Deepfake Bir Tehdit mi Fırsat mı?

Is Deepfake a Threat or an Opportunity?

Hatice Kübra Çiçek[1], Nursel Yalçın[2]

[1]Arş. Gör., Kırşehir Ahi Evran Üniversitesi, Kaman Uygulamalı Bilimler Yüksekokulu, Kırşehir, Türkiye
[2]Doç. Dr., Gazi Üniversitesi, Eğitim Fakültesi, Ankara, Türkiye

** Corresponding author: hkubra.cicek@gazi.edu.tr*

Geliş Tarihi / Received: 22.12.2023
Kabul Tarihi / Accepted: 10.02.2024
Derleme Makalesi/Review Article
DOI: 10.5281/zenodo.11531297

ÖZET

Deepfake teknolojisi, son yıllarda hızla gelişen yapay zeka ve derin öğrenme algoritmalarıyla oluşturulan sahte ses ve görüntü içeriklerini ifade etmektedir. Bu teknoloji, hem olumlu hem de olumsuz yönleriyle dikkat çekmektedir. Bu araştırma, deepfake teknolojisinin potansiyel tehditlerini ve fırsatlarını incelemeyi amaçlamaktadır. Sahte haber ve propaganda, kimlik avı, dolandırıcılık, sahtecilik, kişisel gizlilik ihlalleri gibi olumsuzlukların yanında sanat, film, eğitim, eğlence ve sahte haberlere karşı mücadele gibi siber güvenlik alanlarına ise olumlu anlamda katkı sunmaktadır. Bu araştırma, deepfake teknolojisinin güvenlik, gizlilik ve etik konularıyla birlikte eğitim, eğlence, sanat ve siber güvenlik gibi alanlardaki pozitif etkilerini değerlendirmektedir. Sosyal medya kullanımıyla birlikte bu teknoloji milyonlarca insana hızla ulaşmakta ve toplumu her açıdan etkilemektedir. Sahte ses ve görüntü içeriklerinin hızla yayılabilmesi bu teknolojilerin potansiyel etkilerini daha da artırmaktadır. Toplumu, siyaseti ve iş dünyasını ilgilendiren deepfake teknolojilerinin yükselişi, güvenlik ve etik sorunlarını da beraberinde getirmektedir. Bu doğrultuda bu teknolojilerin tehdit ve fırsatlarının ele alınması, çalışma için önem arz etmektedir. Doküman analizi yöntemiyle ilgili veri tabanlarından (Google Scholar, Wiley, Science Direct) 2018-2023 yılları arasında elde edilen ulusal ve uluslararası makaleler incelenerek değerlendirilmiştir. Bu çalışmadan elde edilen sonuçlar, deepfake ile ilgili potansiyel tehditlere karşı hazırlıklı olunmasına yardımcı olmakla birlikte, etik ve güvenilir bir şekilde kullanılmasına da katkıda bulunabilir.

Anahtar Kelimeler: sahte görüntü, sahte ses, deepfake, deepfake tehdit, güvenlik, siber güvenlik

ABSTRACT

Deepfake technology refers to the creation of fake audio and visual content using rapidly advancing artificial intelligence and deep learning algorithms. This technology has gained attention due to its both positive and negative aspects. This research aims to examine the potential threats and opportunities presented by deepfake technology. While it poses negative implications such as fake news, propaganda, phishing, fraud, and violations of personal privacy, it also contributes positively to areas like art, film, education, entertainment, and the fight against fake news in the field of cybersecurity. This study evaluates the positive effects of deepfake technology in fields such as security, privacy, and ethics, alongside the negative consequences. With the use of social media, this technology rapidly reaches millions of people, impacting society in various ways. The swift dissemination of fake audio and visual content further amplifies the potential effects of these technologies. The rise of deepfake technologies, which concern society, politics, and business, also brings security and ethical issues. Therefore, it is essential to explain the threats and opportunities of these technologies. This research conducts a document analysis by examining national and international articles from databases like Google Scholar, Wiley, and Science Direct, spanning the

years 2018-2023. The findings from this study can help prepare for potential threats related to deepfake technology and contribute to its ethical and reliable use.

Keywords: fake image, fake audio, deepfake, deepfake threat, security, cybersecurity

GİRİŞ

Bilgi çağında yaşanan gelişmeler, yaşamımızın önemli bir kısmını oluşturmaktadır ve bu süreç, bireyleri ve toplumu derinlemesine etkileyerek bilgi toplumuna dönüşüm sürecini başlatmıştır (Çil, 2023). Bu dönüşümle birlikte sınırlar ortadan kalkmış ve internetin geniş bir etki alanına yayılması, birey ve toplumu değişim ve dönüşüm sürecine itmiştir. Ancak, birçok fırsat sunmasının yanı sıra, teknolojik gelişmeler beraberinde çeşitli riskler ve tehditleri de getirmektedir (Baş & Şenol, 2023). Dijitalleşme çağında, yapay zeka, nesnelerin interneti ve deepfake gibi yeni teknolojiler, bu hızlı değişimin öne çıkan örneklerindendir. Özellikle, "deepfake" teknolojisi, dijitalleşme sürecinde hızla büyüyerek yayılmaktadır. Bu bağlamda, bu teknolojinin potansiyel tehditlerini belirlemek, aynı zamanda olumlu yönlerini değerlendirmek, toplum, siyaset ve iş dünyası üzerindeki etkilerini ortaya koymak araştırma için önemlidir.

Teknolojinin hızlı ilerlemesiyle birlikte, yapay zeka ve derin öğrenme algoritmalarının etkileyici bir şekilde evrim geçirdiği günümüzde, deepfake teknolojisi, sahte ses ve görüntü içeriklerinin oluşturulmasında önemli bir role sahiptir. Bu teknoloji, gerçeklikten ayırt edilemeyen sahte içerikler üretme kapasitesiyle dikkat çekerken, aynı zamanda potansiyel tehditleri de beraberinde getirmektedir. Bu araştırma, deepfake teknolojisinin geniş bir perspektifte incelenmesini amaçlayarak, hem olumlu hem de olumsuz yönlerini anlamak için bir zemin oluşturmaktadır.

Deepfake'in, sahte haber ve propaganda gibi olumsuz etkilerinin yanı sıra, sanat, film, eğitim gibi alanlarda da nasıl pozitif bir rol üstlenebileceği önemli bir konudur. Bu bağlamda, teknolojinin sunduğu fırsatlar ve bu fırsatların güvenlik, etik ve gizlilikle nasıl dengelenebileceği bu araştırmada ele alınacaktır.

Araştırmanın temel odak noktalarından biri, deepfake'in siber güvenlik, kişisel gizlilik, etik değerler ve toplumsal etkiler gibi geniş bir yelpazedeki konularda nasıl bir etkiye sahip olduğunu anlamaktır. Ayrıca, bu teknolojinin eğitim ve sanat gibi alanlarda nasıl kullanılabileceği ve bu kullanımların olası katkıları da detaylı bir şekilde incelenecektir.

Bu araştırma, deepfake teknolojisinin hızla evrim geçiren doğasını anlamak, potansiyel risklere karşı farkındalık oluşturmak ve bu teknolojinin etik kurallar çerçevesinde nasıl yönetilebileceği konularında bir perspektif sunmayı hedeflemektedir.

Deepfake Teknolojisinin Tanımı

Deepfake teknolojisi, yapay zeka ve derin öğrenme algoritmalarının entegrasyonuyla gerçekleştirilen, sahte ses ve görüntü içeriklerinin üretilmesini sağlayan bir yenilik olarak tanımlanabilir. "Deep" ve "Fake" kelimelerinin birleşiminden oluşur; "deep" burada derin öğrenme mimarisine vurgu yaparken, "fake" sahte anlamını taşır (Korkmaz & Alkan, 2023).Deepfake teknolojisinin ilk ortaya çıkışı, kimliği bilinmeyen bir kullanıcı tarafından Reddit sosyal medya platformunda gerçekleşti. Reddit, farklı konularda küçük topluluklar oluşturan ve içerik paylaşımı ve tartışmaya olanak sağlayan bir platformdur. Kasım 2017'de, yalnızca "u/deepfakes" olarak bilinen bir kullanıcı, bu teknolojiyi kullanan videoları paylaşmak amacıyla r/deepfakes adında bir topluluk oluşturdu. Bu topluluk, deepfake algoritmasını kullanan ilk yüz değiştiren videoların paylaşılmaya başlandığı yerdir (Fikse, 2018). Bu teknoloji, genellikle gerçek bir kişinin yüz ifadesini, konuşma tarzını veya sesini bir başka kişiye ait olanla değiştirmek amacıyla kullanılmaktadır. Günümüzde, bu tür videoların oluşturulması giderek daha kolay hale geldi ve gerçekte söylemedikleri veya yapmadıkları şeyleri inandırıcı bir şekilde ifade eden, ikna edici ve

tanınması zor tasvirler haline geldi. Teknolojinin ilerlemesiyle birlikte, ünlülerden politikacılara ve genel olarak hükümetlere kadar birçok kişiyi hedef alan sahte ancak oldukça inandırıcı videoların oluşturulma kapasitesi giderek artmaktadır (Pieterse & Botha, 2020).Derin öğrenme algoritmaları, büyük veri setleri üzerinde eğitilerek, hedeflenen kişinin yüz özelliklerini ve konuşma tarzını hassas bir şekilde taklit edebilmektedir. Deepfake teknolojisi ile kişilerin ses ve görüntülerini manipüle etmek, gerçeklikle neredeyse ayırt edilemeyecek kadar inandırıcı hale getirmek mümkündür (Kırık & Özkoçak, 2023). Bu teknolojinin kullanımının artmasıyla birlikte sahte içerikler üretilmekte ve yayılımıyla çeşitli etik, güvenlik ve hukuki sorunlar da ortaya çıkmaktadır. Deepfake teknolojisinin etkileri topluma, şirketlere ve bireylere zarar verebilir, itibar kayıplarına neden olabilir (İdiman, 2021)

Deepfake Teknolojisinin Potansiyel Tehditleri

Polis soruşturmalarında ve mahkemelerde kullanılan fotoğraf ve videolar hukuki davaları neticelendirmek için güvenilir kaynak olarak kabul edilmekte ve delil olarak kullanılmaktadır. Son zamanlarda fotoğraf ve videolar üzerinde bazı değişiklikler gerçekleşmekte ve düzenlemeler yapılmaktadır. Bu değişiklikler sonucunda delillerin de güvenilirliğinin sağlanması gerektiği görülmektedir (Albahar & Almalki , 2019).Teknolojinin hızla gelişmesiyle fotoğraf ve video dosyalarının manipülasyonu oldukça kolaylaşmıştır ve mahkemeye sunulmadan önce kullanılacak fotoğraf ve video gibi delillerin kontrol edilmesi gerektiği gündeme gelmektedir (Chesney & Citron, 2018). Bu alandaki en bilinen manipülasyon tekniği ise "deepfake" olarak adlandırılmaktadır. Bir kişinin başka bir kişinin yüzüyle değiştirilmesi, sesi üzerinde düzenlemeler yapılması, intikam almak, şantaj yapmak veya sahtecilik yaparak suç işlemek gibi birçok amaçla farklı alanlarda deepfake teknolojisinin kullanıldığı görülmektedir. Bu teknolojiler özellikle güvenlik ve sosyal etkileşim gibi alanlarda ciddi potansiyel tehlikeleri beraberinde getirmiştir.

1. Manipülasyon ve Sahtecilik

Bilgi çağı olarak nitelendirildiğimiz bu çağın en önemli aktörü bilgi ve iletişim teknolojileridir. Bu teknolojilerin gelişmesi ve yaygınlaşması ve sunduğu ürün ve hizmetlerin artmasıyla birlikte bilgiye erişim kolaylaşmıştır. Hayatın her alanında yer alan teknolojiler ile bilgiye kolaylıkla erişirken aynı zamanda doğru bilgiye ulaşmak için bir mücadele içine girilmiştir (Özdemir, 2021). Web 2.0 teknolojisi sayesinde, web platformlarında sadece içeriklere erişmekle kalmayıp, bu içerikleri düzenleme, güncelleme, ekleyip kaldırma gibi çeşitli etkileşimlerde bulunmak mümkün hale gelmiştir. Ayrıca, kullanıcılar kendi orijinal içeriklerini oluşturarak web ortamını daha aktif ve etkileşimli bir hale getirebilmektedir. (Güler, Bayzan, & Güneş, 2016).

Bu teknolojiler ile birlikte bilginin artışı, hızlı yayılışı, yanlış ve doğru bilginin ayırt edilmemesinden doğan bilgi kirliliği, bilginin değiştirilerek manipüle edilmesi ve sahte bilgilerin ortaya çıkması gibi sorunlar meydana gelmektedir. Özellikle bilginin manipüle edilmesi ve sahte haberlerin ortaya çıkması toplumsal açıdan bir tehdit oluşturmaktadır (Özdemir, 2021). Kişilerin yüz ifadelerini ve ses tonlarını başka bir kişiye ait gibi manipüle etme yeteneğine sahip olan deepfake teknolojisi, sahte video ve ses dosyalarının üretilmesine olanak tanımaktadır. Şahısların, sözlerinin ve eylemlerinin yanıltıcı bir şekilde değiştirilerek kullanılmasına yol açmaktadır. Ünlülerin, politikacıların veya sıradan bireylerin bu teknolojiler kullanılarak manipüle edilmesi itibar kayıplarına sebep olabilmektedir (Arslan, Deepfake Technology: A Criminological Literature Review, 2023). Deepfake ile üretilen görüntüler gerçekleklik anlayışının sorgulanarak bilgi ekosisteminde güvenilirlik ve inandırıcılık noktasında ciddi bir krize sebebiyet verebilecek düzeydedir. Günlük hayatta sıkça kullandığımız fotoğraf, video, ses gibi veriler de önemli bir rol oynamakta ve çok fazla bilgi içermektedir. Görüntü ve ses gibi dosyaların internet teknolojileriyle hızlı şekilde yayılması ve paylaşılması konusu da önem arz etmektedir. Bu doğrultuda, değişikliğe uğramış görüntülerin paylaşılması ve hızla yayılması hayatı ciddi ölçüde etkileme potansiyeline sahiptir (Kırık & Özkoçak, 2023).

Son yıllarda sosyal medya platformları üzerinden birçok ülkenin manipülasyona uğradığı gözlenmektedir. Modern iletişim teknolojilerinin gelişimi, yapay zekadan sanal gerçekliğe doğru hızlı ilerleyiş yeni ve zor problemleri de beraberinde getirmektedir. Yapay zekadaki ilerlemeler, verileri ayrıştırmak ve sosyal medyadaki kullanıcılar için içeriği önceliklendirmek için daha etkili yöntemler yaratmış olmasının yanı sıra endişe verici bir şekilde, bilginin paylaşılma sürecinde de değişikliklere sebebiyet vermektedir. Dijital araçlarımızdaki ilerlemeler iletişim teknolojisinde ve toplumda büyük değişikliklere neden olmaktadır (Woolley, 2020, s. 20). Bireyin ve toplumun karar verme sürecini etkileyen ve farklı şekillerde algılamasına sebebiyet veren bu teknolojiler, karar alma süreçlerini manipüle etme potansiyelini taşımaktadır (Jacoby, 1984). Herhangi bir yanlış bilgi, manipülatif haber, sahte inceleme veya aldatmaca da insan topluluğunu olumsuz yönde etkilemekte ve bilgi kirliliğine sebebiyet vermektedir (Meel & Vishwakarma, 2020). Teknolojinin gelişimiyle birlikte sahte haberler de farklı bir boyuta taşınmıştır. Daha öncesinde metinsel ve paylaşımlar halinde yayılan haberler yapay zeka ve makine öğrenmesi teknolojileri kullanarak kurgulanmış video içerikleriyle oluşturulmaya başlanmıştır (Özdemir, 2021). Bilgi teknolojileri ve yapay zeka alanındaki gelişmeler önemli zorlukları da beraberinde getirmektedir. Kurumlar, işletmeler, devletler ve sivillerin sanal ağ yapılarına dahil olması ve finansal işlemler, savunma, sağlık, kişiel veri gibi sisstemleri kullanması siber tehditlere yönelik savunmasızlığı artırmaktadır. Siyasal iletişim alanında da teknolojilerin kullanılmasıyla birlikte sahte haberler ve propagandalar yapılmaya başlanmıştır. Özellikle deepfake teknolojilerinin de gelişmesiyle birlikte görüntülerin tespit edilmesi giderek zorlaşmaktadır (Ruiter, The Distinct Wrong of Deepfakes, 2021). Sahte içeriklerin daha gerçekçi ve ikna edici hale gelmesinin sonucunda dijital taklitçilik gündeme gelmektedir. Özellikle siyasi liderlerin ve önde gelen kişilerin yer aldığı görüntüler ve videolar önem arz etmektedir. Çünkü ulusal ve uluslararası alanda krizleri başlatan sahte haberler için kullanılabilir (Caporusso, 2021).

Deepfake teknolojileriyle oluşturulmuş sahte haberler, geleneksel sahte haberlere göre daha büyük bir tehdit oluşturmaktadır. Yeni dijital teknolojilerin gelişimi, gerçek bilgilerle sahte bilgileri birbirinden ayırt etmeyi daha zor bir hale getirmektedir. Ayrıca insanlar sahte olarak yapılan içeriklere gerçek olduğuna inanmaya meyilli olduğu görülmüştür (Westerlund, 2019).

2. Mahremiyet ve Gizlilik

Deepfake teknolojisinin kullanım alanları çok geniş ve çeşitlidir. Bu teknoloji, mahremiyet ve gizlilik alanında ciddi endişelere yol açan bir dizi potansiyel sorunu beraberinde getirmektedir. Bu teknolojinin en belirgin kullanım alanlarından biri, kişilerin ses ve görüntülerinin izinleri olmadan manipüle edilerek sahte içerikler oluşturulmasıdır. İnsanları gerçekte söylemedikleri ve yapmadıkları şeyleri söylüyor ve yapıyormuş gibi gösteren görüntüler tartışmasız mahremiyet ve gizliliklerini ihlal ettiklerini göstermektedir (Ruiter, The Distinct Wrong of Deepfakes, 2021). Sosyal ağların yaygın kullanılmaya başlanmasıyla birlikte tüm veri kümeleri açığa çıktığından, paylaşılan tüm ses ve görüntü dosyalarına ulaşılabilmesi ve kullanılabilmesi bir endişeye sebebiyet vermektedir (Caporusso, 2021).

Geniş kimlik manipülasyon yöntemlerinin varlığı “görmek inanmaktır” gerçeğini karmaşıklaştırmaktadır. Deepfake teknolojisinin kullanım durumu zarar vermekle sınırlı değildir ve eğlence materyali oluşturmak veya model performansını artırmak için etkili bir şekilde kullanılabilse de, yoğun bir şekilde olumsuz amaçlar için kullanılmaktadır. Bu nedenle, sorunun ciddiyetine bakarak mahremiyet ve gizlilik konusunda önlemler almak gerekmektedir. (Chapter 8)

Deepfake'ler, bireylerin görüntülerini ve seslerini sahte bir şekilde kullanarak, gizliliklerini doğrudan ihlal edebilir. Özellikle, kişisel yaşamın mahremiyeti bu teknolojinin kullanımıyla ciddi şekilde tehlikeye girebilmektedir. Dijital teknoloji çağında, açık kaynaklı yazılımlar sayesinde bireyler, teknik bilgiye sahip olmasalar bile deepfake içerikleri oluşturabilmektedirler (Westerlund, 2019).

Deepfake teknolojisi, başlangıçta eğlence amaçlı kullanım için ortaya çıktı ancak zamanla daha karanlık ve tehlikeli alanlara da yöneldi. Sahte videoların oluşturulmasıyla birlikte, bireylerin yüzleri ve sesleri, izinleri olmadan manipüle edilebilmektedir. Bu durum, mahremiyetin ve gizliliğin korunmasını zorlaştırmaktadır. Bu nedenle, tehditlerin farklı olası yönleri hakkında farkındalık kazanmak yalnızca tespit etmek için değil manipüle edilmiş videoların etkisini azaltmak için etkili güvenlik mekanizmaları geliştirmek önemlidir.

3. Dolandırıcılık ve Yargı Sistemlerinde Rolü

Deepfake teknolojilerinin kullanım alanlarından biri de işletmeleri dolandırmak ve kuruluşlara siber güvenlik noktasında problemler oluşturmak amacıyla kullanılabilmektedir. Bu alan özellikle adli bilişimi ilgilendirmekte ve yargı sistemlerinde de yanlış bilgi kaynağı ve delil olarak kullanılmasına sebebiyet vermeketedir (Buo, 2020). Delil manipülasyonu, yargı sistemlerinde deepfake teknolojileri tarafından oluşturulan önemli sorunlardan birini oluşturmaktadır. Hukuk mahkemelerinde deliller, deepfake kullanılarak davayı etkilemek amacıyla manipüle edilebilir. Derin taklitler sonucunda oluşturulan delillerin mahkemeye sunulması ve karşı tarafın itirazı üzerine yeniden incelemeler yapılması zaman ve para kaybına da sebebiyet vermektedir.

Bir video kaydının güvenilir bulunması durumunda mahkemeye sunulması ve delil olarak kabul edilmesinin yeterli olduğu belirtilmektedir. Fakat günümüzde taklitlerin orijinal içerikte ayırt edilmesi giderek zorlaşmaktadır. Örneğin, Birleşik Krallık'ta bir çocuk velayet davasında, bir anne, çocuklarına erişime izin verilmemesi gereken kadar şiddetli olduğunu iddia etmek amacıyla babasının kendisine tehditler içeren bir kayıt gibi ses çalan derin taklit bir ses dosyasını mahkemeye delil olarak sundu. Anne, bu dosyayı oluşturmak için derin taklit teknolojisi ve çevrimiçi eğitimleri kullanmıştı. Ancak, dosya adli olarak incelendikten sonra sahte olduğu kanıtlandı ve mahkeme tarafından reddedildi (Swerling, 2020). Bu sebeple, gelecekteki mahkeme davalarında delil manipülasyonunu önlemek için yeni ve etkili karşı önlemlere ihtiyaç duyulmaktadır.

İşletmeler ve bireyler üzerinde finansal açıdan olumsuz etkiler oluşturabilir. E-posta dolandırıcılığı gibi sosyal mühendislik saldırıları ile birleştirildiğinde, deepfake kullanılarak üretilen ses ve görüntüler işletmeleri dolandırmak için kullanılabilir ve bu işletmenin itibarini etkileyebilir (Buo, 2020). Dolandırıcılık da derin sahte uygulamalarının gelişmesiyle farklı bir boyut kazanmıştır. Geçmişte, telefon operatöründen veya müşteri hizmetlerinden aranıldığında dolandırıcılık olduğundan şüphelenebilir ve işlem yapmadan telefonu kapatılabilirdi. Günümüzde dolandırıcılık senaryoları da gelişmeye başlamıştır. Şirketinizin CEO'sunun sesinin taklit edilerek kullanıldığı ve sizin arandığınız bir senaryoyu hayal ettiğinizde sonuç farklı bir şekilde bitebilir (Arslan, Deepfake Technology: A Criminological Literature, 2023). Örneğin, dolandırıcılar, Birleşik Krallık merkezli bir şirketi CEO'yu taklit ederek dolandırdı. Finans departmanındaki çalışanlara, dolandırıcıların kontrol ettiği bir hesaba 220,000 dolar transfer etmeleri için ikna etti. Yine başka bir örnekte, Hong Kong'daki bir bankanın müdürü, satın alma işlemi sırasında tanıdık bir ses tarafından arandı ve 35 milyon dolarlık bir transfer gerçekleşmesi istendi (Brewster, 2021). Bu teknoloji kullanılarak gerçekleştirilen dolandırıcıkların ciddi mağduriyetler oluşturabileceği görülmektedir (Durmuş & Göztaş, 2023). Geleneksel aldatmacalarla karşılaştırıldığında deepfakelerin oluşturduğu tehditlerin kapsamının firmalar için önemli ölçüde daha büyük olduğunu göstermektedir (Johnson & Diakopoulos, 2021). Bunlar arasında karalama ve sabotaj gibi aşağılayıcı faaliyetlerin yanı sıra firmanın imajına, itibarına ve güvenilirliğine zarar verilmesi de yer almaktadır (Botha & Pieterse , 2020). Deepfake'lerin çoğalmasıyla birlikte şirketler iftira ve sabotaj gibi aşağılayıcı faaliyetlere maruz kalmaktadır, bu da pazardaki aldatma yoluyla şirketin itibarını ve marka imajını tehdit edebilmektedir. Bu doğrultuda müşterilerin ve diğer paydaş gruplarının güveninin kaybolmasına neden olabilmekte ve şirketleri zarara uğratabilmektedir (Domenico & Visentin, 2020).

Deepfake Teknolojisinin Olumlu Yönleri

Eğitim: Yapay zeka ve derin öğrenme tekniklerinin eğitim alanında kullanılması yeni olanakları ve fırsatları beraberinde getirmektedir. Deepfake teknolojisi sayesinde eskiden yaşamış önemli figürlerin derslerinin canlandırılması mümkün hale gelmektedir. Örneğin tarihte önemli yer edinmiş insanların eğitim amaçlı kullanılması tarih derslerine olan ilgiyi artırabilir (Chesney & Citron, 2019). Özellikle video tabanlı uzaktan eğitim sistemleri daha ilginç hale getirilebilir. Örneğin Einstein'ın fizik öğretmesini ya da Immanuel Kant'ın felsefe öğretmesini deepfake teknolojisi kullanarak canlandırılması ilgiyi artırabilir (Temir, 2020). Deepfake teknolojisi, eğitimde yenilikçi yöntemlerin keşfedilmesine olanak tanıyan bir araç olabilir. Ancak, bu teknolojinin etik kullanımı ve öğrenci mahremiyeti konularında dikkatli bir şekilde yönetilmesi önemlidir. Eğitimciler ve teknoloji uzmanları, deepfake'in eğitimdeki potansiyel avantajlarından en iyi şekilde faydalanabilir. Deepfake teknolojileri, eğitimcilere farklı materyal ve metot kullanabilmelerine yönelik fırsatlar sunmaktadır. Öğretmeyi daha eğlenceli ve ilgi çekici haline getiren yöntemlerin kullanılmasına olanak tanımaktadır. Gandhi veya Nelson Mandela gibi tarihi kişiliklerin kendileri ve çalışmaları hakkında eğitim verdiği videoların deepfake ile oluşturularak öğrencilere içerik olarak sunulması buna örnek olabilir (Dagar & Vishwakarma, 2022).

Eğlence: Deepfake teknolojisi, film yapımcılarına klasik sahneleri yeniden oluşturma veya artık yaşamayan aktörleri kullanarak yeni filmler yapma olanağı sunmaktadır. Örneğin, 2019'da küresel ısınma farkındalığı kampanyası, endüstrinin deepfake teknolojisini kullanarak çeşitli dillerde farklı izleyici kitlesini çekmeyi amaçlayan bir video oluşturmasını sağladı. Reklam, David Beckham'ın görüntüsünü ve sesini değiştiren çok dilli bir video sunuyordu (Westerlund, 2019).

Sağlık: Deepfake teknolojisi, sosyal ve sağlık alanlarında yeni olanaklar sunmaktadır. Yapay zeka (YZ) alanındaki ilerlemelerle, deepfake teknolojisi, Alzheimer hastalığı teşhisi konulan bireylerin geçmişlerini potansiyel olarak hatırlamalarına yardımcı olabilir, onları genç yıllarındaki bir videosuyla yeniden bağlantı kurabilmesine olanak tanımaktadır (Westerlund, 2019). Yapay zeka destekli deepfake'ler, yakın zamanda kaybedilen yakınlarının konuşma tarzını video ve ses manipülasyonu yoluyla yeniden oluşturmak için kullanılabilir, böylece bireyler yas süreçlerine yardımcı olmak için onlarla tekrar bağlantı kurabilmelerini sağlar (Ruiter, The Distinct Wrong of Deepfakes, 2021). ALS gibi hastalıklar yüzünden konuşamayan kişilerin sesini yapay olarak yeniden üretebilen yazılımlar bulunmaktadır ve bu araçlarla bu bireylerin sesini kullanmasına olanak tanınmaktadır.

Sanat: Özellikle sinema sektöründeki fırsatlara bakıldığında, kullanılan efektler, canlandırmalar ve doğasında olan sahte görselleştirmeler yeni olmasa da, Deepfake teknolojisinin getirdiği yenilikler, sinema için sınırları belirlenmemiş derin öğrenmeye dayalı bambaşka bir dünyanın kapısını aralayacak ve hatta bu kapıyı çoktan açmış gibi görünüyor. Deepfake teknolojisinin, artık hayatta olmayan oyuncuları canlandırması, kuşkusuz getirdiği en önemli yeniliklerden biridir. Ülkemizde bir banka reklamında Deepfake ile Türk sinema sanatçısı Kemal Sunal'ın canlandırılması, bu teknolojinin geleneksel kitle iletişim araçlarına ve reklam çalışmalarına nasıl katkı sağlayabileceğini göstermektedir. Deepfake'in sinema sektöründeki etkileri sadece ölü oyuncuları canlandırmakla sınırlı değildir. Ayrıca, amatör videoların kalitesini iyileştirme, yüksek maliyetle yapılan üç boyutlu çekimleri daha ekonomik hale getirme ve yarım kalan projelerin tamamlanma maliyetlerini düşürme potansiyeline sahiptir. Geçmişte, Hızlı ve Öfkeli 7 filmi, çekimleri sırasında hayatını kaybeden başrol oyuncusu Paul Walker'ın yüzünün, kardeşinin yüzüyle yer değiştirilerek tamamlanmıştır. Bu örnek, Deepfake'in sinema endüstrisinde nasıl başarılı bir şekilde kullanılabileceğine dair etkileyici bir örnek sunmaktadır (Durmuş & Göztaş, 2023).

Siber Güvenlik ve Sahte Haberlere Karşı Mücadele: Deepfake teknolojisinin yükselişi yalnızca toplum için değil aynı zamanda birçok işletme, anayasal siyasi sistemim ve ulusal güvenlik için de büyük bir siber güvenlik tehdidi haline gelmektedir (Chesney & Citron, 2019). Aynı zamanda

teknolojide yeni gelişmeler ortaya çıktıkça ve siber suçlular yenilikçi saldırı geliştirme konusunda daha karmaşık teknolojileri kullanmaya başladıkları görülmektedir. Siber suçlular, spam ve kötü amaçlı yazılım türlerini kullanarak yıllardır sosyal mühendislik saldırıları yoluyla işletmeleri hedef almakta ve bireyleri gizli bilgileri ifşa etme ve zayıf yönlerden yararlanma konusunda manipüle etmek için çeşitli kimlik avı kampanyaları başlatmaktadır (Arslan, Deepfake Technology: A Criminological Literature, 2023). Deepfake teknolojisi ise bu saldırıların tespit edilmesini zorlaştıran başka bir tehdit yöntemidir. Deepfake video ve ses içeriği, birçok işletme ve toplum için giderek büyüyen bir siber güvenlik tehdidi haline gelmesi bu doğrultuda mücadele için de yeni meslek kollarının oluşmasına imkan tanımaktadır.

Deepfake Teknolojisinin Yayılması ve Toplumsal Etkileri

Son zamanlarda elde edilen kanıtlar, yanlış bilginin yayılmasının yanlış yönlendirmelere sebebiyet verebileceğimi ve toplumda ciddi sorunları tetikleyebileceğini göstermektedir (Fedeli, 2020). Kullanıcıları sanal ve fiziksel etkileşimlere dahil eden sanal, artırılmış ve karna gerçeklik gibi sürükleyici ancak manipüle edilebilir deneyimlerin ortaya çıkmasına neden olmuştur (Sigala, 2018). Deepfake ise gerçek ve sahte bilgi arasındaki ayrımı belirsizleştiren içeriği kolayca üretmek için çeşitli medya biçimlerini birleştirerek değiştirebilmektedir.

Kötü niyetli olarak kullanılan ve internette dolaşan sahte videoların çoğalması yanlış bilginin yayılması noktasında endişeler oluşturmaktadır. Özellikle tanınmış kişilerin, politikacıların veya ünlülerin yapmadıkları şeyleri söylediği veya yaptığı gerçekçi görünümlü videolar veya ses kayıtları oluşturmak için kullanılabilir. Bu, yanlış bilgi yaymak, kamuoyunu manipüle etmek ve seçimleri aksatmak için kullanılabilir.

Yapılan çalışmaların çoğu kaydedilmiş videolar üzerinde yapılan deepfake teknolojisine vurgu yapmaktadır. Ancak makine öğrenimi ile gerçek zamanlı deepfake üretmeyi mümkün kılacak kadar gelişmiştir. Göz bakışının başka bir yere bakıyor olmasına rağmen kameraya doğru çevrilmesini sağlamak buna örnektir. Bu filtreler, bir video konferansın video içeriğinin gerçek zamanlı olarak değiştirilmesine ve optimize edilmesine olanak tanır. Örneğin konferans sırasında göz bakışının başka bir yere bakıyor olmasına rağmen kameraya yönelik görünmesini sağlamak gibi değişiklikleri uygulamak deepfake teknolojisi ile mümkün hale gelmeye başlamıştır (Hancock & Bailenson, 2021). Araçlar daha erişilebilir ve verimli hale geldikçe, kötü niyetli kullanıcıların teknolojiyi kâr amacıyla kullanmanın yollarını bulması kolaylaşmaktadır. Hem şirketleri hem de bireyleri hedef alan derin sahte uygulamaları ile kimlik avı saldırılarında ve dolandırıcılıklarda bir artış görmesi beklenmektedir. Deepfake oluşturma ve tespit teknikleri üzerine yapılan araştırmada, deepfakelerin para kazanma açısından kullanıldığı ve ilerde bunun daha fazla kullanılacağı görülmektedir (Mirsky & Lee, 2020). Hukuk, siyaset, medya, işletmeler gibi birçok alanı etkileyen bu teknolojilerin potansiyel kullanımı ve toplumsal etkileriyle birlikte incelenmesi önemlidir (Karnouskos, 2020). Pornografi alanında sıkça adından bahsedilen bu teknolojilerin çocuk pornografisindeki kullanımına ilişkin endişeleri söz konusudur (Eelmaa, 2021). Kadın fotoğraflarından kıyafetleri çıkarıp, gerçekçi çıplak görüntü elde edilmesini sağlayan DeepNude uygulaması da olumsuz bir etkiyi göstermektedir (Greengard, 2019). Çocuk ve kadın pornografisi gibi rahatsız edici kullanımlarının yanı sıra siyasi alanda da kullanılarak demokrasiyi etkilemektedir. Siyasi anlamda kamuoyu da farkındalık oluşturarak eğitmek gibi olumlu örnekleri de olsa genellikle yanlış bilgi yayma niyetinde olduğu görülmektedir (Gamage, Chen, Ghasiya, & Sasahara, 2022). Yapılan araştırmaya göre, insanların politik deepfakeler aracılığıyla nasıl aldatılabileceğini anlamaya çalışılmıştır ve toplumun deepfakeler tarafından tamamen yanıltılmaktan ziyade belirsizlik hisleriyle karşı karşıya olduğu görüşü ortaya çıkmıştır. Ancak ortaya çıkan belirsizlik sosyal medyadaki haberlere olan güveni de azaltmaktadır (Vaccari & Chadwick, 2020).

YÖNTEM

2018-2023 yılları arasında deepfake teknolojisi üzerine yayımlanan ulusal ve uluslararası makalelerin sistematik bir incelemesi yapılmıştır. Bu inceleme, Google Scholar, Wiley ve Science Direct veritabanlarında "deepfakes" ve ilgili terimler ("yüz manipülasyonu", "sahte video", "video manipülasyonu", "sahte ses", "ses manipülasyonu") kullanılarak gerçekleştirilmiştir. İlk 100 sonuç alaka düzeyine göre sıralanmış, başlık ve özetler bu anahtar kelimelere göre Wiley ve Science Direct veritabanlarında incelenmiştir. 2018-2023 arasında yayınlanan toplam 49 makale seçilmiş ve doküman analizi yöntemiyle detaylı bir şekilde incelenerek içerdikleri bilgiler kategorize edilmiş ve değerlendirilmiştir. Bu süreç, deepfake kavramının 2017 yılında ortaya çıkışını temel alarak gerçekleştirilmiştir.

Sıra	Makale Yazarları ve Yılı	Bulguların Özeti
1	Albahar ve Almalki (2019)	Deepfake videoların oluşturulması ve tespiti için kullanılan tekniklerin detaylı incelemesini gerçekleştirmiştir. Yapay zeka ve makine öğrenimi yöntemleri ile deepfake oluşturma süreçleri ve tespit algoritmaları ele alınmıştır.
2	Botha ve Pieterse (2020)	Sahte haberler ve deepfake'lerin oluşturulması ve tespiti üzerine kapsamlı bir analiz yapılmıştır. Bu çalışma, geniş bir perspektif sunarak deepfake ve diğer sahte haber türlerini ele almaktadır.
3	Carvajal ve Iliadis (2020)	Deepfake teknolojisi üzerine akademik çalışmaların önemli bir literatür incelemesi yapılmıştır. Deepfake teknolojisinin gelişimi ve araştırmalarındaki ana trendler detaylandırılmıştır.
4	Godulla ve ark.	İletişim çalışmaları alanında deepfake konusunda yapılan araştırmaların literatür taraması yapılmıştır. Deepfake'lerin medya ve iletişim alanındaki etkileri ve tartışmaları derinlemesine incelenmiştir.
5	Verdoliva (2020)	Manipüle edilmiş görüntü ve videoların tespiti için kullanılan yöntemlerin detaylı incelemesi gerçekleşmiştir. Deepfake tespiti için kullanılan tekniklerin gelişimi ve etkinliği değerlendirilmiştir.
6	Westerlund (2019)	Çevrimiçi haber makaleleri üzerinden deepfake teknolojisinin analizine yönelik literatür taramasıdır. Deepfake'lerin medya ve haber endüstrisi üzerindeki etkileri incelenmiştir.
7	Ahmed (2021)	Deepfake teknolojisinin sosyal ve demokratik sonuçları üzerine tartışmalar mevcuttur. Deepfake'lerin toplum ve demokrasi üzerindeki etkileri ele alınmıştır.
8	Somoray ve Miller (2023)	İnsanların deepfake'leri tespit etme yeteneğinin incelenmesidir. Deepfake tespitinde insan algısının sınırları ve etkinliği değerlendirilmiştir.
9	Rouiter (2021)	Deepfake teknolojisinin etik kullanımı üzerine bir araştırmadır. Deepfake'lerin etik boyutları ve sorumlu kullanımı için öneriler sunulmuştur.

Bu tablo, deepfake teknolojisi üzerine yapılan çeşitli akademik çalışmaların ana bulgularını özetlemektedir.

BULGULAR

Teknolojik gelişmelerde yaşanan hızlı değişim ve dönüşümlerle, teknolojinin her gün farklı bir boyuta taşındığı görülmektedir. Deepfake oluşturma tekniklerinin gelişmesi ve ilerlemesi, tespit edilmesini zorlaştırmaktadır. İnsanlar bu teknolojiler yüzünden çevrimiçi içeriğe olan güvenlerini kaybetmeye başlamışlardır. Herhangi bir kişinin bu teknolojileri kullanarak kötü amaçlı içerikler oluşturabilmeleri mümkündür (Mahmud & Sharmin, 2021). İnternet ve ağ teknoloji hızının

ilerlemesiyle yayılımı oldukça kolaylaşan ve dünyadaki gelişmeleri etkileyen bir hal almıştır. Birçok olumsuz özelliği olmasının yanında film, eğitim ve diğer alanlarda ise olumlu kullanımları mevcuttur. Deepfake oluşturma yöntemlerini kullanarak sesini kaybetmiş bir bireyin sesini geri döndürmek mümkündür. Deepfake teknolojisinin kullanımının hızla artış gösterdiği ve bu artışın sadece alanda çalışan uzmanlar arasında değil, sıradan insanlar arasında da yaşandığı görülmektedir. Bu doğrultuda deepfake teknolojisinin kontrol edilebilmesi zorlaşmakta ve mağdurlar için bir tehdit oluşturabileceğine dair görüşler bulunmaktadır (Arslan, Deepfake Technology: A Criminological Literature Review, 2023). Özellikle gazetecilik alanında sahte haberlerin yayılmasına yol açan deepfake teknolojilerine önem verilmesi doğrultusunda görüşler bulunmaktadır. Büyük verinin olduğu bu çağda, doğru bilginin güvenli şekilde paylaşımı oldukça zordur. Deepfake teknolojilerine karşı önlem almak ve bu videoları analiz ederek paylaşmak oldukça önemli bir konudur (Temir, 2020). Yapılan çalışmada, gelecekteki teknolojik gelişmelerle birlikte sahte videoların görsel kalitesi ve üretim verimliliğimin artırılacağı tahmin edilmektedir. Günümüzde deepfake oluşturma yöntemlerinin dezavantajlarından biri, cilt ve yüz kılları gibi detayları üretememeleridir. Son çalışmalar ve GAN modellerinin dahil edilmesiyle iyileştirilebileceğine dair tartışmalar mevcuttur. Teknolojinin ve sahtecilik yöntemlerinin ilerlemesinin sonucunda tespit yöntemlerinin de ilerlemesi önemli bir konudur ve bu noktada büyük bir rekabet olduğu görülmektedir (Lyu, 2020). Deepfake teknolojisinden orijinal görüntüleri yeniden oluşturabilen yöntem üzerine yapılan çalışmada, hukuk alanında medya içeriğinin manipüle edildiğinin kanıtlanabilmesinin önemli olduğuna vurgu yapılmaktadır. Mahkeme süreçlerinde ve adli vakalarda karşılaşılan en büyük zorluklardan birinin manipüle edilen unsurların belirlenerek deepfake ile elde edilmiş kanıtının açıklanması olduğu görülmektedir (Guarnera, ve diğerleri, 2022). Deepfake teknolojisinin henüz emekleme evresinde olduğunu belirten bir çalışma, GAN teknolojisinin ortaya çıkışını araştırmakta ve deepfakelerin potansiyel etkisine ilişkin birçok cevaplanmamış soru olduğunu belirtmektedir. Bu teknolojilerin kötüye kullanmalarıyla ile ilgili endişelerin yönetişim ve kimlik doğrulama ile mekanizmaları iyileştirileceğinden bahsedilmektedir (Kwok & Koh, 2021). Manipülasyona dair çok az kanıt bulunan, üzerinde oynanmış fotoğraf ve videoların sosyal medyada hızla yayılması oldukça kolaydır. Özellikle kurumsal ve kamuya mal olmuş kişilerin deepfake medyası, bireylerin itibarlarının zedelenmesine yol açabilir ve ünlülerin marka elçisi olarak hareket ettiği kuruluşlara da yayılabilir. Gerçek ile sahteyi ayırt edememek, yanlış beyana, kafa karışıklığına ve güvensizliğe yol açabilir ve derin sahte uygulamalarının hızlı ilerlemesi, sonuçta güvenlik ve mahremiyet ihlali gibi sadece imaj ve markalamayla sınırlı olmayan diğer benzeri görülmemiş risklere yol açabilir (Bazarkina & Pashentsev, 2020). Bu doğrultuda önerilen aşılama teorisine göre önceden maruz kalma, bireylerin gelecekteki tehditlere karşı korunmasına yardımcı olabilir. Deepfake bağlamında, teknoloji hakkında bilgi sunarak ve toplumu videoları eleştirel bir şekilde yorumlamaya ikna eden girişimler, bireylerin kötü niyetli olarak sahte videolara karşı "aşılanmasına" yardımcı olabilir. Farkındalığı artırmanın yanı sıra, deepfake tespitine yönelik araçların daha da geliştirilmesi önemlidir görüşü ortaya çıkmaktadır (Blauth, Gstrein, & Zwitter, 2022).

Yapılan bir çalışmada, "kimlik avı" (phishing) ve "deepfake" teknolojisinin kullanılması, aynı zamanda akıllı botların bilgi kampanyalarında çeşitli amaçlar için kullanılmasından bahsedilmektedir. Bu tür tekniklerin kullanılmasıyla, bir kişi, kuruluş veya ülkenin itibarı zarar görebilir. Örneğin, rakip bir kişinin, kuruluşun veya ülkenin itibarını zedelemek amacıyla manipülatif bilgi kampanyalarının yapılması söz konusu olabilir. Sonuç olarak, psikolojik yöntemlerin yaygın kullanımı, bilinçli bir şekilde toplumun algısını yönlendirme ve etkilemek için kullanılabilir. Bu durum, özellikle kriz dönemlerinde, insanların duygusal tepkilerini daha etkili bir şekilde yönetmek için kullanılabilir ve bu da toplumda belirli bir politika veya görüşü destekleme amacına hizmet edebilir (Bazarkina & Pashentsev, 2020).

Deepfake teknolojileri, saldırganlara başkalarının kimliğine bürünerek yanlış bilgi yayma yeteneği vermektedir. Suçluların kuruluşun çalışanlarını kandırmak amacıyla şirket yöneticileri gibi davranarak sağlam deepfakeler kullandığı örnekler mevcuttur (Pashentsev & Bazarkina, 2022). Tüm yeni teknolojiler gibi deepfake teknolojisinin de kullanım şeklinin gelişmeye devam edeceği öngörülmektedir. Bu teknolojilerin oluşturduğu tehditlere karşı basit bir çözüm yolu bulunmamaktadır. Teknolojiyi iyiye veya kötüye kullandığına dair tespit edilmesine yönelik geliştirilen bir araç mevcut değildir. Eğitimin artırılması, medya okuryazarlığının desteklenmesi gibi önlemler bu tehditlerle mücadele edilebilir (Brown, 2020). Dezenformasyon yani yanlış bilgilendirme güçlü, ucuz ve genellikle ekonomik açıdan kârlı bir etkileme aracı olarak bilinmektedir. 'Derin sahtekarlıklar' olarak adlandırılan sahte resimler ve görsel-işitsel içerikler oluşturmak için yeni, uygun fiyatlı ve kullanımı kolay teknolojiler artık mevcuttur ve kamuoyunu manipüle etmek için daha fazlasını sunmaktadır (Pashentsev & Bazarkina, 2022). Facebook, Instagram ve WhatsApp gibi sosyal medya platformlarında sahte haberlerin ve kolay üretilebilir olan deepfakelerin hızla yayılması bir endişeye neden olmaktadır. Bu doğrultuda geliştirilen bir mobil uygulama ile internet bağlantısına ihtiyaç olmadan deepfakelerin tespit edilmesi ve gerçekliği ortaya çıkararak tanıyan bir modelin geliştirilmesi planlanmıştır. Böylelikle bu uygulama ile manipüle edilmiş dijital medyanın yayılmasını önlenebilir veya sınırlanarak güvenilir kaynakları kontrol etmelerine yardımcı olabilmesine olanak tanınabilir (Vamsi, ve diğerleri, 2022).

İnsanların deepfakeleri tespit etme yeteneğini inceleyen çalışmada bireylere sahte resim ve videolar gösterilmiştir. Bireylerin deepfakeleri tespit etme yeteneğinin genellikle zayıf olduğunu görülmüştür ve insanlara basit stratejiler sunarak tespit edilmesinin anlamlı bir şekilde iyileşip iyileşmediği gözlemlenmiştir (Somoray & Miller, 2023). Gerçek zamanlı deepfake oluşturma sürecine çok yaklaşıldığını tahmin edilmektedir. Bu noktaya ulaşıldığında, siber uzay bir değişime uğrayabilir. Deepfake oluşturmaya uygun çeşitli medyalara kolay erişim sayesinde, daha ayrıntılı deepfake saldırıları gerçekleşmesi mümkün olabilir. İnsan ve makine tabanlı tespit yöntemleri üzerinde çalışmalar devam etmektedir ve geliştirmeye yönelik çeşitli girişimler bulunmaktadır (Firc, Malinka, & Hanacek, 2023). Deepfake teknolojisinin önümüzdeki yıllarda daha da popüler hale geleceği görülmektedir. Bu teknolojinin olumlu kullanım örnekleri olmasına rağmen birçok olumsuz etkisi de vardır ve sadece kişileri değil sosyal kurumlar üzerinde de büyük olumsuz etkileri olabilmektedir. Bunlar arasında dolandırıcılık, kimlik hırsızlığı, itibar zedelenmesinin yanı sıra medyanın, hukukun üstünlüğünün ve demokrasinin zedelenmesi yer almaktadır (Sloot & Wagensveld, 2022). Yapay zeka teknolojilerinin kötüye kullanıma ilişkin yasal düzenlemeler, Amerika'da bile başlangıç aşamasındadır. Farklı ülkelerde de ilk yasal adımlar atılmaya başlanmıştır. 2018'de AB, yanlış bilgilendirmeyi genel olarak inceleyen ve derin sahtekarlıklara karşı koruma yönergeleri de dahil olmak üzere bir dizi yönerge sağlayan 'Çevrimiçi Dezenformasyonla Mücadele: Avrupa Yaklaşımı' belgesini yayınlamıştır (Pashentsev & Bazarkina, 2022). Toplumun gelişen teknolojiyle birlikte gerçek ve sahte yönleri tespit edememesi ve bunun da paydaşlara, süreçlere ve gazeteciliğe olan güveni azaltmasının sonucunda "her şey sahtedir" sloganı hakim olabilir. Bu tür sahte haberlerin tanımlanması, doğrulanması ve kaldırılması için teknik çözümlere ihtiyaç duyulsa da, sorun yalnızca teknolojik değildir, düzenleyici ve eğitsel yöntemlere de ihtiyaç duyulmaktadır (Karnouskos, 2020).

Yanlış bilgi, özellikle siyasi seçimler ve günlük bilgi akışı üzerindeki ciddi etkisi nedeniyle internet yönetiminin önemli bir sorunu haline gelmektedir (Chin, Park, & Li, 2022). Sosyal medyada dezenformasyon ve uydurma içeriğin oluşturulması, yayılması ve tüketilmesi, özellikle bu tür kaynaklara erişimin kolaylığı ve bu tür yanlış bilgilerin varlığına ilişkin farkındalığın eksikliği nedeniyle giderek artan bir endişe kaynağıdır. Deepfake teknolojisinin etik kullanıma ilişkin bir inceleme yapan araştırmada, doğası gereği ahlaki açıdan bir yanlış olmadığı söylenmektedir. Ahlaki açıdan yanlış olmasında üç temel sebep olması gerektiği söylenmektedir. Bunlar, deepfake yapılan kişinin temsil edilme şekline itirazı, izleyiciyi aldatması ve oluşturulma amacı olarak belirlenmiştir.

Bu teknolojileri ahlaki açıdan yanlış kılan en belirgin husus, kişilerin görüntü ve seslerine karşılık gelen dijital verilerin, onları tasvir edilmek istemeyecekleri şekilde kullanılması şeklinde yorumlanmıştır (Ruiter, The Distinct Wrong of Deepfakes, 2021). Deepfake teknolojileri ile ünlülerin yüzlerinin pornografik videolarda kullanılması, kamuoyunun düşüncelerini değiştirme ve seçim prosedürlerini etkileme potansiyeline sahip sahte politikacı videoları oluşturma gibi çoğunlukla kötü niyetli amaçlarla kullanıldığı görülmektedir. Bir videonun deepfake dönüşümü içim genellikle yüzün orta kısmını farklı bir yüzünkiyle değiştirerek oluşturulmaktadır. Bu dönüşüm sırasında doğal olmayan bazı dizilim farklılıkları ortaya çıkabilir ve bu da tutarsız kafa duruşlarına neden olmaktadır. Bu ince ve küçük ayrıntıların, yanlış hizalamaların bir insan tarafından fark edilmesi zor olabilir, ancak bir model tarafından gerçek ve derin sahte videolar arasında ayrım yapmak öğrenilebilir (Shu, ve diğerleri, 2020).

Tarihsel olarak, deepfake'in azaltılması için öncelikle tespit ve kaldırmaya odaklanmış, deepfake oluşturmada insan faktörünü ele alan önleyici tedbirler büyük ölçüde göz ardı edilmiştir. İçerik oluşturucularının eylemlerinin ardındaki motivasyonları ve sonuçları inceleyerek youtubedaki deepfake eğilimlerini inceleyen araştırmada, içerik oluşturucular arasında endişe verici bir etik farkındalığı eksikliğine işaret ettiği görülmektedir. Bu durum, pornografi, şantaj ve dezenformasyon gibi derin sahte bilgilerin kötü amaçlı kullanımlarını potansiyel olarak kolaylaştırmaktadır (Nick, 2023). Deepfake tespit tekniklerinin iyileştirilmesinin yanı sıra kötü niyetli olmayan derin sahte üretim prosedürlerinin oluşturulmasını ve kullanımını sınırlandırmak için bazı kurallar ve düzenlemeler getirilmesi önemlidir. Dijital içeriğin geçmişte kötü niyetli olarak veya yararlı amaçlarla tahrif edildiği açıktır, bu teknolojilerle sahte içerikteki gerçekçilik düzeyini artırmıştır. Araştırmacılar ve hükümet, konuyu tartışmak ve bu yıkıcı deepfake teknolojisiyle mücadele etmek için birlikte çalışarak daha hızlı yol alabilir, çünkü kendi başına yol alması daha zor görünmektedir (Dixit, Kaur, & Kingra, 2023). Çin'de kötüye kullanıma yönelik en bilinen örnek, deepfake kullanımının yasaklanması olmuştur. 2019'da Çin, internetteki video ve ses içeriğini düzenleyen, yayın ve dağıtım yasağı da dahil olmak üzere yeni kurallar belirlemiştir. Çin'de Siber Uzay İdaresi'nin yayınladığı bir metne göre deepfake teknolojileri, ulusal güvenliği tehlikeye atabilir, sosyal istikrarı ve düzeni bozabilir ve başkalarının meşru haklarını ihlal edebilir. Bu nedenle, bu teknolojilerin her türlü kullanımı açıkça işaretlenmeli ve internet kullanıcıları tarafından da açıkça görülebilmesi gerektiğine dair görüşler sunulmuştur. Aslında burada yasaklanan bu teknolojilerin kullanımı değildir, teknoloji aracılığıyla halkı yanıltmaktır (Pashentsev & Bazarkina, 2022). Artık deepfake teknolojisinin olumlu ve yapıcı uygulamaları arttığı için deepfake üretmeyi ve yaymayı cezalandıran politikaları uygulamak noktasında farklı yaklaşımlar benimsenmesi gerekmektedir. Yapay olarak güçlendirilmiş bir dijital ortamda, insanlar hangi bilgilere güvenilebileceği konusunda bilinçli kararlar verme hakkına sahiptir. Manipüle edilmiş görüntülerde anormal bölgelerin varlığının görsel olarak belirtilmesinden ve bu görüntülerin doğru koşullar altında şüpheli üreticiler tarafından yeniden üretilebilirliğinin gösterilmesi tespit edilmesi noktasında büyük ölçüde fayda sağlayacaktır (Khoo, Phan, & Lim, 2022). Derin sahtekarlıkların doğasını anlamaları ve insanların bunları nasıl tespit edebileceği konusunda eğitim ihtiyacını gerektirmektedir (Zeng, Song, Guo, & Lian, 2023).

TARTIŞMA VE SONUÇ

Günümüzde derin öğrenme ve yapay sinir ağları gibi tekniklerin hızla gelişmesinin sonucunda yapay zeka alanında önemli bir ilerleme yaşanmıştır. "Deepfake" terimi genel olarak yapay zekanın kullanılmasıyla oluşturulan uydurma görsel ve işitsel medyayı ifade eder. Çeşitli araştırmaların bulgularına göre, deepfake videoların ilk ortaya çıkışı 2017'ye kadar izlenebilse de, bunların önemi ve kamu farkındalığının 2019'dan bu yana belirgin şekilde arttığı açıktır. Deepfake teknolojisinin uygulama alanları sanat ve eğlence, reklam ve pazarlama, film endüstrisinin yanı sıra politik

iletişim ve medya gibi çeşitli alanları kapsamaktadır. Deepfake'lerle ilgili mevcut araştırmaların büyük çoğunluğunun, özellikle sahte haberleri güçlendiren bir faktör olarak, bunların aldatma amacıyla kullanılmasına yönelik olduğu görülmektedir. Deepfake teknolojisinin faydalı etkileri incelendiğinde ise bireylerin yaratıcı yeteneklerinin geliştirilmesini kolaylaştırmaktadır. Yapay zeka teknolojileri ve deepfake kullanılarak sanat alanındaki insan yaratıcılığını artırma potansiyelini öne süren araştırmalar mevcuttur.

Araştırma makalelerinin profili, deepfake araştırmalarının 2019'dan bu yana son üç yılda önemli ölçüde ilgi gördüğü ve bu doğrultuda çalışmaların sayısının arttığı gözlemlenmektedir. Mevcut literatüre bakıldığında, deepfake yöntemleri, tehditleri ve risklerine yönelik çalışmaların, verisetleri üzerinden incelemelerin olduğu görülmektedir. Teknik disiplinlerdeki çoğu çalışma, derin sahtekarlıkların neden olduğu potansiyel zararlara odaklanır ve bu nedenle tespit tekniklerinin geliştirilmesine yönelik çalışmalar mevcuttur. Deepfake kullanımının potansiyel faydalarını ele alan birkaç çalışma bulunmaktadır. Bu alanlardan birisi moda sektörüdür. "Derin moda" terimi altında deepfake, kıyafetleri sanal olarak denemeyi veya bunları kişinin kendi vücuduna yansıtmayı mümkün kılmaktadır.

Deepfake teknolojisi, bireyler, işletmeler ve toplum üzerinde önemli tehditler oluşturabilir:

Bireylere Yönelik Tehditler: Deepfake, bireylerin rızası olmadan onları zararlı veya aşağılayıcı durumlarda göstererek psikolojik stres ve somut zarara neden olabilir. Deepfake videoları, şantaj, taciz veya itibar zedeleme amacıyla kullanılabilir.

İşletmelere Yönelik Tehditler: Deepfake teknolojisi, iş liderleri veya CEO'ların seslerini taklit ederek dolandırıcılık yapılmasını mümkün kılar. Piyasa manipülasyonu, marka değerinin düşürülmesi ve şirket itibarına zarar verme gibi riskler barındırır.

Topluma Yönelik Tehditler: Deepfake, gerçek ve sahte bilgi arasındaki sınırları bulanıklaştırarak, gazetecilik ve kurumlara olan güveni zayıflatabilir. Yanlış bilginin yayılması toplumsal kaos ve bölünmeye yol açabilir.

Uluslara Yönelik Tehditler: Deepfake, ulusal ve uluslararası ilişkileri bozabilir, demokratik süreçlere müdahale olasılığını artırabilir ve iç güvenliği tehlikeye atabilir. Bu teknolojinin yayılması, politik tartışmaları ve veri bütünlüğünü etkileyebilir (Dagar & Vishwakarma, 2022).

Deepfake tespit konularının alana hakim olduğunu ve yükseliş eğiliminde olduğunu görülmektedir. Çünkü deepfakeler artık çok yaygın ve daha gerçekçi hale geldikçe, daha sağlam dedektörlerin geliştirilmesine ihtiyaç duyulmaktadır. Görsel deepfakeler daha fazla araştırma ilgisi çekerken, işitsel deepfakeler de yavaş yavaş ilgi görmeye başlamıştır. Özellikle sahte haber üretiminde deepfakelerin tanınmasının erken fark edildiğini ve üzerinde çalışılmaya devam edildiğini görülmektedir. Son olarak, disiplinler arası derin sahte araştırma potansiyeli çok büyüktür. Yasal ve gizlilik konularına odaklanmak önemlidir bu endişelerin hâlâ teknolojik boyutların gerisinde kaldığını ve derin sahtekarlıkların olumsuz kullanımlarına ilişkin daha fazla araştırmaya ihtiyaç olduğunu göstermektedir. Gizlilik ve hukuki konuların yanı sıra diğer teknik olmayan alanlar da önemlidir ve daha fazla çaba gösterilmesi gerekmektedir. Buna, insan odaklı deepfake tespiti ve deepfake'lerin bireyler ve toplum üzerindeki etkisi gibi potansiyel yeni alanlar da eklenerek farklı çalışmalar yapılabilir.

ÖNERİLER

Deepfake teknolojisi, hızla gelişen yapay zeka ve derin öğrenme algoritmalarıyla birlikte sahte ses ve görüntü içeriklerinin oluşturulmasını sağlayan bir alandır. Gelecekte, bu teknolojinin önemli etkileri ve beraberinde getirdiği zorluklar göz önünde bulundurulmalıdır. Deepfake teknolojisinin gelecekte geniş bir kullanıcı kitlesi tarafından benimsenmesi ve popüler hale gelmesi muhtemeldir.

Deepfake üretimindeki teknik gelişmelerle birlikte, oluşturulan sahte içeriklerin gerçekçiliği artacaktır. Deepfake'lerin sosyal, politik ve ekonomik alanlarda daha fazla etkisi olabilir. Bu doğrultuda, gelecekteki çalışmalar, daha etkili deepfake tespit yöntemleri üzerine odaklanarak gelişmiş algoritmalar ve tekniklerle bu tespit süreçleri güçlendirilebilir. Deepfake kullanımının yaygınlaşmasıyla birlikte, güvenlik ve etik standartları daha da önem kazanacaktır. Bu bağlamda, etik yönergelerin ve güvenlik protokollerinin oluşturulması önemlidir. Toplumun, deepfake'lerin varlığı ve etkileri konusunda bilinçlendirilmesi önemli bir husustur ve bu teknolojinin hukuki boyutları ele alınmalı, mevcut yasal çerçeve güçlendirilebilir ve yeni düzenlemeler ortaya konulabilir. Deepfake'lerin potansiyel kötü niyetli kullanımına karşı siber güvenlik alanında yatırımlar artırılarak ve bu alanda çalışan uzmanlara daha fazla destek sağlanarak gelişimine katkı sunulabilir. Bu çalışma önerileri, deepfake teknolojisinin geleceğine dair önemli zorlukları aşma ve toplumu koruma amacını taşımaktadır. Yüksek düzeyde işbirliği ve multidisipliner bir yaklaşım, bu alandaki sorunları çözmek için kritik öneme sahiptir.

Deepfake içeriğinin kullanıcıya açıkça gösterilmesi, insanların bu teknolojinin varlığı konusunda bilinçlenmelerine ve farkındalıklarının artmasına yardımcı olabilir. Bu sayede, insanlar sahte içerikleri daha kolay tanıma ve onlarla başa çıkma konusunda daha bilgili hale gelebilir. Deepfake teknolojisinin nasıl çalıştığının ve oluşturulan içeriğin hangi amaçlarla kullanıldığının açıkça gösterilmesi aynı zamanda eğitim ve bilgilendirme amacıyla kullanılabilir. Kullanıcılar, bu teknolojinin potansiyel etkilerini anlamak için interaktif eğitim materyalleri ve etkileşimli platformlar kullanabilirler. Deepfake içeriğinin şeffaf bir şekilde sunulması, kullanıcılar arasında güven inşa edebilir. Toplum, bu teknolojinin kullanımını denetleyen ve şeffaflığı sağlayan düzenleyici çerçevelere daha fazla güvenebilir. Ayrıca, sanat ve eğlence sektörlerinde yaratıcı bir potansiyelin ortaya çıkmasına olanak tanır. Deepfake içeriğinin açıkça gösterilmesi, bu teknolojinin kontrollü bir şekilde kullanılmasına olanak tanır. Yasal düzenlemeler ve etik kurallar altında, deepfake teknolojisi belirli kullanım alanlarında kullanılabilecek ve kötü niyetli amaçlara karşı önlemler alınabilecektir. Ancak, bu avantajların yanında dikkat edilmesi gereken etik sorumluluklar ve güvenlik konuları da vardır. Yasal düzenlemelerin, kullanım sınırlamalarının ve şeffaf standartların belirlenmesi, deepfake teknolojisinin etik bir çerçevede kullanılmasını sağlamak adına önemlidir.

KAYNAKLAR

Albahar, M., & Almalki , J. (2019). Deepfakes: Threats And Countermeasures Systematıc Revıew. *Journal of Theoretical and Applied Information Technology*, 3242-3250.

Arslan, F. (2023). Deepfake Technology: A Criminological Literature Review. *Sakarya Üniversitesi Hukuk Fakültesi Dergisi*, 701-720.

Aslan, E. (2014). Yabancı Dil Öğretiminde Robot Öğretmenler. *Ondokuz Mayıs Üniversitesi Eğitim Fakültesi Dergisi*, 15-26.

Baş, M. H., & Şenol, A. (2023). Derin Kurgu (Deepfake) Araçları ile Üretilen Resimlerin Adli Analizi ve Derin Kurgu Tespiti. *Fırat Üniversitesi Fen Bilimleri Dergisi*, 97-118.

Bazarkina, D. Y., & Pashentsev, E. N. (2020). Malicious Use Of Artificial Intelligence. *Russia in Global Affairs*, 154-177.

Blauth, T., Gstrein, O. J., & Zwitter, A. (2022). Artificial Intelligence Crime: An Overview of Malicious Use and Abuse of AI. *IEEE Access*, 77110-77122.

Brewster, T. (2021, Ekim 14). *Innovation: Forbes*. Forbes Web sitesi: https://www.forbes.com/sites/thomasbrewster/2021/10/14/huge-bank-fraud-uses-deep-fake-voice-tech-to-steal-millions/?sh=76f853e07559 adresinden alındı

Brown, N. I. (2020). Deepfakes and the Weaponization of Disinformation. *Virginia Journal of Law & Technology*, 1-59.

Buo, S. A. (2020). The Emerging Threats of Deepfake Attacks and Conutermeasures. *arXiv preprint arXiv:2012.07989*.

Caporusso, N. (2021). Deepfakes for the Good: a Beneficial Application of Contentious Artificial Intelligence Technology. *Proceedings of the AFHE 2020 virtual conferences on software and systems engineering, and artifcial intelligence and social computing*, (s. 235-241). Springer.

Chesney, B., & Citron, D. (2019). Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security. *California Law Review*, 1753-1820.

Chesney, R., & Citron, D. K. (2018). Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security. *California Law Review*, 1753-1820.

Chin, Y. C., Park, A., & Li, K. (2022). A comparative study on false informationgovernance in Chinese and American socialmedia platforms. *Policy Internet*, 263-283.

Çil, S. (2023). Türkiye'de Deepfake Teknolojisi: Youtube'da En Çok İzlenen Türkçe Deepfake Videolar Ve İzleyicileri Üzerine İnceleme. *Uluslararası İletişim ve Sanat Dergisi*, 173-195.

Dagar, D., & Vishwakarma, D. K. (2022). A literature review and perspectives in deepfakes: generation,. *International Journal of Multimedia Information Retrieval*, 219-289.

Dixit, A., Kaur, N., & Kingra, S. (2023). Review of audio deepfake detection techniques: Issues and prospects. *Expert Systems*, 1-19.

Durmuş, Y., & Göztaş, A. (2023). Deepfake: Post-Truth Çağı İnsanlarında Fırsat ve Tehditlere Yönelik Algılar. E. Er içinde, *Medya ve Kültür Odağında İletişim* (s. 25-53). Ankara: Akademisyen Kitabevi.

Eelmaa, S. (2021). Sexualization of Children in Deepfakes and Hentai: Examining Reddit User Views. *SocArxiv*, 1-19.

Fedeli, G. (2020). Fake news' meets Tourism: a proposed research agenda. *Annals of Tourism Research*.

Fikse, T. D. (2018). *Imagining Deceptive Deepfakes An ethnographic exploration of fake videos.* Norveç: University of Oslo Master's thesis.

Firc, A., Malinka, K., & Hanacek, P. (2023). Deepfakes as a threat to a speaker and facial recognition: An overview of tools and attack vectors. *Heliyon*, 1-33.

Gamage, D., Chen, J., Ghasiya, P., & Sasahara, K. (2022). Deepfakes and Society: What lies ahead? M. Khosravy, I. Echizen, & N. Babaguchi içinde, *Frontiers in Fake Media Creation and Detection* (s. 3-43). Springer.

Guarnera, L., Giudice, O., Guarnera, F., Ortis , A., Puglisi, G., Paratore, A., & Bui , L. (2022). The Face Deepfake Detection Challenge. *Journal of Imaging*, 1-22.

Güler, N., Bayzan, Ş., & Güneş, A. (2016). İnternette Çocuklara Yönelik Riskler Ve Ailelerin Bilinçlendirme Faaliyetlerindeki Rolü. *10. Uluslararası Bilgisayar ve Öğretim Teknolojileri Eğitimi Sempozyumu* (s. 1-11). Rize: Recep Tayyip Erdoğan Üniversitesi.

Hancock, J. T., & Bailenson, J. N. (2021). The Social Impact of Deepfakes. *Cyberpsychology, Behavior, and Social Networking*, 149-152.

İdiman, E. Ç. (2021). *Sentetik Medya ve Çevrimiçi Gerçeklik.* İzmir: Dokuz Eylül Üniversitesi Güzel Sanatlar Enstitüsü Sanat ve Tasarım Anabilim Dalı Yüksek Lisans Tezi.

Jacoby, J. (1984). Perspectives on Information Overload. *Journal of Consumer Research*, 432-435.

Karnouskos, S. (2020). Artificial Intelligence in Digital Media: The Era of Deepfakes. *IEEE Transactions on Technology and Society*, 138-147.

Karnouskos, S. (2020). Artificial Intelligence in Digital Media:The Era of Deepfakes. *IEEE Transactions on Technology and Society*, 1-10.

Khoo, B., Phan, R., & Lim, C. H. (2022). Deepfake attribution: On the source identification of artificially generated images. *WIREs Data Mining and Knowledge Discovery*, 1-21.

Kırık, A. M., & Özkoçak, V. (2023). Medya Ve İletişim Bağlamında Yapay Zekâ Tarihi Ve Teknolojisi: Chatgpt Ve Deepfake İle Gelen Dijital Dönüşüm. *Karadeniz Uluslararası Bilimsel Dergi*, 73-99.

Korkmaz, Ş., & Alkan, M. (2023). Derin Öğrenme Algoritmalarını Kullanarak Deepfake Video Tespiti. *Politeknik Dergisi*, 855-862.

Kwok , A. O., & Koh, S. G. (2021). Deepfake: a social construction of technology. *Current Issues in Tourism*, 1798-1802.

Lyu, S. (2020). Deepfake Detection: Current Challanges and Next Steps. *IEEE international conference on multimedia & expo workshops (ICMEW)* (s. 1-6). London: IEEE Xplore.

Mahmud, B. U., & Sharmin, A. (2021). Deep Insights of Deepfake Technology : A Review. *arXiv preprint arXiv:2105.00192*.

Meel, P., & Vishwakarma, D. K. (2020). Fake news, rumor, information pollution in social media and web: A contemporary survey of state-of-the-arts, challenges and opportunities. *Expert Systems With Applications*, 1-26.

Nick, V. (2023). Between Realities: Information Sharing Practices of Deepfake Creators. *Proceedings of the Association for Information Science and Technology* (s. 1161-1163). London : ASIS&T Annual Meeting.

Özdemir, Ş. (2021). Yeni Nesil Tehdit: Derin Kurgu (DeepFake). *TRT Akademi*, 904-917.

Pashentsev, E., & Bazarkina, D. (2022). The Malicious Use of Artificial Intelligence Against Government and Political Institutions in the Psychological Area. D. M. Bielicki içinde, *Regulatıng Artıfıcıal Intellıgence In Industry* (s. 36-53). New York: Routledge.

Pieterse, H., & Botha, J. (2020). Fake News and Deepfakes: A Dangerous Threat for 21st Century Information Security. *15t International Conference on Cyber Warfare and SecurityAt: Old Dominion University* (s. 57-67). USA: Council for Scientific and Industrial Research.

Ruiter, A. d. (2021). The Distinct Wrong of Deepfakes. *Philosophy & Technology* , 1311-1332.

Ruiter, A. d. (2021). The Distinct Wrong of Deepfakes. *Philosophy & Technology*, 1311–1332.

Shu, K., Bhattacharjee, A., Alatawi, F., Nazer, T. H., Ding, K., Karami, M., & Liu, H. (2020). Combating disinformation in a social media age. *WIREs Data Mining Knowledge and Discovery*, 1-23.

Sigala, M. (2018). New technologies in tourism: From multi-disciplinary to anti-disciplinary advances and trajectories. *Tourism Management Perspectives*, 151-155.

Sloot, B. v., & Wagensveld, Y. (2022). Deepfakes: regulatory challenges for the synthetic society. *Computer Law & Security Review*, 1-15.

Somoray, K., & Miller, D. j. (2023). Providing detection strategies to improve human detection of deepfakes: An experimental study. *Computers in Human Behavior*, 1-8.

Swerling, G. (2020, Ocak 31). *The Telegraph News*. The Telegraph: https://www.telegraph.co.uk/news/2020/01/31/deepfake-audio-used-custody-battle-lawyer-reveals-doctored-evidence/ adresinden alındı

Temir, E. (2020). Deepfake: New Era in The Age of Disinformation & End. *Selçuk İletişim Dergisi*, 1009-1024.

Vamsi, V. V., Shet, S. S., Reddy, S. S., Rose, S. S., Shetty, S. R., Sathvika, S., . . . Shankar, S. P. (2022). Deepfake detection in digital media forensics. *KeAi Chinese Roots Global Impact*, 74-79.

Westerlund, M. (2019). The Emergence of Deepfake Technology: A Review. *Technology Innovation Management Review*, 39-52.

Woolley, S. (2020). *The Reality Game: How the Next Wave of Technology Will Break the Truth* . New York: Public Affairs.

Zeng, R., Song, S., Guo, Z., & Lian, D. H. (2023). Real or Fake: Eliciting Deepfake Identification Strategies through a Diary Study. *Proceedings of the Association for Information Science and Technology* (s. 1206-1208). London: ASIS&T Annual Meeting.