Deep Learning Substitutes Gadolinium in Detecting Functional and Structural Brain Lesions with MRI

Chen Liu†1 Nanyan Zhu†2 Dipika Sikka3 Xinyang Feng3 Haoran Sun3 Xueqin Liu3 Sabrina Gjerswold-Selleck3 Hong-Jian Wei1 Pavan S. Upadhyayula5 Angeliki Mela5 Peter D. Canoll5 Cheng-Chia Wu6 Andrew F. Laine3 Jeffrey A. Lieberman7 Frank A. Provenzano8 Scott A. Small9 Jia Guo10,11 for the Alzheimer’s Disease Neuroimaging Initiative11

August 21, 2020

While MRI contrast agents such as those based on Gadolinium are needed to enhance the detection of structural and functional brain lesions, there are rising concerns over their safety. Here, we hypothesize that a deep learning model, trained using quantitative steady-state contrast-enhanced MRI datasets in mice and humans, could generate contrast-equivalent information from a single non-contrast MRI scan. The model was first trained, optimized, and validated in mice. It was then transferred and adapted to human data, and we find that it can substitute Gadolinium-based contrast agents for detecting functional lesions caused by aging, Schizophrenia, or Alzheimer’s disease, and, for enhancing structural lesions caused by brain or breast tumors. Since derived from a commonly-acquired MRI, this framework has the potential for broad clinical utility and can be applied retrospectively to research scans across a host of diseases.

Gadolinium-based contrast agents (GBCAs) are intravenously administered with MRI, for mapping, enhancing, and detecting brain lesions. For structural lesions such as tumors, where there is often a breakdown of the blood brain barrier, GBCAs enhance lesion detection and characterization. GBCAs can be also used to identify functional lesions, like those that alter brain metabolism in aging, neuropsychiatric disorders and neurodegeneration. For these functional lesions, GBCAs can be used to generate high-resolution cerebral blood volume (CBV) maps, thereby deriving a quantitative correlate of metabolic dysfunction that is otherwise invisible to non-contrast MRI.

Recent studies have begun to call the safety of GBCAs into question. Even if these concerns are addressed, GBCA administration in any case requires intravenous access, a requirement that places burdens on patients and on healthcare practitioners, as well as limiting its application in cases when contraindicated. Thus, there is a general need to find a ‘GBCA substitute’, one that can generate GBCA-equivalent information from a non-contrast MRI scan. Deep learning, a subset of machine learning,
holds great promise in this regard, and previous studies have shown that non-contrast MRI scans can be used to obtain contrast-equivalent information\(^\text{12}\). A deep learning model should therefore be able to learn how to optimally extract this information, by training the model on previously-acquired MRI datasets where GBCAs were administered. Indeed, a growing number of recent studies have begun validating this assumption\(^\text{13–15}\). Nevertheless, among these, one study managed to use deep learning to reduce the GBCA dose\(^\text{13}\), but not to completely substitute it. Other studies succeeded in obviating the need for GBCA\(^\text{14,15}\), but the deep learning models require the acquisition of an array of either six or ten different MRI sequences, some that are not widely available.

With these issues in mind, we hypothesized that a deep learning model could extract GBCA-equivalent information from a single and commonly-acquired high resolution MRI scan, by training and optimizing the model using a unique GBCA MRI dataset. Previous deep learning studies relied on GBCA datasets generated for radiological purposes, where post-GBCA scans are, by necessity rescaled, facilitating a radiologist’s ability to detect and characterize brain lesions. This rescaling, however, increases intrasubject variability across a dataset. With our specific interest in mapping functional brain lesions that localize to specific regions of the hippocampal formation, over the last couple of decades we have been using GBCAs to generate quantitative, high-resolution CBV maps\(^\text{5,6,16–19}\). By design, these quantitative maps do not re-scale the post-GBCA image. While not the original intent, this large-scale dataset has reduced intrasubject variability, an attribute that should benefit deep learning. In parallel to generating a large-scale and quantitative GBCA dataset in humans, we have also been generating a similar MRI dataset in mice\(^\text{6,20}\). Here again, the original intent was to validate patterns of hippocampal dysfunction observed across disease states, but because in mice studies subjects are siblings with identical genetic backgrounds, this mouse dataset is notable for even less intersubject variability than in humans.

Here, we exploit this distinct cross-species and quantitative GBCA dataset. Beginning with mice, we first designed, optimized, and trained a deep learning model, and then validated that it can substitute GBCA enhancement. We then forward applied this knowledge to humans, by showing that trained deep learning models can localize functional lesions that occur in the hippocampal formation in aging, Schizophrenia, and Alzheimer’s disease, and enhance structural lesions caused by brain tumors. Lastly,

---

**Fig. 1: Overview of the studies conducted.** We first performed proof-of-concept studies in mice to validate our hypothesis that deep learning can extract information equivalent to Gadolinium-based contrast agent (GBCA) contrast enhancement from a single-modal non-contrast MRI scan, and then conducted extensive analyses in humans to scrutinize the capability of this proposed approach.

**Study:** study conducted; **Aim:** purpose of the study; **Source:** where the imaging data come from; **MRI:** modality/type of data used in the study; **Model:** specific DeepContrast model used in the study; **Outcome:** specific utility of GBCA replicated by DeepContrast.

AD: Alzheimer’s disease; ADNI: Alzheimer’s Disease Neuroimaging Initiative dataset; BraTS: Brain Tumor Segmentation dataset; CBV: cerebral blood volume; CHR: clinical high-risk for Schizophrenia; Gd-Uptake: GBCA contrast uptake maps; GBM: glioblastoma multiforme; T2W: T2-weighted scans; T2W-CE: T2-weighted contrast-enhanced scans; T1W: T1-weighted scans; T1W-CE: T1-weighted contrast-enhanced scans.
we demonstrated that the method can be generalized to other organs, by showing that it can enhance structural lesions caused by breast tumors. The deep learning model will be referred to as ‘DeepContrast’. The studies conducted are outlined in Fig. 1.

Results

DeepContrast in the mouse brain

We first designed, optimized and trained the model on wildtype (WT) mice brain scans (37 for training and 6 for validation; see methods), in which we have previously generated quantitative T2-weighted GBCA-uptake brain maps. Similar to previous studies, we compared the similarities between the GBCA-predicted maps and the GBCA-uptake ground truth maps by performing voxel-wise analyses across the whole brain on a test set with 6 scans (Fig. 2a) using metrics that measure signal quality (peak

Fig. 2: Quantitative evaluation of the DeepContrast in the mouse brain. a. DeepContrast prediction (Gd-Predicted) highly concords with the ground truth GBCA-uptake map (Gd-Uptake) in the mouse brain. The non-contrast scans and the contrast-enhanced scans are displayed for reference. A healthy wild type (WT) subject is shown in the left panel while a subject with glioblastoma multiforme (GBM) is shown on the right panel. Color bars indicate the colormap and dynamic range used in the cross-sectional brain images. b-e, Similarity between the model prediction and the ground truth, evaluated on all scans in the test sets (b: WT, 6 scans; c-e: GBM, 4 scans) using quantitative metrics, where the non-contrast (T2W) scans are used as the performance baseline. f, ROC curves for mouse GBM high-enhancement region similarity assessment on the test set: ROC curve for the model prediction in comparison to the ground truth GBCA-uptake map (left) and ROC curve for the non-contrast (T2W) scans in comparison to the ground truth GBCA-uptake map (right). ROC curve for the model prediction (sensitivity = 0.72, specificity = 0.88, AUC = 0.85) outperforms the ROC curve for the non-contrast (T2W) scans (sensitivity = 0.50, specificity = 0.75, AUC = 0.64). The standard deviation is indicated by the shaded area.

For all voxel-based metrics, only the voxels within the brains or subregions are used. SSIM is calculated on the minimum bounding box of the brains or subregions. Asterisks indicate level of statistical significance (* p < 0.05, ** p < 0.01, *** p < 0.001, **** p < 0.0001).

PSNR: peak signal-to-noise ratio; SSIM: structural similarity index; P.R: Pearson correlation coefficient; S.R: Spearman correlation coefficient.
signal-to-noise ratio) and structural similarity (structural similarity index). We further incorporated two
other metrics to represent linear (Pearson correlation coefficient) and monotonic (Spearman correlation
coefficient) relationships across corresponding voxels. Between the maps, the peak signal-to-noise ratio was
24.59±0.60, the Pearson correlation coefficient was 0.695±0.008 \( (p < 0.0001) \), the Spearman correlation
coefficient was 0.606±0.008 \( (p < 0.0001) \), and the structural similarity index was 0.831±0.008 (Fig. 2b
and Table 1). With \( n < 10 \), individual data points were reported in Supplementary Table 1.

Next, we retrained the same network architecture after adding brain MRI scans from 6 mice with
glioblastoma multiforme (GBM) into the training set. Similar to a previous study\textsuperscript{14}, we compared
the similarity between the predicted and ground truth maps by performing a receiver operating
characteristic (ROC) analysis to measure the similarity of the high-enhancement regions (Fig. 2f) in
addition to the quantitative similarity metrics above (Fig. 2c-e) on a test set with 4 scans. Between the maps
across the whole brain, the peak signal-to-noise ratio was 21.07±0.43, the Pearson correlation coefficient
was 0.670±0.060 \( (p < 0.0001) \), the Spearman correlation coefficient was 0.442±0.035 \( (p < 0.0001) \), and
the structural similarity index was 0.737±0.019 (Fig. 2c and Table 1). We further calculated the same
metrics in tumor and non-tumor regions, and the results were illustrated in Fig. 2c-e and reported in
Table 1. The average ROC curve reached a sensitivity of 0.72 and a specificity of 0.88 at the operating
point, whereas the area under the curve (AUC) was 0.85 (Fig. 2f). With \( n < 10 \), individual data points
were reported in Supplementary Table 1.

DeepContrast in the human brain

We adapted the DeepContrast model to human brain MRI datasets by modifying the network architecture,
hyper-parameters and training strategies. First, same as the mouse study, we compared the similarities
between the GBCA-predicted maps and the GBCA-uptake ground truth maps by performing voxel-wise
analyses across the whole brain on a test set with 179 scans (Fig. 3a-b). Between the maps, the peak

Fig. 3: Quantitative evaluation of the DeepContrast in the cognitively normal human brain.

a. DeepContrast prediction (Gd-Predicted) highly concords with the ground truth GBCA-uptake map (Gd-
Uptake) in the cognitive normal human brain. Color bars indicate the colormap and dynamic range used
in the cross-sectional brain images. b. Similarity between the model prediction and the ground truth,
evaluated on 179 scans of cognitively normal (CN) subjects using quantitative metrics, where non-contrast
scans are used as the performance baseline. c. DeepContrast shows higher test-retest reliability than the
experimentally acquired Gd-Uptake ground truth.

For all voxel-based metrics, only the voxels within the brains or subregions are used. SSIM is calculated on
the minimum bounding box of the brains or subregions. Asterisks indicate level of statistical significance
\( (* p < 0.05, ** p < 0.01, *** p < 0.001, **** p < 0.0001) \).

PSNR: peak signal-to-noise ratio; SSIM: structural similarity index; P.R: Pearson correlation coefficient;
S.R: Spearman correlation coefficient.
signal-to-noise ratio was 29.64±0.07, the Pearson correlation coefficient was 0.822±0.002 ($p < 0.0001$), the Spearman correlation coefficient was 0.625±0.003 ($p < 0.0001$), and the structural similarity index was 0.879±0.002 (Fig. 3) and Table 1.

Besides the whole-brain analysis for similarity, we decided to extend our comparisons to two additional analyses. In the tissue of interest (TOI) analysis, we compared the similarities between the maps in white matter, gray matter, and cerebrospinal fluid (CSF). The results were illustrated in Fig. 3, and reported in Table 1. In the region-of-interest (ROI) analysis, we compared the similarities between the maps in 126 distinct ROIs in the whole brain segmented by FreeSurfer. Among the 126 ROIs, 121 had significant Pearson correlation coefficients ($p < 0.001$) and 123 had significant Spearman correlation coefficients ($p < 0.001$) (Supplementary Fig. 4).

Finally, we were also interested in the reproducibility in test-retest acquisitions. We conducted a series of test-retest reliability analyses on the GBCA-predicted maps versus the GBCA-uptake ground truth maps across the whole brain on a test set with 11 repeated scan pairs (Fig. 3). For the GBCA-predicted maps, the peak signal-to-noise ratio was 30.11±0.44, the Pearson correlation coefficient was 0.919±0.005 ($p < 0.0001$), the Spearman correlation coefficient was 0.722±0.010 ($p < 0.0001$), and the structural similarity index was 0.948±0.002. As a comparison, for the GBCA-uptake ground truth maps, the peak signal-to-noise ratio was 27.70±0.24, the Pearson correlation coefficient was 0.815±0.007 ($p < 0.0001$), the Spearman correlation coefficient was 0.415±0.017 ($p < 0.0001$), and the structural similarity index was 0.832±0.007. Among all the analyses, the test-retest reliabilities of the GBCA-predicted maps were consistently higher than the test-retest reliabilities of the GBCA-uptake ground truth maps ($p < 0.0001$) (Fig. 3). Besides the whole-brain analysis, we also performed the TOI analysis, and the results were illustrated in Fig. 3, and reported in Table 1. Among all metrics in all tissue types, the test-retest reliabilities of the GBCA-predicted maps were consistently higher than the test-retest reliabilities of the GBCA-uptake ground truth maps ($p < 0.0001$) (Fig. 3).

DeepContrast visualizes functional lesions

We generated GBCA-predicted maps from non-contrast T1-weighted MRI scans with DeepContrast, and subsequently quantified CBV-predicted maps with sub-millimeter in-plane resolution of 0.68×0.68 mm in the coronal planes and slice thickness of 3 mm (see methods). Then, we conducted voxel-based analyses (VBA) and ROI-based analyses on the CBV-predicted maps to pinpoint the precise sites of dysfunctions in normal aging, Schizophrenia clinical high-risk (CHR) and Alzheimer’s disease (AD).

Normal Aging. The first study we conducted aimed to validate whether DeepContrast can capture the subtle aging effects on basal metabolism. First, we focused on the hippocampal circuit (Fig. 4a). Replicating prior studies, age-related decline in CBV-predicted maps localized primarily to the dentate gyrus (DG) (Fig. 4a). In the complementary ROI analysis of the DG, the CBV-predicted values showed a linear decline with age ($\beta_{age} = -6.36e-4$, $t_{age} = -4.64$, $p_{age} = 6.85e-6$) (Fig. 4a). Second, we included two other regions, namely the aging-vulnerable inferior frontal gyrus (IFG) and the aging-resistant entorhinal cortex (EC) (Fig. 4a). The CBV-predicted maps demonstrated the same age-related trends as the ground truth CBV over these regions (Fig. 4a). Finally, we extended the analysis to the entire cortex and found that the CBV-predicted maps carried similar age-related changes as the ground truth CBV over all cortical ROIs (Supplementary Fig. 2). The multi-class ROC representing the concordance between CBV-predicted and ground truth CBV reached a sensitivity of 0.76 and a specificity of 0.89 at the operating point, whereas the AUC was 0.91 (Supplementary Fig. 2).

Schizophrenia. The second study we conducted aimed to validate whether DeepContrast can capture the regional vulnerability in patients who are clinically high-risk for Schizophrenia. First, we performed a VBA analysis over the hippocampal circuit (Fig. 4a). Replicating previous studies, elevated CBV-predicted values primarily localized to the bilateral CA1 regions, despite that the cluster-level significance did not reach $p < 0.05$ (Fig. 4a). In the complementary ROI analysis of the left anterior CA1 as defined previously, the CBV-predicted values were significantly higher ($p = 0.046$) in CHR patients compared to the normal controls (Fig. 4a). Next, we conducted a slice-based analysis similar to a prior study to inspect the CBV-predicted elevations in the left and right CA1 along the anterior-posterior axis. The slice-based analysis indicated significant elevations ($p < 0.05$) in five consecutive slices (thickness = 0.68 mm, see methods) within the left anterior CA1 region (Supplementary Fig. 3).

Alzheimer’s Disease. The third study we conducted aimed to validate whether DeepContrast can...
Fig. 4: DeepContrast maps differential anatomical patterns of dysfunction in the hippocampal formation. 

a. A three-dimensional rendering of the bilateral hippocampal formation (left panel) consisting of the hippocampus (HC) and the entorhinal cortex (EC) and axial, sagittal, and coronal slices from a group-wise T1-weighted MRI template cutting through the hippocampal formation (right three panels). The hippocampal formation is displayed with the hot-to-cold colormap along the anterior-to-posterior axis.

b. A voxel-based analysis on the CBV-predicted maps of 177 individuals ranging from 20-72 years of age reveals that the greatest age-related decline occurred in the body of the hippocampal circuit (top, color-coded by the degree of significance). A coronal slice (bottom), onto which the hippocampal formation mask is applied, reveals that age-related decline localizes primarily to the dentate gyrus. The voxel-based analysis is conducted using a multiple regression model in SPM12 using sex as a covariate and age as the regressor, and the age-related differences are contrasted using Student’s t test. Multiple comparisons are corrected for, yielding voxel-wise $p < 0.005$ and cluster-wise $p < 0.05$ (see methods).

c. A voxel-based analysis on the CBV-predicted maps of 74 Schizophrenia clinical high-risk (CHR) patients with 18 normal controls reveals CHR-related increase in the body of the hippocampal circuit (top, color-coded by the degree of significance). A coronal slice (bottom), onto which the hippocampal formation mask is applied, reveals that CHR-related increase localizes primarily to the CA1. The voxel-based analysis is conducted using a general linear model in SPM12 using a two-sample student’s t test after controlling for global variables, and the CHR-related differences are contrasted using Student’s t test. Multiple comparisons are corrected for, yielding voxel-wise $p < 0.005$ and cluster-wise $p = 0.3$ (see methods).

d. A voxel-based analysis on the CBV-predicted maps of 50 Alzheimer’s disease (AD) patients compared with 50 normal controls, each with 2 back-to-back scans, reveals AD-related reduction in the entorhinal cortex (top, color-coded by the degree of significance). A coronal slice (bottom), onto which the hippocampal formation mask is applied, reveals that AD-related decline localizes primarily to the transentorhinal cortex. The voxel-based analysis is conducted using a multiple regression model in SPM12 using age, sex and subject identity as covariates and diagnostic class (i.e., cognitive normal vs. dementia) as the regressor, and the AD-related difference are contrasted using Student’s t test. Multiple comparisons are corrected for, yielding voxel-wise $p < 0.005$ and cluster-wise $p < 0.05$ (see methods).

e. A scatter plot shows the association between age and mean CBV-Predicted values in the dentate gyrus after removal of gender effects ($\beta_{\text{age}} = -6.36e-4$, $t_{\text{age}} = -4.64$, $p_{\text{age}} = 6.85e-6$). Shaded area surrounding the regression line indicates the 95% confidence interval.

f. A box plot showing individual-subject mean CBV-Predicted values in the left anterior CA1 indicates a significant difference between patients with Schizophrenia clinical high-risk and healthy controls (two sample t-test one-tailed $p = 0.046$). Center line: median; box limits: upper and lower quartiles; whiskers: $1.5 \times$ interquartile range; points: outliers.

g. A box plot showing individual-subject mean CBV-Predicted values in the right transentorhinal cortex indicates a significant difference between patients with Alzheimer’s disease and healthy controls (two sample t-test one-tailed $p = 0.031$). Center line: median; box limits: upper and lower quartiles; whiskers: $1.5 \times$ interquartile range; points: outliers.

HC: hippocampus; EC: entorhinal cortex; DG: dentate gyrus; CA3: cornu Ammonis 3; CA1: cornu Ammonis 1; Sub: subiculum; Prs: presubiculum; PaS: parasubiculum.
Fig. 5: DeepContrast maps age-related changes in brain regions vulnerable and resistant to aging. a. A three-dimensional rendering of the inferior frontal gyrus (IFG), dentate gyrus (DG) and entorhinal cortex (EC) overlaid on a group-wise T1-weighted MRI template. b. The age-related regressions of cerebral blood volume maps (CBV) over these regions demonstrate the aging-vulnerability of IFG and DG and the aging-resistance of EC. c. The age-related regressions of CBV-predicted maps (CBV-Predicted) over the same regions demonstrated the same vulnerability or resistance to aging. d. The scatter plots of the ROI-mean CBV vs. CBV-Predicted values of the 177 subjects further show the consistency of the two measures.

DeepContrast enhances structural lesions

**Brain Tumor.** In order to accurately capture the high variance present within brain tumors, we retrained a DeepContrast model with a large-scale brain tumor MRI dataset collected from the Brain Tumor Segmentation (BraTS) database. First, we compared the similarities between the GBCA-predicted maps and the GBCA-uptake ground truth maps by performing voxel-wise analyses across the whole brain on a test set with 15 scans (Fig. 6a-b). Between the maps, the peak signal-to-noise ratio was 26.65±0.26, the Pearson correlation coefficient was 0.656±0.016 (p < 0.0001), the Spearman correlation coefficient was 0.439±0.028 (p < 0.0001), and the structural similarity index was 0.802±0.004. Next, we calculated the same metrics in tumor and non-tumor regions, and the results were illustrated in Fig. 6b and reported in Table 1. Finally, we conducted the ROC analysis, and the average ROC curve reached a sensitivity of 0.81 and a specificity of 0.85 at the operating point, whereas the AUC was 0.91 (Fig. 6c).

**Breast Tumor.** We further explored the possibility of extending DeepContrast to other organs. We conducted a study using the Breast-MRI-NACT-Pilot image collection which contains longitudinal dynamic contrast enhanced (DCE) MRI studies on patients undergoing neoadjuvant chemotherapy (NACT) for invasive breast cancer. As breast tumors are difficult to distinguish from healthy tissue without the additional contrast provided by GBCAs, the role of DeepContrast will be especially critical in screening for breast cancer. First, we compared the similarities between the GBCA-predicted maps and the GBCA-uptake ground truth maps by performing voxel-wise analyses across the whole breast on a test set with 16 scans (Fig. 6d-e). Between the maps, the peak signal-to-noise ratio was 27.40±0.64, the
Pearson correlation coefficient was $0.691 \pm 0.034$ ($p < 0.0001$), the Spearman correlation coefficient was $0.630 \pm 0.023$ ($p < 0.0001$), and the structural similarity index was $0.826 \pm 0.021$. Next, we calculated the same metrics in tumor and non-tumor regions, and the results were illustrated in Fig. 6 and reported in Table 1. Finally, we conducted the ROC analysis, and the average ROC curve reached a sensitivity of 0.77 and a specificity of 0.82 at the operating point, whereas the AUC was 0.87 (Fig. 6f).

**Fig. 6:** DeepContrast enhances structural lesions in human brain and breast MRIs. a. DeepContrast prediction (Gd-Predicted) highly concords with the ground truth GBCA-uptake map (Gd-Uptake) of structural lesions in human brain. Color bars indicate the colormap and dynamic range used in the cross-sectional brain images. b. Similarity between the model prediction and the ground truth, evaluated on 15 scans of subjects with glioblastoma multiforme (GBM) using quantitative metrics, where non-contrast scans are used as the performance baseline. c. ROC curves for human GBM high-enhancement region similarity assessment on the test set: ROC curve for the model prediction in comparison to the ground truth GBM GBCA-uptake map (top) and ROC curve for the non-contrast (T1W) scans in comparison to the ground truth GBCA-uptake map (bottom). ROC curve for the model prediction (sensitivity = 0.81, specificity = 0.85, AUC = 0.91) outperforms the ROC curve for the non-contrast (T1W) scans (sensitivity = 0.97, specificity = 0.07, AUC = 0.25). The standard deviation is indicated by the shaded area. d. DeepContrast prediction (Gd-Predicted) highly concords with the ground truth GBCA-uptake map (Gd-Uptake) of structural lesions in human breast. Color bars indicate the colormap and dynamic range used in the cross-sectional breast images. e. Similarity between the model prediction and the ground truth, evaluated on 16 scans of subjects with breast tumor using quantitative metrics, where non-contrast scans are used as the performance baseline. f. ROC curves for breast tumor high-enhancement region similarity assessment on the test set: ROC curve for the model prediction in comparison to the ground truth GBCA-uptake map (left) and ROC curve for the non-contrast (T1W) scans in comparison to the ground truth GBCA-uptake map (right). ROC curve for the model prediction (sensitivity = 0.77, specificity = 0.82, AUC = 0.87) outperforms the ROC curve for the non-contrast (T1W) scans (sensitivity = 0.59, specificity = 0.70, AUC = 0.70). The standard deviation is indicated by the shaded area. For all voxel-based metrics, only the voxels within the brains, breasts or subregions are used. SSIM is calculated on the minimum bounding box of the brains, breasts or subregions. Asterisks indicate level of statistical significance (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$).

PSNR: peak signal-to-noise ratio; SSIM: structural similarity index; P.R: Pearson correlation coefficient; S.R: Spearman correlation coefficient.
Table 1: Quantitative evaluations of DeepContrast models.

Evaluations varied depending on the aspects being assessed for each model. All metrics reported in the form of mean ± standard error of the mean (SEM).

PSNR: peak signal-to-noise ratio; P.R: Pearson correlation coefficient; S.R: Spearman correlation coefficient; SSIM: structural similarity index.

| Model                  | Evaluation          | Region                  | Data                  | PSNR     | P.R         | S.R         | SSIM     |
|------------------------|---------------------|-------------------------|-----------------------|----------|-------------|-------------|----------|
| Healthy Mouse Brain    | Similarity (WT)     | Whole Brain             | Go-Uptake & T2W       | 22.16±0.59 | 0.072±0.002 | 0.074±0.013 | 0.707±0.007|
|                        |                     |                         | Go-Uptake & T2W       | 24.59±0.90 | 0.965±0.008 | 0.606±0.008 | 0.831±0.006|
| Tumor Mouse Brain      | Similarity (GBM)    | Whole Brain             | Go-Uptake & T2W       | 18.00±0.44 | 0.092±0.076 | 0.017±0.026 | 0.629±0.009|
|                        |                     | Tumor                   | Go-Uptake & T2W       | 21.07±0.43 | 0.670±0.060 | 0.442±0.035 | 0.737±0.019|
|                        |                     | Non-Tumor               | Go-Uptake & T2W       | 10.15±1.27 | 0.301±0.053 | 0.357±0.085 | 0.484±0.035|
| Healthy Human Brain    | Similarity (CN)     | Whole Brain             | Go-Uptake & T1W       | 19.92±0.68 | 0.098±0.023 | -0.043±0.044 | 0.641±0.008|
|                        |                     | White Matter            | Go-Uptake & T1W       | 22.48±0.66 | 0.487±0.029 | 0.499±0.035 | 0.736±0.016|
|                        |                     | Gray Matter             | Go-Uptake & T1W       | 15.40±0.09 | -0.194±0.003 | -0.323±0.005 | 0.446±0.002|
|                        |                     | CSF                     | Go-Uptake & T1W       | 29.64±0.07 | 0.822±0.002 | 0.625±0.003 | 0.879±0.002|
|                        | Test-Retest Reliability (CN) | Whole Brain | Go-Uptake of two consecutive acquisitions | 30.11±0.44 | 0.919±0.005 | 0.722±0.010 | 0.946±0.002|
|                        |                     | White Matter            | Go-Predicted of two consecutive acquisitions | 31.18±0.18 | 0.713±0.015 | 0.129±0.008 | 0.943±0.003|
|                        |                     | Gray Matter             | Go-Predicted of two consecutive acquisitions | 35.14±0.40 | 0.899±0.006 | 0.368±0.014 | 0.986±0.000|
|                        |                     | CSF                     | Go-Predicted of two consecutive acquisitions | 30.42±0.44 | 0.937±0.004 | 0.596±0.013 | 0.978±0.001|
| Tumor Human Brain      | Similarity (Brain Tumor) | Whole Brain             | Go-Uptake & T1W       | 24.94±0.27 | 0.762±0.014 | 0.858±0.019 | 0.911±0.003|
|                        |                     | Tumor                   | Go-Uptake & T1W       | 26.33±0.48 | 0.882±0.009 | 0.800±0.008 | 0.966±0.001|
| Tumor Human Breast     | Similarity (Breast Cancer) | Whole Breast            | Go-Uptake & T1W       | 17.17±0.25 | 0.127±0.047 | 0.207±0.056 | 0.546±0.010|
|                        |                     | Tumor                   | Go-Uptake & T1W       | 26.65±0.26 | 0.566±0.016 | 0.439±0.028 | 0.802±0.004|
|                        |                     | Non-Tumor               | Go-Uptake & T1W       | 11.96±0.59 | 0.015±0.060 | 0.033±0.058 | 0.503±0.035|
|                        |                     | Go-Uptake & T2W         | 24.10±0.41 | 0.544±0.035 | 0.372±0.031 | 0.843±0.014 | 0.722±0.010 | 0.946±0.002|
|                        |                     | Go-Uptake & T2W         | 17.22±0.24 | 0.131±0.047 | 0.214±0.059 | 0.566±0.012 | 0.722±0.010 | 0.946±0.002|
|                        |                     | Go-Uptake & T2W         | 26.59±0.25 | 0.660±0.017 | 0.445±0.030 | 0.813±0.005 | 0.722±0.010 | 0.946±0.002|
|                        |                     | Go-Uptake & T1W         | 19.32±0.87 | 0.228±0.056 | 0.184±0.045 | 0.612±0.043 | 0.722±0.010 | 0.946±0.002|
|                        |                     | Tumor                   | Go-Uptake & T1W       | 27.40±0.64 | 0.691±0.034 | 0.630±0.023 | 0.826±0.021 | 0.722±0.010 | 0.946±0.002|
|                        |                     | Go-Uptake & T1W         | 13.29±1.05 | 0.450±0.070 | 0.456±0.072 | 0.753±0.037 | 0.722±0.010 | 0.946±0.002|
|                        |                     | Non-Tumor               | Go-Uptake & T1W       | 18.35±0.77 | 0.632±0.075 | 0.617±0.073 | 0.844±0.037 | 0.722±0.010 | 0.946±0.002|
|                        |                     | Go-Uptake & T1W         | 18.74±1.00 | 0.154±0.046 | 0.153±0.040 | 0.614±0.043 | 0.722±0.010 | 0.946±0.002|
|                        |                     | Go-Uptake & T1W         | 27.16±0.72 | 0.641±0.030 | 0.616±0.021 | 0.829±0.021 | 0.722±0.010 | 0.946±0.002|

Discussion

By using a quantitative GBCA dataset in mice and humans, we demonstrated that deep learning can, in principle, generate GBCA-equivalent information from a single and common MRI scan across an array of lesions.

GBCA’s utility for MRI can be organized around two primary pathophysiologies. The first is a breakdown of the blood-brain barrier that often accompanies many structural lesions, and in which case GBCA extravasates into the parenchyma and enhances lesion detection. The second is alterations in neuronal metabolism, typifying most functional disorders, in which case intravascular GBCA is used to quantify regional CBV, a hemodynamic variable tightly coupled to energy metabolism. As proof-of-principle, we optimized five models for our investigations across two species, two organs and multiple disorders. As GBCA’s utility can be reduced to two pathophysiologies we anticipate that future large-scale studies across a range of diseases might lead to two generalizable models—one for structural disorders that break down the blood-brain barrier, another for functional lesions that alters brain metabolism.

While we have demonstrated that DeepContrast can act as an alternative to GBCA for both pathophysiological categories, GBCA contrast is much subtler for functional compared to structural lesions, and so we consider our findings in aging, Schizophrenia and Alzheimer’s disease the strongest validation of our hypothesis. Nevertheless, since most of the concerns over GBCA’s safety have emerged when cancer patients are imaged multiple times over the course of their disease, validating DeepContrast in tumors was equally important.
DeepContrast’s utility can be organized according to its broad applications. The first is for research. There is an increasing number of brain MRI databases, such as ADNI (see Supplementary Table 2 for an example list of more than twenty open datasets), whose sole purpose is brain imaging and disease research. Standard T1-weighted MRI scans are among the most common acquisition across all of these datasets, typically acquired for mapping regional structural differences, such as regional volume or cortical thickness. DeepContrast can be retroactively applied to these datasets, and can be used to generate functional maps, significantly expanding pathophysiological insight that can be derived across the range of disorders.

DeepContrast’s second application is for patient care. For patient populations with structural lesions, such as cancer patients for example, GBCA administration will always be considered the gold standard, particularly during initial evaluation or for surgical planning. For these patients, however, DeepContrast might substitute GBCA when tracking the course of the disease or treatment response. For patient populations with functional lesions, those with neuropsychiatric and neurodegenerative disorders, a T1-weighted scan may be ordered as part of standard clinical practice, to exclude structural findings. For these patients, deriving CBV maps via DeepContrast potentially obviates the need for ordering other more invasive, burdensome, and expensive neuroimaging studies for mapping metabolic dysfunction.

Methods

Subjects

**Healthy Mouse Brain.** We used 49 healthy adult C576J/BL male mice (12-14 months old).

**Mouse GBM.** We included 10 adult C576J/BL male mice (3-6 months old) which were injected with PDGF (+/+) PTEN (-/−) p53 (−/−) glioblastoma cell. 50,000 cells in 1 µL solution were stereotactically injected into the brain. MRI scans of GBM mice were obtained 10 days after injection.

**Healthy Human Brain and Human Aging.** We aggregated the healthy human MRI data from a collection of previous acquisitions at Columbia University, resulting in 598 subjects (16-94 years old) with single acquisitions, and another 11 subjects with baseline and follow-up acquisitions 14 days apart. The aging study consists of 177 subjects (20-72 years old) that are cognitively normal.

**Human CHR.** We collected scans from a previous study that includes 92 subjects (15-35 years old), among which 74 are Schizophrenia clinical high-risk patients and 18 are normal controls.

**Human AD.** We randomly sampled 50 CN and 50 AD subjects from the Alzheimer’s Disease Neuroimaging Institute (ADNI) [33], resulting in a 100-subject (60-90 years old) dataset. Data used in the preparation of this study were obtained from the Alzheimer’s Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). The ADNI was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI has been to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer’s disease (AD). For up-to-date information, see [www.adni-info.org](http://www.adni-info.org).

**Human GBM.** We selected 316 subjects (36-86 years old) from the original 335 subjects present within the Brain Tumor Segmentation (BraTS) dataset [39–42] based on successful segmentation through the MALPEM segmentation pipeline [53].

**Human Breast Cancer.** We used all 68 subjects from the Breast-MRI-NACT-Pilot dataset [43,44].

Image acquisition protocols

**Healthy Mouse Brain and Mouse GBM.** We used CBV-fMRI to image two independent groups of mice, young and old male WT and GBM mice used in healthy mouse brain and mouse GBM studies, with the imaging protocol as previously described [20]. A Bruker BioSpec 94/30 (field strength, 9.4 T; bore size, 30 cm) horizontal small animal MRI scanner equipped with CryoProbe and software ParaVision 6.0.1 (Bruker BioSpin, Billerica, MA, USA) and a 23-mm 1H circularly polarized transmit/receive capable
mouse head volume coil were used for the imaging. Mice were anesthetized using medical air and isoflurane (3% volume for induction, 1.1-1.5% for maintenance at 1 liter/min air flow, via a nose cone). A flowing water heating pad was used to maintain the body temperature at around 37°C. Sterile eye lubricant was applied before each scan. T2-weighted images were acquired before and 36 min after intraperitoneal injections of the GBCA-based contrast agent Gadodiamide (Omniscan; GE Healthcare, Princeton, NJ, USA) at the dosage of 10 mmol/kg. T2-weighted images were acquired with a Refocused Echoes (RARE) sequence (repetition time (TR) = 3,500 ms, effective echo time (TE) = 45 ms, rapid acquisition and relaxation enhancement (RARE) factor = 8, voxel size = 76×76×450 µm).

**Healthy Human Brain and Human Aging.** The images were acquired under a steady-state CBV-fMRI protocol as previously described. A gradient echo T1-weighted scan (TR = 6.7 ms, TE = 3.1 ms, field of view (FOV) = 240×240×192 mm, voxel size = 0.9×0.9×0.9 mm) was acquired before a pair of un-scaled T1-weighted images (TR = 7 ms, TE = 3 ms, FOV = 240×240×196 mm, voxel size = 0.68×0.68×3 mm), all using a Philips Achievia 3.0-T MRI scanner. The image resolution used results from a systematic exploration of the scan protocol’s parameters. Scans were acquired before and after a bolus injection of a GBCA-based contrast agent (Omniscan, GE Healthcare).

**Human CHR.** The T1-weighted images are acquired using the same scan parameters as mentioned in the studies above (Philips Achievia 3.0-T MRI scanner, TR = 7 ms, TE = 3 ms, FOV = 240×240×196 mm, voxel size = 0.68×0.68×3 mm).

**Human AD (ADNI).** The images included in our studies were acquired using a customized back-to-back 3D magnetization prepared rapid gradient echo (MP-RAGE) protocol, yielding near-isotropic images (voxel size around 1×1×1 mm). For more details, refer to the ADNI paper.

**Human GBM (BraTS).** The images were acquired using different protocols from 19 institutions, the majority of which was acquired with 3D acquisition and voxel spacing of isotropic 1 mm. For more details, refer to the BraTS paper.

**Human Breast Cancer (Breast-MRI-NACT-Pilot).** All breast MRI used in this study were acquired on a 1.5-T scanner (Signa, GE Healthcare, Milwaukee, WI) using a bilateral phased array breast coil. The MR imaging protocol included a 3D localizer and a unilateral sagittal DCE acquisition. The DCE acquisition utilized a high spatial resolution, low temporal resolution, T1-weighted, fat-suppressed 3D fast gradient-recalled echo sequence developed for pre-surgical staging (TR = 8, TE = 4.2, flip angle = 20 degrees; FOV = 18-20 cm, acquisition matrix = 256×192×60, voxel size = 0.7×0.94×2.0 mm). For more details, refer to the dataset description.

Preprocessing and partitioning

**Healthy Mouse Brain.** In total, 49 WT mice were used in this study. Whole brain T2W MRI scans before (T2W) and 35 mins after intraperitoneal injection (T2W-CE) of Gadodiamide at 10 mmol/kg were acquired with identical scan parameters as previously described in CBV-fMRI protocol. The Gd-Uptake ground truth was quantified with the standardized delta-R2, which was derived using the same method as discussed before, followed by a standardization to the dynamic range of [0, 1]. We used 3D PCNN with manual correction to generate brain masks, which we used as training fields over which the model was optimized and performance metrics were calculated. A train-validation-test ratio at 8:1:1 was applied in the healthy mouse brain model training.

**Mouse GBM.** For scans of tumor subjects, the CBV maps and brain masks were derived using the same methods as described in the healthy mouse brain study, and tumor masks were generated in addition to the brain masks using the Fuzzy-C-Means segmentation. 6 GBM subjects were added to the training set while 4 GBM subjects replaced the original testing set of the Healthy Mouse Brain Model.

**Healthy Human Brain.** T1-weighted MRI scans were acquired using the protocols as described previously before (T1W) and 4 minutes after (T1W-CE) intravenous injection of Gadodiamide. During the MRI acquisition for the same session, the receiver gain was kept constant and the offset was set to zero, and as a result the T1W and T1W-CE scans share the same scaling and zero shifting. Each T1W & T1W-CE pair was spatially aligned when provided. For intensity normalization, each
T1W scan was compressed to the dynamic range of [0, 1], and the corresponding T1W-CE scan was scaled accordingly to match the constant scaling. The Gd-Uptake ground truth was quantified with the steady-state MRI method by subtracting the normalized T1W scans from the respective T1W-CE scans. We generated brain masks using the BET function in FMRIB Software Library (FSL), which we used as training fields over which the model was optimized and performance metrics were calculated. We generated tissue label maps using the FAST function in FSL for tissue-of-interest analyses. The ADNI protocol and the CBV-fMRI protocol, thus yielding the mismatch in appearance. The subjects in the AD study cohort, and the resulting dataset contains 100 scans (50 subjects) of normal controls and 100 scans (50 subjects) of patients with dementia. A major challenge we needed to face was that the appearance and anatomy of the scans used in the AD study notably differ from those used to train the DeepContrast Healthy Human Brain Model. They were acquired under the same field strength (i.e., 3T), but the specific scan parameters such as echo time and repetition time are different between the ADNI protocol and the CBV-fMRI protocol, thus yielding the mismatch in appearance. The subjects in the AD study are generally older (60-90 years old) and half of them suffered from Alzheimer’s, thus resulting in the mismatch in anatomy. We approached these issues by first minimizing the between-cohort appearance difference using a dynamic histogram warping (DHW) algorithm as it was demonstrated to be among the best intensity matching methods in medical imaging. Specifically, we calculated the mean normalized-brain-region 2048-bin histogram of each cohort, derived a bin-to-bin mapping between the cohorts, and applied the mapping to each individual scan in the AD study.

Secondly, we minimized the anatomical difference by running a diffeomorphic registration prior to applying the DeepContrast model. After these two steps, we normalized the scans to the dynamic range of [0, 1] and provided them to the model to generate the Gd-Predicted maps. CBV-predicted maps were then generated by applying the same normalization method as we would to quantify CBV maps.

Human Aging. The 177-subject cohort used for the aging study was a subset of the 179 subjects in the test set of the Healthy Human Brain Model, where 2 subjects were dropped due to low segmentation quality through the FreeSurfer (v6.0.0) Parcellation. After normalization to the dynamic range of [0, 1], the scans were directly treated as inputs to the model to generate the Gd-Predicted maps. CBV-predicted maps were then generated by applying the same normalization method as we would to quantify CBV maps.

Human CHR. The 92-scan cohort for the CHR study was acquired using the same scan parameters as those used to train the DeepContrast Healthy Human Brain Model, which ensures minimal discrepancy in scan appearance. CHR patients present very little structural deformation, which ensures minimal discrepancy in scan anatomy. Therefore, no additional measure needs to be taken to deal with appearance or anatomy variances. After normalization to the dynamic range of [0, 1], the scans were directly treated as inputs to the model to generate the Gd-Predicted maps. CBV-predicted maps were then generated by applying the same normalization method as we would to quantify CBV maps.

Human AD. We gathered the two back-to-back repeated baseline scans for each subject in the AD study cohort, and the resulting dataset contains 100 scans (50 subjects) of normal controls and 100 scans (50 subjects) of patients with dementia. A major challenge we needed to face was that the appearance and anatomy of the scans used in the AD study notably differ from those used to train the DeepContrast Healthy Human Brain Model. They were acquired under the same field strength (i.e., 3T), but the specific scan parameters such as echo time and repetition time are different between the ADNI protocol and the CBV-fMRI protocol, thus yielding the mismatch in appearance. The subjects in the AD study are generally older (60-90 years old) and half of them suffered from Alzheimer’s, thus resulting in the mismatch in anatomy. We approached these issues by first minimizing the between-cohort appearance difference using a dynamic histogram warping (DHW) algorithm as it was demonstrated to be among the best intensity matching methods in medical imaging. Specifically, we calculated the mean normalized-brain-region 2048-bin histogram of each cohort, derived a bin-to-bin mapping between the cohorts, and applied the mapping to each individual scan in the AD study. Secondly, we minimized the anatomical difference by running a diffeomorphic registration prior to applying the DeepContrast model. After these two steps, we normalized the scans to the dynamic range of [0, 1] and provided them to the model to generate the Gd-Predicted maps. CBV-predicted maps were then generated by applying the same normalization method as we would to quantify CBV maps.

Human GBM. The data used in the Human GBM study were collected from the BraTS dataset. The BraTS dataset includes T1W, T1W-CE, and tumor region scans. The T1W and T1W-CE pairs were spatially aligned when provided. The train-validation-test split at the subject level resulted 268 for training, 33 for validation and 15 for testing. N4 bias field correction (part of ANTS) and rigid registration was done to register all scans across subjects to a common template space. T1W scans were normalized using the maximum of the non-tumor region present within each scan. Scaling correction was then used to adjust T1W-CE scans. Specifically, T1W and T1W-CE pairs must be in the same scaling system in order to determine Gd-Uptake. This scaling correction can be applied by identifying a region which remains fairly unchanged after contrast enhancement. White matter has shown to demonstrate this property. White matter regions were identified using the MALPEM segmentation pipeline and T1W-CE scans were then scaled using a scaling ratio calculated from the average intensities in the T1W and T1W-CE white matter regions.

Human Breast Cancer. For each DCE acquisition, we included the non-contrast (T1W) scan and the scan acquired at the first time point of the DCE protocol (T1W-CE), totaling a number of 161 pairs. The train-validation-test split at the subject level resulted 129 for training, 16 for validation and 16 for testing. Corresponding tumor masks were also present within the dataset, and processed to ensure proper
spatial matching with the T1W and T1W-CE pairs. Whole breast masks were generated using k-means clustering. T1W and T1W-CE pairs were both normalized using the maximum of the T1W scans before being fed into the DeepContrast model.

DeepContrast and its implementations

All five model variants developed in our studies, as mentioned in Fig. 1, shared the common residual attention U-Net (RAU-Net) architecture (Fig. 7). Model inputs were the non-contrast MRI scans, while the outputs were the corresponding predicted GBCA contrast (Gd-Predicted). The inputs and outputs were in equal dimensions and were either 2D or 3D depending on the nature of the scan protocols (i.e., 2D slices were used for 2D MRI scans, whereas 3D volumes were used for 3D MRI scans).

The RAU-Net is an extension of the arguably most popular deep learning architecture in medical imaging, the U-Net, with the addition of residual blocks and the attention gates. As an example of a convolutional neural network (CNN), the U-Net extracts imaging features by utilizing local convolutions along the entire image or volume. The U-Net consists of multiple encoding layers across which the image dimension shrinks whereas the feature dimension increases so that compact high-level abstractions are generated along the process, and the same number of decoding layers to decipher these abstractions into

![DeepContrast Model Diagram](image_url)

**Fig. 7:** Training strategies of the various DeepContrast models we implemented in our studies. a. This training strategy is shared by the Healthy Mouse Brain Model and the Tumor Mouse Brain Model. Preprocessing includes intensity normalization and brain extraction. Ground truth Gd-Uptake was derived using the standardized delta-R2 equation. Note that there is an additional standardization step that maps the dynamic range of the standardized delta-R2 to the range of [0, 1], before the application of the brain mask. The loss function is calculated between the Gd-Uptake and the predicted version only using the voxels within the brain mask region. b. This training strategy applies to the Healthy Human Brain Model. Preprocessing includes intensity normalization and brain extraction. Ground truth Gd-Uptake was derived using the steady-state delta-R1 equation. The loss function is calculated between the Gd-Uptake and the predicted version only using the voxels within the brain mask region. c. This training strategy applies to the Tumor Human Brain Model. Preprocessing includes intensity normalization and brain extraction. Ground truth Gd-Uptake was derived using the steady-state delta-R1 equation. The loss function is calculated between the Gd-Uptake and the predicted version only using the voxels within the brain mask region. d. This training strategy applies to the Tumor Human Breast Model. Preprocessing includes intensity normalization and brain extraction. Ground truth Gd-Uptake was derived using the steady-state delta-R1 equation. The loss function is calculated between the Gd-Uptake and the predicted version only using the voxels within the breast mask region.
image space information. The add-on residual blocks simplify the entities to be approximated across each layer and therefore enables training of deeper networks, while the attention gates learn to differentially enhance or suppress specific regions in the feature maps so that the downstream outcomes are better represented for targeting objective.

Specifically, the encoding and decoding paths consist of the same number of residual convolution blocks that utilize concatenation, attention mechanisms and skip connections such that layers feed not only into the next layer, but into the layer after the next layer. On the encoding path, each residual block is followed by a max-pooling layer, and the last feature map feeds into a bottleneck layer with 3×3 convolution and batch normalization, connecting the deepest layer to the decoding path with several more blocks alternating one un-pooling layer and one residual block. Skip connections concatenate the output of each dense layer in the encoding path with the respective un-pooled feature map of the same size before feeding it as input to the decoding residual block. The output of the last decoding layer is the input for a 1×1 convolution layer that produces the final Gd-Predicted map.

**Healthy and Tumor Mouse Brain Model.** The model used in mouse studies (Supplementary Fig. 4) was a 2D RAU-Net that consisted of 5 encoding and decoding layers. The model input was a 2D axial slice of the mouse brain scans. Adam optimizer with a learning rate of 0.001 was used in this study. Our batch size was 3 and the loss function was mean squared error (MSE).

**Healthy Human Brain Model.** The model used in the healthy human study and further applied to the Aging, CHR and AD studies (Supplementary Fig. 5) was a 2D RAU-Net that consisted of 6 encoding and decoding layers. The model input was a 2D coronal slice of the human brain scans. SGD optimizer with an adaptive learning rate handle with 0.1 initial learning rate was used in this study. Our batch size was 4 and a robust adaptive loss function was utilized. The robust adaptive loss function is a generalization of the Cauchy/Lorentzian, Geman- McClure, Welsch/ Leclerc, generalized Charbonnier, Charbonnier/pseudo-Huber/L1-L2, and L2 loss functions. By introducing robustness as a continuous parameter, the robust adaptive loss function allows algorithms built around robust loss minimization to be generalized, which improves performance on basic vision tasks like calculating the intensity mapping function in our case.

**Tumor Human Brain Model.** The model used in the human tumor brain study (Supplementary Fig. 6) was a 3D RAU-Net that consisted of 6 encoding and decoding layers. The model input was a 3D human brain volume. SGD optimizer with an adaptive learning rate handle with 0.001 initial learning rate was used in this study. Our batch size was 1 and the robust adaptive loss function was utilized. TorchIO was used for data augmentation to counteract the insufficiency of tumor variance presented in the training data.

**Tumor Human Breast Model.** The model used in the tumor human breast study (Supplementary Fig. 7) was a 2D RAU-Net that consisted of 5 encoding and decoding layers. The model input was a 2D sagittal slice of the human breast scans. SGD optimizer with an adaptive learning rate handle with 0.1 initial learning rate was used in this study. Our batch size was 4 and a robust adaptive loss function was utilized.

**Statistical methods**

**Prediction-GT similarity assessment (applicable to Healthy Mouse Brain, Healthy Human Brain, Mouse GBM, Human GBM, Human Breast Cancer).** Peak signal-to-noise ratio (PSNR), structural similarity index (SSIM), Pearson correlation coefficient (P.R) and Spearman correlation coefficient (S.R) were used to quantify the performance of all the DeepContrast models. PSNR, Pearson correlation coefficient and Spearman correlation coefficient were evaluated within the brains, breasts, or subregions, and SSIM was calculated in the minimum bounding box around the brains, breasts, or subregions.

**Prediction-GT high-enhancement region similarity ROC assessment (applicable to Mouse GBM, Human GBM, Human Breast Cancer).** Each ROC figure contained the same number of individual ROC curves as the number of ground truth-prediction candidate pairs (i.e., 4 for Mouse GBM, 15 for Human GBM and 16 for Human Breast Cancer). The ground truths were the GBCA-uptake maps while the prediction candidate were either the GBCA-predicted maps or the non-contrast scans. The average ROC was shown as a solid black curve while the standard deviation was shown as the shaded
Voxel-based analysis for regional vulnerability localization: Human Aging. Voxel-based analysis (Fig. 4b) was performed by first transforming the non-contrast images using a diffeomorphic registration algorithm with nearest-neighbor interpolation to an unbiased brain template created from the 177 scans in the Aging study. The GBCA-predicted maps were generated by the Healthy Human Brain model using the native-space non-contrast T1W scans as the input and were subsequently used to generate CBV-predicted maps by normalization using mean value among the top 10% brightest voxels within the brain region (representing signal intensity from pure blood). These CBV-predicted maps were then transformed into the template using the same transformation parameters calculated from the registration process, and subsequently smoothed using a 3mm-diameter spherical kernel. Transformed and filtered CBV-predicted maps were analyzed using SPM12. Data were analyzed with a multiple regression model, including sex as a covariate and age as the regressor. Age-related differences were contrasted using Student’s t test. FreeSurfer regional segmentation was then performed on the unbiased template image, and the hippocampal formation mask is generated by binarizing and combining the labels corresponding to the hippocampus and entorhinal cortex. The age-related regression t-map was then projected onto the MNI-152 brain template using diffeomorphic transformation with nearest-neighbor interpolation. The result was thresholded at \( p < 0.005 \) and corrected for multiple comparisons at the cluster level within the hippocampal formation using a Monte-Carlo simulation implemented in AFNI-3dClustSim (10,000 iterations) to yield a corrected \( p < 0.05 \). The final corrected age-related regression t-map was then overlaid onto the MNI-152 template in cross-section using 3DSlicer and also displayed with composite-with-shading volume rendering over semi-transparent models of the hippocampal formation.

ROI-based analysis: DG in Human Aging. The 177 native-space CBV-predicted scans were used to conduct the DG ROI analysis. A multiple linear regression with sex as a covariate and age as the regressor was conducted over the bilateral dentate gyrus (DG), as defined by FreeSurfer parcellation. A scatter plot was drawn (Fig. 4e) with each point representing the DG-mean CBV-predicted value after removal of sex effect for one subject.

Whole Brain Aging Analysis. The GBCA-predicted maps were generated in the native space of each subject and were afterwards used for CBV quantification together with the experimentally acquired ground truth GBCA-uptake maps using the same whole brain top 10% mean normalization. Similarly, the T1W scans were normalized to generate a comparable counterpart. We used T1W scans for comparison because they were the only input to the DeepContrast model to generate GBCA-predicted maps. The CBV (quantified from Gd-Uptake), CBV-Predicted (quantified from Gd-Predicted), and normalized T1W scans were used for age-related regression in the multiple brain regions. Multiple linear regressions with sex as a covariate and age as the regressor were conducted using the mean CBV/CBV-Predicted/T1W values extracted from the region across 177 subjects, over selected regions (Fig. 5) and over all 72 cortical ROIs (Supplementary Fig. 2). The ROIs were parcellated by FreeSurfer over the T1W scans in the native space in order to minimize segmentation errors.

For the ROC analysis, each ROC figure contained 1,000 individual ROC curves. The average ROC was shown as a solid black curve while the standard deviation was shown as the shaded area. All these individual ROC curves were computed using one pair of ground truth (CBV) t-score map and prediction candidate (CBV-predicted or non-contrast T1W) t-score map. Both the ground truth t-score map and the prediction candidate t-score map were binarized into 2 classes at 1,000 different binarization thresholds evenly distributed between the minimum and the maximum value, yielding 1,000 versions for each. Each individual ROC curve was derived using the regular ROC computation method as described above with one of the 1,000 versions of the ground truth and all 1,000 versions of the prediction candidate.
from the 92 scans from the CHR study. The GBCA-predicted maps were generated by the Healthy
Human Brain model using the native-space non-contrast T1W scans as the input and were subsequently
used to generate CBV-predicted maps using the same method as mentioned in the Aging study above.
These CBV-predicted maps were then transformed into the template using the same transformation
parameters calculated from the registration process, and subsequently smoothed using a 3mm-diameter
spherical kernel. Transformed and filtered CBV-predicted maps were analyzed using SPM12. Data were
analyzed with a two-sample t-test after controlling for global variables. CHR-related differences were
contrasted using Student’s t test. FreeSurfer regional segmentation was then performed on the unbiased
template image, and the hippocampal formation mask was generated by binarizing and combining
the labels corresponding to the hippocampus and the entorhinal cortex. The CHR-related regression
t-map was then projected onto the MNI-152 brain template using diffeomorphic transformation with
nearest-neighbor interpolation. The result was thresholded at \( p < 0.005 \) and showed two clusters,
although the two clusters (\( p = 0.3 \)) did not reach \( p < 0.05 \) after cluster level multiple comparisons
using AFNI-3dClustSim (10,000 iterations). The final corrected CHR-related regression t-map was
then overlaid onto the MNI-152 template in cross-section using 3DSlicer, and also displayed with
composite-with-shading volume rendering over semi-transparent models of the hippocampal formation.

**ROI-based analysis: Left anterior CA1 in Human CHR.** The 92 template-space CBV-predicted
scans were used to conduct the left anterior CA1 ROI analysis. A two-sample t-test was conducted over
the left anterior CA1. A box plot overlaid with individual data points was shown in Fig. 4 to indicate
the group-wise difference between the normal controls and the CHR patients.

**Slice-based analysis for regional vulnerability localization: Human CHR.** Slice-based analysis
(Supplementary Fig. 3) was performed by first transforming the non-contrast images using a diffeomorphic
registration algorithm with nearest-neighbor interpolation to an unbiased brain template created from
the 92-scan population. The GBCA-predicted map was generated by the Healthy Human Brain model
using the native-space non-contrast T1W scans as the input and was subsequently used to quantify
CBV-predicted maps by normalizing them by their respective mean value among the top 10% brightest
voxels within the brain region. These CBV-predicted maps were then transformed into that template
using the identical transformation parameters calculated from the registration process, and subsequently
smoothed using a 3mm-diameter spherical kernel. Next we upsampled the unbiased template as well as
these CBV-predicted scans to an isotropic resolution (voxel size = 0.68×0.68×0.68 mm) using cubic
spline interpolation. We parcellated the hippocampal subfields of the template using FreeSurfer, and
further cut the left and right CA1 subregions in the hippocampus into slices along the anterior-posterior
axes of these structures. We computed the slice-mean CBV-Predicted values for each slice, followed by a
3-slice sliding window averaging to smooth the results. Then we ran two-sample t-tests as discussed
before over the smoothed slice-mean CBV-Predicted values to generate the slice-based analysis results.

**Voxel-based analysis for regional vulnerability localization: Human AD.** Voxel-based analysis
(Fig. 4) was performed by first transforming the non-contrast images using a diffeomorphic registration
algorithm with nearest-neighbor interpolation to an unbiased brain template created from the 200
scans (i.e., 100 subjects each with 2 back-to-back repeated scans) in the AD study. We then ran these
non-contrast scans through the DeepContrast Healthy Human Brain Model to generate CBV-predicted
maps, which were subsequently smoothed using a 3mm-diameter spherical kernel. Unlike in the aging
study, the application of DeepContrast was performed after the registration process to help eliminate
major anatomical variances, since the deformations present in the diseased population were not previously
observed by the model trained on healthy data. GBCA-predicted scans, the direct output of the model,
were used to quantify CBV-predicted maps using the same method as described in the Aging study above.
These CBV-predicted maps, already co-registered upon creation, were analyzed using SPM12. Data
were analyzed with a multiple regression model, including age, sex and subject identity as covariates
and diagnostic class (i.e., cognitive normal vs. dementia) as the regressor. AD-related differences were
contrasted using Student’s t test. FreeSurfer regional segmentation was then performed on the unbiased
template image, and the hippocampal formation mask was generated by binarizing and combining the
labels corresponding to the hippocampus and the entorhinal cortex, while an extended hippocampal
formation mask was additionally generated to also include the parahippocampal cortex. The AD-related
regression t-map was then projected onto the MNI-152 brain template using diffeomorphic transformation
with nearest-neighbor interpolation. The result was thresholded at \( p < 0.005 \) and corrected for multiple
comparisons at the cluster level within the extended hippocampal formation using a Monte-Carlo
simulation implemented in AFNI-3dClustSim (10,000 iterations) to yield a corrected \( p < 0.05 \). The final corrected AD-related regression t-map was then overlaid onto the MNI-152 template in cross-section using 3DSlicer, and also displayed with composite-with-shading volume rendering over semi-transparent models of the hippocampal formation.

**ROI-based analysis: Right TEC in Human AD.** The 200 template-space CBV-predicted scans were used to conduct the right transentorhinal cortex (TEC) ROI analysis. A two-sample t-test was conducted over the right TEC, at the boundary between the right entorhinal cortex (EC) and the right parahippocampal cortex (PHC). The region was defined as the intersection between the EC-PHC region and a sphere centered at the middle of the EC-PHC intersection and spanning a diameter of the extent of the EC-PHC boundary (11 mm). A box plot overlaid with individual data points was drawn (Fig. 4g) to indicate the group-wise difference between the normal controls and the AD patients.

**Data Availability**

The trained Healthy Human Brain Model, alongside the test-retest reliability dataset (n = 11, each with two test-retest acquisitions) with both non-contrast scans and ground truth GBCA-uptake maps, is available on GitHub (link to be announced). The scripts that predict GBCA-uptake maps from non-contrast scans, as well as the script performing quantitative evaluations, are included. All code and data (except for those from public datasets) are proprietary and managed by the Columbia Technology Ventures Office of Intellectual Property. The custom training code and large-scale datasets are not publicly available.

**Code Availability**

The authors declare that all data supporting the results in this study are available from the corresponding author J.G. upon reasonable request, after permission from the Columbia Technology Ventures Office of Intellectual Property.

**Reference**

[1] Lohrke, J. *et al.* 25 years of contrast-enhanced mri: developments, current challenges and future perspectives. *Advances in therapy* **33**, 1–28 (2016).

[2] Borges, A., López-Larrubia, P., Marques, J. & Cerdan, S. Mr imaging features of high-grade gliomas in murine models: how they compare with human disease, reflect tumor biology, and play a role in preclinical trials. *American journal of neuroradiology* **33**, 24–36 (2012).

[3] Shen, Q. & Duong, T. Q. Magnetic resonance imaging of cerebral blood flow in animal stroke models. *Brain circulation* **2**, 20 (2016).

[4] Belliveau, J. *et al.* Functional mapping of the human visual cortex by magnetic resonance imaging. *Science* **254**, 716–719 (1991).

[5] Schobel, S. A. *et al.* Imaging patients with psychosis and a mouse model establishes a spreading pattern of hippocampal dysfunction and implicate glutamate as a driver. *Neuron* **78**, 81–93 (2013).

[6] Khan, U. A. *et al.* Molecular drivers and cortical spread of lateral entorhinal cortex dysfunction in preclinical Alzheimer’s disease. *Nature neuroscience* **17**, 304–311 (2014).

[7] Lewandowski, N. M. *et al.* Regional vulnerability in huntington’s disease: fmri-guided molecular analysis in patients and a mouse model of disease. *Neurobiology of disease* **52**, 84–93 (2013).

[8] Quattrocchi, C. C. & van der Molen, A. J. Gadolinium retention in the body and brain: is it time for an international joint research effort? (2017).

[9] Ramalho, M., Ramalho, J., Burke, L. M. & Semelka, R. C. Gadolinium retention and toxicity—an update. *Advances in chronic kidney disease* **24**, 138–146 (2017).
[10] Guo, B. J., Yang, Z. L. & Zhang, L. J. Gadolinium deposition in brain: current scientific evidence and future perspectives. *Frontiers in molecular neuroscience* **11**, 335 (2018).

[11] Dillman, J. R. & Davenport, M. S. Gadolinium retention—5 years later…. *Pediatric Radiology* **50**, 166–167 (2020).

[12] Small, S. A. *et al.* Imaging physiologic dysfunction of individual hippocampal subregions in humans and genetically modified mice. *Neuron* **28**, 653–664 (2000).

[13] Gong, E., Pauly, J. M., Wintermark, M. & Zaharchuk, G. Deep learning enables reduced gadolinium dose for contrast-enhanced brain mri. *Journal of Magnetic Resonance Imaging* **48**, 330–340 (2018).

[14] Klesieck, J. *et al.* Can virtual contrast enhancement in brain mri replace gadolinium?: a feasibility study. *Investigative radiology* **54**, 653–660 (2019).

[15] Liu, J., Gong, E., Christen, T. & Zaharchuk, G. Contrast-free mri contrast enhancement with deep attention generative adversarial network. *ISMRM 27th Annual Meeting and Exhibition* (2019).

[16] Small, S. A., Schobel, S. A., Buxton, R. B., Witter, M. P. & Barnes, C. A. A pathophysiological framework of hippocampal dysfunction in ageing and disease. *Nature Reviews Neuroscience* **12**, 585–601 (2011).

[17] Brickman, A. M. *et al.* Enhancing dentate gyrus function with dietary flavanols improves cognition in older adults. *Nature neuroscience* **17**, 1798–1803 (2014).

[18] Provenzano, F. A. *et al.* Hippocampal pathology in clinical high-risk patients and the onset of schizophrenia. *Biological Psychiatry* **87**, 234–242 (2020).

[19] Pavlopoulos, E. *et al.* Molecular mechanism for age-related memory loss: the histone-binding protein rbap48. *Science translational medicine* **5**, 200ra115–200ra115 (2013).

[20] Moreno, H., Hua, F., Brown, T. & Small, S. Longitudinal mapping of mouse cerebral blood volume with mri. *NMR in Biomedicine: An International Journal Devoted to the Development and Application of Magnetic Resonance In vivo* **19**, 535–543 (2006).

[21] Fischl, B. Freesurfer. *Neuroimage* **62**, 774–781 (2012).

[22] Small, S. A., Tsai, W. Y., DeLaPaz, R., Mayeux, R. & Stern, Y. Imaging hippocampal function across the human life span: is memory decline normal or not? *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society* **51**, 290–295 (2002).

[23] Small, S. A., Chawla, M. K., Buonocore, M., Rapp, P. R. & Barnes, C. A. Imaging correlates of brain function in monkeys and rats isolates a hippocampal subregion differentially vulnerable to aging. *Proceedings of the National Academy of Sciences* **101**, 7181–7186 (2004).

[24] Moreno, H. *et al.* Imaging the αβ-related neurotoxicity of alzheimer disease. *Archives of neurology* **64**, 1467–1477 (2007).

[25] Chawla, M. K. & Barnes, C. A. Hippocampal granule cells in normal aging: insights from electrophysiological and functional imaging experiments. *Progress in brain research* **163**, 661–821 (2007).

[26] Feng, X. *et al.* Brain regions vulnerable and resistant to aging without alzheimer’s disease. *PloS One* **15**, e0234255 (2020).

[27] Shen, X., Liu, H., Hu, Z., Hu, H. & Shi, P. The relationship between cerebral glucose metabolism and age: report of a large brain pet data set. *PloS one* **7**, e51517 (2012).

[28] Feng, X. *et al.* Estimating brain age based on a uniform healthy population with deep learning and structural mri. *Neurobiology of Aging* (2020).

[29] Hoffman, P. & Morcom, A. M. Age-related changes in the neural networks supporting semantic cognition: A meta-analysis of 47 functional neuroimaging studies. *Neuroscience & Biobehavioral Reviews* **84**, 134–150 (2018).
[30] Gómez-Isla, T. et al. Profound loss of layer ii entorhinal cortex neurons occurs in very mild alzheimer’s disease. *Journal of Neuroscience* **16**, 4491–4500 (1996).

[31] Talati, P. et al. Increased hippocampal ca1 cerebral blood volume in schizophrenia. *NeuroImage: Clinical* **5**, 359–364 (2014).

[32] Allen, P. et al. Emerging temporal lobe dysfunction in people at clinical high risk for psychosis. *Frontiers in psychiatry* **10**, 298 (2019).

[33] Mueller, S. G. et al. The alzheimer’s disease neuroimaging initiative. *Neuroimaging Clinics* **15**, 869–877 (2005).

[34] Sperling, R., Mormino, E. & Johnson, K. The evolution of preclinical alzheimer’s disease: implications for prevention trials. *Neuron* **84**, 608–622 (2014).

[35] Braak, H. & Braak, E. Evolution of the neuropathology of alzheimer’s disease. *Acta Neurologica Scandinavica* **94**, 3–12 (1996).

[36] Coughlan, G., Laczó, J., Hort, J., Minihane, A.-M. & Hornberger, M. Spatial navigation deficits—overlooked cognitive marker for preclinical alzheimer disease? *Nature Reviews Neurology* **14**, 496–506 (2018).

[37] Brickman, A. M., Stern, Y. & Small, S. A. Hippocampal subregions differentially associate with standardized memory tests. *Hippocampus* **21**, 923–928 (2011).

[38] Simoes, S. et al. Alzheimer’s vulnerable brain region relies on the vps26b retromer core. *Cell* (2019).

[39] Menze, B. H. et al. The multimodal brain tumor image segmentation benchmark (brats). *IEEE transactions on medical imaging* **34**, 1993–2024 (2014).

[40] Bakas, S. et al. Advancing the cancer genome atlas glioma mri collections with expert segmentation labels and radiomic features. *Scientific data* **4**, 170117 (2017).

[41] Bakas, S. et al. Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the brats challenge. *arXiv preprint arXiv:1811.02629* (2018).

[42] Bakas, S. et al. Segmentation labels and radiomic features for the pre-operative scans of the tcga-gbm collection. the cancer imaging archive (2017).

[43] Clark, K. et al. The cancer imaging archive (tcia): maintaining and operating a public information repository. *Journal of digital imaging* **26**, 1045–1057 (2013).

[44] Newitt, D. & Hylton, N. Single site breast dce-mri data and segmentations from patients undergoing neoadjuvant chemotherapy. *The Cancer Imaging Archive* **2** (2016).

[45] Seely, J. M. Management of breast magnetic resonance imaging-detected lesions. *Canadian Association of Radiologists Journal* **63**, 192–206 (2012).

[46] Monticciolo, D. L. Magnetic resonance imaging of the breast for cancer diagnosis and staging. In *Seminars in Ultrasound, CT and MRI*, vol. 32, 319–330 (Elsevier, 2011).

[47] Garcia, G. C., Bockel, S., Majer, M., Ammari, S. & Smits, M. Imaging of brain metastases: Diagnosis and monitoring. In *Central Nervous System Metastases*, 145–158 (Springer, 2020).

[48] Gonz, R. et al. Functional mri in the evaluation of dementia: correlation of abnormal dynamic cerebral blood volume measurements with changes in cerebral metabolism on positron emission tomography with fludeoxylucose f 18. *American journal of neuroradiology* **16**, 1763–1770 (1995).

[49] Aronen, H. J. et al. High microvascular blood volume is associated with high glucose uptake and tumor angiogenesis in human gliomas. *Clinical Cancer Research* **6**, 2189–2200 (2000).

[50] Østergaard, L. et al. Absolute cerebral blood flow and blood volume measured by magnetic resonance imaging bolus tracking: comparison with positron emission tomography values. *Journal of Cerebral Blood Flow & Metabolism* **18**, 425–432 (1998).
[51] Sugahara, T. et al. Correlation of mr imaging-determined cerebral blood volume maps with histologic and angiographic determination of vascularity of gliomas. *AJR. American journal of roentgenology* 171, 1479–1486 (1998).

[52] Lei, L. et al. Glioblastoma models reveal the connection between adult glial progenitors and the proneural phenotype. *PloS one* 6 (2011).

[53] Ledig, C. et al. Robust whole-brain segmentation: application to traumatic brain injury. *Medical image analysis* 21, 40–58 (2015).

[54] Chou, N., Wu, J., Bingren, J. B., Qiu, A. & Chuang, K.-H. Robust automatic rodent brain extraction using 3-d pulse-coupled neural networks (pcnn). *IEEE Transactions on Image Processing* 20, 2554–2564 (2011).

[55] Bezdek, J. C., Ehrlich, R. & Full, W. Fcm: The fuzzy c-means clustering algorithm. *Computers & Geosciences* 10, 191–203 (1984).

[56] Jenkinson, M. et al. Fsl. *Neuroimage* 62, 782–790 (2012).

[57] Cox, I. J., Roy, S. & Hingorani, S. L. Dynamic histogram warping of image pairs for constant image brightness. In *Proceedings., International Conference on Image Processing*, vol. 2, 366–369 (IEEE, 1995).

[58] Wagenknecht, G., Kaiser, H.-J., Sabri, O. & Buell, U. Dynamic programming algorithm for contrast correction in medical images. In *Nonlinear Image Processing XI*, vol. 3961, 216–226 (International Society for Optics and Photonics, 2000).

[59] Feng, X. et al. Temporal lobe epilepsy lateralization using retrospective cerebral blood volume mri. *NeuroImage: Clinical* 19, 911–917 (2018).

[60] 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).

[61] 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*, 770–778 (2016).

[62] Vaswani, A. et al. Attention is all you need. In *Advances in neural information processing systems*, 5998–6008 (2017).

[63] Oktay, O. et al. Attention u-net: Learning where to look for the pancreas. *arXiv preprint arXiv:1804.03999* (2018).

[64] Barron, J. T. A general and adaptive robust loss function. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 4331–4339 (2019).

[65] Pérez-García, F., Sparks, R. & Ourselin, S. Torchio: a python library for efficient loading, preprocessing, augmentation and patch-based sampling of medical images in deep learning (2020).

[66] Otsu, N. A threshold selection method from gray-level histograms. *IEEE transactions on systems, man, and cybernetics* 9, 62–66 (1979).

[67] Pedregosa, F. et al. Scikit-learn: Machine learning in python. *the Journal of machine Learning research* 12, 2825–2830 (2011).

[68] Avants, B. B., Tustison, N. & Song, G. Advanced normalization tools (ants). *Insight j* 2, 1–35 (2009).

[69] Ashburner, J. et al. Spm12 manual. *Wellcome Trust Centre for Neuroimaging, London, UK* 2464 (2014).

[70] Cox, R. W. Afni: software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical research* 29, 162–173 (1996).
[71] Forman, S. D. et al. Improved assessment of significant activation in functional magnetic resonance imaging (fmri): use of a cluster-size threshold. *Magnetic Resonance in medicine* 33, 636–647 (1995).

[72] Cox, R. W., Chen, G., Glen, D. R., Reynolds, R. C. & Taylor, P. A. Fmri clustering in afni: false-positive rates redux. *Brain connectivity* 7, 152–171 (2017).

[73] Fedorov, A. et al. 3d slicer as an image computing platform for the quantitative imaging network. *Magnetic resonance imaging* 30, 1323–1341 (2012).

**Ethical compliance**

All animal procedures and experiments were performed in accordance with national guidelines (National Institutes of Health) and approved by the Institutional Animal Care and Use Committee of Columbia University Medical Center.

All human data used in preparation of the article were obtained from either public databases or previous studies conducted at Columbia University. All subjects whose MRI scans were acquired at Columbia University provided informed consent. Approval from the Institutional Review Board of the New York State Psychiatric Institute was obtained prior to initiating those research.

**Acknowledgement**

This study was funded by the Seed Grant Program and Technical Development Grant Program at the Columbia MR Research Center. This study was also funded by grants from Matheson Foundation (UR010590) and the cancer center support grant (P30CA013696 NIH/NIC). This study was performed at the Zuckerman Mind Brain Behavior Institute MRI Platform, a shared resource.

Data collection and sharing for this project was partially funded by the Alzheimer’s Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer’s Association; Alzheimer’s Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer’s Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

**Author Contributions**

J.G. conceived, designed, and supervised all studies constituting this article. C.L. and N.Z. participated in the design and optimization of all DeepContrast models, and particularly optimized and trained the Healthy Human Brain Model including data preprocessing. D.S. led the optimization and training of the Tumor Human Brain Model and the Tumor Human Breast Model including data preprocessing. H.S. and X.L. led the optimization and training of the Healthy Mouse Brain Model and the Tumor Mouse Brain Model including data preprocessing. S.A.S. kindly provided all human MRI data in the Columbia University cohort, which was collected, organized and maintained by F.A.P. X.F. kindly provided the organized and preprocessed ADNI data, and the brain parcellations of the 177 healthy subjects used in the Human Aging study. J.G. led and S.G.-S. participated in the acquisition of the mouse MRI data. P.S.U., A.M. and P.D.C. provided the mouse glioma cells. H.J.W., P.S.U. and C.C.W. provided stereotactic injection of mice with glioblastoma. C.L. and N.Z. performed statistical analyses of the Healthy Human Brain Model. C.L. and N.Z. conducted the Human Aging study. C.L. conducted the
Human AD study. N.Z. conducted the Human CHR study. D.S. performed statistical analyses of the Tumor Human Brain Model and the Tumor Human Breast Model. H.S. performed statistical analyses of the Healthy Mouse Brain Model and the Tumor Mouse Brain Model. J.G. verified all statistical results. C.L., N.Z., S.A.S., J.G., and D.S. wrote the manuscript. N.Z. and C.L. created all display items (figures and tables) and supplementary information (figures and tables). All authors reviewed, commented and edited the manuscript. Data used in the Human AD study was obtained from the ADNI database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this manuscript.

Competing Interests

The authors declare the following competing interests. F.A.P. is a consultant for and equity holder of Imij Technologies. S.A.S. serves on the scientific advisory board of Meira GTX, recently came off the scientific advisory board of Denali Therapeutics, and is an equity holder in Imij Technologies. X.F., F.A.P., S.A.S. and J.G. have either granted patents or applications in neuroimaging for which no royalties are received.
Supplementary Table 1: Subject-level evaluations of the Healthy Mouse Brain Model and the Tumor Mouse Brain Model. As a supplement to Fig. 2 and Table 1, individual data points are provided to comply with the Nature Research data presentation policy since the sample sizes (i.e. number of scans in the test set) are smaller than 10 (WT: 6, GBM: 4).

| Model              | Evaluation       | Region         | Data                 | Subject | PSNR  | P.R  | S.R  | SSIM  |
|--------------------|------------------|----------------|----------------------|---------|-------|------|------|-------|
| Healthy Mouse Brain| Similarity (WT)  | Whole Brain    | Gd-Uptake & T2W      | 1       | 23.01 | 0.079| 0.133| 0.709 |
|                    |                  |                |                      | 2       | 21.47 | -0.143| 0.050| 0.694 |
|                    |                  |                |                      | 3       | 19.91 | -0.116| 0.055| 0.709 |
|                    |                  |                |                      | 4       | 23.64 | -0.068| 0.057| 0.733 |
|                    |                  |                |                      | 5       | 21.51 | -0.102| 0.056| 0.681 |
|                    |                  |                |                      | 6       | 23.40 | -0.083| 0.089| 0.718 |
|                    | Gd-Uptake & Gd-Predicted |          |                      | 1       | 25.60 | 0.666| 0.590| 0.847 |
|                    |                  |                |                      | 2       | 25.10 | 0.699| 0.586| 0.828 |
|                    |                  |                |                      | 3       | 23.12 | 0.680| 0.606| 0.814 |
|                    |                  |                |                      | 4       | 23.36 | 0.700| 0.614| 0.809 |
|                    |                  |                |                      | 5       | 23.59 | 0.698| 0.602| 0.824 |
|                    |                  |                |                      | 6       | 26.79 | 0.728| 0.641| 0.862 |
| Tumor Mouse Brain  | Similarity (GBM) | Whole Brain    | Gd-Uptake & T2W      | 1       | 18.84 | -0.040| -0.030| 0.624 |
|                    |                  |                |                      | 2       | 16.79 | 0.245| 0.069| 0.622 |
|                    |                  |                |                      | 3       | 18.37 | 0.199| 0.052| 0.615 |
|                    |                  |                |                      | 4       | 18.02 | -0.038| -0.024| 0.655 |
|                    | Gd-Uptake & Gd-Predicted |          |                      | 1       | 20.41 | 0.497| 0.351| 0.700 |
|                    |                  |                |                      | 2       | 20.28 | 0.775| 0.438| 0.713 |
|                    |                  |                |                      | 3       | 21.58 | 0.711| 0.460| 0.755 |
|                    |                  |                |                      | 4       | 22.00 | 0.699| 0.516| 0.782 |
|                    | Tumor            | Gd-Uptake & T2W |                      | 1       | 12.14 | 0.240| 0.131| 0.428 |
|                    |                  |                |                      | 2       | 6.53  | 0.369| 0.431| 0.467 |
|                    |                  |                |                      | 3       | 11.65 | 0.411| 0.529| 0.575 |
|                    |                  |                |                      | 4       | 10.28 | 0.186| 0.337| 0.486 |
|                    | Gd-Uptake & Gd-Predicted |          |                      | 1       | 14.82 | 0.592| 0.543| 0.566 |
|                    |                  |                |                      | 2       | 11.53 | 0.628| 0.634| 0.706 |
|                    |                  |                |                      | 3       | 13.64 | 0.560| 0.560| 0.654 |
|                    |                  |                |                      | 4       | 15.20 | 0.788| 0.778| 0.726 |
|                    | Non-Tumor        | Gd-Uptake & T2W |                      | 1       | 19.09 | -0.059| -0.036| 0.629 |
|                    |                  |                |                      | 2       | 21.92 | -0.062| -0.042| 0.647 |
|                    |                  |                |                      | 3       | 19.69 | -0.113| -0.040| 0.628 |
|                    |                  |                |                      | 4       | 19.00 | -0.157| -0.055| 0.661 |
|                    | Gd-Uptake & Gd-Predicted |          |                      | 1       | 20.59 | 0.471| 0.345| 0.701 |
|                    |                  |                |                      | 2       | 23.03 | 0.435| 0.351| 0.712 |
|                    |                  |                |                      | 3       | 23.59 | 0.471| 0.406| 0.754 |
|                    |                  |                |                      | 4       | 22.73 | 0.570| 0.497| 0.777 |
Supplementary Fig. 1: Correlation between the GBCA-uptake map and GBCA-predicted map over the entire brain. The Pearson (top) and Spearman (bottom) correlation coefficients are computed for each of the 126 anatomical ROIs defined by FreeSurfer between the array of ROI-mean Gd-Uptake values and the array of ROI-mean Gd-Predicted values extracted from the 177 subjects with successful FreeSurfer parcellation. The analysis demonstrates significant correlation between the DeepContrast-predicted contrast and the corresponding ground truth across a wide range of brain regions (121 ROIs with $p < 0.001$ for Pearson correlation coefficient and 123 ROIs with $p < 0.001$ for Spearman correlation coefficient).
Supplementary Fig. 2: DeepContrast maps age-related changes over the entire cortex.

a. Three-dimensional volume rendering of the age-related t-score maps over the 72 FreeSurfer cortical region-of-interests (ROIs) reveals that the age-related changes in the CBV-predicted maps (CBV-Predicted) is similar to those in the ground truth CBV maps (CBV) despite that the non-contrast T1W scans which serves as the input to the DeepContrast model does not share either the same or the opposite trends.

b. A scatter plot of the age-related t-score over the 72 ROIs demonstrates that the age-related changes in CBV-Predicted are consistent to those in CBV (P.R = 0.816, S.R = 0.786) while the T1W counterparts do not (P.R = -0.131, S.R = -0.122).

c. An analysis of the concordance to CBV t-scores by treating it as a 1000-class classification problem reveals that age-related changes in CBV-Predicted have significant predictive power on those in CBV (sensitivity = 0.76, specificity = 0.89, AUC = 0.91) while the T1W counterparts do not (sensitivity = 1.00, specificity = 0.04, AUC = 0.43).
Supplementary Fig. 3: DeepContrast maps changes in the left anterior CA1 of patients with clinical high risk (CHR) Schizophrenia. a. A three-dimensional rendering of the left CA1 overlaid on the left hippocampus, built from a group-wise T1-weighted MRI template. b. The slice-based two-sample $t$-test result for each coronal slice along the anterior-posterior axis of the left CA1 (isotropic template-space with voxel spacing = 0.68 mm). The 13rd to 17th slices (out of 53) in the left CA1 show significant CHR-related increase. c. The same $t$-test result in the right CA1. None of the 48 slices in the right CA1 show significance beyond $p = 0.05$. 

![Supplementary Fig. 3](image-url)
| Dataset                  | Full Project Name                                      | Disease/Category                        | Modalities | Description Site                              | Download Site                                      |
|-------------------------|--------------------------------------------------------|----------------------------------------|------------|-----------------------------------------------|---------------------------------------------------|
| A4                      | Anti-Amyloid Treatment in Asymptomatic Alzheimer's     | Alzheimer's disease                    | dMRI, fMRI | http://alzstudy.ucsf.edu/                     | https://nda.loni.usc.edu/login/jpp                  |
| ABCD                    | Adolescent Brain Cognitive Development                 | Brain development and child health     | dMRI, fMRI | https://abcdstudy.org                          | https://nda.loni.usc.edu/login/jpp                  |
| ABIDE                   | The Autism Brain Imaging Data Exchange                 | Autism spectrum disorder               | fMRI, sMRI | http://abcdstudy.ucsf.edu/                     | https://nda.loni.usc.edu/login/jpp                  |
| ABV1                    | Aging Brain: Vascularity, Ischemia, and Behavior       | Vascular risk factors and cerebrovascular disease | fMRI, sMRI | http://ncibus.ucsc.edu/                        | https://nda.loni.usc.edu/login/jpp                  |
| ADNI                    | Alzheimer's Disease Neuroimaging Initiative            | Alzheimer's disease, normal aging      | fMRI, sMRI | http://adni.loni.usc.edu/                      | https://nda.loni.usc.edu/login/jpp                  |
| AIBL                    | Australian Imaging, Biomarkers & Lifestyle             | Normal aging                           | sMRI       | https://albci.cairnsresearch/neuroimaging/     | https://nda.loni.usc.edu/login/jpp                  |
| ATLAS                   | Anatomical Tracings of Lesions After Stroke           | Stroke                                 | sMRI       | http://fcon.1000.projects.nitrc.fr/retrotales.html | http://fcon_1000.projects.nitrc.fr/retrotales.html |
| BIND                    | Bipolar Disorder Neuroimaging Database                 | Bipolar Disorder                       | sMRI       | https://sites.google.com/site/bpicaldbase      | https://sites.google.com/site/bpicaldbase          |
| BrnTIS                  | Brain Tumor Segmentation                              | Glioblastoma and low-grade glioma     | fMRI, sMRI | https://www.med.upenn.edu/clock/lbrane2020     | https://www.med.upenn.edu/clock/lbrane2020         |
| Cam-CAN                 | Cambridge Centre for Ageing and Neuroscience           | Healthy cognitive aging                | fMRI, dMRI | https://www.cam-canc.org                        | https://www.cam-canc.org                            |
| CoviR                   | Consortium for Reliability and Reproducibility         | Test-retest reliability and reproducibility for functional and structural connectomics | fMRI, dMRI | http://fcon.1000.projects.nitrc.fr/retro/ReCon  | http://fcon_1000.projects.nitrc.fr/retro/ReCon      |
| DLBS                    | Dallas life-span Brain Study                          | Preservation and decline of cognitive function at different stages | sMRI       | http://fcon.1000.projects.nitrc.fr/retro/dlbs.html | http://fcon_1000.projects.nitrc.fr/retro/dlbs.html |
| GSP                    | Brain Genomics Superfund Project                      | Relationship between brain and behavior | fMRI       | http://www.neuinfo.org                         | https://nda.loni.usc.edu/login/jpp                  |
| ICBM                    | International Consortium for Brain Mapping             | Relationship between brain structure and function | fMRI       | http://fcon.1000.projects.nitrc.fr/retro/ICBM   | http://fcon_1000.projects.nitrc.fr/retro/ICBM       |
| IXI                     | Information extraction from Images                    | Normal, healthy subjects               | fMRI, sMRI | https://brain-development.org/ixi-dataset     | https://brain-development.org/ixi-dataset          |
| MinD                    | Major Depressive Disorder Neuroimaging Database        | Major depressive disorder              | sMRI       | https://sites.google.com/site/depressiondatabase | https://sites.google.com/site/depressiondatabase    |
| NeuromCon               | NeuromCon                                             | Parkinson's disease                    | fMRI, sMRI | http://fcon.1000.projects.nitrc.fr/retro/parkinsons.html | http://fcon_1000.projects.nitrc.fr/retro/parkinsons.html |
| NIFD                   | Neuroimaging In Frontotemporal Dementia               | Frontotemporal lobar degeneration      | fMRI, sMRI | N/A                                            | https://nda.loni.usc.edu/login/jpp                  |
| OASIS                   | Open Access Series of Imaging Studies                 | Alzheimer's disease                    | fMRI, sMRI | https://www.oasis-brains.org                   | https://nda.loni.usc.edu/login/jpp                  |
| PAIN                   | Pain and Intervention Imaging Network                  | Persistent pain disorders              | fMRI, sMRI | https://www.painrepository.org/papaoisories   | https://www.painrepository.org/papaoisories        |
| PPMi                   | Parkinson's Progression Markers Initiative            | Parkinson's disease                    | fMRI, sMRI | https://www.ppmi-info.org                     | https://nda.loni.usc.edu/login/jpp                  |
| SALD                   | Southwest University Adult Life-span Dataset          | Normal aging                           | sMRI       | http://fcon.1000.projects.nitrc.fr/retro/sald.html | http://fcon_1000.projects.nitrc.fr/retro/sald.html |
| SchizoConnect          | Large-Scale Schizophrenia Neuroimaging Data Mediation & Federation | Schizophrenia and bipolar disorder | sMRI       | http://schizocnect.org                         | http://schizocnect.org                             |
| SLM                    | Southwest University Longitudinal Imaging Multimodal Brain Data Repository | Test-retest reliability of brain-behavior correlates | fMRI, sMRI | http://fcon.1000.projects.nitrc.fr/retro/southwestuniv_qb_ind_xiv.html | http://fcon_1000.projects.nitrc.fr/retro/southwestuniv_qb_ind_xiv.html |
| TAIWu                  | TAIWu                                                  | Parkinson's Disease                    | fMRI, sMRI | http://fcon.1000.projects.nitrc.fr/retro/parkinsons.html | http://fcon_1000.projects.nitrc.fr/retro/parkinsons.html |
| TCIA                    | The Cancer Imaging Archive                            | Cancer (30 types)                      | fMRI, sMRI | http://www.cancerimagingarchive.net            | http://public.cancerimagingarchive.net/ncibase-search |
Supplementary Fig. 4: Details of the healthy mouse brain model and the tumor mouse brain model (Network 1 in Fig. 7). These two models are both implemented with a 2D five-layered Residual Attention U-Net architecture. The encoding path (blue blocks in the left half of the architecture) condenses the image dimensions and enriches the feature dimension, shrinking the image size from 208×208 pixels to 13×13 pixels while extracting 1024 channels of features. The decoding path (red blocks in the right half of the architecture) expands these high-level features and returns back a single slice of the predicted image with 208×208 pixels.
Supplementary Fig. 5: Details of the healthy human brain model (Network 2 in Fig. 7). This model is implemented with a 2D six-layered Residual Attention U-Net architecture. The encoding path (blue blocks in the left half of the architecture) condenses the image dimensions and enriches the feature dimension, shrinking the image size from 352×352 pixels to 11×11 pixels while extracting 2048 channels of features. The decoding path (red blocks in the right half of the architecture) expands these high-level features and returns back a single slice of the predicted image with 352×352 pixels.
Supplementary Fig. 6: Details of the tumor human brain model (Network 3 in Fig. 7). This model is implemented with a 3D six-layered Residual Attention U-Net architecture. The encoding path (blue blocks in the left half of the architecture) condenses the image dimensions and enriches the feature dimension, shrinking the image size from $192 \times 192 \times 160$ pixels to $6 \times 6 \times 5$ pixels while extracting 256 channels of features. The decoding path (red blocks in the right half of the architecture) expands these high-level features and returns back a single slice of the predicted image with $192 \times 192 \times 160$ pixels.
Supplementary Fig. 7: Details of the tumor human breast model (Network 4 in Fig. 7). This model is implemented with a 2D five-layered Residual Attention U-Net architecture. The encoding path (blue blocks in the left half of the architecture) condenses the image dimensions and enriches the feature dimension, shrinking the image size from 256×256 pixels to 16×16 pixels while extracting 256 channels of features. The decoding path (red blocks in the right half of the architecture) expands these high-level features and returns back a single slice of the predicted image with 256×256 pixels.