

Computer Aided System for Automated Heterogeneous Cancer Recognition using Google Cloud Platform

Ebanesar.C, Hamsa Vagini R, Sangeetha Bregit.A, Dr. J Dinesh Peter

*M.Tech Information Technology
Karunya Institute of Technology & Sciences
Coimbatore, Tamilnadu.*

Abstract - Cancer is the one of the deadliest disease that leads to the death of many people in the present decades. Earlier recognition of tumour can prevent the patient from developing the cells from benign to malignant stage which leads to cancer. computer tomography(CT) is the prevalent way to detect cancer. Our work is formulated by integrating computer and medical intelligence. A CAS accepts CT images as input and the image undergoes deep analysis. With the help of CAS, the doctor can make the decision regarding diagnosis of the cancer. This work is about stochastic recognition of a cancer by using machine learning in google cloud platform. Images incorporate a few undesirable data and some components that are imperative for processing, pre-processing is used for enhancing the image by evacuating distortion and improve vital features. To increase the quality of CT scan image, the CAS adopt Denoising, Gabor and Edge Filters in the pre-processing stage. The image acquired from the previous stage is given as an input to the model for training and automatically deduct the cancer nodule growing region and return the cancer stage of the person. The proposed approach can reduce the error rate and increase the accuracy of the system.

Key words - Image Processing, Google cloud Platform, Computer aided System, Pattern Recognition.

I.INTRODUCTION

Cancer is one of the leading causes of morbidity and mortality worldwide, Cancer is the second leading cause of death globally of 8.8 million in the year of 2015. Cancer is a generic term for a large group of diseases that can affect any part of the body. Other terms used are malignant tumours and neoplasms. One defining feature of the cancer is the rapid creation of abnormal cells that grow beyond their usual boundaries, and which a then invade adjoining parts of the body and spread to other organs, latter process is referred to

As metastasizing. Metastasizing. Metastases are a major cause of death from Cancer is a leading cause of death worldwide, accounting for 8.8 million deaths in 2015. The most common causes of cancer death are cancers of: Lung (1.69 million deaths), Liver (780 000 deaths), Colorectal (770 000 deaths), Stomach (754 004 deaths), Breast (571 200 deaths) [1].

Cancer emerges from the change of ordinary cells into tumour cells in a multistage procedure that transform from a pre-cancerous lesion to a malignant tumour. These changes cause by three main DNA factors. 1) Physical carcinogens know as ultraviolet and ionizing radiation. 2) Chemical carcinogens know as asbestos caused by tobacco smoke, food contamination. 3) Biological carcinogens caused by infections from certain viruses, bacteria or parasites. This Work contribute on Recognising the type of cancer cell and growing region of the tumour. Computer Aided diagnosis(CAD) is the common way for early identifying the cancer cell using CT images using google cloud platform. CAS is developed by using the intelligences of computer science technology and Medical science. The System accept the CT scanned image as the input and process on the image and outcome result is view by doctor with complete details to make the decision. That help for the making the better decision feature by the CAS system using google cloud platform used for image feature extraction. Combining the feature extracted from the platform used for making the better classification comparing with other conventional models. The CAS system Flow diagram shown in fig.1

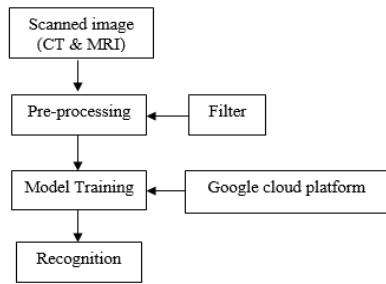


Fig 1: CAS System Flow Diagram

Related Work

Computed Aided Diagnosis for Cancer for medical image was introduced in year 1960 by Lusted L.B. The result is used as the second opinion for the doctor for identification and Recognition using the cellular automata by combining the medical results of the cancer person by training the system [2].

For processing the medical image, important criteria need to maintain is image quality, so medical image used to undergoes, the several stages of pre-processing for making the classification and decision proposed by N.Asadce et.al in 1990.[3]. By Pre-processing the medical image the quality of the image is enhanced by using the filters. Filters are used for removing the noise the image without affecting the information present in the input image.

Tsai et al.,(1995) [4] low pass filter is used for removing the noise in medical image without affecting the quality of the image before processing and identifying the cancer. This filter are used in brain tumour identification using CAD.

H. Aboelsoudm et al.,[5] used the edge filter in medical image processing for detecting the cancer by removing the noise rate by low pass filter, then increasing the image edge level using edge transformation functions by specifying the threshold values.

Zhang jin-yu et al.,(2009) [6] introduce the idea for automatically setting the threshold value based on image we give, that improve the image by sharpening edges of the image using sobel filter. This filter used for medical image for increasing the quality of the image by stochastic way of setting the value that need for sharpening the image.

Rosniza Roslan et al., (2013) [7] used Gabor filter used for image enhancement by detecting the texture of the image after removing the noise and enhance the edge quality using the sobel filter. That allow to make the better decision.

A.H.F Navid et.al, and [8] H.Beigy et al., 2008 used machine learning technique for identification of cancer using CAD. J.P Rennard et al., [9] introduces the idea about Automated cancer identification using CAD using Machine Learning Techniques in year 2000. M.Gomathi et.al in 2010, use the support vector machine for lung cancer detection in CAD system achieve the better result compare with other conventional model. Sarsika Tall et.al (2015) used the SVM machine learning technique in lung cancer

for identification using CAD system. Ming Ni Wu et.al used K-means clustering machine learning method for detection of brain cancer using CAD system. Sajeena A et.al, used sum classification ML method for automated identification of cervical cancer using CAD system.

Neural Network are used in CAD system for better reorganization compare with other system, because neural network uses more hidden layer that can achieve high accuracy, Neural Network is used in lung cancer in CAD system in year 1990 [10] yongjun WV et.al uses 21 feature from CT scanned image for lung achieve 84% of accuracy using neural network for 117 CT image lung.

Punam. S Pawar et.al. use back propagation neural network is used as one of the efficient way for detecting the breast cancer with help of computer aided diagnosis. Ling KH et.al used Deep learning CNN is for automated identification for brain. We used DW-MTM filter for MRI. Muhammad Rukunuddin Ghalibet et.al used hybrid filter for noise removal and image enhancement ANN is used for classification and identification for detecting the renal tumour.

III. Experimental setup

This experimental setup clearly explain step by step procedure to detect the cancer for given CT image using CAS system adapting the google cloud platform. Google cloud provide the platform as service for processing and extracting the feature for give CT image. Our propose CAS detect the cancer image with minimum error rate, input image need to undergo several process for detecting the type of cancer using machine learning techniques. Fig 2. Gives the overall idea about the CAS architecture

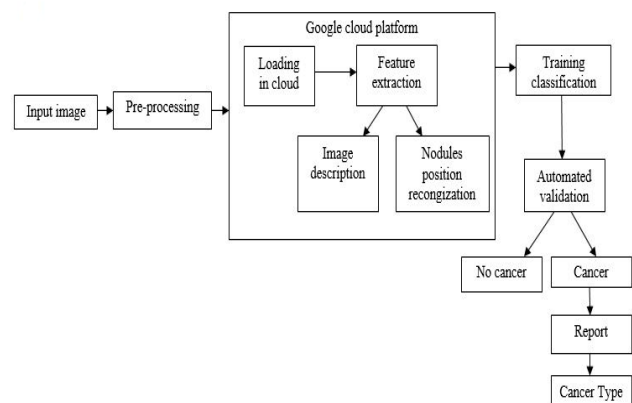


Fig 2 Proposed General Architecture.

From above architecture input medical image need to undergo the several stage for detecting the cancer using google cloud platform. Low pass filter, Sobel filter and Gabor filter are used for image enhancement during the process of input image pre-processing. That help CAS to take better decision compare with other CAD systems and

CAS provide better identification of the input image using machine learning technique.

A. Medical image pre-processing

Image pre-processing play very major role for identification of the cancer in CAS system. In this stage medical image that need be enhanced by removing unwanted noise present in the input image without affecting the quality and information present the input image. Image pre-processing use filter for enhancing the quality of the input image which help the CAS to detect the cancer with low error rate.

1) Filters used in image pre-processing

Filter are used in the process of image enhancement by removing the unwanted information present the input image and sharpen the edges of the input image. Filter is very important because during the input image processing some important information present the image is loosed like blurring of edges and noise etc. To avoid this above short come filter are used in CAS for identifying cancer with low error rate [6]. Edge filter used for sharpening the corner edges of the image using threshold value using transformation function [7]. Sobel filter are used for automatically setting threshold value based on transform function [8]. Gabor filter are used for enhancing the image quality and texture after enhancing the sharpening the edge using above filters [9].

2) Google Cloud platform

Google cloud(gcloud) provide the platform as the service to the developer for developing the application using google cloud platform. CAS adopt the google cloud for platform for Analysing and processing the medical input image for identifying the cancer. After enhancing the image in pre-processing stage outcome of the pre-processed image is given to google cloud platform for extracting the feature of the medical image using google trained model. Features are extracted using the google cloud platform by considering the different criteria for medical image. Google cloud machine learning use GoogLenet(Deep Convolution Neural network). GoogLenet use CNN for extracting the feature, because GoogLenet has 22 hidden layer that accuracy of identifying the extract feature of the cancer [12].

3) gcloud for feature extraction

gcloud used for extracting the feature using GoogLenet using criteria like nodules position and increasing size by comparing previous and current CT image. Gcloud provide the JSON (Java script Object Notation) about the image using Deep CNN with 22 layer hidden layer that increase the identification high accuracy rate. JSON format return the position of tumour with coordinates. By comparing the present and last CT image position we can calculate the tumour growing position rate.

B. Training and classification

After feature are extracted from gcloud. We developed the trained model that will accept the feature and automatically Identify the cancer type and classify the information related to the identified cancer type using our trained model.

IV Work Flow for Automated Recognition of Cancer using CAS

This work flow explain the automated recognition of cancer using CAS by adopting google cloud platform fig 3 explain complete work flow of CAS using gcloud platform.

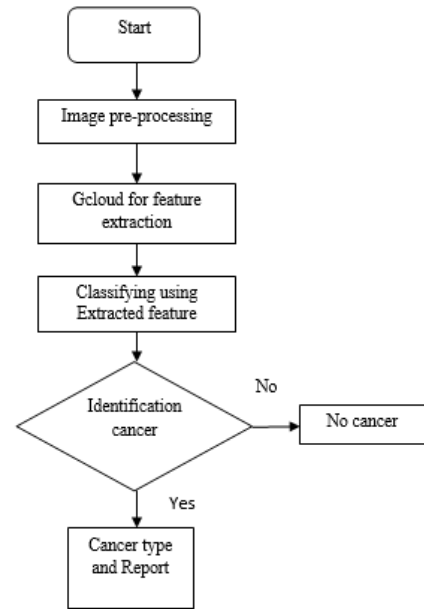


Fig 3 Workflow of CAS

CAS accept the CT image as the input, Before processing the input image, that need to be pre-processed by enhancing the image quality without reducing the input image information. In image enhancement filter are used for reducing the noise ratio and sharpening the image edges by using Sobel and after enhancing the edge of the image texture of the image need to be enhanced by using Gabor filters. This filters are used for image pre-processing without eliminating the information present in the image. After image is enhanced given to gcloud for extracting feature by using google trained machine system. Google uses Deep Convolution neural network that contain 22 hidden layer that increase the level of accuracy by using this algorithm [13]. After extracting feature Reinforcement learning ML[14]for learning the model and Navie Bayes classifier are used for classifying the feature that are extracted from gcloud[15]. Navie Bayes classifier are used to classify based on dependencies. It normally work based on probabilistic way based on feature vector. Finally decision is based on the classified result. The outcome of the CAS system is Cancer and Non Cancer report if the person is affected with the cancer that CAS will give the type of cancer and nodule percentage and Stage of the cancer with other import information related to that Cancer with report that helps the doctor to make the second opinion.

V Machine Learning model used in CAS

Reinforcement learning is machine learning used in CAS for automated identification of the cancer. Reinforcement learning one of the best ML techniques [14] because first it learn the environment then training and testing are made in that environment. Main advantage of using Reinforcement learning is try to finding out the best from the environment using backward propagation technique is also called as exploitation and find out the best from give set is called exploration. We choose to use Q- learning model [16] used in CAS that able to identify Specified thing that from the feature that has ne extracted from gcloud.

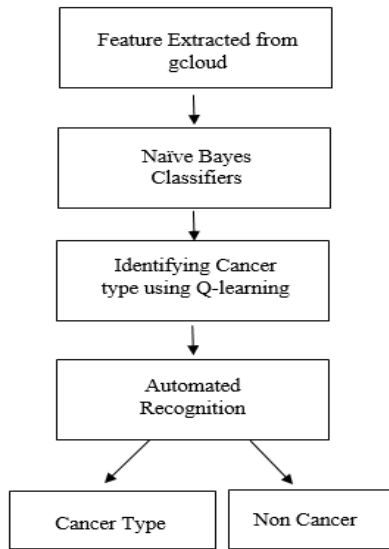


Fig 4 Reinforcement Learning in CAS

Fig 4 explain about the reinforcement learning is used in CAS for diagnosis the cancer and other information extracted from the CT scanned image.

A) Feature extracted from gcloud

Google cloud provide the google vision API platform as a service (Paas) for a developer to application [17]. Using google Paas we extracted medical input image. Google has well trained machine learning model [18].

B) Naïve Bayes Classifiers

Naïve Bayes classifier used for classify the feature that are extracted from gcloud. That find the probabilistically way of finding the result using random probabilistic methods.

$$P(f) = \frac{P(\text{Cancer Feature}) P(\text{Type of Cancer})}{P(\text{Type cancer})}$$

C) Identifying cancer type using Q-learning

Q- Learning used for stochastic way of identifying the cancer type and other information [18]. Q – Learning compare the excepted result with classifier result [19]. This learning method is most stochastic way of learning the environment by performing the action with in the

environment. Based on level of learning environment, action are performed in the environment to make the environment more complex. Making the environment more we can achieve the better results.

VI Proposed Q- learning Algorithm

F ← Feature extracted from gcloud

P (f) ←Classified feature

Q ← Making the Environment to learn using classified feature.

Q ← initialized to an arbitrary value.

T_C ← Testing the environment

T ← Set of feature P(f) = {T}

Nt_n ← New Feature

Δwt_n and ΔwNt_n ← old and new Reward weighted

For all P (f) do

Q ←learn (P (f))

For all T action in T_C where T_C = Q

If t_n in T

T_C ← Action performed (t_n) T = { t₁, t₂.....t_n }

Δwt_n ← calc(t_n)

else

Q ← learn Nt_n Q = T_C

Nt_n ← Action performed in Q

ΔwNt_n ←calc (Nt_n)

end

Above proposed method shows the learning environment and result that came from the environment by calculating the reward achieve from the environment.

VII Result and Analysis

By using the above machine learning technique we made CAS for medical image with help of google cloud platform. Because of this technique rate of execution time and identification accuracy has been increased by using the Q-learning. Automated way of detecting the cancer using medical image that helps the doctors for second opinion. Q-Learning is one of automated learning environment by using that we extracted from gcloud that reduces the rate of execution and increase the level of accuracy of the CAS. Trained model we use help to identify the all the kind of cancer image in CAS. Users need to upload the image in CAS system and system will return type of cancer , region of growth of cancer and other element that found in CT image using Google cloud and Q- learning model in CAS.

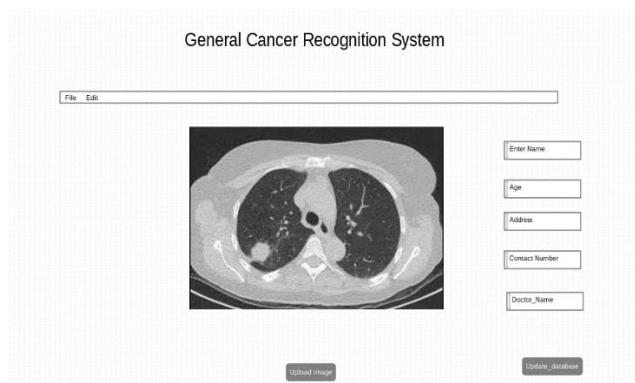


Fig 5 uploading the medical image in CAS

Fig 5 show the how the image is load and detail of the person are filed. We developed desktop application by embedding the google cloud vision Application Programmable Interface. By adopting the cloud platform that reduces the execution rate and increase the Accuracy of the system by GoogLenet(Deep Convolution Network).

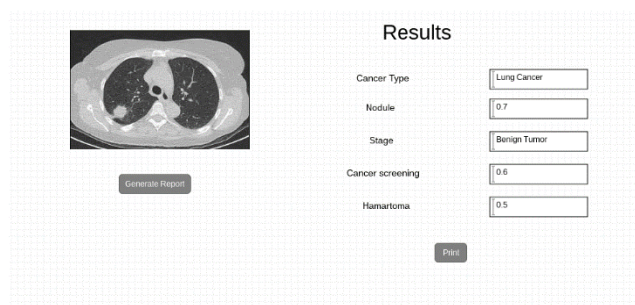


Fig 6 Identified result

Fig 6 gives the clear identified result from the input image. Above result show lung cancer identified result using CAS and Google cloud platform.

Type of Cancer	No of image used training - 20		Image Detected	
	Cancer	Non Cancer	yes	No
Lung	17	3	15	3
Brain	15	5	12	5
Breast	16	4	12	4
Liver	16	4	11	4
Kidney	17	3	14	3
Cervical	13	7	10	7
Skin	15	5	12	5
Prostate	15	5	11	5

Table 1 Result Analysis Table

Above table 1 shows the complete result of our work. 20 medical image is used for detecting the cancer by training and testing image and result achieved from our system. CAS achieve 96 percentage comparing with other computer

aided system. Accuracy got from out CAS system calculated by using precision and recall value

$$\text{Precision} = \frac{\text{Image Detected yes}}{\text{training cancer image} + \text{Non Cancer}}$$

$$\text{Recall} = \frac{\text{Image Detected yes}}{\text{No. cancer image used for training} + \text{Fn}}$$

Fn represent false negative image in table 1 all the Fn value are. Because all non-cancer image has been detected all Fn value zero. By using this accuracy calculated is 96 percent from above results. Above system use less than 0.55 sec to detect the result.

Conclusion

Now a days computer aided system became the common in the area of medicine because of technology that run in the background of the system by combining the medical intelligences and computer technology. CAS is one among the system by using the both the intelligences as the background with help of google cloud platform for faster execution of the system. This help the doctors for making the second opinion by using the CAS and automatic detecting the heterogeneous type of cancer using machine learning model.

Future work

Increasing more accuracy in area of cancer growth area and making the system for readable by adopting the deep neural networks.

References

- [1] <http://www.who.int/cancer/en/>
- [2] L. B. Lusted, "Logical analysis in roentgen diagnosis," Radiology, vol. 74, pp. 178-93,1960.
- [3] N. Asada, K. Doi, H. MacMahon, S. Montner, M. Giger, C. Abe, et al., "Potential usefulness of an artificial neural network for differential diagnosis of interstitial lung diseases: pilot study," Radiology, vol. 177, pp. 857-860, 1990.
- [4] Tsai "Introduction to low pass filter in area of detection in image processing" International conference on Inventive computer Technology (ICICT), vol 2. Pp 1-6, 2000.
- [5] C.S. Avinash, K. Manjunath and S.S. Kumar, "An improved image processing analysis for the detection of lung cancer using Gabor filters and watershed segmentation technique," International conference on Inventive Computation Technologies (ICICT), vol. 3, pp. 1-6, 2016.
- [6] Wenshuo Gao; Xiaoguang Zhang; Lei Yang; Huizhong Liu; "An improved Sobel edge detection" 3rd IEEE

International Conference on Computer Science and Information Technology (ICCSIT), vol. 5, pp. 67-71, 2010.

- [7] Anjali goswami . “For Image Enhancement and segmentation by using evaluation of Gabor filter parameters,” International Journal of Advanced Technology and Engineering research (IJATER).
- [8] A. H. F. Navid and A. B. Aghababa, “Cellular Learning Automata and Its Applications,” Emerging Applications of Cellular Automata edited by Alejandro Salcido, <http://dx.doi.org/10.5772/52953>, 2013.
- [9] J. P. Rennard, “Introduction to cellular automata.” Ed: PhD Thesis, Universitat Trier, 2000.
- [10] W. Yongjun, W. Na, Z. Hongsheng, Q. Lijuan, Y. Zhen, and W. Yiming, “Application of artificial neural networks in the diagnosis of lung cancer by computed tomography,” Sixth International Conference on Natural Computation (ICNC), pp. 147-153, 2010.
- [11] <https://cloud.google.com/vision/documentation>
- [12] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, “Going deeper with convolutions,” In CVPR, 2015
- [13] K. O’Shea and R. Nash, “An introduction to convolution neural networks,” CoRR, abs/1511.08458, 2015.
- [14] K. Das and R. N. Behera, “A Survey on Machine Learning: Concept, Algorithms and Applications,” International Journal of Innovative Research in Computer and Communication Engineering, vol. 5, pp. 1301-1309, 2017.
- [15] D. Lowd, P. Domingos, “Naïve Bayes Models for Probability Estimation,” Proceedings of the 22nd International Conference on Machine Learning (ICML), pp. 529-536, 2005.
- [16] K. Arulkumaran, M. P. Deisenroth, M. Brundage, and A. A. Bharath, “A Brief Survey of Deep Reinforcement Learning,” arXiv: 1708.05866, 2017.
- [17] <https://www.salesforce.com/in/paas>
- [18] cloud.google.com/ml-engine/docs/training-steps