SHORT TERM SCIENTIFIC MISSION (STSM) SCIENTIFIC REPORT

This report is submitted for approval by the STSM applicant to the STSM coordinator

Action number: CA15224
STSM title: Using computer vision to automatically detect hazardous flock behaviour
STSM start and end date: 1/11/2020 to 30/11/2020
Grantee name: Assistant Professor Dan B. Jensen

PURPOSE OF THE STSM:

(max.200 words)
Keel bone damage (KBD) is one of the main welfare issues in laying hens kept in group housing. Several causes of KBD, such as housing and perching design, have been extensively investigated. However, other causes such as hazardous behavior that may lead to KBD are still poorly understood. Smothering events are one hazardous behavior that occurs in flocks of laying hens which may lead to KBD. During a smothering event, chickens crowd together to such a degree that chickens may even be suffocated to death. Naturally, this behavior can also lead to more minor injuries which may include KBD.

In this STSM, we attempt to use computer vision to detect hazardous behaviors that lead to KBD. The use of technology to monitor flock behavior is a successful recent strategy that helps optimize management and prevent events caused by hazardous behavior such as smothering. Successfully detecting hazardous behavior using non-invasive camera monitoring will make it simpler to study causes of hazardous behavior and to intervene as this behavior occurs, preventing KBD from occurring.

DESCRIPTION OF WORK CARRIED OUT DURING THE STSMS

(max.500 words)
The visiting researcher tested the software for obtaining video footage and corresponding labels from the Vetsuisse Faculty Bern, Centre for Proper Housing: Poultry and Rabbits, Zollikofen (ZHTZ), which had previously been made available for use in this project. After gaining access, the researcher selected a subset of videos for use during her stay. These videos were of the same, white flock farm, and contained multiple smothering events for testing the classification. At the same time, the researcher was trained in the use of the R-package *mxnet* and adapting and applying the pre-trained neural network model *VGG16* on novel data.

After selecting the data, the visiting researcher investigated which method was most efficient in transforming the videos into frames which can be used as input for the pretrained neural network. She investigated several different methods for video frame extraction including VLC, FFmpeg and an R-based package using OpenCV. In the end, the best method appeared to be FFmpeg, as while more complex to use this method was able to accurately extract frames at a constant rate. Following extraction, the frames were adapted for use in a pretrained model by resizing the frames.

The visiting researcher created a pipeline detailing the steps from accessing the video data to obtaining the results of the classification. This included adapted code from the host institute as well as novel code written by the researcher, such as the labelling the data automatically from the labels in the provided excel sheet.

In the final week, the visiting researcher presented this pipeline to other researchers at the host institute and oversaw the installation of the relevant software on the computers of the researchers that will continue to be involved in the project. This also included generating some tentative classification results by applying the pipeline on two hours of video footage (see main results).

DESCRIPTION OF THE MAIN RESULTS OBTAINED

The visiting researcher supported local researchers in the use of the VPN client to access the data, as well as how to download, set up and use FFmpeg for use in video extraction. This included a demonstration of how to use the program, as well as tips for use and common pitfalls.
The researcher further created a pipeline which would turn the video data into useable, labeled frames and use this data to build a classifier for smothering behavior from an existing pre-trained neural network. This code was tested by applying the pipeline to a subset of 8 consecutive videos from the same camera.

Through FFmpeg, roughly 900 frames were extracted per video at a rate of 1 frame per second. Of the eight videos, four videos were used for training, while the testing and external validation set were both comprised of 2 videos. The data was labelled into three classes, “Smothering” if smothering occurred on that frame, “No Event” if no smothering occurred, and “30 seconds pre or post event” to flag frames directly preceding or following a smothering event. The pre/post event variable was included because identifying the exact moment a pile begins proved to be subjective. For example, there were often several seconds difference between labels made by researcher from the University of Bern and new labels made by the visiting researcher. Therefore, by labeling the 30 seconds directly preceding or after the event, these frames could be excluded from the training to avoid confusing the model during training. In the future, this class may also be used for prediction, as accurately predicting a smothering event 30 seconds ahead of the pile forming could be useful in preventing smothering behavior and thus keel bone damage. For now, the “pre and post event” frames were discarded.

Both the training and testing set were stratified by under-sampling the majority class (so that both classes had an equal number of examples for training the model), and the validation videos were left as is. For the distribution of classes, see Table 1. The overall accuracy of the model run was 0.908 though a note should be made that the validation video is consecutive to the training/testing example and the accuracy may therefore be overestimated. The distribution for the classes in the classification is shown in Table 2.

Table 1. The distribution of the classes in the data for the final results

	Smothering	**No Event**	**Total**
Training	1175	1175	2340
Testing	623	623	1246
Validation	955	787	1732

Table 2. Table of the early results obtained from an adapted VGG16 model.

		Predicted label	
		Smothering	**No Event**
Actual label	**Smothering**	941	14
	No Event	146	641

The adapted pipeline for use of pre-trained models for image classification to detect smothering was explained and handed over in the form of R-code to the researchers which will continue the project. By generating early results, the visiting researcher was able to give proof of concept for the approach that is intended to be used in the classification of potentially hazardous behavior. Furthermore, the initial pipeline uncovered several practical hurdles and potential solutions to be used in the follow up experiments. The visiting researcher herself gained practical experience in using pre-trained models for image classification.

FUTURE COLLABORATIONS (if applicable)

The goal of this visit was to lay the groundwork for a PhD project starting in the Spring of 2021. It is possible that further collaboration or assistance will occur once this project is under way, but no concrete plans have been made at this moment.