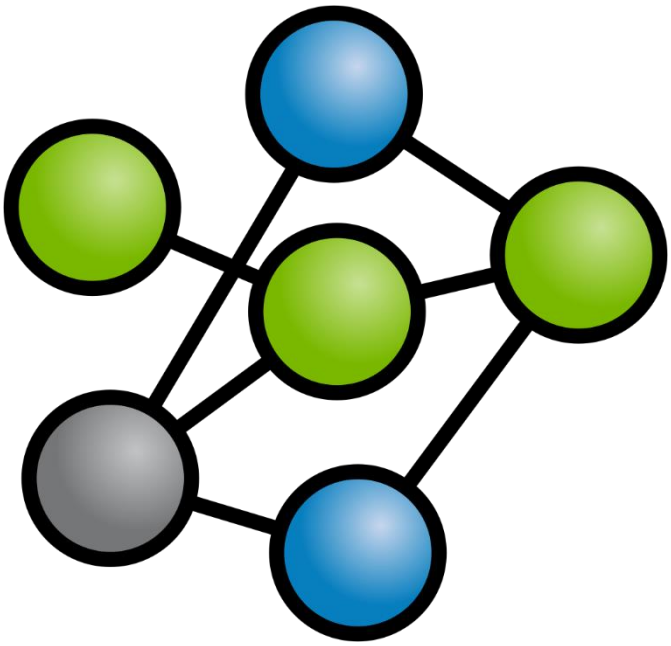

AI Call Detector Model

Performance and Accuracy

bat bio acoustic technology

Version: 1.0.4

1. Introduction:

The AI Call Detector of BATscreen PRO is designed as a three-stage process:

- First Stage: Detect bat calls in sound files.
- Second Stage: Identify species of each call.
- Third Stage: Statistically review species identifications.

To use the AI Call Detector for data set analysis and interpret its results, it is crucial to know the performance and accuracy of the detector stages. There are many metrics and plots that can be used to measure and visualize the performance of a classifier or predictor. We will mainly use "precision" and "recall" to measure the accuracy. These metrics can be easily calculated from the confusion matrix:

	Predicted: Call	Predicted: No Call
Actual: Call	True Positives	False Negatives
Actual: No Call	False Positives	True Negatives

Figure 1: Confusion Matrix for binary bat call detector - like Stage 1 of the AI Call Detector

Precision: *Fraction of relevant instances among the retrieved instances.*

In case of a binary bat call detector, like stage 1 of the AI Call Detector, that means the precision is the fraction of true bat calls among detected bat calls.

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives}$$

Recall: *Fraction of relevant instances that were retrieved.*

Or adapted for our stage 1 bat call detector, that means the fraction of all bat calls that were detected.

$$Recall = \frac{True\ Positives}{True\ Positives + False\ Negatives}$$

To visualize precision and recall depending on an adaptable parameter like the detection threshold of Stage 1 Precision-Recall plots are used. As precision and recall would be 1 for an optimal classifier, the curve of such a Precision-Recall plot should snuggle up to the upper right corner of the plot, ideally.

2. Stage 1 – BAT Call Detection

The Stage 1 of the AI Call Detector is a binary classifier. It detects bat calls within a file by performing multiple Convolutional-Neuronal-Network (CNN) inferences in a sliding-window-fashion on the spectrogram of the file. If a local maximum of the resulting probability curve is above an adjustable detection threshold, a bat call is predicted at the respective temporal position.

Training Data:

The CNN model of Stage 1 was trained with 2875 labeled audio files. The resulting training data consisted of 20745 calls and 28596 noise samples.

Validation Data:

The CNN model of Stage 1 was validated with 319 files containing 3154 calls.

Accuracy:

All results are generated with a “Step” of 1 and a “Minimum-Call-Distance” of 40ms.

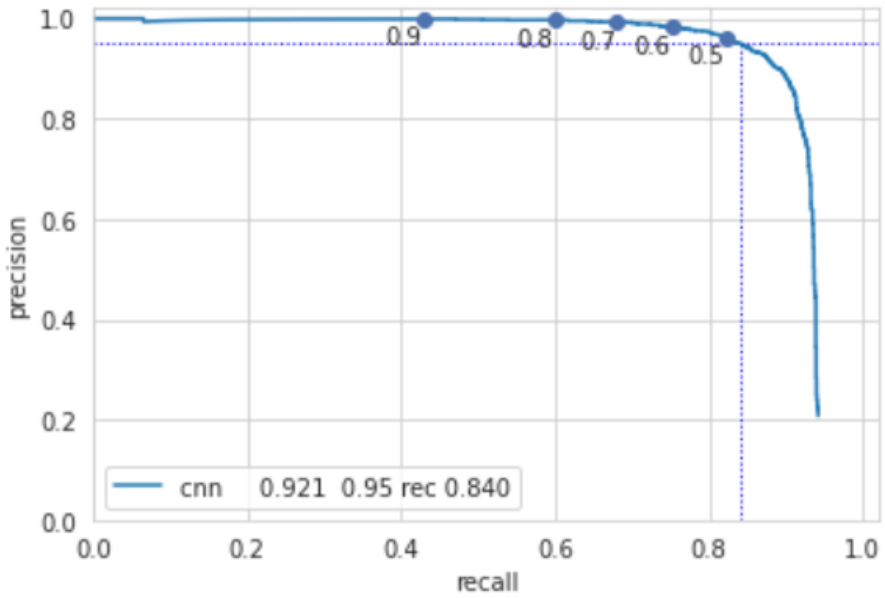

Figure 2: Precision-Recall Plot on Call Level - Curve was obtained by computing the precision and recall depending on the detection threshold. A call is considered as bat call if the detection probability of the local maximum that refers to this call is above the detection threshold. Average Precision: 0.921, Recall at 95% Precision: 0.840.

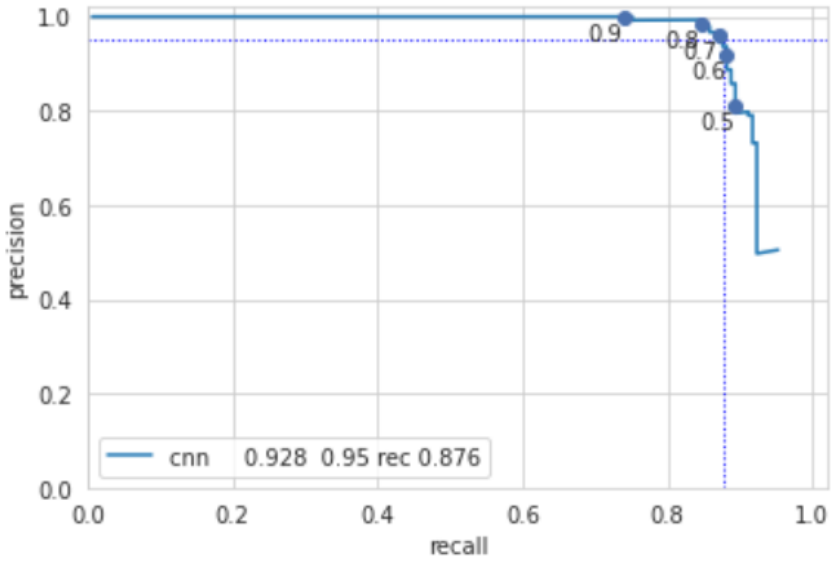

Figure 3: Precision-Recall Plot on File Level - Curve was obtained by computing the precision and recall depending on the detection threshold. A file is considered as bat file as soon as the detection probability is at least once above the detection threshold. Average Precision: 0.928, Recall at 95% Precision: 0.876.

	Predicted: Call	**Predicted:** No Call
Actual: Call	47%	3%
Actual: No Call	6%	44%

Figure 4: Relative Confusion Matrix on File Level for a Detection Threshold of 0.6. Values are normalized based on the total file number.

3. Stage 2 and 3 – Species Identification

In Stage 2 of the AI Call Detector a CNN as multiclass classifier is used for species suggestions of each detected call. Subsequently, in Stage 3 the known confusion data of the multiclass classifier CNN is used for a Maximum-Likelihood-Ratio-Test (MLRT) determining the final species occurrences depending on the number of species suggestions per species in the whole file.

Training Data for CNN model:

The CNN model of Stage 2 was trained with 2501 labeled audio files. The files contained the following number of calls per species:

Species	Nlei	Nnoc	Enil	Eser	Vmur	Ppip	Pnat	Ppyg	Pkuh	Myotis	Plecotus	Bbar	Hsav	Tten
#	1308	1778	1310	957	246	7386	6341	2327	113	4577	4965	4426	2188	1909

Figure 5: Number of Training Calls per Species for Stage 2 CNN

Confusion Data of CNN for MLRT and final Accuracy of Species Identification:

As for some species only very limited data was available for the training and validation process, the confusion data of the Stage 2 CNN as well as the final accuracy were determined by running the whole Stage 2 and Stage 3 training and validation process several times based on a bootstrapped train-test-split. The resulting accuracy of the species identification is depicted in the confusion matrix (see Figure 6).

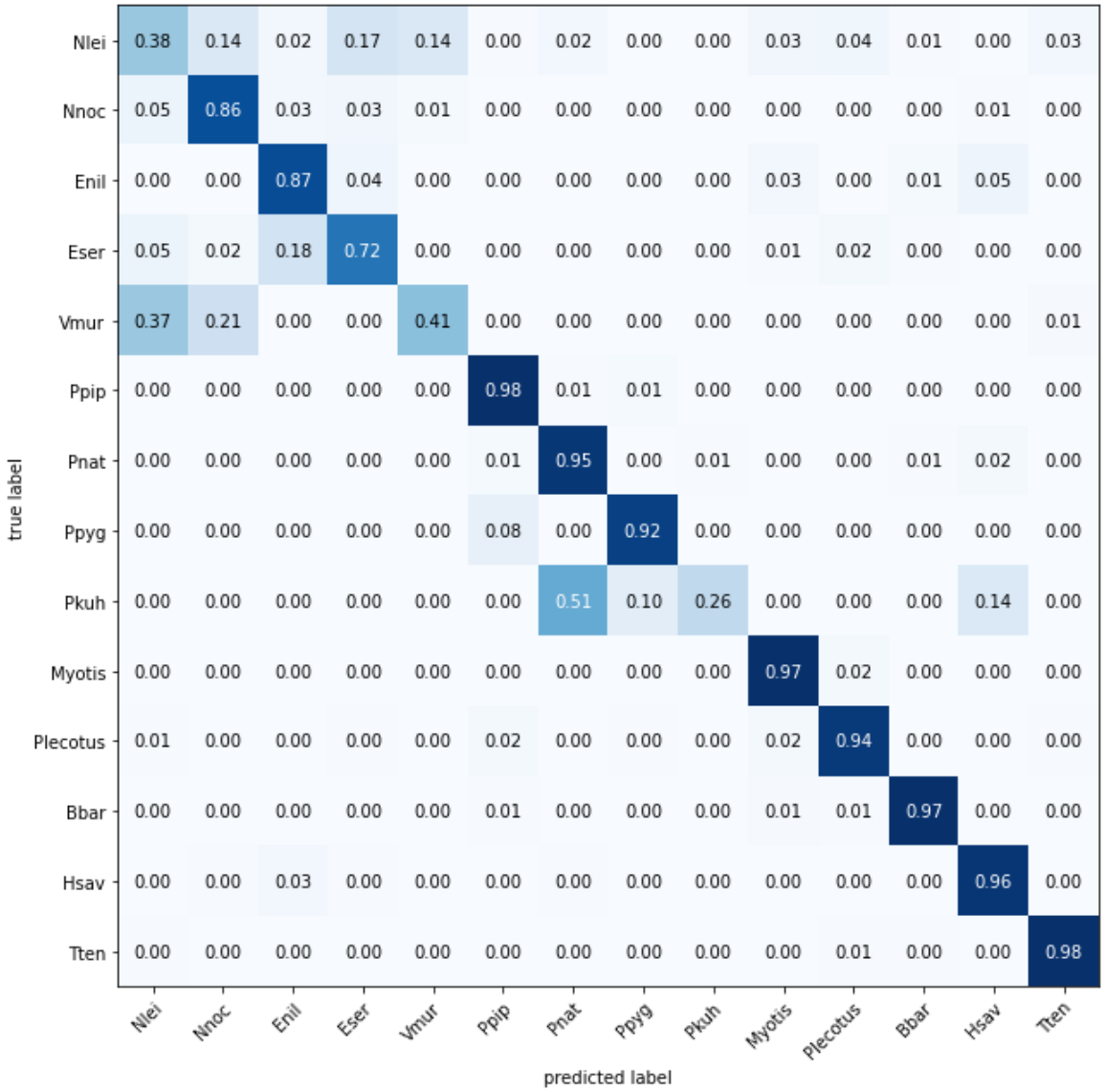

Figure 6: Confusion Matrix for species identification