ARTICLE IN PRESS

Computers and Electronics in Agriculture xxx (xxxx) xxx

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Detection of anomalies in bee colony using transitioning state and contrastive autoencoders

Tymoteusz Cejrowski *, Julian Szymański

Faculty of Electronics, Telecommunications and Informatics, Department of Computer Architecture, Gdańsk University of Technology, Poland

ARTICLE INFO

Keywords:
Hive monitoring
Contrastive autoencoder
Anomaly detection
Bees buzz

ABSTRACT

Honeybees plays vital role for the environmental sustainability and overall agricultural economy. Assisting bee colonies within their proper functioning brings the attention of researchers around the world. Electronics systems and machine learning algorithms are being developed for classifying specific undesirable bee behaviors in order to alert about upcoming substantial losses. However, classifiers could be impaired when used for general honeybee colony state inference. Application of the classifier models for the hazardous situations detection without focusing on the model's genericity could result with systems that are not applicable in the real environment. Furthermore, the detection of a specific phenomenon does not provide researchers with any new conclusions about the honeybee colony life but only with the binary information about hazardous situation presence. In our research we propose a method for inferring the bee colony state using a sensitive contrastive autoencoder and an anomaly detection model. With presented approach, hive's internal state is modeled with the use of an autoencoder's latent vector extended with in-hive temperature dynamics. We test our methodology with a bee feeding experiment where the glucose syrup application was detected and the length of food intake was estimated. As our methodology has been applied successfully, we argue that contrastive autoencoders can be used for precise inference about the behavior of honeybees.

1. Introduction

The honeybees (Apis mellifera) is a species that has assisted humans for many centuries. Their work is considered highly beneficial, and many efforts are currently being made to protect them (Decourtye et al., 2019). However, honeybees' profitable actions are only possible when working in a larger collective, that is a bee colony (Tautz et al., 2008). In the context of current uncontrolled bees number decline (Colony Collapse Disorder) notably, after the winter seasons (Oberreiter and Brodschneider, 2020), a strong need for studying the bees as the whole collective emerges. Researchers should deeply understand reasons for honeybee losses, their acting, or preferable locations through modern tools and advancements in e.g. machine learning or Internet of Things (IoT) fields.

Today's technology provides a variety of sensors capable of monitoring bees in their natural habitat. The IoT systems are getting increasingly smaller and are currently capable of non-invasive measurements that may characterize the state of a honeybee colony. Joined with machine learning algorithms, it is possible to classify the colony in terms of the presence of various phenomena such as swarming (Žgank, 2018) or infestation (Bjerge et al., 2019). However, the classification task within the machine learning methods has some limitations. The classifier training process requires the usage of data originating from periods where an undesirable phenomenon has occurred. That case is relatively rare as in general bees work undisturbed and hazardous phenomenon such as sudden swarming is associated with the loss of most bees. Collecting data from multiple hives may also be ineffective as shown in one of our previous works (Cejrowski et al., 2018) where the bee colonies' characteristics vary so the classifier trained on data originating from multiple hives could not generalize well. Training hive-aware classifier requires data that's size is relatively small thus models may not be accurate for the future hazardous situation's anomaly prediction. Moreover, models are trained with the objective to label the data as one of the defined classes, i.e. swarming or healthy bees. Such an approach could not detect well other hazardous activities such as pest attacks or pesticide infestation.

In this paper, we propose a novel approach for inferring bee hazardous state with the usage of contrastive autoencoders for anomaly detection. We assume that each hive has its unique rhythm, which we attempt to identify using autoencoders and contrastive learning. The autoencoder neural network model consists of two sub-models: an encoder, and a decoder, which task is to reconstruct the input to the output. The model compresses the data into a latent vector, that

* Corresponding author.
E-mail addresses: tymoteusz.cejrowski@pg.edu.pl (T. Cejrowski), julian.szymanski@eti.pg.gda.pl (J. Szymański).

https://doi.org/10.1016/j.compag.2022.107207
Received 19 March 2022; Received in revised form 4 July 2022; Accepted 7 July 2022

Please cite this article as: Tymoteusz Cejrowski, Julian Szymański, *Computers and Electronics in Agriculture*, https://doi.org/10.1016/j.compag.2022.107207

contains all the information needed for proper input reconstruction. Thus, all the essential hive characteristic data for a modeled object are encapsulated into the feature vector. Training special case of autoencoder model, which is the contrastive autoencoder, involves subtracting common features found within the whole apiary hives which is detailed in Section 3.1. The contrastive model's latent vector based on audio samples serves as the features unique for the modeled hive. Together with colony temperature information, the latent's serves as complete colony description that can be trained for anomaly detection to infer hazardous bee colony states.

In the presented work, we focus on detecting the presence of the bee feeding as an anomalous situation. Bees within a specified date were supplied with glucose syrup which could potentially increase their activity. Such a situation should be detected with the use of an Isolation Forest anomaly detection algorithm operating on the autoencoder latent and in-hive temperature dynamics. Moreover, we attempt to estimate the length of the food intake process. We believe that the presented methodology, together with novel experiments setups, could provide new insights into the bees and their unique responses to different conditions and external stimuli.

The paper is organized as follows: Section 3 describes used models, especially contrastive autoencoder and anomaly detection models as they are the core of the presented methodology. With Section 3.3 we present features which were used to characterize colony internal state, evaluation metrics to assess methodology performance was introduced in Section 3.4. Results of anomaly feeding application experiment was presented in 5, at the end we discuss our findings and conclude presented work in Sections 6 and 7.

2. Related work

Remote monitoring systems have gained popularity in recent years. With the advancing sensors, miniaturization, and its affordability more Universities along with professional companies are developing systems for honeybee monitoring. Work in Zacepins et al. (2020) introduced a remote hive monitoring platform equipped with weight and temperature sensors. The system was based on ESP8266 low-cost WiFi chip where authors developed a custom power consumption monitoring tool to assess its sustainability. The hive monitoring energy efficiency issue was also addressed in Cejrowski et al. (2019) where authors presented software-based optimizations for bee colony monitoring firmware. Another custom beehive monitoring system from Ammar et al. (2019) is equipped with weight, humidity, temperature gas, and light sensors. A worthwhile enhancement is a camera that monitors the hive entrance. The authors proposed solar panels as a power supply source charging LiPo batteries. Work in Cecchi et al. (2020) implemented sensor fusion concept as a tool for honeybee colony monitoring. The authors used weight, sounds, temperature, humidity, and ambient weather to assess the status of the bee colony. The system is based on RaspberryPi 3B board, custom weight frame, and some ready-to-use modules for collecting in-hive and ambient measurements.

There are multiple industry systems for colony monitoring like Arnia,[1] OsBeehives with BuzzBox[2] or HiveMind.[3] All the devices collect measurements like temperature, humidity, sound or even number of bees entering or leaving the hive. They are shipped along with web application that allows one to inspect monitored colony and assess state via multiple pre-defined rules or even machine learning algorithms. Despite the existence of various bee monitoring solutions we decided to use custom and already proven in Cejrowski et al. (2020), Cejrowski and Szymański (2021) solution based on RaspberryPi. System overview is described in Section 3.2.

[1] https://www.arnia.co/

[2] https://www.osbeehives.com/

[3] https://hivemind.nz/

Bee monitoring systems are intended to serve as an alarm or early warning system. To meet these requirements, it is necessary to implement logic that works with collected data. Authors in Voudiotis et al. (2021) introduced a hive monitoring system equipped with a camera sensor monitoring beehive frames to prevent swarming phenomena. They utilize and test multiple CNN deep-learning networks which could be embedded both on cloud and edge devices. The proposed swarming detection system was validated for provoked swarming and overpopulated beehives where authors report success with preliminary results. Similarly, the task of swarming detection was addressed in Zacepins et al. (2016) where single point temperature measurements served as input for pre-defined rules on phenomena detection. Work in Ramsey et al. (2020) utilizes accelerometers placed in the center of honeybee hives to predict swarming. The authors used two machine learning algorithms and vibration spectra to assess the possibility of a hazardous situation. The results show that extraction of one-hour logarithmic averaged spectra and the cross-correlation between the spectra and three discriminant functions can predict upcoming swarming.

One could observe a growing interest in the sound analysis as a tool for honeybee colony state assessment. Our research based on sound analysis is presented in Cejrowski et al. (2020), Cejrowski and Szymański (2021) has shown that bees change their behavior according to varying external conditions and time that can be detected using the MFCC features with SVM classifier. Furthermore, work in Terenzi et al. (2021) compared different audio features along with convolutional autoencoder for detection of orphaned colony state. The mel-spectrogram and spectrogram features were described as the most promising for the queenless detection task. Authors in Terenzi et al. (2020) summarize multiple works on sound analysis and different microphone setups for honeybee monitoring. Various algorithms for bees sound feature extraction were reported along with machine learning models used within the bee research community. A similar survey in Hadjur et al. (2022) summarized and classified Internet of Things systems for honeybee monitoring. Authors covered the whole life cycle for precision beekeeping starting from reporting quantities which are most often used by researchers ending with data analysis tools and techniques. So far, we found that work as the most comprehensive survey of the bee monitoring systems and tools.

The inspiration for the presented work is the modeling done in Abid and Zou (2019) where authors proposed Contrastive Variational Autoencoder for enhancing salient features in the dataset. As an example, they identified a scenario with diverse dermatology images where skin color or sex should not affect dominant latent factors distribution. They proved that the Contrastive Variational Autoencoder could discriminate features essential for further classification or inference. Similarly, we believe that each hive generates unique sound characteristics which could not be easily captured by basic feature engineering methods but only with the use of enhanced contrastive models.

The use of basic autoencoder modeling could be observed widely in the Predictive Maintenance field where models are trained to comprehend machine characteristics. For example, authors in Henze et al. (2019) use autoencoder architecture as an anomaly scorer for monitored industrial equipment where audio data serve as the model's input. Furthermore, work from Renström et al. (2020) employed a deep autoencoder for anomaly detection in wind turbines. The model trained on correct operation data would only give a good reconstruction when test data has the same characteristics. All the data which deviates from the normal operation would yield high reconstruction error thus indicating an anomaly.

3. Methodology

The estimation of bee feeding excitation time interval implies the use of a sensitive model characterizing the state of the bee colony. The model should be accurate enough to capture the varying state of the swarm as the glucose syrup is applied. Moreover, the definition of measured quantity which serves as model input should exhibit information about swarm activity. We chose the bee feeding situation as a forced anomaly because of the ease of the implementation and the high likelihood for bees excitement observation.

3.1. Contrastive model

To generate descriptive hive features one should use an accurate model for bees' internal state inference. We use the neural network autoencoder model which provides us with a tool for colony characterization. Autoencoders are trained to minimize the reconstruction error between the input data dataset X and models' prediction $\bar{X}$. They consist of two mirrored sub-models called encoder and decoder. The encoder task is to perform dimensionality reduction on input data with the use of function q and its trainable parameters ϕ. The decoder objective is to reconstruct the encoder's output to match original input data with the use of function f_θ. The bottleneck latent-space z defined with Eq. (1), contains essential spatial information of the training data.

$$\mathbf{z}_i = q_\phi(\mathbf{x}_i), \tag{1}$$

We assume that compressing the hive data with the use of a neural network could exhibit its most distinctive features as it is the objective of training. Such derived features usually perform better for the inference or classification tasks than handcrafted ones (Ma et al., 2021). The core principle of shifting feature generation task to neural network forms the core of deep learning, where the model aims to discover the input combination that minimizes the cost function. We believe that the autoencoder's latent space reflects the state of the bee colony.

Contrastive learning was used to enhance the model's sensitivity to distinctive features originating from the modeled hive. The goal of contrastive modeling is to identify the differences between two data sets. In the presented case, we are interested in making the features of the monitored hive explicit, so that they differ from other hives within the apiary. Having that we believe to emphasize the hive's most distinctive features thus even small disturbances could be detected.

To reach such a goal the background dataset $\{b_j\}_{j=1}^N$ was introduced. It provides common data whose difference with the target dataset $\{s_i\}_{i=1}^N$ would be maximized. The target dataset stands for data collected from the modeled hive and the background dataset contains data from all the remaining hives within the apiary. Autoencoder model is trained with pairs of data (s_i, b_j) over mini-batches. The size of the background set should be equal to the target dataset. Custom loss function incorporating the background and target reconstruction losses with the difference between the salient and background latents was introduced.

Our previous work has shown that Contrastive Convolutional 2-D Variational Autoencoder is the most promising model for colony state inference. Its latent space encode Gaussian distribution which difference with background hives latent distribution is maximized by separate discriminator model and custom loss from Eq. (2b).

$$\begin{aligned}\mathcal{L}_{\mathrm{VAE}}(\mathbf{s}_i, \mathbf{b}_i) \geq \mathbb{E}_{q_\phi}\left[f_\theta(\mathbf{s}_i|\mathbf{z}_{s_i})\right] + \mathbb{E}_{q_\phi}\left[f_\theta(\mathbf{b}_i|\mathbf{z}_{b_i})\right] \\ -\ \mathrm{KL}(q_\phi(\mathbf{z}_{s_i}|\mathbf{s}_i)\ \|\ p(\mathbf{z}_{s_i})) \\ -\ \mathrm{KL}(q_\phi(\mathbf{z}_{b_i}|\mathbf{b}_i)\ \|\ p(\mathbf{z}_{b_i}))\end{aligned} \tag{2a}$$

$$\mathcal{L}_{\mathrm{CVAE}} = \mathcal{L}_{\mathrm{VAE}} + \mathrm{BCE}(\mathcal{D}_\gamma(\mathbf{z}_s, \mathbf{z}_b), \mathbf{y}_D) - \alpha TC \tag{2b}$$

The loss function for contrastive variational autoencoder incorporates $\mathcal{L}_{\mathrm{VAE}}$ components for input compression and reconstruction along with Kullback Leibler divergence for shaping the latent space to follow Gaussians. The $\mathcal{L}_{\mathrm{VAE}}$ component is the base for all the variational autoencoder models. The contrastive part within the $\mathcal{L}_{\mathrm{CVAE}}$ loss is represented by Binary Cross Entropy Loss for discriminator D_γ and *Total Correction* term. The latter is derived from Density Ratio Trick (Sugiyama et al., 2012) where the Kullback Leibler divergence between two unknown distributions of latents is estimated with optimal classifier D_γ. The BCE term is the classifier output fed with autoencoder's latent $\mathbf{z}_b$ for background data and $\mathbf{z}_s$ for target data. The $\mathbf{y}_D$ is true target class for both latents.

The classifier D_γ is trained to differentiate latent for background and target which are the means of encoded Gaussian distributions. In the presented work the D_γ model is implemented with Logistic Regression and trained along with the autoencoder model. Full workflow for training contrastive autoencoder was presented in Fig. 1.

In presented work the Contrastive Convolutional 2-D Variational Autoencoder was used with 3-layer setup: 512, 142, 72. Dropout probability p was set to 0.1 after each layer. The latent space was set to be 2-dimensional, *kernel* = 3, *padding* = 1, *max pool* = 2. The total number of parameters was estimated at $1,552,143$. Such size allows one to transfer trained model into more constrained devices which is also the scope of future work. The Adam optimizer (Kingma and Ba, 2014) with learning rate $\lambda = 0.0001$ was used for the autoencoder model. Similarly, the discriminator model was trained with Adam and learning rate of $\lambda_\gamma = 0.00001$.

3.2. Monitoring system

The autoencoder neural network model requests an explicit definition of the input data that should contain the internal hive's state information. To fulfill that requirement a dedicated Raspberry Pi 3B+ system equipped with a digital MEMS microphone and temperature/humidity sensor was used. The audio data was recorded with the use of SPH0645 module which is I2S MEMS microphone. The in-hive ambient data were acquired through SHT31 module. To collect reliable measurements, the microphone and temperature/humidity sensor were placed in a probe and then inserted into the spacial center of the hive. A custom bee-frame with a tunnel through the center of the frame was designed to ensure reliable recordings as shown in Fig. 2. Monitoring system was designed for langstroth hive but could be easily adapted to any type of beehive. A server application was developed for the collection and storage of the measured quantities with the use of TCP sockets. A web application provides access to the data through dedicated views. Recorded sound data are obtained every 15 min with 2 s duration and sampling frequency of 44100 Hz. Device was rebooted every 15 min in order to save power and server disc space.

3.3. Features

When modeling the state of the hive, the majority of the researchers suggest the bee colony sound as the main source of colony information (Terenzi et al., 2020; Hadjur et al., 2022). However, it is crucial to choose a bees' sound features as we found raw data to be too complex for proposed autoencoder models. We use mel-spectrograms as the most powerful audio description method for the bee colony characterization task. Such a feature is defined as an acoustic time–frequency representation of the sound. The process of obtaining mel-spectrogram features from raw audio is presented in Fig. 3. The mel-spectrogram feature was calculated from an audio signal cropped with 0–3000 Hz window since no significant components at higher frequencies were observed within the dataset. The spectrum was extracted with the use of the Fast Fourier Transform algorithm and *hann* window function with the length of 1024 samples resulting in 50% overlapping. The number of mel-bands was set to 64 which yields mel-spectrogram of shape 64 × 173. Features were generated with torch audio package (Yang et al., 2021).

A mel-spectrogram feature passed to variational autoencoder yields a latent vector consisting of Gaussian distribution mean and variance. In the presented work we use the mean of the encoded distribution as the input for the colony state inference. It was decided to use a two-dimensional mean vector for ease of visualization. However, it is worth noticing that increasing the latent vector dimensionality could enhance the performance of the demonstrated modeling.

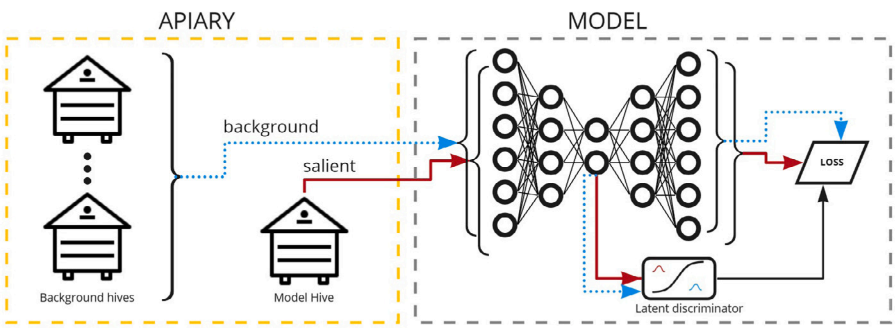

Fig. 1. Forward pass visualization with loss function for Contrastive Model.

Fig. 2. Custom frame for collecting sound recordings placed in the center of monitored hive.

The latent vector is assumed to carry compressed information about the state of the bee colony within the audio context. However, it is well-known that temperature and humidity are extremely vital for bees' health assessment. The use of autoencoder latent space provides a way for using other measurable values in the process of colony condition assessment. The two-dimensional latent vector could be easily expanded to include temperature, humidity, or other quantities like CO_2 level. For example, the extension of the audio latent vector with temperature yields a three-dimensional vector that could serve as a description of the bee colony internal state.

Usage of raw temperature or humidity values seemed like the most natural choice. Many authors have proven the usefulness of these quantities for the bee colony condition assessment (Zacepins et al., 2016; Kviesis et al., 2020). However, in our work, we test an alternative approach where instead of using raw temperature value we use its dynamics. We believe that the difference between the current temperature and the last recorded value should better assess the situation in the hive. Within the feeding process, the forager bees do not necessarily fly out searching for food but could operate directly inside the beehive. The temperature should be more stable within that time compared to normal operation.

In-hive humidity measurements were discarded due to their high correlation with the syrup intake. Supplied food is in the form of liquid thus contains a significant portion of water. The humidity level used as a colony state indicator would lead to wrong conclusions as its variation cannot be connected with the internal colony state. Final feature vector consisting of models' latent vector and temperature t_i difference was presented in Eq. (3). After concatenating features into $\mathbf{Z}_i$ vector it was standardized by removing the means and scaling features to unit variance.

$$\mathbf{Z}_i = \left[q_\phi(\mathbf{x}_i), t_i - t_{i-1}\right] \tag{3}$$

3.4. Metrics

It is believed that the process of supplying the hive with syrup appears as an anomaly in the bees' behavior. The colony, therefore, becomes excited and yields different characteristics than during the regular operation days. The well-trained model is expected to generate data considerably dissimilar for the feeding days than for the normal operation time. Specifically, it is preferable to observe two distinct areas of data corresponding to the feeding and normal operation time. Cluster analysis would be performed for the day of glucose syrup application and its adjacent days to observe the colony's transforming state. Detection of the start and end of the feeding process is based on Silhouette Coefficient plots analysis as further described. The anomaly detection algorithm was used to prove the model's applicability in the real environment.

The Silhouette Coefficient is used as an unsupervised learning performance indicator defined with Eq. (4).

$$\mathrm{SIH} = \sum_{C}\sum_{i=1}^{N}\frac{b_i - a_c}{max(a_c, b_i)}, \tag{4}$$

where C is the anomaly or hive data class, N number of samples, b_i is the cluster distance from a given sample to cluster that b_i is not part of it, a_c stands for the mean intra-cluster distance. A higher number indicates more distinct data clusters formed on data originating from feeding and normal days. Identifying the classification capability using the SIH coefficient serves as a tool for estimating the food intake period. The SIH coefficient value observed the day before fed application serves as the reference while for the successive days, values would change during the glucose syrup intake. The day where SIH coefficient reaches its reference level could be identified as the end of the feeding process.

An anomaly detection algorithm was used to test the applicability of the presented model where data were fed to the Isolation Forest (IF) algorithm (Liu et al., 2008) to identify the anomalies. By using the IF algorithm we want to test if it is possible to identify the feeding process with the use of autoencoder's contrastive learning and in-hive ambient temperature.

The IF algorithm employs decision trees called iTrees and random sampling to label anomaly data. For datapoint x there is an IF anomaly score s calculated with use of Eq. (5).

$$s(x, n) = 2^{-\frac{E(h(x))}{c(n)}}, \tag{5}$$

where n is the size of the sample set used when building iForest, x is a data point, $h(x)$ is the path length of observation x and $c(n)$ is the average path length of unsuccessful search in a Binary Search Tree. Based on the threshold value of 0.5 algorithm assigns a label indicating that the data point corresponds to an anomaly or normal measurement. For the presented case we treat feeding data as an anomaly and model's

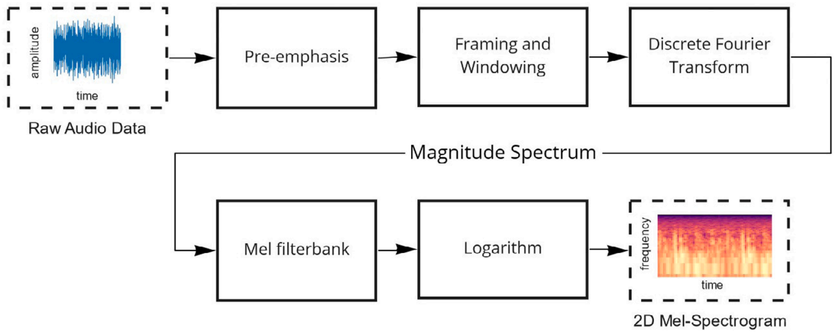

Fig. 3. Mel-spectrogram feature extraction workflow.

training data as normal operation. The desired scenario is the day of food intake begins a cycle of consecutive days where data are marked as an anomaly. We use F1 score to benchmark isolation forest performance.

4. Experiment

The experiment was conducted on an small apiary located in northern Poland, consisting of four langstroth hives. All the bees within apiary were the species of Apis mellifera. One of the colony, which beekeeper intended to feed with glucose syrup was marked as salient hive and the remaining 3 colonies were tagged as background ones. Sound and atmospheric data were collected with a 15-minute interval starting from 20 April 2021 to 01 September 2021 from all the hives. The bees were healthy and the colonies was found by the beekeeper to be strong, estimated as 40000 bees individuals in each of the hive.

The dataset was divided into two subsets: Normal Operation data (NOD) and Feeding Data (FED). The first subset was dedicated for the autoencoder training to generate the most accurate colony characteristics based on audio data. Normal operation's autoencoder latents and the in-hive temperatures from 20 April to 20 August were merged and divided into a training and validation dataset with 80%–20% ratio. Within the NOD dataset nor swarming or queen-removal was observed for all the hives. The bees were healthy and vital. It could be concluded that all the NOF data reflected natural rhythm of the families throughout the summer season. Furthermore, we assume that atmospheric conditions do not drive the hive anomalies. The NOD dataset for salient hive does not contain any previous feeding as the first glucose syrup was delivered on 25 August 2021. The Contrastive Convolutional 2-D Variational Autoencoder was trained on the NOD dataset for 100 epochs with early stopping set to 10 epochs. No overfitting nor gradient explosion was observed during the training phase.

The FED dataset consists of audio collected from August 21 to September 1. The glucose syrup was applied on August 25 around 4 p.m to the salient hive. Days between the August 21–24 should reflect the normal operation similarly to NOD data samples. The data originating from August 25 to September 1 are considered as anomaly out of which feeding final day should be derived.

5. Results

In order to cross-examine bees feeding behavior with normal operation data the autoencoder's latents were plotted in Fig. 4. Green datapoints corresponds to Contrastive Convolutional 2-D VAE latents derived from training dataset. Red points are latent generated from subsequent FED days.

One could notice two data clusters for the anomalous period data on August 25 which is the day of syrup application. Approximately half of the points were shifted downward on the latent 1 and latent 2 axes. The occurrence of a feeding event that directly affects the sounds produced by the bees could be observed. With successive samples, the points corresponding to the FED period move downward reaching a culmination on the next day (26 August). Approximately 5 days after the first food intake, the FED data cluster was observed to return to its position from the day before syrup application.

The SIH coefficient was used to define the syrup intake process duration as its value should return to the level of one day before syrup application. A plot of the SIH coefficient over the FED days is shown in Fig. 5.

The SIH coefficient reaches the maximum level on August 25, where the following day the value decreases by only 10%. It corresponds to syrup application day proving that the autoencoder model can detect feeding applications. The subsequent values head downwards indicating a decrease in excitement among the bees. The SIH value of the day before the day of feeding was reached on 30 August which represents a 5-day food intake period based on contrastive autoencoder latent features.

Nevertheless, the curve shape is not perfect, where the SIH value should asymptotically converge to the value before the day of syrup application. Furthermore, the maximum value of SIH coefficient is 0.2 which could be improved resulting with FED day cluster to be more distinctive. To address these issues, the in-hive temperature dynamics were employed. The autoencoder latent vectors were extended with the colony's internal temperature as described in 3.3.

The SIH curve for extended vectors within the FED days has been plotted on Fig. 6. An Isolation Forest (IF) model was trained to prove the applicability of the bee colony state inference using contrastive autoencoders. The anomaly detection model was trained on data originating from the NOD period. For each day within the FED period, an F1 metric was found to identify the model performance.

Using temperature increase as the information characterizing the bee colony state resulted in a smoothed SIH plot. The coefficients asymptotically converged to a value before the day of glucose syrup application. Moreover, the highest SIH coefficient of 0.32 was observed on August 26, which is more consistent with a fact that only half of the points from the 25th of August corresponded to anomalous events (syrup was applied at 16:00).

The anomaly detection model working with the extended latent vector (autoencoder latent and temperature) successfully detected the glucose syrup application on August 25. On that day there was an increase from 0.16 to 0.8 for the F1 metric while a peak with a value of 1.0 was observed on August 26. The F1 value closest to the day before the syrup application was observed on August 30 coinciding with SIH value.

The August 30 SIH coefficient of 0.32 was the largest observed value for the full dataset, which is the FED and NOD datasets. To verify the anomalous nature of the glucose syrup application phenomenon, the SIH coefficients were calculated for each day within the NOD dataset. Data from subsequent days were set against all the NOD's the remaining days. Model was challenged to estimate anomaly level for every presumed normal operation day. The average SIH coefficient for all the NOD days was −0.02, while the maximum of 0.21 was observed

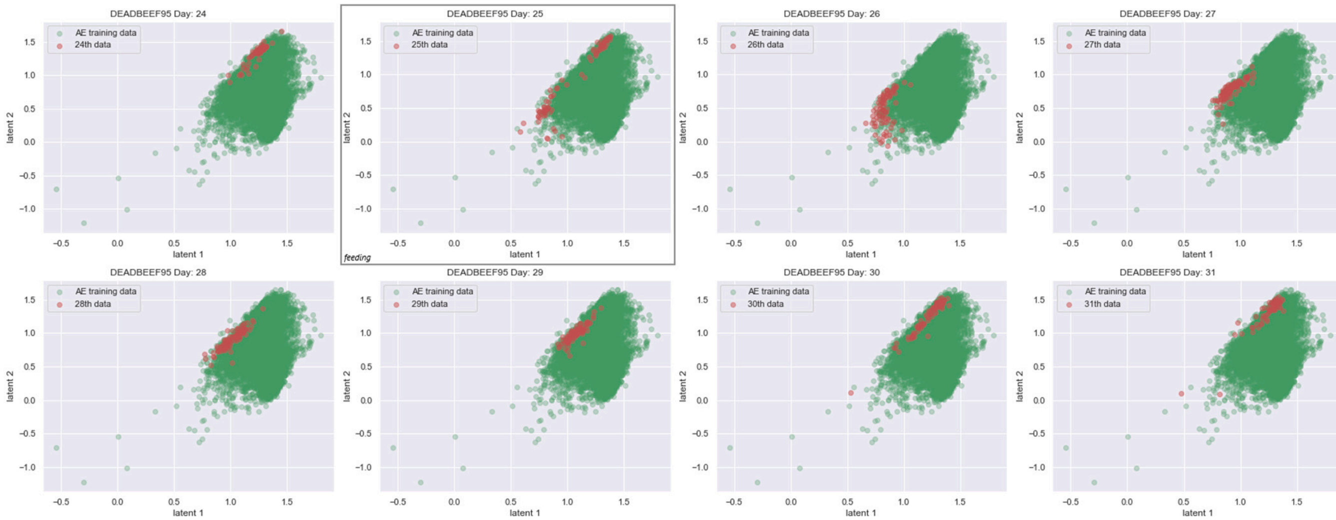

Fig. 4. Autoencoder's latent space (without temperature dynamics) for days between 24–31 August 2021 (red) and training data (green).

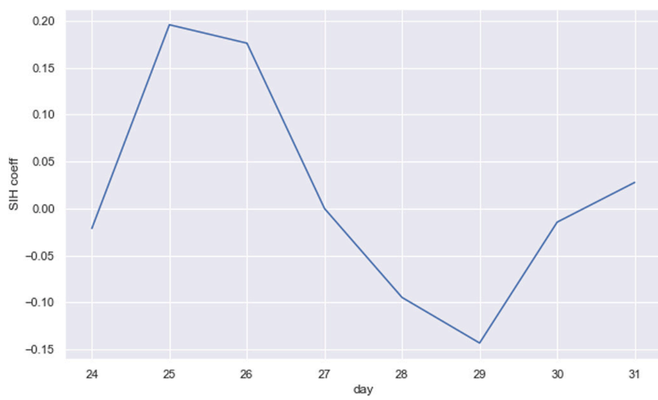

Fig. 5. The SIH coefficient transition through FED anomaly days.

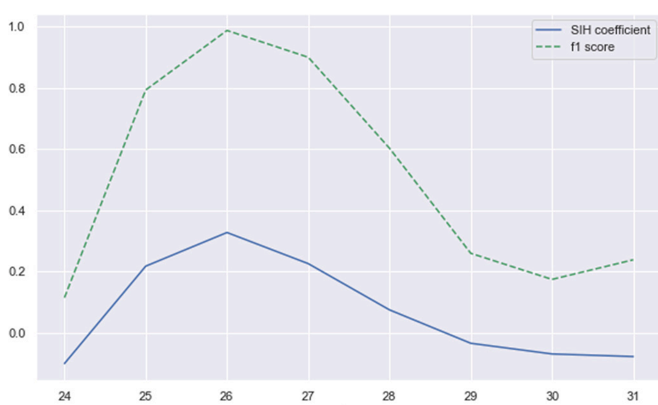

Fig. 6. The SIH coefficient and F1 score for the Isolation Forest Model.

for the June 6. The weather data did not exhibit any anomalies and the beekeeper did not remember any unusual situations for that day. Still, value for June 6 is 34% lower compared to one-day after glucose syrup application.

6. Discussion

Based on the SIH coefficient plot and the Isolation Forest model, it is possible to identify the food supply start time. The use of contrastive autoencoder latent features extended with temperature gives similar conclusions where the day of August 25 was identified as the time of feeding beginning due to the high SIH value increase. However, the plot derived solely from sounds and the autoencoder latents suggests that the days of August 25 and 26 are nearly identical in their buzz characteristic. Such a conclusion is wrong as only some portion of the August 25 samples should be identified as anomalies. Such a scenario might indicate an SIH metric usage impairment for the bee condition inference. However, using temperature increments as the extra input feature overcomes that issue by introducing an additional dimension to the input data. For that case, the highest SIH value is observed on the day following the fed application.

The end of the feeding process was defined as the time where the SIH value returns to the level the day before the feeding start. When considering only sound-related features, the threshold is reached between the 27th and 28th day of August for the first time. However, value is not realized asymptotically as one would expect. The second crossing point corresponds to August 30 which coincides with the plot generated from the concatenated latent vector and temperature. August 30 might be considered as the end of the food intake process for the case of the modeled bee colony.

The inconsistency when defining the end of the feeding process points to the need to use a heterogeneous vector describing the bee colony state. It is preferable to use sound signals as the primary source of information on colony status, but its extension with values such as temperature or humidity can help in proper reasoning. Our conclusions about thermal data usage coincides with similar work (Davidson et al., 2020) where the authors developed LSTM-based deep recurrent autoencoder for in-hive temperature anomaly detection. They state that an autoencoder trained on normal temperature data would fail to accurately reconstruct samples that have not been seen before, i.e., swarming data. Such an approach implies the use of the predefined *alpha* factor, which is the acceptable reconstruction error. If the Mean Square Error exceeds that threshold, the data are labeled as an anomaly. In the presented work, we do not rely on *alpha* factor. Furthermore, we believe that the temperature recordings are meant to support the sound-related features as there are some situations, such as a pest attack, where the in-hive temperature remains constant as the bees do not leave the hive.

The use of the anomaly detection model demonstrates the utility of using the autoencoder's latent vector as a feature for the colony characterization task. Together with its temperature extension, it is possible to detect anomalies such as bee colony feeding. It is worth noting the high model efficiency where the F1 metric reaches the value of 1.0 for the day after the fed start.

7. Conclusions and future work

In the presented work, we address the problem of anomaly detection and bee colony state inference using a novel autoencoder-based approach. A methodology for describing the bee colony state as the concatenation of the Contrastive Convolutional 2-D Variational Autoencoder latents, derived from mel-spectrograms audio data, and standalone temperature increments is proposed. Further processing is composed of the Isolation Forest anomaly detection where the model is trained on data originating from the data considered as normal hive operation. We test our methodology on the glucose syrup application experiment where the model proved its applicability resulting in F1 score of 1.0 on the one day after the syrup application.

The experiment made it possible to estimate the bee's food intake process time range. The F1 value and SIH coefficient plot analysis yield the 5 days period of the food intake. The SIH coefficient asymptotically returned to the level before the food application which is August 30. Such observation could help beekeeper to gauge the colony strength and identify the need of serving greater amount of food. However, the food intake interval estimation is still based on the single hive analysis. In order to test model generalization ability the presented methodology will be tested on more extensive hive setups. We aim to conclude with more general definitions of the bees food intake process length.

Future work involves the usage of the contrastive model to study bee responses with polluted environments as the model's applicability has been proven. Further experiments are planned where sensitive contrastive autoencoders will be used to describe the state of a bee swarm within different environments e.g. urban and rural and various anomalous stimuli. Moreover, we strongly believe that the bees' excitement for the food application could resemble the other bee colony states where e.g. bees accept a new queen. Such hypothesis will be tested with future work. Ongoing research is focused on transferring described models into more constrained embedded devices where inference could be performed in place of device installation.

CRediT authorship contribution statement

Tymoteusz Cejrowski: Methodology, Software, Investigation, Formal analysis, Data curation, Writing – original draft. **Julian Szymański:** Conceptualization, Supervision, Validation, Writing – original draft, Writing – review & editing, Visualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

The work has been supported partially by the founds of Department of Computer Architecture Faculty of Electronics, Telecommunications and Informatics, Gdańsk University of Technology, Poland.

References

Abid, A., Zou, J., 2019. Contrastive variational autoencoder enhances salient features. arXiv preprint arXiv:1902.04601.
Ammar, D., Savinien, J., Radisson, L., 2019. The makers' beehives: Smart beehives for monitoring honey-bees' activities. In: Proceedings of the 9th International Conference on the Internet of Things. pp. 1–4.
Bjerge, K., Frigaard, C.E., Mikkelsen, P.H.g., Nielsen, T.H., Misbih, M., Kryger, P., 2019. A computer vision system to monitor the infestation level of varroa destructor in a honeybee colony. Comput. Electron. Agric. 164, 104898.
Cecchi, S., Spinsante, S., Terenzi, A., Orcioni, S., 2020. A smart sensor-based measurement system for advanced bee hive monitoring. Sensors 20 (9), 2726.
Cejrowski, T., Szymański, J., 2021. Buzz-based honeybee colony fingerprint. Comput. Electron. Agric. 191, 106489.
Cejrowski, T., Szymański, J., Logofătu, D., 2020. Buzz-based recognition of the honeybee colony circadian rhythm. Comput. Electron. Agric. 175, 105586.
Cejrowski, T., Szymański, J., Mora, H., Gil, D., 2018. Detection of the bee queen presence using sound analysis. In: Asian Conference on Intelligent Information and Database Systems. Springer, pp. 297–306.
Cejrowski, T., Szymański, J., Sobecki, A., Gil, D., Mora, H., 2019. Low-power wsn system for honey bee monitoring. In: 2019 Twelfth International Conference on Mobile Computing and Ubiquitous Network. ICMU, IEEE, pp. 1–6.
Davidson, P., Steininger, M., Lautenschlager, F., Kobs, K., Krause, A., Hotho, A., 2020. Anomaly detection in beehives using deep recurrent autoencoders. arXiv preprint arXiv:2003.04576.
Decourtye, A., Alaux, C., Le Conte, Y., Henry, M., 2019. Toward the protection of bees and pollination under global change: present and future perspectives in a challenging applied science. Curr. Opin. Insect Sci. 35, 123–131.
Hadjur, H., Ammar, D., Lefèvre, L., 2022. Toward an intelligent and efficient beehive: A survey of precision beekeeping systems and services. Comput. Electron. Agric. 192, 106604.
Henze, D., Gorishti, K., Bruegge, B., Simen, J.-P., 2019. Audioforesight: A process model for audio predictive maintenance in industrial environments. In: 2019 18th IEEE International Conference on Machine Learning and Applications. ICMLA, IEEE, pp. 352–357.
Kingma, D.P., Ba, J., 2014. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.
Kviesis, A., Komasilovs, V., Komasilova, O., Zacepins, A., 2020. Application of fuzzy logic for honey bee colony state detection based on temperature data. Biosyst. Eng. 193, 90–100.
Liu, F.T., Ting, K.M., Zhou, Z.-H., 2008. Isolation forest. In: 2008 Eighth Ieee International Conference on Data Mining. IEEE, pp. 413–422.
Ma, J., Jiang, X., Fan, A., Jiang, J., Yan, J., 2021. Image matching from handcrafted to deep features: A survey. Int. J. Comput. Vis. 129 (1), 23–79.
Oberreiter, H., Brodschneider, R., 2020. Austrian COLOSS survey of honey bee colony winter losses 2018/19 and analysis of hive management practices. Diversity 12 (3), 99.
Ramsey, M.-T., Bencsik, M., Newton, M.I., Reyes, M., Pioz, M., Crauser, D., Delso, N.S., Le Conte, Y., 2020. The prediction of swarming in honeybee colonies using vibrational spectra. Sci. Rep. 10 (1), 1–17.
Renström, N., Bangalore, P., Highcock, E., 2020. System-wide anomaly detection in wind turbines using deep autoencoders. Renew. Energy 157, 647–659.
Sugiyama, M., Suzuki, T., Kanamori, T., 2012. Density Ratio Estimation in Machine Learning. Cambridge University Press.
Tautz, J., Heilmann, H.R., Sandeman, D., 2008. The Buzz About Bees: Biology of a Superorganism, Vol. 1007. Springer.
Terenzi, A., Cecchi, S., Spinsante, S., 2020. On the importance of the sound emitted by honey bee hives. Vet. Sci. 7 (4), 168.
Terenzi, A., Ortolani, N., Nolasco, I., Benetos, E., Cecchi, S., 2021. Comparison of feature extraction methods for sound-based classification of honey bee activity. IEEE/ACM Trans. Audio Speech Lang. Process. 30, 112–122.
Voudiotis, G., Kontogiannis, S., Pikridas, C., 2021. Proposed smart monitoring system for the detection of bee swarming. Inventions 6 (4), 87.
Yang, Y.-Y., Hira, M., Ni, Z., Chourdia, A., Astafurov, A., Chen, C., Yeh, C.-F., Puhrsch, C., Pollack, D., Genzel, D., Greenberg, D., Yang, E.Z., Lian, J., Mahadeokar, J., Hwang, J., Chen, J., Goldsborough, P., Roy, P., Narenthiran, S., Watanabe, S., Chintala, S., Quenneville-Bélair, V., Shi, Y., 2021. Torchaudio: Building blocks for audio and speech processing. arXiv preprint arXiv:2110.15018.
Zacepins, A., Kviesis, A., Komasilovs, V., Muhammad, F.R., 2020. Monitoring system for remote bee colony state detection. Balt. J. Modern Comput. 8 (3), 461–470.
Zacepins, A., Kviesis, A., Stalidzans, E., Liepniece, M., Meitalovs, J., 2016. Remote detection of the swarming of honey bee colonies by single-point temperature monitoring. Biosyst. Eng. 148, 76–80.
Žgank, A., 2018. Acoustic monitoring and classification of bee swarm activity using MFCC feature extraction and HMM acoustic modeling. In: 2018 ELEKTRO. IEEE, pp. 1–4.