2021 IEEE International Conference on Image Processing (ICIP) | 978-1-6654-4115-5/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICIP42928.2021.9506741

KINET: A NON-INVASIVE METHOD FOR PREDICTING KI67 INDEX OF GLIOMA

Xuhui Li[1], Yong Xu[1], Feng Xiang[2], Qing Liu[2], Weihong Huang[3], Bin Xie[1,4,]*

[1]School of Automation, Central South University, Changsha, China
[2]Department of Neurosurgery, Xiangya Hospital of Central South University, Changsha, China
[3]Mobile Health Ministry of Education-China Mobile Joint Laboratory, Changsha, China
[4]Hunan Xiangjiang Artificial Intelligence Academy, Changsha, China. {xiebin@csu.edu.cn}

ABSTRACT

In this paper, a multimodal magnetic resonance imaging (MRI) and heterogeneous metadata (including age, gender) dataset containing 263 patients was established. Based on this dataset, a new multimodal deep neural network (KiNet) was proposed, aiming to effectively predict the Ki67 index in gliomas in a non-invasive way by fusing multimodal MRI features and metadata. We adopted a five-fold cross-validation approach to verify the performance of the network. KiNet achieved results with an AUC of 0.79 and a kappa coefficient of 0.47. The proposed approach's outperformance indicated the feasibility of predicting the Ki67 index in gliomas in a non-invasive way.

Index Terms— Ki67 index, KiNet, multi-modal, glioma, MRI

1. INTRODUCTION

As the most deadly tumor in central nervous system [1], glioma's several important molecular markers (including isocitrate dehydrogenase mutation status and 1p/19q co-deletion status, etc.) had been identified by the World Health Organization (WHO) through the genomic analysis of glioma. These markers can make great effect on histopathological classification and subtype identification, so as to guide doctors to make clinical decisions for glioma patients [2]. Among them, the Ki67 index can be used clinically to measure the degree of tumor proliferation and predict the patient's survival and recurrence [3]. Generally speaking, high Ki67 index corresponds to shorter survival and higher recurrence rate [4, 5]. Whereas different Ki67 indexes often correspond to different treatment programs, and it will be of great guiding significance if the Ki67 index can be obtained before the surgery [6]. During this process of determining the Ki67 index, it usually requires to sample the tumor tissue, stain the tissue using a Ki67 antibody and then calculate the proportion of stained cells [7]. However, it might be a big challenge to obtain tumor-rich tissue samples to calculate the Ki67 index, as a report from The Cancer Genome Atlas (TCGA) indicated that, only 35% of biopsy samples contain sufficient tumor components which allow accurate molecular characterization experiments [8]. Therefore, it is a really good news for relevant patients to develop a non-invasive Ki67 index identification method.

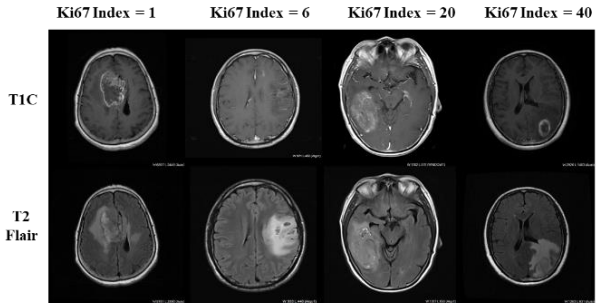

Figure 1. The multi-modal MRI images of patients with different Ki67 indexes.

The most mainstream non-invasive glioma diagnostic method is multi-parameter MRI. Figure 1 showed the multi-modal MRI images of patients with different Ki67 indexes. It is convenient for professional physicians to visually compare the MRI images of T1C and T2Flair to understand the details of the tumor, including its size and location. However, it is difficult for them to distinguish the differences in MRI images of tumors with different Ki67 indexes.

In recent years, the identification of molecular markers based on non-invasive machine learning has developed very rapidly. In 2014, Gevaert *et al*. demonstrated that imaging has the potential to predict the clinical presentation and molecular markers of tumors non-invasively [9]. Chang K *et al*. non-invasively predicted IDH mutation status of gliomas based on conventional MRI images by deep learning with an accuracy of 94% [10]. Chandan *et al*. designed a three-dimensional U-Net for classifying the 1p/19q co-deletion status of gliomas with an accuracy of 93.46% [11]. Matsui *et al*. demonstrated that the simultaneous use of multimodal heterogeneous information including MRI and patient's age and gender can achieve higher accuracy in predicting IDH mutation status [12]. For multi-modal fusion learning tasks, Valentin *et al*. hypothesized that each modality can be processed by a separated deep convolutional network, allowing decisions to be made independently of each modality. The characteristics

of each modality were fused step by step in the central network to improve the accuracy of multi-modal fusion methods by introducing multi-task learning [13]. As few prediction studies on Ki67 could be found, Fuyong Xing *et al*. designed a weakly supervised semantic segmentation network for predicting Ki67 index in tumors based on pathological images after staining [14], but no study has been found to predict Ki67 index using non-invasive information.

In this study, a novel multi-modal deep learning network (KiNet) was proposed for non-invasive prediction of Ki67 index. Our approach has implemented the task of non-invasive prediction of Ki67 index for the first time by effectively fusing multimodal MRI and metadata.

2. DATASETS

The data of this study came from patients who visited the Department of Neurosurgery of Xiangya Hospital of Central South University in recent three years. These medical data contained the basic information metadata of the patients, MRI images of two modalities (T1C and T2Flair combined 11,894 images were chosen) in the horizontal position, and Ki67 index results of postoperative pathological tests of all the included 263 patients who were diagnosed with glioma by postoperative pathological diagnosis. Among all these cases with an age range of 13 to 75 and an average age of 45, 113 patients were males. 140 cases were of high Ki67 indexes and 123 cases were of low Ki67 indexes. Initially, we considered Ki67 index prediction as a regression task, but the clinically acquired Ki67 index itself was a statistical value, so we summarized it as a classification task. According to the WHO guidelines, Ki67 indexes less than 10% are called low Ki67 indexes, which was labeled as 0, and the rest are called high Ki67 indexes, labeled as 1 [15].

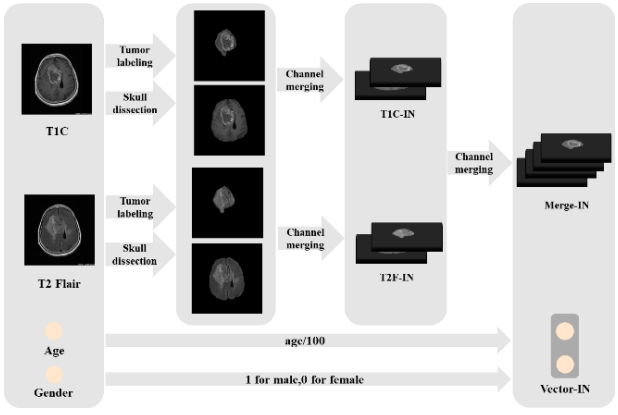

Figure 2. Preprocessing of multi-modal heterogeneous data.

Figure 2 showed the data preprocessing procedures. To eliminate the interference of non-brain tissues such as skull, firstly the brain MRI image was obtained by performing skull stripping on the head MRI image, which was labeled with tumor location by a physician using Labelimg. Then, the brain MRI images of each modality and the labeled tumor images were merged into the two-channel images with shape of (224, 224, 2), which were called T1C-IN and T2F-IN, respectively. In addition, the images of T1C-IN and T2F-IN were merged into a four-channel image with shape (224, 224, 4), called Merge-IN. In order to enhance the generalization performance of the algorithm, the image was translated, rotated, and flipped for data enhancement. Besides, the mean variance normalization operation was also performed before training. For the metadata of the patient, the gender and age were encoded as a 2×1 vector (gender male was recorded as 0, while gender female was recorded as 1, and age was converted into a floating-point number between 0 and 1 by dividing 100), called Vector-IN.

3. METHODS

Multimodal approaches have been used as a key technique when performing correlational studies based on medical images. The fusion of multimodal can correlate different information characteristics from multiple modalities (such as multi-modal images generated by multi-parameter MRI of glioma in this paper), so as to form a better model decision than that based on single modality.

In the existing studies, multimodal fusion is usually divided into early fusion, late fusion and hybrid fusion with in which layer the model is fused. However, scholars have not been able to agree on which layer of the model to be fused can bring the best results. Concatenation, element-wise products and other methods are often used to map multi-modal information features into the same multi-modal model spatial dimension [16], which can be regarded as learning a joint representation. Unlike this, coordinated representations enjoy more complementariness and can maximize the relevance of unearthing different multi-modal data information representations.

CentralNet borrowed on two types of representations at the same time, and the convolutional neural networks in the existing depth which were used to independently process the model of a single modality. The model introduced a central network proposed to connect the modality specific networks, so as to map the representations of different modalities to the same common space [13].

3.1. Framework

In this paper, inspired by the CentralNet, KiNet was proposed to predict Ki67 index. The model contains two independent auxiliary branches as well as a master branch, with each branch containing four feature extraction stages and each of stage being composed of several blocks. Such blocks adopt a similar structure to ResNeSt block [17], as shown in Figure 3. After the fourth stage, the classification results are output through global average pooling and fully-connected layer. Especially, for master branch, multi-scale classification

results are obtained by adding fully-connected layer after the feature map is output in each stage.

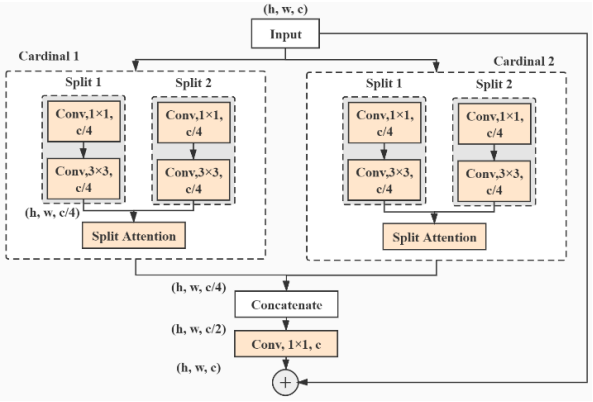

Figure 3. The architecture of KiNet block. Among (h, w, c) is the height, width, channel of the image input

Figure 4 showed the overall framework of the model. The feature maps of T1C-IN, T2F-IN and Merge-IN after passing through the shallow feature extractor are used as input for master branch after passing through concatenate. The input to the second stage of master branch is obtained by the output of the previous stage spliced by channel with the output of the first stage of the auxiliary branch model. The final decision can be obtained by passing the master branch multi-scale output with the metadata vector concatenate through a fully-connected layer. The code of the KiNet is available at https://github.com/XuYongi/KiNet.

3.2. Loss function

A multi-task multi-modal network was constructed for the classification of Ki67 index, including the prediction task of the independent Auxiliary branch as well as the multi-scale prediction task of master branch. Next the loss function in the training phase of the network is introduced.

The ki67 indexes prediction tasks are deemed as a two-classification task by selecting the cross-entropy loss function, as shown in Eq.(1), y is the label of Ki67 index, while p represents the prediction label:

$$loss = -(y\,log(p) + (1-p)\,log(1-p)) \quad (1)$$

The loss of Master Branch is defined, as shown in Eq.(2):

$$loss_{master} = \beta_1 \times loss_{s1} + \beta_2 \times loss_{s2} + \beta_3 \times loss_{s3} + \beta_4 \times loss_{s4} + \beta_5 \times loss_{cat} \quad (2)$$

The total loss function of the model Loss is defined, as shown in Eq.(3):

$$Loss = \alpha_1 \times loss_{T1C} + \alpha_2 \times loss_{T2F} + \alpha_3 \times loss_{master} \quad (3)$$

Where $loss_{T1C}$、$loss_{T2F}$ are the losses of two independent auxiliary branches, while $loss_{si}$ is the loss of scale output extracted by $stage_i$ corresponding to master branch, and $loss_{cat}$ is the loss of output after master branch multi-scale classification and metadata concatenating. In our experiments, α_i (i = 1, 2, 3) = 0.35, 0.25, 0.4, β_j (j = 1, 2, 3, 4, 5) = 0.025, 0.025, 0.2, 0.35, 0.4.

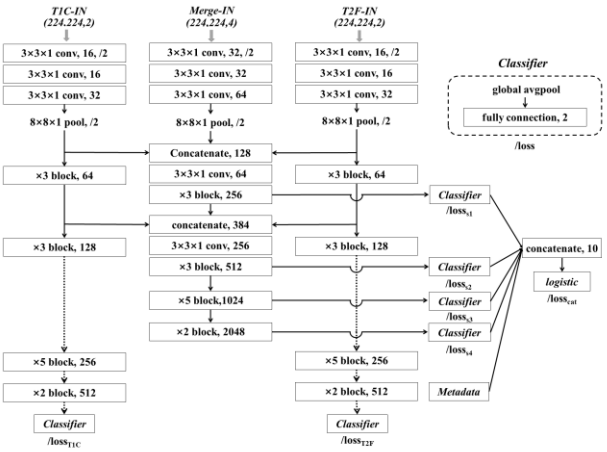

Figure 4. The overall framework of KiNet (dotted markers are not considered during the testing phase).

4. RESULTS AND DISCUSSION

A five-fold cross-validation method was introduced to the experiment on basis of the collated dataset. Table 1 showed the performance comparison between KiNet and mainstream network architecture in the Ki67 index prediction task. ResNet-50 [18] and ResNeSt-50 both used the early fusion mode of image data superposed through channels in different modalities, that is, the input of the network is Merge-IN [19]. Figure 5 showed the ROC curves of different neural networks for Ki67 index.

Table 1. Predictive performance of different neural networks for Ki67 index.

Model	Accuracy	Se	Sp	AUC	Kappa
ResNet-50	0.691	0.693	0.689	0.724	0.422
ResNeSt-50	0.724	0.734	0.715	0.764	0.451
CentralNet	0.725	0.745	0.705	0.762	0.450
KiNet	**0.752**	**0.755**	**0.749**	**0.788**	**0.472**

It can be seen from the table data that, CentralNet performed better for predicting ki67 indexes by using a central network to fuse two auxiliary branch multi-level features compared with simple modal channel superposition. The network architecture proposed by KiNet was more conducive to the expression of relatedness between different levels of feature information, so that it outperformed the mainstream network in this task.

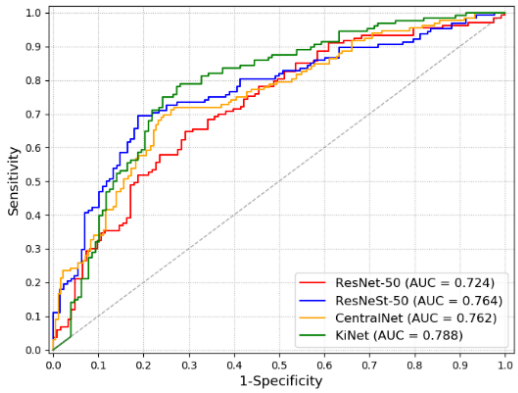

Figure 5. the ROC curves of different model for Ki67 index.

To visually demonstrate the role of each sub-module of KiNet, ablation experiments were designed to verify the effect of sub-modules (multi-scale loss, Merge-IN, Meta-data) in KiNet on model performance, in which a five-fold cross-validation training method was employed. The results of ablation experiments of KiNet are given in Table 2.

Table 2. Predictive performance of KiNet ablation experiments.

Num	Model	Multi-scale loss	Merge-IN	Meta-data	Acc (%)	AUC
1	KiNet	0	0	0	73.5	0.766
2	KiNet	0	0	1	73.6	0.768
3	KiNet	0	1	0	74.2	0.773
4	KiNet	0	1	1	74.3	0.774
5	KiNet	1	0	0	74.6	0.778
6	KiNet	1	0	1	74.8	0.781
7	KiNet	1	1	0	75.0	0.785
8	KiNet	1	1	1	75.2	0.788

0/1 denotes without/with.

According to the data in the table above, Experiment 8 was of the optimal prediction performance, which validated the rationality of the network structure design in KiNet again. Experiments 1 and 5 (2 and 6, 3 and 7, 4 and 8) demonstrated the effectiveness of multi-scale losses. Losses at different scales presented a lot of hidden image noise and feature, which would facilitate the model robustness enhancement and feature learning. After repeated tuning experiments, it was finally determined that the optimal network performance was achieved when the multi-scale loss weights β_1~β_5 were set to 0.025, 0.025, 0.2, 0.35 and 0.4, respectively. To some extent, it made clear that the deep characteristics would exert more effects on the prediction results of the model. Through comparing the results of Experiments 1 and 3 (2 and 4, 5 and 7, 6 and 8), it was found that the addition of additional Merge-IN could be conducive to Ki67 index prediction in KiNet. Merge-IN could effectively realize representation fusion among different modalities by directly combining the original images. By this input, the model could be easier to learn the correlation of shallow feature information such as location and gray level of different modality images. It was discovered in Experiments 1 and 2 (3 and 4, 5 and 6, 7 and 8) that the performance improvement brought about by metadata including the gender and age of the patients was less than 0.2%. The result indicated that the gender and age of the patients made smaller improvement in our model for the specific task of predicting the Ki67 index for glioma patients. Physician explained the reason for this may lie in the wide range of glioma populations, and the gender factor has little effect on the proliferation ability of glioma.

In this paper, the master branch structure was finally established through multiple exploratory experiments. Under the preconceived influence of CentralNet, the element-wise sum operation with weights was firstly considered as the representation fusion method, in which the weights were learnable parameters. To be specific, the weights of element-wise sum were initialized as uniform distribution for the ki67 index prediction task. Parameter visualization was performed on the trained model to determine the channel ratio of our concatenate. As for the input at the last two stages, we abandoned the connection of auxiliary branch to streamline our network. For the loss weight of each branch, the Ki67 index was predicted on the independent auxiliary branch to set a greater loss weight for T1C auxiliary branch, respectively. After multiple debugging parameters, the loss weights α_1~α_3 were finally set as 0.35, 0.4, 025 (T1C auxiliary branch, master branch, T2Flair auxiliary branch). Compared with networks using element-wise sum operation for multimodal feature fusion, KiNet's fusion mode could enhance the performance for more than 1%.

5. CONCLUSION

In this paper, a novel multi-modal deep learning network (KiNet) was proposed to predict Ki67 index of gliomas non-invasively based on heterogeneous information including multi-modal MRI and metadata. Two auxiliary branches were introduced in the network for feature extraction of two modal MRI images. Multi-level auxiliary branch features and metadata are integrated in master branch, and a multi-scale loss was used to supervise the network training process. For the first time, the network predicted the Ki67 index of glioma non-invasively and achieved results with an AUC of 0.79 and a kappa coefficient of 0.47. With this method, physicians are promising to obtain the Ki67 index of gliomas and develop a more effective treatment plan before surgery.

6. ACKNOWLEDGMENTS

This work was supported by the National Key Research and Development Program of China under Grant 2017YFC0909901 and Mobile Health Ministry of Education-China Mobile Joint Laboratory under Grant 2020MHL02009.

7. REFERENCES

[1] Lapointe S, Perry A, Butowski N A. Primary brain tumours in adults[J]. The Lancet, 2018, 392(10145): 432-446.

[2] Eckel-Passow J E, Lachance D H, Molinaro A M, et al. Glioma groups based on 1p/19q, IDH, and TERT promoter mutations in tumors[J]. New England Journal of Medicine, 2015, 372(26): 2499-2508.

[3] Xing F, Su H, Neltner J, et al. Automatic ki-67 counting using robust cell detection and online dictionary learning[J]. IEEE Transactions on Biomedical Engineering, 2013, 61(3): 859-870.

[4] Arshad H, Ahmad Z, Hasan S H. Gliomas: correlation of histologic grade, Ki67 and p53 expression with patient survival[J]. Asian Pac J Cancer Prev, 2010, 11(6): 1637-1640.

[5] Torp S H , Johannesen E , Lindboe C F . Comparison of different Ki67 antibodies in human glioblastomas[J]. Clinical Molecular Pathology, 1995, 48(4):M191.

[6] Li L T, Jiang G, Chen Q, et al. Ki67 is a promising molecular target in the diagnosis of cancer[J]. Molecular medicine reports, 2015, 11(3): 1566-1572.

[7] Reid M D, Bagci P, Ohike N, et al. Calculation of the Ki67 index in pancreatic neuroendocrine tumors: a comparative analysis of four counting methodologies[J]. Modern Pathology, 2015, 28(5): 686-694.

[8] Cancer Genome Atlas Research Network. Comprehensive genomic characterization defines human glioblastoma genes and core pathways[J]. Nature, 2008, 455(7216): 1061.

[9] Gevaert O, Mitchell L A, Achrol A S, et al. Glioblastoma multiforme: exploratory radiogenomic analysis by using quantitative image features[J]. Radiology, 2014, 273(1):168-174.

[10] Chang K, Bai H X, Zhou H, et al. Residual convolutional neural network for the determination of IDH status in low-and high-grade gliomas from MR imaging[J]. Clinical Cancer Research, 2018, 24(5): 1073-1081.

[11] Bangalore Yogananda C G, Shah B R, Vejdani-Jahromi M, et al. A novel fully automated MRI-based deep-learning method for classification of IDH mutation status in brain gliomas[J]. Neuro-oncology, 2020, 22(3): 402-411.

[12] Zhang B, Chang K, Ramkissoon S, et al. Multimodal MRI features predict isocitrate dehydrogenase genotype in high-grade gliomas[J]. Neuro-oncology, 2017, 19(1): 109-117.

[13] Vielzeuf V, Lechervy A, Pateux S, et al. Centralnet: a multilayer approach for multimodal fusion[C]//Proceedings of the European Conference on Computer Vision (ECCV). 2018: 0-0.

[14] Xing F, Cornish T C, Bennett T, et al. Pixel-to-pixel learning with weak supervision for single-stage nucleus recognition in Ki67 images[J]. IEEE Transactions on Biomedical Engineering, 2019, 66(11): 3088-3097.

[15] Louis D N, Perry A, Reifenberger G, et al. The 2016 World Health Organization classification of tumors of the central nervous system: a summary[J]. Acta neuropathologica, 2016, 131(6): 803-820.

[16] Tao C, Wu W, Xu C, et al. Multi-representation fusion network for multi-turn response selection in retrieval-based chatbots[C]//Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining. 2019: 267-275.

[17] Zhang H, Wu C, Zhang Z, et al. Resnest: Split-attention networks[J]. arXiv preprint arXiv:2004.08955, 2020.

[18] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 770-778.

[19] Li S, Kang X, Fang L, et al. Pixel-level image fusion: A survey of the state of the art[J]. information Fusion, 2017, 33: 100-112.