An Investigation into the Application of Neural Networks to Underwater Acoustic Communication Signal Recovery

1st Tim Smith
Faculty of Engineering and Applied Science
Memorial University
St. John's, NL, Canada
tims@mun.ca

***Abstract*—Acoustic waves can travel long distances underwater and are often used for the transmission of signals among networks of underwater devices. Signal recovery in these networks can be challenging as the acoustic channel traversed by the signal can alter it significantly. Phenomena such as background noise, reflections from surfaces, and multi-path propagation need to be identified and filtered at the receiver. Additionally, the channel is not constant and the resultant effects on the signal can vary greatly with things such as ice formation at the surface, changing current speeds, and water temperature.**

Traditionally, such systems use a mixture of methods to decode the transmissions at the receiver. Pilot signals between devices are used to estimate channel interference. Individual packets may contain complex preamble bits to attempt to detect problematic sequences and correct them at the receiver. There can also be multiple stages of filtering, tuned for the specific deployment environment and communication protocol in the receiver. These methods can be effective, but they require complex hardware and must be designed specifically for each communication platform.

The non-linear behaviour of the acoustic channel makes this problem a good candidate for deep-learning methods. A trained neural network that could classify sequences of data within an underwater communications environment could hold many advantages over traditional methods of channel decoding. A neural network could be trained on data from different protocols, modulation schemes, carrier frequencies, and deployment environments. This would create a solution that could be used across multiple platforms without requiring specialized hardware or significant redesign.

This paper includes a design and simulated testing of such a neural network. A bidirectional long short-term memory (LSTM) network is utilized to perform classification on a set of simulated received signals. In comparison to traditional demodulation methods, results show that the neural network performs significantly better on the test set.

***Index Terms*—deep learning, neural networks, long short-term memory networks, ocean acoustics, underwater communications**

I. Introduction

Signal recovery in underwater wireless communication networks has many challenges. Electromagnetic (EM) waves do not travel very far underwater while acoustic waves can travel great distances. For this reason, acoustic waves, particularly those between 100 Hz and 20kHz, are used in most underwater networks. The acoustic receiver must be able to interpret the transmitted signal after the signal has been summed with all the effects of the acoustic channel.

Major sources of channel interference include attenuation, multipath wave travel, the Doppler effect, reflections, and environmental noise. Reference [1] contains a comprehensive explanation of each of these factors and their individual complexities. The ever-changing nature of the marine environment can mean that the effect these parameters have on the signal can vary from one transmission to the next. This makes traditional methods of using pilot signals and channel-estimation less effective in underwater networks.

The idea of using neural networks in digital communication systems is not a new one. Reference [2] contains a survey of all the active areas of research in this field from 1997 and is quite extensive. In the past twenty years, the cost of processing power has decreased significantly making it feasible to apply neural networks to new areas. In 1997 the idea of using neural networks to decode messages in a compact, battery-powered, underwater receiver may have seemed implausible. The modern onset of efficient, high-speed microcontrollers makes it a more realistic option today.

Traditional methods used to decode messages at acoustic receivers often include a combination of techniques. Pilot signals are sent between devices to perform channel estimation while individual packets contain known preamble bits so that problematic sequences can be filtered by the receiver. A recent design of an underwater Orthogonal Frequency-Division Multiplexing (OFDM) network is included in [3]. It includes field testing results using cutting-edge signal recovery techniques . The bit errors are low, but each packet requires four seconds of silence, significantly impacting the transmission rate.

A study and simulation of the use of neural networks to decode transmissions in noisy, electromagnetic, wireless communication systems is contained in [4]. This paper states favourable results with less channel estimation and preamble bits required to achieve similar recovery rates to traditional systems. The benefits of such an improvement would be greater in an acoustic system, in which the travel speeds and bandwidth are much lower than EM networks. Reference [5] applies recursive neural networks to a chemical, pH-based, communication link with very promising experimental results. The chemical communication channel is also non-linear and unpredictable, but its specific behaviour is very different from the acoustic channel.

This work attempts to combine ideas from each of these references into a study on the direct application of neural networks to underwater communication links. Two sets of simulated received data packets are used to train and test a bi-directional long short-term memory network. The network packets simulate data received over an Amplitude Modulated (AM), On-Off Keying (OOK) based network that is suffering from a wide range of common types of acoustic channel distortion. Although OOK is not commonly used in underwater acoustic networks, it serves as a suitable test in determining the ability of the neural network to decode heavily distorted transmissions. Retraining and testing of the network using more commonly used modulation methods, such as OFDM, is reserved for future work.

The paper is organized into six sections. The first two provide a brief background on the effects of the underwater acoustic channel and recursive neural networks, respectively. Section four provides an overview on the simulated network. Section five is a detailed description of the LSTM network design used, including several iterations and parameter changes utilized during development. Section six provides the results obtained and includes a comparison of the classification accuracy rate from each of the different LSTM network designs and a baseline, traditional, AM demodulation method. The conclusion section includes some final thoughts on the results and identifies the potential for future work in this area.

II. Effects of the Underwater Acoustic Channel

In wireless subsea networks, all messages between devices must pass through the underwater acoustic channel. The channel transforms the message to a varying amount depending on the surroundings, distance, depth, height from the ocean floor, current speed, temperature, surface waves and relative receiver/transmitter motion. Many of these factors are time-varying, which can make it difficult to precisely measure the effect the channel will have on the signal.

Acoustic reflections and multipath interference are two of the most difficult phenomena to filter out. Multipath propagations are caused by temperature and pressure variations in the acoustic path. Copies of the original signal will travel at slightly different speeds and appear to the receiver as slightly out-of-phase variations. Multipath propagation waves can cause severe attenuation or cancellation in extreme cases. Fig. 1 shows an example of a single multipath propagation wave.

Reflections consist of attenuated copies of the transmitted signal being seen at the receiver as they bounce off nearby surfaces. These could include the surface of the water, the seabed, surrounding objects, or moving particles in the acoustic path. The timing and amplitude of these reflections will depend on the distance of the receiver from the objects and their acoustic

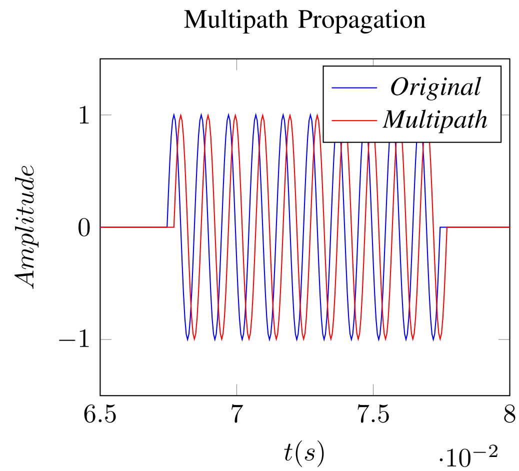

Fig. 1. Multipath Propagation Example with One Additional Path.

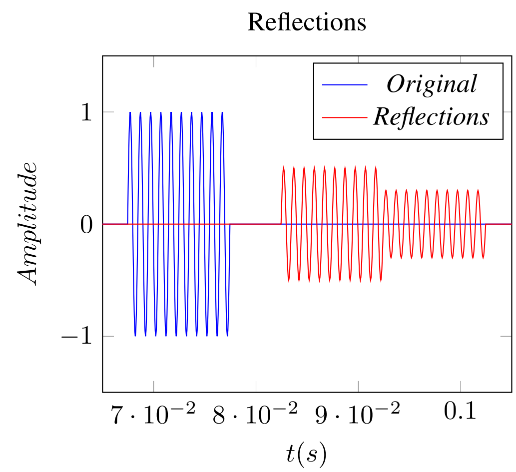

Fig. 2. Reflection Example with Two Reflections.

reflecting properties. Fig. 2 shows an example of the effect of reflections on a signal.

Environmental noise is ever-present in the underwater environment. The ability of acoustic waves to travel long distances means that a non-participating acoustic source many kilometres away can interfere with the signal. In addition to this the marine environment has many naturally occurring sources of sound such as ice-formation/cracking, marine life, and the sounds created by water motion. In addition to this the signal will be attenuated by a factor inversely proportional to its frequency and directly proportional to the distance travelled. The combination of attenuation and signal noise can mean that the receiver is sometimes operating in very low Signal-to-Noise Ratio (SNR) environments. Fig. 3 shows an example of the effect of noise on an acoustic signal.

The Doppler effect can alter the perceived frequency of the signal at the transmitter. It is created by relative motion between the transmitter and receiver. It is a major factor in moving systems, where the receiver may be on a moving vessel and the transmitter is stationary or another moving

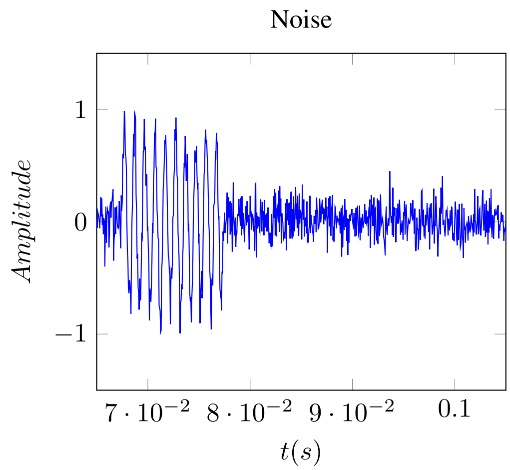

Fig. 3. Noise Example with an SNR of 1.5.

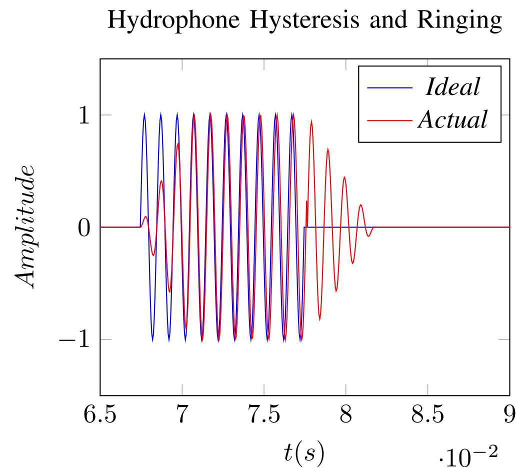

Fig. 5. Hydrophone Reaction Delay

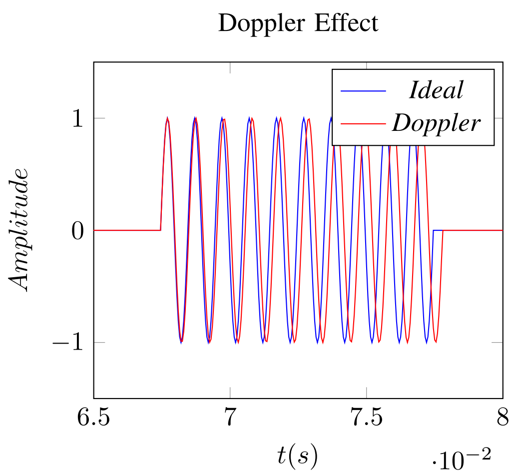

Fig. 4. Doppler Effect with $\Delta v_{TR} = 50m/s$

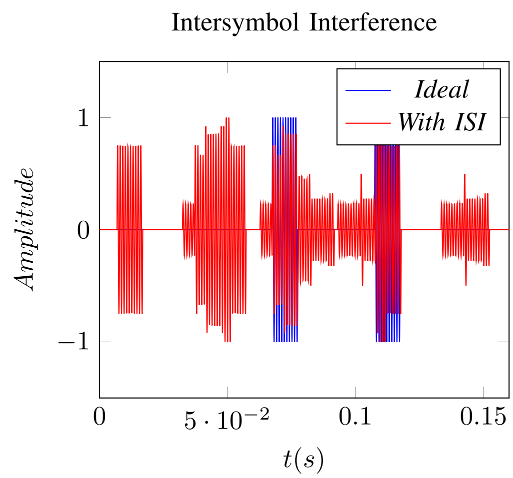

Fig. 6. Intersymbol Interference in a Reflective Environment

vessel. It can also affect the signal in stationary systems, being caused by motion of the receiver or transmitter with the water at its location. The Doppler effect will cause problems in communication configurations with strictly enforced carrier frequencies and tight packet windows as the signal will stretch or compress – effectively changing the carrier frequency and data rate. Fig. 4 shows a severe example of the Doppler effect.

In addition to these common effects of the acoustic channel there are several other, practical factors to consider in an underwater network. The transmissions are typically sent through piezoelectric hydrophones which suffer from hysteresis upon excitation. This means that the transitions from 0 to 1 and 1 to 0 will not be instantaneous but will be enveloped as the hydrophone begins to move from rest or stops after being in motion. Fig. 5 shows an example adjusted to model the response of an actual hydrophone to electrical pulses.

Inter Symbol interference (ISI) is also a major factor in channel distortion. Reflections from the previous packet will often appear in the current packet window, making it appear as a completely different sequence to an untrained receiver. Fig. 6 shows an example of inter symbol interference in a reflective environment.

These are the major components of distortion at the receiver in an underwater network. Some of these factors are unique to marine transmissions while some can also appear in EM messages. The limited bandwidth and long travel times of the acoustic waves make interpreting the correct packet through this distortion even more critical however, as the employment of communication retries is much more costly to the network throughput.

III. Long Short-Term Memory Networks

A LSTM network is a type of neural network often used to classify time-series data. A standard LSTM network consists of input, hidden, fully connected, Softmax, and classification layers. A very basic example is shown in Fig. 7.

The inputs connect to the LSTM layer, which is connected to the fully connected layer. The fully connected layer provides a neuron for each class that uses all information

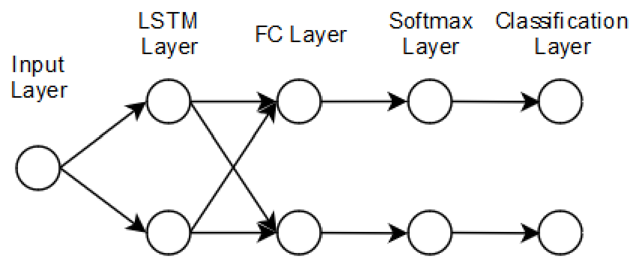

Fig. 7. A Basic LSTM Example

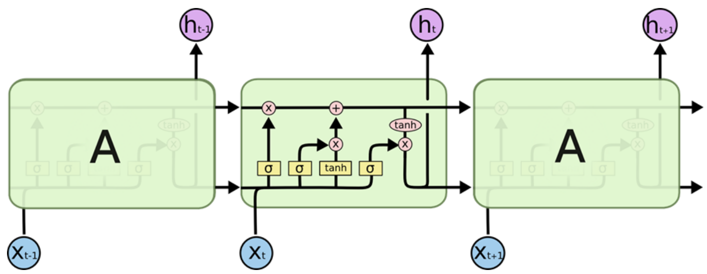

Fig. 8. A Detailed View of a LSTM Unit

from the LSTM layer to make a classification. The Softmax layer converts the output of the fully connected layer into a probability that the input belongs to each class. Finally, the classification layer chooses the class with the highest probability and classifies the input accordingly.

Unique to the LSTM, each node in the LSTM layer contains four interacting cells. These cells work together to form an intelligent memory at each neuron that uses past and present input values to extract its feature. Fig. 8 shows a detailed view of a single LSTM neuron.

These four cells perform the following functions:

1. The input is passed to a sigmoid forget gate that decides what past information to forget based on the new input.
2. A sigmoid input gate cell looks at the current input vector and decides which values will be passed to the new state.
3. A candidate for the updated state is created based on the input, via a tanh function. This is combined with the output from the input gate cell to determine which values from the candidate state will be stored as the new state.
4. Another sigmoid cell decides which values from the current state to output (pass to the fully connected layer).

This complex interaction between past and present inputs make the LSTM network adept at detecting complex patterns in sequences such as time-series signals. If the length of the sequence is known, a bi-directional LSTM network can be used to analyse the sequence in the forward and reverse directions. This was the option chosen for this design because the packet length is always a fixed set of samples.

IV. Simulated Network Design

A simulated network needed to be created to generate the training and test sets of received packets. The goal was to design the simplest, non-trivial, network that could test the concept. For this purpose, the OOK modulation scheme was used. A ‘1’ on the network is represented by transmitting a number of pulses at the carrier frequency (f_c). A ‘0’ on the network is represented as silence from the transmitter. A single packet consists of a start bit, two zeros, two bits of data, and a trailing zero. There are four different packets possible (data bits enclosed in square brackets):

- 0: 100[00]0
- 1: 100[01]0
- 2: 100[10]0
- 3: 100[11]0

The following parameters define the network:

- Data Rate: $f_d = 100Hz(bps)$
- Carrier Frequency: $f_c = 1kHz$
- Sample Frequency: $f_s = 20kHz$

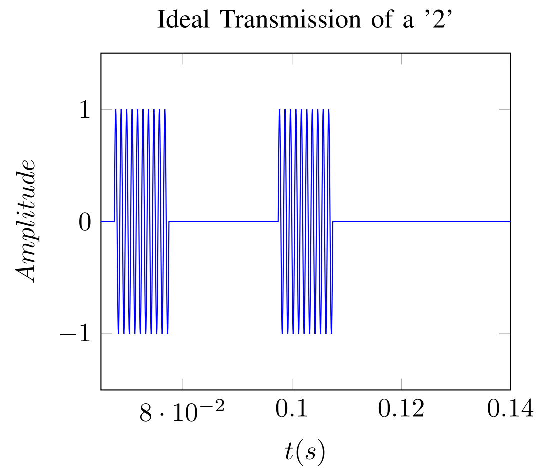

Fig. 9. Single Network Packet Containing a 2

At these values, each ‘1’ contains ten pulses of the carrier frequency. Each packet includes 1200 data points, 400 of which enclose the data. As an example, a transmitted ‘2’ is shown in Fig. 9.

The following transformation is performed on each data packet to simulate the acoustic channel (the values in square brackets are randomized between those two limits):

- Travel Time and Varying Speed of Sound in Water: Delay the signal based on a distance of 100m and a $v_{sound} = [1408\ 1556]\ m/s$.
- Doppler Effect: Alter the carrier of the received signal by the following formula, where the transmitter and receiver velocities are set at $v_T = [-25.0\ 25.0]\ m/s$ and $v_R = [-25.0\ 25.0]\ m/s$.

$$f_{c,real} = f_{c,ideal} \frac{v_{sound} + v_R}{v_{sound} + v_T}$$

- Hydrophone Ringing: Add a fade out time at each ‘1 to 0’ transition of $[2\ 4]\ ms$.
- Hydrophone Hysteresis: Fade in each ‘1’ preceded by a ‘0’ over a period of $[1\ 3]\ ms$.

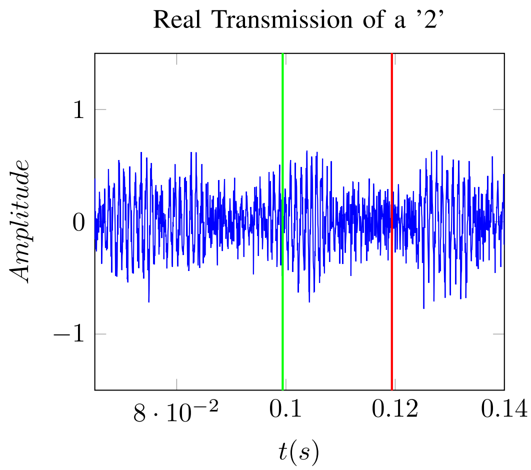

Fig. 10. Single Network Packet Containing a 2 After Transformation

- Inter symbol Interference: Add a preceding data packet containing data: $[0\ 3]$.
- Attenuation: Multiply the signal by an attenuation factor of $[0.3\ 0.7]$.
- Multipath Propagation: Add $[1\ 3]$ copies of the original signal shifted by $[-T_c\ T_c]$ s and attenuated or amplified by a factor of $[0.9\ 1.1]$.
- Reflections: Add $[1\ 3]$ copies of the original signal. Shift each copy by $[nT_d\ 3nT_d]$ s and attenuate it by a factor of $[0.6\ 0.8]$–$0.2n$. Where n is the reflection number.
- Noise: Add white Gaussian noise over the entire signal with an $SNR = [1.5\ 5.0]$.

To generate the training and testing data, two sets of 2000 such transmissions were generated. Each contained an even distribution of five hundred 0s, 1s, 2s, and 3s. These randomized sources of noise generated sets of signals with moderate to severe distortion. A typical result of the signal transformation is shown in Fig. 10. Compared to the ideal signal in Fig. 9, it is clear that the receiver could have difficulty interpreting the input as a '2'

Prior to storing these samples in the test and training sets some pre-processing is performed. First, each sample is scaled to between -1 and 1. Then each sample is clipped to contain only the two data bits. This is done by iterating through the sample from the point after the previous packet ended until the first point greater than 0.5 was detected. The sample is taken as being the 400 data points starting at three data periods after the start bit. The green and red lines in Fig. 10 enclose the data sample, as detected by this method. This sample is stored as vector x to be input into the neural network. If i_{sbd} is defined as the index of the sample at which the start bit is detected and N_{bit} is the number of samples per bit, the formula for x is as follows:

$$x = Received\ Signal[i_{sbd} + 3N_{bit} : i_{sbd} + 5N_{bit}]$$

A single additional feature is stored for each sample. Each clipped sample is divided into two individual bit periods. The energy in each period is divided by the SNR and stored in an array. The values are duplicated 200 times each and stored in an array that aligns each value with its associated data points in x. This is done to prepare it for input into the neural network.

$$E_x = \frac{1}{SNR}\Big[\sum_{i=1}^{N_{bit}}[x_i]^2 \quad \sum_{i=N_{bit}+1}^{2N_{bit}}[x_i]^2\Big]$$

This network makes several simplifications of the effects of the acoustic channel. The Doppler effect is applied uniformly to the original signal, the multipaths, and the reflections. In a real environment each of these things may be affected differently. Additionally, the noise added to the channel is white Gaussian, in a real deployment it is a combination of ambient and site-specific noise.

V. LSTM Network Design and Training

This network design and sample generation reduces the problem to a time-series classification problem. As such, a classification network is required, capable of dividing the input sets into classes based on the two defined features. A long short-term memory neural network is used, as the input data is in time-series form. The data pre-processing means that each input sample is the same, fixed length so the network can be bidirectional. The architecture of the LSTM is as follows:

- Sequence Input Layer (Size: 1 or 2)
- Bi-Directional Long Short-Term Memory Layer (Size: 150)
- Bi-Directional Long Short-Term Memory Layer (Size: 150). Only in 2F2L.
- Fully-connected (Size: 4)
- Softmax Layer (Size: 4)
- Classification Layer (Size: 4)

Three different LSTM network architectures were trained and tested:

- 1F1L: Input Size: 1 (only x is input), 1 Bi-LSTM layer
- 2F1L: Input Size: 2 (x and E_x are inputs), 1 Bi-LSTM layer
- 2F2L: Input Size: 2 (x and E_x are inputs), 2 Bi-LSTM layers

Each network was trained and tested using the same sets of training and test data. The initial weights were seeded with the same values to ensure that they did not affect the outcome. The following parameters were used for training in all cases:

- Training Size: 2000
- Weight Update Method: Adaptive Moment
- Epochs: 25
- Mini Batch Size: 40
- Gradient Threshold: 1

VI. Baseline Method

To compare the results of the neural network classifier, a baseline method was created. This method uses more traditional means to classify the data packets. The data sample is first demodulated using a low-pass filter. After demodulation, the data sample is divided into two sections (one for each

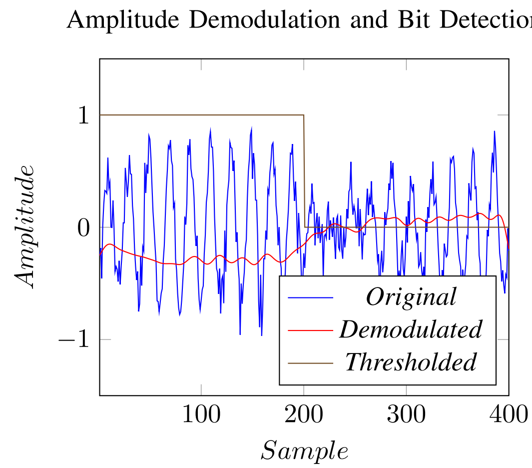

Fig. 11. Noisy Data 2 Packet Decoded Using Baseline Method

bit) and the average of each is calculated. The data samples are then constrained to a binary value based on an iteratively optimized threshold. Fig. 11 shows the demodulation and classification of a noisy data 2 sample.

VII. Results

The results are presented in terms of the Bit Error Ratio (BER) for each network. The BER is defined as the number of bits incorrectly identified divided by the number of bits sent. As this does not include the preamble bits, a total of 4000 bits were sent in the 2000 sample test set. Tab. I shows the classification accuracy and BER for the baseline method as well as the three LSTM networks.

TABLE I
Classification Method Accuracy Comparison

Name	Features	Layers	Accuracy	BER
Baseline	-	-	62.55%	21.05%
1F1L	1	1	31.8%	43.9%
2F1L	2	1	85.8%	8.08%
2F2L	2	2	85.65%	8.18%

The 2F1L network performed the best overall. This network used both input features and had a single Bi-LSTM layer of 150 units. With a BER improvement of 12.97% over the baseline method, it correctly identified 519 more bits, successfully classifying 85.8% of the samples.

More testing was performed once the best network architecture was identified, while fixing some of the distortion parameters. This was included to identify which factors most significantly affect the classifier performance. The parameters tested were the Doppler effect (Fig. 12), number of paths (Fig. 13), number of reflections (Fig. 14), and SNR (Fig. 15).

It is evident from the figures that each of these parameters affects the method performance. The Doppler effect has a drastic effect on the performance of both methods. The Doppler effect stretches or compresses the signal, causing a misalignment between the expected bit windows and the actual

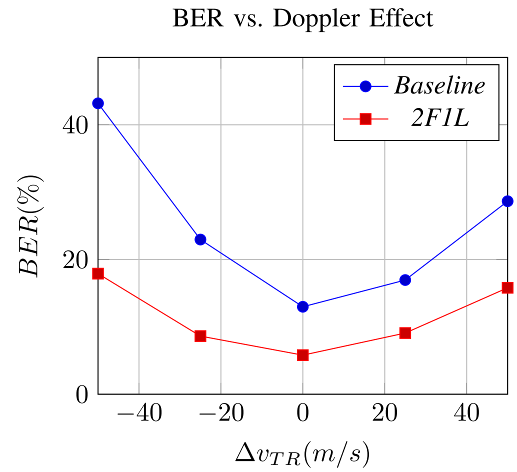

Fig. 12. Baseline and LSTM Performance vs. the Doppler Effect

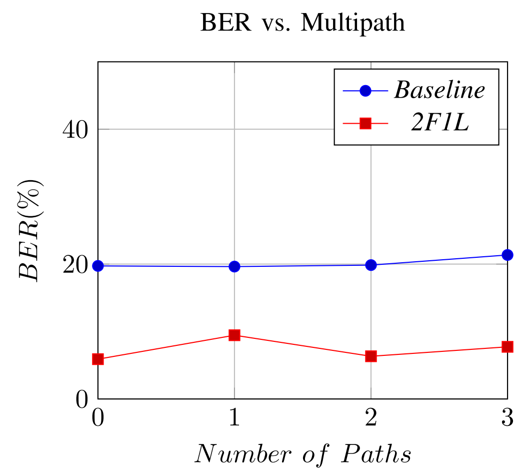

Fig. 13. Baseline and LSTM Performance vs. Multipath Propagation

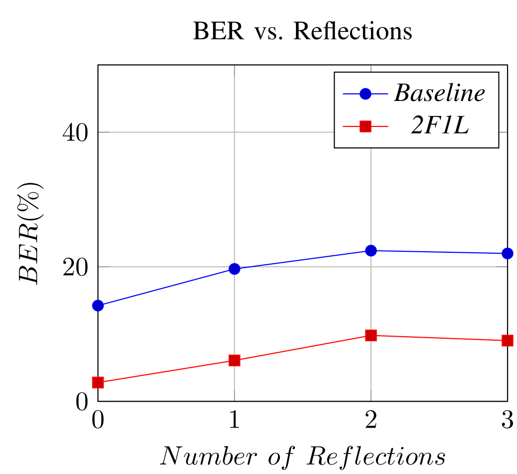

Fig. 14. Baseline and LSTM Performance vs. Reflections

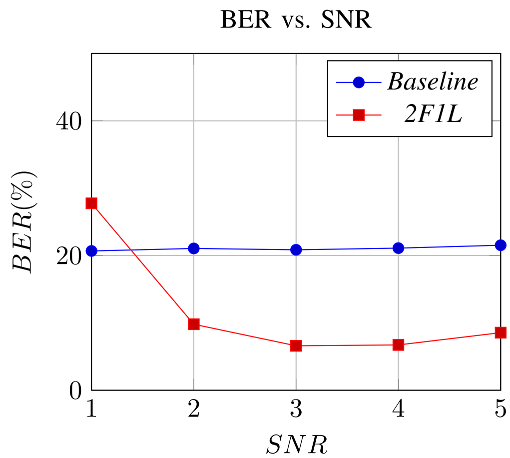

Fig. 15. Baseline and LSTM Performance vs. SNR

transmission. Multipath propagation has little effect on the baseline method, although a single extra path causes a notable increase in the BER of the LSTM classifier. The first two reflections have equally negative effects on both methods while the third reflection has little impact. The LSTM network was more susceptible to changes in the SNR. At 1.0 the BER was very high, although this was outside the possible SNR values in the training data which would have impacted performance. The filters employed in the baseline method make it more resistant to low SNR values although the simplified Gaussian noise model may have contributed to this.

VIII. Conclusion

The LSTM network has displayed a marked improvement over the baseline method in packet identification. This work shows that this may be a feasible method to decode transmissions in an underwater network environment. In addition to the improved performance, it has the added benefit that, with additional training, it could be modified to work with different modulation methods, carrier frequencies, and environments.

The testing used a simple OOK-based network with a data rate of 100 Hz. Subtracting the preamble bits from this data rate gives the network a data throughput of 33 bps. Additional work in this area should expand the network design to faster and more efficient modulation schemes. It would be a simple extension to increase the size of the data frames within each packet and identify larger sets of bits (e.g. an eight output LSTM network could classify packets containing four data bits instead of two). While this would increase the data throughput of the network, there could be a potential reduction in accuracy, as there would be more room in the data window for interference.

References

[1] M. Stojanovic and J. Preisig, ”Underwater acoustic communication channels: Propagation models and statistical characterization,” in IEEE Communications Magazine, vol. 47, no. 1, pp. 84-89, January 2009.

[2] M. Ibnkahla, ”Applications of neural networks to digital communications – a survey”, Signal Processing, vol. 80, no. 7, pp. 1185-1215, 2000. Available: 10.1016/s0165-1684(00)00030-x.

[3] H. Lee, C. Baek, J. Jung and W. Kim, ”A Weighted Turbo Equalized Multi-Band Underwater Wireless Acoustic Communications”, Applied Sciences, vol. 8, no. 10, p. 1711, 2018. Available: 10.3390/app8101711.

[4] H. Ye, G. Y. Li and B. Juang, ”Power of Deep Learning for Channel Estimation and Signal Detection in OFDM Systems,” in IEEE Wireless Communications Letters, vol. 7, no. 1, pp. 114-117, Feb. 2018.

[5] A. Goldsmith and N. Farsad, ”Detection Algorithms for Communication Systems Using Deep Learning”, 2017. Available: https://arxiv.org/pdf/1705.08044.pdf. [Accessed 20 March 2019].