Guide for contributors of species identification models to Horama ID

Alexander N. Schmidt-Lebuhn, alexander.s-l@csiro.au

1. Introduction

Horama ID makes it possible for taxonomists to deliver image classification models for species identification to end-users such as collection curators, other researchers, or the public. It consists of an app available on both Android and iOS and a server that stores the models and species profile data. The system was funded by CSIRO, Australia's national research agency, and coded by 2pi Software, Bega, Australia.

This document provides a step-by-step guide to training and contributing image classification models for Horama ID. Horama ID can currently only make use of Custom Vision models. We intend to expand to other formats, but some degree of standardisation will always be required so that the app knows what format(s) to expect.

Microsoft Custom Vision is a commercial service for training image classification and object detection models. As of writing, the free tier allows the creation of two projects with up to 50 classes and up to 5,000 images each, for one year. Beyond that, the user has to pay fees. However, the advantage of this service is that it is extremely easy to use, allowing a biologist without expertise in Python coding or AI to quickly train and test a model as described below.

Once trained, contribution of a model to Horama ID requires the preparation of a few accompanying files. These contain information that allows an end-user to understand the taxonomic and geographic scope of a model and what it was trained on, either a specific organ or an entire specimen. They also include species profiles that the contributing taxonomist can populate more or less as they would in traditional taxonomic publications or eFloras, so that the end-user can compare the model's identification result against an example photo, a description, geographic range, and other relevant data.

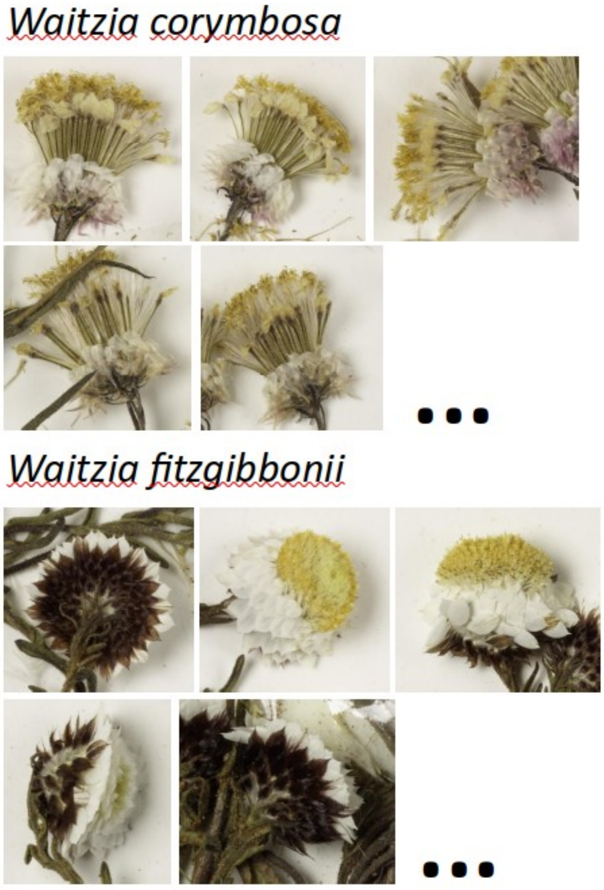
In this guide, the term 'class' will be used not in the sense of taxonomic hierarchy, but in the sense of deep learning / artificial intelligence research, as one of the answers a trained model produces when used for inference. In our case of species identification, these classes will be species or other taxa such as genera or subspecies. Another notable term in this context is 'tag', the name of a model class in Custom Vision.

2. Building an image library for training the model

The design of good image libraries for is critical for training models to high accuracy. There are various aspects to be kept in mind: number of images, maximising identification-relevant characters in the image, avoidance of biases, and how to augment data to make the most of the available images.

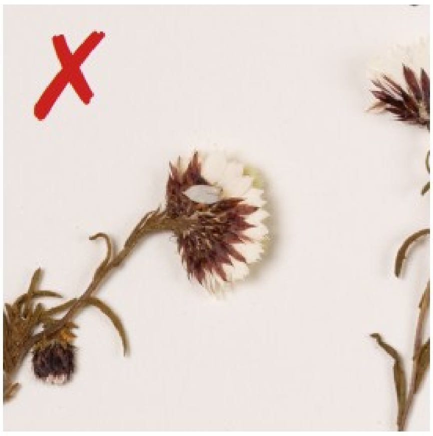
2.1. Minimum size of the library

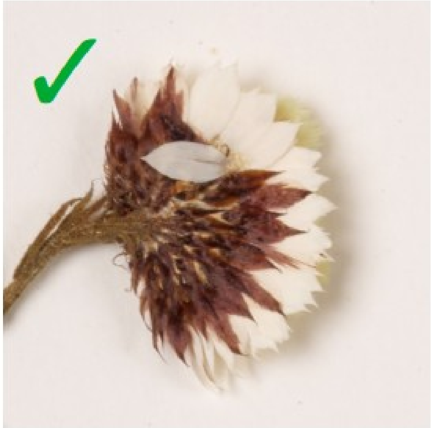
Although Custom Vision only warns if fewer than 50 training images are available for a class, it may be necessary to have higher numbers to achieve high confidence. The required number depends on the difficulty of the identification problem and the degree of variation within the same species. I find that in many cases, good results can be expected with at 200-300 images per class, but the more, the better.



2.2. The model only sees low resolution

The Custom Vision image classification model sees only 224 x 224 pixels, and comparable compact models for use in smartphones are generally similarly restricted in image resolution. This is because larger input image sizes would require larger, slower, and overall more unwieldy models. Because of this limited input image size, it is important to fill out the image with the character that we want the model to be trained on. White space on the image is wasted, and if the object of interest occupies only a small part in the centre of a training image, the model sees it only as a few blurry pixels and cannot discern the key identifying characters. The same applies when the trained model is used for inference: the user has to understand that they need to get up close to the specimen to get a good identification result and cannot expect the model to 'see' sufficient detail from half a meter away.

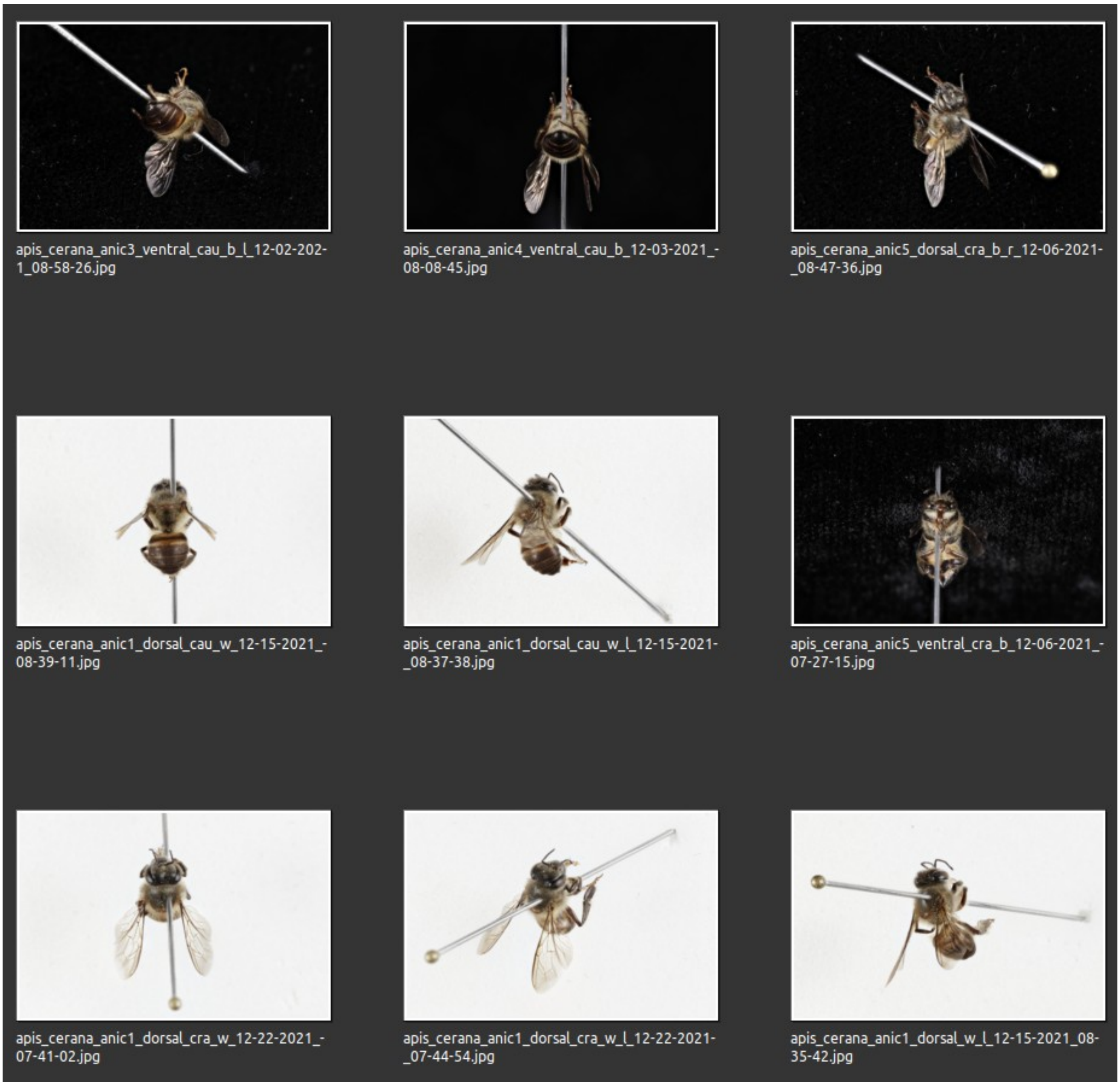




Another implication of the limited resolution of what the model sees is that there is no reason for us to upload unnecessarily large training images to Custom Vision. Storing images in original sizes of, say, anywhere between five and fifty megapixels in a project increases upload times and cloud compute fees for that project while not providing any advantage over images of only a few hundred pixels in edge length. It is therefore useful to bulk resize the training library before upload, e.g., with a Python script or the Bulk Image Manipulation Plugin (https://alessandrofrancesconi.it/projects/bimp/) for the open source software GIMP (https://www.gimp.org/).

2.3. Avoiding bias

AI models will learn to be biased if the training data are biased. For example, if most photos of one species were taken under poor light conditions, but those of other species were taken under good light conditions, the model may learn to recognise the species that way. It is therefore useful to vary the backgrounds and other environmental conditions under which the images of each species are taken, unless the intended use case involves a fairly predictable setting, e.g., parts of herbarium specimens, which will nearly always have a light-coloured, papery background. Below is an example of insects photographed before dark and light backgrounds.

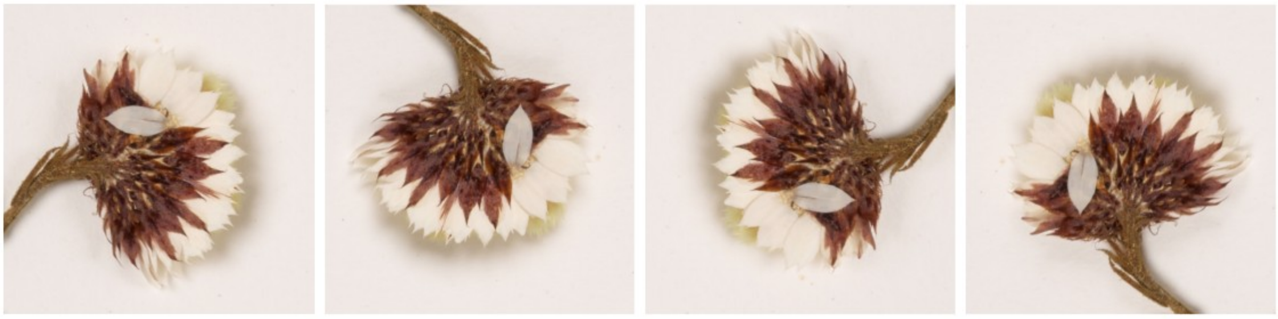

apis_cerana_anic3_ventral_cau_b_l_12-02-202-1_08-58-26.jpg
apis_cerana_anic4_ventral_cau_b_12-03-2021_-08-08-45.jpg
apis_cerana_anic5_dorsal_cra_b_r_12-06-2021-_08-47-36.jpg
apis_cerana_anic1_dorsal_cau_w_12-15-2021_-08-39-11.jpg
apis_cerana_anic1_dorsal_cau_w_l_12-15-2021-_08-37-38.jpg
apis_cerana_anic5_ventral_cra_b_12-06-2021_-07-27-15.jpg
apis_cerana_anic1_dorsal_cra_w_12-22-2021_-07-41-02.jpg
apis_cerana_anic1_dorsal_cra_w_l_12-22-2021-_07-44-54.jpg
apis_cerana_anic1_dorsal_w_l_12-15-2021_08-35-42.jpg

Model classes should have approximately the same number of images, because there is some risk of bias towards classes with many images if their sizes are very unbalanced. A good rule of thumb is that the largest class should not have more than twice the number of images than the smallest.

2.4. Data augmentation

It is unlikely that we have hundreds of specimens of each species we want to include in the model. However, the model does not necessarily have to see that many individual specimens to understand some aspects of variation that we want it to understand. We can take images of the same specimen from different angles, from different distances, or under different conditions (see previous section). A photo can also be processed in a variety of ways to create additional information, e.g., rotated, flipped, or cropped to 'zoom in' a bit more, for example using the GIMP software and its BIMP plugin mentioned previously. In the context of training AI models, these operations are called 'data augmentation'. A dataset augmented in this way teaches the model that the orientation of the specimen doesn't matter, in contrast to other scenarios like character recognition, where, for example, the difference between M and W is important.

There is only so much data augmentation will do for us, however, as becomes clear if we image using two hundreds variations of only a single image for training. Ultimately, covering variation across additional specimens of a species is much more valuable than rotating or flipping a few images over and over.



2.5. Negative class

Custom Vision allows the inclusion of a negative class in addition to the species we want to identify. The advantage is that this allows the model to return something other than one of the species names if it is shown the surface of a desk or some random object, and that improves the user experience. Depending on the use, this class can be a collection of potential backgrounds or random items that would be expected given a particular use case. For example, if our use case is fruits on a herbarium specimen, we can make a negative class of images showing blank cardboard, herbarium labels, leaves, and stem sections, whatever is not one of the fruits in question but would be found around them.

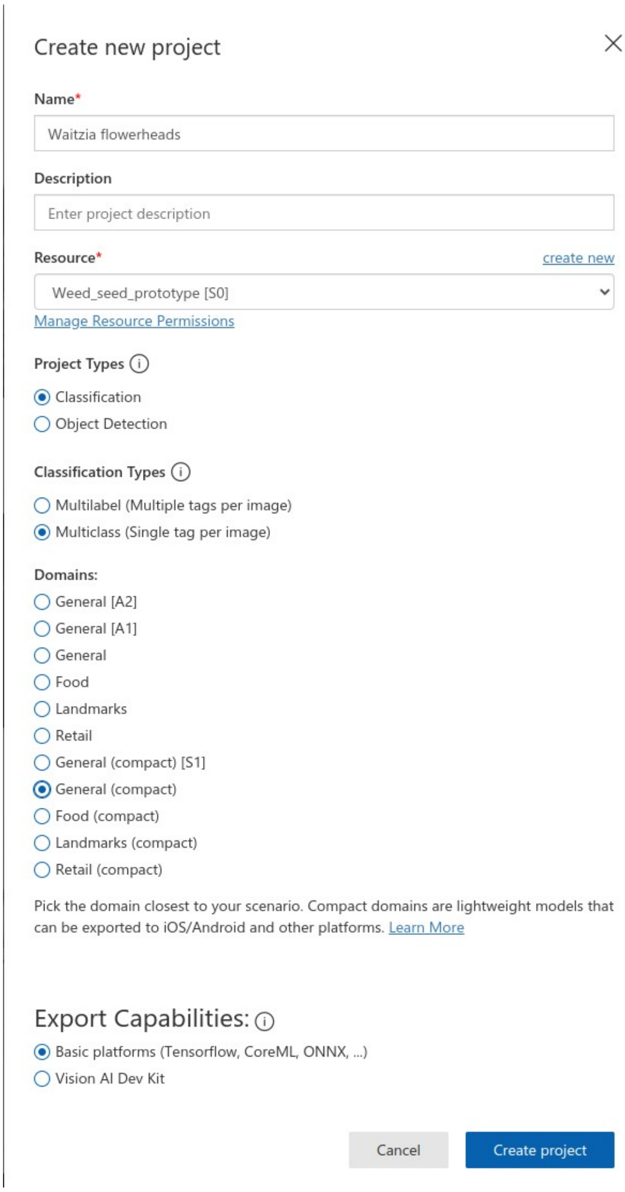
3. Custom Vision

When we have our image library ready, we can train a model in Microsoft Custom Vision.

3.1. Create project

Custom Vision projects intended to be contributed to Horama ID need to be set up with the following options:

- (Image) classification – the model classifies the entire image into a class as opposed to finding an object of interest on a larger image (object detection).
- Multiclass – as opposed to multilabel, as we will assign each image to only one species.
- Compact – this allows exporting of the model after it is trained, at a small cost in performance.



Create new project

Name*

Waitzia flowerheads

Description

Enter project description

Resource* create new

Weed_seed_prototype [S0]

Manage Resource Permissions

Project Types

● Classification
○ Object Detection

Classification Types

○ Multilabel (Multiple tags per image)
● Multiclass (Single tag per image)

Domains:

○ General [A2]
○ General [A1]
○ General
○ Food
○ Landmarks
○ Retail
○ General (compact) [S1]
● General (compact)
○ Food (compact)
○ Landmarks (compact)
○ Retail (compact)

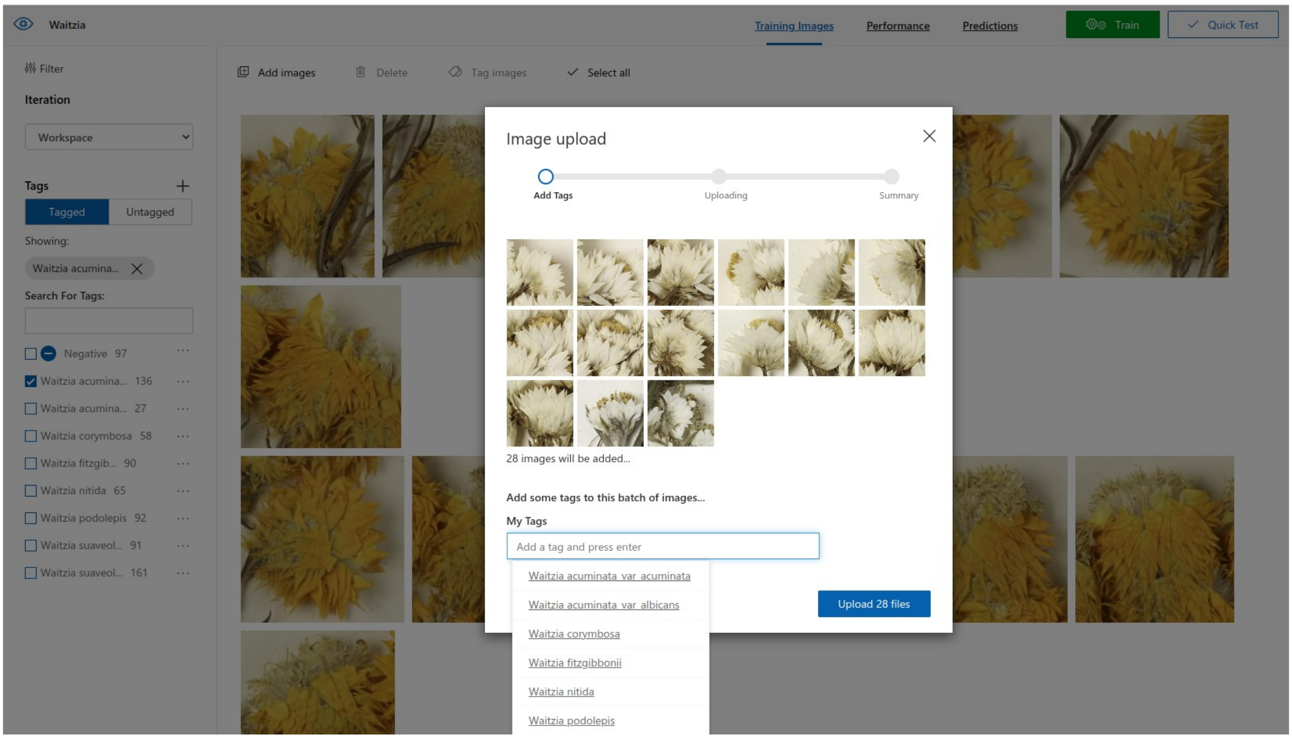
Pick the domain closest to your scenario. Compact domains are lightweight models that can be exported to iOS/Android and other platforms. Learn More

Export Capabilities:

● Basic platforms (Tensorflow, CoreML, ONNX, ...)
○ Vision AI Dev Kit

3.2. Upload images

If not already there, go to the Training Images tab. Click “Add images” in the upper left corner and navigate to a folder where you have stored training images. Select only images of one class and click “Open”. At the bottom of the Image upload window, either select an existing tag to assign the images to a class, or type out a new tag. Click Upload. After these images have been added, repeat the process with the next class until the entire training library is uploaded.



28 images will be added...

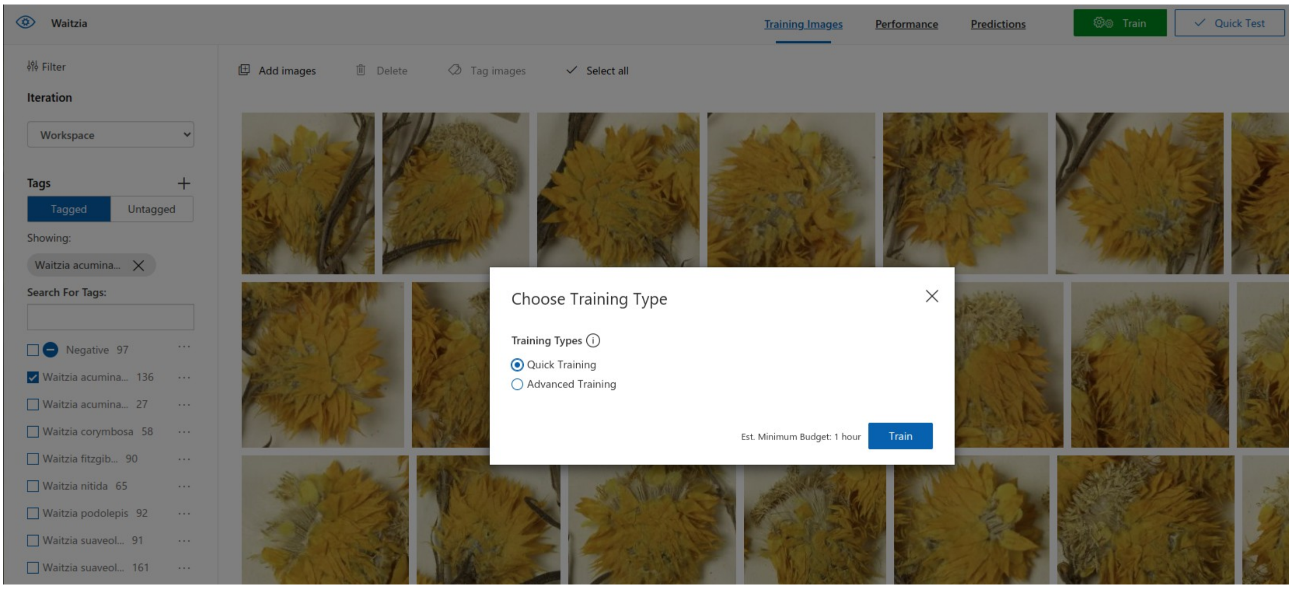
Add some tags to this batch of images...

Add a tag and press enter

Waitzia acuminata var acuminata
Waitzia acuminata var albicans
Waitzia corymbosa
Waitzia fitzgibbonii
Waitzia nitida
Waitzia podolepis

3.3. Train model

Click the green cogwheel and Train button. If uncertain, select Quick Training for a start. Advanced Training allows the specification of training time, but overshooting it risks over-fitting the model. This means it will get very good at identifying the training images but less capable of generalising beyond them.



Training Types

● Quick Training

○ Advanced Training

Est. Minimum Budget: 1 hour

3.4 Evaluate model performance

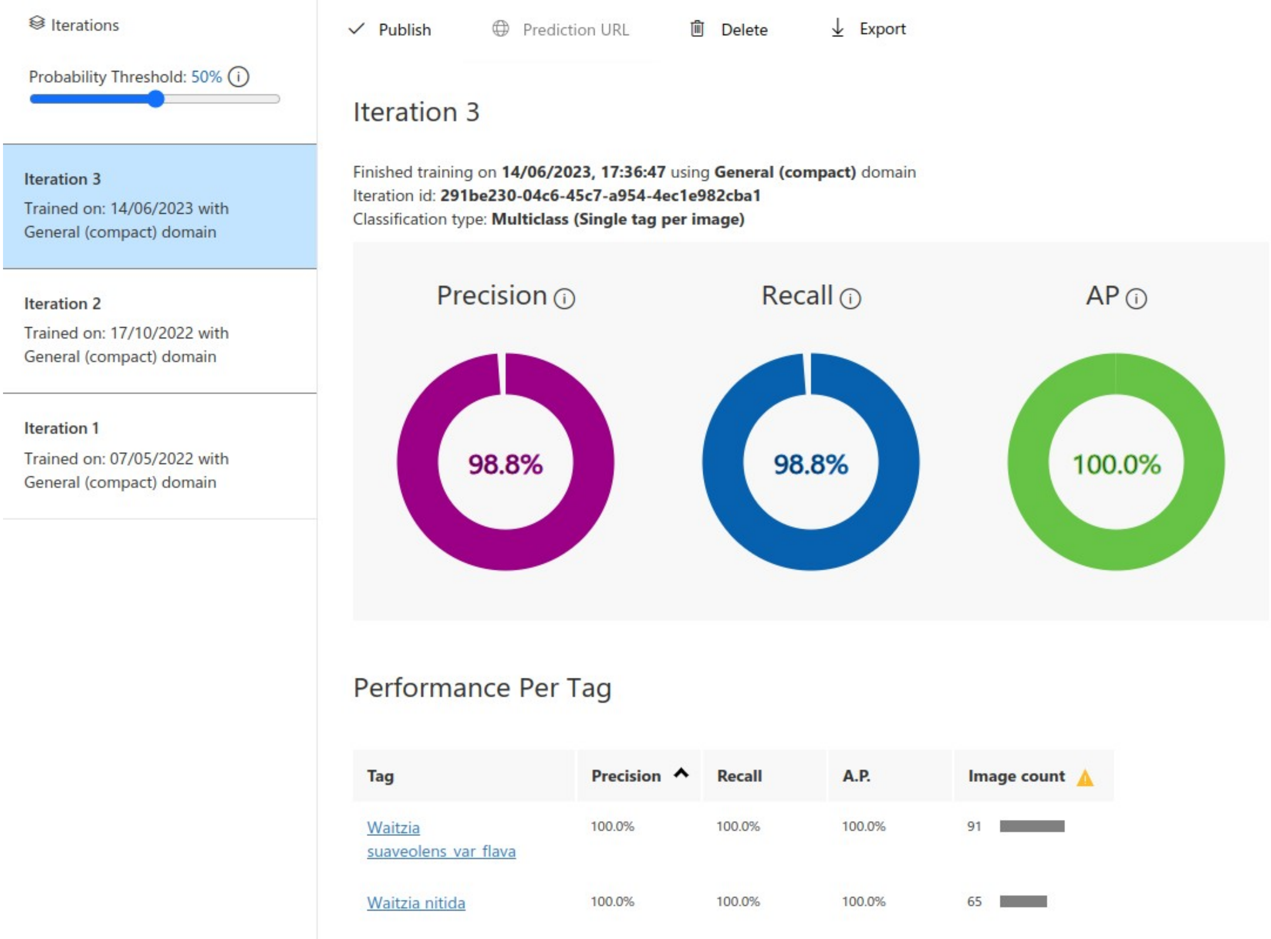
After a model iteration has been trained, examine it in the Performance tab. Custom Vision reports three values to quantify performance both for each individual class and overall:

Precision. A low value means we get many false positives, i.e., images get misidentified as the wrong species.

Recall. A low value means we get many false negatives, i.e., images do not get recognised as the species they belong to.

Although it may at first seem one is the flip-side of the other, the two values can be different, and that is because of the probability threshold, which is by default 50%. This means that images for which no class has a probability of at least that threshold do not count into false positives. Try changing the threshold to see how statistics change.

AP. A summary taking into account both precision and recall.



Probability Threshold: 50%

Iteration 3
Trained on: 14/06/2023 with General (compact) domain

Iteration 2
Trained on: 17/10/2022 with General (compact) domain

Iteration 1
Trained on: 07/05/2022 with General (compact) domain

Iteration 3

Finished training on **14/06/2023, 17:36:47** using **General (compact)** domain
Iteration id: **291be230-04c6-45c7-a954-4ec1e982cba1**
Classification type: **Multiclass (Single tag per image)**

Precision	Recall	AP
98.8%	98.8%	100.0%

Performance Per Tag

Tag	Precision	Recall	A.P.	Image count
Waitzia suaveolens var flava	100.0%	100.0%	100.0%	91
Waitzia nitida	100.0%	100.0%	100.0%	65

Clicking on a model tag below the performance statistics will show the test images of that class that were held back during training. The ones were testing failed to identify the image correctly are marked with a red box. Click on any image to see its test result including as what it was identified instead. This allows exploration of where the model struggles. Add or remove training images and try training again to improve results.

3.5. Test model on other images

Click on the Quick Test button. You can then either upload an image from your computer or enter the URL of an image on the internet. The model will be tested against this one image. This is not a very efficient way of testing it, but it can be of interest to try a few images that are very different from the ones you produced to train your model. In the example below, a model trained on herbarium specimens was tested on an image of a living specimen.

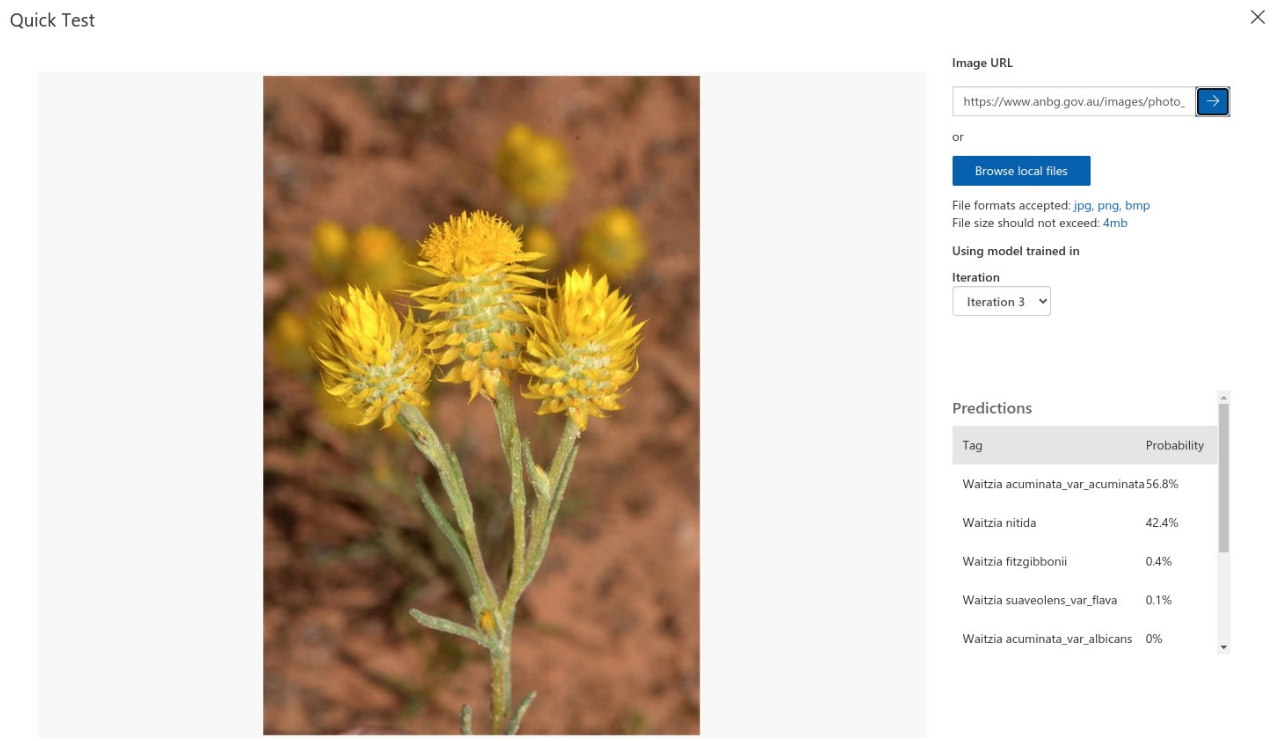


Image URL: https://www.anbg.gov.au/images/photo_

or

File formats accepted: jpg, png, bmp
File size should not exceed: 4mb

Using model trained in

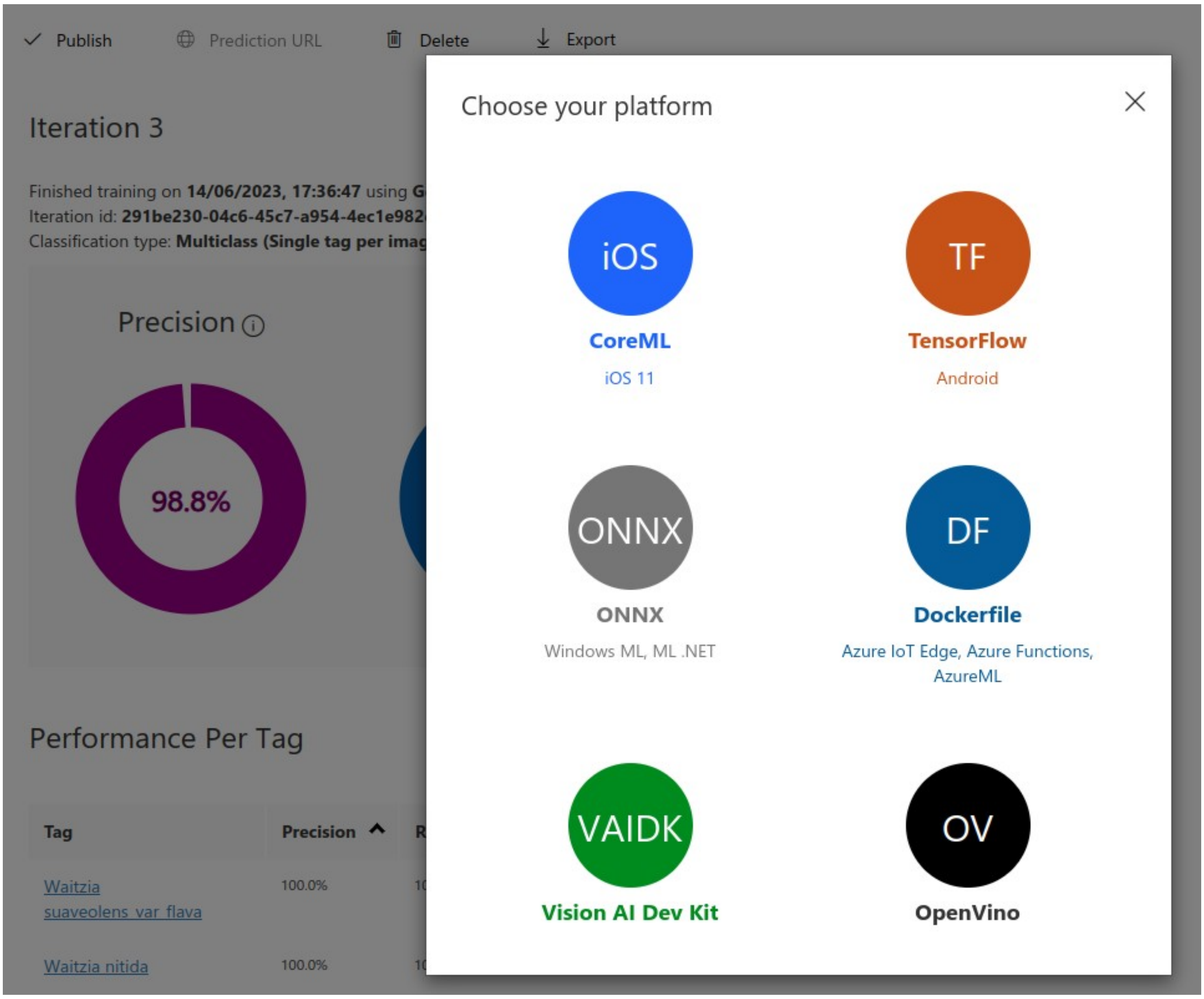
Iteration: Iteration 3

Predictions

Tag	Probability
Waitzia acuminata_var_acuminata	56.8%
Waitzia nitida	42.4%
Waitzia fitzgibbonii	0.4%
Waitzia suaveolens_var_flava	0.1%
Waitzia acuminata_var_albicans	0%

3.6. Export model

Once you consider the model to be performing satisfactorily, click Export in the Performance tab. Select ONNX, then Export. Save the resulting ZIP archive on your local disk and unzip. With the right software libraries and scripts, the model in it can be used to test an entire folder of images, or it can be integrated into a website. For present purposes, we will prepare it to contribute to Horama ID.



Iteration 3

Finished training on **14/06/2023, 17:36:47** using **G**
Iteration id: **291be230-04c6-45c7-a954-4ec1e982**
Classification type: **Multiclass (Single tag per imag**

Performance Per Tag

Tag	Precision	R
Waitzia suaveolens var flava	100.0%	10
Waitzia nitida	100.0%	10

4. Prepare model and other files for Horama ID

Contribution of a model to Horama ID does not only require the model itself but also several accompanying files. Some of these tell the app how to make use of the model; one of them is exported with the model, and the other is part of the example file package, so they are easily prepared. Others provide information to the end-user on the model itself and on the species it can be used to identify, so they will in each case have to be prepared by the contributor of the model.

An example archive is available that demonstrates how all of the files need to look like. It includes the model for flowerheads of the daisy genus *Waitzia* as currently available on Horama ID.

4.1. Convert model to ORT

Models should be converted to ONNX RunTime (ORT) format to optimise them for running on mobile devices. This can be done through a single command using Python’s ONNX RunTime library. For a Python 3 environment, the commands may look as follows:

```
# install ONNX runtime using PIP

pip3 install onnx onnxruntime

# convert a model called “model.onnx”

python3 -m onnxruntime.tools.convert_onnx_models_to_ort model.onnx
```

Commands may use pip and python instead on your system, or you may use a different software to install Python packages than PIP. The author of this document is also happy to do the conversion for you if needed.

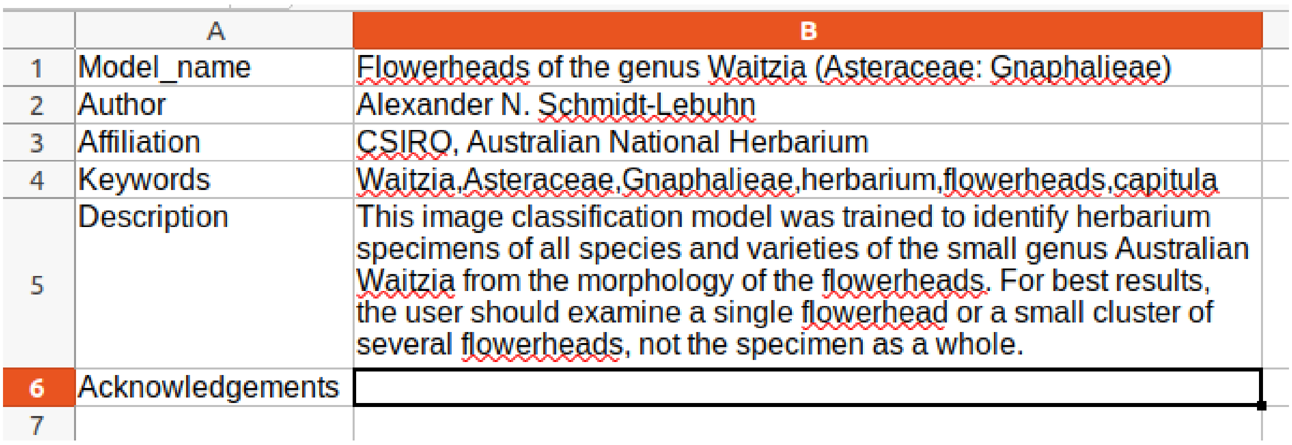
An example model file that has been converted to ORT format is available is part of the example file archive under the name model.with_runtime_opt.ort.

4.2. Model metadata and title image

Prepare a tab separated values (TSV) formatted text file containing information on the model. It will be displayed in the Horama ID app when the user examines a model and tries to decide whether to download it. The fields of the file are:

Model_name	The title of the model that will be displayed in the model list. It should concisely mention the taxonomic group and, if relevant, the geographic scope and the organ, life stage, or morph that the model has been trained on, e.g., “Flowerheads of the genus Waitzia (Asteraceae”).
Author	Names of the author or authors of the model.
Affiliation	Institutional affiliation of the author or authors of the model.
Keywords	Keywords are not yet used but may become useful for searches in future iterations of Horama ID as the model list grows. Keywords should be separated by commas.
Description	To allow the user to decide if this model is of interest to them, this field should specify the taxonomic group, geographic scope, and organ, life, stage morph, preparation, or whatever else is relevant. It should also give the user any other information they may need to get good identification results. Example: “This model is for the identification of herbarium specimens of Australian species of Casuarina from female cones. It will not work on male or sterile specimens. Results on living specimens in the field may not be reliable. For best results, maximise a single cone in the camera view.”
Acknowledgements	This field can be used to acknowledge the help provided by colleagues or funding sources that supported the development of a model.

An example file of model metadata is part of the example file archive under the name project_data.tsv.



	A	B
1	Model_name	Flowerheads of the genus Waitzia (Asteraceae: Gnaphalieae)
2	Author	Alexander N. Schmidt-Lebuhn
3	Affiliation	CSIRO, Australian National Herbarium
4	Keywords	Waitzia,Asteraceae,Gnaphalieae,herbarium,flowerheads,capitula
5	Description	This image classification model was trained to identify herbarium specimens of all species and varieties of the small genus Australian Waitzia from the morphology of the flowerheads. For best results, the user should examine a single flowerhead or a small cluster of several flowerheads, not the specimen as a whole.
6	Acknowledgements	
7		

In addition, every model needs a single title image that is displayed when the user reads the above information. This can be a copy of one of the profile images (see below). In the example archive, the title image is called title_image.jpg, and it comprises a collage of four of the species in the genus.

4.3. Species profiles

To allow the user to confirm a correct identification by the model or to recognise incorrect identifications, species profiles are provided as part of Horama ID models. The user can examine them by tapping an identification suggestion, and they can also be accessed from the model page after the model has been downloaded. A taxonomist contributing a model should provide profiles as a single comma separated values (CSV) formatted text file with the following column headings:

Tag

The exact name/label that the class has in the Custom Vision model.

Linnean_name

The scientific/Latin/Linnean name of this class. This is the primary name that will be displayed in the identification suggestions in Horama ID.

Authority

The taxonomic authority of the name. (optional)

Common_name

One or several common names for the class in whatever languages are relevant to the project. (optional)

Higher_taxon

This field is intended for, e.g., the plant family or insect order that this model class belongs to. (optional)

Description

A short description that allows the user to check if the model's inference is actually correct. It would therefore ideally focus on key distinguishing or spotting characters and maybe mention a similar species that may cause false positives for this species. (optional)

Occurrence

Geographic range of the model class in whatever format is relevant to the overall geographic scope of the project, be it the provinces of a country if the overall scope is that country, entire continents if the scope is global, or even information such as "native" and "introduced". (optional)

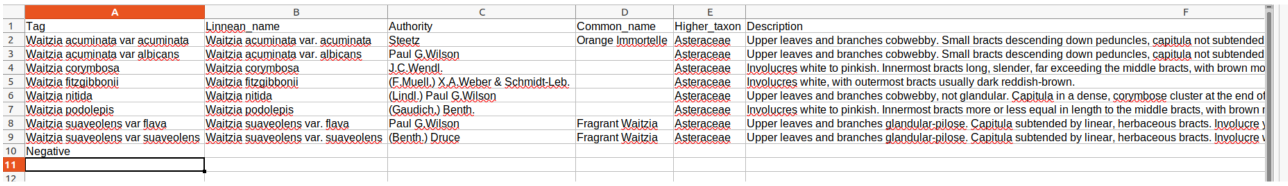
Image

Exact file name of a single profile image of this model class (see also next section). It should make it as easy as possible for the user to confirm if the specimen they are trying to identify has or has not been correctly identified by the model.

Weblink

A single URL that may lead to additional information on this model class. This could be its profile in an electronic flora, a page outlining how to treat an infestation of this species, or a link to its records on GBIF. (optional)

All column headings are required. However, most information is optional. For classes that are identification suggestions, only Tag, Linnean_name, and Image are required. For the Negative class, only its entry under Tag is required. The CSV file can be prepared in Excel or equivalent software. Ideally, it would be saved or exported with quotation marks around the cells.



	A	B	C	D	E	F
1	Tag	Linnean_name	Authority	Common_name	Higher_taxon	Description
2	Waitzia acuminata var acuminata	Waitzia acuminata var. acuminata	Steetz	Orange Immortelle	Asteraceae	Upper leaves and branches cobwebby. Small bracts descending down peduncles, capitula not subtended
3	Waitzia acuminata var albicans	Waitzia acuminata var. albicans	Paul G.Wilson		Asteraceae	Upper leaves and branches cobwebby. Small bracts descending down peduncles, capitula not subtended
4	Waitzia corymbosa	Waitzia corymbosa	J.C.Wendl.		Asteraceae	Involucres white to pinkish. Innermost bracts long, slender, far exceeding the middle bracts, with brown mo
5	Waitzia fitzgibbonii	Waitzia fitzgibbonii	(F.Muell.) X.A.Weber & Schmidt-Leb.		Asteraceae	Involucres white, with outermost bracts usually dark reddish-brown.
6	Waitzia nitida	Waitzia nitida	(Lindl.) Paul G.Wilson		Asteraceae	Upper leaves and branches cobwebby, not glandular. Capitula in a dense, corymbose cluster at the end of
7	Waitzia podolepis	Waitzia podolepis	(Gaudich.) Benth.		Asteraceae	Involucres white to pinkish. Innermost bracts more or less equal in length to the middle bracts, with brown r
8	Waitzia suaveolens var flava	Waitzia suaveolens var. flava	Paul G.Wilson	Fragrant Waitzia	Asteraceae	Upper leaves and branches glandular-pilose. Capitula subtended by linear, herbaceous bracts. Involucre y
9	Waitzia suaveolens var suaveolens	Waitzia suaveolens var. suaveolens	(Benth.) Druce	Fragrant Waitzia	Asteraceae	Upper leaves and branches glandular-pilose. Capitula subtended by linear, herbaceous bracts. Involucre w
10	Negative					
11						
12						

An example file of species profiles is part of the example file archive under the name profiles.csv.

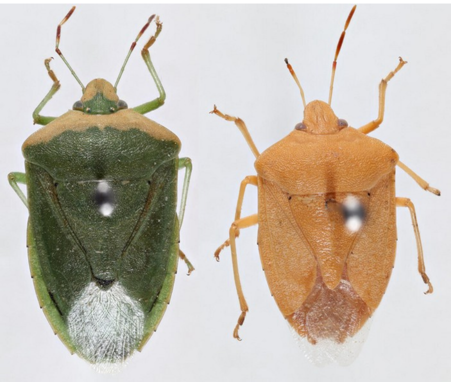
4.4. Profile images

Every species should have a profile image that the user can compare identification results against. Currently, Horama ID only supports a single profile image per model class. If several different views or variation in morphology need to be displayed, a single image could be created showing both of them, as in the example of the polymorphic insect depicted below.

Profile images should ideally be JPG formatted. They would ideally be fairly small in file size, as large numbers of large images would bloat the size of the model metadata that Horama ID has to download onto the end user's phone. It would therefore be good to rescale profile images to a width of no more than 1440 pixels and use at least some degree of JPG compression.

The file names of the profile images need to match exactly the entries in the Image column of the species profile table, see previous section.

Example profile images are part of the example file archive. They can be found in the subfolder called profile_images.



4.4. Other files

In addition to the model file, project metadata TSV, the title image, species profile CSV, and profile images discussed above, Horama ID requires two other files.

First, the labels.txt file that is exported with the model from Custom Vision.

Second, a JSON formatted text file that specifies the image preprocessing parameters for the model. However, as these parameters are the same for all image classification models created with Custom Vision, the same JSON file can be re-used for each such model. It is included in the example file archive under the name metadata_properties_bgr.json.

5. Contribute model to Horama ID

5.1. Publication of models

If you are only contributing a single model that you want to make available to end-users, it is easiest to email the files to the author of this document (alexander.s-l@csiro.au), who will add your model to Horama ID's model pool and can also convert the model file to ORT if required. If you want to become more involved in Horama ID or use the app iteratively for field- or bench-testing models with a smartphone as you develop them, reach out to discuss being given access to the server storing the models.

If you are contributing directly to the server, the steps for uploading a new model are as follows.

Open the CSIRO ID Models tab and click Add A New Bundle. Enter a name and save.

Each bundle of files can be examined in two windows, View and Files.

Under Files, you upload and manage all the files outlined above. Click Add Files To Bundle to upload. You can upload several files together, and you can also upload a ZIP archive and have the server unzip it after upload. After upload, each file shows a Mapping, which the server at first tries to guess, i.e., whether it is the species profiles, the model, a profile image, etc. Click Manage to change the mapping if it is incorrect. The Manage interface also allows you to choose 'blocked', which means the file will be ignored. This is how you resolve a conflict caused by uploading two different files of the same type where only one is allowed, i.e., all except profile images.

Go the View window to see the status of the bundle. If it says "Ready for release" at the top, the model will appear in the Horama ID app. If it isn't ready for release, the system sees and error, which could be missing files, wrong formats or mismatches between files that refer to each other, or multiple files with the same mapping. The Details of Parts section at the bottom should provide information on what file or files have caused the error.

If you are publishing a model for the first time, you will have to accept the Term of Use. See the next section for details.

If you have updated an existing bundle, you should change the text in the Current Version field. This will signal to the Horama ID app that the model has been updated, and it prompts users of the app to re-install this particular model to get the newest version.

If you are only testing a model that isn't yet ready for prime time, tick the field Is Beta Test/Prototype. It will only be visible to users of the Horama ID app who opt into test/beta models. Note, however, that any app user is free to do so.

5.2. Licensing and other conditions

The text below is a copy of the Terms Of Use for contributors of models to Horama ID.

1. CSIRO provides you with the opportunity to upload your content (the Content) to the Horama ID app and the website [https://dev.idapp.cosinecrm.com.au/] (together, Horama ID) so that the Content can be used by other members of the public for non-commercial purposes.

2. CSIRO does not claim ownership of the Content that you upload to Horama ID. However, by uploading your Content, you grant CSIRO permission to use your Content in connection with Horama ID including, without limitation, the license rights to copy, distribute, transmit, publicly display, publicly perform and reproduce your Content; the right to sublicense such rights to any user of Horama ID; and the right to publish your supplied name in connection with your Content.

3. By uploading the Content, you warrant and represent that:

a) you own the Content or have all the rights necessary for you to upload the Content to Horama ID and to grant the permissions to CSIRO at paragraph 2 above;

b) your Content does not infringe any third party intellectual property or other rights; and

c) the Content contains no harmful or offensive material.

4. By uploading the Content, you agree that:

a) all users of Horama ID are permitted to use the Content for non-commercial purposes; and

b) CSIRO will not pay you any compensation for the upload and use of your Content.

5. CSIRO will not be responsible if Content uploaded does not meet terms of copyright, fair use, correctness or safety. You are solely responsible for any third-party claims regarding your Content.

6. You acknowledge that at any time CSIRO may decide to discontinue its support of Horama ID and withdraw Horama ID and your Content from public access.

7. You acknowledge and agree that many of the trade marks or logos displayed on Horama ID such as the CSIRO logo or name are trade marks of CSIRO and cannot be used in a manner that suggests CSIRO endorses or is associated with your business, products or services without the express written permission of CSIRO. Except to the extent that you include your name or logo as part of the Content, CSIRO will not use your name or logo without your express written permission.

8. CSIRO reserves the right to remove your Content from the App if it becomes aware that you have not complied with the conditions above.

9. CSIRO’s Privacy Policy available at https://www.csiro.au/en/About/Access-to-information/Privacy also forms part of these conditions. The information in the Privacy Policy about how CSIRO collects and uses personal information through the CSIRO website also applies to personal information collected or provided by you as part of your upload of Content.

10. CSIRO will not be responsible for any misuse of your Content by a third party user of Horama ID.

11. To the fullest extent permitted by applicable law, CSIRO (including its employees and contractors) excludes all liability to any person for any consequences, including but not limited to all losses, damages (including indirect, special or consequential damages, loss of business, revenue/profit, loss of time etc.), costs, expenses and any other compensation, arising directly or indirectly from your upload of the Content to Horama ID.