Sensitivity Study with a Simple CRNN

The extra data here is to support that the choice of model architecture does not affect the results found in the sensitivity study done in the main report. The CRNN architecture used is shown in Table 1, and Figures 1-4 show the results in the same order as done in the report. These results conclude the same parameters found using the CNN model: the best pre-processing parameters and features to extract for the model training of acoustic UAV detection are one second segments of mel-spectrogram features with 90 mel-bins and no overlapping.

Table 1 CRNN architecture hyperparameters used for sensitivity study.

Layer Type	Kernels	Kernel Size	Kernel Stride Size	Padding	Memory Units	# of Neurons	Rate	Activation
2D Convolutional	8	5x5	-	-	-	-	-	ReLU
2D Max Pooling	-	5x5	2	same	-	-	-	-
Batch Normalization	-	-	-	-	-	-	-	-
Reshape for LSTM	-	-	-	-	-	-	-	-
LSTM	-	-	-	-	16	-	-	-
Flatten	-	-	-	-	-	-	-	-
Dense				-		16		ReLU
Dropout	-	-	-	-	-	-	0.	-
Dense (Output)	-	-	-	-	-	2	-	Softmax

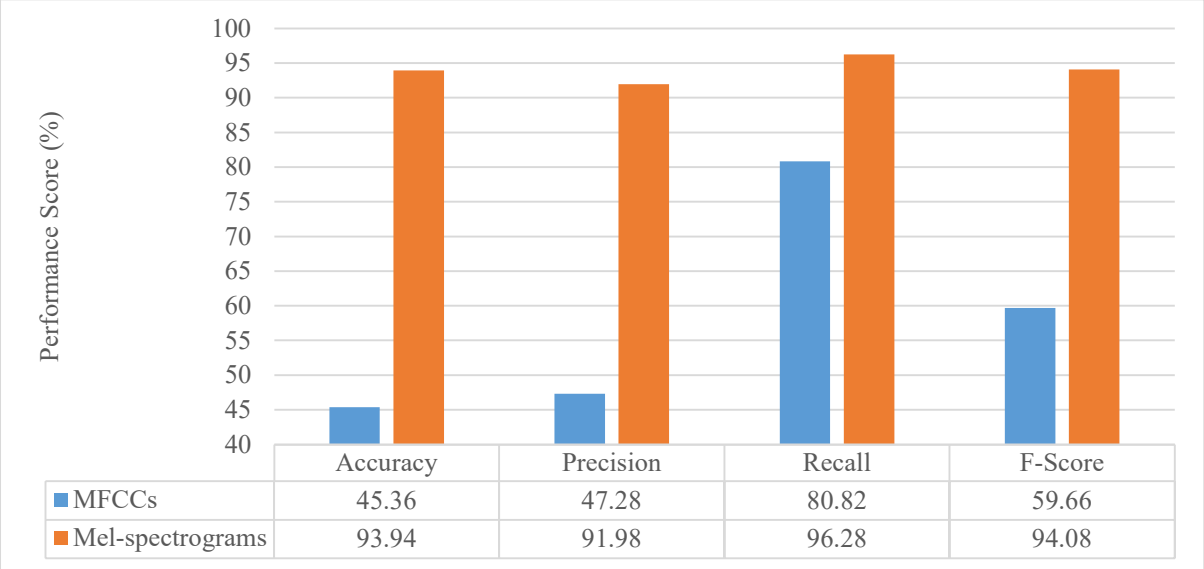
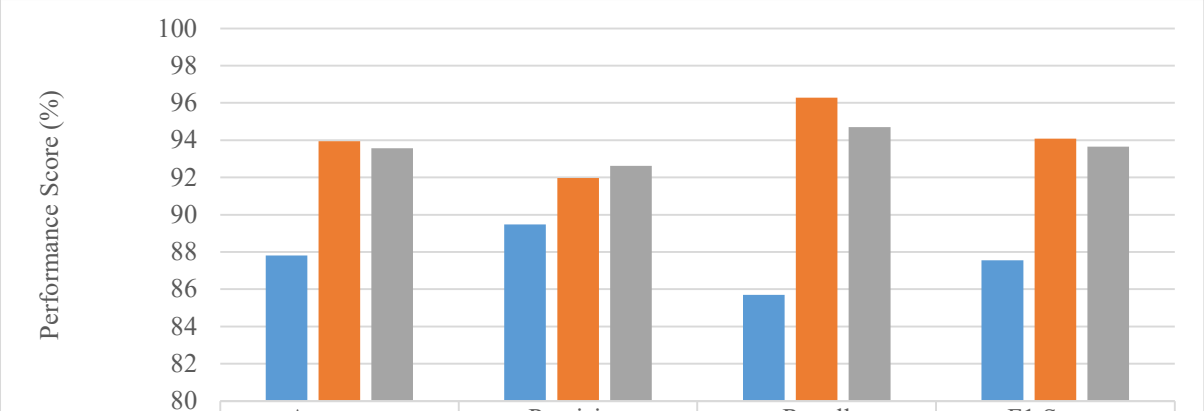
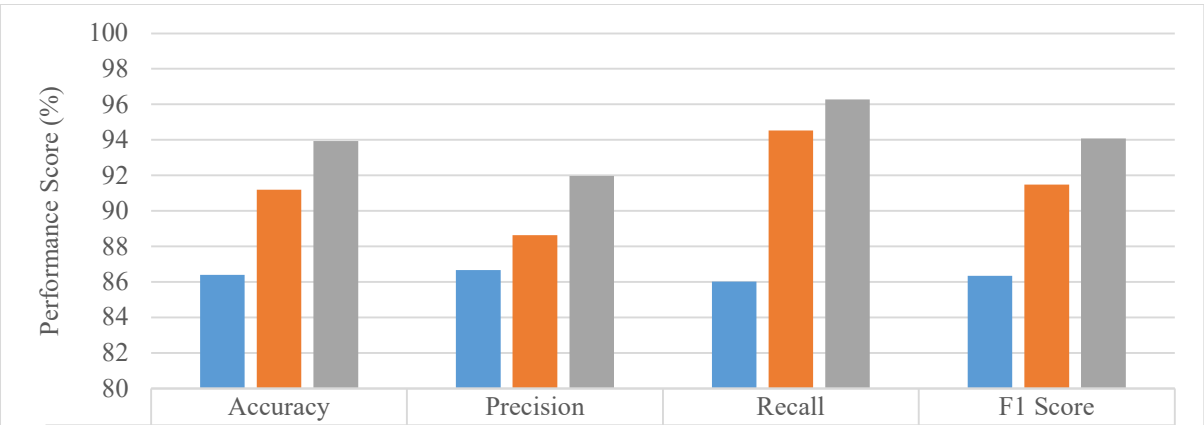


Fig. 1 Performance comparison of MFCC features versus mel-spectrograms, against the unseen dataset.



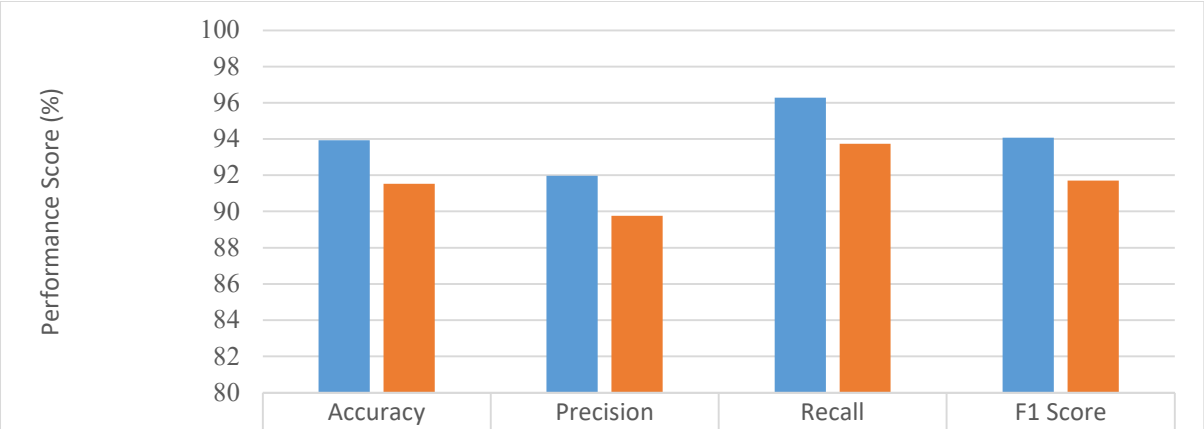
	Accuracy	Precision	Recall	F1 Score
40 mel-bins	87.81	89.48	85.69	87.55
90 mel-bins	93.94	91.98	96.28	94.08
128 mel-bins	93.57	92.62	94.70	93.64

Fig. 2 Performance comparison of three different mel-frequency bin values, against the unseen dataset.



	Accuracy	Precision	Recall	F1 Score
250ms	86.39344262	86.66823974	86.01873536	86.3422661
500ms	91.19633099	88.6286279	94.51990632	91.47951571
1s	93.93832943	91.97553882	96.2763466	94.07681452

Fig. 3 Performance comparison of three different segment lengths, against the unseen dataset.



	Accuracy	Precision	Recall	F1 Score
1s (no overlap)	93.93832943	91.97553882	96.2763466	94.07681452
1s (50% overlap)	91.52029664	89.75595171	93.7392662	91.70437406

Fig. 4 Performance comparison of having overlapping training data or none, against the unseen dataset.