

Variational Auto-Encoder (VAE) ile Denoising ve Representation Generation (Fashion-MNIST)

Fatma Nur Sakal

Giriş

- Bu projede amaç iki şeyi aynı modelle göstermek:
 1. Gürültülü görüntülerden temiz görüntü üretmek (denoising),
 2. Veriden anlamlı bir latent uzay (representation) öğrenip bu uzaydan yeni örnekler üretebilmek (generation).
- Varyasyonel Otoenkoder (VAE) ile hem rekonstrüksiyon kalitesi hem de latent uzayın yapısı analiz edilecek; eğitim eğrileri, test metrikleri ve görsel örnekler üzerinden modelin gerçekten öğrenip öğrenmediği kanıtlanacak.

Problem Tanımı ve Motivasyon

- Gürültü, gerçek dünyada sensörler, düşük ışık, sıkıştırma, iletim hataları gibi birçok kaynaktan gelir ve görsel bilgi kaybına yol açar.
- Bu projede “noisy \rightarrow clean” dönüşümünü öğrenebilen bir VAE tasarlanır.
- Ek olarak VAE’nin sadece kopyalamadığını; latent uzayda sınıfları kısmen ayıran, anlamlı kümeler oluşturan bir temsil öğrendiğini göstermek hedeflenir (PCA/UMAP/t-SNE analizleri).

Veri Seti (Fashion-MNIST)

- Fashion-MNIST; 28×28 tek kanallı (grayscale) kıyafet görüntülerinden oluşan, 10 sınıflı standart bir benchmark'tır.
- Veri seti; hızlı eğitim, net kıyas ve görsel kaliteyi yorumlamayı kolaylaştırdığı için denoising + representation çalışmaları için uygundur.
- Sınıf etiketleri latent uzay görsellerinde (renk/etiket) kullanılacaktır.

Ön İşleme ve Veri Boru Hattı

- Görüntüler $28 \times 28 \times 1$ şekline getirilir ve piksel değerleri $[0,1]$ aralığına normalize edilir.
- Eğitimde denoising için giriş görüntülerine kontrollü gürültü eklenir: `NOISE_FACTOR = 0.45`.
- Model “gürültülü görüntüyü al \rightarrow temiz görüntüyü üret” mantığıyla öğrenir.
- Eğitim/validasyon/test ayrımı; genelleme performansını ölçmek için zorunlu tutulur.

Deney Kurulumu ve Hiperparametreler

- Bu çalışma tek bir koşu üzerinden raporlanır ve tekrar üretilebilirlik hedeflenir.
- Temel hiperparametreler: LATENT_DIM = 16, BATCH = 256, EPOCHS = 20. Latent boyut, hem denoising kapasitesini hem de representation ayrışmasını etkiler; bu yüzden orta boy bir latent (16) seçilerek kalite/ayrışma dengesi hedeflenmiştir.

Neden VAE?

- Klasik autoencoder, latent uzayı “sıkıştırma” yapar ama örneklenebilir (sampling) bir uzay garantisi vermez.
- VAE ise latent değişkeni olasılıksal tanımlar: model hem görüntüyü yeniden kurmayı öğrenir hem de latent uzayı düzenli (regularized) tutar.
- Bu düzenlilik sayesinde “prior’dan örnekleme” ile yeni görüntüler üretilebilir ve latent uzayda anlamlı geçişler (interpolation) denenebilir.

Model Mimarisi (Encoder–Decoder)

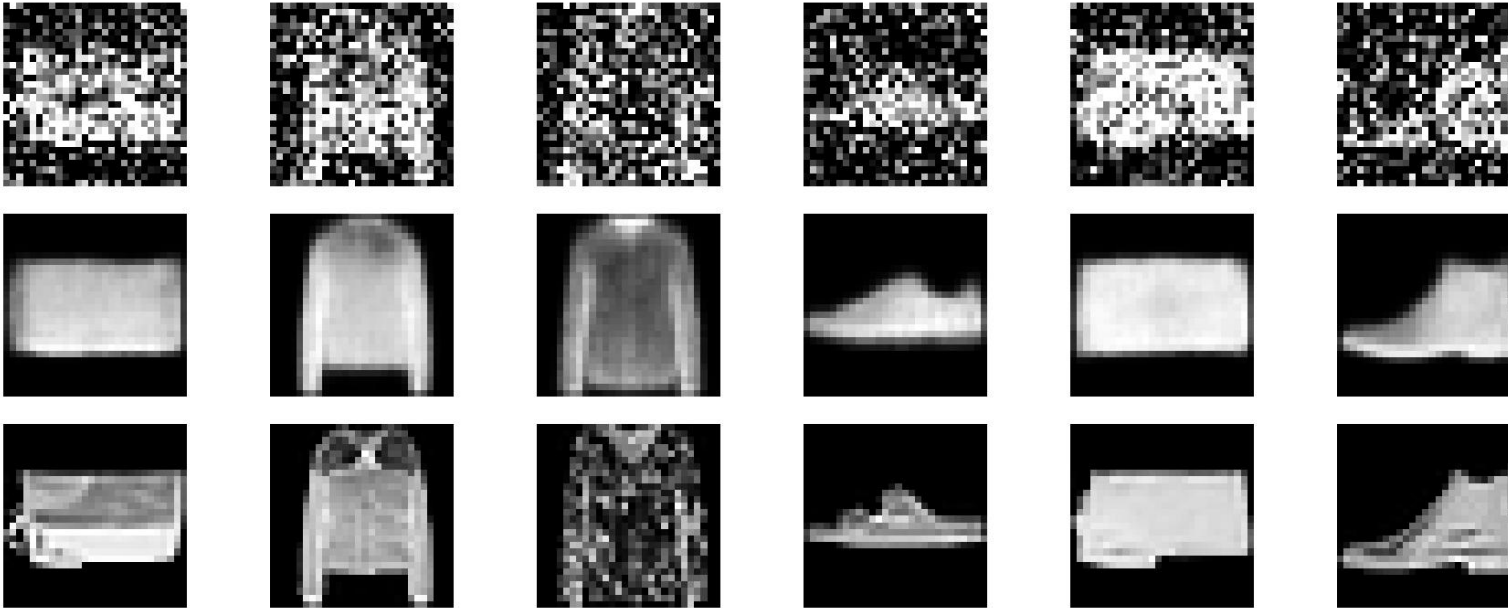
- Encoder: girdi görüntüsünü alıp latent dağılım parametrelerine indirger (z_mean , z_logvar).
- Reparameterization trick ile latent örneği üretilir (z).
- Decoder: z 'den tekrar görüntü uzayına çıkar ve temiz görüntüyü üretmeyi hedefler.
- Denoising'de kritik nokta: giriş gürültülü olsa bile hedefin temiz olmasıdır; bu, decoder'ın "gürültüyü filtreleyen" bir temsil öğrenmesini zorlar.

Kayıp Fonksiyonu ve Gerekçesi

- Toplam amaç üç parçanın dengelenmesidir:
 1. Reconstruction loss: çıktı görüntünün temiz hedefe yakın olmasını sağlar.
 2. KL divergence: latent dağılımı prior'a yaklaştırır (örneklenebilir latent).
 3. Supervised contrastive loss (opsiyonel/ek): latent temsilin sınıf bazlı ayrışmasını güçlendirmeyi hedefler (z_{mean} üzerinde).
- Bu yapı, hem denoising kalitesini hem representation analizini aynı projede ölçülebilir kılar.

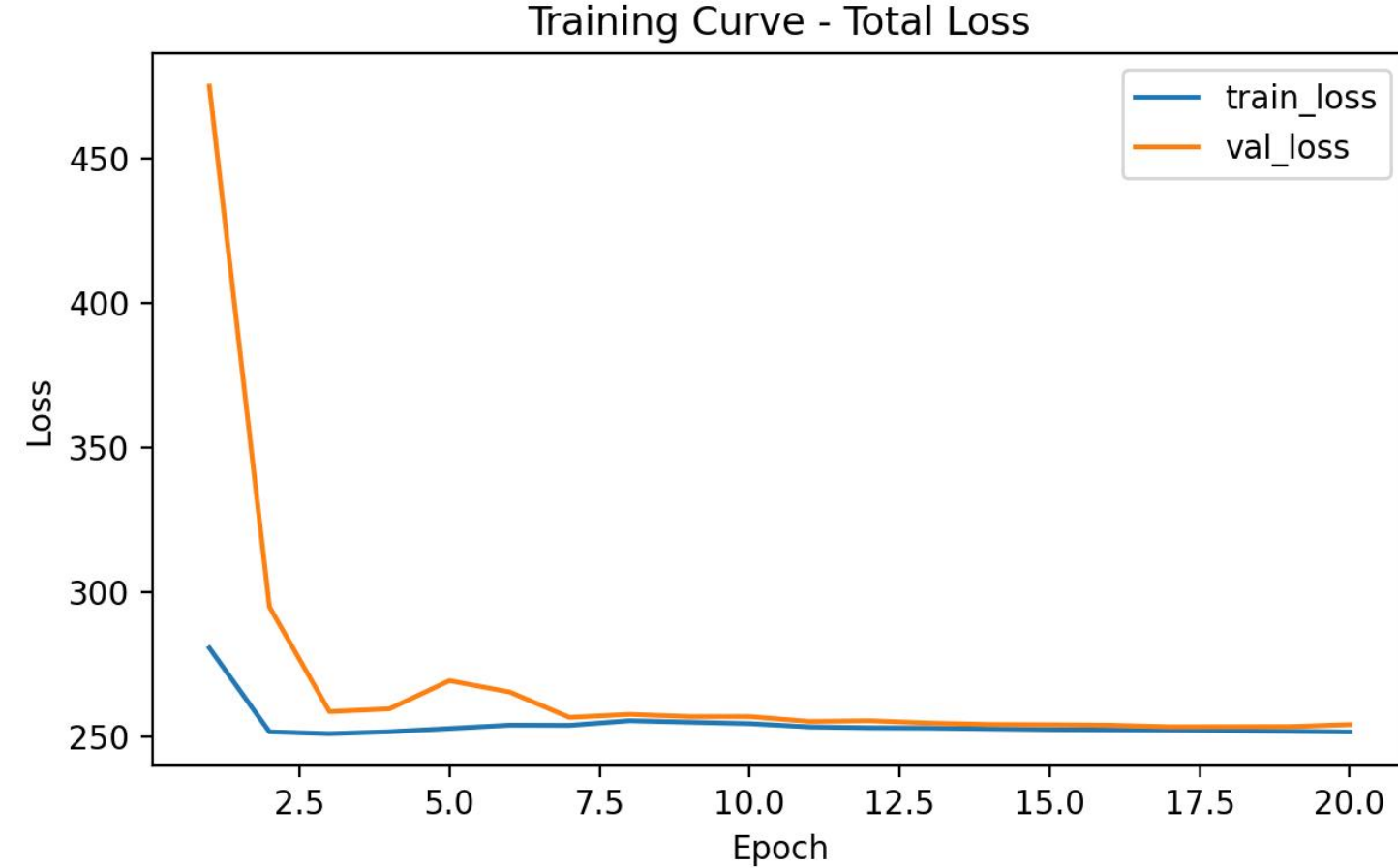
DENOISING BÖLÜMÜ (Giriş)

VAE Denoising (Fashion-MNIST)



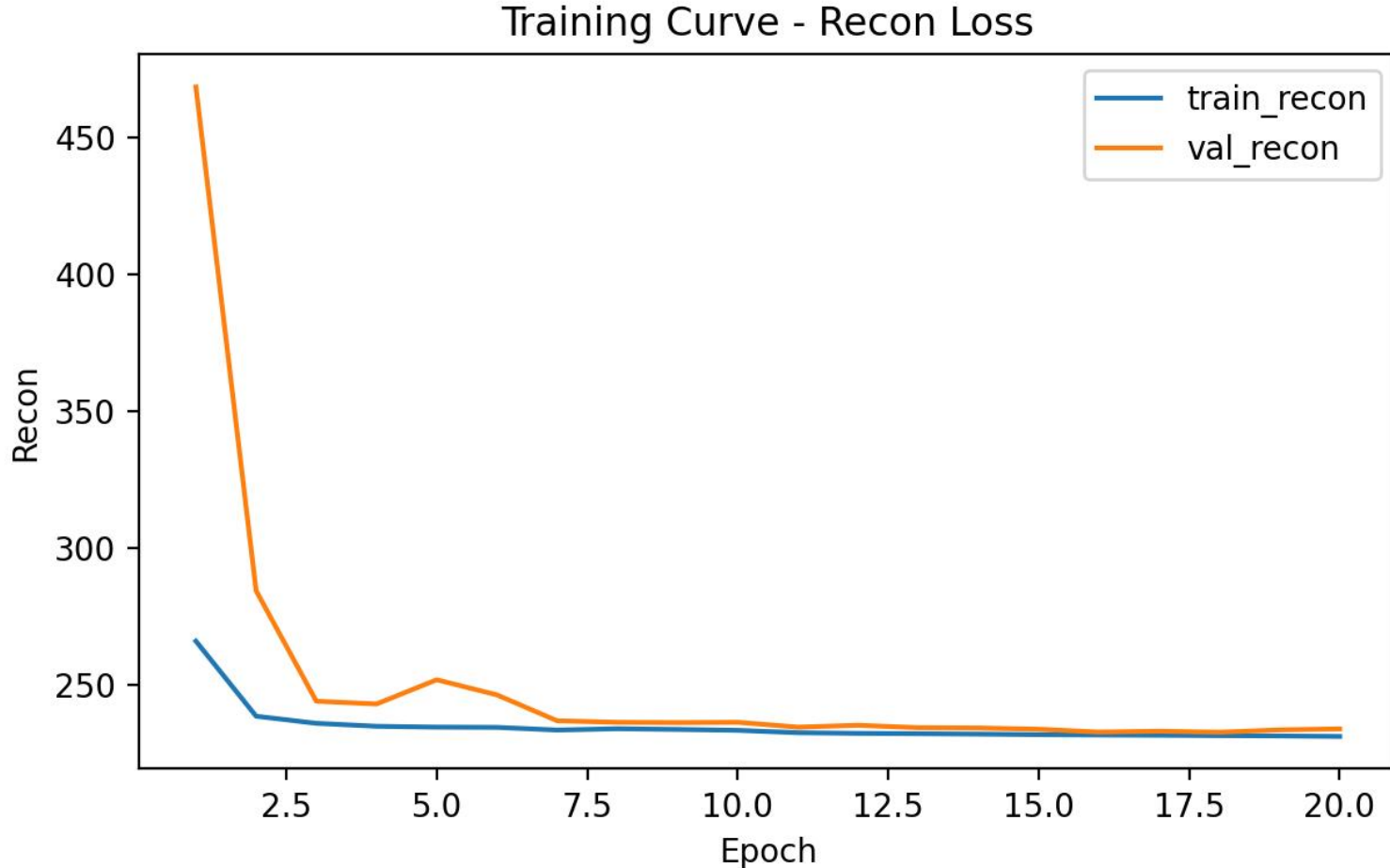
- Bu bölümde VAE'nin “gürültülü görüntüden temiz görüntü üretme” yeteneği hem eğitim eğrileriyle hem de testte nicel + nitel sonuçlarla gösterilir. Amacımız: loss'ların stabil öğrenmeyi işaret etmesi ve test metriklerinde baseline'a göre anlamlı iyileşme görülmesidir.

Eğitim Toplam Kayıp Eğrisi (Stabilite)



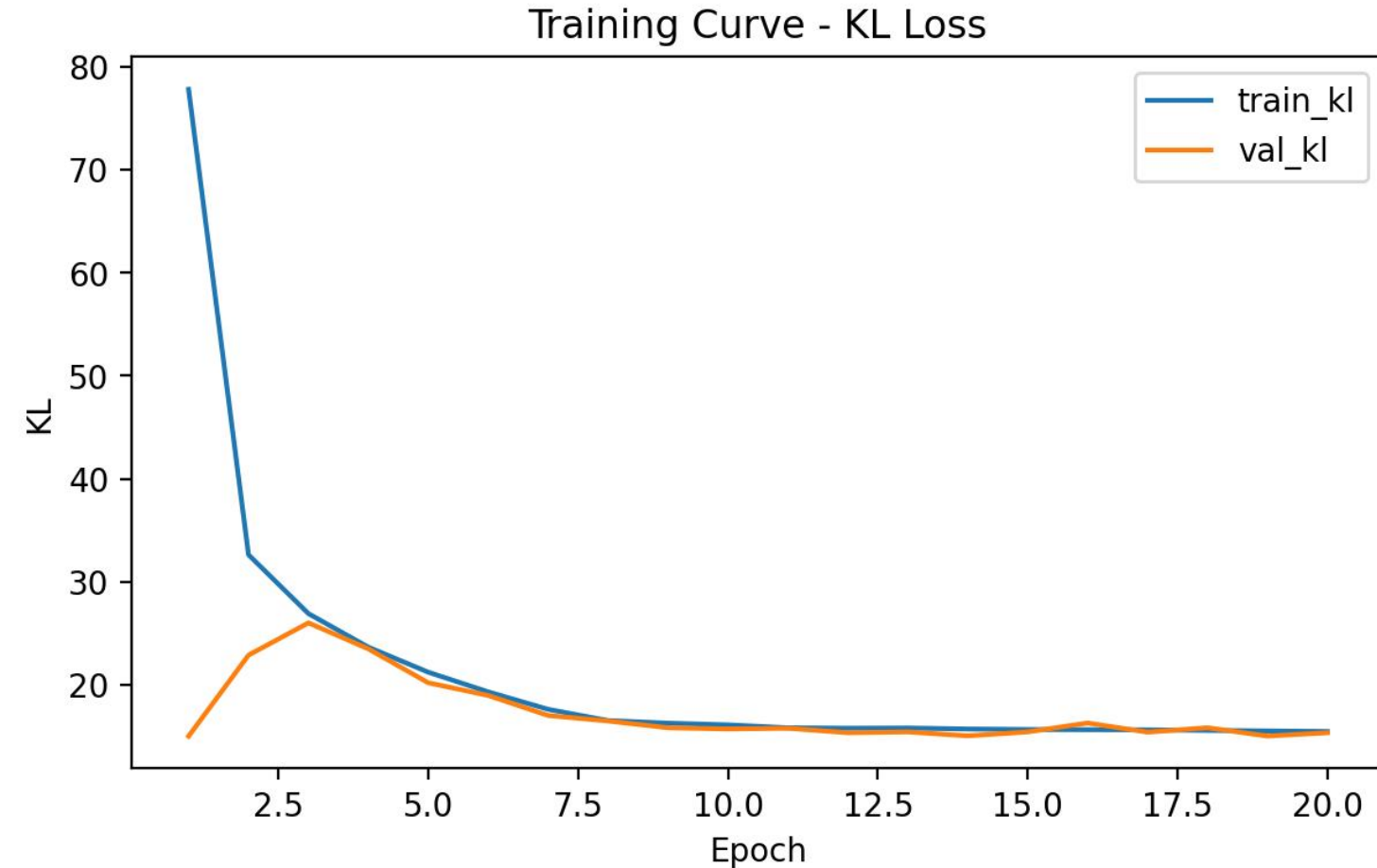
- Toplam loss, modelin tüm hedefleri birlikte optimize ederken nasıl ilerlediğini gösterir. Başta hızlı düşüş; modelin temel rekonstrüksiyon ve latent düzenleme davranışını öğrendiğini, sonlara doğru yavaşlama ise yakınsamayı işaret eder. Eğitim sürecinin dalgalanmadan “oturması”, denoising için güvenilir bir temel oluşturur.

Reconstruction Loss (Gürültüden Temize Yakınsama)



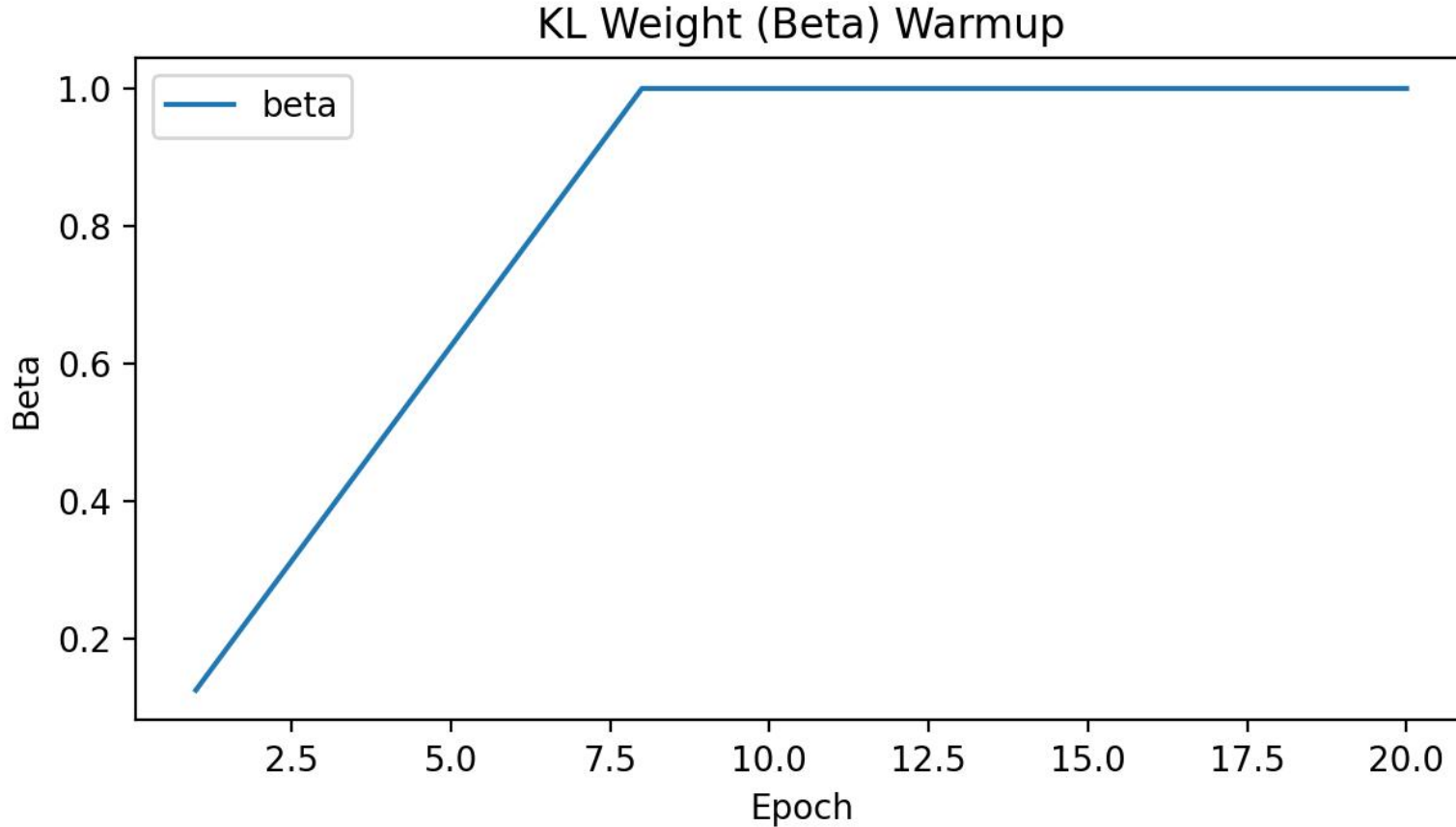
- Reconstruction loss, denoising performansının doğrudan eğitim sinyalidir. Eğrinin düzenli düşmesi; decoder'ın piksel uzayında “temiz hedefe yaklaşmayı” öğrendiğini gösterir. Gürültü faktörü yüksekken (0.45), bu düşüş daha değerlidir çünkü model sadece kopyalama yapamaz; yapısal desenleri öğrenmek zorundadır.

KL Loss (Latent Düzenleme Davranışı)



- KL terimi, latent dağılımının prior'a çok uzaklaşmasını engeller. KL'nin tamamen sıfıra çökmesi "posterior collapse" riskini; aşırı yükselmesi ise rekonstrüksiyonun bozulmasını işaret edebilir. Bu grafik KL'nin eğitim boyunca kontrol altında tutulduğunu ve latent uzayın "örneklenebilir" kalmaya çalıştığını gösterir.

Beta Warm-Up (Dengeyi Kurma Stratejisi)

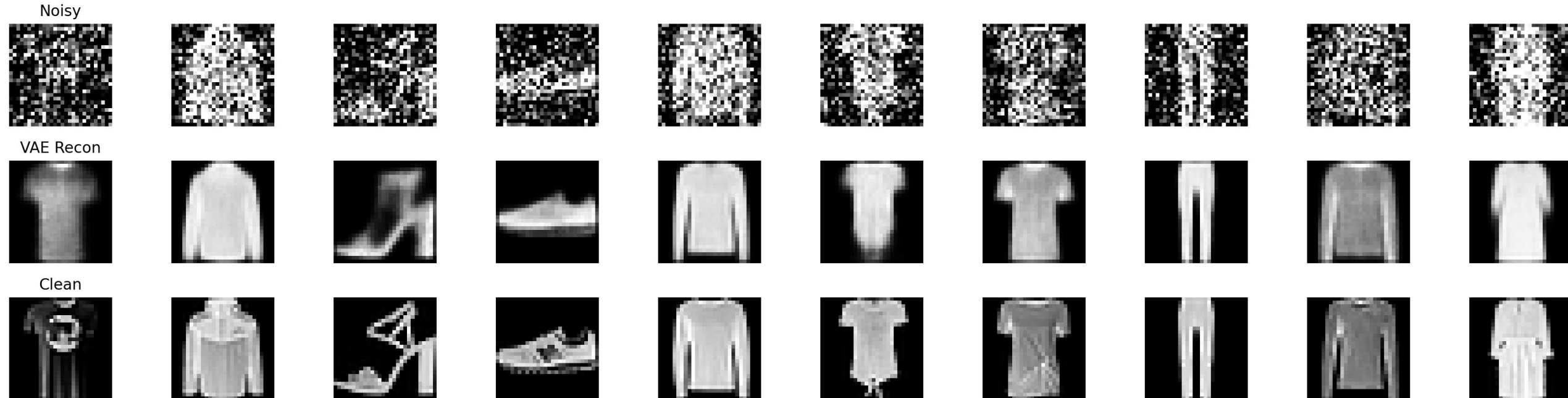


- Beta warm-up; eğitimin başında rekonstrüksiyona ağırlık verip, KL regularizasyonunu kademeli artırma fikridir. Böylece model önce “görüntüyü üretmeyi” öğrenir, sonra latent uzayı düzenlemeye zorlanır. Bu yaklaşım, hem daha iyi denoising hem de daha anlamlı latent yapı üretme açısından pratikte etkilidir.

Denoising Örnekleri (Nitel İference)

- Bu grid, testten örnekler üzerinde üç şeyi aynı anda gösterir: gürültülü giriş, modelin rekonstrüksiyonu ve hedef temiz görüntü. Burada beklenen: VAE çıktısının gürültüyü bastırırken temel formu (kontur, yoğun bölgeler) koruması. Nitel örnekler, metriklerin “gözle görülen” karşılığını sunar.

VAE Denoising (Fashion-MNIST)



Denoising Test Metrikleri (Nicel Kanıt)

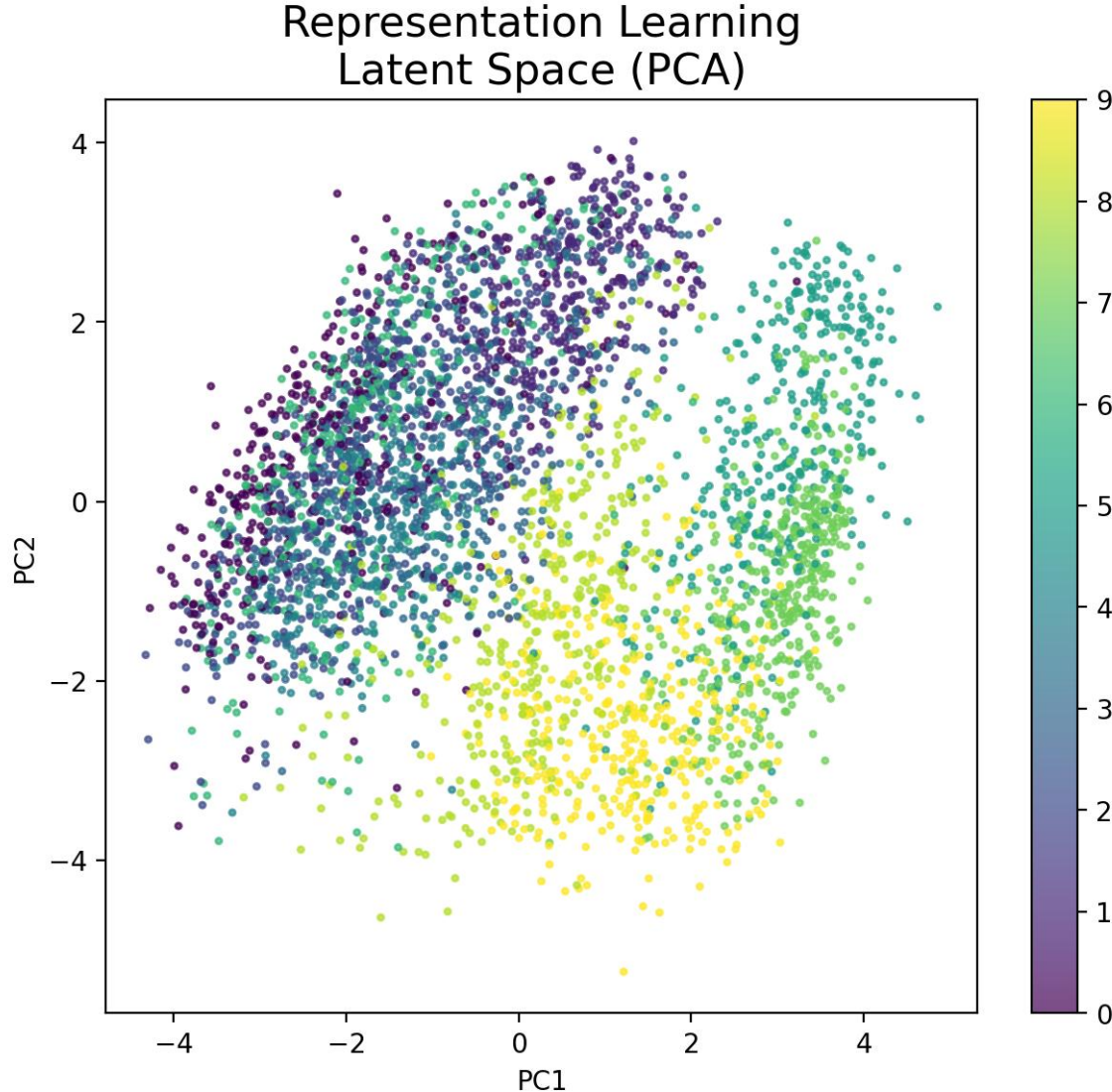
- Test setinde baseline (noisy→clean doğrudan) ile VAE rekonstrüksiyonu karşılaştırılmıştır:
- Baseline: MSE=0.101500 | PSNR=9.945 dB | SSIM=0.2871
- VAE: MSE=0.018064 | PSNR=17.962 dB | SSIM=0.5950
- İyileşme: MSE %82.20 azalma, PSNR +8.017 dB, SSIM +0.3079. Bu; modelin yalnızca “ortalama alma” değil, yapısal benzerliği de ciddi artırdığını gösterir.

```
metric,baseline_noisy_to_clean,vae_recon_to_clean,delta_or_improvement
MSE,0.10150019079446793,0.018063679337501526,82.20330504116836
PSNR_dB,9.945466041564941,17.962064743041992,8.01659870147705
SSIM,0.28713834285736084,0.5950444340705872,0.3079060912132263
```


metrics.txt

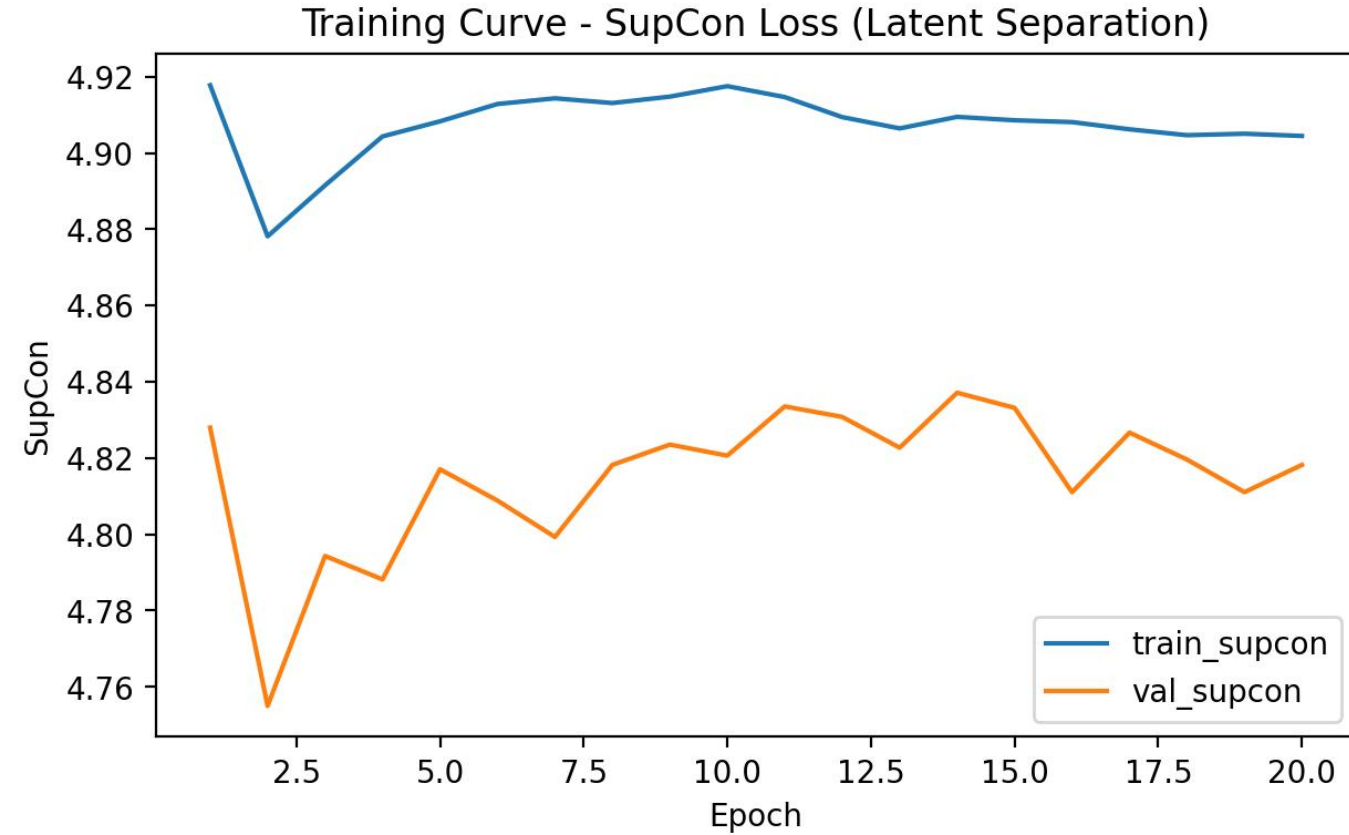
- === Denoising Metrics (Test) ===
- NOISE_FACTOR=0.45
- LATENT_DIM=16
- BATCH_SIZE=256
- EPOCHS=20
- === Improvement ===
- MSE improvement: 82.20%
- PSNR gain : 8.017 dB
- SSIM gain : 0.3079
- Baseline (Noisy->Clean) MSE=0.101500 | PSNR=9.945 dB | SSIM=0.2871
- VAE (Recon->Clean) MSE=0.018064 | PSNR=17.962 dB | SSIM=0.5950

Representation Learning Bölümüne Geçiş



- Denoising iyi olabilir ama “model ne öğrendi?” sorusu representation analizleriyle cevaplanır. Bu bölümde latent uzay; sınıfların kısmi kümelenmesi, ayrışma eğilimleri ve nonlineer indirgeme yöntemleriyle incelenecek. Amaç: VAE’nin latent uzayının rastgele değil, yapı taşıyan bir temsil oluşturduğunu göstermek.

Supervised Contrastive Loss (Latent Ayırışmayı Zorlama)

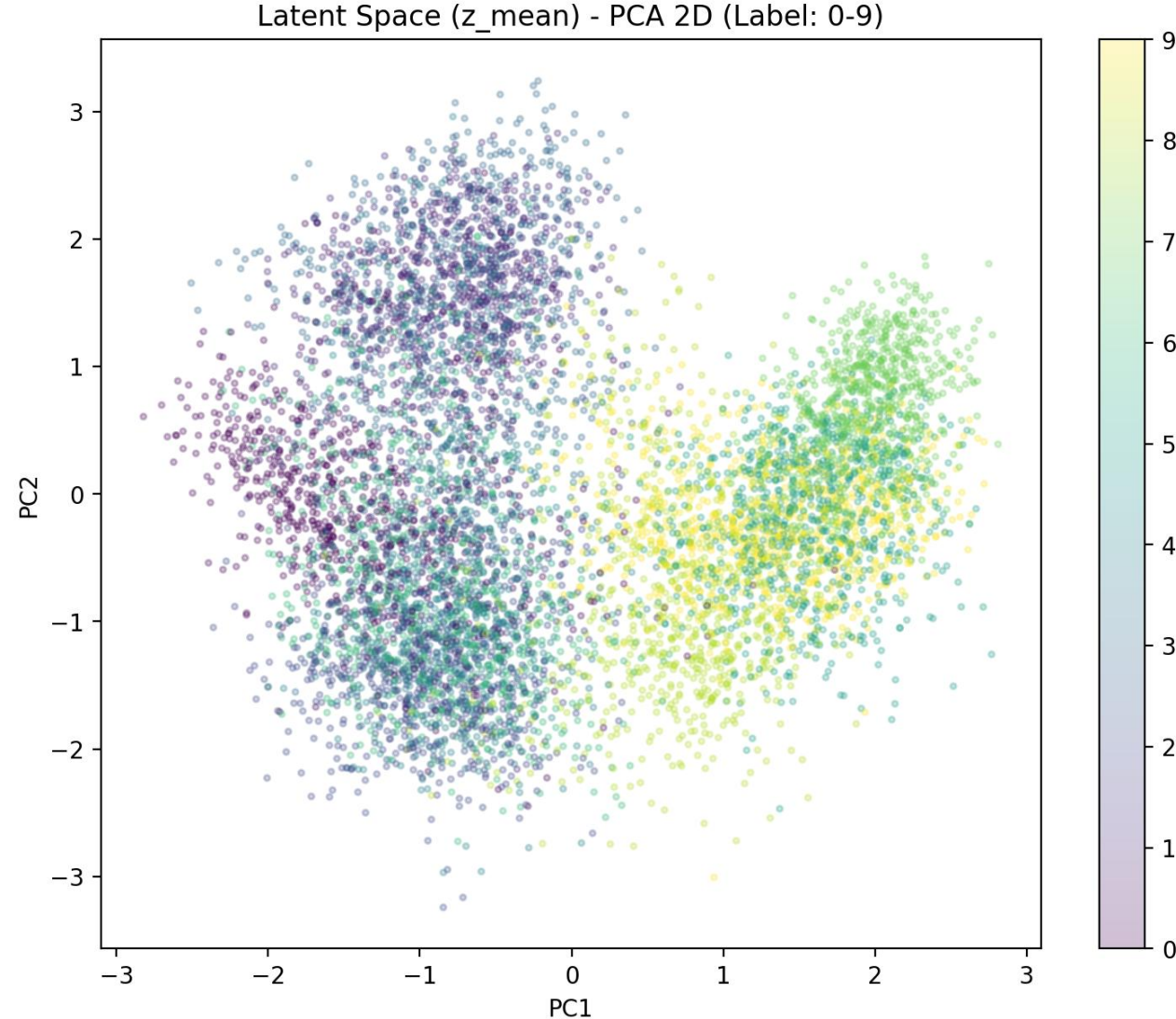


- SupCon loss, aynı sınıftan örneklerin latent uzayda birbirine yaklaşmasını; farklı sınıfların uzaklaşmasını teşvik eder. VAE'nin doğal olarak “tam ayırışma” hedefi yoktur; bu ek terim, representation kalitesini raporlanabilir şekilde güçlendirmek için kullanılır. Eğrinin düşmesi, modelin sınıf farklarını latent uzaya taşımayı öğrendiğini işaret eder.

PCA Varyans Açıklaması (2B'ye Sıkıştırma Sınırı)

- PCA ile latent uzayı 2 boyuta indirirken tüm bilgiyi korumayız; bu yüzden “tam ayrışma yok” diye representation yok sayılmaz. İlk iki bileşenin açıkladığı varyans oranları [0.1424, 0.1166], toplam 0.2590’dır. Yani 2B PCA, latent bilginin yaklaşık %25.9’unu taşır; ayrışma sinyali görülebilir ama sınırlı olması normaldir.
- PCA Explained Variance Ratio (2 components)
- [0.14242163 0.11660043]
- Sum: 0.25902206

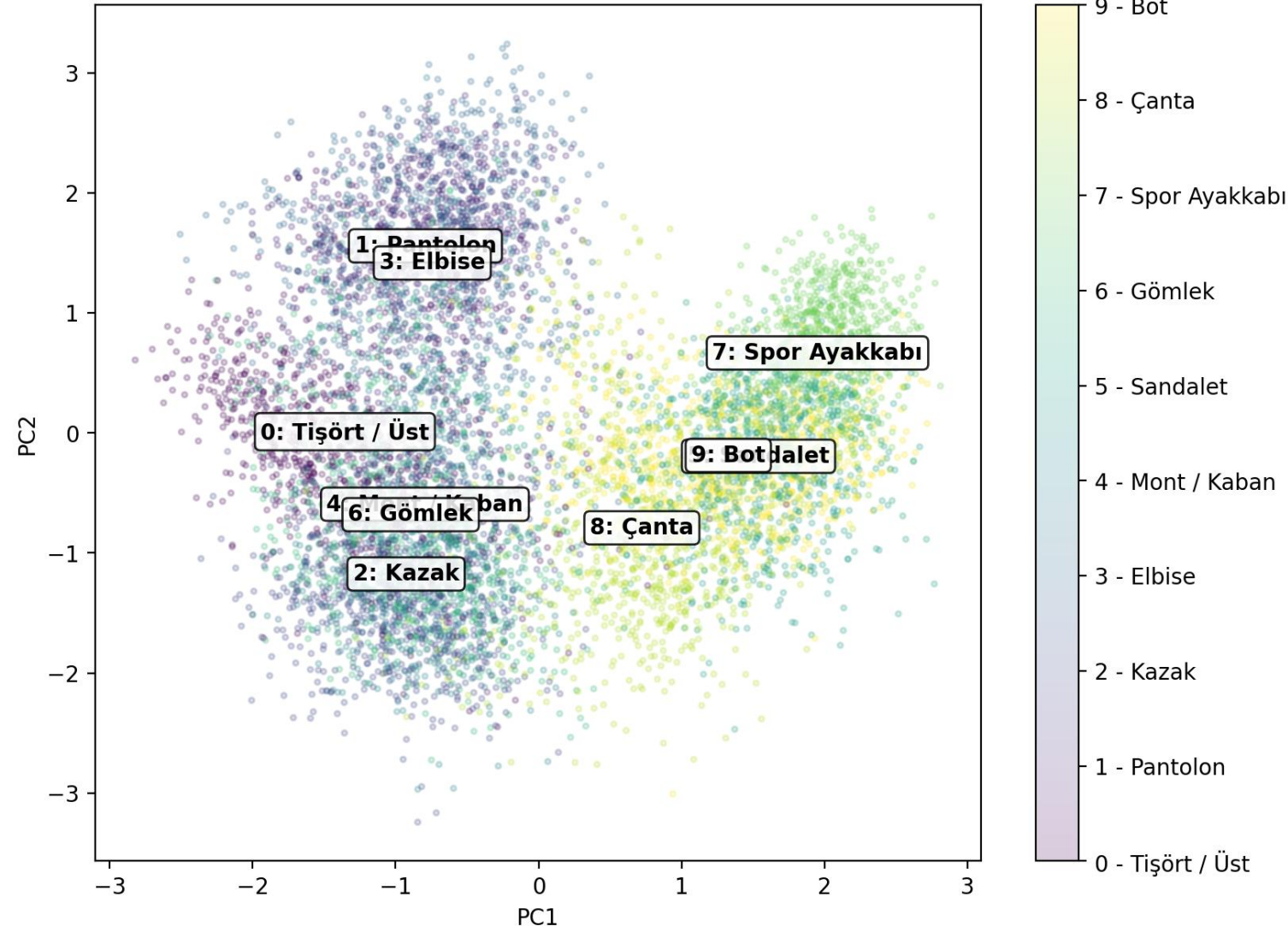
Latent PCA (Sınıf Numaralarıyla Görünüm)



- Bu görsel, latent vektörlerin 2B PCA izdüşümünü ve sınıf etiketlerini (sayısal) gösterir. Beklenen davranış: bazı sınıfların daha sıkı kümelenmesi (ör. benzer kıyafet türleri) ve bazı sınıfların overlap etmesi (görsel olarak benzer sınıflar). Bu, latent uzayın semantik benzerliği kısmen yakalayabildiğine dair ilk işarettir.

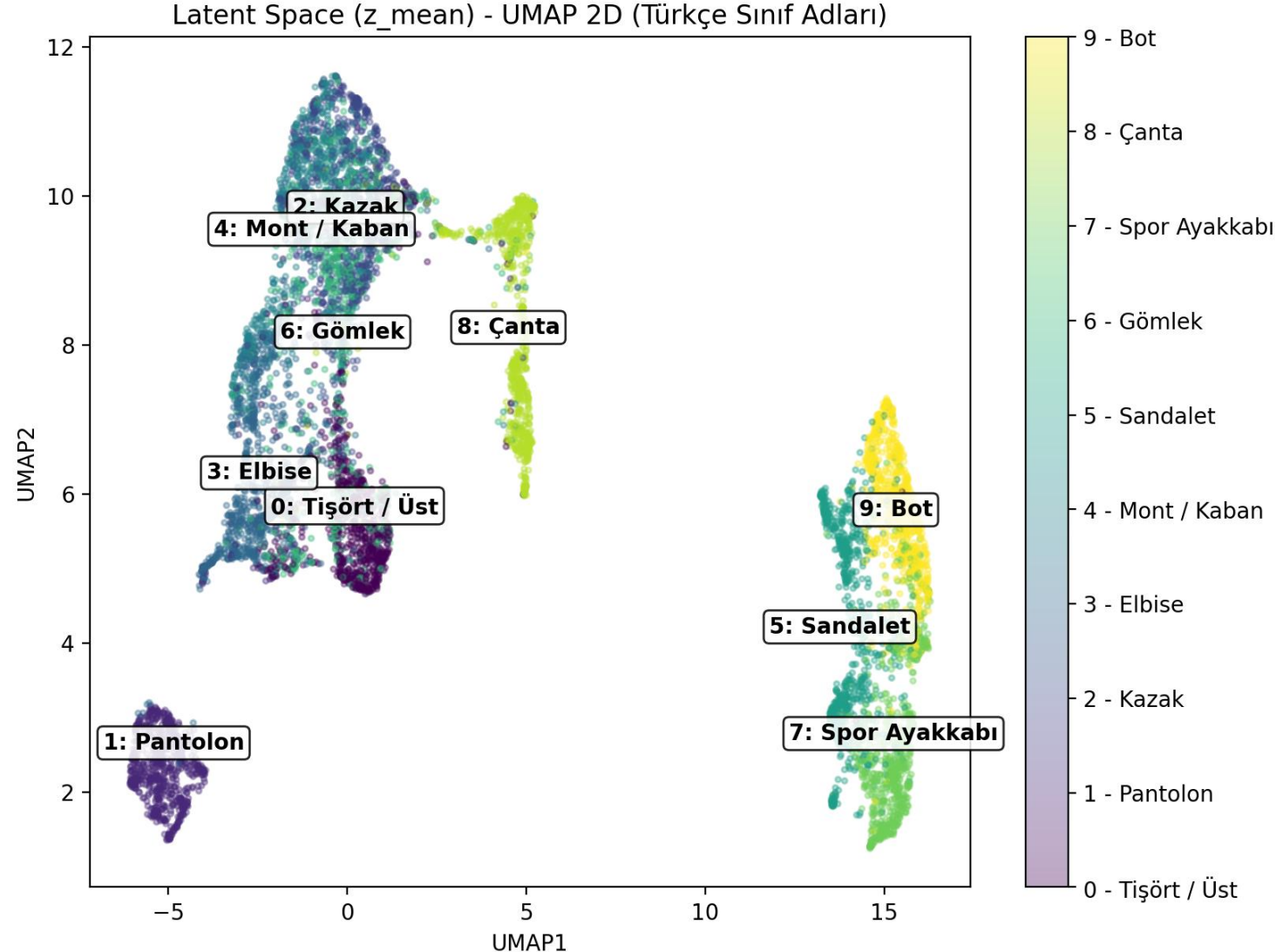
Latent PCA (Türkçe Etiketlerle Yorumlanabilirlik)

Latent Space (z_mean) - PCA 2D (Türkçe Sınıf Adları)



- Sayısal etiket yerine Türkçe sınıf adlarıyla gösterim, yorumlamayı hızlandırır. Burada “yakın duran sınıflar” çoğu zaman görsel özellikleri benzer gruptlardır (ör. üst giyim benzer dokular). Bu slayt, hocaya “model ne öğrendi?” sorusunun daha anlaşılır bir cevabını görsel olarak sunar.

Latent UMAP (Nonlinear Kümelenme Analizi)



- UMAP, yerel komşuluk ilişkilerini koruyarak nonlineer bir gömme üretir; PCA'ya göre kümeleri daha belirgin gösterebilir. Bu görselde sınıflar arası ayrışma sinyalinin güçlenmesi, VAE'nin latent uzayının sadece "sıkıştırma" değil, anlamlı manifold yapısı taşıdığını destekler.

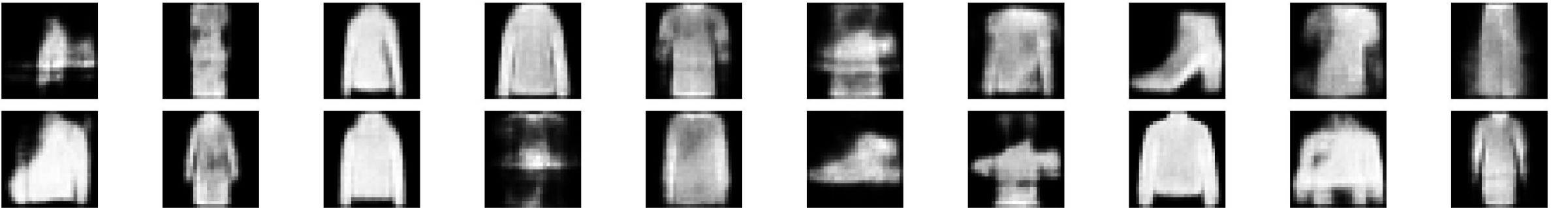
Latent t-SNE (Yakın Komşulukların Vurgusu)

- t-SNE, özellikle yakın komşulukları iyi gösterir ve sınıf kümelerini görsel olarak “ayırt edilebilir” hale getirebilir.
- t-SNE’nin global mesafeleri birebir temsil etmediği unutulmadan yorum yapılır: burada amaç “tam sayısal ayrışma” değil, representation’ın sınıf yapısını taşıyıp taşımadığını görmektir.

GENERATION BÖLÜMÜ (Prior'dan Örnekleme)

- Bu bölümde VAE'nin kritik avantajı gösterilir: latent uzay düzenli olduğu için prior'dan örnekleyip yeni görüntüler üretebiliriz. Denoising başarısı tek başına yeterli değildir; VAE'nin “üretebilen” bir model olduğuna dair kanıt, prior sampling çıktılarıyla verilir.

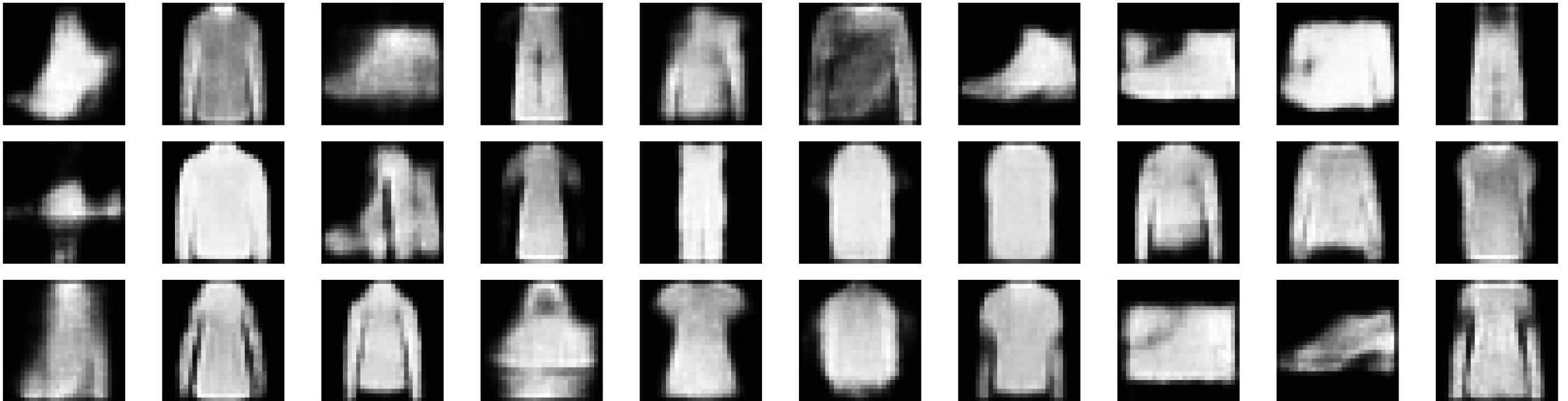
Representation Generation
 $z \sim N(0,1)$



Generated Samples (Prior Sampling Grid)

- Burada latent uzaydan rastgele z örnekleri çekilerek decoder ile görüntü üretilmiştir. Beklenen: üretilen örneklerin Fashion-MNIST'e benzer genel formu taşıması (bulanık ama anlamlı silüetler). Üretim kalitesi, latent düzenleme (KL) ve rekonstrüksiyon dengesinin doğru kurulup kurulmadığına da dolaylı kanıttır.

Generated Samples (Prior $\sim N(0,1)$)



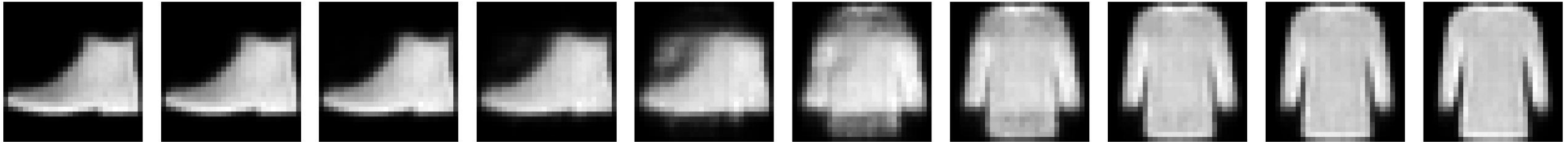
Generation Çıktılarının Yorumlanması (Güçlü/Zayıf Yan)

- VAE çıktıları çoğu zaman GAN'lere göre daha yumuşak/blur görünebilir; bu, olasılıksal rekonstrüksiyon hedefi ve MSE-benzeri kayıplarla ilişkilidir.
- Buna karşın VAE'nin artısı; latent uzayın daha düzenli olması, anlamlı interpolasyonlar ve daha stabil eğitimidir.
- Bu slayt, üretimin “fotogerçekçilik” değil “anlamlı örneklenebilir manifold” kanıtı olduğunu netleştirir.

INTERPOLATION BÖLÜMÜ (Latent Süreklilik)

- Interpolation, latent uzayın sürekliliğini test eder: iki örnek arasında latent uzayda adım adım yürüyüp decoder'dan görüntü alırız.
- Eğer geçişler yumuşak ve semantik olarak tutarlıysa, modelin latent uzayı “boşluklu değil”, süreklilik taşıyan bir temsil öğrenmiştir.

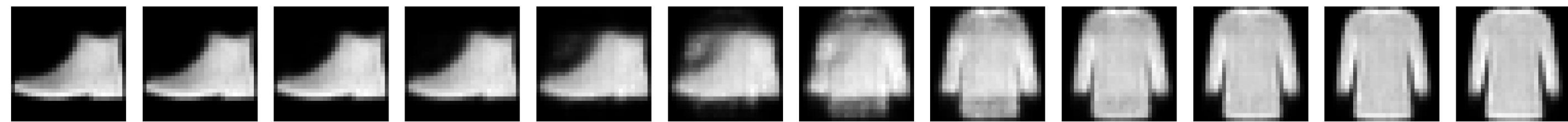
Latent Interpolation ($z_A \rightarrow z_B$)



Latent Interpolation Grid (Süreklilik Kanıtı)

- Bu grid; iki görüntü arasındaki geçişin “aniden sıçramadan” dönüşmesini bekler. Örneğin bir sınıfa benzer formdan diğerine benzer forma kademeli yaklaşma görülüyorsa, latent uzayda ara noktaların da anlamlı olduğu anlaşılır. Bu, generation bölümünü destekleyen en güçlü görsel kanıtlardan biridir.

Latent Interpolation ($z_A \rightarrow z_B$) | $i_1=892$, $i_2=7739$



Eğitim Kayıtları ve Tekrar Üretilebilirlik (History Log)

- Eğitim sırasında loss/metric geçmişisi dosyalanır. Böylece aynı eğitim koşusu; epoch bazında izlenebilir, grafikler yeniden çizilebilir ve raporla birebir örtüşür. Bu slayt, eğitim loglarının repo içinde saklandığını ve sunumdaki eğrilerin kaynağının şeffaf olduğunu gösterir.

```
epoch,beta,gamma_sc,kl_loss,loss,recon_loss,supcon_loss,val_beta,val_gamma_sc,val_kl_loss,val_loss,val_recon_loss,val_supcon_loss,learning_rate
1,0.125,1.0,77.82079315185547,280.66961669921875,266.02435302734375,4.9177470207214355,0.125,1.0,15.010518074035645,474.98992919921875,468.28570556640625,4.82796049118042,0.0010000000474974513
2,0.25,1.0,32.626590728759766,251.65557861328125,238.6208953857422,4.878142833709717,0.25,1.0,22.882312774658203,294.8927307128906,284.4171142578125,4.755013942718506,0.0010000000474974513
3,0.375,1.0,26.912078857421875,251.0050506591797,236.02154541015625,4.891499996185303,0.375,1.0,26.016101837158203,258.6485900878906,244.0982208251953,4.794314861297607,0.0010000000474974513
4,0.5,1.0,23.6556339263916,251.6758270263672,234.9436492919922,4.904326915740967,0.5,1.0,23.487516403198242,259.617431640625,243.0854949951172,4.788168907165527,0.0010000000474974513
5,0.625,1.0,21.222518920898438,252.809326171875,234.6371307373047,4.908282279968262,0.625,1.0,20.195043563842773,269.35858154296875,251.91964721679688,4.817055702209473,0.0010000000474974513
6,0.75,1.0,19.31583023071289,253.94676208496094,234.5470428466797,4.912827491760254,0.75,1.0,18.96605682373047,265.4390869140625,246.4056854248047,4.808846473693848,0.0010000000474974513
7,0.875,1.0,17.618484497070312,253.8922119140625,233.56163024902344,4.9143171310424805,0.875,1.0,17.019672393798828,256.64501953125,236.9535369873047,4.799304485321045,0.0005000000237487257
8,1.0,1.0,16.52889060974121,255.49166870117188,234.04974365234375,4.913078784942627,1.0,1.0,16.488868713378906,257.69940185546875,236.39236450195312,4.818195819854736,0.0005000000237487257
9,1.0,1.0,16.278079986572266,254.9777069091797,233.78488159179688,4.914754390716553,1.0,1.0,15.826177597045898,256.93359375,236.2838897705078,4.823521614074707,0.0005000000237487257
10,1.0,1.0,16.110198974609375,254.49708557128906,233.46925354003906,4.917503833770752,1.0,1.0,15.717117309570312,256.93359375,236.3958282470703,4.820627689361572,0.0005000000237487257
11,1.0,1.0,15.827923774719238,253.34671020507812,232.60411071777344,4.914643287658691,1.0,1.0,15.785748481750488,255.24765014648438,234.62841796875,4.833531379699707,0.0002500000118743628
12,1.0,1.0,15.786408424377441,253.05250549316406,232.3567657470703,4.909390926361084,1.0,1.0,15.346197128295898,255.4827880859375,235.3058319091797,4.830779075622559,0.0002500000118743628
13,1.0,1.0,15.806422233581543,252.96072387695312,232.24783325195312,4.906424522399902,1.0,1.0,15.426411628723145,254.6839141845703,234.43478393554688,4.822707653045654,0.0002500000118743628
14,1.0,1.0,15.71179485321045,252.71011352539062,232.08885192871094,4.9094557762146,1.0,1.0,15.048365592956543,254.2062530517578,234.3207244873047,4.837116241455078,0.0002500000118743628
15,1.0,1.0,15.662786483764648,252.44862365722656,231.87733459472656,4.9085564613342285,1.0,1.0,15.42597484588623,254.11581420898438,233.856689453125,4.833155632019043,0.0002500000118743628
16,1.0,1.0,15.645011901855469,252.3101806640625,231.75698852539062,4.908082962036133,1.0,1.0,16.28913688659668,253.9423828125,232.84219360351562,4.811046600341797,0.0002500000118743628
17,1.0,1.0,15.629648208618164,252.19883728027344,231.6629180908203,4.906189441680908,1.0,1.0,15.404886245727539,253.39157104492188,233.1600799560547,4.826654434204102,0.0002500000118743628
18,1.0,1.0,15.574337005615234,252.00088500976562,231.52183532714844,4.904648780822754,1.0,1.0,15.833056449890137,253.42825317382812,232.7755584716797,4.819612979888916,0.0002500000118743628
19,1.0,1.0,15.520181655883789,251.8400115966797,231.4147491455078,4.905034065246582,1.0,1.0,15.025001525878906,253.4567108154297,233.6206512451172,4.811042308807373,0.0002500000118743628
20,1.0,1.0,15.470368385314941,251.61883544921875,231.2440643310547,4.904445171356201,1.0,1.0,15.348835945129395,254.1532440185547,233.98617553710938,4.818184852600098,0.0002500000118743628
```

Sınırlılıklar ve İyileştirme Fikirleri

- Bu çalışma tek koşuda güçlü bir temel kanıt sunar; ancak daha ileri için: latent_dim taraması (8/16/32), farklı noise_factor seviyeleri, decoder kapasitesini artırma, perceptual loss/SSIM-ağırlıklı loss denemeleri, beta schedule alternatifleri ve daha güçlü representasyon için supervised/metric learning ağırlık ayarı yapılabilir.
- Ayrıca generated örneklerin keskinliği için VAE-GAN hibritleri veya daha gelişmiş decoder kayıpları denenebilir.

Sonuç

- Bu projede VAE'nin iki yönü aynı çatı altında kanıtlandı:
 1. Denoising'de hem nitel hem nicel olarak güçlü iyileşme,
 2. Latent uzayda sınıf yapısını kısmen taşıyan representation ve buradan prior sampling + interpolation ile üretim.