(Special Session)

AlN Sputtering Parameter Estimation Using A Multichannel Parallel DCT Neural Network

Yingyi Luo*, Talha M. Khan†, Emadeldeen Hamdan*, Xin Zhu*, Hongyi Pan*, Didem Ozevin†, A. Enis Cetin*
*Department of Electrical and Computer Engineering, †Department of Civil, Materials, and Environmental Engineering
University of Illinois Chicago
Chicago, Illinois, U.S.A
{yluo52,tkhan41,ehamda3,xzhu61,hpan21,dozevin,aecyy}@uic.edu

2024 IEEE 42nd VLSI Test Symposium (VTS) | 979-8-3503-6378-4/24/$31.00 ©2024 IEEE | DOI: 10.1109/VTS60656.2024.10538771

***Abstract*—In this paper, we present a method for estimating the deposition parameters of the thin film material Aluminum Nitride (AlN) using a deep neural network. The neural network predicts the AlN orientations, which are critical for micromachining Micro-Electro-Mechanical Systems (MEMS) transducers such as accelerometers and acoustic emission sensors. The network features three parallel channels, each equipped with a Discrete Cosine Transform (DCT) based layer that encodes the input parameters into a latent space. This DCT layer applies a hard-thresholding nonlinearity to eliminate noise from the input parameters, resulting in a sparse representation of the latent space. Trained with a dataset comprising AlN orientations parameters and their optimal values, our model is adept at simultaneously extracting and integrating various essential frequency components. Experimental results underscore the effectiveness of our proposed approach in achieving accurate and comprehensive estimation of AlN orientations and MEMS design parameters, thereby providing a promising path for advanced optimization.**

***Index Terms*—Aluminum nitride, sputtering, microfabrication, discrete cosine transform, deep neural network**

I. Introduction

Aluminum nitride (AlN) [1] is prized for its unique properties, such as piezoelectricity [2], electrical insulation [3], chemical resistance [4], and high surface acoustic wave velocity [5], making it crucial for various technological applications including surface acoustic wave devices, Micro-Electro-Mechanical Systems (MEMS) devices [6], and optical devices with a wide band gap. Depositing AlN in a desired orientation requires specific conditions and significant resource investment. Cheng *et al.* used RF magnetron sputtering with Argon and Nitrogen gases to deposit AlN thin films on Silicon substrates, achieving c-axis orientations perpendicular and parallel to the surface [7]. Film orientation shifted from (002) to (100) with changes in substrate-target distance and pressure, indicating the significant role of target-substrate distance and working pressure on AlN orientation.

AlN with varying thicknesses and orientations can be deposited by controlling different input conditions. In Zhang's study [8], eight experiments altered sputtering power while keeping substrate temperature, deposition pressure, and substrate-to-target distance constant. The orientation of the AlN thin films was analyzed via X-ray diffractometer (XRD) measurements. These orientation measurements were further digitized in the output vector. Peak intensities were converted into a scale from 0 to 1, indicating 0 as the baseline and one as the maximum peak at an orientation. Even with input-output conditions in various studies, adapting processes for different manufacturing tools and materials requires extensive optimization to reproduce. Therefore, we used data from four dimensions (sputtering power, substrate temperature, sputtering pressure, and target-to-substrate distance) as inputs to predict the orientation of AlN on Si substrates, which aids in estimating the design parameters for MEMS fabrication. This study aims to employ machine learning to predict input parameters for desired single crystal orientations, thereby refining the microfabrication optimization workflow and substantially reducing the time required for micromachining MEMS transducers.

In this paper, we introduce a parallel fully dense network with a Discrete Cosine Transform (DCT)-based layer [9]–[11] for MEMS design parameter estimation. In this model, multiple fully dense layers operate in parallel paths, allowing the network to learn different representations or features simultaneously. This parallel structure can enhance the model's ability to capture complex patterns and relationships in the input data [12]. Furthermore, at the center of each path, we insert a DCT-based layer to capture frequency-specific features or patterns in the input data. We experimentally show that either the parallel structure or the DCT-based layer effectively improves the accuracy of the network in the MEMS design parameter estimation.

The organization of the paper is as follows: Section II elucidates the diverse neural network architectures that underpin our predictive methodology. Herein shall ensure analysis of the autoencoder neural network structure, DCT neural network structure, the parallel DCT neural network structure, and finally, the multichannel DCT neural network structure, which entails an in-depth analysis of the more sophisticated multichannel parallel DCT neural network. In Section III, empirical validation is left for a complete description of the dataset used, an exposition of the training methodology used in some depth, and a test outcome analysis to highlight the predictive understanding of the models. Section IV presents the concluding remarks, which integrate the findings from the study and deliberate the prospective utility of the proposed predictive framework in the MEMS domain, with specific focus on strategic incorporation of Aluminum Nitride (AlN), a piejsones-electric layer in MEMS transducer.

II. Methodology

In this section, we elucidate the design of a comprehensive, fully dense neural network, incorporating both a Discrete Cosine Transform (DCT) layer and a parallel network architecture, to enhance the precision of the microfabrication data estimation process.

A. Autoencoder

Our initial approach was to embark on machine learning training using an autoencoder [13] neural network. This initial step allowed us to develop a preliminary model and record its accuracy. These findings served as a valuable reference for subsequent model training efforts, setting a foundational benchmark for comparing the effectiveness of more advanced algorithms introduced later in our research process.

This work was supported by the National Science Foundation (NSF) under grant 2303700. Cetin is also funded by Award Number: DE-SC0023715.

As illustrated in Fig. 1, the autoencoder network [14] is divided into two main components: the encoder parts and the decoder parts. The encoder and the decoder consist of a structure that starts with an input or output layer, followed by multiple dense (fully connected) layers, with ReLU activation functions linking these layers. A dense layer transforms an input vector $\mathbf{u} \in \mathbb{R}^p$ into an output vector $\mathbf{v} \in \mathbb{R}^q$ through the following relation:

$$\mathbf{v} = \phi(\mathbf{A}\mathbf{u} + \mathbf{b}), \tag{1}$$

where $\phi(\cdot)$ represents the selected activation function. $\mathbf{A} \in \mathbb{R}^{q \times p}$ denotes the transformation matrix, and $\mathbf{b} \in \mathbb{R}^q$ is the bias vector.

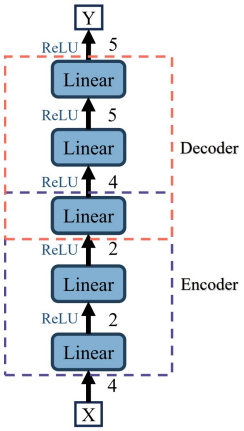


Fig. 1: Structure of the Autoencoder Neural Network.

B. DCT Neural Network

Frequency analysis in neural networks [15]–[20] can help identify which features are more significant. This information can guide feature selection or dimensionality reduction techniques, improving the efficiency and performance of the network. In this work, to obtain a better performance, we design a fully dense network with a DCT layer [9]–[11]. The DCT layer brings frequency-specific analysis and filtering, and it is inserted in the middle of the fully dense neural network as Fig. 2.

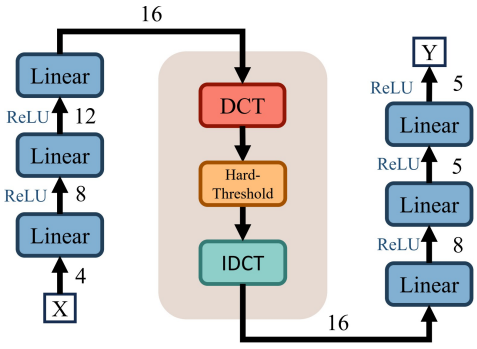


Fig. 2: Structure of the fully-dense neural network with a DCT layer

DCT possesses the ability to decorrelate a signal, meaning that after transformation, the correlation among DCT coefficients is significantly reduced compared to the correlation among the original signal samples. This decorrelation is particularly beneficial for analyzing MEMS data that contains a substantial number of zero values. By reducing the correlation, the DCT simplifies the complexity of the data, making it easier to identify and analyze the underlying patterns and characteristics within MEMS datasets. Furthermore, although the data changes insignificantly as their values are normalized into $[0, 1]$, DCT is relatively robust to minor variations in the signal. Therefore, a suitable DCT analysis can effectively assist the network in applications where the signal changes slightly.

In the DCT family, the type II DCT is the most widely used. For a time sequence $\mathbf{x} = [x_0, x_1, \ldots, x_N]$, its orthogonal DCT representation, $\mathbf{X} = [X_0, X_1, \ldots, X_N]$ is expressed as [21]:

$$\begin{aligned} \mathbf{X}[0] &= \sqrt{\frac{1}{N}} \sum_{n=0}^{N-1} \mathbf{X}[n], \\ \mathbf{X}[k] &= \sqrt{\frac{2}{N}} \sum_{n=0}^{N-1} \mathbf{X}[n] \cos \frac{(2n+1)k\pi}{2nN}, \end{aligned} \tag{2}$$

We apply a trainable hard-thresholding function in the discrete cosine frequency domain to suppress noise. Hard-thresholding [22]–[24] is a nonlinear adaptive filtering. It surpasses conventional linear adaptive filtering in achieving optimal solutions [25]. The hard-thresholding function is computed as:

$$F_\theta(x) = \begin{cases} x & \text{if } |x| > \theta, \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

We assign each entry a trainable threshold parameter θ. These threshold parameters are determined by the backpropagation algorithm [26]. This is how our DCT layer learns the frequency-specific patterns.

Consequently, the inverse DCT is applied on the tensor filtered by the hard-thresholding function $\tilde{\mathbf{X}}$ to convert it back to the time domain:

$$\tilde{x}_n = \sqrt{\frac{1}{N}} \tilde{X}[0] + \sqrt{\frac{2}{N}} \sum_{k=1}^{N-1} \tilde{X}[k] \cos \frac{(2n+1)k\pi}{2N}. \tag{4}$$

C. Parallel DCT Neural Network

To enhance the network's ability to learn different features simultaneously, we further extend the model in Fig. 2 into a parallel structure as Fig. 3. This improves the model's ability to capture complex patterns and relationships in the input data.

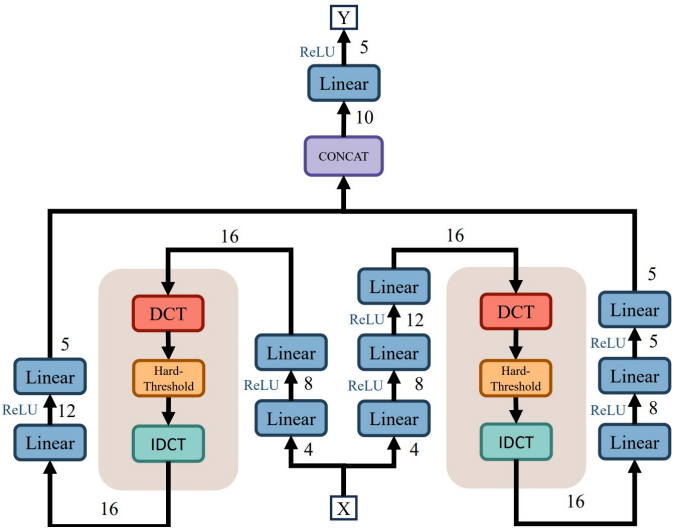


Fig. 3: Structure of the parallel fully-dense neural network with DCT layers

The parallel structure in this study consists of two streams. Each stream is built with several linear layers and a DCT layer. The outputs of the streams are concatenated to feed into a final linear layer to estimate the MEMS design parameter.

D. Multichannel Parallel DCT Neural Network

Given the excellent performance and application of DCT, we include an additional DCT layer to the model in Fig. 3, as shown in Fig. 4. This is our final model. We experimentally show that this additional stream reduces the estimation error.

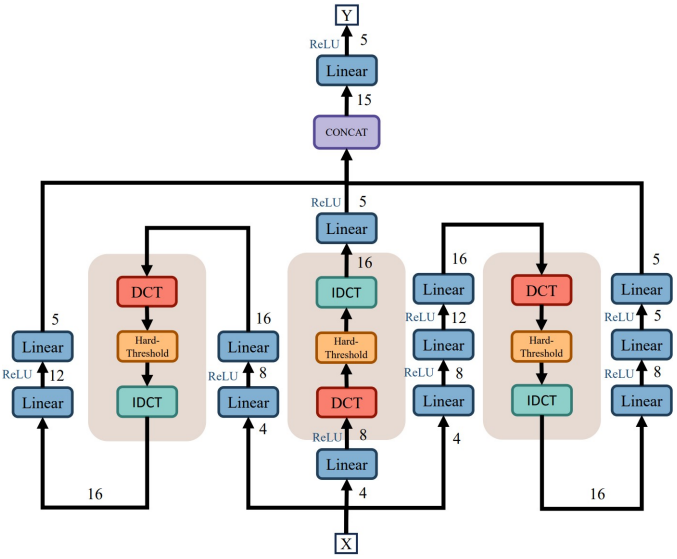


Fig. 4: Structure of the parallel fully-dense neural network with Multichannel Parallel DCT layers

III. Experimental Results

In this section, we provide a brief overview of our data collection methodology. Subsequently, we introduce the training processing and the accuracy calculation. Lastly, we analyze the results of the different models and compare the testing accuracy.

A. Dataset Description

The MEMS data in this study is corrected from the data presented in [27]–[40]. It employs radio frequency (RF) and direct current (DC) sputtering techniques on silicon wafers to deposit AlN. We gather this data for analysis purposes, including sputtering type, substrate type, sputtering power (W), substrate temperature ($^{\circ}C$), sputtering pressure (Pa), target to substrate distance (cm), rotation of substrate during deposition, Ex-situ annealing temperature ($^{\circ}C$), Ar flow rate (sccm), N_2 flow rate (sccm), Ar/N_2 mixture ratio, and AlN orientation.

The input variables include sputtering power, substrate temperature, sput pres, and target to substrate distance. In contrast, the output variables are AlN orientations represented by the 2θ (degrees) angle of 33°, 36°, 38°, 50°, 59° for orientations (100), (002), (101), (102), and (110) respectively, denoted as Variable; f_1, f_2, f_3, f_4 and f_5 as the Table I shows.

TABLE I: Output Variables

Variable Denotation	f_1	f_2	f_3	f_4	f_5
AlN Orientation	100	002	101	102	110
2θ (degrees)	33°	36°	38°	50°	59°

We apply Gaussian data augmentation to the initial dataset, expanding it by a factor of 8:

$$X_{\text{augmented}} = X + \left(\frac{X}{20}\right) \cdot \mathcal{N}(0, 0.1^2), \tag{5}$$

where X and $X_{\text{augmented}}$ are the original and the augmented data points, $\mathcal{N}(0, 0.1^2)$ denotes the Gaussian distribution with the mean of 0 and the standard deviation of 0.1. By introducing variability through Eq. (5), we simulate potential real-world variations in the dataset and increase the amount of our training and testing data. With such an augmentation, the dataset is expanded from 124 samples to 1116 samples.

B. Training Process Description

For the experiment, we first normalize the augmented data into the range from -1 to 1. Then, we divide the total of 1116 data points into 80% training and 20% test subsets. The experiments are implemented using PyTorch and Python 3. We choose the mean absolute error (MAE) as the loss function. Models are trained using the AdamW optimizer [41] under 20,000 epochs with a learning rate of 0.01. During the training, we save the best models based on the best test mean squared error (MSE) and present the results in Table II.

C. Analysis of Test Results

Since orientation is a set of eleven categorical outcomes ranging from 0 to 1 with a minimum of 10% deviation, We employ a threshold-based accuracy calculation method. By comparing the difference between the models' predictions and the target truth values, we calculate accuracy on individual test samples. Here, we extract the index of the highest probability category from each prediction and compare it to the actual labels. If the prediction differs within the minimum output deviation, it is considered a correct prediction. This value of 0.1 degree is determined by the precision of the dataset. The average accuracy is obtained by summing up all the correct predictions and then dividing by the number of samples N:

$$Accuracy = \frac{1}{N} \sum_{i=0}^{N-1} \#(|y_i - \hat{y}_i| \leq 0.1), \tag{6}$$

where $\#(\cdot) = 1$ if the inside condition is true and $\#(\cdot) = 0$ if it is false.

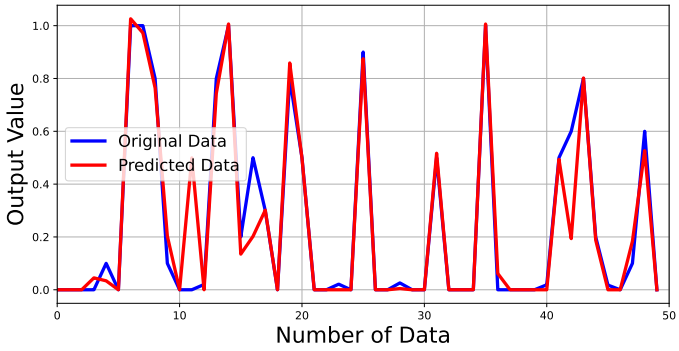
As Table II shows, the highest accuracy is achieved by the proposed multichannel parallel DCT neural network. It marks a significant accuracy improvement of 27.59% over the autoencoder network.

TABLE II: Testing result for ALN Deposition Parameter estimation

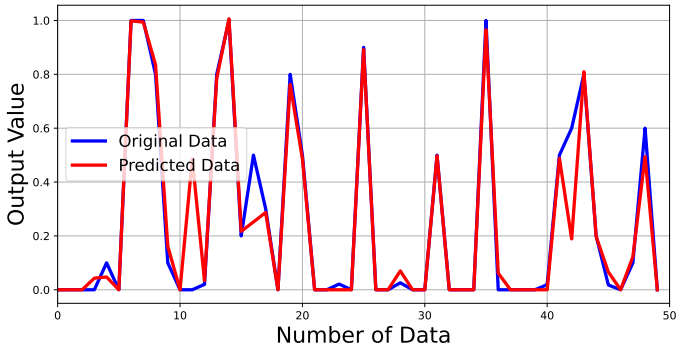
	Autoencoder	**DCT**	**Parallel DCT**	**Multichannel Parallel DCT**
MSE	0.0472	0.0201	0.0064	**0.0051**
MAE	0.1138	0.1008	0.0251	**0.0190**
Accuracy on f_1	0.1964	0.8036	0.9152	**0.9330**
Accuracy on f_2	0.1875	0.3036	0.8170	**0.8304**
Accuracy on f_3	0.9911	0.9911	0.9911	**0.9911**
Accuracy on f_4	1.0000	1.0000	1.0000	**1.0000**
Accuracy on f_5	1.0000	1.0000	1.0000	**1.0000**
Average Accuracy	67.50%	81.97%	94.47%	**95.09%**

Due to the data orientations of f_3, f_4, and f_5 being almost exclusively 0 values, the test accuracy differences among these three directions are insignificant. The orientations f_1 and f_2 contain fewer 0 values and show noticeable variations in detection accuracy across different models. Therefore, we primarily discuss the accuracy differences between the first two orientations, f_1 and f_2. Here, the 0 values in either different orientations or the same orientation are not due to missing data but are precise and necessary values. Hence, we cannot exclude the orientations f_3, f_4, and f_5 in training. Being able to train a high-accuracy network in an environment with a large number of 0 values also serves as a validation of our method's rationality and effectiveness.

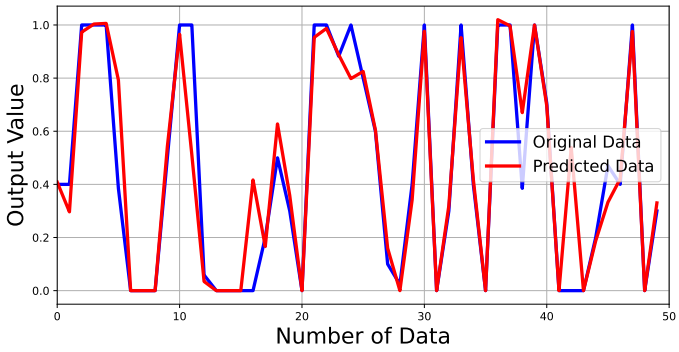
In the multichannel parallel DCT neural network, the addition of the third parallel layer significantly improved the prediction accuracy



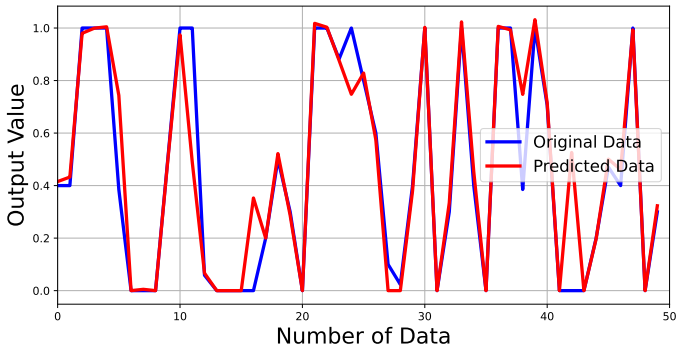
(a) Parallel DCT neural network comparison in f_1.



(b) Multichannel parallel DCT neural network in f_1.



(c) Parallel DCT neural network comparison in f_2.



(d) Multichannel parallel DCT neural network comparison in f_2.

Fig. 5: Comparison of original and predicted data output variables.

for AlN orientation, enhancing the f_2 channel's prediction accuracy to 83%, with an overall accuracy reaching 95%. Additionally, by comparing the predicted images of the f_2 channel, it can be observed from the comparison of the f_2 predicted images from different models and their actual values shown in the following figure that the multichannel parallel DCT neural network is the best predictor.

In the subsequent analysis, we delve into the comparison of model performance, specifically evaluating the results from Fig. 5. It is evident that within the number range of 15-20, the predicted trends by the multichannel parallel DCT neural network more accurately mirror the shape of the original data, with predictions closely aligning with the actual data points. Comparison of the original and predicted data output variable in f_2 shows that multichannel parallel DCT neural network fits the data better around 45.

In summary, experimental testing has demonstrated that the parallel DCT neural network and multichannel parallel DCT neural network models deliver stable and reliable predictions. Among these, the multichannel parallel DCT neural network model stands out for its superior accuracy.

IV. Conclusion

In this paper, we unveiled an advanced method for predicting the orientation changes of aluminum nitride (AlN) micromachining on silicon wafers, addressing the inherent challenges of this complex, time-intensive, and expensive process. Through the analysis of existing literature and data, we crafted a neural network that incorporated a distinctive discrete cosine transform (DCT) layer. This network precisely estimated the thickness of AlN across 5 specific orientations. Our multichannel parallel DCT neural network obtained superior predictive accuracy and minimized mean absolute error (MAE) and mean squared error (MSE). Utilizing machine learning techniques to predict a series of input parameters for achieving the desired single crystal orientation, our research demonstrated that our multichannel parallel DCT neural network could significantly elevate prediction accuracy to 95%. Our findings unlock a multitude of benefits. Aluminum nitride (AlN) is excellent for its piezoelectric qualities vital in creating MEMS transducers, unearthing the potential to control AlN thicknesses and orientations with machine learning approaches, which represents a significant breakthrough in our paper. This breakthrough doesn't just offer a smoother path to active device layer finesse. It also slashes the hefty time and costs tied to these processes, marking a significant leap in AlN micromachining research.

References

[1] Abid Iqbal and Faisal Mohd-Yasin. Reactive sputtering of aluminum nitride (002) thin films for piezoelectric applications: A review. *Sensors*, 18(6):1797, 2018.

[2] Morito Akiyama, Toshihiro Kamohara, Kazuhiko Kano, Akihiko Teshigahara, and Nobuaki Kawahara. Influence of oxygen concentration in sputtering gas on piezoelectric response of aluminum nitride thin films. *Applied Physics Letters*, 93(2), 2008.

[3] Stefan Bengtsson, Mats Bergh, Manolis Choumas, Christian Olesen Christian Olesen, and Kjell O Jeppson Kjell O Jeppson. Applications of aluminium nitride films deposited by reactive sputtering to silicon-on-insulator materials. *Japanese journal of applied physics*, 35(8R):4175, 1996.

[4] M Gillinger, M Schneider, A Bittner, P Nicolay, and U Schmid. Impact of annealing temperature on the mechanical and electrical properties of sputtered aluminum nitride thin films. *Journal of Applied Physics*, 117(6), 2015.

[5] MB Assouar, O Elmazria, L Le Brizoual, and P Alnot. Reactive dc magnetron sputtering of aluminum nitride films for surface acoustic wave devices. *Diamond and related materials*, 11(3-6):413–417, 2002.

[6] Gianluca Piazza, Valeriy Felmetsger, Paul Muralt, Roy H Olsson III, and Richard Ruby. Piezoelectric aluminum nitride thin films for microelectromechanical systems. *MRS bulletin*, 37(11):1051–1061, 2012.

[7] Yiping Chen, Ruzhi Wang, Bo Wang, Tao Xing, Xuemei Song, Mankang Zhu, and Hui Yan. Effects of mean free path on the preferentially orientated growth of aln thin films. *Journal of Crystal Growth*, 283(3):315–319, 2005.

[8] J.X. Zhang, H. Cheng, Y.Z. Chen, A. Uddin, Shu Yuan, S.J. Geng, and S. Zhang. Growth of aln films on si (100) and si (111) substrates by reactive magnetron sputtering. *Surface and Coatings Technology*, 198(1):68–73, 2005.

[9] Diaa Badawi, Agamyrat Agambayev, Sule Ozev, and A Enis Cetin. Discrete cosine transform based causal convolutional neural network for drift compensation in chemical sensors. In *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 8012–8016. IEEE, 2021.

[10] Hongyi Pan, Xin Zhu, Zhilu Ye, Pai-Yen Chen, and Ahmet Enis Cetin. Real-time wireless ecg-derived respiration rate estimation using an autoencoder with a dct layer. In *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1–5. IEEE, 2023.

[11] Xin Zhu, Hongyi Pan, Shuaiang Rong, and Ahmet Enis Cetin. Electroencephalogram sensor data compression using an asymmetrical sparse autoencoder with a discrete cosine transform layer. *arXiv preprint arXiv:2309.12201*, 2023.

[12] Hiroki Shiratori, Hideaki Goto, and Hiroaki Kobayashi. An efficient text capture method for moving robots using dct feature and text tracking. In *18th International Conference on Pattern Recognition (ICPR'06)*, volume 2, pages 1050–1053. IEEE, 2006.

[13] Jin Uk Ko, Kyumin Na, Joon-Seok Oh, Jaedong Kim, and Byeng D Youn. A new auto-encoder-based dynamic threshold to reduce false alarm rate for anomaly detection of steam turbines. *Expert Systems with Applications*, 189:116094, 2022.

[14] Chun Chet Tan and Chikkannan Eswaran. Performance comparison of three types of autoencoder neural networks. In *2008 Second Asia International Conference on Modelling & Simulation (AMS)*, pages 213–218. IEEE, 2008.

[15] Lu Chi, Borui Jiang, and Yadong Mu. Fast fourier convolution. *Advances in Neural Information Processing Systems*, 33:4479–4488, 2020.

[16] Hongyi Pan, Diaa Badawi, and Ahmet Enis Cetin. Fast walsh-hadamard transform and smooth-thresholding based binary layers in deep neural networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 4650–4659, 2021.

[17] Hongyi Pan, Diaa Badawi, and Ahmet Enis Cetin. Block walsh–hadamard transform-based binary layers in deep neural networks. *ACM Transactions on Embedded Computing Systems*, 21(6):1–25, 2022.

[18] Hongyi Pan, Diaa Badawi, Chang Chen, Adam Watts, Erdem Koyuncu, and Ahmet Enis Cetin. Deep neural network with walsh-hadamard transform layer for ember detection during a wildfire. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 257–266, 2022.

[19] Hongyi Pan, Xin Zhu, Salih Furkan Atici, and Ahmet Cetin. A hybrid quantum-classical approach based on the hadamard transform for the convolutional layer. In *International Conference on Machine Learning*, pages 26891–26903. PMLR, 2023.

[20] Xin Zhu, Daoguang Yang, Hongyi Pan, Hamid Reza Karimi, Didem Ozevin, and Ahmet Enis Cetin. A novel asymmetrical autoencoder with a sparsifying discrete cosine stockwell transform layer for gearbox sensor data compression. *Engineering Applications of Artificial Intelligence*, 127:107322, 2024.

[21] Nasir Ahmed, T_ Natarajan, and Kamisetty R Rao. Discrete cosine transform. *IEEE transactions on Computers*, 100(1):90–93, 1974.

[22] Licheng Jiao, Ronghua Shang, Fang Liu, and Weitong Zhang. Chapter 5 - theoretical basis of compressive sensing. In Licheng Jiao, Ronghua Shang, Fang Liu, and Weitong Zhang, editors, *Brain and Nature-Inspired Learning Computation and Recognition*, pages 109–126. Elsevier, 2020.

[23] Michael Elad, Mário A. T. Figueiredo, and Yi Ma. On the role of sparse and redundant representations in image processing. *Proceedings of the IEEE*, 98(6):972–982, 2010.

[24] Xiaojie Jin, Xiaotong Yuan, Jiashi Feng, and Shuicheng Yan. Training skinny deep neural networks with iterative hard thresholding methods. *arXiv preprint arXiv:1607.05423*, 2016.

[25] Xiao-Ping Zhang. Thresholding neural network for adaptive noise reduction. *IEEE transactions on neural networks*, 12(3):567–584, 2001.

[26] Robert Hecht-Nielsen. Theory of the backpropagation neural network. In *Neural networks for perception*, pages 65–93. Elsevier, 1992.

[27] Hiroshi Okano, Yusuke Takahashi, Toshiharu Tanaka, Kenichi Shibata Kenichi Shibata, and Shoichi Nakano Shoichi Nakano. Preparation of c-axis oriented aln thin films by low-temperature reactive sputtering. *Japanese Journal of Applied Physics*, 31(10R):3446, oct 1992.

[28] A. Ababneh, U. Schmid, J. Hernando, J.L. Sánchez-Rojas, and H. Seidel. The influence of sputter deposition parameters on piezoelectric and mechanical properties of aln thin films. *Materials Science and Engineering: B*, 172(3):253–258, 2010.

[29] Harish C. Barshilia, B. Deepthi, and K.S. Rajam. Growth and characterization of aluminum nitride coatings prepared by pulsed-direct current reactive unbalanced magnetron sputtering. *Thin Solid Films*, 516(12):4168–4174, 2008.

[30] Xiao-Hong Xu, Hai-Shun Wu, Cong-Jie Zhang, and Zhi-Hao Jin. Morphological properties of aln piezoelectric thin films deposited by dc reactive magnetron sputtering. *Thin Solid Films*, 388(1):62–67, 2001.

[31] A Ababneh, Ulrich Schmid, J Hernando, JL Sánchez-Rojas, and Helmut Seidel. The influence of sputter deposition parameters on piezoelectric and mechanical properties of aln thin films. *Materials Science and Engineering: B*, 172(3):253–258, 2010.

[32] Hao Jin, Bin Feng, Shurong Dong, Changjian Zhou, Jian Zhou, Yi Yang, Tianling Ren, Jikui Luo, and Demiao Wang. Influence of substrate temperature on structural properties and deposition rate of aln thin film deposited by reactive magnetron sputtering. *Journal of electronic materials*, 41:1948–1954, 2012.

[33] Hao Cheng, Yong Sun, and Peter Hing. Microstructure evolution of aln films deposited under various pressures by rf reactive sputtering. *Surface and coatings technology*, 166(2-3):231–236, 2003.

[34] B Wang, YN Zhao, and Z He. The effects of deposition parameters on the crystallographic orientation of ain films prepared by rf reactive sputtering. *Vacuum*, 48(5):427–429, 1997.

[35] Hwan-Chul Lee, Guen-Hong Kim, Soon-Ku Hong, Ki-Young Lee, Yoon-Joong Yong, Chang-Hwan Chun, and Jai-Young Lee. Influence of sputtering pressure on the microstructure evolution of aln thin films prepared by reactive sputtering. *Thin solid films*, 261(1-2):148–153, 1995.

[36] Qiu-ping Wei, Xiong-wei Zhang, Dan-ying Liu, LI Jie, Ke-chao Zhou, Dou Zhang, and Zhi-ming Yu. Effects of sputtering pressure on nanostructure and nanomechanical properties of aln films prepared by rf reactive sputtering. *Transactions of Nonferrous Metals Society of China*, 24(9):2845–2855, 2014.

[37] HY Liu, GS Tang, F Zeng, and F Pan. Influence of sputtering parameters on structures and residual stress of aln films deposited by dc reactive magnetron sputtering at room temperature. *Journal of crystal growth*, 363:80–85, 2013.

[38] Hsyi-En Cheng, Tien-Chai Lin, and Wen-Chien Chen. Preparation of [0 0 2] oriented aln thin films by mid frequency reactive sputtering technique. *Thin Solid Films*, 425(1-2):85–89, 2003.

[39] Jyoti P Kar, G Bose, and S Tuli. Correlation of electrical and morphological properties of sputtered aluminum nitride films with deposition temperature. *Current Applied Physics*, 6(5):873–876, 2006.

[40] CC Wang, MC Chiu, MH Shiao, and FS Shieu. Characterization of aln thin films prepared by unbalanced magnetron sputtering. *Journal of The Electrochemical Society*, 151(10):F252, 2004.

[41] Rajesh Prasad, Akpan Uyime Udeme, Sanjay Misra, and Hashim Bisallah. Identification and classification of transportation disaster tweets using improved bidirectional encoder representations from transformers. *International journal of information management data insights*, 3(1):100154–, 2023.