

# DERİN ÖĞRENME'NİN SİNEMA SEKTÖRÜNDE KULLANIMINA İLİŞKİN BİR İNCELEME

**Mustafa Evren BERK**

Dr. Öğretim Üyesi, Necmettin Erbakan Üniversitesi Güzel Sanatlar ve Mimarlık Fakültesi, meberk@erbakan.edu.tr, ORCID: 0000-0002-5395-6204

Mustafa Evren Berk. "Derin Öğrenme'nin Sinema Sektöründe Kullanımına İlişkin Bir İnceleme". idil, 101 (2023 Ocak): s. 75-84. doi: 10.7816/idil-12-101-07

## ÖZ

Kameranın icadıyla başlayan sinema sektörü teknolojik gelişmeler ile eşdeğer şekilde değişim ve gelişim göstermiştir. Günümüzde teknoloji, insan beyninin işlevlerini taklit eden yapay sinir ağları ve bilgisayar sistemlerinin kullanıldığı boyutlara ulaşmıştır. Derin öğrenme olarak tanımlanan beyin taklit edilmesine dayalı bu sistem bu çalışmada sinema sektörü bağlamında ele alınmış, bilgisayar destekli Deepfacelab uygulaması kullanılarak uygulamalı olarak denenmiştir. Deepfacelab uygulaması ile 2008 yapımı Kara Şövalye filminde Joker karakterini canlandıran ve hayatını kaybeden Heath Ledger ile 2019 yapımı Joker isimli filmde Joker karakterini canlandıran Joaquin Phoenix'in yüz değişimi işlemi yapılmış ve süreçler bu çalışma kapsamında ele alınmıştır. Uygulama aşaması dokuz (9) aşamadan oluşmaktadır. Videonun seçimi, kaynak videodan yüz karakterinin, kafa hareketlerinin alınması, uygun olmayanların silinmesi, kafa hareketlerinin kontur çizgilerinin oluşturulması, train (öğrenme) işlemi, tekrarlamalar, eşleşmelerin sağlanması ve ince ayarların yapılması, iki görüntünün birleştirilmesi Deepfacelab uygulamasının aşamalarıdır. Bu çalışmada bilgisayar destekli program ile oyuncular arasında yüz değişimi yapılmış, hayatta olmayan karakter geçmişteki videoları temel alınarak hayattaymış algısı oluşturulmuş ve sinema sektöründeki varlığını devam ettirebileceğini gösterilmiştir. Uygulama ile derin öğrenme kavramına dikkat çekilmiş, teknoloji düzeyi sinema sektörü üzerinden aktarılmıştır. Sonuç olarak, derin öğrenme ile hayatta olmayan bir sanatçının, siyasetçinin, sevilen bir insanın görüntüsünün, konuşmasının, mimiklerinin bilgisayar destekli program ile geçmiş gibi algılanabileceği ortaya konmuştur. Bu bağlamda bu çalışma derin öğrenme kavramı, derin öğrenme/ teknoloji ilişkisi, derin öğrenmenin sinemada kullanımı açısından önem taşımaktadır.

**Anahtar Kelimeler:** Sinema, Derin Öğrenme, Deepfacelab

*Makale Bilgisi:*

*Geliş: 19 Kasım 2022*

*Düzeltilme: 27 Aralık 2022*

*Kabul: 12 Ocak 2023*

## Giriş

Hareketli görüntünün ilham kaynağı olarak gösterilen fotoğraf sanatı, Eadweard Muybridge'in deneysel çalışması ile birleşerek sinema tarihinin ilk ayak seslerinin somut bir örneği olarak tarihe kazınmıştır. 1878 yılında Kaliforniya Valisi Leland Stanford, arkadaşı ile girdiği tartışma sonucu atların dört ayaküstünde durup durmadığına dair araştırma yapması için Muybridge'i görevlendirir. Muybridge, 1878'de Kaliforniya Palo Alto'daki bir yarış pistinde, atın pist boyunca koşacağı alana on iki kamera yerleştirilir. Her kameranın çekim mekanizması uzun bir telle birbirine bağlanır. Bu yöntem ile ilk kez farklı görüntülerle her an değişen, ancak nihai sonucu aynı olan bir fotoğraf serisi hareketi yapmak için kullanılır (Onaran, 1994:13). Bu deneysel çalışma hareketli görüntünün temellerinin atılmasında sinemanın önünü açan önemli oluşumlardandır. Hareketli görüntüler üzerinde illüzyonları ile bilinen George Melies, bir gün Paris'teki Opéra Garnier'de çekim yaparken kamerası aniden bozulur ve durur. Kameranın önündeki hayat, kamera tekrar çalıştırılana dek zaman akıp geçer. Méliès film izlerken birdenbire bir kadının bir erkeğe ve bir otobüsün cenaze arabasına dönüştüğünü görür. "Hareketli görüntü" teknolojisinin harika ve doğaüstü güçlerini bu şekilde keşfeder. Kısa sürede bindirme, eritme, çerçeve dondurma, hareketi hızlandırma ve yavaşlatma, soldurma, maskeleye, geriye doğru hareket gibi teknikleri geliştirir. (Abisel, 2003:56). Günümüzde de filmler artık bilgisayar teknolojileriyle birlikte görüntüde manipülasyonun en üst düzeyde gerçekçi olarak seyirciye sunulmasına imkânı sağlar. Ancak yapılan işlemler geçmişten günümüze insan zekâsı ve yeteneği ile birlikte ortaya çıkarken, günümüzde insan eliyle oluşturulmuş bilgisayar kodlamasıyla birlikte yapay zekâ olarak da adlandırabileceğimiz makine öğrenme ve derin öğrenme olarak adı verilen yapay sinir ağları ön plana çıkmıştır. Bu üç kavram da birbiriyle eş zamanlı olarak iletişim halinde olsalar da işlevleri farklıdır. İnsanlar, son zamanlarda "yapay zekâ işimizi elimizden alabilir mi?" sorusuyla sık sık karşılaşmaya başlamıştır. Derin öğrenme çalışmaları son dönemde sinema alanında teknik aletlerin uygulama noktasında ve post prodüksiyon işlemlerinde yazılımların içine girmeyi başarmıştır. Bu türden teknoloji ve yenilikler bir kısım için faydalı bir konumda olurken diğer kısım için ise bir tehlike olarak görülmektedir. Yapay zekâ, insanların işlerini daha da kolaylaştırmak için kullanılmaktadır ve örneğin, James Cameron'ın yönetmenliğini üstlendiği Alita filmi gibi yapımlarda, yapay zeka teknolojisi yüz kaslarının hareketlerini yorumlamak ve gerçekçi animasyonlar oluşturmak için kullanılmıştır. (Seymour, 2019). Derin öğrenme teknolojileri sanat alanında giderek daha fazla kullanılmaya başlanmıştır ve bu sayede sanatçılar, yeni ve ilginç görsel ve işitsel deneyimler yaratma konusunda daha özgür hale gelmektedir.

## Derin Öğrenme

Derin öğrenme, insan beyninin işlevlerini taklit eden yapay sinir ağları ve bilgisayar sistemleri kavramına dayanmaktadır. Derin öğrenmenin tarihi, Warren McCulloch ve Walter Pitts'in 1913'te bir sinir ağı düşünce sürecini taklit etmek için matematiğe ve sinirsel mantık adı verilen algoritmalara dayanan bir bilgisayar modeli oluşturduğu zamana kadar uzanır. 1958'de, iki katmanlı bir bilgisayar yapay sinir ağına dayanan denetimli bir eğitimsel örüntü tanıma algoritması olan Perceptron, Frank Rosenblatt tarafından basit toplama ve çıkarma işlemleri kullanılarak oluşturulmuştur (Kayaalp ve Süzen, 2018:7). 1965 yılında Alexey Ivakhnenko ve Grigorevich Lapa, toplama ve çarpımlara dayalı doğrusal olmayan aktivasyon fonksiyonlarını kullanarak, keyfi olarak birçok nöron benzeri eleman katmanına sahip, denetimli derin ileri beslemeli çok katmanlı algılayıcılar için çalışan öğrenme algoritmasını yayınladılar. Regresyon ve ayrı bir doğrulama seti kullanarak dahili temsilleri öğrenmek için ağ katmanlarını aşamalı olarak eğittiler. Örneğin, Ivakhnenko'nun 1971 tarihli makalesinde, yeni milenyumda, özellikle de makine öğreniminin çoğunun doğduğu Doğu Avrupa'da hala popüler olan, yüksek oranda alıntı yapılan yöntemlerle eğitilmiş 8 katmanlı bir derin öğrenme ağının tanımlanmış olduğunu belirtmiştir (Schmidhuber, 2015). Ivakhnenko'dan sonra ilk "Neokognitron" derin öğrenme mimarisi 1979'da Fukushima tarafından önerilir. Omurgalıların görsel sinir sisteminden esinlenen yapıda, "denetimsiz öğrenme" ile kendi kendini organize eden bir ağ geliştirilir. Fukushima'nın ağı, modern ağa benzer şekilde birden fazla eğrilik katmanı ve küme içerir. Derin mimarinin çoklu katmanlarındaki hataların geri yayılması, en büyük öğrenme kusurudur. Geri yayılım algoritmaları önceki yıllarda tanıtılmış olmasına rağmen, derin sinir ağlarının ilk başarılı uygulaması Yann LeCun ve diğerleri tarafından geliştirilmiştir (Şeker ve diğerleri, 2017:49).

Derin öğrenme, gelişmekte olan bir makine öğrenimi araştırması alanıdır. Yapay sinir ağlarının birden fazla gizli katmanını içerir. Derin öğrenme metodolojisi, büyük veri tabanlarında doğrusal olmayan dönüşümler ve yüksek düzeyde model soyutlamalarını uygular. Çok sayıdaki alanda derin öğrenme mimarilerindeki son gelişmeler, yapay zekaya şimdiden önemli katkılar sağlamıştır (Vargas vd., 2017:1). Derin Öğrenme, Makine Öğreniminin çok popüler ve yaygın olarak kullanılan bir alt kümesidir. Yapay zekâ ve makine öğrenimini bir araya getirir. Derin Öğrenme, çok katmanlı Yapay Sinir Ağı'na (YSA) dayanır.

Konsept, insanlardaki biyolojik sinir ağlarına dayanmaktadır. Bu ağlar, alınan verileri anlamak ve analiz etmek için birbirleriyle iletişim kuran düğümler ve çok katmanlar içerir. Derin Öğrenme, Makine Öğrenimi ile benzer özellikleri paylaşır. Her iki kavram da eğitim modellerine ve verileri sınıflandırmaya dayanır. Bilgisayara örneklerden öğrenme ve örnekleri tanıma yeteneği sunar. Makine Öğreniminin bir alt kümesidir ancak daha karmaşıktır. Derin Öğrenme, verilerle başa çıkmak için yüksek bilgi işlem performansı ile birlikte büyük verileri kullanmalıdır. Görüntü İşleme, Doğal Dil İşleme veya Ses Tanıma ile ilgilendirilir. Derin Öğrenme, veri türüne bağlı olarak verileri eğitmek için farklı yaklaşımlar kullanır (Alaskar ve Saba, 2021:5). İnsan beyninin karmaşık problemleri gözlemlenme, analiz etme, öğrenme ve karar verme yeteneğini taklit eden ve çıkarma, dönüştürme ve manipülasyon gibi işlemleri gerçekleştirebilen bir makine öğrenme tekniğidir (Kayaalp ve Süzen, 2018:7). Çalışmada kullanılan Deepfacelab uygulaması karar verme mekanizması olarak da çekişmeli üretici ağ algoritmasını kullanmaktadır. Çekişmeli Üretici Ağlar (Generative Adversarial Network) modellerinin genel mimarisinde biri üretken diğeri ayrıştırıcı olmak üzere iki farklı derin ağ bulunmaktadır. Bu rakip ağlardan, üretken ağı girdi olarak aldığı gürültü vektörü eğitim setindeki gerçek görüntüleri simüle etmeyi amaçlarken, bölünmüş ağ, oluşturulan bileşik görüntüler, çıktı (yanlış) ve gerçek eğitim görüntüleri arasında ayırım yapmaya çalışır (Çelik ve Talu, 2020:183). Çekişmeli Üretici Ağlar, gerçek dünya veri kümelerinden karmaşık kalıpları öğrenen, enerji tabanlı bir mimaridir. Tartışmalı sentetik ağı temel yapısı ünlü matematikçi John Nash'in oyun teorisine dayanır ve birbirine zıt çalışarak öğrenen iki bütünleştirici modülden oluşur. Sentetik ağ, giderek daha gerçekçi (yanlış) yapay görüntüleri sentezlemeye çalışan bir üretken ağ iken, ayrıştırma ağı, gerçek görüntüleri sahte görüntülerden ayırmaya çalışan basit bir ikili sınıflandırıcıdır. Yapımcı, ayırmacıyı kandırmak için sahte görüntüler oluştururken, ayırmacı zaman içinde öğrenir ve sahte görüntüsünü gerçek görüntüyle karıştırmamaya çalışır (Şahin ve Talu, 2021:1577).

### Derin Öğrenmenin Sinema alanında uygulamaları

Derin Öğrenme ile oluşturulan uygulamalar hayatımızın birçok alanında yer almaktadır. Sağlık sektörü bunların başında gelmektedir. Ancak son dönem post prodüksiyon işlemlerinde hızlı kararlar vermek adına derin öğrenme sistemleri sinemada da kullanılmaya başlanmıştır. 2019 yılında yayınlanan Irishman filminde oyuncuların gençleştirme işleminde yapay zekâ yazılımlarına başvurulmuştur. ILM (Industrial Light and Magic), 2015 ile 2017 yılları arasında yepyeni bir yapay zekâ Face Finder programı geliştirdi. ILM'nin görsel efekt süpervizörü Paul Helman, ILM ekibinin "inanılmaz bir kitaplık" oluşturduğunu söyler. Üç ana aktörün hedeflenen tüm yaşlarındaki kliplerini ve resimlerini Face Finder ile toplarlar. Ardından, ihtiyaç duyulan karakterin göz ifadelerini arayabilmeleri için bölümlere ayırdıkları kitaplıktan faydalanırlar. Örneğin sanatçılar, belirli bir yaşta De Niro'nun gözlerini arayabilir, böylece ihtiyaç duyacakları tüm bu farklı performans ve aydınlatma bileşenlerine erişebilme imkânı sağlanmış olur. Karakterin mevcut zamanda çekilen sahne performansına dayalı olarak bir çekim üç boyutlu olarak işlendiğinde, Face Finder programı Robert De Niro'nun yüz ifade kitaplığından çekime uygun benzer açılara, yaşlara sahip alternatifler resimler (kaplamalar) bulur. Geçmişteki filmlerinden kütüphanesinden seçilen bu resimler üç boyutlu modelin üstüne bindirilir (Seymour, 2019).

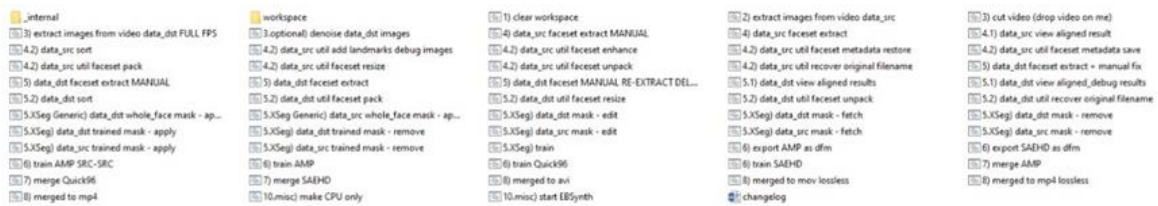
Derin öğrenmenin son zamanlarda yapmış olduğu sinema alanındaki diğer bir örneklerinden bir tanesi de IBM şirketinin geliştirdiği yazılım olan Watson ile bir fragman oluşturulmuştur. Herhangi bir yapay zekâ sisteminde olduğu gibi Watson'un konu alanında eğitilmesi işlemiyle işe başlanmıştır. Yüz (100) korku filmi ile eğitilen Watson görsel, işitsel ve sahne kompozisyonu analizini yaparak filmleri etiketleri ayırır. Watson'u "Morgan" filminin uzun haliyle besleyen IBM ekibi, fragman için en uygun olan on (10) anı Watson tarafından belirlendiğini ifade eder. Doksan dakikalık filmde Watson'un fragman oluşturması yirmi dört saat sürer. Normal şartlarda bir fragmanın oluşturulmasının yoğun bir emek gerektiren bir işi Watson, bunu bir gün içerisinde tamamlamıştır. IBM üyesi ve yöneticisi John Smith, bu çalışmanın yapay zekâ ve insan yaratıcılığının gelecek vaat eden bir alan olmasını umduklarını belirtmektedir (Smith, 2016).

Yapay zekâ araştırmaları sanat alanında yavaş yavaş her alanda gelişmeye başladığını, son dönemdeki çalışmalardan görülmektedir. Hikâyenin başlangıcı olan senaryo yazımının yapay zekâ ile birleşmesi, önümüzdeki yıllarda farklı boyutlara geleceğinin habercisi olarak gözükmektedir. Bunun en somut örneği ise senaryosu yapay zekâ tarafından yaratılan Sunspring filmidir. Sunspring, tamamen yapay bir zekâ tarafından yazılan ve Oscar Sharp tarafından yönetilen 2016 yapımı deneysel bir bilim kurgu kısa filmidir. Sci-Fi London'ın 2016'daki 48 saatlik film yapımı yarışmasına bir giriş filmi olarak yapılan bu kısa film, tamamen bir yapay zekâ tarafından yazılmıştır. Sci-Fi London'ın yarışmasının kurallarında filmin 48 saat içinde çekilmesi gerekmektedir. Yönetmen Oscar Sharp, çoğu 1980'ler ve 90'lardan kalma birkaç yüz bilimkurgu senaryosunu uzun, kısa süreli hafıza tekrarlayan bir sinir ağında (LSTM) besler (Goode, 2018).

Filmde yapay zekâ uzmanı Goodwin, önce belleğe tek cümlelik bir komut gönderilir. Bu komutla bellek, veri tabanından hayali bir sayfaya sahnelere dökerek mevcut sözcük ve deyimden hangisinin birbirini tamamlama eğiliminde olduğunu saptar. Tekrarlayıcı olduğu için sonunda senaryo yapısını taklit etmeyi, sahne kontrolünde ustalaşmayı ve karakter koordinasyonunu öğrenir. Daha sonra Sharp Benjamin adını verdikleri yapay zekâya ona kendi senaryosunu yazmasını söyler. Sonuç ilham vericidir ama mükemmel değildir (Goode,2018). Sanat alanında diğer bir örnek ise Foley işlemi olmuştur. Film yapımlarında Foley sanatçısı, filmin seyirciler için ses canlandırmasına yardımcı olan işten sorumlu kişidir. Sanatçı, sahnede tasvir edilen hem sesin hem de hareketin özünü hangi yapay sesin yakalayacağına karar verir. Elektrik ve Bilgisayar Mühendisliği'nde bölümünde eğitim gören Sanchita Ghose, AutoFoley adlı yazılım ile derin öğrenme algoritmalarıyla destek alarak yaptığı çalışmada, deneklerin %73'ünden fazlasının oluşturulan ses parçasını orijinal olarak kabul ettiğini göstermiştir. Bu durum ses sentezinde modlar arası araştırmalarda dikkate değer bir gelişmedir. Diğer araştırmalarda olduğu gibi Ghose, sessiz filmlere derin öğrenme görüntüye uygun ses yerleştirmesinde birçok ses kütüphanesinde beslenerek görüntüye uygun sesleri vermeyi başarmıştır (Ghose, 2020:1).

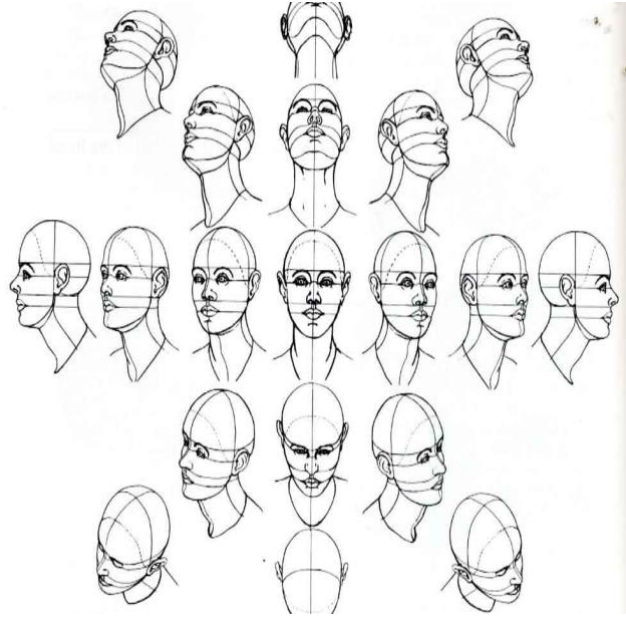
### Derin Öğrenme Kavramının Deepfacelab ile Kara Şövalye Filmi Üzerinden Uygulanması

Çalışmada, 2008 yapımı Kara Şövalye filmine hayat veren Joker karakterini canlandıran ancak genç yaşta vefat eden Heath Ledger oyuncusunun yüzünü, 2019 yapımı "Joker" filminin ana karakterini canlandıran Joaquin Phoenix suretleri ile yer değiştirme işlemleri yapılacak olup, süreçleri detaylandırılacaktır. Çalışmada Quadro RTX 4000 ekran kartı ve Xeon işlemcili bir bilgisayar kullanılmıştır. GPU adı verilen Graphics Processing Unit (Grafik İşlemci Ünitesi) 8 GB olarak bilgisayar konfigürasyonunda yer almaktadır. Deepfacelab, yazılım geliştirme projeleri için web tabanlı depolaması servisi olan GITHUB üzerinden paylaşılan ve "iperov" isimli kullanıcı tarafından sunulan bir yazılımdır. Yazılım, yüz değiştirme işlemleri ve internette sahte içerikler üretmek için kullanılan popüler araçlardan bir tanesidir. Çalışmadaki Heath Ledger'in yüzünün Deepfacelab yazılımı ile beslenmesi için 2005 yapımı "Casanova" filminden sekanslar kullanılmıştır. Joaquin Phoenix'in yüzünün ise 2019 yapımı "Joker" filminden sekanslar kullanılmıştır. Deepfacelab yazılımı Görsel-1 'de görüldüğü itibariyle birtakım sırayla çalıştırılması gereken birçok uygulamadan oluşmaktadır. Görsel-1'de "workspace" adlı klasörün içinde iki tane dosya vardır. Bunlardan bir tanesi source (kaynak) diğeri de destination (hedef) olmak üzere iki dosya barındırmaktadır. Kaynak klasörüne yüzünün değiştirilmesini istediğiniz sanatçı, siyasetçi, oyuncu ya da herhangi birisinin yüzünün video kaynağının kaydedileceği yeri temsil etmektedir. Destination klasörünün temsil ettiği yer ise kaynak videonun değiştirileceği yüze ait videonun bulunduğu klasör olarak gösterilmektedir. Yüz değiştirme işleminin en iyi şekilde yapılabilmesi için yüzleri değiştirilecek karakterlerin her ikisinin de videolarının ilgili klasörlere eklenmesi gerekmektedir. Derin öğrenmeyi kullanan bu uygulama, verilen videolar üzerinden tahminlerde bulunarak en iyi yüzü hangisinin olacağına karar verecektir.



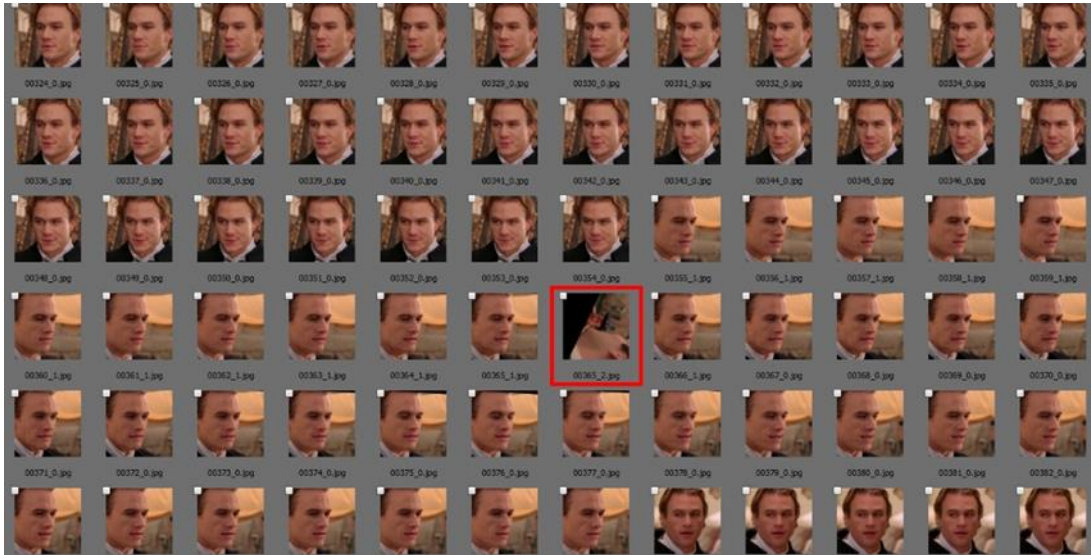
Görsel 1: 1 Deepfacelab uygulama içeriği

Bu açıdan yüzleri değiştirilecek her iki karakterin de bakış açılarının benzer ya da aynı olması gerektiği için videoların seçimi önemli bir yer teşkil etmektedir. Bir oyuncu sağa, sola ya da yukarıya baktığı açının diğer kaynak videoda da olması gerekir. Çünkü derin öğrenme aşamasında uygulamadan kaynaklı yapay zekâ, kaynak videonun kafa hareketlerine göre hedef yüzün referans yüzlerini yerleştirmeye çalışacaktır. Eğer kaynak videodaki baş hareketleri, hedef videodaki baş hareketleri yoksa yapay zekâ oraya uygun baş hareketini yerleştirecektir fakat işlemin başarılı bir şekilde tamamlanabilmesi için referans videonun bakış açıları bu açıdan önem arz eder. Görsel-2'deki resimde de görüldüğü üzere seçilecek videolardaki karakterlerin baş hareketlerinin referans alınması için iyi bir örnek gösterilmiştir.



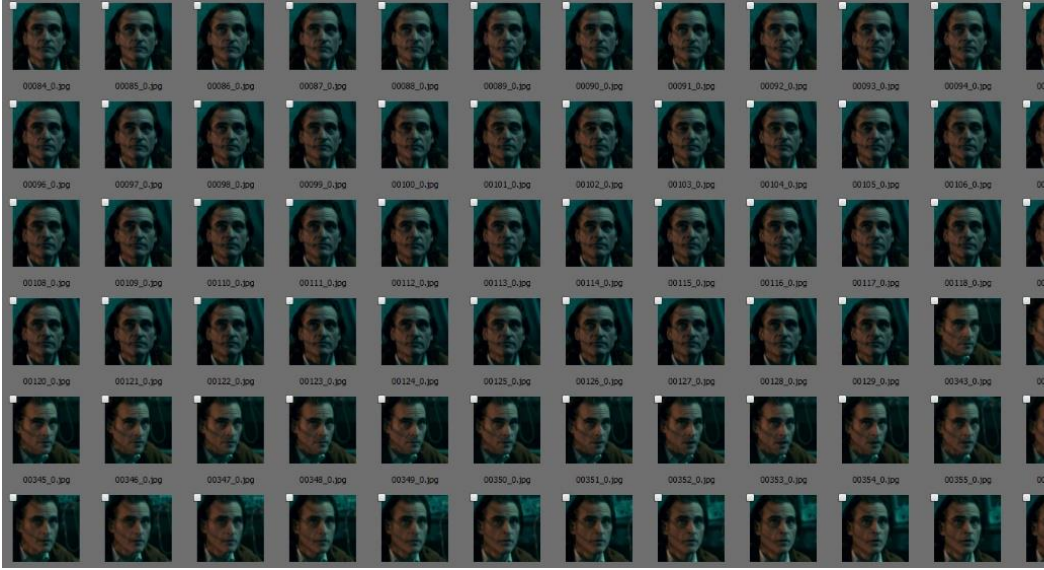
**Görsel 2:** Hedef ve Kaynak videoların seçiminde önem arz eden baş hareketleri.

Deepfacelab uygulamasının ilk aşaması videoların seçimi olarak söylenebilir. İkinci aşama ise kaynak videodan mevcut karakterin yüz kısmının çıkartılması işlemi vardır. Kaynak videodaki insan yüzlerinin teker teker çıkartılıp bir klasöre kaydedilme işlemi bu aşamada yapılmaktadır. Buradaki amaç kaynak videodaki kafa hareketlerinin alınıp hedef videodaki kafa hareketinin üzerine bindirilmesi için ayrıştırılma işleminin yapılmasıdır. Baş hareketlerinin ayrıştırılmasından sonra bunların kontrolünün yapılması gerekir. Çünkü derin öğrenme sırasında uygulama kafa hareketlerini kesin bir şekilde kavrayamamaktadır. Bu yüzden kafa hareketlerinden başka görüntüler de içerebilir. Bu açıdan bu resimlerin yine uygulama yardımı ile hesaplanacak görüntülerden ayrılması gerekir. Görsel-3'te kırmızı ile işaretlenen Heath Ledger'in baş görüntülerinden bir tanesinde tutarsız olduğu görülmüştür. Derin öğrenme sırasında uygulama bu resmi de kullanacağı için bu resmin hesaplanmaması için havuzdan silinmesi gerekir.



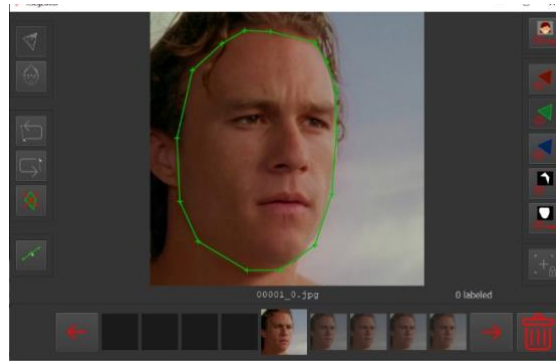
**Görsel 3:** Tutarsız kaynak Heath Ledger'in görüntüsünün gösterimi



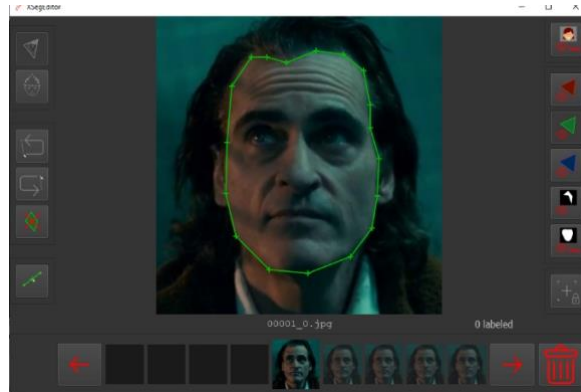


**Görsel 4:** Joaquin Phoenix baş görüntülerine ait havuz resimler

Görsel 4’de Joaquin Phoenix oyuncusunun hedef (destination) klasöründeki resimlere bakıldığında herhangi bir tutarsızlıkla karşılaşılmağıdır. Dördüncü aşama, üçüncü aşamadaki kafa hareketlerinin düzenlenmesi olarak söylenebilir. Bu aşamada yapılacak işlem kafa hareketlerinin kontur çizgilerini oluşturmaktır. Bu sayede derin öğrenmenin işi kolaylaşacak olup, hangi kontur sınırları içerisinde işlem yapılacağını uygulamaya gösterilmiştir. Normal şartlarda bu türden el ile müdahalelerine gerek yoktur ancak yazılımından iyi bir sonuç isteniyorsa bu türden müdahalelere ihtiyaç vardır. Görsel-5 ve Görsel-6’da her iki oyuncunun kontur sınırları gösterilmiştir.



**Görsel 5:** Heath Ledger’in baş sınırlarının gösterimi



**Görsel 6:** Joaquin Phoenix’in baş sınırlarının gösterimi

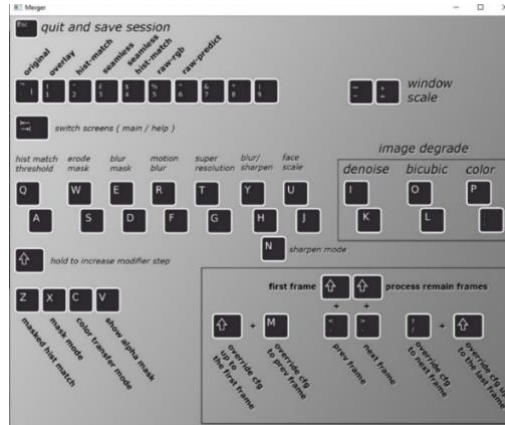
Dördüncü aşamada her iki oyuncunun da baş hizalarının kontur sınırları çizildikten sonra “train” adı

verilen beşinci aşama olan öğrenme işlemine geçmek için hazır hale gelmiştir. Görsel-7'de derin öğrenme işleminin nasıl gerçekleştiğine dair bu görüntü işlem nasıl yapıldığına dair bize fikir vermektedir.



**Görsel 7:**Derin öğrenme işleminin deepfacelab uygulaması üzerinden gösterimi

Görsel 7'de de gösterildiği üzere her kafa görüntüsünün hedef ve kaynak baş hareketlerindeki benzerlik dikkat çekmek gerekir. Derin öğrenme eğitimi sırasında uygulama en iyi eşleşmeyi bulana kadar işleme devam edecektir. Görsel-7'de üst kısımda Iter: 104534 ibaresi gözükmemektedir. Iter'in açılımı Iterations'dan (tekrarlama) gelmektedir. Uygulamanın en iyi eşleşmeyi yaparken, uygulama bize değer ne kadar yüksek olursa gerçekçilik açısından en iyi sonucu vereceğini belirtmektedir. Çalışmada 104534 tekrarlama sonucu ile yeterli eşleşmeye ulaşıldığı görülmüştür. Ancak tekrarlama sayıları ne kadar çok olursa derin öğrenme işlemi daha yoğun olacağından çıkacak olan sonuç 104534 tekrarlama daha iyi bir sonuca ulaşılacaktır. Yapılan çalışma yaklaşık olarak 4 gün sürmüştür. Normal şartlarda bu tarz bilgisayar destekli efekt çalışmalarında bu işlem daha uzun sürmektedir. Çalışmanın altıncı aşamasında, her iki görüntünün birleştirilmesine adına işlem yapılır. Bunun için yine uygulama içindeki özelliği kullanılarak yapılmaktadır. Altıncı bölümde yapılacak uygulama her iki video arasındaki renk farklılıklarını giderilmesi, kontur çizgilerinin olduğu kafa sınırlarının iki video arasındaki geçişlerinin ayarlanması gibi birçok özellik bu aşamada yapılır. Görsel-8 'de altıncı aşamada uygulamanın ince ayarlarını verildiği arayüz gösterilmiştir. Bu aşamada yüzün kenar konturları yumuşatılarak daha gerçekçi bir manipülasyonun oluşmasına imkân vermiştir.



**Görsel 8:** İki yüz hareketinin ince ayarlarının yapılmasını imkân veren pencere görünümü



**Görsel 9:** Yedinci aşama iki yüzün üstü üste bindirilmesine ilişkin görsel

Görsel 9'da en son aşama olan iki görüntünün birleştirilmesi işlemi görülmektedir. Bu aşama, altıncı aşamadaki ince ayarların son halinin verildiği aşamadır. Görüntünün son hali kare kod uygulaması ya da youtube üzerinden bir link ile çalışmanın sonuç kısmında paylaşılmıştır.



**Görsel 10:** Üsteki QR kodunu okutularak çalışmanın sonucu görülebilir.

### Sonuç

Derin öğrenme, insan beyninin işlevlerini taklit eden bilgisayar sistemlerinin kavramsal altyapısıdır. Derin öğrenme işlemi bu çalışma kapsamında deepfacelab uygulaması ile yapılmaktadır. Deepfacelab uygulaması ile Joker karakterine ilham veren Heath Ledger'in yüzü, 2019 yılında yayınlanan Joker adlı filmde rol alan Joaquin Phoenix'e monte edilerek bir manipülasyon işlemi yapılmıştır. Manipülasyon işlemi 9 (dokuz) aşamadan oluşmaktadır. Videonun seçimi, kaynak videodan yüz karakterinin kafa hareketlerinin alınması, uygun olmayanların silinmesi, kafa hareketlerinin kontur çizgilerinin oluşturulması, train (öğrenme) işlemi, tekrarlamalar, eşleşmelerin sağlanması ve ince ayarların yapılması, iki görüntünün birleştirilmesi Deepfacelab uygulamasının aşamalarını oluşturmaktadır.

Çalışmada, oyuncular arasında yüz değişimi yapılmıştır. Yüz değişimi işlemi ile yaşamsal döngüsü sona eren bir insan, aktör olarak filmde yerini almış, hayattaymış algısı oluşturulmuştur. Teknolojinin çağımızda geldiği bu boyutun sinema sektöründe birçok değişim ve dönüşümü beraberinde getireceği öngörülmektedir. Bu bağlamda hayatta olmayan aktör ve aktrislerin belirtilen uygulama ile hayattaymış gibi ekranlarda yerini alabileceği görülmektedir. Normal şartlarda bilgisayar destekli efekt ile yapıldığında çok zaman alacak bir işlem, 96 saatlik kısa bir süre içinde tamamlanmıştır. Bilgisayarların hız kapasitesi her geçen gün hızla arttığı düşünüldüğünde bu türden derin öğrenme işlemlerinin sonuçları daha da kısalacaktır.

Çalışma kapsamında yapay zekâ alanındaki gelişmeler, yapay sinir ağlarının çalışma prensipleri, derin öğrenme uygulamaları ve bu yeni teknolojilerin sinema alanını nasıl etkilediğine dair bilgiler ele alınmıştır. Yapay zekâ çalışmalarının temel amacı, çeşitli görevleri yerine getirebilecek programlar üretmekten çok bu bilinmeyen keşfine yolculuk olarak ifade edilebilir. Yapay zekâ uygulama içeriğinin yükselişiyle birlikte



zaman içinde film yapımını etkileyecek ve içeriğin gerçek kişiler mi yoksa yapay zekâlar tarafından mı üretildiğini insanların anlaması zorlaşacaktır.

Bunun yansısı müzelerde vefat eden sanatçıların eserleri gösterilirken, yine ölen sanatçının ya da siyasetçinin kendi ağzından eserlerini tanıtarak daha inandırıcı bir etki sağlanabilecektir. Yapay zekâ destekli algoritmalar sayesinde yüz değişim işlemi avantajlarının yanı sıra dezavantajları da beraberinde getirecektir. Mavi yakalılarının başında olan beyaz yakalılarının günümüz çağında rolleri ve görevleri mevcut çağa uyum sağlamaya başlamıştır. Derin öğrenme sistemlerinin sanat dalına kadar girmiş olması gelecekte bu sistemlerin insanların iş konusunda kaygı verici bir durum oluşturabilir.

(Aşamaların uygulamasına QR koddan ya da <https://youtu.be/ULT4WSe9ttM> linki üzerinden erişim sağlanabilir).

### Kaynaklar

- Abisel, Nilgün. *Sessiz Sinema*. İstanbul: Om Yayınevi, 2003.
- Alaskar, Hind, ve Tanzila Saba Saba. "Machine Learning and Deep Learning: A Comparative Review." *Proceedings of Integrated Intelligence Enable Networks and Computing*, 143-150, 2021.
- Alpaydin, Ethem. *Introduction to Machine Learning*. Cambridge, Massachusetts: MIT Press, 2020.
- Schmidhuber, Jürgen. "My First Deep Learning System of 1991+ Deep Learning Timeline 1962-2013, 2013.
- Çelik, Gaffari, ve Muhammed Fatih Talu. "Çekişmeli Üretken Ağ Modellerinin Görüntü Üretme Performanslarının İncelenmesi." *Balıkesir Üniversitesi Fen Bilimleri Enstitüsü Dergisi* 22.1 181-192, 2019.
- Ghose, Sanchita, ve John Jeffrey Prevost. "Autofoley: Artificial Synthesis Of Synchronized Sound Tracks For Silent Videos With Deep Learning." *IEEE Transactions on Multimedia* 23 1895-1907, 2020.
- Goode, Lauren. "AI Made a Movie—and the Results Are Horrifyingly Encouraging". 11 Haziran 2018. <https://www.wired.com/story/ai-filmmaker-zone-out>
- Hallström, Lasse, yön. *Casanova*. Sen. Michael Cristofer, Jeffrey Hatcher. Oyun. Heath Ledger. DVD. Buena Vista Pictures Distribution, 2005.
- Onaran, Alim Şerif. *Sessiz Sinema Tarihi*. İstanbul: Kitle Yayınları, 1994.
- Sienna Miller, Jeremy Irons, Oliver Platt vd diğer. DVD. Touchstone Pictures, 2005.
- Phillips, Todd, yön. *Joker*. Sen. Todd Phillips, Scott Silver, Bob Kane. Oyun. Joaquin Phoenix, Robert De Niro, Zazie Beetz, Frances Conroy ve diğer. DVD. Warner Bros. Pictures, 2019.
- Reddy, Poondru Prithvinath. "AI-Generated Film: The Next Phase in Movie Making", 10 Eylül 2020. <https://medium.com/@ppreddy576/ai-generated-film-the-next-phase-in-movie-making-a957f0fc0372>
- Seymour, Mike. "Weta Digital's Remarkable Face Pipeline: Alita Battle Angel", 8 Mart 2019. <https://www.fxguide.com/fxfeatured/weta-digital-remarkable-face-pipeline-alita-battle-angel/>
- Seymour, Mike. "De-aging the Irishman (updated)", 6 Aralık 2019. <https://www.fxguide.com/fxfeatured/de-aging-the-irishman>
- Smith, John R. "IBM Research Takes Watson to Hollywood with the First "Cognitive Movie Trailer", 25 Kasım 2020. <https://www.ibm.com/blogs/think/2016/08/cognitive-movie-trailer/>
- Şahin, Emrullah, ve Muhammed Fatih Talu. "Bıyık Deseni Üretiminde Çekişmeli Üretici Ağların Performans Karşılaştırması." *Bitlis Eren Üniversitesi Fen Bilimleri Dergisi* 10.4 1575-1589, 2021.
- Şeker, Abdulkadir, Banu Diri, ve Hasan Hüseyin Balık. "Derin Öğrenme Yöntemleri Vv Uygulamaları Hakkında Bir İnceleme." *Gazi Mühendislik Bilimleri Dergisi* 3.3, 47-64, 2017.
- Vargas, Rocio, Amir Mosavi, and Ramon Ruiz. "Deep learning: a review.", 2017.

# A STUDY ON THE USE OF DEEP LEARNING IN THE CINEMA INDUSTRY

Mustafa Evren BERK

## ABSTRACT

The cinema industry, which started with the invention of the camera, has changed and developed in line with technological developments. Today, technology has reached the dimensions where artificial neural networks and computer systems that imitate the functions of the human brain are used. This system, which is based on the imitation of the brain, which is defined as deep learning, has been discussed in the context of the cinema industry in this study, and has been tested in practice using the computer-assisted Deepfacelab application. With the Deepfacelab application, the character who lost his life from the actors of the 2008 film The Dark Knight was exchanged with the main character of the 2019 film Joker, and the processes were discussed within the scope of this study. The implementation phase consists of nine (9) stages. Selection of the video, taking the facial character and head movements from the source video, deleting the inappropriate ones, creating the contour lines of the head movements, training (learning), repetitions, matching and fine-tuning, combining the two images are the stages of the Deepfacelab application. In this study, a face swap was made between the actors with a computer-aided program, the perception that the deceased character was alive was created based on his past videos, and it was shown that he could continue his existence in the cinema industry. With the application, attention was drawn to the concept of deep learning, and the technology level was conveyed through the cinema sector. As a result, it has been revealed that the image, speech and facial expressions of a dead artist, politician, loved one can be perceived as real with a computer-aided program with deep learning. In this context, this study is important in terms of the concept of deep learning, deep learning / technology relationship, and the use of deep learning in cinema.

**Keywords:** Cinema, Deep Learning, Deepfacelab