The Effects of Simulated SAR Data Processing Methods and Network Parameter Tuning on Gridding Artifacts and Network Estimation Accuracy

Katherine Blanter, Alix Plumley, Emre Kopanoglu

October 2023

Synopsis

Gridding artifacts in neural network estimated images are common. We test the effects of simulated body models, neural network parameters, and post-processing methods on gridding artifacts, and their effect on overall neural network estimation accuracy in the context of local specific energy absorption rate (SAR) matrices, a patient safety concern for MRI scanning.

Impact

Researchers working with computer vision whose images experience a gridding artifact can inform their neural network parameter tuning efforts with the results of this exploratory study.

Introduction

A pipeline to use conditional generative adversarial networks (cGANs)[1] to predict the effects of patient head motion on local specific energy absorption rate (SAR) distributions was designed and implemented. During the design process, a gridding artifact (GA) (figure 2) was discovered in the network estimated (NE) images. To mitigate the artifact and discover its effects on NE accuracy, a variety of parameters was tested.

The cGAN consists of generator and discriminator convolutional neural networks, where the generator has a U-net architecture.

According to[2] and,[3] the gridding or "checkerboard" artifact could occur during the decoding portion of the U-net which uses transposed convolution. In particular, it could relate to the stride and filter size (FS) used. The combination of these parameters could cause a GA when the FS is not divisible by the stride, leading to uneven overlap. Even overlap causes artifacts too.[2,3] No overlap at all (FS = stride), ie. sub-pixel resolution,[4] was proposed as a solution, though it still causes artifacts.[2,3] Enhancing the training by increasing the epochs can capture more effective generator weights (GW) for testing. Though the present network implementation uses the GW from the epoch yielding the lowest loss, more training epochs may give the network a higher chance of securing a GW with lower loss.

Methods

The cGANs tested in this study are variations of the one used in,[5] which is a marginally altered version of pix2pix.[1]

The local SAR distribution data was derived from electromagnetic (EM) simulations in Sim4Life (ZMT, Zurich, Switzerland) using finite difference time domain (FDTD) calculations performed on Billie, Duke, Ella, Fats and Glenn from the Virtual Population (IT'IS, Zurich, Switzerland[6]) at the center of the coil and posterior 5mm using a generic 8 channel pTx coil at 7 T (295 MHz). This investigates two different head positions due to a different padding selection. The composition of each SAR distribution dataset was 256 x 256 (image resolution) x 140 (slices) x 64 (self- and mutual-interactions of 8 pTx channels).

1. Body model configurations investigated using the cGAN parameters from[5] with a leaky rectified linear unit (LReLu) instead of a ReLu during convolution in the generator:
   (a) Train: Fats, Ella, Glenn; validate: Billie; test: Duke
   (b) Train: Fats, Billie, Duke; validate: Glenn; test: Ella (configuration FBD-G-E)
   (c) Train: Fats, Ella, Glenn; validate: Duke; test: Billie
2. Neural network (NN) parameters evaluated on FBD-G-E:
   (a) Changing the leaky rectified linear unit (LReLu) to a ReLu
   (b) FS=2
   (c) FS=3
   (d) FS=3; stride=1
   (e) FS=1; stride=1
   (f) FS=1; strides=1; change LReLu to a ReLu during convolution in generator
   (g) Epochs=160
3. Implemented Hanning filter (HF) in postprocessing to mitigate the GA (configuration FBD-G-E), FSs: 7x7, 6x6, 5x5, 4x4, 3x3, 2x2 and 1x1.

Results and Discussion

The results were evaluated through the extent of the GA and the pertaining L1 error (L1e) between the NE images and the simulated ground truth (GT).

Figure 1 shows that the GA presented when testing on Duke and Ella. The GT images do not contain gridding, therefore the artifact is not learned. Table 1 indicates that mean and maximum L1e is lower when testing on Ella. This is likely because Ella's anatomy fits within the bounds of the anatomies of the training body models (BMs) (Fats, Billie, and Duke) in the given BM configuration, suggesting that when arising from the preprocessing phase, the GA does not affect overall NN accuracy. To increase versatility, more BMs would have to be used for training to account for anatomical variations.

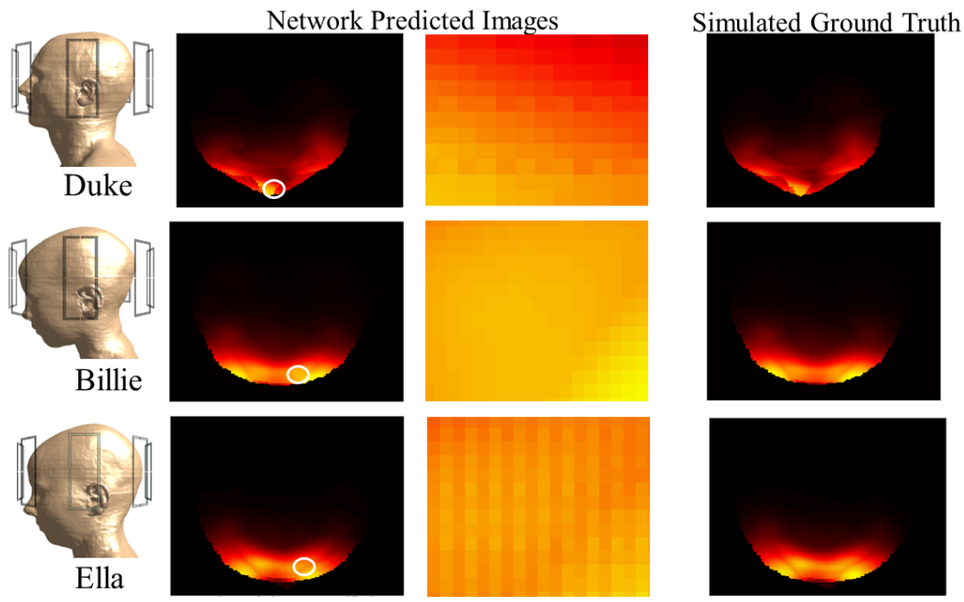

Figure 1: Maximum intensity projections along z. The GA appeared when Ella and Duke were used for testing, but disappeared when testing on Billie. White ellipse shows region of magnification.

Figure 2 displays the effects of NN parameter changes on the GA, which disappeared when stride = 1. The GA was affected by the FS. Increasing the epochs did not improve image quality, indicating that the artifact is not due to underfitting, though the mean L1e improved (table 1). The low L1e for FS=1x1 and stride=1x1 suggests that the GA, when arising from parameter tuning, has an effect on NE accuracy.

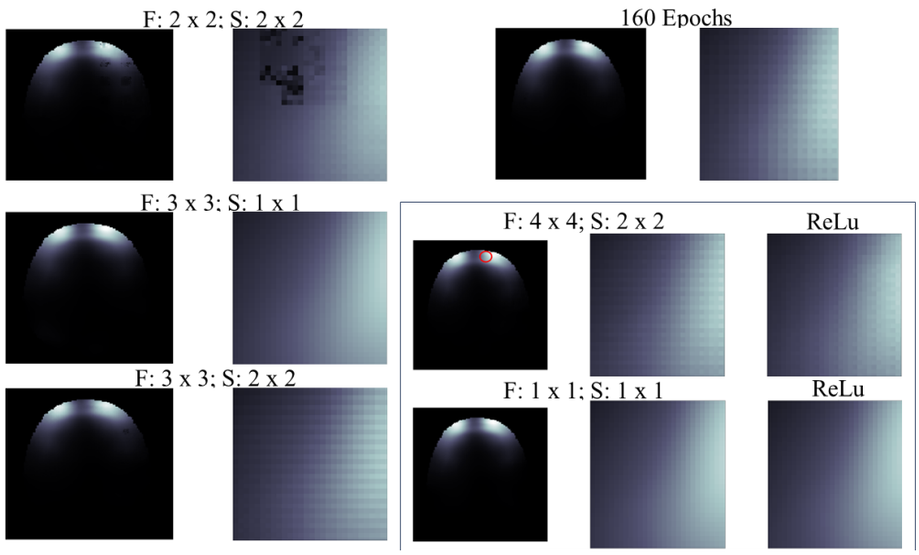

Figure 2: F is FS and S is strides. 'ReLu' indicates the version of the network where the LReLU was replaced by a ReLu. The GA persisted with all parameter changes except when strides=1 x 1. All magnified images are from the area specified by the red ellipse in F:4x4;S:2x2.

Applying the HF during postprocessing improves the accuracy of the network predictions which contained the GA, but the values are higher than when the network parameters include FS=1x1 and stride=1x1. The HF introduced areas of high intensity pixels, which contributed to higher L1e. FS=3x3 was optimal (table 1).

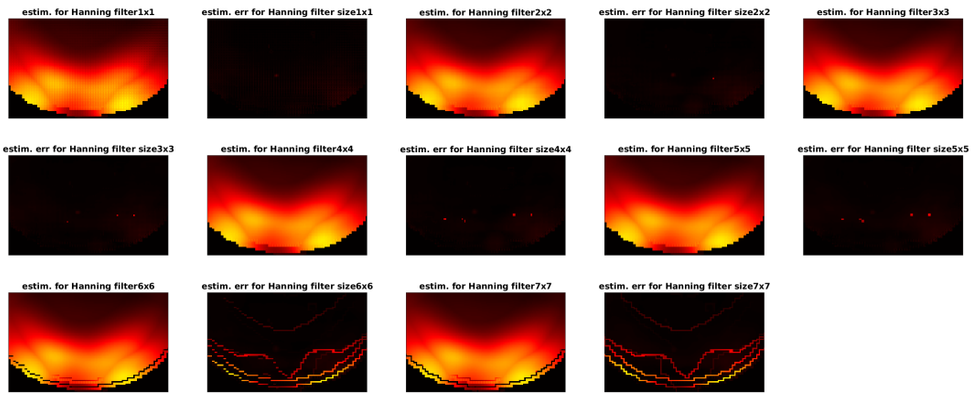

Figure 3: While the HF eliminated the GA in post-processing, it created isolated pixels of magnified error. Smaller FSs yielded smaller areas of isolated high error, but at the expense of a reemerging GA. FSs greater than 5x5 were too large for the images and caused an edge artifact.

	μ Estim.	μ MI	max Estim.	max MI	μ ratio	max ratio
Duke	0.1346	0.1526	0.2205	0.4376	0.8820	0.5038
Billie	0.1092	0.1636	0.1850	0.5771	0.6674	0.3205
F4;S2 (Ella)	0.1047	0.1699	0.1798	0.5718	0.6162	0.3144
F4;S2;ReLu	0.1313	0.1699	0.2375	0.5718	0.7728	0.4153
F3;S2	0.1089	0.1699	0.1709	0.5718	0.6409	0.2988
F2;S2	0.1107	0.1699	0.1718	0.5718	0.6515	0.3004
F3;S1	0.1039	0.1699	0.2661	0.5718	0.6115	0.4653
F1;S1	0.0862	0.1699	0.1502	0.5718	0.5126	0.2667
F1;S1;ReLu	0.1309	0.1699	0.3729	0.5718	0.7704	0.6521
160 ep	0.0978	0.1699	0.2007	0.5718	0.5756	0.3509
H 1x1	0.1047	0.1699	0.1798	0.5718	0.6162	0.3144
H 2x2	0.0950	0.1699	0.1800	0.5718	0.5591	0.3147
H 3x3	0.0929	0.1699	0.1746	0.5718	0.5467	0.3053
H 4x4	0.0952	0.1699	0.1801	0.5718	0.5603	0.3149
H 5x5	0.0934	0.1699	0.1752	0.5718	0.5497	0.3064

Table 1: Overview of motion induced (MI) and NE (Estim.) L1e values from networks run with various parameter changes, where Billie and Duke are the alternative BM configurations, F is NN FS, S is NN strides, ep is epochs, and H is Hanning FS in postprocessing. The ratios account for cases in which the MI L1e differed (eg. different BM configuration or normalization method). A smaller ratio value indicates better NN performance.

The FS and stride had the most effect on the GA and NE accuracy, with a 1x1 FS and stride with LReLu yielding the best outcome. This method improves prediction accuracy and reduces computational cost through so-called pointwise convolution.[7]

Conclusion

We have tested a variety of parameters which contribute to and eliminate cGAN-caused GAs in local SAR matrices. The results of this work can further inform researchers working with computer vision whose results are affected by GAs.

References

[1] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, Alexei A Efros, and Berkeley Ai Research. Image-to-image translation with conditional adversarial networks, 2017.

[2] Augustus Odena, Vincent Dumoulin, and Chris Olah. Deconvolution and checkerboard artifacts. *Distill*, 1(10):e3, 2016.

[3] Yusuke Sugawara, Sayaka Shiota, and Hitoshi Kiya. Checkerboard artifacts free convolutional neural networks. *APSIPA Transactions on Signal and Information Processing*, 8:e9, 2019.

[4] Wenzhe Shi, Jose Caballero, Ferenc Huszár, Johannes Totz, Andrew P Aitken, Rob Bishop, Daniel Rueckert, and Zehan Wang. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1874–1883, 2016.

[5] Alix Plumley, Luke Watkins, Matthias Treder, Patrick Liebig, Kevin Murphy, and Emre Kopanoglu. Rigid motion-resolved prediction using deep learning for real-time parallel-transmission pulse design. *Magnetic Resonance in Medicine*, 87:2254–2270, 5 2022.

[6] Honegger Katharina Zefferer Marcel Neufeld Esra Oberle Michael Szczerba Dominik Kuster Niels Kainz Wolfgang Guag Joshua W Hahn Eckhart G Rascher Wolfgang Janka Rolf Bautz Werner Chen Ji Shen Jianxiang Kiefer Berthold Schmitt Peter Hollenbach Hans-Peter Christ, Andreas and Anthony Kam. The virtual family-development of surface-based anatomical models of two adults and two children for dosimetric simulations, Jan 2010.

[7] Binh-Son Hua, Minh-Khoi Tran, and Sai-Kit Yeung. Pointwise convolutional neural networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 984–993, 2018.