DS261: AI for Medical Image Analysis

Assignment-03

Submitted By: Aman Pawar (SR No: 22761),

MTech 1st Year, Department of Bioengineering

PART-I

Answer 01:

Part (a)

Subject	VA (mL)	VB (mL)	VA - VB (mL)
1	0.50	0.79	-0.29
2	0.58	0.71	-0.13
3	0.90	0.82	0.08
4	1.17	0.82	0.35
5	1.14	0.73	0.41
6	1.25	0.77	0.48
7	0.75	0.72	0.03
8	1.22	0.79	0.43
9	0.74	0.72	0.02
10	0.80	0.91	-0.11

a) We can use the paired t-test statistic, as we have paired data (VA and VB for each subject).

Submitted By: Aman Pawar SR.NO: 22761
M Tech 1st year, Department of
                        Bioengineering.

Ans 01. (a)

Subjects	V_A (ml)	V_B (ml)	$V_A - V_B$	$((V_A - V_B) - \mu_D)^2$
1	0.50	0.79	−0.29	$(-0.29 - 0.127)^2$ = 0.173889
2	0.58	0.71	−0.13	0.066049
3	0.90	0.82	0.08	2.209×10^{-3}
4	1.17	0.82	0.35	0.049729
5	1.14	0.73	0.41	0.080089
6	1.25	0.77	0.48	0.124609
7	0.75	0.72	0.03	9.409×10^{-3}
8	1.22	0.79	0.43	0.091809
9	0.74	0.72	0.02	0.011449
10	0.80	0.91	−0.11	0.056169
			Σ 1.27	

$$\mu_D = \frac{\Sigma(V_A - V_B)}{N}$$

$$= \frac{1.27}{10} = 0.127$$

$$SD = \sqrt{\frac{\Sigma(V_A - V_B - \mu_D)^2}{N-1}}$$

$$= 0.2719$$

$$t_{statistic} = \frac{\mu_D}{\frac{SD}{\sqrt{n}}}$$

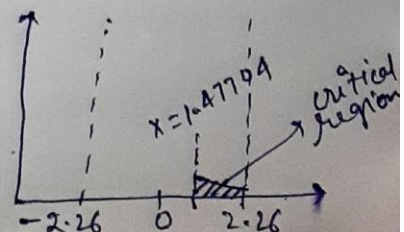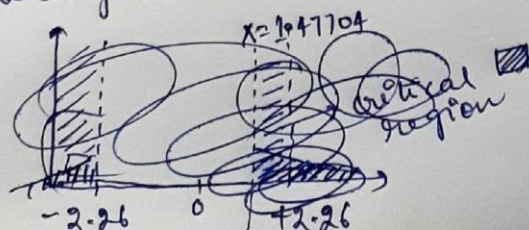$$= \frac{0.127}{\frac{0.2719}{\sqrt{10}}}$$

$$= 1.47704$$

∴ ~~as we accept the null hypothesis~~

$\alpha = 0.05$ and degree of freedom df = n−1 = 9

the critical value comes out to be ± 2.2622
Taking it a double tail t-test ∴ $\alpha = 0.025$.

For a two-tailed test with α = 0.05 and degrees of freedom (df) = n - 1 = 10 - 1 = 9, the critical region is defined by the t-distribution with 9 df and α/2 = 0.025 on each tail.

Mean of differences = 0.127

Sample SD of differences = 0.2719

*T statistic = mean of differences/(Sample SD/ √n) = 0.127/(0.2719/√10) = 1.477*

For the two-tailed test, the critical region is defined by the t-distribution with 9 df and α/2 = 0.025 on each tail. From the t table, we see it is +/- 2.262.
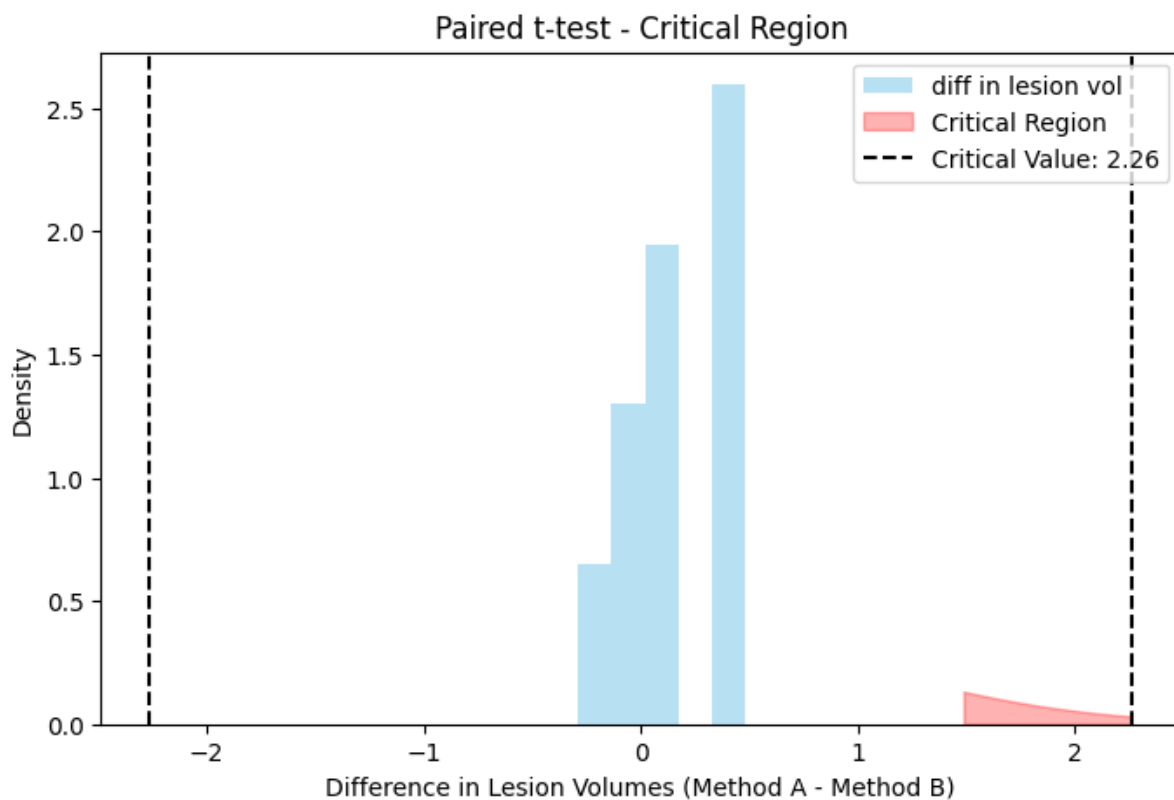
Part (b)

This is a one-tailed paired t-test, as we are only interested in detecting if the mean difference is greater than zero. Because it is greater than 0, it is a right-tailed t-test.

For the one-tailed test, the critical region is defined by the t-distribution with 9 df and α = 0.05. From the t table, we see it is +/- 1.833

Paired t-test - Critical Region
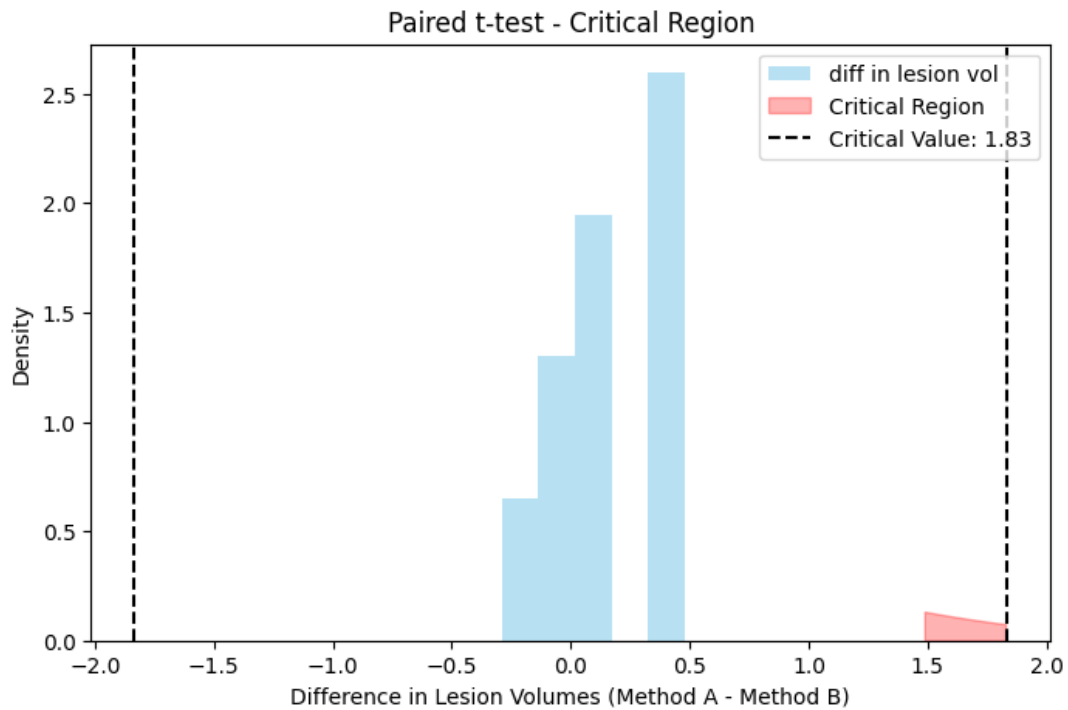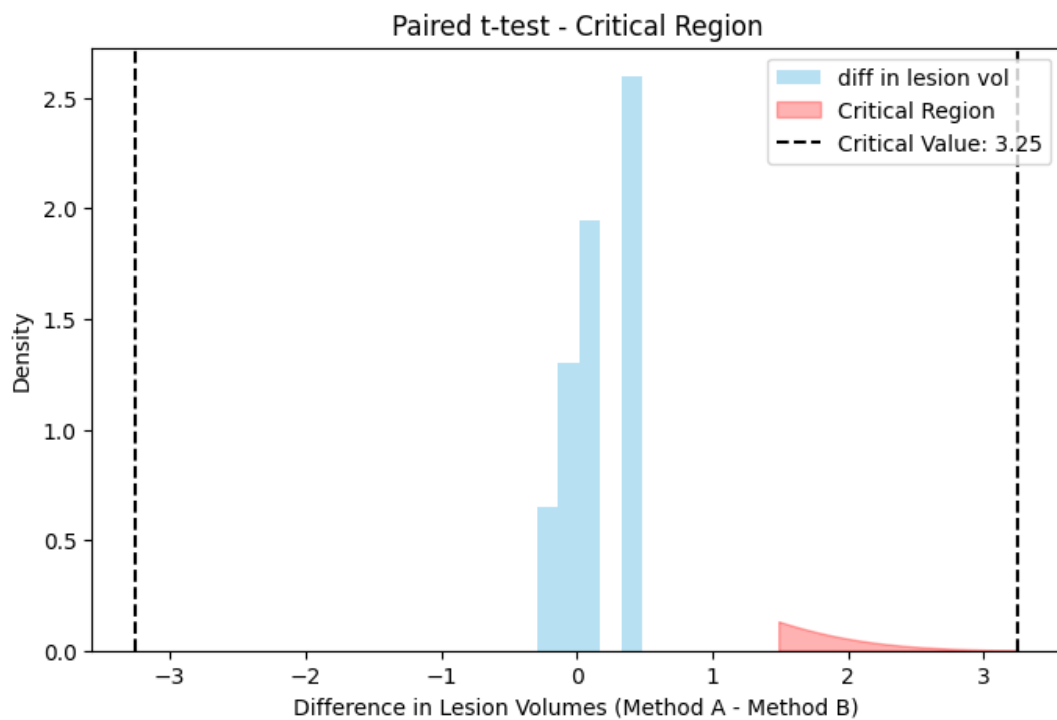
Part (c)

With $\alpha = 0.01$,

For two-tailed t-test, the critical region from the t table is +/- 3.25.

For one-tailed t-test, the critical region from the t table is +/- 2.821.

The calculated t-statistic is less than the critical value for all the cases above, so we would fail to reject H0.

Paired t-test - Critical Region

The p-value for this test is the probability of observing a t-statistic as extreme as 1.477

For the two-tailed test, the p-value is 0.173 (Result not significant at α=0.01 and α=0.05).

For the one-tailed test, the p-value is 0.0868 (Result not significant at α=0.01 and α=0.05).

Answer 02:

Outcome (x)	Frequency (f)	fx	fx^2
1	1	1	1
2	4	8	16
3	13	39	117
4	19	76	304
5	16	80	400
6	15	90	540
7	9	63	441
8	12	96	768
9	7	63	567
10	2	20	200
11	1	11	121

12	*1*	*12*	*144*
Total	*100*	*559*	*3610*

Null Hypothesis (H0): The data follows a Poisson distribution.

Alternative Hypothesis (Ha): The data does not follow a Poisson distribution.

Significance level: α = 0.05

Poisson parameter (λ) =559/100=5.59

For poisson distribution, df = n-2 =10

The critical value for α = 0.05 and df = 10 in the chi-square distribution table is 18.31.

To get the chi-square statistic, For each outcome (x), we calculate the squared deviation between the observed (fi) and expected (ei) frequencies and divide by the expected frequency: $(fi - ei)^2 / ei$.

We then sum the squared deviations for all outcomes to get the chi-square statistic (χ^2):

$\chi^2 = \Sigma((fi - ei)^2 / ei)$ --- 1

ei for each outcome is obtained by using the Poisson parameter value using the formula $\lambda^x * exp(-\lambda) / x!$ where x would take the values 1 to 12.

The ei values are : 2.088,5.836,10.874,15.196,16.989,15.828,12.64,8.832,5.486,3.067,1.558,0.726

*From these values, we can substitute them in Formula 1.*

We get the chi-square statistic as 5.87

As this value is less than the critical value, we fail to reject the Null hypothesis.

Chi-square statistic: 5.87714611574077

Critical value at alpha = 0.05 and df = 10: 18.307038053275146

The observed frequencies fit a Poisson distribution (fail to reject H0)

Observed vs Expected Frequencies
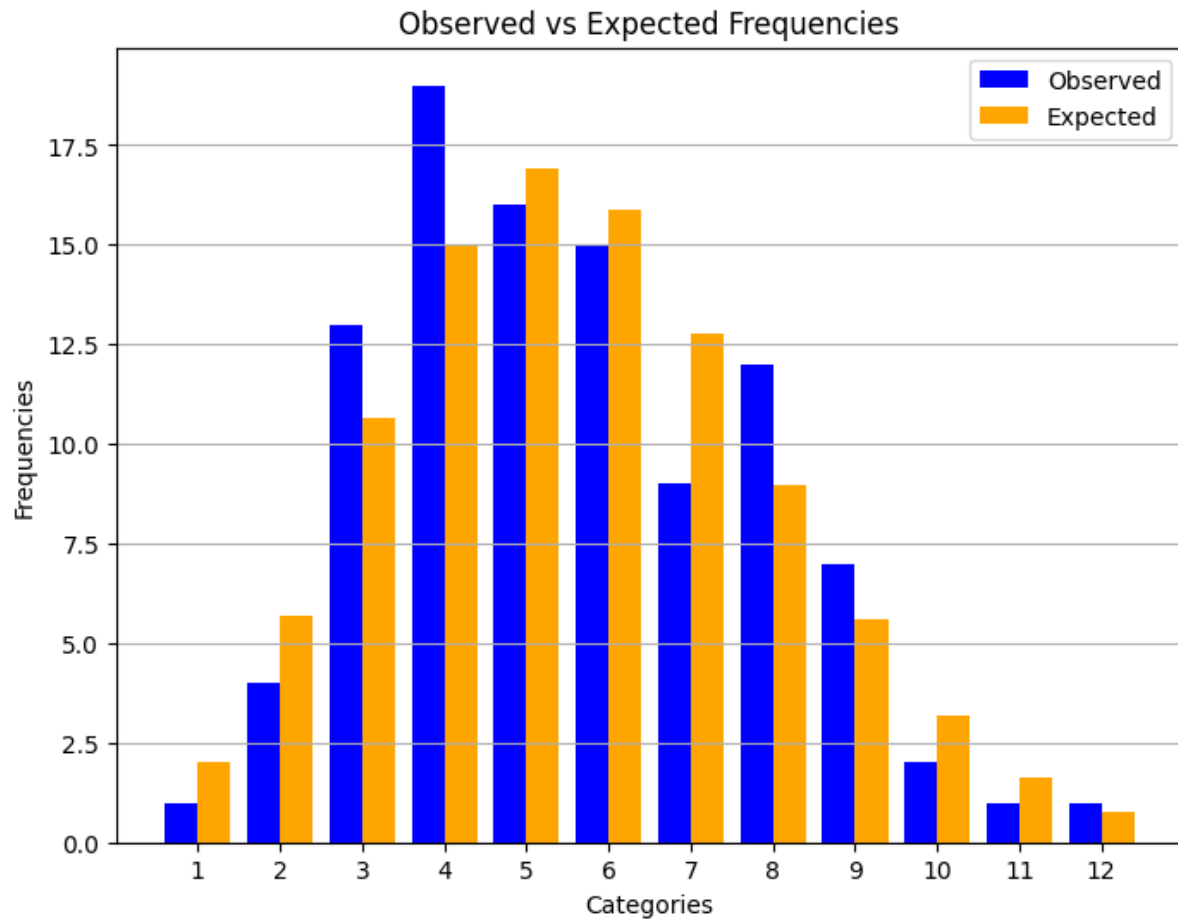
```
                      df       sum_sq      mean_sq           F  PR(>F)
C(Nutrient)          1.0     29.27521    29.275210    5.532832  2.497936e-02
C(Age)               3.0    454.69923   151.566410   28.645106  3.439786e-09
C(Nutrient):C(Age)   3.0     27.34379     9.114597    1.722602  1.820629e-01
Residual            32.0    169.31776     5.291180         NaN         NaN
```

Factor	Degrees of Freedom (df)	Critical F-value (α = 0.05)
Nutrient	1	Approximately 4.11
Age	3	Approximately 3.10
Nutrient:Age	3	Approximately 3.10
Residual (Error)	32	- (Not applicable)

F-statistic for Nutrient: 5.532831995887513, p-value: 0.024979359808804828

F-statistic for Age: 28.64510562861211, p-value: 3.4397861200280045e-09

F-statistic for Interaction: 1.7226018896856075, p-value: 0.18206291537597424

There are 40 samples. The overall total is 794.82. There are a total of 8 groups with 5 subjects.

The row total, column total, and overall total are mentioned in the table.

Given α is 0.05.

The correction term Cx is the square of the sum/number of observations. That is 631738.83/40=15793.47

Sum of squares of total (SSt) = (Sum of squares) - Cx = 16474.10-15793.47=680.62

Sum of squares of columns (SSc) = (Sum of squares of column sum)/number of entries per column - Cx = 16248.17 - 15793.47 =454.7

Sum of squares of rows(SSr) = (Sum of squares of row sum)/number of entries per row- Cx = 15822.74 - 15793.47 = 29.27

Sum of squares within groups(SSg) = (Sum of squares of per group sum)/number of groups - Cx - SSr - SSc = 27.34

a) HA: $\alpha 1 = \alpha 2 = 0$

With the values obtained above, we can reject the null hypothesis for HA

b) HB: $\beta 1 = \beta 2 = \beta 3 = \beta 4 = 0$

With the values obtained above, we can reject the null hypothesis for HB

c) HAB: $\gamma ij = 0$; i=1, 2; j = 1,2,3,4

With the values obtained above, we cannot as the critical value is more than the F statistic reject the null hypothesis for HAB

Part (a)

Train Ratio : 0.8, Val Ratio: 0.19999999999999996

Training Samples: 9567, Validation Samples: 2392, Test Samples: 3421

Visualtizing Data from the DataLoader...

The Shape of Image is : (3, 28, 28)

Part (b)

Variational Auto Encoder

Epoch [96/100] - Train Loss: 1037.7058

Epoch [96/100] - Validation Loss: 1033.2474

Epoch 97/100: 100%|█████████| 10/10 [00:01<00:00, 8.02it/s]

Epoch [97/100] - Train Loss: 1037.6524

Epoch [97/100] - Validation Loss: 1033.0226

Epoch 98/100: 100%|█████████| 10/10 [00:01<00:00, 8.43it/s]

Epoch [98/100] - Train Loss: 1037.5984

Epoch [98/100] - Validation Loss: 1033.0655

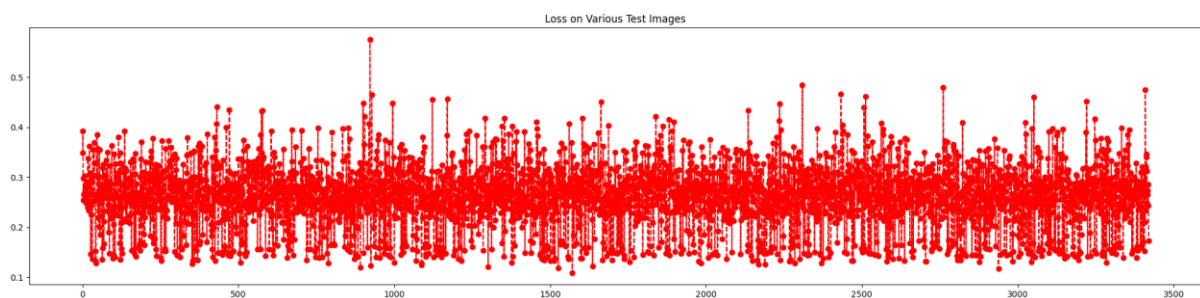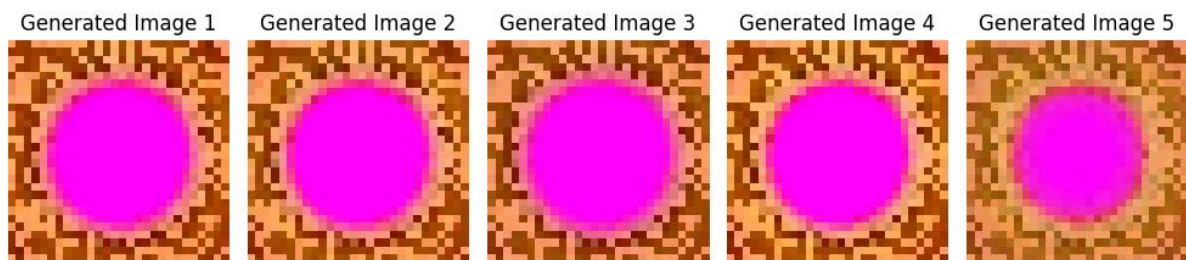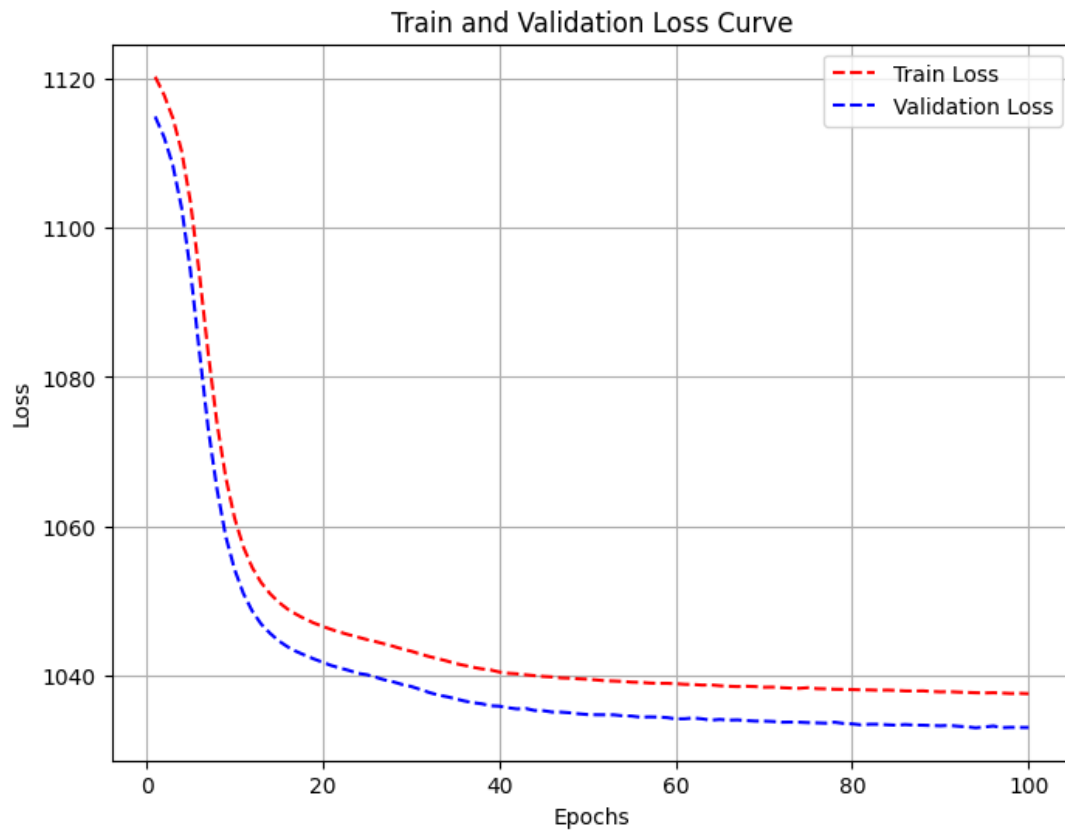*Epoch 99/100: 100%|█████████| 10/10 [00:01<00:00, 7.69it/s]*

Epoch [99/100] - Train Loss: 1037.6158

Epoch [99/100] - Validation Loss: 1033.0473

Epoch 100/100: 100%|█████████| 10/10 [00:01<00:00, 8.12it/s]

Epoch [100/100] - Train Loss: 1037.5720

Epoch [100/100] - Validation Loss: 1033.0410

Train and Validation Loss Curve

Generated Image 1 Generated Image 2 Generated Image 3 Generated Image 4 Generated Image 5

Loss on Various Test Images
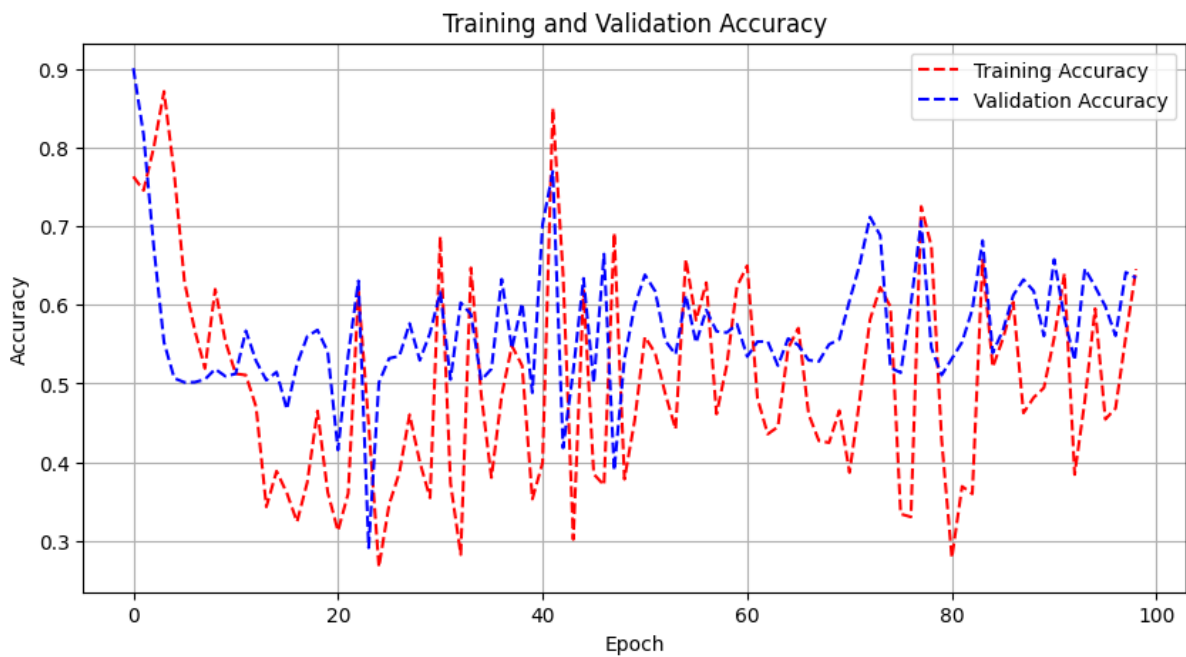
Generative Adversarial Network (GAN):

...

Epoch 99/100 - Training: 100%|███████████| 10/10 [00:03<00:00, 3.28it/s, Generator Loss=1.65, Discriminator Loss=1.62]

Epoch 99/100 - Validation: 100%|██████████| 3/3 [00:00<00:00, 6.37it/s, Validation Loss=1.5]
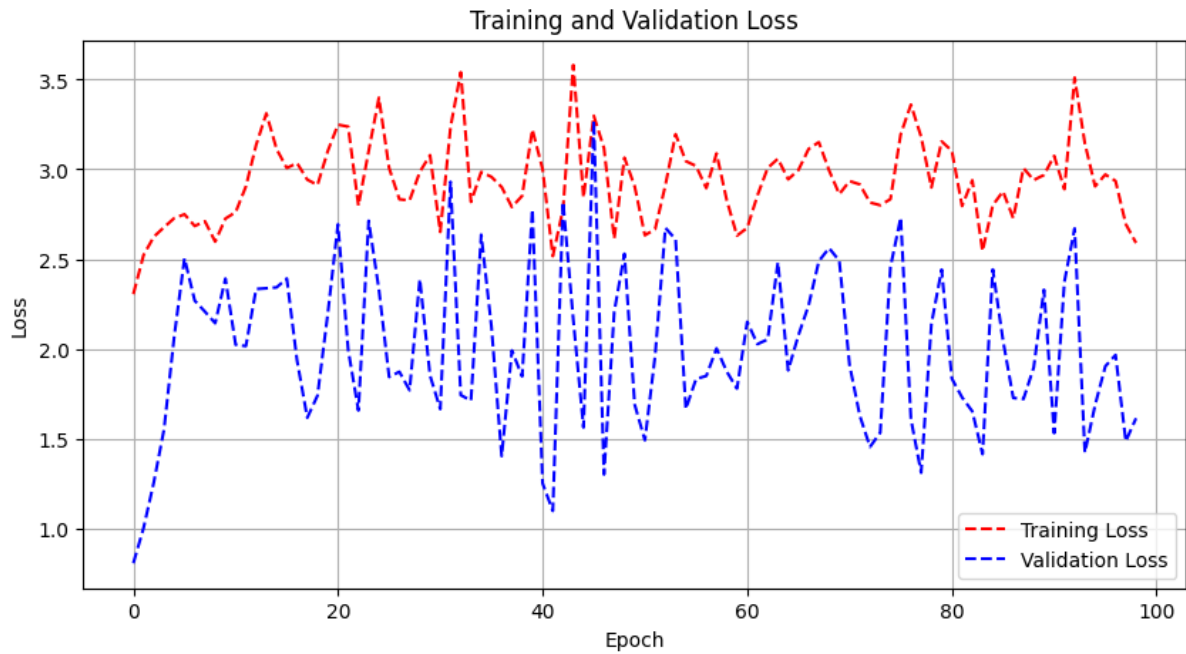
Epoch 100/100 - Training: 100%|███████████| 10/10 [00:03<00:00, 3.30it/s, Generator Loss=1.56, Discriminator Loss=1.62]

Epoch 100/100 - Validation: 100%|███████████| 3/3 [00:00<00:00, 6.33it/s, Validation Loss=1.65]

Training and Validation Loss

Training and Validation Accuracy

Generated Image 1 Generated Image 2 Generated Image 3 Generated Image 4 Generated Image 5

Frechet Inception Distance (FID) between test and generated images: 78.47885692843965

Image: 1:

Real Images vs. Generated Images:

t-statistic (Real vs. Gen1): -8.79722069525412, p-value (Real vs. Gen1): 1.937389471240449e-18

t-statistic (Real vs. Gen2): 6.195650401501249, p-value (Real vs. Gen2): 6.300141414941959e-10

Generated Images (Gen1 vs. Gen2):

t-statistic (Gen1 vs. Gen2): 16.90250816771179, p-value (Gen1 vs. Gen2): 2.8782337074464014e-62

Reject null hypothesis (H0): Real vs. Gen1 - There is a statistically significant difference.

Reject null hypothesis (H0): Real vs. Gen2 - There is a statistically significant difference.

Reject null hypothesis (H0): Gen1 vs. Gen2 - There is a statistically significant difference.

Image: 2:

Real Images vs. Generated Images:

t-statistic (Real vs. Gen1): -17.534743127709927, p-value (Real vs. Gen1): 9.929207409446673e-67

t-statistic (Real vs. Gen2): -0.6720064720997452, p-value (Real vs. Gen2): 0.5016125929860872

Generated Images (Gen1 vs. Gen2):

t-statistic (Gen1 vs. Gen2): 17.4969563822551, p-value (Gen1 vs. Gen2): 1.8519374679600713e-66

Reject null hypothesis (H0): Real vs. Gen1 - There is a statistically significant difference.

Failed to reject null hypothesis (H0): Real vs. Gen2 - There is no statistically significant difference.

Reject null hypothesis (H0): Gen1 vs. Gen2 - There is a statistically significant difference.

Effect Size	Sample Size	Power (Real vs. Gen1)	Power (Real vs. Gen2)	Power (Gen1 vs. Gen2)
0.2	20	0.094567	0.094567	0.060959
0.2	50	0.167675	0.167675	0.078524
0.2	100	0.290646	0.290646	0.108372
0.5	20	0.337939	0.337939	0.120335
0.5	50	0.696893	0.696893	0.23578
0.5	100	0.940427	0.940427	0.420538
0.8	20	0.693404	0.693404	0.234349
0.8	50	0.977279	0.977279	0.508186
0.8	100	0.999878	0.999878	0.803647

Effect Size Impact: As the effect size increases (from 0.2 to 0.8), the statistical power of the tests increases across all comparisons.

Sample Size Impact: Increasing the sample size generally improves statistical power. For a fixed effect size, larger sample sizes lead to higher power. This relationship is evident in most comparisons, where higher sample sizes correspond to higher power values.

Comparing Generations (Gen1 vs. Gen2): Across the different effect sizes and sample sizes, the power differences between Gen1 and Gen2 tend to increase with larger effect sizes and sample sizes. This suggests that larger effect sizes and sample sizes lead to more distinguishable power differences between these generations.

Effect Size Comparisons: For a given sample size, larger effect sizes consistently lead to higher power values. This is evident when comparing the power values within each row for different effect sizes but the same sample size.

In summary, larger effect sizes and sample sizes generally result in higher statistical power, which means a greater ability to detect true effects. Additionally, the power differences between different effect sizes and sample sizes impact the sensitivity of the statistical tests, influencing the ability to detect differences between generations (Gen1 and Gen2).