Advances in Intelligent Systems Research, volume 148
International Conference on Information Technology and Management Engineering (ICITME 2018)

Voxelwise Detection of Cerebral Microbleed in CADASIL Patients by Naive Bayesian Classifier

Fangzhou Bao[1], Meiling Shi[2] and Felix Macdonald[3]
[1]School of Computer Science and Technology, Nanjing Normal University, Nanjing, Jiangsu 210023, China
[2]School of Automation Science and Electrical Engineering, Beihang University, Beijing, 100083, China
[3]Université de Sherbrooke, Sherbrooke, Québec, QC J1K 2R1, Canada

***Abstract*—It is important to detect cerebral microbleed voxels from the brain image of cerebral autosomal-dominant arteriopathy with subcortical infarcts and Leukoencephalopathy (CADASIL) patients. Methods developed by other researchers before have a high variablity of intra-observer and inter-observer. In our study, we collect our dataset from the 20 brain volumetric images, 10 for CADASIL patients and 10 for healthy controls. And we used naive baysian classifier to get the results. We use cross validation to improve the performance of naive Baysian classifier. The results show that the average sensitivity is 74.53±0.96%, the average specificity is 74.51±1.05%, and the average accuracy is 74.52±1.00%.**

Keywords—CADASIL; voxel; naive Baysian classifier; cross validation

I. Introduction

The cerebral autosomal-dominant arteriopathy with subcortical infarcts and Leukoencephalopathy (CADASIL) syndrome is the most common hereditary stroke diseases which can never be ignored.

The subclinical sign of cerebral microbleed (CMB) [1] of CADASIL patients can be detected years prior to clinical manifestation by magnetic resonance imaging (MRI). Susceptibility weighted imaging (SWI) is a 3D flow-compensated T2* imaging technique. Compared to standard MRI protocol, the SWI scans the patient at high resolution (less than 1 mm). It uses the phase image to enhance the contrast, providing 2 to 6 times increased sensitivity in CMB detection.

Manual detection results of CMB location of CADASIL patients were regarded as the ground truth. However, it takes too much time to detect with a high variablity of intra-observer and inter-observer. A large amount of advanced methods based on computer vision and digital image processing were proposed to help manual CMB evaluation. Hou (2017) [2] used leaky rectified linear unit (ReLU). Chen (2016) [3] utilized sparse deep neural network. Hou (2018) [4] employed autoencoder. Jiang (2017) [5] used rank-based average pooling. Lu (2017) [6] presented to use deep convolutional neural network.

Though methods mentioned above are very important attempts in the past, they are not acceptable for 2 reasons: (1) The complexity of these methods are very high. (2) Their detection do not reach voxel-wise resolution.

Our method used naive Bayesian theory to perform the classification task. For general classification tasks, if we have enough training data and the features are well designed, then the accuracy will be high enough to accept. It estimates the posterior probability by priori probability collected for training. And the mathematical foundations for Bayesian theory provides solid explain ability without much complexity. Besides, our purpose it to identify CMB at voxel resolution. In all, our methods is a new computer vision [7-10] and image processing [11-14] approaches.

II. Methodology

Naive Bayesian classifiers are classifiers based on statistical analysis [15]. They can predict class membership probabilities i.e. the probability that the sample with given values of its features belongs to a specific category [16]. Naive Bayesian classifiers assume that the effect of an attribute value on a given category is independent of the values of the other attributes. This assumption is called class conditional independence. It is made to simplify the computations involved and, in this sense, is considered “naive.”

Naïve Bayesian classifier can present a comparable performance to state-of-the-art classifiers, such as support vector machine [17-21], shallow neural network [22-27], etc.

Let X be a sample which is described by measurements made on a set of n attributes. Let H be some hypothesis such as that the data sample X belongs to a specified category C. For classification problems, we want to determine $P(H|X)$, the probability that the hypothesis H holds given the observed sample X. In other words, we are looking for the probability that sample X belongs to category C, given that we know the attribute description of X.

$P(H|X)$ is the posterior probability of H conditioned on X. $P(H)$ is the prior probability. Similarly, $P(X|H)$ is the priori probability of X conditioned on H, and $P(X)$ is the prior probability of X. Bayes’ theorem is useful in that it provides a way to calculate the posterior probability, $P(H|X)$, from $P(H)$, $P(X|H)$, and $P(X)$. Bayes’ theorem is [28]:

$$P(H|X) = \frac{P(X|H) * P(H)}{P(X)}. \tag{1}$$

Let D be a training set of samples and their associated class labels. each sample is represented by an n-dimensional attribute vector, $X = (x_1, x_2, ..., x_n)$, depicting n measurements made on the sample from n attributes, respectively, $A_1, A_2, ..., A_n$.

Suppose that there are m categories, C_1, C_2, ..., C_m. Given a sample X, the classifier will predict that X belongs to the category having the highest posterior probability [29], conditioned on X. That is, the naive Bayesian classifier predicts that sample X belongs to the class C_i if and only if

$$P(C_i|X) > P(C_j|X) \; for \; 1 \le j \le m, j \ne i. \tag{2}$$

By Bayes' theorem, we have

$$P(C_i|X) = \frac{P(X|C_i)*P(C_i)}{P(X)}. \tag{3}$$

Since $P(X)$ is a constant for all categories, only $P(X|C_i)*P(C_i)$ needs to be determined. Note that the class prior probabilities may be estimated by

$$P(C_i) = \frac{|C_{i,D}|}{|D|} \tag{4}$$

where $|C_{i,D}|$ is the number of training samples of class C_i in D and $|D|$ is the total number of training samples of D.

Given data sets with many attributes, it would be extremely computationally expensive to compute $P(X|C_i)$. To reduce computation in evaluating $P(X|C_i)$, the naive assumption of class-conditional independence is made. This presumes that the attributes' values are conditionally independent of one another, given the class label of the sample (i.e., that there are no dependence relationships among the attributes). Thus, we can obtain $P(X|C_i)$ by following formula:

$$P(X|C_i) = \prod_{k=1}^{n} P(x_k|C_i). \tag{5}$$

To predict the class label of X, $P(X|C_i)*P(C_i)$ is evaluated for each class C_i. The classifier predicts that the class label of sample X is the class C_i by formula (2) mentioned above.

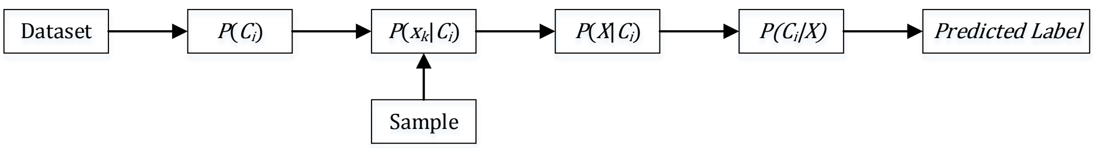


FIGURE I. ILLUSTRATION OF NAÏVE BAYESIAN CLASSIFIER

Figure 1 shows the illustration of naïve Bayesian classifier. We did not use deep learning methods, because this dataset is relatively small, and the convolution operation [30-35] is difficult to handle on an 7x7-size input.

III. MATERIAL AND PREPROCESSING

Figure 2 shows a portion of generated dataset. It can be regarded as a matrix with width of 50 and height of 50. Each row of this matrix represents a sample, the first 49 columns represent input data and the 50-th column represents the target. Input data were generated by vectorizing sliding neighborhood with size of 7*7.

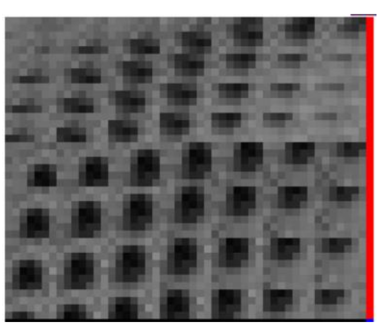


FIGURE II. GENERATED DATASET (COLUMN 1-49 ARE INPUTS, COLUMN 50 IS THE TARGET)

All the pixels in the 7x7 neighbor were regarded as features, and submitted to the naïve Bayesian classifier. We collected in total 20 subjects, and obtained 69,356 CMB voxels, and 124,063,981 non-CMB voxels. For class imbalance problem, we selected randomly 69,327 non-CMB voxels from those 124,063,981 samples. Now we have a dataset as listed in Table 1.

TABLEI. OUR DATASET

Type	Number of samples
CMB	69,365
Non-CMB	69,327
Total	138,692

A 10-fold cross validation was used to segment the dataset into 10 folds, and report the out-of-sample error. For avoiding the effect of randomness, we run the 10-fold cross validation ten times, and report the average and standard deviation of sensitivity, specificity, and accuracy. The definition of those three measures are depicted by following three equations:

$$\text{Sensitivity} = \frac{TP}{TP+FN} \tag{6}$$

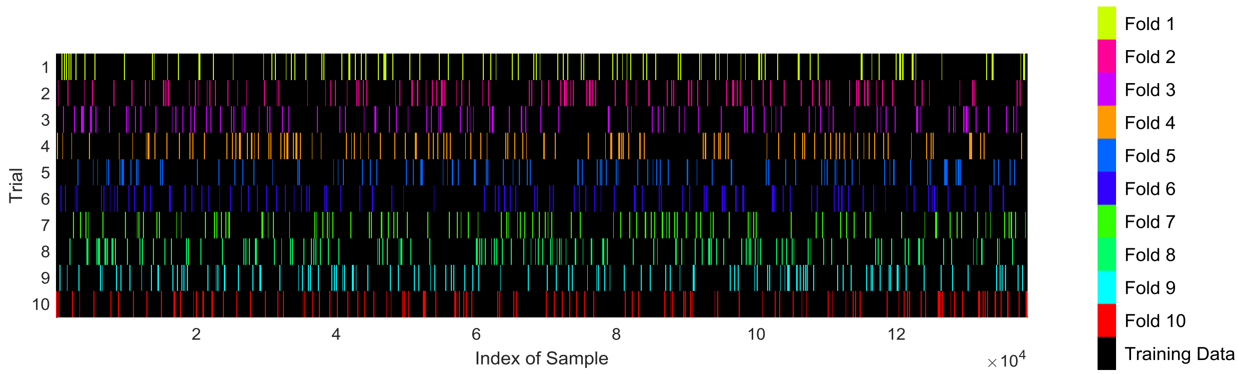
$$\text{Sepcifity} = \frac{TN}{FP+TN} \tag{7}$$

$$\text{Accuracy} = \frac{TP+TN}{TP+FN+FP+TN} \tag{8}$$

where TP *represents* true positive, TN represents true negative, FP represents false positive, and FN represents false negative.

Figure 3 shows the illustration of one run of cross validation. Figure 3(a) shows the index, where the x-axis is the combination of both CMB and non-CMB voxels, in total 138,692. Figure 3(b) shows the legend

.



(a) Index of cross validation (b) Legend

FIGURE III. ONE RUN OF CROSS VALIDATION

IV. Results

The sensitivities, specificities, and accuracies of this 10x10-fold cross validation were shown below in Table 2, The average sensitivity is 74.53±0.96%, the average specificity is 74.51±1.05%, and the average accuracy is 74.52±1.00%. Figure 4 shows the box plot of our algorithm, from which we can observe the three measures can similar results, indicating the effectiveness of our technique to handle the imbalance problem.

TABLEII. PERFORMANCE OF OUR NBC METHOD

Sen	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	Total
R1	77.45	74.31	75.91	76.38	78.71	76.33	75.57	75.35	75.12	79.09	76.42
R2	70.92	77.35	75.58	75.89	76.13	69.45	76.38	75.48	73.40	72.83	74.34
R3	73.94	77.25	70.43	74.67	71.18	75.37	73.85	73.46	71.21	74.80	73.62
R4	73.85	73.18	75.63	74.51	73.14	72.05	73.83	70.75	72.47	71.28	73.07
R5	75.43	74.25	71.61	76.94	75.08	74.35	76.43	74.56	75.12	75.30	74.91
R6	74.41	75.32	73.09	74.86	71.81	75.31	75.84	77.29	76.30	75.55	74.98
R7	75.62	76.64	70.88	76.95	73.59	74.21	78.23	74.54	78.29	72.16	75.11
R8	71.67	73.86	74.17	72.25	74.48	72.47	76.27	71.81	73.66	75.00	73.56
R9	73.23	75.01	73.07	77.41	77.22	73.20	74.92	73.83	72.68	77.84	74.84
R10	70.65	73.92	76.11	73.33	73.86	75.92	72.26	76.89	76.55	74.52	74.40
Spc	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	Total
R1	75.35	76.65	75.84	76.65	78.75	77.40	76.23	74.87	74.64	76.45	76.28
R2	75.38	74.61	75.49	72.62	74.99	74.93	74.73	75.21	75.35	70.11	74.34
R3	74.05	72.59	74.34	73.70	73.01	69.85	71.24	78.28	71.98	74.61	73.37
R4	75.72	74.87	71.03	70.20	75.32	74.01	67.89	72.28	73.76	72.57	72.77
R5	76.43	77.44	72.21	74.63	76.79	76.89	72.94	74.70	72.48	77.60	75.21
R6	74.27	73.82	73.77	75.88	73.10	77.63	74.79	75.08	74.20	78.32	75.08
R7	73.69	76.17	75.06	75.87	76.63	77.37	70.59	75.80	75.21	73.76	75.01
R8	71.60	73.53	75.85	73.98	73.40	72.23	70.75	74.77	74.73	76.47	73.73
R9	76.27	75.05	70.85	76.40	73.01	77.34	73.97	76.42	74.84	78.09	75.22
R10	71.02	73.66	74.27	74.84	73.59	72.52	75.90	76.67	74.87	73.56	74.09
Acc	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	Total
R1	76.40	75.48	75.88	76.52	78.73	76.86	75.90	75.11	74.88	77.77	76.35
R2	73.15	75.98	75.54	74.26	75.56	72.19	75.56	75.34	74.37	71.47	74.34
R3	73.99	74.92	72.39	74.19	72.10	72.62	72.54	75.87	71.60	74.71	73.49
R4	74.79	74.03	73.33	72.35	74.23	73.03	70.86	71.51	73.11	71.93	72.92
R5	75.93	75.85	71.91	75.78	75.93	75.62	74.68	74.63	73.80	76.45	75.06
R6	74.34	74.57	73.43	75.37	72.45	76.47	75.31	76.18	75.25	76.94	75.03
R7	74.66	76.41	72.97	76.41	75.11	75.79	74.41	75.17	76.75	72.96	75.06
R8	71.63	73.70	75.01	73.11	73.94	72.35	73.51	73.29	74.19	75.74	73.65
R9	74.75	75.03	71.96	76.90	75.12	75.27	74.44	75.12	73.76	77.97	75.03
R10	70.83	73.79	75.19	74.09	73.73	74.22	74.08	76.78	75.71	74.04	74.25

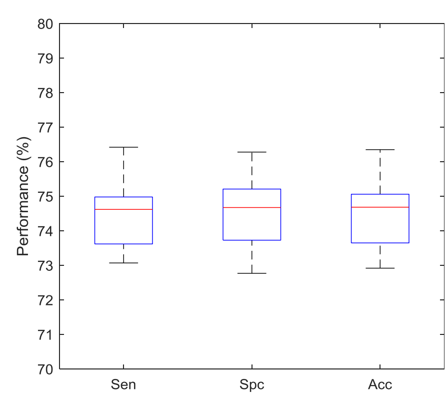


FIGURE IV. BOX PLOT OF OUR ALGORITHM

In the next experiment, we compared NBC with decision tree using ID3 algorithm [36]. The results were shown in Figure 5. Here ID3 obtained a lower performance, with an overall accuracy of 72.19%. We can observe that our native Bayesian classifier can get 2.33% more performance in terms of accuracy.

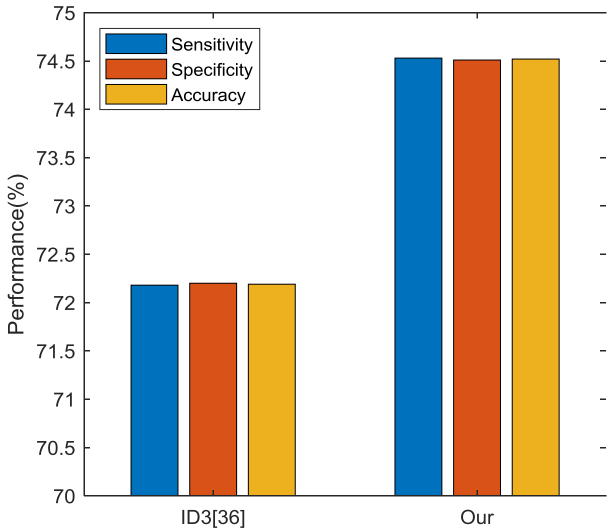


FIGURE V. COMPARISON BETWEEN OUR METHOD WITH ID3 ALGORITHM

V. CONCLUSION

In this study, we proposed a new voxelwise CMB detection system using naive baysian classifier for CADASIL patients. The results showed better result than ID3.

In the future, we will collect more brain images and test other advanced techniques for classification and image preprocessing, to see if we can achieve better accuracy which can meet the requirement of real applications.

We shall test using optimization algorithm [37, 38] to improve the performance of our NBC classifier. In addition, we shall collect more data, and try deep learning methods.

REFERENCES

[1] Imamura, J., et al., *Antithrombotic Therapy Does Not Increase Cerebral Microbleeding.* Stroke, 2013. **44**(2): p. 1

[2] Hou, X.-X., *Voxelwise detection of cerebral microbleed in CADASIL patients by leaky rectified linear unit and early stopping*. Multimedia Tools and Applications, 2017, doi: 10.1007/s11042-017-4383-9.

[3] Chen, H. *Sparse Autoencoder based deep neural network for voxelwise detection of cerebral microbleed*. in *22nd International Conference on Parallel and Distributed Systems*. 2016. Wuhan, China: IEEE. p. 1229-1232

[4] Hou, X.-X., *Seven-layer deep neural network based on sparse autoencoder for voxelwise detection of cerebral microbleed.* Multimedia Tools and Applications, 2018. **77**(9): p. 10521-10538

[5] Jiang, Y.Y., *Cerebral Micro-Bleed Detection Based on the Convolution Neural Network With Rank Based Average Pooling.* IEEE Access, 2017. **5**: p. 16576-16583

[6] Lu, S. *Detection of cerebral microbleeding based on deep convolutional neural network*. in *14th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP)*. 2017. Chengdu, China: IEEE. p. 93-96

[7] Ji, G., *Fruit classification using computer vision and feedforward neural network.* Journal of Food Engineering, 2014. **143**: p. 167-177

[8] Bradley, P.S., *A support-based reconstruction for SENSE MRI.* Sensors, 2013. **13**(4): p. 4029-40

[9] Wu, L., *Crop Classification by forward neural network with adaptive chaotic particle swarm optimization.* Sensors, 2011. **11**(5): p. 4721-4743

[10] Wu, L., *Magnetic Resonance Brain Image Classification by an Improved Artificial Bee Colony Algorithm.* Progress in Electromagnetics Research, 2011. **116**: p. 65-79

[11] Wu, L.N., *Improved image filter based on SPCNN.* Science In China Series F-Information Sciences, 2008. **51**(12): p. 2115-2125

[12] Wu, L.N., *Segment-based coding of color images.* Science in China Series F-Information Sciences, 2009. **52**(6): p. 914-925

[13] Wei, G., *Color Image Enhancement based on HVS and PCNN.* SCIENCE CHINA Information Sciences, 2010. **53**(10): p. 1963-1976

[14] Wang, S., *A novel method for magnetic resonance brain image classification based on adaptive chaotic PSO.* Progress in Electromagnetics Research, 2010. **109**: p. 325-343

[15] Zhou, X. *Detection of pathological brain in MRI scanning based on wavelet-entropy and naive Bayes classifier*. in *International Conference on Bioinformatics and Biomedical Engineering (IWBBIO)*. 2015. Granada, Spain: Springer International Publishing. p. 201-209

[16] Becker, E.J., et al. *Determination of the Defining Features of Texts Written in Isolation with a Naive Bayesian Classifier*. in *8th IEEE Integrated STEM Education Conference (ISEC)*. 2018. Princeton Univ, Princeton, NJ: IEEE. p. 209-210

[17] Liu, A., *Magnetic resonance brain image classification via stationary wavelet transform and generalized eigenvalue proximal support vector machine.* Journal of Medical Imaging and Health Informatics, 2015. **5**(7): p. 1395-1403

[18] Ji, G., *An MR brain images classifier system via particle swarm optimization and kernel support vector machine.* The Scientific World Journal, 2013: Article ID. 130134

[19] Wu, L., *An MR brain images classifier via principal component analysis and kernel support vector machine.* Progress In Electromagnetics Research, 2012. **130**: p. 369-388

[20] Wu, L., *Classification of fruits using computer vision and a multiclass support vector machine.* Sensors, 2012. **12**(9): p. 12489-12505

[21] Dong, Z., *Classification of Alzheimer disease based on structural magnetic resonance imaging by kernel support vector machine decision tree.* Progress In Electromagnetics Research, 2014. **144**: p. 171-184

[22] Rao, R.V., *Abnormal Breast Detection in Mammogram Images by Feed-forward Neural Network trained by Jaya Algorithm.* Fundamenta Informaticae, 2017. **151**: p. 191-211

[23] Lu, Z., *A Pathological Brain Detection System Based on Radial Basis Function Neural Network.* Journal of Medical Imaging and Health Informatics, 2016. **6**(5): p. 1218-1222

[24] Wu, J., *Fruit classification by biogeography-based optimization and feedforward neural network.* Expert Systems, 2016. **33**(3): p. 239-253

[25] Chen, M., et al., *Detection of dendritic spines using wavelet-based conditional symmetric analysis and regularized morphological shared-weight neural networks.* Computational and Mathematical Methods in Medicine, 2015: Article ID. 454076

[26] Wei, L., *Fruit classification by wavelet-entropy and feedforward neural network trained by fitness-scaled chaotic ABC and biogeography-based optimization.* Entropy, 2015. **17**(8): p. 5711-5728

[27] Feng, C., *Feed-forward neural network optimized by hybridization of PSO and ABC for abnormal brain detection.* International Journal of Imaging Systems and Technology, 2015. **25**(2): p. 153-164

[28] Mehmood, A., et al., *ANTSC: An Intelligent Naive Bayesian Probabilistic Estimation Practice for Traffic Flow to Form Stable Clustering in VANET.* IEEE Access, 2018. **6**: p. 4452-4461

[29] Brownjohn, J.M.W., et al., *Bayesian operational modal analysis of Jiangyin Yangtze River Bridge.* Mechanical Systems and Signal Processing, 2018. **110**: p. 210-230

[30] Cheng, H., *Classification of Alzheimer's Disease Based on Eight-Layer Convolutional Neural Network with Leaky Rectified Linear Unit and Max Pooling.* Journal of Medical Systems, 2018. **42**: Article ID. 42

[31] Jia, W.-J. *Ford Motorcar Identification from Single-Camera Side-View Image Based on Convolutional Neural Network*. in *18th International Conference on Intelligent Data Engineering and Automated Learning (IDEAL)*. 2017. Guilin, China: Springer. p. 173-180

[32] Pan, C., *Abnormal breast identification by nine-layer convolutional*

neural network with parametric rectified linear unit and rank-based stochastic pooling. Journal of Computational Science, 2018. **27**: p. 57-68

[33] Lv, Y.-D., *Alcoholism detection by data augmentation and convolutional neural network with stochastic pooling*. Journal of Medical Systems, 2018. **42**(1): Article ID. 2

[34] Zhao, G., *Polarimetric synthetic aperture radar image segmentation by convolutional neural network using graphical processing units*. Journal of Real-Time Image Processing, 2017, doi: 10.1007/s11554-017-0717-0.

[35] Muhammad, K., *Image based fruit category classification by 13-layer deep convolutional neural network and data augmentation*. Multimedia Tools and Applications, 2017, doi: 10.1007/s11042-017-5243-3.

[36] Phu, V.N., et al., *A decision tree using ID3 algorithm for English semantic analysis*. International Journal of Speech Technology, 2017. **20**(3): p. 593-613

[37] Phillips, P., *Pathological brain detection in magnetic resonance imaging scanning by wavelet entropy and hybridization of biogeography-based optimization and particle swarm optimization*. Progress In Electromagnetics Research, 2015. **152**: p. 41-58

[38] Wu, X., *Smart detection on abnormal breasts in digital mammography based on contrast-limited adaptive histogram equalization and chaotic adaptive real-coded biogeography-based optimization*. Simulation, 2016. **92**(9): p. 873-885