

Dijital İmgelerdeki Nesnelerin Konumlarını Bulmak İçin Yeni Bir Yöntem

A Novel Method for Object Localization in Digital Images

Onur Akyüz, Hakan Çevikalp

Elektrik ve Elektronik Mühendisliği Bölümü
Eskişehir Osmangazi Üniversitesi
{onur.accounts,hakan.cevikalp}@gmail.com

Güvenç Usanmaz

Bilgisayar Mühendisliği Bölümü
Orta Doğu Teknik Üniversitesi
gusanmaz@gmail.com

ÖZETÇE

Bu bildiri genel nesne sınıflarına ait görüntü örneklerinin dijital imgelerdeki konumlarının bulunarak, dikdörtgensel çerçeveler ile belirtilmesi problemi ele alınmıştır. Geleneksel nesne konum belirleme yöntemleri, nesne örneklerinin koordinatlarının bulunması problemini nesneye ait örnekleri içeren sınıfı, arka plan sınıfı olarak adlandırılan ve söz konusu nesneye ait görüntüleri içermeyen örneklerden oluşan sınıftan ayırma amacı güden iki-sınıflı bir sınıflandırıcı tasarlama problemi olarak ele alır. Eğitilen sınıflandırıcı imgeler üzerinde kaydırılan farklı boyutlardaki pencerelere uygulanır ve bu sayede nesnelerin imge üzerindeki koordinatları ve ölçeği tespit edilir. Bu çalışmada biz de kayan pencereler yöntemini kullandık. Fakat klasik yöntemlerin aksine biz nesneye ait sınıfı konveks bir model kullanarak yakınsadık ve her bir pencere seçilen konveks modele olan uzaklığına bakılarak nesne yada arka plan olarak belirlendi. Deneysel çalışmalar bu yaklaşımın konum bulmada, doğrusal destek vektör makinelerini de içeren bir ardışıl sınıflandırıcıda kullanılmasının yüksek sınıflandırma başarısından ödün vermeden hızlı önemli ölçüde artırdığını göstermiştir.

ABSTRACT

Here we consider generic object localization in digital images where the goal is to find a tight bounding box enclosing the instances of object of interest. Traditional object localization methods treat this problem as building a binary classification that distinguishes between the object class and the background. The trained classifier is usually turned into a detector by sliding it across the image at different scales and classifying each window. In this study we also use the sliding window approach, but as opposed to the traditional methods, we approximate object class by using a convex class model, and each window is assigned to the object class or background based on the distance to this convex model. Our experiments demonstrate that using such models in a cascade for object localization with linear Support Vector Machines significantly improves the real-time efficiency with maintaining high classification accuracies.

1. GİRİŞ

Günümüzde bilgisayarlı görü uygulamaları gittikçe önem kazanmaktadır ve dijital imgelerdeki nesne görüntülerinin konumlarının bulunması bu uygulamaların içinde önemli bir yere sahiptir. Dijital imgelerdeki nesnelerin konumlarının bulunması güvenlik, robotik, askeri ve ticari uygulamalarda geniş kullanım alanları bulmaktadır fakat bu oldukça zor bir iştir. Bunun en önemli nedeni aynı sınıfa ait veri örneklerinin

görünüş, renk ve doku bakımından büyük farklılıklar gösterebilmesidir. Ayrıca doğal günlük yaşamdan manzaralar içeren dijital imgelerde arka plan son derece karmaşıktır ve bu durum nesnelere arka plandan ayırt etmede büyük zorlukları beraberinde getirir. Aydınlatma, poz ve ölçek farklılıkları ve nesnelerin görüntülerinin başka objeler tarafından kapatılması (occlusion) durumları da problemi zorlaştıran önemli faktörler arasındadır.

Nesne konum bulma sistemlerinin başarısını etkileyen başlıca iki önemli faktör vardır: Örnekleri betimlemek için kullanılan öznitelikler ve konum bulma işlemini gerçekleştiren öğrenme algoritması. Betimleme ile ilgili olarak, iyi bir nesne konum belirleme algoritmasının karmaşık sahneleri içeren imgelerde sınıf-ıçi değişimlerden az etkilenen ve sınıflar arasındaki farklılıkları öne çıkaracak betimleme teknikleri kullanması gerekmektedir. Bu şartları sağlayan imge betimlerini kullanan sınıflandırıcıların işleri daha da kolaylaşacaktır. Gerek nesne sınıflandırmada gerekse nesne konum bulma yöntemlerinde Histograms of Oriented Gradients [2], Scale Invariant Feature Transform (SIFT), dalgacık dönüşüm öznitelikleri, Edge Orientation Histogram, Gabor süzgeçler ve Local Binary Patterns [1] gibi bölgesel betimleme teknikleri başarıları ile ön plana çıkmaktadır. Bölgesel betimleme teknikleri ile elde edilen öznitelikler, nesnelerin karakteristik özelliklerini oluşturan bölgesel biçimleri iyi bir şekilde modelleyebilmekte ve ayrıca belirli bir dereceye kadar dönme ve öteleme etkilerini bastırabilmektedirler.

Birçok nesne konum bulma sistemi, imgelere ait özniteliklerin belirlenmesinden sonra, probleme nesnelere ait örnekleri içeren veri sınıfını, söz konusu nesne haricinde imgelerden oluşan arka plan veri sınıfından ayırt etmeyi amaçlayan iki-sınıflı bir sınıflandırıcı tasarlama problemi olarak yaklaşır. Bu amaçla kullanılan sınıflandırıcılara örnek olarak Destek Vektör Makineleri (Support Vector Machines), Yapay Sinir Ağları, AdaBoost, Rassal Sınıflandırıcı Ağaçlar (Randomized Classification Trees), naive Bayes gibi sınıflandırıcılar sayılabilir. Test imgelerindeki nesnelerin konumlarını belirleme sırasında, eğitilen sınıflandırıcılar imge üzerinden sistematik olarak seçilen farklı boyutlardaki pencerelere uygulanır ve sınıflandırıcının verdiği çıktıya göre nesnelerin imgede olabileceği yerler tespit edilmeye çalışılır. İmge üzerinde sınıflandırıcının uygulanacağı bölgesel pencereler seçilirken genellikle iki yol izlenir: İlkinde imge üzerinde değişik boyutlarda pencereler kaydırılır (kayan pencereler yöntemi) ve bu şekilde tüm imge taranır. İkincisinde ise, ilk önce imge üzerinde belirgin noktaları bulmak için Harris Laplace ve Difference of Gaussians gibi imgelerdeki belirgin noktaları bulan detektör algoritmaları uygulanır ve sınıflandırıcılar bu algoritmaların bulduğu noktalar etrafından seçilen pencerelere uygulanır. Kayan

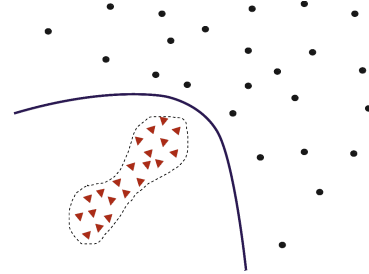
pencereler yöntemi detektör algoritmalarının kullanıldığı yöntemlere göre çok daha iyi sonuçlar verdiğinden, son zamanlarda yüksek başarımları ile ön plana çıkan tüm nesne konum bulma sistemleri kayan pencereleri kullanmaktadır. Kayan pencerelerde onbinlerce pencerenin sınıflandırılması gerektiğinden bu işlem, sınıflandırıcıların ardışıl (cascade structure) olarak kullanılması gerekliliğini ortaya çıkarmıştır. Bu sebeple ardışıl olarak kullanılan sınıflandırıcıların ilk aşamalarında daha çok fazla hesap yükü gerektirmeyen hızlı sınıflandırıcılar seçilir. Bu amaçla seçilen sınıflandırıcıların amacı nesne örneklerine ait olan tüm pencereleri geçirerek, arka plana ait görüntülerin birçoğunu elemektir. Bu sınıflandırıcıları geçen pencereler daha sonraki aşamada sınıflandırma başarımları çok daha yüksek, fakat daha fazla hesap yükü gerektiren sınıflandırıcılar tarafından sınıflandırılarak nesne sınıfına ait olup olmadığı kesin olarak belirlenir. Örnek olarak, nesne konum bulmada en iyi yöntemler olarak kabul edilen Harzallah ve arkadaşları [4] ve Vedaldi ve arkadaşları [7] tarafından geliştirilen sistemlerde ardışıl sınıflandırıcıların ilk aşamasında hızlı sebebiyle doğrusal destek vektör makineleri tercih edilmiş, ardışıl sınıflandırıcının son aşamasında ise başarımları yüksek fakat oldukça yavaş doğrusal olmayan destek vektör makineleri kullanılmıştır. Fakat son aşamadaki sınıflandırıcı sadece ilk aşamaları geçen az sayıda pencereye uygulandığından sistemin hızı çok kötü etkilenmemektedir.

Bu projede biz, geleneksel yöntemlerin aksine, sınıflandırma problemini iki-sınıflı sınıflandırma yerine tek sınıflı sınıflandırıcı tasarlama problemi olarak ele alacağız. Bu amaçla yukarıda bahsi geçen imge betimleme teknikleri kullanılarak elde edilen nesne örneklerine ait veri vektörlerinin örnek uzayında bulunabileceği bölgelerin kestirimi gereklidir. Bu bölgeler doğru bir şekilde modellenebilirse nesnelerin konumlarının belirlenmesi için daha doğru karar sınırları bulunabilir. Sınıflandırıcı tasarımı sırasında arka plan sınıfını kullanmayarak, sadece nesnelerin modellenmesi suretiyle, arka planın sınırlı sayıdaki imgelerle modellenmeye çalışılmasından kaynaklanan sorunlar büyük ölçüde giderilebilir ve konum bulma sisteminin hızı büyük ölçüde iyileştirilebilir.

2. YÖNTEM

Günlük hayattan kareleri yansıtan dijital imgelerde arka plan oldukça karmaşıktır ve büyük farklılıklar gösterebilir. Sınıflandırıcıların eğitilmesi sırasında arka planı oluşturan örnek sayısı genellikle nesneye ait örnek sayısından biraz daha fazladır ve sınırlı sayıdaki örnek, arka planın modellenmesi için yeterli olmayabilir. Bu sebeple eğitilen sınıflandırıcıların bulunduğu karar sınırları doğru olmayabilir. Bu durum Şekil 1’de gösterilmiştir. Şekilde görüldüğü üzere, söz konusu nesneyi temsil eden örneklerin giriş uzay içinde gömülü olduğu bölgeler arka plandan gelen örneklerin azlığından ve bu örneklerin doğru bölgelerde seçilememesinden dolayı iki-sınıflı sınıflandırıcı yaklaşımıyla doğru olarak bulunamamıştır. Bu durumda arka plana ait birçok örnek yanlış olarak nesneye ait görüntüler olarak sınıflandırılacak (false positives) bu da sistemin güvenilirliğini azaltacaktır. Bizce bu sorunu çözmek için araştırmacılar arka plan verilerini modellemeye ziyade ilginin büyük bölümünü nesne örneklerinin giriş uzayında gömülü olduğu bölgelerin kestirimi, yani nesne sınıfının topolojik yapısını modelleme üzerinde toplamalıdır. Başka

bir deyişle sınıflandırma problemi iki-sınıflı sınıflandırma yerine tek sınıflı sınıflandırıcı dizayn etme problemi olarak ele alınmalıdır. Nesnelerin topolojik yapısı uygun modellerle modellendikten sonra, bir test örneğinin bu geometrik modellere olan uzaklığına bakılarak o örneğin nesne sınıfına ait olup olmadığına karar verilebilir. Şekil 1’de görüldüğü gibi nesnelere ait örneklerin gömülü olduğu bölgeler doğru bir şekilde modellenebilirse, nesnelerin konumlarının belirlenmesi için daha doğru karar sınırları bulunabilir.



Şekil 1. Kesik çizgiyle gösterilen sınırlar üçgen sembollerle ifade edilen nesne sınıfının topolojisini belirlemektedir. Bu bölgelerin yakınlarna düşen örneklerin nesne sınıfına ait olması beklenirken arka plan sınıfına ait örneklerin azlığı ve uygun yerlerden seçilmemiş olması sebebiyle iki-sınıflı sınıflandırıcının bulunduğu karar sınırları (koyu renkle belirtilen karar sınırı) doğru olmayabilir.

Biz bu çalışmada nesne konum bulma için 3 katmandan oluşan ardışıl bir sınıflandırıcı sistemi geliştirdik. Ardışıl sınıflandırıcısının ilk katmanında hızlı bir sınıflandırıcı olan doğrusal destek vektör makineleri kullanıldı. Bu katmanın amacı nesneye ait hemen hemen tüm örnekleri geçirerek, arka plana ait örneklerin bir çoğunu elemektir. İkinci aşamasında ise nesne sınıfını doğrusal bir hiper küre ile modelleyen sınıflandırıcı kullanılmıştır. Bu sınıflandırıcı da oldukça hızlı olup destek vektör makineleri ile uyumlu bir şekilde çalışarak ilk katmanda geçen bir çok arka plan görüntüsünü Şekil 2’de görüldüğü gibi elemektedir. Son katmanda ise doğrusal olmayan hiper küre kullanılmıştır. Bu sınıflandırıcı hız bakımından diğer iki katmandaki sınıflandırıcılara oranla çok daha yavaştır. Fakat her iki katmandaki sınıflandırıcıyı da geçen az sayıda örneğe uygulanmaktadır.

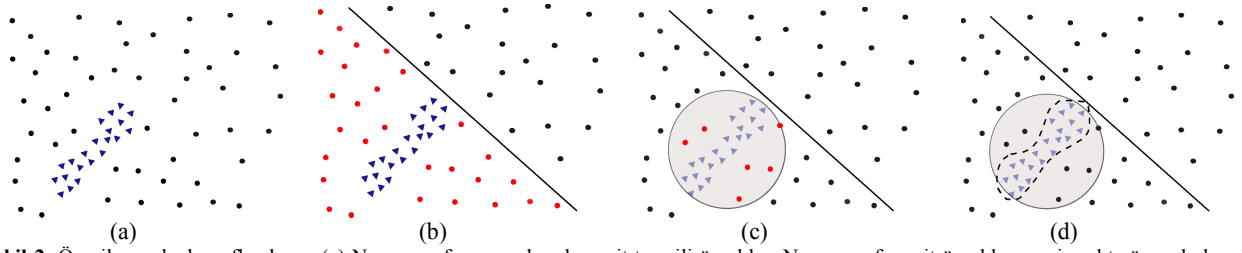
Destek vektör makineleri oldukça bilinen bir yöntem olduğundan bu yönteme ait detaylar atlanarak aşağıda hiper küreleri kullanan sınıflandırıcı anlatılmıştır.

2.1. Hiper Küre Sınıflandırıcısı

Bir sınıfa ait tüm örnekleri içeren en küçük ve kompakt hiper küre kavramı Tax ve Duijn [6] tarafından önerilmiş ve aykırı değer sezimi için başarı ile kullanılmıştır. Bir sınıfa ait d -boyutlu uzaydaki veri örneklerini $\{\mathbf{x}_i \in R^d\}_{i=1, \dots, n}$ ile gösterelim.

Bu sınıfa ait örnekleri içeren en küçük ve kompakt hiper küreyi temsil eden iki parametre vardır: kürenin merkezi c ve kürenin yarıçapı r . Hiper kürenin merkezi ve yarıçapı veri örneklerinin giriş uzayındaki konumlarına göre değişir ve aşağıdaki ikilenik programlama (quadratic programming) probleminin çözülmesiyle bulunurlar.

$$\begin{aligned} \min_{c, r, \xi} \quad & r^2 + \gamma \sum_{i=1}^n \xi_i \\ \text{s.t.} \quad & \|\mathbf{x}_i - \mathbf{c}\|^2 \leq r^2 + \xi_i \end{aligned} \quad (1)$$



Şekil 2. Önerilen ardışıl sınıflandırıcı: (a) Nesne sınıfına ve arka plana ait temsili örnekler. Nesne sınıfına ait örnekler mavi renkte üçgenlerle arka plan ise siyah renkli çemberlerle gösterilmiştir. (b) Birinci katmanda uygulanan doğrusal destek vektör makinelerinin çıktısı. Yanlış olarak nesne sınıfına atanan arka plan görüntüleri (false positives) kırmızı renk ile gösterilmiştir. (c) İkinci katmandaki doğrusal hiper küre sınıflandırıcısı birinci katmanı geçen bir çok arka plan örneğini elemiştir. (d) Son katmanda kullanılan doğrusal olmayan hiper küre sınıflandırıcısı nesne sınıfını çok doğru bir şekilde modeller ve her iki katmandan geçen yanlış örnekleri eleyerek en son kararları verir.

Genellikle bu problemi çözmek yerine bu problemin aşağıda verilen çiftleşliği (dual) çözülür

$$\begin{aligned} \min_{\alpha} \quad & \sum_{i,j} \alpha_i \alpha_j \langle \mathbf{x}_i, \mathbf{x}_j \rangle - \sum_i \alpha_i \langle \mathbf{x}_i, \mathbf{x}_i \rangle \\ \text{s.t.} \quad & \sum_{i=1}^n \alpha_i = 1, \quad 0 \leq \alpha_i \leq \gamma, \quad i = 1, \dots, n. \end{aligned} \quad (2)$$

Yukarıdaki formülde α_i 'ler Lagrange katsayıları olup $\gamma \in [0, 1]$ ise kullanıcının belirlediği ve kompakt modele çok uzak olan aykırı değerleri belirlemede kullanılan bir katsayıdır. Problemin çözümünü veren optimal α_i katsayıları

belirledikten sonra hiper kürenin merkezi $\mathbf{c} = \sum_{i=1}^n \alpha_i \mathbf{x}_i$ formülü ile yarıçapı ise $0 < \alpha_i < \gamma$ şartını sağlayan Lagrange katsayısına karşılık gelen herhangi bir \mathbf{x}_i örneğinin hiper kürenin merkezine olan uzaklığına bakılarak bulunur, $r = \|\mathbf{x}_i - \mathbf{c}\|$.

Hiper küreyi karakterize eden parametreler (kürenin merkezi \mathbf{c} ve kürenin yarıçapı r) bulunduğundan sonra bir test örneğinin, \mathbf{x}_{test} , kürenin merkezine olan uzaklığı bulunur ve uzaklık yarıçapla karşılaştırılarak o nesnenin o sınıfa ait olup olmadığı belirlenir. Matematiksel olarak ifade edilirse

$$\begin{aligned} |(\|\mathbf{x}_{test} - \mathbf{c}\|) - r| \leq \theta & \Rightarrow \mathbf{x}_{test} \text{ nesne sınıfına ait;} \\ |(\|\mathbf{x}_{test} - \mathbf{c}\|) - r| > \theta & \Rightarrow \mathbf{x}_{test} \text{ arkaplan sınıfına ait} \end{aligned}$$

Hiç şüphesiz sınıfların topolojisi çok daha karmaşık ve konveks olmayan yapılar olabilir. Bu durumda sınıfları hiper kürelerle modellemek uygun değildir. Bu problemin çözümü için izlenen yol, örnekleri kernel hilesini kullanarak bu modelin uygun olduğu çok daha büyük boyutlu bir uzaya çıkarmak ve bu yeni uzayda kürenin parametrelerini bulmaktır. Bu işlem destek vektör makinelerinde olduğu gibi eşitlik 2'deki iç çarpımların kernel fonksiyonlarla değiştirilmesi ile kolayca gerçekleştirilir. Daha açık bir şekilde ifade edilirse her bir iç çarpım $\langle \mathbf{x}_i, \mathbf{x}_j \rangle$, kernel fonksiyonu $k \langle \mathbf{x}_i, \mathbf{x}_j \rangle = \langle \phi(\mathbf{x}_i), \phi(\mathbf{x}_j) \rangle$ ile değiştirilmelidir.

Burada $\phi: R^d \rightarrow \mathcal{S}$ girişi uzayındaki örnekleri büyük boyutlu yeni uzaya aktaran fonksiyonu ifade etmektedir. Kullanılan kernel fonksiyonlar arasında m . dereceden polinom fonksiyonu $k(\mathbf{x}_i, \mathbf{x}_j) = (\langle \mathbf{x}_i, \mathbf{x}_j \rangle)^m$, Gaussian kerneli $k(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\|\mathbf{x}_i - \mathbf{x}_j\|^2 / q)$ ve sigmoid kerneli $k(\mathbf{x}_i, \mathbf{x}_j) = \tanh(\kappa \langle \mathbf{x}_i, \mathbf{x}_j \rangle) + \Theta$ sayılabilir. Bu sayede çok daha karmaşık sınıfların topolojileri doğru olarak modellenebilir.

Doğrusal olmayan durumda bir örneğin hiper küreye olan uzaklığının hesaplanabilmesi için sıfırdan farklı Lagrange katsayılarına karşılık gelen destek vektörleri kullanılarak kernel fonksiyonların hesaplanması gerekir. Bu sebeple doğrusal olmayan hiper küre sınıflandırıcısı doğrusal hiper küre sınıflandırıcısına oranla oldukça yavaştır. Fakat deneysel çalışmalar doğrusal olmayan hiper küre sınıflandırıcısının doğrusal olmayan destek vektör makinelerine oranla çok daha az destek vektörü döndürdüğünü ortaya çıkarmıştır. Bu da önerilen ardışıl sınıflandırıcıyı kullanan konum bulma sisteminin hızını doğrusal olmayan destek vektör makineleri kullanan konum bulma sisteminin hızına oranla 20 kata çıkan oranlarda arttırmıştır. Bu sebeple önerilen konum bulma sistemi, literatürdeki doğrusal olmayan sınıflandırıcı kullanan konum bulma sistemlerine oranla çok daha avantajlıdır.

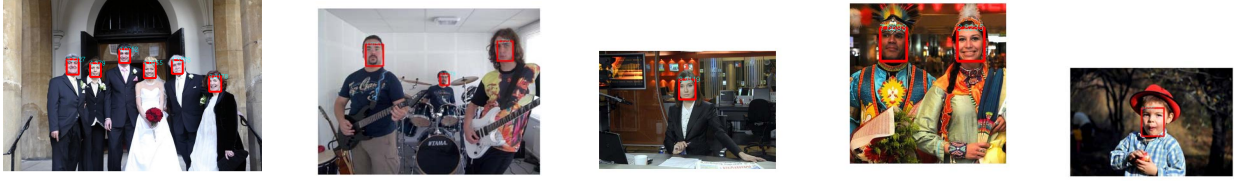
3. Deneysel

Önerilen konum bulma yöntemini dijital imgelerde yüz ve insanların konumlarını bulmada kullandık.

3.1. Yüz Sezme

Dijital imgelerde yüzlerin konumlarının bulunması için ESOGU Face Detection Database (Eskişehir Osmangazi Üniversitesi Yüz Sezme Veri Tabanı) adını verdiğimiz bir veri tabanı oluşturduk. Veritabanı toplam 285 adet renkli dijital imgeden oluşmaktadır ve içinde 970 adet ön cepheden gözükten yüz resmi içermektedir. Dijital imgelerin tamamına yakını günlük hayatı yansıtan kareler olup, arka plan oldukça karmaşıktır. Ayrıca yüzler arasında belirgin aydınlatma ve ölçek farklılıkları vardır.

Konum bulma sistemini eğitmek için ilk olarak web üzerinden 12500'e yakın yüz resmi topladık. Arka plan içinde yüz içermeyen resimlerden rastgele 10000 adet örnek aldık. Seçilen resimler 35×28 boyutlarına sığacak şekilde örneklendikten sonra local binary patterns (LBP) ve histograms of oriented gradients (HOG) öznitelikleri çıkartılmıştır. LBP'ler için her resim 2×2 eşit eş parçaya bölünmüş ve her parçadan yarıçapı 1 ve komşuluğu 8 olan LBP histogramları çıkartılmıştır. Bu histogramlar arka arkaya bağlanmak suretiyle LBP öznitelik vektörü elde edilmiştir. HOG'lar için 6×6 piksel boyutunda hücreler (cell) oluşturularak her bir hücreden elde edilen gradyanlar kullanılarak 19 boyutlu açı histogramları elde edilmiştir. LBP'lerde olduğu gibi her bir hücre için oluşturulan histogramlar arka arkaya bağlanarak HOG histogram vektörü elde edilmiştir. Daha sonra bu iki öznitelik vektörü birleştirilerek imgeleri betimleyen LBP+HOG öznitelik vektörü oluşturulmuştur.



Şekil 3. Önerilen ardışıl sınıflandırıcının ESOGU Yüz Sezme veritabanından seçilen bazı imgeler üzerindeki çıktıları. Yöntem yüzlerin konumlarını başarılı bir şekilde bulmaktadır.

İlk verilerle eğitilen sistem binlerce imge üzerinde denenerek başarısız olduğu zor ve yeni örnekler toplanmış ve bunlarda eğitim setine dahil edilmiştir. Eğitim setindeki son örnek sayısı 120000 civarındadır. Konum bulma algoritmalarının başarımlarını ölçmek için PASCAL VOC metriği kullanılmıştır. Buna göre konum bulma sisteminin döndürdüğü pencere R ile elle işaretlenen yüzlerin konumlarını içeren pencereler Q arasındaki örtüşmenin tüm alana olan oranı, $\frac{alan|Q \cap R|}{alan|Q \cup R|}$, 50%'den büyükse bu doğru bir konum bulma sayılmıştır. Aksi durumda bulunan pencere yanlış (false positive - FP) olarak kabul edilmiştir.

Buna göre elde edilen başarımlar ve yanlış bulunan pencere sayısı (FP) Tablo I'de verilmiştir. Bunlara ilaveten tabloda ayrıca Precision-Recall eğrilerinden elde edilen ortalama kesinlik skorları da (average precision-AP) verilmiştir. Tabloda, Cascade I sadece doğrusal sınıflandırıcıları içeren konum bulma sistemini, Cascade II son katmanında doğrusal olmayan hiper küre sınıflandırıcısını içeren sistemi, Cascade III ise son katmanında doğrusal olmayan destek vektör makinelerini içeren sistemi ifade etmektedir. Sonuçlardan da görüldüğü üzere önerilen yöntemler OPENCV ile birlikte gelen Viola&Jones tarafından önerilen ardışıl sınıflandırıcıdan ve Fdlib algoritmasından (<http://people.kyb.tuebingen.mpg.de/kienzle/fdlib/fdlib.htm>) çok daha iyi sonuçlar vermiştir. Doğrusal olmayan hiper küre sınıflandırıcısı 2393 destek vektör döndürürken, doğrusal olmayan destek vektör makineleri 15657 destek vektör döndürmüştür. Bu sebeple Cascade II, Cascade III'ten ortalama 8 kat daha hızlı çalışmaktadır.

3.2. İnsan Sezme

İnsan sezimi için Inria Person veritabanı kullanılmıştır. Yüz sezmede olduğu gibi imgeleri betimlemek için LBP+HOG öznelikleri kullanılmıştır. Eğitim setindeki insanlara ait kısıtlı örnek sayısını arttırmak için, elle belirlenen doğru koordinatlar rastgele küçük miktarlarda ötelenerek örnek sayısının arttırımı sağlanmıştır. Daha sonra test setindeki imgeler kullanılarak rastgele 12180 arka plan örneği oluşturulmuştur. Bu verilerle eğitilen sınıflandırıcılar eğitim setindeki tüm imgelere uygulanarak zor örnekler toplanmış ve bunlar da eğitim setindeki verilere eklenmiştir. Doğrusal olmayan hiper küreleri kullanan Cascade II sistemi 2818 destek vektör döndürürken, doğrusal olmayan destek vektör makinelerini kullanan Cascade III sistemi 28251 destek vektör kullanmaktadır. Bu da Cascade II'nin Cascade III'e oranla ortalama 20 kat daha hızlı çalışmasını sağlamıştır.

Elde edilen başarımlar Tablo I'de verilmiştir. Önerilen yöntemlerin başarısı şu an en iyi konum bulma yöntemleri arasında gösterilen Hussain&Triggs [5] ve Felzenswalb ve arkadaşları [3] tarafından önerilen yöntemlerle karşılaştırılmıştır. Önerilen tüm ardışıl sınıflandırıcılar bu veri tabanında literatürde en iyi kabul edilen yöntemlerden çok daha başarılı sonuçlar vermiştir. Doğrusal olmayan destek

vektör makinelerini kullanan Cascade III en iyi sonuçları verirken, doğrusal olmayan hiper küreleri kullanan Cascade II ikinci olmuştur. Fakat Cascade II'nin çok daha hızlı çalıştığı gerçeği unutulmamalıdır.

Tablo I. Konum Bulma Yöntemlerinin Başarımları

ESOGU Face Detec. Database	Sezme Başarımı	False Positives	AP
Cascade I	91.24%	550	95.67%
Cascade II	92.47%	18	99.01%
Cascade III	92.27%	10	98.67%
Fdlib	63.81%	344	-
Viola&Jones	75.36%	103	98.60
Inria People Database	Sezme Başarımı	False Positives	AP
Cascade I	78.19%	497	90.43%
Cascade II	75.67%	144	93.46%
Cascade III	82.83%	104	96.03%
Hussain & Triggs [5]	-	-	75.00%
Felzenswalb et al [3]	-	-	86.90%

4. Sonuçlar

Bu çalışmada dijital imgelerdeki nesnelerin konumlarını bulma amaçlı yeni bir yöntem önerdik. Bu doğrultuda nesneye ait sınıfa ait örnekleri yakınsamak için doğrusal ve doğrusal olmayan hiper küreler kullanılmıştır. Daha sonra bu modeller kullanılarak oluşturulan sınıflandırıcılar ardışıl sınıflandırıcılardan oluşan bir sisteme adapte edilmiştir. Ardışıl sınıflandırıcıların ilk katmanlarında oldukça hızlı çalışan fakat başarımları çok yüksek olmayan doğrusal sınıflandırıcılar kullanılırken, son katmanda ise yavaş fakat başarımları çok yüksek sınıflandırıcı kullanılmıştır. Oluşturulan sistem hem yüz hem de insan seziminde literatürdeki en başarılı yöntemlerden çok daha iyi sonuçlar vermiştir.

5. Kaynakça

- [1] T. Ahonen, A. Hadid, M. Pietikainen, "Face description with local binary patterns", *IEEE Tran. on PAMI*, 28, pp. 2037-2041, 2006.
- [2] N. Dalal, B. Triggs, "Histograms of oriented gradients for human detection," *CVPR*, 2008.
- [3] P. Felzenswalb, D. McAllester, D. Ramana, "A discriminatively trained, multiscale, deformable part model," *CVPR*, 2008.
- [4] H. Harzallah, F. Jurie, C. Schmid, "Combining efficient object localization and image classification," *ICCV*, 2009.
- [5] S. Hussain, B. Triggs, "Feature sets and dimensionality reduction for visual object detection," *BMVC*, 2008.
- [6] D. M. J. Tax, R. P. W. Duin, "Support vector data description," *Machine Learning*, 54, pp. 45-66, 2004.
- [7] A. Vedaldi, V. Gulshan, M. Varma, A. Zisserman, "Multiple kernels for object detection," *ICCV*, 2009.
- [8] P. Viola, M. J. Jones, "Robust real-time face detection," *Int. Jour. of Computer Vision*, 57, pp. 137-154, 2004.