Pollinator Recognition Using Transfer Learning

Project Proposal

Unit Coordinator: Dr Greg Falzon

Unit Supervisor: Dr Romina Rader

Co-Supervisor: Dr Tobias Smith

Unit: COSC592 Masters Information Technology Project

Student Name: Sharifah Aldossary

Introduction

One of the main ecosystem services is pollination by wild animals. Hundreds of thousands of animal species play a significant role as pollinators. The majority are insects such as bees, beetles, butterflies, wasps, moths, and ants. Identifying and recognising insect pollinators is a major component in crop pollination research, as well as in the study on natural ecosystems and their processes (Klein, et al., 2007; Wena & Guyer, 2012). In addition to the interest from pollinator researchers, there is great enthusiams for pollinators and their protection in the wider non-scientific community, and increseigly 'citizen scientists' are engaging with pollinator monitoring projects (Smith & Saunders, 2016). For the identification of even course groups of insect pollinators, a basic understanding of insect taxonomy is required by the observer, and as such this is often a limiting factor in the ability of citizen scientists to contribute to insect pollinator research. Therefore, automating the pollinator identification process would have significant benefits in both educating and informing potential citizen scientists and in generating reliable citizen science datasets (Wena & Guyer, 2012).

This research aims to automate the identification of insect pollinators on flowers from photos, in Australia. In this project, we propose to recognise and identify pollinators automatically based on a transfer learning technique. A major and fundamental component in the field of pollination research, both professional and citizen science-based, is the identification of insect pollinators from on-flower observations or photos. Having functional automated pollinator identification systems that are available to researchers and citizen scientists in the field or in the lab could significantly increase the speed at which projects can be undertaken and improve the reliability of the data generated.

Automating identification and recognition from images and videos involves identifying data (i.e. objects in the images and videos), collecting the relevant data, and registering the collected data directly into computer systems (Agarwal, n.d.). Many studies have considered using machine learning, deep learning, and transfer learning techniques to automate recognition and identification processes (Akçay, et al., 2016; Ding & Taylor, 2016; Kandaswamy, et al., 2016; Mortensen, et al., 2007; Wen & Guyer, 2012; Wang, Lin, Ji, & Liang, 2012; Weeks, O'Neill, Gaston, & Gauld, 2003; Larios, et al., 2008; Zhang, et al., 2006; Glick & Miller, 2016; Mohanty, et al., 2016). Thus, it is important to understand the differences between machine learning, deep learning, and transfer learning, as well as the reason why we decided to use the transfer learning in our approach.

Briefly, in machine learning, the desired features, such as corners or edges of the object (a dog, a car, etc.) are manually extracted using functions, to train the machine learning model. After that, the trained model references the extracted features when analysing and classifying new objects. Deep learning is a branch of machine learning. However, deep learning learns features and tasks directly from data, which is why it is called end-to-end learning. In other words, in deep learning the manual feature extraction step is skipped, and the data are directly fed to the deep learning algorithm which then predicts the object as shown in Figure 1. Most deep learning models use neural network architecture as shown in Figure 2. (Patel, & Pingel, n.d.).

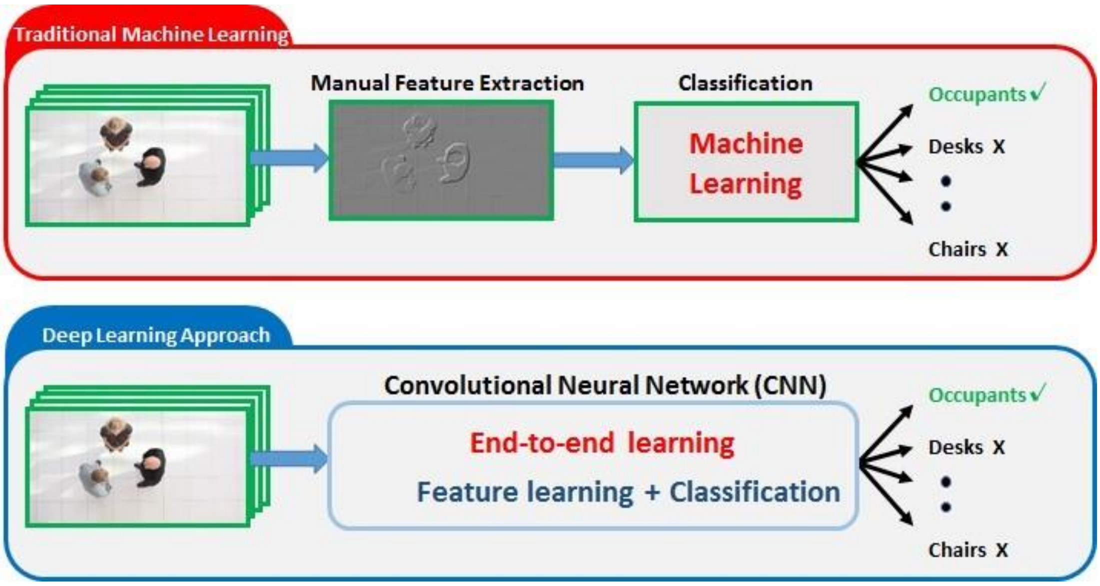


Figure 1. Machine learning and deep learning approaches. (reprinted from PointGrab, 2016).

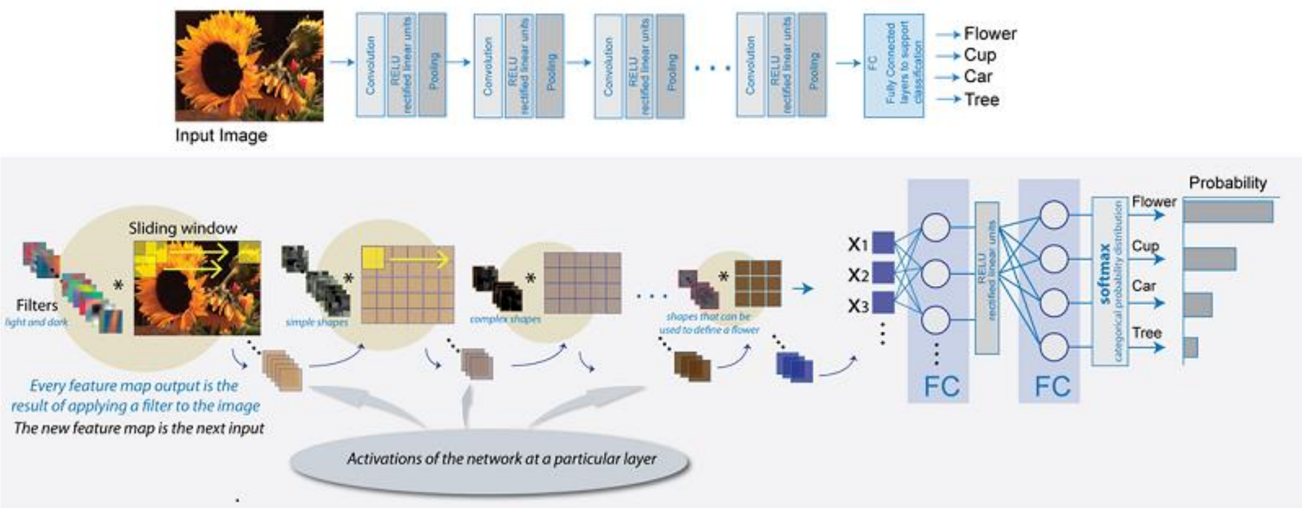


Figure 2. Deep learning with MATLAB. Reprinted from MathWorks, Deep learning with MATLAB, https://au.mathworks.com/discovery/deep-learning.html.

The most commonly used type of deep learning, which is also called ‘deep neural network’, is the convolutional neural network (CNN). A CNN contains several layers of convolutional, nonlinear, pooling, and fully connected layers, that process and transforms an input data to produce an output. (Deshpande, 2016).

CNNs are based on three important concepts: first, respective local fields, second, shared weights and biases, and lastly activation and pooling. In CNNs, small regions of the input layer neurones connect to neurones in the hidden layer, unlike in a typical neural network where each neurone in the input layer is connected to a neurone in the hidden layer. These regions are called local receptive fields and are used map feature from the input layer to the hidden layers. In any particular layer, the weights and biases for the hidden neurones are all the same in, or in other words, the hidden neurones for a particular layer detect the same features in different regions of the data. Activation function implements a transformation to the output of each neurone, a rectified linear unit (ReLU) is one example of an activation function. Pooling is also used to implement additional output transformation, as it helps to decrease the number of parameters that the model needs to learn and thus helps to simplify the following layers (Mathworks, Patel, & Pingel, n.d.).

When the CNN model is trained from scratch, it is called deep learning. In contrast, when a knowledge of one type of problem, gained through deep learning, is used to solve a similar problem, it is called transfer learning. For example a CNN model that has been trained to recognise animals in images can be used as a base to initialize and train a new model to classify transportation vehicles. In addition to transfer learning methodologies, features can be extracted from pre-trained CNNs to train machine learning models (Patel, & Pingel, n.d.).

In this project, we will implement and develop our pollinators' recognition (PR) model to answer the first question using the transfer learning method to classify six different insects pollinators namely; bees, beetles, butterflies, flies, moths, and wasps. Transfer learning models do not require a significant number of data nor a long time to train, and this is the reason why we are using this technique to train our PR model. Firstly, we will recognise and identify insect pollinators from images to coarse-grained taxonomic classifications (e.g. bees, flies, beetles). We will then evaluate the model and benchmark the performance of the algorithm by taking into account such factors as accuracy, sensitivity, specificity, the receiver operating characteristic (ROC) curve, the area under the curve (AUC)[1], and the confusion matrix (Sayad , Model Evaluation, 2010). We will determine the diagnostic accuracy of the model by checking which images were misclassified.(Falzon, pers. Comm, 2017)

[1] According to (Sayad, n.d.), the “Area Under the Curve: used as a measure of quality of the classification model.”

In a more advanced stage of the project, after the algorithm has been evaluated, we plan to use the PR model in the cloud. All major cloud platform such as Microsoft Azure Google Cloud Platform, Amazon web services, and IBM cloud provide GPU cloud computing(Bri, n.d.)[2]. Deep learning in the cloud can be done using MATLAB through the new Amazon EC2 P2 instances (Moulder, Sheridan, Cavallo, & Rossini, 2016).

Background

Object classification is classifying input data into objects or classes based on key features (interesting parts of an image) (Mathworks, n.d.). The purpose and necessity of object classification and recognition vary from one approach to another. In recent years, object recognition has driven computer vision and image processing research (Oquab, Bottou, Laptev, & Sivic, 2014), which have come to recognise images by their extracted features. There are two methods to represent images one uses global features, such as colour, texture, or shape, whereas the other method uses local feature descriptors extracted from interest regions, which are specific regions within the image. First, we apply image feature detector techniques such as Features from Accelerated Segment Test (FAST) which is a corner detector, or the Hessian blob detector which is also called the Hessian matrix describe the local structure in a neighbourhood around a point, and many other detectors. We refer readers to Hassaballah, Abdelmgeid, and A. Alshazly (2016) After detecting and extracting the relevant data from an image, it needs to be encoded in a fitting descriptor using one or a combination of descriptors, such as scale invariant feature transform(SIFT) or the speeded-up robust features descriptor (SURF), and many others. For a more complete description, we refer readers to Hassaballah, Abdelmgeid, and A. Alshazly (2016).

Accurate automated identification of insects has previously been used to help manage the agricultural pests (Weeks, O’Neill, Gaston, & Gauld, 2003). The automated insect identification can provide owners of vast orchards who have limited pest scouting expertise with automated counting enabling them to control harmful pests (Wen & Guyer, 2012. Wen and Guyer (2012) used the Hessian-Affin detector and SIFT to extract the local features, in a global study that used geometrics, contours, textures, colour features and invariant moments (Wen & Guyer, 2012). Zhang, Deng, Dietterich, and Mortensen (2006) proposed nother generic object description system that characterises object class features on both the global and local levrls based on multi-scale affine-invariant image regions.

---

[2] While these are not the only platforms, these names were mentioned on the NVIDIA website along with ten others.

Identifying insects using artificial recognition systems can be a challenging task because of the enormous number of species. Wang and Liang (2012) presented a new system for identifying images of insects to the order level. They used three recognition methods: pattern recognition, artificial neural networks (ANNs) and a support vector machine (SVM) to design several relative features. They found that some feature extraction techniques are suitable for identifying certain insects, but that they cannot be applied to others. Zhang and Dietterich (2008) used an efficient discriminative approach, called iterative discriminative clustering (IDC), to construct a visual dictionary. The researchers reduced the problem of object recognition to the problem of classifying a bag of descriptor vectors into one of the potential object classes.

Several other methods have been proposed. For example, Prince (1996) considered tree-based image classification, and other researchers have used SIFT to classify images (Mortensen, et al., 2007; Wen & Guyer, 2012; Wang, Lin, Ji, & Liang, 2012; Larios, et al. 2008). Larios, et al.(2008) also used in their study a concatenated histogram to develop a fully automated stonefly-larvae classification system using the computer vision and machine learning approach. Deep learning, which develops and trains the neural network from scratch using different processes to extract various features using supervised techniques, has been used in some studies (Mohanty, Hughes, & Salathé, 2016; Ding & Taylor, 2016; Mortensen, et al., 2007; Pingel, 2017).

The use of transfer learning has also grown over the past few years. One particular study implemented a transfer learning model to analyse breast cancer (Kandaswamy, Silva, & Alexandr, 2016), while another used transfer learning to classify objects in baggage security X-ray scans for security purposes (Akçay, Kundegorski, Devereux, & Breckon, 2016). Although some of these studies focused on automating the counting of insects (Mortensen et al., 2007), this study pointed out that providing ongoing high-resolution data is costly because existing methods require human experts to manually perform the collection and identification. The same approach was developed for mechanical devices used for automatically photographing insects and used with algorithms; these methods were applied to environmental and agricultural problems (Larios et al., 2008).

Proposed Methods and Required Resources

Recently, convolutional neural networks (CNNs), have shown a remarkable performance in image classification in a large-scale visual recognition challenge[3] (Oquab, Bottou, Laptev, & Sivic, 2014).

---

[3] This study used a model on mid-level image representation. Mid-level image processing is mainly concerned with extracting descriptions of a scene from the image descriptions extracted at the low

The present project focuses on developing an algorithm using transfer learning to identify pollinator groups from images. This transfer learning is based on the pre-trained AlexNet CNN, which was developed by Krizhevsky, Sutskever, and Hinton (2012). The AlexNet CNN uses deep learning from a diverse range of subject groups, from aeroplanes and cars to beavers and trees. A number of key differences exist among machine learning, deep learning and transfer learning. To train the machine learning model, we can manually select the relevant features (edges–corners). The model then references these features when analysing and classifying new objects as shown in (Figure 1). However, in deep learning, the step of manually extracting the features is not needed. Instead, we feed the images into the deep learning algorithm, which then predicts the object as shown (Figures 1 and 2). Likewise in transfer learning, however, we use an existed pre-trained CNN model to solve a problem similar to the problem that had been trained for, as been mentioned earlier in the paper.

The machine learning technique is applied if we do not have access to a high-performance GPU or a large set of data. In machine learning, there is a large set of classifiers to choose from and features, such as HOG, FAST and Harris, to extract image descriptors to produce the best results. Also, machine learning has the flexibility to combine more than one approach (Pingel, 2017). However, deep learning uses neural networks that combine multiple nonlinear processing layers. Deep learning is used because of its high accuracy, and it is especially suited for image recognition (Moulder, Sheridan, Cavallo, & Rossini, 2016).

A major difficulty in training a neural network is that it is time-consuming, often involving several days of computing time, and requires a large set of data (millions of images) (Moulder, Sheridan, Cavallo, & Rossini, 2016). We can use transfer learning to overcome these limitations. Transfer learning requires a shorter training period, as it needs only hundreds of images, rather than millions, to train (Figure 3) (Pingel, 2017).

---

level (Price, 1996). Low-level image processing is mainly concerned with extracting descriptions from images (Price, 1996).

Training data	1000s to millions of labeled images
Computation	Compute intensive
Training Time	Days to Weeks for real problems
Model accuracy	High (can over fit to small datasets)

Training data	100s to 1000s of labeled images (small)
Computation	Moderate computation
Training Time	Seconds to minutes
Model accuracy	Good, depends on the pre-trained CNN model

Figure 3: A comparison of deep learning and transfer learning. Reprinted from MathWorks, Deep learning with MATLAB, https://au.mathworks.com/discovery/deep-learning.html.

In this project, we will use the pre-trained AlexNet CNN and fine-tuned the last few layers to learn the particular features of the new dataset. We will fine-tune six course-grained categories: bee class, beetle class, butterfly class, fly class, moth class and wasp class.

Three sourcing methods will be used to collect the colour image dataset. The first sourcing method involves using images provided by experts.[4] The second sourcing method involves using images available under a creative commons licence.[5] The dataset will then be labelled and refined by an expert[6] ; to constrain the approach to a specific situation. After this step, the image sizes will be adjusted to match AlexNet's input layer size. Note that it is preferable to standardise the resolution, the distance between the source and the subject, the camera orientation and the background as much as possible for later development.

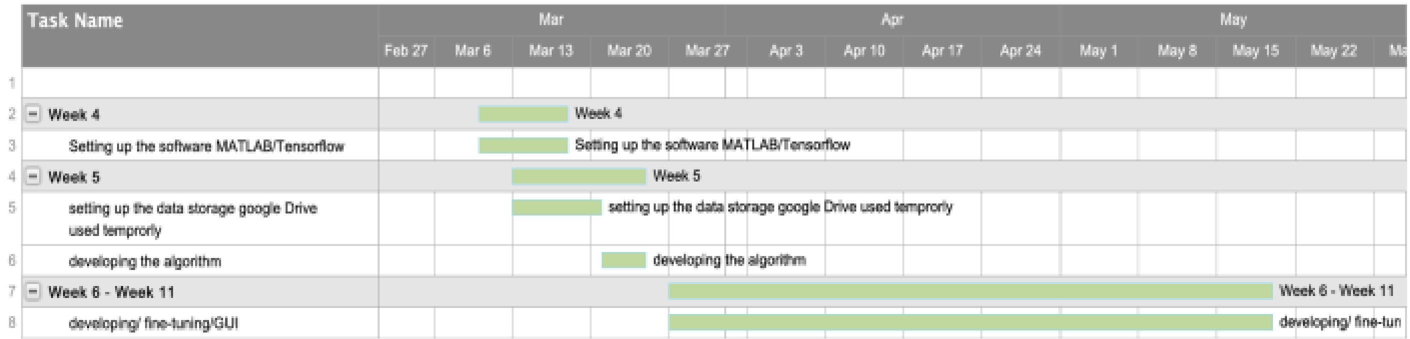
We will use MATLAB software to develop the algorithm. The following set of toolboxes are required to achive this: Parallel Computing Toolbox, Computer Vision System Toolbox, Image Processing Toolbox, Neural Network Toolbox and the Statistics and Machine Learning Toolbox (Mathworks, 2017).This research requires a reasonable large data storage capacity, preferably one terabyte.

[4] Dr Romina Rader and Dr Tobias Smith from the School of Environmental and Rural Science at the University of New England, and other researchers.
[5] Online resources include Flickr, Twitter, Google and Facebook.
[6] Dr Tobias Smith

It is also important to have the proper hardware to develop an algorithm that uses a neural network and a Computer Vision Toolbox. Both require a high-performance GPU and CUDA-enabled NVIDIA with a compute capability of 3.0 or higher. With powerful hardware, the model spends less time analysing the images (Mathworks, 2017).

Project Timeline

The following Gantt chart shows the project timeline.



Bibliography

Agarwal, T. (n.d.). *Automatic Identification and Data Capturing (AIDC) Technology*. Retrieved from ELECTRONICS PROJECTS FOCUS: https://www.elprocus.com/

Akçay, S., Kundegorski, M., Devereux, M., & Breckon, T. (2016). TRANSFER LEARNING USING CONVOLUTIONAL NEURAL NETWORKS FOR OBJECT CLASSIFICATION WITHIN X-RAY BAGGAGE SECURITY IMAGERY. *Image Processing (ICIP), 2016 IEEE International Conference on.* IEEE.

Bring the Power of Deep Learning to Your Data. (n.d.). Retrieved from http://www.nvidia.com/object/gpu-cloud-computing.html

Deshpande, A. (2016). *A Beginner's Guide To Understanding Convolutional Neural Networks*. Retrieved from https://adeshpande3.github.io/adeshpande3.github.io/

Ding, W., & Taylor, G. (2016, Feb 24). Automatic moth detection from trap images for pest management. 1-17.

Glick, J., & Miller, K. (2016, March 13). Insect Classification With Heirarchical Deep Convolutional Neural Networks.

Glover, B. (2011, May 10). Pollinator Attraction: The Importance of Looking Good and Smelling Nice. *21*(9), R307–R309.

Hassaballah, M., Abdelmgeid, A., & A. Alshazly, H. (2016). *Image Features Detection, Description and Matching* (1 ed., Vol. 630). (A. I. Awad, Ed.) Springer International Publishing.

Kandaswamy, C., Silva, L., & Alexandr, L. (2016, January 8). High-Content Analysis of Breast Cancer Using Single-Cell Deep Transfer Learning. *21*(3), 252-259.

Klein, A.-M., Vaissière, B., Cane, J., Steffan-Dewenter, I., Cunningham, S., Kremen, C., & Tscharntke, T. (2007, Feb 7). Importance of pollinators in changing landscapes for world crops. *Proceedings Of The Royal Society B, 274*(1608).

Krizhevsky, A., Sutskever, I., & Hinton, G. (2012). ImageNet Classification with Deep Convolutional Neural Networks. Retrieved from http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks

Larios, N., Deng, H., Zhang, W., Sarpola, M., Yuen, J., Paasch, R., . . . Dietterich, T. (2008). Automated insect identification through concatenated histograms of local appearance features: feature vector generation and region detection for deformable objects. *Machine Vision and Applications*, *19*(2), 105–123.

Mathworks. (2017). *Image Category Classification Using Deep Learning*. Retrieved from https://au.mathworks.com/help/vision/examples/image-category-classification-using-deep-learning.html

Mathworks. (2017). *Transfer Learning Using Convolutional Neural Networks*. Retrieved from https://au.mathworks.com/help/nnet/examples/transfer-learning-using-convolutional-neural-networks.html

Mathworks. (n.d.). *Deep learning with MATLAB*. Retrieved from https://au.mathworks.com/discovery/deep-learning.html

Mathworks. (n.d.). *Object recognition methods in computer vision*. Retrieved from Mathworks: https://au.mathworks.com/

Mohanty, S., Hughes, D., & Salathé, M. (2016, April 15). Using Deep Learning for Image-Based Plant Disease Detection. 1-7.

Mohanty, S., Hughes, D., & Salathé, M. (2016). Using Deep Learning for Image-Based Plant Disease Detection.

Mortensen, E., Delgado, E., Deng, H., Lytle, D., Moldenke, A., Paasch, R., . . . Dietterich, T. (2007). *Pattern Recognition for Ecological Science and Environmental Monitoring: An Initial Report.*

Moulder, S., Sheridan, T., Cavallo, P., & Rossini, G. (2016). *Deep Learning with MATLAB and Multiple GPUs.* Retrieved from MathWorks: https://www.mathworks.com/content/dam/mathworks/tag-team/Objects/d/Deep_Learning_in_Cloud_Whitepaper.pdf

NAPPC. (n.d.). *Primer on Pollination and Pollinators*. Retrieved from Pollinator Partnership: http://www.pollinator.org/Resources/facts.Primer.pdf

Oquab, M., Bottou, L., Laptev, I., & Sivic, J. (2014). *Learning and Transferring Mid-Level Image Representations using Convolutional Neural Networks.* IEEE Conference on Computer Vision and Pattern Recognition.

Pingel, J. (2017, March 24). *Introduction to Deep Learning: Machine Learning vs. Deep Learning*. Retrieved from MathWorks: https://au.mathworks.com/videos/introduction-to-deep-learning-machine-learning-vs-deep-learning-1489503513018.html

Patel, S., & Pingel, J. (n.d.). *Deep learning with MATLAB*. Retrieved from mathworks: https://au.mathworks.com/

PointGrab. (2016, August 30). *Traditional Machine Learning vs. Deep Learning*. Retrieved from https://www.wirelessdesignmag.com/article/2016/08/applying-deep-learning-edge-analytics-sensors-building-automation

Price, S. (1996, July 4). *Low Level Image Processing*. Retrieved from http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/MARBLE/low/low.htm

Price, S. (1996, April 22). *Medium Level Image Processing*. Retrieved from http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/MARBLE/medium/medium.htm

Sayad , S. (2010). *Model Evaluation.* Retrieved from An Introduction to Data Mining: http://www.saedsayad.com/data_mining_map.htm

Sayad, S. (n.d.). *Model Evaluation - Classification*. Retrieved from http://www.saedsayad.com/

Smith, T. J., & Saunders, M. E. (2016, June 14). Honey bees: the queens of mass media, despiteminority rule among insect pollinators. *Insect Conservation and Diversity, 9*(5), 384-390.

Wang, J., Lin, C., Ji, L., & Liang, A. (2012). A new automatic identification system of insect images at the order level. *Knowledge-Based Systems*, 102–110.

Weeks, P., O'Neill, M., Gaston, K., & Gauld, I. (2003). Automating insect identification: exploring the limitations of a prototype system. *Journal of Applied Entomology, 123*(1), 1-8.

Wen, C., & Guyer, D. (2012). Image-based orchard insect automated identification and classification method. *Computers and Electronics in Agriculture, 89*, 110–115.

Zhang, W., & Dietterich, T. (2008). Learning visual dictionaries and decision lists for object recognition.

Zhang, W., Deng, H., Dietterich, T., & Mortensen, E. (2006). A Hierarchical Object Recognition System Based on Multi-scale Principal Curvature Regions. *4*.