

A Pilot Study: Deep Multi-Instance Learning for Origin Tracing of Brain Metastases

Hui Yu

Sichuan University

Zhongzhou Zhang

Sichuan University

Ziyuan Yang

Sichuan University

Tao Wang

Sichuan University

Zhiwen Wang

Sichuan University

Zhongxian Wang

Sichuan University

Lunxin Liu

West China Hospital of Sichuan University

Yan Liu

liuyan77@scu.edu.cn

Sichuan University

Yi Zhang

Sichuan University

Research Article

Keywords: Brain metastasis, Origin tracing, Image classification, Deep learning, multi-instance learning

Posted Date: March 27th, 2024

DOI: <https://doi.org/10.21203/rs.3.rs-4137880/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

A Pilot Study: Deep Multi-Instance Learning for Origin Tracing of Brain Metastases

Hui Yu¹, Zhongzhou Zhang¹, Ziyuan Yang¹, Tao Wang¹,
Zhiwen Wang¹, Zhongxian Wang¹, Lunxin Liu², Yan Liu^{3*},
Yi Zhang⁴

¹College of Computer Science, Sichuan University, No.24 South Section 1, Yihuan Road, Chengdu, 610065, Sichuan, China.

²Department of Neurosurgery, West China Hospital of Sichuan University, No. 37 Guoxue Lane, Wuhou District, Chengdu, 610044, Sichuan, China.

³College of Electrical Engineering, Sichuan University, No.24 South Section 1, Yihuan Road, Chengdu, 610065, Sichuan, China.

⁴School of Cyber Science and Engineering, Sichuan University, Chuanda Road, Shuangliu County, Chengdu, 610207, Sichuan, China.

*Corresponding author(s). E-mail(s): liuyan77@scu.edu.cn;

Abstract

Treatment decisions for brain metastasis heavily rely on identifying the primary site, which is typically accomplished through biomarker-based techniques such as genomics and histopathology. However, limited healthcare resources sometimes can hinder their availability. Therefore, we innovatively transform origin tracing into an image classification task. Based on T1ce-MRI, we develop a non-invasive and cost-effective pipeline, called deep multi-instance learning (DMIL). The DMIL-based pipeline includes three steps: pre-processing, training and testing. Particularly, in pre-processing, mix-modal data decoration is proposed to learn multiple modal knowledge. For DMIL training, center-point-based lesion identification is employed to automatically crop ROIs, eliminating the need for manual intervention. Additionally, self-adaptive lesion classification aims to achieve slice-wise origin tracing. During the inference stage, to address the uncertainty stemming from heterogeneity within a patient's volume, we design a voting majority mechanism to make final patient-wise predictions. Evaluated on the clinical dataset, our DMIL-based pipeline demonstrated promising results. The best

patient-wise results achieved at 87.27% (accuracy), 85.00% (PPV) and 83.33% (sensitivity).

Keywords: Brain metastasis, Origin tracing, Image classification, Deep learning, multi-instance learning

1 Introduction

Brain metastasis (BM) is a type of malignant tumor and has a high incidence ($> 20\%$) in cancer patients [1]. It refers to the spread of cancer cells from other parts of the body, known as primary sites, to the brain. There is a significant variance in the sensitivity of different primary sites to chemotherapy or radiotherapy [2, 3]. Therefore, in clinical practice, the identification of the primary site is crucial for BM patients, which greatly impacts the subsequent treatment and management. However, up to 15% of BM patients have cancers of unknown primary (CUP) because the primary tumor is too small or obscured [3]. It is imperative that the medical community places a high level of emphasis on accurate origin tracing.

Most existing origin tracing methods are biomarker-based techniques deduced from genetics or histopathology. Genomic testing is not always widely adopted in low-resource settings, and it can indeed impose a significant economic burden on patients [4]. In response to the limitations of genomic testing, research institutions explore an alternative approach based on easily acquired histology slides [5, 6]. However, it is challenging for invasive histopathology to obtain tissue samples in complex brain structures. Advances in imaging technology greatly improve the accessibility and effectiveness of radiology. Recently, non-invasive radiology has emerged as a promising solution for origin tracing. **In western world**, positron emission tomography (PET)/computed tomography (CT) is the most commonly used imaging for diagnosis of primary cancer. For example, [7] used PET/CT to trace the origins of BMs with an accuracy of 53%. However, the ^{18}F -FDG tracer is limited and very expensive, which may render the PET/CT exam unavailable in many developing countries. Therefore, we start an attempt to develop a non-invasive BM tracing method based on cost-effective routine imaging examinations.

Actually, we have noted that each metastasis slice contains one lesion at least. The lesion serves as a distinguishing feature between BM patients and normal individuals. Upon observation and drawing inspiration from multi-instance learning (MIL) [8], we introduce a novel resolution to origin tracing. MIL has two core concepts: bag and instance. where a bag comprises of multiple instances. If a bag contains at least one positive instance, the entire bag will be labeled as "positive" regardless of the presence of other negative instances. This indicates that MIL aims to flexibly learn a more discriminative data representation. Therefore, in this study, when we consider a metastasis slice as a positive bag, a lesion appearing in it can be treated as a positive instance. Moreover, we extract the lesion-based region of interest (ROI) and trace its origin, which further represents the primary of corresponding slice.

Tracing the origin on ROIs in radiology images can be formulated as an image classification task. Thanks to the outstanding ability to automatically extract data features, deep learning (DL) techniques, especially convolutional neural networks (CNNs), have achieved significant success in image classification tasks [9]. Some attempts are also made to classify the primary site. To differentiate metastatic lesions in the spine originated from primary lung cancer and other cancers, [10] employed the ROIs from dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) with three parametric maps as independent inputs in a convolutional long short term memory (CLSTM) network. For ROI extraction, [10] combine manual delineation with region growing to segment the lesion. In [11], to classify primary bone tumor, the ROIs in pre-operative radiographs were manually cropped and fed into the EfficientNet-B0 architecture. Recently, [12] applied the ResNet18 network with CAM attention to classify manual ROIs across multiple MRI sequences, aiming to trace the origin of BM. In contrast, our study introduces an automatic algorithm for lesion-based ROI identification. Additionally, the values of different DL models in identifying multiple types of primary sites remain unclear. Therefore, we explore various CNN architectures to find an optimal model for this task.

In this paper, we investigate a clinically applicable pipeline based on contrast-enhanced T1-weighted MRI (T1ce-MRI), named deep multi-instance learning (DMIL). DMIL first automatically identifies lesion-based ROIs that clinicians typically focus on during the feature learning process. Then, DMIL classifies the primary site type of the ROI. The major contributions of this paper are summarized as follows:

- We introduce a center saliency map generation algorithm to locate potential lesion-based ROI. Furthermore, we adopt a self-adaptive threshold algorithm to choose the largest ROI as the positive instance representing the corresponding slice for origin tracing.
- To leverage multi-modal knowledge, we design a data decoration method in data pre-processing that complements our T1ce-MRI dataset with other MRI sequences. This incorporation of multi-perspective information aims to enhance the model performance.
- Our DMIL framework can work with any CNN encoder in both lesion identification and lesion classification stages. We explore different CNN networks to identify the optimal one for this task. This flexibility enhances the practical application of our framework.

2 Approach

This retrospective study was approved by the Sichuan University Medical Ethics Committee in accordance with the Declaration of Helsinki. The informed consent was obtained from all human participants.

As shown in Fig. 1, the origin tracing pipeline involves three stages: pre-processing, training and testing. First, to enrich multi-modal information, we perform mix-modal data decoration (MDD) in pre-processing. Then, during the training phase, the deep multi-instance learning (DMIL) framework is constructed for slice-wise prediction,

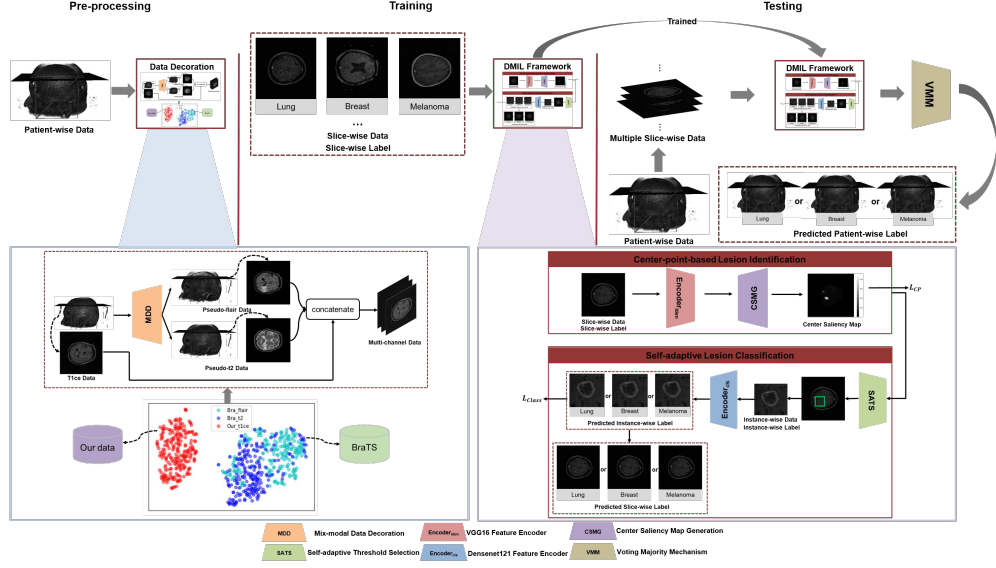


Fig. 1 The deep multi-instance learning based pipeline, which involves pre-processing, training and testing stages.

which includes center-point-based lesion identification and self-adaptive lesion classification. When inferring, a voting majority mechanism (VMM) is considered to avoid heterogeneity misleading and produce the final patient-wise result for origin tracing. The details will be described below.

2.1 Mix-modal Data Decoration

For brain metastases, the acquisition of multiple modalities is costly in terms of time, resources and patient comfort. By leveraging multiple MRI sequences from the BraTS dataset [13], such as FLAIR and T2-weighted (T2), we can obtain complementary information about the brain’s anatomy and pathology. Therefore, inspired by [14], we design a mix-modal data decoration (MDD) method to enhance the richness and complexity of our T1ce data. Specifically, MDD mixes the data statistics of our T1ce dataset with those of FLAIR and T2 datasets in BraTS. Given a BM patient volume X_{T1ce} and a random BraTS patient with two volumes of X_{FLAIR} and X_{T2} . MDD first separately computes the T1ce, FLAIR and T2 data statistics (see Eq. (1) and Eq. (2)). Then, the FLAIR and T2 data statistics are applied to X_{T1ce} to generate the mix-modal data X_{FLAIR} and X_{T2} (see Eq. (3) and Eq. (4)). Finally, they are concatenated with the original X_{T1ce} in the channel dimension to form a new input data (see Eq. (5)).

$$\alpha_k = \mu(X_k) \quad (1)$$

$$\beta_k = \sigma(X_k) \quad (2)$$

where $\mu(\cdot)$ and $\sigma(\cdot)$ represent mean and standard deviation functions. $k \in [T1ce, FLAIR, T2]$.

$$MDD_{FLAIR} = \beta_{FLAIR} \frac{X_{T1ce} - \alpha_{T1ce}}{\beta_{T1ce}} + \alpha_{FLAIR} \quad (3)$$

$$MDD_{T2} = \beta_{T2} \frac{X_{T1ce} - \alpha_{T1ce}}{\beta_{T1ce}} + \alpha_{T2} \quad (4)$$

$$X = cat(MDD_{FLAIR}, X_{T1ce}, MDD_{T2}) \quad (5)$$

2.2 Center-point-based Lesion Identification

Since BM usually has a solid sphere-like shape, the lesion can be identified via its geometric center, called center point. Following our center point detection strategy [15], we present a center saliency map generation (CSMG) algorithm to detect the center point. The center saliency map indicates the probability of each pixel being the center point, which can be calculated from a feature map by any feature encoder. To determine whether a pixel is a center-point or not, CSMG first obtains the maximum values in both horizontal and vertical directions through a max-pooling operation, and then adds them together. In each direction, the max-pooling operation is further divided into two mini-max-pooling operations (i.e., left-maxing and right-maxing for horizontal direction, bottom-maxing and top-maxing for vertical direction). Fig. 2 illustrates an example of CSMG.

2.3 Self-adaptive Lesion Classification

The key idea of MIL is to construct the high-quality instance-wise dataset. Different from the traditional instance selection criteria [8], we introduce a new instance generation algorithm, named self-adaptive threshold selection (SATS). After CSMG, metastatic lesions are identified as positive instances. Among multiple positive instances, SATS selected the most discriminative one. SATS is described as follows: 1) Given a center saliency map $H_{ct}^{H \times W}$, it's up-sampled to $H_{ct'}^{H' \times W'}$ with the original slice resolution. 2) A dynamic threshold T is adopted to obtain the binary mask $H_{ct''}^{H'' \times W''}$. We set $T = 0.5 \times H_{ct'_{max}}$, where $H_{ct'_{max}}$ represents the maximum value in $H_{ct'}$. 3) For $H_{ct''}$, the connected component analysis is utilized to find the largest lesion-based bounding box B . 4) B is amplified to produce the final detected box B' . 5) According to B' , a lesion-based ROI is cropped from the metastatic slice, as the selected instance data to be behalf of the corresponding tag (i.e., metastatic slice). Fig. 3 shows a visual processing of SATS. Once the instance-wise dataset is prepared, it will be fed into the deep classifier in a fully supervised manner.

2.4 Voting Majority Mechanism

BM often exhibits the significant intratumoral heterogeneity, which means that within the tumor, the malignant regions are not uniformly distributed [16]. If the deep feature extracted from a single slice lacks critical information, it may will result in error decisions. Additionally, the initial and final slices of a lesion are also difficult to accurately

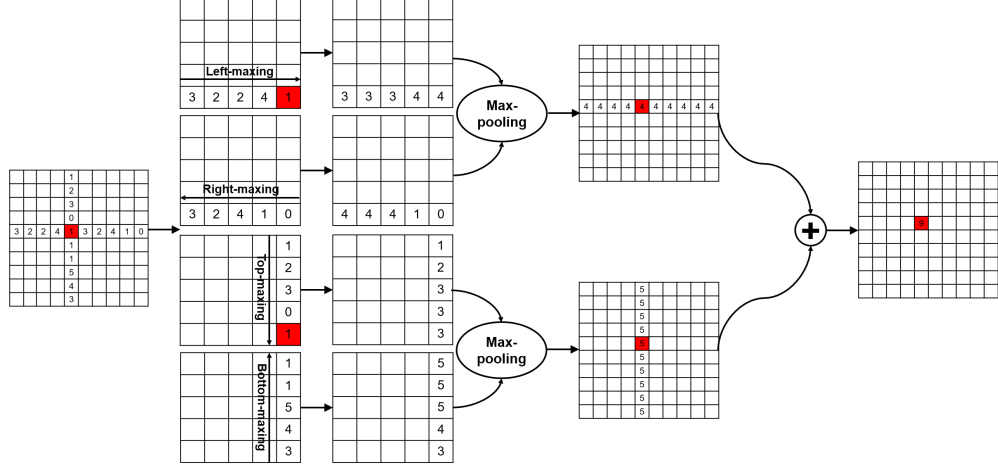


Fig. 2 An example of center saliency map generation. The dynamic programming is implemented as follows: 1) Given a feature map F with the resolution of $H \times W$, the F_{ij} indicates the feature value at location (i, j) ($i \in [0, W-1], j \in [0, H-1]$). 2) We perform left-maxing and right-maxing operations. Particularly, the left-maxing operation takes the maximum values between $(i+1, j)$ and $(W-1, j)$ to find a local feature maximum l_{ij} . Then, the right-maxing operation max-pools the values from $(0, j)$ to (i, j) to generate a local feature maximum r_{ij} . We further apply an element-wise max operation between l_{ij} and r_{ij} to obtain the maximum value in horizontal direction h_{max} . 3) Similarly, we perform top-maxing and bottom-maxing operations to find the maximum value in vertical direction v_{max} . 4) We add h_{max} and v_{max} together to replace the original feature value F_{ij} for center saliency map prediction.

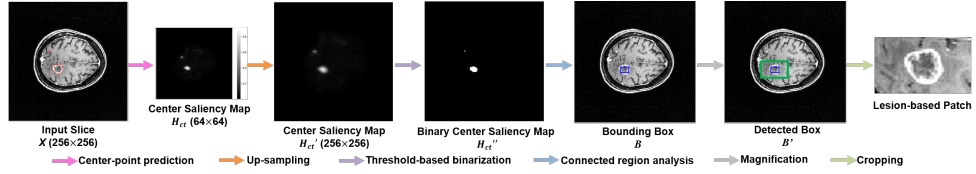


Fig. 3 A visual processing of self-adaptive threshold selection algorithm. The red curves contour the brain metastases. The blue box is the largest lesion-related bounding box after connected region analysis. The green one represents the final detected ROI after applying an amplification operation, which serves as the positive instance.

identify and delineate, and this can have a significant impact on slice-wise classification results. Therefore, we propose a voting majority mechanism (VMM) to make agreements among slices for the final patient-wise inference. Specifically, given a patient volume, we first obtain slice-wise results via DMIL. Then, we count the slice number of each predicted origin class. Finally, we vote for the class with the maximum slice number to determine the patient-wise label. This follows an assumption that each BM patient only has one primary site. In the event of equal voting counts, we will compare the corresponding average confidence scores and choose the highest one.

Table 1 Patients’ demographics and primary tumore types in the training, validating and testing groups.

parameter	training group	validating group	testing group	total
No. of patients	84	10	11	105
Gender (male/ female)	31/ 53	5/ 5	5/ 6	41/ 64
Mean age	56±12	56±14	57±10	56±12
1.5T	81	9	8	98
3T	3	1	3	7
Lung	28	4	4	36
Breast	28	3	4	35
Melanoma	28	3	3	34



Fig. 4 Image examples represent axial metastasis slices originated from lung (63-year-old woman), breast (54-year-old woman) and melanoma (64-year-old man), respectively.

3 Experiments and Results

3.1 Dataset and Processing

We retrospectively collected a total of 240 BM patients treated with Gamma Knife SRS from March 2015 to September 2018. A sequential inclusion criterion was employed to select appropriate cases. First, only patients with brain metastases originating from lung, breast, or melanoma were included, as these cancer types were more likely to metastasize to the brain [2]. Then, we performed random under-sampling to balance the classes of patients. Finally, we collected a dataset of 105 BM patients (lung: 36, breast: 35 and melanoma: 34). They were randomly split into training (84 patients), validating (10 patient) and testing (11 patients) sub-datasets without any overlapping. Tab. 1 summarizes the demographics of the dataset. Here each patient volume was further divided into a sequence of 2D slices along the axial direction to construct DMIL. A total of 3194 slices containing metastatic lesions were collected. Fig. 4 shows some examples of metastasis slices originated from lung, breast and melanoma. The data was normalized to the signal intensity range of 0-1. For data augmentation, we performed random flip and rotation operations.

DMIL includes lesion identification and classification tasks. Thus, there are two kinds of labels. The one is the center saliency map generated from the delineated contours via the unnormalized gaussian function [17]. The other one is the primary tumor label with histologically confirmed.

3.2 Training

We experimentally used Adam optimizer for center-point-based lesion identification and SGD optimizer for self-adaptive lesion classification. The initial learning rates were set to 2.5×10^{-4} and 1×10^{-3} , and they were reduced by 20% if there was no improvement in validation dataset after ten epochs.

3.2.1 Feature Encoder

Our DML framework is compatible with any deep CNN architecture. To establish DMIL, we explored various architectures as the feature encoders in lesion identification and classification, denoted as $Encoder_{iden}$ and $Encoder_{cls}$ in Fig. 1. Most CNNs were pretrained on the ImageNet dataset [18], so we also considered their corresponding pre-trained models. Particularly, lesion identification involved Hourglass [17], VGG16 [19], AlexNet [18], InceptionV3 [20], MobileV3-small [21], ResNet18 [22] and DenseNet121 [23]. Here, for a favorable trade-off between localization and classification, we down-sampled the original slice resolution by a ratio of approximately 1/4. This required us to prune the original network architectures. For lesion classification, we followed the original architectures without modification, including VGG16, AlexNet, InceptionV3, MobileV3-small, ResNet18 and DenseNet121.

3.2.2 Loss Function

We optimized the object functions of center-point-based lesion identification (see Eq. (6)) and self-adaptive lesion classification (see Eq. (7)) in an end-to-end manner.

$$L_{ceter} = \frac{-1}{N} \sum_{i=1}^H \sum_{j=1}^W \begin{cases} (1 - p_{ij})^\alpha \log(p_{ij}) & \text{if } y_{ij} = 1 \\ (1 - y_{ij})^\beta (p_{ij})^\alpha \log(1 - p_{ij}) & \text{otherwise} \end{cases} \quad (6)$$

where L_{ceter} was a variant of focal loss [17]. p_{ij} was the probability at location (i, j) in the predicted center saliency map. y_{ij} was the ground truth of saliency map. N was the number of lesions in a slice. α and β were the hyper-parameters which controlled the contribution of each point (we empirically set α to 2 and β to 4).

$$L_{cls} = - \sum_{i=0}^{c-1} y_i \log(p_i) \quad (7)$$

where L_{cls} is the cross-entropy loss [15]. C was the class number; p_i was the predicted probability for the i -th class; y_i was the i -th class label.

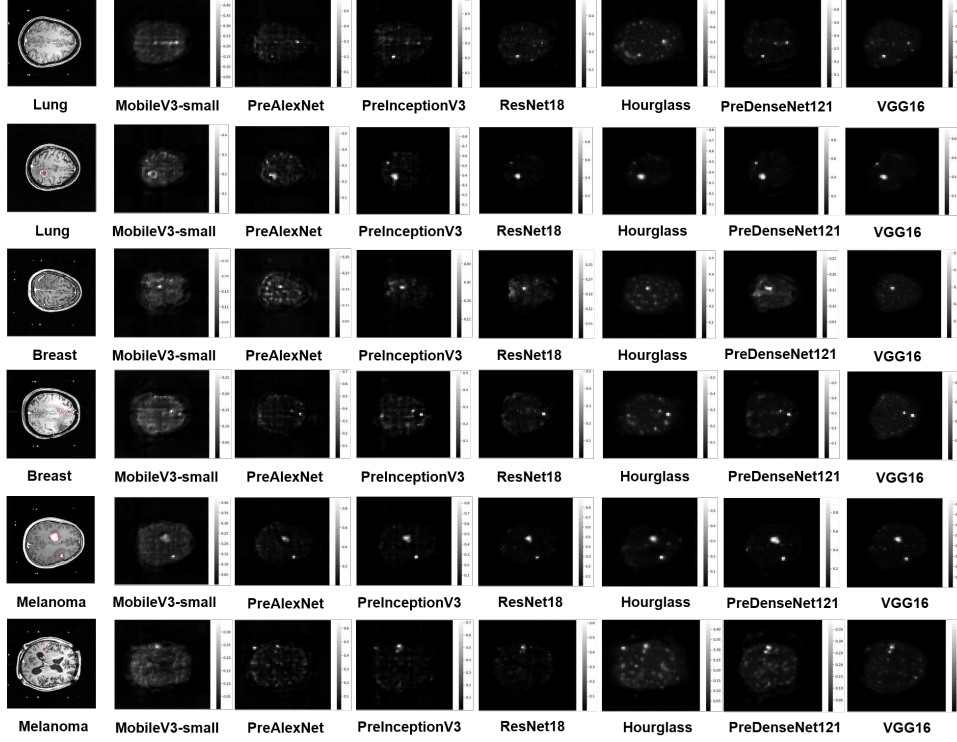


Fig. 5 Visualization results of center-point-based lesion identification. The first column represents the original T1ce-MRI slice, where the red curves contour BMs. The other columns are the predicted center saliency maps by different encoders. We add a prefix “Pre” to denote an encoder using the pretrained model.

3.3 Results

The performance of origin tracing was assessed by the accuracy, positive predictive value (PPV) and sensitivity [24]. First, in lesion identification, VGG16 without pre-trained model yielded the best center point detection performance. Fig. 5 illustrates the visualization results of predicted center saliency maps from different CNN encoders with or without the pretrained model (we only showed the better one for each feature encoder). This also suggested that our CSMG algorithm captured important information to accurately identify the center regions of BM. Accordingly, the subsequent classification model was focused on the lesion-based ROI that clinicians usually prioritize during the feature learning process, also improving the interpretability of the model and demonstrating its potential power in diagnosing BM.

Following lesion identification, we adopted SATS algorithm to crop ROIs as positive instances and performed lesion classification on them. Correspondingly, we obtained the slice-wise classification results, as shown in Table 2. Comparing with other encoders, DenseNet121 achieved the best overall performance of 70.19% (accuracy), 59.88% (PPV) and 59.48% (sensitivity). ResNet18 yielded moderate results of

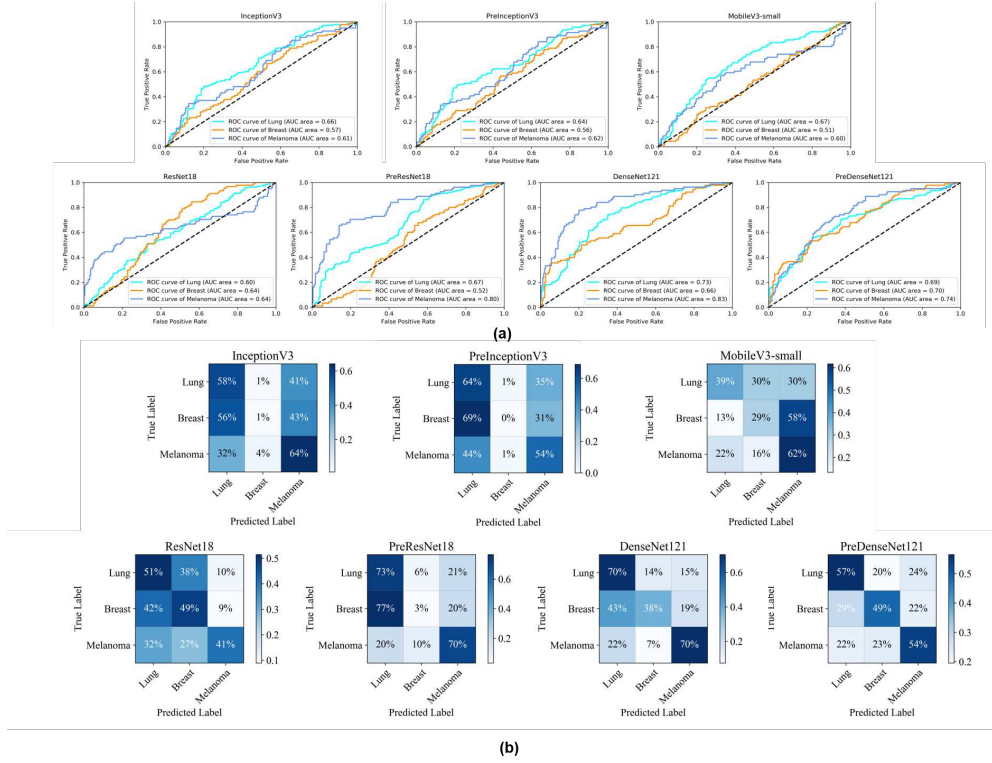


Fig. 6 Statistical results of slice-wise classification from competitive CNN encoders. (a) represents the ROC curve including the AUC value. (b) is the confusion matrix. We add a prefix “Pre” to denote an encoder using the pretrained model.

58.65%, 49.86% and 47.03%, accordingly. Mobilev3-small and InceptopnV3 both had unsatisfactory results, where their average PPVs and sensitivities were even lower than 40%. VGG16 and AlexNet lacking of sophisticated architectures had worst classification performance, and they both tended to classify the lesion into the lung type. Actually, the accuracy, PPV and sensitivity showed nearly no difference for an encoder with or without the pretrained model, such as DenseNet121 ([70.19%, 59.88% and 59.48%] vs. [63.56%, 52.73% and 53.24%]). Also, the prediction quality assessment for competitive classifiers were exhibited in Fig. 6, including the ROC curve and the class confusion matrix. The AUC values of DenseNet121 reached at 73% (lung), 66% (breast) and 83% (melanoma).

Finally, we showed the patient-wise results via VMM in Table 3. To demonstrate the effectiveness of mix-modal data decoration, we conducted a set of ablation experiments based on the optimal model combination for DMIL (VGG16 plus DenseNet128), i.e., DMIL plus MDD (T1ce, pseudo-T2 and pseudo-Flair), and DMIL without MDD (only T1ce). When removing mix-modal data decoration from the pipeline, we observed a significant decline in performance: 87.27% vs. 80.54% (accuracy), 85.00% vs. 80.56% (PPV) and 83.33% vs. 75.00% (sensitivity). Mix-modal data decoration

Table 2 Slice-wise results of different CNN backbones. $\times$ represents an invalid value. – suggests there is no pretrained model. Bold indicates the best result.

Encoder	w/o pretrained model (%)			pretrained model (%)		
	Accuracy	PPV	Sensitivity	Accuracy	PPV	Sensitivity
VGG16	56.07	$\times$	33.33	56.07	$\times$	33.33
AlexNet	56.07	$\times$	33.33	56.07	$\times$	33.33
InceptionV3	51.60	34.44	41.09	53.88	27.99	39.36
MobileV3-small	51.82	43.17	42.88	-	-	-
ResNet18	58.65	49.86	47.03	62.70	42.03	48.96
DenseNet121	70.19	59.88	59.48	63.56	52.73	53.24

Table 3 Patient-wise results of ablation experiment. Bold indicates the best result.

Method	Accuracy (%)	PPV (%)	Sensitivity (%)
DMIL w/o MDD	80.54	80.56	75.00
DMIL w/ MDD	87.27	85.00	83.33

could be treated as a specific data augmentation method in our scenario. Actually, data argumentation studies have proved that even a simple data copying can enhance the model performance.

4 Discussion and Conclusion

Interestingly, we found a contradictory result that VGG16 had the best performance for lesion identification but the worst result for lesion classification. This suggested that the plain features were more suitable for localization rather than semantic classification. In addition, seen from Fig. 6, the classifier was easy to be confused by the breast site. Referring to the histopathologic records, we discovered that the unclassified breast cases exhibited positive mutations, which could pose a challenge in terms of origin tracing.

Our work still has several limitations. First, as a prospective study, we mainly focused on the three most common primary tumor types (i.e., lung, breast and melanoma). However, this is far from sufficient to meet clinical demands. In the future, we need to further expand the range of origin tracing. Second, the amount of data available was limited, especially after down-sampling. We plan to collaborate with multiple hospitals to collect more data. Third, the performance of origin tracing still needed to be improved. We will investigate the correlation between macro (radiology) and micro (genetics) views to further enhance the effectiveness of computer-aided clinical decisions.

In conclusion, the proposed DMIL-based pipeline leverages deep learning techniques for origin tracing. Our work mainly enjoys three major advantages. First, DMIL has the potential to enhance radiologists’ reading capabilities when screening radiology images for brain metastasis. It could reduce the reading time and speed up the overall workflow. Second, the DMIL-aided pipeline provides a non-invasive clue for doctors to identify the primary tumor. Third, the use of a common MRI acquisition protocol reduces the overhead costs and improves the feasibility of clinical applications.

Declarations

- This work is supported by the Sichuan Province Science and Technology Support Program(Grant 2021JDJQ0024, 2021YFS0204 and 2023YFG0274).
- We have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.
- This retrospective study was approved by the Sichuan University Medical Ethics Committee(No. 201968, date: 2019-4-1) in accordance with the Declaration of Helsinki. The informed consent was obtained from all human participants.
- Data and code involved in this study are available from the corresponding author by request.
- H. Y., Z. Z., Z. Y. and Z. W. designed and implemented the algorithm. H. Y., Y. L. and Z. Y. wrote the main manuscript. L. L. collected the data. T. W. and Z. W. processed the data. All authors reviewed the manuscript.
- This work is original research that has not been published previously and not under consideration for publication elsewhere, in whole or in part.

References

- [1] Achrol, A.S., *et al.*: Brain metastases. *Nat Rev Dis Primers* **5**, 1–26 (2019)
- [2] Maesawa, S., *et al.*: Brain metastases in patients with no known primary tumor. *Cancer* **89**, 1095–1101 (2000)
- [3] Polyzoidis, K.S., Miliaras, G., Pavlidis, N.: Brain metastasis of unknown primary: A diagnostic and therapeutic dilemma. *Cancer Treatment Reviews* **31**, 247–255 (2005)
- [4] Penson, A., *et al.*: Development of genome-derived tumor type prediction to inform clinical cancer care. *JAMA Oncology* **6**, 84–91 (2020)
- [5] Lu, M.Y., *et al.*: Ai-based pathology predicts origins for cancers of unknown primary. *Nature* **594**, 106–110 (2021)
- [6] Chuheng, C., *et al.*: Identifying primary tumor site of origin for liver metastases via a combination of handcrafted and deep learning features. *The Journal of Pathology: Clinical Research* **10**, 344 (2024)
- [7] Ambrosini, V., *et al.*: 18f-fdg pet/ct in the assessment of carcinoma of unknown primary origin. *Radiol med* **111**, 1146–1155 (2006)
- [8] Ray, S., Scott, S., Blockeel, H.: Multi-instance learning. *Encyclopedia of Machine Learning* **30**, 701–710 (2011)
- [9] Rawat, W., Wang, Z.: Deep convolutional neural networks for image classification: A comprehensive review. *Neural Computation* **29**, 2352–2449 (2017)

- [10] Ning, L., *et al.*: Differentiation of spinal metastases originated from lung and other cancers using radiomics and deep learning based on dce-mri. *Magnetic resonance imaging* **64**, 4–12 (2019)
- [11] Yu, H., *et al.*: Deep learning-based classification of primary bone tumors on radiographs: A preliminary study. *EBioMedicine* **62**, 103121 (2020)
- [12] Tianyu, J., *et al.*: Deep learning with an attention mechanism for differentiating the origin of brain metastasis using mr images. *Journal of Magnetic Resonance Imaging* **58**, 1624–1635 (2023)
- [13] Baid, U., *et al.*: The RSNA-ASNR-MICCAI BraTS 2021 Benchmark on Brain Tumor Segmentation and Radiogenomic Classification. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–17 (2021)
- [14] Liu, X.Q., Ono, K., Bise, R.: Mixing Data Augmentation with Preserving Fore-ground Regions in Medical Image Segmentation. In: *Proceedings of the 20th IEEE International Symposium on Biomedical Imaging*, pp. 18–21 (2023)
- [15] Yu, H., *et al.*: Deseg: auto detector-based segmentation for brain metastases. *Phys. Med. Biol.* **68**, 025002 (2023)
- [16] Yin, Y., *et al.*: Tumor cell load and heterogeneity estimation from diffusion-weighted mri calibrated with histological data: an example from lung cancer. *IEEE Transactions on Medical Imaging* **37**, 35–46 (2018)
- [17] Duan, K., *et al.*: CenterNet: Keypoint Triplets for Object Detection. In: *Proceedings of the IEEE International Conference on Computer Vision*, pp. 6568–6577 (2019)
- [18] Alex, K., Ilya, S., Geoffrey, E.H.: Imagenet classification with deep convolutional neural networks. *Communications of the ACM* **60**, 84–90 (2017)
- [19] Simonyan, K., Zisserman, A.: Very Deep Convolutional Networks for Large-Scale Image Recognition. In: *Proceedings of the International Conference on Learning Representations*, pp. 1–14 (2015)
- [20] Szegedy, C., *et al.*: Rethinking the Inception Architecture for Computer Vision. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2818–2826 (2016)
- [21] Howard, A., *et al.*: Searching for MobileNetV3. In: *Proceedings of the IEEE International Conference on Computer Vision*, pp. 1314–1324 (2019)
- [22] He, K., *et al.*: Deep Residual Learning for Image Recognition. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778 (2016)

- [23] Huang, G., *et al.*: Densely Connected Convolutional Network. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2261–2269 (2017)
- [24] Watson, S.: The classification of metrics and multivariate statistical analysis. *Topology and its Applications* **99**, 237–261 (1999)