GENERATING MEDULLOBLASTOMA IMAGES USING W-GAN

[1]Parth Jariwala
School of Computer Science & Engineering (SCOPE)
Vellore Institute of Technology, Chennai
Chennai, India
parth10.jariwala@gmail.com

VSS Nishwan
School of Computer Science & Engineering (SCOPE)
Vellore Institute of Technology, Chennai
Chennai, India
vssnishwan@gmail.com

Author
Thomas Abraham
School of Computer Science & Engineering (SCOPE)
Vellore Institute of Technology, Chennai
Chennai, India
thomasabraham.jv@vit.ac.in

***Abstract*— Currently if a clinician is to understand how to diagnose and treat childhood brain tumors or any other disease there is usually need to refer to microscope images but as you well know these are few and far between since childhood brain tumors are thankfully rare. There is a new strategy of solving this problem with the help of artificial intelligence that we have designed. Our system employs deep learning in order to generate clean images from a database of brain tumour images in order to replicate the level of complexity required in the images needed by doctors. Using only 11 patients, we can develop a model that uses AI to generate images of different zoom levels which indicate tissue organization and higher powered images of the cell shapes. Of course, like with any new technology our system is not perfect – for example it often could not distinguish color or when the difference was very small– but it definitely helps to advance the medical training and research tools. By sharing both our successes and the areas that need improvement, we hope to advance the field of medical AI while keeping the focus on what matters most: On the issue of the improvement of Young patient care.**

Keywords - Medulloblastoma, WGAN-GP, Synthetic Images, Deep Learning, Artificial Intelligence, Microscopic Imaging, Data Augmentation, Classification, Histopathology, Neural Networks, Image Generation, Medical Imaging, Wasserstein Distance.

I.INTRODUCTION

Medulloblastoma is the most frequent adult malignant brain tumor and is responsible of 20% of all pediatric gliomas . Derived from the cerebellum and characterised by high-grade and rapid growth- this tumour is quickly capable of metastasising to nearly all other parts within the CNS, hence the need for the correct diagnosis of this condition at an early stage to increase the quality of the outcomes in patients . While medulloblastoma is a rare tumor, it has a strong effect on the quality of child's life in terms of neurological development and therefore early and precise diagnosis remains a clinical imperative.

Medulloblastomas are amongst the most biologically complex tumours and are divided into molecular and histopathological subtypes, whereby their subtype significantly predicts survival and therapeutic response[1].Medulloblastomas are classified into several molecular and histopathological subtypes, each of which has distinct prognostic and therapeutic implications.[1] Echinocandins have been classified into three generation based on structural differences, however, the most useful characteristics that were needed for identification at the species level are revealed at higher magnification such as 10x or 100x.

A major drawback of diagnostic histopathology for medulloblastoma is the lack of high quality microscopic image data [3]. Medulloblastoma is a rare pediatric cancer and, therefore, it does not have similar amount of images as other cancers, although it has enough data to train machine learning classification models. and can rapidly spread to other parts of the central nervous system, making early and accurate diagnosis critical for effective treatment and improved patient outcomes.

One of the main challenges in advancing diagnostic tools for medulloblastoma is the limited availability of high-quality microscopic image datasets. This scarcity, coupled with the expertise required to annotate such data, restricts the training

of AI models, often leading to overfitting and reduced generalization to new cases [3][4]. The publicly available Childhood Medulloblastoma Microscopic Images dataset, while valuable, is too small and lacks diversity in sample representation, further limiting its usefulness [8].

Synthetic data generation offers a promising solution. By using GANs, it is possible to create realistic microscopic images at 10x and 100x magnifications. These synthetic images not only expand existing datasets but also enhance model robustness by incorporating rare tumor patterns and subtype variations [9]. Such advancements could significantly improve the diagnostic capabilities of AI models, making them more reliable in clinical settings. This study aims to address the data scarcity issue through the creation of synthetic images, evaluating their impact on classification accuracy and exploring their clinical potential [10].

II. RELATED WORK

I. Overview of Medulloblastoma and Medical Imaging

Medulloblastoma is the most common malignant primary brain tumor in children, and early diagnosis is crucial due to its aggressive nature. Histopathological analysis identifies tumor subtypes—WNT, SHH, Group 3, and Group 4—each with different clinical implications [1]. Tumor differentiation is done through microscopic imaging, using magnifications like 10x and 100x to assess features such as cell morphology, mitotic figures, and nuclear arrangement. These differences are essential for subtype identification and treatment decisions [2].

II. Advances in AI for Medical Imaging

AI models, especially CNNs, have improved cancer detection and segmentation in medical imaging. These models help train AI systems to perform as expert pathologists [6]. Due to limited datasets for rare diseases, GANs and VAEs are increasingly used to generate synthetic images, augmenting small datasets and addressing data diversity issues [7].

III. Current State of Medulloblastoma Image Datasets

Currently, there is only one open-access medulloblastoma dataset on IEEE Dataport [8]. Limited data makes it challenging to capture diverse subtypes and cellular arrangements, affecting model performance. AI models risk overfitting and poor generalization without sufficient data [3].

IV. Challenges in Classification Model Development Due to Data Limitations

Limited data causes overfitting, where models memorize details instead of generalizing to unseen data. This is problematic for medulloblastoma classification, as subtle morphological differences are critical for accurate diagnosis [4]. The diversity of disease subtypes and their underrepresentation in datasets further contribute to model bias [5].

V. Synthetic Data as a Solution to Data Scarcity in Medical Imaging

GANs help generate synthetic images, addressing data scarcity and improving model stability, especially for rare diseases [9]. Synthetic data must meet high standards, such as expert validation and FID scores, to ensure clinical applicability [10].

VI. Introduction to GANs and Their Role in Medical Imaging

GANs, introduced by Goodfellow et al. [13], have revolutionized synthetic data generation, especially in medical imaging where data is scarce. GANs consist of a generator that creates images and a discriminator that evaluates them, producing high-quality synthetic data for training [9].

VII. Applications of GANs in Medical Imaging

GANs augment medical imaging datasets, creating more diverse samples and improving model accuracy. This is particularly useful for rare diseases and cancers with limited data, as shown by Shin et al.'s use of GANs to augment MRI and CT scans.

VIII. GANs for Modality Translation in Medical Imaging

CycleGANs are effective for translating images between different modalities, helping when data is lacking in one modality. For example, Jiang et al. adapted lung cancer models from CT to MRI, improving diagnostic workflows across modalities [15][16].

IX. Image Enhancement Using GANs

GANs improve image quality through tasks like denoising and super-resolution. SRGANs generate high-resolution images from low-resolution inputs, crucial for fine-detail imaging, such as in retinal imaging [17]. You et

al. used GANs to enhance CT resolution, improving diagnostic clarity [18].

X. Challenges and Limitations of GANs in Medical Imaging

Despite their potential, GANs face challenges like mode collapse and instability, limiting their direct use in medical imaging [10]. Clinical validation is necessary to ensure GAN-generated images are reliable, and ethical considerations must guide their clinical use [19].

III. METHODOLOGY

In this study, we employ a Wasserstein GAN with Gradient Penalty (WGAN-GP) to generate synthetic microscopic images of childhood medulloblastoma at different magnifications. The WGAN-GP model is designed to overcome the training instabilities commonly encountered in GANs by minimizing the Wasserstein distance between real and generated data distributions.

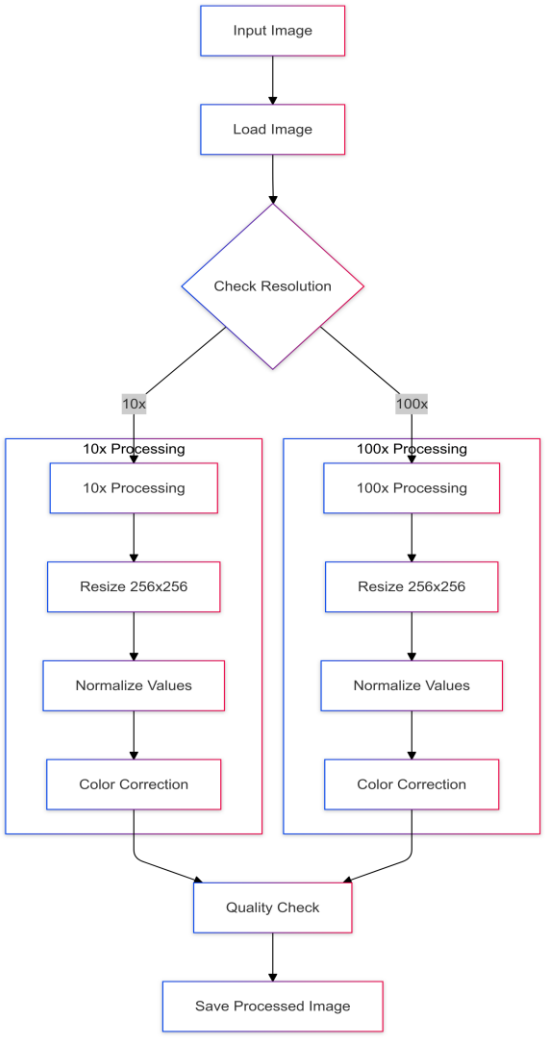

Fig 1 – Proposed System

A. Dataset:

Childhood medulloblastoma is one of the most common malignant brain tumors in children, necessitating highly accurate identification and subtype classification for effective prognosis and treatment. Microscopic image analysis in histopathological diagnostics plays a vital role in providing cellular-level insights, enabling precise disease detection and differentiation between subtypes.

This dataset comprises biopsy samples captured at two magnifications:

- **10x Magnification:** 202 images.
- **100x Magnification:** 153 images.

The images, saved in JPEG format with a 24-bit RGB color scheme, ensure high-quality color representation, making them suitable for accurate machine learning model training, including GANs for synthetic image generation. The dataset is organized into subfolders by magnification level, with separate directories for normal brain tissue and medulloblastoma samples, categorized by WHO-defined subtypes. This diversity ensures the GAN captures essential features of both healthy and tumor-affected brain tissue.

B. Image Processing :

Preprocessing is important for all these microscopic images of childhood medulloblastoma. First, all images are read from the designated directory. Some have been changed from their native BGR format to RGB for color consistency. All images are resized to
a standardized dimension of 256x256 pixels, which would balance computational efficiency with sufficient resolution to capture those histopathological details.

The pixel values were normalized to the range [-1, 1] that scales the original values from the range [0, 255]. It served the stabilization of the process during GAN training, by coincidence with activation functions of the model. Some error handling was present to skip unreadable or in-compatible files during processing, thus ensuring nondisturbed processing of big datasets.

This one will turn all images into a numpy array with a shape of (−1, 256, 256, 3) for preparation into feeding the GAN model.

C. Discriminator:

The discriminator is a CNN which tends to classify the input images as either real images or synthetic images. It includes four convolutional layers where the filters have grown step by step: 64, 128, 256, 512. The kernel size was set to 4 × 4 and the stride used was 2. After every convolutional layer, a Leaky ReLU function with an alpha of 0. 2 is applied as activation. To prevent overfitting, a dropout layer with a rate of 0.3 has been added. After the last convolutional layer, the network is flattened and outputs a single scalar value, that is further used in calculation of the Wasserstein loss.

The architecture of the discriminator is described as:

$$\text{Discriminator}(x) = \text{Flatten}\left(\text{LeakyReLU}\left(\text{Conv}(x, 64, 4, 2)\right)\right)$$

Fig 2: Architecture of Discriminator

Where:

• x is the input image,

• Conv(x, f, k, s) represents a convolutional layer with 'f' filters,
kernel size 'k', and stride 's',

• LeakyReLU(α) applies the Leaky ReLU activation with slope 'α',

• Flatten flattens the output for the final classification.

D. Generator Architecture:

The Generator Network generates microscopic images that are realistic from a random noise vector. The input to the generator is a 128-dimensional noise vector that is forward passed through a fully connected layer reshaping it into a 16 × 16 × 512 tensor. This tensor is then processed using a sequence of transposed convolutions layers which iteratively upscale spatial dimensions up to the desired image size of 256 × 256 pixels. Each of the transposed convolution layers uses batch normalization with momentum 0.8, along with Leaky ReLU activations to train the network more stably. The output layer makes use of the tanh activation function, scaling the pixel values to lie in the range [−1, 1].

The architecture of the generator is described as:

$$\text{Generator}(z) = \text{Tanh}\left(\text{ConvTranspose}(\text{BN}(\text{LeakyReLU}(z)))\right)$$

Where:

- z is the input noise vector,
- ConvTranspose applies a transposed convolutional operation,
- BN applies batch normalization,
- LeakyReLU(α) is the Leaky ReLU activation.

E. Loss Function and Gradient Penalty:

The WGAN-GP framework employed the Wasserstein loss that was defined as the difference between discriminator output corresponding to images in the real world and generated images.
This meant that the discriminator would be optimized to maximize this difference and, on the other hand, the generator would be optimized to minimize it. Furthermore, the Lipschitz constraint was enforced by incorporating a gradient penalty term in the loss of the discriminator.

The total loss function for the discriminator is then as follows:

$$\mathcal{L}_D = \mathbb{E}_{x \sim P_{\text{data}}}[D(x)] - \mathbb{E}_{z \sim P_z}[D(G(z))] + \lambda \mathbb{E}_{\hat{x} \sim P_{\hat{x}}}\left[(\|\nabla_{\hat{x}} D(\hat{x})\|_2 - 1)^2\right]$$

Where:

- $D(x)$ is the output of the discriminator for real data x,
- $G(z)$ is the generator's output for noise z,
- $\hat{x}$ is a random interpolation between real and generated images,
- λ is the gradient penalty weight (typically set to 10),
- The gradient penalty term enforces that the gradient norm is close to 1.

The loss function for the generator is:

$$\mathcal{L}_G = -\mathbb{E}_{z \sim P_z}[D(G(z))]$$

Where $D(G(z))$ is the discriminator's output for generated images. The generator aims to minimize this loss to improve the quality of generated images.

Fig 3: Loss function for the discriminator

F. Training Strategy:

The WGAN-GP model is then trained with Adam optimizer with an exponential decay learning rate schedule from 0. 0002. For each generator update the model sees lots of discriminator updates, which could be improving discriminator performance and delivering a more stable signal to learn from. For the extent of training sampling random noise vectors into images at frequent intervals makes those available to use in visualizing the generated images. The training process consists of the following steps:

• Load and preprocess the 10x and 100x magnification images.

• Normalize the images to the range [−1, 1].

• Feed the images in batches of 128 samples to the network.

• Train the discriminator on real and generated images.

• Update the generator based on the discriminator's feedback.

The model successfully captures the structural differences in the dataset by gradually learning to generate realistic representations of medulloblastoma cells at various magnifications over several epochs.

Algorithm 1 Wasserstein GAN with Gradient Penalty (WGAN-GP) Training

Require: Generator G, Discriminator D, Training data X, Number of epochs N_{epochs}, Steps per epoch N_{steps}, Discriminator steps n_d, GP weight λ_{gp}
Ensure: Trained Generator and Discriminator

function TRAINDISCRIMINATOR($D, G, X_{real}, n_d, \lambda_{gp}$)
  for $i = 1$ to n_d **do**
    $z \sim \mathcal{N}(0, I)$ ▷ Sample noise
    $X_{fake} \leftarrow G(z)$ ▷ Generate fake images
    Compute $D(X_{real})$ and $D(X_{fake})$
    Compute gradient penalty GP:
      $\epsilon \sim U[0, 1]$
      $\hat{x} \leftarrow \epsilon X_{real} + (1 - \epsilon) X_{fake}$
      $GP \leftarrow (\|\nabla_{\hat{x}} D(\hat{x})\|_2 - 1)^2$
    $\mathcal{L}_D \leftarrow \mathbb{E}[D(X_{fake})] - \mathbb{E}[D(X_{real})] + \lambda_{gp} GP$
    Update D using $\nabla_D \mathcal{L}_D$
  end forreturn $\mathcal{L}_D$
end function
function TRAINGENERATOR(G, D)
  $z \sim \mathcal{N}(0, I)$ ▷ Sample noise
  $X_{fake} \leftarrow G(z)$
  $\mathcal{L}_G \leftarrow -\mathbb{E}[D(X_{fake})]$
  Update G using $\nabla_G \mathcal{L}_G$ **return** $\mathcal{L}_G$
end function
for epoch $= 1$ to N_{epochs} **do**
  for step $= 1$ to N_{steps} **do**
    Sample minibatch X_{real} from training data
    $\mathcal{L}_D \leftarrow$ TRAINDISCRIMINATOR($D, G, X_{real}, n_d, \lambda_{gp}$)
    $\mathcal{L}_G \leftarrow$ TRAINGENERATOR(G, D)
    if step % visualize_interval $= 0$ **then**
      Generate and display sample images
    end if
  end for
end for

Fig 4: Algorithm for W-GAN

IV. RESULTS AND DISCUSSION

A. 10x Maginiftication Images:

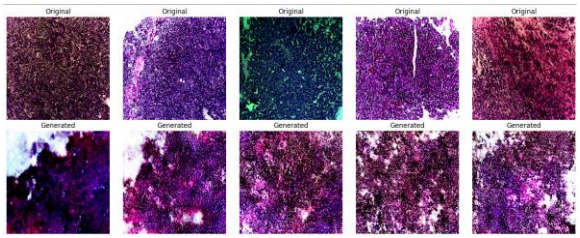

Fig. 5: Generated 10x Magnification Image

• Dense cellular patterns are evident, with a predominant purple pink staining, likely corresponding to H&E (Hematoxylin and Eosin) staining, which is commonly used for visualizing tissue sections.
• The generated images exhibit a good preservation of the tissue architecture but show some variations in contrast and texture when compared to the ground truth images.
• Some artifacts and whitespace are present in the generated images, which were not seen in the original samples.

These variations in contrast and artifacts suggest areas where the model can be refined to improve the authenticity and realism of the generated images.

C. 100x Magnicfication Images:

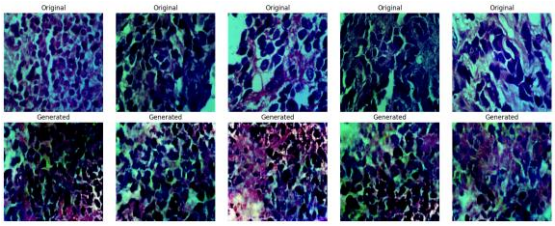

Fig. 6: Generated 100x Magnification Image

Individual cell morphology is clearly visible, with the characteristic "small round blue cell" appearance, typical of medulloblastoma.

• The nuclear-cytoplasmic ratio appears high, consistent with this tumor type.
• The generated images successfully preserve the cellular patterns, though there are slight variations in color intensity, which may affect diagnostic accuracy.

While the generated 100x magnification images reflect the general tissue structure and cellular characteristics well, some inconsistencies in color intensity can be noted, requiring further model adjustments.

We employed two primary metrics—Inception Score (IS) and Fr´echet Inception Distance (FID)—to assess the model's performance. These metrics offer a numerical assessment of image diversity and quality, which are essential for evaluating the produced photos.

• 10x Magnification Images:
  – Inception Score: 1.003 ± 0.001
  – Fr´echet Inception Distance (FID): 4.464
• 100x Magnification Images:
  – Inception Score: 1.005 ± 0.001
  – Fr´echet Inception Distance (FID): 2.383

For reference, an **ideal IS** for well-structured image datasets ranges from 5 to 10, highlighting the current limitations of the model. Conversely, lower FID values (<10) are indicative of closer alignment with real-world data, with **ideal FID scores** being close to 1 for realistic and diverse image generation. The FID scores, particularly for 100x magnification, demonstrate that the generated images are of higher

quality and more closely resemble real data at finer resolutions, capturing crucial morphological details.

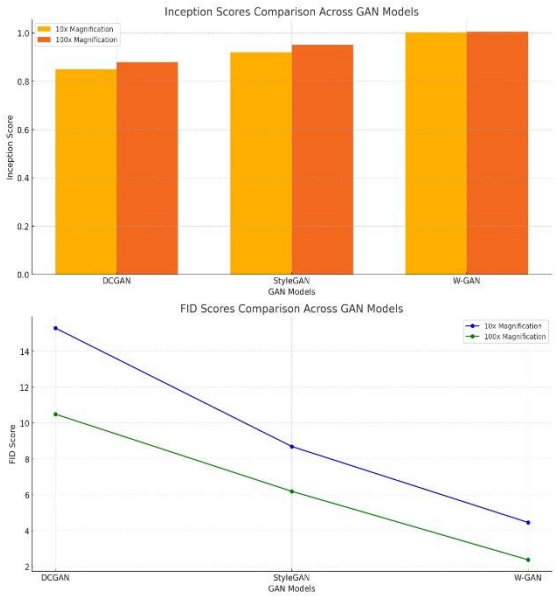

Fig 7: Comparison with other GANs

1. **Bar Chart (Inception Score)**:
    - Compares the performance of DCGAN, StyleGAN, and W-GAN across both 10x and 100x magnifications.
    - W-GAN exhibits the highest Inception Score, indicating superior image quality.
2. **Line Graph (FID Score)**:
    - Highlights FID trends across the GAN models for 10x and 100x magnifications.
    - W-GAN achieves the lowest FID, demonstrating its ability to generate more realistic and diverse images.

These graphs clearly establish W-GAN's improved performance over other models.

V. FUTURE WORK AND CONCLUSION

Preliminary evaluation results showed promising capability in generating synthetic brain tumor images with relatively high Inception Score and low FID values for 10x and 100x magnifications, respectively. Important improvements are still needed with regard to mainly saving fine details, maintaining color consistency, and reducing artefacts, to make the synthetic images applicable for actual clinical purposes.

Several key areas have been identified for improving the quality of
the generated images:

1. Fine Cellular Detail Preservation 100x Magnification: Better preservation of finer details is required. It will further improve the clinical application of the generated images if its resolution improves and the model can capture microscopic details better.
2. Consistency in Color Reproduction: Though the basic tissue is maintained, there seems to be variability in the strength of color intensity. Since color reproduction consistency is an area that significantly impacts medical image authenticity, it is important. This is because it is used as the basis for analysis in pathological assessments.
3. Reduction of Artifacts: It is observed that some of the artifact structures, appearing in the generated images, are not found in the original samples. At 10x magnification images, these artifacts should be minimized to such an extent that images become clinically viable.
4. Better Preservation of Nuclear Details: The nuclearcytoplasmic ratio, a very significant feature for the classification of tumors, should be preserved much better in the generated images, especially at 100x magnification, to increase the diagnostic accuracy.

VI. REFERENCES

Here are the references formatted according to the specified style:

[1] S. A. Shahriyar, M. A. H. Akhand, N. Siddique, and T. Shimamura, “Speech enhancement using convolutional denoising autoencoder,” in 2019 International Conference on Electrical, Computer and Communication Engineering (ECCE), pp. 1-5, 2019.

[2] S. Arakit, T. Hayashi, M. Delcroix, M. Fujimoto, K. Takeda, and T. Nakatani, “Exploring multi-channel features for denoising-autoencoder-based speech enhancement,” in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 116-120, 2015.

[3] X. Chen, “Unsupervised speech denoising method based on deep neural network,” in 2018 11th International Symposium on Computational Intelligence and Design (ISCID), pp. 254-258, 2018.

[4] X. Feng, Y. Zhang, and J. Glass, “Speech feature denoising and dereverberation via deep autoencoders for noisy reverberant speech recognition,” in 2014 IEEE International Conference on Acoustic, Speech, and Signal Processing (ICASSP), pp. 1759-1763, 2014.

[5] Y.-H. Lai, F. Chen, S.-S. Wang, X. Lu, Y. Tsao, and C.-H. Lee, “A deep denoising autoencoder approach to improving the intelligibility of vocoded speech in cochlear implant simulation,” IEEE Transactions on Biomedical Engineering.

[6] Z. Nian, Y.-H. Tu, J. Du, and C.-H. Lee, “A progressive learning approach to adaptive noise and speech estimation for speech enhancement and noisy speech recognition,” in 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 6913-6917, 2021.

[7] I. Potamitis, N. Fakotakis, and G. Kokkinakis, “Speech enhancement using the sparse code shrinkage technique,” in IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), pp. 621-624, 2001.

[8] L. Wang, B. Ren, Y. Ueda, A. Kai, S. Teraoka, and T. Fukushima, “Denoising autoencoder and environment adaptation for distant-talking speech recognition with asynchronous speech recording,” in IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA), 2023.

[9] H. Zhang, C. Liu, N. Inoue, and K. Shinoda, “Multi-task autoencoder for noise-robust speech recognition,” in Proceedings of the 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 5599-5603, 2018.

[10] M. Zhao, D. Wang, Z. Zhang, and X. Zhang, “Denoising autoencoder in speech recognition,” in Proceedings of APSIPA Annual Summit and Conference, pp. 338-342, 2015.