Mosquito wingbeat analysis and classification using deep learning

Eleftherios Fanioudakis
Dept of Music Technology & Acoustics
Techn. Educational Institute of Crete
Rethymno, Greece
eleftherios.fanioudakis@gmail.com

Matthias Geismar
Biogents
Regensburg, Germany
Mathias.Geismar@stud.uni-regensburg.de

Ilyas Potamitis
Dept of Music Technology & Acoustics
Techn. Educational Institute of Crete
Rethymno, Greece
potamitis@staff.teicrete.gr

***Abstract*—We examine the signal and the attributes of mosquitoes' wingbeat. Subsequently we carry on large-scale classification experiments based on optical recordings of mosquitoes' wingbeat of the following species: *Aedes aegypti, Aedes albopictus, Anopheles arabiensis, Anopheles gambiae, Culex pipiens, Culex quinquefasciatus*. We report 96% classification accuracy on the species level for a database of 279,566 flight recording cases using top-tier deep learning techniques. The database and the associated code are offered open. The long-standing goal is to run prediction models, perform risk assessments, issue warnings and make historical analysis based on wingbeats acquired through suction traps deployed in the field.**

Keywords—wingbeat, smart traps, deep learning, Culex

I. Introduction

Mosquitoes infected with bacteria, viruses or parasites may transmit diseases to humans and livestock. Serious diseases and pathogens that can be transmitted by mosquitoes include: malaria, West Nile virus, Zika virus, chikungunya, yellow fever, dengue, lymphatic filariasis, and many forms of encephalitis. It has been reported that nearly 700 million people get a mosquito borne illness each year resulting in greater than one million deaths [1].

Prevention of the vector borne diseases is best achieved by vector control which, today in Africa, relies on the use of insecticides. Surveillance and monitoring mosquito vector populations is an integral component of most vector control programs and a prerequisite for effective interventions.

Public Health departments establish comprehensive surveillance and control programs based on mosquito traps. Due to the worldwide spread of invasive mosquitoes and mosquito-borne pathogens many surveillance programs have been applied that include a monitoring stage using trapping devices [2-5]. A major limitation of current surveillance and monitoring of mosquito vectors is the lack of an automated reporting service of the number and species composition of the captured insects [6]. Manual counting requires highly qualified personnel and is tedious as the pest manager must cover long distances since traps are dispersed at large spatial scale and located in not always easily reachable areas. Manual inspection of dispersed traps can increase the cost of the surveillance program or induce compromises in its design.

There are some attempts to automatize the flow of information and decision process on mosquito species identification and estimation of their age via remote sensors and geographic information system (GIS) analyses, but most of these technologies are still in the research stage [7-9]. A device, called the BG-Counter, which automatically differentiates mosquitoes from other insects entering a trap, counts them, and wirelessly transmits the results to a cloud server has recently been commercially available by Biogents AG [10-11]. The BG-Counter allows to remotely measure the dynamics of nuisance caused by mosquito populations to apply mosquito control measures more focused in time and space. The following step is to be able to differentiate mosquito genera or species.

The first and most important step in the development of new traps endowed with the capability of discriminating mosquito species is therefore to evaluate with carefully designed, large-scale experiments the quality of the discriminative information the wingbeat bears. The aim of this paper is to validate the classification accuracy that can be achieved from optical sensors involving six mosquito species from the three most important genera making it, to the best of our knowledge, the largest reported experiment on mosquito species classification based on wingbeats. Optical sensors are a promising alternative technology to microphones/ vision/ multispectral cameras for the task of discerning species and sex. The wingbeat recording is based on measuring the light amplitude variation from an emitter to a receiver as the insect flying partially occludes the light path with its flapping wings [12-13]. Deep learning (DL) techniques currently revolutionize the way machine learning is applied [14-15]. In this work we use advanced, top-tier DL architectures that are adapted to encompass the 1D nature of the wingbeat recording. The outcome is highly successful and embeddable in commercial mosquito traps.

II. Dataset & Signals

A. *Recordings Acquisition*

The recordings of the database took place at Biogents in Regensburg Germany, in a mosquito breeding establishment at approximately 26-27 ºC and 53-74% humidity for all experiments. Counting from the time of being able to fly all species were 3-10 days old. We placed 200-300 adult insects of both sexes strictly of a single species in cages and the recording of their wingbeat occurs the moment they pass through the

ISBN 978-90-827970-1-5

rectangle of the sensors on a random basis [12-13]. This setting allows us to take effortless and secure recordings of a large number of flight cases in a practical way, even for quarantine pests, as insects are generally difficult to handle in free-flight experiments. We place the same sensor in each cage holding only a single species, in turn and we pick-up several thousands of recordings from the SD card of the device after a day and was recorded using the optical device referenced in [13]. Each snippet has a length of 5000 samples at a 8KHz sampling rate. The details of the 'Wingbeats' database are depicted in Table I.

TABLE I. WINGBEATS DATABASE

Species	#recs
Ae. aegypti	85553
Ae. Albopictus	20231
An. Gambiae	49471
An. Arabiensis	19297
Cu. pipiens	30415
Cu. quinquefasciatus	74599
Total	279566
Train (80% total)	223652
Test (20% total)	55914

B. *The wingbeat signal*

In Fig. 1 you see a typical example of a wingbeat snippet. The mosquito performed a free flight. One can discern a high frequency signal superimposed on a low frequency signal. The low frequency part is due to the main body movement, whereas the high frequency is due to the wingbeat alone.

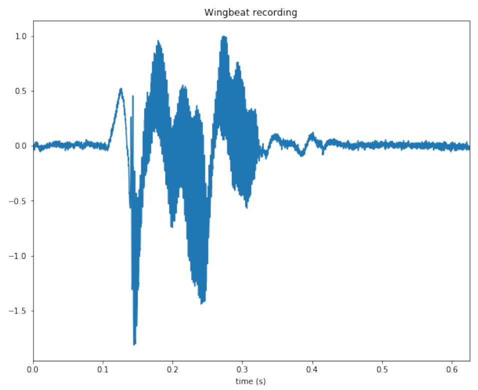

Fig. 1. An optical recording of an *Aedes aegypti* wingbeat.

If Fig. 2 one can see the frequency content of the wingbeat signal through the Welch power spectral density (PSD) estimation method set to average the PSD of 256 sample-chunks, with 192 samples overlap. One can see significant power at the low frequencies close to 0 and up to 80 Hz. This is due to the main body movement. An object falling of the size of the mosquito would also have power around these frequencies. At around 550 Hz one can see the first peak. This corresponds to the wingbeating frequency of the insect (we will call it *f0*). Note that the harmonics appear at integer multiplies of the *f0*. The harmonic structure in practice, is not very thin for two reasons: a) the signal is short in time and, b) the insect performs manoeuvres and modulates its frequencies as better seen in Fig. 3.

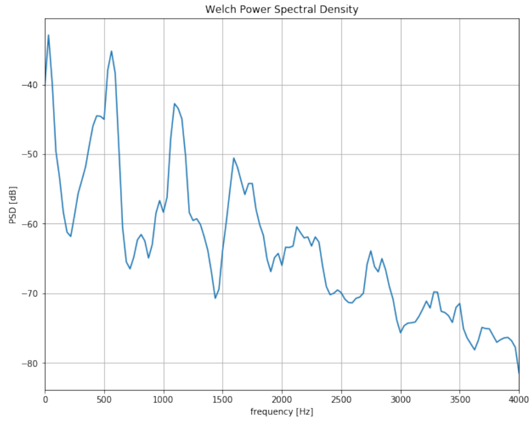

Fig. 2. The frequency content of a wingbeat signal. The main body movement is picked up at frequencies below 100Hz, the *f0* is seen around 550 Hz and the harmonics at multiples of the *f0*.

Mosquitoes in particular, are capable of changing their higher harmonics during courtship [16]. In Fig. 3 we see a spectrogram of the same wingbeat (i.e. the frequency composition of the wingbeat as it changes over time). One can see the characteristic harmonic structure of the wingbeat and the modulation of the harmonics. In the recognition experiments using Deep networks the 2D representation in the form of a spectrogram performed better probably because the convolutional filters picked up the frequency modulated patterns that cannot be picked up by the PSD representation of Fig. 2 as PSD integrates over time.

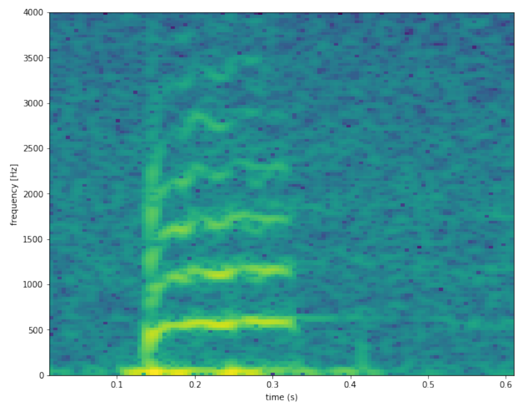

Fig. 3. Spectrogram of the wingbeat signal. Notice the frequency modulation present in the harmonic structure.

III. CLASSIFICATION EXPERIMENTS

In [17] we have shown that the generated optical fluctuations modulated by the wingbeats of insects, though short in time, offer enough information that can be used to discern species adequately. One can get a variety of features out of a recording but, we believe that the unprocessed spectrum and the spectrogram containing the fundamental and its harmonics and possibly certain simple transformations of it (e.g. frequency pooling through a filter-bank, logarithmic amplitude

ISBN 978-90-827970-1-5 © EURASIP 2018

compression) are a better choice than estimating the *f0*, the harmonics, and autoregressive features for this task. Note that class labels are known for the entire dataset as each insectary cage contains strictly insects of the same species and therefore, daily recordings from each cage all receive the same label tag. We have run a number of experiments on a random 20% holdout set. All experiments used the same random split. The results are gathered in Table I. Raw samples means that the time domain recording has been used as input without any feature extraction. PSD stands for a 129-dimensional vector corresponding to the log-power of the frequencies of each recording. Note that Raw samples as well as the PSD are 1D signals. Spectrogram is a time-frequency domain representation (the FFT size is set to 256 samples and the hopping length to FFT_SIZE/6), therefore, a 2D signal.
Regarding data augmentation, we randomly roll the time domain signal up to 15% of its length randomly left or right.
The results out of a number of approaches are gathered in Table II. Shallow learners are applied directly on the PSD. State of the art DL architectures are also applied and boost the recognition performance up to 96%.

TABLE II. CLASSIFICATION RESULTS

Deep Learning architectures		
Table column subhead	**Features**	**(%) acc**
DenseNet121	Spectrogram	96
5 layer CNN+	PSD	92.1
5 layer CNN+	Raw samples	91.2
InceptionV3	Spectrogram	95.20
MobileNet	Spectrogram	95.62
Xception	Spectrogram	92.18
NASNetMobile	Spectrogram	94.85
Shallow Learning approaches		
XGBoost	PSD	81.81
LightGBM	PSD	82.4

\+ Details of the architecture in the Appendix links

In Fig. 4 we present the confusion matrix for the DenseNet case. There are three crucial point in this picture:

a) The structure is strongly diagonal indicating that the confusion rates are small.

b) The classes are arranged so that by two, they belong to the same genus (i.e. *Anopheles, Aedes, Culex*). Notice the coloration in the sidebands of the confusion matrix inside each genus. Most errors appear between species of the same genus.

c) Focus on the cases of *Anopheles gambiae* and *Anopheles arabiensis* that both belong to the *Anophels gambiae* complex and are important malaria vectors in equatorial Africa. In research and control programs, the identification of what species is present is often important. However, adults of both species are morphologically indistinguishable, and the identification can currently only be performed on the molecular level. In this dataset, we report for the first time that we were able to discriminate between *An. gambiae* s.s. and *An. arabiensis*, using an opto-acoustic analysis. Both populations could be distinguished on a highly significant level.

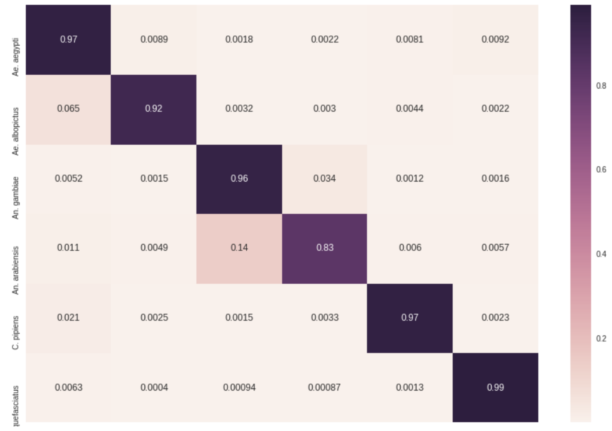

Fig. 4. Confusion Matrix of the DenseNet on the Species on a 20% hold out set. One can see clearly the diagonal structure of the confusion matrix indicating relatively low confusion rates.

IV. CLUSTERING, ATTENTION & TRANSFER LEARNING

Using DL for classification comes along with the benefit of a powerful toolkit for visualization, namely: clustering of features taken from intermediate layers and saliency maps, the latter aiming at identifying the most visually distinctive parts in a wingbeat spectrogram that affect classifier's decision.

A. *Clustering*

We took a DenseNet121 trained on the Wingbeats database and removed the last 2 layers. A balanced subset out of the 20% holdout set excluded from the training of the CNN is propagated through the layers of the truncated DenseNet to predict features instead of classes. We then used t-SNE [18] to compute a 2-dimensional embedding of the predicted features. t-SNE arranges wingbeats that have a similar CNN code nearby in the embedding space. We visualize the embedding and we demonstrate that the six wingbeat mosquito classes are well separated and wingbeats belonging to the same species cluster together (see Fig. 5). Although professional mosquito traps include specific attractants (e.g. CO_2 and specialized scents) it is possible that a non-targeted flying insect species is sucked in. Embedded classifiers will tend to classify the flying-in insect in their predefined classes. Clustering can serve as a general outlier detector that detects strange incidents outside clusters. The detection of atypical situations based on non-conforming wingbeats with predefined clusters of targeted insects reduces false counts.

B. *Salience Mapping*

There is a concern to find methods that quantify on which grounds the CNNs base their decision. In the context of wingbeat classification, we need to ensure that the decision was based on the properties of the information bearing parts of the spectrum. These parts are the lower frequencies close to [0-100] Hz, the fundamental frequency corresponding to the

ISBN 978-90-827970-1-5 © EURASIP 2018

wingbeat frequency of the insect (e.g. around 600 Hz for the first figure on the left) and its harmonics (the almost parallel

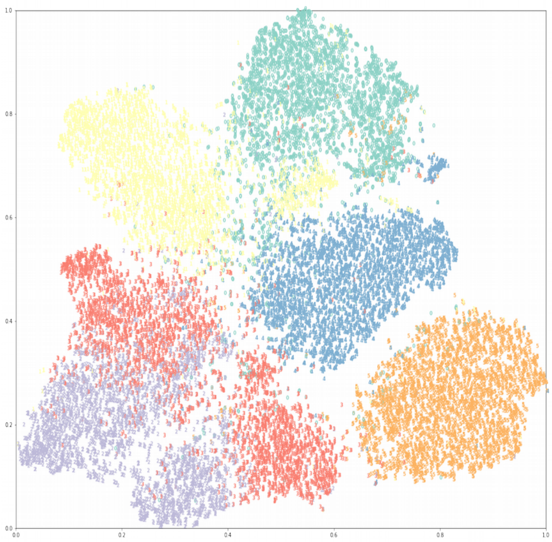

Fig. 5. Clustering the wingbeat features produced by a DenseNet121 using t-SNE. Six clusters for six mosquito species.

spectral lines at integer multiples of the fundamental). We performed an experiment where we used a mixture of two recordings from two different species (i.e. *Ae. aegypti* and *An. arabiensis*, Fig. 6-left and middle figures). The spectrogram of the mixture is Fig.6-right. We used Gradient-weighted Class Activation Mapping (Grad-CAM) [15] to visualize which regions of the spectrogram were important to the CNN and if the CNN shifted its attention to different parts of the spectrogram given different target labels. Note, that the CNN has never encountered in the dataset mixtures of different species, as in each insectary cage there was only one species of both sexes. Still, it clearly switched its attention conditioned on the class label it was given. Grad-CAM receives as input only the spectrogram of the mixture and given that we look for an *Ae. aegypti* it produces the Fig. 7-left, whereas, given the mixture and the fact that we look for an *An. Arabiensis* it produces the figure Fig.7-right. Note the clear shift of attention

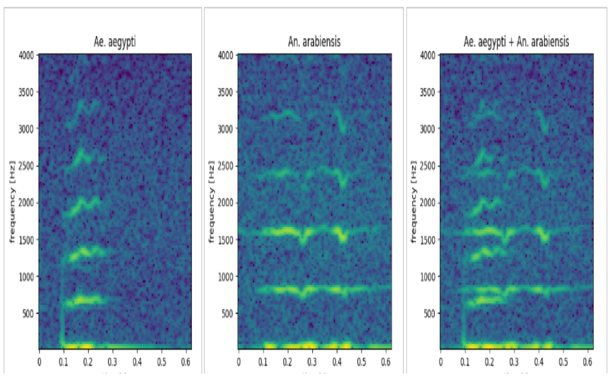

Fig. 6. Spectrograms. (LEFT) An *Ae. Aegypti* recording, (MIDDLE) An *An. Arabiensis* recording, (RIGHT) Adding the samples of both recordings.

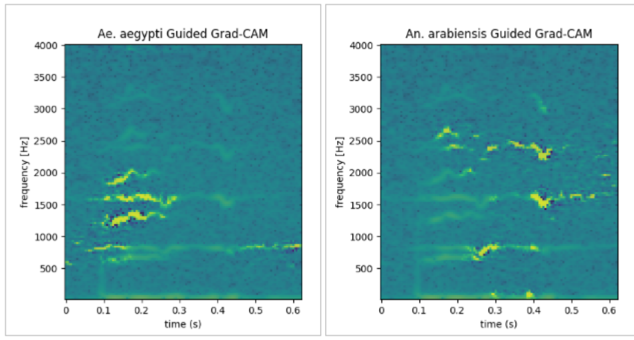

Fig. 7. Grad-weighted Class Activation Mapping for a mixture of recordings from two species. When we change the given label, the CNN changes focus.

that can be verified by comparing the attention map to the spectrogram of the original source.

C. *Transfer Learning*

Transfer learning or model adaptation in DL means that we train a model using one database and we subsequently use this model as the starting point to build a classification model on a different task (i.e. using a different database). The pretrained model of the first task is typically derived after training on a large database (e.g. ImageNet). The model in the subsequent task could be something more specific, not vastly represented in the large database (e.g. vision-based recognition of whale species). Transfer learning is widely applied in deep learning to refine models in several applications. It works because deep learning architectures have a modular layer composition where the layers close to the input learn to extract low-level features and subsequent layers rely on the previous layer(s) to synthesize patterns of higher abstraction (e.g. starting from edges and textures and ending in objects). The last layers of a DL net decide on the classes to be classified. In transfer learning we generally keep the lower levels of the model trained with the large dataset and discard the higher levels that need to learn to classify the classes of the second dataset.

'Wingbeats' is a large database derived from mosquitoes performing free flight in cages. We want to use it for transfer learning on different sensors and species that are not encountered in this specific database. There are several important consequences of this experiment:

a) Obviously, in a specific geographical region only few species of wingbeating insects co-exist. We would like to have a practical approach that would allow us to build region-specific classifiers applicable to any place in the world.

b) Labelled wingbeat recordings coming from traps installed in the field are hard to get. One can acquire them by releasing insects in a room/tent with a suction trap that has the appropriate optical sensor inside and get a smaller, local database.

c) Commercial mosquito traps include a suction mechanism. This means that mosquitoes sucked in perform forced flight in contrast to the Wingbeats database that is composed totally from free flight events.

d) The core idea is to use Wingbeats as a general-purpose database for flying insects and adapt the models using a much

ISBN 978-90-827970-1-5 © EURASIP 2018

lower number of forced flight wingbeat recordings taken from the specific geographic region of interest. This has been confirmed on a dataset not disclosed in Wingbeats.

e) Once the models are adapted to the local species the weights of the CNN are downloaded in a raspberry pi3 performing prediction is then embedded in commercial traps. The traps are then transferred to the open field where they can perform recognition in situ or transmit the wingbeat snippets to a server where they are logged and classified.

We use a Densenet121 to be trained from scratch using an open database of 10 insect species from UCR that is described in Flying Insect Classification Using Inexpensive Sensors website. We demonstrate that there is a small but consistent advantage if we pretrain on Wingbeats and refine on the UCR database over training from scratch using the UCR database (84.65% vs 84.12%).

V. Discussion

Insect Biometrics, in the context of our work, is a behavioral characteristic of flying insects measured by light intensity fluctuations. Biometric identifiers are related to the shape of the body (main body size, wing shape, wingbeat frequency, pattern movement of the wings). We demonstrated that the recordings of optical sensors provide the quality of information to discriminate mosquito species. DL techniques are compatible with optical recordings and demonstrate high classification rates.

Acknowledgments

The Dataset and associated code for all results presented in this work can be found in KAGGLE/Wingbeats database. Deep architectures were developed using the Keras library and all other classifiers were developed with scikit-learn [19] and Anaconda Python 2.7. Funded by GSRT-Greece matching funds for Entomatic and Remosis projects.

References

[1] Caraballo, H. (2014). Emergency Department Management of Mosquito-Borne Illness: Malaria, Dengue, and West Nile Virus. Emergency Medicine Practice 16 (5).

[2] Koella, J., Sorensen A. (1998). The malaria parasite, Plasmodium falciparum, increases the frequency of multiple feeding of its mosquito vector, Anopheles gambiae, Proceedings of the Royal Society B 265 (1398): 763–768.

[3] L'Ambert, G., Ferré, J.-B., Schaffner, F. and Fontenille, D. (2012). Comparison of different trapping methods for surveillance of mosquito vectors of West Nile virus in Rhône Delta, France. Journal of Vector Ecology, 37: 269–275. doi: 10.1111/ j.1948-7134.2012.00227.x

[4] Lühken, R., Pfitzner, W. P., Börstler, J., Garms, R., Huber, K., Schork, N., Krüger, A. (2014). Field evaluation of four widely used mosquito traps in Central Europe. Parasites & Vectors, 7, 268. doi:10.1186/1756-3305-7-268

[5] Hapairai L., et al., (2015). Evaluation of traps and lures for mosquito vectors and xenomonitoring of Wuchereria bancrofti infection in a high prevalence Samoan Village, Parasit Vectors. 2015; 8: 287.

[6] Kampen H, Medlock JM, Vaux AG , Koenraadt CJ , van Vliet AJ , Bartumeus F, Oltra A, Sousa CA, Chouin S, Werner D. (2015). Approaches to passive mosquito surveillance in the EU. Parasit Vectors,8:9. doi: 10.1186/s13071-014-0604-5.

[7] Burkina Faso Marco Pombi, Wamdaogo M Guelbeogo, Maria Calzetta , N'Fale Sagnon, Vincenzo Petrarca , Vincenzo La Gioia and Alessandra della Torre (2015). Evaluation of a protocol for remote identification of mosquito vector species reveals BG-Sentinel trap as an efficient tool for Anopheles gambiae outdoor collection, Malaria Journal 14:161 DOI 10.1186/s12936-015-0674-7

[8] Mayagaya V., Michel K, Benedict M., Killeen G., Wirtz R., Ferguson H., Dowell F, (2009). Non-destructive determination of age and species of Anopheles gambiae s.l. using near-infrared spectroscopy. American Journal of Tropical Medicine and Hygiene. 81(4):622-30. doi: 10.4269/ajtmh.2009.09-0192.

[9] Zou L, Miller S., and Schmidtmann E. (2006). Mosquito Larval Habitat Mapping Using Remote Sensing and GIS: Implications of Coalbed Methane Development and West Nile Virus J. Med. Entomol. 43(5): 1034D1041.

[10] Pruszynski C, Geier M, and Weber M. (2016) Field evaluation of the BG-Counter, a new surveillance tool to remotely measure mosquito densities. AMCA 2016 abstract book, 12 oral presentation.

[11] Geier M, Weber M, Rose A, Obermayr U, Pruszynski C, and Doyle M. (2016). A smart Internet of Things (loT) device for monitoring mosquito trap counts in the field while drinking coffee at your desk. AMCA 2016 abstract book, P-32 poster presentation.

[12] Potamitis I., Rigakis I., (2015). Novel Noise-Robust Optoacoustic Sensors to Identify Insects Through Wingbeats. IEEE Sensors Journal, vol.15, no.8, pp. 4621-4631, doi: 10.1109/JSEN.2015.2424924

[13] Potamitis, I. Rigakis, (2016). Large Aperture Optoelectronic Devices to Record and Time-stamp Insects' Wingbeats. IEEE Sensors J., 2016.

[14] Huang, Gao and Liu, Zhuang and van der Maaten, Laurens and Weinberger, Kilian Q, Densely connected convolutional networks},Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017.

[15] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh and D. Batra, "Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization," 2017 IEEE International Conference on Computer Vision (ICCV), Venice, 2017, pp. 618-626. doi: 10.1109/ICCV.2017.74

[16] Aldersley A, Champneys A, Homer M, Robert D. Quantitative analysis of harmonic convergence in mosquito auditory interactions. Journal of the Royal Society Interface. 2016;13(117):20151007. doi:10.1098/rsif.2015.1007.

[17] Potamitis I. (2014). Classifying insects on the fly, Ecological Informatics, vol. 21, May, pp. 40-49, ISSN 1574-9541

[18] L.J.P. van der Maaten. Accelerating t-SNE using Tree-Based Algorithms. Journal of Machine Learning Research 15(Oct):3221-3245, 2014.

[19] Pedregosa et al., (2011). Scikit-learn: Machine Learning in Python, Journal of Machine Learning Research 12: 2825-2830.