



Yeni Nesil Tehdit: Derin Kurgu (DeepFake)

Şebnem ÖZDEMİR*

İnsanlık tarihi boyunca en önemli olgulardan biri bilgiye erişebilmektir. Özellikle sanayi toplumundan bilgi toplumuna geçişle birlikte, bu olgu daha baskın hâle gelmiştir. İnternetin keşfi, yaygınlaşması, akıllı telefonlar sayesinde zaman ve mekândan bağımsız bir şekilde bağlantılı kalabilmek, bilgiye erişimi daha da kolaylaştırmıştır. Ancak bu erişim kolaylığı zamanla büyük bir zorluğa dönüşmüş, insanlığı daha kaliteli bilgiye kolaylıkla sahip olmaktan çok, bilgi ile mücadele eder bir hâle getirmiştir (Kominiarczuk and Ledzińska, 2014) (Özdemir, 2016).

Bilgiye erişim süreçlerindeki mücadelenin ana kaynağı, aşırı bilgi artışı (information overload), yanlış bilgilenme, bilgi kirliliği, sahte ve manipülatif haberlerdir. Tüm bu etmenler, toplum algısı, sosyal birliktelik ve hatta demokratik süreçler bakımından ciddi tehditlerdir (Özdemir, 2015; Quayyum, ve diğ., 2019;). En basit biçimde bu tehditlerin tamamı birey ve/veya toplum özelinde karar verme süreçlerini sekteye uğratabilir, verilecek kararı pek çok açıdan etkilediği için manipülatif hâle getirebilir (Jacoby, 1984; Malhotra, 1982; Schick, Gorden and Haka,1990; Sparrow, 1999; Bawden ve diğ., 1999; Özdemir and Gülseçen, 2015)

Aşırı bilgi artışı; bilginin ulaştırıldığı kaynakların sayısının, türünün ve barındırdığı bilginin yüksek bir miktara ve şiddete ulaşmasını ifade etmektedir (Bawden ve diğ., 1999; Özdemir, 2016). Bilgi kirliliği (information pollution, infollution) ise bilginin kaynağının alakasız, sapırtıcı, düşük değerli bilgilerle birleştirilerek kirlenmesi, küresel ölçekteki bilgi yayılımının negatif sonuçlar doğuracak şekilde bozulmasıdır (Cai ve Zhang, 1996; Berkan, 2012). Bilgi kirliliğinin bir diğer yüzü sahte haberlerdir. Sahte haber; toplumu kandırmak, manipüle etmek amacıyla oluşturulmuş gerçek dışı içeriklerdir (Aldwairi ve Alwahedi, 2018; Jang ve Kim,

* İstinye Üniversitesi, İktisadi İdari Bilimler Fakültesi, Yönetim Bilişim Sistemleri, Doktor Öğretim Üyesi

ORCID: 0000-0001-6668-6285

2018). Önceleri sadece kitlesel aktarım, paylaşım ve metinler şeklinde ilerleyen sahte haberler, derin-kurgu teknolojileri ile yeni bir boyut kazanmıştır. Göz önünde olan isimlerin yeniden oluşturulmuş videoları ile toplum çok daha hızlı ve kolay biçimde yanlış içerikler, sahte bilgiler ve haberlerle buluşabilir hâle gelmiştir.

Yapay zekânın günümüzdeki en popüler uygulamaları, bir alt kümesi olarak kabul edilen makine öğrenmesinden yükselmektedir. Makine öğrenmesinde, özellikle insan beynindeki sinir hücrelerinin taklidi ve bunlar arasındaki bağlantıları çok katmanlı olacak şekilde yeniden kurgulayan bir başka alt küme, derin öğrenme (deep learning), yer almaktadır. Derin-kurgu özellikle derin öğrenme ile üretildiğinden, bu yöntemler kümesine atıfta bulunmak üzere deep ve kurgusal, gerçek olmayan tarafını vurgulamak için de fake kelimeleri ile temsilenmiştir.

Derin-kurgu, insan yüzünün tamamının veya bir kısmının yeniden üretilmesi ile ortaya çıkan bir dijital manipülasyondur (Korshunov & Marcel, 2018). Bir başka deyişle derin-kurgu, bir bireyin resim veya videolarının gerçek hâlleri kullanılarak yeniden üretilmesi ile ortaya çıkan, dijital ortamlarda beğeni ve yeniden paylaşım yoluyla hızlı yayılan, yapay/sentetik bir medyadır (Somers, 2020).

Yapay Zekânın Bir Alt Çalışma Alanı Olarak Derin-Kurgu

1956 yılında John McCarthy tarafından yapay zekâ (artificial intelligence) olarak adlandırılan çalışma alanı, o dönemde salt bir mühendislik bakışı ile incelenmekteydi. Ancak zaman içerisinde gelişen teknoloji ve matematiksel altyapı, yapay zekânın kapsamını mühendislikten tüm bilimlere ve alanlara etki edebilecek hâle getirmiştir. Özellikle bilgisayarın merkezi işlem birimi (CPU) yerine grafik işlem biriminin (GPU) veri işlemede devreye alınması, yeni nesil GPUların üretilmesi makine öğrenmesi ve alt çalışma alanlarında daha başarılı/performanslı, gerçeğe yakın sonuçlar/çıktılar üretilmesini sağlamıştır.

Makine öğrenmesi, makinenin bir görevi, belli bir performansla ortaya koyabilmek için belli bir sürede, belli bir yığınlıktaki veriden öğrenerek deneyim oluşturmalarıdır. Kullanılan veri, sayısal, metinsel, ses dosyası ve/veya görüntü/video vb. şeklinde olabilir. Yani makine; önceden tanımlanmış bir amaç doğrultusunda ne türlü bir veri yığını üzerinde çalıştırılıyorsa o yönde çıktı/karar kuralları oluşturur. Buradaki makine, salt bir robot gibi düşünülmemeli, bir düşünme/karar üretme birimi barındırabilecek tüm cihazlar olarak kabul edilmelidir.

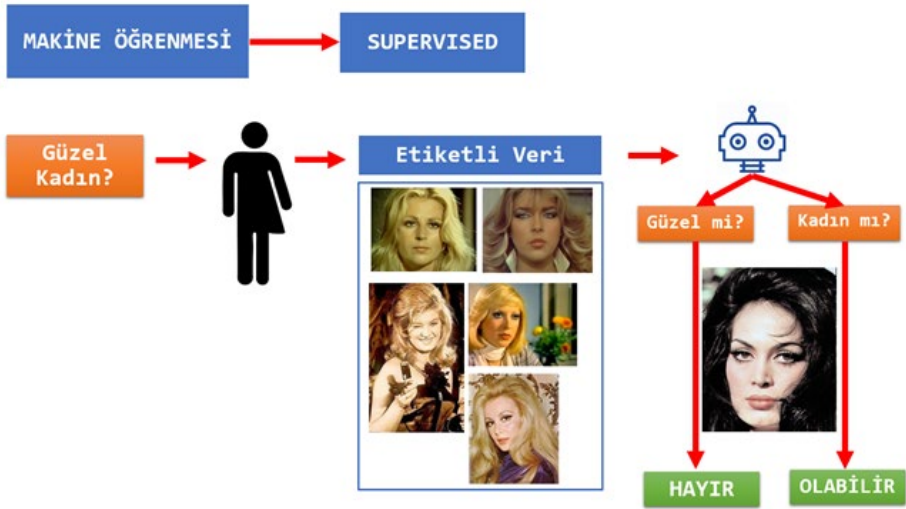
Makinenin öğrenmesinde dört temel yaklaşım bulunmaktadır. Bunlar: denetimli/gözetimli/danışmanlı (supervised), denetimsiz/gözetimsiz/danışmansız (unsupervised), yarı-denetimli/gözetimli/danışmanlı (semi-supervised), pekiştirmeli

(reinforcement) öğrenmedir. İlk üç öğrenme yönteminde yüksek miktarlarda veriye ihtiyaç duyulmaktadır. Özellikle denetimli öğrenmede, öğrenmenin sağlanması için etiketlenmiş veri kritiktir. Bir konuda çıktı üreten bir makine (yapay zekâ) tasarımında, öncelikle algoritmalar kullanılmakta veya mevcut algoritmalar modifiye edilmekte veya yeniden üretilebilmektedir. Buradaki algoritma, bilgisayar programlamada kullanılan terminolojiden biraz farklıdır. Makine öğrenmesinde algoritma, öğrenme yolculuğunun nasıl gerçekleşeceğini tanımlayan sıralı adımlar, matematiksel bağıntı ve formüller bütünüdür.

Bir algoritma, ne türlü bir veri seti üzerinde koşturulursa o verideki ilişkileri, örüntüleri tespit eder, öğrenir. Bu öğrenme süreci sonunda algoritma, artık o veriye göre karar/çıktı üreten bir modele evrilmiştir. Bu makine öğrenmesi (yapay zekâ) modeli, aynı değişkenlere sahip bir başka veri seti üzerinde sonuç/çıktı üretebilme kabiliyetine sahiptir. Bu veri seti, görüntü verilerinden oluşuyorsa algoritma, bu görüntüler üzerinde koşar, görüntüyü belli değişkenler ışığında tespit eden, değiştiren bir modele evrilir.

Örneğin bir Cast ajansına, eski Türk sinemasının güzel kabul edilen oyuncularına benzeyen kadınların oynatılacağı bir reklam teklifi gelsin. Cast ajansı, buradaki ilk elemeyi yapmak için makine öğrenmesi kullanacaksa seçtiği algoritmaya önce eski Türk sinemasındaki kadın oyuncuların yüzlerini öğretmesi gerekecektir. Bu öğrenme için önce o dönem kadın oyuncuların görselleri toplanmalı ve etiketlenmelidir. Yani makinenin öğrenmesi için veri seti oluşturulmalıdır. Şekil 1’de denetimli (supervised) öğrenme için varsayımsal makine öğrenmesi süreci anlatılmaktadır.

Araştırmacı (siyah figür) önce, “Eski Türk sinemasının güzel kadınları kimlerdir?” sorusunun cevabını düşünmelidir. Bu cevaba göre oyuncuların fotoğraflarını derleyecek, etiketleyecektir. Böylece bir veri seti oluşacaktır. Denetimli öğrenmede veri setleri, eğitim ve test kümesi olarak ikiye ayrılmaktadır. Algoritma, eğitim kümesindeki verilerden öğrenme yaşayacak ve bir modele evrilecektir. Test kümesi, modelin hedeflenen doğrultuda uygun/öğrenmiş bir model olup olmadığının testinin yapıldığı kümedir. Özetle makine (algoritma), eğitim kümesindeki fotoğraflardan yola çıkarak bir düşünme birimi oluşturacak, yani modele evrilecektir.



Şekil 1. Denetimli (Supervised) Öğrenme için Varsayımsal Makine Öğrenmesi Süreci

Şekil 1'deki süreç sonunda Türk sinemasının güzel kadın oyuncularının yüzlerinden öğrenmiş, buna göre seçme işlemi yapabilecek bir model varsayımsal olarak elde edilmiştir. Modelin, hedef doğrultusunda istenilen performans ile çalışıp çalışmadığı test sonucuna göre belli olacaktır. Test işleminde, Türkan Şoray'ın resmi kullanıldığında üreteceği etiket, modelin performansına yönelik bilgi verecektir. Bu varsayımsal makine öğrenmesi modeli, test sonucunda Türk Sinemasının Sultanı Türkan Şoray'ı "güzel değil" olarak etiketlemiştir. Bu varsayımsal makine öğrenmesi deneyiminde üretilen sonuç sanıldığı gibi tesadüfi/hatalı değildir. Çünkü makineye (modele) sağlanan eğitim kümesinde sarışınlık özelliği baskındır. Dolayısıyla model, sarışın olmayı güzelliğin temel kriteri olarak kabul ederek test kümesindeki fotoğrafı (Türkan Şoray) değerlendirmiştir. Varsayımsal makine öğrenmesi deneyiminden de anlaşıldığı gibi veri ve verinin etiketlenme biçimi oldukça kritiktir.

Derin Öğrenme

Derin öğrenme, makine öğrenmesinin bir alt çalışma kümesidir. Temel olarak insan beynindeki sinir hücrelerinin öğrenme prensibinin modellendiği, yapay sinir ağları adı verilen teknikleri barındırmaktadır. Deng ve diğ. (2014); derin öğrenmeyi, insan beyninin karmaşık problemleri çözerken kullandığı gözlem, analiz, öğrenme ve karar verme gibi yeteneklerini simüle edecek şekilde modeller/uygulamalar geliştirilmesini sağlayan teknikler olarak tanımlamaktadır.

1940'lı yıllardan beri geliştirilen bu tekniklerle, el yazısından resimlere kadar farklı pek çok alanda örüntünün keşfi, 3 boyutlu nesnelere/varlıkların, ses ve metnin

lerin tespiti yapılabilmektedir (McCulloch ve Pitts, 1943; Fukushima, 1982; Weng ve diğ., 1992; Cortes ve Vapnik, 1995; Hinton ve Salakhutdinov, 2006; Mikolov, ve diğ. 2010; Krizhevsky ve diğ. 2012).

Özellikle çok katmalı seviyedeki yapay sinir ağlarının üretilmesi, milyonlarca bağlantıya sahip bu modellerin örüntü tanımada insan seviyesine ulaşmasını sağlamıştır. Örneğin Facebook'un sahip olduğu DeepFace modeli, kullanıcının yüklediği resimlerdeki kişileri tanıyabilmekte ve otomatik olarak etiketlenmesini sağlayabilmektedir. Modelin bu seviyeye ulaşmasını sağlayan ise insan yüzünü tanımlamak için 120 milyon değişken yardımıyla etiketli veri üretilmesidir (Taigman ve diğ., 2014). 2014 yılında insanın bir diğer insanın yüzünü algılamadaki performansı %97.53 iken, yüksek miktarda veri ile eğitilen DeepFace modeli %97.35'lik bir performans sergileyerek insan seviyesine yetişmiştir (Wang ve Deng, 2018).

En temel seviyede yüzün tanımlanmasında üç ana adım kullanılmaktadır. Bunlar sırasıyla, bir fotoğraf karesindeki ya da videodaki insan yüzünün nerede olduğunu bulma/tespit etme, yüzü görüntünün/videonun içinden ayırma ve ayrılan yüzü tanıma şeklindedir (Ranjan et al., 2018).

Bu süreçte algoritmanın yüzü tanınması demek, yüzün hatlarını işaretlemesi ve farklı açılardan çekilmiş olsa bile kişinin aynı kişi olduğunu tespit edebilmesi demektir.

Derin-kurgu uygulamalarının geliştirilmesinde de başlangıçta benzer bir süreç yürütülmektedir. Yani, makine önce insan yüzünü öğrenmektedir.

Derin-Kurgu ve Yüz Manipülasyon Türleri

Akademik alanda derin-kurgu çalışmalarının ilk işareti sayılabilecek çalışma 1997 yılında yayınlanmıştır. Çalışmada Video Rewrite Program (videonun yeniden oluşturulması programı) tanımlanmış, videodaki konuşma, bambaşka kelimeler eklenerek gerçeğin bir başka versiyonu olacak şekilde yeniden üretilmiştir (Bregler, ve diğ., 1997). Bu çalışma insan yüzünün yeniden canlandırılması noktasındaki ilk örneklerden biri olarak kabul edilebilmektedir. Ancak 90'lı yıllardaki makinelerin GPU kapasitelerinin yetersizliği derin-kurgu uygulamalarının geliştirilmesini önemli ölçüde sınırlamıştır.

2000'li yılların başından itibaren teknoloji ve algoritmalar anlamındaki gelişmeler, derin-kurgu uygulamalarında yeni türlerin doğması ve daha performanslı çıktılar geliştirilmesinin önünü açmıştır. Günümüzde derin-kurgu uygulamaları ile insan yüzünün tamamının, bir kısmının ve mimiklerin transferi yapılabildiği gibi; yüz ifadesinin, ağzın yeniden modellenmesi yoluyla sahte bir konuşma oluşturmak; sıfırdan insan yüzü ve tepkisel bir surat/mimik üretmek de mümkündür.

Yüz sentezlemesi (entire face synthesis), farklı insan yüzlerinden, insan yüzünün neye benzediğini öğrenmiş bir yapay zekâ modelinin, aslında hiç var olmayan insan yüzleri oluşturabilmesidir (Karras ve diğ., 2019). 2014 yılında Ian Goodfellow tarafından sunulan üretici-çekişmeli yapay sinir ağları (generative adversarial neural network – GAN) ile, makinenin var olan örneklerden öğrenerek, gerçekte var olmayan çıktılar üretilmesi mümkün hâle gelmiştir. GAN'ın insan yüzleri ile eğitilmesi, yüzü tanımadaki performanslı öğrenmelere sahip modellerin ortaya çıkmasını sağlamaktadır. Örneğin StyleGAN, gerçekte var olmayan bir insan yüzünü üretebilen bir modeldir. (<https://thispersondoesnotexist.com>)

Bir başka derin-kurgu uygulaması kimlik değişimidir (face swap). Kimlik değişiminde kişinin yüzü, bir başka video/görsel içindeki insan yüzünün yerine transfer edilmektedir. Bu türlü bir derin-kurgu uygulamasında klasik grafik tabanlı işleme teknikleri veya klasik derin öğrenme uygulamaları kullanılmaktadır.

Kimlik değişimi dışında, nitelik manipülasyonu (attribute manipulation) olarak adlandırılan bir yöntem daha bulunmaktadır. Burada nitelik kelimesi ile kastedilen, yüzün makine tarafından öğrenilmesini sağlayan değişkenlerin tamamıdır. Dolayısıyla nitelik manipülasyonunda saçın ve/veya ten renginin değiştirilmesi, cinsiyet ya da yaş ile ilgili değişiklikler ve hatta gözlük gibi aksesuarlar eklenmesi vb. de yapılmaktadır (Generated Media, 2018). Kozmetik ürünleri, makyaj değişiklikleri, saç stili denemeleri ile ilgili tüm uygulamalar da yine bu türdeki derin-kurgulardandır. Nitelik manipülasyonu özetle, yüzü betimleyen değişkenlerin kullanılarak bireyin yüzünde ufak ya da büyük çaplı dokunuş ve değişiklikler yapmaktır.

İnsan yüzüne yapılan bu şekildeki müdahalelerden biri de yüz ifadesi değişimidir (expression swap). Herhangi bir videodaki insan yüzüne başka bir yüz ifadesi kazandırılması, hatta başka birinin yüzündeki ifadenin aktarılması gibi de düşünülebilir. Yüz ifadesi değişimi, sadece duygusal değişim olarak yorumlanamaz. Kişilerin ağız hareketlerine yapılacak müdahale ile konuşmalarında değişiklikler yapılması da bu kategoridedir.

Derin-kurgu ile üretilen bu manipülatif, kurgusal çıktılar, 2017 yılından itibaren ciddi zarar verebilecek güce erişmiş ve yeni nesil bir siber tehdit olarak nitelenmeye başlanmıştır.

Derin-Kurgu Vakaları

Terminoloji olarak derin-kurgu ilk defa Reddit isimli bir sosyal ağda zikredilmiştir. İfadeyi ilk defa ortaya atan kullanıcı, makine öğrenmesi (yapay zekâ) kullanarak ünlü isimlerin yüzlerini porno videolara transfer etmiş, üretilen pornografik vide-

oların paylaşımına açılmasına ortam sağlamıştır (Cole, 2018, Roettgers, 2018; Goggin, 2019; Dickson, 2019; BBC, 2019). Hatta yüzlerce ünlü aktrisin derin-kurgu ile üretilen pornografik içerikler sunan sitelerden en bilindik dördü, dünya genelinde 134 milyondan fazla görüntüleme/izleme almıştır (Ajder v.d., 2019). Makinenin özellikle yaşlandırma ve gençleştirme anlamında başarılı bir performansa erişmesi, çocuk pornosu üretiminde de yeni bir alan oluşturmaktadır. Bu alan, bu anlamda derin-kurgu uygulamaları ile ilgili oluşan kaygıyı perçinlemektedir (Eelmaa, 2021). Özetle, bireylerin eğlence ve sosyalleşme anlamında kendilerine ve kendi hayatlarına dair yaptıkları her görsel paylaşım, herkesçe erişilebilir durumda ise derin-kurgu uygulamaları ile kritik vakalara yol açabilmektedir.

Derin-kurgu ile oluşabilecek kritik vakalar sadece pornografik içerik üretilmesi ile sınırlı değildir. Bu uygulamalar mahremiyetin zedelenmesine yeni bir boyut kazandırmaktadır. Örneğin kolaylıkla indirilebilen bir web uygulaması, DeepNude, kadın fotoğraflarındaki kıyafetleri yok ederek çıplak görüntüler oluşmasını sağlamaktaydı (Ajder, ve diğ., 2019).

Derin-kurgu uygulamalarının bireyler bazındaki oluşturduğu bu tehdit, söz konusu ülkeler olduğunda itibarsızlaştırma, halkı manipüle etme gibi farklı eylemleri de kapsar hâle gelmektedir. Örneğin Arjantin Devlet Başkanı Mauricio Macri'nin yüzü derin-kurgu ile Adolf Hitler'in yüzü ile (YouTube, 2018), Almanya Devlet Başkanı Angela Merkel'in yüzü ise eski ABD Başkanı Trump'ın yüzü ile (YouTube, 2018) değiştirilmiştir. Her iki derin-kurgu videosu da 360.000'den fazla görüntüleme almıştır. Siyasi isimler üzerinden gerçekleştirilen bu çıktılar, mizahın yeni boyutu olarak görülse de kişisel haklar, itibar ve mahremiyet alanında yeni tartışmalar başlatmıştır.

Yüz değişimi ile daha basit düzeyde kabul edilebilecek zararlar oluşturan derin-kurgu çıktıları, sahte konuşmaların üretilmesi ile kritik düzeye ulaşmıştır. 2018 yılında Donald Trump'ın Belçikalılara seslendiği videosu Facebook üzerinden paylaşılmıştır. 114.000'den fazla görüntüleme alan derin-kurgu çıktısında, Trump, hakaret ve küfür içeren sözler sarf ediyormuş gibi gösterilmiştir (Vooruit, 2018). Aynı yıl Barack Obama derin-kurgu uygulamalarının malzemesi olmuş, aslında hiçbir zaman yapmadığı bir konuşmanın video kaydı internette yayınlanmıştır. Bu durum derin-kurgu uygulamalarının yaratabileceği kaos potansiyeline dair ilk işaretlerden biridir (Romano, 2018). Böylece bir mizah ögesi gibi lanse edilen derin-kurgunun, ulusal ve uluslararası arenada ne türlü krizlere neden olabileceği ortaya konulmuştur.

2019 yılında Bill Posters ve Daniel Howe, Marc Zuckerberg'in Eylül 2017'de Rusya'daki seçimler hakkındaki videosunu manipüle etmiştir. Bir derin-kurgu vakası

olarak kabul edilen bu olayda, orijinal videodaki konuşmalar saptrılmış, Zuckerberg'in Facebook ile elde ettiği gücü nasıl demokrasi, mahremiyet ve hür irade üzerinde nasıl kullanabileceğine dair açıklamaları yine kendi sesinden servis edilmiştir (Cole, 2019). Bu vaka, derin-kurgu videolarının kişilerin/kurumların itibarlarını sarsmada ciddi bir tehdit olduğunu bir kez daha hatırlatmıştır.

Yine 2019 yılında Fox TV kanalının sahibi olduğu bir dijital kanal, Donald Trump'ın oval ofiste yüzünü tuhaf şekillere soktuğu, görünüşü ile alay edilebilecek mimikler yaptığı bir derin-kurgu videosunu yayınlamıştır (Swenson, 2019). Aynı yıl dönemin meşhur dizilerinden olan Breaking Bad'in sahnelerinden biri derin-kurgu kullanılarak manipüle edilmiştir. Dizinin ana karakterinin yüzü yerine Donald Trump'ın yüzü transfer edilmiş ve "nasıl para aklanabileceğine" dair ipuçları Trump'ın kendi sesinden servis edilmiştir. Video, Youtube üzerinde 750.000'nden fazla görüntülemeye sahiptir (Face, 2019). Yine 2019'da White House Sözcüsü Nancy Pelosi'yi sanki sarhoşken konuşma yapıyormuş gibi gösteren bir başka derin-kurgu videosu paylaşılmıştır. Derin-kurgu teknolojilerine kıyasla daha basit şekilde oluşturulan bu video Facebook'ta 2,5 milyondan fazla görüntüleme aldı (CBS, 2019).

2020 yılında Donald Trump'ın komik durumlara düşmesine neden olacak, alay edilecek şekilde tasarlanmış bir başka derin-kurgu videosu yayınlandı (Economist, 2020).

Derin-kurgu videoları ağırlıklı olarak ünlüler ve özellikle eski başkan Trump ile ilgili üretilse de diğer ülkelerde de farklı yansımalarla kendini göstermiştir. Örneğin, öldüğüne dair haberler dolaşan Gabon Cumhurbaşkanı Ali Bongo'nun hayatta olduğunu ve ulusuna seslendiğini gösteren bir video paylaşılmıştır. Videodaki yüz ifadesi ve duruş, Bongo'nun gerçekçiliğinden şüphe ettirerek bir derin-kurgu videosu olarak kabul edildi (Breland, 2019). Politik olarak hassas durumda olan ülkede tartışmalı videodan kısa bir süre sonra darbe girişimi meydana geldi (Washington Post, 2020).

2020 yılında Delhi'deki seçimlerde adaylardan birinin asla söylemediği sözleri içeren bir derin-kurgu videosu, farklı dillerde tüm dünya ile paylaşıldı (Christopher, 2020). Bu video tüm politik derin-kurgu vakaları içinde pozitif bir etki yaratmak için kullanılan ilk videodur.

Aynı yıl Belçika Başbakanı Sophie Wilmes'in Facebook'ta, kovid-19 ve ormanlık alanların imhası arasındaki ilişkiyi anlatan bir videosu paylaşıldı. Bu derin-kurgu vakası sadece 24 saatlik zaman diliminde 100.000'den fazla kez görüntülenmiştir (Extinction Rebellion, 2020).

Deepfake, bu şekilde ülke yöneticileri, kritik roldeki bireyleri kullanarak ağırlıklı olarak negatif bir propaganda işlevi kazanmıştır. Ancak bu derin-kurgunun dolandırıcılık amaçlı kullanıldığı örnekleri de mevcuttur.

2019 yılının Mart ayında Birleşik Krallık'taki bir enerji firmasının CEO'su, firmanın Almanya ana merkezindeki direktörü ile yaptığı telefon konuşması üzerine bir tedarikçiye 220.000€'luk para transferi yapılmasına onay vermiştir. Telefonda direktörün ödemenin aciliyetini vurgulaması, transferin bir saat içinde gerçekleştirilmesini sağlamıştır. Konuşma esnasında CEO, direktörün Alman aksanındaki tuhaflığa ve arka fondaki melodiye anlam verememiş olmasına rağmen, yeni nesil bir sahtekârlıkla karşı karşıya olduğunu düşünmemiştir (Stupp, 2019). Bir çeşit derin-kurgu kurgusu olarak kabul edilen bu olay, doğrudan maddi zarar doğuran ilk vakadır.

2021 yılında Birleşik Krallık, Letonya, Litvanya ve Estonya'dan bir grup parlamenter, Putin karşıtı Rus politikacı Alexei Navalny'nin Genelkurmay Başkanı Leonid Volkov ile görüntülü görüşme yaptıklarını zannetmişlerdir. Kendilerinden yardım talep eden Volkov'un yüzünün ardında, derin-kurgu uygulaması ile oluşturulmuş Leonid Volkov'un yüzünü kullanan bir dolandırıcı bulunmaktaydı (Vincent, 2021).

Patel v.d., (2020) tarafından yapılan çalışmada, Kasım-2018'de 7964 adet, Temmuz-2019'da 14.678, Ocak-2020'de 24.263 adet çevrim içi derin kurgu ürünü tespit edildiğine dair bilgi verilmektedir. Aynı çalışmada 2020 yılının ilk 6 ayında 2 kattan daha fazla bir artış meydana geldiği ve deepfake sayısının haziran ayında 49.081'e ulaştığı sunulmaktadır (Patel ve diğ., 2020) (Hofesmann, 2020). Derin kurgu uygulamalarının yaygınlaşması, bu şekilde agresif artışı tesadüfi değildir. Dünya genelinde sadece derin kurgu üretimine odaklanmış yirmi topluluk ve forum ortamı bulunmaktadır. Bu topluluk ve forumlardan erişilebilir durumdaki 13 tanesinin üye sayısı 90.000'den fazladır (Ajder v.d., 2019).

2020 yılı içinde üretilen derin kurgu çıktılarının %62.7'si eğlence amaçlı üretilmiş gibi servis edilmektedir (Patel v.d., 2020). Ancak bu uygulamalar basit bir eğlence unsuru gibi görülse de itibarın sarsılması, mahremiyetin zedelenmesi, şantaj, casusluk ve etkileme yoluyla zarar verme gibi ulusal ve uluslararası çapta kriz yaratma potansiyeline sahiptir. Bu potansiyel ABD Başkanlık Güvenlik Komitesinde (House Intelligence Committee) seçimlerdeki süreçlerin baltalanması açısından değerlendirilmeye alınmıştır (O'Sullivan, 2019). Artan vakalar nedeniyle derin kurgu, 8 Haziran 2021 tarihinde Amerikan Kongresinde bir ulusal güvenlik problemi olarak ele alınmıştır (Kelley ve Laurie, 2021).

Sonuç

Sahte bilgi ve içerik dünya tarihi boyunca çeşitli süreçlerde kendini göstermiş, itibarı arttırmak, sarsmak, rakipleri elemek, toplumları ve/veya bireyleri manipüle etmek amaçlı kullanılmıştır. Yapay zekânın alt çalışma alanlarından olan derin-kurgu, bu sahte bilgi ve içerik üretiminde yeni bir dönemi başlatmıştır. Özellikle kovid-19

döneminde dijital dünyanın fiziksel dünyanın önüne geçmesi, görüntü verisi bakımından sistemleri zenginleştirmiş, derin-kurgu teknolojilerinde daha performanslı uygulamaların ortaya çıkmasını sağlamıştır.

Derin kurgu uygulamalarının başlangıçtan günümüze kadar yolculuğu incelendiğinde gerçeğin sesini kısıbilecek kapasitede olduğu açıktır. Özellikle kitleleri yoğun bir biçimde sürüklerken bunu göze ve kulağa da hitap edebilen, oldukça inandırıcı bir gerçeklikle sunması ne kadar ciddi bir tehlike olduğunu göstermektedir.

Derin kurgu videolarının yayılım sıklığı ve amaçlarına bakıldığında oldukça az bir kısmının iyi niyetli ve zararsız olarak nitelenebilmektedir. Hindistan, Avrupa Birliği ve Amerika tarafında yayınlanan resmî dokümanlarda derin kurgu, yeni nesil bir silah olarak nitelenmektedir. Ancak tüm bu açıklamalara rağmen, derin-kurgu hizmeti veren uygulamaların, uygulama marketlerinden kolaylıkla indirilmesi, en çok bilindik web uygulamalarının da milyonlara varan abonesi olması, toplum genelinde konunun hâlâ eğlence ve mizah bakımından değerlendirildiğini göstermektedir.

Dijital dünya farklı alanlarda farklı yaşam deneyimleri sundukça, yapay zekâ ve beraberinde gelen teknolojiler daha performanslı hâle geldikçe, imkânlar kadar tehditler de artacaktır. Ancak derin kurgu tüm bu gelişim yolculuğu içinde şimdiye kadar gibi manipülasyon potansiyeli en kuvvetli olandır. Bu anlamda yeni filtreleme mekanizmaları ve derin kurguyu tespit eden uygulamalar geliştirilmesine daha fazla yatırım yapılması, bu uygulamaların da derin-kurgu uygulamaları kadar kolay erişilebilir hâle getirilmesi gereklidir.

Kaynakça

- Ajder, H., Patrini, G., Cavalli, F., & Cullen, L. (2019). The State of Deepfakes: Landscape, Threats and Impact. DeepTrace.
- Bawden, D., Holtham, C., & Courtney, N. (1999, October). Perspectives on information overload. In *Aslib Proceedings*, MCB UP Ltd, 51(8), 249-255.
- BBC Bitesize (2019). Deepfakes: What Are They and Why Would I Make One? [Online]. Available: <https://www.bbc.co.uk/bitesize/articles/zfkwcq>
- Bergamaschi, S., Guerra, F., & Leiba, B. (2010). Guest editors' introduction: Information overload. *IEEE Internet Computing*, 14(6), 10-13.
- Bregler, C., Covell, M., & Slaney, M. (1997). Video Rewrite: Driving Visual Speech with Audio. *Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques*, 24, s. 353-360. doi:10.1145/258734.258880
- Breland, J. (2019, March 15). The Bizarre and Terrifying Case of the "Deepfake" Video that Helped Bring an African Nation to the Brink. *Mother Jones*: <https://www.motherjones.com/politics/2019/03/deepfake-gabon-ali-bongo/> adresinden alındı

- CBS. (2019, May 26). Doctored Nancy Pelosi video highlights threat of "deepfake" tech. CBS News: <https://www.cbsnews.com/news/doctored-nancy-pelosi-video-highlights-threat-of-deepfake-tech-2019-05-25/> adresinden alındı
- Christopher, N. (18 February 2020). "We've Just Seen the First Use of Deepfakes in an Indian Election Campaign". Vice. Archived from the original on 19 February 2020.
- Cole, S. (2019, June 11). This Deepfake of Mark Zuckerberg Tests Facebook's Fake Video Policies.
- Cole, S (19 June 2018). Gfycat's AI Solution for Fighting Deepfakes Isn't Working. Vice. Archived from the original on 8 November 2019.
- Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine learning*, 20(3), 273-297.
- Deluca, J., Christodoulou, C., Diamond, B. J., Rosenstein, E. D., Kramer, N., & Natelson, B. H. (2004). Working memory deficits in chronic fatigue syndrome: differentiating between speed and accuracy of information processing. *Journal of the International Neuropsychological Society*, 10(1), 101-109.
- Deng, L. & Yu, D. (2014). Deep Learning: Methods and Applications. *Foundations and Trends® in Signal Processing*: 7(3-4), 197-387. <http://dx.doi.org/10.1561/20000000039>
- Dickson, E. J. (7 October 2019). "Deepfake Porn Is Still a Threat, Particularly for K-Pop Stars". Rolling Stone. Archived from the original on 30 October 2019.
- Economist (28 April 2020). Amabie: the mythical creature making a coronavirus comeback. *The Economist*. ISSN 0013-0613. Archived from the original on 20 May 2021.
- Eelmaa, S. (25 March 2021). Sexualization of Children in Deepfakes and Hentai: Examining Reddit User Views. SocArxiv. doi:10.31235/osf.io/6wuhj. Archived from the original on 1 April 2021.
- Edmunds, A., & Morris, A. (2000). The problem of information overload in business organisations: a review of the literature. *International Journal of Information Management*, 20(1), 17-28.
- Eppler, M. J., & Mengis, J. (2004). The concept of information overload: A review of literature from organization science, accounting, marketing, MIS, and related disciplines. *The Information Society*, 20(5), 325-344.
- Face, C. S. (2019, September 19). Better Call Trump: Money Laundering 101. Youtube: <https://www.youtube.com/watch?v=Ho9h0ouemWQ&t=41s> adresinden alındı
- Farfadi, S. S., Saberian, M. J., & Li, L. J. (2015, June). Multi-view face detection using deep convolutional neural networks. In *Proceedings of the 5th ACM on International Conference on Multimedia Retrieval ACM*, 643-650.
- Farhoom and, A. F., & Drury, D. H. (2002). Overload. *Communications of the ACM*, 45(10), 127.
- Fukushima, K. (1980). Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological cybernetics*, 36(4), 193-202.
- Generated Media. (2018). 100,000 Faces Generated by AI. *Generated Photos*: <https://generated.photos/> adresinden alındı
- Goggin, Benjamin. "From porn to 'Game of Thrones': How deepfakes and realistic-looking fake videos hit it big". Business Insider. Archived from the original on 8 November 2019.
- Himma, K. E. (2007). The concept of information overload: A preliminary step in understanding the nature of a harmful information-related condition. *Ethics and Information Technology*, 9(4), 259-272.
- Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. *Science*, 313(5786), 504-507.

- Hofesmann, E. (2020, November 19). The State of Deepfakes in 2020. SKYNET TODAY: Putting AI News In Perspective: <https://www.skynettoday.com/overviews/state-of-deepfakes-2020> adresinden alındı
- Jacoby, J. 1977. Information load and decision quality: Some contested issues. *Journal of Marketing Research*, 14, 569–573.
- Jacoby, J., Speller, D. E., and Berning, C. K. 1974. Brand choice behavior as a function of information load: Replication and extension. *Journal of Consumer Research* 1, 33–43.
- Jacoby, J. 1984. Perspectives on information overload. *Journal of Consumer Research* 10, 432-436.
- Karras, T., Laine, S., Aila, T. (2019). A Style-Based Generator Architecture for Generative Adversarial Networks. in Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019
- Kelley, T. M., & Laurie, A. H. (2021). Deep Fakes and National Security. Washington: Congressional Research Service.
- Kominiarczuk, N., & Ledzińska, M. (2014). Turn down the noise: Information overload, conscientiousness and their connection to individual well-being. *Personality and Individual Differences*, 60, 76.
- Korshunov, P., & Marcel, S. (2018, December 20). Deepfakes: a New Threat to Face Recognition Assessment and Detection. (C. University, Dü.) <https://arxiv.org/abs/1812.08685> adresinden alındı
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems* (pp. 1097-1105).
- Lee, Dave (3 February 2018). "'Fake porn' has serious consequences". Archived from the original on 1 December 2019.
- Malhotra, N. K. (1982). Information load and consumer decision making. *Journal of Consumer Research*, 8, 419-430.
- McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5(4), 115-133.
- Meyer, J. A. (1998). Information overload in marketing management. *Marketing Intelligence & Planning*, 16(3), 200-209.
- Mikolov, T., Karafiát, M., Burget, L., Černocký, J., Khudanpur, S. (2010). Recurrent neural network based language model. In *INTERSPEECH-2010*, 1045-1048.
- O'Sullivan, D. (2019, June 29). Congress to investigate deepfakes as doctored Pelosi video causes stir. (CNN, Röportaj Yapan)
- Quayyum, A., Quadir, J., Janjua, M. U., & Vira, F. S. (2019). Using Blockchain to Rein in the New Post-Truth World and Check the Spread of Fake News. *IT Professionals*, 21(4), 16-24. doi:10.1109/MITP.2019.2910503
- Özdemir, Ş. (2016). Individual contributions to infollution (information pollution): trust and share. *International Journal on New Trends in Education and Their Implications*, 7(3), 23-33
- Ranjan, R., Sankaranarayanan, S., Bansal, A., Bodla, N., Chen, J. C., Patel, V. M. & Chellappa, R. (2018). Deep learning for understanding faces: Machines may be just as good, or better, than humans. *IEEE Signal Processing Magazine*, 35(1), 66-83.
- Roettgers, J. (21 February 2018). "Porn Producers Offer to Help Hollywood Take Down Deepfake Videos". *Variety*. Archived from the original on 10 June 2019.

- Romano, A. (18 April 2018). "Jordan Peele's simulated Obama PSA is a double-edged warning against fake news". Vox. Archived from the original on 11 June 2019.
- Patel, M., Gupta, A., Tanwar, S., Obaidat, M. (2020). Trans-DF: A Transfer Learning-based end-to-end Deepfake Detector. 10.1109/ICCCA49541.2020.9250803.
- Salakhutdinov, R., & Hinton, G. (2009, April). Deep boltzmann machines. In *Artificial Intelligence and Statistics*, 448-455.
- Sparrow, P. (1999). Strategy and cognition: Understanding the role of management knowledge structures, organizational memory and information overload. *Creativity and Innovation Management*, 8(2), 140-148.
- Somers, M. (2020, July 21). Deepfakes, explained. (M. M. School, Dü.) Boston, Usa.
- Stupp, C. (2019, August 30). Fraudsters Used AI to Mimic CEO's Voice in Unusual Cybercrime Case. USA: Wall Street Journal. Ağustos 11, 2021 tarihinde <https://www.wsj.com/articles/fraudsters-use-ai-to-mimic-ceos-voice-in-unusual-cybercrime-case-11567157402> adresinden alındı
- Swenson, Kyle (11 January 2019). "A Seattle TV station aired doctored footage of Trump's Oval Office speech. The employee has been fired". The Washington Post. Archived from the original on 15 April 2019.
- Taigman, Y., Yang, M., Ranzato, M. A., & Wolf, L. (2014). Deepface: Closing the gap to human-level performance in face verification. In *Proceedings of the IEEE Conference On Computer Vision and Pattern Recognition* (pp. 1701-1708)
- Taslitz, A. E. (2013). Information overload, multi-tasking, and the socially networked jury: Why prosecutors should approach the media gingerly. *The Journal of the Legal Profession*, 37, 89-138.
- "#TellTheTruthBelgium". Extinction Rebellion Belgium. Archived from the original on 25 April 2020.
- Vincent, J. (2021, April 30). 'Deepfake' that supposedly fooled European politicians was just a look-alike, say pranksters. The Verge: <https://www.theverge.com/2021/4/30/22407264/deepfake-european-politicians-leonid-volkov-van-lexus> adresinden alındı
- Vooruit. (2018, May 19). Vooruit. Facebook: <https://www.facebook.com/vooruit.nu/videos/10155618434657151/?t=33> adresinden alındı
- Wang, M. & Deng, W.(2018). Deep face recognition: A survey, ArXiv, 1804.06655.
- Washington Post. (2020, February 13). The suspicious video that helped spark an attempted coup in Gabon. Youtube: <https://www.youtube.com/watch?v=F5vzKs4z1dc> adresinden alındı
- Weng, J., Ahuja, N., & Huang, T. S. (1992, June). Cresceptron: a self-organizing neural network which grows adaptively. In [Proceedings 1992] IJCNN International Joint Conference on Neural Networks (Vol. 1, pp. 576-581). IEEE.
- Wurman, R. S. (1989). *Information Anxiety: What to do When Information Doesn't Tell You What You Need To Know*. New York.
- YouTube. (2018, January 17). Adolf Hitler "Downfall Movie" to Mauricio Macri. (Faceswap Deeplearning). YouTube: <https://www.youtube.com/watch?v=M8t6hGrtDac&t=1s> adresinden alındı
- YouTube. (2018, January 28). Merkel Trump Deepfake. YouTube: <https://www.youtube.com/watch?v=5hZOcmqWKzY&t=2s> adresinden alındı