



OPEN

Automated F18-FDG PET/CT image quality assessment using deep neural networks on a latest 6-ring digital detector system

Moritz Schwyzer^{1,2,3,4,7}, Stephan Skawran^{1,3,7}, Antonio G. Gennari^{1,3}, Stephan L. Waelti^{1,3,5}, Joan Elias Walter^{1,3}, Alessandra Curioni-Fontecedro^{3,6}, Marlena Hofbauer¹, Alexander Maurer^{1,3}, Martin W. Huellner^{1,3} & Michael Messerli^{1,3}✉

To evaluate whether a machine learning classifier can evaluate image quality of maximum intensity projection (MIP) images from F18-FDG-PET scans. A total of 400 MIP images from F18-FDG-PET with simulated decreasing acquisition time (120 s, 90 s, 60 s, 30 s and 15 s per bed-position) using block sequential regularized expectation maximization (BSREM) with a beta-value of 450 and 600 were created. A machine learning classifier was fed with 283 images rated “sufficient image quality” and 117 images rated “insufficient image quality”. The classification performance of the machine learning classifier was assessed by calculating sensitivity, specificity, and area under the receiver operating characteristics curve (AUC) using reader-based classification as the target. Classification performance of the machine learning classifier was AUC 0.978 for BSREM beta 450 and 0.967 for BSREM beta 600. The algorithm showed a sensitivity of 89% and 94% and a specificity of 94% and 94% for the reconstruction BSREM 450 and 600, respectively. Automated assessment of image quality from F18-FDG-PET images using a machine learning classifier provides equivalent performance to manual assessment by experienced radiologists.

Abbreviations

AUC	Area under the receiver operating characteristics curve
CI	95% Confidence interval
CT	Computed tomography
F18-FDG	2-[¹⁸ F]fluoro-2-deoxy-D-glucose
MIP	Maximum intensity projection
Q.Clear	Block sequential regularized expectation maximization reconstruction algorithm by GE Healthcare, Waukesha, WI
PET	Positron emission tomography
SNR	Signal-to-noise ratio

2-[¹⁸F]fluoro-2-deoxy-D-glucose (F18-FDG)-PET is an imaging modality increasingly used for staging various oncological diseases¹. Image quality in F18-FDG-PET, as well as in other medical imaging modalities, significantly impacts its diagnostic value. Assessment of image quality in PET is commonly achieved using phantoms while applying objective measurements such as number of total photon counts, signal-to-noise ratio (SNR), contrast-to-noise ratio or noise equivalent count rate²⁻⁴. Some studies have shown a correlation between subjectively perceived image quality and noise equivalent count rate⁵. In clinical practice however, image quality may differ depending on the target lesion studied and ultimately the evaluation remains subjective. Generally in PET, image quality is negatively affected by shorter acquisition time and/or lower administered activity⁶. On the other

¹Department of Nuclear Medicine, University Hospital Zurich, Rämistrasse 100, 8091 Zurich, Switzerland. ²Institute of Diagnostic and Interventional Radiology, University Hospital Zurich, Zurich, Switzerland. ³University of Zurich, Zurich, Switzerland. ⁴Institute of Food, Nutrition and Health, Health Sciences and Technology, ETH Zurich, Zurich, Switzerland. ⁵Department of Radiology and Nuclear Medicine, Children’s Hospital of Eastern Switzerland, St. Gallen, Switzerland. ⁶Department of Medical Oncology, University Hospital Zurich, Zurich, Switzerland. ⁷These authors contributed equally: Moritz Schwyzer and Stephan Skawran. ✉email: michael.messerli@usz.ch

hand, decreasing injected activity reduces patient and staff radiation exposure and costs, while decreasing scan duration may additionally increase the modality's availability. An optimization of these parameters specifically requires the capability of fast and reliable image analysis, ideally on-the-fly, while the images are being acquired.

The advent of machine learning in medical imaging has delivered encouraging results, e.g. automated detection of pathology, such as Alzheimer's disease in F18-FDG-PET data⁷, by massively accelerating otherwise time-consuming segmentation tasks⁸ or in broader terms by linking data collected from medical devices from multiple centers enabling collaborative machine learning without mutual raw data exchange^{9,10}. In addition to enabling a timely evaluation of image quality, machine learning has the potential to reduce the notorious subjectivity of the parameter. To our knowledge, a machine learning algorithm has not yet been applied for automated classification of image quality in F18-FDG-PET. Maximum intensity projection (MIP) images serve as a daily driver in the clinical routine enabling physicians to visualize the distribution of F18-FDG and even semi-quantitatively determine the uptake¹¹ therefore constituting a primary target for automated image quality classification tasks.

Accordingly, our study aimed to assess the feasibility of using a machine learning algorithm for automated image quality assessment of F18-FDG-PET images in a cohort with simulated decreasing activity.

Methods

Study population. This retrospective study included patients undergoing clinically indicated oncologic F18-FDG PET/CT between March and April 2021. Written informed consent for the scientific use of medical data was obtained from all patients. The study was approved by the local ethics committee (Kantonale Ethikkommission Zürich, Zurich, Switzerland). All methods were carried out in accordance with relevant guidelines and regulations.

Imaging protocol and image reconstruction. Examinations were performed using a latest generation six-ring digital detector PET/CT scanner (Discovery MI Gen 2, GE Healthcare, Waukesha, WI). A clinical 18F-FDG dosage protocol was used, as previously described in detail¹². Two PET reconstructions were generated using block sequential regularized expectation maximization (Q.Clear, GE Healthcare, Waukesha, WI) with beta-values 450 and 600. To generate PET images with simulated reduced injected 18F-FDG activity, for each patient five sets of PET images were reconstructed using the software supplied with the scanner from identical PET emission data by unlisting "list mode" data, resulting in reduced emission counts equivalent to 120 s, 90 s, 60 s, 30 s and 15 s per bed position. Advantage Workstation Version 4.7 (GE Healthcare, Waukesha, WI) was used to generate maximum intensity projection images in anteroposterior orientation.

Reader assessment of image quality. Two readers (M.M. and S.S., 9 and 6 years of experience in diagnostic imaging) reviewed all MIP images per patient and assigned each image a subjective image quality score from 1 to 4 (1: non-diagnostic; 2: limited diagnostic value; 3: good diagnostic value; 4: optimal diagnostic value) emulating the assessment done according to the institute's standard operating procedure and similarly to previous work¹³. One additional reader (A.G.G., 6 years of experience in diagnostic imaging) measured SNRs of PET image datasets by drawing a semi-automated cubicle volume of interest ($2 \times 2 \times 2 \text{ cm}^3$) in the right liver lobe in the SUV images and dividing the mean standardized uptake value by its standard deviation. The SNR served as a surrogate for objective image quality. Readers were blinded to clinical information.

Automated image quality assessment using machine learning. For automated image quality assessment, the fast.ai deep learning library¹⁴ was used in conjunction with a Res-Net-34¹⁵, a 34-layer residual convolutional network pre-trained on the Image-Net dataset (<https://arxiv.org/abs/1512.03385>). The core of the classifier was built using the fast.ai vision methods (<https://docs.fast.ai/tutorial.vision.html>). "CrossEntropyLoss" was chosen as the loss function and "Adam" (<https://arxiv.org/abs/1412.6980>) was used for gradient-based optimization. For training of the algorithm, 15 learning cycles were set. The library was otherwise empirically used with standard settings as suggested in previous work by Schwyzer et al.^{16,17} who found excellent performance of the algorithm. The MIP datasets from PET images were exported with a total of 400 DICOM files. Each image, according to the readers subjective score, was associated with either a label 1 (=image quality sufficient) when both readers assigned a score greater than 1 or a label 0 (image quality insufficient), when at least one reader assigned a score 1. Imaging sets were split into 10 subsets (8 for training, 1 for validation, 1 for testing) to perform tenfold cross-validation. The results were reported on the test set. Class activation maps were calculated showing which areas of the MIP contributed most to classification. For machine learning computation, a consumer-grade personal computer with an Nvidia GeForce GTX 980 graphics processing unit was used. Performance of the classifier was assessed by calculating the area under the receiver operating characteristic curve (AUC), using reader-based score as target. Sensitivity and specificity were calculated at a cut-off maximizing Youden's index.

Statistical analysis. All statistical analyses were performed in the open-source statistics software R (version 4.1.0, R Foundation for Statistical Computing, Vienna, Austria)¹⁸ including the pROC package. Categorical variables are expressed as frequency distribution and were compared using a Chi-Square-test. Continuous variables are presented as mean \pm standard deviation if normally distributed or median (range) otherwise. Assessment of group differences was determined using an unpaired *t*-test after ensuring a normal distribution of the data using the Shapiro–Wilk test. For non-normally distributed data a Wilcoxon-test was used. For all comparisons, a *p*-value of < 0.05 was defined as statistically significant.

Results

Study cohort. Forty patients were retrospectively included. The mean body mass index was 25.8 ± 5.4 kg/m². The mean injected F18-FDG-activity was 240 ± 60 MBq and images were acquired at 58.7 ± 7.8 min after injection. The most frequent indication for imaging was melanoma (12 / 40, 30%) followed by various squamous cell carcinomas (6 / 40, 15%). Demographic data of the cohort are summarized in Table 1. The algorithm was fed with a total of 400 images ($n = 283$ images with sufficient image quality where a rating of at least “limited diagnostic value” was given by both readers; and $n = 117$ images with insufficient quality where at least one of the two readers assigned the rating “non-diagnostic”). A series of PET images with simulated decreasing activity is given in Fig. 1.

Classification performance of the machine learning classifier. The AUC of the machine learning algorithm for the classification of image quality in images reconstructed with BSREM beta 450 was 0.978 (95% Confidence Interval [CI] 0.962–0.994), while for the BSREM beta 600 reconstruction it was 0.967 (CI 0.931–1.000) (Fig. 2). AUCs were not significantly different between the two reconstruction groups ($p = 0.60$). The sensitivities and specificities for the machine learning classifier were 89% and 94% for BSREM beta 450 images and 95% and 94% for BSREM beta 600 images respectively. Among the incorrectly classified images, those with 30 s (i.e., 25%) simulated acquisition dose per bed were most frequently affected (22/80, 15%) while among the images with full (120 s) acquisition time, none were misclassified (Table 2).

SNR measurements in the liver reached an AUC of 0.975 (CI 0.957–0.993) and 0.975 (CI 0.957–0.993) and were similar to the machine learning classifier’s performance ($p = 0.70$). A correlation plot illustrating the relationship between the machine learning classifier’s result and SNR measurements is given in Fig. 3.

To demonstrate targets of the machine learning classifier’s discriminative ability, examples of class activation maps are given in Fig. 4. In cases deemed “non-diagnostic” by the readers, the maps show preferential activation in noisy areas.

Discussion

In this study, we assessed whether a machine learning classifier can automatically assess the image quality of PET images with simulated decreasing acquisition times resulting in poorer quality. The major findings of our study are as follows: First, machine learning reliably evaluates subjective PET image quality. Second, its classification ability is comparable to manual SNR measurements—a commonly used objective surrogate for image quality. Third, the classifier’s performance did not differ between the different reconstruction settings of BSREM with beta values 450 and 600. Fourth, the classifier most frequently misclassified images in the group of images the experienced readers found to be at the border between non-diagnostic and of limited diagnostic value. This finding further underlines the usefulness of a robust automated classifier in the context of highly subjective visual assessments. Fifth, the classifiers activation maps match noisy areas in non-diagnostic images, ascertaining the validity of the otherwise non-intuitive algorithm.

Female/male, n (%)	15 (37.5%)/25 (62.5%)
Age, years	62 ± 14
Body weight, kg	77.9 ± 19.3
Body height, m	1.73 ± 0.11
BMI, kg/m ²	25.8 ± 5.4
Blood glucose level at time of injection, mg/dl	101 ± 19
Injected F18-FDG activity, MBq	245 ± 60
PET/CT scan time post injection, min	58.7 ± 7.8
Reason for referral	
Melanoma and other skin cancers	12 (30%)
Various squamous cell carcinomas	6 (15%)
Lung cancer	5 (12.5%)
Breast cancer	5 (12.5%)
Lymphoma	4 (10%)
Laryngeal cancer	2 (5%)
Esophageal cancer	1 (2.5%)
Seminoma	1 (2.5%)
Pancreatic cancer	1 (2.5%)
Sarcoma	1 (2.5%)
Cervical cancer	1 (2.5%)
Colorectal cancer	1 (2.5%)

Table 1. Demographic data of study subjects ($n = 40$). Values are given as absolute numbers and percentages in parenthesis or mean \pm standard deviation. *BMI* Body mass index, *MBq* Mega-Becquerel, *PET/CT* positron emission tomography/computed tomography.

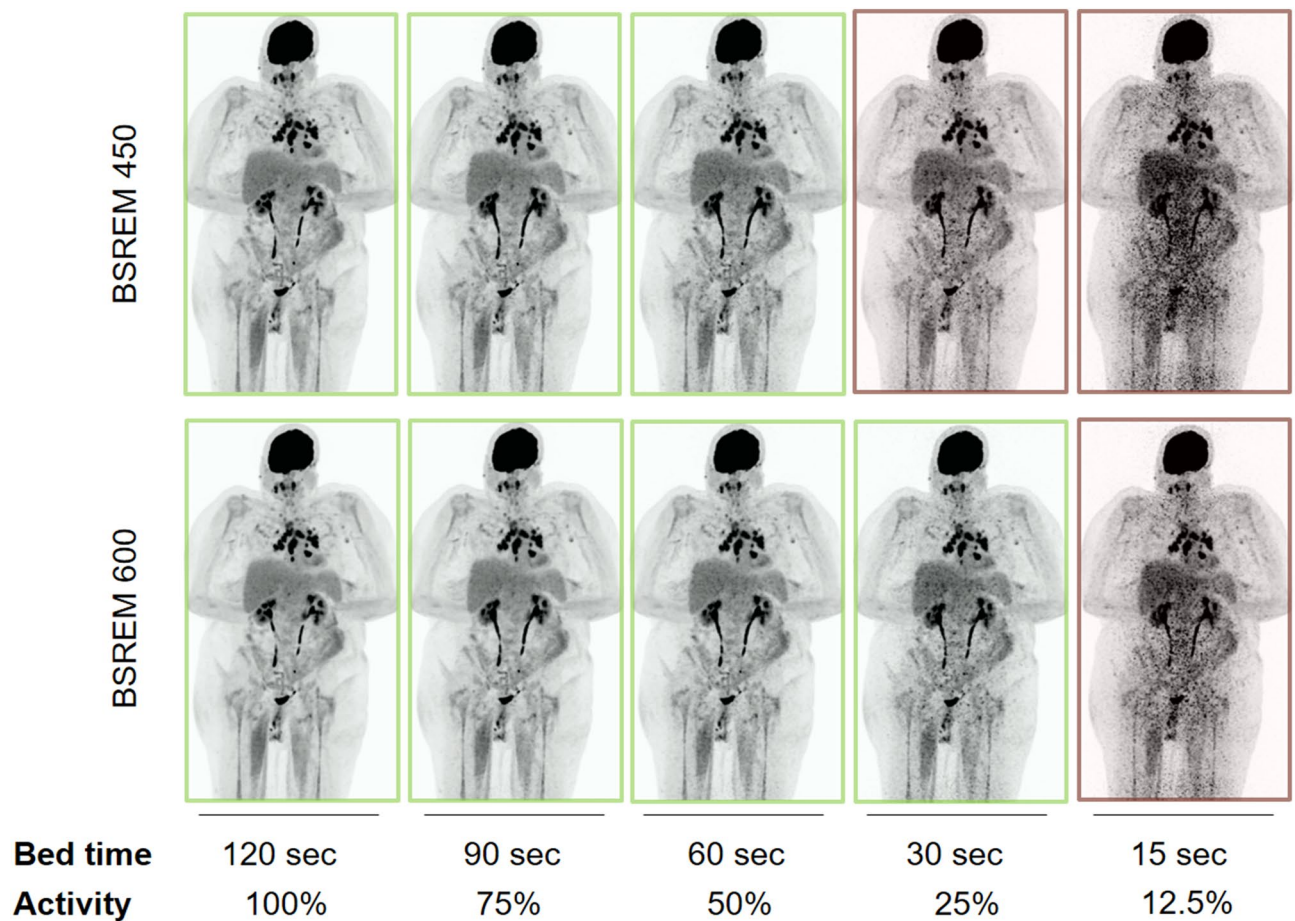


Figure 1. Series of maximum intensity projection PET images with simulated decreasing activity (columns) and the two applied reconstruction setting (rows) using block sequential regularized expectation maximization (BSREM). Images rated sufficient quality by the readers are marked light green, and the images rated as “non-diagnostic” by the readers are marked in light red.

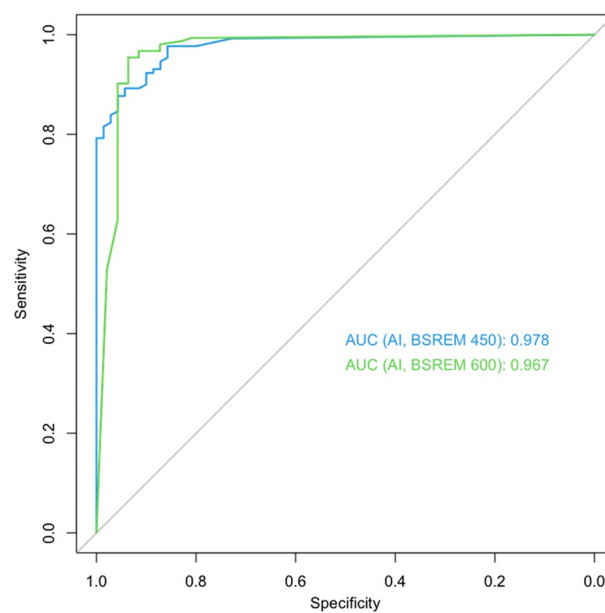


Figure 2. Receiver operating characteristics curves for the performance of image quality classification using the machine learning classifier (AI) grouped by the applied reconstruction setting of block sequential regularized expectation maximization (BSREM) beta 450 and 600, respectively.

Characteristics	Correctly classified ($n = 372$)	Incorrectly classified ($n = 28$)	p value
Patient age, years	62 ± 13	65 ± 14	0.337
BMI	25.7 ± 5.4	27.0 ± 5.0	0.242
Bed time			
120 s	80 (100%)	0 (0%)	
90 s	78 (97.5%)	2 (2.5%)	
60 s	77 (96.3%)	3 (3.8%)	
30 s	68 (85%)	22 (15%)	
15 s	79 (98.8%)	1 (1.3%)	
Reconstruction algorithm			
Q.Clear 450	182 (91%)	18 (9%)	
Q.Clear 600	190 (95%)	10 (5%)	

Table 2. Individual image classification result based analysis of patient and scanning characteristics ($n = 400$). Values are means \pm standard deviations, or frequencies (percentages). *BMI* Body mass index, *Q.Clear* block sequential regularized expectation maximization PET reconstruction algorithm.

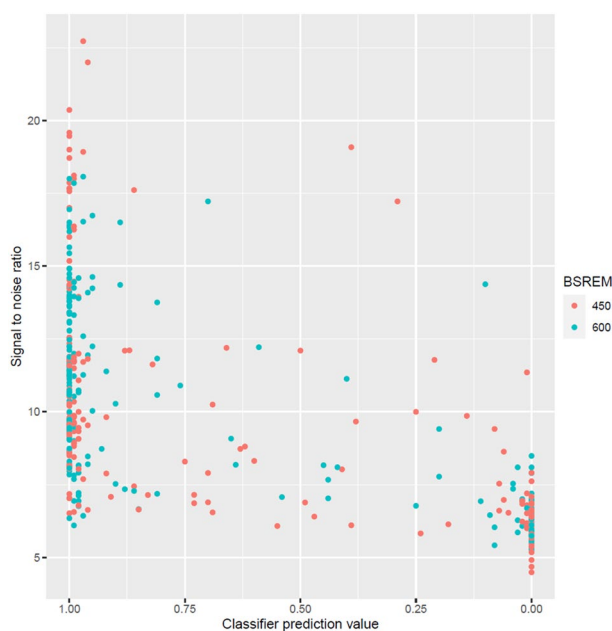


Figure 3. Correlation plot of the machine learning classifiers prediction value and signal to noise ratio measured in liver. Two different reconstruction settings (i.e., block sequential regularized expectation maximization; BSREM) used are marked in red (BSREM beta 450) and green (BSREM beta 600).

Previous studies have shown that acquisition times in PET, within certain limits, can be safely reduced while not compromising image quality^{19–21}, while others proposed a BMI-based dose regimen²² instead of the currently weight-based EANM recommendation²³. All these suggestions at least partly rely on experienced reader-based evaluation of image quality.

The utility of machine learning in PET has been demonstrated by e.g. detection of F18-FDG-PET-positive nodules and lung cancer^{16,17} and—regarding image quality—very recently by denoising F18-FDG-PET-images post-reconstruction²⁴. Yet, it has not been used for estimation of PET image quality itself, a major unmet clinical need in order to perform large-scale image analyses to reduce both dose and scan time while not weakening the diagnostic accuracy of the modality. To our knowledge, our study is the first to exhibit the capability of a candidate machine learning classifier, more specifically deep neural network classifier, aiming to unify a rating task known to be highly subjective. Compared to a radiomics approach used in a previous study including 112 patients the machine learning classifier delivered higher AUCs (AUC 0.978 versus 0.798 in the training and 0.675 in the test dataset³).

If more widely applied, it may have an impact on individualized clinical imaging—e.g., building individualized imaging protocols by iterative on-the-fly evaluation of PET images and adjusting acquisition times until a desired image quality is achieved—and research, establishing an agreed-on standard-of-reference for image

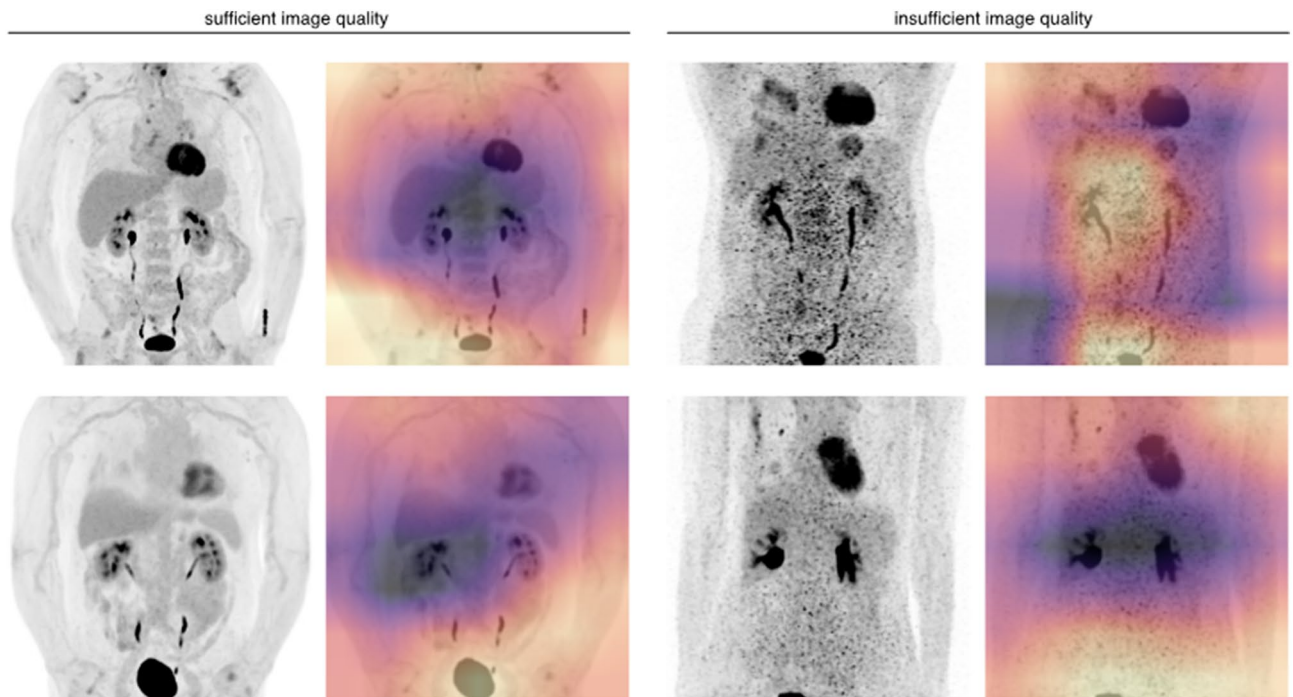


Figure 4. Class activation maps illustrating the machine learning classifier's targets. Maximum intensity projections are shown with color-coded overlays depicting which parts of the images were most important for classification (yellow areas show high class activation). In images with sufficient image quality (first column) the classifier gets preferentially activated in peripheral areas with low noise. In images with insufficient quality (second column), noisy areas are shown to be most relevant for classification.

quality. Moreover, the classifier may be assigned to similar quality assessment tasks in PET imaging whenever qualified subjective judgement is necessary.

Our study has some limitations. Its retrospective scale, the relatively small cohort, and the usage of only one scanner all limit generalizability. Further studies are warranted to ensure wider applicability of our results. Second, image quality assessed in our study is a strictly qualitative and subjective measure. Robustness of quantitative parameters, such as standardized uptake values, needs to be assured by further research. Third, the classifier used in the study was not specifically validated for this task. Further optimization of the classifier might deliver even better performance for evaluation of image quality. Fourth, the standard of reference used to train this algorithm in this study, a reader-based assessment, ultimately remains subjective. To ensure wider applicability and acceptance of an automated assessment, further studies should include a larger number of experienced readers to establish an accepted consensus of image quality.

Conclusion

Our results suggest that machine learning may be used to automatically evaluate PET image quality. An automated, non-reader dependent machine learning based classifier delivers reliable and almost instantaneous PET image quality assessment. This enhances the clinical workflow significantly by potentially enabling on-the-fly evaluated, individually optimized acquisition protocols with reduced activity and acquisition times.

Data availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Received: 1 February 2022; Accepted: 17 June 2023

Published online: 13 July 2023

References

- Hess, S., Blomberg, B. A., Zhu, H. J., Hoiland-Carlson, P. F. & Alavi, A. The pivotal role of FDG-PET/CT in modern medicine. *Acad. Radiol.* **21**, 232–249. <https://doi.org/10.1016/j.acra.2013.11.002> (2014).
- Yan, J., Schaefferkoette, J., Conti, M. & Townsend, D. A method to assess image quality for Low-dose PET: Analysis of SNR, CNR, bias and image noise. *Cancer Imaging* **16**, 26. <https://doi.org/10.1186/s40644-016-0086-0> (2016).
- Reynes-Llompart, G., Sabate-Llobera, A., Llinares-Tello, E., Marti-Climent, J. M. & Gamez-Cenzano, C. Image quality evaluation in a modern PET system: Impact of new reconstructions methods and a radiomics approach. *Sci. Rep.* **9**, 10640. <https://doi.org/10.1038/s41598-019-46937-8> (2019).
- Carlier, T. *et al.* Clinical NECR in 18F-FDG PET scans: Optimization of injected activity and variable acquisition time. Relationship with SNR. *Phys. Med. Biol.* **59**, 6417–6430. <https://doi.org/10.1088/0031-9155/59/21/6417> (2014).

5. Queiroz, M. A., Wollenweber, S. D., von Schulthess, G., Delso, G. & Veit-Haibach, P. Clinical image quality perception and its relation to NECR measurements in PET. *EJNMMI Phys.* **1**, 103. <https://doi.org/10.1186/s40658-014-0103-y> (2014).
6. Karakatsanis, N. A., Fokou, E. & Tsoumpas, C. Dosage optimization in positron emission tomography: State-of-the-art methods and future prospects. *Am. J. Nucl. Med. Mol. Imaging* **5**, 527–547 (2015).
7. Lu, D. *et al.* Multimodal and multiscale deep neural networks for the early diagnosis of Alzheimer's Disease using structural MR and FDG-PET images. *Sci. Rep.* **8**, 5697. <https://doi.org/10.1038/s41598-018-22871-z> (2018).
8. Zhang, H. T. *et al.* Automated detection and quantification of COVID-19 pneumonia: CT imaging analysis by a deep learning-based software. *Eur. J. Nucl. Med. Mol. Imaging* **47**, 2525–2532. <https://doi.org/10.1007/s00259-020-04953-1> (2020).
9. Lian, Z. *et al.* DEEP-FEL: Decentralized, efficient and privacy-enhanced federated edge learning for healthcare cyber physical systems. *IEEE Trans. Netw. Sci. Eng.* **9**, 3558–3569. <https://doi.org/10.1109/TNSE.2022.3175945> (2022).
10. Lian, Z., Zeng, Q., Wang, W., Gadekallu, T. R. & Su, C. Blockchain-based two-stage federated learning with non-IID data in IoMT system. *IEEE Trans. Comput. Soc. Syst.* <https://doi.org/10.1109/TCSS.2022.3216802> (2022).
11. Hotta, M. *et al.* PSMA PET tumor-to-salivary gland ratio to predict response to [(177)Lu]PSMA radioligand therapy: An international multicenter retrospective study. *J. Nucl. Med.* <https://doi.org/10.2967/jnumed.122.265242> (2023).
12. Messerli, M. *et al.* Impact of a Bayesian penalized likelihood reconstruction algorithm on image quality in novel digital PET/CT: Clinical implications for the assessment of lung tumors. *EJNMMI Phys.* **5**, 27. <https://doi.org/10.1186/s40658-018-0223-x> (2018).
13. Delcroix, O. *et al.* Assessment of image quality and lesion detectability with digital PET/CT system. *Front. Med.* **8**, 629096. <https://doi.org/10.3389/fmed.2021.629096> (2021).
14. Howard, J. & Gugger, S. Fastai: A layered API for deep learning. *Information* **11**, 108. <https://doi.org/10.3390/info11020108> (2020).
15. 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* (2016).
16. Schwyzer, M. *et al.* Automated detection of lung cancer at ultralow dose PET/CT by deep neural networks - Initial results. *Lung Cancer* **126**, 170–173. <https://doi.org/10.1016/j.lungcan.2018.11.001> (2018).
17. Schwyzer, M. *et al.* Artificial intelligence for detecting small FDG-positive lung nodules in digital PET/CT: Impact of image reconstructions on diagnostic performance. *Eur. Radiol.* **30**, 2031–2040. <https://doi.org/10.1007/s00330-019-06498-w> (2020).
18. R: A language and environment for statistical computing. (Vienna, Austria, 2019).
19. Hausmann, D. *et al.* The impact of acquisition time on image quality in whole-body 18F-FDG PET/CT for cancer staging. *J. Nucl. Med. Technol.* **40**, 255–258. <https://doi.org/10.2967/jnmt.112.103291> (2012).
20. Hu, P. *et al.* Total-body (18)F-FDG PET/CT scan in oncology patients: How fast could it be?. *Eur. J. Nucl. Med. Mol. Imaging* **48**, 2384–2394. <https://doi.org/10.1007/s00259-021-05357-5> (2021).
21. Tragardh, E. *et al.* Impact of acquisition time and penalizing factor in a block-sequential regularized expectation maximization reconstruction algorithm on a Si-photomultiplier-based PET-CT system for (18)F-FDG. *EJNMMI Res.* **9**, 64. <https://doi.org/10.1186/s13550-019-0535-4> (2019).
22. Xiao, J. *et al.* Can the BMI-based dose regimen be used to reduce injection activity and to obtain a constant image quality in oncological patients by (18)F-FDG total-body PET/CT imaging?. *Eur. J. Nucl. Med. Mol. Imaging* <https://doi.org/10.1007/s00259-021-05462-5> (2021).
23. Boellaard, R. *et al.* FDG PET/CT: EANM procedure guidelines for tumour imaging: Version 2.0. *Eur. J. Nucl. Med. Mol. Imaging* **42**, 328–354. <https://doi.org/10.1007/s00259-014-2961-x> (2015).
24. Ly, J., Minarik, D., Jogi, J., Wollmer, P. & Tragardh, E. Post-reconstruction enhancement of [(18)F]FDG PET images with a convolutional neural network. *EJNMMI Res.* **11**, 48. <https://doi.org/10.1186/s13550-021-00788-5> (2021).

Acknowledgements

Dr. Stephan Skawran and Dr. Moritz Schwyzer are supported by a grant from the Palatin-Foundation, Switzerland. Dr. Moritz Schwyzer is funded by the MedLab Fellowship at ETH Zurich. Dr. Joan Walter received research grants from the Swiss Heart Foundation (FF19097 and F18111) as well as the Swiss Academy of Medical Sciences and the Gottfried and Julia Bangerter-Rhyner Foundation. Dr. Michael Messerli received a research grant from the Iten-Kohaut Foundation, Switzerland. Dr. Michael Messerli and Dr. Martin W. Huellner are supported by a grant from the CRPP AI Oncological Imaging Network of the University of Zurich. Dr. Martin W. Huellner received grants from GE Healthcare and a fund by the Alfred and Annemarie von Sick legacy for translational and clinical cardiac and oncological research.

Author contributions

M.S.: Conceptualization, methodology, software, formal analysis, writing – review and editing, visualization; S.S.: Investigation, formal analysis, methodology, data curation, formal analysis, writing – original draft and review and editing, visualization; A.G.G.: Data curation, writing – review and editing; S.L.W.: Validation, writing – review and editing; J.E.W.: Conceptualization, validation, writing – review and editing; A.C.-F.: Validation, writing – review and editing; M.H.: Resources, data curation; A.M.: Validation, writing – review and editing; M.W.H.: Validation, formal analysis, resources, writing – review and editing; M.M.: Project administration, supervision, resources, conceptualization, formal analysis, writing – review and editing, visualization.

Competing interests

Dr. Martin W. Huellner received grants from GE Healthcare unrelated to the current research. Other than that, the authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to M.M.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023