Separability of Acute Cerebral Infarction Lesions in CT Based Radiomics: Toward Artificial Intelligence-Assisted Diagnosis

Yun Guan, Peng Wang, Qi Wang, Peihao Li, Jianchao Zeng, Pinle Qin, and Yanfeng Meng

North University of China-Taiyuan Central Hospital Joint Innovation Institute, 3 Xueyuan Road, Taiyuan, Shanxi 030051, China
College of Big Data, North University of China, 3 Xueyuan Road, Taiyuan, Shanxi 030051, China
Taiyuan Central Hospital of Shanxi Medical University, 5 Dong San Dao Lane, Jiefang Street, Taiyuan, Shanxi 030009, China
School of Information and Communication Engineering, North University of China, 3 Xueyuan Road, Taiyuan, Shanxi 030051, China

Correspondence should be addressed to Yanfeng Meng; yanfeng.m@163.com

Received 16 September 2020; Revised 2 November 2020; Accepted 3 November 2020; Published 16 November 2020

Academic Editor: Zhiguo Zhou

Copyright © 2020 Yun Guan et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This study aims at analyzing the separability of acute cerebral infarction lesions which were invisible in CT. 38 patients, who were diagnosed with acute cerebral infarction and performed both CT and MRI, and 18 patients, who had no positive finding in either CT or MRI, were enrolled. Comparative studies were performed on lesion and symmetrical regions, normal brain and symmetrical regions, lesion, and normal brain regions. MRI was reconstructed and affine transformed to obtain accurate lesion position of CT. Radiomic features and information gain were introduced to capture efficient features. Finally, 10 classifiers were established with selected features to evaluate the effectiveness of analysis. 1301 radiomic features were extracted from candidate regions after registration. For lesion and their symmetrical regions, there were 280 features with information gain greater than 0.1 and 2 features with information gain greater than 0.3. The average classification accuracy was 0.6467, and the best classification accuracy was 0.7748. For normal brain and their symmetrical regions, there were 176 features with information gain greater than 0.1, 1 feature with information gain greater than 0.2. The average classification accuracy was 0.5414, and the best classification accuracy was 0.6782. For normal brain and lesions, there were 501 features with information gain greater than 0.1 and 1 feature with information gain greater than 0.5. The average classification accuracy was 0.7480, and the best classification accuracy was 0.8694. In conclusion, the study captured significant features correlated with acute cerebral infarction and confirmed the separability of acute lesions in CT, which established foundation for further artificial intelligence-assisted CT diagnosis.

1. Introduction

Globally, stroke is still the leading cause of mortality and disability, and there are substantial economic costs for post-stroke care [1–4]. In practice, CT is the preferred radiologic modality for patients with stroke-like clinical manifestation, since it is immediately available, cost effective, and capable of differentiating brain disorders [5]. CT is very sensitive in detecting intracranial hemorrhagic stroke and chronic ischemic stroke. CT detects acute cerebral infarction (ACI) in terms of decrease of CT attenuation, loss of gray-white matter differentiation, sulcal effacement, and other indirect signs [6]. In practice, radiologists often encounter poor accuracy in diagnosing acute infarct by CT, with accuracy rate ≤67% within 3 hours [5].

Patient, who has stroke-like symptom but CT showed negative findings, needs MRI [7]. MR diffusion-weighted imaging (DWI) can detect ischemic lesions within minutes...
of symptoms, which is an extremely sensitive technology and used for estimating whether thrombolysis is appropriate. However, MRI is still not the primary modality, because it is time-consuming, which may lead to miss time window of thrombolysis, costing expensive, and various contradictions [8, 9]. Studies demonstrated that better clinical outcomes correlated with earlier diagnosis of ischemic stroke [10]. Therefore, it is essential to improve the accuracy rate of early recognizing ACI by CT within the time window.

In this study, we hypothesize that ACI lesions in CT are separable by combining image registration, accurate lesion location, radiomic feature extraction, and information gain calculation, so as to establish a foundation for artificial intelligence-assisted CT diagnosis.

2. Materials and Methods

2.1. Patients. This retrospective study protocol was reviewed and approved by the institutional review board of our hospital. Written informed consent was waived.

Between February 2019 and February 2020, we retrospectively studied 38 patients, who have performed both CT and MRI and diagnosed as ACI by DWI; meanwhile, CT has no positive finding by radiologist. Another 18 patients with no positive finding in either CT or MRI were also enrolled as normal control (Figure 1).

2.2. CT and MRI. CT images were acquired on a 320-MDCT scanner (Aquilion ONE, Toshiba Medical System, Otawara, Japan) with the following parameters: 120 kV and 300 mA, 5-mm slice thickness, 512 × 512 matrix, and 0.6 mm collimation.

MRI was acquired on 1.5 T MR scanner (Sonata, Siemens Healthcare, Erlangen, Germany) and 3.0 T MR scanner (MAGNETOM Skyra, Siemens Healthcare, Erlangen, Germany). The parameters of 1.5 T DWI were as follows: TR/TE = 3800 ms/84 ms, slice thickness = 5 mm, matrix = 128 × 128, FOV = 200 × 220, b value = 1000. The parameters of 3.0 T DWI were as follows: TR/TE = 4950 ms/64 ms, slice thickness = 5 mm, matrix = 164 × 164, FOV = 220 × 220, b value = 1000.

2.3. Registration and Candidate Region Acquisition. The pipeline of our methodology was shown in Figure 2. Since the position and angle of CT were different from MRI for one patient, the DWI had to be registered to CT images (Figure 3). Herein, the CT images were not adjusted to avoid loss of intact information. The DWI were multiple planners reconstructed to get a consistent angle with CT and achieve coarse registration. Then, a series of affine transformations were performed to get a consistent position and achieve fine registration, including translation, rotation, and scaling transformation.

Early cerebral infarction was obvious on DWI which was sensitive to the restricted Brownian movement of water molecules in brain tissue [5]. Immediately after registration, we highlighted the lesion regions in DWI through adjusting the window width and window level. Because of CT and DWI were matched, the salient lesion position of DWI was also used as the lesion label for CT.
Besides delineating exact lesion regions in CT projected from DWI, we took the midline of the brain as the axis of symmetry and depicted the profile of symmetrical regions (Figure 4). Instead of simply comparing the left and right sides of the brain [11–13], the lesion regions and their symmetrical regions were served as candidate regions to reduce redundant information and achieve accurate comparative analysis.

2.4. Feature Extraction and Analysis. Unlike Lo et al. [13] who improved a texture feature of radiomics, our scheme was to extract features from the image firstly, and then used machine learning techniques to learn these features, so that the computer can mine the information of cerebral infarction in CT according to the acquired characteristics and then identify. Radiomic feature extraction and statistical analysis were performed to complete the plentiful features extraction.
and choose features which are contributing to classification, respectively.

High-throughput feature extraction was applied to search abundant information in CT images. In this study, we followed the radiomic method by Lambin et al. [14] for the extraction process, which were divided into two steps: (1) image transformation and (2) feature calculation[15]. For image transformation, feature map was constructed nonlinearly from the original images, including wavelet, square root, and logarithm. For feature calculation, the feature was calculated on the original and transformed images, including first-order statistics and gray level cooccurrence matrix.

The information gain was further introduced as a statistical standard to measure the correlation between radiomic features and ACI, which was devoted to select significant information from a great deal features of above [16–18]. Each feature was calculated out a value in terms of information gain for dichotomy, lesion, or normal regions, by the equation below:

\[
H(X_i) = \sum_{x \in X_i} - p(x) \log p(x),
\]

\[
H(Y | X_i) = \sum_{x \in X_i} p(x)H(Y | X_i = x),
\]

\[
IG(X_i, Y) = H(Y) - H(Y | X_i),
\]

where \(X_i\) represents the random variable of \(i\)th feature value, \(x \in X_i\) denotes the possible value of random variable \(X_i\), \(p(x)\) represents the probability when the random variable \(X_i\) takes the value \(x\), and \(Y\) denotes the random variable of whether or not a cerebral infarction. \(IG(X_i, Y)\) represents the information gain which is used to measure the reduction of uncertainty of event \(Y\) after \(X\) is known.

Theoretically, the feature was effective when the value was greater than zero, but we chose 0.1 as the minimal threshold to prevent calculating error and sampling error [18]. That is, features below the threshold were considered to be insignificant. The higher the information gain value of features, the greater contribution to remove noise and retain significant feature information.

### 3. Results

Demographic characteristics of all the patients in this study were shown in Table 1. For each of the 38 patients, one slice from CT was selected, which had a prominent lesion in MRI correspondingly. A total of 38 slices, which are 38 ACI regions and 38 symmetric noninfarct regions, were obtained. We extracted 1301 radiomic features from the candidate regions; meanwhile, the information gain was calculated to extract key information from abundant features. As shown in Table 2, there were 280 features with information gain greater than 0.1, which were considered to be contributory to classify candidate regions. There were 23 features with information gain greater than 0.2, and 2 features with information gain greater than 0.3, which showed potential
capability for separating lesion regions and their symmetric noninfarct regions. The related features were used to build up 10 classifiers to verify the feature effectiveness in candidate regions. The average classification accuracy was 0.6467, and the best classification accuracy was 0.7748 (Table 3).

For each of the 18 patients with no positive finding in either CT or MRI, three slices from CT were selected to augment data, which depicted by projecting the lesion labels obtained from 38 aforementioned MRI. A total of 54 slices, which are 54 normal brain tissue regions and 54 symmetrical regions, were obtained. As shown in Table 2, there were only 176 features with information gain greater than 0.1, 1 feature with information gain greater than 0.2, and no feature with information gain greater than 0.3. Although the best classification accuracy was 0.6782, from the overall classification results, the classification results were generally low and the average classification accuracy was only 0.5414 (Table 3).

For each of the 56 aforementioned patients, one slice from CT were selected. A total of 56 slices, which are 38 lesion regions and 18 normal brain tissue regions, were obtained. As shown in Table 2, there were 501 features with information gain greater than 0.1, 126 features with information gain greater than 0.2, 51 features with information gain greater than 0.3, 18 features with information gain greater than 0.4, and 1 feature with information gain greater than 0.5. The average classification accuracy was 0.7480, and the best classification accuracy was as high as 0.8694 (Table 3).

Besides, feature map that features reflected on candidate regions were shown to illustrate the effectiveness of feature analysis (Figure 5). We visualized one of the first three features ordered by information gain on the candidate region. Among them, it is a clear distinction on lesion and its symmetrical region, which explains the separability of ACI. The left and right sides of the normal region showed no obvious difference, which confirmed the inseparability of the left and right sides of the normal brain. The difference between lesion and normal region was also prominent, which indicated separability of lesion and normal region.

### 4. Discussion

Sensitively recognizing acute cerebral infarction is a valuable research for clinical treatment, within effective thrombolytic time. To the best of our knowledge, the finding of analyzing the separability of acute cerebral infarction lesions in CT based on image registration, precision positioning, radiomic feature extraction, and information gain calculation has not previously been well established in the literature. The overall concept of the algorithm was to extract and analyze the feature of regions where there is cerebral infarction, and more importantly, to separate lesions from normal regions.

Accurately recognizing acute ischemic stroke by CT remains challenging, due to the low accuracy of radiologist diagnosis. A lot of studies focused on prior knowledge, including decrease of CT attenuation, loss of gray-white matter differentiation, and sulcal effacement. However, 1/3 cases were missed since the sensitivity and specificity were low [5].
On the other hand, acute ischemic stroke is inconspicuous, more complex features including texture, need to be introduced [19, 20]. Later, Rajini [11] proposed a method to separate ischemic stroke regions from normal tissues in CT, which used segmentation, midline offset, and image features. Nevertheless, most of these studies involved acute lesions that were already visible. Instead, Chawla [12] investigated a two-stage classification system by comparing the image intensity differences between the two hemispheres, which can detect hemorrhagic and ischemic stroke. Recently, a predictive model based on Ranklet features to distinguish strokes and normal tissues was proposed, which achieved significantly high accuracy 81% [13]. Compared with MRI, it is still a gap which is not sufficient for practice, and artificial intelligence-assisted CT diagnosis needs more robust features. As Petrou [21] suggested, a few features, which are not sensitive to human vision and tend to be ignored, are needed to be excavated. Therefore, it is potential for improving the accuracy of detection ACI by exploring other features besides texture.

Lesion delineation is the primary premise in medical image analysis. However, defining the entire lesion boundaries in CT might be complicated because of the invisible of lesion. The next practical way is that ischemic tissues can be highlighted by comparing the left and right sides, since inherent anatomical structures in the human brain are symmetric [22, 23]. Nevertheless, it is unreliable by simply comparing both sides of the brain especially the lesion size was small because of normal brain tissue overwhelming the characteristic features from small lesion.

In order to solve the conundrum of candidate region acquisition and feature quantity insufficiency, we matched exact lesion regions from DWI to CT images. The exact lesion regions and their symmetrical regions served as candidate regions for imaging feature extraction. Inspired by Lambin et al. [14] who extracted a large number of radiomic features from medical images and used statistical analysis to identify features that could characterize disease, we extracted 1301 radiomic features through image transformation and feature extraction. Note that we do not claim any novelty in the extraction design. Instead, our contribution lies in the essence of that constructing more complex feature is necessary for selecting certain features which contribute to classification in the next step. Since not all features were effective, the information gain was further introduced as a standard to measure the correlation between features and ACI, which is according to the principle of feature distinction and independence in mathematical description [16–18]. Furthermore, machine learning is often used as a means to evaluate radiomic analysis [24–26]. Hence, 10 classifiers were established with selected features to verify the effectiveness.

The sufficient experimental data showed differences between the cerebral infarction and their symmetrical non-infarct regions, since the effective features extracted had great potential in classify lesions and their symmetrical regions. Simultaneously, to rule out these separable differences probably coming from the inherently separable between the left and right sides of the normal brain, we operated on normal brain tissue and their symmetrical regions. The results confirmed that the left and right sides of the normal brain tissues were inseparable. According to effective features achieved astounding performance in classify lesion regions and same position of another normal brain tissue, the lesions which were separable from normal tissue in CT were further confirmed. More importantly, the classification results proved the necessity and effectiveness of feature extraction and screening.

Some limitations are noteworthy in the current study. We only included 18 patients performed head CT with no positive finding. They were younger compared to 38 patients performed with both CT and MRI and diagnosed as ACI by DWI, since it is difficult to select normal brain tissue in the elderly. Besides, the size of our population might be considered small; further studies that include a larger population are needed to strengthen the statistical power of these investigations.

5. Conclusion

This study analyzed the separability of acute cerebral infarction lesions in CT, which facilitates CT diagnosis directly. Furthermore, the results of the study established a theoretical foundation for artificial intelligence-assisted CT diagnosis, which will bring potential benefits for acute cerebral infarction patients: shortening the waiting time of thrombolysis, saving the cost of examination, and improving the prognosis.
Abbreviations

ACI: Acute cerebral infarction
DWI: Diffusion-weighted imaging.

Data Availability

The data used to support the findings of this study is available from the corresponding author upon request.

Conflicts of Interest

The authors report no conflicts of interest in this work.

Acknowledgments

We thank Dr. Yanpeng Cheng, PhD, from the Chinese Center for Disease Control and Prevention, for his help in Statistics. This work was supported by the National Natural Science Foundation of China (Grant No.: 81671789), the project of Health Commission of Shanxi Province (Grant No.: 2018109), the Construction project of Engineering Technology Research Center of Shanxi Province (Grant No.: 201805D121008), the Graduate Research and Innovation Projects of Shanxi Province (Grant No.: 2020SY381), and the Science and Technology Foundation of North University of China (Grant No.: 20191634).

References

[1] C. O. Johnson, M. Nguyen, G. A. Roth et al., “Global, regional, and national burden of stroke, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016,” *Lancet Neurology*, vol. 18, no. 5, pp. 439–458, 2019.
[2] S. Rajsic, H. Gothe, H. H. Borba et al., “Economic burden of stroke: a systematic review on post-stroke care,” *The European Journal of Health Economics*, vol. 20, no. 1, pp. 107–134, 2019.
[3] J. Kim, T. Thyabaranathan, G. A. Donnan et al., “Global stroke statistics 2019,” *International Journal of Stroke*, vol. 15, no. 8, pp. 819–838, 2020.
[4] Y. J. Wang, Z. X. Li, H. Q. Gu et al., “China Stroke Statistics 2019: A Report from the Chinese Center for Healthcare Quality Management in Neurological Diseases, China National Clinical Research Center for Neurological Diseases, the Chinese Stroke Association, National Center for Chronic and Non-communicable Disease Control and Prevention, Chinese Center for Disease Control and Prevention and Institute for Global Neuroscience and Stroke Collaborations,” *Stroke and Vascular Neurology*, vol. 5, no. 3, pp. 211–239, 2020.
[5] W. J. Powers, A. A. Rabinstein, T. Ackerson et al., “Guidelines for the early management of patients with acute ischemic stroke: 2019 update to the 2018 guidelines for the early management of acute ischemic Stroke: a guideline for healthcare professionals from the American Heart Association/American Stroke Association,” *Stroke*, vol. 50, no. 12, pp. E344–E418, 2019.
[6] C. L. Truwit, A. J. Barkovich, A. Gean-Marton, N. Hibri, and D. Norman, "Loss of the insular ribbon: another early CT sign of acute middle cerebral artery infarction," *Radiology*, vol. 176, no. 3, pp. 801–806, 1990.
[7] M. Mitomi, K. Kimura, J. Aoki, and Y. Iguchi, "Comparison of CT and DWI findings in ischemic stroke patients within 3 hours of onset," *Journal of Stroke and Cerebrovascular Diseases*, vol. 23, no. 1, pp. 37–42, 2014.
[8] J. Emberson, K. R. Lees, P. D. Lyden et al., "Effect of treatment delay, age, and stroke severity on the effects of intravenous thrombolysis with alteplase for acute ischaemic stroke: a meta-analysis of individual patient data from randomised trials," *The Lancet*, vol. 384, no. 9958, pp. 1929–1935, 2014.
[9] A. E. Arch, D. C. Weisman, S. G. Coca, K. Nystrom, C. R. Wira, and J. Schindler, "Missed ischemic stroke diagnosis in the emergency department by emergency medicine and neurology services," *Stroke*, vol. 47, no. 3, pp. 668–673, 2016.
[10] J. L. Saver, M. Goyal, A. van der Lugt et al., "Time to treatment with endovascular thrombectomy and outcomes from ischemic stroke: a meta-analysis," *JAMA*, vol. 316, no. 12, pp. 1279–1288, 2016.
[11] N. Hema Rajini and R. Bhavani, "Computer aided detection of ischemic stroke using segmentation and texture features," *Measurement*, vol. 46, no. 6, pp. 1865–1874, 2013.
[12] M. Chawla, S. Sharma, J. Sivaswamy, and L. T. Kishore, "A method for automatic detection and classification of stroke from brain CT images, international conference of the iee engineering in medicine and biology society," in 2009 *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 3581–3584, Minneapolis, MN, USA, 2009.
[13] C. M. Lo, P. Hung, and K. L. Hsieh, "Computer-Aided Detection of Hyperacute Stroke Based on Relative Radiomic Patterns in Computed Tomography," *Applied Sciences*, vol. 9, no. 8, p. 1668, 2019.
[14] P. Lambin, E. Rios-Velazquez, R. T. H. Leijenaar et al., "Radiomics: extracting more information from medical images using advanced feature analysis," *European Journal of Cancer*, vol. 48, no. 4, pp. 441–446, 2012.
[15] J. J. M. van Griethuysen, A. Fedorov, C. Parmar et al., "Computational radiomics system to decode the radiographic phenotype," *Cancer Research*, vol. 77, no. 21, pp. e104–e107, 2017.
[16] N. Hoque, D. K. Bhattacharyya, and J. Kalita, "MIFS-ND: a mutual information-based feature selection method," *Expert Systems with Applications*, vol. 41, no. 14, pp. 6371–6385, 2014.
[17] A. Jovic, K. Brkic, and N. Bogunovic, "A review of feature selection methods with applications," in 2015 *38th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, pp. 1200–1205, Opatija, Croatia, 2015.
[18] W. Kim, J. Park, H. Sheen et al., "Development of Deep Learning Model for Prediction of Chemotherapy Response Using PET Images and Radiomics Features," in 2018 *IEEE Nuclear Science Symposium and Medical Imaging Conference Proceedings (NSS/MIC)*, pp. 1–3, Sydney, Australia, 2018.
[19] A. Usinskas, R. A. Dobrovolskis, and B. Tomandl, "Ischemic stroke segmentation on CT images using joint features," *Informatica*, vol. 15, no. 2, pp. 283–290, 2004.
[20] I. Rekik, S. Allassonniere, T. K. Carpenter, and J. M. Wardlaw, "Medical image analysis methods in MR/CT-imaged acute-subacute ischemic stroke lesion: segmentation, prediction and insights into dynamic evolution simulation models. A critical appraisal," *NeuroImage: Clinical*, vol. 1, no. 1, pp. 164–178, 2012.
[21] M. Petrou and P. Garcia Sevilla, *Image Processing: Dealing with Texture*, John Wiley & Sons, Ltd, 2006.

[22] Y. Shieh, C. Chang, M. Shieh et al., “Computer-aided diagnosis of hyperacute stroke with thrombolysis decision support using a contralateral comparative method of CT image analysis,” *Journal of Digital Imaging*, vol. 27, no. 3, pp. 392–406, 2014.

[23] J. Minnerup, G. Broocks, J. Kalkoffen et al., “Computed tomography–based quantification of lesion water uptake identifies patients within 4.5 hours of stroke onset: a multicenter observational study,” *Annals of Neurology*, vol. 80, no. 6, pp. 924–934, 2016.

[24] J. Ma, Q. Wang, Y. Ren, H. Hu, and J. Zhao, “Automatic lung nodule classification with radiomics approach,” in *Medical Imaging 2016: PACS and Imaging Informatics: Next Generation and Innovations*, vol. 978906, San Diego, California, United States, 2016.

[25] B. Zhang, X. He, F. Ouyang et al., “Radiomic machine-learning classifiers for prognostic biomarkers of advanced nasopharyngeal carcinoma,” *Cancer Letters*, vol. 403, pp. 21–27, 2017.

[26] R. Shiradkar, S. Ghose, I. Jambor et al., “Radiomic features from pretreatment biparametric MRI predict prostate cancer biochemical recurrence: preliminary findings,” *Journal of Magnetic Resonance Imaging*, vol. 48, no. 6, pp. 1626–1636, 2018.