**Installation and use of CNN for desktop processing with Shiny and Rstudio**

Updated March 2022.

Direct comments and questions to Zack Ruff, [zjruff@gmail.com](mailto:zjruff@gmail.com).

**Note:** If you have used a previous version of this app, please be advised that the setup instructions have changed a bit. This is due to changes in the relationship and interdependency between Keras and TensorFlow, and their corresponding packages in Python and R, as of TensorFlow version 2.0.

If you obtained this document in a zipped folder you will want to extract (unzip) it and copy or move the whole folder into RStudio’s working directory. This is the directory that is returned if you run getwd() from the console window; the default on Windows is C:\Users\[username]\Documents. The folder should contain this document, two R scripts called app.r and functions.r, a file called target\_classes.csv containing descriptions of the CNN target classes, and a file called Owl\_CNN\_v4\_TF.h5 which contains the trained neural network.

For purposes of this guide we assume that you are using a computer running a 64-bit version of Windows on which you have at least some ability to install programs. The app will not function on 32-bit operating systems. Running the app on MacOS or Linux would likely require some modifications to the code which we have not attempted thus far.

**Part 1. Installing required software**

**Step 1.** Install the most recent version of R. If it is already installed, go to step 2.

Go to <https://cran.rstudio.com/> and click “Download R 4.x.x for Windows” (latest version at time of writing was R 4.0.4). Open your Downloads folder and run the installer. It will be called something like R-4.0.4-win.exe. All the default settings should be fine.

**Step 2.** Install Rstudio. If it is already installed, go to step 3.

Go to <https://rstudio.com/products/rstudio/download/#download> and click the big blue button that says “Download Rstudio for Windows”. This file is about 150 MB. Go to your Downloads folder and run the installer. Again, the default settings are fine for our purposes.

**Step 3.** Update R and Rstudio.

If you just installed both programs, skip this step. This also may not be strictly necessary as long as your packages are reasonably up to date, but if you have been using the same version for a year or more, you should probably update regardless. Updating regularly is good practice.

To update R, open the R console. Run the following commands:

install.packages("installr")  
library(installr)  
updateR()

This will install the “installr” package, load the package, and run the updateR function, which will search for the latest version and install it if there is one available. It will generate various popups; just agree to the default options and it should all be fine.

To update Rstudio, open Rstudio. In the menu bar up top, click Help > Check for Updates. If there is a newer version available, click “Quit and download.” Run the installer.

**Step 4.** Install SoX.

Sox is a small command-line program that we use to process audio data. You won’t need to do anything directly with it, but it needs to be installed on your computer. This program has been on version 14.4.2 since 2015. Go to <https://sourceforge.net/projects/sox/files/sox/> and click “Download latest version” (the big green button) to download the executable installer for Windows. Open your Downloads folder and run the installer. Make a note of where it is installed; the default is something like C:\Program Files (x86)\sox-14-4-2.

**Step 5.** Set up the r-reticulate conda environment.

This is the most complicated step, so some explanation might be worthwhile. Our CNN was built and trained using the TensorFlow software library developed by Google. The tensorflow R package provides a way to access this library and thus run our CNN, but this R package is dependent on packages that were written for Python, which is a completely different programming language. So in order to run the CNN from RStudio, we first have to install Python and tell RStudio where to find it. For convenience and consistency, we do this using Miniconda, an environment manager for Python that can be installed from within RStudio using the reticulate package.

Open Rstudio and use the console pane to enter the command

install.packages("reticulate")

This will install the reticulate package and its dependencies. Next, enter the command

reticulate::install\_miniconda()

The default installation path for Miniconda depends on your specific system, but it generally will be somewhere in your user directory, so it should not require any elevated permissions. The installation path on my machine is

C:\Users\zjruf\AppData\Local\r-miniconda

The path will be shown in the console window once the installation is successful, but you can also check it from RStudio by running the command

reticulate::miniconda\_path()

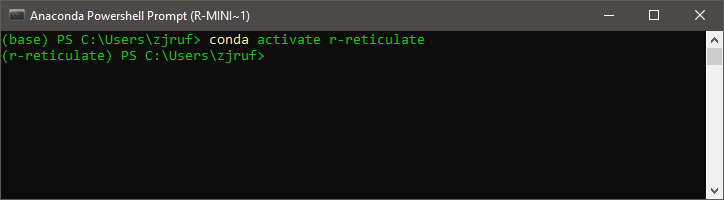
Once Miniconda is installed, it will automatically create a conda environment called r-reticulate. This is essentially a self-contained Python installation that RStudio will use to load and run the CNN. However, we will first need to configure this environment with the right Python packages. Use the Windows search bar to search for "anaconda" and click on the "Anaconda Prompt" option. This should spawn a command prompt window that looks like this:



The (base) indicates that the base conda environment is active, while the C:\Users\zjruf> indicates the current working directory. We can run commands from this window just like from the RStudio console pane. Enter the command

conda activate r-reticulate

This will activate the r-reticulate environment that we just created. The (base) at the beginning of the prompt should change to (r-reticulate), which indicates that this environment is now active.



Once the environment is active, we can install Python packages to it. Confusingly, Miniconda includes two separate tools for doing so, Pip and Conda, and we will need to use both. However, they work very similarly. First we will use the pip tool to install the tensorflow package, like so:

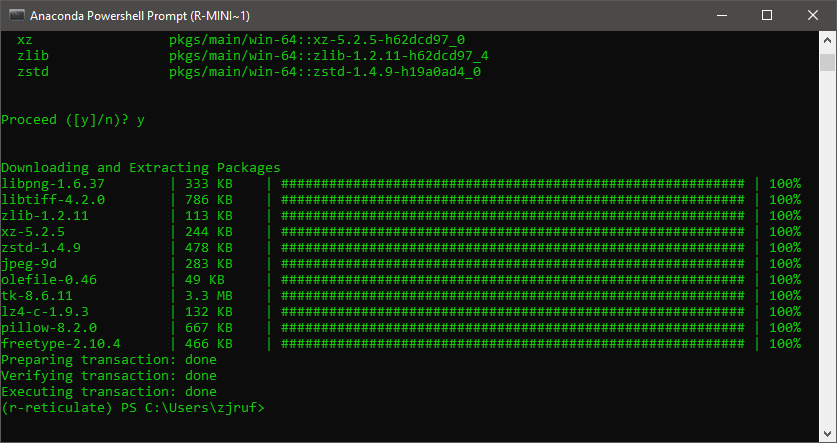
pip install tensorflow-cpu==2.2

(Note that you need *two* equal signs between "cpu" and "2.2".) This will install the tensorflow Python package (specifically version 2.2 of the CPU-only variant) and a number of dependencies. Semi-informative text will flood the console window, progress bars will appear and fill up to indicate that things are happening, and ultimately you should be returned to a prompt with a blinking cursor.

We will then use the conda tool to install the pillow package, like so:

conda install pillow=8.2

Type "y" and hit Enter when you are prompted to proceed. Pillow is the currently maintained fork of the Python Image Library and is used by TensorFlow to handle the spectrograms, which are stored as image files. There will be some text output, though probably not as much as from the last operation. Again, you will be returned to a prompt with a blinking cursor once everything is done.



Once these steps are complete, the r-reticulate environment should be all set up, and you can close the Anaconda Prompt window.

**Step 6.** Install the other necessary R packages.

Back in RStudio, run the command

install.packages(c("tensorflow", "keras", "tidyverse", "shiny", "shinyjs", "tuneR"))

keras and tensorflow are closely related packages that give R an interface for using Google's TensorFlow machine-learning software library. The tidyverse is a family of packages that provide various useful data-handling functions. tuneR is a package for R-based analysis of audio and provides useful functions for dealing with wav files. shiny is a package for building interactive apps and is used to create the graphical user interface for the app. shinyjs adds some useful JavaScript-based functionality to the base shiny app.

**Step 7.** Close Rstudio and open it again.

At this point you should have all the software that is required to use the CNN app.

**Step 8.** (Optional) Install Bulk Rename Utility.

As the name implies, this is a simple utility program for quickly and efficiently renaming many files. The app assumes that filenames follow our naming conventions, in which each filename encodes some important information about where and when the recordings were made; it is therefore important that any files you wish to process be formatted in the same way. The naming conventions are described in the next section.

**Filename formatting**

Your audio files should have filenames structured as follows:

Area\_Site-Station\_Date\_Time.wav

This filename consists of four components separated by underscores: Study area, location, date, and time. The Study area or project ID is typically a three-letter code. Site and station ID should both be alphanumeric and should be separated by a dash (-). The date should be in YYYYMMDD format. The time should be in HHMMSS format in 24-hour time. This time and date format matches the default filename convention used by Wildlife Acoustics ARUs.

For instance, our filenames are formatted as follows:

TYE\_19317-A\_20210516\_201530.wav

The filename indicates that this file was recorded in the Tyee demographic study area (TYE), in hexagon 19317, at station A, and was recorded beginning on May 16, 2021, at 8:15:30 PM. Note that both the spatial and temporal information in this filename is arranged in order from most general to most specific – from study area to hexagon to station within the hexagon, and from year to month to day, and finally from hour to minute to second.

We recommend not using directory names containing spaces, as these can cause issues with processing.

**Part 2. Using the CNN through the Shiny interface**

These instructions assume you have installed all the necessary software as outlined above.

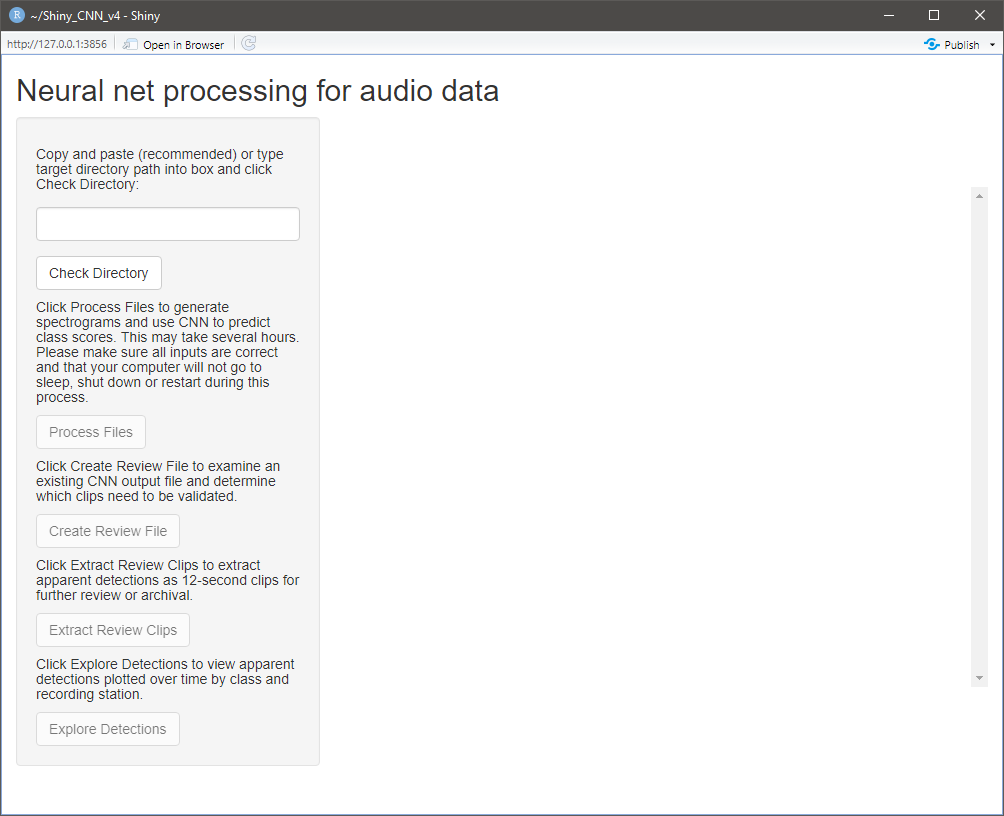
Note: This application works for our purposes, and we have tried to make it somewhat user-friendly, but it was not made by professional software engineers, so its operation is somewhat finicky. It makes some assumptions in order to work efficiently, and it may not work well (or at all) if those assumptions are broken. Be careful!

You will launch this application from within Rstudio by typing a command into the Console pane; further operation will use the Shiny window that appears once the program is running. The Shiny window will take over Rstudio’s input and output as long as it is open; you will not be returned to the normal prompt until this window is closed. However, some of the program’s output (progress bars, messages, etc.) will appear in the console pane, so it is helpful to keep an eye on Rstudio while the CNN application is running.

**Step 1.** Copy this folder to your RStudio home directory, i.e. whatever folder comes up if you open Rstudio and run getwd(). The default in Windows is C:\Users\[username]\Documents. This home directory needs to contain a folder called “Shiny\_CNN\_v4” which should contain two R script files called app.r and functions.r, which contain the source code for the app, a text file called “readme” which lays out some potentially helpful information, and a file called Owl\_CNN\_v4\_TF.h5, which contains the pretrained CNN model.

If you installed SoX in a non-default location in Part 1 (Step 4), you will need to edit the app.r file so that the app knows where to find SoX. Open the app.r script file in RStudio and edit line 17, replacing C:/Program Files (x86)/sox-14-4-2 with the path to the correct folder.

**Step 2.** In the RStudio console window, type shiny::runApp("Shiny\_CNN\_v4") and hit Enter. Rstudio will search the current working directory for a folder called “Shiny\_CNN” containing a file called app.r; if it does not find these items, you will get an error. Assuming the folder and app file are found, a Shiny application window will appear. You will also get some output in the console pane as the app loads some R packages and establishes communication with RStudio. When first launched, the Shiny window will look like this:



Most of the controls are disabled on launch; these will become active as the app locates data to be processed and as various steps in the processing workflow are completed.

**Step 3.** Enter the path to the target directory and click Check Directory.

In the Shiny window, use the text box on the left-hand side to input the directory containing the files that you want to process. Use the Check Directory button to verify that R can find some WAV files. The number that shows up is wav files that are ≥500 KB in size and do not have “\_part\_” in the name. The list of files that were found will appear in the main part of the window. If any valid WAV files were found, the Process Files button will become active.

**Step 4.** Click the button labeled Process Files.

[Note: The first time you attempt to process data, you may get a Windows Firewall pop-up saying that features of the program have been blocked. This should not actually prevent the app from working; you can close this pop-up and the program should continue to run.]

The app will create spectrograms in the form of image files in Portable Network Graphic (.png) format, representing non-overlapping 12-second segments of the audio. The images will be created in a new directory called Temp/images within the folder you specified in step 4. There will be 300 images for every hour of audio, so if you have a substantial dataset there will be a lot of them. We have not implemented a progress bar for this step, but as a crude workaround you can open the Temp/images folder in Windows Explorer and watch as the files are created.

Note that this step is by far the longest; processing time will vary depending on your computer’s hardware capabilities and where the data are stored. We recommend using a computer with multiple cores, although if you have a lot of cores the speed can also be limited by the read and write speeds of your storage device. As a very general rule of thumb, with data stored on external USB hard drives, you should be able to process approximately 100 hours of recordings per hour of processing time.

Once all the images have been generated, the app will load the trained CNN model and start classifying the images in batches. A text-based progress bar will appear in the console pane in Rstudio showing how many batches have been processed. Once the full set of images has been classified, the app will write the results to a comma-separated values (.csv) file in the same folder and will open a dialog box indicating that audio data processing has finished. The results file can be opened in a spreadsheet program or text editor.

In addition to the file containing the raw predictions, the app will create a summary file called [Site]\_detection\_summary.csv showing the number of apparent detections of each target class by station and date across a range of detection thresholds. Apparent detections are defined as images (equivalently, 12-second clips) to which the CNN assigned a score higher than the detection threshold for a given class.

If the CNN predictions were successfully written to file, the Create Review File and Explore Detections buttons will become active.

**Step 5.** [Optional] Click the button labeled Explore Detections.

The interface will change to look something like this:



[From here, click the button labeled Back to inputs to return to the previous view.]

The plot to the right is generated directly from the detection summary file generated in step 4. You can use the controls in the sidebar to visualize apparent detections of any of the 51 target classes at thresholds ranging from 0.05 to 0.99, summarized daily or weekly over the recording period.

Please note that these counts are intended **only** for data exploration purposes and should **not** be treated as objective evidence of species presence or absence without further human review. Put simply, AI is far from perfect and lots of weird stuff happens in the woods. While the false positive rate is quite low for many of these classes, especially at higher thresholds, verification of apparent detections by knowledgeable human reviewers is an important and necessary part of the process. Which brings us to…

**Step 6.** Click the button labeled Create Review File.

This will search the CNN prediction file for lines where either a) the score for one of the spotted owl classes (STOC or STOC\_IRREG) was ≥ 0.25 or b) the score for any of the other target classes was ≥ 0.95, and write these lines to a new file. This is the set of clips that are flagged for human review under our current validation procedure. The app will write a message to the console pane that says Attempting to create review file based on CNN\_Predictions\_[location].csv… Once the review files have been created the console will report the number of lines written to file. This button will actually create two files containing similar information; one ends in \_review.csv and is simply a filtered version of the CNN\_Predictions file, while the other ends in \_review\_kscope.csv and is intended to be used with Kaleidoscope. See the section below titled "Using the review\_kscope file with Kaleidoscope" for details.

**Step 6.** [Optional] Click Extract Wav Files.

[Note: We do not recommend this as part of the normal review process. You can easily review apparent detections of your species of interest using the \_review\_kscope.csv file in combination with Kaleidoscope without having to extract them as short clips first, which saves a fair amount of time and disk space. See the following section for details.]

This will create a new folder called Review in the target directory specified in step 3. The Review folder will have subfolders for each of the target classes, which will contain the clips of that class that warrant further review. The nested directory structure that is created is divided up by target class, then by recording station, then by week.

The app will print a message to the console that says “Extracting wav files beginning at [time]…” Extracting the clips may take some time. When all the new clips have been made there will be another message that says “Finished at [time]. [n] clips extracted.” Once you get the message that review clips have been extracted, you can close the Shiny window and exit Rstudio. Below is an example of console output from the processing of a very small test set.

Clips are extracted as .wav files with the same parameters (bit depth, sample rate, number of channels) as the input data. Depending on the dataset and how much other data is on the same hard drive, storage space may be an issue when extracting clips for review. For instance, our data are single-channel .wav files with a sample rate of 32 kHz and a bit depth of 16 bits, which works out to roughly 225 MB per hour of audio. This means that every 12-second clip takes up roughly 750 KB, and 10,000 such clips will take up about 7.5 GB. You can use the number of lines in the Review file to judge how much free space you will need before extracting clips for review.

Once clips for review have been extracted, you can close the Shiny window by clicking the X in the upper right-hand corner or by pressing Esc from the Rstudio console window.

**Using the review\_kscope file with Kaleidoscope**

*Note: We are not affiliated with Wildlife Acoustics, Inc. This guide has not been reviewed or endorsed by Wildlife Acoustics and may not represent specifically authorized or intended usage of Kaleidoscope.*

*For detailed instructions on using Kaleidoscope, see the User's Manual available at* [www.wildlifeacoustics.com](http://www.wildlifeacoustics.com).

When you click the "Create Review File" button, the app will create two different files. Both have a .csv extension and contain comma-separated text; the name of one file ends with \_review.csv and the other ends with \_review\_kscope.csv. The review\_kscope file is meant to be used with Kaleidoscope, an audio processing program available from Wildlife Acoustics, Inc. Kaleidoscope is available in a free (Lite) or paid (Pro) version; the free version is sufficient to review apparent detections.

The review\_kscope file is a table which points to specific locations within the original audio in a format that Kaleidoscope can understand, allowing you to browse these locations and apply species tags. This approach has a few advantages: You do not need to take the time to extract thousands of short clips or allocate extra disk space for them. You can easily examine the audio surrounding a given clip for additional context. You can sort the review file in different ways, allowing for more flexibility in how you browse the clips. Kaleidoscope also seems to be slightly faster at loading short portions of a long audio file than many shorter clips.

**Setting up**

When you launch Kaleidoscope you will see the main window, which has a number of tabs; the Batch tab will be in focus. This lets you specify an Input and an Output directory. In both of these fields, enter the path to the target directory that you specified when you processed the audio using the CNN app, and it should contain the CNN\_Predictions file, the review file, and the review\_kscope file. Now click the Signal Params tab. In the field marked "(s) Maximum inter-syllable gap," enter 0. By default Kaleidoscope will add "padding" equal to this number of seconds before and after each clip. It is better to avoid this so that you see what the CNN sees.

In the Kaleidoscope main window, click File > Open Results. This will bring up a file selection dialog. By default it will show you the folder specified in the Output directory field. Select the review\_kscope file and click Open. This will bring up two additional windows: the Viewer window, which features waveform and spectrogram representations of the audio, and the Results window, which contains a table representation of the review\_kscope file.

In the Results window, click File > Edit columns. In the window that pops up you can use the Show, Hide, Move Up and Move Down to change which columns are visible and in what order. The files I recommend having visible, and their interpretation, are as follows:

IN FILE: The name of the original audio file.  
PART: The segment of the original audio file that we are looking at. An hour of audio is divided into 300 12-second segments.  
MMSS: Offset from the start of the file in minutes and seconds. Derived from PART – for instance, part\_005 represents the 5th 12-second segment of the audio file, which begins at an offset of 00:48.  
TOP1MATCH: The most likely class according to the CNN. Same as "Sortlabel" in the regular review file. Note that the app prioritizes spotted owls and applies a lower threshold to this class, so clips labeled STOC are those which had a score of 0.25 or more for the STOC class, regardless of the scores for the other classes.  
TOP1DIST: The actual score assigned to the most likely class by the CNN, indicating the strength of the match. Ranges from 0 (no match) to 1 (very close match).  
MANUAL ID: If you apply species tags to a clip, they will appear in this column. When the file is originally created this field will be blank.  
SORT: A combination of apparent label (same as TOP1MATCH), Station, and Week. This is included so the user can quickly confirm the presence of each species at each station within each week, facilitating the efficient generation of weekly encounter histories.

The other columns are mostly irrelevant for our purposes and can be ignored. Click OK in the Edit Columns window to apply the selection to the results window.

In the Viewer window, click File > FFT Settings. Change FFT SIZE to 1024 and WIN SIZE to 256 and click OK. The sliders under the spectrogram control brightness / gain (which also controls playback volume) and contrast / dynamic range; I generally keep them at 0 and -96, respectively.

You can browse through clips in the results\_kscope file by clicking a row in the Results window and using the up and down arrow keys to move to the next and previous row, or by clicking on the spectrogram in the Viewer window and using the up and down arrows, or using the < and > buttons below the spectrogram on the right side. If the Auto next file box is checked, you will automatically advance to the next clip after applying a tag.

Tags can be applied in a few ways as well. In the Viewer window, you can type a tag into the Identification field and hit Enter. You can use the buttons under this field as well; the button on the left will automatically populate with the value in the TOP1MATCH field, while the button on the right will always apply the Noise tag. There are four rows of rectangular buttons at the bottom of the Viewer window; these can be reprogrammed by right-clicking them and typing the tag you want them to apply. Buttons in the top row are linked to number keys 1 through 8 so you can use these keys to apply tags from this row, which is often convenient. You can also select a range of rows in the Results window and click File > Bulk ID selected rows to tag all the selected rows with a particular tag.

You can sort the Results table by clicking on the column headers to browse the clips in different ways. Clicking a column once will sort the table by that column in ascending order, and clicking it again will reverse the order. To sort by multiple fields, click the column headers in order from the finest to the broadest level of the sort. For instance, sorting by TOP1DIST in descending order and then by SORT will organize the table by class-station-week and show you the best matches first within those groupings. Sorting by TOP1DIST in descending order and then by TOP1MATCH will allow you to browse the best matches for each class within the whole dataset. Use the << and >> buttons in the Viewer window to move to the previous or next "folder," defined as the next value of the broadest level of the sort. For instance if you have sorted the table in the manner just described, clicking >> will take you from a row with BUVI in the TOP1MATCH field to the first row with COCO in that field.

Note that you must click File > Save in the Results window for any tags you have applied to be preserved in the review\_kscope file. We recommend doing so frequently to avoid losing work in the event that Kaleidoscope crashes or you lose power.