Comparison of Modelling ASR System with Different Features Extraction Methods Using Sequential Model

Aishwarya Suresh(✉), Anushka Jain, Kriti Mathur, and Pooja Gambhir

Indira Gandhi Delhi Technical University for Women, Delhi, India
aishwarya100btit19@igdtuw.ac.in

Abstract. Speech recognition refers to a device's ability to respond to spoken instructions. Speech recognition facilitates hands-free use of various gadgets and appliances (a godsend for many incapacitated persons), as well as supplying input for automatic translation and ready-to-print dictation. Many industries, including healthcare, military telecommunications, and personal computing, use speech recognition programmes. In our paper, we are including the comparison between the different feature extraction methods (BFCC, GFCC, MFCC, MFCC Delta, MFCC Double Delta, LFCC and NGCC) using neural networks.

Keywords: Feature extraction method · Speech recognition · Neural network · FFT

1 Introduction

Speech recognition is a capability of computer software that allows it to turn human speech into text. In simple words, it means that when humans are speaking, a machine understands it. Speech recognition employs a wide range of computer science, linguistics, and computer engineering research. Speech recognition functions are included into many current gadgets and text-focused programmes to make using them easier or hands-free. Speech recognition, which is commonly mistaken with voice recognition, is concerned with converting speech from a verbal to a text format in a spoken language, whereas the biometric technique of voice recognition focuses only on recognising the voice of a certain individual. Traditional methods of interfacing with a computer, such as textual input through a keyboard, are being replaced by speech recognition. A good system can either eliminate or reduce the need for traditional keyboard input. By analysing the audio, breaking it down into parts, digitising it into a computer-readable format, and matching it to the most appropriate text representation using an algorithm, a computer programme translates the sound acquired by a microphone into a textual language that computers and people can comprehend (Fig. 1).

The most common method for building a speech recognition system is to create a generative model of language. Using language models, we create a certain sequence of words. Then, for each word, there's a pronunciation model that describes how to

A. Dev et al. (Eds.): AIST 2021, CCIS 1546, pp. 47–61, 2022.
https://doi.org/10.1007/978-3-030-95711-7_5

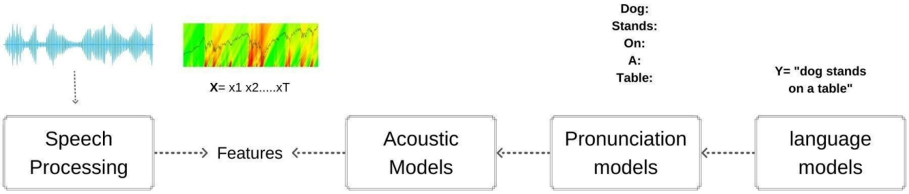


Fig. 1. A simple speech recognition model

pronounce the term. It's usually defined as a series of phonemes — basic sound units — but for the sake of our language, we'll just call it a series of tokens — which represent a collection of objects. The pronunciation models are then fed into an acoustic model, which determines the sound of a token. The data is currently described using these acoustic models. The data is x, which is a sequence of audio feature frames spanning from ×1 to xT in this case. Professionals in signal processing are usually the ones that decide on these characteristics (such as the frequency components of the audio waveforms that are captured).

Speech recognition software should adjust to the widely changeable and context-specific nature of human speech. Software algorithms that translate and organise audio into text are trained using a variety of speech patterns, speaking styles, languages, dialects, accents, and phrasings. The software distinguishes speech audio from the often-present background clatter. Speech recognition systems use two types of models to achieve these requirements: acoustic models and linguistic models. We used the acoustic model in our paper. Speech recognition software employs natural language processing (NLP) and deep learning neural networks. We employed convolutional neural networks, which are a sort of deep learning neural network. The flexibility and forecasting capacity of deep neural networks, which have lately become more accessible, are an advantage of deep learning for voice recognition. Another major issue in voice recognition is latency; in order to translate in real time, the model must properly predict words without knowing the entire sentence. Because of the increased context, some deep learning models profit greatly from using the entire sentence. To reduce latency, integrate restricted context in the model structure by allowing the neural network to access a little amount of data after a given word. Despite its difficulty, speech recognition is always present in a variety of industries. It allows a large number of individuals to readily access whatever material they wish. Speech recognition is a burgeoning field with numerous applications. Speech recognition research will almost certainly continue, and important practical applications will emerge. Despite the fact that speech recognition is a flourishing field, accuracy is a big challenge in this sector. The most accurate machine conceivable is continually being developed through research and development.

Speech recognition is achieved using a convolutional neural network because human speech signals are significantly variable due to various speaker features, speaking styles, and other sounds. Convolutional neural networks are a type of deep neural network that conducts little preprocessing, or learning the filter before doing the classification. When fed a huge number of signals as input, CNNs, which can have one or more layers, can do a lot of things. The Convolutional Neural Network (CNN) uses a subset of the data rather than the entire signal because it is difficult for a computer system to evaluate the entire signal. CNN is a type of neural network in which the input variables are connected spatially. The Xuejiao Li and Zixuan Zhou speech recognition system, which uses Google's Tensorflow's Speech command dataset, shows how CNN improves speech recognition. The models employed were the vanilla single-layer Softmax model, Deep Neural Network, and convolutional neural network, with the convolutional neural network surpassing the other two. CNN outperformed DNN and vanilla in terms of precision value, with an 18.6% relative improvement over DNN and a 72.3% relative improvement over Vanilla. A basic 2-layer ConvLayer CNN network beats Vanilla and DNN, with 31.43% and 66.67% comparitive improvements in test accuracy and 82% and 94.6% in loss, respectively, over DNN and Vanilla.

HMM and GMM are two must-learn voice recognition technologies that existed before the Deep Learning (DL) era. There are now hybrid systems that mix HMM with Deep Learning, as well as systems that do not use HMM. We now have more design options. HMM, on the other hand, is still important for many generative models. A Markov chain contains all of a system's possible states as well as the probability of changing states. The next state of a first-order Markov chain is solely determined by the current state. We call it a Markov chain for simplicity's sake. For sequential tasks like speech recognition, recurrent neural networks (RNNs), particularly long short-term memory (LSTM) RNNs, are effective networks. Because of their excellent learning potential, deeper LSTM models perform well on large vocabulary continuous speech recognition. A deeper network, on the other hand, is more difficult to train. For deeper LSTM models, we present a training architecture that includes layer-wise training and exponential moving average approaches.

MFCC, MFCC Delta, MFCC Double Delta, GFCC, BFCC, and LFCC are the feature extraction methods employed in our paper. For a certain dataset, their performance is compared. In the field of speech processing, the Gammatone Frequency Cepstral Coefficients (GFCCs) are a relatively novel characteristic. The GFCCs work on the basis of an auditory peripheral model that is similar to the human cochlear filtering system. They are a set of Gammatone Filter banks for creating auditory features. A Cochleagram, which is a frequency-time representation of the signal, can be created using the Gammatone filterbank output. The Gammatone filters are designed to emulate the processes of the human auditory system. Windowing the signal, using the DCT, and selecting the log of the magnitude are all part of the GFCC feature extraction approach. LFCC is as robust as MFCC in babbling noise, but not in white noise. LFCC consistently outperforms MFCC in female trials. LFCC has the same qualities as MFCC except for the frequency scale. Linear filter banks provide excellent resolution in higher frequency bands. Linear Frequency Cepstral Coefficient (LFCC) is a feature extraction method. In LFCC, the method for extracting features is the same as in MFCC. LFCC extraction differs from MFCC extraction in that it employs a linear filter bank rather than a mel filter bank. It functions in the same way that the human auditory system does. The linear filter bank has improved resolution in the higher frequency band. To compute LFCC features, first transform a windowed signal with the Fast Fourier Transform (FFT), which converts each frame of N samples from time to frequency dominion. After the FFT block, the power coefficients are filtered using linear frequency filter banks. Signal disintegration with a filter bank is the foundation of the MFCC algorithm. Windowing the signal is part of the MFCC feature extraction approach, also known as Mel Frequency Cepstral Coefficient. On each frame, a window is bored to taper the signal to the frame limits. Hanning or Hamming windows are commonly utilised. DFT is then used to transform the magnitude spectrum of each windowed frame. The Mel spectrum is calculated using the log of the magnitude and a Mel scale to bend the frequencies after passing a Fourier transformed signal through a Mel-filter bank of band-pass filters. A Mel is a measuring unit obtained from the human ear's perceived frequency and modified using the inverse Discrete Cosine Transform. On the Mel frequency scale, the MFCC generates a discrete cosine transform (DCT) of a short-term energy's real logarithm. MFCC is used to identify airline reservations, phone numbers, and speech recognition systems for security reasons. Understanding the dynamics of the power spectrum, or the trajectories of MFCCs over time, is critical for improving speech recognition, which is why delta (differential) and delta-delta (acceleration) coefficients are used. FFTs (fast Fourier transforms) are simple methods for quickly performing the discrete Fourier transform (DFT) on the basis of a finite abelian group. They are among the most important algorithms in engineering and applied mathematics, as well as computer science, with Signal processing and one- and multidimensional systems theory applications.

2 Related Work

In [1], Zhang Wanli and Li Guoxin used Mel frequency cepstral coefficients (MFCC) for speaker recognition. They worked on a study of MFCC-based feature extraction for speaker recognition. MFCCs outperform hidden Markov model-based MFCCs. [2] Dev Amita Agrawal, S.S., "A Novel MFCCs Normalization Technique for Robust Hindi Speech Recognition" 17th International Congress on Acoustics (ICA) Rome September 2–7, 2001. Nagajyothi and P. Siddaiah worked on Speech Recognition Using Convolutional Neural Networks [3]. They investigated the performance of a CNN-based ASR that uses raw speech signals as input to large vocabulary challenges in this research. Their research on wideband signals revealed that the CNN-based system outperforms the traditional Neural Network Techniques-based system. They employed the primary activation function in the first convolutional layer to provide fragments in a 2-D matrix. In [4] Jui-Ting Huang et al. worked upon an analysis of sequential Conv2D layer for speech recognition. They presented a detailed examination of CNNs in this paper. They showed that by analysing the localised filters acquired in the convolutional layer, edge detectors in multiple orientations may be automatically taught. CNNs outperformed FCNNs in four areas: channel-mismatched training-test conditions, noise reliability, remote speech recognition, and small-footprint models, according to the researchers. In [5] Ossama Abdel-Hamid et al. worked on FFN for Speech Recognition. They demonstrated how to apply CNNs in a novel method for speech recognition in which the CNN structure directly accommodates specific sorts of speech variability in this study. Using this strategy, they demonstrated a performance improvement over normal DNNs with equal amounts of weight parameters, in comparison to the more ambiguous findings of convolving along the axis of time, as CNNs have sought to use speech in the past. (about 6–10% relative error reduction). They improved performance on two ASR tasks:

TIMIT phone perception and an enormous-terminology voice search test, using a variety of Conv2D layer parameters and design choices. They discovered that integrating energy information improves feedforward neural network's recognition accuracy significantly. In [6], Phani Bhusan S et al. worked on stuttered speech recognition using Convolutional Neural Networks. They extracted features using the MFCC approach. They examined the scalability of a SSR based on CNN that accepts the raw speech signal as input in the proposed method. They were able to achieve 92% accuracy with only 7% validation loss in the suggested technique using CNN. Taabish Gulzar et al. surveys a correlative analysis of LPCC, MFCC and BFCC as feature extraction techniques and classifier as ANN in [7]. For database purposes, they used Hindi isolated, paired, and hybrid words. The results of their study reveal that MFCC outperforms the traditional LPCC and BFCC approaches. Using the Speech command dataset provided by Google's Tensorflow, the Xuejiao Li and Zixuan Zhou speech recognition system shows an improvement in speech recognition using CNN in [8]. The models used were the vanilla single-layer Softmax model, DNN, and feed forward neural network, with the Conv-2D layer surpassing the other two by obtaining an accuracy of 95.1% for six labels. The work of Mariusz Kubanek et al. proposes a unique technique to speech recognition established on the exact time-domain coding and frequency properties in [9]. Their plan was to combine three convolution layers: classic time convolution, feedforward neural network convolution, and spectrum convolution. Their research found that using the correct sound coding and pictures, effective speech recognition for isolated words may be achieved. To reduce noise impacts and raise robustness against various forms of environmental disturbances, Mohamed Tamazin et al. upgraded the (PNCC) system by combining gammatone channel clarity with channel bias minimization in [10]. At low SNR, the proposed method considerably improves recognition accuracy (SNR). Furthermore, in terms of recognition rate, the suggested method beats the GFCC [11] and PNCC methods. [12] Dr. Amita Dev, Sweeta Bansal, "Emotional Hindi Speech: Feature Extraction and Classification" published in IEEE Explorer, Computing for Sustainable Global Development (INDIACom), 2015 2nd International Conference (989-9-3805–4415-1), page(s): 1865–1868, 11–13 March 2015. In many languages, the accuracy of the identification evaluation for speaker recognition remains a key challenge. So, Ankur Maurya et al. in [13], used Conv-2D layer–vector quantization (MFCC-VQ) and Conv-2D layer–Gaussian mixture model (MFCC-GMM) for text dependent and text independent phrases to develop speaker detection for Hindi voice samples. In terms of text dependent recognition accuracy, MFCC-GMM surpassed MFCC-VQ by a significant margin. In [14], Pooja Gambhir and Amita dev reviewed different frontend feature extraction methods and DNN feature vectors for identification of context independent speaker voice.

3 Experiment Conducted

In this research, a sample from Tensorflow's Speech Commands Dataset [15] was used to compare the suggested method's performance to that of state-of-the-art approaches. Thousands of people contributed 65,000 one-second long utterances comprising 30 small words. Twenty of the words are core terms, while the remaining ten are auxiliary words that could be used as tests for algorithms to determine whether or not a speech contains triggers. A variety of background noise audio recordings is included with the 30 words. It was bifurcated into three subsets. The training set consisted of 51,094 audio clips, validation set consisted of 6,798 audio clips and testing set consisted of 6,835 audio clips. All audio files have a 16 k sample rate which means they capture up to 8 k Hz sound frequency. The feature extraction methods used- MFCC, GFCC, BFCC, LFCC, NGCC.

The MFCC feature extraction technique or Mel Frequency Cepstral Coefficient includes windowing the signal. A window is bored on each frame to taper the signal to the frame boundaries. Typically, Hanning or Hamming windows are used. The magnitude spectrum of each windowed frame is then transformed using DFT. On passing Fourier transformed signal via a Mel-filter bank of band-pass filters, the Mel spectrum is calculated using the log of the magnitude and a Mel scale to bend the frequencies. A Mel is a unit of measurement derived from the perceived frequency of the human ear, then transformed using the inverse Discrete Cosine Transform. Finally MFCCs are calculated as (Fig. 2):-

$$c(n) = \sum\nolimits_{M=0}^{M-1} \log_{10}(s(m)) \cos\left(\frac{\prod n(m-0.5)}{M}\right) \tag{1}$$

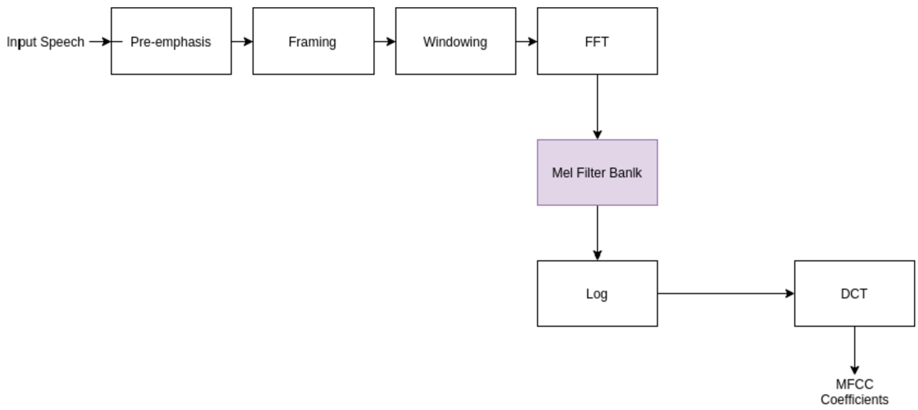


Fig. 2. A block representation of MFCC extraction

The cepstral coefficients are referred to as static features because they only carry information from a single frame. The 1st and 2nd derivatives of cepstral coefficients provide further information about the temporal dynamics of the signal. The first-order derivative is delta coefficients, while the second-order derivative is delta–delta coefficients. Delta coefficients represent the speech pace, while delta–delta coefficients represent the speech acceleration.

$$c(n) = \sum\nolimits_{M=0}^{M-1} \log_{10}(s(m)) \cos\left(\frac{\prod n(m-0.5)}{M}\right) \tag{2}$$

LFCC feature extraction method is Linear Frequency Cepstral Coefficient. The approach for extracting features in LFCC is the same as in MFCC. The difference between MFCC and LFCC extraction is that the latter uses a linear filter bank rather than a Mel filter bank. It works in a similar way to the human auditory system. In the higher frequency band, the linear filter bank has more resolution. To compute LFCC features, first apply the Fast Fourier Transform (FFT) to a windowed signal, which turns each frame of N samples from time to frequency dominion. The power coefficients are filtered by linear frequency filter banks after the FFT block. Finally, utilizing the Discrete Cosine Transform, the log Mel spectrum is altered into time (DCT). The feature vector of LFCC uses 13 coefficients (Fig. 3).

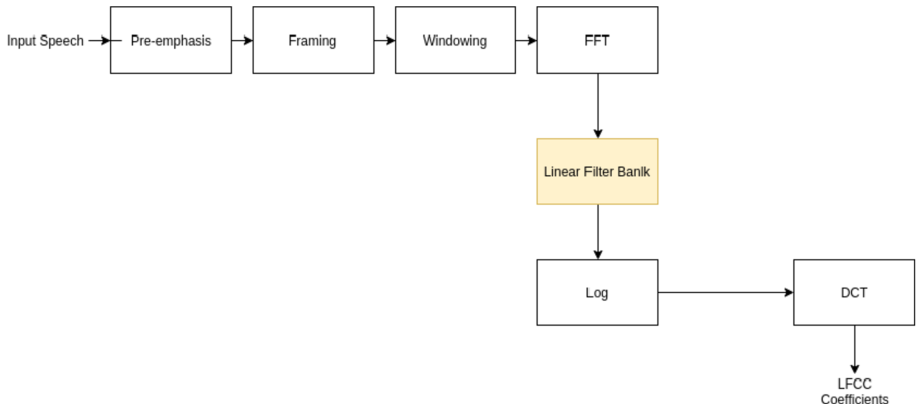


Fig. 3. A block representation of LFCC extraction

The Gammatone Frequency Cepstral Coefficients (GFCC) are a series of Gammatone Filter banks that are used to create auditory features. The Gammatone filterbank output can be used to create a Cochleagram, which is a frequency-time representation of the signal. The Gammatone filters are intended to mimic the human auditory system's processes. The GFCC feature extraction methodology primarily comprises windowing the signal, employing the DCT, picking the log of the magnitude. The below equation represents the GFCC extraction (Fig. 4)

$$\mathrm{g(n;\ u)} = \left(\frac{2}{M}\right)^{0.5} \sum_{M=0}^{M-1} \left\{\frac{1}{3} log\left(\underline{y}(n;i)\right) cos\left[\frac{\Pi u}{2M}(2i-1)\right]\right\} \tag{3}$$

The BFCC [16] extraction method is the Bark Filter Cepstral Coefficient method. To obtain the cepstral coefficients, the PLP processing of the spectra and the cosine transform are merged in the BFCC process. The Bark filter bank was employed instead of the Mel filter bank, and the MFCC-like features were given the same loudness pre-emphasis with an intensity to loudness power law. The implementation of BFCC is very similar to that of MFCC (Fig. 5).

$$f_{bark} = 6ln\left[\frac{f}{600} + \left[\left(\frac{f}{600}\right)^2 + 1\right]^{0.5}\right] \tag{4}$$

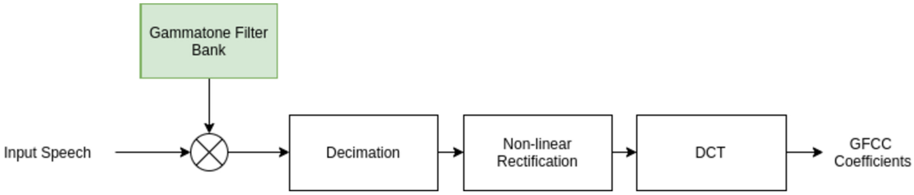


Fig. 4. A block representation of GFCC extraction

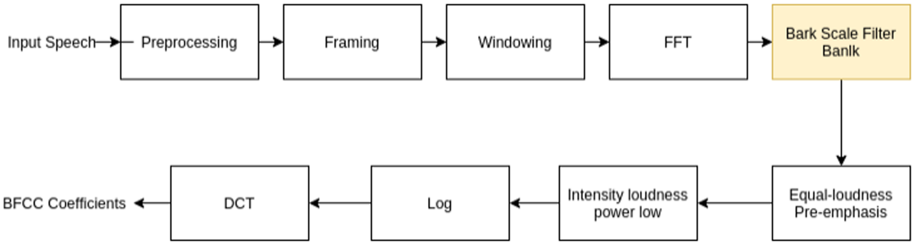


Fig. 5. A block representation of BFCC extraction

NGCC feature extraction method is Normalized Gammachip Coefficient method. To simulate the mechanism, NGCC uses a second order low-pass filter and a normalised gammachirp filterbank. This method is similar to the MFCC computational process. It integrates the features of the peripheral auditory system and uses a Normalized Gammachirp filter bank to improve robustness in noisy speech situations. The feature vector in NGCC used 13 coefficients as well (Fig. 6).

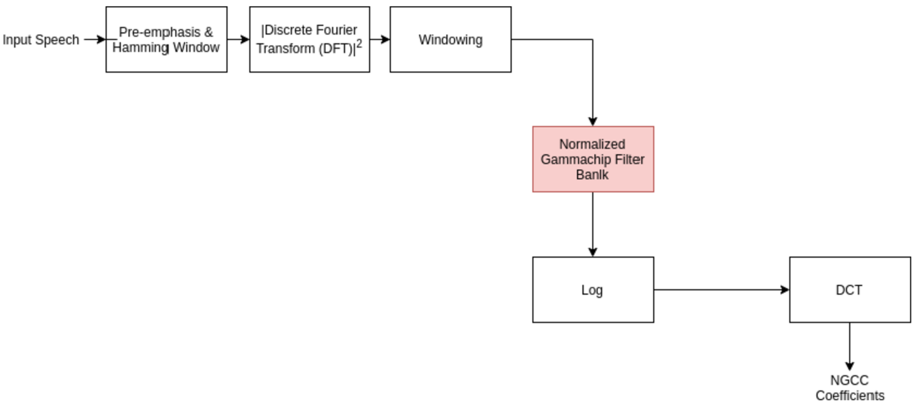


Fig. 6. A block representation of NGCC extraction method

To separate each syllable, 25.6 ms overlapping frames with 10 ms variations amongst frames were employed in the feature extraction method's design. A Hamming window was then applied to each frame. The FFT was then applied with a 256-bit resolution. For each approach, 13 features (cepstral coefficients) were collected in the final stage.

After applying CMN, Δ and ΔΔ features were calculated' the total number of extracted features was 39 features. Despite the fact that higher order coefficients reflect increased spectral information, 12 to 20 cepstral coefficients are often appropriate for speech analysis, depending on the sampling rate and estimate method. The models become more complex when a high number of cepstral coefficients are chosen. For example, in order to effectively estimate the parameters of a Gaussian mixture model (GMM) to represent a speech signal with a large number of cepstral coefficients, we normally require more data.

A Sequential 2D Convolution Neural Network Model (2D CNN Model) was utilised to construct acoustic models for each type of feature extraction approach used. It comprises two completely linked layers and three 2D convolution layers, all with kernels of size 3 3. The number of channels in the first convolutional layer is assumed to be 30. The network is trained using a Soft max Loss function. Before being put to the test with both noise-free and loud utterances, all feature extraction algorithms were skilled on noise-free utterances (Fig. 7).

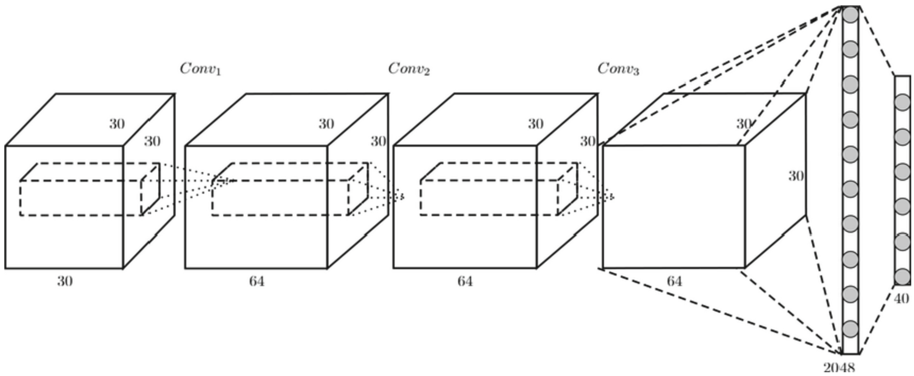
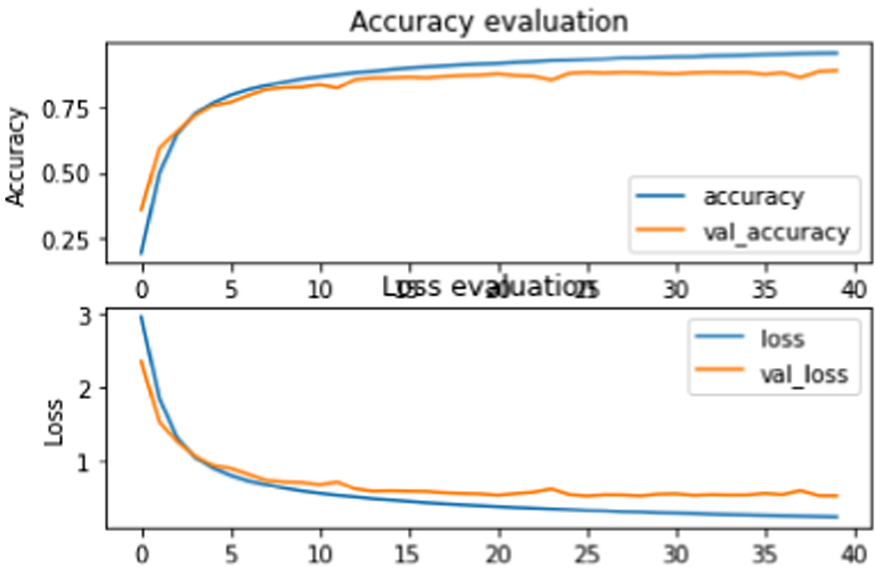


Fig. 7. Network architecture of CNN model

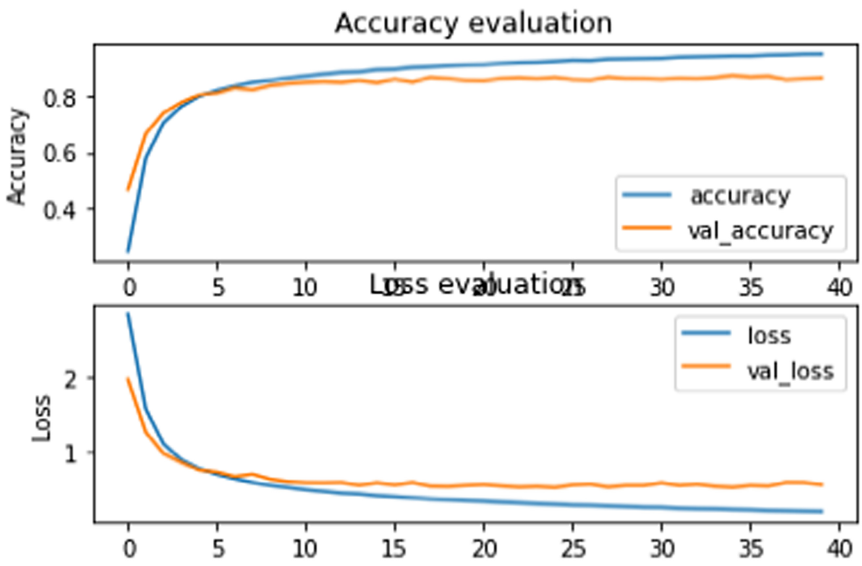
4 Results and Outcomes

The MFCC feature extraction method yielded the maximum accuracy; 88.20% for the sequential model, and the BFCC feature extraction method yielded the lowest accuracy; 68.65%. The best results are obtained with MFCC feature extraction, as shown in the following observations. This suggests that when training CNN models, MFCC is more efficient. With BFCC feature extraction, this is not the case. As a result, BFCC is shown to be the least accurate of all, making it less effective for training the sequential model (Fig. 8 and Table 1).

We can observe the rise and fall about this test validation performed on these features using CNN.

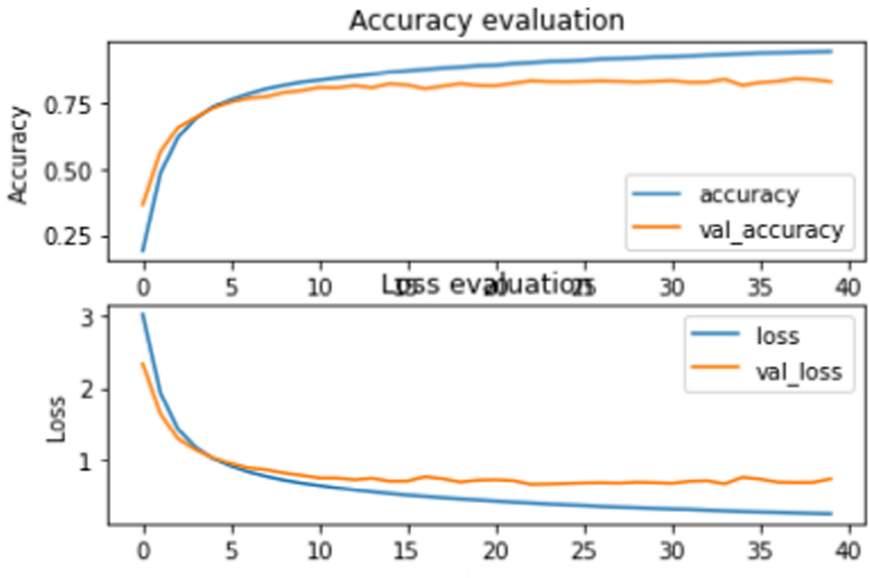


(a) MFCC

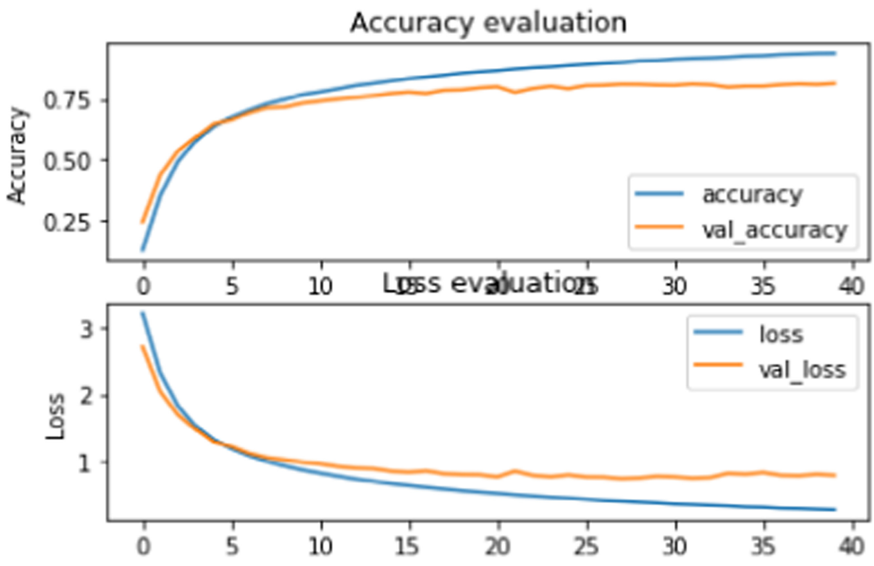


(b) MFCC-Delta

Fig. 8. Percentage accuracy and loss of each feature extraction method namely (a) MFCC, (b) MFCC-Delta, (c) MFCC Double Delta, (d) GFCC, (e) BFCC, (f) LFCC, (g) NGCC

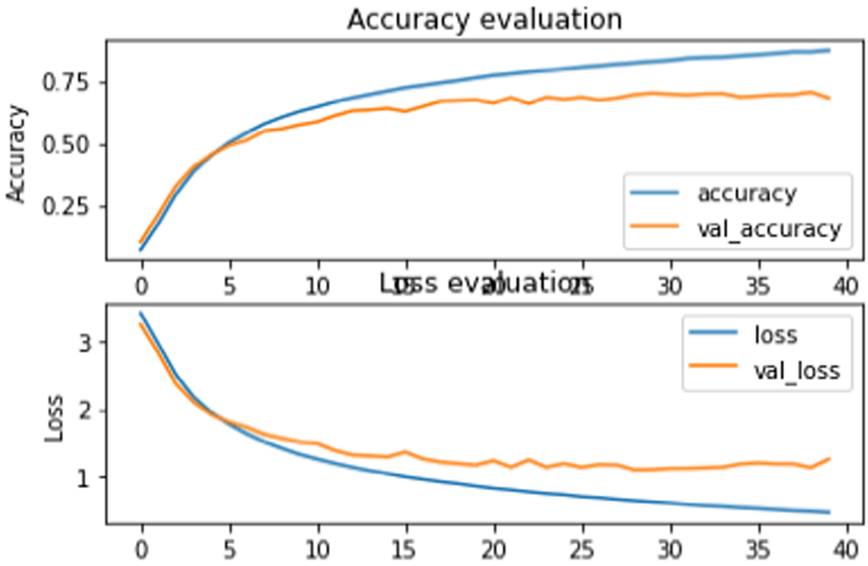


(c) MFCC-Double Delta

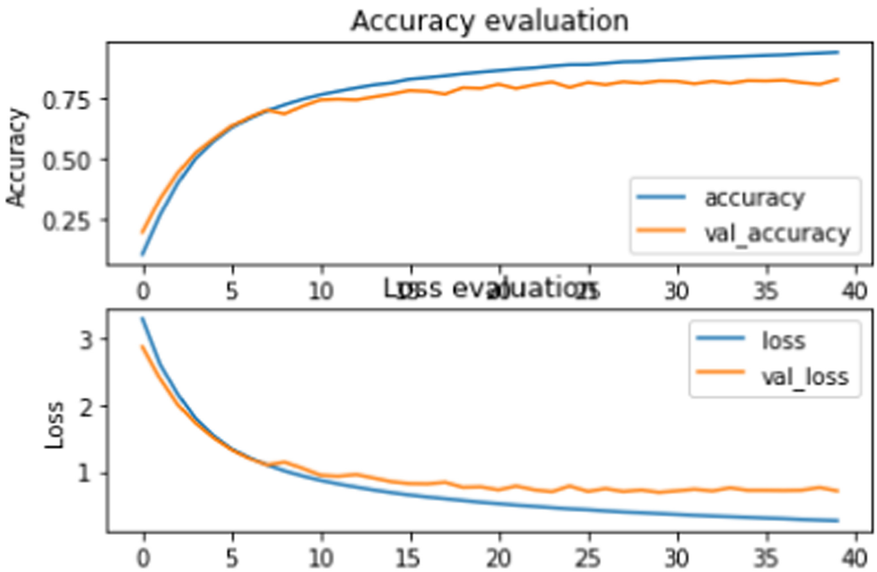


(d) GFCC

Fig. 8. continued

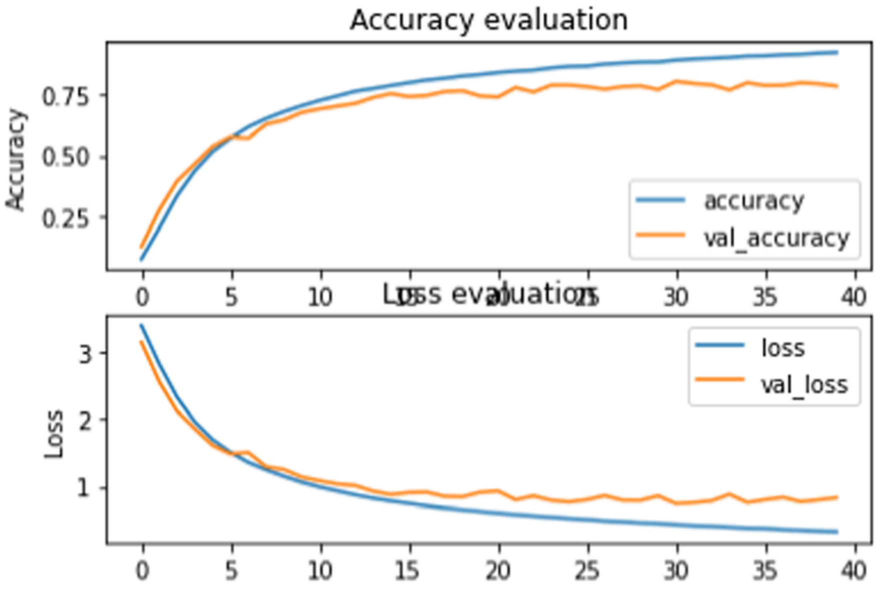


(e) BFCC



(f) LFCC

Fig. 8. continued



(g) NGCC

Fig. 8. continued

Table 1. Comparison between accuracy and loss of different feature extraction methods

Model	Feature extraction method	Test accuracy	Test loss
2D Sequential Convolution Neural Network Model	MFCC	88.20%	0.49
	MFCC-Delta	86.62%	0.57
	MFCC-Double Delta	83.01%	0.73
	GFCC	81.41%	0.77
	BFCC	68.65%	1.21
	LFCC	81.66%	0.73
	NGCC	78.11%	0.84

5 Conclusion

In this paper, a comparison of performance of automatic speech recognition between different feature extraction methods; MFCC, MFCC-$\triangle$, MFCC-Double$\triangle$, GFCC, BFCC, LFCC, NGCC trained for a sequential 2D CNN model was presented. Experimental results were acquired for the Speech Command Dataset by TensorFlow consisting of 65,000 wav files of one-second utterance of 30 short words. The highest accuracy was achieved for MFCC feature extraction method; 88.20% for the sequential model and lowest accuracy was achieved using BFCC feature extraction method; 68.65%. Future research would include a variety of modifications, testing, and experiments. For example, the suggested system's performance can be evaluated using a larger vocabulary and many language datasets. Noisier scenarios must also be evaluated, like resounding noise effects, colourful noises, background music, and mixtures of external noises.

References

1. Wanli, Z., Guoxin, L.: The research of feature extraction based on MFCC for speaker recognition. In: Proceedings of 2013 3rd International Conference on Computer Science and Network Technology, pp. 1074–1077. IEEE (Oct 2013)
2. Dev Amita Agrawal, S.S.: A Novel MFCCs Normalization Technique for Robust Hindi Speech Recognition 17th International Congress on Acoustics (ICA) Rome, 2–7 Sept (2001)
3. Kherdekar, V.A., Naik, S.A.: Convolution neural network model for recognition of speech for words used in mathematical expression. Turkish J. Comput. Math. Educ. (TURCOMAT) **12**(6), 4034–4042 (2021)
4. Huang, J.T., Li, J., Gong, Y.: An analysis of convolutional neural networks for speech recognition. In: 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 4989–4993. IEEE (Apr 2015)
5. Abdel-Hamid, O., Mohamed, A.R., Jiang, H., Deng, L., Penn, G., Yu, D.: Convolutional neural networks for speech recognition. IEEE/ACM Trans. Audio Speech Lang. Process. **22**(10), 1533–1545 (2014)
6. Phani Bhushan, S., Vani, H.Y., Shivkumar, D.K., Sreeraksha, M.R.: Stuttered speech recognition using convolutional neural networks. Int. J. Eng. Res. Technol. (IJERT) NCCDS **9**(12) (2021)
7. Gulzar, T., Singh, A., Sharma, S.: Comparative analysis of LPCC, MFCC and BFCC for the recognition of Hindi words using artificial neural networks. Int. J. Comput. Appl. **101**(12), 22–27 (2014)
8. Li, X., Zhou, Z.: Speech command recognition with convolutional neural network. In: IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (2017)
9. Kubanek, M., Bobulski, J., Kulawik, J.: A method of speech coding for speech recognition using a convolutional neural network. Symmetry **11**(9), 1185 (2019)
10. Tamazin, M., Gouda, A., Khedr, M.: Enhanced automatic speech recognition system based on enhancing power-normalized cepstral coefficients. Appl. Sci. **9**(10), 2166 (2019)
11. Kanthi, A.N., Moinuddin, M.: Speaker Identification based on GFCC using GMM. Int. J. Innov. Res. Adv. Eng. **1**(8), 224–232 (2014)
12. Dev, A., Bansal, S.: Emotional Hindi speech: feature extraction and classification. In: IEEE Explorer, Computing for Sustainable Global Development (INDIACom), 2015 2nd International Conference (989-9-805-4415-1), 11–13 Mar 2015, pp. 1865–1868 (2015)

13. Maurya, A., Kumar, D., Agarwal, R.K.: Speaker recognition for Hindi speech signal using MFCC-GMM approach. Procedia Comput. Sci. **125**, 880–887 (2018)
14. Gambhira, P., Devb, A.: A run-through: text independent speaker identification using deep learning. In: Artificial Intelligence and Speech Technology: Proceedings of the 2nd International Conference on Artificial Intelligence and Speech Technology, (AIST2020), 19–20 Nov 2020, Delhi, India, pp. 139. CRC Press (June 2021)
15. Warden, P.: Speech commands: A dataset for limited-vocabulary speech recognition (2018). arXiv preprint arXiv:1804.03209
16. Kumar, C., Ur Rehman, F., Kumar, S., Mehmood, A., Shabir, G.: Analysis of MFCC and BFCC in a speaker identification system. In: 2018 International Conference on Computing, Mathematics and Engineering Technologies. iCoMET 2018 - Invent, Innovate and Integrate for Socioeconomic Development: Conference Proceedings, vol. Jan–Dec, pp. 1–5 (2018)