Classifying Flies Based on Reconstructed Audio Signals

Michael Flynn and Anthony Bagnall

The University of East Anglia

Abstract. Advancements in sensor technology and processing power have made it possible to create recording equipment that can reconstruct the audio signal of insects passing through a directed infrared beam. The widespread deployment of such devices would allow for a range of applications previously not practical. A sensor net of detectors could be used to help model population dynamics, assess the efficiency of interventions and serve as an early warning system. At the core of any such system is a classification problem: given a segment of audio collected as something passes through a sensor, can we classify it? We examine the case of detecting the presence of fly species, with a particular focus on mosquitoes. This gives rise to a range of problems such as: can we discriminate between species of fly? Can we detect different species of mosquito? Can we detect the sex of the insect? Automated classification would significantly improve the effectiveness and efficiency of vector monitoring using these sensor nets. We assess a range of time series classification (TSC) algorithms on data from two projects working in this area. We assess our prior belief that spectral features are most effective, and we remark on all approaches with respect to whether they can be considered "real-time".

Keywords: Insect classification · Time series classification · Spectral features.

1 Introduction

Over the last century there have been many attempts at solving insect classification problems. Increased interest in classifying insects has been fuelled by a number of factors. Insects are responsible for the pollination of the majority of crop species, but are also vectors for disease and responsible for a massive number of fatalities. Monitoring the presence and abundance of mosquitoes is crucial in understanding the population dynamics and effectiveness of interventions. Quantifying the abundance of an insect specie in a natural setting is challenging. Typically, expert entomologists are required to manually identify species using morphological differences. This can result in a lengthy delay in detection and quantification and the amount of data that can be collected is limited. However, recent advances in sensor technology has made the collection of large datasets more feasible [14] [19] [16] [2]. These approaches, described in more detail in Section 3, record data as an object passes through a target area. This results in a data segment that can be interpreted as audio and used to classify the object. As with most audio problems, standard classification approaches to this problem use features in the frequency domain. We assess a range of standard techniques in addition to recently proposed algorithms for general time series classification (TSC) problems. The most successful algorithms for TSC (the classification of real valued, ordered series such as audio) are based on transformation and ensembling. The effectiveness of these techniques is explored in the Hierarchical Vote Collective of Transformation-based Ensembles (HIVE-COTE) [7]. We perform a thorough experimental evaluation of a range of classifiers, three of which are used in HIVE-COTE, and conclude that in this case TSF is most effective.

The rest of this paper is structured as follows. In Section 2 we discuss what motivates insect classification, and describe current techniques used for this problem. In Section 3 we describe two data sets collected by research groups in USA and Germany which we use in the experimental evaluation. In Section 4 we outline the methods used in our evaluation and in Section 5 we present our results. We conclude in Section 6.

2 Background

Traditionally, insects are classified by their morphological differences. This becomes more difficult as you move down through taxonomic ranks. In some cases, it is not possible to classify species without gene sequencing [4]. Early investigations into the classification of flying insects were focused on wingbeat frequency [1][15] which was collected manually via the use of a stroboscope. The effectiveness of this data as a class predictor was then quantified using standard statistical modelling. These studies also quantified the effects of air temperature and location on wingbeat frequency. However, findings were often inconclusive and lacked robustness due to small sample sizes. As technology advanced, recording insect wingbeats became feasible. This allowed additional spectral information to be utilised, such as harmonics [10][18]. These early studies concluded that wingbeat frequency alone is not an adequate predictor of class. It follows a normal distribution and exhibits significant intra class variability. This results in substantial overlap between wingbeat frequencies of different classes, a problem that is only made worse as the number of classes increase.

The development of robust phototransitor recording techniques, capable of operating for extended periods of time, provided the first medium size datasets to work with [9]. The advent of artificial neural networks also provided a novel approach to classification. Many studies went on to report increases in accuracy when including or using entirely spectral features [11][6]. It was also noted that classification could be confounded by the insect behaviour, such as the relative insect trajectory, the fact that some species tendency to buzz their wings before takeoff and the effect of the circadian rhythm on behaviour.

Researchers at the University of California, Riverside (UCR) have developed a low cost recording system and used it to produce the first large high dimension multiclass insect wingbeat problem [2]. They also went on to establish a benchmark accuracy using a relatively simple Bayesian based approach, augmented with temporal information. A similar performance has been achieved via a combination of power spectral density features and Gaussian mixture models [12] and, via the convolutional neural network AlexNet, which was used to extract features from spectrograms before they were classified, via an support vector machine [20].

3 Datasets

We use two publicly available datasets for insect classification, summarised in table 1. The first, InsectWingbeat, comes from the ongoing project at the University California Riverside (UCR) [2] and is part of the UCR TSC archive[1]. The second, MosquitoWingbeat, was produced during the development of a low cost insect sensor at TEII [14], and was recently used in a Kaggle competition[2].

In the case of both datasets, perspex boxes were used to confine flies of each class for recording. In the case of the InsectWingbeat dataset, the four mosquito species were also separated by sex. Both systems use a combination of infrared LEDs and photodiodes to record fly behaviour. Recordings are made as the wings and body occlude the signal from the LEDs during flight. The signal produced can be interpreted as audio and captures data similar to that of conventional audio recording devices [13].

The InsectWingbeat dataset contains 50,000 one second audio segments recorded with a sample rate of 16kHz. There are ten equally distributed classes, comprised of four mosquito species (differentiated into male and female) from two genera and two other fly species from different genera (not differentiated by sex). These are Ae. aegypti, Cx. stigmatosoma, Cx. tarsalis, Cx. quinquefasciatus, Mu. domestica and Dr. simulans.

The MosquitoWingbeat dataset is comprised of six mosquito species from three genera. These are Ae. aegypti, Ae. albopictus, An. arabiensis, An. gambiae, Cu. pipiens, Cu.quinquefasciatus. There is no differentiation between sexes. It consists of 279,566 instances of 0.625 seconds of audio segments samples at 8kHz. For the purpose of this paper, the number of instances per class has been reduced to 5000, reducing runtime and creating equal class distribution.

Table 1. Summary of attributes for datasets.

Dataset	Instances	Classes	Attributes	Sample rate	% of second
InsectWingbeat	50,000	10	16000	16kHz	100
MosquitoWingbeat	30,000	6	5000	8kHz	62.5

[1] http://www.timeseriesclassification.com
[2] https://www.kaggle.com/potamitis/wingbeats

4 Methods

The UCR approach, laid out briefly in section 2, consists of a prepossessing step and a classification step. During the preprocessing step, instances are transformed into the frequency domain via the Fast Fourier Transform algorithm (FFT). Of the resulting 16,000 attribute output vector, indicies 100 - 2000 are kept and form the data used for classification. In the classification stage, a Nearest neighbour (k-NN) approach is used in conjunction with Euclidean distance. Through leave one out cross validation, k was set as 8.

A second approach, also outlined in section 2 uses the well known neural network AlexNet to derive features that are subsequently classified via a support vector machine (SVM). In this case instances are transformed into spectrogram images prior to classification. The number of FFT bins used is 512 and windows overlapped by 50%.

In section 5, we go on to publish results of approaches that have not been applied to the problem of insect classification. These include: Shapelet Transform (ST) [8] used with the C4.5 decision tree. In this approach, the dataset is transformed under a 48Hr contract, such that it is expressed in terms of intervals which are class discriminant. A C4.5 decision tree is then grown on the training data; Time Series Forest (TSF) [3], in which random intervals are selected and distilled into statistical features that are used to grow C4.5 decision trees; the Bag of SFA Symbols (BOSS) [17], in which instances are first split and compiled into a dictionary of words represented as histograms and classification takes place via a 1-NN used in conjunction with a bespoke distance measure; the contract Random Interval Spectral Ensemble (cRISE), in which random intervals are selected and transformed into spectral and autocorrelation coefficients. These new representations are then combined before being used to grow random decision trees. We also evaluate two ensembles. The first consists of cRISE contracted to 1 hour of training, BOSS contracted to 1 hour of training and TSF. In this approach (CAWPE) we use a scheme in which constituents are weighted by cross-validated accuracy estimates [5]. The second approach is The hierarchical vote collective of transformation-based ensembles for time series classification (HIVE-COTE). This consists of TSF, cRISE contracted to one hour, BOSS contracted to one hour and ST contracted to 48 hours with C4.5.

Furthermore, we evaluate all approaches, other than cRISE which manages transformation internally, in combination with two preprocessing approaches as well as the raw data. The first approach, labelled T-1, sees instances resampled to 6000Hz prior to transformation and the entire output is used. This reduction in sampling is motivated by evidence that these insects have little to capacity to produce frequencies exceeding 3000Hz [14]. The second approach, labelled T-2, is the preprocessing step of UCR approach defined at the beginning of this section.

5 Results

In order to produce robust results from which to draw our conclusions, all experiments were subject to a stratified 10-fold cross validation. In the interest of producing reproducible results, all random functions used to produce data folds were seeded.

The rest of this section is organised as follows. In section 5.1 we evaluate the accuracy achieved by benchmark classifiers using just the fundamental frequency attribute. In section 5.2 we investigate the performance of approaches in conjunction with spectral features and in section 5.3 we present and discuss all approaches in respect to timing.

All code used in these experiments is available from the UEA TSC repository[3] and raw results and analysis is available at[4].

5.1 Fundamental frequency

The fundamental frequencies of the instances in both the MosquitoWingbeat and InsectWingbeat datasets were extracted using a peak finding algorithm in conjunction with the harmonic product spectrum technique. Table 2 displays results from experiments undertaken with these datasets. At just over 50% accuracy, the performance of this feature alone is in-line with results seen in literature evaluating similar datasets.

Table 2. Table showing mean accuracy, the Area under the reciever operator curve (AUROC) and Negative log liklihood (NLL) for 1 Nearest Neighbour with Euclidean distance and Naive Bayes approaches, evaluated over 10 folds on the fundamental frequency attribute of the MosquitoWingbeat (6 classes) and InsectWingbeat (10 classes) datasets.

Dataset	Classifier	Accuracy	AUROC	NLL
InsectWingbeat	ED	**0.56**	0.74	2.93
	NB	0.45	**0.83**	**2.49**
MosquitoWingbeat	ED	**0.56**	0.74	2.95
	NB	0.53	**0.78**	**1.89**

5.2 Spectral approaches

Table 3 shows the results of cRISE, 8-NN, BOSS, TSF, ST and the ensembles of CAWPE and HIVE-COTE. The results are separated by dataset and ordered with respect to accuracy. All transform/classifier combinations are also published.

[3] https://github.com/TonyBagnall/uea-tsc
[4] https://tinyurl.com/yxqgfl9e

The results shown in table 3 confirm our prior belief, "*that spectral features are most effective*". This is most obvious when looking at the results of TSF in respect to InsectWingbeat. In this case, we see an increase of 28% in accuracy between spectral and non-spectral features. However, in all cases other than CAWPE and ST accuracy differs by at least 8%.

The effect T-1 and T-2 have on accuracy are confined to the MosquitoWingbeat dataset. Table 3 shows that for the MosquitoWingbeat dataset TSF differs in accuracy by 8%, HIVE-COTE by 6%, ST by 3% BOSS by 4% and CAWPE by 7%. Physical differences used to produce the MosquitoWingbeat dataset result in a larger target area. This results in insects being recorded for a greater duration and ultimately results in signals containing more low energy information, information which the T-2 approach discards.

Overall, HIVE-COTE is the most accurate. On the InsectWingbeat dataset the HIVE-COTE approach is 14% more accurate than the current state of the art approach, 8-NN+T-2. The most effective approach on the MosquitoWingbeat dataset, HIVE-COTE+T-1, is 16% more accurate than the 8-NN+T-2 combination and 9% more accurate than TSF+T-2. Significantly, these results omit powerful time-of-flight information, an attribute that is reported to have significantly increased the accuracy of the 8-NN+T-2 combination to 79.44% [2] on the InsectWingbeat dataset.

5.3 The relevance of test time.

The successful application of classification algorithms in real world scenarios also require them to be timely. It is commonly accepted that an algorithm is "real time" if it is able to classify an instance in less time than is represented in the data. Instances from the MosquitoWingbeat represent 620 milliseconds and those from the InsectWingbeat represent 100 milliseconds.

Figure 1 plots mean test time per instance averaged over folds for each approach. The timing data was generated during experiments run on the spectral datasets, the results of which were discussed in section 5.2. Results of non-spectral experiments have been omitted in the interest of brevity.

In all cases TSF performs best and in a timely manner with respect to relative instance length. We note, it also exhibited very little variance across folds. In respect to timing, the UCR transformation approach performs best overall. This is most clear when comparing InsectWingbeat+T-1 and InsectWingbeat+T-2 in respect to TSF. Our observation indicates this is down to the reduced number of instances attributes present for classification.

6 Conclusion

In conclusion, we have shown that the combination of simple audio features and HIVE-COTE outperforms all approaches evaluated in this paper on both the MosquiotWingbeat and InsectWingbeat datasets. Whilst omitting powerful time of flight information HIVE-COTE in conjunction with spectral features

Table 3. Table showing mean accuracy, AUROC and NLL over 10 folds for ST+C4.5, TSF, cRISE, 8-NN, BOSS and CAWPE ensembles for T-1, T-2 and no spectral transformation.

Dataset	Classifier	Transform	Accuracy	AUROC	NLL
InsectWingbeat	HIVE-COTE	T-1	**0.7951**	**0.9794**	1.1384
	HIVE-COTE	T-2	0.7821	0.9780	1.1258
	TSF	T-1	0.7540	0.9751	0.9667
	TSF	T-2	0.7526	0.9748	**0.9316**
	CAWPE	T-1	0.7482	0.9731	1.2560
	CAWPE	T-2	0.7442	0.9728	1.2316
	cRISE	n/a	0.7172	0.9642	1.3791
	CAWPE	none	0.7138	0.9616	1.9251
	BOSS	T-2	0.6668	0.9496	1.2531
	8-NN	T-2	0.6626	0.9308	1.9543
	BOSS	T-1	0.6620	0.9474	1.2650
	8-NN	T-1	0.6556	0.9275	2.0613
	HIVE-COTE	none	0.6404	0.9616	1.6395
	ST+C4.5	T-2	0.6239	0.8095	2.3152
	ST+C4.5	T-1	0.6229	0.8076	2.3386
	ST+C4.5	none	0.5805	0.7854	2.6107
	BOSS	none	0.5751	0.8962	1.9126
	8-NN	none	0.5639	0.9009	2.7342
	TSF	none	0.4744	0.8647	2.3057
MosquitoWingbeat	HIVE-COTE	T-1	**0.8141**	**0.9680**	0.9874
	TSF	T-1	0.7956	0.9643	**0.8188**
	CAWPE	T-1	0.7881	0.9606	1.0271
	HIVE-COTE	T-2	0.7505	0.9487	1.1580
	TSF	T-2	0.7225	0.9407	1.0180
	CAWPE	T-2	0.7222	0.9395	1.2056
	CAWPE	none	0.7149	0.9327	1.6154
	cRISE	n/a	0.6808	0.9199	1.4349
	BOSS	T-1	0.6768	0.9193	1.2076
	BOSS	T-2	0.6445	0.9035	1.3160
	HIVE-COTE	none	0.6310	0.9287	1.4489
	ST+C4.5	T-1	0.5937	0.7755	2.5144
	ST+C4.5	none	0.5772	0.7808	2.6137
	ST+C4.5	T-2	0.5600	0.7626	2.6760
	TSF	none	0.5189	0.8363	1.8608
	BOSS	none	0.4632	0.7831	2.0347
	8-NN	T-2	0.3829	0.7257	4.3566
	8-NN	T-1	0.2840	0.6402	5.3337
	8-NN	none	0.2539	0.5885	6.1839

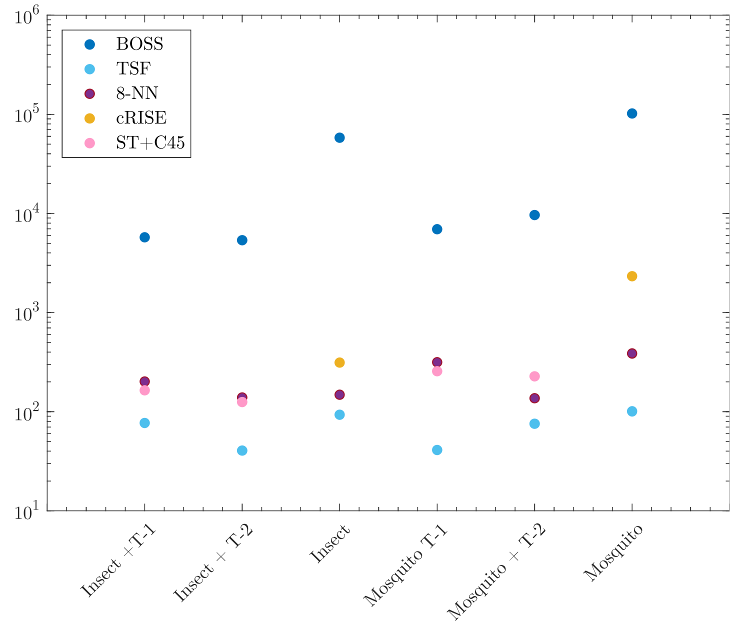

Fig. 1. Figure showing mean test time per instance for all combinations of ST+C4.5, TSF, cRISE, 8-NN, BOSS, CAWPE ensembles, with no spectral transformation, T-1 and T-2 transformations.

is shown to be 14% more accurate than the benchmark approach on the InsectWingbeat dataset and 16% more accurate on the new MosquitoWingbeat dataset. However, as the slowest constituent of HIVE-COTE, BOSS, does not perform in a timely manner we conclude that even if threaded HIVE-COTE is not timely and therefore would need a considerably faster processor to meet the requirements of an application setting.

We conclude that intra class variance in fundamental frequency prevents its use as a discriminant between species. However, in a real world setting this feature is likely to play a key role in determining candidate intervals in an application setting. Our view is that an appropriate algorithm architecture would consist of layers designed to minimise power consumption, by preventing unnecessary computations. In this context, fundamental frequency could prove an adequate method of solving both $insect|noninsect$ and $fly|nonfly$ decisions.

Acknowledgements

This work is supported by the Biotechnology and Biological Sciences Research Council [grant number BB/M011216/1], and the UK Engineering and Physical

Sciences Research Council (EPSRC) [grant number EP/M015807/1]. The experiments were carried out on the High Performance Computing Cluster supported by the Research and Specialist Computing Support service at the University of East Anglia and using a Titan X Pascal donated by the NVIDIA Corporation.

References

1. Chadwick, L.E.: A simple stroboscopic method for the study of insect flight. Psyche **46**(1), 1–8 (1939)
2. Chen, Y., Why, A., Batista, G., Mafra-Neto, A., Keogh, E.: Flying insect classification with inexpensive sensors. Journal of insect behavior **27**(5), 657–677 (2014)
3. Deng, H., Runger, G., Tuv, E., Vladimir, M.: A time series forest for classification and feature extraction. Information Sciences **239**, 142–153 (2013)
4. Greven, H., Kaya, M., Junker, K., Akyuz, L., Amemiya, C.T.: Characterization of tongue worm (pentastomida) chitin supports α-rather than β-chitin. Zoologischer Anzeiger **279**, 111–115 (2019)
5. Large, J., Lines, J., Bagnall, A.: A probabilistic classifier ensemble weighting scheme based on cross-validated accuracy estimates. Data Mining and Knowledge Discovery pp. 1–36 (2019)
6. Li, Z., Zhou, Z., Shen, Z., Yao, Q.: Automated identification of mosquito (diptera: Culicidae) wingbeat waveform by artificial neural network. Artificial Intelligence Applications and Innovations pp. 483–489 (2005)
7. Lines, J., Taylor, S., Bagnall, A.: Time series classification with HIVE-COTE: The hierarchical vote collective of transformation-based ensembles. ACM Transactions on Knowledge Discovery from Data **12**(5) (2018)
8. Lines, J., Davis, L.M., Hills, J., Bagnall, A.: A shapelet transform for time series classification. In: Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining. pp. 289–297. ACM (2012)
9. Moore, A.: Development of a data acquisition system for long-term outdoor recording of insect flight activity using a photosensor. In: 13th Conference on Aerobiology and Biometeorology, American Meteorological Society, Albuquerque, New Mexico (1998)
10. Moore, A., Miller, J.R., Tabashnik, B.E., Gage, S.H.: Automated identification of flying insects by analysis of wingbeat frequencies. Journal of economic entomology **79**(6), 1703–1706 (1986)
11. Moore, A., Miller, R.H.: Automated identification of optically sensed aphid (homoptera: Aphidae) wingbeat waveforms. Annals of the Entomological Society of America **95**(1), 1–8 (2002)
12. Potamitis, I.: Classifying insects on the fly. Ecological informatics **21**, 40–49 (2014)
13. Potamitis, I., Rigakis, I.: Novel noise-robust optoacoustic sensors to identify insects through wingbeats. IEEE Sensors Journal **15**(8), 4621–4631 (2015)
14. Potamitis, I., Rigakis, I.: Large aperture optoelectronic devices to record and timestamp insects wingbeats. IEEE Sensors Journal **16**(15), 6053–6061 (2016)
15. Reed, S., Williams, C., Chadwick, L.: Frequency of wing-beat as a character for separating species races and geographic varieties of drosophila. Genetics **27**(3), 349 (1942)
16. Sarpola, M., Paasch, R., Mortensen, E., Dietterich, T., Lytle, D., Moldenke, A., Shapiro, L.: An aquatic insect imaging system to automate insect classification. Transactions of the ASABE **51**(6), 2217–2225 (2008)

17. Schäfer, P.: Bag-of-sfa-symbols in vector space (boss vs) (2015)
18. Unwin, D., Ellington, C.: An optical tachometer for measurement of the wing-beat frequency of free-flying insects. Journal of Experimental Biology **82**(1), 377–378 (1979)
19. Wen, C., Guyer, D.: Image-based orchard insect automated identification and classification method. Computers and Electronics in Agriculture **89**, 110–115 (2012)
20. Zhang, C., Wang, P., Guo, H., Fan, G., Chen, K., Kämäräinen, J.K.: Turning wingbeat sounds into spectrum images for acoustic insect classification. Electronics Letters **53**(25), 1674–1676 (2017)