Key Role Players in Artificial Neural Networks: An Overview

MINAKSHI PRADEEP ATRE
Department of Artificial Intelligence and Data Science,
PVG's COET & GKPIM Pune,
Maharashtra,
INDIA

ORCiD: 0000-0002-5194-201X

Abstract: - The computational world of machine learning (ML) has been transformed by two simple key role players in the signal processing domain: *sampling* and *convolution*. Sampling and convolution are highly mathematical yet are the most significant signal-processing techniques. Digital signal processing has reduced the complexity of the analog processing with approximations and encouraged intelligent humans to consider these powerful role players, sampling, and convolution. Their applications are diverse and affect the world we see today. Sampling has changed the way the data are processed, stored, and transmitted. The sub-processes of sampling, such as interpolation and decimation, have the advantage of changing the sampling rate within a system and work effectively in the wavelet transform. It helps to store the two versions of the digital image with approximate and detailed coefficients and achieves a remarkable compression of data in this high-resolution world with the help of multi rate sampling. On the extended line, convolution has played a large role in identifying features and has led to a deep understanding of human intelligence and cognitive science. The understanding of the features of deep learning has been strongly affected by convolution. This article focuses on these two signal-processing techniques and their role in the transformation of machine learning algorithms into deep learning techniques.

Key-Words: - Sampling, convolution, machine learning, interpolation and decimation, wavelet transform, multi-rate sampling, deep learning.

Received: June 22, 2024. Revised: January 23, 2025. Accepted: March 24, 2025. Published: May 28, 2025.

1 Introduction

It is an era of artificial intelligence (AI), and the world is fascinated by machine learning (ML) applications and deep learning strategies to understand and represent the world by mimicking human learning. The people are mesmerized by large language models (LLMs), such as ChatGPT, which interact and help humans solve seemingly any problem in the world. Copilot and Gemini are a few of the arrows in the quiver! They are changing the way we converse; we see the world and drive the automobiles!, [1] Lives are surrounded by AI fueled recommender systems, autonomous cars and virtual assistants. This revolution did not occur overnight and started taking shape with the understanding of powerful signal processing role players such as sampling and convolution, [2].This paper focuses on providing an overview of key signal-processing operations and their significance in the domain of artificial intelligence (AI).

Sampling, the first key role player, is a simple process of converting an analog signal into a digital signal. The analog signal is converted into a digital signal because of the noise immunity of the digital signals. The digital signal is filtered and converted back to an analog signal *via the Whittaker–Shannon–Kotelnikov interpolation* formula. [3] The sampling theorem underlines the principle that any band-limited signal can be converted into a digital signal by using twice the maximum frequency available in the signal present and recovered via the *Shannon interpolation* equation. Similarly, *convolution* is the second key role player in the signal-processing domain. The convolution filters and extracts the desired information from speech, image, or video signals. To propose a simple understanding of the signal, an airflow interacts with the vocal tract of the human, and the interaction produces speech, so speech is the convolved signal!!, [4]. These two role players are involved in the success of deep learning, the subdomain of AI.

Following paragraphs discuss the milestones in the field of deep learning.

The ImageNet Large Scale Visual Recognition Challenge (ILSVRC), an annual competition started since 2010, emerged with an award-winning model, VGG-16. It is an object classification model for an image dataset of 1000 classes. It changed the way, people used machine learning for classification purpose. The ML algorithms required feature calculation and were restricted to hundreds to a few thousand samples. VGG-16 model made the feature calculation step redundant and this was due to convolution!! The convolution layers in VGG-16 enhanced the understanding of the objects and resulted in increased validation accuracy. Another mathematical operation, max pooling, variation of sampling, was also used in VGG-16 [5] to reduce the size of the feature vector.

Certainly, sampling and convolution played important roles in the development of deep learning models and increased validation accuracy. The revolutionary deep learning algorithms are balanced with three pillars, namely, data, computational complexity, and equally complex algorithms, using billions of parameters.

Data is oil and researchers started to sense the power of data when every minute 500 hours of videos started getting uploaded on YouTube. The key role players, sampling layers and convolution layers helped these algorithms achieve increasingly high accuracies through the addition of many layers with billions of parameters, [6]. Convolution changed the game of feature extraction and led to the development of deep learning (DL) algorithms. Deep learning algorithms involve science, technology, engineering, and mathematics (STEM) to understand and interpret the data for a variety of applications. To develop a medical database model, *BioBERT*, ML engineers, along with biomedical domain experts, are needed.

This article provides an overview of the key players and their role in the subfields of AI. The sampling pulled the trigger of high compression with great resolution. [7], whereas the convolution completely flipped the traditional "*learning by feature extraction or calculation*" on its head.

The subsection 1.1 discusses the fundamentals of the first role player, sampling with the *Whittaker–Nyquist–Shannon theorem* whereas subsection 1.2 introduces the second role player, convolution along with its mathematical expression. It also explains the role of convolution in time domain analysis. *Section 2* provides an overview of AI, machine learning (ML), and deep learning (DL). *Section 3* summarizes the sampling and convolution techniques and highlights their role in today's powerful domain of artificial intelligence. The operations, especially the convolution, are the game changers in deep learning models.

1.1 Sampling

A signal is any varying quantity with respect to time. The time has two versions: continuous and discrete. The continuous-time signal is represented with respect to the continuum set of time values. The discrete-time signal is represented with reference to integer values of the time index. The sampling theorem converts the continuous-time (CT) signal into a discrete-time (DT) signal. It is a simple switch ON-OFF operation.

The *Whittaker–Nyquist–Shannon theorem* [8] is the bridge between the continuous-time signal and the discrete-time signal. The theorem summarizes this connection as follows: "*When one reduces a continuous function to a discrete sequence and interpolates back to a continuous function, the fidelity of the result depends on the density (or sample rate) of the original samples.*" Sampling changed the way; signal processing operations were implemented earlier. The digital signal arising from sampling introduced a 360-degree shift in signal processing, [9]. Figure 1 depicts the sampling process.

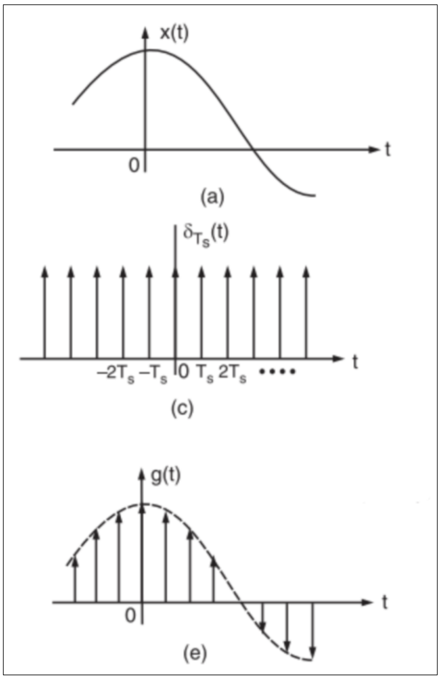

Fig. 1: Visualization of the Sampling Theorem

In Figure 1(a) is the sinusoidal signal and Figure 1(c) is the train of the pulse signal. When the signal is multiplied by the train of pulses, it is a simple ON–OFF operation which results in a discrete wave, as shown in (e). This is alternatively called a digital signal. Digital signal processing changes the method of processing and storing the data. The best examples include the old days' compact discs (CDs) and the digital versatile discs (DVDs). The audio-video recording and playback with subsampled rates was the miracle of that time!

Another important concept, named, subsampling, shown in Figure 2 was used to save the bandwidth of television (TV) signal transmission. The digital TV transmission achieved new dimensions with the help of chroma subsampling process. The chroma subsampling helps transmit the two-color signals instead of the three-color signals. This saves bandwidth but retains quality with the help of digital signal processing. Figure 2, adopted from [10] shows the scanning method and depicts how a subsampled signal is represented by Y, Cb, and Cr signaling components.

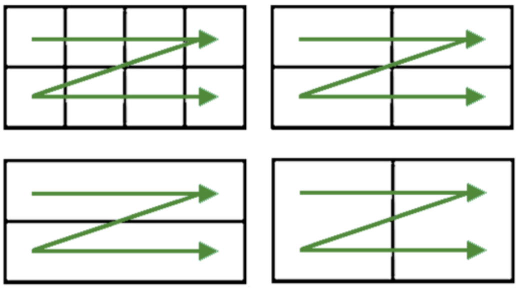

Fig. 2: Subsampling of digital video signals
Top left: *subsampling(Y-Cb-Cr) =4:4:4 (no sampling),*
Top right: *subsampling (Y-Cb-Cr) = 4:2:2;*
Bottom left: *subsampling (Y-Cb-Cr) = 4:1:1;*
Bottom right: *subsampling (Y-Cb-Cr) = 4:2:0*

The signal Y is the *luminance* (or brightness) signal, and Cb & Cr are the chrominance (or color) signals. Cb and Cr are notations for the *digital blue* and *red* chromatic signals, respectively. The same subsampling method, in the form of *pooling,* helps reduce the size of the convolved image and saves the computations of the deep learning models.

1.2 Convolution

Signal analysis is the first step in the highly computational world of signal processing. The continuous-time signal and discrete-time signals present two parallel sets of analysis tools. Figure 3 summarizes the signal analysis techniques for continuous-time signals, and Figure 4 summarizes the signal analysis techniques for discrete-time signals. The analysis provides an in-depth understanding of the behavior of signals. The time domain analysis is carried out with the help of (i) a solution to the difference equations and (ii) convolution.

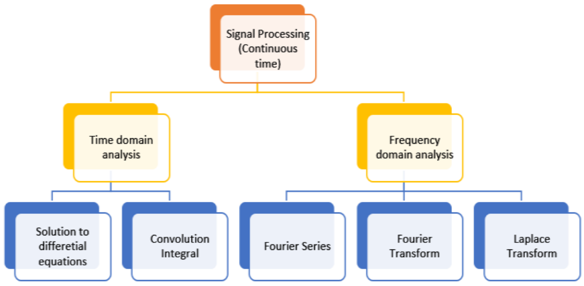

Fig. 3: Continuous-time signal analysis

There are two major domains for signal representation, time and frequency. Solution to differential equations and convolution are time-domain analysis tools. The frequency cannot be observed in time domain analysis; hence, it is transformed into the frequency domain. The transformation gives rise to Fourier tools: Fourier series (used for periodic signal representation), Fourier transforms (mainly for periodic signals but also used for both types of signals) and the Laplace transform (typically used for checking the stability and causality of the signals).

The solution to the differential equation presents an easy way to analyze the signal and express the solution in the form of extremely smooth functions such as sines and cosines. A sample mathematical formula for the differential equation is given by Eq. (1) and the solution to it with an understanding of the exponential term.

$$\frac{dy}{dt} = \frac{d\varkappa^2}{dt^2} + \frac{d\varkappa}{dt} \tag{1}$$

$$y(t) = C_1 \cdot e^{-2t} + C_2 \cdot e^{-t}$$

The solution has exponential terms indicating Euler's identity. The term $e^{-j\Theta}$ exists in the solution to the difference equation. This exponential term has a combination of extremely smooth functions, sines and cosines. This is exactly the same as Fourier's theory proposed in [11]: any signal can be represented in the form of sines and cosines. This was opposed by Sir Laplace but was later approved by Sir Lagrange. This difference of opinion appeared as Gibb's phenomenon on the canvas of the Fourier series. The Fourier series is the transform necessary to understand the behavior of a signal in the frequency domain. This is useful when the time domain analysis is bounded by only time-stamping and does not provide information about the frequency contents. Figure 4 summarizes the same set of tools for discrete-time signals.

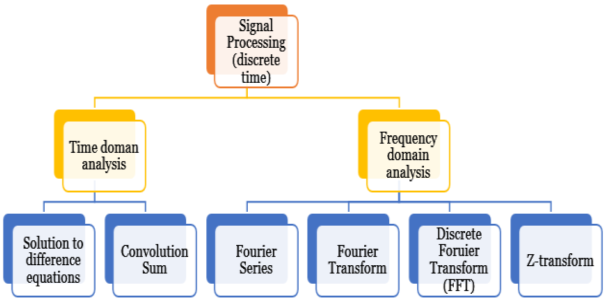

Fig. 4: Discrete-time signal analysis

The differential equation has a counterpart in the discrete domain: difference equations. Eq. (2) is the mathematical representation of the difference equation for a linear time-invariant (LTI) system. y[n] is the output, and x[n] is the input to an LTI system.

$$y[n] = -\sum_{k=1}^{N} a_k y[n-k] + \sum_{m=0}^{M} b_m x[n-m] \tag{2}$$

The solution h(n), is known as the impulse response of that LTI system. This involves the most significant mathematical operation known as convolution. It appeared to be a complex process in the analog domain, but its digital convolution version is much simpler and understandable. The resulting system can be represented in Figure 5 and the system response is represented by h[n]. When the signal δ[n] is applied to a system, the output is h[n]. And when system's response, h[n] is known, the output of that system, to any input, x[n], can be calculated via the convolution of x[n] and h[n]. This demonstrates, '*Convolution is filtering.*'

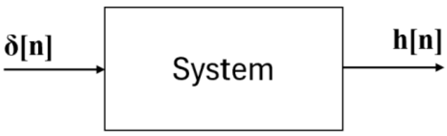

Fig. 5: System response

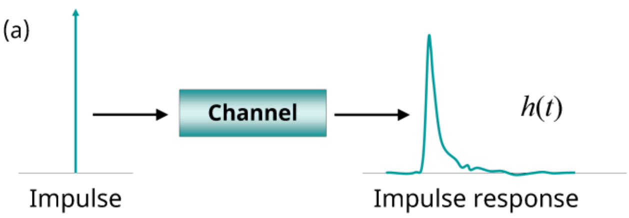

Fig. 6: Impulse and impulse response

Figure 6 depicts the impulse as the input given to a system and the impulse response of a system. This approach is very effectively used in RADAR (radio detection and ranging) to detect obstructions in the sailing path of ships.

Figure 7 depicts the LTI system representation with x[n] as the input and y[n] as the output.

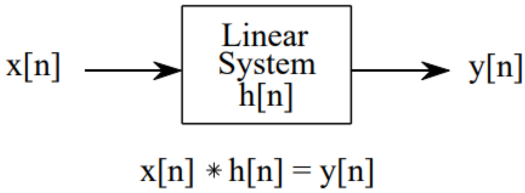

Fig. 7: LTI system representation

$$y[n] = \sum_{k=-\infty}^{\infty} x[k]h[n-k] \tag{3}$$

Eq. (3) is the mathematical representation of the convolution. Precisely, it is called the convolution sum. To illustrate the role of convolution in system representation and understanding, an example of speech is perfect. Speech is a convolved signal. Convolution occurs between vocal tract and an air in the mouth cavity. A music note is generated in similar fashion. Let us dive into the world of music signal processing.

A generated music note, shown in Figure 8, [12] is also the convolution of the excitation signal (plucking of a string) and the impulse response of the guitar body. Convolution can be used to generate a musical note of an instrument that is on the verge of obsolescence. Such instruments can be preserved in terms of the impulse responses of those musical instruments.

Fig. 8: Acoustic guitar player, [12]

The article [13] provides a discussion on how to generate the music note from the restored impulse response. A simple algorithm is shown in Figure 9. A guitar is plucked with either a *finger* or a *plectrum* (short-formed as pick). It produces the excitation signal. It convolutes with the guitar body and generates a beautiful musical note. The guitar's body has a typical response, known as the impulse response.

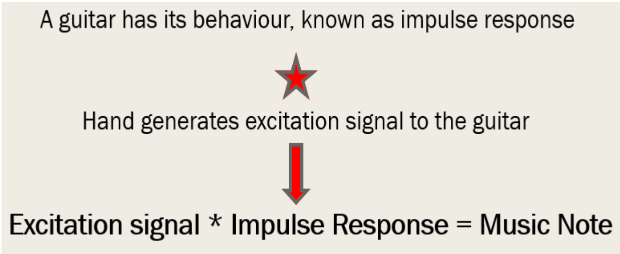

Fig. 9: Generalized model for music note, [13]

Each instrument has an impulse response. Different instruments are played and perceived by their musical notes. Musical notes have frequency and loudness. Further, the understanding of whether a guitar or violin is played, is based on the time envelope of the music note generated. It is known as the timbre of the music signal. Music notes are perceived by their frequencies and Fourier analysis provides the frequency spectrum of these music notes. The fundamental frequency of the music note is known as the 'pitch'. Pitch and timbre are time domain characteristics of music or speech signals.

To summarize this section and to understand the convolution in the simplest way, you are waiting for a friend at the railway platform; he or she comes and you spot him or her correctly from the enormous Crowd—that's convolution. The convolution senses the features of one's friend. The convolution filters your friend from the Crowd. The convolution is an understanding of the desired object. A musical note is a convolved signal. Our speech is a convolved signal. A guitar note is generated by the convolution of a string plucked by the player and the guitar body. A speech signal is generated by the convolution of the vocal tract response and the air blown/thrown off the mouth.

2 Overview of Neural Networks with Reference to Sampling and Convolution: AI, ML and DL

The conceptual understanding of two key role players is essential for their extended work in the artificial intelligence domain. This section provides an overview of the signal processing operations named sampling and convolution in the domains of AI, ML, and DL.

2.1 How AI is Ruling

Machine learning has three categories of algorithms: *supervised, unsupervised, and reinforcement learning*. The supervised algorithms run on the labeled data whereas the unsupervised algorithms run on the unsupervised data. The third category of reinforcement learning occupies the field of game theory. Reinforcement learning (RL) is in the form of rewards and punishments.

Fig. 10: Garry Kasparov vs Deep Blue (IBM), [14]

RL became so powerful that many games, such as Chess and AlphaGo, developed with AI, defeated the champions like Garry Kasparov and Lee Sedol respectively. The defeat of Garry Kasparov was evidence of winning machines over human intelligence. IBM's supercomputer Deep Blue won and set the record. Figure 10 is sourced from the related article [14] for understanding the milestone of AI revolution.

The next section provides an overview of the machine learning algorithms, based on feature calculations, and deep learning algorithms, eliminating the need of feature calculations.

2.2 Machine Learning and Feature Extraction

The Iris flower classification is the "*Hello World*" program in the domain of machine learning. This method is based on simple four features such as petal length, petal width, sepal length, and sepal width. Machine learning flips traditional programming on its head. Although it outperforms traditional programming, ML requires feature calculations.

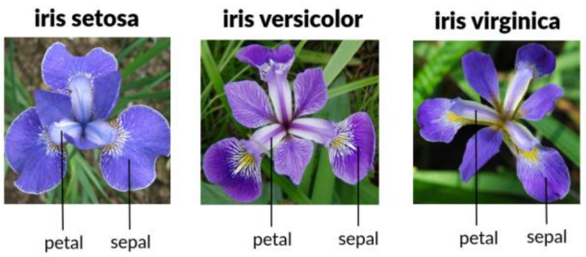

Fig. 11: Iris family of flowers, [15]

Table 1 shows the sample entries of features of the Iris flower classification project. Prof Fisher demonstrated the classification Iris flowers for three sub-classes, *setosa, versicolor and virginica* as shown in Figure 11, [15]. There were only 150 samples of the Iris family with simple features as length and width. It was definitely a milestone in the domain of ML.

Table 1. Features of iris flower classification

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
0	1	5.1	3.5	1.4	0.2	Iris-setosa
1	2	4.9	3.0	1.4	0.2	Iris-setosa
2	3	4.7	3.2	1.3	0.2	Iris-setosa
3	4	4.6	3.1	1.5	0.2	Iris-setosa
4	5	5.0	3.6	1.4	0.2	Iris-setosa

2.3 Role of Convolution in Deep Learning

The Iris flower classification is an exemplary work based on feature calculation. The convolution i.e. filtering solved the complex problem of feature calculation by making it redundant. The MNIST Handwritten digit recognition, recognized as the "*Hello World*" program in the deep learning domain, used convolution layers for the effective object recognition. The architecture [16] of CNN is shown in Figure 12. The figure depicts the role of convolution for feature understanding based on edges and textures. The image of a Koala is fed to the network and the convolution layers will extract the features like eye, nose, ear etc. The task of feature calculation becomes challenging when it comes to different animal categories. This is implemented effectively with CNN.

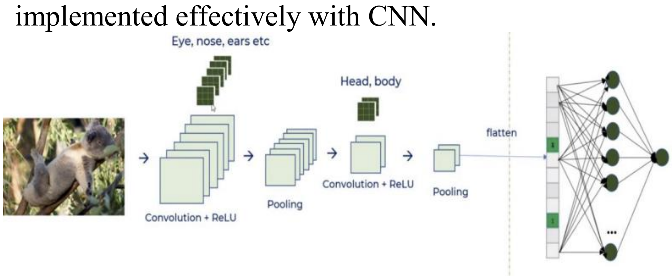

Fig. 12: Deep Learning Architecture [16]

While machine learning requires features to be calculated for the classification of objects, with an abundance of data, deep learning overcomes this need for feature calculation. Thousands of samples are presented to the neural network, and the hidden layers work efficiently to understand the features. A simple artificial neural network (ANN) and a deep neural network (deep NN) are differentiated by the hidden layers. The middle layers in the ANN are called the NN layers, while these layers are called the convolution layers in the deep NN. They extract the features on their own and learn the pattern or class from the fed samples.

In addition, that's where the ***convolution*** comes in! Convolution involves filtering, and convolution senses the data to understand the features. They are not known to the NN designer, but understanding is complete after the networks are exposed to large amounts of sample data. While evaluating the performance of the NN models, training and validation accuracies are checked. If the training accuracy is greater than the validation accuracy, then it is evidence of ***overfitting***, which is a famous term in the domain of deep learning. Overfitting is memorization. To overcome this overfitting, a CNN (convolutional neural network) is used which increases the accuracy to a great degree, as the convolution layers extract the features for the best understanding of the objects.

2.4 Role of Sampling in Neural Networks: Sampling and Subsampling

Pooling in the convolutional network is a down-sampling or subsampling technique. It minimizes the number of neurons in the convolutional layer by extracting important information. Different types of pooling exist in deep CNNs to help this feature extraction of the objects. *Subsampling* also reduces image matrix size by selecting a subset of the original data. It is performed on structured data, as unstructured data have inherent complexity. The block schematic shown in Figure 13 depicts the role of *pooling* in the deep learning algorithm, [17]. The deep learning algorithms need thousands of input samples, as they learn the features on their own and do not need explicit feature calculations. The data are large, and the computational complexity needs to be reduced. This is done with subsampling processes like pooling.

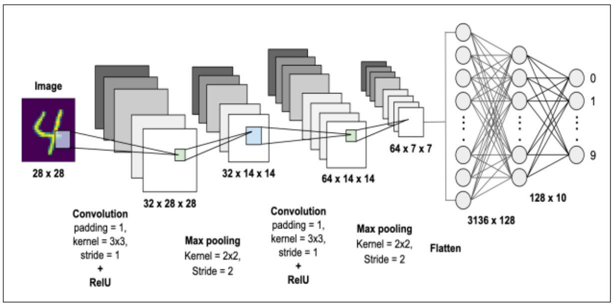

Fig. 13: Concept of convolution neural network (CNN), [17]

3 Summary

The author presented an overview of the key role players in the domain of an artificial neural network with the required mathematical background of sampling and convolution. It is very important to understand the mathematical operations that revolutionized the world of ANNs. The transformation of deep learning occurred because of the convolutional layers. This was evident in the ImageNet Challenge held in 2014. The winning architecture, the VGG-16 model, works on 1000 classes for object classification via a convolutional neural network. Convolution increased the understanding of the texture, edges, and color of the images and increased the validation accuracy. The convolutional neural network (CNN) based models make the feature calculation redundant. For deep learning model deployment, the validation accuracy of the model must be equal to or possibly greater than the training accuracy. This is feasible with convolution and subsampling processes of signal processing. The data, computational power, and complex algorithms are catalysts for deep learning models. Sampling and convolution, play important roles in the fast-changing world of AI, resulting in higher accuracies than simple machine learning algorithms. Thus, the purpose of this study is to provide a quick review of signal processing operations in the domain of ANNs.

References:

[1] Yiheng Liu, Hao He, Tianle Han, Xu Zhang, Mengyuan Liu, Jiaming Tian, Yutong Zhang, Jiaqi Wang, Xiaohui Gao, Tianyang Zhong, Yi Pan, Shaochen Xu, Zihao Wu, Zhengliang Liu, Xin Zhang, Shu Zhang, Xintao Hu, Tuo Zhang, Ning Qiang, Tianming Liu, Bao Ge (2024). Understanding LLMs: A comprehensive overview from training to inference. arXiv preprint arXiv:2401.02038. Available: http://arxiv.org/abs/2401.02038.

[2] S. D. Casey, M. M. Dodson, P. J. S. G. Ferreira, and A. Zayed, Eds., "Sampling, Approximation, and Signal Analysis," 2023, doi: 10.1007/978-3-031-41130-4.

[3] Dokuchaev, Nikolai. "Sampling Theorem and interpolation formula for nonvanishing signals." *arXiv preprint arXiv:2405.10007* (2024).

[4] Apte, S.D., Shridhar (2011). "An Efficient Speech Enhancement Algorithm Using Conjugate Symmetry of DFT." In: Zhu, M. (eds) *Electrical Engineering and Control. Lecture Notes in Electrical Engineering*, vol 98. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-21765-4_87.

[5] Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition." *arXiv preprint* https://doi.org/10.48550/arXiv.1409.1556.

[6] Laith Alzubaidi, Jinglan Zhang, Amjad J. Humaidi, Ayad Al-Dujaili, Ye Duan, Omran Al-Shamma, J. Santamaría, Mohammed A. Fadhel, Muthana Al-Amidie & Laith Farhan, "Review of deep learning: concepts, CNN architectures, challenges, applications, future directions," *Journal of Big Data*, 2021 8:1, vol. 8, no. 1, pp. 1–74, Mar. 2021, doi: 10.1186/S40537-021-00444-8.

[7] W. J. Schroeder and K. M. Martin, "*Overview of Visualization*", Accessed: Aug. 12, 2024, [Online]. https://booksite.elsevier.com/samplechapters/9780123875822/9780123875822.PDF (Accessed Date: December 27, 2024).

[8] P. Dehghani, S. Shirani, and X. Wu, "On interpolation and resampling of discrete data,", *IEEE Signal Process Lett.*, vol. 13, no. 12, pp. 733–736, Dec. 2006, doi: 10.1109/LSP.2006.879854.

[9] H. Al-Hammali and A. Faridani, "Uniform and nonuniform sampling of bandlimited functions at minimal density with a few additional samples", *Sampling Theory, Signal Processing, and Data Analysis*, vol. 21, no. 1, pp. 1–29, Jun. 2023.

[10] Pokorny, P. "Using Chroma Subsampling in Lossy Compression." *Mathematical Models and Computational Methods* 128.1 (2015): 134-137, [Online]. https://inase.org/library/2015/crete/bypaper/AME/AME-19.pdf (Accessed Date: December 27, 2024).

[11] Max E. Valentinuzzi, “Highlights in the History of the Fourier Transform - IEEE Pulse”, [Online]. https://www.embs.org/pulse/articles/highlights-in-the-history-of-the-fourier-transform/ (Accessed Date: August 12, 2024).
[12] Figure 8 Acoustic guitar, [Online]. https://www.shutterstock.com/image-photo/female-hand-playing-acoustic-guitar-playclose-127635929 (Accessed Date: December 3, 2024).
[13] M. P. Atre and S. D. Apte, “Generalized modeling of body response of acoustic guitar for all frets using a response for single fret,” Applied Acoustics, vol. 145, 2019, DOI: 10.1016/j.apacoust.2018.10.017. (Figure 9).
[14] Figure 10 Garry Kasparov vs Deep Blue (IBM) , [Online]. https://dotesports.com/chess/news/world-chess-champions-ranked (Accessed Date: December 3, 2024).
[15] Figure 11 Iris flower family, [Online]. https://www.analyticsvidhya.com/blog/2022/06/iris-flowers-classification-using-machine-learning/ _(Accessed Date: December 3, 2024).
[16] Figure 12 Deep Learning Architecture https://oldsite.jssstuniv.in/wp-content/uploads/2022/11/4_Deep-learning.pdf
[17] Figure 13 Convolution neural network, [Online]. https://towardsdatascience.com/mnist-handwritten-digits-classification-using-a-convolutional-neural-network-cnn-af5fafbc35e9 (Accessed Date: December 3, 2024).

Contribution of Individual Authors to the Creation of a Scientific Article (Ghostwriting Policy)

The article is single-authored who contributed in the present work, at all stages from the formulation of the problem to the final findings and solution.

Sources of Funding for Research Presented in a Scientific Article or Scientific Article Itself

No funding was received for conducting this study.

Conflict of Interest

There is no conflict of interest to declare that is relevant to the content of this article.