With adding data augmentation we could decrease loss from 0.0664 at epoch 298 to ? at epoch ?.

Despite the initial successes, when the model was applied to different radiotracers, the results were suboptimal, indicating a limitation in its ability to generalize across varying tracer types without further adaptation.

In addition to voxel-wise assessments, model performance was further validated through various statistical tests, including the two-sample Wilcoxon test for comparing image-derived metrics between different training models. To account for multiple comparisons, p-values were adjusted using the Benjamini-Hochberg procedure, with a significance threshold set at 0.05. Statistical significance was determined using paired-sample t-tests across different training modalities, adhering to a p-value significance level of 0.05. The joint histogram analysis, including Pearson correlation, visually represented the voxel-wise SUV correlations for a defined SUV range, offering a granular understanding of the model’s accuracy.

Results

Image-based analysis

The two proposed DL algorithms were evaluated in this study for 68 Ga-PET imaging (INCM and ADCM). We tested the trained DL model with two datasets to evaluate the robustness, which include 8 subjests from 4 different centers as Internal test-set and 12 subjects from an external non-seen test-set.

Figure 3 proviseds quantitative accuracy of the DL attenuation scatter corrected PET ((DL ASC-PET) images to the grounth-trith CT-based attenuation scatter corrected PET (CT ASC-PET) images on all centers. Both DL methods were capable of some degree of attenuation and scattering correction for internal and external centers.

Inja ye chizi bayad bashe vase test amari mesle in : , all three DL methods were capable of some degree of attenuation and scattering correction for different scanners, but Decomposition-based DL significantly outperformed the other two on all scanners (p < 0.025). Specifically, in terms of normalized root mean squared error (NRMS

Badesh natayeje adadi:

improved 47.5% over Direct 2D and 49.1% over Direct 3D on Vision 450, and 60.0% over Direct 2D and 58.4% over Direct 3D on Vision 600, while on both scanners maintained a similar level of error (p = 0.88). When applied to DMI and uMI 780, Decomposition-based DL still outscored the other two by more than 20%.

|  |  |  |
| --- | --- | --- |
|  |  |  |

Results of peak signal-to noise ratio (PSNR) and structural similarity index measurement (SSIM) showed the same tendency as the ? results.

|  |  |
| --- | --- |
|  |  |

Figure ? in appendix 2, shows the relative and Absolute relative percentage error of entire image on all centers, calculated in reference to the CT ASC-PET, which demonstrates that IMCM performed the other on all scanners regarding local metrics as well.

In this investigation, two PET imaging methods—ADCM and IMCM—were analyzed across both external and internal centers. Evaluation metrics included Mean Error (SUV), Mean Absolute Error (SUV), Relative Error (SUV%), Absolute Relative Error (SUV%), Root Mean Squared Error (RMSE), Peak Signal-to-Noise Ratio (PSNR), and Structural Similarity Index (SSIM).

For the external center (C5), the ADCM method yielded a Mean Error (SUV) of -0.631±0.965 (CI 95%: -1.23 to -0.03), Mean Absolute Error (SUV) of 3.072±1.012 (CI 95%: 2.815 to 3.329), and a Relative Error of -8.139±27.364% (CI 95%: -21.76 to 5.48). In contrast, the Multi-Center method demonstrated improved consistency with a Mean Error (SUV) of -1.835±1.387 (CI 95%: -2.80 to -0.87) and a Mean Absolute Error (SUV) of 2.588±0.931 (CI 95%: 2.386 to 2.790). Notably, the Structural Similarity Index for the Multi-Center method at C5 was superior, recorded at 0.879±0.020 (CI 95%: 0.871 to 0.887).

Internal centers (C1 to C4) analyzed collectively showed the ADCM method produced a Mean Error (SUV) of 0.373±1.455 (CI 95%: -0.55 to 1.30) and a Mean Absolute Error (SUV) of 2.343±0.768 (CI 95%: 2.191 to 2.495). The Multi-Center method showed lower Mean Error (SUV) of -0.364±0.841 (CI 95%: -0.76 to 0.03) and Mean Absolute Error (SUV) of 1.415±0.327 (CI 95%: 1.360 to 1.470). The Peak Signal-to-Noise Ratio also favored the Multi-Center method, registering at 35.526±2.117 (CI 95%: 34.9 to 36.2) compared to 38.251±1.923 (CI 95%: 37.6 to 38.9) for the ADCM method.

Significant differences were observed between the methods when evaluated using the Wilcoxon test, with the Multi-Center method consistently showing reduced error magnitudes and higher structural similarity across all centers. These results highlight the Multi-Center method's superiority in precision and reliability for clinical PET imaging applications.

This structured and formal analysis underscores the efficacy of the Multi-Center imaging approach over ADCM, providing a clearer understanding of each method's performance across different centers, which is crucial for optimizing PET imaging protocols in clinical settings.

Statistical Analysis

The Wilcoxon test confirmed significant differences between the ADCM and Multi-Center methods across all metrics, with the Multi-Center method showing consistently lower errors, higher PSNR, and higher SSIM values, indicating superior image quality and more reliable estimations.

Conclusion

The Multi-Center method exhibited enhanced performance compared to the ADCM method across various quantitative metrics, highlighting its effectiveness in providing more accurate and reliable PET imaging. This superior performance, especially noted in the consistency of Relative Errors and Absolute Relative Errors, underlines the potential of the Multi-Center method for clinical applications where precise imaging is critical.

In addition the center-wise quantitative evaluations are available in appendix 2.

In addition to the quantitative evaluation, as shown in fig?, a comparison of a representative imaging example of NA, DL-MAC of all methods, and MAC, as well as joint histogram analysis depicting the correlation between activity concentration of outputs in fig?.

The IMCM provided image quality cpmparable with MAC and preserved more detailed information and less noise was observed compared to   
ADCM.

The joint histogram analysis of an exemplary subject (Fig. 4a) exhibited voxel-wise similarity between reference CT ASC-PET and Decomposition-based DL ASC-PET with slopes of 0.94, 0.97, 1.05, and 0.94 for Vision 450, Vision 600, uMI 780 and DMI respectively. Voxel-wise absolute percentage error map of an exemplary subject depicting the difference between DL ASC-PET and reference CT ASCPET are shown in Supplementary Fig. 3.

Test on external tracers

A graph of a diagram

Description automatically generated with medium confidenceA graph of a diagram

Description automatically generated with medium confidence

A diagram of a scatter plot

Description automatically generated

|  |
| --- |
|  |
|  |
|  |
|  |
|  |

|  |  |
| --- | --- |
| A comparison of ct-a-zoster  Description automatically generated | A comparison of a person's body  Description automatically generated |
| A close-up of x-ray images  Description automatically generated |

A comparison of a person's body

Description automatically generatedA comparison of images of a human body

Description automatically generated

A comparison of images of a person's body

Description automatically generatedA comparison of images of a dog

Description automatically generated

|  |  |
| --- | --- |
|  |  |
|  |  |
|  |  |
|  | |

|  |  |
| --- | --- |
| A close-up of a person's body | A close-up of a person's body  Description automatically generated |
| A close-up of a hand scan  Description automatically generated | A close-up of a person's body |
| A close-up of a person's body  Description automatically generated | A close-up of a scan of a person  Description automatically generated |
| A close-up of a person's body  Description automatically generated | A close-up of a person's body  Description automatically generated |

|  |  |
| --- | --- |
| A close-up of a scan of a dog  Description automatically generated | A close-up of a scan of a person  Description automatically generated |
| A close-up of a person's body  Description automatically generated | A comparison of a baby body  Description automatically generated with medium confidence |
| A close-up of x-ray images  Description automatically generated | A close-up of a person's body  Description automatically generated |
| A close-up of a person's body  Description automatically generated | A close-up of x-ray images  Description automatically generated |
|  |  |
|  |  |
|  |  |
|  |  |

|  |
| --- |
| A comparison of a person's body  Description automatically generated |
| A close-up of a pair of images  Description automatically generated |
| A close-up of a scan of a person  Description automatically generated |
| A close-up of a pair of eyes  Description automatically generated |
| A close-up of x-ray images  Description automatically generated |
| A close-up of a pair of eyes  Description automatically generated |
| A close-up of a person's body  Description automatically generated |
| A close-up of a person's body  Description automatically generated |
| A close-up of a pair of eyes  Description automatically generated |
| A comparison of a person's body  Description automatically generated |
| A close-up of a pair of eyes  Description automatically generated |