1. **Instance vs batch normalisation in DNN**

Instance normalization and batch normalization are both techniques used to normalize the activations of neurons in a deep neural network (DNN). However, they differ in their approach to normalization and the scope of normalization that they apply.

Batch normalization (BN) is a technique that normalizes the activations of neurons across an entire mini-batch of training samples. Specifically, BN normalizes each feature channel independently by subtracting the mean and dividing by the standard deviation of the activations within that channel across the batch. This technique can improve the training of deep neural networks by reducing internal covariate shift and helping to maintain a stable distribution of activations throughout the network. BN is often used in convolutional neural networks (CNNs) and has been shown to improve the convergence rate and accuracy of the model.

On the other hand, instance normalization (IN) is a technique that normalizes the activations of neurons within a single sample. Specifically, IN normalizes each feature channel independently by subtracting the mean and dividing by the standard deviation of the activations within that channel within the sample. This technique has been found to be effective in style transfer and other image-to-image translation tasks.

In summary, batch normalization normalizes the activations of neurons across an entire mini-batch, while instance normalization normalizes the activations of neurons within a single sample. Both techniques can be useful for improving the training of deep neural networks, but the choice of which to use may depend on the specific application and architecture of the network.

1. **what is style transfer**

Style transfer is a computer vision technique that involves applying the style of one image to the content of another image. The goal of style transfer is to generate a new image that combines the content of one image with the style of another image in a visually pleasing way. Style transfer is often used in artistic applications to create images that have a unique and aesthetically pleasing look.

The process of style transfer typically involves using a deep neural network, such as a convolutional neural network (CNN), to separate the content and style of the input images. The network is trained on a large dataset of images to learn how to extract the features that represent the content and style of an image. Once the network is trained, it can be used to generate a new image that combines the content of one image with the style of another image.

There are various approaches to style transfer, but a common approach involves using a pre-trained CNN to extract the feature maps of the input images, and then modifying these feature maps to generate a new image that combines the content and style of the input images. This can be done by optimizing a loss function that encourages the output image to match the content of the input image and the style of the reference image.

Style transfer has many practical applications, such as generating novel artwork, video game graphics, and even improving the visual appearance of medical images. It is an active area of research

1. **Use of Normalisation techniques in DNN**

Normalization techniques are widely used in deep neural networks (DNNs) to improve the training process and achieve better performance. Here are some ways in which normalization techniques are used in DNNs:

Reducing Internal Covariate Shift: One of the main reasons to use normalization techniques in DNNs is to reduce internal covariate shift. Internal covariate shift occurs when the distribution of activations changes throughout the layers of a DNN, making it difficult to train the network effectively. Normalization techniques, such as batch normalization (BN), layer normalization (LN), and group normalization (GN), help to mitigate this problem by normalizing the activations within a layer or across a mini-batch of training samples.

Improving Gradient Flow: Normalization techniques can also help to improve the flow of gradients during backpropagation. When the activations are normalized, the gradients are less likely to become too large or too small, which can help to prevent the vanishing gradient problem and improve the convergence rate of the network.

Regularization: Normalization techniques can act as a form of regularization by reducing overfitting. This is because the normalization helps to prevent the model from memorizing the training data and instead learn more general features that can be applied to new data.

Adapting to Different Data Distributions: Normalization techniques can also help the model to adapt to different data distributions, which can be important in situations where the distribution of the training data is different from the distribution of the test data. By normalizing the activations, the model becomes more robust to changes in the input data distribution.

1. **internal covariate shift**

Internal covariate shift is a phenomenon that occurs in deep neural networks (DNNs) when the distribution of the inputs to a layer changes during the training process. This can make it difficult to train the network effectively, as the optimization problem becomes more complex and the learning rate may need to be reduced. Internal covariate shift can be caused by a variety of factors, including changes in the distribution of the input data, changes in the parameters of the network, and changes in the objectives of the network.

Normalization techniques, such as batch normalization (BN), layer normalization (LN), and group normalization (GN), are commonly used to mitigate internal covariate shift in DNNs. These techniques work by normalizing the activations within a layer or across a mini-batch of training samples, so that the mean and variance of the activations are approximately constant. By doing so, normalization can help to reduce the dependence of the gradients on the scale of the activations, which can improve the convergence rate and generalization performance of the network.

Batch normalization is one of the most widely used normalization techniques, and involves normalizing the activations across a mini-batch of training samples. Layer normalization, on the other hand, normalizes the activations within a layer of the network, while group normalization divides the activations into groups and normalizes them independently. Each of these normalization techniques has its own strengths and weaknesses, and the choice of which one to use may depend on the specific application and architecture of the network.

Overall, normalization techniques are an important tool for mitigating internal covariate shift in DNNs, and can help to improve the stability, performance, and convergence rate of the network during training.

**what is FID score**

The Fréchet Inception Distance (FID) score is a metric used to evaluate the quality of generated images in generative adversarial networks (GANs). It measures the distance between the distributions of the generated images and the real images in a high-dimensional feature space.

The FID score is computed by first passing both the real and generated images through a pre-trained Inception-v3 network to extract feature representations. Then, the mean and covariance of these feature representations are calculated for both the real and generated images, and the distance between the means and covariances is measured using the Fréchet distance.

A lower FID score indicates that the generated images are more similar to the real images in terms of their visual quality and diversity. The FID score is often used as a more reliable and accurate measure of image quality compared to other evaluation metrics, such as the Inception Score or pixel-level metrics like Mean Squared Error (MSE) or Peak Signal-to-Noise Ratio (PSNR).

Note that the FID score is generally used for comparing the performance of different GAN models or different configurations of the same GAN model, and not as an absolute measure of image quality.



Dhiraj Hasija

How does FID score work?

The Fréchet Inception Distance (FID) score measures the distance between the distributions of the real and generated images in a high-dimensional feature space. The FID score is computed by following these steps:

1. Preprocess the real and generated images: First, the real and generated images are preprocessed to ensure that they are the same size and format, and they are normalized to have the same pixel value range (usually [-1, 1]).
2. Extract feature representations: The real and generated images are passed through a pre-trained deep neural network, such as the Inception-v3 network, which is often used in image classification tasks. The output of the network at a specific layer is used as a feature representation for each image.
3. Calculate the mean and covariance: The mean and covariance of the feature representations are calculated separately for the real and generated images.
4. Compute the Fréchet distance: The Fréchet distance is calculated between the two distributions of feature representations using the formula:  
   FID = ||mu1 - mu2||^2 + Tr(C1 + C2 - 2 \* (C1 \* C2)^0.5)  
   where mu1 and mu2 are the mean feature representations for the real and generated images, and C1 and C2 are the covariance matrices for the real and generated images.  
   The Fréchet distance measures the distance between the two multivariate Gaussian distributions represented by the mean and covariance of the real and generated feature representations. A lower FID score indicates that the generated images are more similar to the real images in terms of their visual quality and diversity.

The FID score is often used as a more reliable and accurate measure of image quality compared to other evaluation metrics, such as the Inception Score or pixel-level metrics like Mean Squared Error (MSE) or Peak Signal-to-Noise Ratio (PSNR).



Dhiraj Hasija

feature representation for each image means?

Feature representation for each image refers to a set of numerical values that capture the visual features of the image in a high-dimensional space. These features are obtained by passing the image through a pre-trained deep neural network, such as Inception-v3 or VGGNet, up to a certain layer in the network. The output of the layer is used as the feature representation for the image.

The feature representation is a vector of values that describe various visual aspects of the image, such as edges, textures, shapes, and colors. Each element in the vector corresponds to a specific feature or aspect of the image, and the vector as a whole captures the overall visual content of the image.

The feature representation for each image is then used to compare the similarity between the real and generated images in terms of their visual content. In the case of the FID score, the mean and covariance of the feature representations are calculated for both the real and generated images, and the distance between these distributions is used as a measure of similarity between the two sets of images.



Dhiraj Hasija

Help me interpret the distance between distributions using mean and covariance

In the context of comparing two distributions using their mean and covariance, the distance between the distributions is measured using the Fréchet distance. The Fréchet distance is a measure of similarity or dissimilarity between two multivariate Gaussian distributions, represented by their mean and covariance.

To calculate the Fréchet distance between two distributions, first, the mean and covariance matrix for each distribution are computed. The mean is the average of the values in each dimension of the distribution, while the covariance matrix describes the relationship between the different dimensions of the distribution.

Once the mean and covariance matrices are calculated for both distributions, the Fréchet distance is computed using the following formula:

FID = ||mu1 - mu2||^2 + Tr(C1 + C2 - 2 \* (C1 \* C2)^0.5)

where:

* mu1 and mu2 are the means of the two distributions
* C1 and C2 are the covariance matrices of the two distributions
* Tr() is the trace operator, which calculates the sum of the diagonal elements of a matrix
* || || denotes the Euclidean distance between two vectors

The Fréchet distance measures the distance between the two distributions in the feature space, where each dimension corresponds to a different visual feature or aspect. The distance is calculated based on the difference between the means of the two distributions and the difference between their covariance matrices.

A lower Fréchet distance indicates that the two distributions are more similar in terms of their visual features, while a higher distance indicates greater dissimilarity. In the context of comparing real and generated images in deep learning, a lower Fréchet distance means that the generated images are more similar to the real images, and therefore are of higher quality.

what is the mathematical function when we change the contrast of a RGB image

The mathematical function used to change the contrast of an RGB image depends on the method used to adjust the contrast. One commonly used method is to apply a linear transformation to the pixel values of the image.

Let's consider a pixel value in the RGB image represented by the three color channels R, G, and B. The value of each channel can range from 0 to 255. To adjust the contrast, we can apply the following linear transformation to each pixel:

New value = (Old value - Mean) x Scale + Mean

where:

* Old value: the original pixel value for a particular color channel (R, G, or B)
* Mean: the mean value of the pixel values across all the color channels in the image
* Scale: a scaling factor to control the strength of the contrast adjustment

By subtracting the mean from the old value, we center the values around zero. We then multiply this difference by a scaling factor to increase or decrease the contrast, and add the mean back to ensure that the pixel values are in the original range of 0 to 255.

The value of the scaling factor depends on the desired amount of contrast adjustment. For example, if we want to increase the contrast of the image, we can set the scaling factor to a value greater than 1, while if we want to decrease the contrast, we can set the scaling factor to a value less than 1.

This linear transformation can be applied to all pixels in the RGB image to adjust the contrast. Note that other methods, such as histogram equalization or gamma correction, can also be used to adjust the contrast of an RGB image, and may involve different mathematical functions.