Saving Christmas: Preventing Reindeer-Vehicle Collisions with Thermal Imaging

Benjamin Kämä[1], Ella Peltonen[1]

[1]Empirical Software Engineering in Software, Systems, and Services, University of Oulu, Finland

Abstract

Around 4,000 reindeer-vehicle collisions (RVCs) happen annually, leading to losses in reindeer and human lives. The annual vehicular property damage from RVCs is estimated to cost 15-20 million euros. Despite this, there have been only a limited number of attempts to automate RVC prevention. To save Rudolf and the rest of the reindeer, this paper tackles RVCs by proposing a thermal imaging-based solution that would be feasible to implement in a vehicular setting. The system relies on an inexpensive FLIR thermal camera that can be attached to most smartphones. Compared to regular cameras, thermal camera performance is not hindered by low lighting conditions or poor weather, especially winter conditions in the reindeer native lands in Lapland. For recognition of the reindeer, we utilise a Convolutional Neural Network (CNN) algorithm that we evaluate with open-sourced and self-collected FLIR thermal camera datasets. Our solution has high accuracy on differentiating reindeer from other pedestrians, 0.998 at best, and thus compares well to other camera-based image recognition algorithms used for general animal detection. In future work, the model can be integrated into a driver warning system or an autonomous vehicle's control system to avoid the reindeer on the road and prevent collisions.

Keywords

Thermal Imaging, Image Recognition, Smart Vehicles, Animal Vehicle Collision

1. Introduction

In Finland, there are around 4,000 reindeer-vehicle collisions (RVC) per year in which both reindeer (*Rangifer tarandus tarandus*) and human lives are lost and property damage occurs. Annual vehicle damage costs from RVCs are around 15-20 million euros, and reindeer owners claim circa 2.5 million euros in insurance settlements for reindeer lost in road traffic accidents [1]. There is also the emotional cost for the humans involved, as RVCs can cause a loss of feeling safe on the roads. As reindeer are pack animals, many can simultaneously become involved in the same RVC event, making the collisions more dangerous and increasing the monetary and emotional load. Although the costs are high, no successful solutions have been found to automate RVC prevention. The current general method for warning oncoming traffic about reindeer sightings is to flash high-distance beams manually. However, automatically detecting any animals on the road, let alone brown-grey reindeer (see Figure **??**), is challenging, especially in dark and less optimal weather conditions [2]. In daylight, standard camera-based methods can be utilised [3]. Still, their capability in less than optimal lighting conditions, such as in the grey evenings and pitch-black nights when reindeer are the most active, remains an open question. To make it possible to warn the driver and other traffic about reindeer on the road or even automatically avoid them in modern autonomous driving vehicles, there is a need for a reliable, weather and light-tolerant solution for reindeer detection.

This work presents a thermal imaging-based method and a CNN image recognition algorithm for detecting reindeer and differentiating them from other pedestrians and animals, especially people walking on the rural roadside. For the sake of automating evading movements of the vehicle, it is crucial to know when the object is indeed a human or reindeer. We presume that a pedestrian walking the roadside would continue progressing straight [4], compared to reindeer, which are semi-domesticated animals [5] and prone to unexpected movements and running in front of traffic. In contrast to completely wild animals, such as moose, elk, and regular deer, reindeer are not scared of cars and may wander

TKTP 2025: Annual Doctoral Symposium of Computer Science, 2.—3.6.2025 Helsinki, Finland

benjamin.kama@oulu.fi (B. Kämä); ella.peltonen@oulu.fi (E. Peltonen)

0009-0005-9107-9168 (B. Kämä); 0000-0002-3374-671X (E. Peltonen)

on the roadside for prolonged periods of time. Reducing reindeer collisions is an especially pressing problem in Northern Scandinavia and Finland; they are considered culturally essential animals, an expensive delicacy (a herder owns each reindeer), and their semi-domestic nature comes without fear of traffic. In the future, we can expand to other wildlife, such as elk/moose, of which avoidance differs from tightly-packed reindeer.

Regular cameras perform well in daylight and during good weather, but bleak autumns and winters hinder visibility and prediction accuracy in poor weather and low-light conditions. These are also the periods when most RVCs happen. Our proposed solution combines a CNN-based image recognition model, an off-the-shelf FLIR thermal camera, and an Android smartphone. Compared to some existing drone-based reindeer detection methods [6], our solution is lightweight to implement, and the device setup is cheaper. The thermal camera images used to train, test, and evaluate the model are an open-source dataset from Roboflow and manually self-collected FLIR thermal camera images from local reindeer. To summarise, our main contributions are the following:

- We present a CNN-based machine vision algorithm that detects reindeer from thermal images.
- We show that detecting reindeer from thermal camera images is possible in an accurate manner, with an accuracy of 0.998 at best, with similar or higher evaluation results than reported in the previous works.
- By evaluating our solution with both an open-source dataset and self-collected data, we can show that our method generalises over different datasets without losing much of its accuracy.
- We compare our solution to other image-based animal detection and collision avoidance systems and show that we can get equally good results with an inexpensive FLIR camera addition to an Android smartphone.

2. Reindeer-collision (RVC) Prevention Methods

Reindeer, hoofed ruminant mammals, can be classified under the family *Cervidae*. Different methods can be utilised for cervid-collision prevention. Next, we discuss reindeer-vehicle collision (RVC), or in more general terms, cervid-vehicle collision (CVC) prevention methods. Some of them include Advanced Driver Assistance Systems (ADAS) technology, but many are not integrated into the vehicles at all. Table 1 summarises some most popular CVC prevention methods, out of which we focus on digital ones. The column "Still in use?" refers to their current status in Finland.

Method	Source	Still in use?	Note
Relocation	[7]	No	Not used in Finland
Hunting	[7]	No	Hunting not applicable to reindeer in Finland as they are considered livestock
Overpasses/crossings	[7, 8]	Yes	Located in southern parts of Finland only
Vegetation removal	[7]	Yes	Done near the roads
(Street) lights	[7]	Yes	Focused near cities and settlements; thus, rural areas rarely have lights. RVCs are a higher risk in rural areas, and the road speeds are higher on rural roads.
Fencing	[7]	Yes	Used in Finland for other cervids, e.g. elk/moose [9]
Road signs		Yes	In use in Finland country for reindeer, elk, and deer
Porokello	[10, 11]	No	Mobile phone application where professional drivers reported reindeer sightings and users were notified of them
Varoporoa		No	Website unavailable at the moment.
Radio transmitters	[12, 13]	Not clear	These were radio transmitters fitted on the reindeer and LED warning lights installed on the roads. They had a lot of technical issues they were working on.
Night Vision	[7]	Not clear	Thermal camera setup connected to NV ECU. First such system in production cars. Made for Audi, BMW, and Daimler. Used FIR or Far Infrared Camera.
Riistavaroitus	[14]	Yes	Mobile application that warns the user when there is a risk for CVC. Not fully-automatic as it requires human input.

Table 1: Summary of existing cervid-vehicle collision prevention methods.

The first example is a service explicitly provided for aiding drivers to avoid reindeer on the roads, a smartphone application called Porokello. It was a mobile app that crowdsourced reindeer sightings. The professional drivers of trucks and delivery vehicles were asked to log reindeer sightings in the

application, and the application would then notify other users (such as personal vehicle drivers) if there was a heightened risk for RVCs [1]. A notification would be sent to all the users for 30 minutes and around 750m of the observation spot. The app was downloaded over 65 thousand times and had 300-600 daily users during its operational period. During the apps' operation in 2017-2018, RVCs went down 30% from 2011-2015. However, for roads with fewer RVCs, the change was either less than average or the number of RVCs had increased [1]. Thus, it is impossible to say that warnings have had a direct effect on RVCs or a change in the number of accidents. However, nine out of ten questionnaire respondents said they had changed how they drove after getting a Porokello warning, for example, paying better attention to driving and lowering driving speed [1].

A similar application to Porokello is the new Riistavaroitus application [14] that launched only recently, at the beginning of 2025. In the Play Store description[1], the application is claimed to be able to warn about active reindeer areas, as well as about the end of fencing, active hunting sites, and about RVCs that have already occurred. However, such a solution still requires human input from hunters who have been notified about an RVC and who go to handle hurt animals [14]. Thus, there can be significant delays in issuing the RVC warnings, which reduces their information value. Another digital service is the Varoporoa website[2]. It is map-based, showing all reindeer vehicle crashes in the last 21 days, in addition to some historic data tables. The last year shown is 2021, with no data for 2022 or 2023. The map does not show any data at the moment. However, a separate page shows yearly RVC statistics between 2011 and 2024.

A different type of RVC prevention method is radio transmitter collars installed on 235 reindeer in Saltfjellet, Norway [12, 13]. They had receivers on the roads equipped with red LED lights that started blinking when a reindeer with a transmitter collar was within 50-100m of the receiver. They had 41 receivers on a 4.5km long road section. Reindeer deaths were recorded before and during the two-month experiment from February to April 2018. Other transmitters were tested on 200 reindeer on the same road from December 2018 to April 2019, resulting in 15 deaths from December 2017 to February 2018, when their experiment started, but no RVCs. By 2019, there were no deaths of transmitter-collared reindeer. However, there were several difficulties with the transmitters. By the end of their first research period, 25% of the receivers had stopped working, and they suspected this was due to battery shortage. During the second experiment, around mid-February 2019, 8% of receivers and 35% of transmitters were not working [12, 13].

Utilising thermal cameras for CVC prevention has been proposed but rarely experimented with in real-life conditions. For example, Zhou et al. [15] built a thermal camera setup using a large FLIR thermal camera with a frame grabber, two motors, and power sources. The work focused more on the technical production of this thermal camera setup than on the efficiency of the algorithmic solution. Lyo et al. [16, 6] have recently proposed a method for deer identification by using Unmanned Aerial Vehicles (UAVs) equipped with thermal cameras. However, this method, even if successful, requires the use of a UAV and is thus less than optimal for driving situations, as reindeer tend to move in a relatively large area. The official reindeer herding area in Finland covers over 120.000 square kilometres of land and around 80.000km of public roads, making it impossible to cover by UAVs. Thus, a vehicle-based solution seems to be the most feasible.

Night Vision (NV) Forslund and Bjärkefur [7] has been the first animal detection system for cars in customer markets. It uses a far-infrared camera (FIR) that can be mounted in the front on top of the vehicle grille. The FIR was connected to an NV electronic control unit (ECU). The NV ECU was further connected to the display, the instrument cluster, and the Head-Up Display (HUD). The enhanced video with highlights around particular objects was then shown to the driver on a display. Warnings were given to the driver with warning symbols or lights in the instrument cluster or HUD. For the object detection, the NV used boosting [17, 18] and a cascade structure [19]. Using a smaller feature evaluation set, the model combines several weaker classifiers to reject the non-animal images. However, the model's accuracy was not reported beyond the accepted error rate of one false detection per year, and

[1] https://play.google.com/store/apps/details?id=com.innotrafik.riistavaroitus&hl=fi
[2] https://www.varoporoa.fi/

Figure 1: Reindeer in Ranua, Finland.
Image by authors.

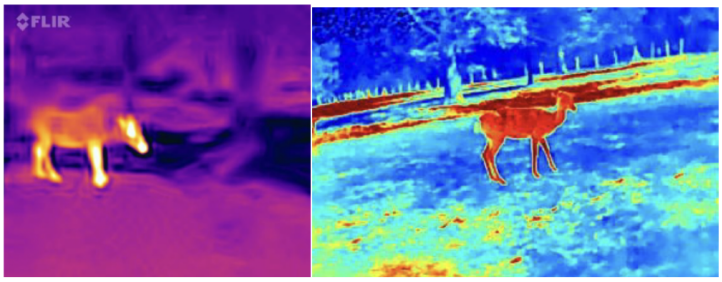


Figure 2: Thermal images of deer: on the left from the self-collected and on the right from the open dataset [21].

vehicle manufacturers had accepted the product. Additionally, the system can detect ”almost all relevant animals” but does not state which animals it can or cannot detect. The reactions of the driver/vehicle are dependent on the animal detected. For example, a horse/rider pair can be slowly overtaken at a safe distance, and sighting a moose requires a complete stop at any cost, as a crash would easily be fatal. Reindeer, conversely, have to be overtaken with caution as they can change their direction at any moment. Due to a lack of reported evaluation metrics, it is hard to evaluate how well the NV system performs against other models. Night Vision was tested with data for eight years through different seasons and weather conditions. They reported having gathered several hundred thousand labelled images and varied animals, but the data is not publicly available.

However, Riistavaroitus is the only technological solution that is still in use today. The Porokello app was terminated, and the Varoporoa website is not available at the moment. It is unclear if NV and the radio transmitter collar project by [12] continue. On the other hand, the radio transmitter project seemed to be successful as no reindeer were killed on the test roads during the test period. It is also possible that Autoliv has continued to develop NV, perhaps integrating it into another project with a different name. The reasoning is that since ADAS is increasing in popularity and importance [20], it makes sense that there is continued interest in sensor technology development [2].

3. Methods and Materials

3.1. Research Context

In this research, we consider animal-vehicle collisions (AVCs), more specifically cervid or reindeer vehicle collisions (RVC), that happen on public roads, including motorways, dual carriageways, and similar. The vehicles considered are all road vehicles, varying from personal cars to trucks and other heavy machinery. AVCs can happen on any road with vehicle traffic, although reindeer are more common in rural settings. Some areas also have more reindeer, and thus, the risk for RVCs is higher.

The current estimation of semi-domesticated reindeer in Finland is 203.700 animals [22]. Reindeer are well accustomed to living in the Arctic areas with feet that have evolved to walk in deep snow. Their movement on the roads can be determined by various factors such as the time of year or day, foraging, or breeding season [10], but may also not follow a specific pattern, including crossing the same roads several times a day. The stag can weigh between 90-180kg and grow to a height of around 110cm. The doe is smaller at 90cm and 60-100kg. As a newborn, they weigh only 4-6kg. The doe can live up to 20 years, and the studs over 10 years. Their colour can differ widely, ranging from black to white, as can be seen in Figure 6a and 1, and both the stag and the doe grow antlers that the stags drop in the Autumn, and the does drop in the Spring [23]. In Fennoscandia, reindeer live in domesticated herds, except in Norway, where there are also mixed herds with reindeer from both wild and domesticated origins [5]. Reindeer is an example of wildlife habituation to roads and vehicles, as they may have flight responses to vehicles, but only in rare observations [5].

3.2. Research Methodology

This research follows, on a general level, the Constructive Research Approach (CRA) [24, 25]. The focus is on creating an artefact specifically to solve a specific real-life problem, in this case, the occurrence of RVCs. As we have identified in Section 2, there is a knowledge gap on existing RVC prevention methods: 1) The current systems are either not in use anymore (most digital services) or ineffective for broader applicability (such as radio transmitter collars), 2) Even if a commercial solution (Night Vision) seems to exist, there are no proofs of its performance for RVC prevention, 3) Camera-based animal-detection algorithms are problematic on nighttime when RVCs are the most common, 4) Thermal-based state of the art solutions focus on UAV perspective, leaving little applicability to real-life driving situations, and 5) No reindeer-specific image recognition algorithm has been presented in the literature that would be evaluated in dark/night environments. To fill this gap, we developed an artefact consisting of a CNN-based algorithmic solution that we validated with actual reindeer image data, combining open-sourced and self-collected images. The fit of the algorithm was evaluated with statistical metrics.

3.3. Open-source Datasets

Only a limited number of thermal image open-source datasets are available, especially those that include deer images. FLIR provides its own ADAS dataset, but no deer images are included. Choi et al. [26] presented a KAIST multi-spectral dataset, but there is no information about the classes used for detection. Roboflow is an image database with 90,000 public datasets with over 220,000 images with 805 class labels [27]; three such datasets were used in this study.

Svarzas Dataset is a deer image dataset (see Figure 2, right-hand side) published by Balzs Sznyi in 2023 [28], and it is available under a CC BY 4.0 license. It contains 2,599 thermal camera images of deer in an enclosure. The images are 416 by 234 pixels. The deer are annotated using Pascal VOC format. According to the README file, some preprocessing is done to the images. The pixel data was auto-oriented, and EXIF orientation was stripped. They were also resized to 416x416 with Fit Within; image augmentation techniques were not used.

Human Detection in IR Images Dataset contains thermal camera images of humans in various poses. It was published by username Karky in 2023 [29] under the MIT license. It contains 895 images, and the humans are annotated using the COCO format. The images are 640 by 640 pixels. Just as with the first dataset by Sznyi, the pixel data was auto-oriented, EXIF-orientation stripped, and the images were resized, although here they were resized to 640x640 with Stretch. Karky did not use any image augmentation techniques, just like Sznyi.

Thermal Human Detection Dataset is published by the username PNUSafetyNet in 2022 [30]. It contains 229 thermal images of humans in different poses. The images are sized at 640 by 480 pixels. These images were black and white instead of the coloured other datasets. The human images were annotated using the COCO format. The only pre-processing applied was auto-orienting the pixel data and stripping EXIF orientation. Image augmentation was not used.

3.4. Self-collected Dataset

We collected thermal camera images of reindeer to ensure that the algorithm could generalise over more than one deer thermal image dataset, as there was a limited number of open-source datasets with deer images available and no datasets with reindeer. We used the Teledyne FLIR One Pro LT (generation 3) thermal camera attached to the OnePlus 7t Android phone. The setup is shown in Figure 3. The resolution of One Pro LT is 80 x 60 or 4800 pixels. The horizontal field of view (HFOV) is 50° ±1°, and the vertical field of view (VFOV) is 38° ±1°. There are three different image modes: infrared, visual, and MSX or Multi-Spectral Dynamic Imaging mode. The file formats are radiometric JPG and MPEG-4 in iOS and MP4 in Android.

We used FLIR's own FLIR One application. The images were handheld without a tripod and in landscape and portrait orientation (see Figure 2, left-hand side). The scan was infrared only, and the colour was iron. The images were taken in Ranua Zoo. The animals were photographed in their

Figure 3: OnePlus 7t with Teledyne FLIR One Pro LT (Gen 3).

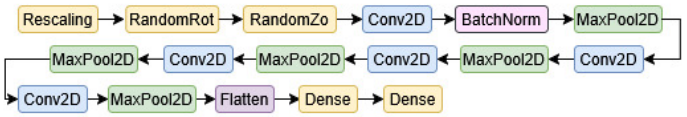


Figure 4: Layers of the CNN model 25.

enclosure from varying distances and angles. There were images of both single and multiple animals. The reindeer were all adults, but one juvenile was in the European Forest Reindeer enclosure. The pictures were taken during one day in October of 2023, when the weather was cloudy without rain. There were altogether 408 images, of which 135 were of the European Forest Reindeer (*Rangifer fennicus fennicus*), and 273 of the similar-looking domesticated reindeer (*Rangifer tarandus tarandus*). The size of the images was 640 by 480 pixels. The self-collected dataset is small for a machine-learning dataset, but data augmentation was used to correct this issue. We also used the open-source deer dataset from Roboflow to increase the number of deer images available for training, testing, and evaluation.

3.5. Model Creation

The Roboflow dataset had 842 images of deer and humans, a total of 1684 images. In the Ranua dataset, the deer images were self-collected, and the human images were collected from an open-source dataset from Roboflow. It had 281 images of deer and humans, totalling 562 images. In the mixed dataset, the deer images were a combination of images from the self-sourced and Roboflow sources, and all the human images were from Roboflow. It had 654 images of deer and humans, totalling 1308 images. The images used for model training, testing, and validation were selected from the Roboflow and Ranua datasets randomly in a 1:1 ratio. This ratio was selected to create a balanced dataset so that the model would avoid overfitting on a particular class. The image recognition models were created using Keras and KFold from Scikit-learn. 27 models were created and then compared against each other. They were created iteratively so that the evaluation metrics were looked at to see which models performed the best and how adding each layer influenced the metrics. All models except model 8 were Convolutional Neural Networks (CNNs). Model 8 was a simple Neural Network (NN) but was eventually left out of further testing as it did not perform well. Figure 4 shows model number 25.

Rescaling and data augmentation: The first layer of all 27 different models was a Rescaling layer, which rescaled all the images to 1. / 255. The last three layers were the Flatten layer, then two Dense layers, first with unit size 128 and activation relu, or rectified linear unit activation, and then one with unit size one and activation sigmoid. The sigmoid activation was chosen as the last one, as this was a binary classification task, and sigmoid activation provides a value between 0 and 1 for the likelihood of the image being of the chosen class. For the data augmentation, models from 9 to 27 used image augmentation layers after the Rescaling layer (see Figure 4). Image augmentation layers improve the training with small datasets [31]. The layers were RandomCrop, RandomFlip, RandomRotation, and RandomZoom, varying in different models. Model 9 used RandomCrop at sizes 200 and 200, RandomFlip at horizontal, RandomRotation with a factor of -0.5 and 0.5, and RandomZoom at 0.25. Model 10 used RandomRotation with a factor of -0.5 and 0.5 and RandomZoom at 0.25. Model 11 used RandomCrop at sizes 200 and 200 and RandomFlip horizontally. They both had five blocks of two Conv2D layers followed by MaxPooling2D. The rest of the models from 12 to 27 used RandomRotation with a factor of -0.5 and 0.5 and RandomZoom at 0.25.

Convolution and maxpooling: All of the convolutional layers used were square Conv2D layers, and they used padding with the same value. The activation used was relu in all of them. The number of filters and the kernel size varied between models and sometimes between different convolutional layers in the same model. This was done to discover the best combination of values for identifying reindeer from thermal images. The models mostly followed the pattern of having a Conv2D layer, or several,

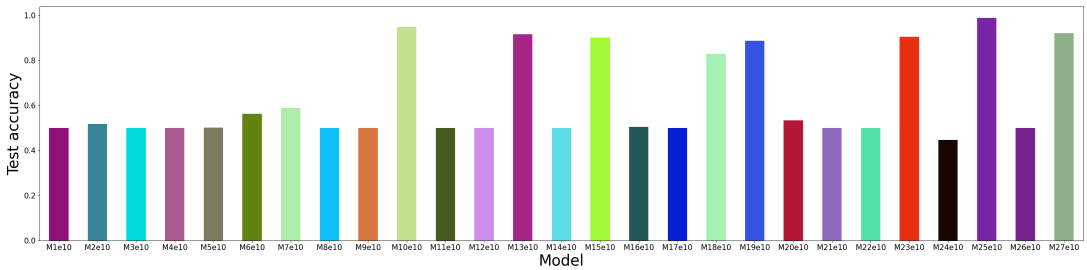


Figure 5: Test accuracy for all the models when trained with 10 epochs on Roboflow, tested on Ranua.

followed by a MaxPooling2D layer. Model 1 had a Conv2D layer with 16 filters and a kernel size of 3, followed by a MaxPooling2D layer. Model 2 added another convolutional layer with 32 filters, a kernel size of 3, and a MaxPooling2D layer. Model 3 added another convolutional layer with 64 filters and a kernel size of 3 and a MaxPooling2D layer, and model 4 further added a convolutional layer with 128 filters and a kernel size of 3 and a MaxPooling2D. The Conv2d layers in models 5-7, 9-24, and 26 had three filters and a kernel size of 3. Model 25 had the first layer with 11 filters, a kernel size of 11, five filters, and a kernel size of 5, followed by three with three filters and a kernel size of 3 convolutional layers. Models 5-25 had five blocks of Conv2D and MaxPooling2D. In each block, models 6 and 9-11 had two Conv2D layers followed by a MaxPooling2D. The rest of the models had one Conv2D layer followed by a MaxPooling2D layer. Model 26 had seven blocks of Conv2D and MaxPooling2D combinations, and model 27 had six blocks of that layer combination.

Normalisation : In models 7, 13-20, and 25-27, normalisation layers were explored. Models 7, 13, and 25-27 used a single BatchNormalization layer after the first Conv2D layer (see Figure 4). Model 14 had the BatchNormalization layer between Conv2D and MaxPooling2d for every five blocks. Models 15 and 16 similarly used LayerNormalization, so in model 15, there was only one normalisation layer in the first convolutional block. In model 16, it was in all five blocks. Models 17 and 18 repeated the same process with UnitNormalization, and models 19 and 20 similarly used GroupNormalization. Different Dropout layers were explored in models 21 to 24. Model 21 used Dropout at a rate of 0.25 after each of the five Conv2D layers, and model 22 was similar, except the dropout rate was 0.5. Models 23 and 24 used SpatialDropout2D similarly; this layer came after each of the five Conv2D layers, and the dropout rate was 0.25 in model 23 and 0.5 in model 24.

3.6. Model Evaluation Setup

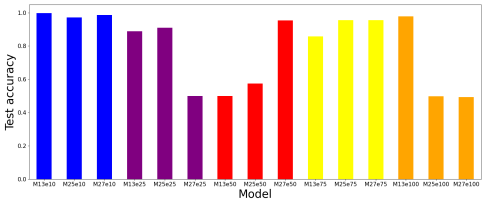
After the model was created, model.compile was called with optimized adam, the loss function was set to BinaryCrossEntropy, and the chosen metrics were also given as attributes. Then model.fit was called, and the history information was taken. Test accuracy for the whole model was recorded, and precision, recall, and F1-score were recorded separately for the deer and the human classes. 10-fold cross-validation was used with varying numbers of epochs. All 27 models were initially tested on 10 epochs while training with the Roboflow and testing with the Ranua data. Out of all the models, models 13, 25, and 27 were selected for further testing. They were tested with epochs 10, 25, 50, 75, and 100. After the tests were run, the model was trained on Roboflow and tested on Ranua data. Models 13, 25, and 27 were run with the same epoch numbers while training with Ranua and testing with Roboflow data, as well as with a mixed dataset. 25% of the data was reserved for testing.

4. Results

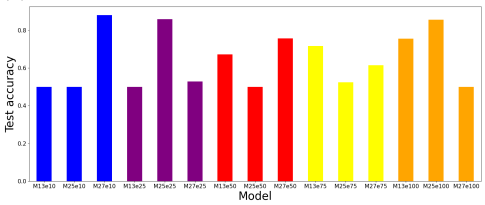
The test accuracy results for all models with 10 epochs are shown in Figure 5. The results vary widely between 0.447 and 0.998. When considering that 0.5 means that the accuracy is 50%, which is the accuracy of a guess, it can be seen that some of these models are as good as a guess. However, there are several models with very good accuracy results. Out of the 27 models, as mentioned in the previous

Model	Epochs	Acc.	Name	Prec.	Recall	F1-score
13	10	0.998	DEER	1	0.996	0.998
13	10		HUMAN	0.996	1	0.998
25	10	0.972	DEER	1	0.943	0.971
25	10		HUMAN	0.946	1	0.972
27	10	0.986	DEER	0.993	0.979	0.986
27	10		HUMAN	0.979	0.993	0.986
13	25	0.888	DEER	1	0.776	0.874
13	25		HUMAN	0.817	1	0.899
25	25	0.909	DEER	1	0.819	0.900
25	25		HUMAN	0.846	1	0.917
27	25	0.5	DEER	0	0	0
27	25		HUMAN	0.5	1	0.667
13	50	0.5	DEER	0	0	0
13	50		HUMAN	0.5	1	0.667
25	50	0.575	DEER	1	0.149	0.260
25	50		HUMAN	0.540	1	0.702
27	50	0.954	DEER	1	0.907	0.951
27	50		HUMAN	0.915	1	0.956
13	75	0.858	DEER	1	0.715	0.834
13	75		HUMAN	0.778	1	0.875
25	75	0.955	DEER	1	0.911	0.953
25	75		HUMAN	0.918	1	0.957
27	75	0.955	DEER	1	0.911	0.953
27	75		HUMAN	0.918	1	0.957
13	100	0.979	DEER	1	0.957	0.978
13	100		HUMAN	0.959	1	0.979
25	100	0.498	DEER	0	0	0
25	100		HUMAN	0.499	0.996	0.665
27	100	0.493	DEER	0	0	0
27	100		HUMAN	0.496	0.986	0.660

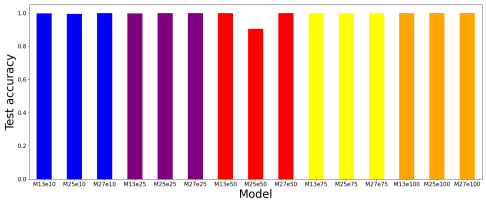
Table 2: Trained on Roboflow, tested on Ranua.



(a) Trained on Roboflow and tested on Ranua.



(b) Trained on Ranua and tested on Roboflow.



(c) Train/test with mixed data.

Figure 6: Testing accuracy results.

section, models 13, 25, and 27 were chosen for further tests. Some model-epoch combinations might have poor test accuracy, poor recall, and F1-score, but precision is 1, or poor test accuracy, precision, and F1-score, but recall is 1. Thus, it is necessary to look at all the metrics. Our best accuracy results are close to 0.998, which aligns with animal recognition algorithms presented in the state of the art, even if they use an ”easier” case of the camera instead of thermal images [3].

4.1. Training with Roboflow and Testing with Ranua Data

Table 2 and Figure 6a show the test accuracy for training with the Roboflow and testing with the Ranua data. The test accuracy varies between 0.493 (model 27 with 100 epochs) and 0.998 (model 13 with 10 epochs). However, 10 of the 15 different models and epoch combinations had an accuracy above 0.85. Table 2 also has the precision, recall, and F1-score. Most models with low test accuracy also had low precision, recall, and F1-scores. The models with better accuracy perform better and best with ten epochs. For example, model 13 at 10 epochs had the best test metrics out of model and epoch combinations with an accuracy of 0.998, with deer precision of 1, recall of 0.996, and F1-score of 0.998, and the values for the human class of 0.996, 1, and 0.998, respectively.

Increasing the epoch count seems to be somewhat inconsistent with the metrics. The test accuracy drops when the epochs increase from 10 to 25 for all three models. They further drop for models 13 and 25 when the epochs are increased from 25 to 50, although for model 27, the accuracy improves. Accuracy seems to improve or stay the same when the epochs are further increased to 75. Yet, they drop again for models 25 and 27 when the epoch count is increased to 100, but model 13 performed well. It had an accuracy of 0.979, which is better than model 25 at epoch 10. Our results compare well with previous work on animal recognition. Kalla et al. [32] presented Nasnet, a CNN-based Nasnet-Mobile model, for identifying horses on the roads. Their training accuracy was approaching 1 at 6 epochs, and for testing accuracy, it was approaching 1 at 45 epochs. Their precision was around 86.5%. When comparing these results to our results shown in Table 2, ours is higher for all the models with 10 epochs when the range is between 0.946 and 1, but there is more variation with higher epoch counts.

Model	Ep.	Acc.	Name	Prec.	Recall	F1
13	10	0.5	DEER	0	0	0
13	10		HUMAN	0.5	1	0.667
25	10	0.5	DEER	0	0	0
25	10		HUMAN	0.5	1	0.667
27	10	0.880	DEER	0.900	0.855	0.877
27	10		HUMAN	0.862	0.905	0.883
13	25	0.5	DEER	0	0	0
13	25		HUMAN	0.5	1	0.667
25	25	0.859	DEER	1	0.718	0.836
25	25		HUMAN	0.779	1	0.876
27	25	0.528	DEER	1	0.057	0.108
27	25		HUMAN	0.515	1	0.679
13	50	0.672	DEER	1	0.344	0.512
13	50		HUMAN	0.604	1	0.753
25	50	0.5	DEER	0	0	0
25	50		HUMAN	0.5	1	0.667
27	50	0.756	DEER	1	0.512	0.678
27	50		HUMAN	0.672	1	0.804
13	75	0.716	DEER	1	0.433	0.604
13	75		HUMAN	0.638	1	0.779
25	75	0.524	DEER	1	0.049	0.093
25	75		HUMAN	0.512	1	0.678
27	75	0.614	DEER	1	0.229	0.373
27	75		HUMAN	0.565	1	0.722
13	100	0.754	DEER	1	0.509	0.675
13	100		HUMAN	0.671	1	0.803
25	100	0.855	DEER	1	0.709	0.829
25	100		HUMAN	0.775	1	0.873
27	100	0.5	DEER	0	0	0
27	100		HUMAN	0.5	1	0.667

Table 3: Trained on Ranua, tested on Roboflow.

Model	Ep.	Acc.	Name	Prec.	Recall	F1
13	10	0.997	DEER	1	0.995	0.997
13	10		HUMAN	0.995	1	0.997
25	10	0.995	DEER	0.989	1	0.995
25	10		HUMAN	1	0.989	0.995
27	10	1	DEER	1	1	1
27	10		HUMAN	1	1	1
13	25	0.997	DEER	1	0.995	0.997
13	25		HUMAN	0.995	1	0.997
25	25	1	DEER	1	1	1
25	25		HUMAN	1	1	1
27	25	1	DEER	1	1	1
27	25		HUMAN	1	1	1
13	50	1	DEER	1	1	1
13	50		HUMAN	1	1	1
25	50	0.905	DEER	0.84	1	0.913
25	50		HUMAN	1	0.809	0.895
27	50	1	DEER	1	1	1
27	50		HUMAN	1	1	1
13	75	1	DEER	1	1	1
13	75		HUMAN	1	1	1
25	75	1	DEER	1	1	1
25	75		HUMAN	1	1	1
27	75	1	DEER	1	1	1
27	75		HUMAN	1	1	1
13	100	1	DEER	1	1	1
13	100		HUMAN	1	1	1
25	100	1	DEER	1	1	1
25	100		HUMAN	1	1	1
27	100	1	DEER	1	1	1
27	100		HUMAN	1	1	1

Table 4: Trained and tested on a mixed dataset.

Antônio et al. [33] presented K-Nearest Neighbors (KNN) and Random Forest (RF) algorithms for animal detection. They reported their results using F-measure. They had two results for KNN (0.611 and 0.624) and two for RF (0.560 and 0.589). 25 out of 30 of the F1-scores shown in Table 2 are higher than the results from Antônio et al. 20 of the results have an F1-score above 0.8. Nguyen et al. [34] presented a CNN for detecting and identifying wild animals from images. They used a simplified version of AlexNet called Lite AlexNet, VGG-16, and ResNet-50. They tried both imbalanced and balanced datasets. For the binary task of animal recognition, the accuracy results were for VGG-16 96,6% for imbalanced and 95.9% for balanced, for ResNet-50 they were 96.1% and 95.7% respectively, and for Lite AlexNet they were 94.9% and 92.7% respectively. For wildlife identification tasks, accuracy varied between 82.5% for Lite AlexNet for identifying the six most common species and 90.4% for ResNet-50 for identifying the three most common species with an imbalanced dataset [34]. Table 2 shows that all three of our best models with epoch 10 have better accuracy.

4.2. Training with Ranua and Testing with Roboflow Data

The accuracy metrics for the model epoch combinations when training with Ranua and testing with Roboflow data are shown in Figure 6b and Table 3. The test accuracy varied between 0.5, which several model epoch combinations had, and 0.88, with model 27, with 10 epochs. From Table 3 we can see that the best accuracy is for model 27 with 10 epochs at 0.88, and the following best are for model 25 with 25 epochs at 0.859 and model 25 with 100 epochs at 0.855. The same happens in test accuracy, and performance correlates with other metrics. However, several model epoch combinations might have perfect precision or recall for one class but a much lower value for the other.

Compared our results to previous works, Kalla et al. [32] have better accuracy than any model or epoch combination in Table 3. When comparing the precision of 86.5% from Kalla et al. [32], we have better precision in model 27 with 10 epochs for deer (0.9) and almost equal for humans (0.862). Despite the lower accuracy, the precision can compare well for model 27 at 10 epochs. When comparing our F1-scores to the F-measure reported by Antônio et al. [33] (ranging between 0.5601 and 0.6243), from

our results, 20 out of 30 F1-scores are higher. The F1 scores, however, are not as high for the model here as for the one trained with Roboflow and tested with Ranua. The accuracy values for Nguyen et al. [34] with a range of 92.7% to 96.6% for the binary identification were all better than in Table 3.

4.3. Training and Testing with a Mixed Dataset

The results for the mixed dataset are very good overall, as shown in Table 4 and Figure 6c. The test accuracy varies between 0.905 (model 25 with 50 epochs) and 1, with 11 of the 15 model epoch combinations having this. As the combined dataset is quite large, it is probably one cause for the better performance compared to the results from training with Ranua and testing with Roboflow. The results compare well with related literature. The test accuracy is 1 for 11 out of the 15 model and epoch combinations; at worst, it is 0.905. It compares well to the test and training accuracy 1 by Kalla et al. [32]. The precision results are superior compared to the 86.5% by Kalla et al. The F1-scores are all higher by a large margin than the F-measure results for Antônio et al. [33]. When compared to Nguyen et al. [34], it can be seen that all but one of our accuracy results in Table 4 are higher.

5. Discussion

As can be seen from the results, the mixed dataset performed the best out of the three, and the models that were trained on Roboflow and tested on Ranua material were the 2nd best. One possible reason for the lower results in the case of models trained on Ranua and tested on the Roboflow data could be the small size of the training material. On the other hand, size is not necessarily the only thing, as the mixed dataset performed better than the one trained with Roboflow, tested with Ranua, despite the latter being a larger dataset (2 times 842 images vs. 2 times 654). As the results varied between the models trained on different datasets, it is hard to recommend the best epoch counts. However, it seems that high epoch counts are not needed to attain well-performing models. The limitation of this work is that it only has two classes. This makes it easier to identify reindeer from images, as only reindeer and human images are included. However, our future work includes adding more classes to the model. At least elk or moose would be an essential addition, but other animal classes could include fox, rabbit, cat, and dog. If the detection of different Arctic wildlife, including birds, were enabled, then the data collection could serve a secondary purpose of collecting critical information about the occurrence of rare species. This could be used to help conservation efforts. However, including additional classes creates a need to gather a larger dataset, as thermal images of varying species are sparse in the open-source datasets.

Another future work is to gather more thermal camera images of reindeer and humans. These could also be taken in a driving situation, providing the most realistic images. Increasing the dataset size would improve the model, as machine learning applications usually perform better with larger datasets. Increasing the dataset with additional reindeer and human pictures, as well as pictures for additional classes, would then likely affect the model training time. However, model complexity and epoch counts could potentially be optimised for better time performance, as shown in our work. Our best evaluation metrics for the models were so good with epoch 10 and reasonable processing times. As such, our future work will focus more on the datasets and real-time animal detection situations. In the future, our goal is to also do realistic testing of the model in a vehicular setting. This would provide the most realistic testing results, but would require a test setup where the thermal camera can be safely positioned outside of the test vehicle. An additional challenge is where to find reindeer for testing purposes while driving, as they tend to appear only at times when you would wish they did not.

6. Conclusions

In this paper, we presented a CNN model that utilised thermal images to identify reindeer and differentiate them from other pedestrians, specifically people. The aim was to create a system that, in the future, could be integrated into a driver warning system or into an autonomous vehicle's control systems

to avoid reindeer-vehicle collisions. The model was trained on an open-source dataset and tested on the self-sourced one, and vice versa, as well as trained and tested on a mixed dataset. The evaluation shows good comparison against and, at times, exceeds the evaluation metrics of previous research. The best accuracy of 0.998 was a promising result for reindeer identification from thermal images. Most importantly, we can showcase that thermal imaging is equally accurate to standard camera-based methods when it comes to the identification of reindeer. However, typical cameras do not perform well on dark nights, and our solution with thermal imaging overcomes this restriction.

Acknowledgments

The work has been supported by the EU HORIZON projects CHIPS-JU CIA FEDERATE (grant 101139749) and CHIPS-JU RIA HAL4SDV (grant 101139789), Business Finland HAL4SDV national funding (grant 7655/31/2023), and the Finnish Research Council project Northern Utility Vehicle Laboratory Consortium GO!-RI (grant 352726). The authors are thankful to Ranua Resort.

Declaration on Generative AI

The authors have not employed any Generative AI tools.

References

[1] S. Kotituomi, M. Huju, R. Kulmala, A. Lahtela, Selvitys porokello-varoitusjärjestelmän vaikutuksista, Traficom (2019).

[2] E. Marti, M. A. De Miguel, F. Garcia, J. Perez, A review of sensor technologies for perception in automated driving, IEEE Intelligent Transportation Systems Magazine 11 (2019) 94–108.

[3] C. Carl, F. Schönfeld, I. Profft, A. Klamm, D. Landgraf, Automated detection of european wild mammal species in camera trap images with an existing and pre-trained computer vision model, European journal of wildlife research 66 (2020) 62.

[4] S. A. Useche, M. Faus, F. Alonso, "cyclist at 12 o'clock!": a systematic review of in-vehicle advanced driver assistance systems (adas) for preventing car-rider crashes, Frontiers in Public Health 12 (2024). doi:10.3389/fpubh.2024.1335209.

[5] E. Reimers, J. E. Colman, Reindeer and caribou (rangifer tarandus) response towards human activities, Rangifer 26 (2009) 55–71. doi:10.7557/2.26.2.188.

[6] H. Lyu, F. Qiu, L. An, D. Stow, R. Lewison, E. Bohnett, Deer survey from drone thermal imagery using enhanced faster r-cnn based on resnets and fpn, Ecological Informatics 79 (2024) 102383.

[7] D. Forslund, J. Bjärkefur, Night vision animal detection, in: IEEE Intelligent Vehicles Symposium Proceedings, IEEE, 2014, pp. 737–742. doi:10.1109/IVS.2014.6856446.

[8] M. Niemi, Vihersillat eläinten kulkureittinä tien yli: Eläinyhteyksien riistakameramonitorointi, 2021. URL: https://urn.fi/URN:ISBN:978-952-317-864-9.

[9] Tielaitos, AIDAT. Teiden suunnittelu V. Tiehen kuuluvat laitteet 4, TIEHALLINTO Tie- ja liikennetekniikka, Helsinki, 1998.

[10] S. Kotituomi, M. Huju, R. Kulmala, A. Lahtela, Selvitys porokello-varoitusjärjestelmän vaikutuksista, 2019. URL: https://traficom.fi/sites/default/files/media/publication/Porokello_Traficom_tutkimuksia_9_2019.pdf.

[11] V.-P. Hiltunen, M. Sivula, Porokello kilkattaa pian viimeisen kerran – poroista varoittava sovellus kaatuu rahoittajan puutteeseen juuri pahimman kolaririskin aikaan, 2022. URL: https://yle.fi/a/74-20006187.

[12] I. Hansen, S. M. Eilertsen, G. H. M. Jørgensen, J. Karlsson, Can a new electronic warning system prevent collisions between vehicles and semi-domestic reindeer?, in: Precision Livestock Farming 2019, ECPLF 2019, Cork, Ireland, August 26-29, 2019, pp. 268–271.

[13] I. Hansen, S. M. Eilertsen, G. H. M. Jørgensen, J. Karlsson, The “animal sense” warning system. low-cost technology to prevent collisions between semi-domestic reindeer and vehicles, 2021.
[14] Yle, Uusi mobiilisovellus varoittaa hirvistä reaaliaikaisesti – hälyttää myös käynnissä olevasta metsästyksestä, Yle Uutiset (2025). URL: https://yle.fi/a/74-20130962, accessed: 2025-04-04.
[15] D. Zhou, M. Dillon, E. Kwon, Tracking-based deer vehicle collision detection using thermal imaging, in: Int Conf on Robotics and Biomimetics, 2009, pp. 688–693. doi:10.1109/ROBIO.2009.5420589.
[16] H. Lyu, F. Qiu, L. An, D. Stow, R. Lewision, E. Bohnett, Advancing wild deer monitoring through uav thermal imaging and modified faster rcnn: A case study in nepal’s chitwan national park, Sensing and Imaging 25 (2024) 50.
[17] J. Friedman, T. Hastie, R. Tibshirani, Additive logistic regression: a statistical view of boosting, The annals of statistics 28 (2000) 337–407.
[18] R. E. Schapire, The boosting approach to machine learning: An overview, Nonlinear estimation and classification (2003) 149–171.
[19] P. Viola, M. Jones, Rapid object detection using a boosted cascade of simple features, in: Computer Society Confon Computer Vision and Pattern Recognition, volume 1, IEEE, 2001.
[20] L. Masello, G. Castignani, B. Sheehan, F. Murphy, K. McDonnell, On the road safety benefits of advanced driver assistance systems in different driving contexts, Transportation research interdisciplinary perspectives 15 (2022) 100670.
[21] B. Sznyi, szarvas dataset, https://universe.roboflow.com/balzs-sznyi-vqd7w/szarvas, 2023. URL: https://universe.roboflow.com/balzs-sznyi-vqd7w/szarvas, visited on 2024-08-28.
[22] Paliskunnat, Poromäärien kehitys, 2024. URL: https://paliskunnat.fi/py/materiaalit/tilastot/poromaarien-kehitys/.
[23] Paliskunnat, Poro (rangifer tarandus tarandus), 2024. URL: https://paliskunnat.fi/poro/poro/.
[24] K. A. Piirainen, R. A. Gonzalez, Seeking constructive synergy: Design science and the constructive research approach, in: Design Science at the Intersection of Physical and Virtual Design: 8th International Conference, DESRIST 2013, Helsinki, Finland, June 11-12, 2013. Proceedings 8, Springer Berlin Heidelberg, 2013, pp. 59–72.
[25] K. Lukka, Konstruktiivinen tutkimusote: luonne, prosessi ja arviointi, in: K. Rolin, M.-L. Kakkuri-Knuuttila, E. Henttonen (Eds.), Soveltava Yhteiskuntatiede ja Filosofia, Gaudeamus, Helsinki, Finland, 2006, pp. 111–133.
[26] Y. Choi, N. Kim, S. Hwang, K. Park, J. S. Yoon, K. An, I. S. Kweon, KAIST multi-spectral day/night data set for autonomous and assisted driving, IEEE Trans on Intelligent Transportation Systems 19 (2018) 934–948. doi:10.1109/TITS.2018.2791533.
[27] F. Ciaglia, F. S. Zuppichini, P. Guerrie, M. McQuade, J. Solawetz, Roboflow 100: A rich, multi-domain object detection benchmark, arXiv preprint arXiv:2211.13523 (2022).
[28] B. Sznyi, ”szarvas dataset”, 2023. URL: https://universe.roboflow.com/balzs-sznyi-vqd7w/szarvas.
[29] Karky, Human detection in ir images dataset, 2023. URL: https://universe.roboflow.com/karky-bk9o4/human-detection-in-ir-images.
[30] PNUSafetyNet, Thermal human detection dataset, 2022. URL: https://universe.roboflow.com/pnusafetynet/thermal-human-detection-00eyg.
[31] C. Shorten, T. M. Khoshgoftaar, A survey on image data augmentation for deep learning, Journal of big data 6 (2019) 1–48.
[32] H. Kalla, B. Ruthramurthy, S. Mishra, G. Dengia, R. Sarankumar, A practical animal detection and collision avoidance system using deep learning model, in: Int conf for Convergence in Technology (I2CT), IEEE, 2022, pp. 1–6.
[33] W. H. Antônio, M. Da Silva, R. S. Miani, J. R. Souza, A proposal of an animal detection system using machine learning, Applied Artificial Intelligence 33 (2019) 1093–1106.
[34] H. Nguyen, S. J. Maclagan, T. D. Nguyen, T. Nguyen, P. Flemons, K. Andrews, E. G. Ritchie, D. Phung, Animal recognition and identification with deep convolutional neural networks for automated wildlife monitoring, in: Int Conf on Data Science and Advanced Analytics (DSAA), IEEE, 2017, pp. 40–49.