OAKLAND UNIVERSITY WILLIAM BEAUMONT

Development of a Deep Neural Network for Synthesis of Non-Contrast Cranial T1-Weighted Magnetic Resonance Imaging

Agueda M. Taylor, B.S.[1], Evan Porter, B.A.[3], Thomas Guerrero, M.D., Ph.D.[2]

[1]Class of 2022 M.D. Candidate, Oakland University William Beaumont School of Medicine
[2]Department of Radiation Oncology, Beaumont Health System
[3]Beaumont Research Institute

Introduction

Although 122 out of 1000 people in the United States have MRI's taken each year, there are over 4 million people in the United States with contraindications that forgo the diagnostic benefits of this imaging modality [1,2]. Cranial CT's are often chosen in cases of brain injury because of the speed and ability to detect surgically important lesions [3], whereas MRI carries the advantages of no radiation exposure, and more sensitive brain imagining detection.

A growing gap exists between the advancing field of machine learning, a subset of artificial intelligence (AI) and current applications in medical imaging, particularly due to medical imaging modality limitations [4]. Advanced AI algorithms known as deep neural networks are promising in terms of bridging this gap [4]. Specifically, generative adversarial networks (GANs) are a popular model design for producing synthetic medical images across all modalities. The two major components that comprise GAN models are the generator and the discriminator, that act in an opposing manner to maintain an equilibrium that will train each in the best possible manner.

This project sought to develop a deep neural network that would perform synthetic Cranial T1 weighted magnetic resonance imaging (MRI) from non-contrast computed tomography (CT) imaging. Our hypothesis was that the synthetic cranial T1 weighted images would be readily accurate, and allow for automatic contour evaluation, when compared to original non-contrast MR images.

Aims and Objectives

Develop a deep neural network that will perform synthetic Cranial T1 Weighted MRI from non-contrast CT.

Aim I: Train a convolutional deep neural network with non-contrast Cranial CTs previously gathered from an SRS Gamma Knife case set to produce non-contrast Cranial T1 Weighted MRIs.

Aim II: Evaluate model quality with an internal dataset.

The results of this project will potentially provide growth of knowledge and be a clinically expandable addition to the use of Computed Tomography; through the subsequent production of synthetic MR images. Additionally, this project will further help bridge the gap that exists between imaging and the expanding field of deep learning in all fields of medicine that utilize CT and MRI imaging [5].

Methods

Acquire Data Set

Compile a cohort of 390 sets of non-contrast MR and CT images from Beaumont Hospital's Gamma Knife Center

Process Data using Machine Learning

Train Generative Adversarial Network , written with TensorFlow for production of synthetic non-contrast MRI images

Evaluate Accuracy of Synthetic MRI non Contrast images

For the evaluation of image similarity between synthetic Non-contrast Cranial T1-Weighted MRI images, and ground truth T1-Weighted MRI images, SSIM (structural similarity metric), MAE (mean average error), and NRMSE (normalized root mean square error) were calculated.

Figure 1: Methods for Development of a Deep Neural Network for Synthesis of Non-Contrast Cranial T1-Weighted Magnetic Resonance Imaging. The chronologic methodology for development of synthetic non-contrast MRI images utilizing retrospective non-contrast CT data.

Network and Training Details

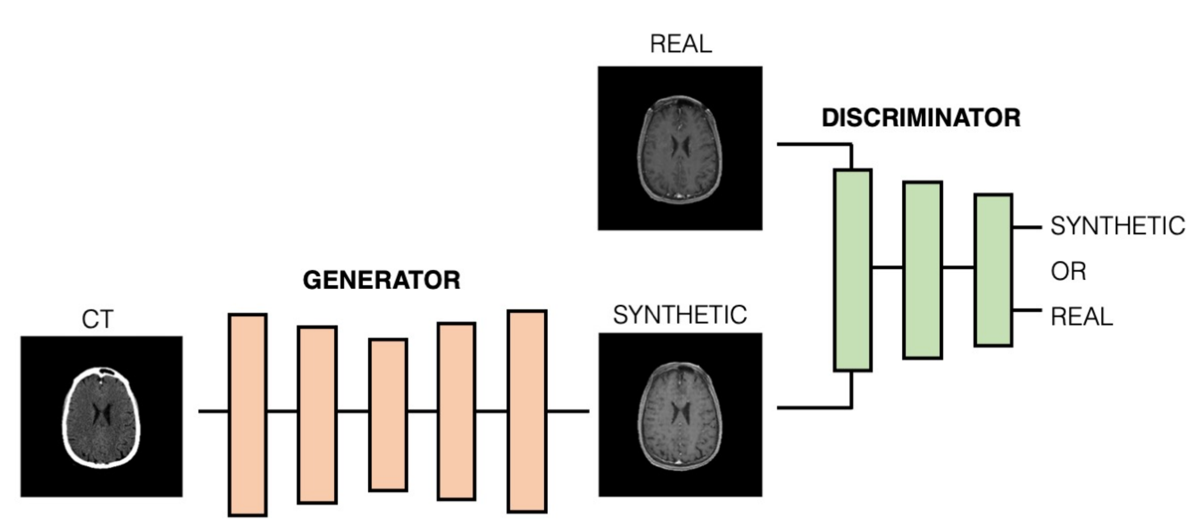

Figure 2: Generative Adversarial Network Diagram

A Generative Adversarial Network (GAN) was used to generate the synthetic MR images. A GAN is fundamentally comprised of two competing systems: a generator and a discriminator. The role of the generator is to become increasingly proficient at generating synthetic MR images from the provided CT image. Then, the synthetic and original MR image are provided to a second network, the discriminator, who's role is to determine if the provided image is real or synthetic. Through this competition, the generator becomes increasingly good at creating realistic synthetic images.
The generator portion of the GAN was based on a 2D-UNet with residual connections. The GAN was created using TensorFlow (version 2.3) and trained on an RTX 8000 GPU. Training was conducted for 20 epochs with an initial learning rate of 0.001.

Results

Table 1: Statistical Evaluation of MR images from GAN

STAT	MEAN	STD. DEV.
SSIM	0.881	0.0437
MAE	0.0259	0.0135
NRMSE	0.0069	0.0077

SSIM (Structural Similarity Metric) of 1 indicates signals are exactly structurally similar, whereas a value of 0 indicates no structural similarity.
MAE (Mean Average Error) is the average of absolute errors used to assess quality of machine learning model. Lower values indicate higher quality.
NRMSE (Normalized Root Mean Square Error) measures the error of a model in terms of predicative performance.

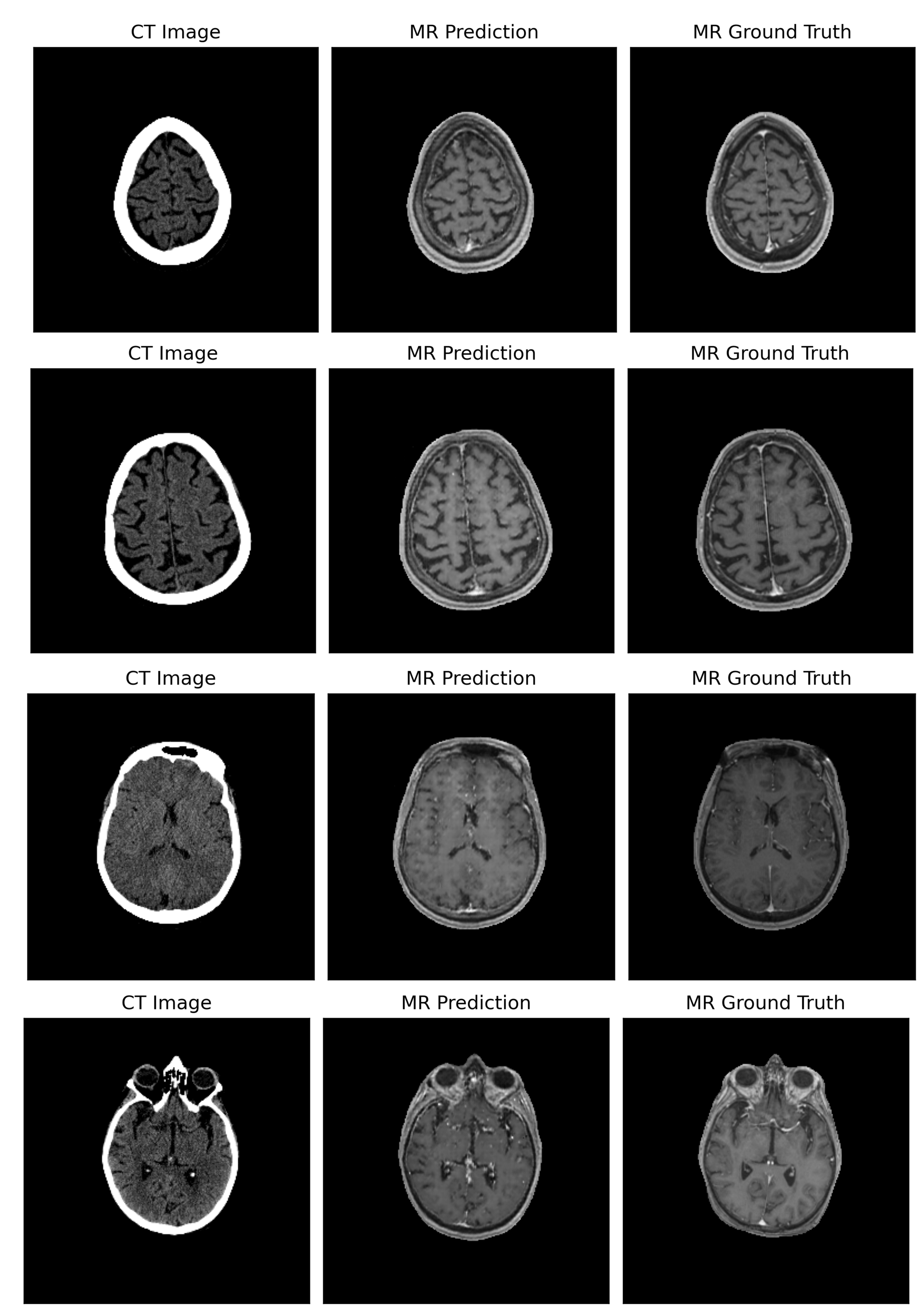

Figure 3: Comparison of ground truth CT & MR images and predictive synthetic MR images.

Conclusions

The result of our research was the production of a Generative Adversarial Network (GAN) model that was successfully able to produce synthetic non-contrast cranial T1-Weighted MR images with well-defined ventricles, bone and meninges. The mean calculated SSIM was 0.881, a value near 1, indicating a strong structural similarity. The MAE was 0.0259, with a smaller value is indicative of higher model quality. The NRMSE of 0.0069 was indicative of the error in the model's predictive performance.
Areas of improvement for the model, include white/grey matter differentiation, interpreting artifacts, and correctly predicting relative intensities. The difficulty with white/grey matter differentiation, resulted in overprediction of white matter. When utilizing ground truth CT images with substantial artifacts caused by the Gamma Knife stereotactic frame, the model had difficulty predicting an MR image that would appropriately be similar to the ground truth MR image. Additionally, the ground truth images were normalized as a volume, thus, the relative intensities of the ground truth vary compared to the prediction.
The GAN developed in this project may be modified to be more discriminative. Furthermore, other projects that may stem from this project include utilizing different imaging modalities, and datasets from other regions of the body. With the continuing advances and reliability of AI technology, the aim would be to enhance current imaging by having of method of translating one type of imaging modality into another for diagnostic benefit. The impact this preliminary research will have on the community, is further strengthening the knowledge of AI implementation in medical imaging and inspire other projects that will also implement AI technology in the medical field, as it is necessary explore what machine learning can bring to clinical care and research.

References

1. Health care use - Magnetic resonance imaging (MRI) exams - OECD Data. (n.d.). Retrieved January 5, 2020
2. First nationwide prevalence study of hip and knee arthroplasty shows 7.2 million Americans living with implants - Mayo Clinic. (n.d.). Retrieved December 31, 2019
3. Beauchamp, M. H., Ditchfield, M., Babl, F. E., Kean, M., Catroppa, C., Yeates, K. O., & Anderson, V. (2011). Detecting traumatic brain lesions in children: CT versus MRI versus susceptibility weighted imaging (SWI). Journal of Neurotrauma, 28(6), 915–927. https://doi.org/10.1089/neu.2010.1712
4. Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts HJWL. Artificial intelligence in radiology. Nat Rev Cancer. 2018;18(8):500-510. doi:10.1038/s41568-018-0016-5
5. Ueda D, Shimazaki A, Miki Y. Technical and clinical overview of deep learning in radiology. Jpn J Radiol. 2019;37(1):15-33. doi:10.1007/s11604-018-0795-3