International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Morse Code Decoding Using Sound Waves in Machine Learning

G. Bhavani[a], G. Ashok Kumar[b], J. Sharat Kumar[c], K. Tarun[d]*

[a]Department of Infromation Technology, GMRIT, Rajam, India - 532127
[b]Department of Infromation Technology, GMRIT, Rajam, India – 532127
[c]Department of Infromation Technology, GMRIT, Rajam, India – 532127
[d]Department of Infromation Technology, GMRIT, Rajam, India - 532127

ABSTRACT :

Morse code, a time-tested method of communication, it remains relevant in certain modern applications such as telecommunication, aviation, military communication, and emergency signalling. It uses dots and dashes (on/off tones) that can be interpreted by a trained operator to represent letters and numbers. The code was designed to be useful as shorthand symbols for common English letters. This paper proposes a novel approach to decode Morse code from audio signals using a combination of Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks. The model combines CNNs for pattern recognition and LSTMs for sequence understanding to handle complex tasks effectively. CNNs excel at extracting local features from the time-frequency representation of audio signals, while LSTMs are adept at capturing long-range dependencies in sequential data. By combining these techniques, the model aims to accurately identify and classify Morse code symbols, even in noisy or distorted audio environments. This involves recognizing the patterns in morse code and translating them into text characters. The models are used to enhance the accuracy of morse code decoding by effectively processing audio signals and translating them into text. LSTM which is particularly effective for processing sequential audio signals and can remember the information for longer period making them to recognize patterns in morse code audio sequences. Combining CNN and LSTM is expected to yield the best accuracy in decoding the code and also improving the research gaps in which it failed in recognizing some information.

Keywords — Morse Code, CNN, LSTM, Audio Signal, Emergency Signaling, Machine Learning, Shorthand Symbols.

1. Introduction :

Morse code, a vital communication method dating back to the 1830s, has retained its importance across sectors like emergency response, aviation, and military communication due to its reliability in noisy or challenging environments. Traditionally transmitted through on-off signals via electrical, radio, or visual mediums, Morse code is highly effective when other communication forms might fail. The auditory transmission of Morse code, encompassing dots and dashes representing characters and numbers, is generated by various sounds like taps, beeps, and honks, making it a versatile medium for communication. However, decoding Morse code from diverse sound sources poses significant challenges, especially in noisy environments or with fluctuating signal intensity. This project, which uses machine learning and signal processing techniques, aims to develop a robust Morse code decoding model that can identify Morse signals in various sound sources, translating them into English text accurately. This project falls within the domains of machine learning, deep learning. To address the complexities of decoding Morse code from multiple sound sources, a combination of Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks was used. CNNs are known for their feature extraction capabilities, especially in handling spatial data such as images or sound waves transformed into spectrograms. In this case, CNN layers help capture features within the amplitude-time graphs of audio signals, enabling the model to identify the distinct features of Morse code (dots and dashes). Meanwhile, LSTMs, a type of recurrent neural network, excel in sequential data processing, crucial for Morse code patterns that require understanding dependencies over time. LSTMs capture the rhythm and timing of signals, distinguishing between short and long tones (dots and dashes) and ensuring sequential accuracy in translating audio waves into text. Decoding Morse code from diverse sound sources presents several challenges, primarily due to variation in sound source characteristics and noise interference. Different sound sources, like tapping on a table or honking a car horn, produce unique acoustic signatures that alter the amplitude and frequency of dots and dashes, making it challenging for a model to generalize across sources. Additionally, human-generated Morse code introduces temporal and amplitude randomness—for instance, a dot may vary slightly in duration or loudness across transmissions, requiring a model that can generalize rather than memorize exact patterns. Furthermore, background noise and signal distortion complicate the decoding process, as extraneous sounds can interfere with the model's ability to identify Morse code patterns. Traditional machine learning models often struggle with these variations due to overfitting on specific sound characteristics. For instance, models trained exclusively on clean, beeping Morse code signals may perform poorly when exposed to distorted or noisy audio inputs, as the model cannot adapt to new sound sources or noise levels. This gap necessitates a model capable of handling diverse sound inputs while remaining resilient to background noise. The project aims to develop a model that accurately decodes Morse code from various sound sources by using dataset augmentation and noise simulation during training. The dataset includes Morse code recordings from taps, honks, and bells to expose the model to different acoustic contexts. Temporal and amplitude randomness is added to mimic real-world variations in tone and spacing, improving the model's adaptability. Connectionist Temporal Classification (CTC) loss is applied in the LSTM-CNN architecture to handle variable-length sequences without requiring precise alignment. CTC loss helps the model recognize Morse code patterns despite timing or signal amplitude inconsistencies. Experiments show that amplitude-time graphs improve the model's accuracy in distinguishing Morse code features. In contrast, frequency spectrograms are less reliable due to noise overlap.

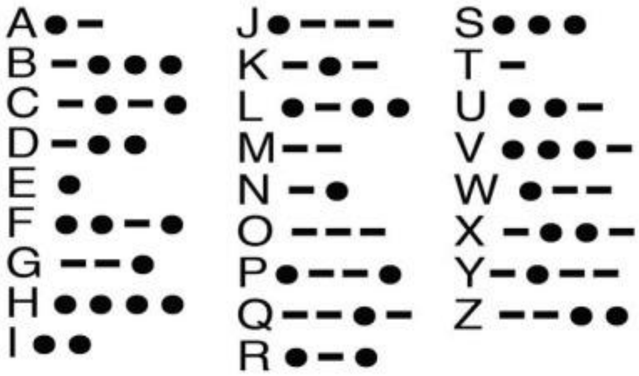

Fig-1: Morse Code Alphabet Chart

2. Literature Survey :

Here we have gathered several journals that have conducted research on our connected work, and we have separately summarized each work as shown below. To date, many authors have proposed various approaches to address these issues, each presenting their own methodologies. In the following chapter, we review existing methodologies and evaluate their effectiveness in addressing these challenges.

The dataset includes 2300 wave files for training and 500 for validation and testing, created by converting English text into Morse code audio signals, each file representing one or two words. The preprocessing involves splitting data into inputs (wave features) and targets (labels). Mel-frequency Cepstral Coefficients (MFCC) are extracted as audio features and structured for compatibility with the Long Short-Term Memory (LSTM) network. The primary model, LSTM, is enhanced with Connectionist Temporal Classification (CTC) for aligning outputs with Morse code labels. Performance is evaluated using metrics like accuracy, precision, recall, and F1-score. The paper focuses on improving Morse code recognition in noisy conditions. Results show the LSTM-CTC model enhances Morse code audio readability effectively [1].

The paper does not specify the dataset used but implies that input data is derived from real-time webcam video capturing eye blinks as Morse code signals. Eye blinks are categorized as short (dots) or long (dashes), with preprocessing analyzing blink durations to convert them into Morse code. A 40-second closed-eye window filters extra signals, though some limitations with extra dots and dashes remain. Four machine learning models are compared, with the Decision Tree Regressor achieving the highest accuracy and Random Forest noted for efficiency in handling large datasets. Evaluation focuses on accuracy, with the Decision Tree Regressor significantly outperforming other models and achieving better results than OpenCV techniques, which had only 51.16% accuracy. Graphs highlight the Decision Tree Regressor's superior prediction alignment with actual data. The study demonstrates that integrating machine learning significantly enhances Morse code translation accuracy [2].

The study collected data from four datasets using a real-world testbed with a 400W shortwave communication platform, multiple transmitters, and a single wideband receiver to test the Deep Morse model under varying conditions like frequencies, distances, seasons, and times. Due to the rarity of Morse signals, data augmentation techniques, such as rotations and adding Gaussian noise, were used to expand training samples. The Deep Morse model, designed for blind detection of Morse signals in wideband spectrum data, includes a multi-signal sensing module to identify signal candidates and a CNN-based module to extract features. Performance was evaluated using accuracy, true positive rate, and precision rate, with ROC and Precision-Recall curves for visualization. The model achieved an accuracy of 0.9718 on the 7M datasets, outperforming HSVM (0.9539) and SVM (0.8658). It demonstrated robustness in wireless environments, leveraging locally connected neural networks to retain spatial information. Experimental results showed that Deep Morse effectively distinguished Morse signals, with a rapid increase in true positive rate and gradual decrease in precision rate compared to other models[3].

The Wi Morse system experiments involved seven volunteers (one female, six male) performing Morse code finger gestures for 54 characters, including letters, numbers, and punctuation. Data was collected in three scenarios: home, office, and meeting room. Preprocessing normalized signals to a range of 0 to 1, unifying patterns for accurate recognition. A novel WiFi signal transformation mechanism addressed location dependency, enabling gesture recognition across varying environments. Performance was evaluated using characters per minute (CPM), words per minute (WPM), and recognition accuracy (92%-96%). Two WPM methods were used: one assuming equal character probability and another using statistical frequency. Results showed high accuracy but low text input speed, with potential for improvement in future designs [4].

The study investigates translating Morse code through eye blinks, using video input to capture blink sequences. Blink patterns are represented as binary values, with short blinks ("0") signifying dots and long blinks ("1") signifying dashes. OpenCV was used in preprocessing to decode eye blinks from video, converting visual data into a machine-learning-compatible format. Four machine learning models were developed for comparative analysis, with the Decision Tree Regressor achieving the highest accuracy in translating Morse code. Evaluation included comparing actual and predicted Morse sequences, with visual graphs highlighting the results. The Decision Tree Regressor demonstrated a strong correlation between predictions and actual values. The system's accuracy improved significantly, from 51.16% without machine learning to much higher with the proposed models [5].

The paper explores datasets for recognizing and decoding Morse code from eye blink patterns, highlighting available datasets for model training. Several preprocessing steps are performed, such as removing punctuation, word tokenization, and forming word sequences, with tokenization playing a crucial role in data structuring. To detect blinks, the eye aspect ratio (EAR) is calculated, registering a blink when it crosses a specific threshold, identifying dots and dashes in Morse code. The study uses an LSTM network built with TensorFlow to capture temporal dependencies in blink patterns. During training, the model achieves a 70% accuracy, showing promise but with potential for improvement with a larger dataset. The findings indicate positive results, demonstrating the model's viability in decoding Morse code. Future plans include further training to boost performance and improving the user interface for better usability [6].

The paper discusses implementing Morse code in underwater acoustic communication, covering system design, modulation techniques, and applications. It highlights that acoustic signals from deep-sea divers and vehicles are the main data source for communication, though no specific dataset is mentioned. The Morse code converter is a vital preprocessing step, transforming text into Morse code for transmission. An acoustic transducer converts sound waves into electrical signals, which are demodulated to retrieve Morse code. The system includes components like the transmitter, modulator, and receiver, which work together to encode and decode the signals. While no quantitative evaluation metrics are provided, the paper emphasizes signal clarity and communication reliability. The results suggest the system is effective for long-distance transmission, especially in emergencies, and supports applications like remote control and scientific data collection [7].

The study uses a dataset of synthetic Morse signals mixed with real-world wireless signals to simulate background noise for Morse detection and recognition tasks. The synthetic signals mimic various skill levels, incorporating variations like speed, frequency drift, and jitter. Real-world signals are combined at different signal-to-noise ratios (SNR) to enhance the dataset. Preprocessing includes generating spectrograms (749 × 512 images) representing 15 seconds of audio within a 4.5 kHz frequency range. The primary model, MorseNet, uses a shared convolutional architecture with detection and recognition branches. The verification unit analyzes the connected text boxes, while the recognition unit uses the convolutional recurrent neural network (CRNN) to identify Morse code segments.MorseNet outperforms traditional methods in both speed and model size, processing 109.5 seconds of signals per second with a GPU, demonstrating its effectiveness for Morse detection and recognition [8].

The study on Morse signal detection and recognition uses simulated spectrograms and real-time audio data for the recognition network. It involves 10,560 simulated spectrograms, divided into training, validation, and test sets, with parameters like frequency, code rate, and signal-to-noise ratio (SNR) randomly assigned. Additionally, 447 spectrogram segments from 400 real-time audio samples are used for testing. Preprocessing includes energy sorting with adaptive thresholds, digital filtering, short-time Fourier Transform for time resolution, and pseudo-color mapping for clearer identification. A Random Forest classifier is used for detection, and an improved CRNN with a convolutional attention mechanism is used for recognition. Model performance is evaluated using character accuracy, sample accuracy, and decoding time. The ED-FE-CCBC algorithm achieves real-time performance and 90% accuracy under optimal conditions, though recognition accuracy declines at lower SNR levels, showing superior performance in complex environments [9].

The paper "Morse Code to Text Converter for Marine Communication" describes a system for transmitting and decoding Morse code in marine environments. It receives Morse code signals based on time intervals: dots (5 ms to 0.5 s), dashes (0.5 to 1 s), inter-element spacing (0.5 s), and short gaps between letters (over 1 s). The system uses an Arduino Atmega328 microcontroller to convert signals into text, displayed on an LCD screen and saved to a PC via HDMI to USB. No specific evaluation metrics are provided; effectiveness is assessed by reducing human error and enhancing data transmission speed.Results show accurate transmission and decoding, ensuring message retention. The system improves communication efficiency in marine settings, focusing on practical implementation rather than theoretical model evaluation [10].

The paper evaluates Morse code decoding efficiency using varying lengths of audio signals, such as one-minute and seven-minute recordings, without a specific training dataset. Traditional methods require preprocessing, but the proposed system minimizes this with a PyTorch-based model, incorporating a Convolutional Recurrent Neural Network (CRNN) and a Bidirectional Gated Recurrent Unit (BiGRU).The model is to improve on-the-fly decoding efficiency, it was converted to RKNN format and sent to Rockchip's RK3568 SoC. Evaluation focuses on decoding time, with significant improvements noted. For a one-minute signal, the pt-format model takes 12.63 seconds, while the RKNN model reduces it to 1.67 seconds, a 7.5-fold improvement. For a seven-minute signal, the pt-format model averages 81.5 seconds, and the RKNN model reduces this to 6.94 seconds, an 11-fold improvement. The paper highlights efficiency gains through hardware acceleration on mobile platforms [11].

The paper explores covert data transfer using internal speakers to transmit information through sound waves, encoded in Morse or binary code. The computer's internal speaker generates inaudible frequency signals, recorded by a smartphone for analysis. This method is ideal for secure environments with restricted communication channels. The effectiveness is evaluated by measuring the bit error rate, influenced by distance and beep duration. The maximum data transfer rate achieved is 20 bits per second at 1.5 meters. Results demonstrate successful data transfer with internal speakers, highlighting potential security risks. The study emphasizes the need for protective measures in organizations [12].

The paper introduces a wearable triboelectric nanogenerator (TENG) made from recycled carbon-coated paper wipes (C@PWs) to address electronic waste and promote sustainability. The carbon nanoparticles are applied to enhance triboelectric properties, with optimal coating thickness for improved electrical output. Characterization techniques like SEM, XRD, and Raman spectroscopy are used to analyze material properties. The study uses the DDEF theoretical model to simulate charge density and output characteristics. Performance is measured through short-circuit current, open-circuit voltage, and other metrics. Results show that carbon loading improves output, with the maximum power rising from 0.1 mW to 0.58 mW under mechanical stress. Durability tests indicate stable performance, highlighting the device's potential for low-power electronics and sustainable energy harvesting [13].

The paper explores underwater communication using Morse code transmitted via a piezo transducer and received by a wireless hydrophone. The study focuses on mitigating noise and interference using digital filtering in MATLAB to restore the transmitted signals. The primary model relies on Morse code's on-off tones (dots and dashes) to encode information, with MATLAB analysis applied to the received data.Although specific evaluation metrics are not detailed, effective transmission is indicated by clear signal retrieval after filtering. Results show that, despite challenges in accuracy, Morse code remains effective for underwater communication. Digital filters are crucial for signal restoration. The paper highlights the importance of reliable transmission techniques for unmanned and autonomous underwater vehicles (UUVs and AUVs), offering insights into preprocessing and signal processing for effective data transmission [14].

The study monitors VHF radio signals from six stations across Japan, operational since May 2019, capturing aeronautical navigation (NAV) signals and broadcasting channels. Signal strength is recorded every 2 seconds, and an algorithm is being developed to identify mid-points of anomalous propagation (EsAP) signals through point concentration analysis. Transmission and reception station mid-points are mapped to visualize EsAP spatial distribution. Although no machine learning models are used, GPS-TEC and InSAR data enhance the understanding of sporadic E layer (Es) distribution. The system detects EsAP signatures, linking them to electron density increases in the Es layer. Real-time monitoring provides valuable insights for aeronautical navigation, with EsAP signatures persisting for hours, affecting signal propagation. This study highlights the significance of real-time monitoring for VHF NAV signal anomalies caused by sporadic E layers [15].

3. Methodology :

This architecture represent a typical end-to-end deep learning model for morse code decoding using sound waves. Here's a breakdown of each component and step involved in the process:

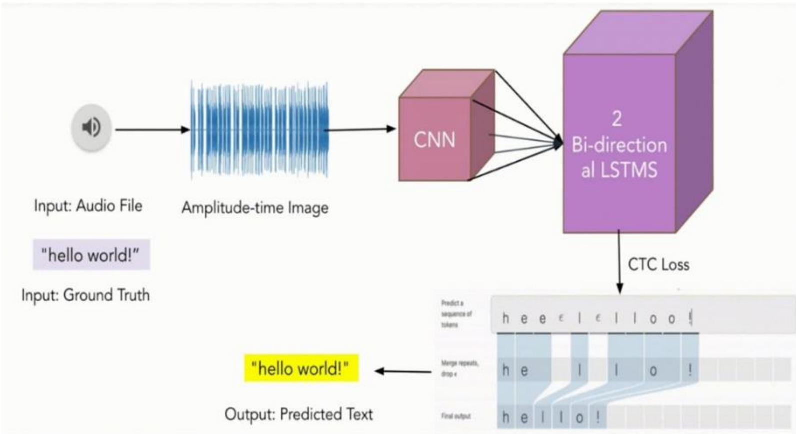

Fig-2: Architecture of Morse Code Decoding Using Soundwaves

Input-Audio File:

The process starts with an audio file, which contains a spoken phrase, like "hello world!".

The audio data is represented as a waveform that captures amplitude (volume) over time, where each point shows the intensity of sound at a specific moment.

Amplitude-Time Image:

The raw audio file is converted into an amplitude-time image (or spectrogram), which is essentially a visual representation of the sound wave.

This conversion helps the neural network model to see the audio in a way that makes it easier to analyze patterns and features within the audio signal.

Convolutional Neural Network (CNN):

The amplitude-time image is passed through a Convolutional Neural Network (CNN). CNNs are effective at identifying features in visual data.

Here, the CNN extracts relevant features from the audio's visual representation. These features might represent specific patterns in the audio that relate to certain sounds or phonemes (basic units of speech).

The CNN essentially acts as a feature extractor, transforming the raw audio representation into higher-level features that can be understood by the following layers.

Bi-Directional Long Short-Term Memory (Bi-directional LSTMs):

The extracted features are then passed to two Bi-Directional LSTM (Long Short-Term Memory) layers.

LSTMs are a type of Recurrent Neural Network (RNN) designed to remember information from previous inputs, making them well-suited for sequential data like speech.

In this case, the bi-directional LSTM can process data both forward and backward, allowing it to understand context from the entire audio sequence, which improves its ability to accurately recognize words.

Connectionist Temporal Classification (CTC) Loss:

The Bi-directional LSTM outputs a sequence of predicted characters or phonemes.

Connectionist Temporal Classification (CTC) Loss is then applied. CTC is a technique used to handle the issue of varying lengths in sequences.

The CTC layer attempts to align the output sequence of characters with the ground truth text (like "hello world!") It learns the best way to match input with output, allowing you to skip or repeat characters. It effectively helps the model output the correct sequence of characters even when the input and output lengths are mismatched.

Decoding Process:

The output of the CTC layer is a sequence of tokens (characters), with some characters repeated and others possibly including blank tokens (representing pauses or silence in speech).

The decoding process merges repeated characters and removes blank tokens. For example, if the output sequence is "hheeellloo wworrldd!!", the decoding process refines this into "hello world!".

Output-Predicted Text:

After decoding, the final output is a cleaned-up text that represents the original spoken phrase, like "hello world!".

This predicted text is compared to the ground truth text, and the CTC Loss is used to improve the model during training, minimizing the difference between the predicted text and the actual spoken words over time.

4. Conclusion :

In Conclusion, this project showcases the powerful application of machine learning in audio signal processing. By utilizing deep learning models such as Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks, the system successfully decodes Morse code from sound waves, enabling effective communication when visual inputs are not possible. The integration of feature extraction techniques like Mel Frequency Cepstral Coefficients (MFCC) further enhances the accuracy and reliability of the model, making it suitable for real-world scenarios. The project emphasizes the importance of preprocessing techniques to improve input data quality, which in turn boosts model performance. By training the model on diverse datasets with varying noise levels and sound qualities, the system demonstrates strong generalization capabilities, minimizing errors in decoding.

REFERENCES :

[1]. Win, A., & San, K. K. (2023). Morse code audio recognition using LSTM-CTC model. 2023 IEEE Conference on Computer Applications (ICCA), 393-398.

[2]. Deepak, G. S. N., Rohit, B., Akhil, C., Bharath, D. S. S. C., & Prakash, K. B. (2021, May). An approach for Morse code translation from eye blinks using tree-based machine learning algorithms and OpenCV. Journal of Physics: Conference Series, 1921(1), 012070.

[3]. Yuan, Y., Sun, Z., Wei, Z., & Jia, K. (2019). DeepMorse: A deep convolutional learning method for blind Morse signal detection in wideband wireless spectrum. IEEE Access, 7, 80577-80587.

[4]. Niu, K., Zhang, F., Jiang, Y., Xiong, J., Lv, Q., Zeng, Y., & Zhang, D. (2019). WiMorse: A contactless Morse code text input system using ambient WiFi signals. IEEE Internet of Things Journal, 6(6), 9993-10008.

[5]. Sumanth, G., Deepak, N., Rohit, B., Akhil, C., Surya, S., Bharath, C., & Prakash, K. B. (n.d.). An approach for Morse code translation from eye blinks using tree based machine learning algorithms and OpenCV.

[6]. Zaveri, D., Joshi, Y., Shah, Y., Shah, B., Parekh, N., & Mangrulkar, R. (2023). Transformation of text to decode Morse code and its detection using eye blinks.

[7]. Hariharan, A., Krishnan, S. S., Karthikeyan, V. M., Dhaneshvar, E., & Raffik, R. (2023). Implementation of Morse code in underwater acoustic communication. 2023 2nd International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA). IEEE.

[8]. Li, W., Wang, K., & You, L. (2020). MorseNet: A unified neural network for Morse detection and recognition in spectrogram. IEEE Access, 8, 161005–161017.

[9]. Gao, Z., Zhang, Y., Su, S., & Liu, J. (2022). Detection and recognition based on machine learning and deep learning for Morse signal in wide-band wireless spectrum. Research Square.

[10]. Suryavanshi, P., Dhavale, V., Shinde, N., Jadhav, K. V., & Ainapure, N. (2019). Morse code to text converter for marine communication. International Journal for Research in Applied Science & Engineering Technology, 7(IV), 1755–1760.

[11]. You, X., & Weng, W. (2024). Real-time Morse code decoding: Exploring domestic SoC applications. 2024 4th IEEE International Conference on Software Engineering and Artificial Intelligence (SEAI), 296–300. IEEE.

[12]. Choi, H., Jung, J. H., & Yoon, J. W. (2023). CASPER: Covert channel using internal speakers. Sensors, 23(6), 2970.

[13]. Dudem, B., Dharmasena, R. D. I. G., Riaz, R., Vivekananthan, V., Wijayantha, K. G. U., Lugli, P., Petti, L., & Silva, S. R. P. (2022). Wearable triboelectric nanogenerator from waste materials for autonomous information transmission via Morse code. ACS Applied Materials & Interfaces, 14(4), 5328-5337.

[14]. Dur-e-Jabeen, D., Shaukat, H., Siddiqui, M., Ahmed, S., Zaheen, M. Y., Din Ahmad, A. U., Mirza, H. A., & Sami, M. (2023). Efficient underwater wireless data transmission technique and signal processing. Engineering Proceedings, 46(1), 43.

[15]. Hosokawa, K., Sakai, J., Tomizawa, I., Saito, S., Tsugawa, T., Nishioka, M., & Ishii, M. (2020). A monitoring network for anomalous propagation of aeronautical VHF radio waves due to sporadic E in Japan. Earth, Planets and Space, 72, 88.

[16]. Yang, C.-H., Chuang, L.-Y., Yang, C.-H., Luo, C.-H., & Yang, C.-H. (2003). Morse code application for wireless environmental control systems for severely disabled individuals.

[17]. Ji, Z., Fu, Y., Li, J., Zhao, Z., & Mai, W. (2021). Photoacoustic communication from the air to underwater based on low-cost passive relays. IEEE Communications Magazine, 59(1), 140–143.

[18]. Geetha, M. C. S., & Sarubala, M. (2024). Dots and dashes: Understanding the language of Morse code using convolutional neural network.
[19]. Regmi, P., Dahal, A., & Joshi, B. (2019). Nepali speech recognition using RNN-CTC model. International Journal of Computer Applications, 178(31), 1–6.
[20]. Kumar, K. K., Srikar, V. S., Swapnika, Y., Sravani, V. S., & Aditya, N. (2020). A novel approach for Morse code detection from eye blinks and decoding using OpenCV. International Journal for Research in Applied Science & Engineering Technology (IJRASET), 8(5), 2096-2101.