

Research Article

Feature Extraction Approach for Speaker Verification to Support Healthcare System Using Blockchain Security for Data Privacy

Shrikant Upadhyay ¹, **Mohit Kumar** ², **Ashwani Kumar** ³, **Ramesh Karnati**,⁴
Gouse Baig Mahommed ⁴, **Sara A. Althubiti** ⁵, **Fayadh Alenezi** ⁶, and **Kemal Polat** ⁷

¹Department of Electronics & Communication Engineering, Cambridge Institute of Technology, Ranchi 834001, India

²Department of Computer Science & Engineering, Cambridge Institute of Technology, Ranchi 834001, India

³Department of Computer Science & Engineering, Sreyas Institute of Engineering and Technology, Hyderabad 500068, India

⁴Department of CSE, Vardhaman College of Engineering, Hyderabad, India

⁵Department of Computer Science, College of Computer and Information Sciences, Majmaah University, Al-Majmaah 11952, Saudi Arabia

⁶Department of Electrical Engineering, College of Engineering, Jouf University, Saudi Arabia

⁷Department of Electrical and Electronics Engineering, Faculty of Engineering, BoluAbantIzzetBaysal University, Bolu, Turkey

Correspondence should be addressed to Kemal Polat; kpolat@ibu.edu.tr

Received 19 May 2022; Revised 19 June 2022; Accepted 12 July 2022; Published 25 July 2022

Academic Editor: Xue Fei Deng

Copyright © 2022 Shrikant Upadhyay et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Speech is one form of biometric that combines both physiological and behavioral features. It is beneficial for remote-access transactions over telecommunication networks. Presently, this task is the most challenging one for researchers. People's mental status in the form of emotions is quite complex, and its complexity depends upon internal behavior. Emotion and facial behavior are essential characteristics through which human internal thought can be predicted. Speech is one of the mechanisms through which human's various internal reflections can be expected and extracted by focusing on the vocal track, the flow of voice, voice frequency, etc. Human voice specimens of different ages can be emotions that can be predicted through a deep learning approach using feature removal behavior prediction that will help build a step intelligent healthcare system strong and provide data to various doctors of medical institutes and hospitals to understand the physiological behavior of humans. Healthcare is a clinical area with data concentrated where many details are accessed, generated, and circulated periodically. Healthcare systems with many existing approaches like tracing and tracking continuously disclose the system's constraints in controlling patient data privacy and security. In the healthcare system, majority of the work involves swapping or using decisively confidential and personal data. A key issue is the modeling of approaches that guarantee the value of health-related data while protecting privacy and observing high behavioral standards. This will encourage large-scale perception, especially as healthcare information collection is expected to continue far off this current ongoing pandemic. So, the research section is looking for a privacy-preserving, secure, and sustainable system by using a technology called Blockchain. Data related to healthcare and distribution among institutions is a very challenging task. Storage of facts in the centralized form is a targeted choice for cyber hackers and initiates an accordant sight of patients' facts which will cause a problem in sharing information over a network. So, this research paper's approach based on Blockchain for sharing sufferer data in a secured manner is presented. Finally, the proposed model for extracting optimum value in error rate and accuracy was analyzed using different feature removal approaches to determine which feature removal performs better with different voice specimen variations. The proposed method increases the rate of correct evidence collection and minimizes the loss and authentication issues and using feature extraction based on text validation increases the sustainability of the healthcare system.

1. Introduction

The sharing of healthcare information in different institutions is quite complex, which requires strong agreement with the promising model as there is a remarkable increase in the exchange of medical data in the network [1]. The main reason not to share data in network is privacy worry [2], creating panic while sending information may give an ambitious choice [3].

Blockchain technology arises as a key technology in the digital uprising of various verticals like healthcare, supply chains, and automobile industry. Many techniques are available for information sharing either using extraction or Blockchain approach but the integration of both found to be less from data privacy point of view. Creating a stable and more secure platform using updated dispersal blockchain increases the potential of healthcare system and that motivates highly. Blockchain utilization for healthcare management of data creates advantages for doctors, patients, and healthcare institutes for proper management of medical records, payment and claims management, data validation, and financial auditing. This also permits healthcare units to restrict the unsanctioned user to access sensitive data and restrict for any modification to patient's data for its sustainability. The primary goal of such technology is to deploy by integrating other existing technology to make an efficient health fact system with infrastructure to fulfil the healthcare system necessity.

There is no comprehensive agreement on the particular infrastructure required to hold up such a job [4]. Knowledge of the meaning and structure of data is a primary requirement for the health while sending information among different institutions. Considering that facts are transferred securely and efficiently, these coordination businesses remain unreached to control facts' value [5]. For a far talking or speaking scenario, clean speech is somehow disturbed by reverberation. Due to this disturbance, an individual's speaker features also get distorted. There is a quarrel between the outline trial and instructed talk model getting recognized, which reduces the talked membrane's production [6]. To minimize this conflict, we present three steps involved methods for far speaker recognition from talk development, model, and attribute compensation opinion as depicted in Figure 1. Feature extraction from a speech has its own importance. Every person has its unique voice quality, voice throw, speaking style, talking, and pitch, and gathering the information from voice plays a crucial role for any model. GMM compared to other modeling approaches adds unique features because its training is relatively fast, and models can be updated and scaled to newly added speech easily. It is composed of a finite composition of multivariate Gaussian mixtures and calculates density function $\delta_\lambda(x)$ as per:

$$\delta_\lambda(x) = \sum_{i=1}^P m_i \delta_i(x), \quad (1)$$

where "P" is the Gaussian term and m_i is the prior probability of i^{th} D-variate density function given as:

$$\delta_i(x) = \frac{1}{(2\pi)^{Q/2} |\Sigma_i|^{1/2}} e^{-(1/2)(x-\mu_i)^T \Sigma_i^{-1} (x-\mu_i)}, \quad (2)$$

where Σ_i and μ_i denotes the covariance matrix and mean vector of multidimensional Gaussian model [7]. Extraction of speech means taking some special characteristics which makes the identity unique for the person and distinguishing different from others in term of voice quality and known features will help to collect its original identity.

One of the major limitations of conventional GMM is its super vector presentation and successive factor model investigation which does not be counted of the fact that original acoustic characteristics [8]. The main contribution is the modeling of training and testing that verifies the said text using threshold level which collects the correct evidence. At the same time, the Blockchain contributes to monitor the sensitive data and maintain its security and privacy at each level while transferring in network.

2. Blockchain in Healthcare

A patient's medical status requires complete monitoring by the hospital and doctors for its better treatment while maintaining privacy and data security of sensitive data of patients to share them with medical institutes and leading hospitals for expert consultation for gathering better information about the related cases. As per accountability law and health insurance, law enforcement and many other public agencies access medical information legally. It is approximated that around 200 to 500 individuals may have the right to read the health records of any patient without any authentication and permission [9]. When information is distributed extensively and kept in multiple outbreaks, securing data is one of the more crucial issues. As per Ponemon Institute, in the year 2016, about 112 million information related to medical were negotiated and such break of data tempered and attacks raised by 162% in the year 2017 [10, 11]. Blockchain ideal in healthcare arises when there is a need for interoperability and security comes into the picture. Many healthcare gadgets and healthcare-related mobile applications boom with the evolution of and a large number of medical facts and important recorded data are transferred in today's life from various healthcare units. So, much facts traffic requires better management resources for its security and privacy. Blockchain is a technology that supports and provides a proper solution that will help to keep the recording. Secure and broadcast of medical information that maintains every patient's privacy to access their medical reports [12–22]. It enables an auditable path and creates transparency for individuals using their encryption key and unique credentials, and allows third parties to provide a grant to access individual medical records. It includes authority grants in terms of health competence, resource contributors, social care, and investigator to execute their medical facts and descriptions for the motive of healthcare delivery directly, statistical analysis, grant research, etc. Blockchain technology uses wise agreement to maintain records of patient interactions. The supplier generates a document, and once it is verified, the

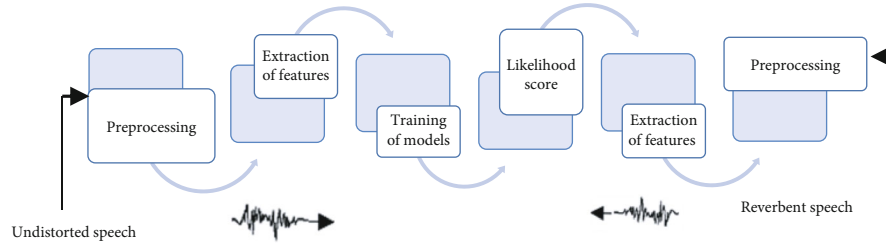


FIGURE 1: Method for speaker identification with undistorted and reverberant speech.

sufferer may have the right to view the reports [23–25]. An automatic notification will be generated after receiving new facts, and a coded indicator reflects to latest medical history. A chain is used to store the permission, which allows the patient to control access to their information. Areas where it is applicable will be vaccination records, blood trials, medicinal, therapeutic instruction, etc. [26–36].

Example 1. In the COVID-19 scenario for the healthcare system, it is challenging to acquire medical apparatus and medical supplies due to excessive demand. Issue of trust always arises for the collapse of stock sequence with authentic vendors. Various issues like the delivery deadline, payment, standards, customer verification, and cheating. This technology possibly assures such cases and maintains the credibility of distributors. So, by deploying such technology, support deadline and monitoring medicine requirements where there is excessive demand for medical needs. Blockchain technology adds transparency, security, and authentication by putting consent standards and protocols into a good flow of any aid. This is one of the primary concerns for the healthcare unit to become competent, and this can create a bridge between supply and demand chains for future advancement in the healthcare system.

Example 2. It can be suitable for wireless sensor network where nodes are dependent on each other for information sharing and dynamic nature make it unsecure for that data transmission.

Example 3. It can also be suitable for natural language processing for the text authentication in speech processing using different extraction approaches.

3. Facts Addition to Blockchain

The fundamental occupation for addition of facts to the Blockchain is depicted in the figure below. Similar to bit coin concept, a bank is attached to Blockchain of uniform duration as this will be in bit coin identified by work function evidence. For network as depicted below, there is a uniform duration for generating a block. Inside this duration of block duration network goes through different phases of pursuit. The first phase is the agreement digger initialize at duration D_1 and this undertaking transferred to the matching node, and this will continue until D_α when matching node no longer accepting updated undertaking for block. The phase of Blockchain involved in attaching chunk to care unit is depicted in Figure 2.

The coupling coefficient helps in attaching chunk to healthcare system that sign an agreement with differ and generates an effect of access to pass between the block. It can be modified by granting permission to each node and updating locally Blockchain for new dispersal. This grant permission to every node acts as a contributor once in undertaking to digitally mark to block, specifying that they inscribe actuality. The block new comes back to digger in return phase of signed block. The digger node added to the locally blockchain and in final phase dispenses the updated blockchain known as new dispersal blockchain. Algorithm first describes the way of generating new chunk and its addition to Blockchain.

4. Extraction and Interfacing

Extraction means to put chunk (block) to blockchain and similar concept like bitcoin but we try to follow some different technique where overall model bypass work model evidence where power calculation is spanned without putting some intrinsic value. Our aim is to create a network agreement by setting nodes to give evidence by reference of transaction with correct interpretation also at the same time node will verify the proof of potency. Such measures maintain consistency in Blockchain but also encourage better inferring among nodes [37]. Second algorithm discussed the steps involved by assigning a specific profile which is further compared with familiar set of admissible description.

If the description is identified, it will acknowledge with its confirmation using function Descp_Check and this will use the IP_URL to authenticate the server and the result request which will be response with IP_Outcome which will be further verified for conformity by Conf_Function.

Interfacing evidence does not need network to get arrangement with the set profiles which includes the indulge of the associated sets value. So, such agreement does not reach declaratively. Consensus in network is completely depending on process of human where network contestant can handshake and negotiate using both expression profession as well as specialist. Such collaboration is the necessary part for well familiar repository [38].

5. Digger Voting

For work evidence, digger runs to right to sum a chunk for Blockchain. Here, we put a sharing of digger same as multi-chain [39]. Adding these nodes will identify the duration of

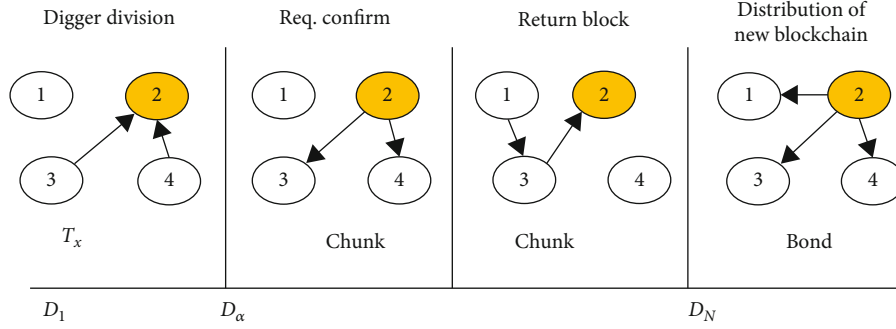


FIGURE 2: Blockchain phase involved in attaching chunk to healthcare system.

Input-Bank of 'N' nodes take part in network.
 Input-Blockchain '2' denotes a sequence $2_0, 2_1, \dots, 2_n$ where 2_n is present chunk on tie.
 Input- D_α is the finish of the distributed information.
 $\alpha \rightarrow \text{selectdiggernode}(2_n, N)$.
 $A \rightarrow []$; //select with vacant set undecided agreement
 While presently $() < D_\alpha$ do
 Forevery $n \in N (-1)$ do
 $A \leftarrow AU$ received agreement from node (n);
 Block assemble (A) $\rightarrow b_\gamma$
 $N^\circ \{n \rightarrow \in N | (\exists q)[q \in Aiscreator(n, q)]\} || N^\circ$ for all nodes ≥ 1
 For every $n \in N^\circ$ do
 Block sign (b_γ, n) ;
 $A^I \rightarrow \text{block add}(A, b_\gamma)$
 For every $n \in NI$
 Blockchain distribution (A^I, n)

ALGORITHM 1

Input-'UA' set of undecided agreement
 Input-'AG' set of consent agreement with network profile
 Input- Present chunk C_β being collected
 Output- Rational agreement with set 'w'
 1. $w \rightarrow []$; // Start with vacant set rational
 2. For every $t \in UA$ do
 3. $v \rightarrow \text{Get Descp_Check}(t)$;
 4. $UA \rightarrow \text{Get Descp_Check}$;
 5. If $UA \in AG$ then
 6. $\text{Result} \rightarrow \text{Conf_Descp}(v, UA)$;
 7. If obeys (Result) then
 8. $w \rightarrow w \cup [t]$;

ALGORITHM 2

Input- Assign contestant participation input 'N'
 Input- The updated chunk on Blockchain, b_x
 Output- Random digger voting β
 1. $R \rightarrow \text{Obtain random grain}(z_n)$
 2. $p \rightarrow \text{Obtain Hash}(\#) \text{ block}(z_n)$
 3. for every $r \in R$ do
 4. $p \rightarrow \#(p + r)$
 5. $\beta \rightarrow \text{where } | \text{obtain public key}(h) - p$

ALGORITHM 3

starting of chunk and who will be next digger. So, agreement may be sent directly without distribution to the whole network. Also, it assures that digger work requires to maintain the network smoothness dispense steadily.

Finally, by removing the contention of work evidence, we minimize and remove misused mathematical attempt. The third algorithm discussed the digger voting process. Recalling first algorithm where and step of collecting a

chunk to Blockchain is for the node take part to enter. During this, allowing phases every node requires to deposit a random code in form of number for digger voting which is collected in line number 1 and hashed (#) with block to generate an updated new number.

The next digger becomes the node whose key (public) is nearest to that value and this process did two works

- (a) Probability to become digger in network for any node will be $1/N_n$ ($N = \text{network}$, $n = \text{number of nodes}$) and

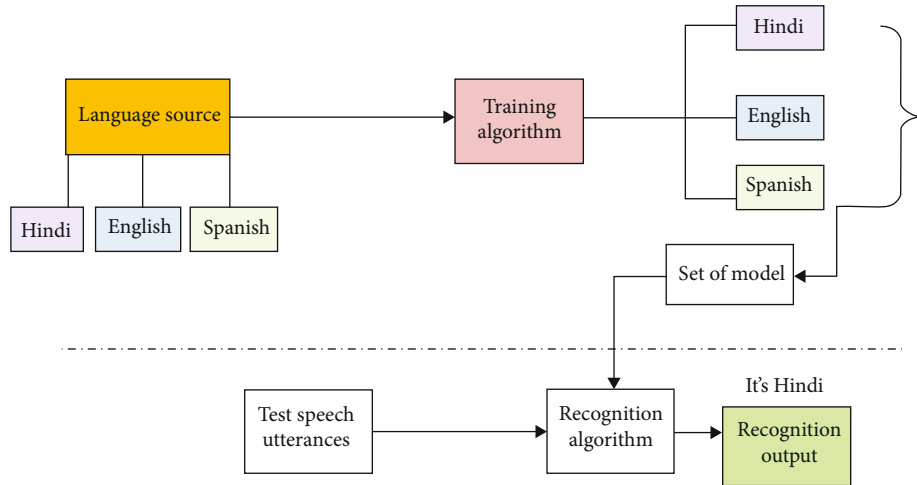


FIGURE 3: Phases involved in language recognition.

TABLE 1: Mean and deviation of different group.

Smoothened cepstral peak prominence	Group	Standard deviation	Mean
Sustained phonation	Choral singers	1.52	6.81
	Nonsingers	2.30	3.47
	Trained singers	1.66	7.10
	Overall	2.48	5.70
Reading task	Choral singers	0.48	3.62
	Nonsingers	0.45	2.20
	Trained singers	2.27	3.81
	Overall	1.44	3.17

- (b) Number that is generated randomly used for voting is planted by all contestant nodes in network and this will protect a node to issue a random code and selecting itself or with collaborator
- (c) Randomly generated key provides a secure code to the participants

6. Improvement in Data Privacy

As discussed, facts security in terms of invisibility and privacy is the basic preference of any system. Security level in our proposed system based on several features includes special encryption, protection of privacy in term of keyword, and agreement in a smart way.

Encryption in Blockchain includes that whatever it stored is not a plain book public facts calculation for every node is predicted as encrypted using key shared in network and awareness facts should be encoded by activated node in network. Ease of fact exploration and availability secretly protected antonyms digging methods are used [40]. Following such techniques, outer institution may request for agreement from Blockchain to match some benchmark both for any query and agreement encryption.

After this, the validation and identification process starts for the talker using Hindi speech for 50 speakers which was conducted using Cool tool and using MATLAB as a simulation tool.

7. Validation and Identification (Speaker)

Taker recognition is very crucial to understand how a person understands the variables and voices that affect the hearing capabilities and the performance of listeners is firmly affected by speaking scenario and acoustic. Various investigations have been done to compare visual and aural talker recognition which reflects that aural inequity error is quite less than discrimination spectrogram error. In earlier, an attempt to a speaker recognition using automatic process was done that will help to analyze spoken utterance acoustically [41, 42]. Talker recognition spells the congruence of the talker where talker validation involves a method of rejecting or to obtain the identity of asserts users. Talker’s validation is further divided into part which involves session training when model for user’s talk is built up and real authentication is done. A model is first trained for new user’s voice and be authenticated by mean of spectral investigation from which feature removal was done to create a talker model. Then, users’ talk can be confirmed by weigh up against the instructed dataset of models. After comparison, the system will take decision where the claim’s naming is one which is designed by teaching material available or not. The identification of different language is depicted in Figure 3.

7.1. Example: Trained and Untrained Talk. Both choral and trained signers have high cepstral compared to untrained singer and singing involves regular practice which involves scale, matching, notes, and its rhythms. The table shown below reflects that trained singer is found to be efficient compared to untrained in term of range, stability, control, etc. and have better control over loudness and pitch independent of frequency and intensity and nonsinger found with low proficiency in voice. In this research paper, we try

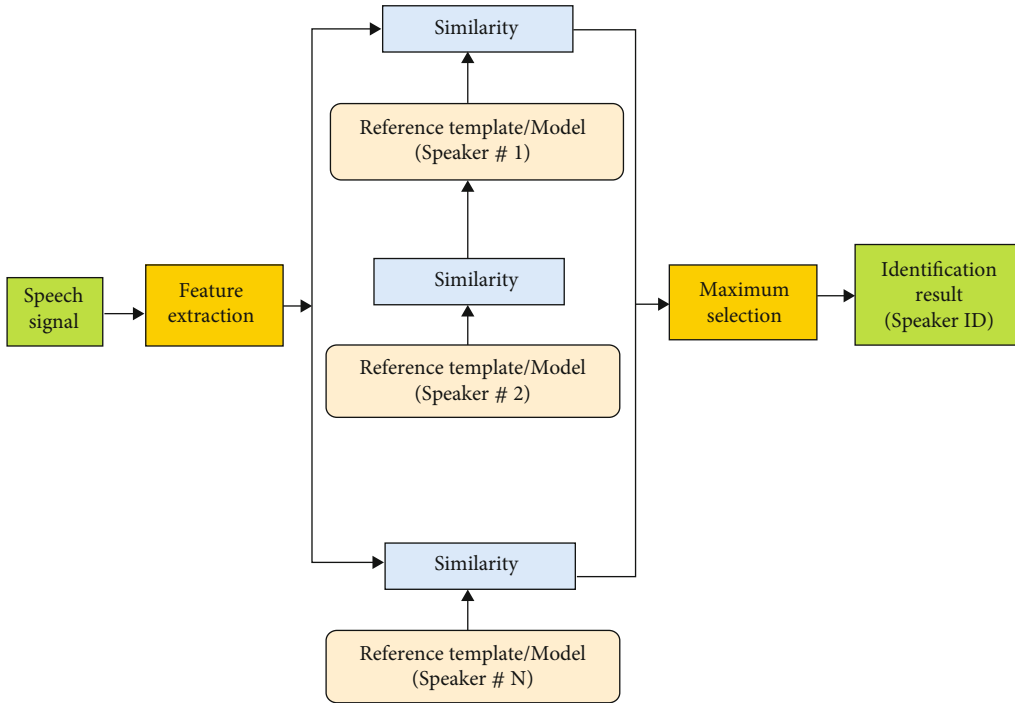


FIGURE 4: Recognition system consists of close set.

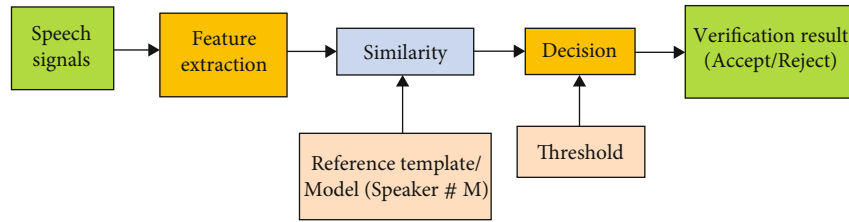


FIGURE 5: Basic structure for talker validation.

TABLE 2: Dataset of 50 speakers.

D:\DB_50\0\00	D:\DB_50\0\01	D:\DB_50\0\02	D:\DB_50\0\03	D:\DB_50\0\04
D:\DB_50\0\05	D:\DB_50\0\06	D:\DB_50\0\07	D:\DB_50\0\08	D:\DB_50\0\09
D:\DB_50\1\00	D:\DB_50\1\01	D:\DB_50\1\02	D:\DB_50\1\03	D:\DB_50\1\04
D:\DB_50\1\05	D:\DB_50\1\06	D:\DB_50\1\07	D:\DB_50\1\08	D:\DB_50\1\09
D:\DB_50\2\00	D:\DB_50\2\01	D:\DB_50\2\02	D:\DB_50\2\03	D:\DB_50\2\04
D:\DB_50\2\05	D:\DB_50\2\06	D:\DB_50\2\07	D:\DB_50\2\08	D:\DB_50\2\09
D:\DB_50\3\00	D:\DB_50\3\01	D:\DB_50\3\02	D:\DB_50\3\03	D:\DB_50\3\04
D:\DB_50\3\05	D:\DB_50\3\06	D:\DB_50\3\07	D:\DB_50\3\08	D:\DB_50\3\09
D:\DB_50\4\00	D:\DB_50\4\01	D:\DB_50\4\02	D:\DB_50\4\03	D:\DB_50\4\04
D:\DB_50\4\05	D:\DB_50\4\06	D:\DB_50\4\07	D:\DB_50\4\08	D:\DB_50\4\09

to put all effort on information sharing with total control on facts by agreement concept between nodes and generation of random codes for authenticated users. Patient is your prime focus whose various health-related information shared to various medical institutions for proper analysis and fast recovery.

Privacy is one of the major concerns for any department and institutions. Healthcare-related data or facts shared

through network are very challenging one. Sending of information using any network must have capability only to share information from source to destination and its security with privacy is real issue that has to be taken care of.

From Table 1, it is found that choral singer have a mean of 6.80 (± 1.53) for comfort phonation and for reading it is 3.63 (± 0.47). It shows that cepstral smoothed prominence cost of nonsinger are less than trained and group singers.

TABLE 3: Feature vector value for 50 speakers and its parameters (a–d)

(a)				
-0.836654	1.119128	-0.949380	0.946176	-0.008731
-0.846651	1.181489	-0.929819	0.911423	0.110117
-0.961810	1.390200	-1.052452	1.069067	-0.045332
-0.795792	1.283035	-1.010187	0.892773	-0.047438

(b)			
-0.607713	0.164727	-0.584664	-0.161334
-0.732032	0.258739	-0.537848	-0.110984
-0.748946	0.325589	-0.532755	0.015159
-0.557250	0.247957	-0.373091	-0.029203

(c)				
0.256184	0.334360	0.162019	-0.110454	0.276337
0.074978	0.460313	0.044143	-0.198783	0.314915
0.081630	0.343054	0.008088	-0.249718	0.358225
0.037462	0.342918	-0.066260	-0.160299	0.235376

(d)	
Parameters	Values
Talkers	0-9
Channel (single)	5KHz
Rate of sampling	8 bit (used for low noise)
Dialects	Text (Hindi)
Each speaker repetition rate	5 digits/minute

Recognition using speech in general used to include different way for discriminating human depending upon their voices. In recognition, a speech observation of unrecognized talker is analyzed and differentiates with all familiar talkers. The unrecognized talker is identified as talker whose model fit best the input declaration. The fundamental structure of talker recognition is depicted in Figure 4.

Probability concept involved for the generation of random codes for every user and mathematical complexity reduction algorithm for fast processing of incoming information. Keeping facts and records safely is a need of an hour and every healthcare institution focus on such technology for flawless exchange of information inside network system. Revolution and large amount of health-related gadgets involved for bulk amount of medical record transfer. So, security of shared data is very important and must have much more fighting capabilities against the cloud thief to make the facts clean.

A closed set or unlocked set is termed for the talker recognition where in unlocked set assumption of test observation contain one the 'N' register talker and in open set recognition there will be addition determination to identify whether verified observation was declared by one of 'N' reg-

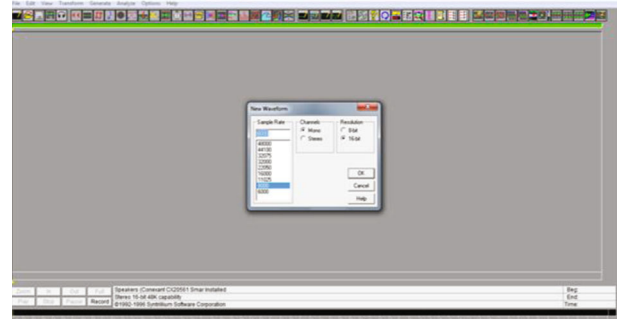


FIGURE 6: Parameters selection procedure as per specimen.

ister talker or not i.e. N_{+1} determination. The basic structure of talker validation is depicted in Figure 5.

The aim of this structure is to allow or reject the talker who claim based on speech specimen. Matching among reference level and test is above certain threshold limit the assert is welcomed. A very high threshold arises a problem for fake to accept by system but always arises an issue of rejecting authenticate one. Low threshold make sure to accept the genuine one but also arise of accepting a fake one. Such issues can be solved by correlation metrics obtained from the speech spectra and feature vector modeling following parametric techniques. Here we also try to consider text-dependent stem as this will have high degree of control over speech conditions.

8. Feature Removal

The principle behind of this approach is to low down the average of total difference that arises between calculated speech and original speech over a definable period. The coefficient prediction can be used to find distinctive set which is generally 20 milliseconds lengthy. This approach is known as linear predictive technique [43].

Gain is an important parameter and time variable transfer function of filter is given as-

$$G(z) = \left(\frac{h}{1} - \sum P_c - s \right), \quad (3)$$

where P_c = coefficients forecasts and $s=1$ to n .

The decision taking in term of analysis of every frame of unspoken and spoken signal was done using this approach. It also helps to employ pitch identification algorithm to detect best possible frequency of pitch and re-emphasis it as gain is varying frame to frame [43].

The next important approach is coefficient acoustic measure of mel frequency and considered to be a standard approach for feature removal. At least twenty coefficients are required in automatic recognition of speech where maximum ten to twelve coefficients are found to be sufficient for speech coding. This coding is quite sensitive to distortion or noise due to its spectral nature [44]. Procedure that employ facts in revolution of speech can be considered to minimize such issues as speech contains causal captivates.

Frequency nonlinear in nature used to approximate the mel component which is approximately below 1

TABLE 4: Error rate in less acoustic scenario (noise).

Dialects	Train (sec)	Test (sec)	Feature extractions		
			LPC (%ER)	MFCC (%ER)	LPCC (%ER)
Hindi					
lnk	10	10	10.01	8.01	9.11
/kU; okn	10	5	11.88	7.77	10.02
gkWa	10	20	9.90	10.03	11.73
okyk	5	10	12.74	10.75	11.88
ebZ	5	5	9.88	8.75	9.78
vdcj	5	20	7.71	9.77	8.33
iz t	20	10	9.46	6.56	7.19
ia[kk	20	5	10.89	9.77	7.36
eka; k	20	20	9.22	5.90	9.79
vua	20	10	8.66	6.57	12.34

TABLE 5: Efficiency rate in (%) of individual feature extraction.

Hindi dialects	LPC (%)	MFCC (%)	LPCC (%)
Lnk	90.96	92.86	91.29
/kU; okn	89.73	93.38	88.43
GkWa	91.26	88.96	89.12
Okyk	88.89	87.23	87.23
EbZ	92.95	90.12	91.56
Vdcj	89.93	91.48	92.89
iz t	90.28	92.71	88.62
ia[kk	89.86	93.18	91.93
eka; k	90.86	95.34	86.46
Vua	93.48	92.14	87.46

kilohertz and above 1 kilohertz for frequency in logarithmic. Person auditory structure has minimum selective of frequency above 1 kilohertz. The feature of this approach follows energy of log filter storage to analyze the output [45–48]:

$$(t) = \ln \left(\sum_n^{N-1} |S_t(n)|^2 G_t(n) \right) 0 \leq t < T, \quad (4)$$

where $S_t(n)$ is the Fourier transform of n^{th} speech input frame and $G_t(n)$ is the response of t^{th} filter storage, T is total filter used, and N is size of window used to transfer. Then, at log energy, cosine transformation will be:

$$(t) = \sum_{l=0}^{L-1} R_l(n) \cos \left(\pi m \left(\frac{n-0.5}{T} \right) \right) 0 \leq t < T. \quad (5)$$

So, to gather fluctuation in coefficient over duration

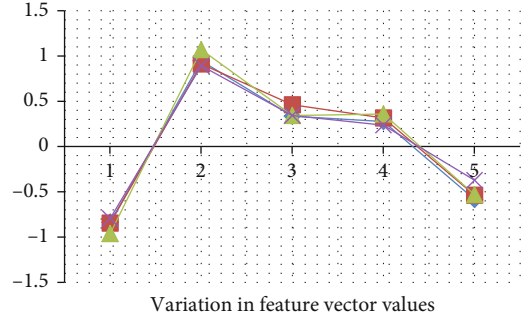


FIGURE 7: Feature vector variation for talkers.

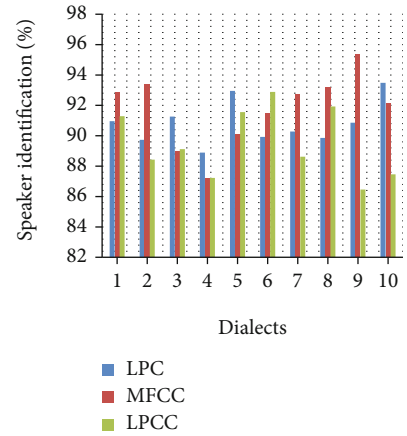


FIGURE 8: Efficiency of different feature extraction for speaker identification.

1st and 2nd coefficient differences are calculated as:

$$\begin{aligned} \Delta(\vec{f}_t) &= (\vec{f}_t) + 2 - \vec{f}_t - 2, \\ \Delta\Delta(\vec{f}_t) &= \Delta\vec{f}_{t+1} - \Delta\vec{f}_{t-1}. \end{aligned} \quad (6)$$

Linear prediction approaches are used for the estimation of fundamental component of speech. In this method, input is first prehighlighted using filter of high pass and its transfer function will be [49]:

$$Z_p(n) = 1 - bh^{-1}. \quad (7)$$

Now, windowing is carried out to minimize the discontinuities in each frame at edge using hamming to smooth side lobe and is given as:

$$v(w) = 0.55 - 0.45 \cos \left(2\pi \frac{m}{M} \right); 0 \leq m \leq M. \quad (8)$$

The shape of spoken span basically determines the temper of produced sound and to model the vocal region

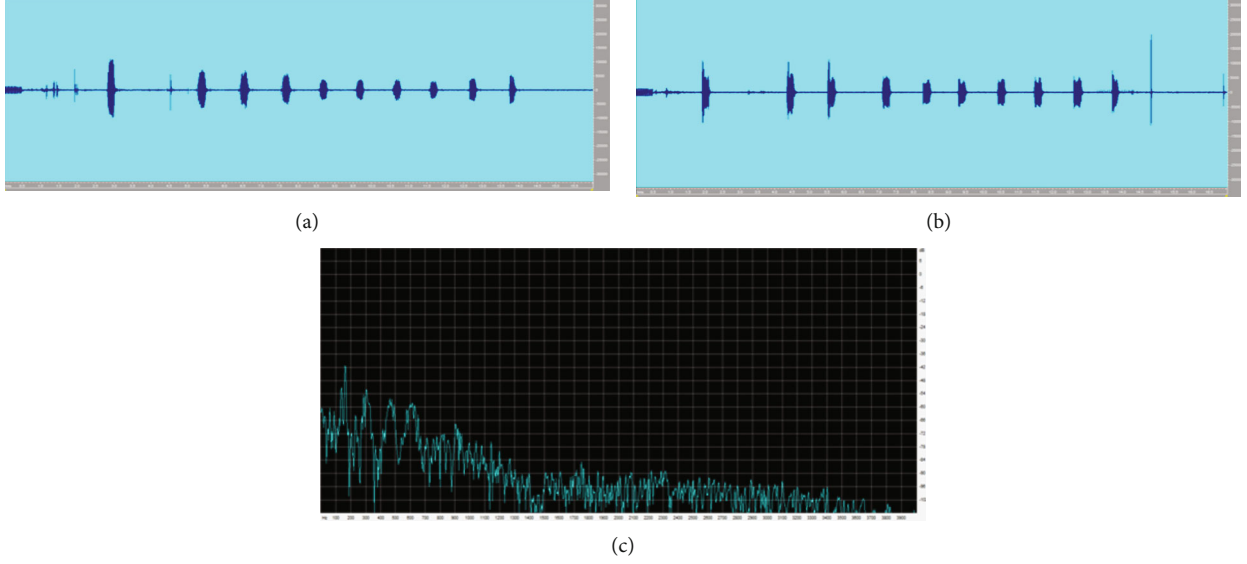


FIGURE 9: Variation in the spectrum for different feature of speech (a–c).

we require transfer function in domain:

$$X(z) = \frac{H}{1 - \sum_{n=1}^r b_k z^{-k}}, \quad (9)$$

where $X(z)$ represents the spoken span transfer function, H is the gain, and b_k is the throwback coefficients. Last phase which includes cepstral evaluation of sequenced speech was done.

$$p(n) = \frac{1}{2\pi} \int_{-\pi}^{+\pi} \ln v(w) e^{jnw} dw. \quad (10)$$

Another method to estimate cepstral is linear perception cepstral coefficients. The feature of this approach follows energy of log filter storage to analyze the output:

$$R_p(t) = \ln \left(\sum_n^{N-1} |S_t(n)|^2 G_t(n) \right) \quad 0 \leq t < T. \quad (11)$$

The linear scale equation will be:

$$LS = \sqrt{\frac{ns^2}{(a_1)^2 + (a_2)^2 + \dots + (a_n)^2}}. \quad (12)$$

Suppose “ S ” is the target person’s speech represented by $S = \{s_1, s_2, \dots, s_T\}$ and “ V ” is the unobservable part of speech dependent represented by $V = \{v_1, v_2, \dots, v_T\}$; then for statistically independent S and V probability, $P(x_t, j, k)$ can be calculated as [47]:

$$P(x_t, j, k) = \int_{-\infty}^{\infty} M \left(s; \sum_{s,j} \mu_{s,j} \right) M \left(x_t - s; \sum_{v,k} \mu_{v,k} \right), \quad (13)$$

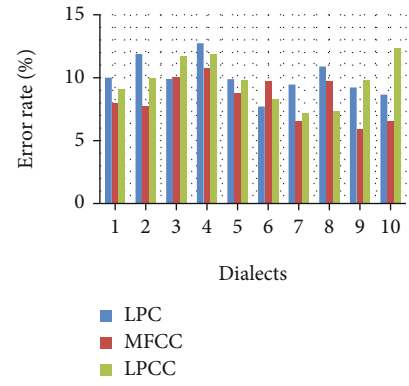


FIGURE 10: Error detection of different feature removal approach for Hindi dialects.

where $\sum_{s,j}$ and $\sum_{v,k}$ are weight covariance matrix and $\mu_{s,j}$ and $\mu_{v,k}$ are mean vector and M denotes the Gaussian density matrix.

A linear predictor of order “ p ” with prediction coefficient (α_k) is defined as a system whose output is defined as [50–53]:

$$\hat{s}(n) = \sum_{k=1}^p \alpha_k s(n-k). \quad (14)$$

The system function is p^{th} order polynomial and it follows:

$$P(z) = \alpha_k z^{-k}. \quad (15)$$

The prediction error $e(n)$ is defined as:

$$e(n) = s(n) - \hat{s}(n) = s(n) - \sum_{k=1}^p \alpha_k S(n-k). \quad (16)$$

The transfer function of prediction error sequence is:

$$A(z) = 1 - \sum_{k=1}^p \alpha_k z^{-k}. \quad (17)$$

9. Result and Discussions

The dataset for fifty speakers with its value achieved is shown in Table 2. This dataset consists of fifty talked persons of different ages and group which help to match the pattern of the said voice and authentication will be done using this dataset. Text-dependent text was considered for the analysis of our research which will help the model to judge the quality of speech with better accuracy. The dataset for fifty speakers is shown in Tables 2 and 3 which represents the feature vectors for fifty speakers with their values from (a) to (c) and (d) denotes the parameters considered for its simulation with its channel, rate of sampling, repetition rate, etc. The parameter section procedure is shown in Figure 6.

This feature value obtained reflects the quality of talker that will help in validation of speakers using different dialects and measure the specimen best possible choice in term of its quality and accuracy. The error rate and efficiency rate are shown in Tables 4 and 5.

The feature vector variation for different talkers is depicted in Figure 7 where LPCC reflects the higher value compared to other said feature extraction. The efficiency of considered extraction is depicted in Figure 8 which is used for the identification of speakers. The error identification rate for the considered feature extraction is depicted in Figure 9 where LPC reflects the higher percentage of error rate compared to others. The variations of different features of talked speech are represented in Figures 9(a)–9(c). The error detection of different feature removal approach for Hindi dialects is shown in Figure 10.

Voice signal while sending network may degrade its original version so its original feature is very crucial to reach destination for matching. Calculating the efficiency rate of different feature removal approach of LPC, MFCC, and LPCC using Hindi dialects, we observed that the MFCCC will have better efficiency of 95.34% compared to LPC and LPCC with 92.95% and 92.895.

Training and validation under different simulation time up to 20 seconds have been conducted to calculate the error rate where we find that MFCC will have the least error rate of 5.90% compared to LPC and LPCC, i.e., 7.71% and 7.19%.

10. Conclusion

Data security and privacy are the primary concerns of this research article. Three algorithms play a crucial role by updating the node chain and adding it to the local Block-

chain. The evidence collection procedure is also highlighted with some sort of agreement procedure to minimize the evidence loss and authentication construction, also putting systematic limitations on information access. A familiar set of admissible concepts is also added to differentiate the unknown group.

The confirmation function is used to authenticate the server. Healthcare vital record facts can be significantly protected after this validation is done, and finally voting approach is applied in the last algorithm to minimize the complex mathematical attempt and to make this algorithm fast which will also help to generate a new block of every time for the patients in term of probability. Whenever a sufferer wants to access their information, a random number is generated using this approach and protect to issue an arbitrary unauthorized code for users.

Such innovative agreement will provide more security to medical records and maintains a secure path in the chain supply process. Also, this verification process was done using three feature removal approach considering Hindi dialects which are text-dependent and validation was done in term of efficiency and error rate using a suitable model. Feature extraction is vital as information is transmitted using text or voice. If the agent is a choice, then MFCC will be the first choice for its better efficiency and minimum error rate. The integration of speech identification with the platform controls the talked dialects and gathers accuracy of high predictive, which is crucial for any system. This platform enhances the bright healthcare facility and provides a better way to exchange information to medical and research-related institutions for better care and evaluation of medical records. Blockchain maintains the security in data/facts while transmission creates a security layer for authentication of the correct user by extracting their voice quality using the feature extraction approach.

Data Availability

We can send the datasets at the request of the authors.

Ethical Approval

This article does not contain any studies with human participants. No animal studies were involved in this review.

Conflicts of Interest

The authors declare no conflict of interest.

Authors' Contributions

All authors contributed equally to this work. In addition, all authors have read and approved the final manuscript and given their consent to publish the article.

References

- [1] Y. Ge, D. K. Ahn, B. Unde, H. D. Gage, and J. J. Carr, "Patient-controlled sharing of medical imaging data across unaffiliated

- healthcare organizations,” *Journal of the American Medical Informatics Association*, vol. 20, no. 1, pp. 157–163, 2013.
- [2] C. Clifton, M. Kantarcioğlu, A. H. Doan et al., “Privacy-preserving data integration and sharing,” in *Proceedings of the 9th ACM SIGMOD Workshop on Research Issues in Data Mining and Knowledge Discovery*, pp. 19–26, Paris, France, 2004.
 - [3] J. R. Vest and L. D. Gamm, “Health Information Exchange: persistent challenges and new strategies: Table 1,” *Journal of the American Medical Informatics Association*, vol. 17, no. 3, pp. 288–294, 2010.
 - [4] P. C. Tang, J. S. Ash, D. W. Bates, J. Marc Overhage, and D. Z. Sands, “Personal health records: definitions, benefits, and strategies for overcoming barriers to adoption,” *Journal of the American Medical Informatics Association*, vol. 13, no. 2, pp. 121–126, 2006.
 - [5] J. Walker, E. Pan, D. Johnston, J. Adler-Milstein, D. W. Bates, and B. Middleton, “The value of health care information exchange and interoperability,” *Health Affairs*, vol. 24, Supplement_1, pp. W5-10–W5-18, 2005.
 - [6] J. Ye, T. Zhenmin, and W. Longbiao, “Identification of a distant speaker and its robustness,” *Chinese Journal of Electronics*, vol. 20, no. 2, pp. 278–282, 2011.
 - [7] A. Fazel and S. Chakrabartty, “An overview of statistical pattern recognition technique for speaker verification,” *IEEE Circuit and System Magazine*, vol. 11, no. 2, pp. 62–81, 2011.
 - [8] K. Kinoshita, M. Delcroix, T. Nakatani, and M. Miyoshi, “Suppression of late reverberation effect on speech signal using long-term multiple-step linear prediction,” *IEEE Transactions on Audio, Speech and Language Processing*, vol. 17, no. 4, pp. 534–545, 2009.
 - [9] R. C. Barrows and P. D. Clayton, “Privacy, confidentiality, and electronic medical records,” *Journal of the American Medical Informatics Association*, vol. 3, no. 2, pp. 139–148, 1996.
 - [10] P. Littlejohns, J. C. Wyatt, and L. Garvican, “Evaluating computerised health information systems: hard lessons still to be learnt,” *BMJ*, vol. 326, no. 7394, pp. 860–863, 2003.
 - [11] A. K. Jha, D. Doolan, D. Grandt, T. Scott, and D. W. Bates, “The use of health information technology in seven nations,” *International Journal of Medical Informatics*, vol. 77, no. 12, pp. 848–854, 2008.
 - [12] M. B. Hoy, “An introduction to the Blockchain and its implications for libraries and medicine,” *Medical Reference Services Quarterly*, vol. 36, no. 3, pp. 273–279, 2017.
 - [13] T.-T. Kuo, H.-E. Kim, and L. Ohno-Machado, “Blockchain distributed ledger technologies for biomedical and health care applications,” *Journal of the American Medical Informatics Association*, vol. 24, no. 6, pp. 1211–1220, 2017.
 - [14] X. Yue, H. Wang, D. Jin, M. Li, and W. Jiang, “Healthcare data gateways: found healthcare intelligence on blockchain with novel privacy risk control,” *Journal of Medical Systems*, vol. 40, no. 10, p. 218, 2016.
 - [15] F. Alenezi, “Image dehazing based on pixel guided CNN with PAM via graph cut,” *CMC-Computers, Materials & Continua*, vol. 71, no. 2, pp. 3425–3443, 2022.
 - [16] F. Alenezi, A. Armghan, S. N. Mohanty, R. H. Jhaveri, and P. Tiwari, “Block-greedy and CNN based underwater image dehazing for novel depth estimation and optimal ambient light,” *Water*, vol. 13, no. 23, p. 3470, 2021.
 - [17] G. P. Joshi, F. Alenezi, G. Thirumoorthy, A. K. Dutta, and J. You, “Ensemble of deep learning-based multimodal remote sensing image classification model on unmanned aerial vehicle networks,” *Mathematics*, vol. 9, no. 22, p. 2984, 2021.
 - [18] F. Alenezi and K. C. Santosh, “Geometric regularized hopfield neural network for medical image enhancement,” *International Journal of Biomedical Imaging*, vol. 2021, Article ID 6664569, 12 pages, 2021.
 - [19] F. Alenezi and E. Salari, “A fuzzy-based medical image fusion using a combination of maximum selection and gabor filters,” *International Journal of Engineering Science*, vol. 9, pp. 118–129, 2018.
 - [20] F. S. Alenezi and S. Ganesan, “Geometric-pixel guided single-pass convolution neural network with graph cut for image dehazing,” *IEEE Access*, vol. 9, pp. 29380–29391, 2021.
 - [21] S. Majid, F. Alenezi, S. Masood, M. Ahmad, E. S. Gündüz, and K. Polat, “Attention based CNN model for fire detection and localization in real-world images,” *Expert Systems with Applications*, vol. 189, article 116114, 2022.
 - [22] Q. Wang, J. Zhao, S. Xu et al., “ExHIBit: Breath-based augmentative and alternative communication solution using commercial RFID devices,” *Information Sciences*, vol. 608, pp. 28–46, 2022.
 - [23] W. Qiang and L. Zhang, “Auditory sparse representation for robust speaker recognition based on tensor structure,” *EURASIP Journal on Audio, Speech and Music processing*, vol. 2008, article 578612, pp. 1–9, 2008.
 - [24] R. De Mori, F. Bechet, D. Hakkani-Tur, M. McTear, G. Riccardi, and G. Tur, “Spoken language understanding,” *IEEE Signal Processing Magazine*, vol. 25, no. 3, pp. 50–58, 2008.
 - [25] T. Hasan and J. H. L. Hansen, “Acoustic factor analysis for robust speaker verification,” *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 21, no. 4, pp. 842–853, 2013.
 - [26] S. Wang, A. Wang, M. Ran et al., “Hand gesture recognition framework using a lie group based spatio-temporal recurrent network with multiple hand-worn motion sensors,” *Information Sciences*, vol. 606, pp. 722–741, 2022.
 - [27] M. Lin, Q. Zhan, and Z. Xu, “Decision making with probabilistic hesitant fuzzy information based on multiplicative consistency,” *International Journal of Intelligence Systems*, vol. 35, no. 8, pp. 1233–1261, 2020.
 - [28] M. Lin, Z. Chen, R. Chen, and H. Fujita, “Evaluation of startup companies using multicriteria decision making based on hesitant fuzzy linguistic information envelopment analysis models,” *International Journal of Intelligence Systems*, vol. 36, no. 5, pp. 2292–2322, 2021.
 - [29] M. Lin, X. Li, R. Chen, H. Fujita, and J. Lin, “Picture fuzzy interactional partitioned Heronian mean aggregation operators: an application to MADM process,” *Artificial Intelligence Review*, vol. 55, no. 2, pp. 1171–1208, 2022.
 - [30] S. A. Parah, J. A. Kaw, P. Bellavista et al., “Efficient security and authentication for edge-based Internet of Medical Things,” *IEEE Internet of Things Journal*, vol. 8, no. 21, pp. 15652–15662, 2021.
 - [31] Q. Hu, F. N. B. Gois, R. Costa et al., “Explainable artificial intelligence-based edge fuzzy images for COVID-19 detection and identification,” *Applied Soft Computing*, vol. 123, p. 108966, 2022.
 - [32] S. Pirbhulal, W. Wu, K. Muhammad, I. Mehmood, G. Li, and V. H. C. de Albuquerque, “Mobility enabled security for optimizing IoT based intelligent applications,” *IEEE Network*, vol. 34, no. 2, pp. 72–77, 2020.

- [33] C. Huang, S. Chen, Y. Zhang, W. Zhou, J. J. P. C. Rodrigues, and V. H. C. de Albuquerque, "A robust approach for privacy data protection: IoT security assurance using generative adversarial imitation learning," *IEEE Internet of Things Journal*, 2021.
- [34] W. Wu, H. Zhang, V. H. C. de Albuquerque, and L. Xu, "Hyper-noise interference privacy protection framework for intelligent medical data-centric networks," *IEEE Network*, vol. 35, no. 1, pp. 333–339, 2021.
- [35] A. F. Hussein, N. ArunKumar, G. Ramirez-Gonzalez, E. Abdulhay, J. M. R. S. Tavares, and V. H. C. de Albuquerque, "A medical records managing and securing blockchain based system supported by a Genetic Algorithm and Discrete Wavelet Transform," *Cognitive Systems Research*, vol. 52, pp. 1–11, 2018.
- [36] F. Alenezi, Ş. Öztürk, A. Armghan, and K. Polat, "An effective hashing method using W-shaped contrastive loss for imbalanced datasets," *Expert Systems with Applications*, vol. 204, article 117612, 2022.
- [37] F. Angeletti, I. Chatzigiannakis, and A. Vitaletti, "The role of blockchain and IoT in recruiting participants for digital clinical trials," in *2017 25th international conference on software, telecommunications and computer networks (SoftCOM)*, pp. 1–5, Split, Croatia, 2017.
- [38] O. Bodenreider, D. Nguyen, P. Chiang et al., "The NLM value set authority center," *Studies in Health Technology and Informatics*, vol. 192, p. 1224, 2013.
- [39] S. D. Norberhuis, *MultiChain: a cybercurrency for cooperation*, [M.S. thesis], TU Delft, DelftUniversity of Technology, 2015.
- [40] Y.-C. Chang and M. Mitzenmacher, "Privacy preserving keyword searches on remote encrypted data," in *International Conference on Applied Cryptography and Network Security*, pp. 442–455, Berlin, Heidelberg, 2005.
- [41] D. O'Shaughnessy, "Automatic speech recognition: History, methods and challenges," *Pattern Recognition*, vol. 41, no. 10, pp. 2965–2979, 2008.
- [42] A. Upadhyay, S. K. Sharma, and A. Upadhyay, "Speaker identification and verification using different model for text dependent," *International Journal of Applied Engineering Research*, vol. 12, no. 8, pp. 1633–1638, 2017.
- [43] D. A. Reed, D. B. Gannon, and J. R. Larus, "Imagining the future: thoughts on computing," *Computer*, vol. 1, pp. 25–30, 2012.
- [44] S. Upadhyay, S. K. Sharma, P. Kumar, and A. Upadhyay, "Comparative study and analysis of different feature extraction techniques for speaker verifications and identification: a review," *INROADS- An International Journal of Jaipur National University*, vol. 4, no. 2, pp. 101–106, 2015.
- [45] N. Dehak, P. Kenny, R. Dehak, P. Dumouhcel, and P. Ouellet, "Front-end factor analysis for speech verification," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 19, no. 4, pp. 788–798, 2011.
- [46] B. Zhou and J. H. L. Hasen, "Rapid discriminative acoustic model based on eigenspace mapping for fast speaker adaptation," *IEEE Transactions on Speech and Audio Processing*, vol. 13, no. 4, pp. 554–564, 2005.
- [47] S. O. Sadjadi and J. H. L. Hansen, "Blind spectral weighting for robust speaker identification under reverberation mismatch," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 22, no. 5, pp. 937–945, 2014.
- [48] T. H. Falk and W. Y. Chan, "Modulation spectral features for robust far-field speaker identification," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 18, no. 1, pp. 90–100, 2010.
- [49] M. Kumar, P. Mukherjee, K. Verma, S. Verma, and D. B. Rawat, "Improved deep convolutional neural network based malicious node detection and energy-efficient data transmission in wireless sensor networks," *IEEE Transactions on Network Science and Engineering*, 2021.
- [50] S. Upadhyay, M. Kumar, A. Kumar, K. Z. Ghafoor, and S. Manoharan, "SmHeSol (IoT-BC): smart healthcare solution for future development using speech feature extraction integration approach with IoT and blockchain," *Journal of Sensor*, vol. 2022, pp. 1–13, 2022.
- [51] A. Kumar, "A cloud-based buyer-seller watermarking protocol (CB-BSWP) using semi-trusted third party for copy deterrence and privacy preserving," *Multimedia Tools and Applications*, vol. 81, no. 15, pp. 21417–21448, 2022.
- [52] A. Kumar, "Design of secure image fusion technique using cloud for privacy-preserving and copyright protection," *International Journal of Cloud Applications and Computing (IJCAC)*, vol. 9, no. 3, pp. 22–36, 2019.
- [53] A. Kumar, "A review on implementation of digital image watermarking techniques using LSB and DWT," in *Information and Communication Technology for Sustainable Development*, M. Tuba, S. Akashe, and A. Joshi, Eds., vol. 933 of *Advances in Intelligent Systems and Computing*, pp. 595–602, Springer, Singapore, 2020.