

# İNSAN VE BİLGİSAYARDA YÜZ TANIMA

Albert Ali Salah

Boğaziçi Üniversitesi

Algısal Zeka Laboratuvarı, Bilgisayar Mühendisliği Bölümü

salah@boun.edu.tr

## ÖZET

Yüz tanıma, bilgisayar bilimi ve imge işleme alanlarında özellikle insan-bilgisayar iletişimi ve biyometri amaçlı bir çok araştırmaya konu olmuş bir problem. İnsanlar yüz tanıma konusunda bilgisayarların çok üstünde bir başarı sergiliyorlar. Bu çalışmada bilişsel bilim, psikofizik ve psikoloji alanındaki bulgulardan yola çıkarak bu başarının nereden kaynaklandığına bakıyoruz ve bilgisayarlar için önerilen yüz tanıma modellerini insanlarda yüz tanıma için geliştirilmiş teorilerle karşılaştırıyoruz.

## ABSTRACT

Face recognition has been an active area of study for both computer science and image processing communities, especially for purposes of better human-computer interaction and biometrics applications. Computers have yet to match humans on the face recognition task. We inspect the findings in cognitive science, psychophysics and psychology for an understanding of how humans recognize faces, and contrast these with different approaches proposed for implementing face recognizers in computers.

**Anahtar Kelimeler:** Yüz tanıma, gelişim psikolojisi

## 1. GİRİŞ

Hayvanlar ve insanlar evrim süreçleriyle iyileştirilmiş karmaşık görsel sistemleri sayesinde zorlu algılama problemlerini başarıyla çözerler. İnsan gözü öylesine karmaşık ve girift bir yapıya sahiptir ki, bu yapının parça parça oluşmasının mümkün olmadığı argümanı uzun zaman Darwin'in evrim teorisine karşı kullanılmıştır. Yine de mühendisler yarım yüzyıldır insanların algısal marifetlerini fizyolojik ve psikolojik bulguların da yardımıyla bilgisayarlarda modellemek için uğraşıyorlar.

Bu yazıda birbiriyle ilintili iki soruya cevap vermeye çalışacağız. Birincisi insanlarda yüz tanımının nasıl gerçekleştiği, ikincisi de bilgisayarlarda bu problemin nasıl çözülebileceği. Birinci soru beynin nasıl çalıştığını araştıran bilişsel bilim açısından önemlidir. Ama bu soruya verilecek yanıt ikinci soru açısından da önem taşıyor; beyinde bulduğumuz süreçler bizi bilgisayar ortamında da etkili bir çözüme götürebilir. Bilgisayarda bu problemi çözmek ise hem daha güçlü bir insan-bilgisayar etkileşimi sağlamak için, hem de

son zamanlarda çok önem verilen biyometrik güvenlik uygulamaları için gereklidir.

Bir bilgisayarda başarılı bir yüz tanıma sistemi yaratabilmek için, bir yapay öğrenme modeline ihtiyaç vardır. Yapay öğrenmeden anladığımız matematiksel ifadesi olan bir programın parametrelerinin belli bazı girdiler için (mesela yüz imgeleri) belli bazı çıktılar verecek şekilde (mesela yüzün bir erkeğe mi yoksa kadına mı ait olduğu) ayarlanmasıdır. Bu şekilde "öğrenen" bir sistem daha önceden görmediği yüzleri de doğru olarak sınıflandırabilir, yani genelleme yapabilir. Sistemin iyi genelleme yapabilmesi, girdiye ne tür dönüşümler (transformation) uyguladığına ve girdinin hangi özniteliklerini (feature) öğrendiğine bağlıdır (bkz. Şekil 1). Başarılı bilgisayar modelleri değişik özniteliklerin ve dönüşümlerin yüz tanıma problemine ne kadar katkı sağladığını sayısal olarak ortaya koyabilir, ve bu şekilde bilişsel hipotezlerin desteklenmesinde veya çürütülmesinde rol oynayabilir.



Şekil 1. Değişik dönüşümler sonucu imgeden değişik özniteliklerin elde edilmesi [2].

Yazının ikinci bölümünde insanlarda yüz tanıma üzerine geliştirilmiş teorileri ele alacağız. Bu teorilerin ortak noktası beynin bir bilgisayar gibi veriyi aşama aşama işleyen bir organ olduğunu varsaymalarıdır. Üçüncü bölümde gelişim psikolojisi ve nöropsikolojinin bulgularını bu teorilerin ışığında değerlendireceğiz. Özellikle cevap aradığımız sorular yüz tanıma yetisinin ne kadarının doğuştan geldiği ve ne kadarının öğrenildiği, beyinde yüz tanımaya özgü

bir bölge olup olmadığıdır. Dördüncü bölümde yüz tanımanın holistik (bütünsel) olup olmadığını tartışacağız; bu ayrım özellikle bilgisayar modelleri açısından önem taşır. Beşinci bölümde günümüzde yüz tanıma deyince akla gelen belli başlı bilgisayar modellerine geçeceğiz ve bilişsel bilimden esinlenmiş modellerin hangi süreçleri ne kadar başarıyla kullandığına bakacağız. Son bölümde genel bir değerlendirme yapmaya çalışarak başarılı bir yüz tanıma sisteminin sahip olması gereken özelliklerin bir listesini çıkartacağız.

## 2. İNSANLARDA YÜZ TANIMA

Yüz tanıma, bilgisayarlıları bugüne kadar en çok uğraştırmış örüntü tanıma problemidir, ama bu problemin ne kadar zor olduğunu ilk psikologlar fark etmiş ve insan beyninde bu problemin nasıl çözüldüğü pek çok araştırmaya konu olmuştur. Bir iki kere gördüğümüz bir insan yüzünü aradan uzun bir zaman geçtikten sonra, farklı bir ışık altında, değişik bir pozda görsek bile tekrar tanıyabiliyoruz. Henüz bilgisayar sistemleri bu başarıyı gösteremiyorlar.

### Yüz tanıma problemi

Aynı insan yüzünün farklı pozlardan ve değişik ışıklandırma koşullarında çekilmiş, farklı yüz ifadelerine sahip fotoğraflarını yan yana koyalım. İnsanlar bu yüzlerin aynı kişiye ait olduğunu zorluk çekmeden söyleyebiliyorlar. Oysa bir bilgisayar için bu çok büyük bir problem, çünkü imgeleri üst üste koyduğunuzda hiçbir piksel diğerini tutmuyor. Hatta iki kişinin aynı ışık altında, aynı açıdan çekilmiş fotoğrafları birbirlerine tek bir kişinin farklı ışıklandırmayla başka açılardan çekilmiş fotoğraflarından daha çok benziyorlar. Bu yüzden öznelikleri istatistiksel olarak modelleyen bir öğrenme modeli doğrudan yüz imgeleri üzerinde çalışırsa başarısız oluyor.

Yapay öğrenme modelleri belli şekillerde genelleme yapmaya eğilimlidir. Bu eğilim, yahut yanlılık (bias) modelin nasıl kurulduğuna göre farklılık gösterir. Beynin doğuştan gelen yüz tanıma eğilimi öğrenmeyi kolaylaştıran bir yanlılıktır. Eğer beyni öğrenen bir makina olarak düşünürsek ve amacımız yüz tanıma problemini beyin kadar iyi çözmek ise, girdinin doğasını ve beynin bu probleme ne kadar ve ne şekilde yanlı yaklaştığını öğrenmemiz gerekiyor. Acaba genel bir öğrenme sistemi bu iş için yeterli midir? Yoksa bu problemi ancak doğuştan sahip olduğumuz ve yüz tanıma için özelleşmiş bir altsistem sayesinde mi çözebiliriz? Yüz imgesi beyinde hangi aşamalardan geçerek işleniyor? Bu probleme ilişkin ne kadar ön bilgiyi öğrenme modelimize koymamız gerekiyor? Bu soruların her biri cevap bekleyen araştırma konularıdır.

## Yüz tanıma, nesne tanımadan farklı mıdır?

Psikologlar insanlarda yüz tanımanın diğer nesnelere tanınmasından daha farklı bir biçimde gerçekleştiğini düşünmek için pek çok sebep olduğunu söylüyorlar. Örneğin beyindeki bazı nöronlar sadece görüş alanında bir yüz belirlediğinde harekete geçiyorlar. Birçok deney beyinde sadece yüz tanıma sırasında aktive olan bölgeler olduğunu ortaya koyuyor. Bazı beyin hasarlarından sonra hasta bildiği yüzleri tanıyamaz hale geliyor, oysa diğer nesnelere tanımakta güçlük çekmiyor (prosopagnosia). Bunu tam tersi bir durum da var, nesnelere tanıyamaz hale gelen hastalar bildikleri yüzleri rahatça tanıyorlar (agnosia). Fonksiyonlardaki bu ikili ayrılma (double dissociation) yüz tanımanın özel bir sistem tarafından gerçekleştirildiğinin en büyük göstergesi olarak görülüyor.

Benzer bir bulgu beyin hasarı sonucunda sadece ve sadece yeni yüzleri öğrenme becerisinin kaybedilebileceğini gösteriyor (prosopamnesia). Capgras sendromunda ise hastalar önceden bildikleri yüzleri tanıyorlar, ama tanıdık birini gördüklerinde duymaları gereken aşinalık hissi kayboluyor. Bunun sonucunda hasta, yakınının aslında gerçekten tanıdığı insan olmadığını, birilerinin onu taklit ettiğini iddia ediyor. Bu hastalarda yapılan araştırmalar sonucunda yüz tanımanın beyinde iki farklı sürecin paralel çalışmasıyla gerçekleştiği hipotezi ortaya atılmıştır. Bu süreçlerden birinin duygusal (affective), diğerinin bilişsel (cognitive) olduğu öngörülmüştür.

Beyin üzerinde yapılan araştırmalar özellikle fusiform girus (fusiform face area) denilen bir bölgede yüz tanıma sırasında belirgin bir aktivite olduğunu gösteriyor. Özellikle Mooney imgeleri ile yapılan deneylerin ilginç sonuçları var. Mooney imgeleri dengelenmiş siyah ve beyaz bölgelerden oluşan yüz resimleridir (bkz. Şekil 2). Bu imgelere bakanlar öncelikle bir yüz algılamayabilirler. Ancak yüz bir kere algılandıktan sonra, bütün yüz hatları yerine oturur ve tekrar bakıldığında yüz bu sefer kolayca bulunur. Mooney yüzlerine bakan kişilerde eğer yüz algılanmışsa fusiform girus bölgesinde aktivite gözleniyor. Bu önemli bir bulgu, çünkü imge değişmiyor, sadece algı değişiyor. Yani beyindeki aktivite girdinin doğasından değil, algının doğasından kaynaklanıyor.

Beyin araştırmalarını yüz tanımanın ayrıcalıklı bir konumu olduğu iddiasında destekleyen davranışbilimsel bulgular da var. Örneğin bir nesneyi ters çevirdiğinizde nesnenin tanınmasını zorlaştırmış olursunuz. Ters çevrilmiş yüzlerde bu etki çok daha belirgin bir biçimde ortaya çıkar. Tanaka ve Farah bu etkiyi şöyle açıklıyorlar: Nesnelere tanımak için onların görsel özelliklerinden yararlanırız. Ama yüz tanımda bu özelliklerin ne şekilde dağıldığı,

hangisinin altta, hangisinin üstte olduğu, yani özneliklerin konfigürasyonu, özneliklerin kendisinden daha önemli bir bilgi [10]. Bu yüzden de çok basit iki nokta ve bir çizgiyi bile yüz olarak algılayabiliriz (örneğin ☺).



Şekil 2. Mooney yüzleri [1].

Biyolojik öğrenmenin önemli süreçlerinden biri uzmanlaşmadır. Beyinde öğrenme aşama aşama gerçekleşir. İleri aşamalarda beyin, problemin doğasına ait bazı kısayollar öğrenir, yeni ve daha etkin gösterimler kullanmaya başlar ve daha zengin kavramsal ayrımlar yapar. Bu şekilde problemi daha hızlı ve daha doğru şekilde çözmeye başlar, ama bazı ayrımları yapma gücünden de feragat eder. Görsel illüzyonların çoğu beyin bu özelliğinden yola çıkar. Konfigürasyonun önem kazanması da uzmanlaşmanın bilinen etkilerinden biridir; satranç ve go ustalarının oyun tahtasına baktıklarında tek tek taşları değil taşların durumlarına ait, daha yüksek seviye konfigürasyonlar gördükleri -örneğin hücum cepheleleri, zayıf bölgeler, vs.- biliniyor. Ayrıca satranç ustaları gerçek satranç oyunlarına ait taş dizilimlerini kolayca ezberleyebilirken, rasgele dizilimlerde acemilerden bir farklılık göstermiyorlar.

Yüz tanımanın da aslında özel bir sistem gerektirmediğini, bütün saydığımız bulguların aslında insanın yüz tanıma uzmanlaşmasının sonucu olduğunu söyleyen çalışmalar var. Gauthier ve Tarr bir deney yaparak bu hipotezi deniyorlar. Denekleri "Greeble" adını verdikleri yapay yaratıkların türlü cinslerini ayırtma konusunda eğitiyorlar [6]. Denekler Greeble uzmanı olduktan sonra bir Greeble gördüklerinde, önceden beyin yüz tanıma bölgesi

olarak düşünülen fusiform girus bölgesinde (fusiform face area) yoğun faaliyet gözleniyor. Dahası Greeble'larda konfigürasyon değişiklikleri yapıldığında uzmanlar bundan daha çok etkileniyorlar. Bu da yüz tanıma ve diğer uzmanlıklarda görülen ortak bir özellik.

Köpekleri hiç tanımayan biri, bir labrador gördüğünde onu bir "köpek" olarak algılar. Köpeklere aşina biri ise "labrador" olarak algılayacaktır. İnsanlar da bir yüz gördüklerinde, onu herhangi bir yüz olarak algılamazlar, belli bir yüz olarak, "A'nın yüzü" veya "B'nin yüzü" olarak algırlar. Bu da bir uzmanlık göstergesidir.

## Nesne tanıma modelleri

İnsan beyninin nesnelere nasıl tanıdığı uzun zamandır merak edilen bir konudur. Bu teorileri kısaca özetleyerek yüz tanıma açısından ne kadar geçerli olduklarına bakalım.

Seksenlerde önem kazanan yapısal tanımlama (structural description) teorisine göre beyin cisimleri parça ve bütün ilişkisi içinde algılar. Bu yaklaşımda bütün nesnelere bazı temel algı ünitelerinin birleşmesiyle ifade edilirler. Irving Biederman'ın "geon" adını verdiği bir dizi üç boyutlu şekil, David Marr'ın "codon" dediği basit öznelikler bu temel algı ünitelerine örnek olarak gösterilebilir. Zamanla bu teori işlerliğini kaybetmiştir.

Bu analitik yaklaşıma karşı holistik, veya bütünsel diyebileceğimiz bir yaklaşım kaynağını daha eskilerden, Gestalt psikolojisinden alır. Bu yaklaşıma göre nesnelere parçalarına ayrılmaz, bir bütün olarak algılanırlar. Böyle bir algının mümkün olabilmesi için beyin gibi paralel çalışabilen bir sistem gereklidir.

Bir diğer teoriye göre beyin, nesnelere değişik açılardan, değişik büyüklüklerde ve farklı ışıklandırma şartları altında görür, bu görüntüleri fotoğraflarını çeker gibi ezberler ve sonradan kullanmak üzere saklar. Bu yaklaşıma göre nesne tanıma bilgisayarla örüntü tanıma çok kullanılan "en yakın komşu" metoduna benzer bir şekilde yapılır. Bu yaklaşıma şablon modeli diyebiliriz.

Şablon modeliyle tutarlı bir diğer önemli hipotez de nesne tanımanın duruş-temelli (view-based) olduğu hipotezidir. Bir üç boyutlu nesneye farklı açılardan baktığımızda farklı görüntülerle karşılaşırız. Bu görüntülerin her birine "duruş" diyoruz. Nesnelere günlük hayatta karşılaştığımız hangi durumlarda çıktığına bağlı olarak bazı duruşlar o nesne için daha belirleyicidir. Mesela çoğu insan "bir at düşünün" denildiğinde atı 3/4-duruşta, yani sağ üstten sol alta doğru düşünür. Bu tip doğal duruşlar (canonic view) az görülen duruşlardan daha çabuk tanınır ve daha çok ayırdedici bilgi içerir. Yapılan deneylerde nesnelere

özgün duruşlarını tanımaya çalışan deneklerin de başarılı olduğu, ama tanınmanın daha yavaş gerçekleştiği gözlenmiştir. Yani beyin bu tür bir genellemeyi yapmak için bir dizi ara işleme gereksinim duyar.

Yüz tanıma açısından bakacak olursak, şablon teorisi her tanıdığımız yüz için beynimizde çeşit çeşit şablonlar olduğunu söyler. Yeni bir yüz gördüğümüzde, varolan şablonlardan hangisine daha yakınsa o kişiye ait olduğu sonucuna varırız. Fakat olası şablonların sayısı çok yüksek olduğu için beynin bu gösterimi tercih ettiğini söylemek çok güç.

Pek çok deney hafif yan dönmüş yüzlerin tam karşıdan bakılan yüzlerden daha rahat tanındığını gösteriyor. Bu da yüz tanınmanın duruş-temelli olduğu savını destekliyor. İki bilinen duruş arasındaki duruşların tanınabilmesi için beynin ne yaptığı tam olarak bilinmiyor, ama bu iş için bazı süreçlerin olduğu da şüphesiz.

Yapısal tanımlama teorisi her ne kadar popülerliğini yitirmiş olsa da, yapısal ilişkilerin beyinde ifade edildiğini söylenebilir. Çok basit şematik yüzleri bile yüz olarak tanıyabildiğimize göre beyinde basit şablonlardan daha fazlası olmalı. Yüz tanımda duruşların önemli olması yüz tanınmanın tamamen şablonlara dayandığı anlamına da gelmiyor. Şu anki veriler yetişkinlerde yüz tanınmanın ağırlıklı olarak bütünsel işleme dayandığını gösteriyor.

### 3. GELİŞİM PSİKOLOJİSİ

Dilbilim camiasında dilin öğrenilmesine dair Chomsky ve karşıtları arasında yıllardır süren bir tartışma var. Tartışmanın özünde dil yetisinin ne kadarının doğuştan var olduğu yatıyor. Yüz tanıma konusunda da böyle bir tartışma var; argümanlar neredeyse aynı, ama olay henüz topyekün bir savaşa dönüşmedi. Bir tarafta "Bu kadarcık yüz görerek yüzleri böylesine iyi tanımak mümkün değildir, yüz tanıma doğuştan var olan, kendine has bir sistem tarafından gerçekleştirilir" diyen bir grup var. Bu gruba göre yüz tanıma Jerry Fodor'un zihnin modülleri olarak adlandırdığı izole fonksiyonlardan biri [5]. Diğer tarafta da "Hayır, beyin düşündüğünüzden daha güçlü bir öğrenme organıdır, azıcık evrimsel yardımla genel görme sistemi yüz tanımayı da başarır" diyen bir grup var. Bu seçenekleri değerlendirebilmek için soruna gelişim psikolojisi açısından bakmak gerekli.

Bilişsel yetiler (ya kısmen veya tamamen) doğuştan gelir, ya da bir öğrenme süreci ile geliştirilirler. Öğrenme çeşitli aşamalar ve gösterimler gerektirebilir. Bazı yetilerin öğrenilmesi diğerlerine bağlıdır. Gelişim psikolojisi açısından sorduğumuz sorular yüz tanınmanın ne kadarının doğuştan geldiği, ne kadarının

tecrübeyle öğrenildiği, bu öğrenmenin nasıl bir yol izlediği, dikkat, duygu ve iletişim gibi diğer fonksiyonlara ne kadar bağlı olduğu, ve hangi gösterimlerin kullanıldığıdır. Gelişim perspektifinden bakılınca yüz tanıma aşama aşama öğrenilen, her aşamada bir önceki aşamada kullanılan gösterimlere yeni fonksiyonların ve gösterimlerin eklendiği karmaşık bir sistem gerektiren bir problemdir [7].

Burada bir parantez açıp, beynin işleyişine dair bir noktaya dikkat çekelim. Basit bir öğrenme sistemi karmaşık bir problemi çözemez. Fakat karmaşık (çok parametresi olan) bir öğrenme sistemi için de öğrenme kolay bir iş değildir. Karmaşık problemleri çözmek için beynin stratejisi öncelikle daha basit bir problemi çözmek, ondan sonra da bu çözümü daha karışık olan problemde kullanmaktır. Problemi basite indirgemek için iki temel mekanizma vardır. Birinci mekanizma kısa süreli hafızadır. Tıpkı bir bidona su doldururken kullanılan huni gibi girdinin az ve düzenli olmasını sağlar. İkinci mekanizma da seçici dikkattir. Girdinin işlenebilecek kısmını öğrenme sistemine geçirir, karmaşık kısımları atar. Hafızanın gelişmesi ve seçici dikkatin seçtiği özniteliklerin yavaş yavaş değişmesiyle karmaşık algılama sistemlerinin aşamalı eğitimi gerçekleşir.

Seçilen özniteliklerdeki değişime bir örnek verelim. Yeni doğmuş bir çocuk annesinin yüzünü çok erken, henüz birkaç günlükken tanımaya başlar. Ama annesi saçını farklı bir renge boyarsa tanıyamaz. Yani bebek saç rengi ve şeklini yüz tanımda kullanır. Oysa saç sürekli değişen ve -yetişkinler için- güvenilmez bir özniteliktir, o yüzden büyüdükçe saçın tanımadaki önemi azalır. Yapılan araştırmalar iki-üç yaşındaki çocukların yüzlerde tek tek özniteliklere baktığını, bunlara sırayla dikkat ettiklerini saptamıştır. Çocuk beş yaşına geldiğinde bazı özniteliklerin öne çıktığı ve tutarlı bir biçimde tanımda kullanıldığı görülür. Bu yaştan sonra bütünsel algı gittikçe gelişir, analitik algıda gerileme olur. Yetişkinlerde bütünsel algı tamamen ön plandadır, sadece bu bütünsellik bozulduğunda yetişkin özniteliklere dikkat ederek yargıya varmaya çalışır. Bu şekilde yüz tanıma sistemi önce basit, yavaş ve özniteliklere dayalıyken, gittikçe daha hızlı, karmaşık ve bütünsel bir sistem haline gelir. Bütünselliğin bozulduğu durumlarda (mesela ağız, burun, ve gözlerin yerleri değiştirilince) bile yüz tanınmanın mümkün olması, ilk gelişen öznitelik tabanlı altsistemlerin kaybolmadığını, karmaşık sistemin birer parçası olarak hayatlarına devam ettiklerini gösteriyor.

Beynin aşamalı öğrenme sistemi yapay sinir ağlarında da denenmiş, yüz tanıma ve dilin modellenmesinde başarıyla kullanılmıştır [4]. Bu yaklaşımda sistemin önce küçük bir hafıza birimi kullanması sağlanır. Sistem bu şekilde sadece basit bağlantıları öğrenebilir. Daha sonra hafıza kapasitesi yavaş yavaş artırılır.

Sistem önceden öğrendiği bağlantıları unutmaz, bunlara ilaveten daha karışık bağlantılar da öğrenir. Aynı sistemi en baştan, bütün hafıza kapasitesini serbest bırakarak eğittiğimizde sistemin başarılı olmadığını görürüz.

Bilgisayar modellerine genel olarak baktığımızda, gösterimlerin bu şekilde değiştiği aşamalı eğitim süreçlerinin nadiren kullanıldığını görürüz. Bunun yerine girdinin çeşitli ön işleme aşamalarından geçirildiği ve problemi zorlaştıran bazı özelliklerin ortadan kaldırılmasının amaçlandığı sistemler tercih edilir. Örneğin yüz tanıma ışıklandırmanın etkisini ortadan kaldırmak için histogram eşitlemesi kullanılır. Bu yöntem aslında genel ışıklandırma probleminin sadece bazı varsayımlar altında, bir dereceye kadar çözülmesini sağlar; zira çoğu bilgisayar modeli, sınırları dikkatle çizilmiş bir alanda yüksek başarıyla çalışmak üzere tasarlanır.

Gösterimdeki aşamalı değişimin uzmanlaşmanın sonucu olduğunu belirtmiştik. Bu değişim rasgele değildir, yeni gösterimler uzmanın işini kolaylaştıracak niteliktedir. Yani gösterimin ne olacağını, problemin doğası belirler. Her insan bir yüz tanıma uzmanı olduğuna göre, yüzlerin insan beyninde nasıl gösterildiği yüz tanımanın ne için yapıldığına bağlıdır diyebiliriz. İnsanlar yüzlere göz kontağı kurma, iletişim sağlama, duygu tanıma, konuşma takibi için dudak okuma gibi amaçlarla bakarlar. O halde beyin uzmanlaşırken bu işleri kolaylaştıracak gösterimleri tercih edecektir. Seçici dikkat çalışmaları insanların yeni yüzlerde gerçekten de göz ve dudaklara özellikle baktığını gösteriyor.

Yeni doğmuş bebekler yüzlere ilgi duyarlar. Bu tercih genetik olarak kodlanmıştır. Bu durumun doğuştan geldiğini belirlemek için yapılan bir deneyde bebeklere üç değişik resim gösteriliyor. Birinci resimde normal bir insan yüzü var. İkinci resimdeki yüzün ağız, burun ve gözlerinin yerleri değiştirilmiş. Üçüncü resimdeyse sadece dış hatları belirgin olan içi boş bir yüz imgesi var. Bebekler doğduktan on dakika sonra birinci tip resimleri diğerlerine tercih ediyorlar.

Aslında birçok hayvan kendi cinsine karşı benzer tercihlerle doğar. Yavruyu annesine bağlayacak doğuştan gelen bir mekanizma yavrunun hayatta kalma şansını artıracığı için evrimsel değeri çok yüksektir, o yüzden doğal seçim böyle bir mekanizmayı destekleyecektir. Yüzlere duyulan erken ilgi, bebeğin diğer bireylerle iletişiminin ve dilin gelişiminin de başlangıç noktası olarak görülebilir (bkz. Şekil 3). Yüz tanımanın engellenmesi halinde, örneğin otistik bireylerde yüzlere dikkat edememe ve göz teması kuramama sonucu iletişim son derece düşük bir seviyede kalır, dil yetisi gelişmez.



**Şekil 3 Bebek üç haftalıkken yetişkinin yüz hareketlerini taklit edebilir. Bu beceriyle gösterdiği algı ve kontrol, bebeğin iletişim kurmasına ve öğrenmesine temel oluşturacak, doğuştan gelen bir sisteme işaret ediyor [8].**

Beynin yüzlere erken ilgi duyma özelliğine bakarak yüz tanıma doğuştan gelmektedir (innate) diyebilir miyiz? Bu soruya verilen en güçlü yanıt Jeff Elman, Anette Karmiloff-Smith ve UCSD'deki araştırmacılarından geliyor: Hayır, yüz tanıma konusunda beyin genetik temelli basit bir eğilimi olsa da, asıl öğrenme doğumdan sonra, daha genel bir sistem tarafından gerçekleştirilir [7], [4]. Çocuk uzun süre çeşitli açılardan insan yüzlerine bakar, ve genel amaçlı sinir ağlarını bu girdinin istatistiksel özelliklerini öğrenmek için kullanır. Öncelikle öznitelikler öğrenilir, daha sonra çocuk yüz tanıma uzmanlaşmaya başlar. Çocuğun görme sistemi doğumu takiben bir süre daha yetişkinlerdeki hassasiyete ulaşamadığı için, öncelikle (saçlar, ten rengi gibi) çok belirgin öznitelikler öğrenilir. Uzmanlaşmayla birlikte bütünsel algı gelişir. Uzmanlaşma genel yeteneklerin kısmi kaybını da beraberinde getirir. Altı aylık çocuklar maymun yüzlerini birbirinden ayırdedebilirken, dokuz aylık çocuklar bunu başaramıyorlar, çünkü insana özgü özniteliklerdeki ayrımlara dikkat etmeyi, diğer ayrımları dikkate almamayı öğrenmişler. Yetişkinlere baktığımızda, başka bir ırkın yüzlerini ayırmakta zorlandıklarını görüyoruz (other race effect). Bu tip bir uzmanlaşma başka alanlarda da görülür, mesela Japonlar "l" ve "r" seslerini ayırtma yetisini zamanla kaybederler, çünkü Japonca'da fonetik olarak bu sesler bir ayrıma karşılık gelmez. Halbuki bütün sağlıklı bebekler bu ayrımı yapabilir halde doğar.

#### **4. HOLİSTİK MODELLER, ÖZNETELİK TABANLI MODELLERE KARŞI**

Yüz imgelerini ters çevirmenin yüz tanımayı diğer nesnelere çok daha fazla etkilediğini söylemiştik. Bu etkinin sebebinin özniteliklerin ters çevirmeden fazla etkilenmemesi, buna karşın bütünsel bilginin

bozulması olduğunu düşünüyoruz. Bu düşünceyi destekleyen başka bulgular da var. Örneğin yüze ait öznitelikler başka özniteliklerin yanında olduğu zaman tek başına olduklarından daha kolay tanınıyor. Mesela A'nın burnunu tek başına görünce tanıyamayan denekler, B'nin ağzı ve gözlerinin yanında tanıyabiliyor. Yani oluşan yüz konfigürasyonu tek tek özniteliklerin tanınmasını da kolaylaştırıyor. Buna yüzün üstünlüğü etkisi (face superiority) deniyor. Başka bir çalışma ise tek tek özniteliklerin ters çevrilmesinin yüz tanımayı fazla etkilemediğini gösteriyor.

Bu çalışmalara dayanarak bütünsel (holistic) ve öznitelik-tabanlı (feature-based) yüz tanıma modelleri arasında bir ayırım yapıyoruz. Bu ayırımın bir ucunda yüz tanımanın tamamen bütünsel olduğunu, yüzlerin tek bir parça olarak algılandığını söyleyen araştırmacılar var [10]. Yüz tanımanın beynin daha çok sağ tarafında gerçekleşmesi, yüz tanımanın bütünsel olduğu varsayımıyla uyumlu görülüyor, zira beynin sağ tarafı daha çok bağlantıları kodluyor. Buna karşılık bir grup araştırmacı bu bulguların başka şekilde açıklanabileceğini savunuyor.

1980'lerde Thompson Margaret Thatcher'ın yüzündeki bazı öznitelikleri ters çevirerek ilginç bir deney yapmıştır [11]. Bu şekilde işlenen yüzler (Thatcher yüzleri) normalde garip görünüyorlar, ama ters çevrildiklerinde insanlar bunu fark etmiyor (bkz. Şekil 4). Thatcher ilüzyonu bize özniteliklerin ayrı ayrı işlenip sonradan birleştirildiğini düşündürüyor. Ters çevirme sırasında bütünsel bilgi bozulduğu için tek tek özniteliklerden gelen bilgi kullanılıyor. Bu bilgi de ters çevirmeden nispeten az etkilendiği için yüzdeki gariplik fark edilmiyor.



**Şekil 4 Thatcher ilüzyonu. Bu resimlere bir de sayfayı ters çevirip bakın [11].**

Bu noktada durup insan gözünün çalışma şeklini hatırlamakta fayda var. İnsanda görme bir anda gerçekleşmez. Göz bütün algı alanını bir anda bütün detayıyla yakalayamaz, çünkü gözün detaylı bilgi alabilen bölümü fovea dediğimiz, göz sinirlerinin çok büyük bir bölümünün toplandığı küçük bir pencereden ibarettir. Bir yere baktığımızda foveadan detaylı bilgi gelir, onun dışındaki bölgelerden ise bulanık diyebileceğimiz çözünürlükte bilgi alırız. Bu yüzden

gözün fovea bölgesi seçici dikkat mekanizması yardımıyla algı alanında gezinir, çeşitli yerlerden aldığı detaylı bilgi birleştirilir. Beyin gördüğü kısımları hemen işlemeye başlar ve bu işlem foveanın bir sonraki adımında nereye yöneleceğini de belirler.

Araştırmacılar ters çevrilmiş Thatcher yüzlerinin ters çevrilmiş yüzlerden daha hızlı tanındığını bulmuşlar. Bunun sebebi insanlarda göz ve ağız bölgelerinin iletişimdeki önemleri sebebiyle seçici dikkat mekanizmasının ilk hedefleri arasında yer alması ve ters Thatcher yüzlerinde bu bölgeler düz oldukları için buradan daha sağlıklı bilgi gelmesidir. Bu bulgulara bakarak öznitelik-tabanlı tanımanın bütünsel tanımadan daha hızlı gerçekleştiği sonucuna varabiliriz. Eğer foveanın yönlendiği ilk birkaç bölgede tanımaya yetecek bilgi varsa, o zaman tanıma çok hızlı bir biçimde gerçekleşir. Yüz üstünlüğü etkisi de bu çerçeveden kolayca açıklanır; bütünsel bilgi foveanın sağlıklı bir şekilde göz ve ağız bölgelerine yönlendirilmesini sağlar.

Karikatürler üzerine yapılan çalışmalar da yüz tanıma konusunda ilginç açılımlar getiriyorlar. Karikatür sanatçıları bir yüzü karikatürize etmek için yüzün ortalamadan farklı özelliklerini abartır, önemsiz detaylarıysa ortadan kaldırır. Örneğin burun normalden biraz büyükse, karikatürde dev gibi çizilir. Bu şekilde yaratılan karikatürler çok az öznitelik içerirler bile tanıma için yeterlidirler. Bilinen yüzlerin çizimlerden tanınmasında abartılmış yüzlerin gerçeğine sadık yüzlerden daha iyi sonuç verdiği bulunmuştur. Karikatür etkisi sadece çizimlerde değil, fotoğraflarda da elde edilebilir. Ama çizimlerde gölgelerin olmaması bu etkinin daha belirgin olmasını sağlar.

Bu bulgular yüz tanıma sisteminde hem öznitelik-tabanlı, hem de bütünsel veri işleme olduğunu ortaya koyuyor. Karikatürleştirme özniteliklerin daha iyi tanınmasını sağlıyor. Bu iyileştirme bütünsel bilginin zayıf olduğu durumlarda çok daha belirgin hale geliyor.

## 5. BİLGİSAYARDA YÜZ TANIMA

Bilgisayarda yüz tanıma dediğinde anladığımız şey bir bilgisayar programının bir yüz imgesini girdi olarak alması ve bu yüzün kime ait olduğunu belirlemesidir. Biyometrik uygulamalarda bazen yüz ile birlikte kimlik bilgisi de verilir, o zaman bilgisayarın verilen yüzün gerçekten de o kişiye ait olup olmadığını belirlemesi istenir. Bu biraz daha basit bir problemdir.

Genel bir yüz tanıma sistemi üç parçadan oluşur:

1) *Veritabanı*: Sistemin tanıyacağı yüzlerle ilgili bilgileri saklar. Bu bilgi yüzlerin imgeleri de olabilir, bu imgelerden çıkartılmış bazı öznitelikler de olabilir.

Veritabanının büyüklüğü için bir üst sınır yoktur, milyonlarca yüz içerebilir, fakat literatüre bakıldığında bu sayının 50 ile 5.000 arasında değiştiği, genelde 1.000 civarında olduğu görülür.

2) *Öğrenme sistemi*: Yüzlerin tanınması için bir yapay öğrenme sistemi eğitilir. Veritabanındaki yüzler bu sistemin *eğitim kümesini* oluşturur. Parametrik bir sistem kullanılıyorsa, eğitim bu parametrelerin değerlerinin bulunması ile olur.

3) *Önişleme sistemi*: Genellikle zor problemlerde bir önişleme aşaması ile girdinin daha kolay tanınması sağlanır. Yüz tanımadaki önişleme sistemi ışıklandırmanın ayarı, poz normalizasyonu, imgenin standart bir büyüklüğe getirilmesi, yüzün ortalanması, arka fonun temizlenmesi, tanımayı zorlaştıracak saç, sakal ve gözlüklerin ortadan kaldırılması gibi aşamalar içerebilir. Önişleme de aslında bir tür öğrenme gerektirir ve önişleme ne kadar iyiyse, öğrenme sisteminin işi o kadar kolaylaşır.

Sistem eğitildikten sonra yeni bir yüz imgesi verildiğinde önce önişleme sistemi devreye girer, imgeden bir dizi öznitelik çıkartır. Bir başka deyişle, imgeyi öznitelik uzayına taşır. Bu dönüşümden çıkan öznitelikler öğrenme sistemine geçer. Öğrenme sistemi de veritabanındaki yüzlerden hangisinin (ya da hangilerinin) bu yüze en yakın olduğunu bulur.

Chellappa, Wilson ve Sirohey yüz tanıma ile ilgili çalışmaları özetledikleri makalede araştırmacıları uyarıyorlar:

*"Yüz tanıma algoritmaları ve sistemleri tasarlayanlar psikofiziksel ve nörofizyolojik bulgulardan haberdar olmalı, ama sadece pratik anlamda işe yarayacak olanları modellerinde kullanmalıdırlar."* [3]

Bu uyarının anlamlı olması için, bilgisayarda yüz tanıma araştırmalarını üç gruba ayırmamız gerekir. Birinci grup yüz tanımayı bir mühendislik problemi olarak ele alır ve amacı bu problemi mümkün olan en az kaynakla, gerektiği kadar iyi çözmektir. Yukarıdaki uyarı daha çok bu grubu ilgilendirir. İkinci grup ağırlıklı olarak psikoloji ve bilişsel bilim tarafında, bilgisayar modellerini beyinde yüz tanıma ile ilgili geliştirilen hipotezleri sınamakta kullanan gruptur. Bu grubun amacı problemi daha iyi bir şekilde çözmek değil, beyindeki süreçlerin modellerini yapmaktır. Üçüncü bir grup da bu ikisinin arasında, henüz bilişsel bilimin pratik anlamda işe yarayıp yaramadığı sınamamış bulgularını modelleyerek mühendislere yeni araçlar sunmayı hedefler. Aslında amacı daha iyi yüz tanımak olan, ama modellerinde biyolojik sistemlere göndermeler yapan ve beynin işleyişiyle tutarlı olmaya çalışan araştırmacıları da bu gruba koyabiliriz.

Biyolojik yüz tanıma sistemleriyle tutarlı olma iddiası taşıyan ilk yüz tanıma modeli Turk ve Pentland'ın özyüz (eigenface) modelidir [12]. Yayınlandığı tarihten bu yana iki binin üzerinde atf alan bu

çalışmada yüz tanıma için anabileşenler analizi (principal components analysis, PCA) önerilir. PCA yönteminde eğitim kümesindeki imgelerin varyansının en yüksek olduğu boyutlar seçilir ve imgelerin bu boyutlara izdüşümü alınır. Bu boyutların her birine özyüz adı verilir. İzdüşüm alma, yüzün bu özyüzlerin ağırlıklı toplamı olarak ifade edilmesi anlamına gelir. Tanıma bu dönüşümün ardından veritabanındaki en yakın şablonu bularak yapılır. Özyüz metodu bütünsel yaklaşıma örnektir. İmgelerin büyüklüklerine, pozlarına ve konumlarına karşı hassas olduğundan, her yüz için değişik pozlarda şablonlar saklayan bir çeşitlemesi de geliştirilmiştir.

Bir diğer biyolojik yaklaşım da elastik çizge (elastic graph) modelidir [14]. Bu modelde Gabor dalgacıkları ile yüzlerden öznitelikler çıkartılır. Gabor dalgacıkları beyindeki görme sisteminin ilk aşaması olan V1'deki basit hücrelere benzer bir işlem yaparlar. İmgelerden gelen öznitelikler daha önceden eğitim kümesinde bulunup saklanmış olan özniteliklerle karşılaştırılır ve bir yüz konfigürasyonuna uyan en iyi öznitelik seti bulunur. Bu metod öznitelik-tabanlı yaklaşıma örnektir.

İki modelin başarısının da insanlardaki yüz tanıma sonuçlarıyla benzeştiği gözlenmiştir. Genel olarak elastik çizge modeli öznitelikler arasındaki yapısal ilişkiyi ayrıca modellediği için daha başarılıdır. Bu iki modele temelde benzeyen, fakat farklı öznitelikler kullanan birçok çalışma vardır. Bunların dışında insanlarda yüz tanıma ile ilgili değişik savları denemek amacıyla da modeller geliştirilmiştir. Bu modellere birkaç örnek verelim.

Ramasubramanian ve Venkatesh yüz imgelerine ayrık kosinüs değişimi (discrete cosine transform, DCT) uyguluyorlar [9]. Kullandıkları bütünsel yöntem yüksek bir başarı yüzdesi yakalıyor. Bu modelde de Gabor dalgacıkları gibi insandaki görme sisteminde var olduğu düşünülen dönüşümlere benzer bir dönüşüm var. (İnsanların görme sistemi düşük frekanslardaki bilgiye karşı daha hassastır. Yüz imgelerinde de enerji daha çok düşük frekanslarda bulunur.)

Almanya'da Max Planck Enstitüsü'nün Biyolojik Sibernetik konusunda çalışan Tübingen araştırma merkezinde insanlarda ve bilgisayarlarda yüz tanımayı karşılaştıran çalışmalar yapılıyor. Bunlardan birinde yüzlerden parçalar kesilip karıştırılıyor ve bütünsel bilgi ortadan kaldırılıyor [13]. Deneklerin bu şekilde yüzleri ne kadar başarıyla tanıdığına bakılıyor. Daha sonra parçaları karıştırılmış yüzler tanıma tamamen başarısız olana kadar bulanıklaştırılıyor. Yani hem öznitelik, hem de bütünsel bilgi ortadan kalkmış oluyor. Daha sonra aynı bulanıklaştırma normal yüz imgelerine uygulanıyor. Böylece öznitelik bilgisi siliniyor, sadece bütünsel bilgi kalıyor. Bu şekilde

öznitelik bilgisiyle bütünsel bilginin yüz tanıma ne kadar katkı sağladığını karşılaştırmak mümkün oluyor. Benzer bir tanıma testi öznitelikler ve bunların arasındaki ilişkileri ayrı ayrı modelleyen bir bilgisayar programına veriliyor ve programın tanıma performansının deneklerinkine son derece benzediği görülüyor.

## 6. SONUÇLAR

Bilgisayarda geliştirilen yüz tanıma modelleri biyolojik savları desteklemek için kullanılabilir. Bilgisayar modelinin başarılı olması beynin de önerilen şekilde işlediğini göstermez, sadece kurulan soyut modelin işlerliğini, verinin öngörülen şartlarda problemi çözmek için yeterli olduğunu gösterir. Bilgisayar modeli başarılı olmazsa, o zaman bundan çok net bir sonuç çıkartmak kolay değildir. Belki programcı modeli iyi eğitememiştir, yahut bazı sınırlandırıcı varsayımları dikkate almamıştır. Alınan negatif sonuçlar çoğu zaman bilimsel yayınlara dönüşmezler. Bu yüzden de literatürde birbiriyle çelişen savları destekleyen bilgisayar modelleri bulmak da mümkündür. Fakat deneyler iyi tasarlanmışsa, bilgisayar modelleri önemli sonuçlar üretebilirler.

Biyolojik sistemlerin incelenmesi de daha başarılı yüz tanıma sistemlerinin kurulmasını sağlayabilir. Eğer yukarıda saydığımız bulguları bu açıdan özetlersek:

- 1- Karmaşık bir sistemi ortaya koyup yüz tanıma problemini bir anda çözmeye çalışmak yerine, doğru ara gösterimler seçmek ve sistemi aşamalı olarak eğitmek daha sağlıklı bir yaklaşımdır.
- 2- Bütünsel ve öznitelik tabanlı sistemlerin yüz tanıma da birlikte çalışması daha gürbüz (robust) çözümler üretecektir.
- 3- Birçok problemde olduğu gibi sorunun çözümü büyük ölçüde gösterime dayanmaktadır. Bu açıdan bakıldığında duruş-temelli modeller diğer alternatiflerden daha iyi görünmektedir.

Beyin her zaman en doğru olanı yapmaz, bazen de imkanları sınırlı olduğu için belli bir yöntem izler. Eğer amaç iyi bir model kurmaksa, mutlaka beynin taklit edilmesi gerekmez. Bu yüzden kızılötesi kameralar ve üç-boyutlu algılama cihazları gibi yeni teknolojiler yüz tanıma kullanılmaya başlanmıştır.

## KAYNAKLAR

[1] Andrews, T.J., D. Schluppeck, 'Neural responses to Mooney images reveal a modular representation of faces in human visual cortex,' *NeuroImage*, vol.21, pp.91-98, 2004.

[2] Bruce, V., A. Young, *In the eye of the beholder: The science of face perception*, Oxford University Press, New York, 1998.

[3] Chellappa, R., C.L. Wilson, S. Sirohey, 'Human and machine recognition of faces: a survey,' *Proceedings of the IEEE*, vol.83, no.5, 1995.

[4] Elman, J.L., E.A. Bates, M.H. Johnson, A. Karmiloff-Smith, D. Parisi, K. Plunkett, *Rethinking innateness: A connectionist perspective on development*, Cambridge, MA: MIT Press, 1996.

[5] Fodor J.A., *The Modularity of Mind*, MIT Press, Cambridge, MA, 1983.

[6] Gauthier, I., M.J. Tarr, 'Becoming a 'Greeble expert': Exploring the face recognition mechanism,' *Vision Research* vol.37, pp.1673-1682, 1997.

[7] Karmiloff-Smith, A., *Beyond modularity: A developmental perspective on cognitive science*, Cambridge MA: MIT Press, 1992.

[8] Meltzoff, A.N., N.K. Moore, 'Imitation of facial and manual gestures by human neonates,' *Science*, vol.198, pp.75-78, 1977.

[9] Ramasubramanian, D., Y.V. Venkatesh, 'Encoding and recognition of faces based on the human visual model and DCT,' *Pattern Recognition*, vol.34, pp.2447-2458, 2001.

[10] Tanaka, J.W. and M.J. Farah, 'Parts and wholes in face recognition,' *Quarterly Journal of Experimental Psychology: Human Experimental Psychology* vol.46A, 225-245, 1993.

[11] Thompson, P., 'Margaret Thatcher - A new illusion,' *Perception*, vol.9, pp.483-484, 1980.

[12] Turk, M., A. Pentland, 'Eigenfaces for recognition,' *Journal of Cognitive Neuroscience*, vol.3, no.1, pp.71-86, 1991.

[13] Wallraven, C., A. Schwaninger, H.H. Bülhoff, 'Learning from humans: computational modeling of face recognition,' *Proceedings of ECVW*, 2004.

[14] Wiskott, L., J.-M. Fellous, N. Krüger, C. von der Malsburg, 'Face recognition by elastic bunch graph matching,' *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.19, no.7, pp.775-779, 1997.