Zhang, Y., Gawade, M., Lin, P.S. and McPherson, T. (2013). "Educational campaign for improving pedestrian safety: a university campus study." *Procedia-social and behavioral sciences*, *96*, pp.2756-2766.

Zhao, J., Malenje, J.O., Tang, Y. and Han, Y. (2019). "Gap acceptance probability model for pedestrians at unsignalized mid-block crosswalks based on logistic regression." *Accident Analysis & Prevention*, *129*, pp.76-83.

Traffic Warning System for Wildlife Road Crossing Accidents Using Artificial Intelligence

Ikenna D. Uwanuakwa, Aff.M.ASCE[1]; Ugochukwu G. Isienyi[2]; John Bush Idoko[3]; and Shaban Ismael Albrka, Ph.D.[4]

[1]Dept. of Civil Engineering, Near East Univ., Nicosia, Turkey. Email: ikenna.uwanuakwa@neu.edu.tr
[2]Dept. of Civil Engineering, Near East Univ., Nicosia, Turkey. Email:isienyiug@gmail.com
[3]Applied Artificial Intelligence Research Centre, Near East Univ., Nicosia, Turkey. Email:john.bush@neu.edu.tr
[4]Dept. of Civil Engineering, Near East Univ., Nicosia, Turkey. Email:shabanismael.albrka@neu.edu.tr

ABSTRACT

Wildlife-vehicle collision (WVC) is a major problem associated with regions with high-density wildlife. Urban designers have in the past introduce overpasses, underpasses fence, reflectors, and sensors to aid safe wildlife road crossing, but these have not been able to reduce the wildlife-vehicle collision. This research focused on the automated warning system to vehicle users to minimise wildlife-vehicle collision which could integrate computer vision in the detection of features on the road together with the location-time information feed. The proposed system was trained using AlexNet, GoogelNet, ResNet-50, and VGG-16 algorithm on a deep convolutional neural network (CNN) using 20,964 images of 25 variables consisting of 21 animals and four different vehicle body type. The dataset was divided into training and validation set. The results show that CNN algorithms could identify objects on real-life traffic data with noise background at a reliable accuracy. The GoogelNet, ResNet-50, and VGG-16 model outputs were found to have a better prediction accuracy than the AlexNet model in detecting the object features on the traffic images.

Keywords: Traffic; Wildlife-vehicle collisions; Deep Learning; AlexNet; GoogleNet; Accident.

INTRODUCTION

Wildlife vehicle collisions are not just a threat to endangered species, it is also a threat to vehicle users and the wildlife wellbeing which may result in the‘loss and fragmentation’ of the wildlife habitat (Ha and Shilling 2018; Wei and Zhang 2010). According to (Centers for Disease Control and Prevention (CDC) 2004), annually, 200 human deaths are estimated results from wildlife-vehicle collision (WVC) in the United States. Furthermore, the report stated that between 2000-2001, an estimate of 26,647 WVC cases was reportedly treated for nonfatal injuries. According to Rowden et al., (2008), in Queensland, Australia, WVC accounts for 5.1 % of all on-road serious casualties with 51 % resulting in fatality or casualty.

According to road ecologist, wildlife road crossing does not occur at a random location but within the locality with a spatial cluster of vertebrate species influenced by several behavioural a geometric factors.The challenges associated with WVC have been a century-long problem. The urban designers have sought ways to minimize WVC which include; underpass, fencing and signs warning to road users of the potential wildlife crossing (Ha and Shilling 2018; Wilkins et al. 2019). These approaches according (Mammeri et al. 2016) are passive methods and mostly ineffective. Even though this might appear on a policy level to have improved the situation, recent researches show that the rate of wildlife-vehicle collisions has increased in the last decades (Cserkesz et al. 2015; Seiler 2004; Sullivan 2011). The alternatives to the physical passive infrastructures are the sensors and reflectors which could be road-based or vehicle-based (Saleh et al. 2018) and have also be found to be inefficient in mitigating WVC (Benten et al. 2018).

Besides the abovementioned approaches, it is observed (Knapp et al. 2004) that vision detection systems are more reliable and efficient in minimizing WVC. However, the contemporary camera-based infrared systems are prone to false activation due to air or human movement. The system sometimes could fail in the detection of animals on the curve-lanes (Mammeri et al. 2016) which could lead to false detection and misclassification.

This research work seeks to improve upon the available systems with adequate artificial intelligent techniques which have the potential to provide warning systems that might reduce the accident associated with a wild-life vehicle collision.

RELATED WORKS

Previous research work has been carried out for effective wildlife-vehicle avoidance system. These works are either focused on a select animal or group of animals such as large animal detection system(Mammeri et al. 2016) On the other hand, the infrared detection system has also been researched (Zhou and Wang 2012),

Sharma & Shah (2017) proposed a model capable of alerting drivers in a vehicle accelerating a maximum speed up to 35 km/h with an accuracy of 82.5%. The model was trained with 2200 images. The research focused on reducing vehicle collision with domestic animals (cow). Input acquisition for the detection system comes from a monocular camera mounted on the vehicle. This approach requires hardware installation in all vehicles and similar to Mammeri et al., (2016) and Zhou & Wang, (2012). The proposed models were trained with support vector machine classifiers from the histogram of oriented gradients (HOG) features. The signal warnings are to be sent from a reflector visible to vehicles within the line of sight distance. The drawback on models which focused on a single animal or a particular species could lead to a fatal accident where the system fails to detect object beyond their scope. On the other hand, it is efficient to warn vehicle users who are capable of making a calculative judgement based on information provided.

A further approach in related work is the implementation of smart camera output for a vision system. The concept was employed in Amato et al., (2017) in a deep CNN architecture using AlexNet for a decentralised vision detection of the parking status. The deep CNN architecture was designed to run on an embedded system of the parking smart cameras. The camera detects the status of a parking lot and sends only a binary output to the central server. The results show the reliability of AlexNet vision detection in the presence of noise-induced from "light intensity and variation, shadows and partial conclusions" (Amato et al. 2017).

In other to eliminate the drawback of models of (Mammeri et al. 2016; Sharma and Shah 2017; Zhou and Wang 2012), the concept used in (Amato et al., 2017) which could integrate image recognition and classification in the detection of features on the road together with the location-time information feed. This approach has the potential to enhance the efficiency of the available camera-based detection systems in preventing road accidents at a higher level.

METHODOLOGY

Data set used in this research comprises 20, 964 images of 25 variables from Google Image search and (Pingel and Ha 2017; Xian et al. 2018). Details of the images are presented in Table 1. Images extracted were in their raw state. Images with less than 24-bit depth were removed. The MATLAB Image Batch processor application was used to transform all images to conform to AlexNet image standard of 227 x 227 x 3 images pixel and 224 x 224 x 3 images pixel for GoogelNet standard.

Table 1. Image training dataset description

Object	Number of data	Object	Number of data
Antelope	1046	Rhinoceros	696
Bear	868	Sheep	1420
Cow	1338	Skunk	188
Deer	2008	Snake	828
Elephant	1038	Squirrel	1200
Fox	664	Tiger	877
Giraffe	1202	Big Truck	201
Gorilla	872	Car	200
Horse	1645	SUV	200
Leopard	720	Truck	200
Lion	1019	Van	200
Moose	704	Weasel	272
Rabbit	1088		

ALEXNET AND GOOGLENETTRANSFER LEARNING ALGORITHM

AlexNet was first present at the ImageNet LSVRC-2010 contest for the classification of 1000 object with 1.2 million images with the deep learning convolutional neural network (Krizhevsky et al. 2012). The developed algorithm can also be applied as a pre-trained Convolutional Neural Networks (CNNs) and re-trained with new images (Pingel and Ha 2017). This research used the pre-trained existing network to train the 25 variables. In this regard, existing AlexNet learning experience was transferred into the new training which increased the training efficiency.

MATLAB script (Pingel and Ha 2017) for pre-trained AlexNet was used to train the new network by adjusting the 23, 24 and 25th layer to suit the research modelling. The model was trained on Intel core i7-9700 CPU equipped with GeForce RTX 2080 Ti GPU.

Network initialisation: The trial test shows that an initial learning rate of 0.0001 was effective in training the model. The Glorot initializer (Glorot and Bengio 2010) was used to initialize the biases to be zero and the weights Wijat each layer independently.

$$W_{ij} \sim U\left[-\frac{1}{\sqrt{n}}, \frac{1}{\sqrt{n}}\right], \quad (1)$$

Where "U[−a, a] is the uniform distribution in the interval (−a, a) and n is the size of the previous layer" (Glorot and Bengio 2010)

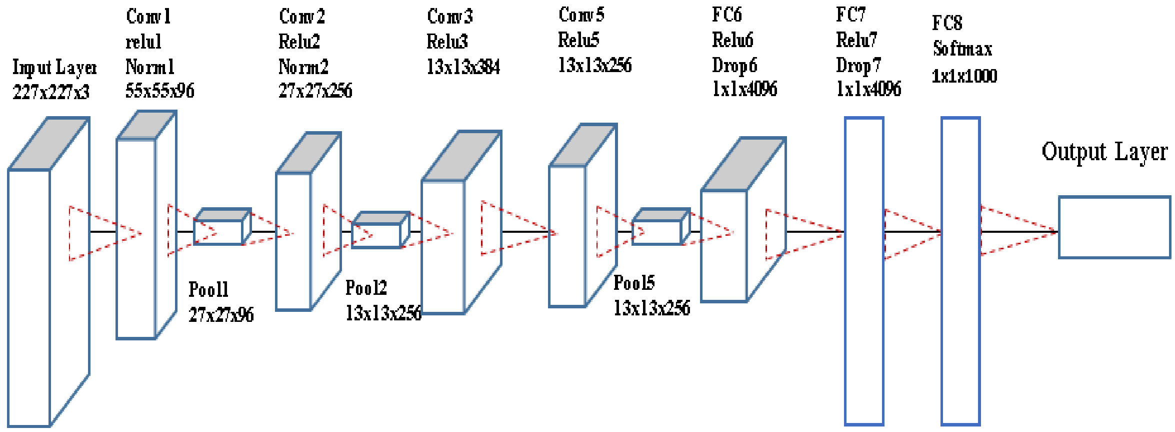


Figure 1. AlexNet Architecture

Training algorithm: The Stochastic Gradient Descent with Momentum (SGDM) training algorithm was employed. The presence of a momentum parameter reduces the oscillation within the path of the stochastic gradient descent algorithm as it moves toward the optimum (Murphy 2012).

$$\theta_{\ell+1}=\theta_{\ell}-\alpha\nabla E(\theta_{\ell})+\gamma(\theta_{\ell}-\theta_{\ell-1}), \quad (2)$$

"where ℓ= iteration number, $\alpha > 0$ is the learning rate, θ is the parameter vector, $\nabla E(\theta)$= the gradient of the loss function where γ determines the contribution of the previous gradient step to the current iteration."

$$\nabla E(\theta) = -\sum_{i=1}^{N}\sum_{j=1}^{k} t_{ij} \ln y_{ij} \quad (3)$$

Where N is the number of cases; K= number of class; yij= is the output case i for class j computed by the softmax function.

THE ARCHITECTURE OF THE AUTOMATED WILDLIFE-VEHICLE COLLISION AVOIDANCE SYSTEM

The proposed automated system design to run on an embedded system of a smart camera (Amato et al. 2017), in which only binary output signals will be sent to the central server. The model consists of two phases; training and simulation. The training section involves the training processed dataset with the proposed AlexNet model. The dataset was divided into training and the validation dataset. The workflow of the proposed model is as shown in Figure 2.

The proposed model output could be integrated into a web mapping application which would provide real-time information (location and estimated distance from the receiver) on detected animal crossing road. Web map applications such as Google Map, Moovit etc, is capable of predicting traffic flow using triangulated signals from GPS machines of road users. This approach has reduced travel time and total delay which improves the level of service on a given route. The proposed model could also be incorporated into these web mapping applications. The advantage of the system lies within the fact that the system signal sent to the road users identifies the animal and its approximate location. This would, in fact, give motorist ample time to adjust on the speed and take appropriate measurements to ensure their safety and the that of the animal.

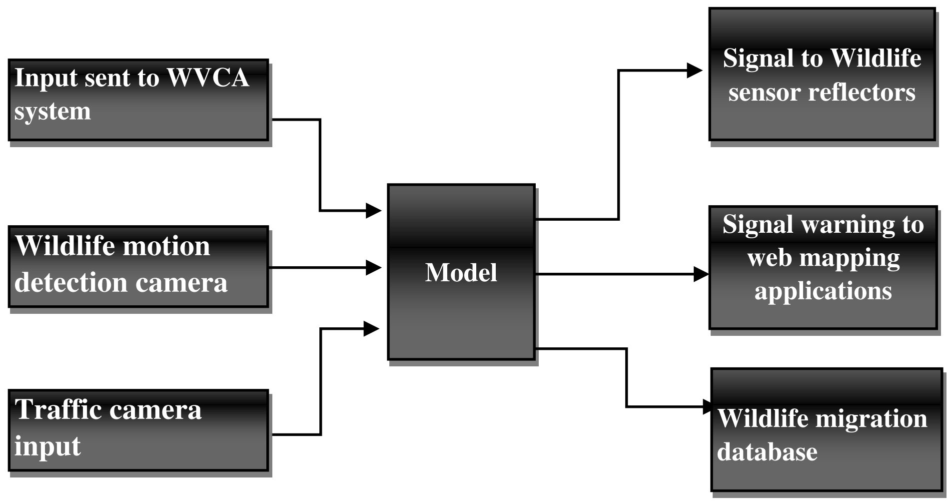
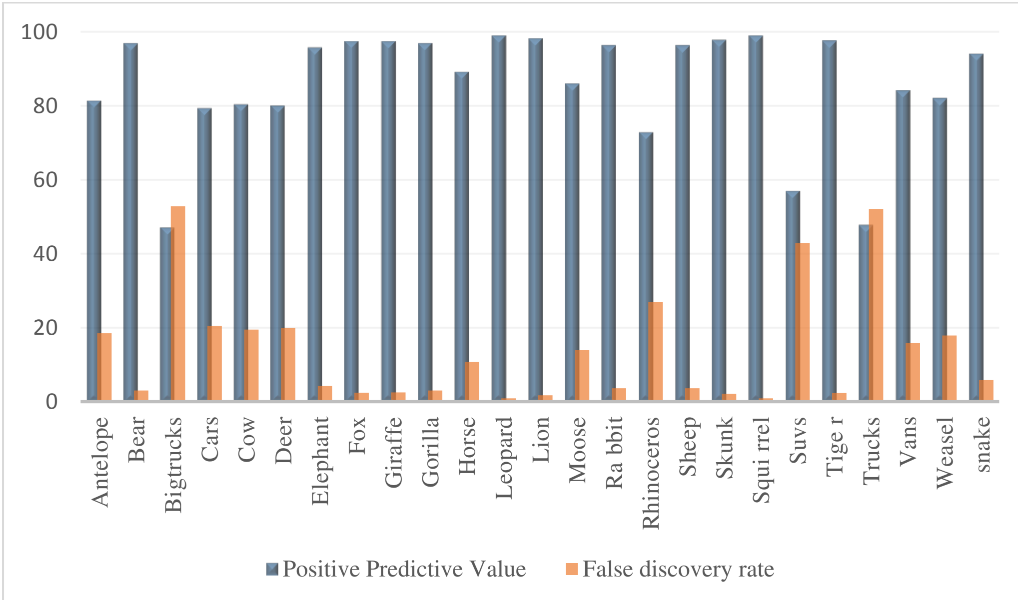


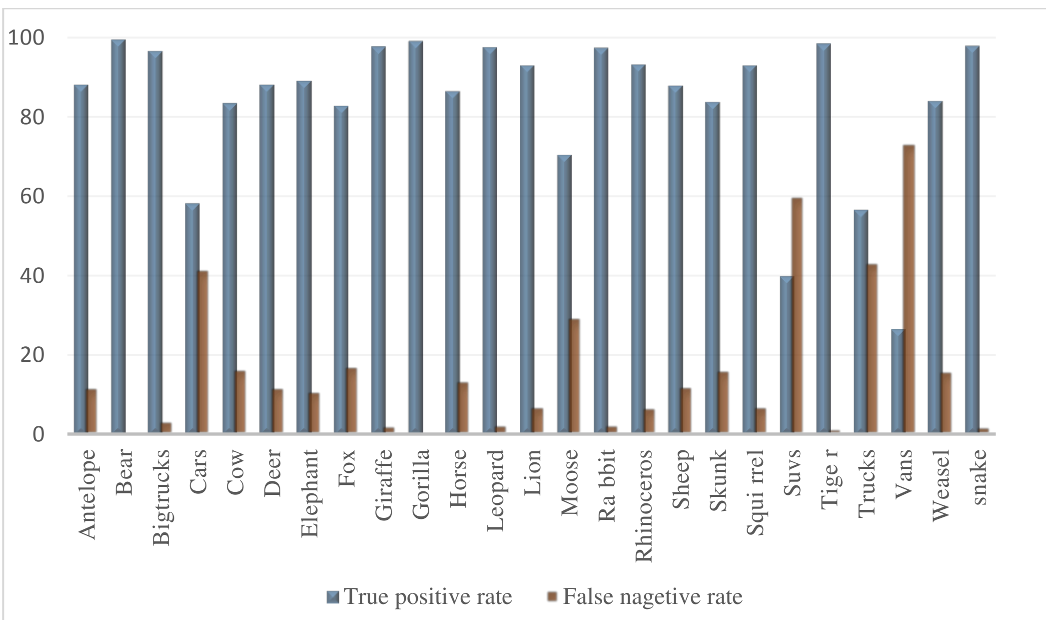
Figure 2. The architecture of the proposed automated wildlife-vehicle collision avoidance system

RESULTS AND DISCUSSIONS

The summary of the confusion matrix plot is presented in Figure 3 (a) the target class metrics output shows the "true positive rate" and "the false-negative rate". The metrics explained the percentage of all the examples correctly and incorrectly classified which belongs to each variable class in the model. The results of the metrics of the output class are shown in Figure 3(b); it evaluates the percentage of all the examples within each variable that are correctly and incorrectly predicted to belong to the class of the variable. The confusion matrix was plotted from the validation set to correctly show the performance of the model. The variables car, SUV, truck and big truck have the highest incorrect classification and predicted rate. This could be as a result of small dataset and similarities among them. For example; out of the 44 and 42 examples used for the validation set of car and SUV, 9 cars were incorrectly classified as SUV while 18 SUVs were also incorrectly labelled as a car. These two variables share their misclassification alone. Furthermore, all incorrect classification of vehicles were among the vehicle input group. It confirms the ability of the model to differentiate a group of variables correctly. Similarly, even-toed ungulates such as deer, antelope, cow, moose and sheep were incorrectly classified among themselves with the deer and the antelope having the highest error shared within each other. This illustrates the similarity in appearance (Artiodactyla) which differentiate them from the other even-toed ungulates contained in the input dataset.



(a)



(b)

Figure3. Summary of the predicted and the target class confusion matrix: (a) percentage of the predicted class classification, (b) percentage of target class classification.

PROPOSED MODEL PERFORMANCE VALIDATION

The results of AlexNet was compared with three other pre-trained deep learning algorithm; GoogLeNet, ResNet-50 and VGG-16. Six variables were selected randomly for accessing the performance of the models. The variables which were not included in the training or validation dataset. The results presented in Figure 4 shows that in general, the other three algorithms have higher detection accuracy than the AlexNet as shown in Tabe 2.

Table 2. Model characteristics and performance evaluation

Model	Total training time (mins)	Number of layers	Validation accuracy (%)
AlexNet	61	25	89.30
GoogLeNet	189	144	95.83
ResNet-50	469	177	96.63
VGG-16	2856	41	96.81

Further, several examples presented in Figure 4 shows the prediction scores of the test dataset. The results of the comparison show that the proposed model of AlexNet was able to predict the variable names based on the combined features from the dataset and the pre-trained experience from the library. However, the prediction scores were lower than the other comparing models. In Figure 4, despite assigning a name class of Rabbit to the variables, the proposed model (AlexNet) was able to classify it as a hare with 93.8% accuracy based on the extracted features whereas the other models predicted it as Rabbit with 100% accuracy. The variables in Figure 4b and 4c true class category was Antelope and Sheep respectively, there was a misclassification from the AlexNet model in classes and ResNet-50 in 4b, while GoogLeNet and VGG-16 true class prediction score was 60.4 % and 79 % respectively.

Further, it can be observed that models with complete replacement of the pre-trained library have high true class prediction scores as compared to AlexNet. AlexNet model performance could have been affected by other variables with similar features. On the other hand, one the advantages of the pre-trained library was observed in Figure 4e were the other models misclassified the variable as Moose due to the position. However, despite the noise introduced by the position of the variables AlexNet true prediction score was 90%. It is obvious that more dataset is required to improve the performance of the models.

CONCLUSIONS

This research proposed a model to detect wildlife animal crossing, for an automated traffic warning system application. The model objective is to minimise the wild-life vehicle collision through the implementation of an avoidance system. The network was trained by modifying a pre-trained AlexNet algorithm with 20, 694 images of 25 variables comprising of cats, ungulates, reptile, small mammals and vehicles. The dataset was built by transforming the images download from open-source databases and stored in a folder with subfolders containing the images of individual variables. The effective of the model was verified by three other convolutional neural networks (CNN) models used in vision computing. The results show that the proposed model has a high validation accuracy of 89.30 %, but the comparing models have higher validation accuracy. In the future, we will increase the dataset and apply the model on an embedded camera to simulate real-life experience.

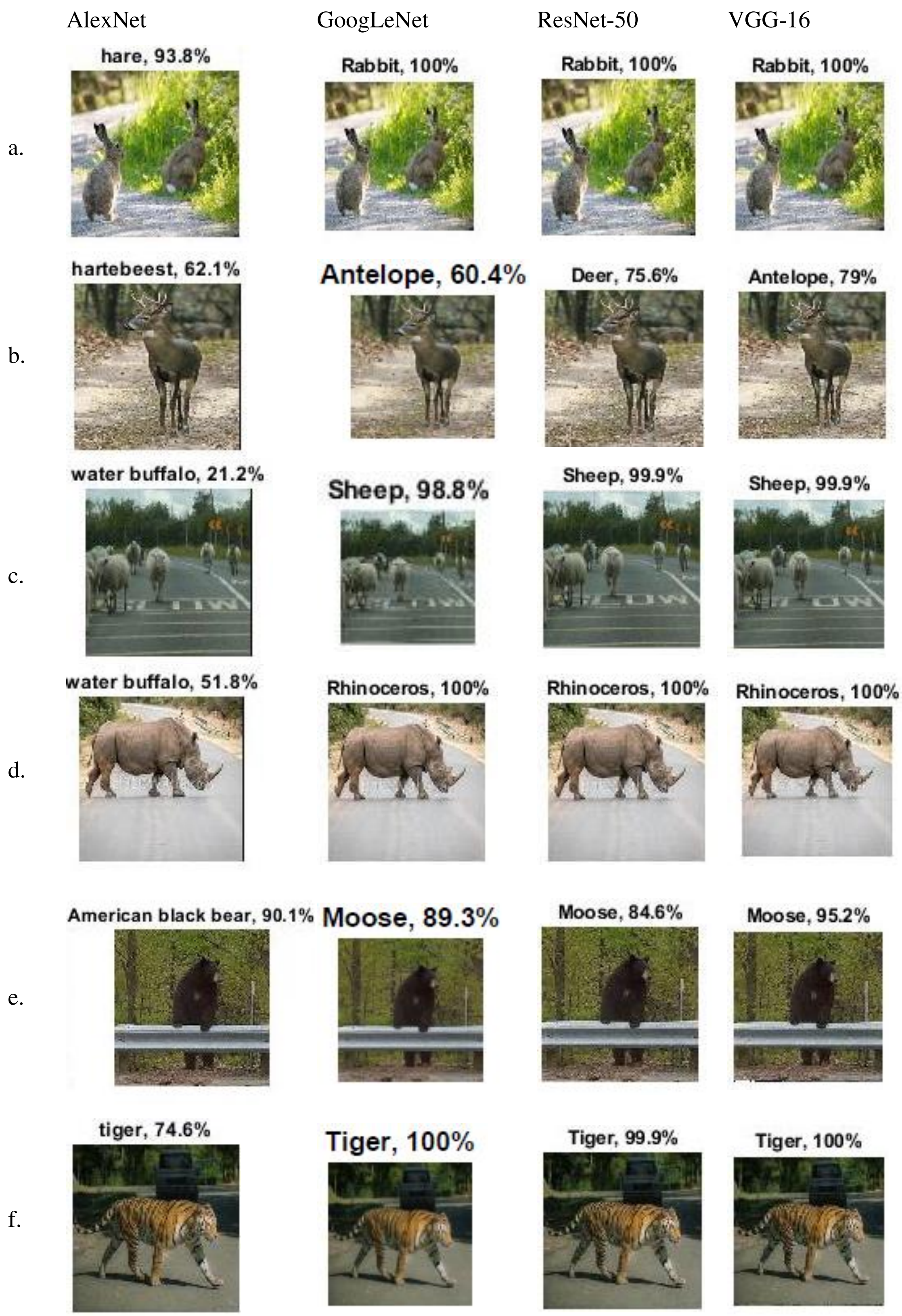


Figure 4. Results of the test dataset of AlexNet, GoogLeNet, ResNet-50 and VGG-16

CONFLICT OF INTEREST

The author declares herewith, that there I have no conflict of interest.

REFERENCES

Amato, G., Carrara, F., Falchi, F., Gennaro, C., Meghini, C., and Vairo, C. (2017). "Deep learning for decentralized parking lot occupancy detection." *Expert Systems with Applications*, Pergamon, 72, 327–334.

Benten, A., Hothorn, T., Vor, T., and Ammer, C. (2018). "Wildlife warning reflectors do not mitigate wildlife–vehicle collisions on roads." *Accident Analysis & Prevention*, Pergamon, 120, 64–73.

Centers for Disease Control and Prevention (CDC). (2004). "Nonfatal motor-vehicle animal crash-related injuries--United States, 2001-2002." *MMWR. Morbidity and mortality weekly report*, 53(30), 675–8.

Cserkesz, T., Zoology, J. F.-N.-W. J. of, 2015, U., Cserkész, T., and Farkas, J. (2015). "Annual trends in the number of wildlife-vehicle collisions on the main linear transport corridors (highway and railway) of Hungary." *North-Western Journal of Zoology*, 11(1), 41–50.

Glorot, X., and Bengio, Y. (2010). "Understanding the difficulty of training deep feedforward neural networks." *Journal of Machine Learning Research*.

Ha, H., and Shilling, F. (2018). "Modelling potential wildlife-vehicle collisions (WVC) locations using environmental factors and human population density: A case-study from 3 state highways in Central California." *Ecological Informatics*, Elsevier, 43, 212–221.

Knapp, K. K., Yi, X., Oakasa, T., Thimm, W., Hudson, E., and Rathmann, C. (2004). *Deer-Vehicle Crash Countermeasure Toolbox: A Decision and Choice Resource*. Madison.

Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). "ImageNet Classification with Deep Convolutional Neural Networks." ImageNet Classification with Deep Convolutional Neural Networks.

Mammeri, A., Zhou, D., and Boukerche, A. (2016). "Animal-Vehicle Collision Mitigation System for Automated Vehicles." *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 46(9), 1287–1299.

Murphy, K. P. (2012). *Machine learning : a probabilistic perspective*. MIT Press.

Pingel, J., and Ha, G. (2017). "Deep Learning for Computer Vision." MathWorks, <https://www.mathworks.com/videos/deep-learning-for-computer-vision-120997.html>(Jun. 29, 2019).

Rowden, P., Steinhardt, D., and Sheehan, M. (2008). "Road crashes involving animals in Australia." *Accident Analysis & Prevention*, Pergamon, 40(6), 1865–1871.

Saleh, K., Hossny, M., and Nahavandi, S. (2018). "Effective Vehicle-Based Kangaroo Detection for Collision Warning Systems Using Region-Based Convolutional Networks." *Sensors (Basel, Switzerland)*, Multidisciplinary Digital Publishing Institute (MDPI), 18(6).

Seiler, A. (2004). "Trends and spatial patterns in ungulate-vehicle collisions in Sweden." *Wildlife Biology*, Nordic Board for Wildlife Research, 10(1), 301–313.

Sharma, S. U., and Shah, D. J. (2017). "A Practical Animal Detection and Collision Avoidance System Using Computer Vision Technique." *IEEE Access*, 5, 347–358.

Sullivan, J. M. (2011). "Trends and characteristics of animal-vehicle collisions in the United States." *Journal of Safety Research*, Pergamon, 42(1), 9–16.