

DERİN ÖĞRENMENİN ÜRETKEN MODELLERİ – ÜRETKEN ÇEKİŞMELİ AĞLAR (GANLAR) VE VARYASYONLU OTOKODLAYICI (VAE)

GENERATIVE MODELS OF DEEP LEARNING – GENERATIVE ADVERSARIAL NETWORKS (GANS) AND VARIATIONAL AUTOENCODER (VAE)

Ömer Faruk EREKEN*
Çiğdem TARHAN**

Öz

Derin öğrenme ile gerçekçi görsel üretimi popüler yapay zeka çalışmalarından bir tanesidir. Üretken modellerin geliştirilmesi konusunda her geçen gün daha fazla çalışma yapılmaktadır. Günümüzdeki en başarılı görsel üretim algoritmaları ise Üretken Çekişmeli Ağlar (Generative Adversarial Network – GANlar) ve Varyasyonlu Otokodlayıcı (Variational Autoencoder – VAE)'dir. Aynı yıllarda literatüre kazandırılmalarına rağmen GANlar öne geçerek çok daha başarılı sonuçlar ortaya koymaya başlamıştır. Bu makalede her iki model çalışma prensipleriyle beraber incelenmiş ve ürünleri kıyaslanmıştır.

Anahtar Kelimeler: Üretken Çekişmeli Ağlar, Varyasyonlu Otokodlayıcılar, Derin Öğrenme, Üretken Modeller.

Abstract

Realistic image production with deep learning is one of the popular artificial intelligence studies. More and more work is being done on the development of productive models. Today, the most successful visual production algorithms are Generative Adversarial Networks (GANs) and Variational Autoencoder (VAE). Although they were brought to the literature in the same years, GANs came to the fore and started to show much more successful results. In this article, both models are examined together with their working principles and their products are compared.

* Dokuz Eylül Üniversitesi İİBF Yönetim Bilişim Sistemleri, Yönetim Bilişim Sistemleri, omerfarukereken@gmail.com

** Dokuz Eylül Üniversitesi İİBF ve DEÜ-BİMER Yönetim Bilişim Sistemleri Bölümü, cigdem.tarhan@deu.edu.tr

Keywords: Generative Adversarial Networks, Variational Autoencoders, Deep Learning, Generative Models.

1. GİRİŞ

Son yıllarda sosyal medyada paylaşılan; ünlü kişilerin yüzlerinin farklı insanların görüntülerine yerleştirildiği, tek bir fotoğraftan kişinin yüzünün hareket ettirildiği, konuşturulduğu videolar çok fazla ses getirdi. Aynı şekilde bilgisayar tarafından üretilen, hiç var olmayan insan yüzlerinin, gerçek yüzlerden ayırt edilememesi pek çok insan tarafından şaşkınlıkla karşılandı. Bazı insanlar ise bu teknolojinin yanlış ellerde can sıkıcı sonuçlar doğurabileceğinden yakındı ki haklılık payları da vardı. Ancak doğru şekilde kullanıldığı takdirde bu teknolojiler insanlık adına büyük faydalar sunabilmektedir. Peki bu teknolojiler nasıl geliştiriliyor? Bu makalenin konusu bilgisayar tarafından üretilen bu görsellerin arkasında yatan algoritmalarıdır.

Bahsi geçen teknolojiler yapay zeka alanındaki gelişmelerin ürünüdür. Yapay zeka alanındaki gelişmeler özellikle son on yılda büyük ivme kazanmıştır. Günümüzde neredeyse her alanda makine öğrenmesi ve derin öğrenme algoritmaları çalıştırılmaktadır. Makine öğrenmesinin denetimli öğrenme algoritmaları halihazırdaki etiketli verilerden sonuçlar elde ederken, denetimsiz öğrenme algoritmaları etiketlenmemiş ham verinin ardındaki gizli matematiksel modelleri çıkararak onu anlamlandırabilmektedir. Bir verinin arkasındaki matematiksel yapının ortaya konması ise o veri gibi yeni verilerin üretilebilmesini sağlayarak bilgisayara gerçek hayatta yer alan figürleri, verileri üretme imkânı tanımaktadır.

Verilerin yapılarını öğrenerek yeni örnekler üretebilen algoritmalara derin öğrenmenin üretken modelleri denmektedir. Üretken Çekişmeli Ağlar (GANlar), Varyasyonlu Otokodlayıcı (VAE), PixelRNN/CNN, Gürültü Giderici Otokodlayıcı (GSN) ve Boltzmann Makineleri üretken modellerden bazılarıdır. Bu makalede en popüler üretken modellerden olan GANlar ve VAE incelenecektir. Birbirleriyle yakın zamanlarda keşfedilen bu üretken modeller zamanla gelişerek birbirlerinden farklı seviyede başarılar elde etmiştir.

Üretken çekişmeli ağlar ilk defa Goodfellow ve arkadaşları tarafından 2014 yılında bir makaleyle ortaya konmuştur. Bu makalede çekişmeli bir süreç ile üretken modeller ortaya koyabilen yeni bir sistem geliştirildiği belirtilmektedir. GAN'da bir tanesi denetimsiz öğrenme diğeri ise denetimli öğrenme metodlarını kullanan iki farklı derin öğrenme modeli eğitilmektedir. Birincisi denetimsiz öğrenme ile görüntülerdeki veri dağılımını öğrenen üretken modeldir. İkincisi ise bir örneğin üretken modelden değil de eğitim verisinden yani gerçek veriden gelme ihtimalini hesaplayan denetimli öğrenmenin sınıflandırma algoritmalarını kullanan ayırıcı modeldir. Üretken model adeta ayırıcı modeli kandırmaya, hata yapma ihtimalini artırmaya çalışan bir sahtekardır. Ayırıcı model ise üretken modelin hatalı ürün ortaya koyduğunu ispatlamaya çalışan bir dedektiftir (Goodfellow vd., 2014). İşte birbirleriyle kapişan bu iki model GANların çok başarılı sonuçlar elde etmesini sağlamaktadır. GANlar ortaya çıktıktan sonra pek çok farklı modeli üretilmiş ve süreç içerisinde büyük gelişmeler kaydedilmiştir.

Varyasyonlu otokodlayıcı ise 2014 yılında Diederik P. Kingma ve Max Welling tarafından literatüre kazandırılmış üretken model olan bir derin öğrenme yöntemidir (Kingma ve Welling, 2013). VAE veri üzerinde aynı anda bir kod çözücü ve kodlayıcı tanımlamaktadır. Kod çözücü aracılığıyla örtülü bir yoğunluk modeli tahmin edilirken aynı zamanda kodlayıcı aracılığı ile saklı bir temsil hesaplamak için bir çıkarım prosedürü sağlanmaktadır. Özetle VAE’de kodlayıcı girdiyi verisini bir temsil uzayına eşlerken kod çözücü orijinal girdinin benzerini temsilden tekrar inşa etmektedir (Cemgil vd., 2020; Tschannen vd., 2018).

2. YÖNTEM

Bu çalışmada yapay zeka tarafından gerçekleştirilen görsel veri üretimi teknolojileri ve bunların çalışma prensipleri araştırılmıştır. Bu alanda en başarılı sonuçlar veren VAE ve GANlar üzerinde detaylı inceleme yapılmıştır. Makale yazılırken konuya yabancı okurların da anlayabilmesi için başlangıç seviyesi konulara da değinilmiştir.

Öncelikle yapay zeka, makine öğrenmesi ve derin öğrenmenin temellerinden bahsedilmiş, sonrasında üretken modellerin dayandığı yapay sinir ağları ve onun altındaki evrimsel sinir ağları anlatılmıştır. VAE ve GANların detaylarından bahsedildikten sonra bu algoritmalar tarafından ortaya konan ürünlerin örnekleri sunulmuştur.

3. BULGULAR

Yapay zeka kümesinde yer alan derin öğrenmenin bir metodu olan üretken modellere giriş yapmadan önce yapay zeka ve makine öğrenmesinden bahsetmek faydalı olacaktır. Yapay zeka, makinelerin öğrenmesi ve onlara zeka kazandırılması anlamına gelmektedir. Yapay zeka altında makine öğrenmesini ve o da altında derin öğrenmeyi barındırmaktadır (Bini, 2018; Buduma, 2017).

Makine öğrenmesi matematiksel formüller kullanılarak bilgisayara bir öğrenme modeli oluşturma çalışmasıdır. Makine öğrenmesi, sunulan eğitim verilerinin türüne, eğitim verilerinin alınma sırası ve yöntemine ve öğrenme algoritmasını değerlendirmek için kullanılan test verilerine göre farklı türlere ayrılabilir. Literatürde farklı pek çok tür bulunmakla birlikte genel anlamda 3 farklı öğrenme yönteminden bahsedilebilir. Bunlar denetimli öğrenme, denetimsiz öğrenme ve pekiştirmeli öğrenmedir. Günlük hayatta farklı öğrenme yöntemleri bulunduğu gibi makinelerle öğretme işlemi de farklı yöntemlerle gerçekleştirilebilmektedir (Mohri, 2012).

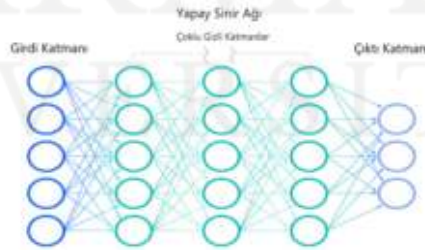
Örneğin bir çocuk sobaya dokunduğunda bir acı hissederek sobaya dokunmaması gerektiğini öğrenir. Her ne kadar sözlü olarak sizi anlamasa da sobanın verdiği acı onda öğrenme eylemine dönüşmektedir. Pekiştirmeli öğrenme de bu yöneme benzemektedir. Sistem ortamda bir yönlendirici bulunmadan, deneme yanılma yoluyla en yüksek ödülü kazanabilmek için ne yapmak gerektiğini keşfederek öğrenmektedir (Sutton ve Barto, 2018).

Denetimli öğrenme ise sistemin sebep-sonuç ile önceden beslendiği ve aradaki ilişkiyi belirleyerek öğrenmenin gerçekleştiği yöntemdir (Hastie vd., 2009). Örneğin bir sınıftaki öğrencilerin uyku

saati, çalışma saati, çözdüğü test sayısı, IQ seviyesi ve sınavdan aldığı puan bilgilerinin tutulduğu bir tablo düşünelim. Buradaki sonuç yani etiket verisi sınav puanıdır. Geri kalan özellikler ise sonucu oluşturan nitelikleri göstermektedir. İnsan gözüyle incelendiğinde dahi bir süre sonra hangi özelliğin sınav sonucunu yükseltip alçalttığını görebilir ve bize verilen niteliklere göre bir sınav sonucu tahmininde bulunabiliriz. İşte bilgisayar sonuçlara bakarak verilen nitelikleri belirli ağırlık değerleri ile matematiksel olarak modelleyerek, nitelikler üzerinden yeni sonuç tahminleri yapabildiğinde denetimli öğrenme gerçekleşmektedir.

Denetimsiz öğrenme ise etiketin bulunmadığı sadece niteliklerin bulunduğu ve eldeki verilerden anlamlı bilgiye ulaşılması istenen öğrenme şeklidir. Denetimli öğrenmenin tanımı belirgin iken denetimsiz öğrenmenin tanımı daha muğlak kalabilmektedir. Özetle denetimsiz öğrenmede eldeki bir veriye ait niteliklerin ardındaki saklı yapı keşfedilerek anlamlı sonuçlar çıkarılması hedeflenmektedir. Kümeleme ve boyut indirgeme denetimsiz öğrenmenin yaygın örneklerindendir. Denetimsiz öğrenmenin diğer bir yaklaşımı ise bu yazının ana temasını oluşturan üretken modellerdir (Goodfellow vd., 2020). Üretken modeller girdi verilerindeki düzenlilikleri ve kalıpları otomatik olarak keşfederek, öğrenerek orijinal veri kümesinin içerisinde çıkarılmış olabilecek şekilde yeni örneklerin üretilmesini hedeflemektedir (Brownlee, 2019).

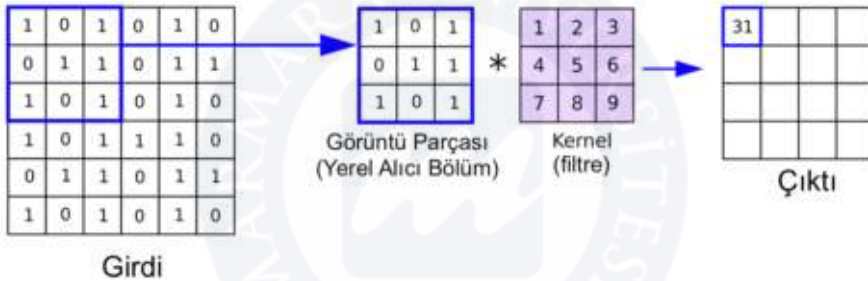
Üretken modeller derin öğrenmede yapay sinir ağları algoritmalarını kullanmaktadırlar. Yapay sinir ağları algoritmaları insan beyninden esinlenerek geliştirilmiştir. Biyolojik nöronların birbirine sinyal gönderme şeklini taklit etmektedirler. Yapay sinir ağları Şekil-1'de görüldüğü üzere bir girdi katmanı, bir veya daha fazla gizli katman ve bir çıktı katmanı içeren bir düğüm katmanından oluşur. Her düğüm veya yapay nöron diğerine bağlanır ve ilişkili bir ağırlık ve eşik sahiptir. Herhangi bir düğümün çıktısı belirtilen eşik değerinin üzerindeyse, o düğüm etkinleştirilir ve ağırlık bir sonraki katmanına veri gönderilir. Aksi takdirde, ağırlık bir sonraki katmanına veri iletilmez. Sinir ağları, doğruluklarını öğrenmek ve gelişmek için eğitim verilerini kullanır. Algoritmalar doğru ayarlara ulaştıktan sonra verileri yüksek bir hızda sınıflandırmamıza, kümelememize, anlamlı sonuçlar çıkarmamıza ve tahminler yapabilmemize olanak tanımaktadırlar (IBM, 2020).



Şekil 1. Yapay Sinir Ağı (IBM, 2020)

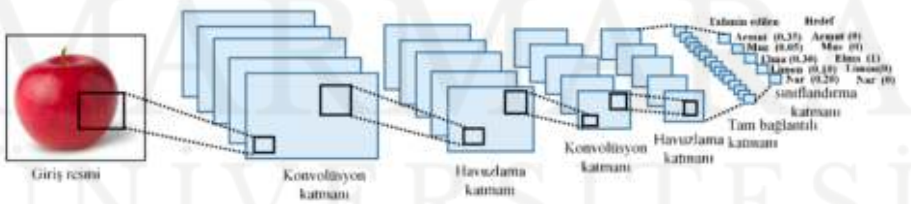
Görüntü işleme ve üretme konusunda çok katmanlı yapay sinir ağlarının bazı dezavantajları bulunmaktadır. Klasik yapay sinir ağları her bir girdi için bir algılayıcı kullanmaktadır. Fakat görüntülerde fazlasıyla girdi değeri bulunmaktadır. Büyük görüntüler için kullanılan ağırlıkların

miktarı yönetilemez hale gelebilmektedir. Bu durum eğitim sırasında zorluk çıkarmakta ve aşırı öğrenme ile neticelenebilmektedir. Diğer bir yaygın sorun, sinir ağlarının bir girdiye (görüntüye) ve onun kaydırılmış versiyonuna farklı tepki vermesidir. Örneğin bir fotoğraftaki figür görüntünün bir köşesinde yer alıyorsa sinir ağı figürün hep o bölgede yer alması gerektiğini öğrenebilmektedir. İşte bu noktada evrişimsel sinir ağları (Convolutional Neural Network – CNN) bir kurtarıcı olarak ortaya çıkmaktadır. CNN'ler bu sorunlara filtreler ile çözüm üretmektedir. Filtreler Şekil-2'de görüldüğü gibi resim piksellerine belirli filtreler uygulayarak sinir ağının öğrenmesi gereken ağırlık miktarını düşürmekte ve resim içerisinde aranan figürün konumunu önemsiz hale getirmektedir. Filtreler literatürde kerneller olarak isimlendirilmektedir (Stewart, 2019).



Şekil 2. Filtreler (Reynolds, 2019)

Şekil-3'te CNN ile bir görüntünün sınıflandırılması süreci görülmektedir. Bu örnekte model tahmini sonrasında toplam hata değeri alınarak geri yayılım algoritmaları ile model eğitilmektedir. Nihayetinde modelin resimlere göre doğru tahminler gerçekleştirmesi beklenmektedir (İnik ve Ülker, 2017).



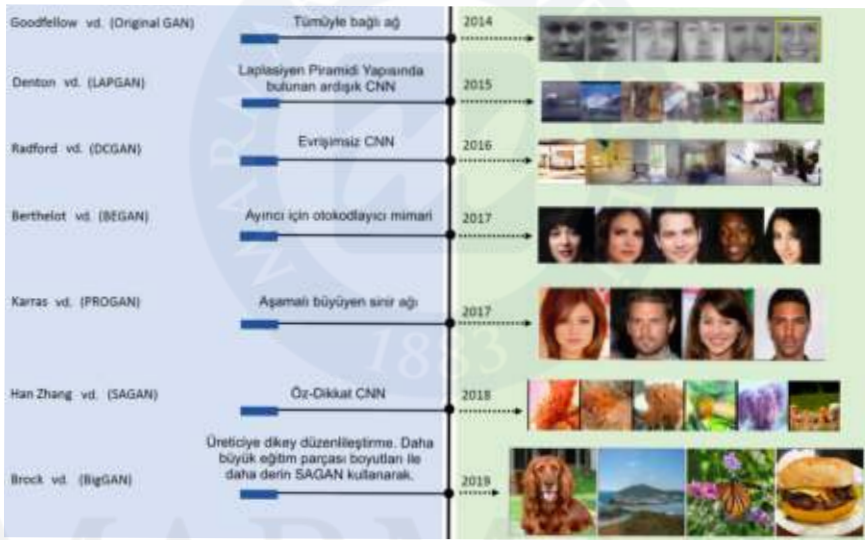
Şekil 3. CNN Yapısı (İnik ve Ülker, 2017, 95)

Buraya kadar anlatılanlar üretken modellerin öğrenme yöntemlerinin temelleriyle alakalıydı. VAE ve GANlar bahse konu yöntemleri farklı tekniklerle bütünleştirerek öğrendikleri bilgileri yeni görsel üretimlerinde kullanmaktadır. VAE ve GANlar haricinde pek çok derin üretken model çalışması bulunmaktadır. Kısıtlı Boltzmann Makineleri (RBM), Derin İnanç Ağları (DBN), Derin Boltzmann Makineleri (DBM), Gürültü Giderici Otokodlayıcı (DAE) ve Üretken Stokastik Ağ (GSN) gibi derin üretken modeller ses, görüntü veya video verileri içerisindeki zengin dağılımları yakalayarak yeni örnekler sentezleyebilme konusunda yapılan bazı popüler çalışmalardır. Bu derin

üretken modeller Markov Zinciri Monte Carlo (MCMC) tabanlı algoritmalar ile modellenmektedir (Sexena ve Cao, 2021). PixelRNN ve PixelCNN de diğer üretken modellerden bazılarıdır. Bu modeller yazımızın konusunu oluşturmadığından VAE ve GANların detaylarıyla devam edilecektir.

3.1. Üretken Çekişmeli Ağlar (GANlar)

GANların literatüre kazandırıldığı 2014 yılından itibaren üzerlerinde pek çok çalışma yapılmış ve bu da birçok farklı çeşidinin doğmasına yol açmıştır (Wang vd., 2020). Şekil-4'te sağ tarafta görülen resimler sol taraftaki çalışmalar sonucunda GANlar tarafından üretilmiş resimlerdir. Özellikle 2017 sonrasındaki çalışmalarda açıkça görüldüğü üzere sonuçlar olağan üstü boyutlara ulaşmıştır.



Şekil 4. GANların Gelişimi (Wang vd., 2020)

Üretken çekişmeli ağlar örnek görüntüler ile beslenerek kendi yapısı içerisinde benzer görüntü örnekleri oluşturabilen bir derin öğrenme tekniğidir. GANlar tarafından üretilebilen temsiller, görüntü sentezi de dahil olmak üzere anlamsal görüntü düzenleme, stil aktarımı, görüntü süper çözünürlüğü ve sınıflandırma gibi çeşitli uygulamalarda kullanılabilir (Creswell vd., 2018).

Üretken modelin amacı bir örneklem kümesini inceleyerek küme elemanlarını oluşturan olasılık dağılımını ortaya koymaktır. Böylece GANlar tahmini olasılık dağılımından edindikleri bilgiye göre yeni örnek resimler oluşturabilir. GANlar oyun teorisi ışığında yapay sinir ağları ile çalıştırılan iki farklı makine öğrenmesi modeli arasındaki 0'a toplam oyununa dayalıdır. Birinci oyuncu ağının adı üreticidir. Diğer oyuncunun adı ise ayrıcıdır (Goodfellow vd., 2020).

Üretici ağ, olasılık modeli ile bir veri setinin nasıl üretildiğini tanımlar. Bu olasılık modeli üzerinden yeni veriler üretilebilir. Ancak başlangıçta üretici, $G(z; \theta^{(G)})$ üretici fonksiyonuna hizmet

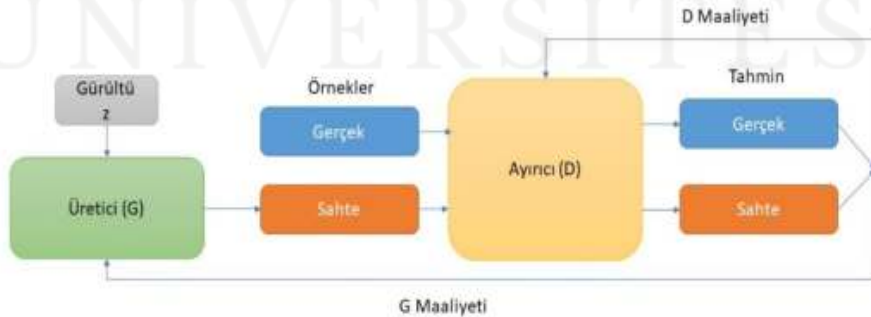
eden bir z gürültü vektörünün $p(z)$ dağılımı ile tanımlanır. $\theta^{(G)}$ üreticinin oyundaki stratejisini tanımlayan bir dizi öğrenebilen parametreyi ifade etmektedir. Girdi vektörü ve z örneği, $p(z)$ dağılımından gelen bir rastgelelik tanımlayan ifade olarak düşünülebilir. Üreticinin esas amacı yapılandırılmamış z gürültüsünü gerçekçi örneklerle dönüştürebilen $G(z)$ fonksiyonunu öğrenmektir (Goodfellow vd., 2020).

İkinci ağ; ayırıcı, üretici tarafından üretilen x örneklerini inceleyerek x 'in gerçek mi yoksa sahte mi olduğuna dair $D(x; \theta^{(D)})$ tahminini döndürmektedir. Her iki oyuncu da bir maliyet (hata) döndürmektedir. Üretici $J^{(G)}(\theta^{(G)}, \theta^{(D)})$ ve ayırıcı $J^{(D)}(\theta^{(G)}, \theta^{(D)})$ döndürmektedir. Her ikisi de kendi maliyetini düşürmeye çalışmaktadır. Her iki oyuncu da birbirinin tersi yönde bir amaç gütmektedir. Aynı amaca hizmet eden farklı maliyet fonksiyonları formülleri bulunmaktadır (Goodfellow vd., 2020).

Üretilen maliyet fonksiyonlarından birisi ikili çapraz entropi fonksiyonudur. Bu fonksiyon bize GANların hatasını çıkararak sistemin kendisini eğitmesini sağlamaktadır. Burada bir eğitim grubu için ortalama hata miktarı elde edilmektedir. Maliyet fonksiyonu aşağıda görülen formül ile ifade edilmektedir (DeepLearning.AI, 2021).

$$J(\theta) = -\frac{1}{m} \sum_{i=1}^m [y^{(i)} \log h(x^{(i)}, \theta) + (1 - y^{(i)}) \log (1 - h(x^{(i)}, \theta))]$$

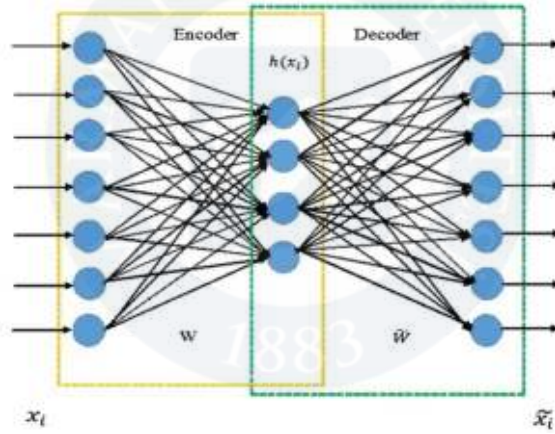
GANların çalışma prensibini özetlemek gerekirse; üretici yapay sinir ağına bir z gürültüsü girdi olarak sokulur. Üretici z gürültüsünden rastgele ağırlıklarda bir resim üretir. Gerçek resimlerin bulunduğu veri setinden bir resim ve sahte resim girdi olarak ayırıcıya sokulur. Ayırıcı sınıflandırma yapan bir yapay sinir ağıdır. Çıktı olarak sahte resmin gerçeklik oranı konusunda bir tahminde bulunur. Ortaya çıkan sonuçlar maliyet fonksiyonunda değerlendirilir. Sonrasında üretici ve ayırıcı yapay sinir ağlarının ağırlık parametreleri buna göre güncellenir. Üretici ve ayırıcının eğitimi başlangıçtan itibaren eş seviyelerde ilerleme ile gerçekleşmektedir. Çalışma prensibi Şekil-5'de gösterilmiştir.



Şekil 5. GAN Algoritması (Sexena ve Cao, 2021)

3.2. Varyasyonlu Otokodlayıcı (VAE)

VAE'nin temeli otokodlayıcılardan gelmektedir. Otokodlayıcı bir sinir ağı, çıktı değerlerini girdilere eşit olacak şekilde ayarlayarak geri yayılım yapan bir denetimsiz öğrenme algoritmasıdır. Otokodlayıcı girdi değerini çıktıya eşitlemek için tanımlayıcı fonksiyonu bulmaya çalışmaktadır. Bunu Şekil-6'da görülen kodlayıcı ve kod çözücü olarak işlev gören iki farklı yapay sinir ağı ile gerçekleştirmektedir. Aslında otokodlayıcı girdi verisini iki aşamada tekrardan inşa etmektedir. Birinci adım f kodlayıcı fonksiyonu üretmekte ve ikinci adım kod çözücü g fonksiyonunu üretmektedir. Öyle ki girdi durumundaki veri dağılımdan örnek alınan çoğu x için $g(f(x)) \approx x$ elde edilmektedir (Alain ve Bengio, 2014).

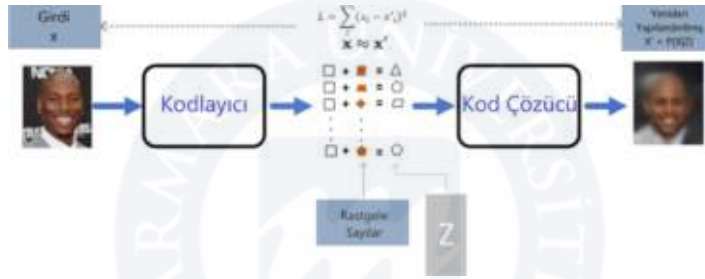


Şekil 6. Otokodlayıcı Yapay Sinir Ağı (Ahmed vd., 2017)

Otokodlayıcı sinir ağlarında, her kodlama boyutu için tek bir değer çıktısı alınmaktadır. Kod çözücü ağı daha sonra bu değerleri alır ve orijinal girişi yeniden oluşturmaya çalışır. Varyasyonlu otokodlayıcılar ise gizli uzayda bir gözlemi tanımlamak için bir olasılık yöntemi sağlamaktadır. Yani gizli durumdaki her bir nitelik için tek bir değer çıktısı alan bir kodlayıcı inşa etmektense kodlayıcı her bir gizli nitelik için olasılık dağılımı tanımlayacak şekilde formüle edilmektedir (Jordan, 2018). Bu modelin varsayımları zayıf ve eğitim geri yayılım yoluyla hızlıdır. VAE'ler bir tahminde bulunur ancak bu tahminin getirdiği hata yüksek kapasiteli modeller olmalarına rağmen küçük sayılabilir. Bu özellikler modelin popülerliğinde hızlı bir yükselişe sonuçlanmıştır (Doersch, 2016).

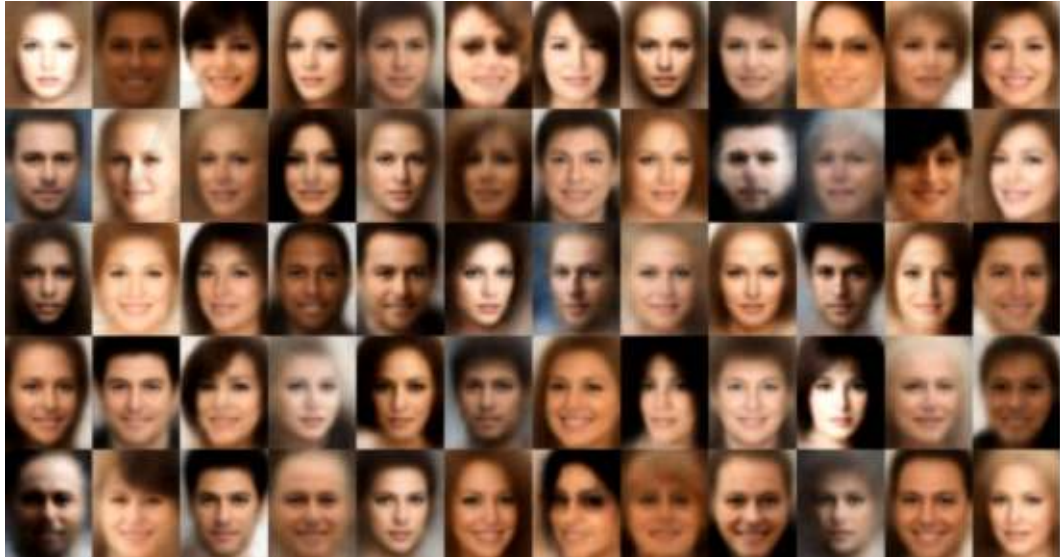
Daha basit olarak ifade etmek gerekirse; otokodlayıcılar girdiyi sıkıştırılmış bir forma (gizli uzaya) öyle bir şekilde kodlamaktadır ki yeniden yapılandırılmış versiyon girdiye mümkün olduğunca benzesin. Kodlayıcı girdiyi küçük bir yoğunluk temsiline dönüştürmektedir. Otokodlayıcılar eğitim esnasında kodlayıcı gizli uzayın sınırlandırılmasında ihtiyaç duyulan ilgili bilgilerin çoğunu korumayı ve alakasız parçaları atmaya öğrenir. Örneğin; gürültü, ses. Ancak kodlayıcı tarafından üretilen gizli uzay seyrek olarak doldurulur ve bu o alandaki değerlerin dağılımını tahmin etmenin

zor olduğu anlamına gelir. Bu durum üretken model olarak görev alan kod çözücünün yeni bir görüntü üretebilmek için gerekli gizli değerleri bulmasını neredeyse imkansız hale getirmektedir. Varyasyonlu otokodlayıcı ise gizli uzayı daha öngörülebilir, daha sürekli ve daha az seyrek hale getirerek çalışır. Gizli değişkenlerin normal dağılıma zorlanmasıyla gizli uzay üzerinde kontrol kazanılır. Gizli değerleri doğrudan kod çözücüye iletmek yerine, VAE bunları bir ortalama ve standart sapma hesaplamak için kullanır. Daha sonra kod çözücüye sunulacak girdi değeri karşılık gelen normal dağılımdan örneklenir (Kana, 2020).



Şekil 7. Varyasyonlu Otokodlayıcı (Kana, 2020)

Şekil-8'deki insan yüzleri VAE tarafından üretilmiş yapay yüzlerdir. Burada da açıkça görüleceği üzere VAE'lerin gerçekleştirdikleri üretimlerde bir bulanıklık yer almaktadır. GANlarda bu bulanıklığın bulunmaması onları daha tercih edilir kılmıştır.



Şekil 8. VAE ile Üretilmiş Resimler (Mormul, 2018)

4. TARTIŞMA ve SONUÇ

Yapay zeka alanındaki gelişmeler tüm dünyada hız kesmeden devam etmektedir. Özellikle görüntü işleme ve yapay zeka üretkenliği revaçta olan konulardır. Son yıllarda şahit olduğumuz bilgisayar tarafından üretilmiş insan resimleri, ünlülerin kafalarının başka görüntülere gerçekçi bir şekilde aktarılması, tek bir kare fotoğraftan video oluşturulması, geçmiş görüntülerin kalitelerinin yükseltilmesi gibi pek çok yeniliğin ardında derin öğrenmenin üretken modelleri yatmaktadır.

Derin öğrenme ile görsel üretiminin şanslı olduğu alanlardan bir tanesi veriye ulaşım kolaylığıdır. Normal şartlarda makine öğrenmesi ve derin öğrenmede belki de en zorlu alan veriye hatta temiz veriye ulaşabilmektir. Çünkü genellikle geçmişe ait düzenli tutulmuş veriye ihtiyaç duyulmaktadır. Ancak böyle bir veri ya hiç tutulmamış olmakta ya da tutulduysa bile dağınık, düzensiz ve istenilen yapıya dönüştürülmesi çok uzun süren bir yapıda bulunmaktadır. Veriye ulaştıktan sonraki aşama genellikle çok daha hızlı gerçekleşmektedir. Ancak görsel verisine ulaşmak internet ve üzerinde bulunan görseller sayesinde çok daha kolay olmaktadır. Bu nedenle bu alanda çalışma yapmak cezbedici niteliktedir.

Üretken modeller özellikle 2014 yılında VAE ve GANların geliştirilmesiyle ivmeli bir başarı öyküsü çizmeye başlamıştır. Halihazırda en başarılı modelin ise GANlar olduğu söylenebilir. Kısıtlı Boltzmann Makineleri (RBM), Derin İnanç Ağları (DBN), Derin Boltzmann Makineleri (DBM), Gürültü Giderici Otokodlayıcı (DAE) ve Üretken Stokastik Ağ (GSN) gibi markov zincirini kullanan üretken modelleri üretmek yavaş ve zorlu bir süreci gerektirebilmektedir. VAE ise gizli uzaydaki bir veri noktasını temsil etmek için istatistiksel çıkarımla derin öğrenmeyi kullanmaktadır. Ancak bu model de zorlu olasılık hesaplamalarının tahminindeki karmaşıklığı deneyimlemektedir. Bu modeller kullandıkları metodlar gereği görsel veri setlerinin yapısından dolayı zorluklarla karşılaşmaktadır. Ek olarak büyük boyutlu uzayda Markov Zinciriden örneklem üretince buğulu, yavaş ve kusurlu sonuçlar elde edilmektedir. GANlar ise tüm bu üretken modellere oranla çok daha başarılıdır (Sexena ve Cao, 2021).

Çoğunluk tarafından kabul edilen bir kaide; adil koşullarda gerçekleştiği takdirde rekabetin iyi olduğu ve gelişimi sağladığıdır. Gerçek hayatta yer alan bu yasa ilginç bir şekilde insan beyninden esinlenilerek geliştirilen yapay zekada da uygulama alanı bulmuştur. Rekabet anlayışıyla çalışan algoritmalar en başarılı sonucu vermektedir. GANlarda eşit eğitim seviyelerinde olduğu takdirde birbirleriyle kapışan iki yapay ağ birbirlerini geliştirmekte ve sonuçta inanılmaz ürünler ortaya koymaktadır. Kim bilir belki de yapay zekanın gelişimi gerçek hayat yasalarının bu alanda uygulanmasına bağlıdır.

KAYNAKLAR

Ahmed, H.O.A., Wongb, M.L.D. ve Nandi, A.K. (2017). Intelligent condition monitoring method for bearing faults from highly compressed measurements using sparse over-complete features. *Mechanical Systems and Signal Processing*. 99: 459-477.

- Alain, G. ve Bengio, Y. (2014). What Regularized Auto-Encoders Learn from the Data-Generating Distribution. *Journal of Machine Learning Research*. 15: 3743-3773.
- Bini, S.A. (2018). Artificial Intelligence, Machine Learning, Deep Learning, and Cognitive Computing: What Do These Terms Mean and How Will They Impact Health Care? *The Journal of Arthroplasty*. 33(8): 2358-2361.
- Brownlee, J. (2019). A Gentle Introduction to Generative Adversarial Networks (GANs). <https://machinelearningmastery.com/what-are-generative-adversarial-networks-gans/>. (20.08.2021).
- Buduma N. (2017). *Fundamentals of Deep Learning*. O'Reilly Media, Inc.
- Cemgil, A.T, Ghaisas, S., Dvijotham, K., Gowal, S. ve Kohli, P. (2020). The Autoencoding Variational Autoencoder. *34th Conference on Neural Information Processing Systems (NeurIPS 2020)*, Vancouver, Canada.
- Creswell, A., White, T., Dumoulin, V., Arulkumaran, K., Sengupta B. ve Bharath, A.A. (2018). Generative Adversarial Networks: An Overview. *IEEE Signal Processing Magazine*. 35(1): 53-65.
- DeepLearning.AI, (2021). *BCE Cost Function*. <https://www.coursera.org/specializations/generative-adversarial-networks-gans>. (20.08.2021)
- Doersch, C. (2021). Tutorial on Variational Autoencoders. *arXiv*: 1606.05908.
- Goodfellow, I.J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. ve Bengio, Y. (2014). Generative adversarial nets. *arXiv*:1406.2661.
- Goodfellow, I.J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. ve Bengio, Y. (2020). Generative Adversarial Networks. *ACM*, 63(11): 139-144.
- Hastie, T., Tibshirani, R. ve Friedman, J. (2009). *The Elements of Statistical Learning*. New York: Springer.
- İnik, Ö. ve Ülker, E. (2017). Derin Öğrenme ve Görüntü Analizinde Kullanılan Derin Öğrenme Modelleri. *Gaziosmanpaşa Bilimsel Araştırma Dergisi*. 6(3): 85-104.
- IBM, (2020). *Neural Networks*. <https://www.ibm.com/cloud/learn/neural-networks>. (20.08.2021)
- Jordan, J. (2018). *Variational Autoencoders*. <https://www.jeremyjordan.me/variational-autoencoders/>. (20.08.2021)
- Kana, M. (2020). *Variational Autoencoders (VAEs) for Dummies – Step By Step Tutorial*. <https://towardsdatascience.com/variational-autoencoders-vaes-for-dummies-step-by-step-tutorial-69e6d1c9d8e9>. (20.08.2021)
- Kingma, D. P. ve Welling, M. (2013). Auto-encoding variational bayes. *arXiv*:1312.6114.
- Mormul, W. (2018). *Variational Autoencoder*. <https://github.com/wojciechmo/vae>. (20.08.2021).
- Tschannen, M., Bachem, O. ve Lucic, M. (2018). Recent Advances in Autoencoder-Based Representation Learning. *arXiv*:1812.05069v1
- Mohri, M., Rostamizadeh, A. ve Talwalkar, A. (2012). *Foundations of Machine Learning*. Cambridge: The MIT Press.
- Reynolds, A.H. (2019). *Convolutional Neural Networks (CNNs)*. <https://anhreynolds.com/blogs/cnn.html>. (20.08.2021).
- Sexena, D. ve Cao, J. (2020). Generative Adversarial Networks (GANs): Challenges, Solutions, and Future Directions. *arXiv*: 2005.00065.
- Stewart, M. (2019). Simple Introduction to Convolutional Neural Networks. <https://towardsdatascience.com/simple-introduction-to-convolutional-neural-networks-cdf8d3077bac>. (20.08.2021).
- Sutton, R.S. ve Barto, A.G. (2018). *Reinforcement Learning: An Introduction*. Cambridge: The MIT Press.
- Wang, Z., She, Q. ve Ward, E.T. (2020). Generative Adversarial Networks in Computer Vision: A Survey and Taxonomy. *arXiv*:1906.01529.