International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-9 Issue-3, January 2020

Medical Image Synthesis using Computer Vision Techniques

Rahil Sarvaiya, Chirag Patel, Karthik Shetty, Somil Shah, Pankaj Sonawane

Abstract: Magnetic Resonance Imaging (MRI) is a type of scan that produces comprehensive images of the inside of the body using a steady magnetic field and radio waves. On the other hand, Computed Tomography (CT) scans, is a combination of a series of X-ray images, which are a type of radiation called ionizing radiation. It can be harmful to the DNA in your cells and also increase the chances that they'll turn cancerous. MRI is a safer option compared to CT and does not involve any radiation exposure. In this paper, we propose the use of Generative Adversarial Networks (GANs) to translate MRI images into equivalent CT images. We compare it with past techniques of MRI to CT scan conversion and elaborate on why GANs produce more realistic CT images while modeling the nonlinear relationship from MRI to CT.

Keywords*: *MRI, CT, GANs, CNNs, FCN, Segmentation based method, Atlas based methods.

I. INTRODUCTION

Computed Tomography (CT) produces cross-sectional images of the body using X-rays and a computer. On the contrary, Magnetic resonance imaging (MRI) produces meticulous images of the inside of the body using strong magnetic fields and radio waves. Comprehensive research has been carried out into whether the MRI scans could pose a risk to the human body due to the magnetic and radio waves. However, no indication has been found to suggest there's a risk, which means MRI scans are one of the safest medical procedures available. CT and MRI scans produce cross-sectional images of the body, allowing the doctor to look at it from the inside. MRI scans produce images using a magnetic field and radio waves, while CT scans produce images using X-rays. However, high doses of radiation are involved in CT scanning - chest CT scan is equivalent to 350 chest X-rays; CT abdomen to 400 chest X-rays and CT pulmonary angiography 750 chest X-rays.

Revised Manuscript Received on January 30, 2020.
***** Correspondence Author
Rahil Sarvaiya*, Student, Department of Computer Engineering, Dwarkadas J. Sanghvi College of Engineering, Mumbai, India.
Chirag Patel, Student, Department of Computer Engineering, Dwarkadas J. Sanghvi College of Engineering, Mumbai, India.
Karthik Shetty, Student, Department of Computer Engineering, Dwarkadas J. Sanghvi College of Engineering, Mumbai, India.
Somil Shah, Student, Department of Computer Engineering, Dwarkadas J. Sanghvi College of Engineering, Mumbai, India.
Pankaj Sonawane, Assistant Professor, Department of Computer Engineering, Dwarkadas J. Sanghvi College of Engineering, Mumbai, India.

Childhood cancer and leukemia in mothers who have imaging during pregnancy are a few other risks related to CT scans. Therefore, precautionary steps should be considered before the use of a CT scan.

To reduce futile imaging dose for patients, it is scientifically desired to estimate CT images from MR images in various applications. In this paper, we describe the existing methodologies in the domain of computer vision that have been implemented for converting MRI to CT scan.

Fully Convolutional Networks is a method of generating CT images from MR images using standard convolutional networks with some modifications in the architecture such as the replacement of fully connected layers with convolutional layers, avoiding the use of pooling layers, use of batch normalization and so on. Therefore, with the help of such changes, FCN based methods produce promising results along with better prediction and worthy of deploying with the help of various deep learning tools.

Another method, **Image Segmentation based methods** are used to segment MR images into different tissue categories using the Fuzzy clustering method. It focuses on the classification of voxels using probability to generate CT images from various MRI contrasts.

Atlas-based methods convert a dataset MRI image into the input MRI image and then use the attenuation map to create a corresponding CT image.

Use of Convolution Neural Network in **Generative Adversarial Networks (GANs)** gives them an additional superiority in terms of accuracy.

II. EXISTING METHODOLOGIES

A. Fully Convolutional Networks

Fully Convolutional Networks(FCN) are nothing but regular Convolutional Neural Networks(CNN) in which the fully connected layers are replaced by the convolutional layers. Therefore, Fully Convolutional Networks consist of only the Convolution layers and pooling layers and there are no Dense layers. As a result of this architectural difference, instead of predicting a single pixel value as in the case of CNN, FCN can predict an entire patch. In addition, as compared to standard CNN, FCN can not only maintain the neighborhood information but can also generate an entire patch by single forward propagation with great efficiency. In [15] by Yuan Zhou, they have used 3D FCN to convert MRI to CT images. In this paper, different loss functions were considered and the performance of FCN for the image synthesis task was evaluated based on these loss functions. Loss functions that were taken into the study are l1, l2, and DSSIM (Dissimilarity Structural Similarity Index) based on SSIM and contextual loss cl2.

The data used for prediction was gathered from fourteen patients undergoing prostate brachytherapy. The patients undergo a 3T MRI scan and each MRI scan included an axial 3D dual-echo gradient-echo sequence. In order to deal with geometrical distortions, different fields of view generated by different scanners and also resolving problems associated with air changing the location at different times, the images were non-rigidly registered to MRI volumes using Elastix, which is a toolbox/open source software for rigid and non-rigid registration of images. Thus, after data preprocessing, the MR and CT images for training are single-channel 3D data with relative voxel scales already approximately equal.

In this paper, they have also used batch normalization. In order to increase the stability of the network, batch normalization performs rescaling of data so that, the mean is zero and the variance of the data is 1. The data used for the synthesis of CT images from MR images contained only 14 samples of MR and CT images. The shape of each sample image was 128X128X40. As there were only 14 samples, overlapping patches were extracted from the whole volume so that sufficient data was available for training and for each MR patch corresponding, CT patch was generated by the FCN. As compared to other metrics such as Mean Absolute Error (MAE) and Peak Signal to Noise Ratio (PSNR), SSIM is better at understanding the structure difference to evaluate the images and at the same time presents more similarity to human evaluation and judgment. Contextual l2 loss (cl2) is obtained by modifying SSIM in order to make use of local mean information.

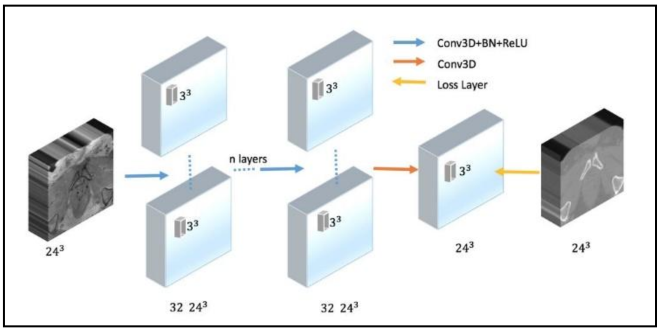

Fig. 1. FCN architecture used to generate CT images

The architecture used is 3 hidden-layer fully convolutional network with the convolutional kernel size set to 3x3x3 and zero paddings are set 1. ReLU activation function was used as the performance of FCN gradually improves with depth provided the ReLU activation function is used. Each layer also called a basic block consists of a 3D convolutional layer followed by Batch Normalization and ReLU except for the last layer which consists of only the layer. The number of filters was 32 and the Pooling layer was omitted since it is not suitable for image synthesis task. Taking SSIM into consideration, the loss for each patch can be given as structure dissimilarity(DSSIM) loss. Finally, the entire network was trained by optimizing DSSIM. Different learning rates were used with different loss functions. The optimal learning rate used for training was obtained by using coarse to fine search. The evaluated results that were based on four different loss functions showed that FCN8 gives the best results for each of the four-loss functions. Finally, the results concluded that as the depth of the network increases the performance of the network also improves as optimizing loss functions such as DSSIM requires a deeper network that is more layers to learn the complex mappings. In addition, a comparative analysis of the evaluation results of the four-loss functions on 7-fold cross-validation, both DSSIM and cl2 performs better than traditionally used l1 and l2 loss. As DSSIM provides good results, for future work this paper has proposed the use of another metric named multi-scale structure similarity(MS-SSIM) which is basically an extension of SSIM and includes a weighting scheme.

B. Segmentation based methods

Tissue Segmentation techniques primarily center around the segmentation of tissues into various tissue classes such as air, fat, soft tissue, bone, and other similar classes. Hsu et al.[16] Proposed a method that consists of 4 steps - (1) collecting multiple MRI volumes, (2) tissues are classified into different classes by using fuzzy c-means clustering, (3) assigning attenuation properties with weights according to the probability of different tissue classes in each voxel and (4) CT samples are obtained from the sum of attenuation properties in a given voxel. The data collected for generating synthetic CT images from MR images were regular MR images including T1 weighted images, T2 weighted images, two echoes from the ultra-short TE sequence. For training purposes, seven MR samples and one CT sample were procured from the patient. Furthermore, different image pre-processing techniques were used to transform all training samples into the required format for further processing and analysis.

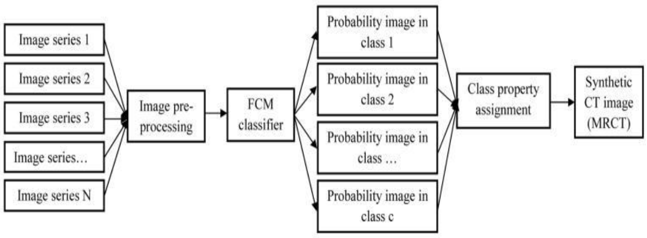

Fig. 2. Overall approach of using MR images to generate CT images

Region of Interests (ROI) was specified within the tissues to identify the tissue categories/materials namely: (1) bone, (2) fat, (3) fluid, (4) white matter, (5) grey matter and (6) air. Then, various contrast MRI volumes corresponding to the volumes of interest are put together in the same frame of reference to differentiate major types of tissues. Tissue classification based on probability is accomplished by using fuzzy C-means clustering (FCM) with a spatial constraint. In FCM clustering, image volumes are classified into various classes in order to obtain probability memberships. and class centroid vectors. Spatial constraints are used to the reduced effect of noise and improved spatial connectivity. Membership probabilities are then utilized for assigning properties of interest. The attenuation properties of a given voxel are then added to obtain the probability-weighted sum that provides an attenuation map, which can be used to generate corresponding CT volume.

Six MR samples were used for generating MRCT image volumes.

The following table demonstrates the capacity of each image type to separate two tissue classes for each patient.

Table 1. Images of 10 patients that have tissue contrasts are enough to separate material pairs

Material	T1-weighted	T2-weighted	Dixon_Fat	Dixon_Water
Differentiation from bone				
Air	0	0	0	0
Fat	10	10	10	4
Fluid	0	10	0	10
GM	10	10	0	10
WM	10	8	0	10
Differentiation from fat				
Bone	10	10	10	4
Air	10	10	10	10
Fluid	10	8	10	10
GM	10	10	10	10
WM	10	10	10	10
Differentiation from fluid				
Bone	0	10	0	10
Air	10	10	1	10
Fat	10	8	10	10
GM	10	10	2	9
WM	10	10	4	1

Atlas based Methods

Atlas-based methods deform the anatomical dataset to the patient's MRI and create an attenuation map. The attenuation map is then used to transform the corresponding dataset CT image to the synthesized CT image. Single atlas-based methods rely heavily on the registration accuracy of the image registration technique and suffer from estimation errors at the boundaries such as bones or air in the nasal cavity which have low intensity in MRI images but are distinctly visible in CT images. Performance can be improved by using Atlas fusion methods where instead of a single atlas image, multiple atlas images are used.

One method involves recognizing the atlas MRI image which is most similar to the target MRI image. Another method involves taking the average of multiple atlases in the dataset which is used as the final atlas. After this, the same single atlas-based deformation methodology is used.

A multi-atlas and multi-channel registration approach have also been proposed to tackle registration errors. A patch-based image synthesis algorithm is also used to refine the synthetic CT image quality at the boundaries where a single atlas-based method performs poorly.

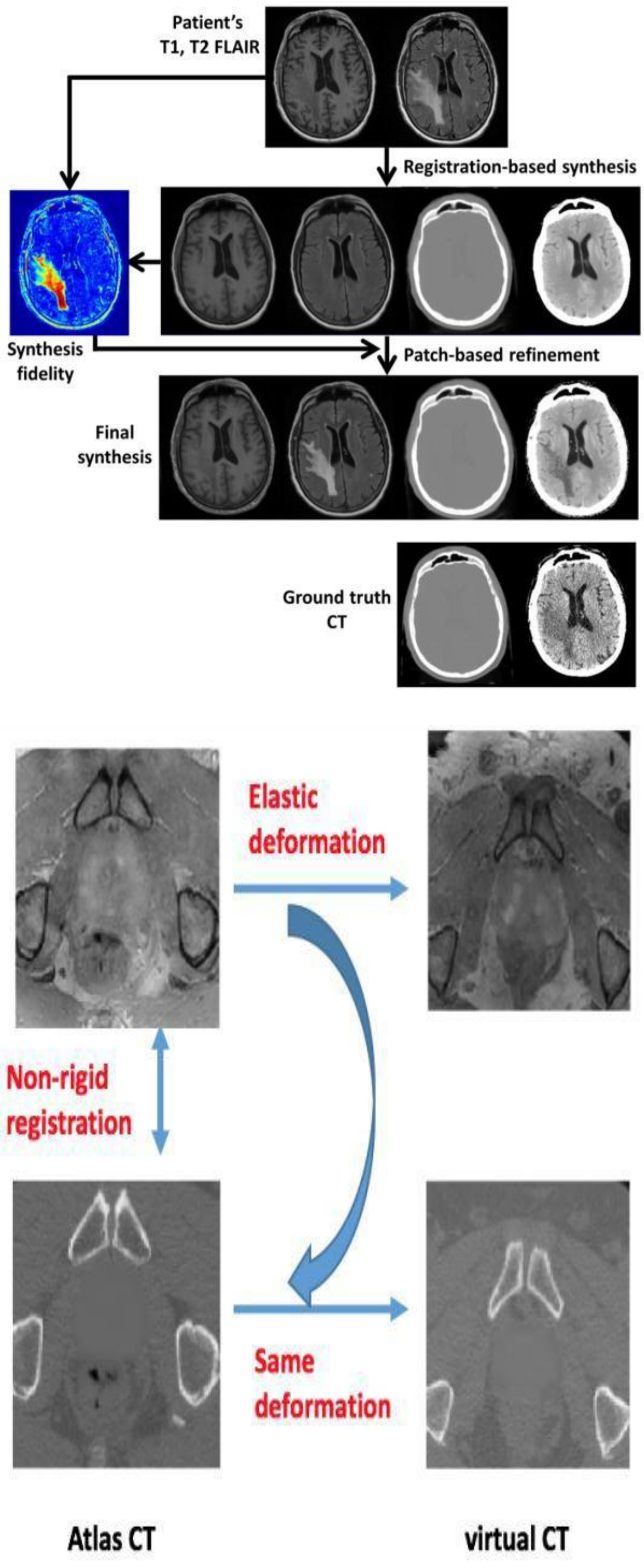

FIG. 3. ATLAS BASED METHODS – PATCH BASED AND FLOW

III. PROPOSED METHODOLOGY

Generative Adversarial Networks

GAN's have recently made great strides in various image synthesis applications. GAN's are essentially two neural networks, one generator and the other a discriminator, which are competing with each other. The discriminator tries to distinguish between real and fake images while the generator tries to synthesize images that can pass as real images when tested by the discriminator.

We synthesize realistic CT images from MRI images using GAN's. The success of the previous methods depended on the quality of feature extraction and their ability to understand the natural properties of the MR image. The non-linear mappings from CT to MRI are preserved with the help of GAN's which lead to realistic images. We use a fully convolutional neural network(FCN) as the generator so that it is better equipped to preserve the neighborhood information of the image.

GANs have an edge over other methodologies due to the efficiency and robustness of adversarial training. Adversarial training helps to generate sharper, "discrete" results, whereas, on the other hand, MSE results are mere blurry averages. Due to this powerful behavior, applications such as super-resolution GANs have come into picture to overcome the existing drawbacks.

Distance-based losses omit complex structures as they only look at pixel-wise deviations whereas GAN'S allow adversarial losses to be represented by many complex functions that attribute local structures.

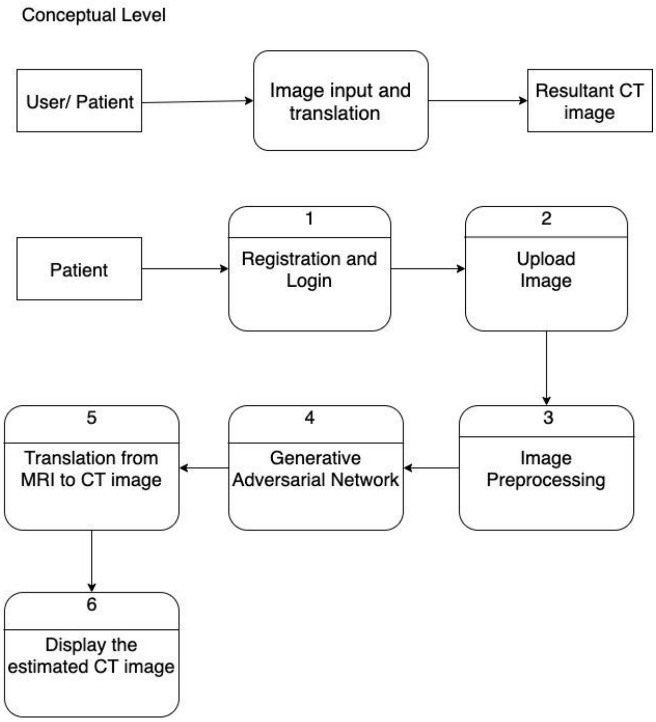

Fig. 4. Data Flow Diagram of Proposed Methodology

IV. RESULTS

Computed tomography (CT) is essential for various medical applications, such as PET attenuation correction and also radiotherapy treatment planning. However, CT exposes radiation during the procedure, which may cause detrimental reactions for patients.

FCNs	Segmentation based method	Atlas-based method	GANs
It provides good results when the network is deeper.	Accuracy decreases as MR images suffer from non-uniform intensity across the same tissue type.	Results degrade due to registration errors.	GANs have proven to be successful in modeling and generating high dimensional data and images.
Training time increases as the depth of the network increases.	Movement of body parts and longer scanning time makes it inefficient.	Single atlas based method performs poorly at the boundaries.	Training is easier as only backpropagation is needed to obtain gradients.
The solution to longer training time can be a pre-trained network.	Differentiates well between certain tissue classes but the system is not sturdy enough.	Other alternatives are required to resolve registration errors.	GANs are better at preserving non-linear features as compared to other techniques.

In Fully Convolutional Networks, the convolution is a significantly slower operation than, say maxpool, both forward and backward. The training step takes a long time if the network is pretty deep. The network is a little slow and complicated for just a good pre-trained model. This is the reason why the researches still use AlexNet and VGGNet for experiments. They are simple in terms of architecture, are well-known and provide good performance.
The Segmentation based method acquires multiple images of different type in order to classify tissue components. Therefore, the length of scanning, although relatively small may raise new issues related to the movement of the patient. In order to solve this problem, future methods can focus on making the scanning process faster. Operation in real environments is not possible as the system is not sturdy. Therefore, tools that synthesize CT images from multiple images will need to be handled by sufficient robustness in busy clinical environment. Although this method differentiates well between tissue and air, the researchers did not find a single MRI image volume that could sufficiently differentiate all tissue types. Single atlas-based methods rely heavily on the registration accuracy of the image registration technique and suffer from estimation errors at the boundaries such as bones or air in the nasal cavity which have low intensity in MRI images but are distinctly visible in CT images. The synthetic CT image quality is refined by a patch-based image synthesis algorithm at the boundaries where the single atlas-based method performs poorly. The tackling of registration errors is handled by a multi-atlas and multi-channel registration approach

V. CONCLUSION

As compared to ATLAS and segmentation based methods, the use of Fully Convolutional Neural Networks gives additional superiority in terms of accuracy.

We can infer that FCNs prove to be a prominent technique to model the nonlinear relationship between MRI and CT to produce more realistic images. In our method, we use an FCN as a generator in a Generative Adversarial Network so as to create realistic CT images from MR images that perform better than all the other methods mentioned above. This is made possible by the ability of GAN's to better preserve the non-linear features from MRI to CT.

REFERENCES

1. G. K. von Schulthess, F. P. Kuhn, P. Kaufmann, and P. Veit-Haibach, "Clinical positron emission tomography/magnetic resonance imaging applications," Seminars Nucl. Med., vol. 43, no. 1, pp. 3–10, 2013.
2. I. Bezrukov, F. Mantlik, H. Schmidt, B. Schölkopf, and B. J. Pichler, "MR-based PET attenuation correction for PET/MR imaging," Seminars Nucl. Med., vol. 43, no. 1, pp. 45–59, 2013.
3. A. Salomon, A. Goedicke, B. Schweizer, T. Aach, and V. Schulz, "Simultaneous reconstruction of activity and attenuation for PET/MR," IEEE Trans. Med. Imag., vol. 30, no. 3, pp. 804–813, Mar. 2011.
4. R. Boellaard, M. Hofman, O. Hoekstra, and A. Lammertsma, "Accurate PET/MR quantification using time of flight MLAA image reconstruction," Mol. Imag. Biol., pp. 1–9, 2014.
5. A. Rezaei, M. Defrise, G. Bal, C. Michel, M. Conti, C. Watson, and J. Nuyts, "Simultaneous reconstruction of activity and attenuation in time-of-flight PET," IEEE Trans. Med. Imag., vol. 31, no. 12, pp. 2224–2233, Dec. 2012
6. Berker Y, Franke J, Salomon A, Palmowski M, Donker HCW, Temur Y, Mottaghy FM, Kuhl C, Izquierdo-Garcia D, Fayad ZA, Kiessling F, Schulz V. MRI-based attenuation correction for hybrid PET/MRI systems: a 4-class tissue segmentation technique using a combined ultrashort-echo-time/Dixon MRI sequence. J Nucl Med. 2012;53:796–804. [PubMed] [Google Scholar]
7. Burgos N, Cardoso MJ, Thielemans K, Modat M, Pedemonte S, Dickson J, Barnes A, Ahmed R, Mahoney CJ, Schott JM, Duncan JS, Atkinson D, Arridge SR, Hutton BF, Ourselin S. Attenuation correction synthesis for hybrid PET-MR scanners: application to brain studies. IEEE Trans Med Imag. 2014;33(12):2332–2341. [PubMed] [Google Scholar]
8. Roy S, Carass A, Jog A, Prince JL, Lee J. MR to CT registration of brains using image synthesis. Proc of SPIE Medical Imaging. 2014;9034:903419. [PMC free article] [PubMed] [Google Scholar]
9. Yannick Berker et al. Mri-based attenuation correction for hybrid PET/MRI systems: a 4-class tissue segmentation technique using a combined ultrashort-echo-time/dixon MRI sequence. J. Nucl. Med., 53(5):796–804, 2012.
10. David J Brenner and Eric J Hall. Computed tomographyłan increasing source of radiation exposure. N. Engl. J. Med., 357(22):2277–2284, 2007.
11. Ciprian Catana et al. Toward implementing an MRI-based pet attenuation-correction method for neurologic studies on the Mr-pet brain prototype. J. Nucl. Med., 51(9):1431–1438, 2010.
12. Chao Dong et al. Learning a deep convolutional network for image super-resolution. In Computer Vision–ECCV 2014, pages 184–199. Springer, 2014.
13. Xavier Glorot and Yoshua Bengio. Understanding the difficulty of training deep feedforward neural networks. In International conference on artificial intelligence and statistics, pages 249–256, 2010.
14. Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. In Advances in Neural Information Processing Systems, pages 2672–2680, 2014.
15. Yuan Zhou. Generating virtual CT from MRI using fully convolutional neural networks with improved structure quality
16. Shu-Hui Hsu, Yue Cao, Ke Huang, Mary Feng, and James M. Balter. Investigation of a method for generating synthetic CT models from MRI scans of the head and neck for radiation therapy, Physics in medicine and biology 58, 8419, 2013.

AUTHORS PROFILE

Rahil Sarvaiya
He is currently a Computer Engineering student at Dwarkadas J. Sanghvi College of Engineering, Mumbai. He has interests in the fields of Machine Learning, Natural Language Processing and Computer Vision and has recently authored a paper on Neural Text Generation. He has gained experience through his recent internship as a Machine Learning Intern & will be pursuing his Masters in the field of Computer Science.

Chirag Patel
He is currently studying Computer Engineering from Dwarkadas J. Sanghvi College of Engineering. His field of interests includes Machine Learning and Deep Learning. He is planning to pursue his Masters where he looks forward to more research opportunities in the field of Computer Science.

Karthik Shetty
He is currently a Computer Engineering student at Dwarkadas J. Sanghvi College of Engineering, Mumbai. He has interests in the fields of Machine Learning, Deep Learning, and Computer Vision. He has gained experience through his R&D internship as an algorithm developer for autonomous cars and will be pursuing his Masters in the future.

Somil Shah
Being a Computer Science major, he is proficient in Algorithms and Software Development with over 3 years of experience in Applied Programming. He has won 2 hackathons and been a finalist for many. He has been working in the field of Machine Learning and Deep Learning for more than 1.5 years and has developed projects in the said domain.

Pankaj Sonawane
He is currently an Assistant Professor of the Computer Engineering Department at Dwarkadas J. Sanghvi College of Engineering. He has more than 10 years of teaching experience. He has expertise in the domain of Image Processing, Computer vision and Machine learning. He has guided various projects in the field of Machine Learning.