

Sağlık Görüntülerinde farklı Sentetik Veri Üretimi Yaklaşımlarının Denenmesi

1- Amaç ve Genel Bilgiler

Bu raporun amacı, özellikle sağlık görüntüleri alanında, farklı sentetik veri üretim tekniklerini incelemek ve analiz etmektir. Sentetik veri, maliyet, gizlilik endişeleri ve veri çeşitliliğine duyulan ihtiyaç gibi sınırlamaların üstesinden gelmek için tıbbi araştırmalarda giderek daha fazla kullanılmaktadır. Özellikle, bu rapor, zatürre, tüberküloz ve COVID-19 gibi çeşitli akciğer hastalıklarının teşhisi için kritik öneme sahip sentetik göğüs röntgen görüntüleri üretmeye odaklanmaktadır.

Bunu başarmak için üç güçlü makine öğrenmesi modeli eğitildi: Üretici Çekişmeli Ağlar (Generative Adversarial Networks - GANs), Değişken Otokodlayıcılar (Variational Autoencoders - VAEs) ve Gürültü Giderici Difüzyon Olasılıksal Modelleri (Denoising Diffusion Probabilistic Models - DDPMs). Bu modeller, mevcut veri dağılımlarını öğrenerek orijinal verilere istatistiksel olarak benzeyen yeni örnekler üreten, gerçekçi görüntüler oluşturmak için yaygın olarak kullanılmaktadır.

2- Veri Seti Hakkında Bilgi

Bu proje için kullanılan veri seti, **Chest X-Ray Images (Pneumonia)** başlıklı bir Kaggle veri setinden alınmıştır. Bu veri seti, toplamda 5.216 adet göğüs röntgen görüntüsünü içermekte olup, iki sınıfa ayrılmıştır: **Normal** (sağlıklı akciğerler) ve **Pneumonia** (enfekte akciğerler). Veri seti, özellikle sağlık alanında modellerin eğitilmesi için ve pnömونيyi tespit edebilecek sistemlerin geliştirilmesi için son derece faydalıdır.

Ön İşleme Adımları: Veri seti, modellerin mimarileri ile uyumluluğunu ve birliğini sağlamak için eğitilmeden önce birkaç ön işleme adımına tabi tutulmuştur:

1. **Yeniden Boyutlandırma:** Tüm görüntüler **64x64 piksel** boyutlarına yeniden boyutlandırıldı. Bu boyut küçültme, hesaplama verimliliğini optimize etmek ve eğitim süresini hızlandırmak için gereklidi; görüntü kalitesini önemli ölçüde düşürmeden bu sağlandı.
2. **Gri Tonlamaya Dönüştürme:** Orijinal göğüs röntgen görüntüleri, RGB (renkli) formatından **gri tonlama** formatına dönüştürüldü. X-ray gibi tıbbi görüntülerde ilgili bilgiler genellikle renkten ziyade kontrast tabanlı olduğu için, bu adım veriyi basitleştirirken temel özelliklerin korunmasını sağladı.
3. **Etiketleme:** Veri seti iki sınıfa ayrılmıştır:
 - **Sınıf 0: Normal** göğüs röntgenlerini (sağlıklı bireyler) temsil eder.
 - **Sınıf 1: Pnömoni** göğüs röntgenlerini (pnömوني teşhisi konmuş hastalar) temsil eder.

Veri Seti Özeti:

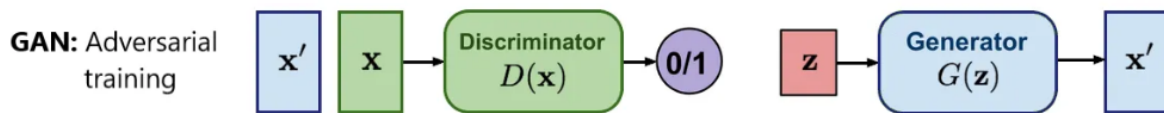
- **Toplam Görüntü Sayısı:** 5.216
- **Normal (Sınıf 0):** 1.341 görüntü
- **Pnömoni (Sınıf 1):** 3.875 görüntü
- **Görüntü Boyutları:** 64x64 piksel (yeniden boyutlandırmadan sonra)
- **Renk Modu:** Gri tonlama (tek kanal)

3- Modeller ve Mimarileri Hakkında Genel Bilgi

3.1 Generative Adversarial Networks (GANs)

GAN'lar, iki rekabetçi ağı eğitildiği bir tür sinir ağı sınıfıdır: **Üretici (Generator)** ve **Ayırt Edici (Discriminator)**. Üretici, sentetik görüntüler üretirken, ayırt edici, bu görüntülerin gerçek mi (eğitim veri setinden mi) yoksa sahte mi (üreticiden mi) olduğunu değerlendirir.

- **Mimari:**
 - **Generator:** Rastgele bir gürültüyü (gizli vektör) girdi olarak alır ve upsampling (örnekleme) ve convolution (evrişim) katmanlarından geçirerek bir görüntü üretir.
 - **Discriminator:** Bir görüntüyü girdi olarak alır ve convolution katmanlarıyla gerçek mi yoksa sahte mi olduğunu sınıflandırma katmanında tahmin eder.
- **Nasıl Çalışır:** Eğitim sırasında, üretici ayırt ediciyi yanıltmaya çalışarak kendini geliştirirken, ayırt edici ise gerçek ve sahte görüntüleri ayırt etmeye çalışarak gelişir. Bu çekişmeli süreç, üreticinin ayırt ediciyi yanıltacak kadar gerçekçi görüntüler üretebildiği noktaya kadar devam eder.

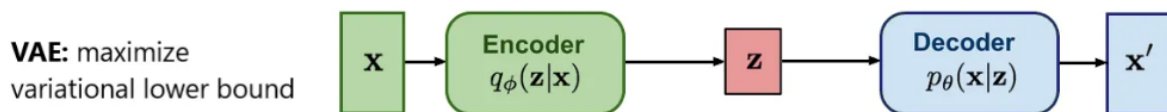


3.2 Variational Autoencoders (VAEs)

VAEs, yeni veri üretmek için kullanılan bir otokodlayıcı türüdür. İki ana parçadan oluşur:

Kodlayıcı (Encoder) ve **Kod Çözücü (Decoder)**. Kodlayıcı, giriş verisini gizli uzaya haritalarken, kod çözücü bu gizli uzay temsilinden veriyi yeniden oluşturur.

- **Mimari:**
 - **Encoder:** Giriş görüntüsünü, evrişimsel ve tam bağlı katmanlardan geçirerek düşük boyutlu gizli uzaya sıkıştıran bir sinir ağıdır.
 - **Decoder:** Gizli uzay temsilinden görüntüyü yeniden oluşturan ve örnekleme ile dekonvolüsyon katmanlarını kullanan bir sinir ağıdır.
- **Nasıl Çalışır:** VAEs, giriş görüntüsünü gizli uzayda sabit bir noktaya değil, ortalama ve varyans olarak kodlayarak olasılıksal bir öge ekler. Bu, modelin gizli uzaydan örnekleme yaparak giriş görüntüsünün farklı varyasyonlarını üretmesini sağlar ve çeşitlilik ile gerçekçiliği artırır.



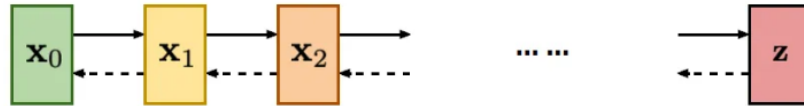
3.3 Denoising Diffusion Probabilistic Models

DDPM'ler, görüntüleri başlangıçta gürültülü bir görüntüden kademeli olarak gürültüyü kaldırarak üreten yeni bir modelleme yaklaşımıdır. Rastgele bir gürültülü görüntünün kademeli olarak tersine çevrilme sürecini öğrenirler.

- **Mimari:**
 - **Forward Process:** Bu süreç, bir görüntüye çok sayıda adımda küçük miktarlarda gürültü ekler ve görüntü tamamen tanınmaz hale gelene kadar (saf gürültü) devam eder.
 - **Backward Process:** Bir sinir ağı, her adımda eklenen gürültüyü tersine çevirmek ve orijinal görüntüyü gürültülü versiyonundan kademeli olarak yeniden oluşturmak için eğitilir.
- **Nasıl Çalışır:** DDPM'ler, bir görüntüyü adım adım gürültüsüzleştirmeyi öğrenir. Eğitim sırasında model, gürültünün nasıl eklendiğini ve kaldırıldığını öğrenir, bu sayede rastgele bir gürültü ile başlayarak difüzyon sürecini tersine çevirip yeni görüntüler üretebilir.

Diffusion models:

Gradually add Gaussian noise and then reverse



4- Model Parametreleri

Bu projede kullanılan modeller, sentetik göğüs röntgen görüntülerinin etkili ve verimli bir şekilde üretilmesini sağlamak amacıyla dikkatlice ayarlanan belirli parametrelerle eğitilmiştir. Aşağıda her model için kullanılan parametrelerin detayları verilmiştir.

4.1 GAN

- Batch Size: 32
- Epochs: 62.000
- Latent Space Dimension: 100
- Learning Rate: 0.00002
- Weight Decay: 6e-9
-

4.2 VAE

- Batch Size: 32

- Epochs: 350
- Latent Space Dimension: 16
- Learning Rate: 1e-3

4.3 DDPM

- Batch Size: 64
- Epochs: 50
- Learning Rate: 0.001
- Number of Diffusion Steps: 1.000

Epochs

Epochs, modelin eğitim sırasında tüm veri setinden kaç kez geçtiğini gösterir. Her modelin öğrenme sürecinin karmaşıklığına bağlı olarak farklı sayıda dönem gereklidir:

- GAN (62.000 epochs): GAN'lar, üretici ile ayırt edici arasında rekabetçi bir eğitim sürecine dayanır. Her iki ağın da kademeli olarak gelişmesi gerekir, bu da eğitim sürecini daha yavaş hale getirir ve yüksek sayıda dönemi zorunlu kılar. Üretici, ayırt ediciyi kandıracak görüntüler oluşturmak için birçok yinelemeye ihtiyaç duyar ve ayırt edici geliştikçe üreticinin çıktıları daha da incelemesi gerekir. Bu nedenle, modelin yakınsaması ve dengelenmesi için büyük bir dönem sayısı gereklidir.
- VAE (350 dönem): VAEs, daha az rekabetçi ve daha basit yapıda oldukları için daha az dönem gerektirir. Kodlayıcı ve kod çözücü, GAN'ların rekabetçi eğitimine kıyasla daha basit bir yeniden yapılandırma görevini optimize eder. Bu nedenle, VAE daha hızlı yakınsar ve iyi performansa ulaşmak için daha az dönem yeterlidir.
- DDPM (50 dönem): DDPM'ler, adım adım gürültü giderme süreciyle eğitilir ve bu süreç, birkaç yineleme boyunca gürültü ekler ve ardından bunu tersine çevirmeyi öğrenir. Gürültü giderme adımları (bu durumda 1.000) hesaplama açısından yoğun olduğundan, model, GAN'lar veya VAE'lere kıyasla daha az dönemle yüksek kaliteli görüntüler oluşturmayı öğrenebilir. Her dönem, binlerce gürültü giderme adımından oluşur, bu yüzden daha az dönem eğitim için yeterlidir.

2. Gizli Uzak Boyutları (Latent Space Dimensions)

Gizli uzak boyutu, verinin içsel temsilinin ne kadar sıkıştırılmış veya soyut olduğunu belirler. Farklı modeller, veriyi nasıl işlediklerine ve yeni örnekler oluşturduklarına bağlı olarak farklı gizli uzak boyutlarına ihtiyaç duyar:

- GAN (latent space = 100): GAN'lar, göğüs röntgen görüntülerinin karmaşık dağılımını yakalamak için nispeten büyük bir gizli uzak kullanır. Daha büyük bir gizli uzak, üreticinin daha çeşitli ve yüksek kaliteli görüntüler üretmesine olanak tanır, bu da tıbbi görüntülerdeki ince farklılıkları öğrenmek için gereken esnekliği sağlar.
- VAE (latent space = 16): VAEs, görüntüleri doğrudan daha düşük boyutlu bir uzaya kodladıkları için genellikle daha küçük bir gizli uzak kullanır. Amaç, sadece görüntü

bilgisini sıkıştırmak değil, aynı zamanda yeni örnekler oluşturmak için olasılıksal bir yapı da sağlamaktır. Daha küçük bir gizli uzay, VAE'yi temel özelliklere odaklanmaya zorlar ve aynı zamanda modeli daha hesaplama açısından verimli hale getirir.

- DDPM (açık bir gizli uzay yok): DDPM'ler, geleneksel bir gizli uzaya ihtiyaç duymazlar. Bunun yerine, rastgele gürültüden nihai görüntüye kademeli bir gürültü giderme işlemi gerçekleştirirler. Gürültü giderme süreci doğrudan görüntü üzerinde çalıştığı için, GAN'lar veya VAE'lerde olduğu gibi bir gizli uzay gerekmez, ancak model difüzyon adımları aracılığıyla karmaşık bir temsil öğrenir.

3. Öğrenme Hızları (Learning Rates)

Öğrenme hızı, modelin eğitim sırasında yaptığı hatalar doğrultusunda parametrelerini ne kadar hızlı güncellediğini kontrol eder. Farklı modeller, istikrarlı ve etkili öğrenmeyi sağlamak için farklı öğrenme hızlarına ihtiyaç duyar:

- GAN (öğrenme hızı = 0.00002): GAN'ların eğitimi genellikle zordur çünkü üretici ve ayırt edici arasındaki denge korunmalıdır. Küçük bir öğrenme hızı, eğitim sırasında büyük güncellemeler yapılmasını önleyerek bir ağın diğerini ezmesine neden olabilecek kararsızlığı önler. Yavaş ve sabırlı bir öğrenme süreci, GAN'ın yakınsama sağlarken üretici ve ayırt edici arasındaki dengenin korunmasına yardımcı olur.
- VAE (öğrenme hızı = 0.001): VAEs, GAN'lara göre genellikle daha kolay eğitilir, bu yüzden daha yüksek bir öğrenme hızı uygundur. VAE'nin optimizasyon süreci, rekabet eden ağlar içermediğinden daha istikrarlıdır. Bu da model ağırlıklarının daha hızlı güncellenmesini sağlar ve modelin daha hızlı yakınsamasına olanak tanır.
- DDPM (öğrenme hızı = 0.001): DDPM'ler de nispeten yüksek bir öğrenme hızı kullanır çünkü rekabetçi bir süreç yerine sürekli bir gürültü giderme sürecine dayanırlar. Gürültü giderme adımları daha istikrarlı olduğundan, model daha yüksek bir öğrenme hızıyla etkili bir şekilde öğrenebilir. Bu, özellikle yüksek difüzyon adım sayısı göz önüne alındığında daha hızlı eğitim sağlar.