**Supplementary for Preoperative Prediction of MVI from Multi-sequence MRI based on Deep Fusion Representation Learning**

**Section V RESULTS AND ANALYSIS**

1. Parameter Settings

The hyperparameters were determined using a grid search method.

Taking the random double sequence fusion as an example, we consider the fusion of the T1WI sequence and the HBP sequence. We varied the batch\_size from 8 to 64, and the epoch from 100 to 200. The search process and related results are presented in Table 1.

This resulted in (16, 100) combination of hyperparameters and the one with the best performance was selected as the hyperparameters of the best model.

Table 1: The parameter grid search process of T1WI+HBP sequences by ResNet-34 model.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Batch\_size | Epoch | Accuracy | AUC | F1-score | Specificity | Sensitivity |
| 8 | 100 | 0.950 | 0.953 | 0.946 | 0.938 | 0.967 |
| 8 | 150 | 0.943 | 0.943 | 0.938 | 0.956 | 0.931 |
| 8 | 200 | 0.933 | 0.932 | 0.925 | 0.952 | 0.912 |
| 16 | 100 | 0.952 | 0.950 | 0.944 | 0.949 | 0.950 |
| 16 | 150 | 0.886 | 0.896 | 0.877 | 0.896 | 0.889 |
| 16 | 200 | 0.928 | 0.932 | 0.923 | 0.909 | 0.954 |
| 32 | 100 | 0.943 | 0.950 | 0.946 | 0.985 | 0.914 |
| 32 | 150 | 0.936 | 0.950 | 0.942 | 0.950 | 0.947 |
| 32 | 200 | 0.912 | 0.920 | 0.911 | 0.945 | 0.895 |
| 48 | 100 | 0.903 | 0.913 | 0.898 | 0.902 | 0.927 |
| 48 | 150 | 0.874 | 0.894 | 0.882 | 0.808 | 0.981 |
| 48 | 200 | 0.889 | 0.908 | 0.895 | 0.836 | 0.979 |

The hyperparameter settings of the DL model are shown in Table 2.

Table 2: Network settings of DFFResNet

|  |  |
| --- | --- |
| **Hyperparameters** | **value** |
| Learning rate | 0.0001 |
| Epoch | 16 |
| Batch\_size | 100 |
| Optimizer | Adam |
| Loss function | Categorical cross entropy |
| Maxpool | 3×3, stride=2, padding=1 |
| Avgpool | AdaptiveAvgPool |
| Manual\_seed | 42 |
| FC | out\_feature: 2, in\_feature: pre-trained ResNet’s in\_feature |
| Convolutional layers | 7×7, stride=2, padding=3 |
| 3×3, stride=1, padding=1 |
| 1×1, stride=2, padding=0 |
| 3×3, stride=2, padding=1 |
| 3×3, stride=1, padding=1 |

1. Performance evaluation on different MRI sequences by different ResNet models

To select a suitable ResNet model for our feature fusion model, we conducted comparative experiments with respect to different ResNet models and MRI sequences. The results are summarized in the Table 3. For example, the values corresponding to ResNet-18 and T1WI represents the performance of ResNet-18 on the T1WI sequence. We calculated 95% CI for all the results.

Table 3: Comparative experimental results on different sequences by using different ResNet models

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| model | sequence | Accuracy (95% CI) | AUC (95% CI) | F1-score (95% CI) | Specificity (95% CI) | Sensitivity (95% CI) |
| **ResNet-18** | **T1WI** | 0.848 (0.839-0.858) | 0.848 (0.827-0.868) | 0.832 (0.798-0.865) | 0.848 (0.827-0.868) | 0.849 (0.839-0.858) |
| **ADC** | 0.818 (0.792-0.845) | 0.817 (0.798-0.836) | 0.799 (0.778-0.828) | 0.817 (0.798-0.836) | 0.818 (0.792-0.844) |
| **AP** | 0.830 (0.799-0.862) | 0.824 (0.788-0.861) | 0.802 (0.754-0.849) | 0.824 (0.788-0.861) | 0.831 (0.799-0.862) |
| **DP** | 0.821 (0.797-0.844) | 0.821 (0.790-0.853) | 0.804 (0.764-0.845) | 0.821 (0.790-0.853) | 0.821 (0.797-0.845) |
| **HBP** | 0.834 (0.808-0.860) | 0.836 (0.809-0.863) | 0.822 (0.792-0.853) | 0.836 (0.809-0.863) | 0.834 (0.809-0.859) |
| **PVP** | 0.824 (0.757-0.892) | 0.826 (0.764-0.889) | 0.811 (0.748-0.874) | 0.826 (0.764-0.889) | 0.824 (0.756-0.893) |
| **T2** | 0.806 (0.777-0.835) | 0.802 (0.784-0.820) | 0.778 (0.739-0.817) | 0.802 (0.784-0.820) | 0.806 (0.777-0.835) |
| **ResNet-34** | **T1WI** | 0.923 (0.909-0.936) | 0.923 (0.911-0.935) | 0.914 (0.899-0.929) | 0.923 (0.911-0.935) | 0.922 (0.906-0.938) |
| **ADC** | 0.851 (0.793-0.909) | 0.844 (0.769-0.920) | 0.822 (0.723-0.921) | 0.844 (0.769-0.920) | 0.851 (0.793-0.909) |
| **AP** | 0.879 (0.858-0.900) | 0.877 (0.857-0.898) | 0.865 (0.842-0.887) | 0.877 (0.857-0.898) | 0.879 (0.857-0.900) |
| **DP** | 0.832 (0.786-0.878) | 0.835 (0.802-0.869) | 0.822 (0.795-0.849) | 0.869 (0.758-0.979) | 0.833 (0.787-0.878) |
| **HBP** | 0.860 (0.837-0.883) | 0.858 (0.830-0.885) | 0.841 (0.800-0.882) | 0.858 (0.830-0.885) | 0.860 (0.837-0.883) |
| **PVP** | 0.805 (0.756-0.853) | 0.809 (0.759-0.860) | 0.796 (0.741-0.851) | 0.809 (0.757-0.861) | 0.805 (0.756-0.853) |
| **T2** | 0.783 (0.771-0.796) | 0.787 (0.769-0.804) | 0.771 (0.737-0.805) | 0.787 (0.769-0.804) | 0.783 (0.771-0.796) |
| **ResNet-50** | **T1WI** | 0.871 (0.855-0.886) | 0.869 (0.840-0.897) | 0.855 (0.817-0.892) | 0.869 (0.838-0.901) | 0.873 (0.846-0.901) |
| **ADC** | 0.830 (0.801-0.859) | 0.825 (0.788-0.862) | 0.803 (0.753-0.852) | 0.825 (0.788-0.862) | 0.830 (0.801-0.859) |
| **AP** | 0.853 (0.819-0.887) | 0.851 (0.825-0.877) | 0.836 (0.809-0.862) | 0.851 (0.825-0.877) | 0.852 (0.818-0.886) |
| **DP** | 0.857 (0.849-0.864) | 0.857 (0.842-0.871) | 0.842 (0.814-0.871) | 0.857 (0.842-0.871) | 0.856 (0.849-0.864) |
| **HBP** | 0.855 (0.848-0.863) | 0.858 (0.838-0.878) | 0.848 (0.804-0.891) | 0.858 (0.838-0.878) | 0.855 (0.847-0.863) |
| **PVP** | 0.845 (0.817-0.873) | 0.841 (0.816-0.866) | 0.823 (0.797-0.850) | 0.841 (0.816-0.866) | 0.845 (0.816-0.874) |
| **T2** | 0.833 (0.808-0.858) | 0.831 (0.804-0.857) | 0.812 (0.775-0.850) | 0.831 (0.804-0.857) | 0.833 (0.809-0.857) |

1. Performance evaluation by different fusion strategies
2. Determining the optimal number of sequences k

Our proposed deep feature fusion method is not limited to , and can be extended to fuse multiple sequence features (). Therefore, we first set and then explored calculation for and in Table 4.

Table 4: Comparative results obtained using different three-sequence or four-sequence fusion strategies by DFFResNet.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Sequence | Accuracy (95% CI) | F1-score (95% CI) | AUC (95% CI) | Sensitivity (95% CI) | Specificity (95% CI) |
| **T1WI+ADC+AP** | 0.760 (0.748-0.772) | 0.731 (0.719-0.743) | 0.758 (0.746-0.770) | 0.726 (0.714-0.738) | 0.789 (0.777-0.801) |
| **T1WI+ADC+DP** | 0.736 (0.724-0.748) | 0.743 (0.731-0.755) | 0.748 (0.736-0.760) | 0.924 (0.912-0.936) | 0.748 (0.736-0.760) |
| **T1WI+ADC+PVP** | 0.745 (0.733-0.757) | 0.727 (0.715-0.739) | 0.747 (0.735-0.759) | 0.756 (0.744-0.768) | 0.738 (0.726-0.750) |
| **T1WI+ADC+T2** | 0.708 (0.696-0.720) | 0.608 (0.596-0.620) | 0.689 (0.677-0.701) | 0.507 (0.495-0.519) | 0.872 (0.860-0.884) |
| **T1WI+AP+DP** | 0.727 (0.715-0.739) | 0.610 (0.598-0.622) | 0.708 (0.696-0.720) | 0.470 (0.458-0.482) | 0.945 (0.933-0.957) |
| **T1WI+AP+PVP** | 0.782 (0.770-0.794) | 0.736 (0.724-0.748) | 0.773 (0.761-0.785) | 0.677 (0.665-0.689) | 0.869 (0.857-0.881) |
| **T1WI+AP+T2** | 0.755 (0.743-0.767) | 0.674 (0.662-0.686) | 0.738 (0.726-0.750) | 0.563 (0.551-0.575) | 0.913 (0.901-0.925) |
| **T1WI+DP+PVP** | 0.837 (0.825-0.849) | 0.823 (0.811-0.835) | 0.839 (0.827-0.851) | 0.843 (0.831-0.855) | 0.935 (0.923-0.947) |
| **T1WI+DP+T2** | 0.707 (0.695-0.719) | 0.640 (0.628-0.652) | 0.670 (0.658-0.682) | 0.582 (0.570-0.594) | 0.809 (0.797-0.821) |
| **T1WI+HBP+ADC** | 0.912 (0.900-0.924) | 0.902 (0.890-0.914) | 0.912 (0.900-0.924) | 0.913 (0.901-0.925) | 0.912 (0.900-0.924) |
| **T1WI+HBP+AP** | 0.831 (0.819-0.843) | 0.783 (0.771-0.795) | 0.817 (0.805-0.829) | 0.677 (0.665-0.689) | 0.958 (0.946-0.970) |
| **T1WI+HBP+DP** | 0.781 (0.769-0.793) | 0.753 (0.741-0.765) | 0.778 (0.766-0.790) | 0.743 (0.731-0.755) | 0.814 (0.802-0.826) |
| **T1WI+HBP+PVP** | 0.727 (0.715-0.739) | 0.634 (0.622-0.646) | 0.709 (0.697-0.721) | 0.921 (0.909-0.933) | 0.709 (0.697-0.721) |
| **T1WI+HBP+T2** | 0.758 (0.746-0.770) | 0.758 (0.746-0.770) | 0.767 (0.755-0.779) | 0.844 (0.832-0.856) | 0.690 (0.678-0.702) |
| **T1WI+T2+PVP** | 0.709 (0.697-0.721) | 0.671 (0.659-0.683) | 0.705 (0.693-0.717) | 0.710 (0.698-0.722) | 0.772 (0.760-0.784) |
| **T1WI+HBP+ADC+DP** | 0.656 (0.644-0.668) | 0.675 (0.662-0.687) | 0.689 (0.677-0.702) | 0.694 (0.672-0.713) | 0.752 (0.730-0.775) |
| **AP+HBP+ADC+T2** | 0.592 (0.538-0.646) | 0.597 (0.445-0.749) | 0.491 (0.211-0.771) | 0.493 (0.000-1.000) | 0.692 (0.000-1.000) |
| **DP+HBP+ADC+PVP** | 0.585 (0.436-0.735） | 0.579 (0.317-0.830) | 0.536 (0.347-0.724) | 0.584 (0.000-1.000) | 0.587 (0.000-1.000) |
| **PVP+AP+T2+DP** | 0.578 (0.442-0.714) | 0.893 (0.530-0.656) | 0.558 (0.366-0.749) | 0.506 (0.000-1.000) | 0.765 (0.132-1.000) |

To identify the most effective combination of two-sequence feature fusion, we explored seven different combination methods with T1WI as the main branch, and the results are presented in Table 5. Regardless of which ResNet model is used, we find that combination of T1WI+HBP yields the most favorable outcomes.

**2) Determining the optimal fusion combinations**

Table 5: Comparative results obtained using different two sequence fusion strategies via different residual network architectures.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Sequence** | **Model** | **Accuracy (95% CI)** | **AUC (95% CI)** | **F1-score (95% CI)** | **Specificity (95% CI)** | **Sensitivity (95% CI)** |
| T1WI+ADC | ResNet-18 | 0.886 (0.783-0.989) | 0.899 (0.920-0.978) | 0.889 (0.802-0.977) | 0.904 (0.758-1.000) | 0.943 (0.771-1.000) |
| ResNet-34 | 0.942 (0.929-0.956) | 0.940 (0.926-0.954) | 0.934 (0.927-0.941) | 0.946 (0.925-0.967) | 0.948 (0.926-0.969) |
| ResNet-50 | 0.823 (0.798-1.000) | 0.797 (0.697-1.000) | 0.718 (0.698-1.000) | 0.979 (0.941-1.000) | 0.616 (0.600-1.000) |
| T1WI+AP | ResNet-18 | 0.881 (0.734-1.000) | 0.884 (0.770-0.998) | 0.877 (0.769-0.985) | 0.994 (0.978-1.000) | 0.905 (0.706-1.000) |
| ResNet-34 | 0.918 (0.911-0.926) | 0.915 (0.906-0.924) | 0.906 (0.895-0.917) | 0.915 (0.906-0.924) | 0.922 (0.915-0.930) |
| ResNet-50 | 0.881 (0.727-1.000) | 0.869 (0.694-1.000) | 0.845 (0.604-1.000) | 0.993 (0.978-1.000) | 0.746 (0.390-1.000) |
| T1WI+DP | ResNet-18 | 0.923 (0.897-0.949) | 0.928 (0.910-0.946) | 0.920 (0.898-0.943) | 0.999 (0.998-1.000) | 0.902 (0.788-1.000) |
| ResNet-34 | 0.928 (0.909-0.947) | 0.931 (0.914-0.947) | 0.924 (0.905-0.944) | 0.931 (0.914-0.947) | 0.931 (0.912-0.950) |
| ResNet-50 | 0.949 (0.888-1.000) | 0.947 (0.884-1.000) | 0.942 (0.872-1.000) | 0.975 (0.936-1.000) | 0.918 (0.832-1.000) |
| T1WI+HBP | ResNet-18 | 0.932 (0.906-0.958) | 0.937 (0.915-0.959) | 0.929 (0.904-0.954) | 0.897 (0.826-0.968) | 0.977 (0.945-1.000) |
| ResNet-34 | 0.952 (0.945-0.960) | 0.950 (0.943-0.956) | 0.944 (0.937-0.952) | 0.949 (0.944-0.954) | 0.950 (0.943-0.957) |
| ResNet-50 | 0.933 (0.894-0.971) | 0.938 (0.899-0.976) | 0.926 (0.887-0.971) | 0.899 (0.871-0.927) | 0.976 (0.921-1.000) |
| T1WI+PVP | ResNet-18 | 0.900 (0.826-0.973) | 0.903 (0.850-0.956) | 0.894 (0.835-0.953) | 0.884 (0.636-1.000) | 0.922 (0.776-1.000) |
| ResNet-34 | 0.933 (0.903-0.962) | 0.936 (0.910-0.962) | 0.931 (0.904-0.958) | 0.936 (0.910-0.962) | 0.937 (0.907-0.966) |
| ResNet-50 | 0.938 (0.828-1.000) | 0.909 (0.822-0.996) | 0.900 (0.805-0.996) | 0.881 (0.681-1.000) | 0.937 (0.911-0.964) |
| T1WI+T2 | ResNet-18 | 0.913 (0.812-1.000) | 0.907 (0.788-1.000) | 0.895 (0.748-1.000) | 0.975 (0.930-1.000) | 0.841 (0.565-1.000) |
| ResNet-34 | 0.891 (0.875-0.907) | 0.899 (0.876-0.921) | 0.874 (0.853-0.895) | 0.899 (0.876-0.921) | 0.909 (0.887-0.931) |
| ResNet-50 | 0.765 (0.499-1.000) | 0.780 (0.336-1.000) | 0.638 (0.051-1.000) | 0.971 (0.884-1.000) | 0.522 (0.000-1.000) |

1. Performance evaluation on CROI v. s. PROI

To evaluate the algorithm's performance, we conducted repeated experiments on both the CROI test set and the PROI test set. As shown in Table 6 and Fig. 1, the result of T1WI fusion with other sequences is better than T1WI single sequence prediction. Among them, despite using parts of ROI regions, the combination of T1WI+HBP in the ResNet-34 model still performed best, with an Accuracy of 94.2%, AUC of 0.948, and F1-score of 0.943, which powerfully met the high-precision requirements of clinical needs.

Table 6: Comparative results obtained using different two sequence fusion strategies on CROI and PROI testing datasets.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Sequence | Testing set | Accuracy (95% CI) | AUC (95% CI) | F1-score (95% CI) | Specificity (95% CI) | Sensitivity (95% CI) |
| T1WI | CROI | 0.923 (0.909-0.936) | 0.923 (0.911-0.935) | 0.914 (0.899-0.929) | 0.923 (0.911-0.935) | 0.922 (0.906-0.938) |
| PROI | 0.885 (0.871-0.899) | 0.882 (0.870-0.895) | 0.869 (0.855-0.884) | 0.882 (0.870-0.895) | 0.885 (0.872-0.898) |
| T1WI+ADC | CROI | 0.942 (0.929-0.956) | 0.940 (0.926-0.954) | 0.934 (0.927-0.941) | 0.946 (0.925-0.967) | 0.948 (0.926-0.969) |
| PROI | 0.915 (0.904-0.927) | 0.927 (0.916-0.938) | 0.918 (0.905-0.931) | 0.927 (0.916-0.938) | 0.928 (0.915-0.940) |
| T1WI+AP | CROI | 0.918 (0.911-0.926) | 0.915 (0.906-0.924) | 0.906 (0.895-0.917) | 0.915 (0.906-0.924) | 0.922 (0.915-0.930) |
| PROI | 0.909 (0.899-0.920) | 0.922 (0.911-0.932) | 0.914 (0.903-0.925) | 0.922 (0.911-0.932) | 0.927 (0.918-0.937) |
| T1WI+DP | CROI | 0.928 (0.909-0.947) | 0.931 (0.914-0.947) | 0.924 (0.905-0.944) | 0.931 (0.914-0.947) | 0.931 (0.912-0.950) |
| PROI | 0.904 (0.895-0.914) | 0.917 (0.906-0.927) | 0.906 (0.895-0.918) | 0.917 (0.906-0.927) | 0.915 (0.903-0.927) |
| T1WI+HBP | CROI | 0.952 (0.945-0.960) | 0.950 (0.943-0.956) | 0.944 (0.937-0.952) | 0.949 (0.944-0.954) | 0.950 (0.943-0.957) |
| PROI | 0.942 (0.939-0.945) | 0.948 (0.944-0.952) | 0.943 (0.939-0.947) | 0.948 (0.944-0.952) | 0.948 (0.945-0.952) |
| T1WI+PVP | CROI | 0.933 (0.903-0.962) | 0.936 (0.910-0.962) | 0.931 (0.904-0.958) | 0.936 (0.910-0.962) | 0.937 (0.907-0.966) |
| PROI | 0.918 (0.904-0.931) | 0.936 (0.926-0.947) | 0.928 (0.915-0.940) | 0.936 (0.926-0.947) | 0.934 (0.921-0.948) |
| T1WI+T2 | CROI | 0.891 (0.875-0.907) | 0.899 (0.876-0.921) | 0.874 (0.853-0.895) | 0.899 (0.876-0.921) | 0.909 (0.887-0.931) |
| PROI | 0.881 (0.870-0.892) | 0.895 (0.884-0.906) | 0.882 (0.869-0.895) | 0.895 (0.884-0.906) | 0.900 (0.889-0.912) |

While the results of T1WI+HBP on CROI do not significantly differ from other combinations on CROI, the results remain relatively concentrated and stable. We also calculated the standard deviation of the results as Table 7 shown.

Table 7: Comparative standard deviations obtained using different two sequence fusion strategies on CROI and PROI testing datasets.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Testing set | Sequence | Model | Accuracy(SD) | AUC(SD) | F1-score(SD) | Specificity (SD) | Sensitivity(SD) |
| CROI | T1WI+ADC | ResNet-18 | 0.003215 | 0.008718 | 0.003786 | 0.005292 | 0.004726 |
| ResNet-34 | 0.01874 | 0.02003 | 0.0100 | 0.0299 | 0.0299 |
| ResNet-50 | 0.004359 | 0.004933 | 0.005196 | 0.002646 | 0.01212 |
| T1WI+AP | ResNet-18 | 0.01058 | 0.04592 | 0.04335 | 0.1711 | 0.07991 |
| ResNet-34 | 0.01055 | 0.01273 | 0.02572 | 0.03662 | 0.01043 |
| ResNet-50 | 0.06198 | 0.07045 | 0.097 | 0.005859 | 0.1432 |
| T1WI+DP | ResNet-18 | 0.01054 | 0.007211 | 0.008963 | 0.07514 | 0.04607 |
| ResNet-34 | 0.02608 | 0.02302 | 0.03827 | 0.01273 | 0.02637 |
| ResNet-50 | 0.02438 | 0.02551 | 0.02816 | 0.01595 | 0.03478 |
| T1WI+HBP | ResNet-18 | 0.01054 | 0.009 | 0.01002 | 0.02858 | 0.0155 |
| ResNet-34 | 0.01059 | 0.009175 | 0.01065 | 0.006667 | 0.009207 |
| ResNet-50 | 0.01537 | 0.0155 | 0.01686 | 0.01124 | 0.02227 |
| T1WI+PVP | ResNet-18 | 0.02972 | 0.02117 | 0.02381 | 0.09996 | 0.05877 |
| ResNet-34 | 0.04108 | 0.03662 | 0.03827 | 0.03662 | 0.04156 |
| ResNet-50 | 0.04424 | 0.03516 | 0.03837 | 0.08033 | 0.01079 |
| T1WI+T2 | ResNet-18 | 0.041 | 0.04792 | 0.05918 | 0.01804 | 0.1113 |
| ResNet-34 | 0.02259 | 0.03127 | 0.02955 | 0.03127 | 0.0308 |
| ResNet-50 | 0.1071 | 0.1788 | 0.2364 | 0.0351 | 0.2856 |
| PROI | T1WI+ADC | ResNet-18 | 0.0415 | 0.0318 | 0.03508 | 0.05877 | 0.06938 |
| ResNet-34 | 0.01585 | 0.0159 | 0.01833 | 0.0159 | 0.01731 |
| ResNet-50 | 0.1306 | 0.161 | 0.2496 | 0.01501 | 0.3128 |
| T1WI+AP | ResNet-18 | 0.007767 | 0.004041 | 0.005033 | 0.01677 | 0.06188 |
| ResNet-34 | 0.01452 | 0.01482 | 0.01566 | 0.01482 | 0.01337 |
| ResNet-50 | 0.06192 | 0.06352 | 0.09288 | 0.1018 | 0.1179 |
| T1WI+DP | ResNet-18 | 0.01758 | 0.01908 | 0.06093 | 0.04065 | 0.03988 |
| ResNet-34 | 0.0134 | 0.01509 | 0.01611 | 0.01509 | 0.01654 |
| ResNet-50 | 0.09265 | 0.1095 | 0.2226 | 0.0132 | 0.2299 |
| T1WI+HBP | ResNet-18 | 0.003215 | 0.005568 | 0.001528 | 0.002082 | 0.04949 |
| ResNet-34 | 0.004326 | 0.005312 | 0.005481 | 0.005312 | 0.004686 |
| ResNet-50 | 0.08496 | 0.1026 | 0.1306 | 0.01929 | 0.1857 |
| T1WI+PVP | ResNet-18 | 0.014 | 0.01308 | 0.01572 | 0.002082 | 0.02443 |
| ResNet-34 | 0.01918 | 0.01494 | 0.01808 | 0.01494 | 0.01871 |
| ResNet-50 | 0.07271 | 0.08173 | 0.225 | 0.003606 | 0.1671 |
| T1WI+T2 | ResNet-18 | 0.02972 | 0.028 | 0.02836 | 0.04362 | 0.013 |
| ResNet-34 | 0.01503 | 0.01557 | 0.01811 | 0.01557 | 0.01554 |
| ResNet-50 | 0.1781 | 0.2094 | 0.3721 | 0.07731 | 0.3876 |

图示, 工程绘图

描述已自动生成

Fig. 1: Comparative ROC curves obtained using different two sequence fusion strategies on PROI (A) and CROI (B) testing datasets.

**F. Performance evaluation at patient level**

To provide meaningful guidance for doctors, we further integrate results at the image-level to obtain predictive outcomes at the patient-level, and the detailed results are presented in Table 8 and Fig. 2.

Table 8: Comparative results obtained using different two-sequence fusion strategies on CROI testing datasets at patient level are presented.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Sequence | Accuracy (95% CI) | F1-score (95% CI) | AUC (95% CI) | Sensitivity (95% CI) | Specificity (95% CI) |
| **T1WI+ADC** | 0.941 (0.934-0.949) | 0.941 (0.935-0.947) | 0.936 (0.924-0.948) | 0.939 (0.932-0.946) | 0.949 (0.942-0.955) |
| **T1WI+AP** | 0.917 (0.912-0.923) | 0.920 (0.902-0.937) | 0.924 (0.906-0.941) | 0.914 (0.900-0.928) | 0.931 (0.915-0.948) |
| **T1WI+DP** | 0.931 (0.920-0.942) | 0.933 (0.920-0.946) | 0.927 (0.913-0.941) | 0.933 (0.921-0.946) | 0.936 (0.924-0.948) |
| **T1WI+HBP** | 0.952 (0.945-0.960) | 0.951 (0.945-0.957) | 0.945 (0.938-0.952) | 0.952 (0.946-0.957) | 0.950 (0.941-0.959) |
| **T1WI+PVP** | 0.933 (0.924-0.942) | 0.937 (0.930-0.943) | 0.934 (0.926-0.943) | 0.936 (0.930-0.943) | 0.938 (0.930-0.945) |
| **T1WI+T2WI** | 0.891 (0.869-0.912) | 0.898 (0.871-0.924) | 0.885 (0.854-0.916) | 0.897 (0.870-0.924) | 0.908 (0.889-0.928) |

图表, 折线图

描述已自动生成

Fig. 2: Comparative ROC curves obtained using different two sequence fusion strategies at the patient-level