A comparison of machine-learning assisted optical and thermal camera systems for beehive activity counting

Morton Williams, Samuel; Bariselli, Sara; Palego, Cristiano; Holland, Richard; Cross, Paul

Smart Agricultural Technology

DOI:
10.1016/j.atech.2022.100038

Published: 01/12/2022

Publisher's PDF, also known as Version of record

Smart Agricultural Technology 2 (2022) 100038

Contents lists available at ScienceDirect

Smart Agricultural Technology

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

A comparison of machine-learning assisted optical and thermal camera systems for beehive activity counting

S.M. Williams [a,*], S. Bariselli [b], C. Palego [a], R. Holland [b], P. Cross [b]

[a] College of Environmental Sciences and Engineering, Bangor University, Wales LL572DG, United Kingdom
[b] School of Natural Sciences, Bangor University, Wales LL572DG, United Kingdom

ARTICLE INFO

Keywords:
Thermal camera
Optical camera
Machine learning
Insect tracking
Apis mellifera

ABSTRACT

There is a documented shortage of reliable counting systems for the entrance of beehives. Movement at the entrance of a hive is a measure of hive health and abnormalities, in addition to an indicator of predators. To that end, two camera systems have been designed to provide a comparative analysis for a thermal camera system. The first, a visible spectrum camera, competed directly with the thermal camera. Machine learning is used to address the narrower field of view of the thermal camera, in addition to lost extracted tracks from both cameras. K-nearest-neighbour, support vector machine, random forest, and neural networks are used to classify flights as arriving, departing, or hovering bees. A hierarchical system is used to determine the nature of any flights where a clear label is not feasibly assigned based on the information from either test camera. A third camera at distance from the hive served as the end authority. After three iterations of training and validating, a test case is evaluated between both camera systems. Results from the test are compared to those from a human observer, showing that the thermal camera can perform with the same success as the visual camera despite a smaller field of view, fewer pixels and lower frame-rate, while both systems achieve greater than 96% accuracy and both camera systems are 93% successful at extracting flights. This is advantageous as a thermal camera will work in a wider range of environments, keeping the accuracy of an optical camera, and predicting based on movement characteristics will allow expanded uses such as predicting the presence of predators.

1. Introduction

Honeybee (*Apis Mellifera*) hive health is a targeted area of research given the recent decline in populations [1]. Insect pollinators are critical for food security and there is a need to carefully monitor population resilience to maintain them as a commercial pollinator [2]. Counting the incoming and outgoing traffic of beehives provides highly contextual information regarding a colony's health. Automated, remote technologies can provide an early warning detection capacity.

Proposed techniques to count bee activity at hive entrances work on optical sensors, radar, mechanical apparatus, and capacitive sensors [3–6]. Each of these technologies feature drawbacks that either affects the behaviour of the study organism or limit the range of environments they can be placed.

For example, optical cameras require good lighting and contrast to accurately count activity and are susceptible to shadows. During periods of insufficient natural light bulbs can be used but the heat and light generated by these may impact bee behaviour [7]. Similarly, infrared sensors require the modification of a hive entrance by adding an emitter [8].

Radar has proven efficient in studies using machine learning to classify insect activity [9]. However, radar operates in frequencies that are of concern for the welfare of the target species [10]. In addition, radar devices have a narrow detection radius, which limits how they might be deployed.

Thermal cameras have the potential to overcome all of these drawbacks. Capturing the body heat of the bees as a contrast image versus a background, they will operate in any environment where the temperature is lower or higher than that of the bee. As shown in Fig. 1, the bee's body is not uniform in temperature and they will appear contrasted against most backgrounds. The thorax of the bee is warmer than the abdomen and legs, therefore should the environmental temperature match the thorax the abdomen and legs will still appear against nearby objects.

Thermal cameras have been used to locate wild bumblebee nests [11]. In addition, they have been used to show that bees increase their body temperature before swarming, map where they sleep relative to hive temperature, and measure the heat dissipation in hives after artificial warming [12–14]. To date, no thermal camera has been used to count the incoming and outgoing activity at a hive entrance.

* Corresponding author.
E-mail address: eeu816@bangor.ac.uk (S.M. Williams).

https://doi.org/10.1016/j.atech.2022.100038
Received 19 December 2021; Received in revised form 24 January 2022; Accepted 24 January 2022

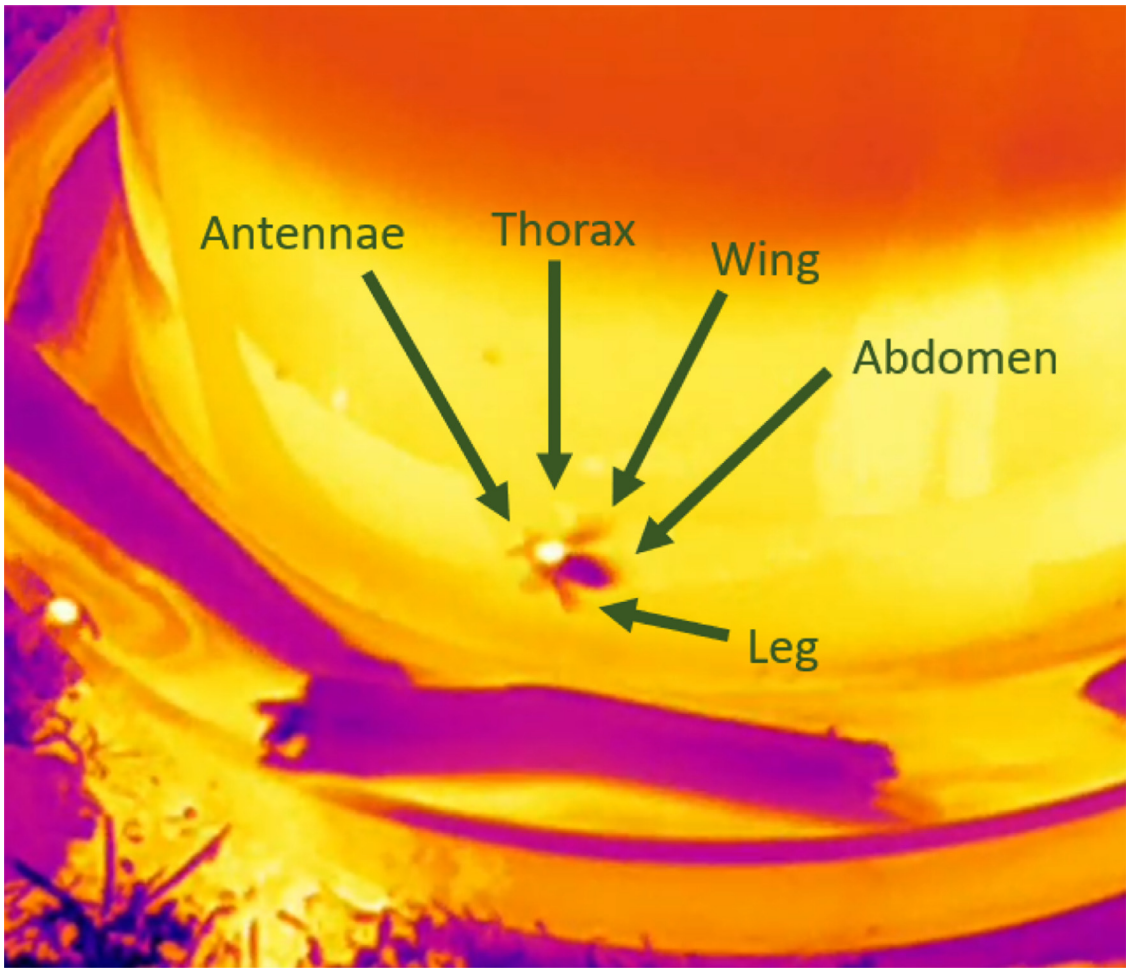

Fig. 1. Anatomy of a honeybee under thermal camera, flying in front of a warm sucrose solution feeder. The thorax was 21°C (±1°C).

Thermal cameras are expensive, have a narrow field of view, low resolution, and low frame rate. Recording of bees leaving a hive requires additional processing to overcome limitations. The cost increase of thermal cameras versus optical is reducing as technology improves and demand for more capable, cheaper devices increases [15]. These costs are expected to decrease further [16]. Fever detection is a recent, high demand area where thermal cameras are being used more frequently [17,18].

Recent developments in uncooled microbolometer technology are generating higher resolution, faster frame rate, and cheaper thermal cameras [19]. As such, this technology may provide a contender for accurate, reliable bee counting systems.

We implemented a hierarchical camera system where a higher resolution, wider field of view optical camera can provide information as to whether a bee is leaving, entering, or hovering near the hive entrance. A third camera positioned at some distance handled ambiguous cases. These are cases when the information from either test camera did not provide a clear flight category determinant. Similar to our previous work, machine learning is based upon the motion characteristics of the bee, which is used to classify each flight [20]. The use of a camera rather than radio-frequency technology, as in the previous study, allowed for shape metrics to be extracted, increasing the depth of machine learning used.

To summarise, this work is an investigation into whether machine learning would allow a thermal camera, with lower base specifications than a competing optical camera, to operate with the same efficiency counting beehive entrance activity. Showing that this is the case by using machine learning to boost the thermal camera, the work demonstrates the versatility of thermal cameras and is an argument for their use considering that they do not rely on lighting conditions, suffer from the effects of shadows, or need modification to hive entrances. The use of movement characteristics for machine learning creates an opportunity for future work to predict the presence of predators such as the Asian Hornet (*Vespa Velutina*) by using an expanded version of the system developed.

2. Materials and methods

The thermal camera used was the HT-301 designed by HTI-Instruments. It is a microbolometer camera with a resolution of 384 by 288 pixels, operating at 25 Hz with a field of view of 28.2° by 21.3°. The primary optical camera used was a GoPro Hero 7 Black model with a resolution of 1920 by 1440 pixels, operating at 30Hz with a field of view of 94.4° by 72.2°.

The 25 Hz framerate on the thermal camera represents how frequently the resistance is measured from the pixel sensors, which are affected by an increase in temperature caused by incoming infrared radiation at the target frequencies. Larger, more sudden shifts in resistance take longer to dissipate resulting in variable data change rates within the frame, causing fast-moving objects to appear as streaks.

Finally, a Panasonic compact system camera (DC-GX800) using a resolution of 1920 by 1080 pixels, 50 Hz framerate, and variable field of view was used as the reference camera. This camera was used when a flight could not be confidently labelled solely on the information of the two test cameras. The two primary cameras were suspended above the entrance of the hive facing downwards at a distance of one meter,

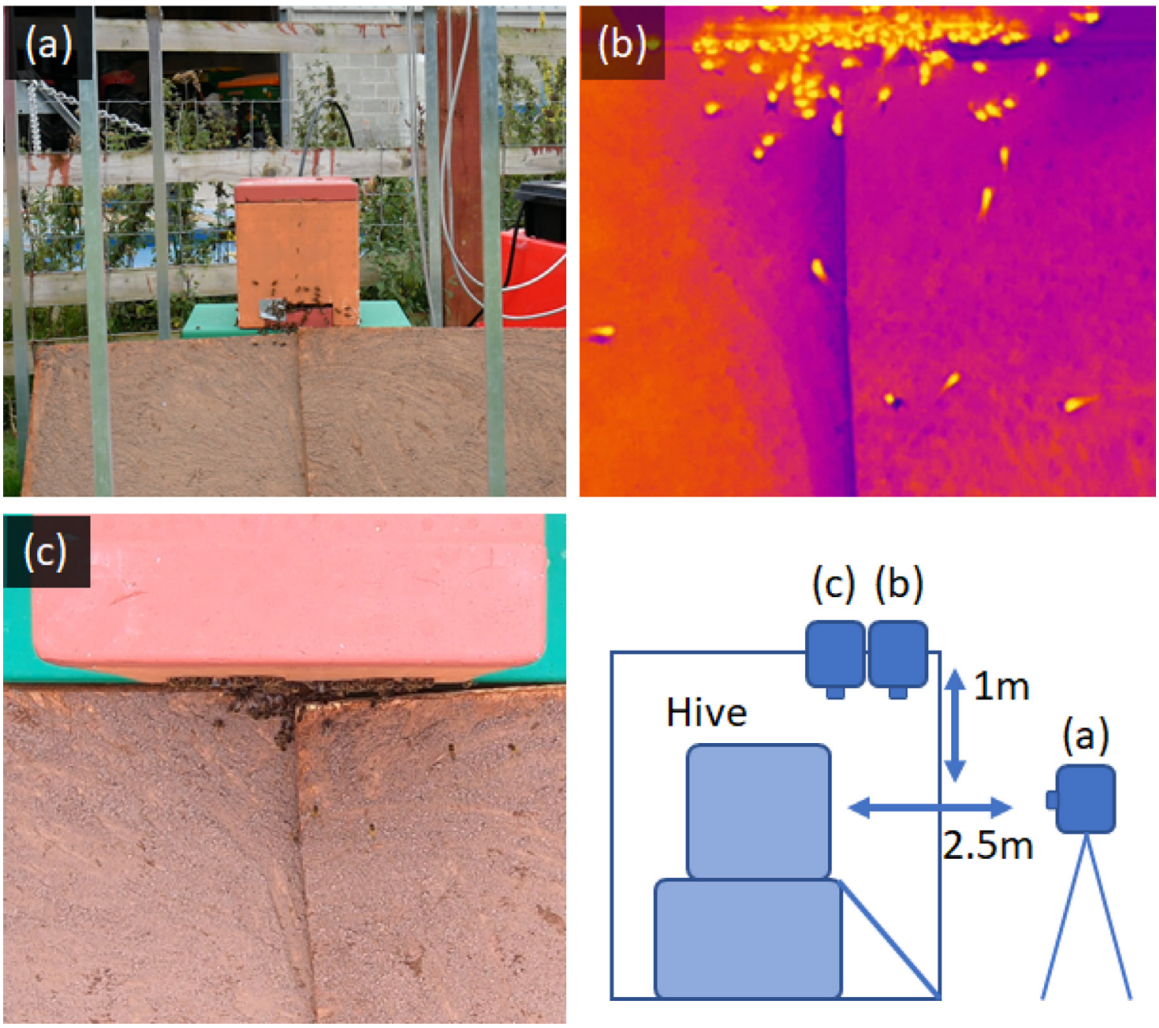

Fig. 2. Cropped views from (a) Ambiguous decision camera (b) Thermal camera (c) Optical camera, and a diagram of the camera arrangement.

whereas the ambiguous decision camera was fixed to a tripod approximately 2.5 m from the hive directly facing, and parallel, to the hive entrance. Cropped views from each camera are shown by Fig. 2, indicating their relative positions. Raw pixel values from the thermal camera are interpreted through a colour map embedded in the software provided by the manufacturer. The particular colour map used made no difference to the tracking software and was chosen to aid with the label correction process. The thermal camera had the option to embed minimum, maximum, and centre-point temperatures as labels into the video frames. The labels would act as moving objects within the frame and would interfere with object tracking tools. Therefore, this overlay was disabled resulting in only the contrast images being recovered.

Videos were recorded across six days, between 10 am and 4 pm, with between 30 min and an hour of footage per day. Ambient temperature varied between 12 °C and 16 °C. Bees were given warm sucrose feed at least two hours before recording to encourage activity. Artificially increasing bee activity allowed an improved understanding of the model performances under high load. The feeder was placed more than fifty meters away from the hive to reduce unnecessary hovering activity. Overcast days were favoured to produce a neutral environment for the optical camera. In bright conditions, this camera would be susceptible to bee shadows creating false detections that would impact the optical track extraction and labelling algorithms, but also required sufficient light to detect the bees. Techniques exist in literature to compensate for other conditions but thermal cameras do not need these, necessitating keeping a fair environment for both cameras.

Flights were extracted from the raw video files using software written in MATLAB [21]. Gaussian mixture models generated a foreground detector by comparing each frame in the raw videos with the learned background model [22]. The parameters for this foreground detector were selected by hand for each video from each camera. The detected shapes were analysed based on their size to determine eligibility for being considered bees, aiming to remove small shapes resulting from minor movements such as blades of grass.

Each eligible shape was assigned a Kalman filter unless otherwise attributed to an existing flight [23]. Between frames, assigning objects to existing flights was handled by computing the costs as the pixel-wise distance between the predicted locations generated by the Kalman filter from the previous frame's objects and the new detections from the current frame. The Munkres variant of the Hungarian algorithm was used to minimise the total cost of assigning each object to an existing flight [24]. A base cost was used, and tuned much like the foreground detector, to determine non-assignment. Unassigned objects could go on to become new flights and flights that had no new object assignments in three frames were flagged as completed and no longer considered.

At the end of processing, the video resulted in several flights consisting of continued x and y pixel coordinates, the longest and shortest axis of the extracted shapes per frame, the frame time of each detection, and a unique ID. Predicted Kalman coordinates were not included.

Initially, a simple naive filter was attempted. Similar in principle to that found in [25] it used a set of exit points around the edge of the camera frame. Flights that started nearest the hive point and ended nearest

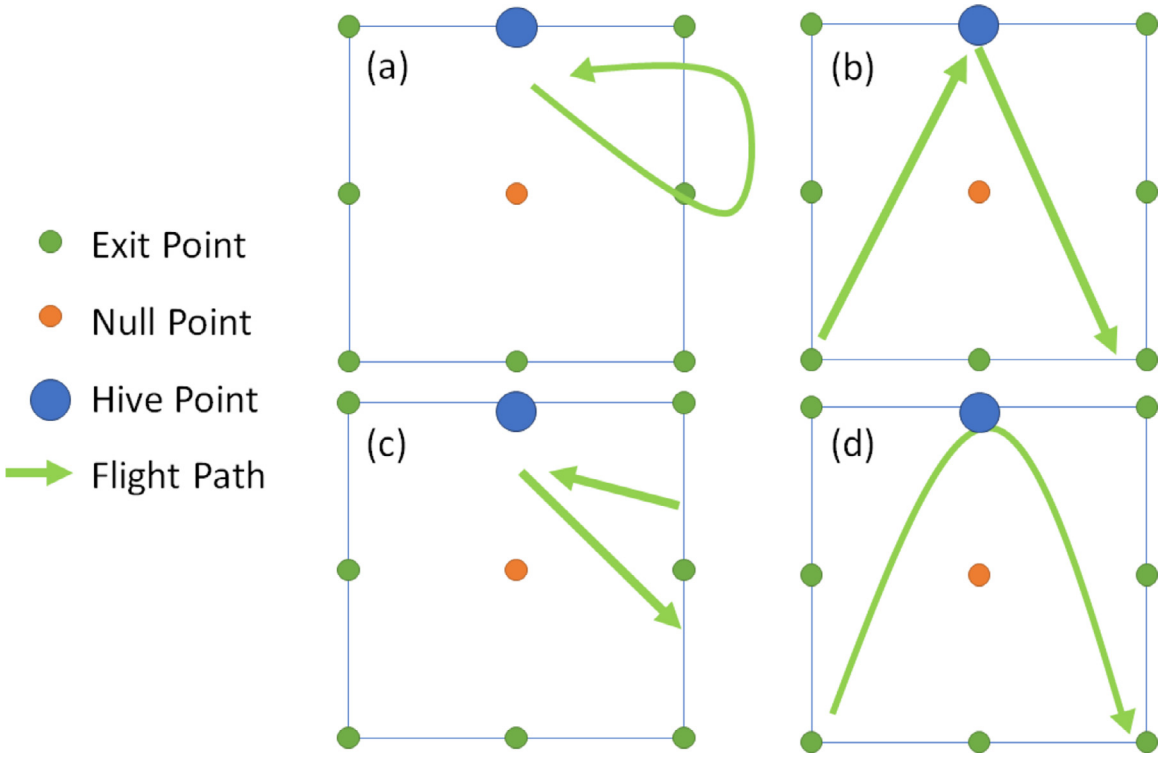

Fig. 3. Naive filter and visualised problematic flights. Squares represent camera frames and their borders. Flights (a) and (c) show flights that exited the frame border limit, and outgoing and incoming flights may not be the same individual. Trajectories (b) and (d) show how two flights may be fused if the end of the incoming flight occurs within the same frame as the outgoing second flight.

the exit point could be labelled as ”out” whereas those that started near an exit and entered could be labelled “in.” Finally, null points were used near the centre of the frame to allow for labelling of invalid hover flights, those that neither started nor ended near the entrance to the hive.

This filtering system is visualised in Fig. 3, however, emergent issues were impacting the quality of extracted flight information when considering the field of view of the cameras.

For example, flight (a) in Fig. 3 starts and ends near the entrance of the hive. It should be discarded as an invalid flight as the bee did not leave the vicinity. However, it crosses the edge of the frame and appears as in (c). This means one exit and one entrance flight would be recorded. Bees were noted to fly out as far as 1.5 m only to return to the hive, putting them out of view of both primary cameras.

The filter produced issues not unique to the field of view. Flights in (b) are one valid ‘in’ flight and one valid ‘out’ flight. However, if the two flights overlapped it would be possible for the extraction algorithm to fuse the flights, resulting in the flight present in (d) meaning one hover flight was recorded rather than one incoming and one outgoing flight.

One way to discern the difference between these cases was by analysing the dimensions of the extracted shapes. Fig. 4 shows four flights captured by the camera systems. In (a) and (c) the bee is flying at speed in a straight line, whereas in (b) and (d) it is flying to the side at a slower speed. At the edge of the frame, this may be enough to ascertain the difference between a bee leaving the area and one soon to return as part of an invalid flight. However, this is much less clear on the optical camera given the lower contrast and higher frame rate.

Also of concern was flights involving a loop where the bee turned around shortly after leaving the hive to face the entrance before continuing its flight as shown in Fig. 5. As such, classifying these correctly would mean that some bees moving to the side were classifiable as outward flights. This would be true where the flight was truncated by the edge of the frame and a full loop was not captured.

To address these issues machine learning techniques were explored. These aimed to correctly label flights when the edge of the frame caused truncated flights and where other issues with the flight extraction software emerged. Cases such as overlapping shapes breaking the tracking of a flight between frames and fused flights were both examples of issues with the tracking system.

Incoming flights were not subject to machine learning. This was because the naive filter was adequate for counting inward flights and valid incoming flights would have the same profile by being at the edge of the frame and returning to the hive entrance.

The machine learning was based on 43 features, typically focused on the minimum, maximum, average, and last three measures of the following variables;

- Shape longest and shortest axis
- Shape ratio and growth
- Track life in seconds
- Distances of the first detection to the hive and the nearest exit
- Distances of the last detection to the hive and the nearest exit
- Closest distance to the hive and the nearest exit
- Current angle versus average bearing angle
- Current angle versus perpendicular to the hive entrance
- Change in angle over time
- Number of loops and loop lifespans
- Duration and size of dropouts (frames where tracking failed)
- Acceleration and speed

Most of the more valuable information of a flight was encoded during the last three measurements taken of the location of the bees, thus the inclusion of minimum, maximum, and average of these ‘end’ variables in addition to the general copies.

Learning and testing took place in three stages. First, the naive filter results were hand-corrected by a human observer. Second, machine learning algorithms were trained on the labelled dataset. Finally, the algorithms were tested on a new video. This process was repeated multi-

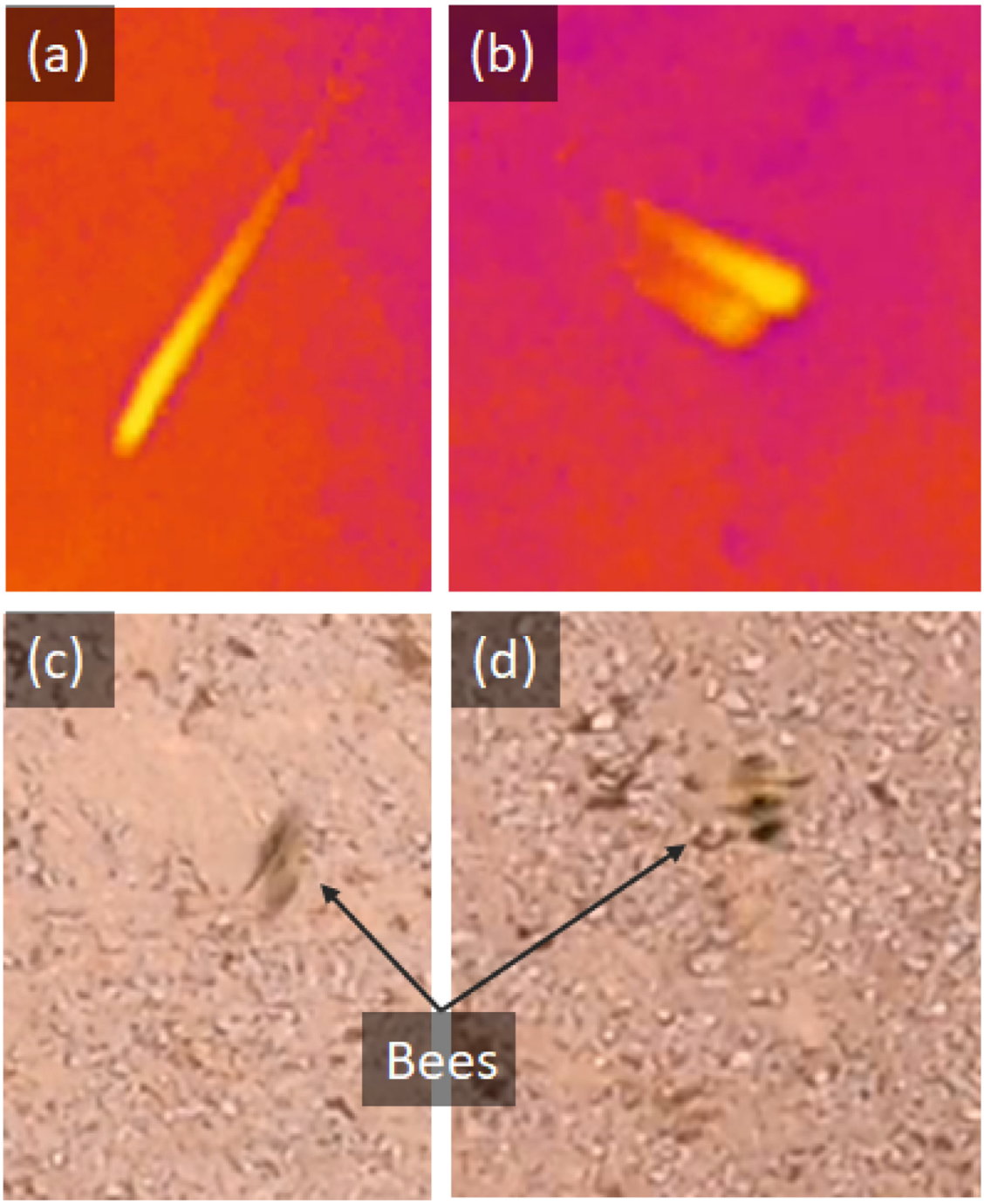

Fig. 4. Flights showing different exit strategies captured by the cameras. Images (a) and (c) show a bee flying straight ahead, whereas (b) and (d) show a bee flying more horizontally. Note the smudged nature of these high speed flights on the thermal camera, caused by the resistance change delay associated with microbolometer technology.

ple times, with the learning algorithms taking place of the naive filter in subsequent iterations. This continuous process allowed for the monitoring of test accuracy and identification of weaknesses in both the features used and the underlying flight extraction algorithm.

Once the test accuracy for the algorithms began to plateau after three training and testing phases, a final test was arranged with a new recording and the procedure assessed for veracity against a human counter.

The algorithms used were K-nearest neighbour (KNN), neural network (NN), support vector machine (SVM), and random forest classifiers (RF) [26–28]. A Bayesian search was enacted to tune the hyperparameters for the KNN, SVM, and RF approaches choosing from the possibilities as follows [29];

- For KNN: Number of neighbours and nearest neighbour algorithm.
- For SVM: Regularisation parameter, kernel, kernel degree, and kernel coefficient.
- For RF: Number of estimators, split criterion, and the number of features to consider for a split.

For the neural network, the model used composed of three hidden layers of 64 neurons each using the rectified linear unit activation (ReLU) function [30]. The shape of this network was chosen by manual testing to find the best hyperparameters, investigating the width and depth of the network compared to the final accuracy. Minor changes to this structure did not yield significant change across either dataset.

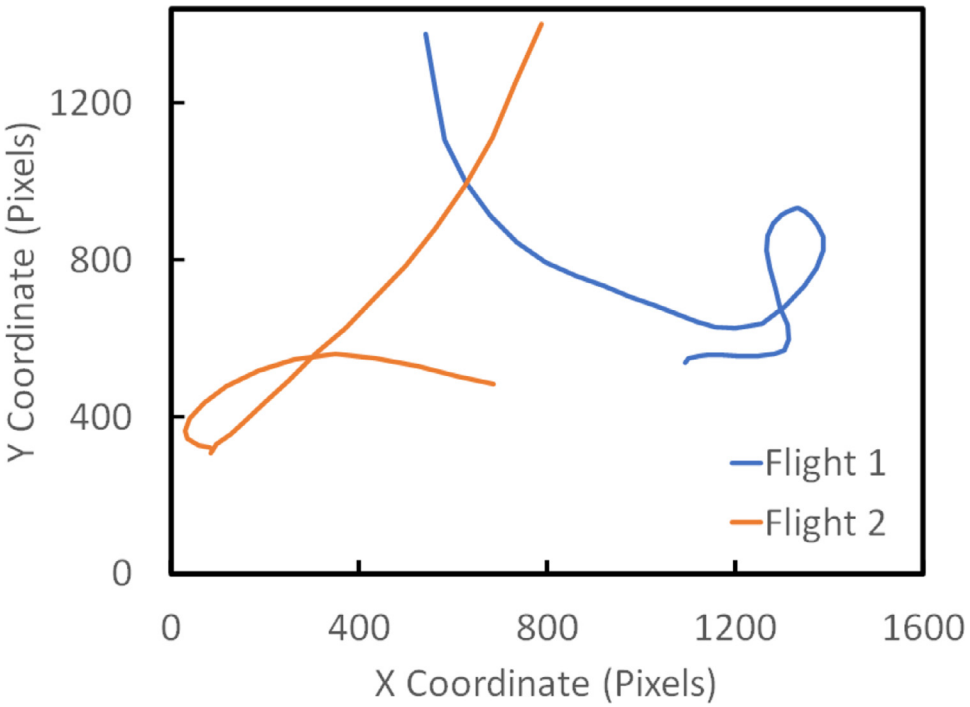

Fig. 5. Two flights where the bee turned back to face the hive before flying away, behaviour associated with new foragers orientating themselves with respect to the hive. Flight 1 went on to complete a second loop but this was not captured fully by the tracking algorithm.

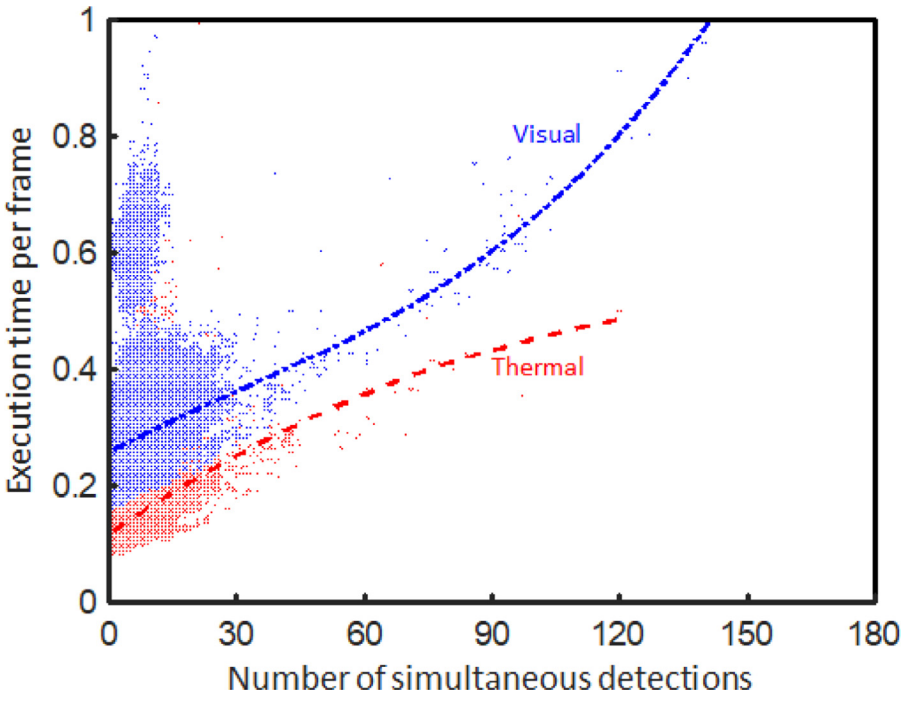

Fig. 6. Time cost per frame to extract flights from both thermal and visual camera systems, with a simple 3rd degree polynomial fit.

3. Results and discussion

Extraction times for both camera systems is shown in Fig. 6, showing that the thermal camera system demonstrated faster extraction and did not grow exponentially, explained by fewer pixels and clearer shape extraction. The exponential growth in the visual camera system is attributed to visual clutter such as debris and raw computational complexity due to resolution and field of view. The greater field of view included more objects that require foreground extraction, filtering, and labelling. This includes undesirable extractions such as blades of grass increasing the number of artefacts requiring attention.

The labelling of individual flights encountered issues when tracks split. Fig. 7(a) shows two flights where labelling could become ambiguous. In the Figure, a loop is formed as the bee flies away from the hive but it is broken by the edge of the frame. Deciding which section to label as an outward flight and which to disregard is challenging when the second section was missing for other flights. Labelling the first section as an outward flight would add a flight into the learning pool that, by itself, is not statistically clear as an outward flight because it is curving back

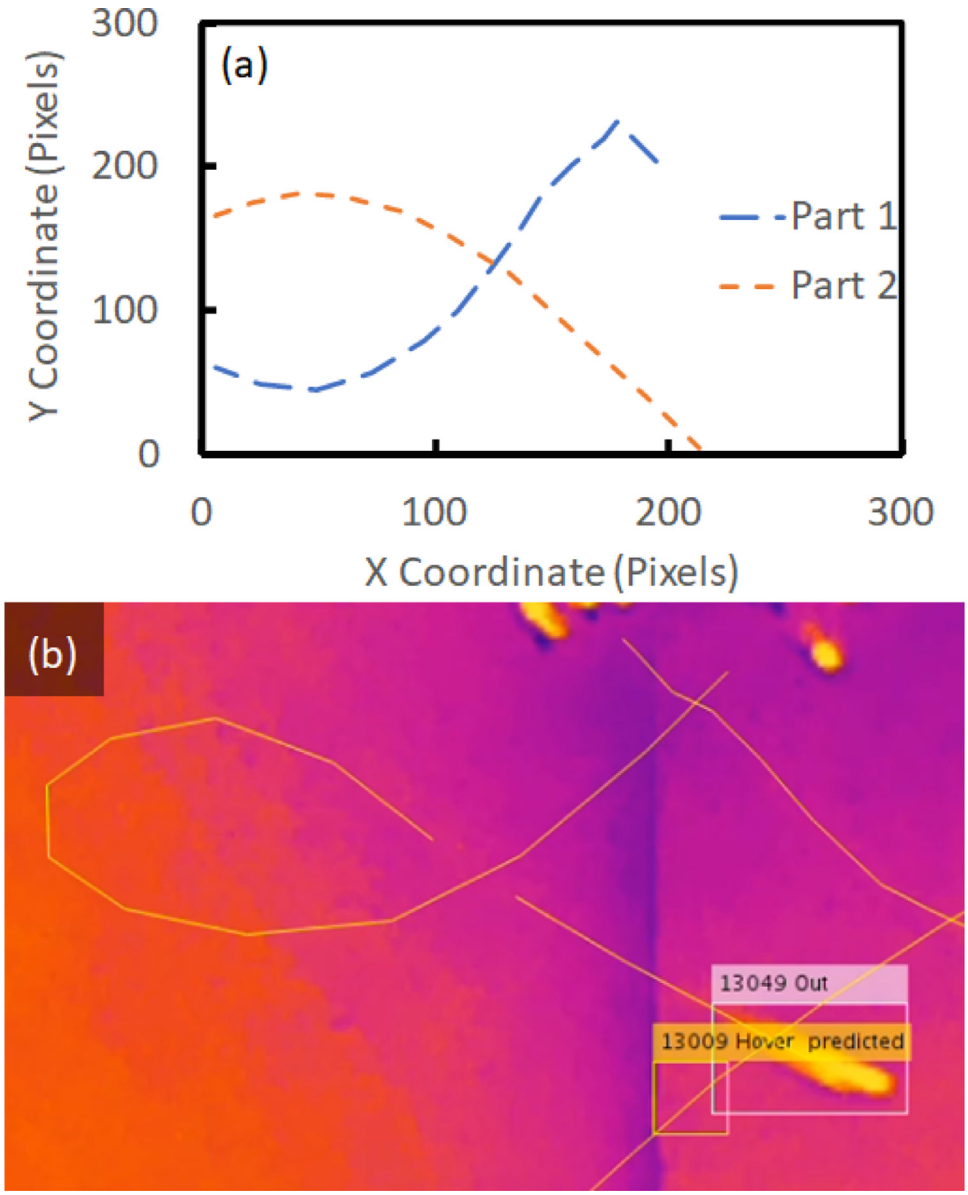

Fig. 7. (a) Trajectory of an unclassifiable flight caused by the edge of the frame and (b) a screenshot illustrating an interrupted flight.

towards the hive. This would increase the residual error in the models generated.

Similarly, in the screenshot flight Fig. 7(b), the track has been interrupted when two bees passed close together. The yellow bounding box, showing the predicted location of the first section as obtained by the Kalman filter, has been correctly labelled as invalid. The white bounding box has been correctly labelled as an out flight. In other cases, the second section of the flight may not be recovered and the remaining first section is unclassified. It could either be removed from the data set, meaning that a flight is labelled as missing, or labelled as an out flight and add error to the generated models.

The first stage of the training process involved using the naive filter to count outgoing flights. In the video recorded as a sample, using the naive filter resulted in 3220 outgoing flights counted. Of these, 1687 visual out flights were extracted correctly with 1533 flights later rejected by human expertise. This means that the naive filter achieved an accuracy of 52.39% and almost doubled the true number of out flights counted.

For the thermal video, 3735 outgoing flights were counted by the naive filter. This was then corrected to 1685 outgoing flights and 2050 incorrect flights, resulting in an accuracy of 45.11%.

As these two videos were aligned by hand, a comparison of flight links between systems was possible. The mode difference in time between thermal and visual extracted flights was 0 s, with an average difference of 0.0468 s. 75.1% of flights occurred across systems within 0.5 s of each other. The larger differences can be explained by the fused, split, and frame-edge flights as discussed above. Fused flights may cause significant time variance across systems as an outgoing flight may be fused with a longer hover flight due to extraction errors, and this hover flight may have been present for many frames before the outgoing flight. This increases the discrepancy between the two systems.

The machine learning models were trained from these initial results (Table 1). With thermal data, the KNN and NN are equally efficient algorithms, followed by RF and SVM. Thermal data outperformed optical data by <2%.

The most accurate models from this training dataset were then used to label a new video. For the thermal data, this is the neural network. For the visual data, the RF was used with the hyperparameters chosen by Bayesian search as entropy for the criterion, $\log_2$ features considered for a split and 1611 estimators used.

Under test conditions, the optical results were that 4053 flights were predicted as out, with only 1149 being correctly labelled. In addition,

Table 1
Results of the first training. Precision is measured as true positive count over true positive count plus false positive count. Recall is measured as true positive over true positive plus false negative. F1 score is a combination of recall and precision, known as the harmonic mean.

Video Type	Algorithm	Accuracy	Precision	Recall	F1-score
	NN	92.40%	0.912	0.904	0.908
Thermal	RF	89.93%	0.884	0.874	0.879
	SVM	90.55%	0.887	0.881	0.884
	KNN	92.40%	0.913	0.903	0.908
	NN	90.50%	0.892	0.891	0.891
Optical	RF	91.07%	0.894	0.893	0.893
	SVM	90.69%	0.900	0.901	0.900
	KNN	89.25%	0.871	0.897	0.884

Table 2
Results of the two following training iterations and the second test.

Type	Algorithm	Accuracy	Precision	Recall	F1-score
Thermal	NN	93.06%	0.907	0.905	0.906
Train	RF	93.02%	0.921	0.898	0.909
2	SVM	93.02%	0.920	0.889	0.904
	KNN	91.72%	0.886	0.894	0.889
Thermal Test 2	RF	95.56%	0.902	0.919	0.911
Thermal	NN	94.35%	0.931	0.887	0.909
Train	RF	94.70%	0.927	0.904	0.916
3	SVM	94.41%	0.916	0.908	0.912
	KNN	93.89%	0.908	0.900	0.904
Optical	NN	92.29%	0.889	0.912	0.900
Train	RF	92.34%	0.896	0.899	0.898
2	SVM	92.19%	0.894	0.899	0.897
	KNN	89.75%	0.865	0.875	0.870
Optical Test 2	RF	95.93%	0.923	0.955	0.939
Optical	NN	94.23%	0.919	0.915	0.917
Train	RF	94.37%	0.926	0.917	0.922
3	SVM	93.81%	0.917	0.912	0.915
	KNN	92.56%	0.893	0.909	0.901

5471 flights were predicted as invalid flights with 221 of these being incorrectly labelled. The overall accuracy of these predictions was 67.19%, however, precision was 0.84 and recall 0.28. These results were because the test dataset was several times larger than the training set, increasing the possibility of flights that bore no resemblance to any flight seen prior, making classification more difficult. The wide field of view of the optical camera meant many more variations of flights were possible. Lastly, the removal of the naive filter meant many flights that were previously removed for consideration by this filter now needed classification by the trained model. This was because some outgoing flights were filtered out by the naive system as the underlying tracking algorithm did not successfully track them to the edge of the frame.

Results for the thermal camera showed 1727 flights were labelled as outgoing of which 1328 were correctly labelled. 6647 flights were predicted as invalid flights and 111 of these were incorrectly labelled. This gave an accuracy of 93.91%, a precision of 0.77, and a recall of 0.92.

The data from this test was then added to the training set and further train/test iterations as in Table 2. Iterations identified features likely to aid in classification. The number and duration of loops were subsequently added, as was the nearest and furthest distance measure from both hive and closest exit. In addition, refinements were made to the parameters controlling the track extraction in MATLAB, particularly as the test results from the first optical test showed that flights were missing from the collected data. These refinements improved the selection of control parameters for the Gaussian models used in flight extraction based on the observations of the human labeller.

Accuracy started to plateau by the third training iteration. The more complex learning algorithms became noticeably more accurate and precise than the simpler KNN. In both cases, the algorithm selected for the second test was random forest, which became the strongest model in the final training iteration. For thermal data, the RF used entropy as the criterion, the square root of feature count to consider for splitting and 316 estimators. For optical data, the RF used the same parameters except for 2500 estimators, significantly higher than the thermal data, owing to the need to rely on more features to make decisions. The issues from the first optical test were no longer present by the second test, explained by the growth in the dataset covering more complex flights.

3.1. Feature importance

To confirm that the need for more complex models to classify optical data was due to the shape metrics, feature importance was extracted from the random forest and the support vector machine using a linear kernel. For thermal data, max shape ratio, max shape length, and final growth were the most important features across both models. Max shape ratio for SVM and max shape length for RF were the most valuable variables for classification and twice as important as any other variable.

For optical data, the smallest change in bearing and overall bearing versus exit angle was most important for the SVM and RF respectively. Only two shape metrics appeared in the five most important features for each algorithm and no metric was twice as predominant as the next. This showed that, for optical data, learning was more nuanced, involving greater use of the features available.

3.2. Wasp detection

In addition to bee tracks, wasps (*Vespula Vulgaris*) were also detected. There were too few of these detections for machine learning, at most four detections in a 20 min period. The presence of wasps supports that future development of this technique could include models able to count and monitor other insects that interact with a hive. For example, in Fig. 8 a wasp was observed to fly close to a hive and hover near the entrance, before finding an isolated bee on the hive platform. There, it flew in proximity to the bee until that bee then fled off the frame. The wasp then returned to the entrance of the hive before flying away when it was clear that there was no access to the hive.

Other interactions were observed, such as wasps flying directly to the entrance of the hive to be met by guard bees that then discouraged further investigation and tailed the wasp for a short distance as it left the area. Wasps will probe honeybee hives to test whether ingress to steal honey is feasible [31]. Thermal images of a wasp did not differ significantly from those of bees, especially when moving and the signature was blurred. Wasps appeared characteristically brighter than honeybees with optical cameras which will be useful for their identification. Their movement patterns are discernible from incoming, outgoing, and hovering bees. Interactions between these two species create more possible data permutations than just bee flights alone, which will necessitate more comprehensive data gathering.

3.3. Test stage

Once the final training iteration was completed, a test was conducted using a final recording. For this recording, human expertise was used to attain the true number of in, out, and invalid flights working with a recording as demonstrated in Fig. 9. Video times were cropped across camera systems to be an exact match, beginning and ending in synchronicity. Fused flights were identified and counted as lost flights, as were any double-counted flights as per split flights. Flights lost to failings in the underlying track extraction algorithm were also counted (Table 3).

The underlying tracking algorithm was 95.08% effective at recovering flights from optical recordings and 93.88% with thermal recordings. These are the useful flights (those that are not invalid.) The overall accuracy of the algorithms was 96.16% for optical data and 97.92% for thermal data, though the accuracy metrics were inflated by the greater

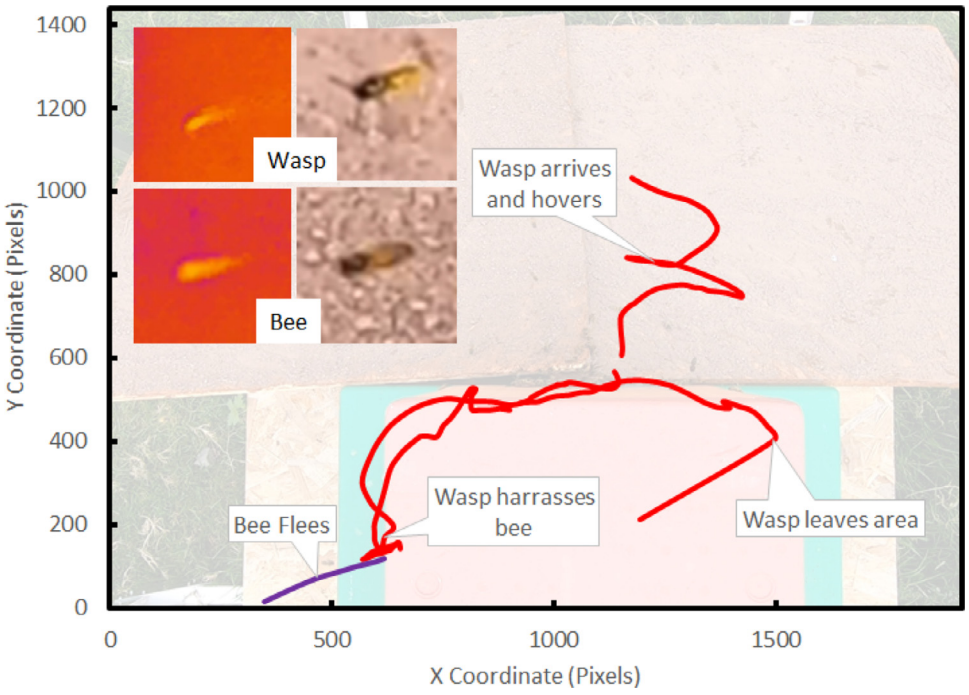

Fig. 8. The flight of a wasp captured investigating the hive and harassing bees, including image of the wasp taken by thermal and optical cameras compared with similar images of bees.

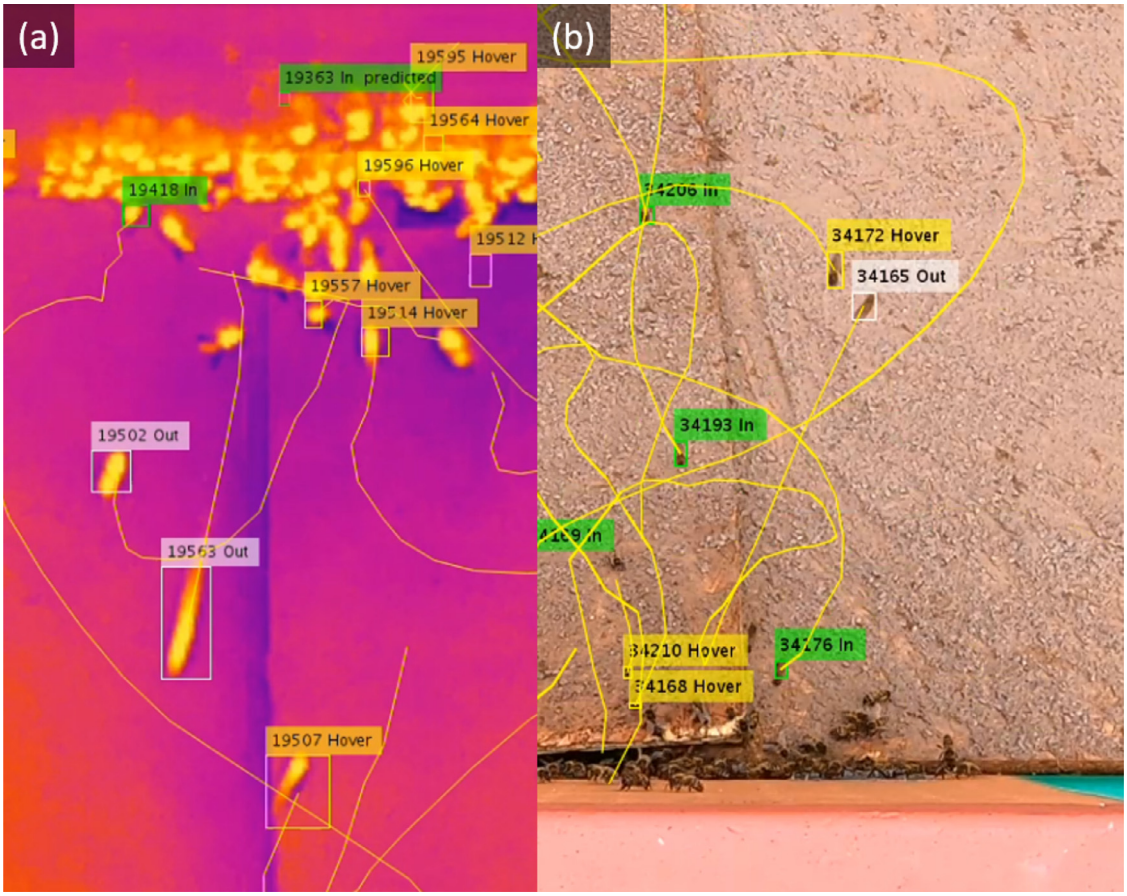

Fig. 9. Cropped screenshot of each camera system during the test phase, showing (a) thermal camera results and (b) optical camera results.

Table 3
Results of the test recording.

Type	Missed Flights	Counted In	Counted Out	False Positives	False Negatives	True Negatives
Visual	82	902	681	65	62	2562
Thermal	102	907	668	64	55	4992

size of the true negative (invalid/hovering flight) class. More useful measures would be a precision of 0.91 and recall of 0.90 for optical data. For thermal, precision was 0.92 and recall 0.90. For completeness, the original naive filter correctly labelled 549 flights with optical data and 286 with thermal data, far below the models created here.

Both cameras performed well in classification and flight extraction, however, the thermal camera missed an additional 20 flights because they occurred outside of the frame. This is an unavoidable 24.39% increase in missed flights over the optical camera system. The optical camera system achieved an accuracy of 95.53%, with precision of 0.92 and recall of 0.96, in its second to last test stage. This is a noticeably better result than the thermal camera at any of the test or train stages. The efficacy of each system will vary according to the environmental conditions on the days that they are used. The thermal camera can function under a wider range of lighting conditions compared to the optical system, making them more responsive under variable environmental conditions.

Further data is required to fully develop the dataset to account for all possibilities associated with seasonal changes affecting ambient temperature, wind conditions, and particularly rain. Rain would appear as a darker object moving within a frame before colliding with the hive entrance or ground and could be mistaken for bees. Wind will make the flights of bees more erratic. These issues could be addressed in future work. During our study we noted that bees were quick to return to the hive and reduce activity when conditions became too windy or wet.

4. Conclusion

We have demonstrated that thermal cameras are a contender for bee counting applications. Their flexibility in deployment is greater than that of optical cameras, working in poorly lit conditions and without visual aids. They require less modification of beehive entrances and are passive devices that will not disturb bee behaviour.

The results suggest that despite the smaller field of view, fewer pixels, and lower frame rate the thermal camera is comparable to the optical camera with the aid of machine learning. Both systems are at least 93% successful at extracting flights and 96% successful at classifying these flights. Recall and precision metrics demonstrate that the thermal system is equal and often better for classification due to the improved shape metrics gathered and the models are not as complex. Flights are more cost-efficient to recover from thermal footage due to fewer pixels to consider. As thermal cameras continue to decrease in price and increase in capabilities this advantage will grow.

Expanding the dataset for this work will allow the inclusion of variable weather conditions and environmental temperatures. This would create a robust model for the system described to function in all weather conditions.

Further work in this developmental pathway could include a more intelligent flight extraction algorithm, able to reconstitute broken flights from frame-edge and collision losses. A reinforcement model, able to tune the extraction parameters, would increase the fidelity of the data and permit improved analyses.

Consideration could also be made to spot other insects near the entrance of the hive and classify them based on shape and movement metrics similar to those here. An example would be the Asian Hornet which is known for hawking behaviour at the entrance to beehives.

Lastly, investigation to correlate flight behaviour with younger bees, particularly investigating orientation loops, would allow for the monitoring and counting of fresh foragers from a hive as a health metric for managed hives.

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.

Acknowledgments

This work was supported by KESS 2. Knowledge Economy Skills Scholarships (KESS 2) is a pan-Wales higher level skills initiative led by Bangor University on behalf of the HE sector in Wales. It is part funded by the Welsh Government's European Social Fund (ESF) convergence programme for West Wales and the Valleys.

References

[1] G. Espregueira Themudo, A. Rey-Iglesia, L. Robles Tascón, A. Bruun Jensen, R.R. da Fonseca, P.F. Campos, Declining genetic diversity of european honeybees along the twentieth century, Scientific Reports (2020), doi:10.1038/s41598-020-67370-2.
[2] V. Patel, N. Pauli, E. Biggs, L. Barbour, B. Boruff, Why bees are critical for achieving sustainable development, Ambio (2021), doi:10.1007/s13280-020-01333-9.
[3] S. Bermig, R. Odemer, A.J. Gombert, M. Frommberger, R. Rosenquist, J. Pistorius, Experimental validation of an electronic counting device to determine flight activity of honey bees (apis mellifera l.), Journal fur Kulturpflanzen (2020), doi:10.5073/JfK.2020.05.03.
[4] A.C. Fabergé, Apparatus for recording the number of bees leaving and entering a hive, Journal of Scientific Instruments 20 (1943) 28–31, doi:10.1088/0950-7671/20/2/302.
[5] A. Shaout, N. Schmidt, Bee hive monitor, in: 2019 International Arab Conference on Information Technology (ACIT), 2019, pp. 52–57, doi:10.1109/ACIT47987.2019.8990982.
[6] A. Souza Cunha, J. Rose, J. Prior, H. Aumann, N. Emanetoglu, F. Drummond, A novel non-invasive radar to monitor honey bee colony health, Computers and Electronics in Agriculture 170 (2020) 105241, doi:10.1016/j.compag.2020.105241.
[7] R. Odemer, Approaches, challenges and recent advances in automated bee counting devices: A review, Annals of Applied Biology n/a (2021), doi:10.1111/aab.12727.
[8] M.H. Struye, H.J. Mortier, G. Arnold, C. Miniggio, R. Borneck, Microprocessor-controlled monitoring of honeybee flight activity at the hive entrance, Apidologie (1994), doi:10.1051/apido:19940405.
[9] N. Aldabashi, S. Williams, A. Eltokhy, E. Palmer, P. Cross, C. Palego, Integration of 5.8ghz doppler radar and machine learning for automated honeybee hive surveillance and logging, in: 2021 IEEE MTT-S International Microwave Symposium (IMS), 2021, pp. 625–628, doi:10.1109/IMS19712.2021.9574826.
[10] A. Thielens, M.K. Greco, L. Verloock, L. Martens, W. Joseph, Radio-frequency electromagnetic field exposure of western honey bees, Scientific Reports 10 (2020), doi:10.1038/s41598-019-56948-0.
[11] B.R. Roberts, J.L. Osborne, Testing the efficacy of a thermal camera as a search tool for locating wild bumble bee nests, Journal of Apicultural Research (2019), doi:10.1080/00218839.2019.1614724.
[12] R.E. Bonoan, R.R. Goldman, P.Y. Wong, P.T. Starks, Vasculature of the hive: Heat dissipation in the honey bee (apis mellifera) hive, Naturwissenschaften (2014), doi:10.1007/s00114-014-1174-2.
[13] B.A. Klein, M. Stiegler, A. Klein, J. Tautz, Mapping sleeping bees within their nest: Spatial and temporal analysis of worker honey bee sleep, PLoS ONE (2014), doi:10.1371/journal.pone.0102316.
[14] T.D. Seeley, M. Kleinhenz, B. Bujok, J. Tautz, Thorough warm-up before take-off in honey bee swarms, Naturwissenschaften (2003), doi:10.1007/s00114-003-0425-4.
[15] Y. Abdelrahman, P. Wozniak, P. Knierim, N. Henze, A. Schmidt, Exploration of alternative vision modes using depth and thermal cameras, 2018, doi:10.1145/3282894.3282920.
[16] R. Gade, T.B. Moeslund, Thermal cameras and applications: A survey, Machine Vision and Applications 25 (2014), doi:10.1007/s00138-013-0570-5.
[17] R.S. Maguire, M. Hogg, I.D. Carrie, M. Blaney, A. Couturier, L. Longbottom, J. Thomson, A. Thompson, C. Warren, D.J. Lowe, Thermal camera detection of high temperature for mass covid screening, medRxiv (2021).
[18] M.N. Mohammed, H. Syamsudin, S. Al-Zubaidi, A.K. Sairah, R. Ramli, E. Yusuf, Novel covid-19 detection and diagnosis system using iot based smart helmet, International Journal of Psychosocial Rehabilitation 24 (2020).
[19] E. Villa, N. Arteaga-Marrero, J. Ruiz-Alzola, Performance assessment of low-cost thermal cameras for medical applications, Sensors 20 (2020), doi:10.3390/s20051321.
[20] S.M. Williams, N. Aldabashi, C. Palego, J.L. Woodgate, J.C. Makinson, P. Cross, Early prediction of bumblebee flight task using machine learning, Computers and Electronics in Agriculture (2021), doi:10.1016/j.compag.2021.106065.
[21] T.M. Inc, MATLAB 2018a, 2018.

[22] C. Stauffer, W.E. Grimson, Adaptive background mixture models for real-time tracking, Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (1999), doi:10.1109/cvpr.1999.784637.
[23] G. Welch, G. Bishop, An introduction to the kalman filter, In Practice (2006).
[24] J. Munkres, Algorithms for the assignment and transportation problems, Journal of the Society for Industrial and Applied Mathematics (1957), doi:10.1137/0105003.
[25] B. Magnier, G. Ekszterowicz, J. Laurent, M. Rival, F. Pfister, Bee hive traffic monitoring by tracking bee flight paths, in: VISIGRAPP 2018 - Proceedings of the 13th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications, 2018, doi:10.5220/0006628205630571.
[26] L. Breiman, Random forests, Machine Learning (2001), doi:10.1023/A:1010933404324.
[27] H. Drucker, C.J. Surges, L. Kaufman, A. Smola, V. Vapnik, Support vector regression machines, Advances in Neural Information Processing Systems, 1997.
[28] F. Rosenblatt, The perceptron: A probabilistic model for information storage and organization in the brain, Psychological Review (1958), doi:10.1037/h0042519.
[29] J. Mockus, On bayesian methods for seeking the extremum and their application, 1977.
[30] V. Nair, G.E. Hinton, Rectified linear units improve restricted boltzmann machines, in: ICML 2010 - Proceedings, 27th International Conference on Machine Learning, 2010.
[31] B.J. Donovan, Occurrence of the common wasp, vespula vulgaris (l.) (hymenoptera: Vespidae) in new zealand, New Zealand Journal of Zoology 11 (1984), doi:10.1080/03014223.1984.10428256.