collections process, related to various assets in the state, an automated assets management data collection was proposed. As an example, the automated traffic barriers asset dataset would collect geometric characteristics, and barriers’ materials’ conditions, e.g., being rusty or not. The information would be stored and accessed for asset-management-decision-making and optimization process to fulfill various objectives such as traffic safety improvement, or assets’ enhancement. For instance, the State of Wyoming has more than a million feet of roadside barriers, worth more than 100 million dollars. One-time collection of various characteristics of those barriers has cost the state more than half a million dollars. Thus, this study, as a first step for comprehensive data collection, proposed a novel approach in identification of roadside barrier types. Pre-trained inception v3, denseNet 121, and VGG 19 were implemented in this study. Transfer learning was used as there were only 250 images for training of the dataset for each category. For that method, the topmost layers were removed, along with adding two more new layers while freezing the remaining layers. This study achieved an accuracy of 97% by the VGG 19 network, training only the few last layers of the model along with adding two dense layers for top layers. The results indicated that although there are not enough observations related to traffic barrier images, a transfer learning application could be considered in data collection. A simple architecture non-transfer model was also implemented. This model achieved an accuracy of 85%, being better that the two other transfer learning techniques. It should be reiterated that although non-transfer learning technique outperformed inception and denseNet networks, it comes short significantly when it come to the VGG network.

Keywords: data collection; transfer learning; image recognition; convolutional neural network; traffic barriers

1. Introduction

A well-functioning and maintained transportation infrastructure is a key to roadway users’ safety and comfort. The United States has invested more than 1 trillion dollars in its highway infrastructure in the past 100 years [1]. There are approximately 4 million miles of roadway system in the nation [2], and In the State of Wyoming alone, there is a total of more than 27,000 miles of roadways owned and maintained by federal, state, and local entities.

The conditions of these assets are determined regularly by the Wyoming Department of Transportation (WYDOT) to decide whether the assets need to be maintained, replaced, or enhanced to meet safety, comfort and convenience of road users. The first step in fulfilling the above objective is to collect data. For roadway barriers, for instance, data should include an accurate and complete inventory of various aspects of traffic barriers such as: barrier types, geometric characteristics, locations of those barriers (e.g., highway system and mile post), and condition assessment (e.g., damaged, rusted, or under designed). This
is because barriers could pose a high crash severity in case of being below standard and as much of these assets are outdated in the state in terms of geometric characteristics such as heights, they could result in a high severe crashes. For instance, it has been discussed that recommended heights of guardrail barriers should be greater than 27 and less than 35 inches to prevent override and underride barriers’ crashes, respectively [3].

The mountainous state of Wyoming has more than 1 million feet of traffic barriers in the state. Optimization of the limited resources of the state department of transportation (DOTs) is of paramount importance [4]. WYDOT is working to maintain and upgrade its assets in a most efficient and cost-effective way within their limited budget. Due to the complexity of asset management data collection in the state, a project has been initiated to optimize data collection cost. The following stages, in order, were identified to automate the data collection of roadside barriers in Wyoming:

1. Pilot study: Roadside barriers identification based on available barriers data set
   a. Distinguish various barriers from one another.
   b. Distinguish rusty barriers from well-maintained barriers.
   c. Identify damaged barriers from other barriers.
2. Measure barriers dimensions based on available dataset.
3. Implement the technique on other roadway assets.

Usually deep convolutional neural network (DCNN) has been conducted in the literature on large amount of labeled dataset. However, with the help of transfer learning, the advantage of using that method for a small number of labeled datasets will be demonstrated in this study. A very small number of observations related to various barrier types were used to evaluate the application of transfer learning for barrier types’ detection. In addition, a simple DCNN with few layers of CNN was used as a base model for a comparison.

In summary, the contributions of this study are as follows:

1. For the first time, images related to the roadside barriers were used for object detection, with applications of various transfer and non-transfer learning.
2. Although this study incorporates very limited number of images, about 250 images for each category, the model reaches a high accuracy with the help of transfer learning.
3. At the end, a model with a very simple architecture, with 4 convolutional layers was trained on images. It was found that although the model performs worse than VGG 19, it outperforms the other transfer learning techniques.
4. Confusion matrix and some statistics were presented about a best performed model to give some insights about the model performance.
5. Extensive funding has been spent on collection and measurement of the dimensions of various roadside barriers in Wyoming. After successful implementation of the proposed techniques and after conducting the next phases of this project, the trained models on new dataset could be used for automated barrier geometric measurements so a significant data collection funding could be saved.

2. Literature Review

The algorithms have been implemented for solving complex optimization problems in the state-of-the-art of day-to-day life not only on image processing but in other areas as well: From feature selection [5], to self-adaptive evolutionary algorithm for the berth scheduling problem [6]. The objective of this study is tied with object detection studies: a sector of computer vision related to detecting instances conducted in the literature review. The section is structured as follows: The first subsection would go over several studies that used various objects for image identification, and then, the second subsection would present application of convolutional neural network (CNN) methods in other fields. A vision-based system is a method that has been used widely for detecting roadway environment for various objectives such as road safety improvement. Low-overhead collision algorithm was presented, which was able to perform successfully in real time under various weather conditions [7]. In another study, a system is proposed relying on cellular communication to
detect probable vehicle to pedestrian crashes [8]. The results showed promising results on potential of safety improvement.

The CNN technique has become one of the state-of-the-art techniques for object detection in the field of machine learning. For instance, a study conducted to evaluate the pedestrian localization accuracy with the help of convolutional neural network model [9]. This system displays the driver real-time locations, and also types of hidden road users at traffic intersection. The data collected and transferred to roadside unit consists of a monovision camera streaming videos to a computing unit, which perform analysis to detect objects and measure distances. The CNN was used in another study for roadside vegetation detection [10]. The data contains vegetation in various conditions.

Detecting various traffic signs was achieved by the CNN [11]. This analysis conducted on Malaysian traffic sign database with two types of training: Incremental and batch training. The results indicated that the analysis could reach 99% accuracy. The extensive research also has been conducted for evaluation of the CNN in other fields, which the following paragraphs would outline.

A study conducted for traffic scene classification with an objective of improvement of the vehicle environment [12]. They highlighted how a limiting size of image descriptors impact the traffic scene performance classification.

High resolution images of synthetic aperture radar (SAR) was used [13]. Different pertained CNN models were used as feature extractors in combination with support vector machine (SVM). AlexNet performed better than other considered methods including VGG and GoogleNet.

Thyroid ultrasound images were taken including the total of 1037 thyroid ultrasound images, including two categories: Benign and malignant images [14]. The images were resized to 224 × 224. Due to limited number of images, a pre-trained CNN learned from ImageNet was also considered. A pre-trained VGG model for transfer learning was compared with feature subset selection and classification, which was based on extracting the features from CNN. The results indicated that proposed method perform better than the transfer learning technique.

Two-phase method combining CNNs transfer learning and web data augmentation was proposed [15]. The original data was augmented with the internet images, expanding the training dataset. The method was proposed when the training data is very limited. The web data augmentation was used to add diversity to the datasets with the purpose of generalization ability and alleviating the overfitting. Various transfer learning techniques were implemented on various datasets. For instance, traditional models outperformed the considered transfer learning of Alexnet, VGG-16, ResNet on events dataset.

The transfer learning with CNN for abdominal ultrasound images were used [16]. The images were categorized into 11 categories based on their text annotation. Also, irrelevant images were excluded. The number of images for 11 categories range from 137 to 576. The pre-trained model based on CaffeNet and VGGNet were used. The accuracies of the models for CaffeNet and VGGNet were 77.3% and 77.9% respectively. In another study, a deep learning-based solution for large-scale extraction of the secondary road was used [17]. The transfer learning was employed to start from pre-trained weights instead of random initialized weights.

It should be noted that most of the studies in the literature review used considerable amount of images for training various application of transfer or nontransfer learning. Also, to the best knowledge of the authors of this study, no study in the literature review used traffic barriers as the image dataset. Thus, this study was conducted with the help of transfer learning, and a model with a simple architect, to classify various traffic barriers with a very low frequency. The findings of this study would contribute to future automated data collection to assist asset management system in the state. This project would be extended for various traffic barrier dimensions measurements.
2.1. Dataset

The traffic barrier inventory was collected from 2016 to 2017 as part of the WYDOT traffic barrier inventory project no. HPR1217 [18]. In this period over 1.3 million linear feet of traffic was measured along with approximately 4176 miles of roadway. Also, the study incorporated 7700 images, which were taken during the field survey. The images were taken by Trimble GeoExplorer 6000 series.

The objective of that field analysis was to measure various roadway and traffic barriers' geometric features like barriers' height, length, offset, and road slope rate. The collected data included 55% of all non-interstate roads of Wyoming [18]. After filtering the dataset to include only relevant images, more than 2000 very high-quality images were included in the analysis. The images were in 25,952 × 1944 pixels, and with horizontal/vertical dots per inch (DPI) of 192.

Some of images used for model training, as an example, are presented in Figure 1. Although most of the images are self-explanatory, a description should be given for hybrid barriers. The bottom two images, images g and h in Figure 1, are related to hybrid barriers. Hybrid barriers are a combination of guardrail and box beam barriers being implemented when a guardrail barrier transfer to a box beam due to being located at the bridge. Those barriers would be located on a step of a concrete to assure enough height for preventing vehicles for any possible override over the bridge. Images e and f are related to concrete barriers but they were taken from a bad position.

Figure 1. Cont.
Figure 1. Some of the images used in this study, from top rows down: Box beam (a,b), cable barrier (c,d), concrete barrier (e,f), and a hybrid (g,h).

The original image sizes were very large and even small frequency of those high-quality images would be very expensive in terms of time, resources, and storage for a model to be trained. Thus, all images were resized to 150×150 for training, validating and testing dataset. This resizing should not be much of a concern regarding the methodological approach due to the following:

- Most part of the background images present information unrelated to the barriers, e.g., sky, grass, and thus would be not beneficent to the training.
- CNN by itself would pick up important information through filters. For instance, barriers consists of simple geometric characteristics, vertical lines would be picked by 3×3 filter of \( \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ -1 & -1 & -1 \end{bmatrix} \) and horizontal line by \( \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \). Thus, it can be assumed using original or resizing the Images would not make much of a difference,
- From practical point of view, it would take a tremendous time for the models to be trained if large images were used.
- The objective was to compare across various methods and as we used similar size across the methods, the size would be of no issue.

The Figure 2 presents some of the resized imaged in Figure 1 for a comparison.
As the primary objective of this study was to identify various barrier types based on images, images were manually labeled. These images were classified into 4 categories, box beam, cable barriers, hybrid and concrete barriers. The frequencies of these images are presented in Table 1. As noted from Figure 1, although there were instances where it was hard to read the images, but those images were included in the evaluation model. For instances, this could be observed from images d, cable barrier, and f, concrete barrier, in Figure 1, which are not distinguishable easily by an ordinary individual. Out of 316 images for each category, 250, 56, and 10 images were used for training, validation and test dataset, respectively. The percentages of validation, test and training are based on the investigators discretion. Normally 8% of images should be reserved for validation and 82% for training. Training is an important part of our methodological steps. So, we allocate a high number to this category, especially because our data suffer from low number of observations.

Table 1. Distribution of various barrier types.

| Type of Images          | Frequency |
|-------------------------|-----------|
| Cable barriers          | 316       |
| Box beam                | 491       |
| Concrete barriers       | 322       |
| Hybrid of box beam and guardrail | 448 |

The description of the filtered image data frequencies, being considered in the analysis, is presented in Table 1. As can be seen from Table 1, the dataset is slightly imbalanced across various barrier types and there are especially fewer observations for cable barriers. Most of the machine learning techniques work in favor of overrepresented categories. Thus, to address this issue and to make the size of four classes comparable, 316 images were selected randomly based on the lowest category frequency being related to cable image category. As a result, there were 316 images for each category used in this study.

2.2. Methodology

This section would first briefly talk about the methodological process/steps taken for applying the model. Then, it would outline the transfer learning techniques and the three models implemented in this study. It then would discuss the architecture of the non-transfer learning method implemented as a base method.

3. Process

There are few important steps in image processing which the next few lines would outline. The method was implemented in Keras with TensorFlow backend in R. Images
were categorized in different folders manually and uploaded. The images, then, were converted into a 4D tensor with single element in the batch dimension. The 3D arrayed images include height, width, and channel. The images have 3 channels called RGB (red, blue, and green). Batches of image data would be generated with image generator. The images would be normalized, centered, resized and scaled to boost image classifier performance by generalizing the images for the model.

Up to this point, the steps would be implemented for both transfer and non-transfer learning. Now a transfer learning, pertained models’ weights would be uploaded. The customs dense layers based on the number of classification and possible drop-out would be added to top of the pretrained model. The model should be trained for few epochs just for the top added layers to have some initialized weights: So, all the layers would have weights. Now few more layers on top would be unfrozen and the remaining would be left frozen. Like non-transfer learning, the model would be compiled and fit.

4. Transfer Learning

In this study, insufficient available labeled dataset makes it difficult to train a deep CNN. In order to tackle this issue, the transfer learning could be used to address a challenge of low number of images. This method uses the knowledge learned from significant number of image datasets, which is transferable to the dataset used in this study. In other word, instead of starting from scratch, the weight of already-trained models would be used, and only few last layers would be left to be optimized. It should be noted the distance between the dataset used in this study compared to the datasets used for ImageNet, e.g., balloons or strawberry, is very large. However, although it was shown that transferability of features decreases as the distance between base and target models increases, transferring features from distant datasets outperform using random features [19].

For a small dataset, transfer learning could be considered for using the CNN using scars training data without overfitting. The literature discussed the transferability of deep CNNs trained with ImageNet dataset, and then adapting those features for a new target task (i.e., fine-tuning), and goodness of results [20]. Three transfer learning models: Inception_V3, VGG 19, and denseNet 121 models with pretrained weights on ImageNet dataset we used [21]. ImageNet is a dataset organized based on the WordNet hierarchy. These images are quality-controlled and human-annotated images [22].These images contain more than a million images distributed over 5247 categories.

The first layer features of network are common across all dataset, regardless of dataset being used or cost function being implemented. However, feature extracted in the last layer greatly depends on the datasets being used, and also the implemented cost/objective function. In the literature review, the first and last layers are called general and specific layers [23]. This method is suitable when the condition of having a target dataset significantly smaller that the base dataset. The method is based on the intuitive concepts of using the weights of the first $n$ layers for the base model, and then using the weights of those layers for training of the target network. The other layers would be initialized randomly and trained based on the classes of the target dataset. For this study, as the target sample size was small and the number of parameters in the base model was large, fine-tuning only implemented on the last few blocks, while the remaining layers were left frozen to prevent over fitting.

In computer vision domain, various architectures of CNNs have been built, which reached impressive performance. These include models such as AlexNet [24], Googlenet [25], VGG-16-Net [26], and res-Net [27] being performed on 1000-class object classification on ImageNet. However, for above models as there are millions of parameters to be verified in the network, large-scale dataset is indispensable when training CNN [19]. The framework of the steps taken in this study for transfer learning is presented in Figure 3.
During the process of categorizing the dataset, instances were encountered that either a decision was hard for the research team to come up with a category, or pictures were taken from an inappropriate spot. So, in order to prevent confusion, these pictures were removed from the dataset (see Figure 3). For instance, Figure 4a,b are unclear from an expert point of view to decide whether those barriers are cable or box beam. Having said that, there is a solution for this challenge. The images could have been cut manually or by the help of some algorithm so each part would be allocated to cable and box beam category. However, that would cause a problem for an automated process as the algorithm needs to know exact images so it can identify them. So, we need to have enough images of cable and box beam barriers so when a machine observes them it knows what step it should take.

After performing the above steps, four folders were populated for various datasets. Due to lack of dataset for each category, 22% of images were assigned to validating dataset while only 4% of the dataset were allocated for testing dataset, and the remaining images were allocated to the training dataset. Two dense layers were added on top of the models so the base model output could pass through this layer before prediction. Also, as the model initially suffered from overfitting, an additional dropout layer of 0.5 was added before a dense layer for prediction. The model was compiled with loss of categorical cross-entropy, optimizer of adam and accuracy as metrics, and learning rate of 0.001.

It should be noted that what has been already learned by models in various transfer learning were combined with new created layers to learn how to categorize new images. The dense layer has the rectified linear unit (ReLU) activation function and the last layer contain as many neurons as the number of classes (which is 4), with softmax activation function, or normalized exponential function. After the last two layers obtained some primarily weight, only the last two blocks of the model would be left to be free to be optimized while all other layers would be left frozen, and weight would not be optimized. In this step, the number of steps for training dataset and validation dataset would be clarified by dividing train sample size by batch size. For more information about fine tuning the model, the reader are referred to the literature review [20,23].
Signals 2021, 2

Figure 4. Some invalid images that were removed from the dataset, (a–f) are not linked to any specific barriers’ types.

4.1. VGG 19

The number of 19 stands for the number of weight layers in the network. The network has approximately 138 million parameters being trained on more than a million images from ImageNet database. The very deep network consists of 16 CNN layers incorporating 5 polling layers, and 3 fully connected layers, and the images would be passes through convolutional layers with a 3 × 3 filter. Also, by going deeper, the volume would be reduced by max pooling layers. It should be noted that top of the CNN layers followed by three fully connected layers. The ImageNet include over 1000 object categories such as mouse, pencil, and some animals [25]. The hidden layers are equipped with ReLU non-linearity.

4.2. Inception v3

There are 4 versions of inception model. First, inception-v2 improved over inception v1 by the introduction of normalization, and later improved by additional factorization to inception v-3. Inception v-3 includes more than 21 million parameters with 48 layers. The latest inception model, v-3, was implemented in this study for transfer learning. Inception v3 is an image recognition model on 1000 classes that has attained more than 78% accuracy on the ImageNet dataset [28].
4.3. DenseNets

Densely connected convolutional network (DenseNets) works by expanding the depth of neural networks [28]. This network solved the vanishing gradient occurred in deep network by simplifying the connectivity. Pattern between layers found in the previous studies including highway networks [26,29], residual networks [26], and fractal networks [30]. This was achieved by ensuring the maximum information/gradient flow.

All convolutional layers are followed by batch norm and ReLU activation function. The network is motivated to avoid representational bottlenecks, reducing the input dimensions of the next layer, resulting in more computational efficiency. One of the unique aspects about this method was the implementation of batch normalization (BN). The BN results in a higher accuracy and faster training speed. The BN normalize the value distribution before going to a next layer.

4.4. Non-Transfer Learning

To have a base model for performance comparison along with transfer learning model, a simple non-transferred model was used in addition to transfer learning. This model has a simple architecture with 4 convolutional layers. Each convolutional layer is accompanied by a max pooling layer with pooling size of 2 by 2. Kernel size or number of filters were chosen as 32, 64, 128, and 128, increasing from a lower level layer to a higher layer, and being a factor of 2. These values were chosen based on trial and error, and based on common practice in the literature review.

Before passing the convolutional layers to the dense layers, flatten and dropout layers were added. A dense layer was added before doing prediction at the last dense layer with 512 neurons (see Table 2).

Table 2. Design scheme of classification pathway, non-transfer learning.

| Layer         | Kernel Size | # of Filters |
|---------------|-------------|--------------|
| Conv1         | 3 × 3       | 32           |
| Max pooling   |             |              |
| Conv2         | 3 × 3       | 64           |
| Max pooling   |             |              |
| Conv3         | 3 × 3       | 128          |
| Max pooling   |             |              |
| Conv4         | 3 × 3       | 128          |
| Max pooling   |             |              |
| Layer flatten + layer dropout (0.5) + layer dense (512) + layer dense (4) | | |

#epochs = 30, steps per epochs: 12, validation steps: 10, lr = 1 × 10^{-3}, #batch = 32

5. Results

In this study, transfer learning as functional application programming interface (API) was used to distinguish among 4 different barriers. In addition to having transfer learning models for performance comparison, a simple CNN model was used and compared. For transfer learning three pre-trained models were considered. Primarily, the transferred learning models were overfit, as can be seen from a huge gap between training and validation accuracy, see Figure 5b for instance. Thus, to address this issue a dropout layer is added before doing a prediction at a last dense layer. A drop out amount, a fraction of the input units to drop, was set as 0.5. After adding a drop out layer, no sign of overfitting was observed from model accuracy plots, see Figure 5c. However, still a small overfitting could be observed for DenseNet 121 even after adding dropout layer.
signals 2021, 2

(a) Non transferred learning

(b) Inception model with no drop out layer at the dense layer

(c) Inception transfer learning with a drop out layer at the dense layer

(d) VGG 19

(e) Densenet 121

Figure 5. A plot of model accuracy on train and validation datasets, transfer learning (left) versus non-transfer learning (right).

Based on Table 3, although the four models were performed on such a small dataset, VGG-19 transfer learning perform significantly better than all transfer learning and non-transfer learning. However, non-transfer learning still performs better than denseNet and inception V3 models. It should be noted that the main advantage of non-transfer learning was related to much lower cost compared to transfer learning, in terms of time spend for training. The models were implemented on unseen test dataset by evaluating generator function. Table 3 presents a summary of generated class prediction, confusion matrix of the 4 included categories over the test dataset. Also Figure 5 presents accuracy and loss plots of the incorporated model. As an example, Inception model with and without a drop out are presented, Figure 5b,c.

Table 3. Performance of the model on test dataset.

| Metric (Accuracy)   | %   |
|---------------------|-----|
| Transfer Learning   |     |
| Inception V3        | 78  |
| Densenet 121        | 65  |
| VGG 19              | 97  |
| Non-transfer learning | 85  |

A can be seen from Figure 5, all training figures have higher accuracy and lower loss for training than the validation datasets. That is expected as the model is already familiar with the training dataset, while the validation data is a collection of new input to the model. The validation dataset, as a set of a sample data, would be used to provide an estimate of model skill during model’s hyperparameters’ tuning. The validation dataset would be used while the model is fitting on the training dataset for evaluation of the prediction power, qualify the model performance so it can adjust the performance while being trained.
So, the model already turned its hyperparameters on the validation so it would not be a fair evaluation to present the validation accuracy.

The confusion matrix for VGG 19 and non-transfer learning, as the best performed models, were included in Table 4 for the test dataset. As can be seen from this table, although in many cases the image of concrete was not clear and many barriers included concrete step at the bottom of the barriers the models were able to predict these barriers in 100% of cases for both transfer and non-transfer learnings.

**Table 4.** Confusion matrix of 4-class of barrier types in test data, non-transfer learning and transfer learning.

|               | Non-Transfer Learning | Transfer Learning VGG 19 |
|---------------|-----------------------|-------------------------|
|               | Predicted             | Predicted               |
| Category      | Box Beam | Cable | Concrete | Hybrid | accuracy | Box Beam | Cable | Concrete | Hybrid | Accuracy |
| Box beam      | 6        | 0     | 0        | 4      | 60%      | 10       | 0     | 0        | 0      | 100%      |
| Cable         | 2        | 7     | 1        | 0      | 70%      | 0        | 9     | 1        | 0      | 100%      |
| Concrete      | 0        | 0     | 10       | 0      | 100%     | 0        | 0     | 10       | 0      | 100%      |
| Hybrid        | 1        | 0     | 0        | 9      | 90%      | 0        | 0     | 0        | 10     | 100%      |

The non-transfer learning model was unable to distinguish between hybrid and box beam barriers in 10% of cases, misclassified hybrid barriers with box beam, and also misclassification was 30% of cases for cable barriers. However, the transfer learning accuracy for box beam and hybrid were 100%, and the model was only not able to predict cable barrier in a single case, confusing with concrete barriers. In summary, the main concern of non-transfer learning was due to confusing box beam with hybrid, due to similarity of these two types. However, the transfer learning was able to address the issue by reaching 97% accuracy.

Few points could be highlighted regarding a better performance of VGG 19. Although there is no clear-cut argument about why one architecture performs better than the other, it could be speculated that the VGG 19 performed better due to having a larger number of weights. Although the higher number of weight has the shortcoming of the higher cost of inference time, it might result in a better performance. In other words, the larger numbers of weights make the architecture to learn more complex images. In addition, for instance compared with Inception, the VGG only flattens the last pooling layer while preserving the information of the locations, while the inception uses global average pooling (GAP).

Also, it should be noted the performances of various model significantly depend on the adjustment of various hyperparameters, image complexity and the number of unfrozen layers. So, the above is just some speculations we could put forward regarding the reasons behind the performance of the models.

**6. Discussions and Future Directions**

There are approximately 800 miles of interstates, and over 4000 miles of state and U.S. highways in the State of Wyoming [31]. The WYDOT needs to monitor and manage the assets they own. The automated method would help them to manage and maintain those assets in a systemized manner. In addition, it would help them to prioritize the maintenance of their resources in a most efficient way. Assets maintenance and optimization has been conducted for various purposes such as safety, comfort and convenience of road users.
Assets maintenance needs an extensive and up-to-date condition data. However, securing the information requires an extensive amount of funding. For instance, collecting information related to the traffic barriers alone cost WYDOT more than half a million dollars. It makes sense that the data collection to be automated in the future to reduce cost. This study was conducted by using limited barrier images to determine if a trained algorithm could distinguish across various barrier types. The main objective of this study was to evaluate image recognition related to the roadside object detection, specifically roadside barriers. Due to the low frequency of images in this study, various transfer learning techniques first were implemented to address the lack of images. A pre-trained weight related to inception v3, VGG 19, and Densenet 121 models, being trained on imagenet, were considered in this study. In those models, two blocks were added, and the majority of pre-trained layers were left frozen while the last top few layers were left to be unfrozen. In summary, only few last layers and a dense layer for prediction were left to be trained on traffic barriers dataset.

In addition, to have a base model for comparison, a simple model with few convolutional layers was considered and trained to be used for barrier images identification. The transfer learning of VGG 19 was able to reach an accuracy of 90% by just misclassifying a single image. That is despite the use of only 250 images for training the algorithm. The results showed that although the VGG 19 outperformed all the models, non-transfer learning performed promising compared with the other two transfer learning techniques. The high performance was despite resizing the images to much smaller dimensions. The resizing was due to the future application of the method on the roadway: The thin clients with limited processing capabilities would send the images to server with a limited bandwidth, and the data would be stored in the database.

So it is impractical to transmit the images descriptions to the server with huge considered size. Also another objective was to test if we could reach an acceptable accuracy before moving to the next phase of the project.

One of the shortcomings of the trained model for non-transfer learning was inability to have a good prediction between hybrid barriers (box beam + guardrail), and box beam. This issue might be addressed by having more images and constructing a deeper CNN. The non-transfer learning was an important part of this study due to its low associated costs in terms of time and resources, it took much lower time to be trained compared with transfer learning techniques.

Artificial intelligence plays a pivotal role in various aspects of asset management systems. Automated data collection is especially important for traffic safety improvement. For future studies, images should be taken from a fixed distance with a fixed angle and the measurements could be recorded and labeled along with the related images. After collecting an adequate number of images, automated vehicles can take new images and the new images could be evaluated and labeled automatically, based on the barrier geometric characteristic without the help of a workload. Thus, the proposed data collection technique will save a significant amount of resources, making a tremendous amount of data available for researchers for data analysis.

As discussed in the content of the manuscript, some of the limitations of this study was lack of enough images for some of the categories, see Figure 4a,b. More images would be taken from unique areas so the algorithm would distinguish and capture various geometric characteristics of those features.

Future studies would be implemented also to utilize machine learning techniques for distinguishing among rusted, damaged or below design standard barriers. In the long run, the proposed approach might be used to help drivers with object detection in autonomous vehicles even with a very low frequency of images. Although CNN layer could extract the most important information, the cropping of images would help for reaching an optimum accuracy as most of the available images were related to the background and were unnecessary.
Author Contributions: K.K. supervision and funding; M.R. methodology. All authors have read and agreed to the published version of the manuscript.

Funding: The authors would like to acknowledge that this work is part of project #RS03218 funded by the Wyoming Department of Transportation (WYDOT).

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References
1. Federal Highway Administration. Transportation Asset Management; Federal Highway Administration: Washington, DC, USA, 2001.
2. Federal Highway Administration. Flexibility in Highway Design; Publication no. FHWA-PD-97-062; Federal Highway Administration: Washington, DC, USA, 2012.
3. Molan, A.M.; Ksaibati, K. Developing the New Barrier Condition Index (BCI) to Unify the Barrier Assessments—A Case Study in Wind River Indian Reservation, Wyoming. Open Transp. J. 2018, 12, 182–191. [CrossRef]
4. Larson, C.D.; Skrypczuk, O. Comprehensive Data Collection to Support Asset Management at Virginia Department of Transportation. Transp. Res. Rec. 2004, 1885, 96–103. [CrossRef]
5. Brezočnik, L.; Fister, J.I.; Podgorelec, V. Swarm Intelligence Algorithms for Feature Selection: A Review. Appl. Sci. 2018, 8, 1521. [CrossRef]
6. Dulebenets, M.A.; Kavoosi, M.; Abiyo, O.; Pasha, J. A Self-Adaptive Evolutionary Algorithm for the Berth Scheduling Problem: Towards Efficient Parameter Control. Algorithms 2018, 11, 100. [CrossRef]
7. Atev, S.; Arumugam, H.; Masoud, O.; Janardan, R.; Papanikolopoulos, N.P. A Vision-Based Approach to Collision Prediction at Traffic Intersections. IEEE Trans. Intell. Transp. Syst. 2005, 6, 416–423. [CrossRef]
8. Artail, H.; Khalifeh, K.; Yahfoufi, M. Avoiding car-pedestrian collisions using a VANET to cellular communication framework. In Proceedings of the 13th International Wireless Communications and Mobile Computing Conference (IWCMC), Valencia, Spain, 26–30 June 2017; pp. 458–465.
9. Ojala, R.; Vepsäläinen, J.; Hanhivuo, J.; Hirvisalo, V.; Tammi, K. Novel Convolutional Neural Network-Based Roadside Unit for Accurate Pedestrian Localisation. IEEE Trans. Intell. Transp. Syst. 2019, 21, 3756–3765. [CrossRef]
10. Harbaš, I.; Prentašić, P.; Subašić, M. Detection of roadside vegetation using Fully Convolutional Networks. Image Vis. Comput. 2018, 74, 1–9. [CrossRef]
11. Lau, M.M.; Lim, K.H.; Gopalai, A.A. Malaysia traffic sign recognition with convolutional neural network. In Proceedings of the 2015 IEEE International Conference on Digital Signal Processing (DSP), Singapore, 21–24 July 2015; pp. 1006–1010.
12. Sikirić, J.; Brkić, K.; Bevandić, P.; Krešo, I.; Krapac, J.; Segvić, S. Traffic Scene Classification on a Representation Budget. IEEE Trans. Intell. Transp. Syst. 2020, 21, 336–345. [CrossRef]
13. Al Mufti, M.; Al Hadhrami, E.; Taha, B.; Wergbi, N. Automatic target recognition in SAR images: Comparison between pre-trained CNNs in a transfer learning based approach. In Proceedings of the International Conference on Artificial Intelligence and Big Data (ICAIIBD), Chengdu, China, 26–28 May 2018; pp. 160–164.
14. Liu, T.; Xie, S.; Zhang, Y.; Yu, J.; Niu, L.; Sun, W. Feature selection and thyroid nodule classification using transfer learning. In Proceedings of the 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017), Melbourne, Australia, 18–21 April 2017; pp. 1096–1099.
15. Han, D.; Liu, Q.; Fan, W. A new image classification method using CNN transfer learning and web data augmentation. Expert Syst. Appl. 2018, 95, 43–56. [CrossRef]
16. Cheng, P.M.; Malhi, H.S. Transfer Learning with Convolutional Neural Networks for Classification of Abdominal Ultrasound Images. J. Digit. Imaging 2017, 30, 234–243. [CrossRef] [PubMed]
17. Cira, C.; Alcarria, R.; Manso-Callejo, M.; Serradilla, F. A Deep Learning-Based Solution for Large-Scale Extraction of the Secondary Road Network from High-Resolution Aerial Orthoimagery. Appl. Sci. 2020, 10, 7272. [CrossRef]
18. Molan, A.M.; Rezapour, M.; Ksaibati, K. Modeling traffic barriers crash severity by considering the effect of traffic barrier dimensions. J. Mod. Transp. 2019, 27, 141–151. [CrossRef]
19. Agrawal, P.; Girshick, R.; Malik, J. Analyzing the Performance of Multilayer Neural Networks for Object Recognition. In Proceedings of the European Conference on Computer Vision, Zurich, Switzerland, 6–12 September 2014; pp. 329–344.
20. Szegedy, C.; Vanhoucke, V.; Ioffe, S.; Shlens, J.; Wojna, Z. Rethinking the Inception Architecture for Computer Vision. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 2818–2826.
21. ImageNet. Large Scale Visual Recognition Challenge (ILSVRC): Competition. 2012. Available online: http://www.image-net.org/challenges/LSVRC/ (accessed on 27 December 2016).
22. Yosinski, J.; Clune, J.; Bengio, Y.; Lipson, H. How transferable are features in deep neural networks? In Proceedings of the Advances in Neural Information Processing Systems 27 (NIPS 2014), Montreal, QC, Canada, 8–13 December 2014; pp. 3320–3328.
23. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. In Proceedings of the Advances in Neural Information Processing Systems (NIPS), Lake Tahoe, NV, USA, 3–8 December 2012; pp. 1097–1105.

24. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going deeper with convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015; pp. 1–9.

25. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. arXiv 2014, arXiv:1409.1556. Available online: https://arxiv.org/abs/1409.1556 (accessed on 2 January 2021).

26. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 26 June–1 July 2016; pp. 770–778.

27. Huang, Z.; Pan, Z.; Lei, B. Transfer Learning with Deep Convolutional Neural Network for SAR Target Classification with Limited Labeled Data. Remote. Sens. 2017, 9, 907. [CrossRef]

28. Huang, G.; Liu, Z.; Van Der Maaten, L.; Weinberger, K.Q. Densely Connected Convolutional Networks. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 4700–4708.

29. Srivastava, R.K.; Greff, K.; Schmidhuber, J. Highway networks. arXiv 2015, arXiv:1505.00387. Available online: https://arxiv.org/abs/1505.00387 (accessed on 2 January 2021).

30. Larsson, G.; Maire, M.; Shakhnarovich, G. Fractalnet: Ultra-deep neural networks without residuals. arXiv 2016, arXiv:1605.07648. Available online: https://arxiv.org/abs/1605.07648 (accessed on 2 January 2021).

31. Wyoming Department of Transportation. Traffic Volume and Vehicle Miles Book; Wyoming Department of Transportation: Cheyenne, WY, USA, 2018.