





Düzce Üniversitesi Bilim ve Teknoloji Dergisi

Araştırma Makalesi

Bölütleme Kullanarak Doğal Görüntülerde Metin Tanıma

 Yeliz ŞENKAYA ^{a,*},  Çetin KURNAZ ^b

^a Bilgisayar Kullanımı Bölümü, Akkuş Meslek Yüksekokulu, Ordu Üniversitesi, Ordu, TÜRKİYE

^b Elektrik-Elektronik Mühendisliği, Mühendislik Fakültesi, Ondokuz Mayıs Üniversitesi, Samsun, TÜRKİYE

* Sorumlu yazarın e-posta adresi: yelizsenkaya@odu.edu.tr

doi: 10.29130/dubited.1107625

ÖZ

OCR olarak da bilinen optik karakter tanıma, taranan görüntülerdeki bir kelimeyi ya da bir cümleyi tanımak için kullanılan bir yöntemdir. Uzun yıllara dayanan araştırmalarla geliştirilmiştir. Taranan görüntüler üzerindeki metni tespit etmede büyük başarı sağlamıştır. Ancak doğal görüntüler üzerinde istenilen sonucu vermemektedir. Bu nedenle, doğal görüntülerdeki metinleri tespit edebilmek için özel yaklaşımların geliştirilmesi gerekliliği doğmuştur. Bu çalışmada, doğal görüntüler üzerinde metin olan bölgeleri algılamak için Otsu ve maksimum kararlı ekstrem bölgeler (MSER) görüntü bölütleme yöntemleri kullanılmıştır. Görüntü bölütleme, bir görüntüyü daha iyi analiz edebilmek için görüntüyü anlamlı bölgelere ayırma işlemidir. Otsu modelinde görüntü için en uygun eşik değeri belirlenerek, görüntü bu eşik değerine göre ön plan ve arka plan olmak üzere iki sınıfa ayrılmaktadır. MSER yöntemi ise metin olmayan bölgeleri engelleyerek, metin olduğu düşünülen bölgeleri sınırlayıcı kutu içerisine almaktadır. Gerçekleştirilen çalışmada, Otsu metodu ve MSER yöntemi ile ICDAR 2013 veri setinden seçilen 20 doğal görüntü üzerinde metin olan bölgelerinin tespit edilmesi amaçlanmıştır. Doğal görüntü üzerinde bölütleme işlemleri yapıldıktan sonra görüntülere OCR uygulanarak doğal görüntüler üzerindeki metnin tespit edilmesi sağlanmış ve doğruluk oranları karşılaştırılmıştır.

Anahtar Kelimeler: Otsu modeli, Maksimum Kararlı Ekstrem Bölgeler, Optik Karakter Tanıma

Text Recognition in Natural Images Using Segmentation

ABSTRACT

Optical character recognition, also known as OCR, is a method for recognizing a word or a phrase in scanned images. It has been developed through years of research. It has had great success in detecting text on scanned images. However, it does not give the desired result in natural images. Therefore, it is necessary to develop special approaches to detect texts in natural images. This study used Otsu and The Maximum Stable Extremal Regions (MSER) image segmentation methods to detect regions with text on natural images. Image segmentation is dividing an image into meaningful regions to analyze it better. In the Otsu model, the most appropriate threshold value is determined for the image, and the image is divided into two classes, foreground, and background, according to this threshold value. On the other hand, the MSER method blocks non-text regions and encloses regions thought to be text in bounding boxes. The study carried out aimed to determine the text areas on 20 natural images selected from the ICDAR 2013 data set with the Otsu method and the MSER method. After segmentation on the natural image, OCR was applied to the images to detect the text on the natural images, and the accuracy rates were compared.

Keywords: Otsu's method, Maximally Stable Extremal Regions, Optical Character Recognition

I. GİRİŞ

Günümüzde dijitalleşen dünya ile birlikte görüntü üretme ve paylaşma hızı her geçen gün artmaktadır. Üretilen bu görüntüler taranan bir belge olabileceği gibi, doğal bir görüntüde olabilmektedir. Bu görüntülerden doğru ve anlamlı metin çıkarma ihtiyacı araştırmacılar için büyük önem kazanmıştır. Optik karakter tanıma (optical character recognition, OCR), görüntülerdeki bir karakteri, bir kelimeyi ve hatta bir cümlenin tamamını tanımak için kullanılan bir yöntemdir. OCR, taranan belgelerdeki metni tanıırken yüksek başarımlı orana sahipken, doğal görüntü metinlerini tanıırken yetersiz kalmaktadır. Taranan belge görüntüsündeki metinler ile doğal görüntülerdeki metinler birbirinden çok farklı özelliklere sahiptir. Taranan belge görüntüsündeki metin, beyaz bir arka plan ve tüm belge boyunca uzanan yazı tipi stili gibi bir dizi özellikten oluşurken, doğal görüntüdeki metinde ise kenarları etkileyen farklı açılar, farklı yazı tipi stilleri, arka plan ve ön plan renkleri gibi farklı birçok özellik bulunmaktadır [1]. Doğal görüntülerdeki metinler, bu özelliklerden dolayı taranan belge görüntüsündeki metinlere göre OCR tarafından daha zor tespit edilmektedir. Bu sebeplerden dolayı doğal görüntü metinlerinin tespit edilmesi için özel yaklaşım geliştirilmesi ihtiyacı doğmaktadır.

Literatürde doğal görüntülerdeki metinleri daha iyi tanımak için farklı yöntemler kullanılarak birçok çalışma yapılmıştır. Gerçekleştirilen çalışmaların büyük bir kısmında veri setlerinin içerdiği metinlerde İngilizce kullanılmıştır. Buradan yola çıkarak geliştirilen bir çalışmada, Türkçe manzara veri metni kümesi (scene text recognition in Turkish, STRIT) oluşturulmuştur. Oluşturulan veri kümesi üzerinde iki taban çizgisi yöntemi ile Türkçe metin tanınması gerçekleştirilmiştir [2]. Araştırmacıların, görüntülerde sıklıkla kullandığı yöntemlerden biri Otsu modelidir. Görüntüde metin tespiti için yapılan çalışmada, Otsu modeli renkli görüntünün R, G, B kanallarına ayrı ayrı uygulanmıştır [3]. Benzer bir çalışmada, doğal görüntülerdeki metinlerin tespiti için OCR uygulanmadan önce k-means ve Otsu algoritmaları kullanılmış ve iki yöntemin sonuçları karşılaştırılmıştır [4]. Farklı bir çalışmada, Otsu yöntemini temel alan iyileştirilmiş bir eşik görüntü bölütleme algoritması geliştirilmiştir. Geliştirilen algoritmanın Otsu modeline ve diğer yöntemlere göre yüksek bölütleme hassasiyetine ve hızlı hesaplama hızına sahip olduğu gösterilmiştir [5]. Doğal görüntülerdeki metinler perspektif bozulması, gölgeler ve eğri karakter yerleşimi gibi nedenlerden dolayı düzensiz şekillere sahiptir. Bu durum görüntüdeki metin tespitini zorlaştırmaktadır. Bu sorunun üstesinden gelmek için yapılan bir çalışmada, düzensiz metinlere karşı dayanıklı bir tanıma modeli olan özel olarak tasarlanmış bir derin sinir ağı RARE (Otomatik Düzeltme özelliğine sahip sağlam metin tanıyıcı) önerilmiştir [6]. Görüntü üzerindeki gölge sorunun çözümü için yapılan çalışmada gölge tespit yöntemlerinden Bayes Sınıflandırma Yöntemi, Otsu modeli ve Histogram dağılımı yöntemi incelenerek görüntü seti üzerinde test edilmiştir ve başarımlı oranları sunulmuştur [7]. Doğal görüntü metni tanıma ile yapılan başka çalışmada maksimum kararlı ekstrem bölgeler (maximally stable extremal regions, MSER) kullanan bir metin algılama algoritması önerilmiştir. MSER'in görüntü bulanıklığına ilişkin hassasiyetinin üstesinden gelmek ve çok küçük harfleri bile algılamasını sağlamak için, Canny kenar tanıma filtresinin özelliklerinden yararlanan zenginleştirilmiş bir MSER geliştirilmiştir [8]. Benzer bir çalışmada, MATLAB uygulaması üzerinde MSER görüntü bölütleme kullanılarak doğal görüntü üzerinde metin tanımlaması yapılmıştır [9]. Farklı bir çalışmada, OCR'nin doğal görüntülerde tek bir karakteri tanımakta güçlük çektiği gösterilmiştir. Sorun, sınıflandırma yöntemleri gibi makine öğrenmesi yöntemleri kullanılarak çözülmeye çalışılmıştır. Sınıflandırma yöntemleri, doğal görüntülerdeki özellikleri çıkarmak için kullanılmıştır [10]. Gerçekleştirilen farklı bir çalışmada, OCR motoru ve genel nesne tanıma dayanan iki farklı sistemin doğal görüntü metinleri tanımadaki başarısı karşılaştırılmıştır. Genel nesne tanıma dayalı sistemin, OCR sisteminden daha iyi sonuçlar ortaya çıkardığı gözlemlenmiştir [11]. OCR ile ilgili bir derleme çalışmasında ise, metin tanımanın çeşitli aşamaları, metin türüne göre el yazısı OCR sistemlerinin sınıflandırılması, Çince ve Arapça metin tanıma çalışmaları ve OCR'de uygulamaya yönelik son araştırmalara değinilmiştir [12]. Yapılan başka bir çalışmada, parça tabanlı ağaç yapılı karakter tespiti kullanan yeni bir doğal görüntü metin tanıma yöntemi önerilmiştir. ICDAR 2003, ICDAR 2011, SVT (singular value thresholding) veri setlerine ilişkin deneysel sonuçlarda, önerilen yöntemin hem karakter tespitinde hem de kelime tanımadaki mevcut yöntemlerden daha iyi performans gösterdiği açıklanmıştır [13]. Diğer bir çalışmada, görme engellilere yardımcı olmak amacıyla doğal görüntü metinlerini tespit eden sistem geliştirilmiştir.

Sistem ICDAR 2003 Robust Reading Competition veri seti üzerinde test edilmiştir [14]. Farklı bir çalışmada, güç dağıtıcılarının yükünü azaltmak hafifletmek amacı ile elektrik faturalandırması için OCR tabanlı otomatik sayaç okuma (automatic meter reading, AMR) geliştirilmiştir [15].

Doğal görüntülerde ki metinler hayatın her alanında karşımıza çıkmaktadır. Kullandığımız ilahtan, okuduğumuz kitaba, gezdiğimiz sokağa kadar paylaştığımız her şey doğal görüntü formatındadır. Doğal görüntülerde ki metinler farklı birçok özelliğten dolayı karakter tanınmasında hala istenilen sonucu vermemektedir. Bu durum günümüzün güncel çalışma konularından biri olmasını sağlamaktadır. Yapılan farklı çalışmalarda yöntemler birbiri ile birlikte kullanılmış ya da karşılaştırılmıştır. Ancak hiçbir çalışmada Otsu Metodu ve MSER Yöntemi karşılaştırılmamıştır.

Bu çalışmada, ICDAR 2013 veri setinden seçilen 20 doğal görüntü üzerindeki metni tanımlamak için, Otsu metodu ya da MSER yöntemi kullanılarak görüntü bölütleme yapılması, sonrasında OCR kullanılarak metnin doğru tespit edilmeye çalışılması ve her iki yöntemin başarı oranlarının karşılaştırılması amaçlanmaktadır.

II. GÖRÜNTÜ BÖLÜTLEME VE OPTİK KARAKTER TANIMA

A. GÖRÜNTÜ BÖLÜTLEME

Görüntü bölütleme, bir görüntüyü daha iyi analiz edebilmek için görüntü içerisindeki bazı karakteristik özelliklere göre görüntüyü anlamlı bölgelere bölme işlemidir [16], [17]. Görüntü bölütleme, görüntü işlemenin ilk ve en önemli basamağıdır [18]. Görüntü üzerinde daha yüksek seviyeli görüntü işleme yöntemleri kullanılmadan önce görüntü bölütleme uygulanır. Farklı sebeplerden dolayı görüntüde gürültü ve bozulmalar meydana gelebilir. Gürültü ve bozulmaların giderilmesi, görüntünün daha iyi analiz edilebilmesi ve görüntüyü bölgelere bölme amacıyla görüntü bölütleme yöntemleri geliştirilmiştir. Geliştirilen yöntemler görüntüdeki pikseller arasındaki gri seviye değerlerinin süreksizlik (discontinuity) ve benzerlik (similarity) özelliklerine dayanır [16]. Literatürde görüntü bölütleme yöntemleri farklı şekillerde sınıflandırılmıştır [18]-[21]. Yaygın olarak kullanılan bölütleme yöntemleri: eşikleme tabanlı bölütleme, kenar tabanlı bölütleme, bölge tabanlı bölütleme ve çizgi tabanlı bölütlemedir [18]. Bölütleme yöntemleri genellikle hedef takibi, nesne sınıflama ve görüntü tanıma gibi alanlarda yaygın olarak kullanılmaktadır.

A. 1. Otsu Modeli

Otsu modeli, Japon bilim adamı Nobuyuki Otsu tarafından 1979 yılında geliştirilmiştir [22]. Eşikleme tabanlı bölütleme yöntemi sınıfında yer alan Otsu modelinde amaç görüntü içerisinde yer alan nesnelere görüntü arka planından ayırmaktır. Eşikleme görüntü piksel değerlerinin siyah-beyaz renge dönüştürülerek görüntünün en sade haline ulaşmasını sağlamaktadır. Eşik olarak kullanılacak değerin doğru tespit edilmesi gerekmektedir. Eğer eşik değeri yüksek veya küçük olursa görüntüde kayıplar meydana gelecektir [23]. Otsu modeli görüntü için optimum eşik değerinin belirlenmesini sağlamaktadır. Otsu modelinde eşik belirleme işlemi için öncelikle renklerin görüntü üzerinde var olma sayısı (histogramı) hesaplanır. Her bir eşik değeri görüntüyü ön plan ve arka plan olmak üzere iki sınıfa ayırır. Arka plan (background) için Eş. 1, ön plan (foreground) için Eş. 2 ile gösterilen ağırlık (weight) değerleri hesaplanır.

$$w_b(t) = \sum_{i=0}^t p(i) / \sum_{i=0}^L p(i) \quad (1)$$

$$w_f(t) = \sum_{i=t+1}^L p(i) / \sum_{i=0}^L p(i) \quad (2)$$

Her iki sınıf için Eş. 3 ve Eş. 4 ile gösterilen ortalama (mean) hesapları yapıldıktan sonra ve Eş. 5 ve Eş. 6 kullanılarak varyans değerleri hesaplanır.

$$\mu_b(t) = \sum_{i=0}^t ip(i) / \sum_{i=0}^t p(i) \quad (3)$$

$$\mu_f(t) = \sum_{i=t+1}^L ip(i) / \sum_{i=t+1}^L p(i) \quad (4)$$

$$\sigma_b^2(t) = \sum_{i=0}^t (i - \mu_b)^2 p(i) / \sum_{i=0}^t p(i) \quad (5)$$

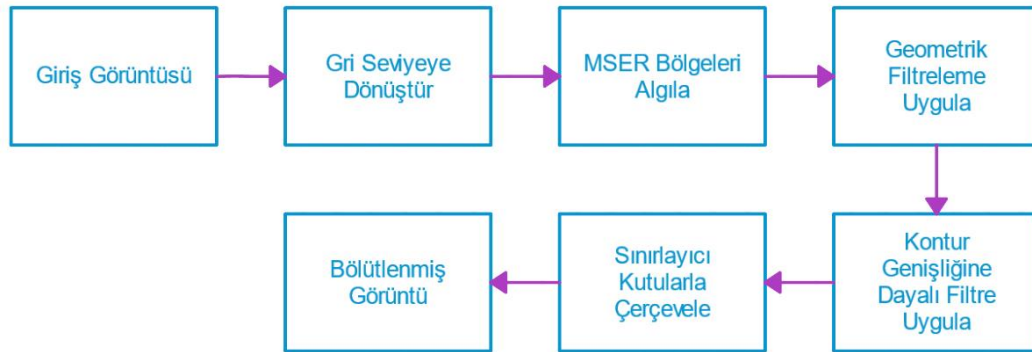
$$\sigma_f^2(t) = \sum_{i=t+1}^L (i - \mu_f)^2 p(i) / \sum_{i=t+1}^L p(i) \quad (6)$$

Bu değerlerden çıkan sonuçlardan Eş. 7 ile sınıf içi varyans değeri hesaplanır. Bütün eşik değerleri için hesaplanan sınıf içi varyanslar arasında en düşük sınıf içi varyans eşik değeri olarak kabul edilir. Belirlenen eşik değeri görüntüdeki her piksel ile karşılaştırılır. Eğer piksel eşik değerinden büyükse ön plan değilse arka plan olarak belirlenir [7], [18], [23].

$$\sigma_w^2(t) = w_b(t) \sigma_b^2(t) + w_f(t) \sigma_f^2(t) \quad (7)$$

A. 2. Maksimum Kararlı Ekstrem Bölgeler

Maksimum kararlı ekstrem bölgeler (MSER) bölgeleri algılamak için kullanılan bir yöntemdir. Matas ve arkadaşları tarafından geliştirilen MSER, görüntülerde kullanılan Blob Analizi yöntemlerinden biridir [24]. Temelleri Lindeberg tarafından atılan Blob Analizi (Blob Coloring) algoritması, görüntü içerisindeki farklı nesnelere tespit edilmesini veya tanınmasını sağlamaktadır [25]. Şekil 1'de MSER yönteminin blok diyagramı verilmiştir. MSER algoritması, karakter olmayan bölgeleri engelleyerek, karakter olduğu düşünülen bölgeleri sınırlayıcı kutu içerisine almaktadır.



Şekil 1. MSER Blok diyagramı

B. OPTİK KARAKTER TANIMA

Optik Karakter Tanıma (OCR), taranmış veya el yazısı ile yazılmış bir belgeyi metne dönüştürme işlemine dayanan bir görüntü tanıma tekniğidir [26]. Belgeye ya da görüntüye OCR tekniği uygulanmak isteniyor ise ilk yapılması gereken belgenin tarayıcı ya da kamera yardımı ile dijital hale dönüştürülmesidir. Belge dijital ortama aktarıldıktan sonraki aşamada, görüntü üzerinde konum bölütleme yapılarak nesne ve karakterlerin birbirinden ayrılması sağlanır. Görüntü bölütlemenin daha iyi sonuç vermesi için dijital ortama aktarılan belgenin kalitesinin iyi olmasına dikkat edilmesi gerekir. Dikkat edilmediği takdirde görüntü bölütlere ayrılırken zorluklar yaşanabilir. Sonraki aşamada görüntüye binarizasyon işlemi uygulanır. Binarizasyon, görüntünün gri tonlamalı görüntüye ardından

ikili görüntüye dönüştürülmesidir [27], [28]. Binarizasyon işlemi karakterlerin tanınmasını kolaylaştırmak için görüntü kalitesini arttıran bir işlemdir. Binarizasyon yöntemi aynı zamanda bölütleme yöntemlerinden biridir. İşlenmiş görüntüyü girdi olarak alan ve bu görüntüden her bir karakteri çıkaran karakter çıkarma işlemi uygulanır. Son aşamada ise çıkarılan her karakter tanıma işlemine tabi tutulur. Her bir karakter üzerinde sınıflandırma işlemi uygulanarak tanıma işlemi gerçekleştirilir [1]. Sınıflandırma, hangi sınıfa ait olduğu bilinmeyen bir örüntünün öznelikleri kullanılarak karar mekanizması yardımı ile o örüntünün hangi sınıfa ait olduğunun belirlenmesidir. Sınıflandırma işlemi için birçok makine öğrenimi yaklaşımı vardır. Makine öğrenimi, insan müdahalesi veya bir talimat olmadan verilerden ve deneyimlerden otomatik olarak öğrenebilen bilgisayar modelleri ve algoritmalar olarak tanımlanabilir [29]. Makine öğrenimi yaklaşımlarından Olasılıksal Sinir Ağı (probabilistic neural network, PNN), destek vektör makineleri (support vector machines, SVM) ve K-en yakın komşu (K- nearest neighbor, K-NN) yaklaşımları öne çıkmaktadır [30]. Bektaş ve arkadaşları karakter tanıma ile ilgili yaptıkları çalışmada üç farklı makine öğrenimi yaklaşımı olan Naive Bayes, K-NN ve Lineer Çekirdek Fonksiyonunu (linear kernel SVM, LibSVM) test ederek karşılaştırmışlardır. Yapılan çalışma ile Naive Bayes %65, LibSVM %81 ve K-NN %89 doğruluk oranlarını elde etmişlerdir [31]. OCR taranan belgelerde istenen doğruluk oranı performansına yaklaşırken, doğal görüntü metinlerinde ise istenen performansı henüz gösterememiştir. Bu sebeple, doğal görüntülerdeki metni daha doğru tespit etmek için farklı çalışmalar yapılmaktadır.

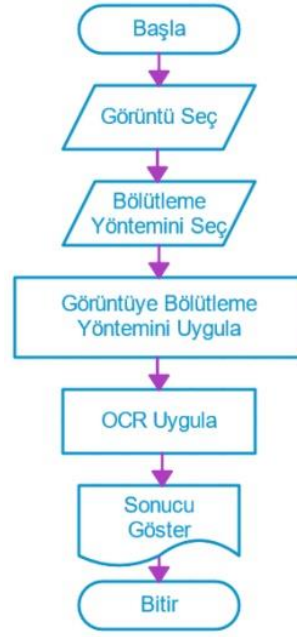
III. MATERYAL VE METOT

Gerçekleştirilen çalışmada Şekil2’de gösterildiği gibi ICDAR 2013 veri setinden 20 görüntü kullanılmıştır. Görüntüler seçilirken her iki yöntemde sonuçlar üretmesi dikkate alınmıştır. Dijital ortamda olan bu görüntüler Şekil3’te verilen akış diyagramı ile MATLAB geliştirme ortamında gerçekleştirilen kod bloğu ile tespit edilmeye çalışılmıştır. Akış diyagramında da gösterildiği gibi kullanıcı görüntüyü ve uygulamak istediği bölütleme yöntemini seçtikten sonra metin bulma işlemi MATLAB içerisinde yer alan hazır OCR fonksiyonu ile gerçekleştirilmiştir.



Şekil 2. 20 adet doğal görüntü

Kullanıcılara kullanım kolaylığı sağlaması için grafik ara yüzü geliştirilmiştir. Kullanıcı Şekil 4’de gösterildiği gibi ara yüz ile istediği görüntüyü ve bölütleme yöntemini seçerek metin tanıma işlemi gerçekleştirebilmektedir. Görüntüde var olan ve Otsu modeli ile MSER yönteminin bölütlediği görüntülerin OCR ile tespit ettiği metinler ara yüzde gösterilmiştir.

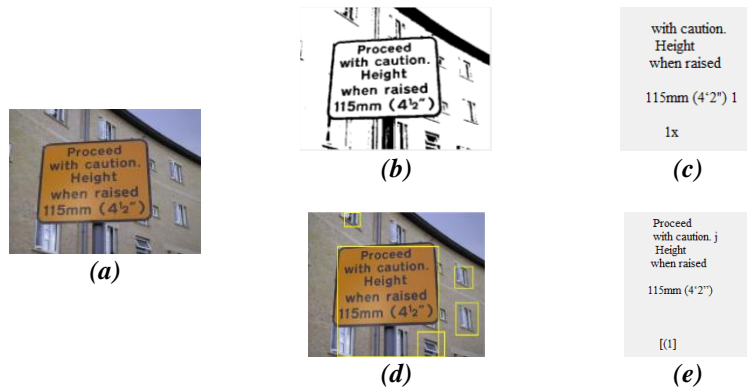


Şekil 3. Akış Diyagramı



Şekil 4. Grafik ara yüzü

Şekil 5 (a)'da gösterilen doğal görüntüye her iki bölütleme yöntemi de ayrı ayrı uygulanmıştır. Şekil 5 (b) 'de Otsu modelinin uygulanmış hali, Şekil 5 (c)'de ise Otsu uygulandıktan sonra OCR uygulanması ile elde edilen çıktı gösterilmiştir.



Şekil 5. (a) Orijinal img4, (b) Otsu uygulanmış hali, (c) Otsu uygulanmış OCR çıktısı, (d) MSER uygulanmış hali, (e) MSER uygulanmış OCR çıktısı

MSER yönteminde Otsu modelinde olduğu gibi öncelikle görüntü gri tonlamalı seviye dönüştürülmüştür. Sonraki aşamada görüntüye MSER tekniği uygulanmıştır. MSER tekniğinin kullanımı sadece metin bölgelerini değil, aynı zamanda metin olmayan bölgeleri de algılayacaktır. Bu sorunun üstesinden gelmek için MSER iki yöntem ile birlikte uygulanmıştır. Kullanılan ilk yöntem geometrik filtrelemedir. Geometrik filtrenin Aspect ratio, eccentricity, solidity, extent ve Euler number özellikleri kullanılarak metin olmayan bölgeler azaltılmaya çalışılmıştır. Metin olmayan bölgeleri azaltmak için kullanılan diğer bir yöntem ise Kontur genişliğine dayalı filtrelemedir. Kontur genişliğine dayalı filtreleme, görüntüdeki filtrelenmemiş metin olmayan bölgeleri azaltmıştır. Filtreleme işlemlerinden sonra kalan bölgeler, bölgeleri çevreleyen sınırlayıcı kutularla çerçevelenmiştir. Sınırlayıcı kutular birbiriyle örtüşecek şekilde büyütülmüştür. Birbirine yakın sınırlayıcı kutuların örtüşme oranı hesaplanarak, yan yana iki veya daha fazla sınırlayıcı kutu birleştirilmiştir. MSER yöntemi kullanıldıktan sonra, karakterleri tanımak için bölütlere ayrılmış görüntüye bir metin tanıma tekniği olan OCR uygulanmıştır. OCR tekniği algılanan ve sınırlayıcı kutularla çerçevelenen görüntü bölgeleri içerisinde tanınan karakterlerden oluşan bir metin üretmiştir. Şekil 5 (d) ve (e)'de MSER yönteminin uygulanmış hali ve sonrasında OCR ile tespit edilen çıktı gösterilmiştir.

IV. BULGULAR VE DEĞERLENDİRME

ICDAR 2013 veri setinden 20 adet doğal görüntü üzerinde Otsu metodu ve MSER yöntemi kullanılarak metin tanıma işlemi gerçekleştirilmiştir. Her yöntem için her görüntünün ayrı ayrı doğru karakter sayısı yüzdelik olarak hesaplanmıştır. OCR metin tanıma işlemi yaparken alt satıra inme durumunda da karakter üretmektedir. Doğru karakter sayısı yüzdelik olarak hesaplanırken alt satıra inme karakterleri hesaplamaya dahil edilmemiştir. Aynı zamanda doğruluk oranları hesaplanırken kelimelerin sıralı olarak bulunması dikkate alınmamıştır. Tablo 1'de her görüntüdeki Otsu ve MSER yöntemleri uygulandıktan sonra OCR ile bulunan toplam karakter sayıları, doğru karakter sayıları ve doğruluk oranları verilmiştir.

Tablo 1. Görüntülerdeki Otsu ve MSER toplam karakter sayıları, doğru karakter sayıları ve doğruluk oranları

Görüntü	Otsu Toplam Karakter Sayısı	Otsu Doğru Karakter Sayısı	Otsu Doğruluk Oranı (%)	MSER Toplam Karakter Sayısı	MSER Doğru Karakter Sayısı	MSER Doğruluk Oranı (%)
img1	35	23	65,71	40	30	75,00
img2	70	30	42,86	41	29	70,73
İmg3	51	36	70,59	46	43	93,48
İmg4	70	42	60,00	56	48	85,71
img5	39	28	71,79	27	27	100,00
img6	33	14	42,42	20	16	80,00
img7	38	23	60,53	35	24	68,57
img8	38	37	97,37	37	36	97,30
img9	18	16	88,89	14	13	92,86
img10	31	24	77,42	23	22	95,65
img11	32	10	31,25	27	25	92,59
img12	37	35	94,59	36	36	100,00
img13	19	8	42,11	20	17	85,00
img14	10	7	70,00	8	8	100,00
img15	29	28	96,55	28	28	100,00
img16	26	13	50,00	13	13	100,00
img17	39	29	74,36	32	29	90,63
img18	18	15	83,33	16	15	93,75
img19	44	21	47,73	62	39	62,90
img20	33	19	57,58	26	18	69,23

Tablo 2’de ICDAR 2013 veri setinden alınan 20 doğal görüntüde Otsu modeli kullanılan metin tanıma ortalama doğruluk oranı %66,25 iken MSER yöntemi kullanılan metin tanıma ortalama doğruluk oranı %87,67 olarak bulunmuştur. Sonuçlardan yola çıkılarak MSER yönteminin Otsu modeline göre doğal görüntü metin tespitinde daha başarılı olduğu sonucuna ulaşılmıştır.

Tablo 2. Görüntülerdeki Otsu ve MSER ortalama doğruluk oranları

Yöntem	Ortalama Doğruluk Oranı (%)
Otsu	66,25
MSER	87,67

V. SONUC

Doğal görüntülerde metin tanıma, bilgisayar bilimlerinde çalışılan önemli konulardan biridir. Yapılan çalışmalar incelendiğinde, değerlendirme yöntemlerine ve kullanılan veri setlerine bakıldığında, doğal görüntülerde metin tanıma için bölütleme yöntemlerinin iyi sonuçlar verdiği ve üzerinde çalışılarak geliştirilebileceği gözlemlenmiştir. İyi bir performans veren ancak mükemmel olan bir sistem hala geliştirilememiştir. Bu sebepten pek çok araştırmacı, bu konu üzerinde çalışmakta ve oluşturdukları sistemlerde en iyi performansı elde etmeyi amaçlamaktadır. İlerleyen çalışmalarda üzerinde çalışılan veri kümesinin genişletilmesi ve başarı oranının artırılması hedeflenmektedir.

VI. KAYNAKLAR

- [1] L. Eikvil. (2022, February 26). *Optical character recognition* [Online]. Available: <http://home.nr.no/~eikvil/OCR.pdf>.
- [2] N. Erdoğan, “Türkçe manzara metni veri kümesi,” 25th Signal Processing and Communications Applications Conference’ında sunuldu, Antalya, 2017.
- [3] M. Thodaskar, and R. Devi, “Segmentation and detection of text in natural scene images,” *International Journal of Engineering Research & Technology (IJERT)*, vol. 4, no. 6, pp. 1272-1277, 2015.
- [4] O. Granlund, and K. Böhrnsen. (2022, March 5). *Improving character recognition by thresholding natural images*, [Online]. Available: <https://www.diva-portal.org/smash/get/diva2:1108666/FULLTEXT01.pdf>.
- [5] M. Huang, W. Yu, and D. Zhu, “An improved image segmentation algorithm based on the Otsu method,” in *13th ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing*, Japan, 2012, pp. 135-139.
- [6] B. Shi, X. Wang, P. Lyu, C. Yao, and X. Bai, “Robust scene text recognition with automatic rectification,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 4168-4176.
- [7] B. Kır Savaş, S. İlkin, S. Hangişi, ve S. Şahin, “Gölge tespitinde kullanılan Bayes sınıflandırma, Otsu bölütleme ve histogram dağılımı yöntemlerinin karşılaştırılması,” *Düzce Üniversitesi Bilim ve Teknoloji Dergisi*, c. 5, s. 2, ss. 345-355, 2016.

- [8] H. Chen, S. S. Tsai, G. Schroth, D. M. Chen, R. Grzeszczuk, and B. Girod, "Robust text detection in natural images with edge-enhanced maximally stable extremal regions," in *2011 18th IEEE International Conference on Image Processing*, Brussels, Belgium, 2011, pp. 2609-2612.
- [9] G. Chakraborty, S. Panda, and S. Roy. (2022, March 9). *Text extraction from image using MATLAB* [Online]. Available: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3525969
- [10] T. E. De Campos, B. R. Babu, and M. Varma (2009), "Character recognition in natural images," in *Proceedings of the International Conference on Computer Vision Theory and Applications*, 2009, 273-280.
- [11] K. Wang, B. Babenko, and S. Belongie, "End-to-end scene text recognition," in *2011 International Conference on Computer Vision*, IEEE, Barcelona, Spain, 2011, pp. 1457-1464.
- [12] K. Karthick, K. B. Ravindrakumar, R. Francis, and S. Ilankannan, "Steps involved in text recognition and recent research in OCR; a study," *International Journal of Recent Technology and Engineering*, vol. 8, no. 1, pp. 3095-3100, 2019.
- [13] C. Shi, C. Wang, B. Xiao, Y. Zhang, S. Gao, and Z. Zhang, "Scene text recognition using part-based tree-structured character detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, USA, 2013, pp. 2961-2968.
- [14] N. Ezaki, M. Bulacu, and L. Schomaker, "Text detection from natural scene images: towards a system for visually impaired persons," in *Proceedings of the 17th International Conference on Pattern Recognition (ICPR)*, IEEE, Cambridge, England, 2004, pp. 683-686.
- [15] K. Karthick, and S. Chitra. "Novel method for energy consumption billing using optical character recognition," *Energy Engineering*, vol. 114, no. 3, pp. 64-76, 2017.
- [16] R. C. Gonzalez, and E. R. Woods, *Digital Image Processing*, 3rd ed., New Jersey, USA: Pearson Education, 2008.
- [17] T. Asano, D.Z. Chen, N. Katoh, and T. Tokuyama, "Polynomial-time solutions to image segmentation," in *Proceedings of the Seventh Annual ACM-SIAM Symposium on Discrete Algorithms*, Atlanta, Georgia, 1996, pp. 104-113.
- [18] Y. P. Zhu, and P. Li, "Survey on the image segmentation algorithms," in *Proceedings of the International Field Exploration and Development Conference 2017*, Singapore, 2019, pp. 475-488.
- [19] M. Sridevi and C. Mala, "A survey on monochrome image segmentation methods," *Procedia Technology*, vol. 6, pp. 548-555, 2012.
- [20] N. R. Pal, and S. K. Pal, "A review on image segmentation techniques," *Pattern Recognition*, vol. 26, no. 9, pp. 1277-1294, 1993.
- [21] M. W. Khan, "A survey: image segmentation techniques," *International Journal of Future Computer and Communication*, vol. 3, no. 2, pp. 89-93, 2014.
- [22] N. Otsu, "A threshold selection method from gray-level histogram," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62-66, 1979.
- [23] A. Çelik, ve S. Demirel, "Otsu ve Ridler-Calvard görüntü işleme yöntemlerinin zatürre tespitinde kullanılması," *Muş Alparslan Üniversitesi Fen Bilimleri Dergisi*, c. 10, s. 1, ss. 917-923, 2022.

- [24] J. Matas, O. Chum, M. Urban, and T. Pajdla, "Robust wide-baseline stereo from maximally stable extremal regions," *Image and Vision Computing*, vol. 22, no. 10, pp. 761-767, 2004.
- [25] T. Lindeberg, "Detecting salient blob-like image structures and their scales with a scale-space primal sketch: A method for focus-of-attention," *International Journal of Computer Vision*, vol. 11, no. 3, pp. 283-318, 1993.
- [26] Y. Alginahi, *Character Recognition*, 1st ed., Rijeka, Croatia: InTech, 2010, ch. 1, pp. 1-19.
- [27] N. H. Barnouti, M. Abomaali, and M. H. N. Al-Mayyahi, "An efficient character recognition technique using K-nearest neighbor classifier," *International Journal of Engineering & Technology*, vol. 7, no. 4, pp. 3148-3153, 2018.
- [28] P. M. Manwatkar, and K. R. Singh, "A technical review on text recognition from image," in *2015 IEEE 9th International Conference on Intelligent Systems and Control (ISCO)*, IEEE, 2015, pp. 1-5.
- [29] M. A. Luján, M. V. Jimeno, J. Mateo Sotos, J. J. Ricarte, and A. L. Borja, "A survey on EEG signal processing techniques and machine learning: applications to the neurofeedback of autobiographical memory deficits in schizophrenia," *Electronics*, vol. 10, pp. 3037-3055, 2021.
- [30] R. Mittal, and A. Garg. "Text extraction using OCR: a systematic review," in *2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA)*, IEEE, 2020, pp. 357-362.
- [31] B. Bektaş, S. Babur, U. Turhal, ve E. Köse, "Makine öğrenmesi yardımıyla optik karakter tanıma sistemi," *5. Uluslararası Matbaa Teknolojileri Sempozyumu*, 2016, ss. 487-494.