

DİJİTAL MANİPÜLASYON 'DEEFAKE' TEKNOLOJİSİ VE OLMAYANIN İNANDIRICILIĞI*

Türker ELİTAŞ

Doç. Dr., Hatay Mustafa Kemal Üniversitesi, Hatay, Türkiye

turkerelitas@gmail.com

Orcid ID: 0000 0001 8018 1208

Makale Geliş Tarihi: 12/02/2022

Makale Kabul Tarihi: 12/04/2022

Makale Türü: Araştırma Makalesi

Atıf: Elitaş, T. (2022). Dijital manipülasyon 'deepfake' teknolojisi ve olmayanın inandırıcılığı. *Hatay Mustafa Kemal Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 19(49), 113-128.

Öz

İçinde bulunduğumuz çağ, yeni iletişim teknolojilerinin fiziksel katılım sağlanan nesnel yaşama entegre olması ile birlikte hibrit yaşam pratiklerinin kendini gösterdiği çağdır. Özellikle hem sanal hem de nesnel gerçekliğin kurucu unsuru olan sosyal medyanın, nesnel gerçekliğe bir alternatif olduğu bu çağ, aynı zamanda yapay zekâ algoritmalarının nesnel gerçeğin sanal gerçeklikte yeniden inşa edildiği gelişmeyi de kapsamaktadır. Yapay zekânın yeni iletişim teknolojilerinde etkinliği özellikle içerik üretme ve dağıtma süreçlerinde kendini göstermektedir.

Deepfake teknolojisi, yapay zekâ teknolojisinin derin öğrenme algoritmalarını kullanarak canlı objelerin yeteneklerini öğrenme ve taklit etme hatta manipüle etme kapasitesi ve yetilerine sahip bir çıktıdır. İlk olarak 2018 yılında kötü niyetli çevrimiçi aktörler tarafından kullanılmaya başlayan bu teknoloji, özellikle sinema sanatına görsel ve anlamsal olarak farklılıklar katarken, dijital uzamlarda daha çok derin sahtekârlıklar için kullanılmaktadır. Dijital uzamlar dijital kimliklere fiziksel yaşam alanlarında pek fazla bulamadığı saklanabilme ve sınırsız hareket etme özgürlüğü sunmaktadır. Dijitalleşmenin kullanıcılarına sunduğu bu imkan ve fırsatlar aynı zamanda deepfake gibi teknolojilerin çevrimiçi alanlarda daha fazla kullanılmasına neden olurken, bu teknolojiler ile oluşturulan içeriklerin dijital uzamlarda dolaşıma sokulması, gerçeğin sorgulanması gibi algısal tutarsızlıklara işaret eden derin sorunları da ortaya çıkarmaktadır.

Çalışma, deepfake teknolojisinin dijital uzamlardaki etkinliği ve kullanım amaçlarına odaklanmaktadır. Özellikle sosyal medya içerik üretim sürecinde yoğun bir şekilde kullanılan bu teknolojinin dijital kimlikler üzerinde oluşturduğu etki ve sonuçlar çalışma kapsamında detaylıca tartışılmıştır.

Anahtar Kelimeler: Deepfake, Sosyal medya, Dijital Manipülasyon, Yapay zekâ, Derin öğrenme

* Bu çalışmada "Yükseköğretim Kurumları Bilimsel Araştırma ve Yayın Etiği Yönergesi" kapsamında uyulması belirtilen tüm kurallara uyulmuştur. Yönergenin ikinci bölümü olan "Bilimsel Araştırma ve Yayın Etiğine Aykırı Eylemler" başlığı altında belirtilen eylemlerden hiçbiri gerçekleştirilmemiştir.

DIGITAL MANIPULATION 'DEEFAKE' TECHNOLOGY AND THE COGENCY OF THE NONEXISTENT

Abstract

The era we live in is in which hybrid life practices manifest themselves with the integration of new communication technologies into objective life with physical participation. This era, in which social media, which is the founding element of both virtual and objective reality, is an alternative to objective reality, also includes the development of artificial intelligence algorithms in which objective reality is reconstructed in virtual reality. The effectiveness of artificial intelligence in new communication technologies is especially evident in content production and distribution processes.

Deepfake technology is an output that has the capacity and abilities to learn, imitate and even manipulate the abilities of living objects by using deep learning algorithms of artificial intelligence technology. This technology, which was first used by malicious online actors in 2018, while adding visual and semantic differences to the art of cinema, is mostly used for deep frauds in digital spaces. Digital spaces offer digital identities the freedom to hide and move freely, which they cannot find much in physical living spaces. While these facility and opportunities offered by digitalization to its users also cause technologies such as deepfake to be used more in online areas, it also reveals deep problems that point to perceptual inconsistencies such as circulating the content created with these technologies in digital spaces and questioning the truth.

The study focuses on the effectiveness and usage of deepfake technology in digital spaces. The effects and results of this technology, which is used extensively especially in the social media content production process, on digital identities are discussed in detail within the scope of the study.

Keywords: *Deepfake, Social media, Digital manipulation , Artificial intelligence, Deep learning*

Giriş

Dijitalleşme sürecinde ortaya çıkan sanal ortamlar kullanıcıları için melez bir yaşamın oluşmasında etkin ve kuşatıcı bir rol üstlenmektedir. Bu ortamlarda dijital kimlikleri ile alternatif bir yaşam pratiği kurgulayan bireyler her geçen gün teknolojinin iletişim olgusuna entegre olarak neler yapabileceğini bizzat deneyimlemektedir. Özellikle yazılımların ve uygulamaların ekseninde kendini ifade etme yöntemini benimseyen ve etkileşim sürecine dâhil olan kullanıcılar bu alternatif yaşam pratiği içinde nesnel gerçeğinden edindiği normlarından uzaklaşma ve sanallığın vermiş olduğu sonsuz gizem ile yaşamaya başlamaktadır.

Yapay zekâ teknolojisinin özellikle çevrimiçi kültürde kendini göstererek iletişim sürecini bambaşka bir evreye taşıması nesnel kimliğinden uzaklaşarak dijital kimliğin melez yaşama yön vermesine neden olmaktadır. Bu süreçte özellikle yapay zekânın bir çıktısı olan deepfake teknolojisinin görmeye ve işitmeyen dayalı etkin manipülasyonu nesnel gerçeklikte olma ihtimalinin düşük olduğu eylem ve olaylar silsilesini kurgulamakta ve melez yaşam formunda sanala olanın güvenilirliğine ket vurmaktadır.

Bugün dijital platformları birçok sebepten ötürü kullanan insanların önünden nesnel gerçeklikte olmayacak kadar veri ve içerik geçmektedir. Bu içeriklerle günlük eğlence, sosyalleşme ya da haber alma ihtiyacını gidermeye çalışan kullanıcılar, görerek ya da işiterek öğrenme güdülerini dâhil oldukları çevrimiçi kültürde de devreye sokmaktadırlar. Deepfake

teknolojisi her ne kadar sinema da önemli bir teknoloji olsa da dijital mecralarda ki bu önemi daha çok derin sahtekârlıkların başat aktörü olduğu için silikleşmektedir. Özellikle kötü niyetli internet aktörleri tarafından bu teknolojinin dijital uzamlarda kullanılması sonucunda ortaya çıkan gerçek olmayan ama sahte olmayacak kadarda inandırıcı olan içerikler görerek ve işiterek öğrenen ya da haber alan kullanıcıların gerçeklik algılarını yönlendirmektedir. Gerçeklik algısının deepfake teknolojilerinin çıktısı olan içeriklerle sürekli manipüle edildiği dijital uzamlarda, gerçeğin arayışında uzaklaşan dijital kimlikler, içerikleri ait olduğu sosyal sınıf, ideolojik duruş ya da fiziksel katılım sağladığı yaşam normları ekseninde yorumlamaktadır.

Birey-Araç İlişkisinde Manipülasyonun Kurguladığı İlişki Pratikleri

Köken olarak batı dillerinden doğan ve Türkçe 'de hemen her disiplinde kullanılan manipülasyon kavramı, kurmaca bir gerçeğe ya da gerçeğin yeniden inşa edilerek dolaşıma sokulmasına atıf yapmaktadır. Kavramın kelime anlamı olarak hile ve aldatmak şeklinde olumsuz bir özelliğe sahiptir. Özellikle insanlık kadar eski olan iletişim olgusu içinde bu kavram sorunlu bir alana işaret etmektedir.

Manipülasyon iletişim sürecinde alıcı ile verici arasında mesaj yapısına ve içeriğine bilinç bir müdahaledir. Bu müdahale özellikle kaynak tarafından ve kimi zamanda araç yardımıyla gerçekleşmekte ve alıcıda istenilen düşünce ve fikir üretme süreçlerine etki etme şeklinde gerçekleşmektedir. Bu bağlamda manipülasyonun temel amacı fikir ekmek, fikri değiştirmek veya fikri yönlendirmek şeklinde değişik amaçları gütmektedir. Bu noktada manipülasyon rıza üretimi için bilinç olarak kurgulanmış bir tür iknadır. Bu kurguda kaynak ikna eden (manipülatör) sosyo-ekonomik olgulara göre hedef üzerinde zihinsel baskı kuran ve bilinçli olarak yönlendirmeye çalışan kurucu unsur, alıcı ise ikna edilen istediği dışında davranmak zorunda bırakılan (Chapaux- Morelli ve Couderc, 2017, s. 8-9) ya da gönüllü olarak etkiyi benimseyen şeklinde konumlanmaktadır.

Tek taraflı bir iletişim etkinliği olan manipülasyon, özellikle özgürlükçü bir yönetim anlayışı benimsemiş toplumlarda oldukça etkin bir argümandır. Totaliter bir anlayışın hâkim olduğu ve doğrudan baskı gören bir toplumda manipülasyona zaten gerek görülmez. Bu bağlamda herhangi erkin özellikle siyasal ve elit kesimin doğrudan baskı ve zorlamaya dayalı bir yönetim anlayışını benimsemediği uyanmış ya da yarı uyanmış toplumsal organizasyon içinde varlığını idame ettiren birey zihinsel güdülemeye her an hazırdır. Bu durum ise manipülasyon için en elverişli ortamdır (Schillerin,1993,s. 10).

Uyanmış ya da yarı uyanmış toplumsal organizasyon içinde medya siyasal erkin ya da elitist kesimin doğrudan kuramadığı baskıyı, hedef kitlenin rızasını alarak kurmaktadır. Medyanın görevi, toplumu eğlendirmek, bilgilendirmek ve halka, toplumun kurumsal yapısıyla bütünleştiren inançları, davranış kodlarını dayatma amacı güder ve son kertede topluma imgeler ve mesajlar ulaştırır (Chomsky, 1993, s. 9). Bu imge ve mesajlar ise hedef kitlenin fikirlerine dolaylı bir şekilde müdahale etme gayesi içindedir. Bu bağlamda özellikle uyanmış ya da yarı uyanmış toplumlarda en güçlü manipülatör medyadır.

Manipülasyondaki en önemli kural gerçek amacın saklı tutulmasıdır. Bu şiardan hareketle gerçek amacı örüntülenmiş bu ikna argümanında yalan söyleme, duygu sömürüsü yapma, kurban rolünü oynama, aşağılama ve provakasyon, duygusal şiddet, mahrum bırakma, suçluluk duygusu yaratma ve egoyu okşama gibi stratejiler en sık kullanılan enstrümanlardır (Björnsson, 2016, s. 637-653). Bu argümanların hepsi aynı ayna kullanılabilirdiği gibi hedefteki birey veya kitlenin taşıdığı karaktere uygun olarak da değişkenlik gösterebilirler.

Manipülasyonun en derinlerinde başkasının davranışlarını kontrol etme ve düşüncesini şekillendirme arzusu yatmaktadır. Bu bağlamda bireylerin sürekli katılım sağladığı iş yaşamından

sosyal çevresine kadar, kendisi gibi düşünmesini isteyen ya da kontrol edilebilir bir fikir aşlamaya çalışan kişi, kurum ya da erk ile karşılaşmaktadır. Özellikle iletişim sürecinde kendini pasif bir şekilde konumlandıran birey, karşılaştığı bu gizli baskı ile mücadele etmek zorunda kalmaktadır.

Kitle iletişim araçları manipülasyon stratejilerinin en yaygın kullanıldığı araçlardır. Manipülasyonun birey odağından çıkıp kitlesel bir yönlendirme işlevi görmesinde şüphesiz ki kitle iletişim araçlarının önemli bir etkisi bulunmaktadır. Bu noktada özellikle hem görsel hem de işitsel bir kitle iletişim aracı olan televizyon karşısında pasif bir şekilde tek yönlü bir iletişim sarmalına giren birey gerçeği aradığı haber programlarından zamansal olarak en kısa süreli olan reklamlara kadar sürekli manipülatif içeriklere maruz kalmaktadır. Bu bağlamda kitle iletişim araçları bir taraftan dolayımlanmış gerçeği kurgu süzgecinden geçirerek yeniden inşa ettiği gerçeğin temsili haberlerle ve 'sürekli alın mutlu olun' ya da 'bu tam da sizin ihtiyacınıza göre' şeklindeki ekme mesajlarla kurban olarak nitelendirilen hedef kitlede arzu edilen duygu, düşünce veya davranış kalıpları oluşturmaktadır. *"Kitle iletişim araçlarının istenen mesajlar, tarafsız bir kaynaktan geliyormuş gibi gözükür oysaki sistematik bir şekilde devletin veya çıkar guruplarının propagandası yapılır. Kitle iletişim araçları bunu yaparken kimi zaman haberleri, reklamları, dizileri kimi zamanda filmleri kullanırlar"* (Özer, 2019, s. 216).

İletişim teknolojilerinin sürekli gelişim evresinde olması son kertede bireyin tüm yaşamsal parametrelerine müdahale etmesine ve bireyin bu araçları önemli bir bilgi, haber, eğlence vd. kaynağı olarak görmesine neden olmuştur. Araç ile kurulan bu denli bir bağ, araçtan gelen her türlü mesaja inanma eğilimini de peşi sıra getirdiği gibi enformasyonun sorgulanabilirlik durumunu da ortadan kaldırmaya neden olmaktadır. Netice itibarıyla kitle iletişim araçları çarpıtılarak, eksiltilecek, çoğaltılarak, ya da abartılarak ve çeşitli mitler kullanılarak ve bu mitler çeşitli dayanaklarla desteklenerek topluma empoze edilmesinde önemli bir görev üstlenmektedir (Özer, 2019, s.215). Bu bağlamda kitle iletişim araçları imgeler ve mesajların en önemli taşıyıcısıdır. Bu noktada özellikle yeni iletişim teknolojileri ile birlikte sürekli çevrimiçi bir yaşam süren bireyin artık hayata dair tüm imge ve kodlarının oluşturucusu araç ile kurduğu bağ ile sağlanmaktadır.

Manipülasyonda arzu edilen duygu, düşünce veya davranışların istenilen şekilde oluşturulabilmesi için her yol mubahtır anlayışı kavramın aslında tehlikesine ve oluşturacağı bireysel ve toplumsal hasarına odaklanmayı elzem kılmaktadır. Manipülasyonda hedef genellikle kitleler olduğu için oluşturacağı hasarda kitlesel olacaktır. Bu bağlamda yeni iletişim teknolojileri sayesinde her an çevrimiçi kalan birey sürekli etkileşimli bir ortam içinde dolaşıma sokulan manipülasyon amaçlı görsel ve işitsel mesajlara maruz kalmakta ve enformasyon yığını içinden şuurlu bir seçim yapma yetisini kullanamayıp, manipülatif mesajların art penceresini görememektedir.

Manipülasyon 'olmayanın olmuş gibi düşündürülmesinin' yanı sıra 'olanında kusurlu gösterilmesi' şeklinde zihinsel etkiye yönelik bir hileli yönlendirmedir. Bu hileli yönlendirme sürecinde gerek konvansiyonel iletişim teknolojileri gerekse de yeni iletişim teknolojileri hedef ile kurduğu bağ ile önemli misyon üstlenmektedir.

Kitle iletişim araçlarında bireyin araç ile kurduğu bağ aracın konvansiyonel veya yeni olmasına göre değişkenlik göstermektedir. Konvansiyonel iletişim araçlarında öncelikle pasif bir konumda olan birey manipülasyona sürekli açık halde beklemekte ve Chomski'nin elit kesimlerin aracı olarak nitelediği bu araçlardan gelen mesajlara karşı gönüllü katılım sağlamak üzeredir. Ancak yeni iletişim teknolojileri ile birlikte ortaya çıkan etkileşim unsuru bireyin araç ile kurduğu bağı dinamik bir yapıya büründürmekte ve aynı zamanda kullanıcıya manipülasyon yapabilme kabiliyetini de kazandırmaktadır. Her kullanıcı artık yeni iletişim teknolojileri ile hem

manipülasyona açık alıcı hem de manipülasyonu kurgulayan bir potansiyel bir manipülatör adaydır.

Etkileşim unsuru yeni iletişim teknolojilerin en önemli tanımlayıcı özelliklerinden biridir. Yeni iletişim teknolojileri ile kültür arasındaki ilişkiyi irdeleyen Manovich'e (1996, s. 2) göre etkileşim batı için demokrasi ve eşitlik fikirlerinin savunuculuğu için mükemmel bir araç iken, doğu toplumlarında totaliter iradenin halka empoze edilmesi için kullanılan manipülasyonun güç kaynağıdır.

Gerçeğin Klonlanması ve Gerçeğin Tehdidi 'Deepfake'

Gelişen teknoloji artık içselleştirilmesi kolay olmayan bazı çıktıları ile nesnel gerçekliğe dair işleyişe sürekli müdahale etmekte ve algılanan formlarda sürekli güncelleme yapılmasını elzem bir hale getirmektedir. Özellikle nesnel gerçeklik ile örüntülenmiş yaşamsal aktivitelerde sanal gerçekliğin müdahalesinin oldukça derinden hissedildiği bu çağ da teknolojinin çıktılarının neyin gerçek neyin sahte olduğunu dair algıları zorlaştırmaktadır. Bu bağlamda yapay zekâ ile bilgisayarın buluşması gerçeğe yönelik arayışta ortaya çıkan büyük veri kümeleri arasında bireyi algıladığına yöneltmektedir. Bu durum teknolojinin iletişime entegre olması ile ortaya çıkan yeni iletişim teknolojileri ile nesnel gerçekliğini sanal gerçeklik ile hibritleştiren birey için sorunlu bir alan olarak kendini göstermektedir. Gerçeğin 'görmek ve izlemek' eyleminden beslendiği nesnel yaşam kültürünü, giderek nesnel ve sanalın karışımı olarak ortaya çıkan melez bir yaşam alanında da devam ettirmeye çalışan bireyler nesnel gerçekliğini sanal ortamlarda sürekli gördüğü içerikler ile kurgulamaktadır. Bu noktada özellikle sanal uzamların erişim ve hız cazibesi, görmek ve izlemek üzerine kurulu gerçeklik arayışında ikna edici derin sahtekârlıkların ortaya çıkmasına mahal vermekte, bu derin sahtekârlıklar dijitalleşmenin kapsayıcı özelliğinden dolayı milyonlarca insana aynı anda ve hızla ulaşabilmekte ve takipçisini bilişsel olarak zoraki rıza göstermeye mecbur etmektedir.

Sanal uzamların en cazip noktalarından biride hiç şüphesiz ki anlık veri akışıdır. Bu anlık veri akışı içerisinde özellikle ses ve görüntü üzerinden içeriklerini hazırlayan ve kullanıcılarından da bu bağlamda içerik oluşturmasını bekleyen bu uzamlar talepkârlığını artırmak için sürekli kendini güncellemekte ve kullanıcısının da güncel kalması için dijital vaatler ve telkinler sunmaktadır. Bu güncellemelerin temel paradigmalarından biri de sese ve görüntüye müdahale etme noktasındadır. Kullanıcı, dijital ortamlarda dolaşıma soktuğu nesnel gerçekliğinin bir temsilcisi olan ses ve görüntü argümanlarını sanal mecralarda ilgi çekebilmek, beğeni alabilmek ya da takipçi sayısını artırabilmek gibi bireysel nedenlerden ötürü gerçeği yeniden kurgularken aynı zamanda pasif manipülatörlük sınırları içinde de gezindiğinin farkında olamamaktadır.

Dijital uzamlar gerçeğin birçok şekilde dolaşıma sokulduğu ve en fazla manipüle edildiği alanlardır. Özellikle sosyal medya ve içerik paylaşım ağlarında gerçeğin delili olan gerek görüntü ve gerekse ses unsurlarının yeniden inşa edilmiş halleri giderek artmaktadır. Çevrimiçi yaşamda yeniden anlam bulan orijinallik sürekli sahtenin tehdidi altındadır. Bu noktada özellikle yapay zekânın algoritmalarından yararlanan yazılımlar bu tehdidin oluşturucu güçlerinden biridir. Dijitalliğin şafağında her kullanıcı potansiyel bir manipülatör olabilmektedir. Bu noktada gerçekle sahtenin ayrımının herkesçe anlaşılabilirliği basit uygulamalar ile her kullanıcı bu potansiyelini kullanabildiği gibi, profesyonel teknoloji ve uygulama bilgisi ile de bu potansiyel kullanılabilir. Bu bağlamda deepfake videolarda yapay zekâ (AI) ile gerçeği klonlayan teknolojinin çıktılarıdır.

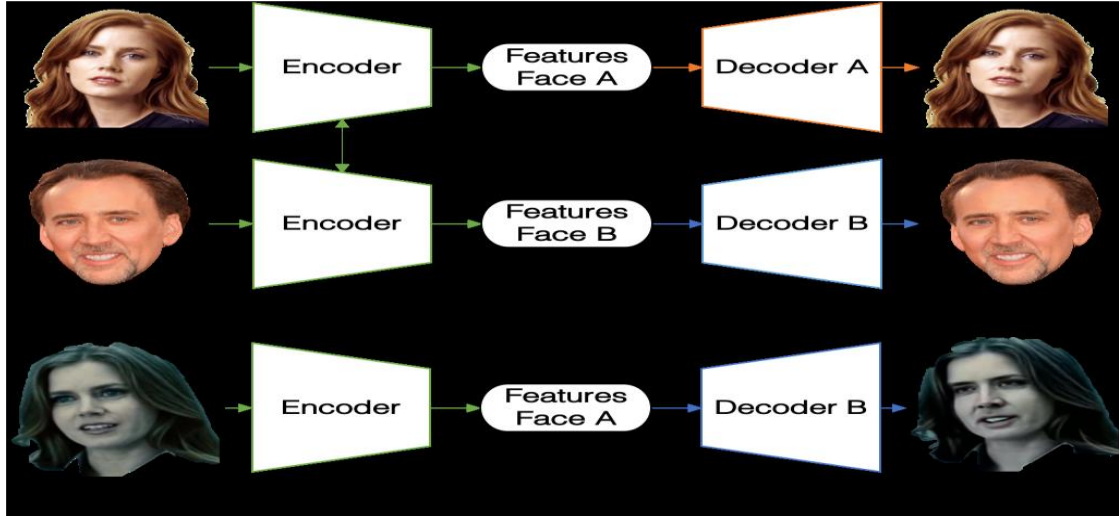
Deep (derin) ve fake (sahte) kelimelerinin birleşimi deepfake literatürde derin öğrenme şeklinde karşılık bulmaktadır. Özellikle son dönemlerde dijital mecralarda siyasetten ünlü film yıldızlarına kadar birçok kesimin şaşırtıcı görüntü ve ses kayıtlarının şoke edici etkisi sonucunda

kamuoyunca bilinirliği artan deepfakeler genel olarak, insanların yüzlerinin ve seslerinin başka yüzler ve sesler tarafından izinsiz bir şekilde takas usulünce kullanılıp dijital olarak manipüle edilen hipergerçekçi videolardır (Westerlund,2019, s. 40).

Deepfake sinema gibi görsel sanatlarda işlerin kolaylaştırılması için yapay zekâ algoritmalarının kullanıldığı bir yazılım türüdür. Hollywood sinemasında çok sık olmasa da kullanılmaya başlanılan bu yapay zekâyâ dayalı makine öğrenimi çıktıları ile hayatını kaybeden oyuncuların yeniden sinema salonlarında izleyici ile buluşması sağlanmıştır. Oluşturduğu gerçekçi etkisi sayesinde dijital uzamlarda da kullanılmaya başlanılan bu teknoloji ile görüntü ve ses taklidinin yanı sıra karakterin istenildiği gibi hareket ettirilmesi ya da konuşturulması gibi orijinallikten uzak ama gerçeğin kendisi gibi manipüle içerikler çevrimiçi dolaşımında sıklıkla karşılaşılan paylaşımlar olmaya başlamıştır.

Deepfakeler, bir kişinin yüz ifadelerini, tavırlarını, sesini ve tonlarını taklit etmeyi öğrenmek için büyük veri örneklerini analiz eden sinir ağlarına güvenir. Süreç, iki kişinin görüntülerini veya yüzleri değişik tokuş etmek için derin bir öğrenme algoritması beslemeyi içerir (Westerlund,2019, s. 40). *“Temel fikir, derin öğrenme algoritmalarının ana bileşeni olan bir dizi yapay sinir ağlarını, aktör ve hedef yüzlerin çoklu örnekleri üzerinde eğitmektir. Yeterli eğitimle, sinir ağları her yüzün özelliklerinin sayısal temsillerini oluşturabilecektir. O zaman tek yapmanız gereken, aktörün yüzünü hedefle eşleştirmek için sinir ağlarını yeniden düzenlemek”* (Dickson, 2022).

Resim 1. Deepfake yüz kod çözücü işlemi



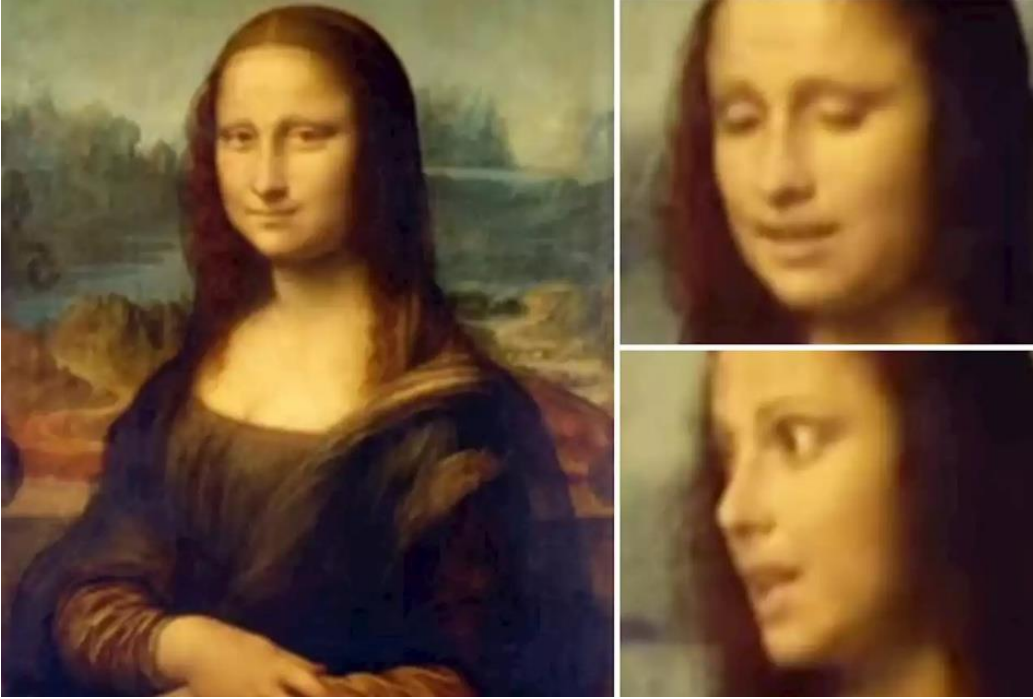
Kaynak: Güera ve Delp (2018)

Deepfake kavramı çok basittir. Ancak onu eğitmek büyük çaba gerektirir. Şayet bu durumu siyasi figürler örneği ile açıklamak gerekirse, Amerika Birleşik Devletleri Başkanı yerine Rusya Başkanının televizyondan Amerikalılara seslendiği derin sahne versiyonunu oluşturmak pekâlâ mümkündür. Öncelikle kaynak olarak kullanılacak Rusya Başkanının ve hedef olarak kullanılacak Amerika Başkanının otomatik kodlayıcıları için eğitim veri setini birleştirmek gerekmektedir. Bu durum deepfake için en zor ve en uzun süreçler biridir. Bu süreçte hem Rusya hem de Amerika Başkanının binlerce dakika video karesini toplayarak yalnızca yüzünü gösterecek şekilde kırpma işlemi uygulanmaktadır. Kırpma işlemi sonrası yapay zekânın algoritmaları devreye girerek yüz haritalama işlemine başlanmaktadır.

İdeal yüz ve ses takası için yüzlerin ve ortamların farklılıklarının kodlanması ve yazılımın aktarmayı öğrenebilmesi için farklı açılardan birçok veriye ihtiyaç duymaktadır. Bu bağlamda deepfake için kullanılması arzu edilen yüz hakkında derinlemesine görüntü hafızası oluşturulmalıdır. Kaynak görüntüyü hedef görüntüye senkronize etmek ya da hedefi kaynağın istediği şekilde yönlendirmek için hedefin her hareketi, jestleri ve mimikleri derin öğrenme algoritması tarafından hafıza alınmaktadır. Bu nedenle gerçeğe yakın sahte videoların hipergerçekçi olabilmesi için kaynağın elinde hedefin farklı ortamlardan alınmış mümkün olduğunca fazla görüntü ve ses içerikleri olmalıdır. Aksi takdirde gerçeklik algısını oluşturma başarısı algoritmadaki veri azaldıkça azalacaktır.

Deepfake video üretme sürecinde hedef her zaman bir hareketli görüntü olmayabilir. Derin sinir ağlarını donuk kare üzerinde kullanmak mümkündür. Örneğin Mona Lisa'yı konuşturmak, ağlatmak veya Mona Lisa'nın yüzünü başka bir donuk kare ile değiştirmek de olası kullanım pratikleri içerisinde yer almaktadır.

Resim 2 : Deepfake ile donuk kare manipülasyonu



Kaynak: Hodge (2022)

Deepfake gerçeği manipüle etme noktasında sadece görüntü üzerinde veri işlememekte aynı zamanda gerçeğin işitsel yansıtıcısı olan ses unsuru üzerinde de derin sinir ağları kullanılabilir. Başka birinin sesini öğrenme noktasında ses çözücü kod işlemi de tıpkı görüntü öğrenme sürecinde olduğu gibi uzun ve detaylı süreçtir. Bu süreçte sesi taklit edilecek hedefin ağız hareketleri, vurgulamaları ve tonlamaları gibi belirleyici özellikleri için oldukça fazla veriye ihtiyaç duyulmaktadır.

Yapay zekâ teknolojisi ile (AI) yüz haritalama işlevini kullanan deepfakeler görüntünün performansını artırarak görsel sanatlara farklı bir ambiyans katarken, görüntü ve sese yönelik komplike bir manipülasyon aracı olması bakımından da tehlike arz etmektedir. Yapay zekâ algoritmalarının ilk kez bu denli etkin ve görünürlük kazandığı deepfake videolar dijital platformlarda ilk olarak 2018 yılında içerik paylaşım sitelerinde kendini göstermiştir. Özellikle Hollywood aktörlerinin ve siyasetçilerin yetişkin sitelerde kullanılması tüm dikkatleri bu yeni

teknolojiye yöneltmiştir.

Deepfake İçeriklerin Sürdürülebilir Mekânları: Sanal Ağlar

Geleneksel iletişim araçlarında olmayan, görüntüyü yeniden oluşturma ve gerçek ile sahtenin ayırım çizgisinin ortadan kaldırılarak orijinalin yani gerçeğin varlığına yönelik tehditlerin ortaya çıkmasına neden olacak dijital manipülasyonlar, özellikle yapay zekânın sürece dâhil olmasıyla birlikte çevrimiçi kültürün en yeni trendleri arasında olmaya başlamıştır. Görüntü ve videoların üretimi ve manipülasyonu aslında çok yeni bir gelişme değildir, özellikle görüntüye dayalı kitle iletişim araçlarının varlığından beri görüntü üzerindeki eklemeler, kırpımlar veya kesmeler gibi oynamalar dijital platformlarda daha profesyonel bir şekilde yapılmaya başlandığı için yeni bir gelişme gibi algılanmaktadır.

Nesnel gerçekliğin yapay zekâ ile manipüle edildiği deepfake videolar, gerçeğe dair yeni algılar oluşturmaktadır. Bu gerçeklik bir taraftan aslına uygun olarak klonlanmış bir gerçek iken diğer taraftan gerçeğin simülasyonu olarak iki farklı şekilde algısal oyun oynamaktadır. Ancak her iki gerçekte de sorunlu olan nokta gerçeğin artık hangisi olduğudur. Deepfake'ler mevcut argümanlar yardımıyla gerçeği yeniden inşa etmek üzere kullanılan bir teknolojidir. Yeniden inşa edilen bu gerçeğin dolaşıma sokulduğu mecralar ise multimedya özelliklerinden dolayı sürekli veri akışının gerçekleştiği sanal ağlardır.

Sanal ağlarda kullanıcısı ile buluşan deepfake teknolojisinin son yıllarda kullanım alanı ve oluşturulan içerik sayısı bakımından bir ivme kazandığı aşikârdır. Dijital görüntü üretimi ve manipülasyonu DNN'lerin (derin sinir ağları) hızlı gelişimi daha kolay ve inandırıcı hipergerçek videoların ortaya çıkmasına zemin oluşturmuştur. *"DNN tarafından oluşturulan sahte videolar, ilk olarak, Deepfakes adlı bir Reddit hesabının DNN tabanlı bir yüz değiştirme algoritması kullanılarak oluşturulan sentetik pornografik videoları yayınlamaya başladığı 2017 yılında kamuoyunu dikkatini çekmiştir"*(Lyu,2020, s. 1). Yapay zekâ algoritmalarının sanal uzamlara entegre olması ile gerçeğin sahte ile silikleştiği bu çağ deepfake videoların çağıdır. Sanal ağlarda önemli bir izler kitleye ulaşan bu sentetik videolar özellikle sanatçı ve siyasetçi gibi ekran görünürlüğü fazla olan kitleyi hedef almakta ve söylenmemiş sözlerden gitmediği mekânlardan onları sorumlu tutmaktadır. *"2017 yılında Fakeapp uygulamalarıyla, aralarında Michelle Obama, Ivanka Trump ve Emma Watson'ın da olduğu bazı ünlülerin yüzleri, yetişkin filmlerinde kullanılınca, teknolojinin karanlık yüzü ortaya çıktı"* (Birkan ve Ulukan, 2019).

Deepfake videoların ortaya çıkmasında kuşkusuz DNN'lerin önemli bir payı bulunmaktadır. Şu ana kadar üç tür deepfake video türü bulunmaktadır.

- Kafa kuklası, hedefin başınının ve üst omuzunun kaynak kişinin kafası ve üst omuzu ile yer değiştirmesi ile gerçekleşen videolardır. Böylece yapılan sentez sonucunda hedef kaynağın davranışlarını sergiler.
- Yüz değiştirme, aynı yüz ifadeleri korunurken kaynağın sentezlenmiş yüzleriyle değiştirilen yüzlerle hedefin bir videosunu oluşturmayı içerir.
- Dudak senkronizasyonu, hedefin gerçekte konuşmadığı bir şeyi konuşuyormuş gibi görünmesi için yalnızca dudak bölgesini manipüle ederek tahrif edilmiş bir video oluşturmaktır (Lyu,2020, s. 1).

Çevrimiçi ortamlarda oldukça fazla dolaşıma sokulan deepfake'ler, DNN'lerin yardımıyla inandırıcı bir hal almaktadır. Sahtenin gerçekten ayrılması için profesyonel yazılımların ve bilirkişilerin kullanılması gereken bu hipergerçekçi vidoların oluşturucuları ise deepfake hobi toplulukları, hükümet temsilcileri, aktivistler, sivil toplum örgütleri, medya yapımcıları gibi meşru aktörler ve son olarak da kötü niyetli aktörlerdir. Kullanıcı skalasının oldukça geniş olduğu

deepfake içerik üreticileri özellikle hobi toplulukları ve dolandırıcı gibi kötü niyetli aktörlerin çevrimiçindeki içerikleri sosyal ağlardaki tüm kullanıcılar için bir tehdit oluşturmaktadır.

Deepfake videoların hobi toplulukları tarafından üretilip dolaşıma sokulduğu mecralar özellikle sosyal medya ortamları ve içerik sağlayıcı platformlardır. Yarı profesyonel bir üretim sürecinden geçen bu deepfakeler de temel amaç eğlenmek ve eleştirmek olduğu kadar siyasilerin veya sinema oyuncularının yüzlerinin pornografik içeriklere yerleştirilmesi gibi yasal olmayan içeriklerde hobi topluluklarının dijital eylemleri arasında sıklıkla görülmektedir. Sosyal medya üzerinden buluşan deepfake hobi toplulukları ve takipçileri dijitalliğin önlenemez özgürlüğü ve esnekliği ile inşa edilen yeni gerçeği yine sanal ağlar üzerinden yorumlamakta ve dolaşım kanallarını artırmaktadırlar. Bu noktada özellikle sosyal medya üzerinden örgütlenen deepfake kullanıcıları ve takipçileri üretilen videoların gerçek olmadığı bilmekte ancak henüz dijital ortamlara yönelik ciddi bir yaptırımın olmamasının vermiş olduğu öz güvenle sürekli sınırları zorlamakta ve başkalarının yapmadığı eylemlerden sorumlu tutulmalarına sebep olmaktadır.

Deepfake teknolojisi ile içerik oluşturan ve bunu dijital platformlarda dolaşıma sokan bir diğer kötü niyetli aktörler dolandırıcı ve suç şebekeleridir. Nesnel varoluşsallıklarını dijital platforma taşıyan bu aktörler oluşturdukları deepfakeler ile kripto para hırsızlığından, kimlik hırsızlığına kadar çok yönlü bir suç eğilimi sergilemektedirler. Sahtenin kanıtlanamayışı ve orijinal ile sahtenin ayrımının silikleşmesi gibi gerçeğin değerinin sıfırlandığı ve sahtenin kutsandığı çevrimiçi yaşam kotu niyetli aktörlerin nesnel gerçeklikteki varlıklarını dijitalle kaydirmalarına sebep olmuştur. Sanal ağlarda teknoloji odaklı yasadışı faaliyetler ile tüm kullanıcılar için potansiyel tehdit olan bu aktörlerin, anlık konum değiştirme yazılımları ile takip edilmesi ve yakalanması oldukça güçtür. Bu durumun vermiş olduğu cesaret ile özellikle sosyal medyada bu aktörler ve oluşturmuş oldukları deepfake videolar günden güne artmaktadır.

Gerek deepfake hobi grupları ve art niyetli kullanıcılar gibi sanal ağlarda yasadışı kullanım için bulunanlar olsun ve gerekse nesnel gerçekliği tamamıyla yasal düzenlemeler üzerine kurulu olan hükümet temsilcileri, sivil toplum örgütleri veya medya kuruluşları olsun sosyal medya ve içerik sağlayıcı platformlar geniş kullanıcı kitlesine hitap etmesi bakımından çevrimiçi yaşamdaki en fazla tercih edilen uzamlardır. Ayrıca deepfake teknolojisinin özellikle sosyal medya da bu denli yaygın olmasının en önemli sebeplerinden biri geniş bir uygulama yelpazesi sunmasının yanı sıra hem profesyonel hem de acemi uygulayıcılar için uygun öneriler sunmasıdır (Mahmud, Sharmin,2020, s. 15).

Özellikle deepfake ile yapılabilecek kışkırtma, karalama, hiciv, propaganda, toplumsal hareketlilik, dolandırıcılık ve hırsızlık gibi tüm yönlendirici, tahrip edici ve zarar verici eylemlerin en fazla karşılık bulduğu uzamlar sosyal medya ve içerik sağlayıcı platformlardır. Tüm bu kullanım amaçların yanı sıra sosyal medya resmi ya da gayri resmi tüm kişi ve grupların nesnel gerçekliklerini korumak ve kamuoyu oluşturmak için de önemli bir platformdur. Bu noktada özellikle kamuoyu oluşturmak için sıklıkla başvurular yöntemlerden biri de deepfake videolardır. Günümüzün dezenformasyon savaşlarında en ciddi yöntemlerinden biri olan deepfake videolar, komplo üretmek, sahte bilgiyi yetkili birinin ağzından çıkmışçasına dolaşıma sokmak gibi birçok etik olmayan hatta suç sayılabilecek eylemlerin görsel ve işitsel şölen ile hazırlanmasında başlıca şüphelilerden biridir. Yapay zekâ tarafından oluşturulan istenmeyen e-postalar ve bağınaz metinler, sahte videolar ve çok sayıda komplo teorisi üzerine kurulu sahte haberlerin giderek artması gerçeğe ile sahtenin yer değiştirmesine neden olduğu gibi sahte bilgi ve videolar ile sürekli temas eden insanların gerçeğe karşı ilgisizlik sendromunu tetiklemesi bakımdan tehlike arz etmektedir.

Gerçeğe Karşı İlgisizlik Sendromu ve Trolleme

Yeni iletişim teknolojilerinin nesnel gerçekliğe kanalize olmaya başlaması ile birlikte dijital kodlar ile örüntülenmiş bambaşka bir kültür ortaya çıkmıştır. Çevrimiçi kültür olarak nitelendiren bu kültürde özellikle iletişim unsuru kendi kalıplarının dışına çıkarak yepyeni bir forma bürünmüştür. Zamansızlık ve mekânsızlık ile kutsanan bu yeni iletişim formunda dijital kimlikleri ile iletişim süreçlerine katılan kullanıcılar fiziksel yaşama dair tüm aktivitelerinde alternatif bir etkileşim içine girmeye başlamışlardır.

Çevrimiçi kültür dijitalleşmenin şafağında sosyal medya ve türevleri içinde olgunlaşan sınırsız, kontrolsüz ve yüksek etkileşim sarmalı ile örüntülenmiş bir oluşumdur. Bu oluşum içinde varlık gösteren her katılımcı dilediği gibi ya da nesnel gerçeklikte hayal ettiği gibi her türlü organizasyonel oluşumlara katılım sağlayabilmekte ve sanal uzamların vaat ettiği her türlü aktivite ve söylemlere iştirak edebilmektedir. Bu katılım sürecinde kullanıcılar bazen nesnel gerçekliğine dair kendi normlarına uygun platform ve grupları tercih edebilirken bazen de sanallığın vermiş olduğu gizliliğin normlarına uygun platform ve gruplara iştirak etmektedir. Dijital kimlik oluşturma süreciyle başlayan gizlilik unsurunu nesnel kimliğine entegre edip gizlilik kalkanını ortadan kaldıran kullanıcıların yanı sıra, gizlilik olgusunu tamamen nesnel gerçekliğinden uzak yepyeni bir yaşam döngüsü üzerine kuran ve bu yaşam döngüsü için de sınırsızca ve fütursuzca hareket eden kullanıcılar da dijital mecranın önemli konakları arasındadır. Özellikle nesnel gerçekliğinden uzak bir yaşam süren bu kullanıcılar dijital mahallelerin görünmeyen yüzleridir ve bu görünmemenin vermiş olduğu cesaretle mahalle sakinlerine olası bir tehdit olmaktadır. Dijital mahallelerin en önemli sorunlarından biri tanımadığı, görmediği biri tarafından sürekli rahatsız edilmek, aşağılanmak ve dijital tacize uğramaktır.

Rahatsız edici söylem ve eylemler literatürde trollemek ya da trollük olarak kabul edilmektedir. Çok geniş bir ifade olan rahatsız edici söylem ve eylemler alanyazın incelendiğinde kışkırtıcı davranıştan istismara kadar pek çok olumsuz davranışın genel olarak ifade edilmesi şeklinde bir karşılık bulmaktadır. Bu bağlamda genel olarak gizlilik çadırında kendini saklayan çevrimiçi konaklarının daha çok başvurduğu bu rahatsız edici söylem ve eylemler özellikle yeni iletişim teknolojilerini içselleştiren kültürlerde yaygın bir kullanımı bulunmaktadır.

Trolleme kavramının dijital ortamlarda ilk kullanımı e-posta uygulamaları ve 'Usenet' olarak bilinen en eski heterojen grup iletişim sistemleri ile başlamıştır. Dijital ortamdaki ilk grup iletişim sistemlerinden biri olan Usenet'te eski üyelerin yeni üyelere karşı almış oldukları tavırların dijital dışavurumu olan trolleme, ilk olarak lakap takma ve aşağılama ile kendini göstermiştir.

Trolleme kavramı dijital ortamda lakap takma, ağa düşürme ve tuzak kurma gibi geniş bir anlam muhtevasını içeren dijital eylem yöntemidir. Konuya dair Barlet (2016, s. 263) trolleme yöntemini 'trailing' adı verilen balık tutma tekniğine benzetmektedir. Söz konusu teknikle oltanın ucuna bir yem takılır ve yeme takılan av yavaşça çekilir. Usenet gibi topluluk iletişimine dayalı bir sistemle başlayan trolleme eylemi, bugün çevrimiçi kültürdeki topluluk iletişimine dayalı tüm platformlarda aktif bir şekilde kullanılmaya devam etmektedir. Özellikle oluşturulan dijital kimliklerin nesnel gerçeklikten kopuk bir yaşam vaat ettiği formlar ve sosyal medya gibi çevrimiçi ortamlarda teknolojinin gelişime paralel olarak güncellenmiş ve daha inandırıcı, daha kırıcı, daha aşağılayıcı ve daha organize bir hal almıştır.

Çevrimiçi kültürün ortaya çıkışından beri var olan trolleme eylemi, herhangi bir tutum ve davranışa karşı gösterilecek tepkinin bir yöntemi olabileceği gibi, aşağıda belirtilen amaçlar içinde yöntem olarak kullanılabilir.

- Tuzaga düşürmek
- Faydasız ve sonucu olmayan tartışmaları başlatmak
- Yalan ve yanlış enformasyon paylaşmak
- Aşağılamak ve alay etmek
- Siyasi provokasyon yapmak
- Tahrik etmek
- Eğlenmek

Yeni iletişim teknolojileri ile iletişim unsurunun yeniden dizayn edildiği içinde bulunduğumuz dijital çağ, kullanıcıların sürekli bilgi bombardımanına tutulduğu ve sürekli akışkan enformasyon ile karşılaştığı bir dönemi kapsamaktadır. Bu dönemde özellikle kullanıcılar dijitalleşmenin sunmuş olduğu eğlence, sosyalleşme, alışveriş yapma gibi edimlerinin yanı sıra nesnel gerçekliğine yön verecek bilgi alma ve öğrenme gibi ihtiyaçları içinde dijital platformlarda vakit geçirmektedirler. Ancak nesnel gerçekliği tanımlayan geleneksel yaşam normlarının oluşturulan dijital kimlikler ile dijital çıktılarda esnemesi, silikleşmesi ya da yok olması dijital ortamlarda yeni norm standartlarının ortaya çıkmasına neden olmaktadır. Bu norm standartları dijital kimliklerin aynı zamanda kimliksizlik, tanımsızlık ya da belirsizlik ile kutsanmasına öncelik verirken aynı zamanda dijitalle yönelik yeni aktörlerin ortaya çıkmasına ve iletişim unsurunun şeffaflıktan uzaklaştırarak güvensiz, tedirgin edici bir hal almasını da kurgulamaktadır. Bu bağlamda nesnel gerçekliğini gizleyip dijital kimliğin vermiş olduğu tanımsız kimliğin en önemli temsilcilerinden biri de trollerdir.

Trollerin en önemli amacı gerçeği saptırmak ve hedefte yeni bir gerçeklik algısı oluşturmaktır. Bu amacı gerçekleştirirken görme ve işitme duyularına hitap edecek içerikler hazırlamak en fazla kullandıkları yöntemlerdir. Bu durumda özellikle deepfake gibi teknolojiler gerçeği çarpıtmak ya da yeniden kurgulamak için tam da trollerin ihtiyacı olan görerek ve duyarak şahit olma mantığını hayata geçirme noktasında önemli bir açığı kapatmaktadır. Görerek ve işiterek elde edilen verilerin güvenilir ve doğru olduğu geleneksel yaşam normlarının aslında dijital platforma yansımalarıdır. Nesnel gerçekliğini sanal gerçeklikle bütünleştirip melez bir yaşam pratikleri oluşturan sanal ağların kullanıcıları, bu gelenekseldeki alışkanlıklarını çevrimiçi yaşamında devam ettirmektedir.

Deepfake teknolojisi hali hazırda sanal uzamlarda trolleme teknolojisi olarak kullanılmaktadır. Kanıt isteyen kullanıcıya kanıt, şahit isteyen kullanıcıya şahit sunma potansiyelindeki bu teknoloji dijital platformlardaki gerçeği biçimlendirme şekillerinin gelenekselden farklı olabileceğinin de önemli bir göstergesidir. Bu bağlamda gerçeğin varlığını gizleyerek oluşturduğu yeni ve tanımsız kimliği ile troller dijital ortamlarda etkin bir yönlendirici odak olarak bulunmaktadır.

Troller aracılığı ile çevrimiçi kültürde dolaşıma giren içerikler gerçeğin kıstasların uzak dezenformasyon şeklindedir. Bu durum dijital ortamlarda geleneksel normları ile okuma yapan kullanıcılar için dijital ortamlardaki içeriklere güvenmeme sorununu zaman sonra ortaya çıkarmaktadır. Nihayetinde gerçeklik arayışında yaşadığı güvensizlik artık kullanıcıların içerikleri sorgulamamasına ve gerçekten kopuşuna neden olmaktadır.

Deepfake ile dijital platformlarda trolleme kültürünün, giderek yaygınlaşması ileri düzeyde kullanıcıların gerçekleri bile sahte olarak algılamasına neden olmaktadır. Gerçeğin ve sahtenin hangisi olduğu çıkmazı ile dijital platformlarda gezinen kullanıcılar inanmak istemedikleri olaylar karşısında gerçeğe karşı önyargılar oluşturmaktadırlar. Bu bağlamda trolleme kültürünün deepfake teknolojisi ile gerçeği tahrip etmesi ve yeni gerçekler inşa etmesi dijital platformlarda

varlık gösteren tüm kullanıcıların her şeyi aldatma olarak görmesine neden olmaktadır. Başka bir deyişle, en büyük tehdit insanların aldatılacakları değil, her şeyi aldatma olarak görmeye başlamalarıdır.

Gerçeğe karşı ilgisizlik sendromu dijitalleşme ile sorunlaşan gerçek ve sahtenin iç içeliği karşısında bireyin eylemsizliği olarak ifade edilebilir. Bu sendromun en önemli sebeplerinden biri deepfake teknolojisinin dijital uzamlarda, yapılabılır yetilerin sınırlılığının her geçen gün ortadan kalkmasıdır. Bu gerçekten hareketle deepfake teknolojilerinin dijital ortamlardaki en önemli tehlikesi derin sahtekârlıklar ya da kendi başına dezenformasyonların kurucu unsur olması değildir, kötü niyetli aktörlerin bu teknolojiye kullanmaları neticesinde yanlış ya da sahte ile sürekli yüzleşen insanların video ve sesler de dâhil bu kadar bilgiye güvenilmeyecek hissetmelerine neden olmasıdır (Westerlund,2019, s. 43).

Öneriler ve Sonuç

Günümüzde dijital uzamların bir çıktısı olan sosyal medya içeriğın dağıtımını, kopyalanması, depolanması ve çoğaltılması bakımından önemli bir işleve sahiptir. Özellikle bu mecralarda kullanılan yapay zekâ teknolojisi ile nesnel gerçekliğe sanal dokuların eklenerek dolaşıma sokulması, içeriklerin zararsın eğlenceli ya da sanatsal bir çalışma olarak değerlendirmesine yol açsa da, derin sahtekârlığın dijital mecrada görünür olması gibi bir tehlikeyi de ortaya koymaktadır.

Yapay zekânın bir ürünü olan deepfake teknolojisi, nesnel gerçeklikte sahtekârlık olarak değerlendirilen ve gerek sosyal yaşam formlarında ve gerekse de hukuk karşısında cezai müeyyideler gerektiren eylemlerin, sosyal medyadaki başlıca faillerinden biridir. Dijital platformlarda sahtekârlığın meşrulaştırılmasında başlıca faillilerden biri olan deepfake teknolojisindeki sahtekârlıklar, gerçeğın manipüle edilmesi, kaynak olarak kullanılan gerçek kişilerin izinsiz olarak görüntülerini başka amaçlar için kullanılması ve bu kişiler üzerinden olmayan bir gerçeği üretip kitleleri aldatması şeklinde çok boyutlu bir şekilde kendini göstermektedir.

Deepfake ile dijital uzamlarda sürekli sahte ve sentetik olarak oluşturulan içerikler ile karşılaşan kullanıcılar, sahtekârlığa karşı fiziksel yaşamdan edindiği algılarını dijital kimliği ekseninde yeni baştan kurgulamakta ve uzun vadede meşrulaşan sahtekârlık eğlenceli bir video ya da sanatsal bir eser olarak algılanmaktadır. Derin sahtekârlığın kurucu unsuru olan deepfakeler içerik üretim aşamasında takipçi sayısı fazla olan ya da izlenebilirliği yüksek olan ünlülere, siyasilere ya da kurumsal liderlere odaklanır. Kaynak olarak kullanılan bu ünlü, siyasi ya da popüleritesi yüksek liderler üzerinden derin sahtekârlık içeriklerini dolaşıma sokan kötü niyetli sosyal medya kullanıcıları yapay zekâ teknolojisini sınırlarını sonuna kadar kullanarak diğer kullanıcıların gerçeklik algılarını yönlendirilebilir bir düzeye getirmektedir.

Dijital platformlarda deepfake teknolojisi ile içeriklerini desenleyen kullanıcı sayısı her geçen gün artmaktadır. Bu durumun birçok sebebi bulunmaktadır. Ancak özellikle çevrimiçi kültür araçlarının deepfake teknolojisinin içeriklerini kutsaması ve kullanıcılarına bu konuda özgürlükçü ve sınırsız bir imkân tanınması bu sebeplerin başında gelmektedir. Ayrıca asgari düzeyde bilgisayar kullanma becerisi olan herkesin amatör bir şekilde deepfake içeriği üretebilmesi de dijital platformlardaki deepfake kullanan kullanıcı sayısının artmasına neden olmaktadır.

Bir tür gerçeklik inşası olan deepfake teknolojisinin kullanıcıları sadece eğlenmek amacı güden amatör olarak nitelendirilen yarı uzmanlar değildir. Zaten bu kitle depfake teknolojisinin masum yüzünü temsil etmektedir. Asıl sorun deepfake teknolojisi ile profesyonelce uğraşan kitledir. Amatör kullanıcılar tarafından üretilen içeriklerin gerçekle olan bağıını çözmek oldukça

basit iken, derin öğrenme algoritmalarını kullanan profesyonel kötü niyetli aktörlerin oluşturdukları içeriklerin gerçekle olan organik bağı sıradan sosyal medya kullanıcılarının çözmesi oldukça zordur. Temel olarak sorunda bu noktada kendini göstermektedir. Gerçek ile sahtenin hangisi olduğu düşüncesi ile gerçeği aramaktan vazgeçen dijital kimliklerin gerçeği istediği gibi yorumlaması ve bu yorumladığı gerçek üzerinden fikirlerini savunması neticede gerçeğin artık önemsiz bir argüman olduğunun ve değerinin dijital kanılar üzerinden ölçüldüğünün bir göstergesidir.

Teknoloji insanlığın yararı için kullanıldığı zaman hayatı kolaylaştıran ve medeniyetlerin çağ atlmasına vesile olan önemli dinamiklerden biridir. Ancak kötü niyetli insanların teknolojiyi zarar verici bir silah haline dönüştürdüğü zaman bunun mücadelesini vermek ve ortaya çıkacak olası zarar ve mağduriyetlerin önüne geçmek de gerekmektedir. Bu bağlamda deepfake teknolojisinin de derin sahtekârlıklar için kullanılan bir teknoloji olmaktan çıkarmak oldukça zor ve zaman gerektiren bir süreçtir.

Teknoloji ile mücadele etmenin en iyi yolu yine teknolojiyi kullanabilmektir. Bu bağlamda öncelikle deepfake teknolojisinin özellikle sosyal medya ve içerik sağlayıcı platformlarda karakter ya da kimlik hırsızlığına yönelik kullanımı engelleyebilecek anti-deepfake yazılımları geliştirilmeli ve bu yazılımlar dijital kimliği olan her kullanıcı tarafından aktif kullanmalarına olanak sağlanmalıdır.

Yapay zekâ teknolojisi ile evrim geçiren iletişim araçları içinde bulunduğumuz çağda çok yönlü bir kullanım kapasitesine sahip olmuştur. İnternet bağlantısı sayesinde bu araçların genişleyen kullanım alanları içinde her kullanıcı bu araçlar ile farklı düzeylerde etkileşim içine girmektedir. Özellikle kullanıma hazır yazılımlar ya da uygulamalar satın alarak ya da ücretsiz bir şekilde indirerek etkileşim düzeylerini artıran kullanıcılar, yazılımlardan ve uygulamalardan aldıkları çıktıları dijital kimlikleri ile dahil oldukları sosyal topluluklarda dolaşıma sokmaktadır. Derin sahtekârlığın ilk basamağı olan bu yazılım ve uygulamalar, belli ücret karşılığında ya da sadece üye olmak şartıyla indirebilen kullanıcıların bu yazılımlardan ve uygulamalardan elde ettikleri içerikleri eğlence ya da sanatsal bir veri olarak dolaşıma sokmaları henüz çevrimiçi kültürde olmasa bile geleneksel yaşamsal normlarda suç olduğunun farkında olmamalarının bir göstergesidir. Bu bağlamda özellikle eğitim yaşına gelmiş her bireyin medya okur-yazarlık ya da yeni iletişim teknolojilerine yönelik yaşam boyu bir eğitime tabi tutulmaları gerekmektedir.

Ayrıca bu yazılım ve uygulamaların ciddi bir denetimden geçirilmesi ve kullanım alanlarının sınırlandırılması gerekmektedir. Herkes tarafından kolayca indirilip kullanılmasının önüne geçilmeli ve kullanmak isteyenlerin hangi maksatla kullanmak istediğinin tespitinin yapılması ve kullanım süresi boyunca takip edilmesi gerekmektedir. Ayrıca hayata dair henüz tam anlamıyla bilinç ve deneyim kazanamamış belli yaşın altındaki kullanıcılar için internet kullanma sorumluluğunun aileye verilerek hem devletin denetim yükünü hafifletmesi hem de gizliliğin ve mahremiyetin korunmasının yakın takibe alınması sağlanacaktır.

Hükümetler ve devletler vatandaşlarını dijital platformlarda doğabilecek her türlü mağduriyetten korumak zorundadır. Bu mağduriyetlerin en önemlilerinden biri de karakter, yüz ve ses hırsızlıklarına imkân veren deepfake teknolojilerinden doğan mağduriyetlerdir. Bu durumda özellikle geliştirilecek anti-deepfake yazılımların küresel boyutta destek görmesi ve savunma politikaların dijital uzamdaki tehditlere yönelik güncellemeleri içermesi oldukça önemlidir. Öte yandan deepfake teknolojisinin zarar verici ve dijital medyalarda özellikle sosyal medyalarda hesapları olan kullanıcıları tedirgin edici kullanımının önü geçilmesi için ülkelerin dijital suçla mücadelesini destekleyecek yasalar oluşturmalıdır.

Dijital suçla mücadele bugün dünyada oldukça güç bir hal almıştır. Özellikle internetin

henüz tam anlamıyla anlaşılabilen 'Derin ve Karanlık Web' katmanlarındaki suç örgütlerinin, kartellerin, kaçakçıların vd. diğer bireysel suçluların güvenli bir limanı haline gelmiştir. Kimlik hırsızlığında yapıldığı ve kimliklerin suça ve ahlaksızlığa bulaştırıldığı internetin bu derin ve karanlık katmanları yapay zekâ teknolojisi yardımıyla artık kontrol edilebilir bir alan olması özellikle ünlülere odaklanan deepfake teknolojisi gibi kötü amaçları gerçekleştirmeye açık teknolojilerin kontrolü için önem arz etmektedir.

Kaynakça

Bahar, U. M. ve Afsana, S. (2020). Deep Insights of Deepfake Technology : A review. *DUIJASE*, 5,1-2,13-23

Bartlett, J. (2016). *Dark Net: İnternetin Yeraltı Dünyası*. (Çev: Y. Konyalı). İstanbul: Timaş.

Björnsson, G. (2016) Outsourcing The Deep Self: Deep Self Discordance Does Not Explain Away Intuitions in Manipulation Arguments, *Philosophical Psychology*, 29,5, 637-653

Chapaux-Morelli, P. ve Couderc, P. (2017). *İkili İlişkilerde Duygusal Manipülasyon: Narsist Bir Partnerle Yüzleşmek*. İstanbul: İletişim

Chomsky, N. (1993). *Medya Gerçeği* (Çev. Yılmaz, Abdullah) İstanbul: Tüm Zamanlar

Güera, D. ve Delp, E. J. (2018) Deepfake Video Detection Using Recurrent Neural Networks, *2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, 2018, s. 1-6, doi: 10.1109/AVSS.2018.8639163.

Iyu, S. (2020). Deepfake detection: Current challenges and next steps. In *2020 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)* s. 1-6. IEEE.

Manovich, L. (1996) . *On Totalitarian Interactivity* (notes from the enemy of the people) http://manovich.net/content/04-projects/017-on-totalitarian-interactivity/14_article_1996.pdf.

Özer, N. P. (2019). Gazete Haberlerinde Manipülasyon 28 Şubat Örneği. *Medeniyet ve Toplum Dergisi*. 3, 2, 213-236.

Schiller, H. (1993). *Zihin Yönlendirenler*. (Çeviren: Cevdet Cerit). İstanbul: Pınar

Westerlund M. (2019). The Emergence of Deepfake Technology: A Review. *Technology Innovation Management Review*, 9, 11, 39-52.

İnternet Siteleri

Biber, A.B. ve Ulukan, B. (2018) Yapay zekanın korkutan teknolojisi: Deepfake. Erişim adresi: <https://www.trthaber.com/haber/dunya/yapay-zekanin-korkutan-teknolojisi-deepfake-432631.html>. Erişim tarihi: 30.12.2021

Dickson, B. (2020) What are Deepfakes. Erişim adresi: <https://bdtechtalks.com/2020/09/04/what-is-deepfake/> Erişim tarihi. 10.01.2022

Hodge,M. (2022) Samsung's 'deepfake' video of a TALKING Mona Lisa painting. Erişim adresi: <https://headtopics.com/uk/samsung-s-deepfake-video-of-a-talking-mona-lisa-painting-6054772>. Erişim tarihi: 12.01.2022

Extended Abstract

New communication technologies to take an active and inclusive role in creating a hybrid life for users of virtual environments that emerged in the digitalization process. Individuals who construct an alternative life practice in these environments with their digital identities personally experience what technology can do by integrating with the phenomenon of

communication. In particular, users who adopt the method of self-expression on the axis of software and applications and are involved in the interaction process begin to live with the eternal mystery of virtuality and moving away from the norms they have acquired from the objective reality within this alternative life practice.

The fact that artificial intelligence technology shows itself especially in online culture and carries the communication process to a completely different stage causes digital identity to direct hybrid life by moving away from its objective identity. In this process, the effective manipulation of deepfake technology, which is an output of artificial intelligence, based on seeing and hearing, constructs a series of actions and events that are unlikely to be in objective reality, and hinders the reliability of what is virtual in the hybrid life form.

Today, those who use digital platforms for many reasons encounter a lot of data and digital content. This situation causes a break from objective reality. Users, who try to meet their daily need for entertainment, socialization or news with these contents, activate their instincts of learning by seeing or hearing in the online culture they are involved in. Although deepfake technology is an important technology in cinema, its importance in digital media becomes more obscure because it is the leading actor of deep frauds. In particular, the untrue, but convincing content that is not fake, which is created as a result of the use of this technology in digital spaces by malicious internet actors, directs the perceptions of reality of users who learn by seeing and hearing or receiving news. In digital spaces, where the perception of reality is constantly manipulated by the outputs of deepfake technologies, digital identities walking away in search of the truth interpret the content in the axis of the social class they belong to, ideological stance or the norms of life in which they participate physically.

Today, social media, which is an output of digital spaces, has an important function in terms of distribution, copying, storage and reproduction of content. Even though the artificial intelligence technology used in these channels and the articulation of virtual textures to objective reality causes the content to be evaluated as an entertaining or artistic work, it also poses a danger such as the visibility of deep dishonesty in digital media.

Deepfake technology, a product of artificial intelligence, is one of the main perpetrators in social media of actions that are considered fraudulent in objective reality and require criminal sanctions both in social life forms and against the law. Deepfake, which is one of the main perpetrators of the legitimization of fraud on digital platforms, manifests itself in a multidimensional way in the form of using images of real people for other purposes without permission and manipulating the truth.

Users who constantly encounter fake and synthetically created content in digital spaces with deepfake reconstruct their perceptions of physical life against fraud on the axis of their digital identity, and fraudulent in the long run is perceived as an entertaining video or an artistic work. Deepfakes, which are the founding elements of deep fraud, focus on celebrities, politicians or corporate leaders who have a large number of followers or high traceability during the content production stage. Malicious social media users, who circulate deep dishonesty content through these famous, political or popular leaders, used as a source, use artificial intelligence technology to the fullest, bringing other users' perceptions of reality to a manageable level.

Users of deepfake technology, which is a kind of reality construction, are not semi-experts who are described as amateurs just for fun. This mass already represents the innocent face of deepfake technology. The main problem is the audience dealing with deepfake technology professionally. While it is quite simple to decouple the content produced by amateur users with reality, it is very difficult for ordinary social media users to untie the organic connection of the

content created by professional malicious actors using deep learning algorithms with reality. Basically, it shows itself at this point in the problem . The fact that digital identities, who have given up searching for the truth with the thought of what is real and fake, interpret the truth as they wish and defend their opinions on the truth they interpret, as a result, is an indication that the truth is now an insignificant argument and its value is measured through digital opinions.