SATNAC: Exploring CNN-based automatic modulation classification using small modulation sets

Andrew Oosthuizen

Overview

- Related work
- Experimental setup
- Analysis and results
- Conclusion
- Questions

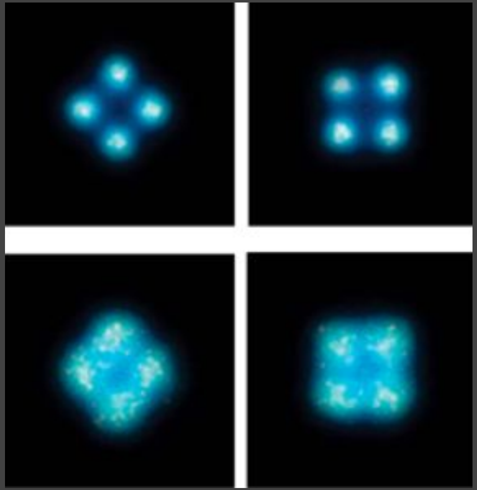
Related work

Presentation methods:

- Constellation diagrams
- Data points

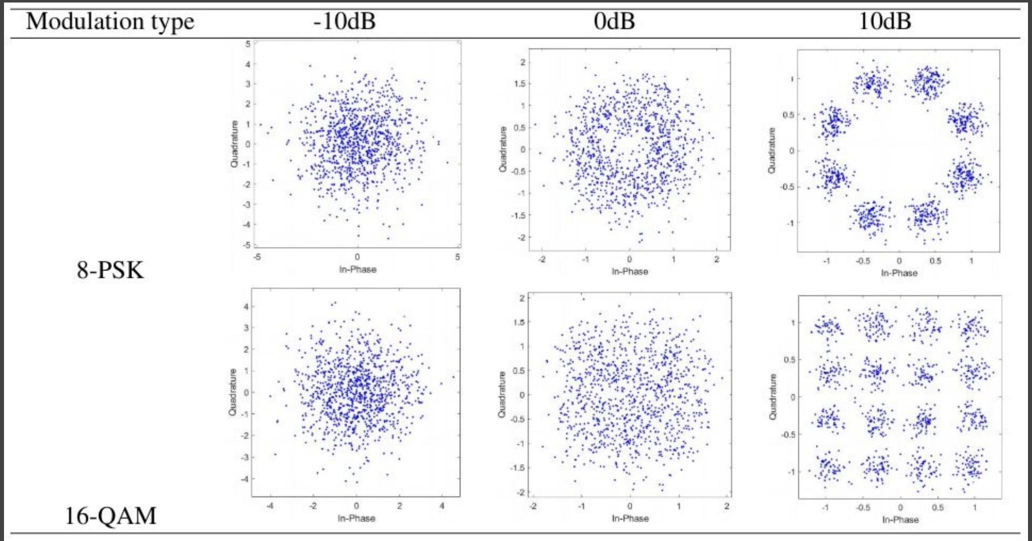
SNR input:

- No SNR input
- SNR estimation
- Multi-label classification



Data set

- Phase shift keying (m=8)
- Quadrature amplitude modulation (m=16)
- -15dB to 5dB bit normalised range
- Each sample contains 1 024 data points
- Presented in a 2 x 1 024 array



Modulation	**Training**	**Validation**	**Evaluation**
8-PSK	1 000	500	1 000
16-QAM	1 000	500	1 000
Total	**42 000**	**21 000**	**42 000**

Approach

Architecture:

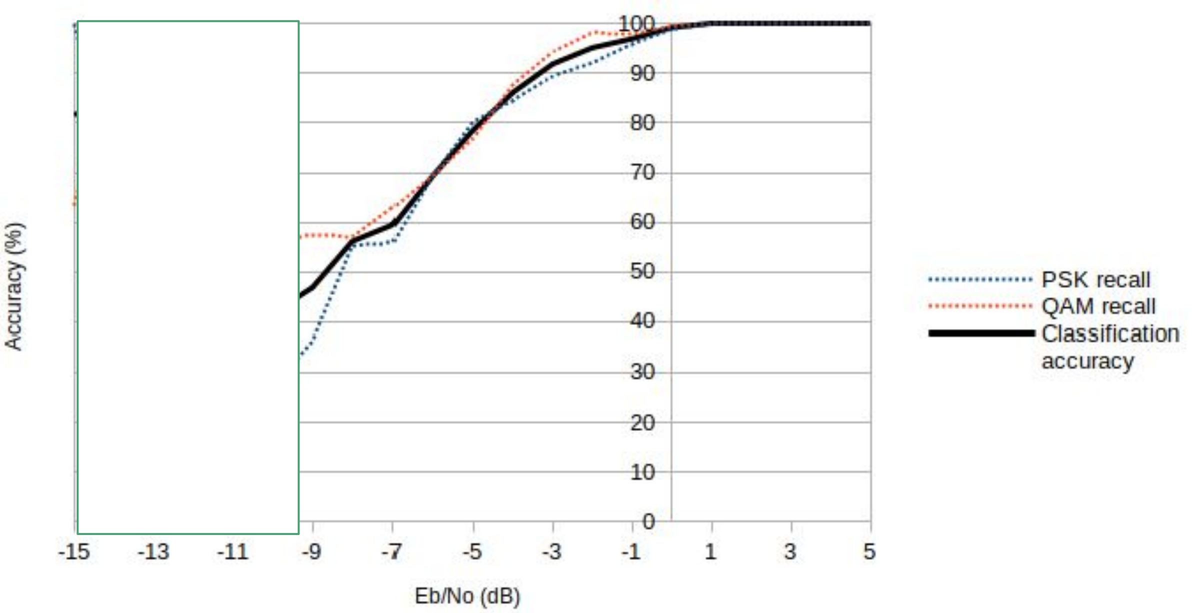
- CNN (stride=1), Dropout, Batchnorm, Maxpool (stride=2)
- CNN (stride=1), Dropout, Batchnorm
- Linear layer (width=100)
- Output (2 classes)

Sweep:

Learning rate, dropout, kernel size, batch size, weight decay

From the search results the best performing hyperparameters will be tested on the evaluation set.

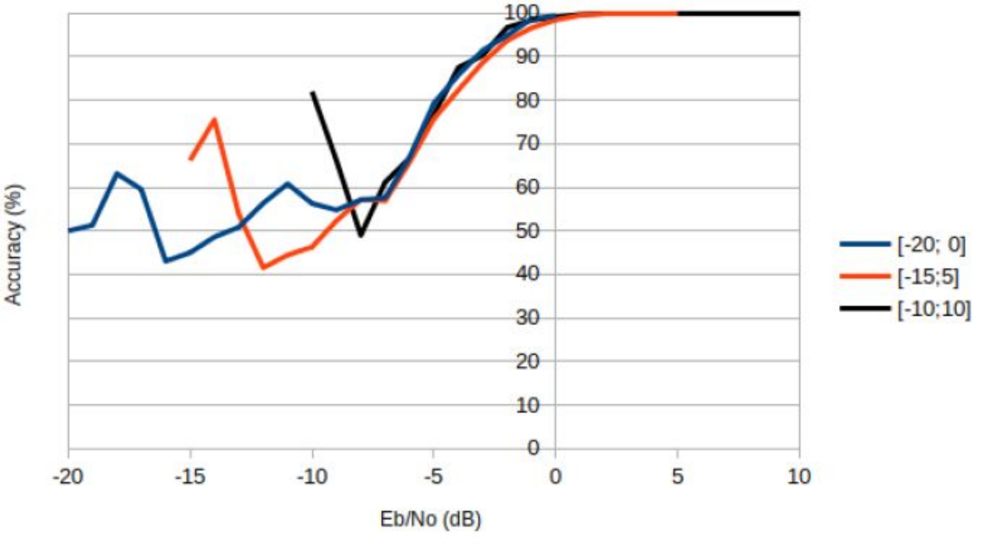
Baseline



Parameter	Value
Learning rate	0.0001
Batch size	32
Dropout	0.5
Weight decay	0.01
Kernel width	36

Range analysis

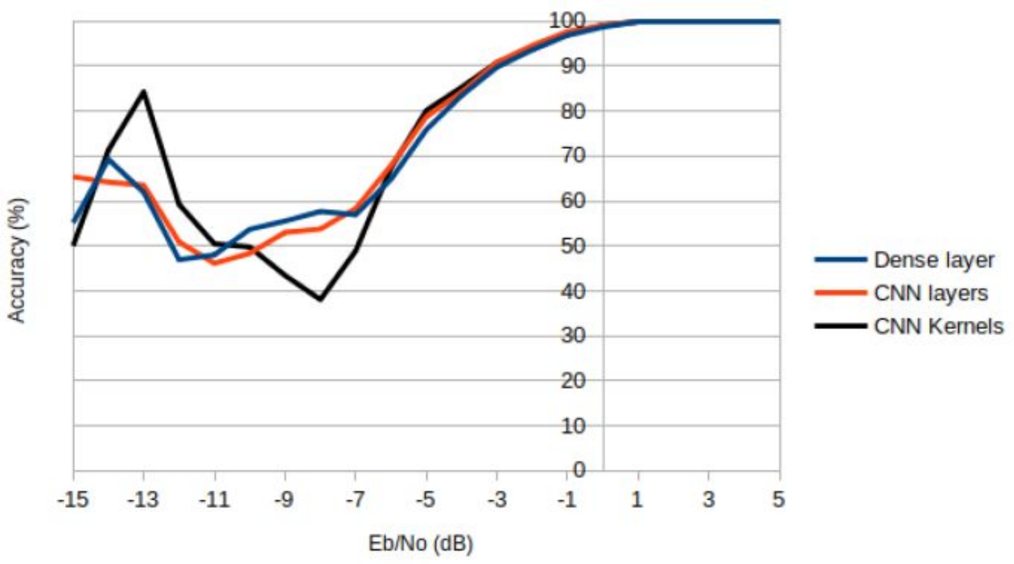
To ensure that the artifact does not occur at a specific SNR level the baseline network was trained and tested over three different SNR ranges.



Architecture analysis

The architecture was adapted by:

- Increase dense layer to 1 000
- Add extra convolution layer
- Increase kernels to 512



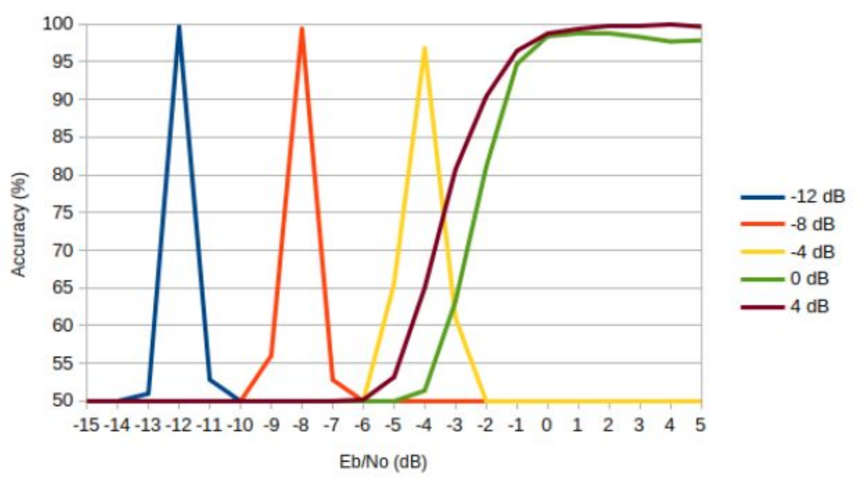
What does this analysis tell us?

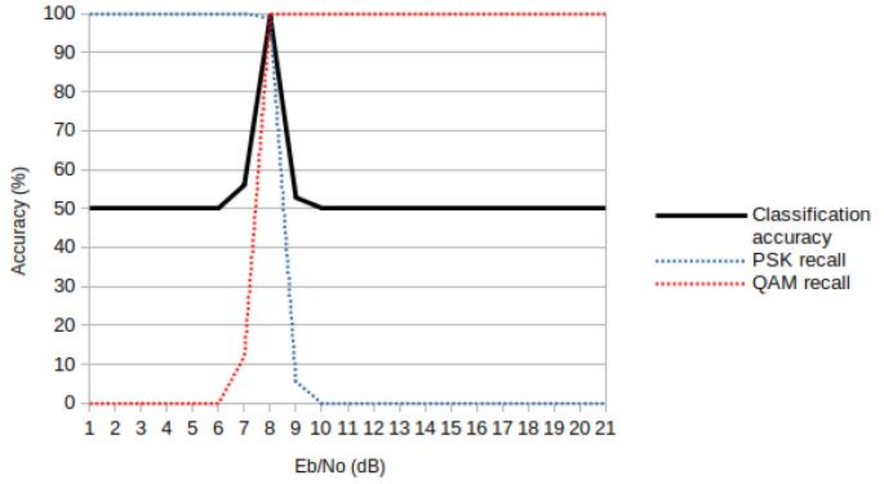
The artifact is still observed after changing the noise range and adapting the architecture which means:

- The artifact is not connected to any specific SNR level, thus not data related.
- The artifact is unrelated to the complexity of the network.

SNR specific training

The networks classification ability is tested by training on a specific SNR level and testing the network over the entire range.



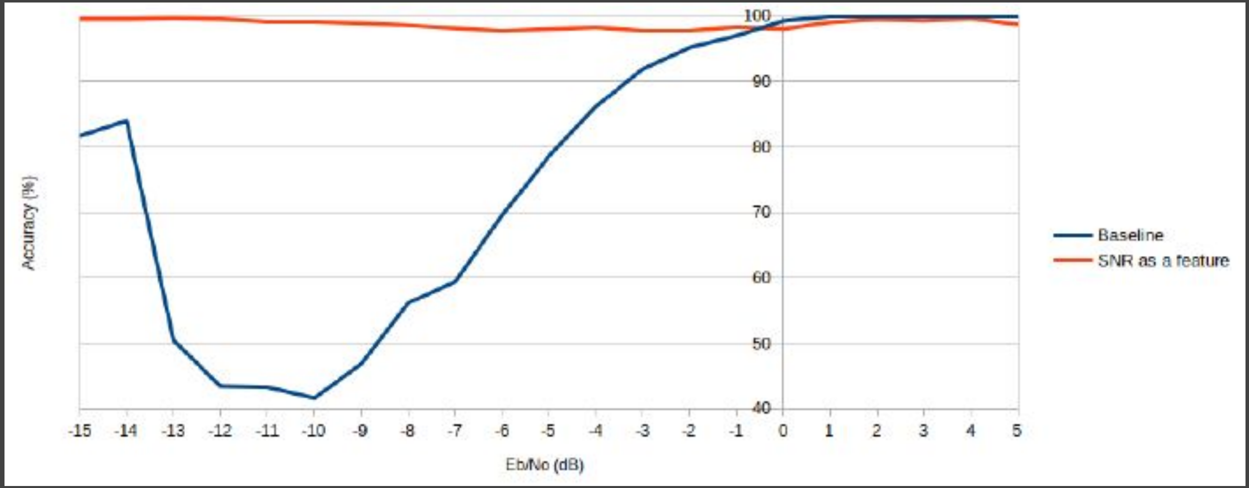


SNR as a feature

- After identifying that the network is able to classify at low SNR levels it was decided to change the architecture to a multi label classifier.
- Oracle SNR data would be presented as labels.
- The convolution layers would act as feature extractors, while the linear layer would be responsible for extracting necessary features for each SNR level.

Improved network

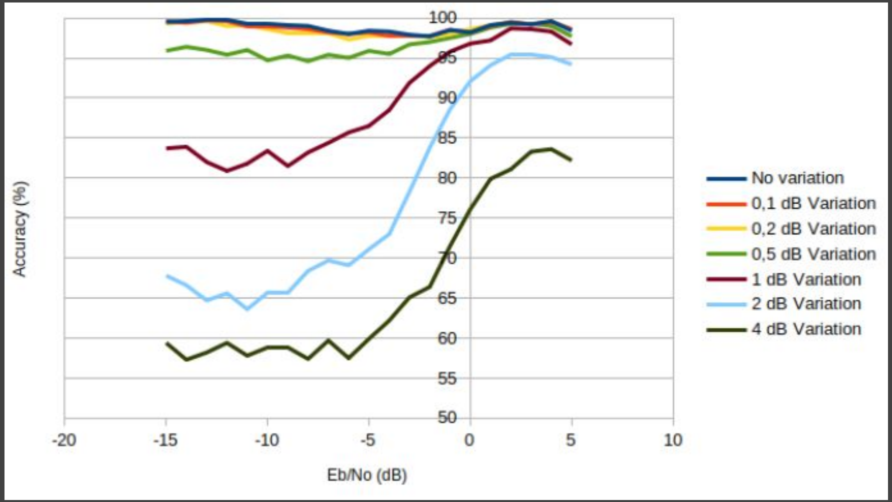
By providing the oracle SNR values at the linear layer as a feature we see a significant improvement in overall accuracy of our network over the entire range.



Sensitivity analysis

Inaccurate noise estimation greatly impacts the network when the error grows larger than 1 dB as the provided noise level moves into the neighboring class.

Even with inaccurate estimations the artifact is no longer observed.



Conclusion

- An artifact was discovered in high noise environments for binary modulation classifiers.
- The discovery of high classification accuracies on single SNR level trained networks prompted the inclusion of SNR information as a feature.
- It was found that the SNR as a feature network performed exceptionally well if accurate noise data was provided.
- We found that CNNs can classify at very high noise levels, but generalises poorly when the amount of noise changes.

Questions

1. Could this high classification accuracy between two modulation types be used to create a decision tree for a larger modulation class pool?
2. Is the knowledge that deep learning can classify such noisy data at the cost of generalisation known or useful?
3. Is there a training method that minimises self optimisation?