AlexNet and Convolutional Neural Network analysis in Thermogram Images for Automobile Application using High Performance Computing

Yuvaraj Munian
Dept. of Computational, Engg. and Mathematical Science,
Texas A&M University, San Antonio
Texas, USA
ymunian@tamusa.edu

Izzat Alsmadi
Dept. of Computational, Engg. and Mathematical Science,
Texas A&M University, San Antonio
Texas, USA
ialsmadi@tamusa.edu

Garrett Crumrine
Dept. of Computational Engg.and mathematical Science,
Texas A&M University, San Antonio
Texas, USA
gcrumrine@tamusa.edu

2024 International Symposium on Networks, Computers and Communications (ISNCC) | 979-8-3503-6491-0/24/$31.00 ©2024 IEEE | DOI: 10.1109/ISNCC62547.2024.10759037

***Abstract*— Lately, Deep learning and machine learning models have been vividly used in image classification to measure the accuracy of the proposed model. Deep learning networks always require high computing networks to increase the model's performance. Computer vision plays a vital role in polishing the dataset by removing unwanted information, without the loss of important features. As per certain applications, two machine learning and deep learning algorithms with different system requirements are compared and summarized. In this paper, instead of the RGB image dataset, the thermogram image dataset has been considered for the process. The main concern in considering thermal images is due to the particular application. Here the thermal images have been collected during the nocturnal hours, especially on the dawn and dusk of the day. Convolutional neural networks are the machine learning models used in CPU and GPU-cluster configurations. AlexNet will be the deep learning model used here to compare with the CNN supervised learning algorithm. Computer vision technique, HOG is used to reduce the complexity and intricacy of the dataset before getting into the network models. Because of the usage of deep network models' performance computing is expended to increase the speed of the whole process. Deer will be the spotted animal for most road accidents during the nocturnal hours in the United States. The results gaudily conclude the classification of machine learning and deep learning model detection for thermal and HOG images with the highest accuracy of 92% for the AlexNet.**

Keywords*— *high performance computing (HPC), thermal images, histogram of oriented gradients (HOG), machine learning and deep learning, - CNN, AlexNet .

I. INTRODUCTION

Animal vehicle collision is a perilous and unpredictable problem in our everyday life cycle. More experiments associated with this carcass will be the prerequisite for rapid globalization. Using digital images like binary, grayscale, and RGB, more experiments have been formerly piloted, but only very few studies are considered for thermal images. This paper focuses more on the thermal image captured during the nocturnal hours. The prediction of animal behavior is highly impossible, especially the deer characteristics, is highly complicated. This is the reason why all the roadway designs, like bridges between the roads, underpasses, fencing, reflectors, and alert sign notifications, are less effective in road accidents. The traditional method reduces the number of accidents but does not match the expected ratio. Roadkill is a constant threat not only to humankind but also to the economy and animal life. Due to these types of accidents, many dollars have been claimed by insurance providers. Hideout is also a big concern in collecting the real data due to the real reason behind it. As per the U.S. Department of Transportation, five percent of overall accidents are due to the animal's unpredictable behavior during road cross. The insurance industry approximates around 4 million is used for property damage, and the mortality rate is around 200 / year [1][18]. As per the reporting and known data, these numbers have been approximated; this is not the final value because of hideouts. Additionally, around 4000 cases have been filed every year; each case will claim a minimum of 1000 dollars per case [Insurance Journal, Oct 2012]. Our goal is to present the ability of machine learning and deep learning convolutional neural networks (CNN) to discriminate the presence of deer in the real-time image or video, utilizing images collected using thermal cameras during nocturnal hours.

For image processing, like classification, detection, and prediction, convolutional neural network is necessitous artificial intelligence technique. Notably, in applications like biomedical, Transportation, architecture, power systems, and cybersecurity, the machine learning, deep learning techniques and its state-of-the-art methods are essential. To improve the performance of the image analysis, these deep learning methods are introduced and implemented. Predictably the thermal images have more details and millions of trainable and nontrainable information's due to the measure of infra-red radiation through the object. Based on the temperature of the object, the clarity of the image will be changed. Feature maps will be determined by the shared weight architecture of kernels in the network model. The unique feature of CNN is the pre-processing will be considered inside the network model itself, whereas other networking models need pre-processing separately. In addition to the built-in pre-processing, more preprocessing techniques will be used to optimize the accuracy of the detection. CNN is a supervised

979-8-3503-6491-0/24/$31.00 ©2024 IEEE

learning algorithm because it learns from the previous iterations. The neurons in each layer will have different weights based on the tunning parameters, the number will change. Therefore, it is called a multilayer perceptron (fully connected) network.

The neural network evolution has started in 1950s and 1960s era, then the necognitron (convolutional layer + sampling) was introduced by Kunihiko Fukushima in 1980s. AlexNet is designed by Alex Krizhevsky collaboration with llya Sutskever and Geofrey Hinton in the year 2012. In the competition, AlexNet proved that it is much better in image classification than other neural networks. Also, it shows the importance of GPU than CPU. The performance in terms of speed has been achieved. Compared to the CPU, the usage of GPU increases the speed and makes the network iterations faster than other system requirements. In general, AlexNet has five convolution layers, three max pooling layers, two normalized layers, two fully connected layers and one SoftMax layer. The replacement of sigmoid into ReLu increases the analysis of features or parameters much better. The fine-tuning of those parameters will increase the speed and performance of the network model.

GPUs with more clusters and nodes will help the network model predict or detect the features from the image input faster than the CPU. The operating systems combined with these types of clusters will produce better accuracy than the existing system configurations. Deep neural networks will be highly complicated, and these types of complex networks need more space to process the image and determine the features of the image. This knowledge paves the way for current research with image datasets in image analysis.

In summary, the objective of this paper includes:

- Deep Learning algorithm is used for the prediction/detection.
- Benchmark the deep learning techniques with the existing machine learning algorithm to identify the better detection/classification algorithm.
- Compare the GPU and CPU system requirements to identify the better performance system.

The following paper is organized as follows: Section II explains the existing methodologies and literature, which strengthen the ideology of the proposed research. Section III describes the different neural network algorithms and their possibilities in the research. Section IV deliberates the experimental results, their comparisons, and findings of the proposed technique in pictorial form and tabulation for easy understanding, and section V concludes the findings and outcomes of this research paper.

II. Literature Review

A. *Vehicle Collision - Animal Detection*

To optimize the reduction in number of accidents, this section explains about the existing methodologies and techniques towards the vehicle collision animal detection research.

Traditional methods like overpasses, underpasses, and bridges across the highways for animal crossing help the drivers drive their vehicles without any disturbances. The construction cost for those traditional method implementations is high compared with other technical ideas. The next level of warning reflectors and sign boards gives some caution for the drivers, but predicting animal behavior is highly possible. The numbers shown in the introduction section motivate the researchers to move on with this application and kindle the young researchers toward vehicle collision and animal detection.

The data collection and the pre-processing are done by many computer vision techniques which leads path to this research. The existing pre-processing techniques in computer vision like circular transformation, cough transformation, HoG transform and other transformation techniques are existing in the field for so many years. The ideology of selecting the variant and transformation techniques is the most important tasks for the researcher.

B. *Machine Learning for Detection*

Machine learning is an emerging trend in detection and prediction of the future based on the previous and current data. The training of the designed model will get all the information's from existing datasets.

Paper [2] talks about the different postures of wildlife animals, which helps the researcher to fine-tune the parameters in the mathematical model. The SVM technique is used to detect human objectives with pre-processing techniques proposed by [3]. Article [4] explains the active and passive infrared sensors used in the vehicle, which alert the driver when it senses the animal movement. The CNN algorithm detects the animal movement based on the image and video as an input, as proposed in [5]. Logistics regression is used to detect the risk evaluation of the real-time crashes as described in Yang. The detection of animals has given an alert to drivers to slow down increase the speed of the vehicle or change the direction of the vehicle to avoid accidents[7]. [9] The comparison of machine learning classifiers has been explained for both animal detection and texture descriptors in [11][12]. Histogram of oriented of Gradients (HOG) techniques [10] is used to pre-process the given thermal image and to reproduce as the same quality is explained in [13][16]. Vgg16 and Vgg19 models are used to identify the animal in the thermal image, as explained in [7]. A pre-trained convolutional neural network can be used to detect or classify the images [8].

C. *Deep Learning for Detection*

In [23] author describes the comparison of recurrent neural networks, deep neural networks, and deep Q networks. By fine-tuning the CNN computer-aided applications, they have been implemented in several scientific areas. In [14], the animal behavior has been classified, and the feature extraction is extracted by the deep learning algorithm. The movement of grazing cattle is measured using the IOT and deep learning techniques. EfficientNetB4 is used to classify the animal image dataset and also to recognize animals in the image. A deep CNN network is used for medical image analysis, and feature detection is discussed in [8]. The highly complicated medical images will have more information, and deep learning models will be used to identify the features.

In an eccentricity from the research described above, to the authors' knowledge, none of the existing techniques fulfil the need of the nocturnal hours thermal image detection. The uniqueness of this research will increase the research interest further and also it paves the path to the young researcher. The next section describes the requirements of dataset acquisition, dataset details, and system requirements.

III. Dataset and System Requirements

This section discusses the dataset collection procedure and the system specifications used for the research. The images are collected from hospitals with high-security conditions.

A. Dataset Acquistion and Details

The data acquisition is one of the individualities of this research work. The dataset was collected in real-time while the car was driving. A thermal camera has been fixed close to the mirror on the driver's side, and the vehicle has been driven at different speeds to capture the different postures of the deer. Timing plays a vital role in data collection because the nocturnal image has to be captured only at night. After the sunset, most of the objects will have high temperatures. For example, the road, stone, trees, iron materials, etc., has high temperatures, the same as the hot-blooded animals. These measures have to be considered when capturing nocturnal infrared thermal images.

The dataset includes 1500 images captured during both dawn and dusk of the day. These images are raw images, and before filtering, they all have the camera logo, the date, and the capture time of the image. If all the images have the same information, then the overfitting error will occur. To avoid those problems, the data have to be filtered using computer vision techniques. The resolution of the images is 640x480, 640x520, 480x480 pixel size. The data collection was performed at night during all seasons, such as spring, fall, and summer. FLIR ONE Pro [6] is the thermal used to capture all the images; the future can be extended with a high-resolution thermal camera that will give high-resolution images and will increase the information in the image. Both the roadside and highways have been considered for the research experiments.

The Important challenges faced during the data acquisition are distraction due to another vehicle, hibernation due to natural calamities, animals' unavailability due to rainy nights, and habitat noises. Different postures will give different information; while doing research, all types of postures have to be considered for training the model. The complexity of the proposed system is based on the complexity of the dataset; the dataset complexity increases based on the temperature of that particular day. Usually, wildlife animals have hot blood in their bodies, especially the deer, which will come under the hot-blooded animal category; that is the reason it can be captured using the infrared camera.

The dataset was randomly split into 60% or 70 % training, 20% validation, and 20% testing (60+20+20), with labeled images with deer as one and without deer as 0. The network has to be trained with the data input and output, validated, and tested by using the same processed data.

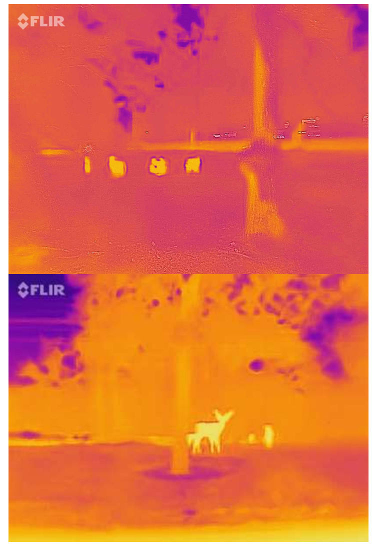

Fig. 1. Thermal images using FLIR thermal camera. Raw image before preprocessing using the computer vision technique.

Fig. 1. describes the raw image download from the camera before filtering or cropping. In both the image the presence of deer is seen clearly. Due to hot blood the brightness of the animal is high than any other object in the same image. This is the important feature of the thermal image.

TABLE I. Dataset Details

Descriptions	No. of Images
Total Images	1500 (Raw)
Images used	1068 (Filtered)
With animal	714
Without animal	354
Training & Validation	854
Testing	214

Table I shows the summary of the image dataset used for the proposed model. Out of 1500 images, the final filtered image for the whole process is 1068. From this tabulation, the number of images from the 1068 category, 854 images, are considered for validation and training. The remaining 214 images will be used for testing purposes. The images of the different scenarios are considered for the model's training. The standard size of the image has to be maintained for the execution. The size of the image has to be resized based on the algorithm considered in this research.

B. System Specifications (High Performance Computing)

Here, the comparison of both CPU and GPU will be shown. To implement the CNN because of the smaller number of images in the dataset, the CPU has been chosen, and the output will be executed. The open-source platform Python 3.9, with compatible TensorFlow and CUDA, is used for the GPU requirements. Windows 10 operating system with 8GB GPU with multiple clusters and nodes has been considered [15]. Both storage spaces are considered to check which one is faster compared with the others.

TensorFlow and Keras are versatile open-source libraries developed for machine learning and deep learning applications. They are fundamental for detection and identification, and are commonly used for large datasets and highly complicated networks, making them a comprehensive choice for a wide range of applications.

Recurrent neural networks, image recognition, classification, natural language processing, and partial differential equations are the different possibilities of the Keras and TensorFlow-based simulations to train and validate the given dataset. The machine learning network model is designed for the TensorFlow 2.5 version and is used to design the machine learning network model efficiently. It is a powerful, robust experimentation model building for the research. A combination of tools and libraries includes the extension toolkit called Compute Unified Device Architecture (CUDA).

For CPU, these configurations, the device setup used for the complete training, validation, and testing is a MacBook Pro with specifications as follows,

- Processor - 2.3 GHz
- Dual-Core - Intel Core i5.
- Memory 8 GB
- 2133 MHz LPDDR3.
- Graphics - Intel Iris Plus Graphics 640 1536 MB.

Aforesaid methods all have been experimented in our research. Mainly to check the speed and performance, all the configuration and dataset will help to complete the whole process. The next section explains about the proposed algorithm and methodology.

IV. Methodology

This section explains the different methodologies used to detect animals in the dataset. The confusion matrix depicts the detected and missed information after using the trained model. Based on the algorithm, the image size has to be adjusted using Python code. In CNN, the input pixels (height, width) various respectively based on the type of the algorithm [20]:

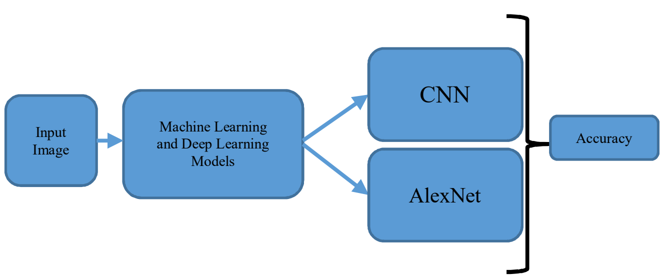

Fig. 2. Network Model – Block Diagram

Figure 2 includes the machine learning networks, CNN and deep learning network, AlexNet. This section elaborately explains all the basic principles of Artificial neural network models. The designed models are trained and tested with the provided data set, and the trained model is used to classify with and without animal.

A. Convolution Neural Network Model

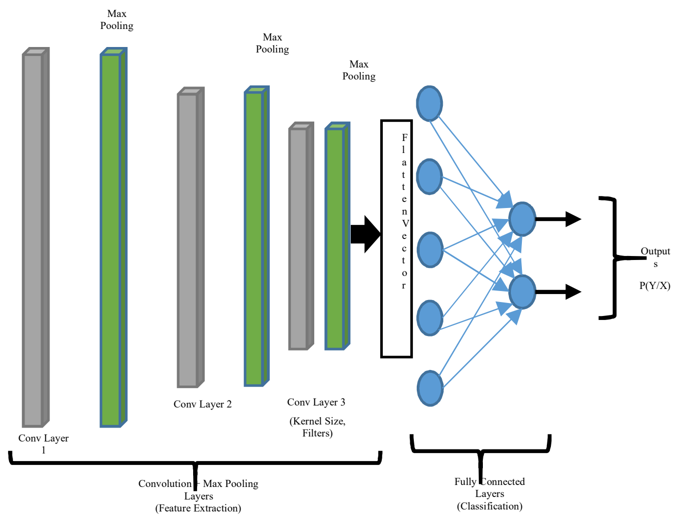

Fig. 3. Convolution Neural Network – Block Diagram

Fig. 3. Shows the detailed structure of the convolutional neural network, which includes convolution layers, filters, kernel size, maxpooling layers, flatten vector, fully connected layers, and final accuracy output. All these parameters help to identify the features of the input image. Fine-tuning these parameters can change or increase the accuracy efficiency [17]. The output is optimized by adjusting the filter and kernel size, with the activation function playing a crucial role in this optimization process.

The algorithm used in this convolution neural network is as follows,

Step 1: Data Collection by FLIR ONE camera.
Step 2: Preprocess the data to remove unwanted information without any loss in the original information.
Step 3: Append all the feature vectors in 1D form.
Step 4: Split the data for training, validating, and testing.
Step 5: Train the model using the processed dataset.
Step 6: Test the model and find the confusion matrix
Step 7: Testing accuracy and loss calculation.

These are the general steps for the convolutional neural network. The image resize is a critical step in the process, as the algorithm requires the image size to be adjusted for an output to be determined. If the image size is not changed as per the algorithm's requirements, no output will be determined.

B. Deep Learning Methoddology

Fig. 4. depicts the detailed architecture of the AlexNet which includes 5 convolution layer, 3 max pooling layers, 2 normalization layers, 2 fully connected layers, 1 SoftMax layer

[22]. This is the first architecture used in the GPU environment. The usage of GPU will increase the performance much faster than CPU; the storage space increases the time takes for every epoch in iteration. It has both the trainable and non-trainable parameters. The convolution layer and max pooling layer extracts the information from the image and forward it to normalization, the normalized value will be connected to the fully connected layer and then to SoftMax. Here the ReLu activation function is used to increase the accuracy with less loss.

The size of the input image to this model is 227x227x3, based on the number filters and strides the size of the image will be reduced and given to the maxpooling layer. The size of the max pooling layer will generate the feature map of the input image. If needed data augmentation can be used to extend the number of input images. In data augmentation, process like stretching, shrinking, flipping, cropping, rotating is used to generate the new image with the loss of information.

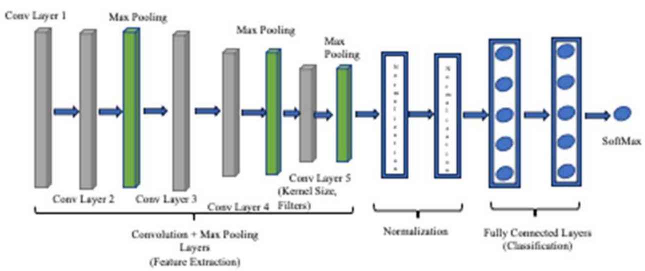

Fig. 4. AlexNet Detailed Architecture

V. Experimental Results

This section shows the pictorial representation of the outputs generated in the research, and a comparison will be shown to identify the best network for detecting animals. The tabulation will give the values; the bar chart will show an accuracy graph in pictorial form, and a confusion matrix for training, validating, and testing the machine learning and deep learning network model results will be described. The accuracy of both the proposed network testing, validation, and training has been shown in this section. The sample confusion matrix describes the number of correct predictions and predictions or detections from the full dataset.

Fig.5. illustrates the practical application of the histogram of oriented gradients (HOG) image output with the original image, both with and without a logo [21]. This comparison is crucial in preventing overfitting errors. The presence of a camera logo creates similarities with the image, leading the network model to interpret the logo as a feature of the input, thereby causing overfitting errors. This practical example underscores the importance of our research in the field of animal detection.
The histogram of oriented gradients processes as follows,

- Step 1: Normalization.
- Step 2: Gradient measure.
- Step 3: In each 8x8 cell, generate the HOG.
- Step 4: Global L1-Norm(16x16).
- Step 5: Generate the HOG feature vector.

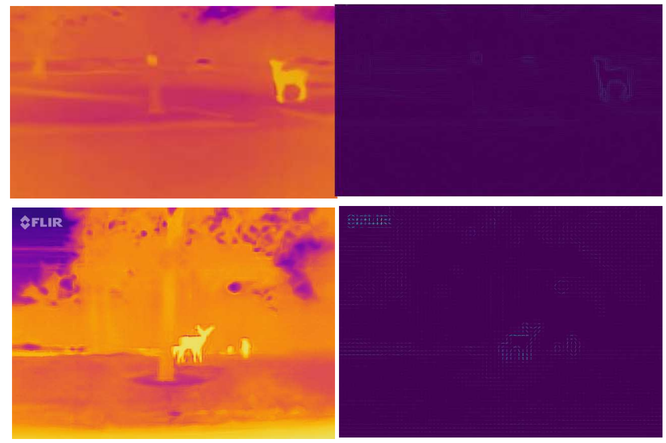

Fig. 5. Two Sample Processed Images With and Without the Camera Logo, top – without logo, bottom – with logo

$$\text{Accuracy} = \frac{\text{True Positive+True Negative}}{\text{Total}} \qquad (1)$$

Accuracy of the network for CNN will be calculated by the equation in (1). From the confusion matrix the detection and mis detection will used to calculate the accuracy.

A. Model Parameter of the Designed Networks

Table II depicts the parameter specification of the designed model. The benchmark for CNN is shown in Table II. Two convolution layers with one global max pooling and dense layer are shown here. The trainable and non-trainable parameters are also shown in the table. Based on the input thermal or HOG image, the parameters are defined in Table II. Here, the non-trainable parameters are zero. The conversion of the thermal image into HOG makes the non-trainable parameters zero [19]. To strengthen the model performance, reduce the non-trainable parameter so that the model training will be better if the model training process is better than the accuracy during testing.

TABLE II. MODEL PARAMETERS WITH SPECIFICATION FOR CNN

Layer (type)	Output shape	Parameters
Conv1d_7 (Conv1D)	(None, 2696, 125)	2000
Conv1d_8 (Conv1D)	(None, 2692, 125)	78250
Global_max_pooling1d_4(Global)	(None, 125)	0
Dense_4(Dense)	(None, 1)	126
Total Parameters: 80,376		
Trainable Parameters: 80,376		
Non-Trainable Parameters: 0		

TABLE III. MODEL PARAMETERS WITH SPECIFICATION FOR ALEXNET

Layer (type)	Output shape	Parameters
Conv2d (Conv2D)	(None, 110, 110,192)	34955
Conv2d_1 (Conv2D)	(None, 55, 55,96)	715466
Conv2d_2 (Conv2D)	(None, 27, 27,256)	1267488
Global_max_pooling2d_4(Global)	(None, 13,13,256)	0
Conv2d_3 (Conv2D)	(None, 27, 27,256)	36734953
Global_max_pooling2d_4(Global)	(None, 6,6,256)	0
Conv2d_4 (Conv2D)	(None, 27, 27,256)	615466544
flatten	(None, 8612)	0
Dense_1(Dense)	(None, 4000)	4758986402
Dropout(Dropout)	(None, 4069)	0
Dense_2(Dense)	(None, 1000)	3746998
Total Parameters: 3480376		
Trainable Parameters: 3480376		
Non-Trainable Parameters: 0		

Table III explains the model parameters of the AlexNet deep learning algorithm. Here, as with the machine learning algorithm, the convolution, max pooling, dense, and dropout layers are shown with their parameter values. The trainable and non-trainable parameters are shown. The preprocessing of the raw image makes the image more informative, and all the features in the image are trainable only. This is why the camera brand logo has been cropped from the image.

B. *Accuracy of the benchmark and proposed netowrk model (CNN & AlexNet)*

Table IV. enlightens the accuracy of the benchmark and proposed network model. Here, three types of epochs have been considered for testing the proposed and benchmark network model. The data split for training validation and testing will be considered as 70% - 30% or 80% -20%. For the benchmark model, if the number of epochs is 50 and the activation function is ReLu, sigmoid and SoftMax are considered. The testing percentage accuracy is 89%, for 100 epochs – 90%, and for 150 epochs – 91%.

For the proposed AlexNet neural network model, the same three different epochs such as 50, 100, and 150 is considered. Here the activation used is ReLu. The ReLu function will reduce the loss of information of the original image without any problem. Here the validation accuracy is 88%, 91%, and 92%. As same as benchmark algorithm, the split of data for training, validation ,and testing is same. Different convolution layer and dense layers are considered, and those values are shown in the parameter specification. The conclusion section will provide a comprehensive comparison of both the machine learning and deep learning models, offering insights into their respective strengths and weaknesses.

TABLE IV. ACCURACY OF THE BENCHMARK AND PROPOSED MODEL

Methods	No. of Epochs	Validation Accuracy (Round-off)
HOG + Convolutional Neural Network	50	89%
	100	90%
	150	91%
AlexNet (Proposed Method)	50	88%
	100	91%
	150	92%

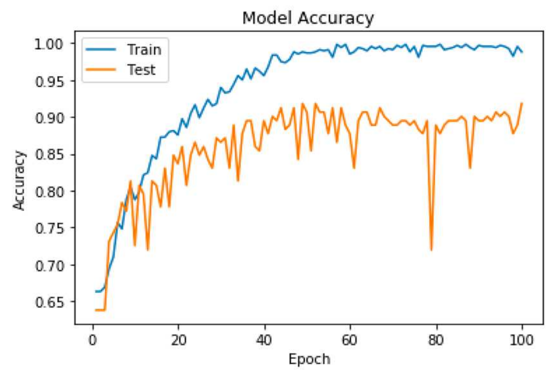

Fig. 6. Sample Accuracy Graph for CNN and AlexNet for 100 epochs

Fig.6. shows the testing and training accuracy graph of the CNN and AlexNet model for 100 epochs. The y axis defines the percentage and x axis defines the number epochs for the testing and training. The graph shows around 91% of accuracy in testing and around 100 in training. To avoid the overfitting error in both the cases the computer vision preprocessing have to be used.

C. *Training time analysis for benchmark and proposed netowrk model (CNN & AlexNet)*

Training time analysis is an important factor if the dataset is much bigger in size. The CPU and GPU will be used differently, and the operating system will make a difference in time analysis. Table V shows the values for all ten experiments conducted, and only three samples are shown in the accuracy section. For all ten experiments with different layer combinations and different epochs, the time taken will be different. For the total number of input images, 1068 214 images will be the ratio used for the testing, and the remaining will be for training and validation. The upgradation of GPU is considered because of the complexity of the deep learning network model. It is much faster than the CPU, not only for CNN but also for AlexNet. The minimum number of seconds is 0.21 seconds for the whole process to complete. The maximum is almost one minute because of the smaller number of images and the parameter fine-tuning setup.

TABLE V. TRAINING TIME ANALYSIS

No. of Experiments	Time / 214 frames (sec)	
	CNN (CPU)	AlexNet (GPU)
1.	2.11	0.27
2.	2.01	0.35
3.	2.05	0.45
4.	1.06	0.21
5.	2.08	0.48
6.	2.06	0.42
7.	3.01	0.59
8.	2.05	0.41
9.	2.07	0.43
10.	1.05	0.20

Fig.7. is the pictorial representation of 10 experiments for both the CNN and AlexNet in both the CPU and GPU. From the bar chart, it is clear that the performance of the GPU is much better than the CPU. The blue color in the chart represents the CNN and CPU usage, and the orange color represents the AlexNet with GPU configurations mentioned in the System requirement section.

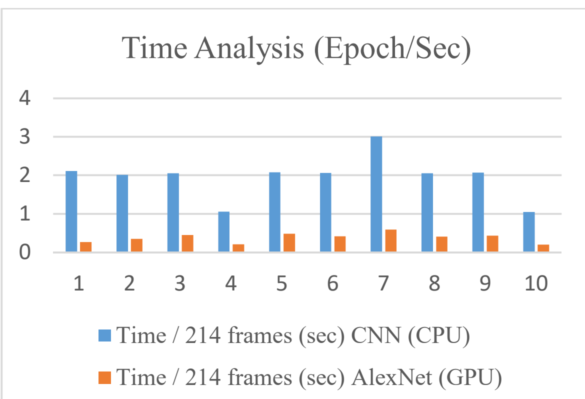

Fig. 7. Bar chart representation of time taken in secs per epoch for testing in proposed and benchmark network model

D. Confusion Matrix

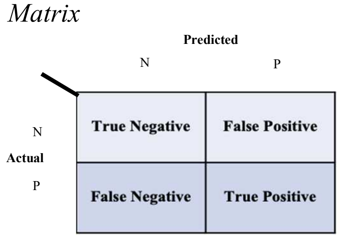

Fig. 8. Confusion Matrix

<u>Confusion Matrix</u>

[[50 13]

[8 143]]

True Positive – 143 (Animal Detected as yes)

True Negative – 50 (Animal Detected as no)

False Negative – 8 (False Detection as no)

False Positive – 13 (False Detection as yes)

Fig. 9. Sample Confusion Matrix of CNN and AlexNet Network Model

As shown in Fig. 9., the diagonal elements from left to right are the mis detection values, and the opposite diagonal from right to left are the detection values. The Fig. 9. shows the sample confusion matrix generated from the proposed model. Fig. 8. Will be the general description of the confusion matrix, where the true positive and true negative are the detected values, and false positive and false negative are the mis detected values.

VI. CONCLUSION AND FUTURE WORK

In the era of large data, the exponential growth of machine learning and deep learning network models, as well as the need for technologies for transportation application, applied Artificial Intelligence technologies provide promising solutions. The deep learning – AlexNet and machine learning - convolutional neural network were compared in all parameters like time analysis and model accuracy to identify the prominent performance model. In this research, two goals have been discussed. The first one is the usage of GPU instead of CPU and how it increases the speed and performance of the model, and the second is deep learning network model comparison with the benchmark CNN. From the comparison of this research, the GPU produces five to eight times faster than the CPU. Compared with CNN, AlexNet produces one percentage higher accuracy. Also, the HOG preprocessing technique is used before getting into the network model. Computer vision technologies filter unwanted noise, especially for thermal images collected during the nocturnal hours. So that the image going into the network model will have all the required information for the training and testing, the fast-growing scientific data will help in the transportation world to autonomous vehicles. Humans have a chance to escape from dangerous roadkill. By avoiding roadkill, the insurance industry lowers the economic expenses spent on refurbishing the car. In the future, the seasons will also be considered for the whole process. These AI technologies will make a big difference in the transportation market, and utilizing artificial intelligence systems will fill the gap between traditional methods. The future work can be extended by considering more images and comparing all deep-learning network models to find the optimized system. In order to serve the community and future transportation, our current research will focus on developing neural network models for complex data and complicated application.

REFERENCES

[1] https://www.fhwa.dot.gov/publications/research/safety/08034/exec.cfm

[2] Christiansen, P., Steen, K. A., Jørgensen, R. N., & Karstoft, H. (2014). Automated detection and recognition of wildlife using thermal cameras. Sensors (Switzerland), 14(8), 13778–13793. https://doi.org/10.3390/s140813778

[3] Guo, L., Ge, P. S., Zhang, M. H., Li, L. H., & Zhao, Y. B. (2012). Pedestrian detection for intelligent transportation systems combining AdaBoost algorithm and support vector machine. Expert Systems with Applications, 39(4), 4274–4286. https://doi.org/10.1016/j.eswa.2011.09.106

[4] Sibanda, V., Mpofu, K., Trimble, J., & Zengeni, N. (2019). Design of an animal detection system for motor vehicle drivers. Procedia CIRP, 84, 755–760. https://doi.org/10.1016/j.procir.2019.04.175

[5] Yuvaraj. M, A. Martinez-Molina, and M. Alamaniotis, Comparison of Image Segmentation, HOG and CNN Techniques for the Animal Detection using Thermography Images in Automobile Applications, 12th Int. Conf. Information, Intell. Syst. Appl. IISA 2021, 2021.

[6] Cilulko, J., Janiszewski, P., Bogdaszewski, M., & Szczygielska, E. (2013). Infrared thermal imaging in studies of wild animals. European Journal of Wildlife Research, 59(1), 17-23.

[7] Yuvaraj. M, A. Martinez-Molina, Dimitrios Miserlis, Hermilo Hernandez and M. Alamaniotis, Intelligent System Utilizing HOG and CNN for Thermal Image-Based Detection of Wild Animals in Nocturnal Periods for Vehicle Safety, Applied Artificial Intelligence, DOI: 10.1080/08839514.2022.2031825, 2022.

[8] P. Dutta, P. Upadhyay, M. De, and R. G. Khalkar, "Medical Image Analysis using Deep Convolutional Neural Networks: CNN Architectures and Transfer Learning," *Proc. 5th Int. Conf. Inven. Comput. Technol. ICICT 2020*, pp. 175–180, 2020.

[9] Yuvaraj. M., A. Martinez-Molina, & M. Alamaniotis, Design and Implementation of a nocturnal animal detection intelligent system in Automobile Applications. International Conference on Transportation and Development 2021 – Transportation operations technologies and safety, American Society of Civil Engineers (ASCE), 2021. 438-449, https://ascelibrary.org/doi/ 10.1061/9780784483534.038

[10] Qiang Zhu; Mei-Chen Yeh; Kwang-Ting Cheng; Avidan, S, "Fast Human Detection Using a Cascade of Histograms of Oriented Gradients," Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on , vol.2, no., pp. 1491- 1498, 2006.

[11] Yuvaraj. M, A. Martinez-Molina, and M. Alamaniotis, Active Advanced Arousal System to Alert and Avoid the Crepuscular Animal Based Vehicle Collision, Intelligent Decision Technologies, vol. 15, no. 4, pp. 707-720, 2021.

[12] Affonso, C., Rossi, A. L. D., Vieira, F. H. A., & de Carvalho, A. C. P. de L. F. (2017). Deep learning for biological image classification. Expert Systems with Applications, 85, 114–122. https://doi.org/10.1016/j.eswa.2017.05.039

[13] Mowen, D., Munian, Y., & Alamaniotis, M. (2022). Improving Road Safety during Nocturnal Hours by Characterizing Animal Poses Utilizing CNN-Based Analysis of Thermal Images. Sustainability, 14(19), 12133. https://doi.org/10.3390/su141912133.

[14] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in *NIPS*, 2012.

[15] "High performance computing lab" – Texas A&M University San Antonio, June 2024, https://hprc.tamusa.edu/

[16] Yuvaraj. M, A. Martinez-Molina, and M. Alamaniotis, Intelligent System for Detection of Wild Animals Using HOG and CNN in Automobile Applications, 11th Int. Conf. Information, Intell. Syst. Appl. IISA 2020, 2020.

[17] Liu, S., & Liu, Z., "Multi-Channel CNN-based Object Detection for Enhanced Situation Awareness," 2017.

[18] San Antonio News, Accessed 2020, https://www.mysanantonio.com/news/local/article/Here-s-where-people-most-often-crash-into-animals-13266975.php#item-85307-tbla-1

[19] Islam KTo, Raj RG, Al-Murad A, "Performance of SVM, CNN, and ANN with BOW, HOG, and image pixels in face recognition," 2nd International Conference on Electrical and Electronic Engineering,. https://doi.org/10.1109/ CEEE.2017.8412925, 2017.

[20] https://blog.goodaudience.com/introduction-to-1d-convolutional-neural-networks-in-keras-for-time-sequences-3a7ff801a2cf

[21] https://www.learnopencv.com/histogram-of-oriented-gradients/

[22] Xin, M., Wang, Y. Research on image classification model based on deep convolution neural network. J Image Video Proc. 2019, 40 (2019). https://doi.org/10.1186/s13640-019-0417-8

[23] Wang, Y., Zhang, D., Liu, Y., Dai, B., & Lee, L. H. (2019, February 1). Enhancing transportation systems via deep learning: A survey. Transportation Research Part C: Emerging Technologies, Vol. 99, pp. 144–163. https://doi.org/10.1016/j.trc.2018.12.004