

Nitelik Tabanlı Sınıflandırıcılar ve Koşullu Rastgele Alan ile Dikkat Çeken Görsel Bölge Tespiti

Visual Saliency Estimation via Attribute Based Classifiers and Conditional Random Field

Berkan Demirel¹, Ramazan Gökberk Cinbiş², Nazlı İkizler-Cinbiş¹

¹Bilgisayar Mühendisliği Bölümü, Hacettepe Üniversitesi, Ankara, Türkiye
{n14125263,nazli}@cs.hacettepe.edu.tr

²Bilgisayar Mühendisliği Bölümü, Bilkent Üniversitesi, Ankara, Türkiye
gcinbis@cs.bilkent.edu.tr

Özetçe—Dikkat çeken görsel bölge tahmini, resimlerde ya da sahnelerde insan gözünün öncelikli olarak odaklandığı bölgeleri bulmayı amaçlayan bir bilgisayarlı görü problemidir. Pek çok bilgisayarlı görü problemi bir sahnedeki arkaplan öğelerini yoksaymayı gerektirdiği için, bu tür problemlerde dikkat çeken görsel bölge tahmini bir ön işlem adımı olarak kullanılabilir. Bu çalışmada yukarıdan aşağıya dikkat çeken bölge tahmini probleminin çözümüne yönelik olarak nitelik tabanlı sınıflandırıcılar ve Koşullu Rastgele Alan (KRA) yöntemlerinin bir arada kullanıldığı bir yöntem sunulmaktadır. Deneysel sonuçlar nitelik tabanlı sınıflandırıcı sonuçlarının görsel bilgiyi alt seviye özelliklere göre daha iyi kodlayabildiğini göstermiştir ve geliştirilen yöntemin, Graz-02 veri kümesi üzerinde en iyi yöntemlerle karşılaştırıldığında umut verici sonuçlar ürettiği gözlenmiştir.

Anahtar Kelimeler—yukarıdan-aşağıya dikkat çeken bölge tahmini, nitelik, koşullu rastgele alan, ayırt edici sözlük.

Abstract—Visual Saliency Estimation is a computer vision problem that aims to find the regions of interest that are frequently in eye focus in a scene or an image. Since most computer vision problems require discarding irrelevant regions in a scene, visual saliency estimation can be used as a preprocessing step in such problems. In this work, we propose a method to solve top-down saliency estimation problem using Attribute Based Classifiers and Conditional Random Fields (CRF). Experimental results show that attribute-based classifiers encode visual information better than low level features and the presented approach generates promising results compared to state-of-the-art approaches on Graz-02 dataset.

Keywords—top-down saliency estimation, attribute, conditional random field, discriminative dictionary.

I. GİRİŞ

Dikkat çeken görsel bölge tahmini probleminde amaç, çeşitli varsayım ve modellerle insanların resimlerde hangi bölgelere odaklandıklarını tespit etmektir. Bu problem, özellikle son yıllarda bilgisayarlı görü ve görüntü işleme araştırmalarında önemli bir yer tutmaya başlamıştır. Bu ilginin sebebi görsel dikkat çeken bölge tahmini probleminin hem insan algısı üzerine araştırmalar hem de çeşitli bilgisayarlı görü problemleriyle ilişkili olmasıyla açıklanabilir.

Bu problem görsel verilerin yaygınlaşması ile birlikte giderek önem kazanma potansiyeline sahiptir. Bunun önemli bir sebebi, görsel çözümlerle resimlerde sadece daha dikkat çekici bölgelere odaklanılması sağlanabilir. Böylece, dijital görsel verilerin daha hızlı ve daha otomatik analizi yapılabilir. Dolayısıyla, dikkat çeken bölge tahminleri tek başına kullanılmasının ötesinde diğer problemlerle birlikte, onların çözümüne yönelik öncül olarak da kullanılabilir.

Dikkat çeken bölge tespiti problemi kapsamında ortaya konulan modeller alttan-üste (*bottom-up*) ve yukarıdan-aşağıya (*top-down*) olmak üzere iki grupta toplanabilir. Alttan-üste yaklaşımlarda resimlerdeki yerel ipuçlarından faydalanarak problem çözümlenmektedir. Bu ipuçları renk, parlaklık, yoğunluk ya da doku bilgileri olabilir ve genel olarak yaklaşımlar [3] [4] [5] bu ipuçlarının yerel komşuluktaki ilişkisine bakarak karar vermektedir. Yukarıdan aşağıya yöntemler ise spesifik hedef tespiti yapmaya çalışmaktadır ve bu açıdan nesne tanıma problemi ile yakın bir ilişki içindedir.

Bu çalışmada görsel dikkat çeken bölge tespiti problemini çözmeye yönelik olarak yukarıdan-aşağı bir yaklaşım sunulmaktadır. Bu yaklaşımda, temel olarak nitelik (*attribute*) tabanlı sınıflandırıcı çıktıları öznelik (*feature*) vektörleri olarak kullanılmıştır. Nitelikler, nesne ya da sahneye ait görsel (renk, şekil, doku) ve fonksiyonel bilgi parçalarını kodlayabilen orta düzey öznelikler (*middle level features*) olarak tanımlanabilir. Daha sonra bu öznelikler üzerinden koşullu rastgele alan (KRA) ve ayırt edici sözlük eş zamanlı olarak öğrenilmiştir. Görsel niteliklerin dikkat çeken bölge tespiti problemi üzerine etkileri literatürde daha önce araştırılmamıştır. Çalışmamız bu problem üzerinde nitelik tabanlı özneliklerin kullanması bakımından yenilikçi bir yaklaşımdır. Graz-02 veri kümesi üzerinde yaptığımız deneylerde literatürdeki en iyi yöntemlerle karşılaştırıldığında umut verici sonuçlar verdiği gözlemlenmiştir.

II. İLGİLİ ÇALIŞMALAR

Dikkat Çeken Bölge Tahmini. Son dönemlerde bu problem üzerine yapılan çalışmalarda yukarıdan-aşağıya yaklaşımlar ön plana çıkmaktadır. Yang ve Yang [1] alt seviye özellikler ile ayırt edici sözlük ve koşullu rastgele alan kullanarak problemi çözmeye çalışmıştır. Koçak vd. [2] parça tabanlı

olan bu yöntemi süperpiksel tabanlı olacak şekilde geliştirmişlerdir. Liu vd. [6] dikkat çeken bölge bulma problemini resim bölütleme problemi olarak tanımlamıştır. Bu amaçla çok ölçekli kontrast ve mekansal renk dağılımı öznitelikleri çıkarılmış ve koşullu rastgele alan yöntemiyle dikkat çeken bölge tespit edilmeye çalışılmıştır. Çalışmamızın yukarıdaki çalışmalardan farkı, alt-düzye öznitelikler yerine, nitelik sınıflandırıcılarının tahmin vektörlerinin öznitelik olarak kullanılması ve bu sayede daha etkin sonuçların elde edilmesidir.

Borji ve Itti [3] farklı renk uzaylarında oluşturulmuş dikkat çeken bölge haritalarını birleştirmişlerdir. Bu çalışmada RGB renk uzayında objelerin daha dikkat çekici olduğunu ancak bazı objeler içinse Lab renk uzayının daha etkili olduğu gösterilmiştir. Jiang vd. [7] üç farklı ve önemli görsel ipucunu (teklik, dikkat çekicilik, nesne olabilirlik) tek bir modelde toplamıştır. Geliştirilen modelde teklik, görsel kontrast bilgisini; dikkat çekicilik, görsel olarak dikkat çeken bölgenin odakta olacağı bilgisini; nesne olabilirlik ise dikkat çeken bölgenin bütünlüğü bilgisini ele almaktadır. Zhang ve Sclaroff [4] resimleri renk kanallarına uyguladığı farklı eşik değerleri üzerinden elde ettiği ikili değerlerle tanımlamıştır. Bu tanımlama ile Gestalt ayırım prensibi yardımıyla resimler üzerinde topolojik bir yapı kurulmuş, dikkat çeken bölge tahmini yapılmaya çalışılmıştır.

Zhu vd. [5] önplan yerine güçlü bir arkaplan modellemeye çalışmıştır. Ardından geliştirilen model ile alt düzey ipuçları birlikte kullanılmıştır. Kim vd. [8] az boyutlu RGB renk uzayını çok boyutlu bir uzaya taşımışlardır. Taşınan uzayda önplanda bulunan obje ile arkaplan daha sağlıklı ayırt edilebilir hale gelmiştir. Erdem ve Erdem [9] parça tabanlı kovaryans tanımlayıcıları kullanarak alttan-üste dikkat çeken bölge tahmini yapabilecek bir model hazırlamışlardır.

Nitelik Tabanlı Sınıflandırıcılar. Nitelik tabanlı sınıflandırıcı problemlerin çözümüne yönelik sağladığı ek bilgiler nedeniyle araştırmacıların üzerinde durduğu konuların başında gelmektedir. Farhadi vd. [10] nesne tanıma problemini nitelikler üzerinden tanımlamıştır. Bu nitelikler nesnelerin mantıksal ve ayırt edici özelliklerini temsil etmektedir. Lampert vd. [11] sıfır-bilgi (*zero shot*) problemini nitelikler üzerinden tanımlamış ve çözmeye çalışmıştır. Ferrari vd. [12] nitelikleri resim bölgelerinde ortaya çıkan kalıplar olarak ifade etmiştir.

III. YÖNTEM

Geliştirdiğimiz yöntem görsel dikkat çeken bölge tahmini problemini yukarıdan-aşağıya bir yaklaşımla çözmeye çalışmaktadır. Yöntem Yang ve Yang [1] tarafından ortaya konulan yöntemin geliştirilmiş halidir. Onlardan farklı olarak SIFT gibi alt-seviye öznitelikler kullanmak yerine, bu çalışmada nitelik tabanlı sınıflandırıcı sonuç vektörleri öznitelik olarak kullanılmıştır.

Modelde, eğitim kümesi dahilinde bulunan resimler alt parçalara bölünmüş ve her bir parça üzerinde nitelik tabanlı sınıflandırıcılar çalıştırılmıştır. Nitelik tabanlı sınıflandırıcı çıktıları birer özellik vektörü olarak kullanılarak k-ortalamlar algoritması yardımıyla görsel sözlük oluşturulmuştur. Ardından, resim parçaları yardımıyla koşullu rastgele alan çizgesi oluşturulmuş ve ağırlıklar görsel sözlük yardımıyla belirlenmiştir. Sonrasında eğitim kümesinde bulunan resimlere ait yer doğruluk haritaları ve koşullu rastgele alan çizgesi

çıktıları kullanılarak daha ayırt edici ve doğru bir görsel sözlük oluşturulmaya çalışılmaktadır. Bu adımlar belli sayıda iterasyon dahilinde tekrarlanmaktadır ve her bir iterasyonda gerek çizge parametreleri gerekse görsel sözlük güncellenmektedir. Böylece şartlı rastgele alan ve görsel sözlük, ardıl ve tekrarlı olarak nitelik tabanlı sınıflandırıcılardan elde edilen öznitelikler üzerinden öğrenilmektedir.

Eğitim kümesi üzerinde işlemler tamamlandıktan, şartlı rastgele alan ve görsel sözlük öğrenildikten sonra test resimleri de parça tabanlı olarak ifade edilip her bir parça üzerinde nitelik tabanlı sınıflandırıcılar çalıştırılmaktadır. Elde edilen özellikler şartlı rastgele alan ve görsel sözlük üzerinden değerlendirilmekte ve görsel dikkat çeken bölge haritası oluşturulmaktadır. Geliştirilen yöntem Algoritma 1 üzerinden özetlenmiştir. Aşağıda, bu algoritma dahilinde anlatılmakta olan adımlar detaylandırılmaktadır.

A. Şartlı Rastgele Alan ve Sözlük Öğrenmesi

Yang ve Yang [1] tarafından geliştirilen yöntemde, şartlı rastgele alan ve ayırt edici görsel sözlük sıralı ve beraber öğrenilmektedir. Resim üzerinden elde edilen p -boyutlu parçalar $X = [x_1, x_2, x_3, \dots, x_m]$, her bir parçanın hedef obje içerip içermediğini gösteren ikili etiketler $Y = [y_1, y_2, y_3, \dots, y_m]$ olarak ifade edilip, bu etiket ve parçalar üzerinden hedef objeyi en iyi ifade ve ayırt edebilecek sözlük $D \in R^{p \times k}$ öğrenilmektedir. Sözlük hedef objeyi temsil edecek elemanlardan oluşmaktadır ve bu elemanlar $D = [d_1, d_2, d_3, \dots, d_k]$ olarak ifade edilebilir.

Sözlükte hedef objeye ilişkin daha bilgi verici bir gösterim elde etmek amacıyla (1) numaralı denklem üzerinden, ℓ_1 düzenleme fonksiyonu etkisiyle, her bir parçanın seyrek kodlaması gerçekleştirilmektedir. Denklemdeki optimizasyon hesabında her bir X parça kümesi, seyrek kodlama sonrası, $S(X, D)$ gizli değişkenleri ile ifade edilmektedir:

$$S(X, D) = \underset{S}{\operatorname{argmin}} \frac{1}{2} \| X - DS \|^2 + \lambda \| S \|_1 \quad (1)$$

Bu denklemde ilk terim X parçalarının iyi bir gösterimini sağlamaya çalışırken, ikinci terim seyreklik derecesini belirlemektedir. λ parametresi ise sabit bir seyreklik derecesi katsayısını temsil etmektedir. Optimizasyon sonrası her bir X parçası için elde edilen gizli değişkenler aşağıdaki gibi temsil edilebilir:

$$S(X, D) = [S(x_1, D), S(x_2, D), S(x_3, D) \dots S(x_m, D)] \quad (2)$$

Böylece görsel bilgi sözlükten gizli değişkenlere aktarılmıştır. Ardından KRA modeli aşağıdaki gibi oluşturulmaktadır.

$$P(Y | S(X, D), W) = \frac{1}{Z} e^{-E(S(X, D), Y, W)} \quad (3)$$

Burada Z normalizasyon fonksiyonunu, $E(S(X, D), Y, W)$ ise enerji fonksiyonunu temsil etmektedir. Ayrıca W , Koşullu Rastgele Alan parametrelerini temsil etmektedir. Yukarıdaki formül üzerinde verilen $S(X, D)$ ve W değerlerine göre Y etiketinin olasılığı elde edilmektedir. Bu olasılık değerlerinin yer doğruluk haritasına uygun olmasını sağlamak amacıyla $S(X, D)$ ve W üzerinde optimizasyon yapılmaktadır.

Model üzerinde öğrenme tamamlandıktan sonra, test aşamasında da parça tabanlı çalışılmaktadır. Bir parça alanı hedef

objeyi içeriyorsa, komşu alanların da hedef nesne içerme olasılığı yükselmektedir. Bu nedenle komşu parçalar enerji fonksiyonunun ikili potansiyeli üzerinde etkilidir. Bu varsayım üzerinden, resim parçaları üzerinden hedef obje bilgisi aşağıdaki gibi hesaplanabilir.

$$P(Y_i | s_i, w) = \sum_{Y_{N(i)}} P(y_i, Y_{N(i)} | s_i, w) \quad (4)$$

Burada $Y_{N(i)}$ ifadesi çizge üzerinde, ilgili parçanın 4-yönlü komşuluğunu temsil etmektedir. Tüm bu hesaplamaların ardından bir parçanın görsel dikkat çeken bölge olma olasılığı aşağıdaki olasılık ile ifade edilebilir:

$$u(s_i, w) = P(y_i = 1 | S_i, w)$$

B. Nitelik Tabanlı Sınıflandırıcılar

Nitelik tabanlı sınıflandırıcıları elde etmek için Farhadi vd. [10] tarafından oluşturulan yaklaşım kullanılmıştır. Bu yaklaşımda her bir nitelik için birer Destek Vektör Makinesi (DVM) öğrenmesi gerçekleştirilmiştir. Bu öğrenme esnasında renk, şekil ve doku bilgisini kodlamaya yönelik öznelikler kullanılmıştır. Renk ve doku yapısı, materyal bilgisini, görsel kelimeler, bölüm bilgisini, kenar yapısı da şekil bilgisini kodlayabilmek için kullanılmıştır.

Doku tanımlayıcıları her bir piksel için texton filtre bankası üzerinden hesaplanmış ve $k=256$ olacak şekilde k -ortalamalar yardımıyla küme merkezleri bulunmuştur. Her bir piksel değeri en yakın küme merkezinin değerine nicelenmiştir. Görsel kelimeler Yönelimsel Dağılım Histogramı (HOG) [14] özneliği uzamsal piramidi üzerinde 8×8 'lik bloklar, 4 piksel adım boyu ve 2 ölçek kullanılarak elde edilmiştir. $k=1000$ olacak şekilde k -ortalamalar algoritması ile küme merkezleri bulunmuş ve HOG tanımlayıcılar bu merkezlere göre nicelenmiştir. Kenarlar standart Canny Kenar Bulma Algoritması yardımıyla tespit edilip, 8 bölmeye nicelenmiştir. Renk tanımlayıcılar her bir piksel üzerinde çalıştırılmış ve k -ortalamalar algoritması yardımıyla 128 küme merkezi elde edilmiştir. Ardından her bir piksel değeri bu küme merkezleri üzerinden nicelenmiştir. Bu renk değerleri LAB renk uzayından elde edilmiştir.

Her bir resim için Destek Vektör Makinelerinde kullanılmak üzere bu öznelikler çıkarılmış ve resimler, bu öznelikler birleşimi şeklinde temsil edilmiştir. Nitelikleri doğru bir şekilde öğrenebilmek amacıyla ℓ_1 -düzenleştirilmiş lojistik regresyon (ℓ_1 -regularized logistic regression) kullanılarak öznelik seçimi gerçekleştirilmektedir. Bu amaçla ilgili niteliğin ilişkili olduğu her bir sınıf için, o sınıf üzerinden niteliğin ayrımını yapabilecek öznelikler lojistik regresyon ile belirlenmektedir. Ardından sınıflar üzerinden elde edilen öznelikler birleştirilerek (pooling), ilgili niteliğe ait sınıflandırıcı öğrenilmektedir. Bu öğrenme her bir nitelik için doğrusal DVM ile gerçekleştirilmektedir.

IV. DENEYLER

Geliştirdiğimiz yaklaşımı Graz-02 [13] veri kümesi üzerinde analiz edilmektedir. Graz-02 veri kümesi dört farklı kategoriden (bisiklet, insan, araba, arkaplan) resimler içermektedir. Veri kümesi dahilindeki bulunan 365 resim bisiklet, 311 resim insan, 420 resim araba ve 380 resim arkaplan

Algoritma 1 Dikkat Çeken Görsel Bölge Modeli

```

1:  $T_r$ : Eğitim resimleri kümesi
2:  $T_e$ : Test resimleri kümesi
3: for all  $i \in T_r$  do
4:    $P \leftarrow$  ResimBölütüle( $i$ )
5:    $C_i \leftarrow$  NitelikTabanlıSınıflandır( $P$ )
6: end for
7:  $D \leftarrow$  SözlükÖğren( $C$ )
8: for all  $i \leq T$  do
9:   // Sözlük ve nitelikleri kullanarak KRA ağırlıklarını düzenle
10:   $KRA_w \leftarrow$  KRAÖğren( $D, C$ )
11:  // KRA ağırlıkları ve nitelikleri kullanarak sözlük düzenle
12:   $D \leftarrow$  SözlükGüncelle( $KRA_w, C$ )
13: end for
14: for all  $i \in T_e$  do
15:   $P \leftarrow$  ResimBölütüle( $i$ )
16:   $C \leftarrow$  NitelikTabanlıSınıflandır( $P$ )
17:  DikkatÇekenBölgeHaritasıÇıkar( $KRA_w, D, C$ )
18: end for

```

Yöntem	Hedef			
	İnsan	Araba	Bisiklet	Ortalama
SIFT Tabanlı [1] ²	50.6	54.3	58.2	54.5
Nitelik Tabanlı	58.7	60.6	64.9	61.4

Tablo I: Eşit Hata Oranları (EER) Metriği Üzerinden Sınıflandırma Başarıları.

sınıflarına aittir. Değerlendirme kriteri olarak da eşit hata oranlarında piksel düzeyi kesinlik oranlarına bakılmıştır.

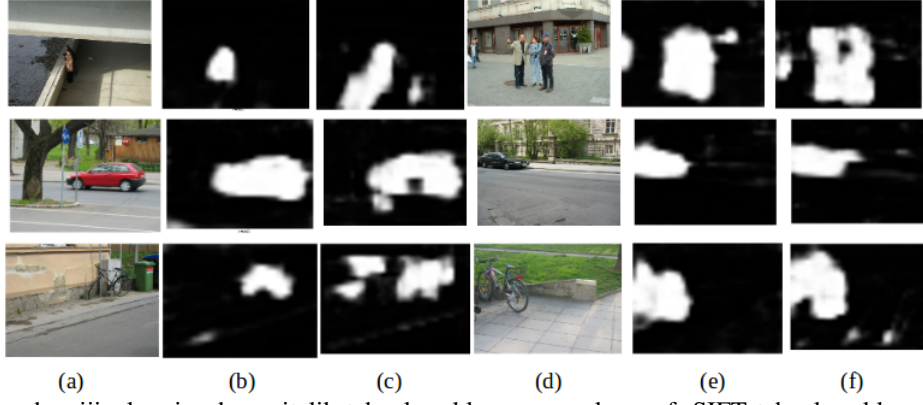
Nitelik tabanlı sınıflandırıcıların eğitimi için aPascal veri kümesi¹ kullanılmıştır. Bu veri kümesi PASCAL VOC 2008 veri kümesi üzerinden oluşturulmuştur ve PASCAL dahilinde bulunan 20 farklı nesneyi tanımlayan 64 farklı nitelik etiketlerine sahiptir. Bu nitelikler nesnelerin şekil, materyal ve bölüm bilgisini ifade etmektedir; "Occluded", "Tail", "Beak", "Head", "Ear" ve "Snout" bu niteliklerden bazılarıdır.

Deneyleerde, her bir nesne sınıfı için farklı koşullu rastgele alan ve sözlük öğrenmesi gerçekleştirilmiştir. Graz-02 veri kümesinde her bir sınıfta 300 görüntüye ait yer doğruluk haritası bulunduğu için eğitim ve test işlemleri bu görüntüler üzerinden gerçekleştirilmiştir. Veri kümesinde tek sayı ile isimlendirilmiş resimler eğitim seti için kullanılırken, diğer görüntüler test aşamasında kullanılmıştır. Dolayısıyla eğitim esnasında her bir sınıfa ait 150 resim pozitif, arkaplan sınıfına ait 150 resim de negatif örnek olarak kullanılmıştır. Her bir resimden, 64×64 piksellik alt pencere boyutu ve 16 piksel pencere kaydırma miktarı kullanılarak, 999 alt alan çıkarılmıştır. Nitelik tabanlı sınıflandırıcılar bu parçalar üzerinde çalıştırılmıştır ve nitelik vektörleri üzerinden seyrek kodlama gerçekleştirilmektedir. Bu seyrek kodlar üzerinden de koşullu rastgele alan ağırlıkları güncellenmektedir. Bu işlem 20 iterasyon boyunca devam etmekte ve daha ayırt edici bir sözlük ile birlikte çizge ağırlıkları öğrenilmektedir.

Modelde 512 görsel kelime kullanılmıştır. Yang ve Yang [1] çalışmasına paralel olarak λ parameresi 0.15 olarak belirlenmiştir. Ayrıca empirik olarak, istatistiksel gradyan azaltma

¹<http://vision.cs.uiuc.edu/attributes/>

² [1] için yazarlarınca sağlanan kod ile elde edilen sonuçlar gösterilmektedir.



Şekil 1: a-d: orijinal resim, b-e: nitelik tabanlı yaklaşım sonuçları, c-f: SIFT tabanlı yaklaşım sonuçları

algoritmasında öğrenme oranı 1e-3, ağırlık maliyet değeri ise 1e-1 olarak belirlenmiştir. Öğrenme işleminin ardından test resimleri üzerinde parça tabanlı gösterim yapılmış ve nitelik tabanlı sınıflandırıcılar çalıştırılmıştır. Ardından öğrenilen koşullu rastgele alan modeli uygulanarak dikkat çeken bölge tespiti yapılmaya çalışılmaktadır.

Tablo I’de verilen sonuçlara göre geliştirdiğimiz yöntem Yang ve Yang [1] tarafından gerçekleştirilen yöntemden daha iyi sonuçlar üretmektedir. [1] tarafından alt düzey özellik olarak SIFT tanımlayıcılar kullanıldığı düşünüldüğünde, bu sonuçlar ışığında, nitelik tabanlı sınıflandırıcıların görsel bilgiyi alt düzey özelliklere göre daha iyi kodladığı söylenebilir. Kocak vd. [2] bu veri kümesi üzerinde, ortalama 70.17% başarı raporlamıştır. Geliştirilen yöntemin bu çalışmaya göre daha düşük performans göstermesinin nedeni [2] tarafından kullanılan yöntemde, resimlerin süperpiksel tabanlı olarak gösterilmesidir. Bu nedenle sınır bilgisini daha iyi kodlayabilmektedirler. Dolayısıyla, nitelik tabanlı sonuçlarımız [2]’in sonuçlarıyla direkt karşılaştırılabilir olmamakla birlikte, süperpiksel tabanlı yaklaşım ile birleştirmeyi ileri çalışmalar için planlamaktayız.

Çeşitli görsel sonuçlar Şekil 1 üzerinde gösterilmiştir. Bu görsellere göre geliştirdiğimiz yöntemin hedef objeler üzerinde [1]’e göre daha net ve sürekli skorlar ürettiği görülmektedir. Ayrıca yanlış alarm (*false alarm*) oranının da daha düşük olduğu gözlenmektedir. Bu gözlemler kullandığımız özneliklerin daha başarılı olduğu savını desteklemektedir.

V. SONUÇLAR VE DEĞERLENDİRME

Bu çalışmada yukarıdan-aşağıya dikkat çeken bölge tespiti problemi için nitelik sınıflandırıcılarını kullanan bir yaklaşım geliştirilmiştir. Geliştirdiğimiz yaklaşıma göre resimlerdeki alt alanlar üzerinde çalıştırılan nitelik tabanlı sınıflandırıcı sonuçları koşullu rastgele alan modeli tarafından öznelik olarak kullanılmaktadır. Deney sonuçları, geliştirilen yaklaşımın umut verici sonuçlar ortaya koyduğunu göstermiştir.

Yapılan deneyler, nitelik tabanlı yöntemlerin görsel bilgiyi alt-düzye özneliklerden daha iyi kodlayabildiğini göstermektedir. Ayrıca nitelik tabanlı yöntemde kullanılan öznelikler ile alt düzey öznelikler ortak bir modelde kullandığında başarı oranının yükselmesi beklenmektedir. Çünkü her iki yaklaşımda kodlanan görsel bilgi farklıdır ve bu bilgilerin birbirini tamamlaması olasılık dahilindedir. Bu gözlem, ileride gerçekleştirilecek çalışmalarda ele alınacaktır.

TEŞEKKÜR

Bu çalışma, 112E149 no’lu TÜBİTAK kariyer projesi tarafından desteklenmiştir.

KAYNAKÇA

- [1] Yang, J., & Yang, M. H. (2012, June). Top-down visual saliency via joint crf and dictionary learning. In Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on (pp. 2296-2303). IEEE.
- [2] Kocak, A., Cizmeciler, K., Erdem, A., & Erdem, E. (2014, September). Top down saliency estimation via superpixel-based discriminative dictionaries. In Proceedings of the British Machine Vision Conference. BMVA Press.
- [3] Borji, A., & Itti, L. (2012, June). Exploiting local and global patch rarities for saliency detection. In Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on (pp. 478-485). IEEE.
- [4] Zhang, J., & Sclaroff, S. (2013, December). Saliency detection: A boolean map approach. In Computer Vision (ICCV), 2013 IEEE International Conference on (pp. 153-160). IEEE.
- [5] Zhu, W., Liang, S., Wei, Y., & Sun, J. (2014, June). Saliency optimization from robust background detection. In Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on (pp. 2814-2821). IEEE.
- [6] Liu, T., Yuan, Z., Sun, J., Wang, J., Zheng, N., Tang, X., & Shum, H. Y. (2011). Learning to detect a salient object. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 33(2), 353-367.
- [7] Jiang, P., Ling, H., Yu, J., & Peng, J. (2013, December). Salient region detection by ufo: Uniqueness, focusness and objectness. In Computer Vision (ICCV), 2013 IEEE International Conference on (pp. 1976-1983). IEEE.
- [8] Kim, J., Han, D., Tai, Y. W., & Kim, J. (2014, June). Salient region detection via high-dimensional color transform. In Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on (pp. 883-890). IEEE.
- [9] Erdem, E., & Erdem, A. (2013). Visual saliency estimation by nonlinearly integrating features using region covariances. Journal of vision, 13(4), 11.
- [10] Farhadi, A., Endres, I., Hoiem, D., & Forsyth, D. (2009, June). Describing objects by their attributes. In Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on (pp. 1778-1785). IEEE.
- [11] Lampert, C. H., Nickisch, H., & Harmeling, S. (2009, June). Learning to detect unseen object classes by between-class attribute transfer. In Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on (pp. 951-958). IEEE.
- [12] Ferrari, V., & Zisserman, A. (2007). Learning visual attributes. In Advances in Neural Information Processing Systems (pp. 433-440).
- [13] Marszatek, M., & Schmid, C. (2007, June). Accurate object localization with shape masks. In Computer Vision and Pattern Recognition, 2007. CVPR’07. IEEE Conference on (pp. 1-8). IEEE.
- [14] Dalal, N., & Triggs, B. (2005, June). Histograms of oriented gradients for human detection. In Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on (Vol. 1, pp. 886-893). IEEE.