

A Novel Approach to Alzheimer's Disease Stage Classification using Supervised Learning Approach

¹Shaik Basheera and ²M. SatyaSai Ram

¹Department of ECE, Acharya Nagarjuna University College of Engineering, Acharya Nagarjuna University, Guntur

²Department of ECE, Chalapathi Institute of Engineering and Technology, LAM, Guntur, India

Key words: Support vector machines, gray level co-occurrence matrix, Alzheimer's, dementia, Naive Bayes classifier, kNN

Corresponding Author:

Shaik Basheera Department of ECE, Acharya Nagarjuna University College of Engineering, Acharya Nagarjuna University, Guntur, India

Page No.: 137-141 Volume: 19, Issue 7, 2020 ISSN: 1682-3915 Asian Journal of Information Technology Copy Right: Medwell Publications Abstract: Medical imaging playsa major role in diagnosis of diseases, machine learning play major role in diagnosis of medical images using computer-aided diagnosis. As India's urban population is goon increasing Neurological disorders also increases, Alzheimer's is one of the major dementia of neurons and make the death tally as high next to cancer. Estimating the stage of the Alzheimer's is a challenging task. We use T2 Weighted Brain MRI and Extract the Texture Features from those images. Train the classifier and perform crossover validation using those features. Support Vector Machines (SVM) give the good classification accuracy than comparing to the Naïve Bayes classifier and KNN. Test the classifiers with unknown images. The result is compared with clinical information SVM gives 100% accuracy.

INTRODUCTION

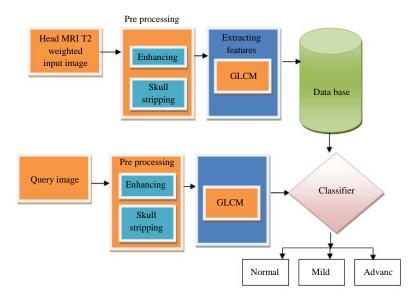
The modern days working environment has changed due to globalization. Working stress is increasing on persons which cause neurological disorders such as dementia. Alzheimer's disease is due to dementia it is a progressive neurological disorder. Generally it appear at the age of 60's. But due to the change in food habits and working environment it appearing in young person's below 40 years of age.

Alzheimer's is the 3rd in major cause of death next to heart diseases and cancer. As India is the second largest in population and developing rapidly in economic. Change in the working Environment and food habits causes Alzheimer's.

Memory loss is the major symptom of the Alzheimer's. It is due to loss the connection between the nerves. White matter of the brain get decreases because of losing the nerve cell connection. Hippocampus volume reduction is observed as primary symptom of Alzheimer's using MRI. Radiologist diagnosis the disease using Imaging technology. But due to visual impairments, less frequent and un characterized imaging features, radiologist face problem to correctly diagnosis the disease.

The texture of the brain is different from one stage to another stage. We come to know that due to the Alzheimer's brain get shrinks, Enlarges its ventricles, change in white and gray matter volume, change in hippocampus size.

Biomarkers are used to evaluate the biological changes carried due to Alzheimer's disease^[1-5]. Which are used to measure β -amyloid, total tau and phospho-tau-181 in Cerebrospinal Fluid (CSF). This technique is the most acceptable method to diagnose AD with high specificity and sensitivity. But Biomarkers are not useful for early diagnosis of the disease. Moreover, it need to use intracerebral-ventricular injection. To collect



Asian J. Inform. Technol., 19 (7): 137-141, 2020

Fig. 1: Work flow

the CSF the clinical staff need to take utmost care without damage brain tissues and spinal card. This is one of the most critical mechanism to estimate the Alzheimer's where they estimate the tangle and plague of the brain tissues for Alzheimer's analysis.

As CSF Biomarkers have remarkable drawbacks. Body Fluid Based Biomarkers are needed to diagnosis the Alzheimer's^[6]. Although, saliva or urine can be easily collected, blood analysis is the gold standard it is still unknown how the concentration of analyses in the blood directly correlates with pathological changes in the brain, especially in AD.

Brain Imaging is used for different brain disorders such as Tumors, subdural hematomas and stroke but not used for severity of the AD^[7-9]. Volumetric analysis is used by analyzing manually or semi-automatic techniques using SPM-5 using MATLAB environment. There the neurologist need to calculate the total volume of the different regions such as white matter, gray matter, CSF and sum together come to conclusion about the stage of the disease.

MATERIALS AND METHODS

Proposed system: Morphological changes in the brain is used as attributes. To collect those attributes texture information is used. The images collected from Harvard Medical School are initially preprocessed using enhancement and Skull stripping algorithm. Work flow is shown in Fig. 1.

Feature extraction: Gray Scale Co-occurrence Matrix (GLCM) is used for feature extraction. The matrix that

gives the probability successive pixels of same intensity in the image. It is a 256×256 square matrix. Preprocessed image verifies the pixel intensities in horizontal direction and write the number representation in the GLCM Matrix.

The procedure continues for all the pixels and note down how many times the same pixels are come side by side based on this we calculate the probability used to find the texture parameters of the image.

Features and database generation: The GLCM used to generate statistical texture parameters such as mean in x direction, mean in y direction, correlation, homogeneity, energy, entropy, standard deviation in x direction, standard deviation in y direction, Angular secondary moment, variance, cluster shade, inertia, skew and skew coefficient, etc., all these features are used to generate the required a data base. This data base is used to train and validate the classifiers.

Classifier: Classifier is used to bifurcate the stage of Alzheimer's disease based on the features that extracted by using GLCM, the features are stored in data base. We use support vector machines classifier having its path in statistics give good results even having small data set. SVM is used as binary classifier used to separate two classes. But our data set is having multi classification for this One Versus all (OVS) or one to one mapping method is used.

Let the training set is having x data set points as (a_1, b_1) , ..., (a_n, b_n) . Where, a_n is the p dimensional feature vector b_n represent the class that be labeled to train the classifier. A hyper plane with great variance represented

Past	A Cut Là Cop Fer Chipboan	nat Painter	Segoe UI B	• s 12 • 12 • Port	• A' A'		- (***) - (* (*	Stand St. 4	Test] e & Center -	1000	• 54 a		ional Form ting - as Tab		100000000	elete Forma	∑ Auto	. Z.	A Find &		
	D1		G	f. gicm	autoCorre	alation	0075														_
1	A	в	c	D	E	F.	6	н	1		к	L	M	N	0	p	a	R	5	т	
,	tage	glcm_mean	glom_varia	glcm_auto Correlation		glcm_cSha de	glcm_cTen dency	glicm_contr ast	gicm_correl ation	gicm_differ enceEntrop	glcm_dissi milarity	gicm_ener gy	gicm_entro py	glcm_hom ogeneity1	glcm_hom ogeneity2	gicm_IDM N	gicm_IDN		gicm_maxP rob	gkm_sum Average	gion ntro
,	ormal1	3.315365	30.65644	39.67646	89358.49	2323.285	120.6825	1.94326	0.9683058	0.9262098	0.3695772	0.663189	1.418031	0.9113272	0.9018703	0.9982663	0.9899616	0.0748576	0.8139859	8.630729	1.
	2	3.344378	30.88506	40.16095	89647.4	2328.958	121.7223	1.817953	0.9705691	0.9409449	0.3746936	0.6596878	1,43541	0.9091445	0.8993428	0.9983346	0.9897461	0.0755556	0.8119638	8.688756	1
	3	3.594332	27.40551	18.299	46924.62	1319.636	107.5706	2.051501	0.9625714	1.047089	0.410386	0.586475	1.65445	0.894164	0.883555	0.9981796	0.9687665	0.1001158	0.7650735	9.188664	1
	4	4.270665	40.80274	56.46408	116280.4	2758.514	160.0565	3.154473	0.9613448	1.254902	0.5666054	0.5351355	2.011934	0.8681066	0.8537157	0.9972072	0.9647203	0.1071057	0.7309743	10.53627	
	5	4.382384	42.62500	\$9.30276	120371.8	2860.473	167.0511	3.461213	0.9595918	1.326303	0.6217831	0.5329202	2.064006	0.8583437	0.8420526	0.9969218	0.9832441	0.1062864	0.7295803	10.75365	
	6	4.338205	43,05906	59.17407	125107.7	2968.579	168.8262	3.410018	0.960403	1.301612	0.6112745	0.5449233	2.016163	0.8622029	0.8461718	0.9969619	0.9835375	0.1007222	0.7377911	10.65619	
	7	3.919507	35.21868	48.101	85506.87	2192.5	137.9096	2.965135	0.9579039	1,159165	0.5187194	0.5786096	1.790248	0.8820765	0.8691232	0.9973841	0.9860623	0.0930639	0.7601103	9.839614	
	8	4.008839	36.12222	50.02762	89370.3	2261.156	142.1582	2.330862	0.9677362	1,141948	0.4722426	0.5676245	1.818183	0.8819416	0.8691806	0.9979113	0.9870708	0.1015645	0.7528339	10.01566	
	9	3.787592	33.90593	46.25606	79638.27	2083.889	133.6339	1.909828	0.9706566	1.036385	0.4095282	0.6050224	1.643574	0.8960474	0.8851587	0.9982085	0.9687689	0.0931066	0.7773284	9.575184	
	10	3.955362	35.37765	48.975	79777.88	2077.346	139.4156	2.095159	0.9703687	1.101557	0,4464767	0.5858685	1.734102	0.8871938	0.8748234	0.998091	0.9877195	0.0960511	0.7648897	9,910723	
	11	4.244393	38-40114	54.24059	93349.26	2302.796	151,4137	2.510846	0.9673756	1.174238	0.4925551	0.5477451	1.091262	0.877562	0.5646233	0.9977645	0.9865608	0.1075566	0.7393995	10.48879	
	12	4.213695	39,93776	55.57237	106668.4	2570.019	157,5098	2.241238	0.9719409	1.117392	0.4575674	0.5701661	1.783963	0.8852924	0.8732201	0.9979759	0.9874695	0.1012988	0.7544118	10,42031	
	13	4.241131	40.67894		115029	2718.732	160.5645	2.151317		1.090213	0.4379596		1.76915	0.8886701	0.5772457	0.9980657	0.9660031	0.1018247	0.7531556	10.46024	
	14	4.281097	42.04179	58.33422	131123.1	3007.471	166.0764	2.090717	0.9751353	1,070129	0.4221814	0.5654514	1.766126	0.8908252	0.8801215	0.998126	0.9554402	0.1052953	0.7511642	10.56219	
	15	4.320098	41.77043		127577	2925.945	164.6311	2.450613		1.134069			1.854918	0.8532134		0.9978023	0.9671631	0.1082002	0.7396293	10.63919	
	16	4.382935	43.65160		134650.3	3121.656	172.7759	2.630821		1,144668			1.862026		0.8701877	0.997677	0.9868552			10.76284	
	17	4.232598	41.90993		123506.2		165.273	2.366728		1.070893			1.748359		0.8810947	0.9979158	0.9581042			10.45913	
	18	4,111501	39.13231		101890.1	2524.108	154.2022	2.267065		1.078593			1.722913		0.8798059		0.9879406			10.21912	
	19	4.115487	39,03141		100967.2	2511.723	153.6261	2.49951		1.12571			1.792719		0.8730932		0.9869535		0.7573683	10.22693	
	20	3.928845	38.69731		109824.7	2741.667		2.150582		1.023568			1.620964		0.8879626	0.9900695	0.9586953			9.054657	
	21	3.809436	35.75358		97466.77	2456.382	140.8353	2.179044		1.043143			1.650045	0.896167	0.8852035	0.9980213	0.9882519		0.7777267	9.618873	
	22 + + She	3.612377 pet1 She	34.228 Mt2 5h		97410.43	2485.477	134.7764	2.135358	0.9687637	1,024702	0.4282475	0.6294556	1.574428	0.8957841	0.8876568	0.9960457	0.9883233		0.7929534		

Asian J. Inform. Technol., 19 (7): 137-141, 2020

Fig. 2: Data base in CSV file

Table 1: Statistical parameters of training classifiers

Classifier	AUC	CA	F1	Precision	Recall
SVM	1.00	0.981	0.951	0.971	1.00
KNN	0.983	0.981	0.957	0.971	1.00
Naive bays classifier	1.00	0.907	0.815	0.688	1.00

with a real vector \vec{q} which is perpendicular to the hyper plane and a offset parameter p is used to make the margin (Table 1 and 2):

 $\vec{q}\vec{a}_i + p = 0$

If the offset become zero then the hyper plane pass through the origin. As the SVM is risk minimizing algorithm it is interested in maximum margin between two classifiers. Two parallel planes are taken if the data set is lineally separable. And no data is placed in-between the hyper plane. As the SVM is binary classifier then we can apply this algorithm by considering one versus all or one to one mapping. In one versus all algorithm one class data is taken as +ve and remaining all classes are labeled as –ve and train the classifier and repeat this procedure for all the class. The algorithm is carried using the given formula:

- $\overrightarrow{qa_i} + p \ge 1$ the given data set is belong the bi = 1 class
- $\vec{qa_i} + p \le -1$ the given data set is belong the bi = -1 class
- $\vec{b}_i(\vec{q}\vec{a}_i+p)-1 \ge 0$ use as a classifier kernel

 $H_{\theta}(a)$, let, we have three classes as normal, medium, advance impairments as $i = \{1, 2, 3\}$ then find the probability of getting the given data set belong to that

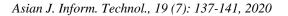
Table 2: Clinical information of Un labelled images

Stages
Mild
Normal
Normal
Advanced
Advanced

class as $H_{\theta}^{i}(a) = p(i/a; w)$ by using this the probability and using maximum value of $H_{\theta}^{i}(a)$, we can come to knowledge that the data set belong to the particular class. This give soft max approach gives the probability as a metric decide the classification of the particular data.

RESULTS AND DISCUSSION

We are taken 54 MRI Slices data base to train our classifier. Which are already labeled as normal, mild and advanced. These images are collected from Harvard medical school. They are total 54 images 21 images are normal, 11 images are mild and 22 images are advanced images. Extracted features are converted as. CSV file as shown in Fig. 2. To compare the performance of the classifiers, we use 5 folded Cross validation and generate confusion matrix of KNN, Navy Baye's classifier and SVM as shown in Fig. 3. Using the respective confusion matrix different parameters are calculated such as Classification accuracy, F1 measure, precision, recall and Area Under Curve (AUC). From all the classifiers are



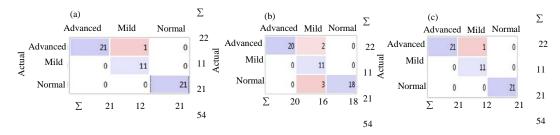


Fig. 3(a-c): Confusion Matrix of individual classifier, (a) SVM, (b) Naive based and (c) kNN classifier

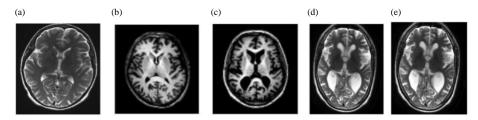


Fig. 4(a-e): Unlabelled test images

	Naive Bay			KNN cla	assifier			SVM classifier				
Images	Advance	Mild	Normal	Status	Advance	Mild	Normal	Status	Advance	Mild	Normal	Status
a	0.00	1.00	0.00	М	0.00	1.00	0.00	М	0.06	0.74	0.00	М
b	0.00	0.00	1.00	Ν	0.00	0.00	1.00	Ν	0.00	0.00	1.00	Ν
с	0.00	0.00	1.00	Ν	0.00	0.00	1.00	Ν	0.00	0.00	1.00	Ν
d	0.03	0.97	0.00	М	0.20	0.80	0.00	Μ	0.20	0.80	0.00	А
e	0.99	0.01	0.00	А	0.40	0.60	0.00	М	0.52	0.46	0.01	А

**Status: N-Normal; M-Mild; A-Advanced

trained by folded 5 cross validation approach and we conclude from the parameters that SVM gives good AUC and accuracy F measurement precision and recall. And the classifier is also test using different unknown images collected from internet source and verified using our algorithm. Unlabelled images are shown in Fig. 4. Clinical stage of these images is given in Table 2.

On predicting the diseases SVM classifier gives the perfect classification while compare with the clinical information as shown in Table 3.

CONCLUSION

We conclude that the Alzheimer's stages are get classified perfectly by the texture attributes collected from the brain using SVM classifier. The predicted results are compared with the clinical information. The Proposed method gives 100% classification accuracy. This can further be expanded to the Schizophrenia as the texture information of the brain is also play the main role in the classification.

REFERENCES

- Frey, H.J., K.M. Mattila, M.A. Korolainen and T. Pirttila, 2005. Problems associated with biological markers of Alzheimer's disease. Neurochem Res., 30: 1501-1510.
- Henley, S.M., G.P. Bates and S.J. Tabrizi, 2005. Biomarkers for neurodegenerative diseases. Current Opin. Neurol., 18: 698-705.
- 03. Sprott, R.L., 2009. Biomarkers of aging and disease: Introduction and definitions. Exp. Gerontol., 45: 2-4.
- 04. Cedazo-Minguez, A. and B. Winblad, 2010. Biomarkers for Alzheimer's disease and other forms of dementia: Clinical needs, limitations and future aspects. Exp. Gerontology, 45: 5-14.
- 05. Ho, L., H. Fivecoat, J. Wang and G.M. Pasinetti, 2009. Alzheimer's disease biomarker discovery in symptomatic and asymptomatic patients: Experimental approaches and future clinical applications Exp. Gerontol., 45: 15-22.

- Brettschneider, S., N.G. Morgenthaler, S.J. Teipel, C. Fischer-Schulz and K. Burger *et al.*, 2005. Decreased serum amyloid β1-42 autoantibody levels in Alzheimer's disease, determined by a newly developed immuno-precipitation assay with radiolabeled amyloid β1-42 peptide. Biol. Psychiatry, 57: 813-816.
- 07. Jack, C.R., D.W. Dickson, J.E. Parisi, Y.C. Xu and R.H. Cha *et al.*, 2012. Antemortem MRI findings correlate with hippocampal neuropathology in typical aging and dementia. Neurology, 58: 750-757.
- Bobinski, M., J. Wegiel, H.M. Wisniewski, M. Tarnawski and M. Bobinski *et al.*, 2004. Neurofibrillary pathology-correlation with hippocampal formation atrophy in Alzheimer disease. Neurobiol. Aging, 17: 909-919.
- 09. Killiany, R.J., T. Gomez-Isla, M. Moss, R. Kikinis and T. Sandor *et al.*, 2004. Use of structural magnetic resonance imaging to predict who will get Alzheimer's disease. Ann. Neurol. Official J. Am. Neurol. Assoc. Child Soc., 47: 430-439.