Long-Short Term Memory-Based Application on Adaptive Cross-Platform Decoder for Bit Patterned Magnetic Recording

TEANCHAI CHANTAKIT, CHAIWAT BUAJONG, (Graduate Student Member, IEEE), AND CHANON WARISARN, (Member, IEEE)

College of Advanced Manufacturing Innovation, King Mongkut's Institute of Technology Ladkrabang, Bangkok 10520, Thailand

Corresponding author: Chanon Warisarn (chanon.wa@kmitl.ac.th)

ABSTRACT Dynamic bit encoding and decoding of the magnetic recording process remain a challenge in that the process is restrained by the balance between reading and writing performance of the decoder's bit error rate (BER). Sequential neural networks offer data streamflow for processes to reproduce recoded bits from signal distribution, overcoming the limitation of codeword mapping designed for each specific bit-patterned magnetic recording (BPMR) channel. Here, we implement the vanilla long short-term memory (LSTM) for adaptive modulation decoders in various BPMR channel designs within a single network, which benefits multi-channel decoder calibration tools with the same standardization. Signal information from media readback, a two-dimensional (2D) equalizer, 2D Viterbi, and a 2D soft-output Viterbi algorithm (SOVA) detector is arranged as a tensor that enables sequence-to-sequence bit prediction even with a highly complex data arrangement. Our adaptive model can predict recorded bits from readback with accuracies of approximately 97% for rate 4/5 decoding and 75% for crossing platforms, using a recently proposed single-reader/two-track reading (SRTR) system at an areal density of 4 Tb/in^2 in a signal-to-noise ratio range of 1 to 8 dB. We conducted a BER simulation with the relevant results from conventional decoders and the LSTM model. Ultimately, our approach may demonstrate the limitation of supervised learning designed for BPMR systems and reveal a sequence data focus on LSTM that paves the way for sequential-type, unsupervised, mechanism-based, next-generation magnetic recordings.

INDEX TERMS Long-short term memory (LSTM), supervised learning, deep learning, bit-patterned media recording (BPMR), channel decoding.

I. INTRODUCTION

As a high-storage-density technology within a compact containment, magnetic recording technology has rapidly developed and paved the way to new bit encode-decode, media and material fabrication techniques [1], [2]. In this context, the development of bit-patterned media recording (BPMR) addresses high areal density (AD) with its proximate bit island distance. BPMR also offers substantial inter-symbol interference (ISI), inter-track interference (ITI), and additive white Gaussian noise (AWGN) to limit performance [3]. Aside from a media environment, all of these schemes' improvements benefit from read/write signal path characteristics [4]. Since readback media signals relate to a crucial channel element for minimizing bit error rates (BER), a sufficient technique for high-precision decoding channels with multiple modulation code rates maintains a pivotal role in the future of magnetic recording platforms [5].

Recently, utilizing deep neural networks (DNN) for magnetic recordings [6]–[8] has often been proposed in order to improve BER performance. Several works have attained deep learning sufficient for feasible encoding and decoding channels—for example, a multilayer perceptron (MLP) compared to an equalizer filter, Viterbi detectors, and decoders [8]–[10]. Moreover, many convolution neural network (CNNs) contrivances have also improved iterative decoding systems with parity checks of log-likelihood ratio (LLR) modulators [6], [11]–[14], lowering BER.

The associate editor coordinating the review of this manuscript and approving it for publication was Xi Peng.

DNN decoders have used the sequential feedforward of recurrent neural networks (RNNs) for turbo decoders [15] and LDPC-based long-short term memory (LSTM) to minimize the complexity of neuron parameters and noises [16], rapidly reducing computational time [17]. Moreover, a DNN modification with adaptive training has been proposed for detection channels [8], offering sufficient BER with data training at a specific SNR [18]. However, codeword/network modifications defined for particular processing has limited these designs' flexibility for use with various systems, sometimes preventing real hard disk drive integration. Also, all existing DNN in magnetic recordings has relied on supervised learning restricted to each BPMR channel, which offers no notable difference from a conventional magnetic recording. Codewords are still necessary for new systems despite channels that offer high BER in previous systems. To the best of our knowledge, studies have never before investigated a network's use in different systems with a single network for various channel designs and BPMR global tools for channel calibrations.

Here, we present a transparent LSTM application for modulation decoders that not only can condense AWGN, ISI, and ITI effects but also learns to interpret any signal in reading channels, overcoming the limitation of all BPMR channels that rely on codeword mapping. Thus, this approach offers benefits as a tool to calibrate all proposed channel systems with the same standardization as real BPMR media in the near future. Instead of mapping a codeword to minimize noise, we address the LSTM strategy to reveal an overall signal distribution throughout an entire sector by every read channel. With this application, neural networks can learn any channel processing across different platforms via backup training for desired signals within a backend channel in order to adapt to such a system.

Furthermore, this backend channel has demonstrated both BER enhancements and backend training using the same LSTM network to decode other completely different BPMR channels. The bit prediction from the readback signal utilizes a potentially versatile BPMR with a decoding accuracy of ∼97% for rate-4/5 modulation codes. It continues with the same network used for single-reader two-track reading (SRTR) systems [19], [20] with a validity of ∼75% from an SNR of 1 to 8 dB. Ultimately, our study offers twofold contributions: a new decoder method that freely adapts to any BPMR channel designs from a hard drive processing perspective, and a sequential data scheme for future magnetic recording channel designs in unsupervised learning approaches.

II. RATE-4/5 MODULATION NETWORK DECODER

The header covers five island tracks within a single reader. Three inner tracks are not affected by ITI due to a write header having fully covered bit islands during the read process. The two outer tracks, however, have been partially covered by the header, and they require another two-side track encoding bit in order to rectify bits with the codeword reference table. The signal pattern's compatibility between conventional decoders

and sequence data in an LSTM model addresses a similar capture correlation to other channels [21]. In this way, the configuration of the sequence-to-sequence channel surpasses a signal level at each timestep with only user bit recording, which includes ISI, ITI, media, and electronic interferences.

A. LSTM NETWORK MODEL

The LSTM network is a helpful and straightforward tool to confirm BER performance since the strategy does not rely on a specific input length dimension, addressing a real-world problem. The generalized strategy can reasonably confirm the limitation of channel development based on a supervised mechanism. Hence, a sequential data flow scheme's sufficient bit decoder versus a codeword-dependent design allows for novel magnetic recording channels to expand unsupervised learning–based BPMR designs.

Interestingly, the LSTM offers a continuous data flow independent of the particular data length. Thus, it can handle not only substantial data from a continual readback signal but also memorize multi-track signal signatures in a similar word translation concept [22], which is suitable for classifying any signal pattern in the read channels, such as equalizers as well as hard and soft detectors. In this way, the LSTM-based channel has the potential to simultaneously learn various input signatures and predict only the same encoded bits pattern back into the LSTM cell to the next sequence. This model signature also enables the LSTM to adapt to any channel in such a BPMR system within a single, fully trained network.

Based on the sequential flow concept, the bit classification depends on neurons' decisions at each timestep that neglects all sector sizes in BPMR. Accordingly, we determined a simple LSTM model since it is more convenient to replicate feature engineering using different BPMR systems than to use a highly complex model unable to learn data signatures from such a new channel. As Fig. 1a depicts, the applied network-based decoder relies on RNNs architecture, followed by an LSTM model [23]. Each LSTM cell in Fig. 1b contains a memory cell state which enables a nonlinear activation function for input, output, and forget gates, as Fig. 1c shows. The LSTM cell components can be expressed by [23]–[25]:

$$i^t = \sigma \left(W_i x^t + U_i \left[h^{t-1}, C^{t-1} \right] + b_i \right) \qquad (1)$$

$$f^t = \sigma \left(W_f x^t + U_f \left[h^{t-1}, C^{t-1} \right] + b_f \right) \qquad (2)$$

$$o^t = \sigma \left(W_o x^t + U_o \left[h^{t-1}, C^{t-1} \right] + b_o \right) \qquad (3)$$

$$C^t = \tanh \left(W_c x^t + b_c \right) \qquad (4)$$

where i^t, f^t, and o^t are the controllable outputs, labeled by input gate, forget gate, and output gate, respectively. C^t is the cell input activation vector or modulation gate that defines the cell state to forget memory. x^t is the input tensor at the time instant t with a size of $5 \times 8{,}192$ from R_k, Z_i, Viterbi, or a soft detector. σ is the sigmoid function, W is the synaptic input weight, U is the recurrent weight, and b is the bias vector. h^t and h^{t-1} are the prediction output tensors at the present time and the recurrent input, respectively. y^t is the output tensor

(a) LSTM-based BPMR model

(b) Recurrent neural network decoding architecture

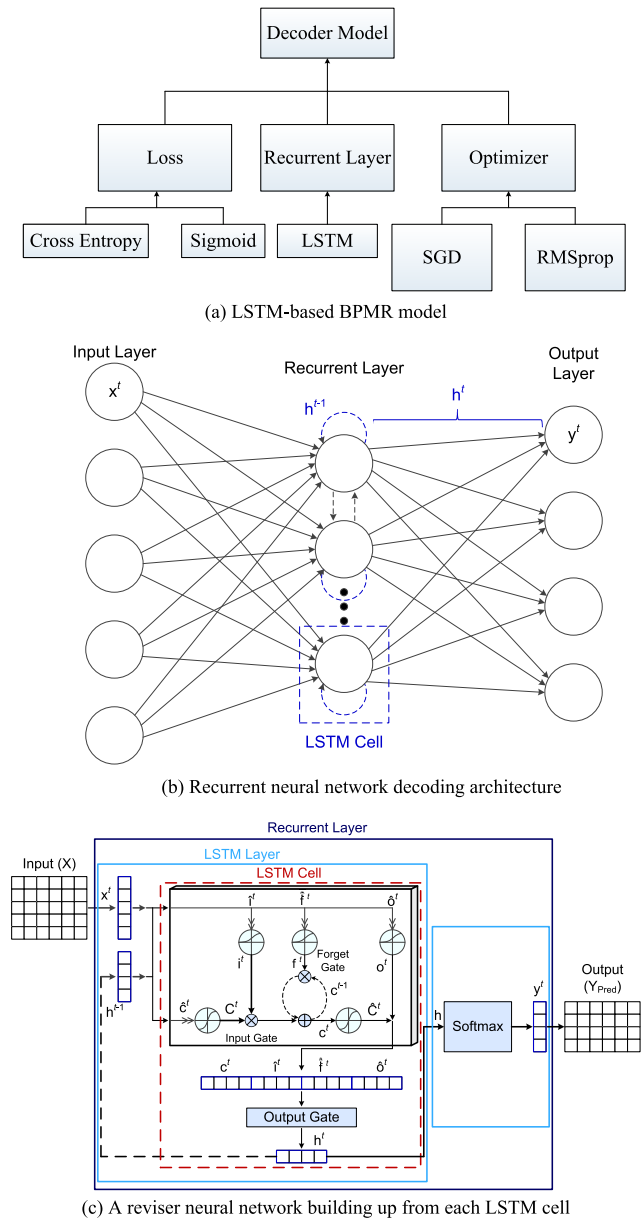(c) A reviser neural network building up from each LSTM cell

FIGURE 1. Implemented five feedforward tracks scheme used for rate-4/5 modulation code. The complete process passes all sequences out to the softmax layer for four-track output.

at each timestep that converts to Y_{Pred} of $4 \times 8{,}192$ via $y^t = \tanh\left(Vh^t\right)$ in a softmax layer, where V is the output weight. The final result is then compared with the encoded user bit for supervised learning.

The RNN output layer contains both the LSTM output vector (h^t) and the cell state vector (c^t) that determines the normalized vector prediction for the corresponding class of binary-shaped data at the output layer as:

$$c^t = f^t \odot c^{t-1} + i^t \odot C^t \qquad (5)$$
$$h^t = o^t \odot \tanh c^t \qquad (6)$$

where $\odot$ is the element-wise multiplications. An overall predicted bit error has evaluated the gradient, affecting the

neuron weight adjustment via a backpropagation algorithm. Thus, model optimization accumulates the error and weight updates through a backpropagation through time (BPTT) algorithm. Moreover, model optimization can be classified more effectively with a recorded bit y^t_{target} for pattern recognition problems when it accounts for parameters in previous timesteps [26]. In this work, BPTT has employed the simple error for a certain time sequence $\delta y = y^t_{target} - y^t_{Pred}$, with respect to the error of the output tensors $\delta h = Vy^t_{Pred} \bullet \delta y$.

For the readback signal, a generalized statistical measurement would address an appropriate change in gradient for various signal signatures from different BPMR designs, allowing for an impartial BER evaluation for each channel. Nevertheless, the same class of input and output data has overfitted a LSTM model. Several attempts at additional training for a default trained network using the same class data in new BPMR systems caused gradient explore within a few iterations, underrating a model for use with a particular system in the same way as conventional codeword mapping.

Since BPTT stores an entire sequence for every timestep, it contains a pre-minimized error for weight gradient, which is denoted as a loss function (L). Specifically, every signal from such a channel is required to categorize each bit island since we must manifest a signal for every channel design. Thus, we chose a well-known loss function over time sequence (T) for an encoded bit (N), as [27]:

$$L = -\frac{1}{N} \sum_{i=1}^{N} \sum_{t=1}^{T} \left[y^t_{target} \log \left(\frac{1}{1 + \exp\left[-y^t_{pred}\right]} \right) \right.$$
$$\left. + \left(1 - y^t_{target}\right) \log \left(1 - \frac{1}{1 + \exp\left[-y^t_{pred}\right]}\right) \right] \qquad (7)$$

The vanilla BPTT offers a more natural way to value overall performance in handling a variety of signals. The gradient $\partial L / \partial y^t = \tanh\left(y^t_{Pred} - y^t_{target}\right)$ is set at the softmax layer.

An RMSProp optimizer displays an appropriate method for extensive evaluation [28] and sufficient computational processing related to prediction accuracy [29]. This selection can equally iteratively improve parameters for the cross-entropy loss function relevant to the BPMR channels with the same standard. The RMSProp optimizer gives an update step that normalizes the gradient as [30]–[32]:

$$\theta_{t+1} = \theta_t - \frac{\alpha}{\sqrt{S_t} + \varepsilon} \frac{\partial Y}{\partial \theta_t} \qquad (8)$$

where

$$S_t = \gamma S_{t-1} + (1 - \gamma) \left[\frac{\partial Y}{\partial V_t}\right]^2 \qquad (9)$$

We defined values $\alpha = 0.001$, $\gamma = 0.9$, and $\varepsilon = 10^{-9}$, which provided the most potent computational speed and a precise classification for our BPMR datasets.

As we mentioned in the introduction above, our study on sequence data for magnetic recordings denotes an alternative

BPMR channel design. The codeword-independent method could be further developed with algorithms that match our approach to BPMR, such as a multi-view linear discriminant analysis network [33]. This cross-view feature classifies all data samples through the supervised mechanism for projecting recorded bits, which relates to our focus on various signal signatures for every readback BPMR channels. In addition, the clustering-based model—which includes the structured autoencoders for deep subspace clustering [34]—feeds a direct nonlinear signal and globally reconstructs each bit track locally. It would be as well to cluster unlabeled data with the deep clustering of prior sample-assignment invariance [35]. Furthermore, occurrences of ISI, ITI, jitter, and media noise can be compacted as adversaries when applied to a real hard drive. The multimodal adversarial network can utilize general discriminative recorded bit projection for cross-modal retrieval [36] instead of a random design code rate (e.g., 4/5, 5/6, and 7/8) for the most sufficient BER performance.

B. CONVENTIONAL AND LSTM-BASED SYSTEM

In this work, we implemented the LSTM with our previous work on rate-4/5 modulation codes [37] without a low-density parity-check (LDPC). Fig. 2 shows a comparison between conventional and LSTM channels schematically. The $1 \times 32{,}768$ data length of one sector of sequence $\hat{a}_k \in \{-1, 1\}$ is split into the user bit track $\hat{a}_k \in \{-1, 1\}$ of $4 \times 8{,}192$, which are later encoded into the bit encoder of $5 \times 8{,}192$ (Raw C_k) for recording on BPMR media. The readback signal contrives the same encoded bit of $5 \times 8{,}192$ (R_k) with AWGN from a single header reading. Through an equalized sequence, a 2D equalizer (Z_i) iterates the rendered sequence with 2D Viterbi and 2D-SOVA detectors. By decoding information, the hard information is decrypted directly with the Viterbi decoder (VB), followed by soft information with the LLR decoder (SO), to reproduce the original $\hat{a}_k$ sequence. We investigated the system with 4 Tb/in^2 AD. A bit island of 9.75 nm × 10 nm in size is arrayed on a bit period of $T_x = 12.5$ nm and a track pitch of $T_z = 12.5$ nm.

To decode with the LSTM channel, the finalized trained network replaces the conventional 2D equalizer channel (RNN) as well as the hard-4/5 (RV) and soft-4/5 (RS) decode channels. In addition, we obliterated five inner tracks for readback signal and used the comprehensive network to replicate an authentic $\hat{a}_k$ sequence directly from the readback signal decoder (RM) without using a codeword in any part of our model.

III. EXPERIMENT AND ANALYSIS

The memory-based network can perform effectively without high processing resources, making LSTM implementation suitable within a hard drive package. However, LSTM is a complicated architecture that involves considerable model parameters relating to the number of hidden cells. We used a computer resource required for generating and training networks with core i5-9500 4.10 GHz DDR4 16GB.
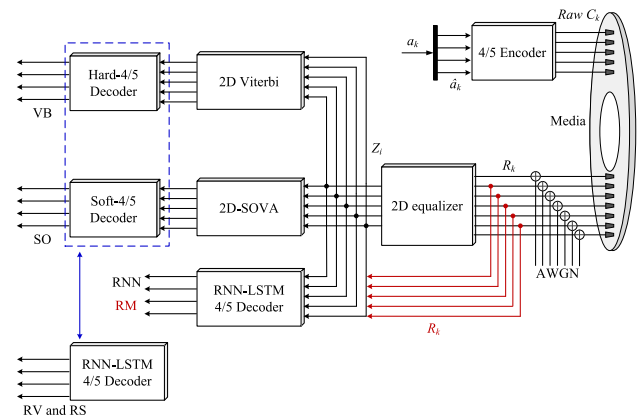

FIGURE 2. Schematic of rate-4/5 modulation code for a 2D BPMR system. BER comparison: the conventional decoder from Viterbi (VB) and soft (SO) decoders, and the adaptive LSTM-based decoder used for hard (RV), soft (RS), equalizer (RNN), and direct media readback (RM).

The independent graphics processing unit (GPU) used for LSTM computation has illustrated the best fit method for embedding deep learning in the hard disk drive.

A. EVALUATION METRICS

Training data was generated from conventional hard and soft information with a user bit of $4 \times 8{,}192$ sector size for every training in this work. Reasonably, we simulated the datasets based on the available FePt nanopost for a bit island of BPM media with AD $= 4$ Tb/in^2 [38]–[40], which can be commercially fabricated by lithography [41], [42].

To find a crucial number of LSTM cells, the relation between accuracy and time for different LSTM cell numbers was first evaluated in order to optimize epochs and other hyperparameters with respect to the sizeable memory required for the network. Subsequently, we tested only the data of the Viterbi (hard) and SOVA (soft) detectors. The input dimension was $5 \times 8{,}192$ (row × column) as a recorded bit, the tensor was $5 \times 1 \times 8{,}192$ (input × timestep × length) of feedforward, and the user bit was $4 \times 8{,}192$ in size (real information) with output prediction from a rate-4/5 modulation code. For data complexity, we divided the $5 \times 8{,}192$ hard and soft information into $5 \times 2{,}048$ each, and we combined those fractions into ten datasets using the same user bit answer of $4 \times 8{,}192$, as Fig. 3 shows.

We performed rough training with a sweeping number of LSTM cells from 1 to 200 and an increasing cell step of four, corresponding to the prediction accuracy and computation time that Fig. 4a illustrates. The condition was fixed for 500 epochs. The best range, in terms of accuracy and duration, was around 121 LSTM cells, resulting in learning performance that tended to increase during training and perish at around 81–86% for lower training durations. We then ran the fine sweep for every cell number from 111 to 131. Nevertheless, 124 cells provided the most effective result in terms of accuracy and time computation, as Fig. 4b displays.

Dataset 1	Hard Information			
Dataset 2	Soft Information			
Dataset 3	Hard Information		Soft Information	
Dataset 4	Soft	Hard		Soft
Dataset 5	Hard Information			
Dataset 6	Soft Information			
Dataset 7	Hard	Soft		Hard
Dataset 8	Soft Information		Hard Information	
Dataset 9	Hard	Soft	Hard	Soft
Dataset 10	Soft	Hard	Soft	Hard

$$\hat{a}_k \in 5 \times 8192$$

FIGURE 3. Training dataset arrangement with the same correct answer for the supervised learning algorithm.
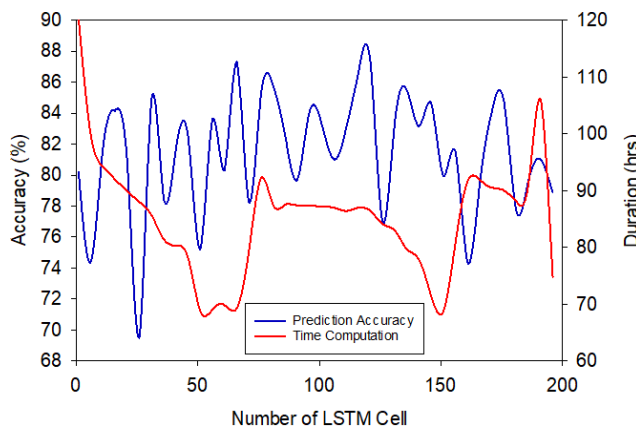
(a) Rough sweep of every four cells
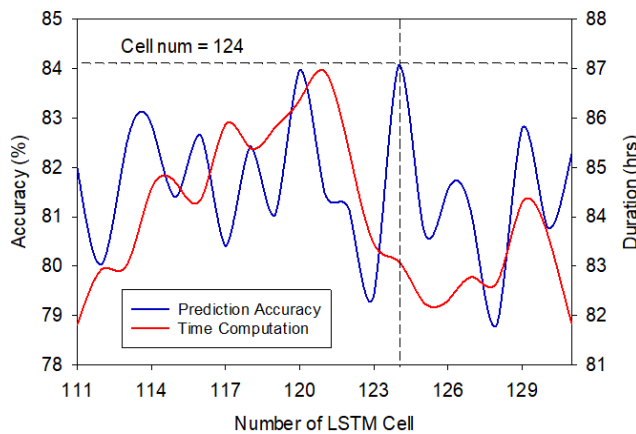
(b) Fine sweep from 111 to 131 cells

FIGURE 4. Training performance investigation between reasonable high prediction accuracy and training duration.

B. DATA TRAINING FOR RATE-4/5 DECODER

While hyperparameter changes did not offer a significant change in estimation accuracy, the arrangement of training datasets had an impact on prediction performance. In the first BPMR on rate-4/5 modulation code, real training was consistent with readback signal, 2D equalizer, 2D Viterbi, and 2D SOVA from random bit patterned media simulations in which bit island position uncertainty was very close to the

real media. For a readback signal, we selected five encoded island tracks and, instead, gathered all seven tracks because our approach did not require the two additional sidetracks as a model looking at signal distribution, whereas the conventional approach required another two sidetracks to complete the codeword for the mapping algorithm.
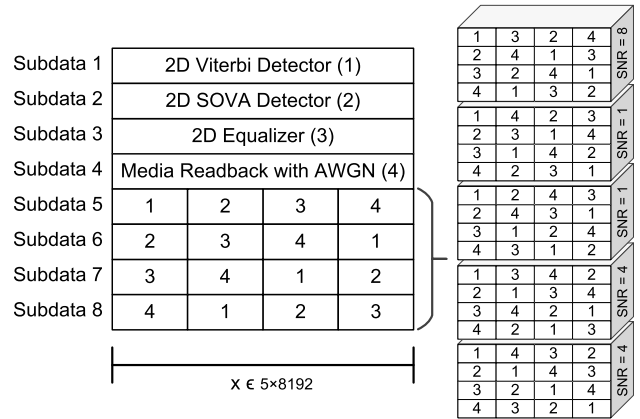

FIGURE 5. Datasets for LSTM-based 4/5 modulation decoder training. The source of data was labeled by 1: 2D Viterbi detector output, 2: 2D SOVA detector output, 3: 2D Equalizer output, and 4: Readback signal with AWGN. The data arrangement was ordered with 4! and combined with four original sources for each dataset. The training was performed with generated 48 datasets in total, at SNR 8, 1, and 4 dB, respectively.

Subsequently, we arranged for the input data to be more complicated by ordering the signal from a conventional system represented for 1 (VB), 2 (SOVA), 3 (equalizer), and 4 (media readback). The total training dataset of these four signal types was arranged by four factorial (4!) order, which resulted in 24 datasets. As Fig. 5 depicts, the networks learned data from three SNRs. We determined two datasets for each SNR since the rapid variation of statistical variability from one SNR to another would cause strong cross-entropy loss fluctuation over the training process. Each iteration accommodated four original channel signals and four integrated signals, which were arraying without duplicating a signal type for each section $X \in 5 \times 2048$ from subdata 5 to 8 in Fig. 5. The model conducted training at 250 epochs with 124 LSTM cells. The complete network from this training was defined as the default network for any other related binary classification with rate-4/5.

In addition, the model prediction involved no decimal number limit, which affected the optimal conditions in a long-time training process. The receiving operating characteristic (ROC) curve [34] was used to define the threshold decision from a true positive rate or sensitivity (100%-exact prediction for each bit) and a false-positive rate or specificity (the evaluated bit number at the fourth decimal number). The 85% threshold was defined by estimating the learning process when a loss function tends to steady. We double-checked the prediction accuracy—both from this statistical bias measurement and from ensuring that the direct number of predicted bits matched with the recorded bits in good agreement.

C. BER ENHANCEMENT VIA BACKEND CHANNEL

Since we performed a large number of readback signal-level distribution related to different SNR to address various modifications of parameters' effect on decoding accuracy, identifying specific training conditions was challenging. Virtually, we found that the best way to train a default network was familiar with every data pattern in such a BPMR system. Thus, the network could perform additional training with particular case functionality for a specific channel replacement. Hence, we needed a new channel to handle such a situation and optimize the network appropriately.

Instead of creating a codeword to map bit correction in each system separately, we integrated the default network within the backend channel of the decoder system for the desired channel type (e.g., equalizer, detector, decoder). The backend channel is a simple modification of the data training process in Fig. 1c that only supports the exact number of input and output nodes corresponding to the default network. To initiate the training, this LSTM backend channel was designated with the threshold condition of BER prediction that Fig. 6 displays.
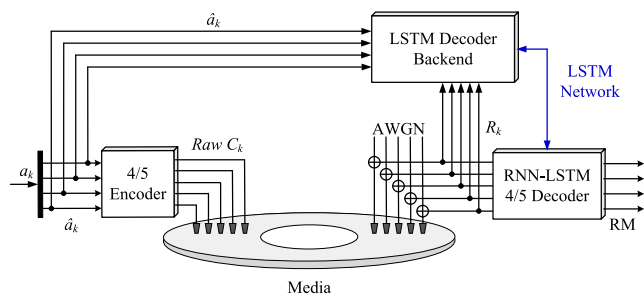

FIGURE 6. The backend channel for LSTM network enhancement. This channel shares the same default network as the system decoder channel (RM). The further training mechanism is defined by the threshold condition of a desired BER of RM.

We defined the mechanism such that the backend responded when BER performance was lower than 90%. The channel stored data only on sector $\hat{a}_k$ and R_k for a backend channel, while it shared the network with the system decoder channel. Since the default network was training at a fixed epoch to generalize the different source patterns in the rate-4/5 modulation code system, the optimized network was trained at the optimal point where the loss function provided a model, neither underfitting nor overfitting. The optimal value was identified by the turning point of cross-entropy loss and prediction accuracy from the ROC curve at a very long epoch pre-training.

Nevertheless, we note that backend training was optimized manually and that we did not successfully implement the channel with real-time processing due to the data pattern's complexity at various SNRs, which addressed a very sophisticated new input data stream. Ultimately, this approach would benefit hard disk drives in real situations rather than simulations since the data flow in the presented scenario only matched the drive environment with all factors (e.g., ISI, ITI, media noise, media fabrication), which resulted in one condition for backend training.
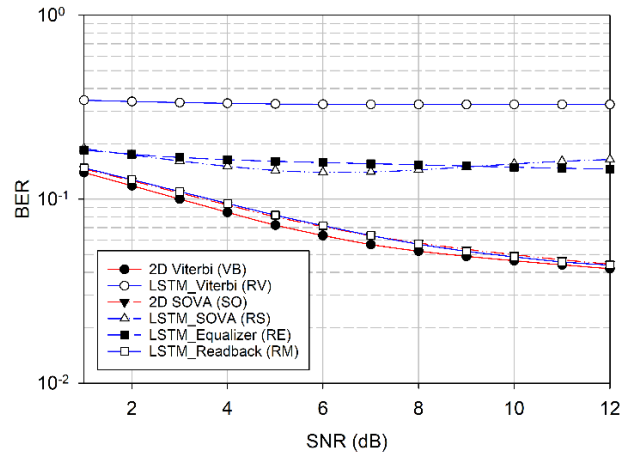

FIGURE 7. BER performance of the default network (blue lines) from every channel in a reading side. The decoding network with a direct readback signal input addresses BER compatible with the conventional 2D SOVA decoder at AD = 4 Tb/in^2.

D. RATE 4/5 DECODER PERFORMANCE COMPARISON

The evaluated BER performance of both the conventional and LSTM-based decoding channels was simulated on a staggered BPMR channel. The decoder results used the readback (RM), equalizer (RE), 2D Viterbi detector (RV), 2D SOVA detector (RS) represented in Fig. 7. In this work, we were interested in decoding performance at a low SNR range since the network was designed to adapt to real scenarios as much as possible—and in such real scenarios, the network would be affected by various noises and media defects resulting from the fabrication process. This consideration led us to focus on adaptive capability with existing BPMR channel processing models, rather than on a proposed breakthrough in the BER performance model. The LSTM model was trained by the data distribution with step datasets of SNR = 8, 1, and 4 dB, respectively, over a simulation range from SNR 1 to 12 dB, whereas the 2D Viterbi and 2D SOVA decoders separately considered the data streams in each case.

Interestingly, the RV result provided the highest BER value, while the conventional Viterbi decoder (VB) resulted in the lowest BER value. This difference in results was because the data training from the 2D Viterbi detector was in the pattern of $\{-1, +1\}$, which the network failed to classify as positive or negative values as it was more fit for codewords. In this simulation, we found that more complex data in a training process would result in sufficient accuracy for a readback signal but less accuracy in a hard decoder. This finding was because the model learned the chaos pattern, not only -1 and 1 in hard cases. The BER performance of RM (the blue line with a shallow square in Fig. 7) was compatible with conventional 2D SOVA (SO) in the order of 10^{-1} over the SNR = 5 to 12 dB. The default network demonstrated that this technique was directly suitable for the high variance in signal levels, suggesting a compact solution for several channels and a direct bit reading from the BPMR media.

Since the backend channel was employed to the system, the LSTM illustrated that the significant BER surpassed all the conventional decoders. And since all the existing hyperparameters were fixed with the default network, we performed a large number of epoch simulations with one dataset in order to find the optimal value with four additional training datasets. This process was updated only with the readback signals at SNR = 1 and 8 dB, respectively. The prediction accuracy (ACC) related to BER was evaluated by [30] ACC = (TP + TN) / (TP + TN + FP + FN), where TP is the true positive, TN is the true negative, FP is the false positive, and FN is the false negative. A turning point of both loss function and prediction accuracy of the data at SNR = 1 dB was found at an epoch of 1,800 with a maximum validity of 87.86%. The model was further trained with the SNR = 8 dB dataset, which provided 2,010 epochs with 99.95% ACC. After the default network was updated, the BER of RM improved significantly, as Fig. 8 shows. The prediction at the lower SNR coming in slightly better than Viterbi and SOVA could demonstrate potential flexibility in replacing any channel within the hard disk drive.
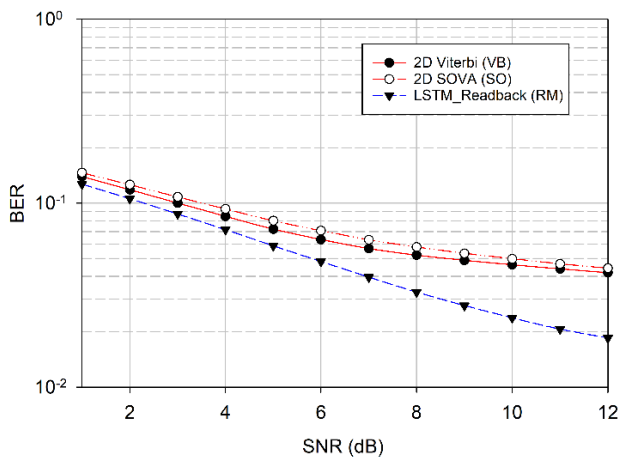

(a) The multiple SRTR system decoded by Viterbi (VB) and the SOVA (SO)

(b) LSTM-based SRTR decoder using the same 4/5 decoding network

FIGURE 9. Schematics of the simulated channels for the cross-platform decoder. SRTR provides six readback tracks from the four-track bit island reading, which is standardized before feed input to the network model.

FIGURE 8. BER performance of the LSTM network improved with additional readback signals via the backend channel at the AD = 4 Tb/in^2.

IV. CROSS-PLATFORM WITH OTHER BPMR DESIGNS

Various proposed codeword-based channels in both conventional and neural networks-based systems have addressed adequate BER performance. Nevertheless, their flexibility for use with different media has been limited by specific codeword mapping designed separately for each channel. Neural networks have taken advantage of many possible applications of the simulation, realizing information within individual scenarios. Thus, we performed a cross-platform application by transferring a fully trained rate 4/5 decoder network for use with another different system using the recently proposed system comprising multiple SRTR system within the same output track.

Fig. 9 shows the multiple SRTR system schematically, providing both the 1-D Viterbi (VB) and 1-D SOVA (SO) decoding results. The staggered BPMR was determined by
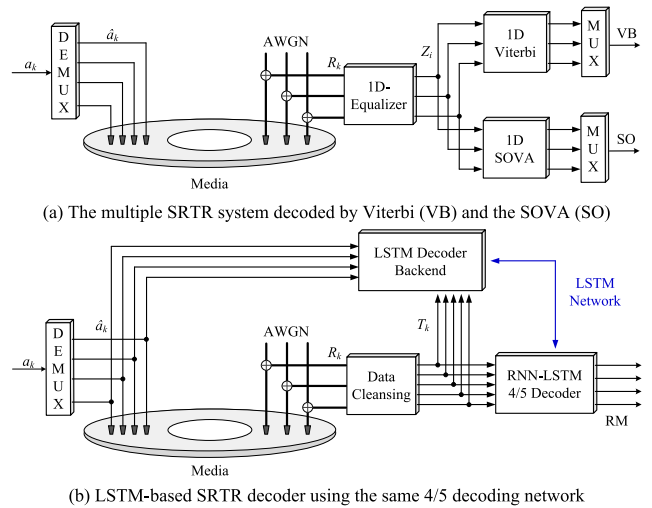
two BPMR media, defined by areal density (AD). For AD = 3 Tb/in^2, the media comprise a 9.75 nm × 10 nm size of bit island arrays with a bit period of T_x = 14.5 nm and a bit island track pitch of T_z = 14.5 nm. Furthermore, the bit island was arrayed with T_x = 12.5 nm and T_z = 12.5 nm for AD = 4 Tb/in^2. The system consisted of three headers across four rows of the bit island. Each header recorded readback signals directly from two tracks of bit islands. Thus, the first header recorded the first and second rows as the first and second data tracks while the second and third rows of the bit island were measured as the third and fourth data tracks for the second header. As a result, the second and the third tracks' measurement overlapped with the same bit island row.

A. ABLATION STUDIES RELATED TO BPMR SYSTEMS

Systematic readback signal checks at each track versus timestep revealed uncertain values. Data validation usually indicated an irrelevant readback input for training in this system. Typically, the signal level from the same bit island should be in the same positive or negative trend when it is measured with a different header. The multiple SRTR provided several variances in signal at the same timestep, read by two headers so that the neural network's gradient update could not decide whether such input data should be in positive or negative values. For a conventional 4/5 decoder or other systems, this decoder addressed the use of the same header for readback with an encoded bit record in media, which provided the same signal distribution standardization. Multiple SRTR, however, used three headers for two tracks' readings, which readback data did not match within the same standardization.

Importantly, the readback signals were replicated with the new standard—other than a vanilla LSTM in a rate-4/5 decoder. In this section, we performed an ablation study of the input LSTM channel to verify the non-equalized headers. The input layer modification with the vanilla LSTM,

the self-organizing map (SOM), and the data cleansing were measured and compared by BER performance in order to determine the LSTM application for magnetic recordings. The global condition of the LSTM input in this investigation was set for SRTR readback from media noise 5% with AD = 4 Tb/in^2.

The first evaluation compared the plain, vanilla LSTM with the fully trained rate-4/5 decoder network by directly replacing all the read-side SRTR channels in Fig. 9a with the LSTM channel. At SNR = 8 dB, the LSTM-based network diminished in prediction accuracy from $\sim$ 99.95% for the latest 4/5 network training to $\sim$ 67% for the LSTM-based SRTR data. As the result, further network updates were always unsuccessful after a certain iteration due to the input values' progression to cross-entropy loss falling outside the function domain.
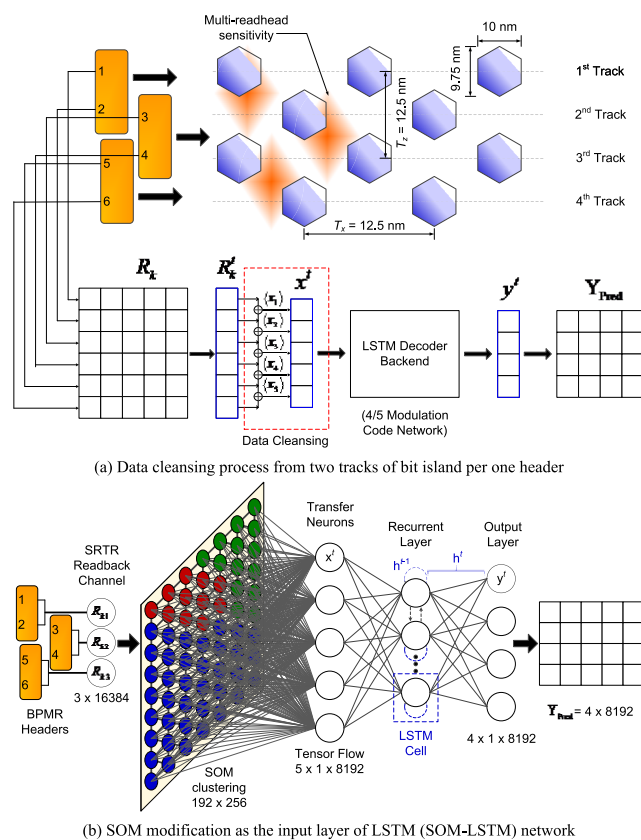

(a) Data cleansing process from two tracks of bit island per one header

(b) SOM modification as the input layer of LSTM (SOM-LSTM) network

FIGURE 10. The multiple overlap readback signal from R_k of $6 \times 8{,}192$ to $5 \times 8{,}192$ for one sector with the same standardization for systematizing the readback signal in a multiple-SRTR system.

For the next step, we modified the LSTM model's input layer to investigate the significant improvement in BER. Although the LSTM-based channel for rate-4/5 can evaluate the SRTR readback over SNR = 1 to 8 dB with an overall accuracy $\sim$ 67%, the formatting inconsistencies of this combination should be identified through the data cleansing process. As Fig. 10a depicts, R_k standardized the irrelevant signals to ensure that all datasets conform to the single header compatibility before feeding into the LSTM input layer,

labelled T_k. The data cleansing for three header readings accumulated a readback set of $6 \times 8{,}192$ and was standardized to a single compatible header of $5 \times 8{,}192$. The pre-processing data was determined by the expected value of mean standard derivation (MSD) between two nearby tracks, as follows:

$$\langle X_{\text{track}} \rangle = \frac{\overline{X}_{a_k, a_{k+1}}}{\sigma_{\overline{X}}^2} \qquad (10)$$

where $\overline{X}_{a_k, a\_1}$ is the mean value of signal level between the upper track (a_k) and the lower track (a_{k+1}), and $\sigma_{\overline{X}}^2$ is the variance of the mean value between a_k and a_{k+1}.

Next, we implemented the SOM to characterize three R_k signals for the LSTM model (SOM-LSTM). We replaced the LSTM input nodes with the SOM to cluster signals corresponding to each bit island track. The input vector of $3 \times 16{,}384$ in one readback sector was transformed into the SOM with a 192×256 weight dimension instead in order to arrange the new data tracks of $6 \times 8{,}192$. Then, the winner weight from the SOM was selected to map raw readback to a transfer neuron in order to transform data into the tensor, as Fig. 10b depicts. Importantly, we pointed out the data pattern before and after attaining the SOM training with the rate-4/5 and the SRTR system, as Fig. 11 illustrates.
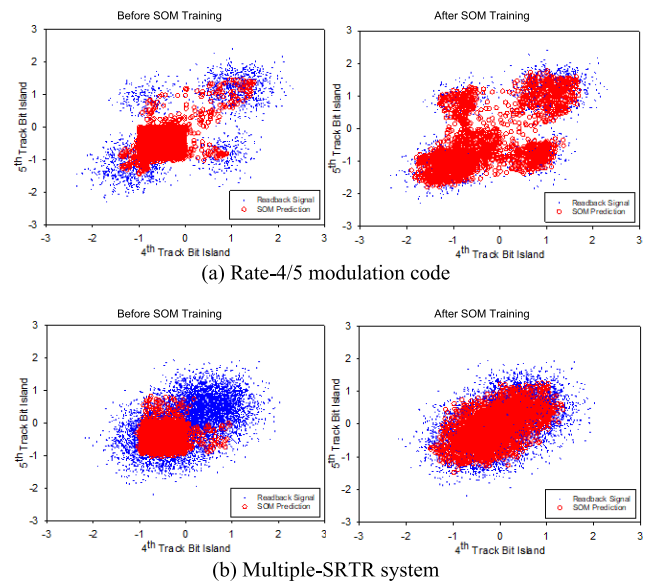

(a) Rate-4/5 modulation code

(b) Multiple-SRTR system

FIGURE 11. Clustering implementation with the cross-track of readback signals from bit island rows measured directly.

Finally, we performed data cleansing with SOM-LSTM by manipulating the cleansing process before it was fed into the SOM. The BER summarized the compared feature engineering results, as Fig. 12 shows. SOM-LSTM enhanced the BER so that it was more sufficient than a vanilla model, while SOM-LSTM reinforced with data cleansing did not affect any change in BER. However, only the data cleansing with a vanilla LSTM attained an improvement for using a single trained network from a rate-4/5 decoder with the SRTR
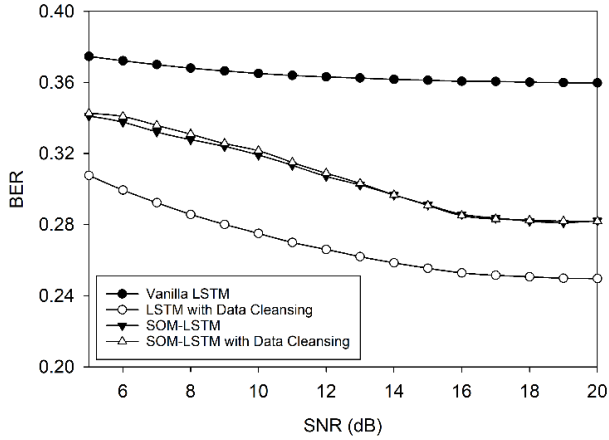
FIGURE 12. Linear scale of BER performance to compare the ablation study's results.

(a) AD = 3 Tb/in^2

(b) AD = 4 Tb/in^2

FIGURE 13. BER performance of LSTM-based 4/5 decoding network used for the SRTR system, with and without data cleansing.

system, and vice versa. Similar works on an ablation study of clustering within LSTM have reported similar results on the vanilla LSTM's effectiveness [43], [44].

B. TRAINING DATA IN THE SRTR ENVIRONMENT
Since the LSTM-based rate-4/5 decoding network remains an instalment, we utilized a gradient update with a similar data arrangement as Fig. 5 to match the network to an SRTR environment. The training process was divided into two iterations. First, we rendered the $6 \times 8{,}192$ R_k readback signal into a conventional SRTR with media noise of 0% and 5% for both AD = 3 Tb/in^2 and 4 Tb/in^2 at SNR = 9 dB. We evaluated an optimal 480 epochs for all the subdata in one dataset. Second, we performed another iteration using only the three datasets with an individual 1,080 epochs of media noise at 5% for AD = 4 Tb/in^2 at SNR = 16 dB.

The ablation study has investigated the simulation with data cleansing under two conditions: the latest trained LSTM network for rate-4/5 modulation used with SRTR directly and the updated version network with SRTR training data. Fig. 9b demonstrates the pre-evaluation of the crossing application at the rate-4/5 network for use with the SRTR without additional training with SRTR data. Both the SRTR Viterbi (VB_0 for media noise = 0% and VB_5 for media noise = 5%) and the SOVA (SO_0 and SO_5) results were identical for all the SNR ranges of both AD = 3 and 4 Tb/in^2. As Fig. 13a and Fig. 13b depict, the rate-4/5 network with data cleansing (RM_0 and RM_5) introduced BER of about $2–3 \times 10^{-1}$, which confirmed an irrelevant readback from the SRTR.

After updating the new training within the SRTR environment, the LSTM network offered slightly better BER performance than the conventional SRTR channels (VB and SO), around 0.5 to 0.7×10^{-1} over the SNR range from 1 to 8 dB, as Fig. 14a and Fig. 14b exhibit. Subsequently, the vanilla LSTM model was insufficient at high SNR levels due to our generalized decoder approach's inability to classify the encoded bit in the media in the same way as the
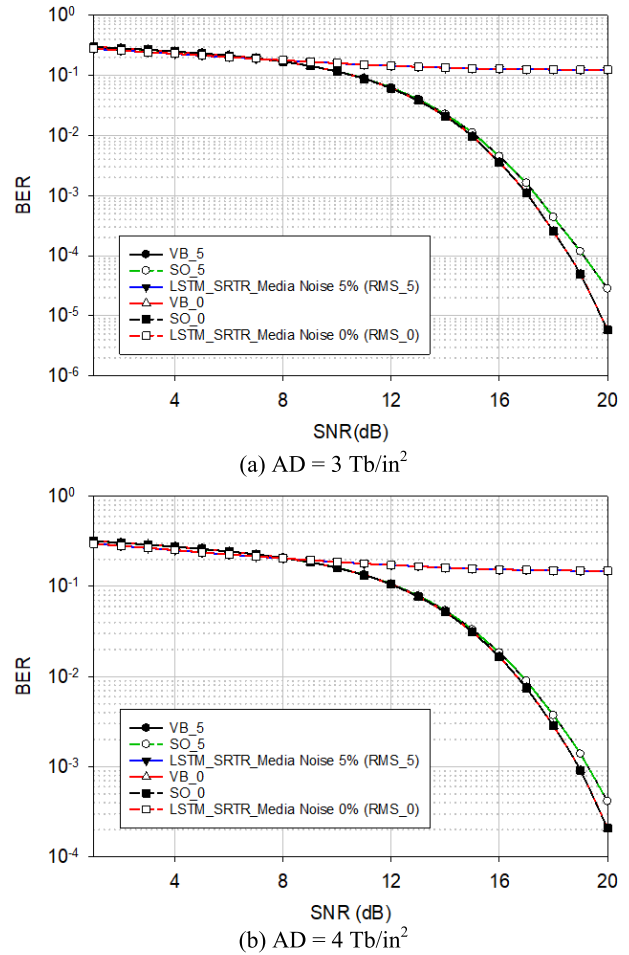
well-defined mapping algorithm designed for such a particular system.

C. THE USE OF THE SAME LSTM NETWORK ACROSS RATE-4/5 MODULATION CODE AND SRTR SYSTEMS
Finally, we again investigated backward decoder performance utilizing the SRTR data-based LSTM with a rate-4/5 system in order to demonstrate cross-platform functionality. The analogous accuracy during the training process of the first use of rate-4/5 decoder-based data and the SRTR-based data was monitored by the ROC curve, as Fig. 15 displays. A systematic simulation of the BPMR media with AD = 4 Tb/in^2 over an SNR range from 1 to 20 dB was considered to observe the BER change when the adaptive network was based on SRTR data with conventional data. The backend channel still manipulated the gradient adjustment with the new readback data of AD = 4 Tb/in^2 and media noise 0% at the SNR = 8 dB, which required a ~7.1 times lower training epoch than the first rate-4/5 backend update performed.

A comparable BER on the rate-4/5 system was evaluated by the first backend update on rate-4/5 data, and the network used for SRTR in Fig. 14 with rate-4/5 data is
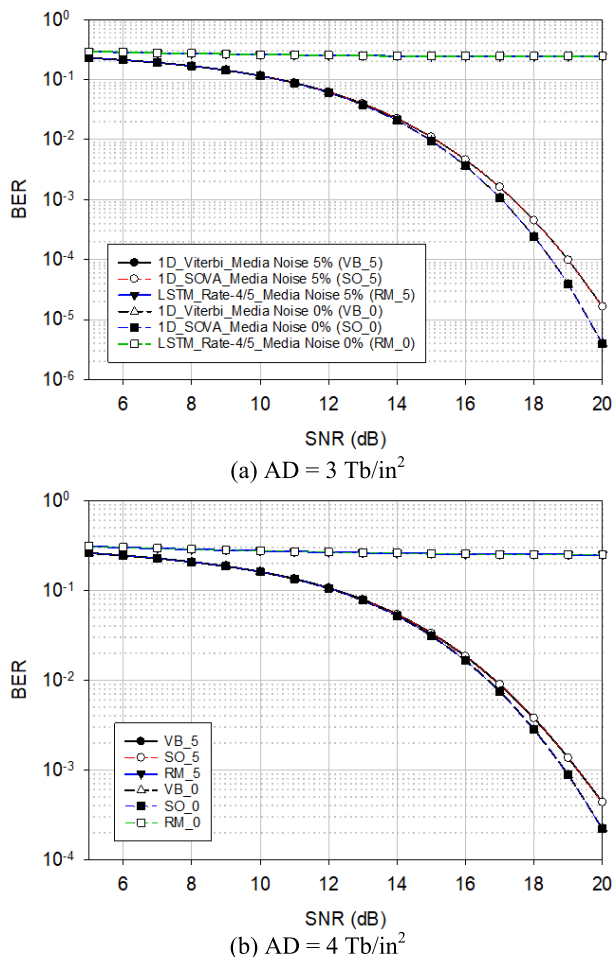
(a) AD = 3 Tb/in^2

(b) AD = 4 Tb/in^2

FIGURE 14. BER performance of the LSTM network enhanced by the readback signals of the SRTR system with SNR of 9 and 16 dB at the AD = 4 Tb/in^2.
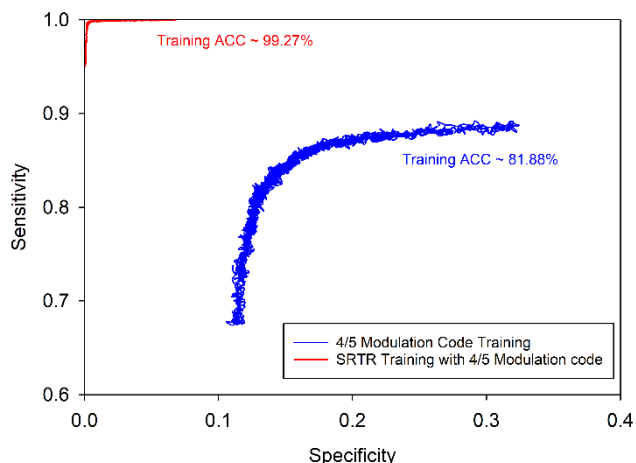

FIGURE 15. ROC curve for data training with the readback signal from a 4/5 decoder system (blue) and an LSTM-based SRTR network on rate-4/5 modulation code system.

illustrated in Fig. 16. As a result, no significant difference occurred between LSTM training with the rate-4/5 data and the SRTR data when performing simulations between two
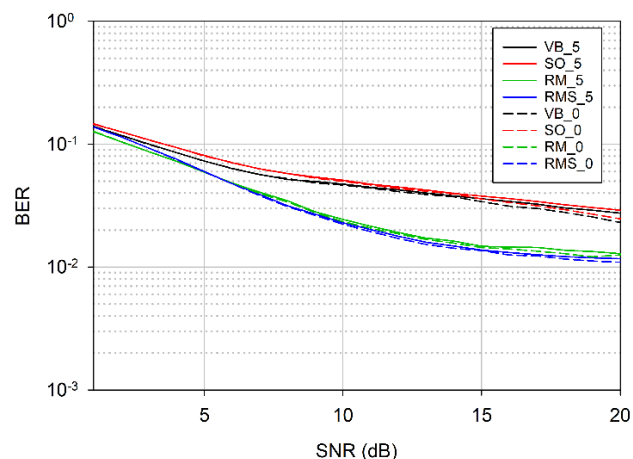

FIGURE 16. Backward usage of an LSTM-based SRTR network with a rate 4/5 modulation code system, comparing with the conventional system and an LSTM-based 4/5 decoding network.

different systems. Overall, the LSTM of both trained network versions had an appreciable BER over the two channels mentioned. The SRTR data-based LSTM with a traditional 4/5 decoder had a slightly better BER than a pure rate-4/5 data training, from approximately 4.05% to 0.79% for media noise 0% and 4.3% to 0.69% for media noise 5% over SNR = 1 to 6 dB.

This comparable reliability demonstrated the possibility of minimizing such new processing channels with upcoming hard disk drive designs and how such new designs could be appropriated to the existing platform and tested for quality compared to the conventional system.

V. CONCLUSION AND FUTURE WORK

In summary, we have demonstrated a neural network-based modulation code made of vanilla LSTM for focusing an application on cross-platform capability between different BPMR systems. The study contributes both calibration tools for different BPMR channel designs' standardization and potential focuses for sequence data with BPMR channels. Also, our results illustrate the limitation of supervised learning-based channels and a first attempt to implement clustering studies with a BPMR channel. We utilized signal levels as data distribution in all read-side channels and optimized the LSTM update with respect to system environment, using the backend channel to enable efficient bit predictions. With this concept, we overcame codeword restrictions with less complexity within each channel design. As a result, we were able to classify the recorded bit from a readback signal more sufficiently than a conventional system. Moreover, our ablation study demonstrated an advantage of the LSTM with data cleansing compared to the vanilla and semi-unsupervised models when focusing on multi-BPMR channel calibration. Furthermore, we demonstrated the use of the same 4/5 decoding network with different SRTR systems, and vice versa. Although the BER results are not as promising at the low SNR as many

conventional systems, our LSTM-based channel revealed the versatility of adequate BER performance in a robust interfered signal environment. Thus, our results pave the way for various BPMR channels' calibration with the usual channels based on a transferable decoding network—as well as further developments in the clustering model with future BPMR channel designs.

REFERENCES

[1] B. Chen and J.-Y. Chang, "Effect of flattening cracked medium on positioning accuracy of a linear magnetic encoder," *IEEE Trans. Magn.*, vol. 56, no. 2, pp. 1–7, Feb. 2020.

[2] J. Um, M. R. Z. Kouhpanji, S. Liu, Z. N. Porshokouh, S.-Y. Sung, J. Kosel, and B. Stadler, "Fabrication of long-range ordered aluminum oxide and Fe/Au multilayered nanowires for 3-D magnetic memory," *IEEE Trans. Magn.*, vol. 56, no. 2, pp. 1–6, Feb. 2020.

[3] T. A. Nguyen and J. Lee, "Error-correcting 5/6 modulation code for staggered bit-patterned media recording systems," *IEEE Magn. Lett.*, vol. 10, pp. 1–5, 2019.

[4] K. Harada, "High-efficient adaptive modulation for PWM-based multi-level perpendicular magnetic recording on insufficient resolution channel," *IEEE Trans. Magn.*, vol. 55, no. 11, pp. 1–8, Nov. 2019.

[5] S. Jeong and J. Lee, "Signal detection under multipath intersymbol interference in staggered bit-patterned media recording systems," *IEEE Magn. Lett.*, vol. 10, pp. 1–5, 2019.

[6] M. Nishikawa, Y. Nakamura, Y. Kanai, H. Osawa, and Y. Okamoto, "A study on iterative decoding with LLR modulator by neural network using adjacent track information in SMR system," *IEEE Trans. Magn.*, vol. 55, no. 12, pp. 1–5, Dec. 2019.

[7] T. Gruber, S. Cammerer, J. Hoydis, and S. T. Brink, "On deep learning-based channel decoding," in *Proc. 51st Annu. Conf. Inf. Sci. Syst. (CISS)*, Baltimore, MD, USA, Mar. 2017, pp. 1–6.

[8] Y. Qin and J.-G. Zhu, "Deep neural network: Data detection channel for hard disk drives by learning," *IEEE Trans. Magn.*, vol. 56, no. 2, pp. 1–8, Feb. 2020.

[9] S. Han, G. Kong, and S. Choi, "A detection scheme with TMR estimation based on multi-layer perceptrons for bit patterned media recording," *IEEE Trans. Magn.*, vol. 55, no. 7, pp. 1–4, Jul. 2019.

[10] C. Cao, D. Li, and I. Fair, "Deep learning-based decoding for constrained sequence codes," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Abu Dhabi, United Arab Emirates, Dec. 2018, pp. 1–7.

[11] A. Sayyafan, B. J. Belzer, K. Sivakumar, J. Shen, K. S. Chan, and A. James, "Deep neural network based media noise predictors for use in high-density magnetic recording turbo-detectors," *IEEE Trans. Magn.*, vol. 55, no. 12, pp. 1–6, Dec. 2019.

[12] F. Liang, C. Shen, and F. Wu, "Exploiting noise correlation for channel decoding with convolutional neural networks," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Kansas City, MO, USA, May 2018, pp. 1–6.

[13] K. Luo, S. Wang, G. Xie, W. Chen, J. Chen, P. Lu, and W. Cheng, "Read channel modeling and neural network block predictor for two-dimensional magnetic recording," *IEEE Trans. Magn.*, vol. 56, no. 1, pp. 1–5, Jan. 2020.

[14] M. Nishikawa, Y. Nakamura, H. Osawa, Y. Okamoto, and Y. Kanai, "A study on iterative decoding with LLR modulator using neural network in SMR system," *IEEE Trans. Magn.*, vol. 55, no. 7, pp. 1–4, Jul. 2019.

[15] W. Lyu, Z. Zhang, C. Jiao, K. Qin, and H. Zhang, "Performance evaluation of channel decoding with deep neural networks," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Kansas City, MO, USA, May 2018, pp. 1–6.

[16] Y. Hu, L. Zhao, Z. Yan, A. Kaushik, Y. Hou, and J. Thompson, "GatedNet: Neural network decoding for decoding over impulsive noise channels," *IEEE Commun. Lett.*, vol. 23, no. 8, pp. 1381–1384, Aug. 2019.

[17] J. Shen, A. Aboutaleb, K. Sivakumar, B. J. Belzer, K. S. Chan, and A. James, "Deep neural network a posteriori probability detector for two-dimensional magnetic recording," *IEEE Trans. Magn.*, vol. 56, no. 6, pp. 1–12, Jun. 2020.

[18] S. Jeong and J. Lee, "Modulation code and multilayer perceptron decoding for bit-patterned media recording," *IEEE Magn. Lett.*, vol. 11, pp. 1–5, 2020.

[19] C. Warisarn, "Mitigating the effects of track mis-registration in single-reader/two-track reading BPMR systems," *IEEE Trans. Magn.*, vol. 55, no. 7, pp. 1–6, Jul. 2019.

[20] W. Busyatras, C. Warisarn, S. Koonkarnkhai, and P. Kovintavewat, "A simple 2D modulation code in single-reader two-track reading BPMR systems," *Digit. Commun. Netw.*, Nov. 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2352864819300689, doi: 10.1016/j.dcan.2019.11.003.

[21] C. Lu, W. Xu, H. Shen, J. Zhu, and K. Wang, "MIMO channel information feedback using deep recurrent network," *IEEE Commun. Lett.*, vol. 23, no. 1, pp. 188–191, Jan. 2019.

[22] H. Palangi, L. Deng, Y. Shen, J. Gao, X. He, J. Chen, X. Song, and R. Ward, "Deep sentence embedding using long short-term memory networks: Analysis and application to information retrieval," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 24, no. 4, pp. 694–707, Apr. 2016.

[23] C. Li, Z. Wang, M. Rao, D. Belkin, W. Song, H. Jiang, P. Yan, Y. Li, P. Lin, M. Hu, N. Ge, J. P. Strachan, M. Barnell, Q. Wu, R. S. Williams, J. J. Yang, and Q. Xia, "Long short-term memory networks in memristor crossbar arrays," *Nature Mach. Intell.*, vol. 1, no. 1, pp. 49–57, Jan. 2019.

[24] X. Li, C. Yu, F. Su, T. Quan, and X. Yang, "Novel training algorithms for long short-term memory neural network," *IET Signal Process.*, vol. 13, no. 3, pp. 304–308, May 2019.

[25] S. Sen and A. Raghunathan, "Approximate computing for long short term memory (LSTM) neural networks," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 37, no. 11, pp. 2266–2276, Nov. 2018.

[26] P. J. Werbos, "Backpropagation through time: What it does and how to do it," *Proc. IEEE*, vol. 78, no. 10, pp. 1550–1560, Oct. 1990.

[27] K. P. Murphy, *Machine Learning: A Probabilistic Perspective*, 1st ed. Cambridge, MA, USA: MIT Press, 2012.

[28] V. F. Torres and F. S. Torres, "Resilient training of neural network classifiers with approximate computing techniques for hardware-optimised implementations," *IET Comput. Digit. Techn.*, vol. 13, no. 6, pp. 532–542, Nov. 2019.

[29] R. V. Kumar Reddy, B. Srinivasa Rao, and K. P. Raju, "Handwritten Hindi digits recognition using convolutional neural network with RMSprop optimization," in *Proc. 2nd Int. Conf. Intell. Comput. Control Syst. (ICICCS)*, Madurai, India, Jun. 2018, pp. 45–51.

[30] Y. Wang, J. Liu, J. Mišic, V. B. Mišic, S. Lv, and X. Chang, "Assessing optimizer impact on DNN model sensitivity to adversarial examples," *IEEE Access*, vol. 7, pp. 152766–152776, 2019.

[31] A. Graves, "Generating sequences with recurrent neural networks," 2013, *arXiv:1308.0850*. [Online]. Available: http://arxiv.org/abs/1308.0850

[32] Y. N. Dauphin, H. de Vries, and Y. Bengio, "Equilibrated adaptive learning rates for non-convex optimization," 2015, *arXiv:1502.04390*. [Online]. Available: http://arxiv.org/abs/1502.04390

[33] P. Hu, D. Peng, Y. Sang, and Y. Xiang, "Multi-view linear discriminant analysis network," *IEEE Trans. Image Process.*, vol. 28, no. 11, pp. 5352–5365, Nov. 2019.

[34] X. Peng, J. Feng, S. Xiao, W.-Y. Yau, J. T. Zhou, and S. Yang, "Structured AutoEncoders for subspace clustering," *IEEE Trans. Image Process.*, vol. 27, no. 10, pp. 5076–5086, Oct. 2018.

[35] X. Peng, H. Zhu, J. Feng, C. Shen, H. Zhang, and J. T. Zhou, "Deep clustering with sample-assignment invariance prior," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Dec. 31, 2020, doi: 10.1109/TNNLS.2019.2958324.

[36] P. Hu, D. Peng, X. Wang, and Y. Xiang, "Multimodal adversarial network for cross-modal retrieval," *Knowl.-Based Syst.*, vol. 180, pp. 38–50, Sep. 2019.

[37] C. Warisarn and P. Kovintavewat, "Soft-output decoding approach of 2D modulation codes in bit-patterned media recording systems," *IEICE Trans. Electron.*, vol. E98.C, no. 12, pp. 1187–1192, 2015.

[38] B. Xu, H. Wang, Z. Cen, and Z. Liu, "4–5 Tb/in² heat-assisted magnetic recording by short-pulse laser heating," *IEEE Trans. Magn.*, vol. 51, no. 6, pp. 1–5, Jun. 2015.

[39] D. Weller, G. Parker, O. Mosendz, E. Champion, B. Stipe, X. Wang, T. Klemmer, G. Ju, and A. Ajan, "A HAMR media technology roadmap to an areal density of 4 Tb/in²," *IEEE Trans. Magn.*, vol. 50, no. 1, pp. 1–8, Jan. 2014.

[40] Y. Kamata, T. Maeda, H. Hieda, R. Yamamoto, N. Kihara, and A. Kikitsu, "5 Tdot/inch2 bit-patterned media fabricated by directed self-assembling polymer mask," in *Dig. APMRC*, Singapore, 2012, pp. 1–2.

[41] T. R. Albrecht *et al.*, "Bit-patterned magnetic recording: Theory, media fabrication, and recording performance," *IEEE Trans. Magn.*, vol. 51, no. 5, pp. 1–42, May 2015.

[42] Y. Shiroishi, K. Fukuda, I. Tagawa, H. Iwasaki, S. Takenoiri, H. Tanaka, H. Mutoh, and N. Yoshikawa, "Future options for HDD storage," *IEEE Trans. Magn.*, vol. 45, no. 10, pp. 3816–3822, Oct. 2009.

[43] M. Wang, D. Shi, N. Guan, T. Zhang, L. Wang, and R. Li, "Unsupervised pedestrian trajectory prediction with graph neural networks," in *Proc. IEEE 31st Int. Conf. Tools with Artif. Intell. (ICTAI)*, Portland, OR, USA, Nov. 2019, pp. 832–839.

[44] V.-H. Nguyen, K. Sugiyama, M.-Y. Kan, and K. Halder, "Neural side effect discovery from user credibility and experience-assessed online health discussions," *J. Biomed. Semantics*, vol. 11, no. 1, pp. 1–16, Dec. 2020.
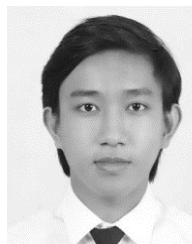
CHAIWAT BUAJONG (Graduate Student Member, IEEE) received the B.Eng. degree in manufacturing system from the College of Advanced Manufacturing Innovation, King Mongkut's Institute of Technology Ladkrabang (KMITL), Bangkok, Thailand, in 2018, where he is currently pursuing the Ph.D. degree with the College of Advanced Manufacturing Innovation. His research interests include coding, read channel modeling, and magnetic recording receiver.

TEANCHAI CHANTAKIT received the B.Sc. degree in physics from Ramkhamhaeng University (RU), Bangkok, Thailand, in 2011, the B.Sc. (Hons.) and M.Sc. degrees in applied physics from the King Mongkut's Institute of Technology Ladkrabang (KMITL), Bangkok, in 2013 and 2015, respectively, and the Ph.D. degree in physics from Kasetsart University (KU), Bangkok, in 2019.

In 2017, he was a Research Assistant in collaborate with the Western Digital (Thailand) Company Ltd., Ayutthaya, Thailand. From 2018 to 2019, he was a Visiting Researcher with the Ultrafast Nanophotonics Group, Universität Paderborn, Paderborn, Germany. He is currently a Postdoctoral Research Fellow with the College of Advanced Manufacturing Innovation, KMITL. His research interests include micro/nano-scale integrated optical devices, metasurfaces, nanophotonic applications, and all-optical neural networks.

CHANON WARISARN (Member, IEEE) received the B.Eng. degree (Hons.) in electronics engineering technology from the King Mongkut's Institute of Technology North Bangkok (KMITNB), Thailand, in 2006, and the Ph.D. degree in electrical engineering from the King Mongkut's Institute of Technology Ladkrabang (KMITL), Bangkok, Thailand, in 2011.

He is currently working with the College of Data Storage Innovation, KMITL. His current research interests are in the areas of communications and signal processing for data storage systems.

● ● ●