# DCGAN Based Brain Metastases Detection Using Limited Labeled Dataset

\*Note: Sub-titles are not captured in Xplore and should not be used

# 1st Given Name Surname

*dept. name of organization (of Aff.) name of organization (of Aff.)*

City, Country

email address or ORCID

# 4th Given Name Surname

*dept. name of organization (of Aff.) name of organization (of Aff.)*

City, Country

email address or ORCID

# 2nd Given Name Surname

*dept. name of organization (of Aff.) name of organization (of Aff.)*

City, Country

email address or ORCID

# 5th Given Name Surname

*dept. name of organization (of Aff.) name of organization (of Aff.)*

City, Country

email address or ORCID

# 3rd Given Name Surname

*dept. name of organization (of Aff.) name of organization (of Aff.)*

City, Country

email address or ORCID

# 6th Given Name Surname

*dept. name of organization (of Aff.) name of organization (of Aff.)*

City, Country

email address or ORCID

***Abstract*—** **Computer-Assisted Diagnosis (CAD) with high sensitivity helps in early diagnosis of Brain tumors. Due to the confidentiality of the medical data, it is not easily accessible. To overcome this problem, Data Augmentation (DA) is one such technique, which helps to generate synthetic data. This data when used along with the training data, helps to handle the small medical image dataset collected from various scanners. Generative Adversarial Networks (GANs) is one of the DA techniques. GAN trained on images can generate new images that contain many authentic characteristics and look realistic to human observers. Therefore, this paper focuses on overcoming the problem of limited labeled dataset, using Deep Convolutional GANs (DCGANs). After augmenting the synthetic images to the training data, there was an increase in the accuracy by a factor of 13.16. To analyze the closeness between the original and synthetic images, a visualization tool called ImageJ was used. In order to validate the CAD model, a visual turing test was conducted with the help of expert physicians.**

***Index Terms*—component, formatting, style, styling, insert**

1. **INTRODUCTION**

A mass or growth of abnormal cells in the brain leads to brain Tumor. The brain is one of the largest and most complex organs in the human body. Any unexpected growth may affect human function and may spread into other body organs and affect human functions. Detection of brain tumor is very complicated and difficult due to the size, shape, location and type of tumor in the brain, and hence early detection and classification of brain tumor helps in treatment methods. Diagnosis is usually done by medical examination, with Computer Tomography (CT) or Magnetic Resonance Imaging (MRI). MRI is one of the commonly used techniques due to its superior image quality and using no ionizing radiation during the scan. According to the article by The Hindu on June 12 2016, Brain Tumor Foundation of India says that brain tumor is the second most common cancer among children after leukemia. In India, every year 40,000-50,000 persons are diagnosed with brain tumors, out of which 20 percent are children. Until 2015, the figure was only somewhere around 5 percent. According to the official data, currently only six per cent of the children suffering from brain tumors are able to get the proper treatment. Accurate Computer-Assisted Diagnosis, associated with proper data wrangling, can alleviate the risk of overlooking the diagnosis in a clinical environment. However in many situations either data is limited or labeling is limited. Towards this, as a Data Augmentation (DA) technique, Generative Adversarial Networks (GANs) can synthesize additional training data to handle the small/fragmented medical imaging datasets collected from various scanners; those images are realistic but completely different from the original ones, filling the data lack in the real image distribution.

**Contributions**: Our main contributions are as follows:

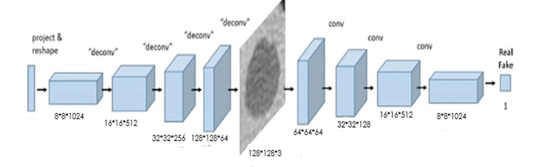
* We developed a model using Deep Convolutional Generative Adversarial Network (DCGAN) for the synthesis of brain image dataset.
* Evaluate and validate the model

1. **ALGORITHMS & TECHNIQUES**

***A. Generative Adversarial Network (GAN)***

Generative adversarial networks (GANs) are algorithmic architectures that use two neural networks, pitting one against the opposite (thus the “adversarial”) so as to get new, synthetic instances of data that can pass for real data. GAN comprises of two type of neural networks, the generator model and the discriminator model. One neural network, called the generator, generates new data instances, while the opposite the discriminator, evaluates them for authenticity; i.e. the discriminator decides whether each instance of knowledge that it reviews belongs to the particular training dataset or not.

The generator is creating new, synthetic images that it provides as input to the discriminator. It does so within the hopes that they, too, are going to be deemed authentic, albeit they're fake. The goal of the generator is to get passable hand-written digits: to lie without being caught. The goal of the discriminator is to spot images coming from the generator as fake.



*Figure 1: Generative Adversarial Network Flow Diagram*

***B. Convolutional Neural Network***

Computers see images using pixels. Pixels in images are related to some specific properties. For example, a particular group of pixels may signify a foothold in a picture or another pattern. Convolutions use this to help identify images. A convolution layer multiplies a matrix of pixels with a filter matrix and sums up the multiplication values. Then the convolution slides over to subsequent pixel and repeats an equivalent process until all the image pixels are covered.

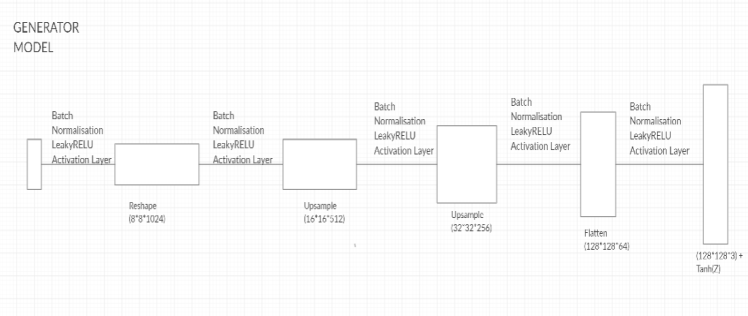
1. **MATERIALS AND METHODS**
2. ***Brain Metastases Dataset***

Dataset is retrieved from Kaggle data repository. Dataset contains 253 brain MRI images in which 98 tumor images 155 non-tumor images. For GAN training the entire dataset is used which has 253 images each of size 200\*200. For Tumor detection the whole dataset is divided into training, testing and validation set.

* Training Set: 261 Images
* Testing Set: 39 Images
* Validation Set: 38 Images

1. ***DCGAN based image generation***

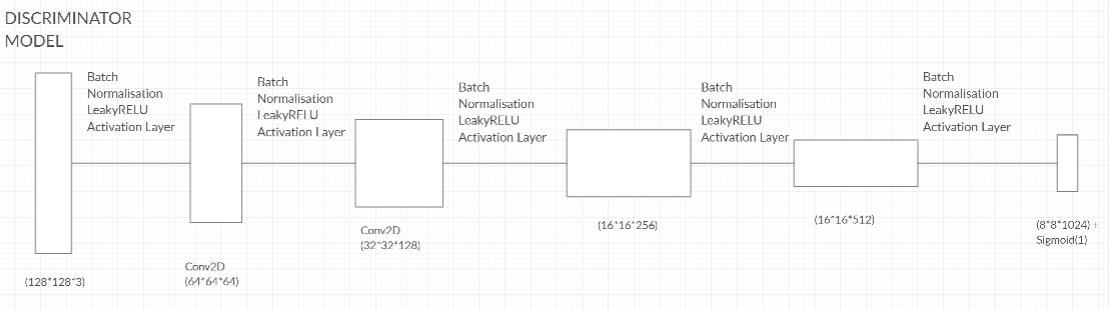
**The generator model** starts with the image size of 8\*8 with the 1024 number of filters, and then the batch normalization is done using the LeakyReLU activation function. The image is up scaled to16\*16 with 512 filters, and then batch normalization is applied using LeakyReLU activation function. This cycle continues for 2 more times the images is up scaled to32\*32 and then to 128\*128 and the batch normalization is applied using LeakyReLU activation function at last the hyperbolic tangent activation function is applied.



*Figure 2: Generator Model*

**The discriminator model** breaks down the images in reverse order as the generator model constructed it from 128\*128 to 8\*8 with batch normalization using LeakyReLU function with the same number of filters at each step.

**Activation functions** such as ReLU are used to address the vanishing gradient problem in deep convolutional neural networks and promote sparse activations (e.g. lots of zero values).The generator uses the hyperbolic tangent (tanh) activation function in the output layer and inputs to the generator and discriminator are scaled to the range [-1, 1].For the discriminator, the last convolution layer is flattened and then fed into a single sigmoid output.



*Figure 3: Discriminator Model*

1. ***Convolutional Neural Network based detection***

**Sequential model** in supervised learning can be used for many application that deals with detection.

The model uses “ReLu” activation function and pooling layer “MaxPooling2D” in each convolutional layer.

MaxPooling2D : It down samples the input representation by taking the maximum value over the window defined by pool size for each dimension along the features axis. The window is shifted by strides in each dimension.

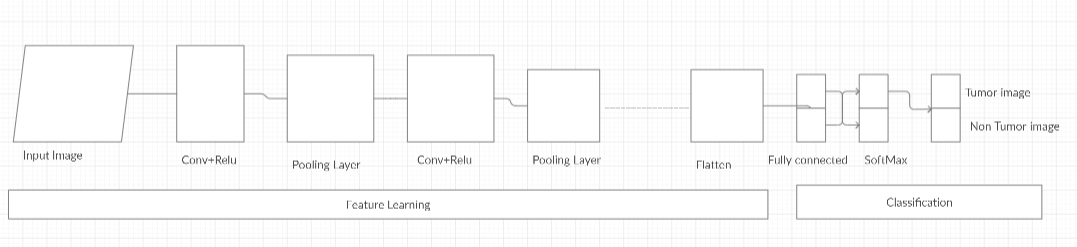
Equation 1 represents the Loss Function used in the final layer that is Binary Cross Entropy loss function.

|  |
| --- |
|  |

This type of loss function involves sigmoid activation and Cross-Entropy loss. It does not depend on each vector component, it elaborates that the loss computed for every CNN output vector component is not affected by other component values. This was the reason for it to be used for multi-label classification, were the insight of an element belonging to a certain class should not influence the decision for another class.

Where ti (target vector) and si (softmax function) are the ground truth and the CNN score for each class i in Class C.

**The optimizer** used in this sequential model is Rmsprop optimizer. These Optimizers are used to change the attributes of your neural network such as weights and learning rate in order to reduce the losses. Thisoptimizer utilizes the magnitude of recent gradients to normalize the gradients.



*Figure 4: CNN Sequential Model Architecture.*

Now the next step is computation of metrics such as specificity, sensitivity, accuracy, etc. which are essential in calculating the accuracy and making predictions.

1. ***Clinical Validation via Visual Turing Test***

To validate the model, we took help of two expert physicians. We combined 100 original and 87 synthetic images and gave them to the physician to differentiate between original and synthetic images. This experiment was done mainly because the validation via technical methods might provide good results, whereas the validation through experts who have immense experience in that field would provide much more scope in order to improve the model efficiency. This experiment was conducted in a timely manner and the outcome of the experiment was used to update the existing model and improve the efficiency.

1. ***Visualization via imageJ***

ImageJ is image processing tool used to compare profiles of original and synthetic images. It can create histograms and line profile plots. It supports image processing functions like contrast manipulation, convolution, Fourier analysis, sharpening, smoothing and edge detection.

Histogram. Calculates and displays a histogram of the distribution of grey values within the active image or selection. The x-axis represents the possible grey values and the y-axis shows the number of pixels found for each grey value

Plot Profile

This type of plot displays a two-dimensional graph of the intensities of pixels along a line within the image. The x-axis represents distance along the line and the y-axis is the pixel intensity.

1. **Results**

This section shows how DCGANs generate brain MR images.

The results include instances of synthetic images and their influence on tumor detection, along with DCGAN generated images evaluation via Visual Turing Test and plot profile comparison using ImageJ.

4.1 MR Images generated by DCGAN

Here the DCGAN has generated two sets of images, i.e The first set for batch size 64 and the second one is for batch size 4.

|  |
| --- |
|  |
|  |

***Fig 5: (i) The first row represents the GAN images generated with batch size 64 (ii) The second row represents GAN images generated with batch size 4.***

The set of images clearly maintains the realism of the original images with less odd artifacts, including the shapes and sizes of tumor inside the skull.

4.2 Brain Metastases Detection Results

The main problem in training the network described above is the lack of a large labelled training dataset. To enlarge the training data and improve the classification results in the brain tumor classification task, we augmented the data in two ways: 1) Classic augmentation that includes varieties of known image manipulations on given data examples; 2) Synthesis of new examples which are learned from the data examples using generative models.

In each case we have recorded Accuracy, Sensitivity, Specificity, Efficiency, F1-score-Testing data, F1-score-validate data and ROC and are presented in Table 3. The following table presents the record of observation when batch size is 64 and 4 respectively. Below figure represent the graph that is plotted with changes in accuracy with respect to change in number of epochs

*Figure 6: Accuracy of brain tumor classification with and without GAN synthetic images*

**4.3 ImageJ Profile Plot Results**

This experiment is carried out to assess the similarity between original images and GAN generated synthetic images. Parameter SSIM (structural similarity index matrix) is not employed as we don’t have one-to-one image mapping.We have used imageJ software tool for this experiment. Following figure presents the original and GAN generated synthetic images as well as Histogram and its Profile respectively.

|  |  |
| --- | --- |
| Image Type | Histogram |
| C:\Users\Sony\Desktop\img94.png  Original Brain MRI |  |
| C:\Users\Sony\Desktop\samples_1155_45.png  GAN Generated Synthetic(Batch Size=64) |  |

***Fig 7: Results of original and GAN generated***

***Synthetic images plotted Histogram.***

|  |  |
| --- | --- |
| Image Type | Plot Profile |
| C:\Users\Sony\Desktop\img94.png  Original Brain MRI |  |
| C:\Users\Sony\Desktop\samples_1155_45.png  GAN Generated Synthetic(Batch Size=64) |  |

***Fig 8: Results of original and GAN generated synthetic images profile plots.***

|  |  |  |  |
| --- | --- | --- | --- |
| **Original Brain MRI Image** | | **GAN generated synthetic(Batch Size=4)** | |
| **Mean** | **Standard Deviation** | **Mean** | **Standard Deviation** |
| 68.054 | 49.841 | 66.888 | 53.243 |
| 100.837 | 56.007 | 72.481 | 65.935 |
| 54.378 | 30.299 | 89.326 | 66.916 |
| 106.046 | 77.610 | 36.092 | 39.807 |
| 54.485 | 47.142 | 60.154 | 48.207 |

***Table 1: similarity study between histogram and image profile for images of batch size 4.***

|  |  |  |  |
| --- | --- | --- | --- |
| **Original Brain MRI Image** | | **GAN generated synthetic(Batch Size=64)** | |
| **Mean** | **Standard Deviation** | **Mean** | **Standard Deviation** |
| 80.727 | 51.178 | 68.054 | 49.841 |
| 44.633 | 42.996 | 100.837 | 56.007 |
| 54.605 | 47.269 | 54.378 | 30.299 |
| 39.614 | 45.698 | 106.046 | 77.610 |
| 105.88 | 77.673 | 54.485 | 47.142 |

***Table 2: similarity study between histogram and image profile for images of batch size 64.***

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | | | | | |
|  |  | **Batch size=64** | | **Batch size=4** | |
| **Parameters** | **Without GAN** | **With-GAN**  **(60 Images)** | **With-GAN**  **(85 Images)** | **With-GAN (60 Images)** | **With-GAN (85 Images)** |
| Accuracy | 81.57 | 94.73 | 84.21 | 81.57 | 84.21 |
| Sensitivity | 100 | 100 | 100 | 100 | 100 |
| Specificity | 53.33 | 86.67 | 60 | 53.33 | 60 |
| Efficiency | 78.30 | 93.80 | 81.40 | 78.30 | 81.40 |
| F1-score-Testing data | 86.79 | 95.83 | 88.46 | 86.79 | 88.46 |
| ROC | 0.767 | 0.933 | 0.791 | 0.80 | 0.80 |

***Table 3: Record of observation (i) CNN without GAN (ii) CNN with GAN***

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Test | **Accuracy**  (In %) | **TP**  **(Real as real)** | **FN**  **(Real as Synthetic)** | **FP**  **(Synthetic as Real)** | **TN**  **(Synthetic as Synthetic)** |
| Physician 1 | 96.25 | 98 | 2 | 5 | 82 |
| Physician 2 | 96.25 | 97 | 3 | 4 | 83 |

***Table 4: Results of Visual Turing tests two expert physicians***

4.3 Visual Turing Test Results

Table 4 shows the confusion matrix for the Visual Turing Test. The expert physicians easily recognize 256×256 synthetic images due to the lack of sharp edges and brain details. However, when GANs is trained and synthetic images are of size 200x200, the experts classify a considerable number of synthetic tumor images as real; it implies that the GAN generated synthetic images remarkably facilitate the realism of both healthy and pathological brain parts while they do not include abnormality.

Thus, GANs might perform as a tool to train medical students and radiology trainees when enough medical images are unavailable, such as abnormalities at rare positions/sizes. Such GAN applications are clinically prospective, considering the expert physicians’ positive comments about the tumor realism.

* Original images = 100
* GAN generated = 87
* Batch size: 64

1. **Model Evaluation**

Table 5 shows the comparison between the proposed model and existing models. The experiment was made to check for the change that we got in the results after applying the DA to the existing model. We researched for the literary papers that worked on DA on brain tumor dataset and got 3 papers who worked on it. We then compared the accuracy they achieved before applying DA and after applying DA. Then we found the increment factor by finding the difference between them. The maximum and minimum increment factor as observed from the table is 8.74 and 1.02 respectively. It can be inferred that the proposed model has better performance with increment factor of 13.16.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | | | | |
| **Paper** | **Method used** | **Results without DA (%)** | **Results with DA (%)** | **Increment Factor** |
| Inﬁnite Brain MR Images: PGGAN-based Data Augmentation for Tumor Detection,  Changhee Han | Progressive Growing of GANs (PGGANs) | 90.06 | 91.08 | 1.02 |
| Brain Tumor Classiﬁcation Using ResNet-101 Based Squeeze and Excitation Deep Neural Network,  Palash Ghosal | The transformations were Flip, Rotate, Elastic transform and Shear with variable degrees of transformations | 89.93 | 98.67 | 8.74 |
| Multi-Grade Brain Tumor Classiﬁcation using Deep CNN with Extensive Data Augmentation,  Muhammad Sajjad | Various augmentation techniques used: rotation, flipping, skewness, and shears for geometric transformations | 93.34 | 96.12 | 2.78 |
| **Proposed Model** | **Deep Convolutional GANs (DCGANs)** | **81.57** | **94.73** | **13.16** |

***Table 5 Comparison of proposed Model with Existing Model***

1. **Conclusion**

In recent days Computer Aided Diagnosis (CAD) is one of the blooming technologies in the field of medical research. However, there are complications that lead to the research getting less efficient or less accurate. One such complication is the unavailability of medical images due to the confidentiality of patient data in the hospitals and limited public data access. . We used the dataset available from Kaggle wherein out of the 255 images, 98 were tumoured and 157 were non-tumored. The batch size used for training the GAN should be less than or equal to 64,

we have trained it for two batch sizes of 4 and 64. With respect to both the batches we had generated 60 and 85 synthetic images, and it was found that 60 synthetic images with batch size 64 had better accuracy of 94.73%.%. We conducted 3 different experiments in order to assess and validate the model.

1. Performance evaluation of CNN with and without DCGAN generated images: The results showed accuracy of 81.57% when the model was trained on

original images. However, when the synthetic images were combined with the original images the accuracy was increased to 94.73% with an increment factor of 13.16.

1. Similarity study using histogram and image profile: To obtain and observe the histogram and image profiles we used the ImageJ tool kit facilitated us to study image similarity by providing different methods. The ImageJ tool calculates the mean and standard deviation of individual images which makes easier by comparing the numbers than by comparing the histograms and image profile graphs.
2. Visual Turing Test: in this experiment, the synthetic images are combined with real MR Images and give 2 Physicians to identify the synthetic and original images, the results were quite impressive the physicians marked a considerable number of synthetic images as real MR Images.

The results obtained from the above experiment were impressive. By experiment 1 we can say that the system gives far better results with DCGAN generated images; we got 13.16 % of the increase in the result. By experiment 2 we can say that the images look a lot more similar when we compare the histogram and image profiles. By experiment 3 we can interpret that the DCGAN generated images are a lot more similar not only in mean and standard deviation but also visually the DCGAN generated are well generated to be considered as real, these images can be used to train the medical exports in future. Through this systematic experimentation and analysis we conclude that DCGAN is found to be effective tool to address the issue of limited labelled dataset problem. As a future scope one shall experiment with other medical cases such as breast cancer classification or other cases.

1. **References**

[1] Frid-Adar, Maayan, et al. "Synthetic data augmentation using GAN for improved liver lesion classification." *2018 IEEE 15th international symposium on biomedical imaging (ISBI 2018)*. IEEE, 2018.

[2] Jiang, Jue, et al. "Tumor-aware, adversarial domain adaptation from ct to mri for lung cancer segmentation." *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, Cham, 2018.

[3] Onishi, Yuya, et al. "Automated pulmonary nodule classification in computed tomography images using a deep convolutional neural network trained by generative adversarial networks." *BioMed research international* 2019 (2019).

[4] Dey, Raunak, Zhongjie Lu, and Yi Hong. "Diagnostic classification of lung nodules using 3D neural networks." *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*. IEEE, 2018.

[5] Milletari, Fausto, Nassir Navab, and Seyed-Ahmad Ahmadi. "V-net: Fully convolutional neural networks for volumetric medical image segmentation." *2016 Fourth International Conference on 3D Vision (3DV)*. IEEE, 2016.

[6] Article by “The Hindu” regarding the statistics of Brain Tumor Cases in India: https://www.thehindu.com/sci-tech/health/Over-2500-Indian-kids-suffer-from-brain-tumour-every-year/article14418512.ece

[7] ImageJ- The tool used to visualize the closeness between the original and synthesized images: https://imagej.net/Welcome

[8] The dataset from Kaggle on which the project was worked on: https://www.kaggle.com/navoneel/brain-mri-images-for-brain-tumor-detection

[9] Information about GAN and DCGAN:https://medium.com/@jonathan\_hui/gan-dcgan-deep-convolutional-generative-adversarial- networks-df855c438f

[10] Sajjad, Muhammad, et al. "Multi-grade brain tumor classification using deep CNN with extensive data augmentation." Journal of computational science 30 (2019): 174-182.

[11] Han, Changhee, et al. "Infinite brain MR images: PGGAN-based data augmentation for tumor detection." Neural Approaches to Dynamics of Signal Exchanges. Springer, Singapore, 2020. 291-303.

[12] Ghosal, Palash, et al. "Brain Tumor Classification Using ResNet-101 Based Squeeze and Excitation Deep Neural Network." 2019 Second International Conference on Advanced Computational and Communication Paradigms (ICACCP). IEEE, 2019.

[13] Han, Changhee, et al. "Learning more with less: Conditional PGGAN-based data augmentation for brain metastases detection using highly-rough annotation on MR images." Proceedings of the 28th ACM International Conference on Information and Knowledge Management. 2019.