Data Augmentation with Multi-armed Bandit on Image Deformations Improves Fluorescence Glioma Boundary Recognition

Anqi Xiao[1,2], Keyi Han[1,2], Xiaojing Shi[1,2], Jie Tian[1,2,3,4,5]✉, and Zhenhua Hu[1,2,5]✉

[1] CAS Key Laboratory of Molecular Imaging, Beijing Key Laboratory of Molecular Imaging, Institute of Automation, Chinese Academy of Sciences, China
[2] School of Artificial Intelligence, University of Chinese Academy of Sciences, China
[3] Beijing Advanced Innovation Center for Big Data-based Precision Medicine, School of Engineering Medicine, Beihang University, China
[4] Engineering Research Center of Molecular and Neuro Imaging of Ministry of Education, School of Life Science and Technology, Xidian University, China
[5] National Key Laboratory of Kidney Diseases, China
tian@ieee.org, zhenhua.hu@ia.ac.cn

Abstract. The recognition of glioma boundary is challenging as a diffused growthing malignant tumor. Although fluorescence molecular imaging, especially in the second near-infrared window (NIR-II, 1000-1700 nm), helps improve surgical outcomes, fast and precise recognition remains in demand. Data-driven deep learning technology shows great promise in providing objective, fast, and precise recognition for glioma boundaries, but the lack of data poses challenges for designing effective models. Automatic data augmentation can improve the representation of small-scale datasets without requiring extensive prior information, which is suitable for fluorescence-based glioma boundary recognition. We propose Explore and Exploit Augment (EEA) based on multi-armed bandit for image deformations, enabling dynamic policy adjustment during training. Additionally, images captured in white light and the first near-infrared window (NIR-I, 700-900 nm) are introduced to further enhance performance. Experiments demonstrate that EEA improves the generalization of four types of models for glioma boundary recognition, suggesting significant potential for aiding in medical image classification. Code is available at https://github.com/ainieli/EEA.

Keywords: Glioma · NIR-II fluorescence imaging · Multi-modal imaging · Image classification · Automatic data augmentation

1 Introduction

Glioma is a fatal primary brain tumor that threatens human health [1]. Microsurgery under white light (WL) is the primary treatment for glioma. However, due to its diffuse nature, identifying the boundary of glioma is challenging for

neurosurgeons [26]. Consequently, the recurrence rate of glioma is high, leading to poor prognosis [22]. Recognizing the boundary of glioma is crucial for improving surgical outcomes and prognosis for glioma patients.

Fluorescence molecular imaging is a promising technology in both pre-clinic and clinic [3,4,13], which can mark the glioma through fluorescent molecular probes during surgery. It guides neurosurgeons to achieve complete resection and further reduce the gliomas recurrence [13,21]. Indocyanine green (ICG) is a commonly used probe for the first near-infrared window (NIR-I, 700-900 nm) imaging that reveals intricate biological structure and functional details of glioma that are not easily discernible under white light WL [13]. Furthermore, the emission spectrum of ICG also has a tail in the second near-infrared window (NIR-II, 1000-1700 nm), where fluorescence imaging offers higher contrast, deep tissue penetration, and sharper boundaries compared to NIR-I imaging [3]. Therefore, NIR-II imaging holds great promise for improving outcomes in glioma resection.

Although fluorescence molecular imaging shows advantages over traditional WL imaging, precise resection is still difficult to achieve. In clinical practice, the recognition of boundaries relies on the subjective experiments of surgeons, or on time-costly intraoperative frozen section biopsies [10]. There lacks a fast and objective diagnostic method for clearly recognizing glioma boundaries. Thanks to the development of deep learning (DL) technology, this goal becomes possible [20,27]. However, research on NIR-II imaging in clinical settings is still in its nascent stages [2,11,21], resulting in limited prior knowledge of its application. Additionally, the scarcity of NIR-II glioma boundary data presents challenges in designing effective data-driven DL methods.

Data augmentation (DA) can reduce the negative impact of small-scale dataset on training DL models by expanding the dataset [7,30], among which automatic DA can automatically adjust its augmentation policy to suit the data [5,6,28,29]. Therefore, automatic DA is expected to improve the performance of DL models for NIR-II imaging-based glioma boundary recognition, even without extensive priors of NIR-II imaging in clinic. However, current automatic DA methods, such as the widely adopted AutoAugment (AA) [5], are primarily studied on natural images. Meanwhile, they either adopt complicated designs for implementation [29] that set barriers to medical image analysis, or extra time-consuming search process [5,6]. Although the searched policies can be transferred to new tasks, the performance is usually decreased compared to searching from scratch [5]. Some automatic DA, like RandAugment (RA) [6], employ simpler implementations based on heuristics to reduce optimization costs, but this may sacrifice flexibility and performance [28].

To strike a balance between easy implementation and performance gain, we propose an intuitive automatic DA that can dynamically adjust the image deformations during training. We call it Explore and Exploit Augment (EEA). The proposed EEA improves the performance of glioma boundary recognition based on NIR-II fluorescence imaging. It is inspired by the exploration and exploitation in multi-armed bandit problem, which dynamically adjusts the image deformations to achieve better generalization on the small-scale NIR-II dataset.

In addition, we further introduce WL and NIR-I imaging of the same glioma tissue to utilize the complementary information for better performance [27]. We emphasize that the contribution of EEA is *modeling deformation selection as multi-armed bandit problem*, providing new insights to DA for better glioma boundary recognition. To the best of our knowledge, *we are the first to study the impact of DA on the diagnosis based on NIR-II imaging.*

2 Method

2.1 Formalize the Automatic Data Augmentation

We decouple the proposed EEA into two parts: augmentation operators $\mathcal{O}$ for controlling how to deform images, and magnitudes for controlling the extent of the deformation. In addition, we refer to the previous work RA [6] and adopt the operators from it as the basic candidate operator set $\mathcal{O}$. We simultaneously augment the WL, NIR-I, and NIR-II images within one transformation. Note that NIR-I and NIR-II images each have only one channel that represents fluorescence intensity, which differs from WL images that have RGB channels. Therefore, operators such as *color* and *contrast*, which require a conversion to grayscale images during transformation, are removed from $\mathcal{O}$ because multi-modal images do not have a grayscale representation. The remaining operators are used in the following experiments.

Magnitude ranges of candidate operators are first normalized to range [0, 1], with 0 indicating no augmentation and 1 indicating the maximum. Then, these magnitudes are uniformly discretized at N levels. Note that operators without magnitudes are always applied when sampled.

Furthermore, we introduce another parameter augmentation depth D that allows the combination of different augmentation operators for the same image. We choose the empirical value used in natural image augmentation [5,6] to reduce the number of learnable parameters. Note that *identity*, which keeps the input unchanged ,is within in the candidate operator set, thus augmented images with fewer than D transformations applied are also included. In general, the augmented image can be formalized as

$$I_i' = O_i^{1,m_i^1} \circ \cdots \circ O_i^{D,m_i^D}(I_i), 1 \leq i \leq B, \tag{1}$$

where I_i and I_i' are images before and after augmentation, B is the batch-size of the image batch $\mathcal{B}$, $\circ$ indicates sequentially applying the operators, and $O_i^{d,m^d}(1 \leq d \leq D)$ is the sampled augmentation operator for the dth transformation O^d with magnitude m^d for I_i. Therefore, data augmentation becomes a problem of determining how to obtain O_i^d and m_i^d.

2.2 Define the Multi-armed Bandit Problem in Data Augmentation

The multi-armed bandit is a classical problem in reinforcement learning (RL), aiming to find a proper balance between exploring the unknown and exploiting

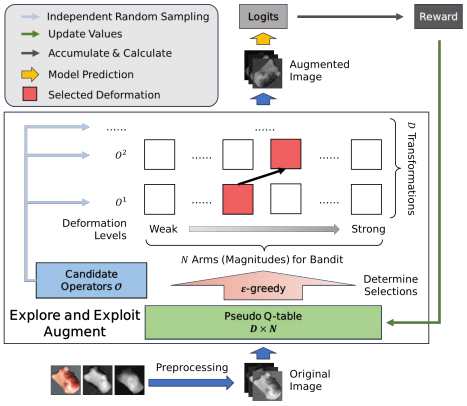

Fig. 1. Pipeline of the proposed Explore and Exploit Augment (EEA) for glioma boundary recognition with white light, NIR-I, and NIR-II imaging.

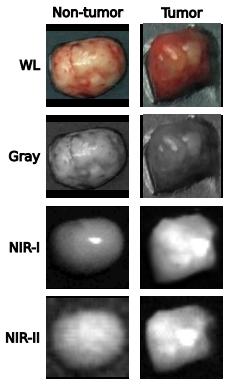

Fig. 2. Cases of multi-modal images for glioma boundary recognition.

the current best option. We formulate the selection of m_i^d in DA as the selection of arms in a multi-armed bandit, where each level can be treated as an arm. With different selections of magnitudes, the augmented image is expected to exhibit different deformations, potentially improving the generalization of DL models. Note that previous works have demonstrated the sensitivity of automatic DA to the selection of O^d [6], which also increases the complexity of these methods. However, a random selection baseline can also achieve satisfactory performance with limited performance decrement [6]. Therefore, for simplicity, the selection of operators O^d in our work is based on random selection. The optimization objective in EEA then becomes the selection of m_i^d. This reduction in parameters for optimization also eases the burden of EEA, making it more intuitive for understanding and implement.

2.3 Choose Image Deformations

SARSA [19] is a widely applied optimization strategy in RL, which involves maintaining a Q-table for selecting the optimal action in a specific state and updating the policy based on action taken. Inspired by this design, and considering that the multi-armed bandit is a one-state problem, we slightly modify the Q-table to a pseudo version (pQ-table) used for selecting the optimal m^d values. Therefore, the pQ-table has a size of $D \times N$.

To update the pQ-table, we measure the reward $r_i\,(1 \leq i \leq B)$ on one image batch $\mathcal{B}$. The final reward r has the same size as pQ-table, with the value at a specific position (d, m^d) denoted as r^{d,m^d}. For each image I_i', we sample D transformations $O_i^{d,m^d}\,(1 \leq d \leq D)$ and accumulate the reward of taking O_i^{d,m^d} in r^{d,m^d}. Therefore, the average r^{d,m^d} can be calculated on $\mathcal{B}$. For position (d, m^d) where the magnitude m^d is not sampled at augmentation depth d in the batch, the reward is assigned 0.

Algorithm 1: Training with Explore and Exploit Augment (EEA)

Input: $D_{taining}$: Training dataset; $\mathcal{B}_1$, $\mathcal{B}_2$: Glioma image batch; I_i, I_j: Single image with WL/NIR-I/NIR-II modalities; $\mathcal{M}$: DL model; O^d: Augmentation operator for the dth transformation; $\mathcal{O}$: Candidate operator set; D: Number of transformations for one image; N: Number of magnitude levels; m^d: Magnitude of the dth transformation; Q^d: Value of pQ-table for the dth transformation; r: Reward; ϵ: Predicted probability for randomly sampling m^d; γ_ϵ: Decay factor for ϵ

Output: $\mathcal{M}$: Well-trained DL model

```
Initial pQ-table with 0 at each position
for epoch in epochs do
    ε ← γ_ε · ε
    for B_1, B_2 in D_train do
        for I_i in B_1 do
            Randomly sample O_i^d s (1 ≤ d ≤ D) from O
            Sample m_i^d s (1 ≤ d ≤ D) based on Q^d through ε-greedy
            Augment I_i to I_i' with Eq. 1
        end
        Stack I_i' s to form B_1', and train M to M' using B_1'
        Calculate r according to Eq. 2
        for I_j in B_2 do
            Randomly sample O_j^d s (1 ≤ d ≤ D) from O
            Sample m_j^d s (1 ≤ d ≤ D) based on Q^d through ε-greedy
            Augment I_j to I_j' with Eq. 1
        end
        Stack I_j' s to form B_2', and train M' to M'' using B_2'
        Calculate Q' using Eq. 4
        Update pQ-table according to Eq. 3
        M ← M''
    end
end
```

To update the pQ-table, we use the following equations:

$$r^{d,m^d} = \begin{cases} 0, \, if \; \forall i, O : O_i^{d,m^d} \; is \; not \; sampled, \; 1 \le i \le B \\ \hat{p}, \, otherwise. \end{cases} \tag{2}$$

$$Q^{d,m^d} \leftarrow Q^{d,m^d} + \alpha(r^{d,m^d} + \gamma Q' - Q^{d,m^d}), \tag{3}$$

where $\hat{p}$ is the average probability of the correct class for I_i's with O_i^{d,m^d} sampled, Q^{d,m^d} is the value of pQ-table at position (d, m^d), α is the learning rate for updating the pQ-table, γ is the discount factor, and Q' is the estimated value of current strategy for selecting m^ds.

We use two batches as a unit for exploiting the current policy and evaluating its value, respectively. This design leverages the observed value directly from the

real fluorescence glioma boundary images to estimate the performance of the current pQ-table and further adjust the policy to better represent the dataset's features. Therefore, Q' in Eq. 3 is evaluated using the average predicted probability of the correct class on the second image batch, which is formulated as

$$Q' = \hat{p_j},\ 1 \leq j \leq B, \tag{4}$$

where $\hat{p_j}$ is the average predicted probability of the correct class on the second image batch. Note that the DL model parameters are updated after each image batch, while the pQ-table in EEA is only updated after utilizing the second image batch. The general pipeline of the proposed EEA is shown in Fig. 1.

2.4 Explore and Exploit in Glioma Boundary Recognition

The magnitude sampling using the pQ-table exploits the priors learned during training. Meanwhile, an ϵ-greedy strategy is adopted for exploration, allowing a probability ϵ to randomly sample m^ds without considering pQ-table. To balance exploration and exploitation, we further introduce a decay factor γ_ϵ for the ϵ-greedy strategy. At the beginning of training, DL models require more image variants to improve generalization [9]. As training progresses, the models become capable of extracting general features from images, thus needing to exploiting a refined data distribution. Therefore, the value of ϵ is expected to gradually decrease with training. We take an exponential decay strategy, where ϵ is decayed at the beginning of each epoch. The pseudo-code for training a glioma boundary recognition model with EEA is shown in Algorithm 1.

3 Experiments

Dataset. A homemade dataset of 2025 (non-tumor:tumor=562:1463) ex-vivo glioma tissues at boundaries from 34 patients, including corresponding WL, NIR-I, and NIR-II images, is established for evaluating EEA (Fig. 2). The fluorescence molecular images represent fluorescence intensity, with each pixel normalized to range [0, 1]. The gold standards for tumor and non-tumor are determined by postoperative pathology. Note that this dataset is small-scale for DL models. Therefore, we adopt 5-fold cross-evaluation at the patient-level to measure the performance of different settings. We utilize the multi-modal data in 3-channel and 9-channel settings according to the model design. For the 3-channel data, all WL images are first transferred into grayscale images, then concatenated with NIR-I and NIR-II images. For the 9-channel data, specially used in DLS-DARTS [27], all NIR-I and NIR-II images are first processed into RGB format and then concatenated with WL images.

Evaluation Metrics. We use accuracy (Acc), specificity (Spe), sensitivity (Sen), F1 score, and area under receiver operating characteristic curve (AUC) to measure the performance of different settings. Since performance evaluation across multiple metrics is usually non-intuitive and we aim to improve the general

performance of DL models, we additionally introduce the harmonic mean (HM) of all the metrics to compare the performance between different DA methods.

Implementation Details. In our experiments, the number of magnitude levels N is set to 11, augmentation depth D is set to 2, learning rate for pQ-table α is set to 0.01, and the discount factor γ is set to 0.98. ϵ starts from 1.0 and gradually decays to a minimum value of 0.1, with the decay factor γ_ϵ set to 0.95, 0.975, and 0.985 for epochs 50, 110, and 200, respectively. More details are provided in the Supplementary Materials.

We select four different DL architectures to evaluate the performance of EEA: ResNet-based architecture (ResNet-18) [8], MobileNet-based architecture (EfficientNet-B0) [24], vision Transformer-based architecture (DeiT-Tiny) [25], and a directly transferred multi-modal neural architecture search-based architecture (DLS-DARTS) [15,27]. All models use Adam as the optimizer, with a cosine annealing learning rate decay schedule [16]. For models apart from DLS-DARTS, Dropout [23] with a 0.5 probability is introduced to avoid overfitting on the training data. DeiT-Tiny and DLS-DARTS further introduce DropPath [12] with a 0.2 probability for regularization. Since the dataset is imbalanced, focal loss [14] with α_{focal} set to [1.0, 0.33] and γ_{focal} set to 2.0 is used. Hyperparameter settings for training DL models are selected based on the original papers and tuned on our data to achieve satisfactory performance due to the feature gaps between natural images and fluorescence molecular images.

3.1 Comparison to Other Data Augmentations

All models are trained from scratch, with the former three using 3-channel data while DLS-DARTS uses 9-channel data. Basic augmentation, including random crop, horizontal flip, and vertical flip, is applied to all experiments at the beginning of augmentation. We compare EEA with two widely adopted automatic DA methods, AA [5] and RA [6], in new tasks, as well as with a manually designed DA method CutMix [30], which is widely used in medical image recognition [17]. Note that the usage of CutMix and other automatic DA is agnostic, allowing them to be combined.

As shown in Table 1, models training with EEA achieve the best performance across most settings and metrics, indicating its advantage in improving the generalization of DL models over the compared methods. It is noteworthy that automatic DA negatively affect the recognition of glioma boundaries, consistent with previous reports [6,18] due to over transformation. Only EEA shows positive effect in DLS-DARTS. DLS-DARTS prefers automatic DA with CutMix in AA and RA, but not in basic augmentation and EEA. We attribute this contradiction to the improvements in the representation of the augmented samples in AA and RA with soft labels generated by CutMix. However, this improvement cannot compensate the negative effect of these two methods. Conversely, in basic augmentation and EEA, the representation of the augmented samples is adequate for recognition, where the mixture of two images introduces over-transformation and bias to the original representation.

Table 1. Performance of different DA using different types of DL architectures on fluorescence glioma boundary recognition. CM: CutMix. **Bold**: The best performance.

	ResNet-18						EfficientNet-B0					
Augmentation	Acc	Spe	Sen	F1	AUC	HM	Acc	Spe	Sen	F1	AUC	HM
Basic	0.8175	0.7841	0.8381	0.8656	0.8596	0.8319	0.7968	0.6694	0.8468	0.8564	0.8248	0.7924
CutMix	0.8323	0.7299	**0.8748**	0.8825	0.8669	0.8331	0.7925	0.7362	0.8183	0.8486	0.8366	0.8044
AA	0.8036	0.7400	0.8362	0.8566	0.8495	0.8148	0.8010	**0.7377**	0.8356	0.8552	0.8553	0.8144
AA+CM	0.8138	0.7696	0.8402	0.8658	0.8597	0.8283	0.7980	0.7238	0.8348	0.8534	0.8284	0.8050
RA	0.8030	0.7291	0.8417	0.8569	0.8560	0.8143	0.7668	0.6909	0.7995	0.8301	0.8149	0.7771
RA+CM	0.8097	0.7473	0.8391	0.8614	0.8609	0.8214	0.7835	0.7005	0.8216	0.8419	0.8296	0.7919
EEA(Ours)	0.8146	0.8023	0.8331	0.8606	**0.8855**	0.8381	0.8046	0.7123	0.8474	0.8601	0.8600	0.8127
EEA(Ours)+CM	**0.8419**	**0.8041**	0.8684	**0.8851**	0.8785	**0.8545**	**0.8224**	0.7245	**0.8678**	**0.8728**	**0.8752**	**0.8282**
	DeiT-Tiny						DLS-DARTS (use architecture in [27])					
Augmentation	Acc	Spe	Sen	F1	AUC	HM	Acc	Spe	Sen	F1	AUC	HM
Basic	0.8039	0.6356	**0.8627**	0.8610	0.8144	0.7855	0.8816	**0.7963**	0.9136	0.9192	0.8995	0.8796
CutMix	0.7795	0.7171	0.8128	0.8389	0.8282	0.7927	0.8704	0.7889	0.8974	0.9106	0.8807	0.8674
AA	0.7871	0.6449	0.8419	0.8485	0.7788	0.7726	0.8389	0.7933	0.8603	0.8844	0.8786	0.8498
AA+CM	0.7795	**0.7293**	0.7957	0.8339	0.7958	0.7854	0.8483	0.7889	0.8724	0.8927	0.8777	0.8544
RA	0.7991	0.6914	0.8283	0.8537	0.8074	0.7917	0.8601	0.7499	0.8982	0.9040	0.9017	0.8584
RA+CM	0.7837	0.7076	0.8124	0.8423	0.8155	0.7894	0.8776	0.7840	0.9098	0.9163	0.8913	0.8730
EEA(Ours)	0.7942	0.7006	0.8312	0.8516	**0.8411**	0.7997	**0.9127**	0.7863	**0.9514**	**0.9416**	**0.9042**	**0.8950**
EEA(Ours)+CM	**0.8099**	0.6944	0.8523	**0.8647**	0.8400	**0.8071**	0.8631	0.7851	0.8902	0.9048	0.8954	0.8654

3.2 Impact of the Number of Magnitude Levels

The number of magnitude levels determines both the difficulty of learning EEA and the diversity of the augmented images, which is especially important when limited training data of fluorescence glioma boundary tissues is available. This value can be intuitively understood as the number of choices for the bandit. We therefore explore the impact of this parameter, with Ns set to 2, 3, 11, and 31, respectively. The selections of these values are based on previous works [5,6] (value 11 and 31) and represent EEA with or without different deformations (value 2 and 3). As shown in Table 2, both too few and too many arms severely hurt performance. Setting the parameter $N = 11$ provides a satisfactory compromise between difficulty and diversity. However, the optimal value may require further exploration.

3.3 Ablation Studies

We further conduct ablation studies to show the contribution of each design in EEA. As shown in Table 3, some metrics achieve the best results when using cross entropy loss, which is influenced by the imbalanced fluorescence glioma boundary dataset. Specifically, specificity is significantly improved when using focal loss. We also find that the decay strategy of ϵ contributes to both specificity and the final HM, indicating that more exploration at the beginning of training is valuable. Additionally, we introduce a baseline that randomly selects m_i^ds during training without using the pQ-table, which shows a noticeable performance drop across almost all metrics. These experiments demonstrate that the synergistic effect of designs in EEA contributes to the performance in glioma boundary recognition.

Table 2. Performance of different Ns with CutMix using ResNet-18.

N	Acc	Spe	Sen	F1	AUC	HM
2	0.8154	0.7108	0.8599	0.8704	0.8652	0.8196
3	0.8317	0.7251	**0.8751**	0.8819	0.8685	0.8320
11	**0.8419**	**0.8041**	0.8684	**0.8851**	**0.8785**	**0.8545**
31	0.8182	0.6909	0.8667	0.8728	0.8466	0.8129

Table 3. Ablation studies on settings of EEA using ResNet-18. CE: Cross entropy.

Ablation part	Acc	Spe	Sen	F1	AUC	HM
Original	0.8419	**0.8041**	0.8684	0.8851	**0.8785**	**0.8545**
Use CE loss	**0.8558**	0.6451	**0.9410**	**0.9037**	0.8711	0.8284
W/o ϵ decay	0.8457	0.6852	0.9086	0.8943	0.8701	0.8320
Random m_i^d	0.8191	0.6694	0.8835	0.8741	0.8537	0.8114

4 Conclusion

In this paper, an automatic DA method EEA is proposed to improve the performance of different types of DL models in fluorescence glioma boundary recognition. The adjustment of image deformations using a multi-armed bandit shows satisfactory performance improvements in small-scale fluorescence molecular images. The proposed EEA can be integrated into the traditional training pipeline for better generalization, demonstrating great potential to aid surgeons in precise glioma resection.

Acknowledgments. The study is approved by the Ethics Committee of Beijing Tiantan Hospital, Capital Medical University (ChiCTR2000029402), and supported by the National Natural Science Foundation of China (92059207, 92359301, 62027901, 81930053, 81227901), CAS Youth Interdisciplinary Team (JCTD-2021-08). Authors acknowledge Hongyuan Yu, Weichen Yu, Xinjian Wu, and Zeyu Lei for their help.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. van den Bent, M.J., Geurts, M., French, P.J., Smits, M., Capper, D., Bromberg, J.E., Chang, S.M.: Primary brain tumours in adults. The Lancet **402**(10412), 1564–1579 (2023)
2. Cao, C., Jin, Z., Shi, X., Zhang, Z., Xiao, A., Yang, J., Ji, N., Tian, J., Hu, Z.: First clinical investigation of near-infrared window iia/iib fluorescence imaging for precise surgical resection of gliomas. IEEE Transactions on Biomedical Engineering **69**(8), 2404–2413 (2022)
3. Carr, J.A., Franke, D., Caram, J.R., Perkinson, C.F., Saif, M., Askoxylakis, V., Datta, M., Fukumura, D., Jain, R.K., Bawendi, M.G., et al.: Shortwave infrared fluorescence imaging with the clinically approved near-infrared dye indocyanine green. Proceedings of the National Academy of Sciences **115**(17), 4465–4470 (2018)
4. Chang, B., Li, D., Ren, Y., Qu, C., Shi, X., Liu, R., Liu, H., Tian, J., Hu, Z., Sun, T., et al.: A phosphorescent probe for in vivo imaging in the second near-infrared window. Nature Biomedical Engineering **6**(5), 629–639 (2022)
5. Cubuk, E.D., Zoph, B., Mane, D., Vasudevan, V., Le, Q.V.: Autoaugment: Learning augmentation strategies from data. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 113–123 (2019)

6. Cubuk, E.D., Zoph, B., Shlens, J., Le, Q.V.: Randaugment: Practical automated data augmentation with a reduced search space. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. pp. 702–703 (2020)
7. Garcea, F., Serra, A., Lamberti, F., Morra, L.: Data augmentation for medical imaging: A systematic literature review. Computers in Biology and Medicine **152**, 106391 (2023)
8. 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)
9. He, Z., Xie, L., Chen, X., Zhang, Y., Wang, Y., Tian, Q.: Data augmentation revisited: Rethinking the distribution gap between clean and augmented data. arXiv preprint arXiv:1909.09148 (2019)
10. Hollon, T.C., Pandian, B., Adapa, A.R., Urias, E., Save, A.V., Khalsa, S.S.S., Eichberg, D.G., D'Amico, R.S., Farooq, Z.U., Lewis, S., et al.: Near real-time intraoperative brain tumor diagnosis using stimulated raman histology and deep neural networks. Nature Medicine **26**(1), 52–58 (2020)
11. Hu, Z., Fang, C., Li, B., Zhang, Z., Cao, C., Cai, M., Su, S., Sun, X., Shi, X., Li, C., et al.: First-in-human liver-tumour surgery guided by multispectral fluorescence imaging in the visible and near-infrared-i/ii windows. Nature Biomedical Engineering **4**(3), 259–271 (2020)
12. Larsson, G., Maire, M., Shakhnarovich, G.: Fractalnet: Ultra-deep neural networks without residuals. In: International Conference on Learning Representations (2016)
13. Lee, J.Y., Thawani, J.P., Pierce, J., Zeh, R., Martinez-Lage, M., Chanin, M., Venegas, O., Nims, S., Learned, K., Keating, J., et al.: Intraoperative near-infrared optical imaging can localize gadolinium-enhancing gliomas during surgery. Neurosurgery **79**(6), 856 (2016)
14. Lin, T.Y., Goyal, P., Girshick, R., He, K., Dollár, P.: Focal loss for dense object detection. In: Proceedings of the IEEE International Conference on Computer Vision. pp. 2980–2988 (2017)
15. Liu, H., Simonyan, K., Yang, Y.: Darts: Differentiable architecture search. In: International Conference on Learning Representations (2018)
16. Loshchilov, I., Hutter, F.: Sgdr: Stochastic gradient descent with warm restarts. In: International Conference on Learning Representations (2016)
17. Nurgazin, M., Tu, N.A.: A comparative study of vision transformer encoders and few-shot learning for medical image classification. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 2513–2521 (2023)
18. Qin, J., Fang, J., Zhang, Q., Liu, W., Wang, X., Wang, X.: Resizemix: Mixing data with preserved object information and true labels. arXiv preprint arXiv:2012.11101 (2020)
19. Rummery, G.A., Niranjan, M.: On-line Q-learning using connectionist systems, vol. 37. University of Cambridge, Department of Engineering Cambridge, UK (1994)
20. Shen, B., Zhang, Z., Shi, X., Cao, C., Zhang, Z., Hu, Z., Ji, N., Tian, J.: Real-time intraoperative glioma diagnosis using fluorescence imaging and deep convolutional neural networks. European Journal of Nuclear Medicine and Molecular Imaging **48**(11), 3482–3492 (2021)
21. Shi, X., Zhang, Z., Zhang, Z., Cao, C., Cheng, Z., Hu, Z., Tian, J., Ji, N.: Near-infrared window ii fluorescence image-guided surgery of high-grade gliomas prolongs the progression-free survival of patients. IEEE Transactions on Biomedical Engineering **69**(6), 1889–1900 (2021)

22. van Solinge, T.S., Nieland, L., Chiocca, E.A., Broekman, M.L.: Advances in local therapy for glioblastoma—taking the fight to the tumour. Nature Reviews Neurology **18**(4), 221–236 (2022)
23. Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R.: Dropout: A simple way to prevent neural networks from overfitting. The Journal of Machine Learning Research **15**(1), 1929–1958 (2014)
24. Tan, M., Le, Q.: Efficientnet: Rethinking model scaling for convolutional neural networks. In: International Conference on Machine Learning. pp. 6105–6114. PMLR (2019)
25. Touvron, H., Cord, M., Douze, M., Massa, F., Sablayrolles, A., Jégou, H.: Training data-efficient image transformers & distillation through attention. In: International Conference on Machine Learning. pp. 10347–10357. PMLR (2021)
26. Weller, M., van den Bent, M., Preusser, M., Le Rhun, E., Tonn, J.C., Minniti, G., Bendszus, M., Balana, C., Chinot, O., Dirven, L., et al.: Eano guidelines on the diagnosis and treatment of diffuse gliomas of adulthood. Nature Reviews Clinical Oncology **18**(3), 170–186 (2021)
27. Xiao, A., Shen, B., Shi, X., Zhang, Z., Zhang, Z., Tian, J., Ji, N., Hu, Z.: Intraoperative glioma grading using neural architecture search and multi-modal imaging. IEEE Transactions on Medical Imaging **41**(10), 2570–2581 (2022)
28. Xiao, A., Shen, B., Tian, J., Hu, Z.: Differentiable randaugment: Learning selecting weights and magnitude distributions of image transformations. IEEE Transactions on Image Processing (2023)
29. Xu, J., Li, M., Zhu, Z.: Automatic data augmentation for 3d medical image segmentation. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd International Conference, Lima, Peru, October 4–8, 2020, Proceedings, Part I 23. pp. 378–387. Springer (2020)
30. Yun, S., Han, D., Oh, S.J., Chun, S., Choe, J., Yoo, Y.: Cutmix: Regularization strategy to train strong classifiers with localizable features. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 6023–6032 (2019)