RESEARCH

"*A net for everyone*": fully personalized and unsupervised neural networks trained with longitudinal data from a single patient

Christian Strack[1,2,3*], Kelsey L. Pomykala[1], Heinz-Peter Schlemmer[2,4], Jan Egger[1,5] and Jens Kleesiek[1,4,5,6]

Abstract

Background With the rise in importance of personalized medicine and deep learning, we combine the two to create personalized neural networks. The aim of the study is to show a proof of concept that data from just one patient can be used to train deep neural networks to detect tumor progression in longitudinal datasets.

Methods Two datasets with 64 scans from 32 patients with glioblastoma multiforme (GBM) were evaluated in this study. The contrast-enhanced T1w sequences of brain magnetic resonance imaging (MRI) images were used. We trained a neural network for each patient using just two scans from different timepoints to map the difference between the images. The change in tumor volume can be calculated with this map. The neural networks were a form of a Wasserstein-GAN (generative adversarial network), an unsupervised learning architecture. The combination of data augmentation and the network architecture allowed us to skip the co-registration of the images. Furthermore, no additional training data, pre-training of the networks or any (manual) annotations are necessary.

Results The model achieved an AUC-score of 0.87 for tumor change. We also introduced a modified RANO criteria, for which an accuracy of 66% can be achieved.

Conclusions We show a novel approach to deep learning in using data from just one patient to train deep neural networks to monitor tumor change. Using two different datasets to evaluate the results shows the potential to generalize the method.

Keywords Neural networks, Personalized, Wasserstein-GAN, Unsupervised, Machine learning, Privacy-safe, Zero-training data, Longitudinal, Brain Tumor, MRI

*Correspondence:
Christian Strack
c.strack@dkfz-heidelberg.de
[1]Institute for AI in Medicine (IKIM), University Hospital Essen (AöR),
Girardetstraße 2, 45131 Essen, Germany
[2]Division of Radiology, German Cancer Research Center (DKFZ),
69120 Heidelberg, Germany
[3]Medical Faculty Heidelberg, Heidelberg University, 69120 Heidelberg,
Germany
[4]German Cancer Consortium (DKTK), Partner Site Essen, Hufelandstraße
55, 45147 Essen, Germany
[5]Cancer Research Center Cologne Essen (CCCE), University Medicine
Essen, Hufelandstraße 55, 45147 Essen, Germany
[6]Department of Physics, TU Dortmund University, Otto-Hahn-Straße 4,
D-44227 Dortmund, Germany

Strack *et al. BMC Medical Imaging*   (2023) 23:174

Page 2 of 9

Introduction

One key difference between human and artificial intelligence is the number of training examples needed to generate knowledge. Whereas humans can learn to recognize new objects with only a few examples, most machine learning tasks require hundreds of examples for the same task. In fact, increasing the dataset size is often a key step in improving the performance of a machine learning model. ImageNet [1], the most famous dataset in computer vision, now consists of over 14 million training examples. The state-of-the-art models in computer vision are often trained on large datasets such as ImageNet and may not transfer well to smaller datasets of different tasks. Getting large datasets may not always be a feasible approach though, especially in the medical domain.

Gathering large datasets is one of the key challenges of medical deep learning applications. Keeping a patient's medical information safe is critical and there are laws protecting it in most countries. This makes it more difficult to get the data and leads to the medical datasets being much smaller compared to traditional computer vision datasets. Additionally, deep neural networks themselves offer another privacy threat. It has been shown that training examples of fully trained networks can be recovered with a model inversion attack [2]. This makes it more difficult to publish medical deep learning applications as the patient's privacy can not be guaranteed. These two reasons give a big incentive to find ways to train neural networks with smaller datasets or even just one patient's data.

There have been several models proposed to challenge the task of reducing the number of training examples. One-shot learning is a method of learning a class from only one labeled example [3]. Siamese neural networks are able to determine if two images show the same person, even if they have never seen images of that person before [4]. They have also been used in medicine to distinguish between chronic obstructive pulmonary disease and asthma [5]. Whereas new classes can be learned from as little as one example, one-shot learning still requires thousands of training examples of other classes beforehand. Furthermore, anomaly detection can be used to detect classes of rare occurrence. This is a technique used to recognize items which do not lie in the usual data distribution and makes use of unsupervised learning in most cases [6]. Anomaly detection usually makes use of learning the data distribution in a healthy population and identifying the anomalies, i.e. a disease, of a new class. Another method to handle small datasets is transfer learning, where networks trained on large datasets are used as a starting point to train on training examples of new classes. Transfer learning makes use of the fact that features learned on the large dataset can be reapplied to new data.

In this paper, we introduce personalized neural networks, which use only one patient's data for training. Our proposed method only needs two MRIs from the same patient and no additional pretraining. This also results in a privacy-safe processing of the data, because the data "stays" within the same patient. Our model is based on generative adversarial networks (GANs) [7]. GANs have gained in popularity in recent years in the medical AI community. Originally used for image synthesis, there have been applications to generate medical images [8, 9]. Other studies focus on classification or segmentation tasks [10, 11]. We apply the personalized neural networks on subjects with brain tumors.

Brain tumors belong to the most devastating diagnoses, in particular for a confirmed glioblastoma multiforme (GBM) [12]. Despite massive research efforts and advancements in other cancer types, like breast cancer [13] or prostate cancer [14], the life expectancy of a confirmed GBM with treatment, including chemotherapy, radiotherapy and surgery, is still only around one year [15]. Nevertheless, disease progression and treatment decisions are strongly dependent on maximum tumor diameter and tumor volume, as well as the corresponding morphological changes during a treatment period. The imaging method of choice here is magnetic resonance imaging (MRI). However, MRI does not provide any semantic information for brain structures or the brain tumor per se. This has to be done manually, semimanually or automatically, in a post-processing step, commonly referred to as a *segmentation*. Manually performed, however, a segmentation is very time-consuming and operator-dependent, especially when performed in a three-dimensional image volume [16], which needs slice-by-slice contouring. Hence, an automatic (algorithmic) segmentation is desired, especially when large quantities of data volumes have to be processed. Even if it is still considered an unsolved problem, there has been steady progress from year to year; and data-driven approaches, like deep neural networks, currently provide the best (fully automatic) results. However, a segmentation with a data-driven approach, like deep learning [17], comes with several burdens: Firstly, the algorithm generally needs massive annotated training data. Additionally, for inter-patient disease monitoring, several segmentations have to be performed, and in addition, these scans have to be registered to each other (which also adds uncertainty to the overall procedure, especially when deformed soft-tissue comes into play [18]). In this regard, we want to tackle these problems with a personalized neural network that needs just the patient's data, no annotations and no extra registration step.

We apply the personalized networks to longitudinal datasets of glioblastoma. To the best of our knowledge, this is the first study using this little training data to

train a deep neural network in the medical domain. The method addresses the issues of gathering big datasets in medicine and producing a privacy-safe network. The approach is considered as unsupervised learning as no data annotation is necessary. Using a Wasserstein GAN, the model creates a map showing the differences between images from two timepoints. We evaluate the model with an receiver operating curve (ROC) analysis as well as a modified RANO criteria on two different datasets of longitudinal MRI images of patients with glioblastoma.

Methods

Model architecture and training

The neural network architecture used in this study is based upon Wasserstein GANs [19]. This is a modified version of GANs [7]. These are a form of deep neural networks in which two sub-models are trained adversarily in a sum-zero game. A generator is trained to create new images, whereas a discriminator is trained to distinguish between real and synthetic images. In Wasserstein GANs, the discriminator is modified to a critic function which leads to more stable training [19].

Our network architecture is similar to the model used by Baumgartner et al. [20]. The aim of the network is to create a map which transforms an image from the first timepoint (t1) to the second timepoint (t2). This will make the model learn to represent the changes between the images, more specifically tumor growth/reduction in our case. To do this, augmented versions of the image at t1 are used as input to the generator. The generator will try to create a map that, when added to the input image creates an image of t2. The critic will try to distinguish these generated synthetic t2 images from the real t2

images. Thereby forcing the generator to learn the differences between the two timepoints.

The generator is based on the U-Net [21] structure. The U-Net is a fully convolutional network consisting of a contracting path (encoder) and an expanding path (decoder) with skip connections at each resolution level. It produces an output image of the same size as the input image. The network structure is shown in more detail in Fig. 1. A random slice of the third dimension was taken during each training step, such that the network received an input size of 256×256 pixels. For the ultimate prediction after training, the result for each of the 128 slices was calculated, saved and concatenated to the final 256×256×128 pixels volume. The critic function is also a fully convolutional network. Like in Baumgartner et al. [20], we used an architecture similar to the C3D network [22]. This is an encoder type architecture which produces a single value output (Figure S1 in the Supplementary Materials).

The network was trained for 1000 epochs. In every epoch we updated the critic five times before updating the generator. In the first 25 epochs and every 100 epochs, the critic was updated 100 times. We used gradient penalty and the ADAM optimizer during training [23, 24]. Figure 2 gives an overview over the whole training process.

During training, we discovered that the training process could be unstable when the two images were too similar or even identical. It could lead to the critic not being able to distinguish the real and the fake images at all and thus not providing any valuable feedback to the generator. We therefore added a small square of 10×10 pixels of noise to a fixed position in one of the images. The noise was created by transforming gaussian noise
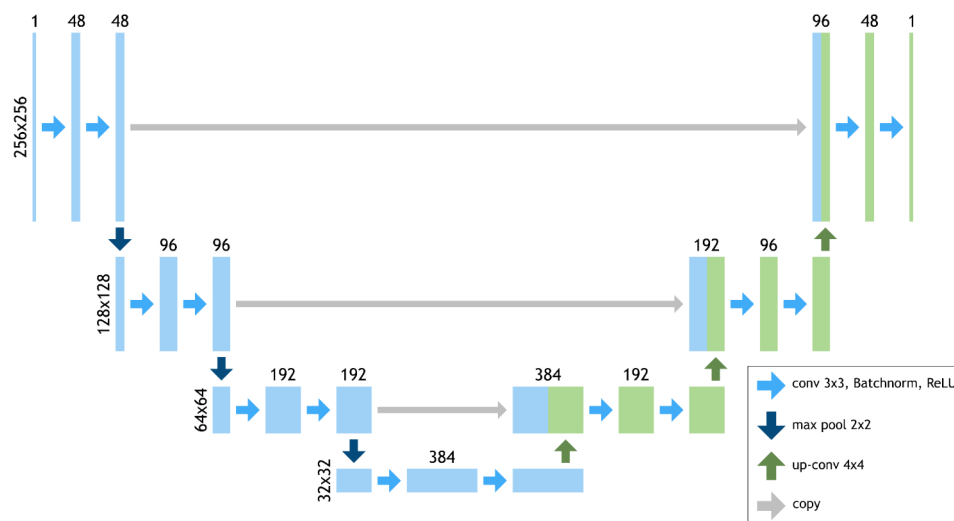

Fig. 1 Architecture of the generator network. A U-Net structure is used. At each level there are two blocks of 3×3 convolution, batch normalization (BN) and ReLU. 2×2 Max pool functions are used for downsizing. 4×4 transposed convolutions with stride = 2 are used for upsizing. The size of the image at each level is shown on the left. The number of features in each block is shown on the top of the block
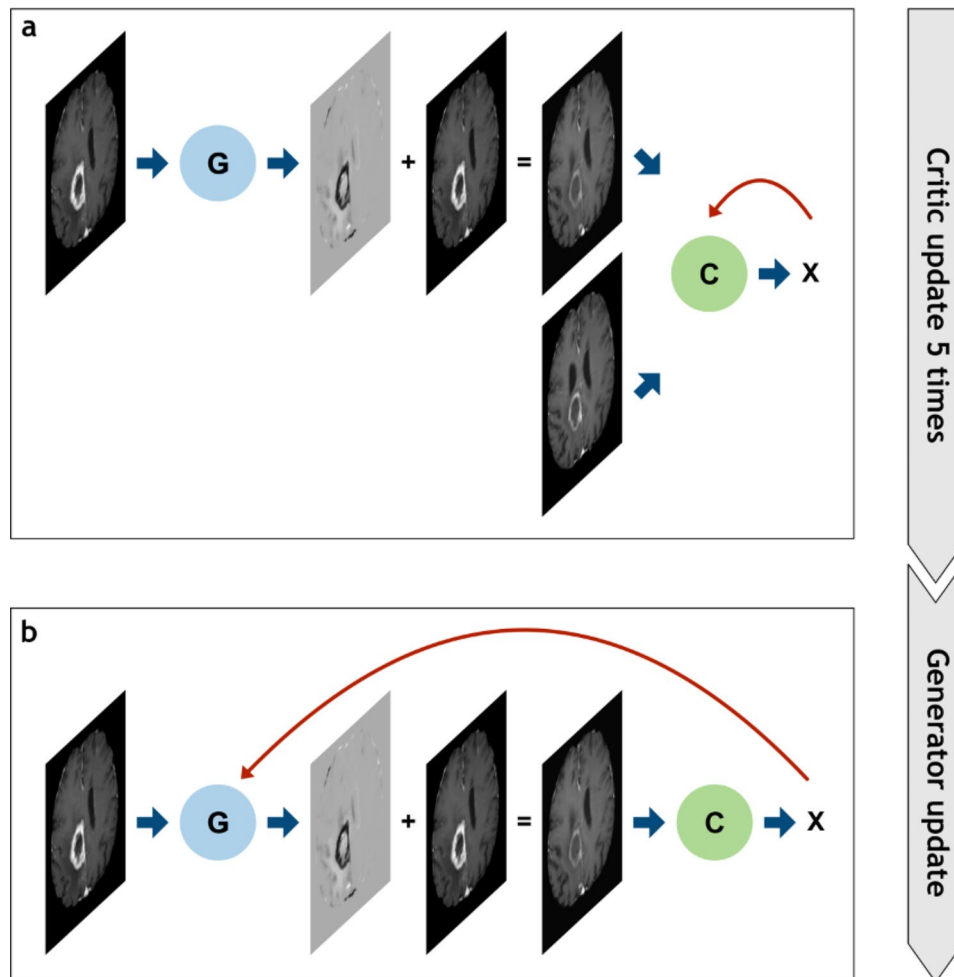
Fig. 2 Overview over one training epoch. In (**a**) the critic function is trained. A t1 image is passed through the generator. The generator's output is a map which gets added to the t1 image. This produces the fake t2 image. The real and the fake t2 images are then passed to the critic. The output of the critic is incorporated into a loss function and backpropagated to update the weights of the critic network. In (**b**) the generator is trained. Again, a t1 image is passed to the generator. The output is added to the t1 image to create the fake t2 image. This is passed to the critic. The output is incorporated into the generator loss function and backpropagated through both networks to update the generator network

with a 2D gaussian filter. The position of the noise was changed twice during the training (after 40% and 60% of all the training epochs). The concrete positions were at 50%, 35% and 65% of the size of the input image in both dimensions. Finally, after 80% of the epochs, the noise was removed completely for the rest of the training. After each change of the position of the noise, a model was saved. After the training finished, we took an ensemble of all the models, averaging over the results, disregarding those pixels that had been artificially changed in that part of the training.

Preprocessing
There were several preprocessing steps in this study. First, all images were resampled to 256×256×128 pixels. In MRIs, the pixel values obtained differ for identical tissues when different scanners are used. To deal with this problem, we histogram matched the images to each other.

This was done using the histogram matching tool of 3D Slicer [25]. Next, the images were normalized to a range between 0 and 1. The brain of the patient was centered in the image. Lastly, we skull-stripped the scans using the HD-BET tool to remove any non-brain tissue [26].

Augmentation
GANs usually take a lot of data to train effectively [27, 28]. However, in this study, only two images of size 256×256×128 pixels were used. The use of data augmentation was therefore crucial. We used the batchgenerators framework for this task [29]. Since our model does not require co-registered images, this had to be accounted for in the data augmentation. Hence, we shifted and rotated the images in all three dimensions such that the network learns the representation of the brain in space. Each training image was randomly rotated between −15° and 15° and shifted between 0 and 10 pixels in all

three dimensions. Lastly, gaussian noise was added to all images with zero mean and the variance ranging uniformly between 0 and 0.1.

Data

In this study two different datasets were used. The first was a local dataset including longitudinal follow-up scans from 15 patients diagnosed with recurrent Grade IV glioblastoma. As described in Kleesiek et al. [30], the baseline scan was defined as the scan before de novo treatment after tumor recurrence. The image resolution was $256 \times 256 \times 128$ pixels. There were 13 male and 2 female patients with a mean age of 55.1 years. Image acquisition was performed on a 3 Tesla MRI scanner (Magnetom Verio, Siemens Healthcare, Erlangen, Germany).

The second was a publicly available dataset from the Cancer Imaging Archive (TCIA) [31], called Brain-Tumor-Progression [32]. This dataset includes two multi-channel MRIs each for 20 patients newly diagnosed with glioblastoma. The resolution of the images varied between $260 \times 320 \times 21$ and $512 \times 512 \times 24$ pixels. The parameters of the model were fine tuned solely on the first three patients of this dataset, therefore only the last 17 patients were included in the evaluation. For both datasets only the T1-contrast-enhanced (T1ce) channels were used in this study.

Segmentation network for ground truth

To evaluate the proposed model's performance, ground truth segmentations were created. We used the neural network of the winner of the 2020 BraTS challenge for brain tumor segmentation for this task [33]. The segmentations contain three classes: enhancing tumor, edema and necrosis. Only the enhancing tumor class was used in this paper.

RANO classification

To further evaluate our model, we predicted a modified RANO classification. The RANO criteria for glioma is a radiological classification used to evaluate the treatment of glioblastoma [34]. We slightly modified this grading to allow for a classification using just the total enhancing tumor volume and disregarding any clinical information. The two classes, complete and partial response, were combined into one class called response. This class is defined as a reduction in tumor volume of more than 50%. Progression is defined as a growth in tumor volume of 25% or more. Consequently, stable disease is a change in tumor volume not corresponding to response or progression. The tumor volume was calculated in voxels.

The segmentations created by the BraTS network were again used to calculate the ground truth. Since the maps often showed a lot of noise at the edge of the brain, as shown in Fig. 3, the outer 10 pixels in each dimension were disregarded. While this is potentially harmful for tumors at the edge of the brain, the advantages of removing the noisy regions outweigh the disadvantages. We created additional ternary maps from our network with just the three classes -1, 0 and 1. Voxels with a value smaller than -0.15 were defined as -1, showing tumor reduction and voxels with a value bigger than 0.15 were defined as 1, showing tumor growth. Classifications with a connected voxel count of 30 or less were set to 0 to remove some noise. The ternary map of each patient was added up to get the absolute change in tumor volume. This was added to the total tumor volume of the first timepoint to predict the volume of the second timepoint.

Results

Qualitative Assessment and Heatmaps

Figure 3 displays representative examples from both datasets. The map shows the changes in contrast-enhancing tumor in a reliable manner. The regions of tumor growth are represented as black (values<1 in the map). The regions of tumor reduction are represented as white (values>1 in the map). Converted to heatmaps they can be used to highlight the key regions of tumor growth/ reduction.

As one can see, there are some recurring regions of noise in the maps. For example, the region next to the ventricular system is incorrectly noted as changed in either direction in most cases. Additionally, the edge of the brain often contains a lot of noise, as highlighted in Fig. 3C. This can be a problem for tumors located at the edge of the brain or the ventricles.

ROC analysis

An ROC analysis was performed to evaluate the model's prediction accuracy. The segmentations created by the BraTS network were used to calculate the ground truth. To get the classes tumor growth and reduction, the segmentation of the first time point was subtracted from the second time point.

The 2-class ROC analysis is shown in Fig. 4. The area-under-the-curve (AUC) for tumor growth and reduction is 0.72 and 0.94 respectively for the public dataset. The AUC is 0.94 and 0.94 for growth and reduction for the private dataset. The total AUC for both datasets combined is 0.87 and 0.86 respectively (see Figure S2 in the Supplementary Materials). The micro-average AUC is 0.87.

RANO classification

The results for the RANO classification are shown in Table 1. The overall sensitivity and specificity for the modified RANO classes were 65.5% and 82.8% respectively. The total accuracy was 65.5%. The accuracy was calculated in a one-vs-all approach with regards to a
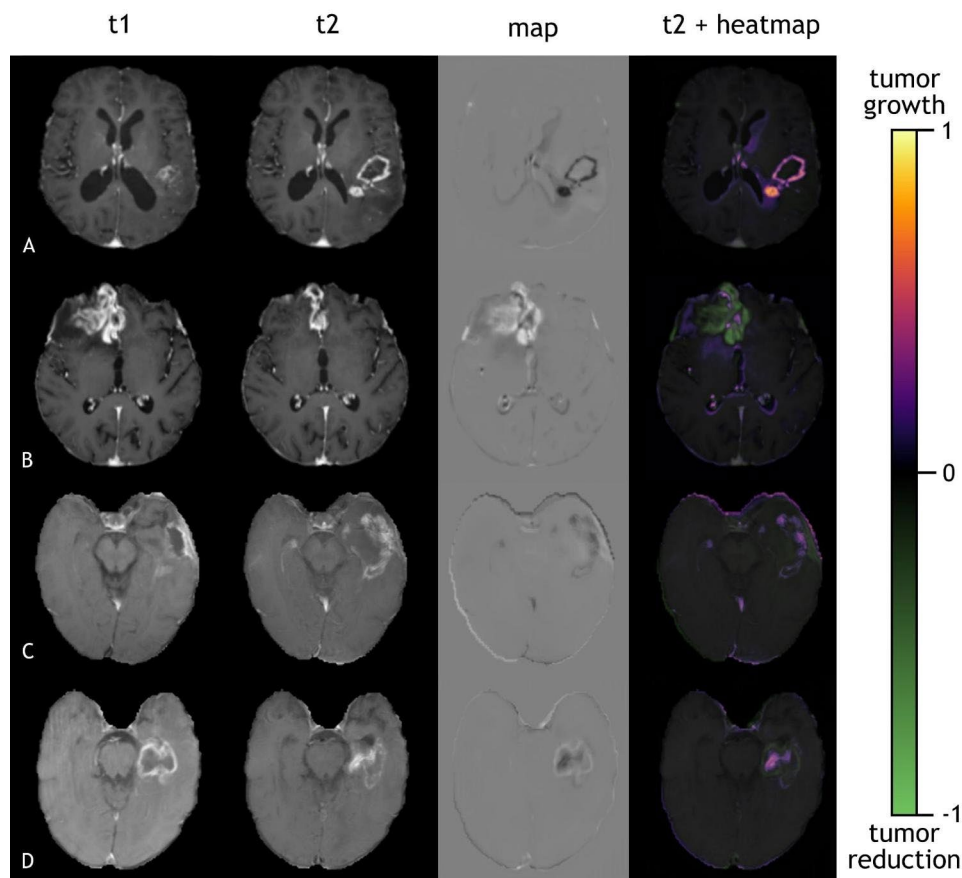
Strack *et al. BMC Medical Imaging*   (2023) 23:174

Page 6 of 9

Fig. 3 Examples of the T1ce images at different timepoints along with the calculated map. The last column shows heatmaps on top of the second time point to highlight key regions of change. A and B are from the local dataset, C and D are from the public dataset

multi-label classification. The overall scores were calculated as a micro-average of all the classes. The performance for the two datasets was comparable (see Table S1 in the Supplementary Materials).

Discussion

In this contribution, we propose "*A net for everyone*", a personalized neural network that is trained with longitudinal data from a single patient. We designed and implemented a Wasserstein-GAN-based approach that works with only two scans from the same patient without any extra training data in an unsupervised fashion. That means, our method does not need any small or large quantities of datasets, and also does not need any manual or semi-manual annotations for training.

Alongside a qualitative evaluation, we show that the model achieves a high AUC in an ROC analysis, when compared to a state-of-the-art deep learning model. It also shows that the model's performance for tumor growth and tumor reduction is very similar. The accuracy for the local dataset was significantly larger than for the public dataset. This can be explained by the difference in quality, as the public data was older and had a lower

resolution, especially in the third dimension. Additionally, there were artifacts in some of the images, like parts of the brain were cut off. We implemented a modified RANO criteria, resulting in a combined accuracy of 66%. The generated heatmaps can aid in the diagnostic process to quickly find the key regions of interest.

It should be noted that the performance of deep learning models usually scales with the size of the dataset [35]. Therefore, this approach has an inherent disadvantage compared to classical supervised learning models with big datasets. However, using only the data of one patient comes with some advantages. First, our method is a privacy-safe approach. Medical records and medical image data are very sensitive and our approach stays within the same patient for the algorithmic training and execution. Second, getting large datasets in medical imaging has proven to be a challenging task due to these privacy concerns, and our method does not rely on this.

Furthermore, no registration is necessary for the training of our approach, which is a mandatory and crucial step in most approaches [36]. There are different methods for image registration, with some being completely automatic and others needing some manual input [37].
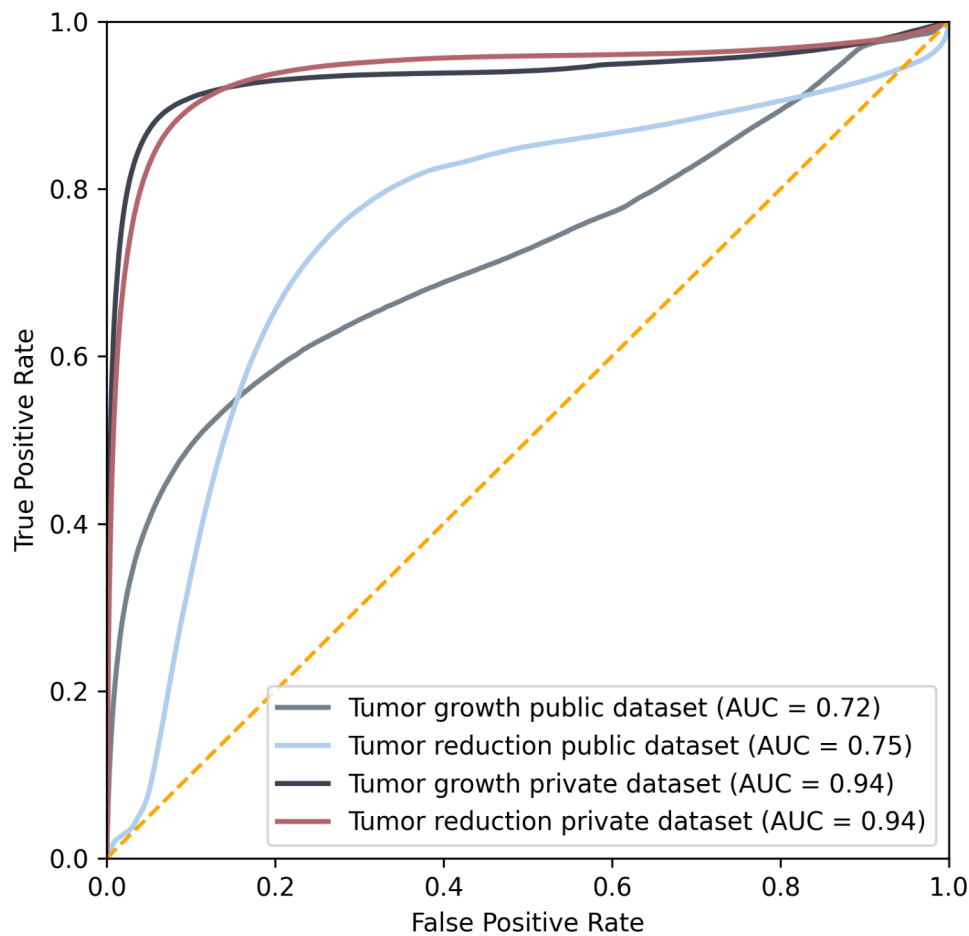
Fig. 4 ROC Analysis for the prediction of tumor change compared to the ground truth of BraTS winning network nnUNet.

Table 1 Sensitivity, Specificity, Accuracy of the prediction of modified RANO criteria for glioblastoma

RANO category	Sensitivity	Specificity	Accuracy
Response	70.0%	100.0%	90.6%
Stable disease	80.0%	63.6%	68.8%
Progression	50.0%	85.0%	71.9%
Total	**65.6%**	**82.8%**	**65.6%**

While these registration methods can be accurate for scenarios, like rigid registrations, especially deformable registrations are still challenging and there are problems with outliers [38]. These include post-surgery scans or patients with a different anatomy due to a large tumor. Both could lead to registration artifacts, which would compromise the further training. Our model does not need a separate registration step, avoiding these potential sources of errors.

The model does not explicitly learn to recognize changes in the tumor, but learns to recognize any changes between two images. However, since the contrast enhancing regions of the tumor are typically amongst the most intense regions in a T1ce scan, changes in these regions are particularly visible in the created maps, highlighting changes in tumor enhancement patterns. However, the proposed approach comes with two disadvantages that can be addressed in future research. First, any structural change in the brain not lying in the tumor will be recognized by the model. For example, a midline shift caused by tumor growth will cause changes in healthy regions of the brain and might be interpreted as growth or reduction of contrast enhancing tumor. This can also be interpreted as an advantage to point out all changes to the reader. Second, the model is prone to noise at the edge of the brain and next to the ventricles. The ventricles differ between two scans depending on the current cerebral spinal fluid volume. At the edge of the brain, the two scans also differ slightly due to the skull stripping. Another reason is the variance in size of the dural venous sinuses. To account for the noise at the edge of the brain, we disregarded the outer pixels in the calculation of the modified RANO criteria. This is obviously a concern for tumors located in the cortex of the brain as it might cut out regions of the tumor. However, glioblastoma are typically located in the centrum semiovale, thus in most cases this should not be a problem [39].

Strack *et al. BMC Medical Imaging*   (2023) 23:174

Page 8 of 9

It should be noted that the ground truth from this work was not created by medical experts but by a neural network. However, the network used achieved a Dice Score for the enhancing tumor of 82% [33]. This lies within the range of the inter-rater variability of human raters of 74–85% [40], suggesting that medical experts would not change the ground truth significantly.

However, despite the above-mentioned limitations, this study is a proof of concept that personalized neural networks can serve as a privacy-safe method to analyze longitudinal imaging data of a single patient in an unsupervised fashion. It has been shown that tumor growth tends to get underestimated on average and overestimated for very small tumors in brain tumor measurements in the current RANO criteria [41, 42]. Therefore, having an efficient method for measuring the 3D tumor volume is necessary for treatment monitoring and surgical planning [43, 44]. Lastly, the produced heatmaps can be a big help in the diagnosis of the MRI images, as they lead the reader directly to the key regions of changes.

Summarized, we proposed a deep learning architecture to create personalized neural networks. This study serves as a proof of concept to show that training data from just one patient can be used to monitor tumor change in longitudinal MRI scans. Areas of future work include the application to other pathologies, such as aortic aneurysms and aortic dissections [45], where disease monitoring over several image acquisitions plays an important role.

Supplementary Information

> Supplementary Material 1

Authors' contributions
CS and JK designed the methodology. CS and KP did the experiments. CS wrote the software. JK, JE and HPS supervised the project. CS, KP, JE, HPS and JK drafted the article. All authors read and approved the final manuscript.

Data Availability
The publicly available datasets analyzed in this study can be found here (accessed on 5 October 2022):
https://wiki.cancerimagingarchive.net/display/Public/Brain-Tumor-Progression#339481197db235d0cc7b490388fdb9be671371bb.
The source code will be uploaded to the following GitHub-repository: (https://github.com/cstrack/pn_vagan).

Declarations

References
1. Deng J, Dong W, Socher R, Li L-J, Li K, Fei-Fei L. ImageNet: A large-scale hierarchical image database. In: 2009 IEEE Conference on Computer Vision and Pattern Recognition. 2009. p. 248–55.
2. Fredrikson M, Jha S, Ristenpart T. Model Inversion Attacks that Exploit Confidence Information and Basic Countermeasures. In: Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security. Denver Colorado USA: ACM; 2015. p. 1322–33.
3. Vinyals O, Blundell C, Lillicrap T, Kavukcuoglu K, Wierstra D. Matching Networks for One Shot Learning. 2017.
4. Taigman Y, Yang M, Ranzato M, Wolf L, DeepFace. Closing the Gap to Human-Level Performance in Face Verification. In: 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus, OH, USA: IEEE; 2014. p. 1701–8.
5. Zarrin PS, Wenger C. Implementation of siamese-based few-shot learning algorithms for the distinction of COPD and Asthma subjects. In: Farkaš I, Masulli P, Wermter S, editors. Artificial neural networks and machine learning – ICANN 2020. Cham: Springer International Publishing; 2020. pp. 431–40.
6. Tschuchnig ME, Gadermayr M. Anomaly Detection in Medical Imaging - A Mini Review. In: Haber P, Lampoltshammer TJ, Leopold H, Mayr M, editors. Data Science – Analytics and Applications. Wiesbaden: Springer Fachmedien; 2022. pp. 33–8.
7. Goodfellow IJ, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S et al. Generative Adversarial Networks. ArXiv14062661 Cs Stat. 2014.
8. Kwon G, Han C, Kim D. Generation of 3D Brain MRI Using Auto-Encoding Generative Adversarial Networks. 2019.
9. Chuquicusma MJM, Hussein S, Burt J, Bagci U. How to fool radiologists with generative adversarial networks? A visual turing test for Lung Cancer diagnosis. 2018.
10. Rubin M, Stein O, Turko NA, Nygate Y, Roitshtain D, Karako L, et al. TOP-GAN: stain-free cancer cell classification using deep learning with a small training set. Med Image Anal. 2019;57:176–85.
11. Lei B, Xia Z, Jiang F, Jiang X, Ge Z, Xu Y, et al. Skin lesion segmentation via generative adversarial networks with dual discriminators. Med Image Anal. 2020;64:101716.
12. Holland EC. Glioblastoma Multiforme: the terminator. Proc Natl Acad Sci U S A. 2000;97:6242–4.
13. Harbeck N, Gnant M. Breast cancer. Lancet Lond Engl. 2017;389:1134–50.
14. Litwin MS, Tan H-J. The diagnosis and treatment of Prostate Cancer: a review. JAMA. 2017;317:2532–42.
15. Adamson C, Kanu OO, Mehta AI, Di C, Lin N, Mattox AK, et al. Glioblastoma Multiforme: a review of where we have been and where we are going. Expert Opin Investig Drugs. 2009;18:1061–83.
16. Egger J, Kapur T, Fedorov A, Pieper S, Miller JV, Veeraraghavan H, et al. GBM Volumetry using the 3D Slicer Medical Image Computing platform. Sci Rep. 2013;3:1364.
17. Egger J, Pepe A, Gsaxner C, Jin Y, Li J, Kern R. Deep learning—a first meta-survey of selected reviews across scientific disciplines, their commonalities, challenges and research impact. PeerJ Comput Sci. 2021;7:e773.

18. Fu Y, Lei Y, Wang T, Curran WJ, Liu T, Yang X. Deep learning in medical image registration: a review. Phys Med Biol. 2020;65:20TR01.
19. Arjovsky M, Chintala S, Bottou L, Wasserstein GAN. ArXiv170107875 Cs Stat. 2017.
20. Baumgartner CF, Koch LM, Tezcan KC, Ang JX, Konukoglu E. Visual Feature Attribution using Wasserstein GANs. ArXiv171108998 Cs. 2018.
21. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. ArXiv150504597 Cs. 2015.
22. Tran D, Bourdev L, Fergus R, Torresani L, Paluri M. Learning Spatiotemporal Features with 3D Convolutional Networks. In: 2015 IEEE International Conference on Computer Vision (ICCV). 2015. p. 4489–97.
23. Gulrajani I, Ahmed F, Arjovsky M, Dumoulin V, Courville A. Improved Training of Wasserstein GANs. ArXiv170400028 Cs Stat. 2017.
24. Kingma DP, Ba J. Adam: A Method for Stochastic Optimization. 2017.
25. Fedorov A, Beichel R, Kalpathy-Cramer J, Finet J, Fillion-Robin J-C, Pujol S, et al. 3D slicer as an image Computing platform for the Quantitative Imaging Network. Magn Reson Imaging. 2012;30:1323–41.
26. Isensee F, Schell M, Pflueger I, Brugnara G, Bonekamp D, Neuberger U, et al. Automated brain extraction of multisequence MRI using artificial neural networks. Hum Brain Mapp. 2019;40:4952–64.
27. Nuha FU. Afiahayati. Training dataset reduction on generative adversarial network. Procedia Comput Sci. 2018;144:133–9.
28. Ferreira A, Li J, Pomykala KL, Kleesiek J, Alves V, Egger J. GAN-based generation of realistic 3D data: A systematic review and taxonomy. 2022.
29. Isensee F, Jäger P, Wasserthal J, Zimmerer D, Petersen J, Kohl S et al. batchgenerators - a python framework for data augmentation. 2020.
30. Kleesiek J, Petersen J, Döring M, Maier-Hein K, Köthe U, Wick W, et al. Virtual raters for reproducible and objective assessments in Radiology. Sci Rep. 2016;6:25007.
31. Clark K, Vendt B, Smith K, Freymann J, Kirby J, Koppel P, et al. The Cancer Imaging Archive (TCIA): maintaining and operating a Public Information Repository. J Digit Imaging. 2013;26:1045–57.
32. Schmainda K, Prah M. Data from Brain-Tumor-Progression. 2019.
33. Isensee F, Jaeger PF, Full PM, Vollmuth P, Maier-Hein KH. nnU-Net for Brain Tumor Segmentation. arXiv; 2020.
34. Wen PY, Macdonald DR, Reardon DA, Cloughesy TF, Sorensen AG, Galanis E, et al. Updated response Assessment Criteria for High-Grade gliomas: Response Assessment in Neuro-Oncology Working Group. J Clin Oncol. 2010;28:1963–72.
35. Hestness J, Narang S, Ardalani N, Diamos G, Jun H, Kianinejad H et al. Deep Learning Scaling is Predictable, Empirically. 2017.
36. Erdt M, Steger S, Sakas G, Regmentation. A New View of Image Segmentation and Registration. 2012;:23.
37. Wyawahare MV, Patil DPM, Abhyankar HK. Image Registration techniques: an overview. Image Process Pattern Recognit. 2009;2:18.
38. Qin B, Gu Z, Sun X, Lv Y. Registration of images with outliers using Joint Saliency Map. IEEE Signal Process Lett. 2010;17:91–4.
39. Rees JH, Smirniotopoulos JG, Jones RV, Wong K. Glioblastoma Multiforme: radiologic-pathologic correlation. Radiographics. 1996;16:1413–38.
40. Menze BH, Jakab A, Bauer S, Kalpathy-Cramer J, Farahani K, Kirby J, et al. The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS). IEEE Trans Med Imaging. 2015;34:1993–2024.
41. Berntsen EM, Stensjøen AL, Langlo MS, Simonsen SQ, Christensen P, Moholdt VA, et al. Volumetric segmentation of glioblastoma progression compared to bidimensional products and clinical radiological reports. Acta Neurochir (Wien). 2020;162:379–87.
42. Dempsey MF, Condon BR, Hadley DM. Measurement of Tumor size in recurrent malignant glioma: 1D, 2D, or 3D? AJNR. Am J Neuroradiol. 2005;26:770–6.
43. Fyllingen EH, Stensjøen AL, Berntsen EM, Solheim O, Reinertsen I. Glioblastoma segmentation: comparison of three different Software packages. PLoS ONE. 2016;11:e0164891.
44. Sorensen AG, Batchelor TT, Wen PY, Zhang W-T, Jain RK. Response criteria for glioma. Nat Clin Pract Oncol. 2008;5:634–44.
45. Pepe A, Li J, Rolf-Pissarczyk M, Gsaxner C, Chen X, Holzapfel GA, et al. Detection, segmentation, simulation and visualization of aortic dissections: a review. Med Image Anal. 2020;65:101773.

Publisher's Note