Learning-Based Detection of MYCN Amplification in Clinical Neuroblastoma Patients: A Pilot Study

Xiang Xiang[1(✉)], Zihan Zhang[1], Xuehua Peng[2], and Jianbo Shao[2]

[1] School of Artificial Intelligence and Automation, Huazhong University of Science and Technology, Wuhan, China
[2] Wuhan Children's Hospital Affiliated with Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China
xex@hust.edu.cn

Abstract. Neuroblastoma is one of the most common cancers in infants, and the initial diagnosis of this disease is difficult. At present, the MYCN gene amplification (MNA) status is detected by invasive pathological examination of tumor samples. This is time-consuming and may have a hidden impact on children. To handle this problem, in this paper, we present a pilot study by adopting multiple machine learning (ML) algorithms to predict the presence or absence of MYCN gene amplification. The dataset is composed of retrospective CT images of 23 neuroblastoma patients. Different from previous work, we develop the algorithm without manually segmented primary tumors which is time-consuming and not practical. Instead, we only need the coordinate of the center point and the number of tumor slices given by a subspecialty-trained pediatric radiologist. Specifically, CNN-based method uses pre-trained convolutional neural network, and radiomics-based method extracts radiomics features. Our results show that CNN-based method outperforms the radiomics-based method.

Keywords: Neuroblastoma · MYCN amplification · CT · Radiomics · Convolutional neural network

1 Introduction

Neuroblastoma is one of the most common extracranial solid tumors in infant patients [1]. Despite a variety of treatment options, patients with high-risk neuroblastoma tend to have poor prognoses and low survival. MYCN gene amplification (MNA) is detected in 20% to 30% of neuroblastoma patients [2]. MNA is an important part of the neuroblastoma risk stratification system. It has been proved to be an independent predictor and is related to aggressive tumor behavior and poor prognosis [3]. The MYCN gene with higher amplification multiple

---

indicates that the neuroblastoma may be a more invasive type and its prognosis may be worse. Therefore, MNA patients of any age are "high-risk" groups [1], and the detection of MNA is an essential part of the evaluation and treatment interventions of neuroblastoma.

MYCN gene amplification status is generally detected by invasive pathological examination of tumor samples which is time-consuming and may have a hidden impact on children. Therefore, it is significant to develop a fast and non-invasive method to predict the presence or absence of MNA.

Radiomics [4] is a method to rapidly extract innumerable quantitative features from tomographic images. This allows the transformation of medical image data into high-dimensional feature data. Radiomics is composed of a set of first-order, second-order, and higher-order statistical features on images. Previous studies [6–8] have shown significant relationships between image features and tumor clinical features. For example, Wu et al. [6] first segment primary tumors and extract radiomics features automatically from the ROI. An ML model is then trained with selected features.

Convolutional neural networks (CNN) are under-explored in the prediction of outcomes in neuroblastoma patients. CNN has shown incredible success in image classification tasks [9], and it is a potential approach for processing medical images. Although CNN is primarily driven by large-scale data, transfer learning has shown its effectiveness in training models with small amounts of data [10]. The number of our CT data is limited, and we use a pre-trained CNN model to handle the challenge of lack of medical image data.

In this work, we investigate the radiomics-based method and CNN-based method on a limited dataset. Specifically, we feed radiomics features into multiple ML models to predict the status of MNA. For CNN-based method, we use pre-trained ResNet [5] to extract deep features and predict the label of the data end-to-end. To the best of our knowledge, our method is the first study to try to simplify the annotation process. Specifically speaking, we do not need a pediatric radiologist to manually segment primary tumors which is time-consuming and not practical in clinical applications. Instead, we only need a pediatric radiologist to point out the center point of the tumor and the number of tumor slices in CT images. We crop the ROI images with fixed size and feed them into the model to predict the MNA status. This can greatly reduce the evaluation time of new CT data. Our results demonstrate comparable performance of previous segment-tumor method.

In summary, the contribution of this paper are as follows.

– We propose a novel CNN-based method to predict the presence or absence of MYCN gene amplification of the CT images.
– We greatly simplify the annotation process which makes the prediction process fast and practical, and we have achieved comparable performance with previous works while the evaluation time is greatly reduced.

In the following, we first review related work and the clinical data preparation, then elaborate on radiomics-based and CNN-based methods, and further empirically compare them, with a tentative conclusion followed in the end.

2   Related Work

There is an increasing interest in the prediction of patient outcomes based on medical images [16–21]. Wang et al. [11] propose a CNN-based method to predict the EGFR mutation status by CT scanning of lung adenocarcinoma patients. By training on a large number of CT images, the deep learning models achieve better predictive performance in both the primary cohort and the independent validation cohort. Wu et al. [6] combine clinical factors and radiomics features which are extracted from the manually delineated tumor. The combined model can predict the MNA status well. However, the annotation process makes the evaluation time-consuming. When evaluating new patient images, the method has to annotate the tumor ROI at first. Similarly [6], Liu et al. [7] extract radiomics feature from tumor ROI and apply pre-trained VGG model to extract CNN-based feature. Angela et al. [12] sketch the ROI on the CT images of neuroblastoma, then extract the radiomics features on ROI. With the extracted feature, they develop the radiology model after feature selection to predict the MNA status.

3   Clinical Data Preparation

Dataset. From the medical records, a total of 23 patients with pretreatment CT scans who have neuroblastoma are selected. Each patient has three-phase CT images. Inclusion criteria are (1) age ≤18 years old at the time of diagnosis, and (2) histopathologically confirmed MNA status detection. The number of presence of MNA in the enrolled patients is only two. The rest 21 patients do not have MNA.

Data Preprocessing. As shown in Fig. 1, the unit of measurement in CT scans is the Hounsfield Unit (HU). We first transform it into the gray level. In the CT scans, a pixel spacing may be [2.5, 0.5, 0.5], which means that the distance between slices is 2.5 mm. And the pixel spacing of different CT scans may vary. As a result, we resample the full dataset to a certain isotropic resolution. Then we transfer the CT scans into image format.

The proportion of MNA and non-MNA in the training cohort is highly imbalanced (2:21). The imbalance harms the generalizability and fairness of the model [15]. To tackle this problem, we adopt re-sampling method to augment the MNA CT image data. Specifically, we apply rotation, flipping, noise injection, and gamma calibration transformation techniques to CT images. For non-MNA images, we randomly select transformation techniques to augment the images, and for MNA images, we apply all transformation techniques to balance the dataset.

With the annotation information, we use a fixed-size filter (128 × 128 size) to crop the tumor out of each slice image around the center point of the tumor, and the cropped slice number is identical to the annotated tumor slice number. That ensures the extracted features correspond to the same spatial information across all images.
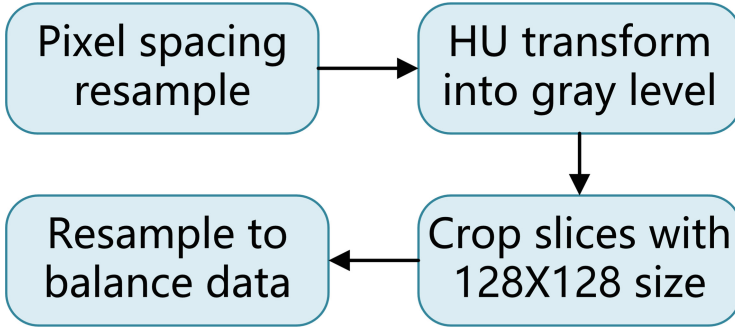
Fig. 1. The pre-processing of the CT data.

4   Proposed Methods

4.1   Radiomics-Based Method

As shown in Fig. 2, primary tumors are annotated from initial staging CT scans using open-source software package ITK-SNAP [13] by a subspecialty-trained pediatric radiologist. We use pyradiomics [14] to extract radiomics features, which is implemented based on consensus definitions of the Imaging Biomarkers Standardization Initiative (IBSI). We extract three kinds of radiomics features as shown in Table 1, 107 features in total. In summary, the first-order statistical features capture the intensity of the images. The shape features describe the geometric shape of the tumor. In our setting, the shape of the tumor is a cube (we do not precisely segment the tumor ROI), which may make the feature not separable because each tumor shape is similar. The gray level features represent the spatial relationship of the voxels.

After the feature extraction, we select the features which are highly correlated to the label. In specific, we apply LASSO linear regression to select the proper features, which reduces the dimension of the data and the number of features and it attenuates over-fitting. Note that we only apply the LASSO linear regression on the training set excluding the test set to prevent data leakage.

Finally, We adopt multiple ML methods to predict the MNA status including SVM, logistic regression, KNN, random forest and AdaBoost. The selected features and the label are the input of the model, and we train the model on CT images of 18 patients while other CT images are used for validation. We use stratified four-fold outer cross-validation to analyze the performance of our models. Note that the slices belonging to one patient could only b divided into training or test set, preventing slices of the same subject are used both for training and testing which invalidates the results.
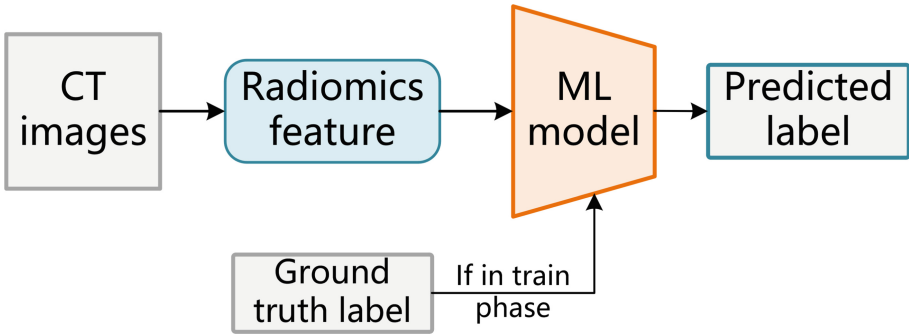
4.2   CNN-Based Method

ResNet is widely used in computer vision. We adopt ResNet34 which is pretrained on ImageNet to predict patient MNA status end-to-end from the CT

Table 1. Extracted radiomics features

First-order statistics	Shape-based	Gray level
Range	2D shape features	GLCM
Maximum		GLDM
Minimum	3D shape features	GLRLM
Mean		GLSZM
Variance		NGTDM
Total 18	Total 14	Total 75

images. We retrain the final layer of ResNet34 to predict the MNA status. To fit in the input size of the model, we crop the images into 128×128 size based on the center point of the tumor to ensure the tumor is at the center of the cropped images. The CT images are gray images while the model requires RGB images which are three-channel. We study three approaches to transform gray images into three-channel images. In specific, the first approach inputs the gray images into the model. The second approach transforms identical gray images into three-channel images. The third approach transforms gray images of adjacent slices into three-channel images.

Fig. 2. Illustration of radiomics-based method process.

5 Experiments

We compare the performance of radiomics-based ML methods and CNN-based methods on our dataset.

Table 2. Accuracy of machine learning models for MNA status prediction. The number in brackets is the variance and the outer number is the mean of four-fold cross-validation accuracy. The best ROC-AUC value is 0.84 (95% CI: (0.81, 0.86))

ML methods	Accuracy
SVM	0.73 ($\pm$0.11)
Logistic regression	**0.74 ($\pm$0.11)**
KNN	0.72 ($\pm$0.09)
Random forest	0.71 ($\pm$0.09)
AdaBoost	0.70 ($\pm$0.12)

5.1   Experimental Results

For all experiments, we split the dataset into training set and validation set and do four-fold cross-validation. The training set contains 18 patients CT images while the validation set contains 5 patients CT images. The accuracy is reported on the total validation set images.

Radiomics-Based Methods. Among the ML techniques we experiment with, logistic regression model over radiomics features outperforms other models predicting MNA status as shown in Table 2. The mean accuracy of logistic regression model is 0.74, 0.01 higher than SVM model. We further present the ROC-AUC value of the best logistic regression model, which is 0.84 (95% CI: (0.81, 0.86)).

CNN-Based Methods. As shown in Table 3, the CNN-based methods outperform the best result of radiomics-based method. The second CNN-based method whose input is synthesized by three identical gray images achieves the best performance 0.79 accuracy and ROC-AUC value 0.87. The ACC result is 0.05 higher than the radiomics-based methods.

Table 3. Mean accuracy of radiomics-based methods and CNN-based methods for MNA status prediction. FS: Feature Selection. ACC: accuracy

Methods-based	ACC	AUC
Radiomics-based	0.72	0.81
FS + Radiomics-based	0.74	0.84
1_{st} CNN-based	0.73	0.85
2_{nd} CNN-based	**0.79**	**0.87**
3_{rd} CNN-based	0.73	0.83

5.2   Discussion

In this study, we investigate multiple methods to predict the MNA status based on the CT scans of neuroblastoma patients. A total of 23 patients are enrolled

with MNA detection report. To the best of our knowledge, there is no such study in the analysis of CT images in neuroblastoma.

Radiomics-Based Methods. In Table 2, we notice that there is no significant difference in the performance of different ML models. The mean accuracy of the logistic regression model is just 0.04 higher than the AdaBoost model. The results reported in [6] are higher than ours because we report the results on the total validation images rather than the patients. Specifically speaking, we test our model on each tumor slice image and report the accuracy rather than test the model on each patient. If the output of our model is the same as the validation patients number, the mean accuracy of our radiomics-based methods is 0.882 which is 0.06 higher than the 0.826 reported in [6]. In addition, we observe a performance promotion of feature selection as shown in Table 3. When using radiomics-based methods without feature selection, the mean accuracy is 0.72 while with feature selection, the mean accuracy is 0.74. That demonstrates the effectiveness of feature selection which helps the model to focus on the important features.

CNN-Based Methods.   Compared to radiomics-based methods, the CNN-based methods achieve higher performance both on the accuracy and AUC. The mean accuracy of CNN-based methods is 0.79 which is 0.05 higher than the best radiomics-based methods.

We notice that, in [7], the results are partly opposite to the results drawn from our experiments. This is likely because of the following reason. [7] uses precisely annotated tumor ROI to extract 3D radiomics features to predict patient outcomes. The 3D features contain more information including the size and shape of the tumor, which helps much to the prediction process. Instead, we do not need the time-consuming segmentation of primal tumors, and CNN focuses on the information of 2D images and performs better on fixed-size images.

As shown in Table 3, we study three approaches to transform gray images into three-channel images. The second approach performs best, and the mean accuracy is 0.06 higher than other approaches. We use the third approach that synthesizes gray images of adjacent slices to three-channel images to capture inter-slices information. However, the performance of this method is worse than the second one. This may be because the original image contains enough information to predict the MNA status.

6   Conclusion

Our study provides insight into that the CNN model has the capability to perform well in the prediction of MNA status of neuroblastoma patient CT scans. In our experiments, the CNN model outperforms multiple radiomics-based ML methods. Different from previous works, we study a much less time-consuming annotation approach which greatly reduces the validation time without manually segmenting primary tumors. We also investigate different approaches to synthesize three-channel images by the original gray images and we find that duplicating the slice image into three-channel images performs better.

Limitation. Due to the computational limitations, we could not perform a study to investigate more CNN models including 3D CNN which may capture inter-slices information better. Also, the radiomics-based methods in our setting is not fully explored. We hope these will inspire future work.

References

1. Cohn, S.L., et al.: The international neuroblastoma risk group (INRG) classification system: an INRG task force report. J. Clin. Oncol. **27**(2), 289 (2009)
2. Ambros, P.F., et al.: International consensus for neuroblastoma molecular diagnostics: report from the International Neuroblastoma Risk Group (INRG) biology committee. Br. J. Cancer **100**(9), 1471–1482 (2009)
3. Campbell, K., et al.: Association of MYCN copy number with clinical features, tumor biology, and outcomes in neuroblastoma: a report from the children's Oncology Group. Cancer **123**(21), 4224–4235 (2017)
4. Lambin, P., et al.: Radiomics: extracting more information from medical images using advanced feature analysis. Eur. J. Cancer **48**(4), 441–446 (2012)
5. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778 (2016)
6. Wu, H., et al.: Radiogenomics of neuroblastoma in pediatric patients: CT-based radiomics signature in predicting MYCN amplification. Eur. Radiol. **31**(5), 3080–3089 (2020). https://doi.org/10.1007/s00330-020-07246-1
7. Liu, G., et al.: Incorporating radiomics into machine learning models to predict outcomes of neuroblastoma. J. Digital Imaging **2**, 1–8 (2022). https://doi.org/10.1007/s10278-022-00607-w
8. Huang, S.Y., et al.: Exploration of PET and MRI radiomic features for decoding breast cancer phenotypes and prognosis. NPJ Breast Cancer **4**(1), 1–3 (2018)
9. He, K., Zhang, X., Ren, S., Sun, J.: Delving deep into rectifiers: surpassing human-level performance on imagenet classification. In: Proceedings of the IEEE International Conference On Computer Vision, pp. 1026–1034 (2015)
10. Weiss, K., Khoshgoftaar, T.M., Wang, D.D.: A survey of transfer learning. J. Big Data **3**(1), 1–40 (2016). https://doi.org/10.1186/s40537-016-0043-6
11. Wang, S., et al.: Predicting EGFR mutation status in lung adenocarcinoma on computed tomography image using deep learning. Eur. Respir. J. **53**(3) (2019)
12. Di Giannatale, A., et al.: Radiogenomics prediction for MYCN amplification in neuroblastoma: a hypothesis generating study. Pediatr. Blood Cancer **68**(9), e29110 (2021)
13. Yushkevich, P.A., et al.: User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. Neuroimage **31**(3), 1116–28 (2006)
14. Van Griethuysen, J.J., et al.: Computational radiomics system to decode the radiographic phenotype. Can. Res. **77**(21), e104–e107 (2017)

15. Zhang, Z., Xiang, X.: Long-tailed classification with gradual balanced loss and adaptive feature generation. arXiv preprint arXiv:2203.00452 (2022)
16. Hosny, A., Parmar, C., Quackenbush, J., Schwartz, L.H., Aerts, H.J.: Artificial intelligence in radiology. Nat. Rev. Cancer **18**(8), 500–10 (2018)
17. Coroller, T.P., et al.: CT-based radiomic signature predicts distant metastasis in lung adenocarcinoma. Radiother. Oncol. **114**(3), 345–50 (2015)
18. Mackin, D., et al.: Measuring CT scanner variability of radiomics features. Invest. Radiol. **50**(11), 757 (2015)
19. Thawani, R., et al.: Radiomics and radiogenomics in lung cancer: a review for the clinician. Lung Cancer **1**(115), 34–41 (2018)
20. Liu, Z., et al.: Radiomics analysis for evaluation of pathological complete response to neoadjuvant chemoradiotherapy in locally advanced rectal cancer. Clin. Cancer Res. **23**(23), 7253–62 (2017)
21. Liu, Z., et al.: The applications of radiomics in precision diagnosis and treatment of oncology: opportunities and challenges. Theranostics **9**(5), 1303 (2019)