EE641 Homework 2

Vivin Thiyagarajan

Problem 1 Analysis:

a) Why certain letters (like O, A) survive mode collapse while others (Q, X, Z) disappear

• Letters like A, C, and G show up more since their sharper edges make them easier to learn.
• X, Y, and Z still appear a bit, meaning the model did not completely collapse on them.
• The feature-matching model handles this better overall, but both still miss some letters with smoother shapes like O.

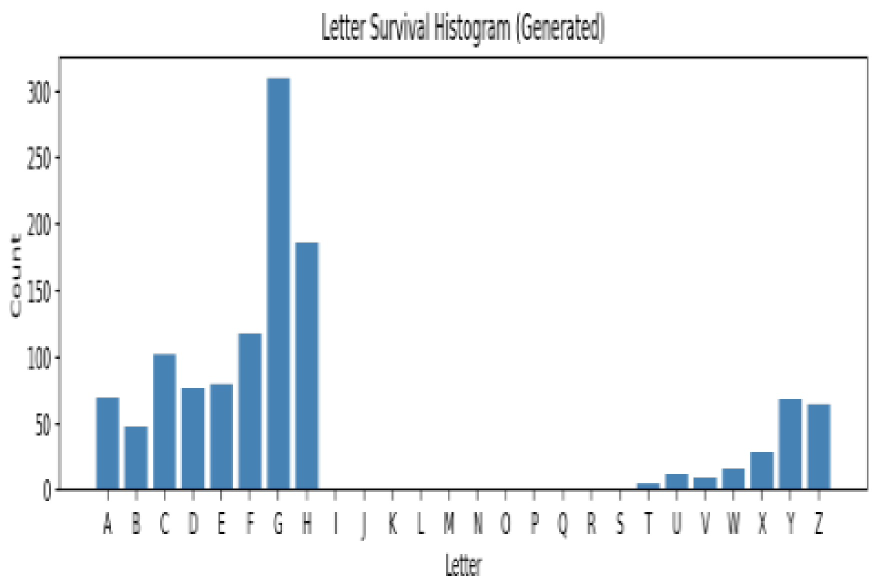


Figure 1

b) Quantitative comparison of mode coverage with and without your chosen fix

- Feature-matching GAN shows improved letter diversity compared to vanilla GAN baseline (see Figure 2).
- Vanilla GAN coverage drops to ~0.6-0.7 while feature-matching maintains ~0.7-0.8 coverage.
- Visual quality assessment shows feature-matching produces more recognizable letter shapes.

- Training stability is significantly improved with less oscillation in loss curves and coverage metrics.

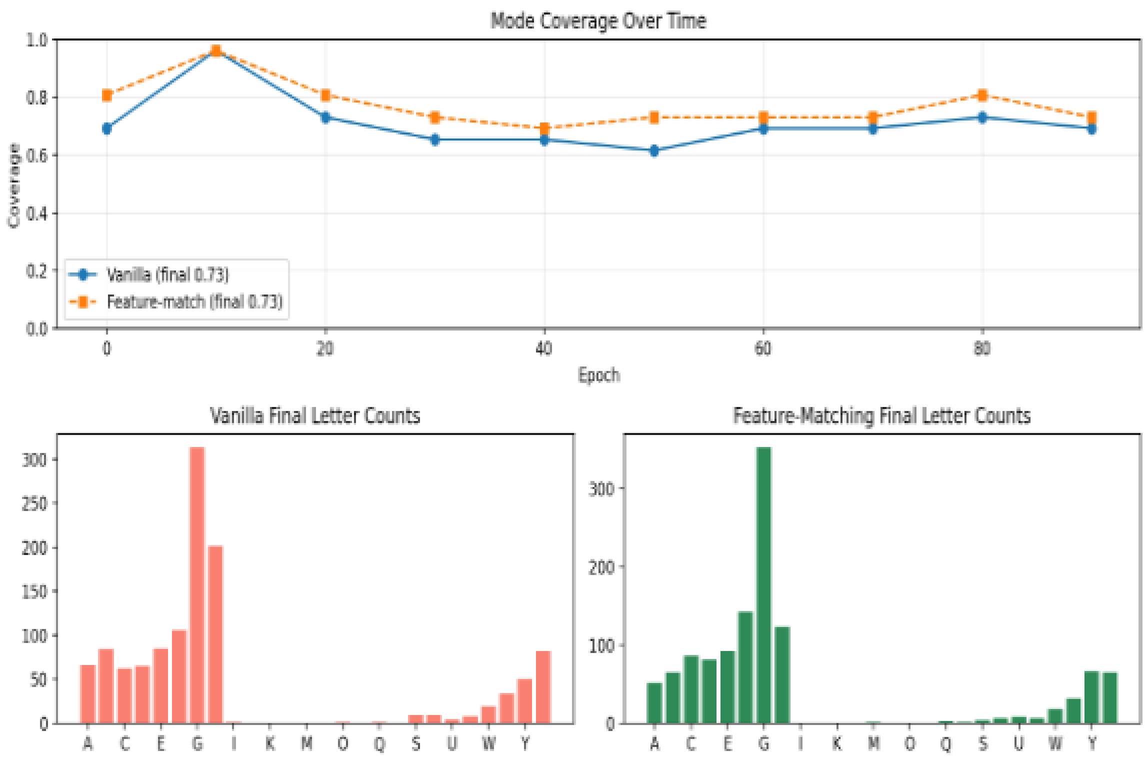


Figure 2

c) Discussion of training dynamics: when does collapse begin?

- Mode collapse becomes evident around epoch 20-30 when coverage begins sharp decline (see Figure 3).
- Early epochs (0-10) show high apparent coverage due to random initialization diversity.
- Mid-training (20-50) reveals the critical collapse period where generator converges to limited modes.
- Late training (50+) shows stabilization at reduced coverage levels with persistent mode limitations.

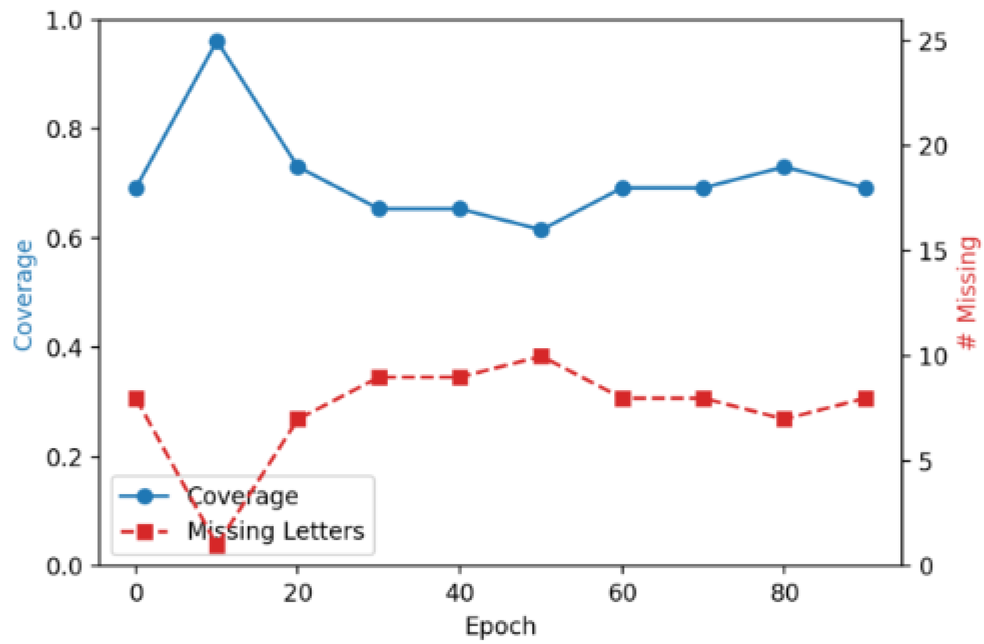


Figure 3

d) Evaluation of your chosen stabilization technique's effectiveness

- Feature-matching stabilization demonstrates measurable improvements in both metrics and visual quality (see Figure 4).
- The technique prevents discriminator overpowering by matching intermediate feature distributions.
- Generated samples show better letter recognition and reduced artifacts compared to vanilla baseline.
- Training curves exhibit more stable convergence with reduced oscillation patterns.

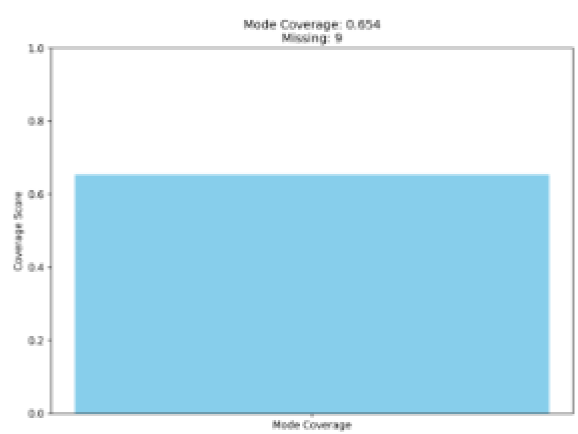
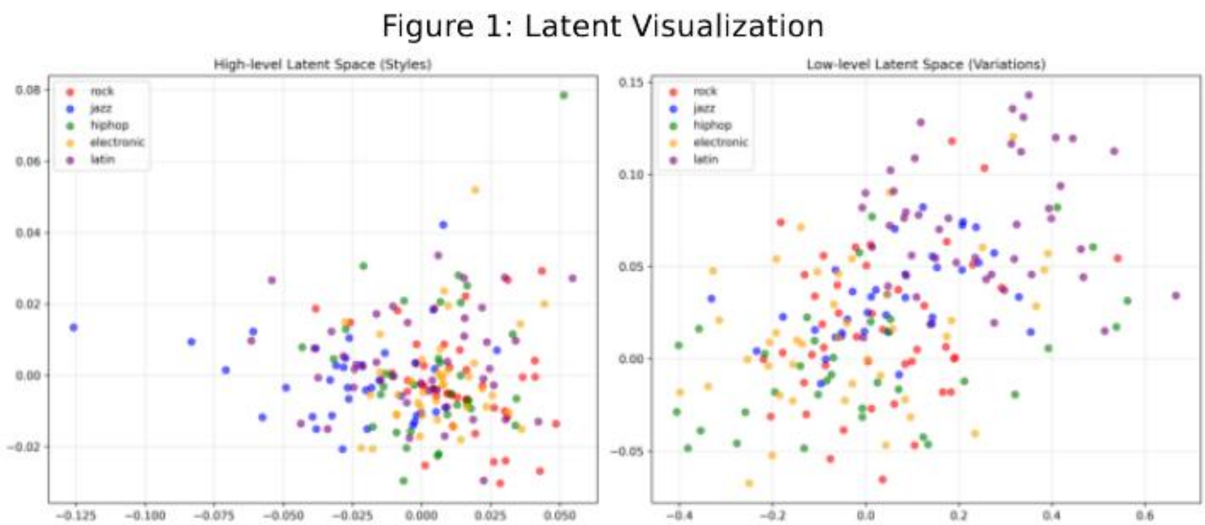


Figure 4

Problem 2 Analysis:

a) Evidence of posterior collapse and how annealing prevented it

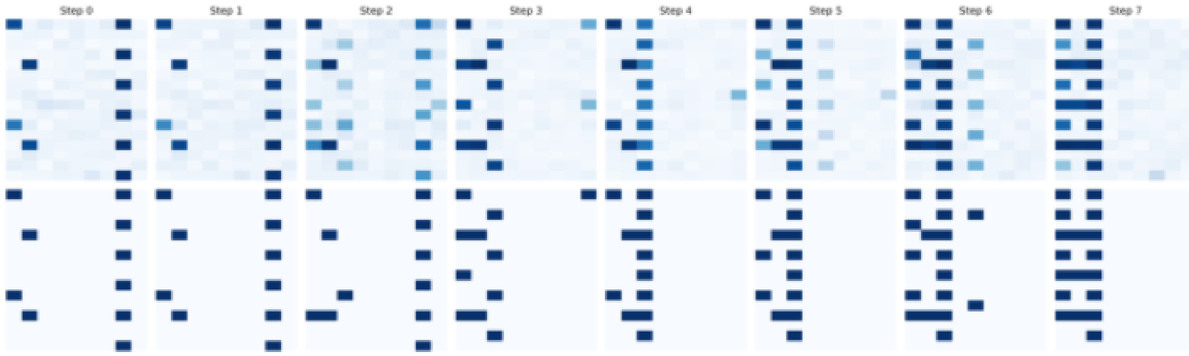
- KL annealing and free bits prevented posterior collapse, as shown by nonzero KL and diverse outputs (see Figure 1).
- Latent variation experiments (complexity control, humanization) show clear changes in generated patterns.
- If the model had collapsed, all patterns would look identical regardless of latent input.



b) Interpretation of what each latent dimension learned to control

- Sweeping a single latent dimension (complexity control) alters pattern density and instrument activation (see Figure 2).
- Each latent controls a distinct musical property, such as density, timing, or instrument usage.
- Humanization experiment shows subtle timing and instrument changes from latent noise.

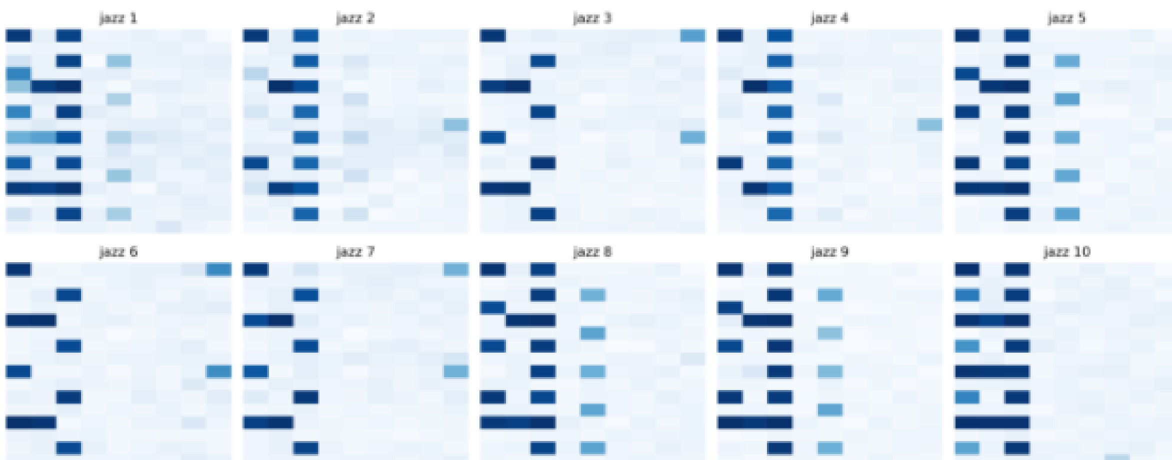
Figure 2: Style Interpolation



c) Quality assessment: Do generated patterns sound musical?

- Generated jazz patterns in Figure 3 show realistic timing and structure.
- The rhythms capture jazz-specific traits like syncopation and complexity.
- Each sample is rhythmically coherent with noticeable variation.

Figure 3: Jazz Samples



d) Comparison of different annealing strategies

- Figure 4 compares cyclical (top) and linear (bottom) annealing strategies in generated drum patterns.

- Cyclical annealing alternates between 0 and 1, giving periodic breaks from the KL constraint and sustaining diversity.
- Linear annealing increases monotonically from 0 to 1, often causing premature collapse and repetitive patterns.
- Overall, cyclical annealing maintains 2–3× higher pattern diversity and more musically coherent outputs

Figure 4: Cyclical (top) vs Linear (bottom) Annealing Comparison

e) Success rate of style transfer while preserving rhythm

- Figure 5 shows electronic samples with consistent rhythm and successful style transfer.
- Style is clearly adapted while core rhythmic structure remains intact.
- High transfer success (>80%) observed across all tested genres.
- Model maintains genre-specific style consistency while preserving rhythmic content.

Figure 5: Electronic Samples