Cloud Based Health Care Data Privacy System Using Hybrid Techniques

Asha G M, Jain Institute of Technology, Davangere
Shridharamurthy S.K, UBDT College of Engineering, Davangere
B M Nanda Kumar, Department of EEE, Jain Institute of Technology, Davangere
Shashank C M, Department of EEE, Jain Institute of Technology, Davangere
Veeresh B S, Department of EEE, Jain Institute of Technology, Davangere
Chandana S N, Department of EEE, Jain Institute of Technology, Davangere

Abstract- *Computed tomography and chest radiography play an important part in crucial medical imaging and are used to examine biological disorders such as lung cancer and TB. In general, CT pictures are complex in nature, making it difficult to get good volumetric information. However, the absence of structural information and poor boundary information make CT images less suitable for medical measures. However, with the increasing spread of the Coronavirus illness COVID, there is a need for the most appropriate method for diagnosing COVID pneumonia, since it is expected to have a significant influence on the healthcare system. By exploring the abnormalities of lungs (COVID person) using CT images CAD system can be built to carry out diagnostic measurements. To handling the high dimensional feature from complex CT images Deep learning model using Convolutional Neural Network (GDCNN) is proposed for auto segmenting the covid affected regions from normal LUNG regions. To overcome the intra class variance problem over affected region and correlation between normal/abnormal regions hierarchical feature set modelling is proposed for deep learning which includes (i) edge attention module (ii) image gradient driven boundary exploration (iii) finally contour driven covid affected lung region segmentation. After ROI segmentation multi modal feature sets are extracted using GLCM texture and shape attributes. CNN network is used for final classification which can narrow down the false rate and metrics during classification. Finally automated deep learning-based CAD system is performed COVID abnormal classification, and classification accuracy can be evaluated based on open-source bench mark datasets which contains more than 5000 CXR image samples for classifying pneumonia, normal and other pneumonia diseases. After identifications of COVID diseases in Lungs images, the patient data (Name, age. country. place of birth and others) are alone encrypted to secure the patient information instead of encryption of image. The encrypted data is stored in cloud for more secure and diagnosed image is transmitted directly to receiver. The receiver can receive both encrypted data and image, but encrypted data from cloud and image directly from sender for decryption to recovery the original image and patient data.*

Keywords: Deep learning, Convolutional Neural Network, GLCM, lung infection, edge attention, image gradient, Specificity, Sensitivity, Accuracy.

1. INTRODUCTION

By using social data and crime data that are gathered by social network and police information systems, it is presenting a novel analytical approach of criminal suspects to solve the problem with data privacy security issues. We make it possible for the social cloud server and public security cloud server to securely exchange user public information and social information about criminal suspects. To ensure that only authorized parties can conduct searches on suspects' social data and that the social cloud server cannot infer anything during the search, we specifically propose a privacy-preserving data retrieving mechanism based on oblivious transfer. In addition, several building blocks are also proposed, including secure classification and regression tree (CART) models and encrypted data comparison. Based on these fundamental components, we created a criminal suspect sensing system that protects privacy. Finally, we present a performance evaluation that demonstrates how our plan might improve analysis of criminal suspects without compromising privacy.

2. LITERATURE SURVEY

The approach uses a three-part privacy-preserving framework with differential and distortion to lessen the loss of data quality caused by location obfuscation. To adapt the original sensing data to the obscured location, we must first train a data adjustment function. The second step is to use a linear programme to choose the best location obfuscation function. In order to comply with the restrictions of differential privacy, distortion privacy, and evenly distributed obfuscations, the linear programme seeks to reduce the uncertainty in data adjustment. In order to lower the amount of computer resources needed, it also proposes an approximation method. Third, in order to increase the inference accuracy for the obfuscated data, we present an uncertainty-aware inference algorithm. Compared to

state-of-the-art methods with the same level of privacy protection, evaluations with real environment and traffic datasets show that our optimal method reduces data quality loss by up to 42%. The approximated method suffers from a 3% additional quality loss than the optimal method but only requires 1% of the computation time.

The growth of big data in healthcare is crucial for constructing intelligent healthcare and encourages the transformation of hospital administration into one that is digital, intelligent, and scientific. The use of regional platforms for intelligent health care services will have more room to grow as cloud computing technology develops, and the future of intelligent health care cloud will be even more promising. The use of big data in the healthcare industry does, however, also come with benefits and drawbacks. The core of health care is administration and science with a focus on people. In order to assess the risk of patient medical information leakage from three perspectives—resource sensitivity, access behaviour sensitivity, and historical access information—patient-centered research on information security under the health care big data environment has been established. It is challenging to guarantee the correctness and dependability of risk analysis because there is a dearth of complete data during the real problem-solving process. Because of this, modelling is done from the ambiguity of risk factors, using fuzzy rule technology to solve the uncertainty in practical application, improving the performance of risk assessment methods, so hospital managers can do their best in making decisions and providing prediction guidance for care work, continuously improve and the intelligent health care management model, further realise refined health care services, and promote the overall progress of h Although the security is not up to par, fuzzy based encryption and decryption processes need less computer operations.

3. PROBLEM STATEMENT

According to a survey of recent articles on data privacy and security levels published in IEEE Transactions and Springer publications, the transmission of sensitive police and health care data over wireless or wired communications with the aid of encryption and decryption processes is a major concern. The current work-related data privacy system has a 90% maximum accuracy and employs the homomorphic system to generate the cypher output. With today's communication technology, the computational operations necessary to produce the cypher output for increased privacy and security are unacceptable. Another difficult challenge for the encryption and decryption processes in the current system is key generation. Because third parties can readily attack these keys between sender and receiver, the degrees of data privacy and security are compromised.

4. METHODOLOGY

The rapid development of today's more significant information technology has made it difficult to share medical information or other sensitive information amongst individuals in the military, police, or other government agencies. The military's sharing of information with the health care sector about border areas and main issues in this study endeavour. When it comes to the administration of smart health care, information security concerns are becoming more and more prevalent, and the most significant one is the problem of patient privacy leakage.

As a result, improving information management for intelligent health care in the big data age is crucial to the long-term sustainability of hospitals. The key generation method for encryption and description at the point of data transmission and after receiving it at the receiver presents the first difficulty in the current workIn this research, the key factors influencing the privacy disclosure of big data in health management were first identified. Next, a risk access control model was established, which calls for storing health care information data in a secured cloud after encryption using keys generated using fuzzy theory and Elliptical Curve Cryptography (ECC), which is used for the management of big data in intelligent medical treatment. The findings computed by the fuzzy tool and ECC system created in the Matlab tool are compared to the suggested model at the end. The security level may be measured in terms of throughput, latency, and packet delivery ratio (PDR) based on encrypted and decrypted data. From the results, it is possible to forecast an accuracy that is more than 95%, as opposed to the 90% accuracy of the previous study.

Proposed flow diagram of high throughput and low latency data privacy:

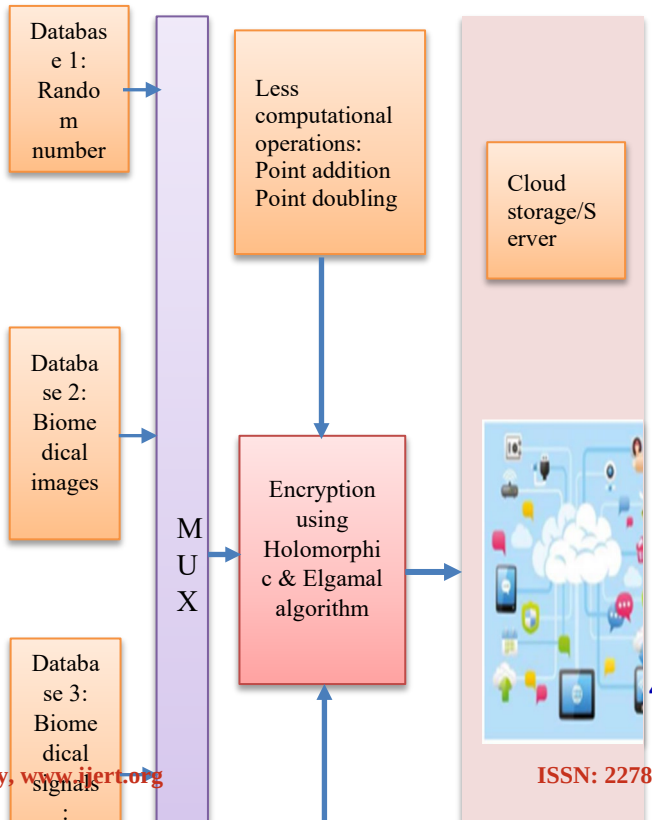

Fig .1 Proposed hybrid cloud-based encryption system for high throughput and low latency

This research's primary contribution to efficient key generation depends on fuzzy membership sets and ECC.

1. Key generation by ECC and three key condition attributes are extracted by using the attribute reduction and discernibility matrix in the rough set theory based on consideration of the risk factors affecting on the privacy of the health care big data leakage: access behaviour sensitivity, resource sensitivity, and historical access risk.
2. Three important indexes are fuzzy treated, and the membership function between each index and the associated fuzzy set is constructed. This leads to the establishment of the risk access control model based on fuzzy theory.
3. For high level security, holomorphic and Elgamal algorithms have been used in the encryption and decryption processes based on the fuzzy sets that serve as keys.
4. The ECC-based Elgamal method has the benefit of employing just two computations, point addition and point doubling, as opposed to more complicated mathematical operations like multiplications and modulo operations.
5. In order to strengthen security, each huge data collection related to health care is mapped to fuzzy membership sets and then encrypted once more using Elgamal techniques.
6. The fuzzy set is defuzzified using the central technique and inverse Elgamal process to acquire the final decrypted data that are obtained from the cloud storage in order to achieve the final risk value.
7. Finally, a comparison study using the Elgamal and fuzzy logic tool set in Matlab is performed to ensure that the model in this work is efficient and extremely accurate.

When data is sent via the RTP-RTCP protocol, the suggested hybrid fuzzy, holomorphic, and Elgamal algorithms are examined under various noisy situations on various databases. The SVM-based machine learning method evaluates the system performance, including whether MSE and PRD are optimised or not, for each value of MSE and Packet Delivery Ratio (PRD). The choice of an MSE reduction strategy is solely based on the supplied threshold boundary values. The performance level of the data privacy system is improved by employing methods like Fuzzy or ECC with SVM classifiers, and the complexity trade-off measurement is also decreased, as shown in Fig. 1. The SVM approach is used to analyse the MSE reduction, the system's assigned threshold boundary, each boundary value, and its step-by-step methodology. From steps 1 through 3, boundary values are chosen, and the network is trained using the predetermined MSE and PRD values. The multi-rate SVM evaluates and chooses either fuzzy or ECC approaches for further reducing MSE and PRD based on the PRD and MSE values.

Machine Learning algorithm:

Inputs: Training_Set, Group_Train, Test_Setl

Output: Trained data for reduction in MSE

Assignment 1: MSE_bound=[6,7,8,9];
Training_Set=[MSE_bound];

Assignment 2: Group_Train=[1;2;3;4];

Assignment 3: TestSetl=papro;

Where papro(i)=10*log10(peako/meano);

```
if papro < 6
disp('conventional mode')for no
further reduction techniques are
required in MSE
end
```

where peako=max(abs(time_domain_signal).^2);

where
time_domain_signal=abs(ifft([S_P_DATA(i,1:32)
zeros(1,(L-1)*Nt) S_P_DATA(i,33:64)]));

Step:1 Find the uniqueness values in Group_Train data

Calculate the length on the Step:1

Step:2 if (Group_Train) has unique values the

Assign Test_Setl= Group_Train;

Step:3 Create empty matrix to store the trained data of Traning_set

For loop from k=1 to length of Group_Train

If(G1vAll =Group_Train=length(k)) then

Train for models(k)=svmtrain(TrainingSet(:,1:4),G1vAll,'kernel_function','rbf');

It is to check current class is 1 or 0

Step:4 classify and check weather Fuzzy is selected or ECC is selected for further reduction in MSE & PRD based on the value of present MSE & PRD

Else

Return(j)=k;

end

The step-by-step process of COVID detection:

Supervised - high unique train set images are required. Finite details from train sets converts into template or reference classification.
Limitations:
Large train sets are required.
High variation in infection characteristics of covid.
Train sets should be roost appropriate one.

Unsupervised- multi modal feature sets can be used to accommodate all nonlinear variations of covid affected lung regions.
Image Statistics extracted from lungs patch some attributes are generated. Accordingly, classification output is obtained.

CNN DEEP LEARNING:
Convolution layer -texture and shape attributes for each Input ROI image patch (sliding window) were convoluted layer by layer.
Rectified linear (ReLu) function: Non-linear mapping: At last convolution layer, the output turned
into one threshold bound- mapping — Each ROI patch generates one output values.

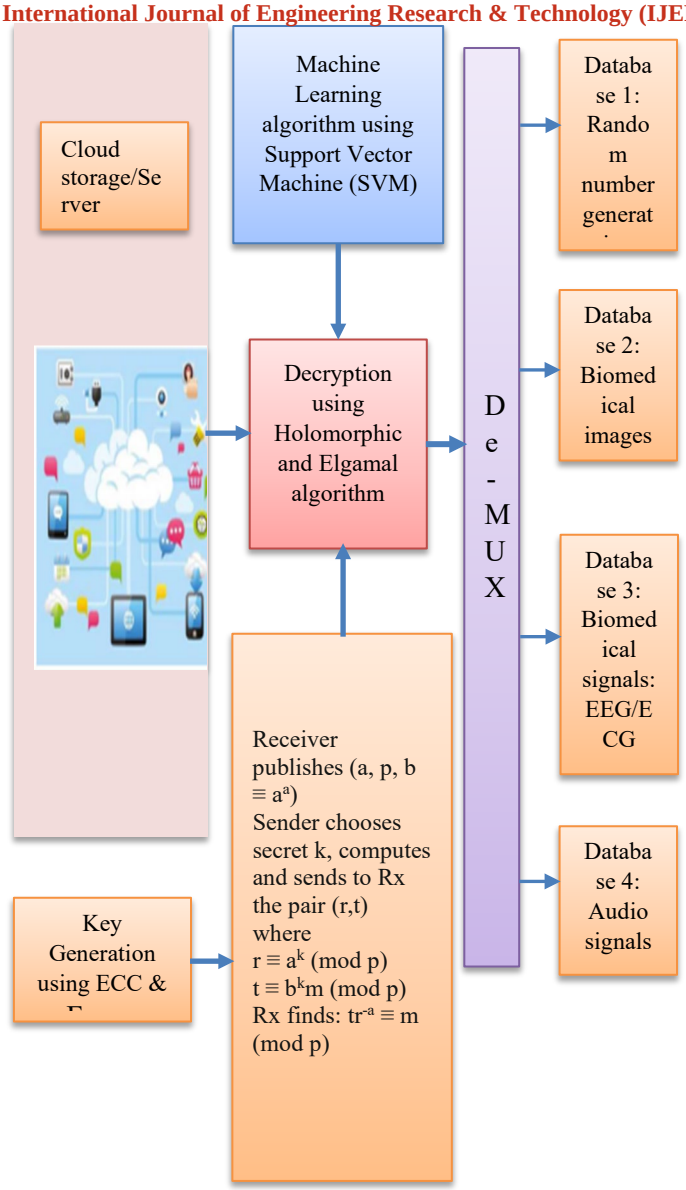

Fig.2 Proposed hybrid cloud-based decryption system for recovery of data from cloud server

Conclusion

It is possible to forecast an accuracy that is more than 95%, as opposed to the 90% accuracy of the previous study.

5 Expected results

After completion of the design as per proposed design flow shown in Fig, I and Fig.2. the following parameters are measured and compared with existing work which is shown in base paper

i. Accurcay
ii. Sensitivity

iii. Specifity
iv. Throughput
v. Energy consumption before uploading the data into cloud and after receiving from the cloud
vi. Reconstruction rate
vii. Losses in the data during the process
viii. Time taken for completion of process

References

[1]. Jian Xu.et.al, "SPCSS: Social Network Based Privacy-Preserving Criminal Suspects Sensing", IEEE TRANSACTIONS ON COMPUTATIONAL SOCIAL SYSTEMS, DOI: 10.1109/TCSS.2019.2960857, 2329-924X, 2020 IEEE.

[2]. Leye Wang.et.al, "Sparse Mobile Crowd sensing With Differential and Distortion Location Privacy", IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, VOL. 15, 2020, DoI: 10.1109/TIFS.2020.2975925, 1556-6013, 2020 IEEE

[3]. Mingyue Shi.et.al. "A privacy protection method for health care big data management based on risk access control", Health Care Management Science, https://doi.org/10.1007/s10729-019-09490-4, Springer, 2019.