Original Article

Osteoporosis Risk Prediction Using Enhanced Support Vector Machine over Artificial Neural Network

Jagadeesh A[1], Senthil kumar S[2]

[1]Research Scholar, Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, Tamilnadu. India, Pincode: 602105.
[2]Project Guide, Corresponding Author, Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, Tamilnadu. India, Pincode: 602105.

Abstract

Aim: The aim of this study is to predict the osteoporosis risk by using the proposed Support Vector Machine(SVM) Algorithm over Artificial Neural Network(ANN) Algorithm.
Materials and Methods: Sample groups that are considered in this project is CT Scan dataset that can be classified into two, one for training data and other for testing data, Dataset are tested using 233.9s for G-power to determine the sample size and for train set analysis.Nearly 215 CT Scan images have been used in each group for testing of cancer.
Results: The Enhanced Support Vector Machine algorithm has better efficiency with 83% accuracy when compared to Artificial Neural Network algorithm's 71%. Statistical significance difference (two-sided) is 0.01 ($p<0.01$).
Conclusion: Support Vector machine algorithm performed significantly better than the Artificial Neural network algorithm.

Keywords: Osteoporosis risk prediction, Enhanced Support Vector Machine, Artificial Neural Network, Bone Cancer, CT Scan.

DOI: 10.47750/pnr.2022.13.S04.191

INTRODUCTION

The aim of this research is to predict Osteoporosis using CT Scan data using Support Vector Machine Algorithm and to compare proposed algorithm with Artificial Neural Network Algorithm. In this case the research aims to improve the rate of efficiency in predicting bone cancer in early times. (Pandiangan, Bali, and Silalahi 2019)Osteoporosis is a disease characterized by accelerated bone turnover and decreased bone mass, as well as skeletal fragility, which increases the risk of fracture. (Heymann 2009)It typically goes unnoticed until it's too late, when fragility fractures have already developed.It occurs when the body loses minerals like calcium faster than the bones can restore them, resulting in a loss of bone thickness. (Mechria, Gouider, and Hassine 2019; Rajagopalan and Babu, n.d.) The disease is malicious and malignant, and it is caused by uncontrolled cell division in the bone.Medical image analysis is one of the most rapidly expanding and inventive fields for improving the accuracy of disease diagnosis and providing improved patient care. (Nikolova et al. 2021)Fractures, arthritis, bone tumors, and osteoporosis are all frequent bone illnesses all over the world. A bone tumor, also known as a neoplasm, is a development of aberrant tissue that is structurally distinct from the surrounding tissue.(Beil et al. 2022) Tumors can grow in a variety of human bone structures, including the fingers, palms, wrists, forearms, and shoulders.(Amini 2012)To predict osteoporosis risk, machine learning models were built utilizing approaches such as the k-nearest neighbours (KNN), decision tree (DT), random forest (RF), gradient boosting machine (GBM), support vector machine (SVM), artificial neural networks (ANN), and logistic regression (LR).(Mechria, Gouider, and Hassine 2019)The main goal of this paper is to extract bone X-Ray or CT scan pictures in order to predict osteoporosis.
There are around 56 IEEE papers and 121 google scholar papers have been published over the past 5 years. The most cited article is "Cancer Gene Detection using artificial neural network".("CANCER GENE DETECTION USING ARTIFICIAL NEURAL NETWORK" 2017) When the body's cells go rogue, cancer develops. Almost each cell in the body has the potential to turn into a tumor and spread to other parts of the body. Essential bone malignancy starts in the cells of the bones. Tumor cells are bone cells that have shown to be harmful(Park et al. 2021; Tang, Han, and Yin 2022) The information in this segment is all about the importance of bone growth. The vast majority of persons who have tumor cells in their bones do not build their bones properly. They have cancer

cells in their bones that have spread from a tumor somewhere else in their bodies.(Park et al. 2021) This disorder is known as auxiliary or metastatic bone disease.Finally, we discuss future research to improve accuracy using the ANN model. In my opinion the above most cited article is best compared to the other articles.
Our institution is passionate about high quality evidence based research and has excelled in various fields (Devarajan et al., 2021; Dhanraj & Rajeshkumar, 2021; Kamath et al., 2020; Nandhini et al., 2020; Parakh et al., 2020; Perumal et al., 2021; Pham et al., 2021; Sathiyamoorthi et al., 2021; Tesfaye Jule et al., 2021; Uganya et al., 2021). In the existing research they didn't identify efficient accuracy for calculating time. The drawbacks of the above cited paper are to create the bone tumor detection system with superpixel segmentation. The offered set of photos can also be used to identify brain cancer. The proposed technology is designed specifically for the identification of bone tumors.(Gühne et al. 2021) The similar approach can be used to determine cancer stagesThe main aim of our project is to predict osteoporosis by using Artificial Neural Network over Support Vector Machine Algorithm.

MATERIALS AND METHODS

The research work was performed in the Image Processing Laboratory, Department of Computer Science and Engineering, Saveetha School of Engineering, SIMATS. The proposed work contains two groups. Group 1 is taken as Enhanced Support Vector Machine and Group 2 is taken as Artificial Neural Network. The Enhanced Support Vector Machine(SVM) algorithm and Artificial Neural Network(ANN) algorithm were evaluated a different number of times with a sample size of 10 with confidence interval of 95%, and with pretest power of 81% and maximum accepted error is fixed as 0.05.
After dataset collection,the null values and unimportant content in the datasets were removed by preprocessing and data cleaning steps.After cleaning and preprocessing the data,an ideal input for the detection model is produced, the input images depends on the mean pixel density value and image id. Which are processed into the detection model using opencv library and efficiency of both Support Vector Machine over Artificial Neural Network algorithm is calculated.
Testing setup for this proposed system used a Kaggle notebook and pycharm.Kaggle notebook is a software which is used for creating the Osteoporosis Prediction with SVM model and ANN model. Hardware configuration for this proposed system is Intel core i5 8th gen processor and requires 4GB random access memory and 256GB Solid state drive used. The configuration of the system is windows 10 operating system and kaggle notebook software and python programming language 3.8.3.

Testing Procedure for Osteoporosis Prediction with SVM and ANN

Step 1: Preprocessing In most circumstances, noise affects almost all captured photos, resulting in low quality. Preprocessing processes such as filtering, histogram equalization, and so on must be performed to remove the relevant portion of the images from the photos without noise and blurriness.In Fig.3. represents the cancer analysis graph of CT Scan Dataset in that 987 images contain cancer and 1569 images do not contain cancer.The photos are preprocessed using Python software. In Fig.4. represents the block diagram of the testing procedure to detect cancer. Preprocessing images has the primary objective of removing the excesses that can be detected in scanned images. To decrease noise and increase quality, each image is preprocessed.Fig.5.represents the normal and preprocessed image.

Step 2: Training and Testing data

The CTscan dataset is used as input to the classifier in order to create indices during the training phase. We can load various and more datasets into the test classifier to test the classifier's accuracy. The method of testing is done to determine whether or not the bone has malignancy. In this stage, the classifier displays the cancer outcome in images.

Step 3: Feature Selection and Extraction

This technique chooses a small selection of useful qualities to be used in the future. Genetic algorithms are used to choose features from the preprocessed image after it has been preprocessed.From a large amount of data, feature extraction can properly forecast the amount of resources required. Features must be extracted after the attributes have been chosen. It serves an important purpose by employing algorithms and methodologies to identify the various portions and qualities that must be deleted.

Step 4: Predicting Cancer using machine learning algorithms

Using the Bone Cancer Detection with Enhanced Residual Network algorithms and Traditional CNN to predict the images and identify the cancer based on the mean pixel density. Fig.6.describes the dataset of the CTscan images.

Support Vector Machine Algorithm

The Support Vector Machine, or SVM, is a linear model that can be used to solve classification and regression issues. It can solve both linear and nonlinear problems and is useful for a wide range of applications. SVM is a basic concept, The method divides the data into classes by drawing a line or hyperplane.

Inputs: CT scan data set
Output: Selected features and Accuracy.

```
Get CTScan()
CTS_slices=pd.read_csv("cnn1.csv")
read_Img=CTS_slices
Img<-exp(Img)
for Img i to n
                    img_ids = Img.str.split('.').str[0]
                    assert df_centers.img_id.equals(img_ids)
                    df_train = pd.DataFrame(mat_images, columns=['pxl' + str(i) for i in
                                                        range(img_ids)])
                    df_train = pd.concat([df_train, df_train_hflip, df_train_vflip],
                                                    ignore_index=True)
                    X = df_train.drop(columns=['img_id', 'cx', 'cy']).values.reshape((-1,
          IMG_HEIGHT, IMG_WIDTH, IMG_CHANNELS))
                    Y = df_train[['cx', 'cy']].values
X_train,X_test,y_train,y_test<-split features set and labels into train subset and test
                                                                    subset
history = model.fit(X_train,y_test<-split features set and labels into train subset and
                                                                    test subset
V<-SVM(X_train,y_train)
score<-evaluate(i,y_test,v)
return score
```

Artificial Neural Network Algorithm

Artificial neural networks (ANNs), also known as neural networks (NNs), are computer systems that are modeled after the biological neural networks that make up animal brains. Artificial neurons are a set of connected units or nodes in an ANN that loosely replicate the neurons in a biological brain. ANN are algorithms that are based on brain function and can be used to model complicated patterns and predict problems.

Inputs: CT scan dataset
Output: Selected features and Accuracy.

```
Get CTScan()
CTS_slices=pd.read_csv("cnn1.csv")
read_Img=CTS_slices
Img<-exp(Img)
for Img i to n
                    img_ids = Img.str.split('.').str[0]
                    assert df_centers.img_id.equals(img_ids)
                    df_train = pd.DataFrame(mat_images, columns=['pxl' + str(i) for i in
                                                        range(img_ids)])
                    df_train = pd.concat([df_train, df_train_hflip, df_train_vflip],
                                                    ignore_index=True)
                    X = df_train.drop(columns=['img_id', 'cx', 'cy']).values.reshape((-1,
          IMG_HEIGHT, IMG_WIDTH, IMG_CHANNELS))
                    Y = df_train[['cx', 'cy']].values
X_train,X_test,y_train,y_test<-split features set and labels into train subset and test
                                                                    subset
history = model.fit(X_train,y_test<-split features set and labels into train subset and
                                                                    test subset
V<-ANN(X_train,y_train)
score<-evaluate(i,y_test,v)
return score
```

Dataset is collected from CT Scan Dataset. In the dataset 70% is used for training and 30% used for testing. The Enhanced SVM and ANN algorithms were evaluated with respect to training, and tests were conducted with the required parameters to improve the accuracy percentage.

Statistical Analysis

Statistical software used in the study is IBM SPSS version 26. The independent sample T-test calculation for analyzing equal variance, standard error, and levene's test are evaluated. Attributes like image id, mean density value, pixel density, detection and class are dependent variables. Independent sample T-test has been carried out for evaluating the accuracy.

RESULTS

In this proposed system it was identified that the Support Vector Machine Algorithm appears to have better accuracy than the Artificial Neural Network algorithm. Table 1 represents the attributes of CTscan Dataset. In Fig.1. represents the architecture for Osteoporosis prediction estimation for each CT Scan image using Enhanced Support Vector Machine algorithm.Table 2 shows the sample accuracy of SVM and ANN algorithms.SVM uses optimization procedures to deliver high accuracy. Table 3 shows the statistical calculation such as mean, standard deviation and standard error mean for SVM and ANN algorithms. It is inferred that the deviation for T-test is far lesser than the comparison algorithm. Moreover, the accuracy value of SVM is around 83.3. while the loss is around 15.40, which seems to be superior to the ANN classifier. In Table 4, it was observed that the Levens test for equality of variance and its significance for SVM is 0.033 and 0.857, respectively and standard error difference and confidence interval are lower than ANN classifiers. Mean accuracy and mean loss graph is depicted in Fig. 2. SVM seems to appear better for the given CT Scan dataset of Osteoporosis risk prediction.

DISCUSSION

The proposed system provides better osteoporosis risk prediction using a support vector machine with a count vectorizer with over 83.3% accuracy compared to artificial neural network algorithms.

There are similar papers on Osteoporosis risk prediction using machine learning algorithms.In this research paper uses a screening and identifying method to predict cancer.Cancer is a deadly disease that affects people of all ages.(Jabarpour, Abedini, and Keshtkar 2020) More than one in three people will experience cancer at some point in their life. By evaluating diagnostic medical techniques such as X-rays, CT scans, and positron emission tomography(PET) scans, the general purpose is to identify the affected area in the bone tract, i.e., the abnormal growth and phase of the disease. Since the scanned images may not have a high resolution due to the large number of layers per pixel and noise, (Bumrungkun, Chamnongthai, and Patchoo 2018)it is necessary to pre-process the images with a medium filter to remove the noise. Certain features of the preprocessed image are evaluated using a genetic approach and retrieved by CNN. The recovered images are classified and recorded with a CNN classifier to determine the stage of the disease, which helps the doctor make treatment recommendations.(Bach-Mortensen 2002) The results of the proposed method show a higher incidence of early prediction of bone cancer.

In the fields of medical imaging diagnosis, illness prediction, and bone cancer diagnosis, as well as risk assessment, the machine learning technique has proven to be effective. We conclude in this research that there are numerous scientific difficulties that must be addressed.(Sharma et al. 2021) Computerized health care systems for diagnosing early-stage bone cancer, for example, have been demonstrated to be quite effective, particularly in countries like India, where the mortality rate is high and the doctor-patient ratio is low. As a result of this scenario, medical imaging is used to diagnose cancer.(Garnero 2000)

CONCLUSION

The Enhanced Support Vector Machine algorithm detects cancer with better accuracy of 83% compared to Artificial Neural Network Algorithm with 71%.

Declarations

Conflict of Interests

No conflict of interest in this manuscript.

Author Contribution

Author Jagadeesh Atthipatla was involved in data collection, data analysis, and manuscript writing. Author RSK was involved in conceptualization, guidance and critical review of manuscript.

Acknowledgments

The authors would like to express their gratitude towards Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences (Formerly known as Saveetha University) for providing the necessary infrastructure to carry out this work successfully.

Funding

We thank the following organizations for providing financial support that enabled us to complete the study.

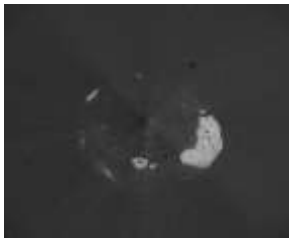
1. STEP UP Technologies, Chennai
2. Saveetha University
3. Saveetha Institute of Medical and Technical Sciences.
4. Saveetha School of Engineering.

REFERENCES

1. Amini, Maryamsadat. 2012. Thermographic Image Analysis Method in Detection of Canine Bone Cancer (osteosarcoma).
2. Bach-Mortensen, Pernille. 2002. Diagnosis of Osteoporosis and Prediction of Fracture Risk Using a New Bone-Mass Estimating Technique: Digital X-Ray Radiogrammetry.
3. Beil, Frank Timo, Julian Stürznickel, Tim Rolvien, Michael Amling, and Ralf Oheim. 2022. "[Tumor localization and treatment of tumor-induced osteomalacia]." Zeitschrift fur Rheumatologie, February. https://doi.org/10.1007/s00393-022-01160-1.
4. Bumrungkun, Prachya, Kosin Chamnongthai, and Wisarn Patchoo. 2018. "Detection Skin Cancer Using SVM and Snake Model." 2018 International Workshop on Advanced Image Technology (IWAIT). https://doi.org/10.1109/iwait.2018.8369708.
5. "CANCER GENE DETECTION USING ARTIFICIAL NEURAL NETWORK." 2017. International Journal of Recent Trends in Engineering and Research. https://doi.org/10.23883/ijrter.2017.3384.anwnb.
6. Garnero, P. 2000. "Markers of Bone Turnover for the Prediction of Fracture Risk." Osteoporosis International. https://doi.org/10.1007/s001980070006.
7. Gühne, Falk, Stefanie Radke, Thomas Winkens, Christian Kühnel, Julia Greiser, Philipp Seifert, Robert Drescher, and Martin Freesmeyer. 2021. "Differences in Distribution and Detection Rate of the [Ga]Ga-PSMA Ligands PSMA-617, -I&T and -11-Inter-Individual Comparison in Patients with Biochemical Relapse of Prostate Cancer." Pharmaceuticals 15 (1). https://doi.org/10.3390/ph15010009.
8. Heymann, Dominique. 2009. Bone Cancer: Progression and Therapeutic Approaches. Academic Press.
9. Jabarpour, Efat, Amin Abedini, and Abbasali Keshtkar. 2020. "Osteoporosis Risk Prediction Using Data Mining Algorithms." Journal of Community Health Research. https://doi.org/10.18502/jchr.v9i2.3401.
10. Mechria, Hana, Mohamed Gouider, and Khaled Hassine. 2019. "Breast Cancer Detection Using Deep Convolutional Neural Network." Proceedings of the 11th International Conference on Agents and Artificial Intelligence. https://doi.org/10.5220/0007386206550660.
11. Nikolova, Petya N., Valeria H. Hadzhiyska, Kiril B. Mladenov, Mihaela G. Ilcheva, Stefani Veneva, Svetla E. Dineva, and Boris S. Mladenov. 2021. "Detection of Ureteral Stump Transitional Cell Carcinoma, Presenting as Bone Metastases from Unknown Primary by 18F-FDG PET/CT: A Case Report with Review of Literature." The Indian Journal of Radiology & Imaging 31 (4): 1065–69.
12. Pandiangan, T., I. Bali, and A. R. J. Silalahi. 2019. "Early Lung Cancer Detection Using Artificial Neural Network." Atom Indonesia. https://doi.org/10.17146/aij.2019.860.
13. Park, Ki Bum, Chul Hyo Jeon, Han Hong Lee, Hyungmin Chin, and Kyo Young Song. 2021. "Prediction of Risk of Osteoporosis after Gastrectomy for Gastric Cancer." BJS Open 5 (6). https://doi.org/10.1093/bjsopen/zrab123.
14. Rajagopalan, Kishore, and Suresh Babu. n.d. "The Detection of Lung Cancer Using MTANN (Massive Training Artificial Neural Network) Based Soft Tissue Technique." https://doi.org/10.21203/rs.3.rs-25307/v1.
15. Sharma, Ashish, Dhirendra P. Yadav, Hitendra Garg, Mukesh Kumar, Bhisham Sharma, and Deepika Koundal. 2021. "Bone Cancer Detection Using Feature Extraction Based Machine Learning Model." Computational and Mathematical Methods in Medicine 2021 (December): 7433186.
16. Tang, Hongwei, Qingtian Han, and Yong Yin. 2022. "Screening of Important Markers in Peripheral Blood Mononuclear Cells to Predict Female Osteoporosis Risk Using LASSO Regression Algorithm and SVM Method." Evolutionary Bioinformatics Online 18 (January): 11769343221075014.

TABLES AND FIGURES

Table 1: Attributes of CT Scan Dataset

S.No	Images	Mean Intensity	Experimental	Prediction
1		248	Not Cancer	False
2		237	Cancer	True

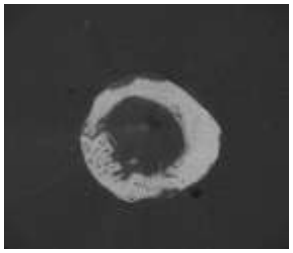
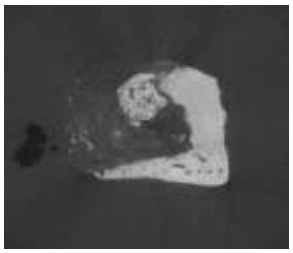
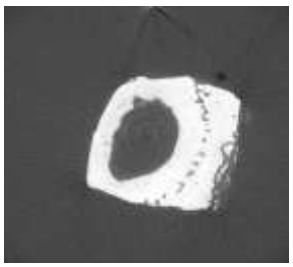
3		244	Cancer	True
4		236	Cancer	True
5		242	Not Cancer	False

Table 2. Efficiency of Enhanced Support Vector Machine and Artificial Neural Network. The Support Vector Machine algorithm is 11% more efficient than the Artificial Neural Network algorithm.

Sample(N)	**Support Vector Machine(SVM)**	**Artificial Neural Network(ANN)**
	Accuracy(%)	**Accuracy(%)**
1	83	71
2	82	69
3	80	68
4	79	67
5	78	65
6	76	64
7	74	62
8	73	60
9	71	59
10	69	57

Table 3. Comparison of the accuracy of Osteoporosis risk prediction using Support Vector Machine and Artificial Neural Networks algorithms. The SVM algorithm had the highest accuracy (83.%) over SVM had the lowest accuracy (71%).

T-Test:

Group Statistics

GROUP		**N**	**Mean**	**STD Deviation**	**STD Error mean**
ACCURACY	**Support Vector Machine**	10	83.30	4.990	1.578
	Artificial Neural Network	10	71.70	4.762	1.506

Table 4. Independent Sample T-Test is applied for the sample collections by fixing the level of significance as 0.05 with confidence interval as 95%. After applying the SPSS calculation, the Support Vector Machine has accepted a statistically significant value ($p<0.05$).

	Equal Variances	**Levene's Test for Equality of Variance**		**Levene's Test for Equality of Variance**						
		F	Sig.	t	df	Sig.(2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Accuracy	Assumed	.003	.857	5.318	18	.000	11.600	2.181	7.017	16.183
	Not Assumed			5.318	17.961	.000	11.600	2.181	7.017	16.183

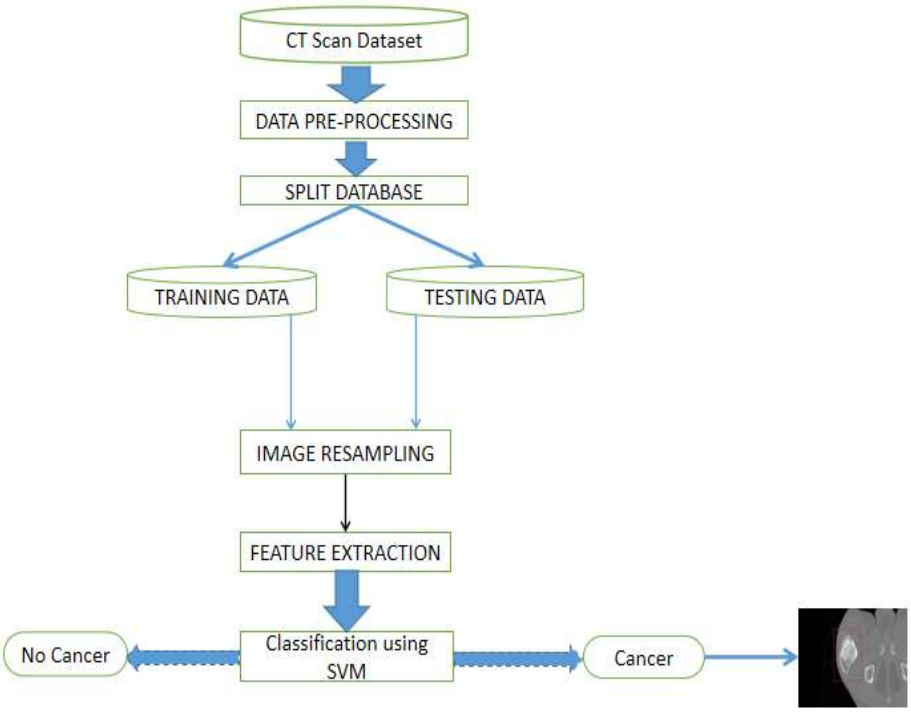


Fig. 1. Architecture for Osteoporosis prediction estimation for each CT Scan image using Enhanced Support Vector Machine algorithm, from dataset processing to output of each image.

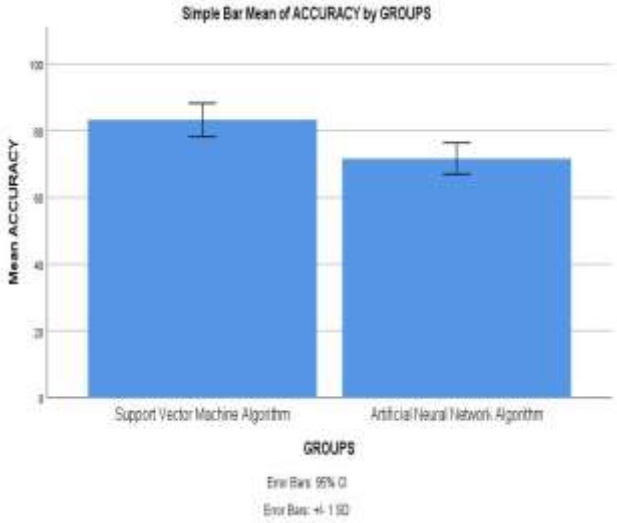


Fig. 2. Bar graph analysis of Support Vector Machine algorithm and Artificial Neural Network algorithm. Graphical representation shows the mean efficiency of 83% and 71% for the proposed algorithm SVM and ANN respectively. X-axis :Support Vector Machine vs Artificial Neural Network , Y-axis : Mean precision ± 1 SD.

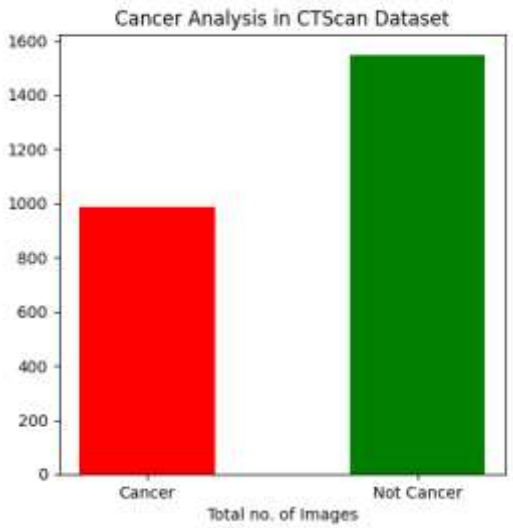


Fig.3. The above graph represents the cancer analysis of CT Scan Dataset in that 987 images contain cancer and 1569 images do not contain cancer.

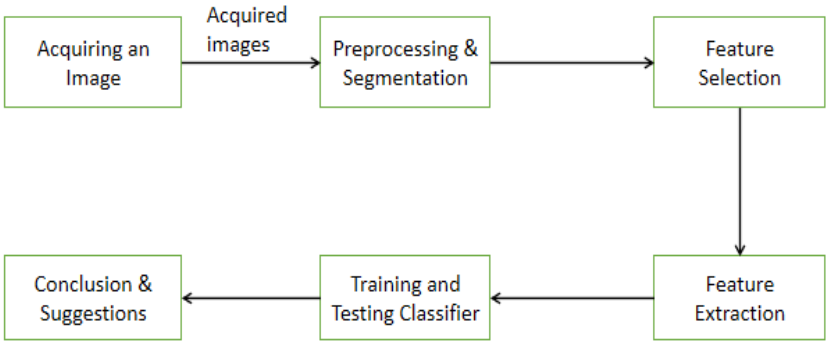


Fig.4.Block Diagram

Fig.5.a)Normal Image b) Preprocessed Image

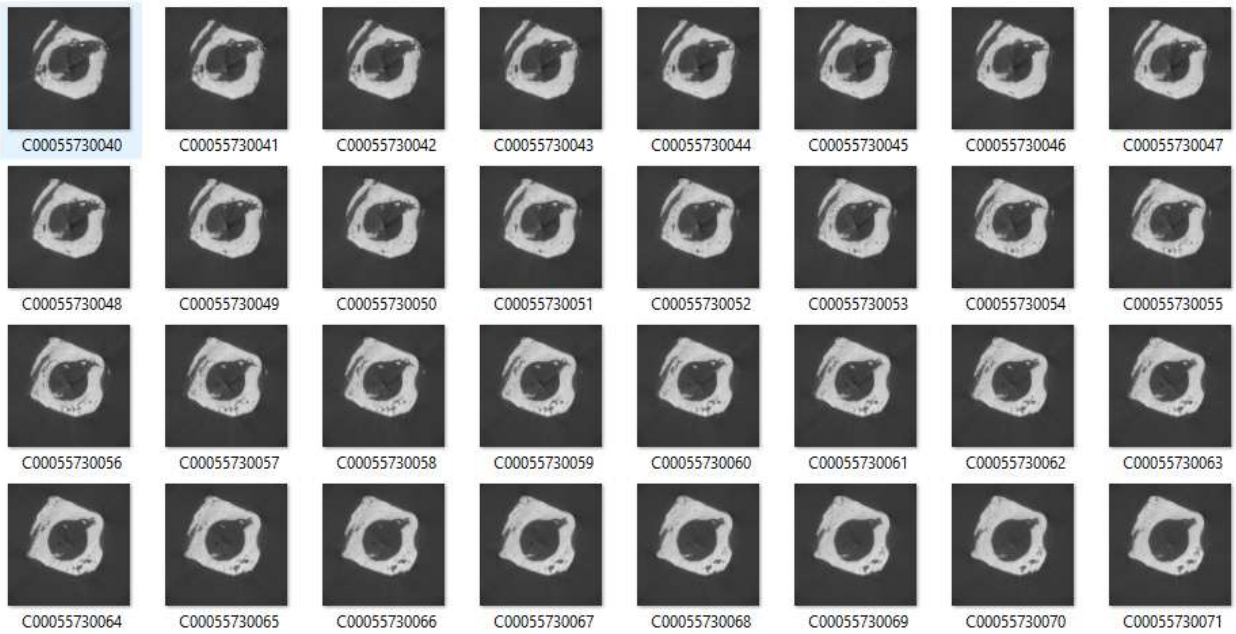


Fig.6. Data stored in dataset of CTscan