OPEN

Machine Friendly Machine Learning: Interpretation of Computed Tomography Without Image Reconstruction

Hyunkwang Lee [1,2], Chao Huang[1], Sehyo Yune [1], Shahein H. Tajmir[1], Myeongchan Kim [1] & Synho Do [1*]

Recent advancements in deep learning for automated image processing and classification have accelerated many new applications for medical image analysis. However, most deep learning algorithms have been developed using reconstructed, human-interpretable medical images. While image reconstruction from raw sensor data is required for the creation of medical images, the reconstruction process only uses a partial representation of all the data acquired. Here, we report the development of a system to directly process raw computed tomography (CT) data in sinogram-space, bypassing the intermediary step of image reconstruction. Two classification tasks were evaluated for their feasibility of sinogram-space machine learning: body region identification and intracranial hemorrhage (ICH) detection. Our proposed SinoNet, a convolutional neural network optimized for interpreting sinograms, performed favorably compared to conventional reconstructed image-space-based systems for both tasks, regardless of scanning geometries in terms of projections or detectors. Further, SinoNet performed significantly better when using sparsely sampled sinograms than conventional networks operating in image-space. As a result, sinogram-space algorithms could be used in field settings for triage (presence of ICH), especially where low radiation dose is desired. These findings also demonstrate another strength of deep learning where it can analyze and interpret sinograms that are virtually impossible for human experts.

Continued rapid advancements in algorithms and computer hardware have accelerated progress in automated computer vision and natural language processing. By combining these two factors with the availability of well-annotated large datasets, significant advances have emerged from automated medical image interpretation for the detection of disease and critical findings[1–3]. The application of deep learning has the potential to increase diagnostic accuracy and reduce delays in diagnosis and treatment for better patient outcomes[4]. Deep learning techniques are not limited to image analysis, but they also can improve image reconstruction for magnetic resonance imaging (MRI)[5,6], computed tomography (CT)[7,8], and photoacoustic tomography (PAT)[9]. In particular, deep learning approaches have been used to improve image quality for low-dose CT reconstruction by interpolating sparse CT projection data[10,11], denoising sparse-view reconstructed image[7,8], or both[12]. These prior works demonstrated that deep learning now is a feasible alternative to well-established analytic and iterative methods of image reconstruction[13–17].

However, most prior work using deep learning algorithms has focused on image analysis of reconstructed images or as an alternative approach to image reconstruction. Despite this human centric approach, there is no reason that deep learning algorithms must function in image-space. Since all the information in the reconstructed images is present in the raw measurement data, deep learning models could potentially derive features directly from raw data in sinogram-space without intermediary image reconstruction, with possibly even better performance than models trained in image-space.

In this study, we determined the feasibility of analyzing computed tomography (CT) projection data — sinograms — through a deep learning approach for human anatomy identification and pathology detection. We

[1]Department of Radiology, Massachusetts General Hospital and Harvard Medical School, Boston, MA, 02114, USA. [2]John A. Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, 02138, USA. *email: sdo@mgh.harvard.edu

Fully sampled	Moderately sampled	Sparsely sampled
I1: *recon360x729* (original CT)	**I3**: *recon120x240*	**I5**: *recon40x80*
S1: *sino360x729*	**S3**: *sino120x240*	**S5**: *sino40x80*
I2: *wrecon360x729* (windowed original CT)	**I4**: *wrecon120x240*	**I6**: *wrecon40x80*
S2: *wsino360x729*	**S4**: *wsino120x240*	**S6**: *wsino40x80*

Table 1. Summary of the 12 different models evaluated in this study. Inception-v3 was used for analyzing reconstructed images (I1–I6), and SinoNet was used for interpreting singograms (S1–S6). '*sino360x729*', '*sino120x240*', and '*sino40x80*' represent sinograms with 360 projection views by 729 detector pixels, 120 projection views by 240 detector pixels, and 40 projection views by 80 detector pixels, respectively. '*wsino360x729*', '*wsino120x240*', and '*wsino40x80*' represent sinograms created from windowed CT images. '*recon360x729*', '*recon120x240*', and '*recon40x80*' are images reconstructed from the corresponding sinograms, and '*wrecon360x729*', '*wrecon120x240*', and '*wrecon40x80*' were reconstructed images from the windowed sinograms.

proposed a customized convolutional neural network (CNN) called SinoNet, optimized it for interpreting sinograms, and demonstrated its potential by comparing its performance to pre-existing system based on other CNN architectures using reconstructed CT images. This approach accelerates edge computing by making it possible to identify critical findings rapidly from the raw data without time-consuming image reconstruction processes. In addition, this could enable us to develop simplified scanner hardware for the direct detection of critical findings through SinoNet alone.

Results

Experimental design. We retrieved 200 contiguous whole-body CT datasets from combined positron emission tomography-computed tomography (PET/CT) examinations for body part recognition and 720 non-contrast head CT scans for intracranial hemorrhage (ICH) detection with IRB approval from the picture archiving and communication systems (PACS) at Massachusetts General Hospital. Axial slices in the 200 whole body scans were annotated as sixteen different body regions by a physician, and slices of the 720 head scans were annotated with the presence of ICH by a panel of five neuroradiologists by consensus (Methods). We evaluated twelve different classification models developed by training Inception-v3[18] on reconstructed CT images and SinoNet with sinograms (Table 1, Methods). The reconstructed CT images containing Hounsfield units (HU) were converted to scaled linear attenuation coefficients (LAC). Two-dimensional (2D) parallel-beam Radon transform was applied to the LAC slices (512 × 512 pixels) to generate a fully-sampled sinogram with 360 projections and 729 detector pixels ('*sino360x729*'), which was then uniformly subsampled in the horizontal direction (projection views) and averaged in vertical direction (detector pixels) by factors of 3 and 9 to obtain moderately sampled sinograms with 120 views by 240 pixels ('*sino120x240*') and sparsely sampled sinograms with 40 views by 80 pixels ('*sino40x80*').

Original CT images were used as fully sampled reconstructed images (*recon360x729*), and images reconstructed from the sparse sinograms ('*recon120x240*' and '*recon40x80*') were generated using a deep learning approach (FBPConvNet[8]) followed by a conversion from LAC to HU. Reconstructed CT images and sinograms with predefined window-level settings were created to evaluate the effect of windowing: '*wrecon360x729*', '*wrecon120x240*', '*wrecon40x80*'; and '*wsino360x729*', '*wsino120x240*', '*wsino40x80*' (Methods). Based on the scanning geometries and window-level settings described above, 12 CNN models were evaluated: 6 were developed by training Inception-v3[18] with reconstructed CT images and the other 6 were obtained by training SinoNet with sinograms (Table 1, Methods). Data for body part recognition was randomly split into training, validation, and test sets with balanced genders: 140 scans (female: n = 70; male: n = 70) in training, 30 (female: n = 15; male: n = 15) in validation, and 30 (female: n = 15; male: n = 15) in testing. A dataset split was also performed for ICH detection with 478 scans in training, 121 in validation, and 121 in testing. Details of data preparation, CNN architecture, sinogram generation, and image reconstruction are described in Methods.

Results of body part recognition. Figure 1 shows test performance of the twelve different models for body part recognition. Models trained on fully sampled images had accuracies of 97.4% in image-space (I1), 96.6% in sinogram-space (S1), 97.9% in windowed-image-space (I2), and 97.4% in windowed-sinogram-space (S2). Moderately sampled images had model accuracies of 97.4% in image-space (I3), 96.3% in sinogram-space (S3), 97.9% in windowed-image-space (I4), and 97.4% in windowed-sinogram-space (S4). Sparsely sampled images had model accuracies of 97.1% in image-space (I5), 96.2% in sinogram-space (S5), 97.2% in windowed-image-space (I6), and 97.1% in windowed-sinogram-space (S6). These results imply that models trained and operating in image-space performed slightly better than sinogram-space (SinoNet) models for body part recognition, regardless of scanning geometry. Additionally, windowed input images consistently outperformed the ones with full-range images/sinograms.

Results of intracranial hemorrhage detection. Figure 2 depicts receiver operating characteristic (ROC) curves, and the corresponding areas under the ROC curves (AUC) for the twelve different models of ICH detection. Models trained on fully sampled images had AUCs of 0.898 in image-space (I1), 0.918 in sinogram-space (S1), 0.972 in windowed-image-space (I2), and 0.951 in windowed-sinogram-space (S2). Moderately sampled images had model accuracies of 0.893 in image-space (I3), 0.915 in sinogram-space (S3),
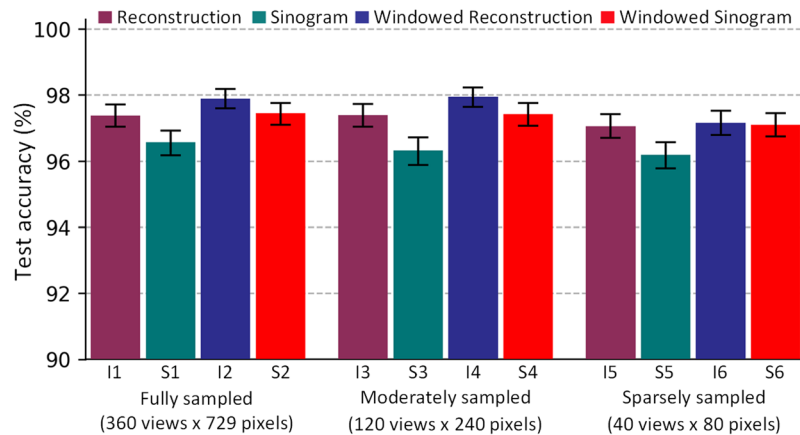
Figure 1. Performance of 12 different models trained on reconstruction images and sinograms with varying numbers of projections and detectors for body part recognition. 95% confidence intervals (CIs) are indicated in black error bars. The purple and blue bars (I1–I6) compare the test accuracy of Inception-v3 trained with full dynamic range reconstructed images with abdominal window setting reconstructed images (window-level = 40 HU, window-width = 400 HU). The green and red bars (S1–S6) compare the performance of SinoNet models trained with sinograms generated from full-range and windowed reconstructed images, respectively.

0.953 in windowed-image-space (I4), and 0.947 in windowed-sinogram-space (S4). Sparsely sampled images had model accuracies of 0.885 in image-space (I5), 0.899 in sinogram-space (S5), 0.909 in windowed-image-space (I6), and 0.942 in windowed-sinogram-space (S6).
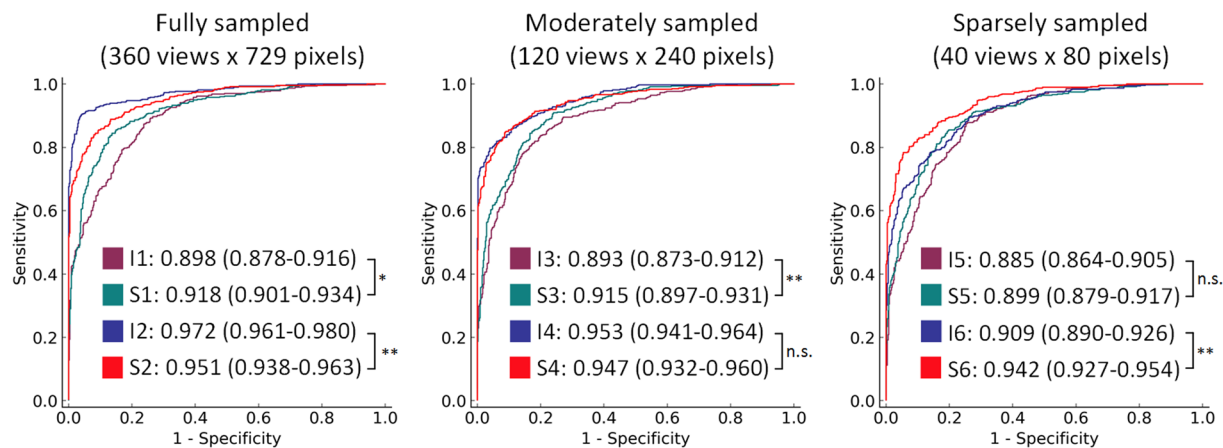
Comparison of SinoNet and Inception-v3 for analyzing sinograms. Table 2 details performance comparisons of Inception-v3 and SinoNet for interpreting fully-sampled sinograms (360 projection views and 729 detector pixels) for both body part recognition and ICH detection. SinoNet models significantly outperformed Inception-v3 models in both tasks.

Discussion

We have demonstrated that models trained on sinograms can achieve similar performance when compared to models using conventional reconstructed images for body part recognition and ICH detection in all three scanning geometries, despite the fact that the raw measurement data are not interpretable to humans. SinoNet, when trained with sinograms, has comparable performance with that of Inception-v3 when trained with reconstructed CT images for body part recognition, regardless of the number of projection views or detectors. For ICH detection, SinoNet trained with full-range sinograms outperformed Inception-v3 trained with full dynamic range reconstructed images for all three sampling densities, with SinoNet significantly outperforming Inception-v3 when using windowed, sparsely sampled images. By applying window settings similar to what a radiologist would use, network performance increased significantly due to the improved contrast of target to background (Fig. 3) in both reconstructed images and in sinogram-space. As depicted in Fig. 3(b), not only are the key features relevant to hemorrhage enhanced in the windowed CT image, but also in the windowed sinogram.

SinoNet, a proposed convolutional neural network, was developed for analyzing sinograms through customized Inception modules with multi-scale convolutional and pooling layers[18]. In SinoNet, the square convolutional filters in the original Inception module were replaced by various sized rectangular convolutional filters which include width-wise (projection dominant) and height-wise (detector dominant) filters. The customized architecture of SinoNet allowed for significantly improved performance in both body part recognition and ICH detection when compared with Inception-v3 models trained with sinograms, regardless of sampling density. These results imply that non-square filters may be effective in enabling models to learn the interplay between projection views and detector pixels from sinusoidal curves and to extract salient features from the sinogram domain for classification, a task thought to be impossible for human experts to grasp. This approach is similar to the one proposed for learning temporal and frequency features using rectangular convolution filters in spectrograms[19].

SinoNet, by operating in sinogram-space, can accelerate image interpretation for pathology detection as complex computations for image reconstruction are not required. SinoNet also excels when the projection data was moderately or sparsely sampled, maintaining its AUC at 0.942 on the hemorrhage detection task, while Inceptionv3 dropped from 0.972 to 0.909. Sparsely sampled datasets suggest that radiation dose could be markedly decreased with only a slight degradation in performance for sinogram-space algorithms. The number of projections linearly correlates with radiation dose, theoretically achieving 33% and 89% dose reductions for moderately and sparsely sampled data respectively. Similarly, by reducing the size and number of detectors required for diagnostic CT data, cheaper and simpler CT scanners can be created. At our institution, the average head CT has a CTDIvol of 50 mGy. Sparsely sampled data could have CTDIvol between 6 and 16 mGy. One possible use of this technique would be to use the sinogram model as a first-line screening tool in the field setting without image reconstruction, subsequently prioritizing a patient for potential stroke therapy given no evidence of intracranial hemorrhage. Subsequent full-dose CT could be used to confirm the interpretation from the sinogram method.

Figure 2. ROC curves for performance of 12 different models trained with reconstruction images and sinograms with various sparsity configurations in numbers of projections and detectors. The purple and blue curves (I1–I6) correspond to performance of Inception-v3 trained with reconstruction images with a full dynamic range of HU values and brain window setting (window-level = 50 HU, window-width = 100 HU), respectively. The green and red curves (S1–S6) show performance of SinoNet models trained with sinograms generated from full-range and windowed reconstruction images, respectively. The areas under the curve (AUCs) for the 12 models are present in legends with their 95% CIs. Statistical significance of the difference between AUCs of paired models (Ix - Sx) was evaluated. n.s., $p > 0.05$; *$p < 0.05$; **$p < 0.01$.

Input	Body part recognition (Test accuracy)		ICH detection (AUC)	
	Inception-v3	SinoNet	Inception-v3	SinoNet
sino360x729	93.9% (93.4%–94.4%)	96.6% (96.2%–96.9%)	0.873 (0.849–0.895)	0.918* (0.899–0.935)
sino120x240	93.5% (93.0%–94.0%)	96.3% (95.9%–96.7%)	0.874 (0.851–0.896)	0.915* (0.897–0.932)
sino40x80	93.4% (92.9%–93.9%)	96.2% (95.8%–96.6%)	0.852 (0.828–0.876)	0.899* (0.879–0.917)

Table 2. Comparison of Inception-v3 and SinoNet network performance when both networks are trained on full-range sinograms are varying sampling densities for body part recognition and intracranial hemorrhage (ICH) detection. Body part recognition is reported in accuracy. ICH detection as AUC. 95% CIs in parentheses. *$p < 0.0001$.

Another possible use for this technique would be to create "smart-scanners" which allow the CT scanner to adjust the protocol and field of view based on the intended region of the body.

Although these results demonstrate the power of the sinogram based approach, several important areas of future investigation remain. Due to their unavailability, the sinograms used in this study were simulated by applying the 2D parallel-beam Radon transform to the reconstructed CT images rather than actual measurement data acquired from CT scanners. Improved simulation data could be acquired by accounting for other advanced projection geometries — cone-beam or fan-beam — and considering Poisson noise when generating projection data. Although SinoNet trained with windowed sinograms achieved comparable or better performance compared with windowed reconstructed images, windowed sinograms were generated from reconstructed images that were post-processed with predefined window settings; generation of windowed sinograms directly from CT measurement data is not straightforward, but it could be implemented by using energy-resolving, photon-counting detectors from multi-energy CT imaging to acquire measurements in multiple energy bins[20]. Our work will need to be further validated by using raw data from clinical scanners as well as raw data from actual low-dose image acquisitions to see if performance remains robust despite increased image noise.

In conclusion, sinogram-space deep learning with our proposed CNN called SinoNet is feasible for human anatomy identification and pathology detection (presence of ICH) on sinograms acquired using different scanning geometries in terms of projections and detectors which are not virtually interpretable to human experts like sinograms. In particular, this study showed SinoNet performed better for pathology detection directly from sparse sinograms than reconstructed images, indicating the potential of deep learning to identify critical findings from raw data without expensive image reconstruction processes in field settings for triage, especially where low dose radiation is required.

Methods

All the images were fully de-identified in compliance with the Health Insurance Portability and Accountability Act (HIPAA). This retrospective study was conducted with the approval of the Institutional Review Board (IRB) of Massachusetts General Hospital and under a waiver of informed consent. All experiments were performed in accordance with relevant guidelines and regulations.
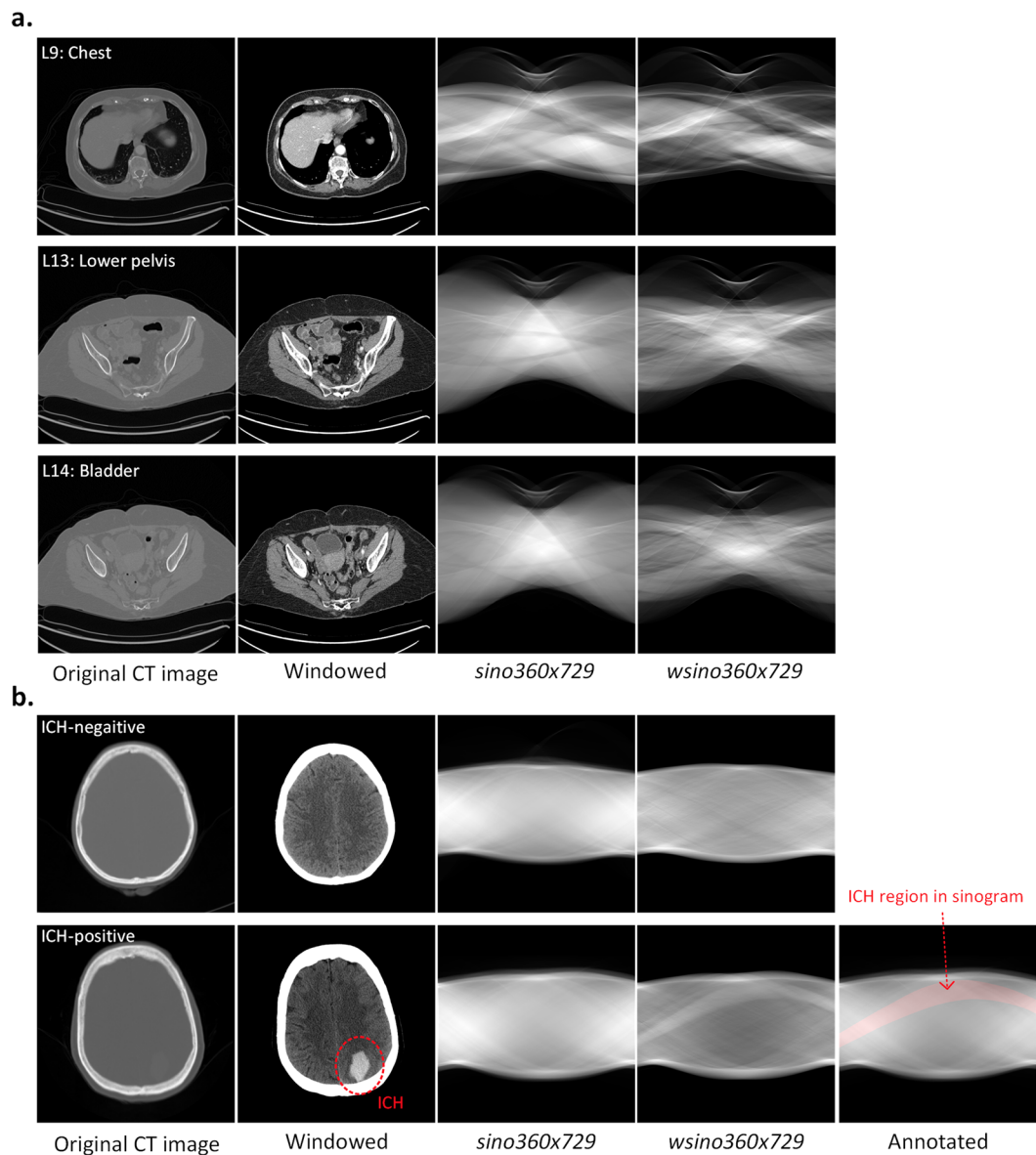
Figure 3. Examples of reconstructed images and sinograms with different labels for (**a**), body part recognition and (**b**), ICH detection. From left to right: original CT images, windowed CT images, sinograms with 360 projections by 729 detector pixels, and windowed sinograms 360×729. In the last row, an example CT with hemorrhage is annotated with a dotted circle in image-space with the region of interest converted into the sinogram domain using Radon transform. This area is highlighted in red on the sinogram in the fifth column.

Data collection and annotation. *Body part recognition.* A total of 200 contrast-enhanced PET/CT examinations of head, neck, chest, abdomen, and pelvis for 100 female and 100 male patients were retrieved from our institutional PACS between May 2012 and July 2012. A total of 56,334 axial slices in the CT scans were annotated as one of sixteen body regions by a physician (see Supplementary Fig. S1). 30 cases (Female: n = 15; Male: n = 15) were randomly selected for use as validation data for hyperparameter tuning and model selection, another 30 cases (Female: n = 15; Male: n = 15) as test data for performance evaluation, and the rest of 140 cases (Female: n = 70; Male: n = 70) as training data for model development (Table 3).

Intracranial hemorrhage (ICH) detection. A total of 720 5-mm non-contrast head CT scans were identified and retrieved from our PACS between June 2013 and July 2017. Every 5-mm thick axial slice (3,151 slices without ICH and 2,895 slices with ICH) was annotated by five board-certified neuroradiologists (blinded for review, 9 to 34 years experience) according to presence of ICH by consensus. The examinations included 201 cases without ICH and 519 cases with ICH, which were randomly split into train (141 cases), validation (30 cases), and test (30 cases) datasets at the case-level to ensure slices from the same case were not split across different datasets (Table 4).

	Train	Validation	Test
No. Cases	140	30	30
No. Images	39,472	8,383	8,479
L1: Head	1,980	483	435
L2: Eye lens	878	189	188
L3: Nose	1,449	309	323
L4: Salivary gland	1,803	361	349
L5: Thyroid	1,508	312	333
L6: Upper lung	1,632	345	392
L7: Thymus	3,213	727	672
L8: Heart	3,360	707	762
L9: Chest	4,647	914	935
L10: Upper abdomen	4,943	1,008	1,103
L11: Lower abdomen	1,736	342	368
L12: Upper pelvis	2,524	617	545
L13: Lower pelvis	2,230	563	422
L14: Bladder	3,144	609	766
L15: Upper leg	2,607	563	532
L16: Lower leg	1,818	334	354

Table 3. Distribution of training, validation, and test datasets for body part recognition.

	Train		Validation		Test	
	No. Cases	No. Images	No. Cases	No. Images	No. Cases	No. Images
No ICH	141	2,202	30	474	30	475
ICH	337	1,915	91	490	91	475
Total	478	4,117	121	964	121	950

Table 4. Distribution of training, validation, and test datasets for ICH detection.

Sinogram generation. Simulated sinograms were utilized in this study instead of raw data obtained by commercial CT scanners as this was a retrospective analysis and access to raw projection data from patient CT scans could not be retrieved. To generate simulated sinograms, the pixel values of 512×512 CT images stored in DICOM file were first converted into scaled linear attenuation coefficients (LACs). Any calculated negative LAC was leveled to zero under the assumption that it is physically impossible to have negative LACs, so this result must represent random noise. Subsequently, three different sinograms were generated based on the scaled LAC images. First, we computed sinograms with 360 projection views over 180 degrees and 729 detectors ('*sino360x729*'), using the 2D parallel-beam Radon transform. '*sino360x729*' were then used to produce sparser sinograms by uniformly subsampling projection views (in the horizontal direction) and averaging projection data from adjacent detectors (in the vertical direction) by factors of 3 and 9 to obtain sinograms with 120 projection views and 240 detectors ('*sino120x240*') and sinograms with 40 projection views and 80 detectors ('*sino40x80*'), respectively (Fig. 4). Sparser sinograms ('*sino40x80*', '*sino120x240*') were resized to 360×729 pixels using a bilinear interpolation to have a uniform resolution with the corresponding full-view sinograms ('*sino360x729*').

Image reconstruction. Reconstructed images were generated from the synthetic sinograms for models I1–I6. Original CT images were used as the reconstructed images for '*recon360x729*' as fully sampled sinogram data could be completely reconstructed into images using filtered back projection (FBP). However, other complex algorithms are needed to reconstruct high-quality images from sparser datasets, such as model-based iterative reconstruction. Rather than employing complex iterative algorithms, we implemented a deep learning approach to reconstruct sparsely sampled sinograms as this technique has been demonstrated to compare favorably to state-of-the-art iterative algorithms for sparse-view image reconstruction[7,8]. We implemented *FBPConvNet*, a modified U-net[21] with multiresolution decomposition and residual learning as proposed by a prior work[8]. *FBPConvNet* takes FBP reconstructed images from sparser sinograms ('*sino120x240*' or '*sino40x80*') as inputs and is trained for regression between the input and the original CT image (converted into LACs) with mean square error (MSE) as the loss function (see Supplementary Fig. S2). Since the output images of *FBPConvNet* were LACs, they were converted into HU as the final reconstructed images. Sparser sinograms were resized to 360×729 pixels using bilinear interpolation in order to make the corresponding FBP images have the uniform resolution of 512×512 pixels, resulting in final reconstructed images of 512×512 pixels. The best *FBPConvNet* models selected based on root mean square error (RMSE) values on the validation data were employed on '*sino120x240*' and '*sino40x80*' to generate '*recon120x240*' and '*recon40x80*' respectively (Fig. 5). The RMSE of reconstructed images obtained from the *FBPConvNet* in validation dataset are much smaller than that of conventional FBP images (see Supplementary Table S1).
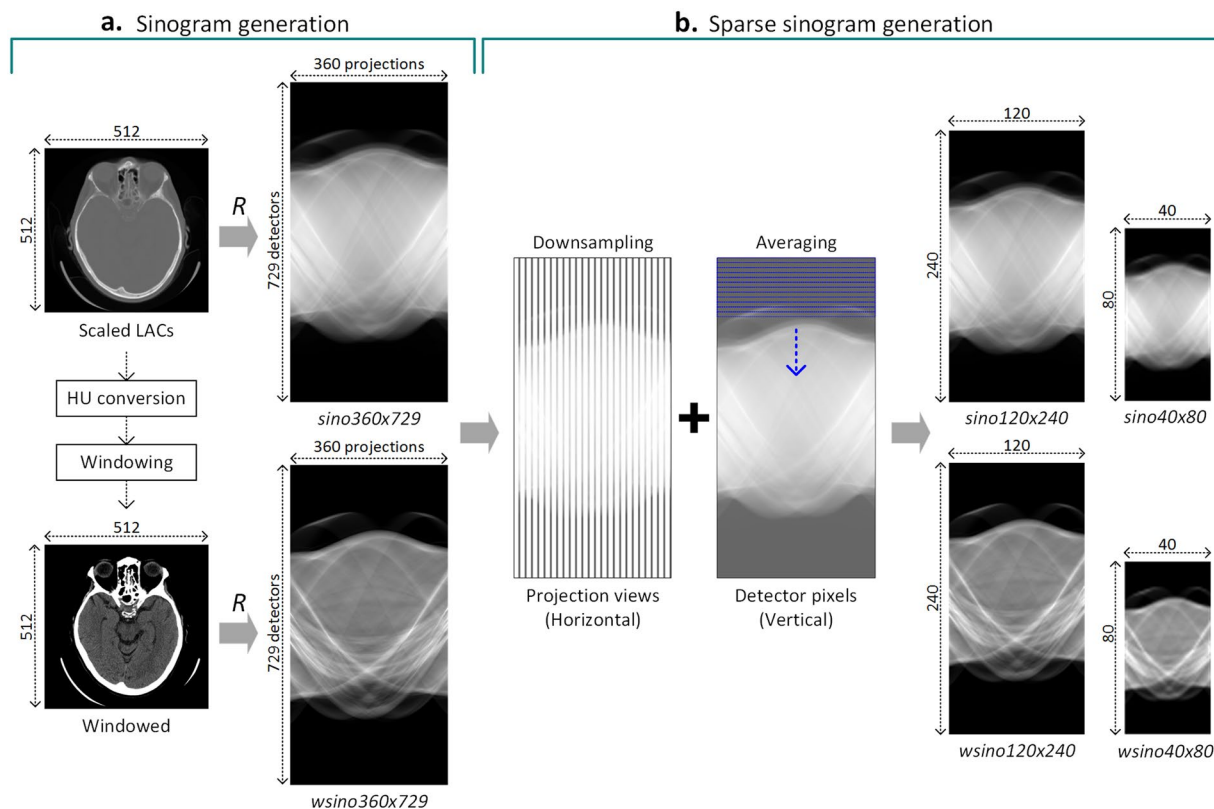
Figure 4. (**a**) Schematic of sinogram generation with 360 projection views and 729 detectors ('*sino360x729*') from original CT images (converted into linear attenuation coefficients). (**b**) Sparse sinograms were created from '*sino360x729*' by downsampling in the horizontal dimension and signal averaging in the vertical dimension to simulate the effect of acquiring an image with 120 projection views and 240 detectors ('*sino120x240*') or an image with 40 projection views and 80 detectors ('*sino40x80*'). *R*, Radon transform.

Windowed images and sinograms.   We utilized full-range 12-bit grayscale images and windowed 8-bit grayscale images with different window-levels (WL) and window-widths (WW) suitable for each task: abdominal window (WL = 40 HU, WW = 400 HU) for body part recognition and brain window (WL = 50 HU, WW = 100 HU) for ICH detection. The windowed sinograms were generated from corresponding windowed CT images. Examples of windowed images and sinograms are shown in Supplementary Fig. S3.

Convolutional neural network for sinograms: SinoNet.   A customized convolutional neural network, SinoNet, was designed for analyzing sinograms using customized Inception modules with multiple convolutional and pooling layers and dense connection for efficient use of model parameters[18,22]. As shown in Fig. 5, the Inception module was modified with various sized rectangular convolutional filters in SinoNet. The non-square filters include height-wise (detector dominant) and width-wise (projection dominant) filters to enable efficient extraction of features from sinusoidal curves. Two Inception modules were densely connected to form a Dense-Inception block, which was followed by a Transition block to reduce the number and dimension of feature maps for computational efficiency, as suggested in the original report[22]. In this study, SinoNet was used only for interpreting sinograms.

Baseline convolutional neural network: Inception-v3.   Inception-v3[18], a validated CNN for object recognition in the ImageNet Large Scale Visual Recognition Challenge (ILSVRC)[23], was selected as the network architecture to develop classification models trained on reconstructed images. We modified Inception-v3 by replacing the last fully-connected layers with a sequence of a global average pooling (GAP) layer, a fully-connected layer, and a softmax layer with outputs of the same number of categories: 16 multi-class outputs for body part recognition and a binary output for ICH detection. Inception-v3 was also used to classify sinograms when evaluating SinoNet performance at body part recognition and ICH detection when using sinograms as the input data.

Weight initialization.   All models developed using Inception-v3 and SinoNet for body part recognition task were initialized with He normal initialization[24]. For the ICH detection task, models were initialized with corresponding pre-trained weights on the body part recognition with full-view scanning geometry. For example, the Inception-v3 model trained with '*recon360x729*' for body part recognition was used as the initial weights for Inception-v3 models trained with reconstructed images for ICH detection for all scanning geometries and
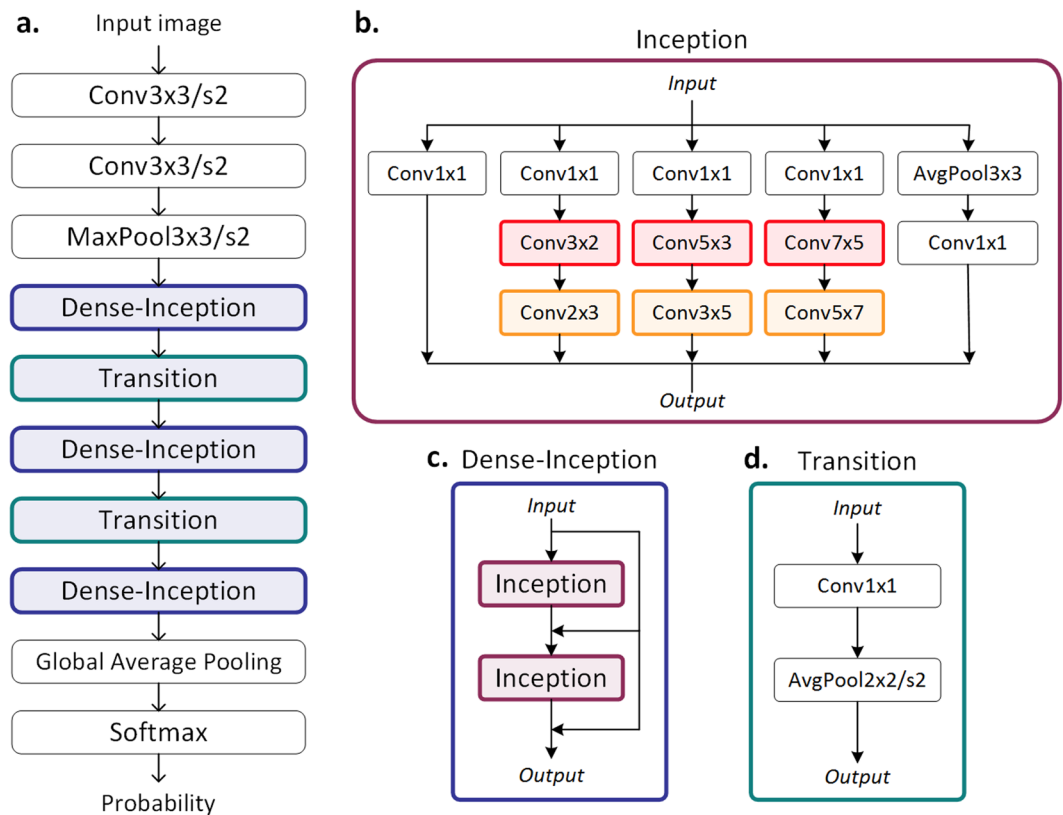
Figure 5. (**a**) Overall network architecture of SinoNet. (**b**) Detailed network diagram within the Inception modules that include rectangular convolutional filters and pooling layers. The modified Inception module contains multiple rectangular convolution filters of varying sizes: height-wise rectangular filters (projection dominant) in red; width-wise rectangular filters (detector dominant) in orange; "Conv3x3/s2" indicates a convolutional layer with 3×3 filters and 2 stride, and "Conv3x2" means a convolution layer with 3×2 filters and 1 stride. (**c**) Dense-Inception layers contain two densely connected Inception modules. (**d**) Transition modules situated between Dense-Inception modules reduce the size of feature maps. Conv = convolution layer, MaxPool = max pooling layer, AvgPool = average pooling layer.

window levels. Similarly, SinoNet ICH detection models were initialized using the weights from the body part recognition SinoNet model trained with '*sino360x729*'.

Performance evaluation and statistical analysis. Test accuracy was used as the performance metric for comparing body part recognition models, and ROC curves with AUC were used for evaluating performance of models for detection of ICH. All performance metrics were calculated using scikit-learn 0.19.2 available in python 2.7.12. A non-parametric approach (DeLong[25]) was used to assess the statistical significance of the difference between AUCs of ICH detection models trained with reconstruction images and sinograms using Stata version 15.1 (StataCorp, College Station, Texas, USA). We employed a non-parametric, bootstrap approach with 2,000 iterations to compute 95% CIs of the metrics including test accuracy and AUC[26].

Network training. Classification models for body part recognition and ICH detection were trained for 45 epochs using the Adam optimizer with default settings[27] and a mini-batch size of 80. *FBPConvNet* models were trained for 100 epochs using the Adam optimizer with default settings and a mini-batch size of 20. The base learning rate of 0.001 was decayed by a factor of 10 every 15 epochs for the classification models and every 33 epochs for *FBPConvNet*. The best classification and *FBPConvNet* models were selected based on the validation loss.

Infrastructure. We used radon and iradon functions in Matlab 2018a for generating sinograms and obtaining FBP reconstructed images, respectively. We used Keras (version 2.1.1) with a Tensorflow backend (version 1.3.0) as the framework for developing deep learning models, and performed experiments using an NVIDIA Devbox (Santa Clara, CA) equipped with four TITAN X GPUs with 12 GB of memory per GPU.

References

1. Esteva, A. *et al*. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* **542**, 115 (2017).
2. Gulshan, V. *et al*. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA* **316**, 2402–2410 (2016).
3. Chilamkurthy, S. *et al*. Deep learning algorithms for detection of critical findings in head ct scans: a retrospective study. *Lancet* (2018).
4. Thrall, J. H. *et al*. Artificial intelligence and machine learning in radiology: opportunities, challenges, pitfalls, and criteria for success. *J. Am. Coll. Radiol.* **15**, 504–508 (2018).
5. Wang, S. *et al*. Accelerating magnetic resonance imaging via deep learning. In *Biomedical Imaging (ISBI), 2016 IEEE 13th International Symposium on*, 514–517 (IEEE, 2016).
6. Zhu, B., Liu, J. Z., Cauley, S. F., Rosen, B. R. & Rosen, M. S. Image reconstruction by domain-transform manifold learning. *Nature* **555**, 487 (2018).
7. Xie, S. *et al*. Artifact removal using improved googlenet for sparse-view ct reconstruction. *Sci. reports* **8** (2018).
8. Jin, K. H., McCann, M. T., Froustey, E. & Unser, M. Deep convolutional neural network for inverse problems in imaging. *IEEE Transactions on Image Processing* **26**, 4509–4522 (2017).
9. Antholzer, S., Haltmeier, M. & Schwab, J. Deep learning for photoacoustic tomography from sparse data. *Inverse Probl. Sci. Eng.* 1–19 (2018).
10. Lee, H., Lee, J. & Cho, S. View-interpolation of sparsely sampled sinogram using convolutional neural network. In *Medical Imaging 2017: Image Processing*, vol. 10133, 1013328 (International Society for Optics and Photonics, 2017).
11. Lee, H., Lee, J., Kim, H., Cho, B. & Cho, S. Deep-neural-network-based sinogram synthesis for sparse-view ct image reconstruction. *IEEE Transactions on Radiat. Plasma Med. Sci.* **3**, 109–119 (2018).
12. Yuan, H., Jia, J. & Zhu, Z. Sipid: A deep learning framework for sinogram interpolation and image denoising in low-dose ct reconstruction. In *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, 1521–1524 (IEEE, 2018).
13. Wang, G., Ye, J. C., Mueller, K. & Fessler, J. A. Image reconstruction is a new frontier of machine learning. *IEEE transactions on medical imaging* **37**, 1289–1296 (2018).
14. Do, S. *et al*. High fidelity system modeling for high quality image reconstruction in clinical ct. *PloS one* **9**, e111625 (2014).
15. Do, S. & Karl, C. Sinogram sparsified metal artifact reduction technology (ssmart). In *The Third International Conference on Image Formation in X-ray Computed Tomography*, 798–802 (2014).
16. Do, S., Näppi, J. J. & Yoshida, H. Iterative reconstruction for ultra-low-dose laxative-free ct colonography. In *International MICCAI Workshop on Computational and Clinical Challenges in Abdominal Imaging*, 99–106 (Springer, 2013).
17. Do, S. *et al*. A decomposition-based ct reconstruction formulation for reducing blooming artifacts. *Phys. Med. Biol.* **56**, 7109 (2011).
18. Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J. & Wojna, Z. Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2818–2826 (2016).
19. Pons, J., Lidy, T. & Serra, X. Experimenting with musically motivated convolutional neural networks. In *Content-Based Multimedia Indexing (CBMI), 2016 14th International Workshop on*, 1–6 (IEEE, 2016).
20. McCollough, C. H., Leng, S., Yu, L. & Fletcher, J. G. Dual-and multi-energy ct: principles, technical approaches, and clinical applications. *Radiology* **276**, 637–653 (2015).
21. Ronneberger, O., Fischer, P. & Brox, T. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, 234–241 (Springer, 2015).
22. Huang, G., Liu, Z., Van Der Maaten, L. & Weinberger, K. Q. Densely connected convolutional networks. *CVPR* **1**, 3 (2017).
23. Russakovsky, O. *et al*. Imagenet large scale visual recognition challenge. *Int. journal computer vision* **115**, 211–252 (2015).
24. He, K., Zhang, X., Ren, S. & Sun, J. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In *Proceedings of the IEEE international conference on computer vision*, 1026–1034 (2015).
25. DeLong, E. R., DeLong, D. M. & Clarke-Pearson, D. L. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 837–845 (1988).
26. Efron, B. & Tibshirani, R. J. *An introduction to the bootstrap* (CRC press, 1994).
27. Kingma, D. P. & Ba, J. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980* (201 4).

Author contributions

Competing interests

S.D. is a consultant of Nulogix Health and Doai and received research supports from ZCAI, Tplus, and MediBloc.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41598-019-51779-5.

Correspondence and requests for materials should be addressed to S.D.

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.