



SÜPERPİKSEL TABANLI SATIR BÖLÜTLEME

Ali Alper DEMİR, Ufuk ÖZKAYA*

Süleyman Demirel Üniversitesi, Mühendislik Fakültesi, Elektrik-Elektronik Mühendisliği Bölümü, Isparta,
Türkiye

Anahtar Kelimeler

*Süperpiksel,
Satır Bölütleme,
El Yazımı Belge,
Doküman İmgesi Analizi,
Havza Bölütleme.*

Öz

Satır bölütleme tarihi doküman analizi uygulamalarının en temel aşamalarından birisidir. Satır bölütleme başarısı, daha sonraki doküman analizi yöntemlerinin başarısını doğrudan etkilemektedir. Matbu belgelerde kayıpsız satır bölütleme işlemi kolaylıkla yapılabilmektedir. Ancak, el yazımı belgeler için satır bölütleme işlemi metin satırlarının eğik, eğri, dalgalı olması, satırlar arası boşlukların darlığı, örtüşen ve temas eden bileşenlerden dolayı hala zorlayıcı bir problemdir. Bu çalışmada, el yazımı dokümanlar için süperpiksel tabanlı yeni bir satır bölütleme yöntemi önerilmiştir. Yöntem ardışık satırları bölütleyebilen en güvenli sınırın elde edilmesini hedeflemektedir. Önerilen yöntem 853 adet Çince el yazımı doküman imgesi içeren HIT-MW veri seti üzerinde uygulanmıştır. Veri setinin en önemli özelliği eğik, temas eden ve örtüşen satır davranışlarına sahip imgelerden oluşmasıdır. Önerilen yöntem ile % 98.03 tespit oranı, % 97.66 tanıma doğruluğu elde edilmiş ve yöntemin başarısı literatürde bulunan diğer yöntemlerle karşılaştırılmıştır. Elde edilen sonuçlar ışığında önerilen yöntemin el yazımı metinlerde satır bölütleme uygulamalarındaki başarısı ve potansiyeli ortaya konmuştur.

SUPERPIXEL BASED TEXT LINE SEGMENTATION

Keywords

*Superpixel,
Text Line Segmentation,
Handwritten Document,
Document Image Analysis,
Watershed Segmentation.*

Abstract

Text line segmentation is one of the essential stages of historical document analysis applications. The accuracy of text line segmentation affects directly the success of following document analysis steps. For printed documents, lossless text line segmentation can be done readily. But, for handwritten documents, unfortunately it is still a challenging problem because of the skewed, curved, fluctuated text lines, narrow gaps between the text lines, overlapping and touching components. In this paper, a novel superpixel-based text line segmentation method for handwritten documents is proposed. This method aims to extract the most reliable boundary to segment consecutive text lines. This method is implemented HIT-MW dataset containing 853 Chinese handwritten document images. The most important feature of this dataset is to be composed of documents having skewed, overlapping and touching text lines. A detection rate of 98.03% and a recognition accuracy of 97.66% is obtained and these results are compared with the ones of existing state of the art methods. With these results, segmentation success and potential of our method for handwriting text line segmentation is pointed out.

Alıntı / Cite

Demir, A. A., Özkaya, U., (2019). Süperpiksel Tabanlı Satır Bölütleme, Mühendislik Bilimleri ve Tasarım Dergisi, 7(4), 854-868.

Yazar Kimliği / Author ID (ORCID Number)

A. A. Demir, 0000-0001-5250-0590
U. Özkaya, 0000-0002-3520-1975

Makale Süreci / Article Process

Başvuru Tarihi / Submission Date	31.01.2019
Revizyon Tarihi / Revision Date	08.06.2019
Kabul Tarihi / Accepted Date	16.06.2019
Yayın Tarihi / Published Date	19.12.2019

* İlgili yazar / Corresponding author: ufukozkaya@sdu.edu.tr, +90-246-211-1910

1. Giriş

Dünyanın birçok yerinden tarihçiler ve araştırmacılar tarihi belgeler üzerinde araştırmalar, çalışmalar yapmaktadırlar. Arşivlerde çok sayıda el yazımı ile yazılmış tarihi doküman bulunmakta ancak bu dokümanların indekslenmesi ve çevirisi insan gücüyle yapılan ve oldukça zaman alan bir işlemdir. Bu işlemlerin otomatik bir şekilde yapılmasının zorluğundan dolayı bu tip dokümanlar araştırmacılar tarafından etkili bir şekilde kullanılamamaktadır. Zaman geçtikçe tarihi belgeler yıpranarak eskimektedir. Bu durum ise belgelerin okunmasını zorlaştırmaktadır. Bu problemin önüne geçmek amacıyla son yıllarda el yazımı ve tarihi dokümanlar taranarak imge formatında dijital ortama aktarılmaktadır. Dijital ortama aktarılan bu belgelerdeki kelimeleri otomatik bir şekilde teker teker ayrıştırmak bu belgelere hızlı ve kolay bir şekilde erişilerek araştırmacılar için büyük bir kolaylık sağlayacaktır. Böylece araştırmacılar istedikleri anahtar kelimeleri doküman imgeleri üzerinde aratarak belgelerdeki ilgili kısmı bularak araştırmalarını hızlandıracaktır.

Günümüzde özellikle tarihi ve el yazımı metinlerin taranıp indekslenmesi ayrıca bu belgelerin transkripsiyonlarının yapılması oldukça güncel bir problemidir. Optik karakter tanıma, doküman imgelerinin bilgisayar ortamında düzenlenebilecek şekilde sayısal bir metine çevrilmesi işlemidir. Optik karakter tanıma sistemlerinde genellikle sırasıyla ön işleme, bölütleme, öznitelik çıkarımı, eğitim ve tanıma şeklinde aşamalar bulunmaktadır. Ön işleme, bölütleme için kolay veri üretmeyi amaçlamaktadır. Ön işleme aşamasında gürültü azaltma, verilerin normalleştirilmesi, renk dönüşümü, histogram eşitleme gibi işlemler bulunmaktadır. Bölütleme, doküman imgesinin paragraflara, metin satırlarına, cümlelere, kelimelere veya harfler gibi daha küçük parçalara ayrılması işlemidir. Öznitelik çıkarma aşamasında, tanıma aşamasında metni tanıyabilmek için doküman imgesinden en alakalı bilgileri çıkarma işlemi yapılmaktadır. Bu bilgiler her bir sembolü diğer sembollerden ayırmak için gerekli öznitelikleri içerir. Tanıma aşaması, karakter tanıma sisteminde son ve ana karar verme aşamasıdır. Bilinmeyen her bir sembolü tanımlayan ve daha önceden tanımlanmış bir sınıfa atayan bir sınıflandırma işlemi yapılmaktadır. Bu sınıflandırma, bir önceki aşamanın çıktısı olan özniteliklere dayanmaktadır.

Doküman analizi uygulamalarından biri olan optik karakter tanıma sistemlerinde bölütleme aşaması en

önemli aşamalardan biridir. Satır bölütleme bu işlemler arasında ilk sıralarda yer aldığı için hataların minimize edilip metin satırlarının yüksek doğrulukta bölütlenmesi gerekmektedir. Böylece devam eden diğer işlemlere daha doğru bir giriş verisi sağlanmış olur. Metin satırı bölütlenmesi işleminin karakterin tanıma başarısı üzerinde doğrudan bir etkisi vardır. Yazarların yazma şekillerindeki farklılıklardan dolayı karakter tanıma sistemindeki hataların çoğu bölütleme aşamasında meydana gelmektedir. Bu nedenle daha iyi bir karakter tanıma sisteminin oluşturulabilmesi için bölütleme aşamasında doküman imgelerinin satırlara bölütlenmesi işleminin yüksek doğrulukta yapılması gerekmektedir.

Matbu doküman imgelerinde satırlar arası örtüşme olmadığından metni oluşturan satırlar kayıpsız bir şekilde kolaylıkla elde edilebilmektedir. Ancak el yazımı doküman imgelerinde satırlar düz bir şekilde gitmemekte; eğiklikler, dalgalanmalar, satırlar arası örtüşmeler, iki komşu satırın birbirine temas etmesi gibi durumlar olabilmektedir (Şekil 1). Bu nedenle, standart satır bölütleme yöntemleri el yazımı doküman imgelerinde başarılı sonuçlar elde edememektedir.

Literatürde el yazımı dokümanların satırlarının bölütlenmesi için birçok algoritma bulunmaktadır. Bunlardan en yaygın kullanılanları iz düşüm tabanlı yöntemler (Ataer ve Duygulu, 2006; Adıgüzel vd., 2012), hough dönüşümü tabanlı yöntemler (Likforman-Sulem vd., 1995; Louloudis vd., 2006; Louloudis vd., 2008) ve bağlı bileşen tabanlı (Ryu vd., 2014) yöntemlerdir. İzdüşüm tabanlı yöntemler genellikle matbu doküman imgeleri üzerinde uygulanmakta ve bu tip dokümanlarda başarılı sonuçlar elde etmektedir. İzdüşüm tabanlı yöntemler bir metin satırının düz bir şekilde olduğu varsayımına dayanmakta, komşu metin satırları arasındaki boşluk bilgisinden faydalanarak bölütleme işlemini gerçekleştirmektedir. İzdüşüm tabanlı yöntemler doküman imgesinin dikey hatlara bölünüp metin satırlarının o dikey hat boyunca da düz bir şekilde ilerlediği varsayımını kullanarak çok fazla eğikliğin bulunmadığı doküman imgelerinde de kullanılabilir. Ancak izdüşüm tabanlı yöntemler kısıtlı sayıda dokümana uygulanabilmektedir. İzdüşüm tabanlı yöntemler el yazımı metinlerde bölütlemeyi zorlaştıran eğimli ve temas eden satırlar nedeniyle başarılı sonuçlar vermemektedir.

yazılarından oluşan toplam 11581 metin satırı içeren 720 doküman üzerinde test edilen yöntem % 97.31 başarı oranıyla metin satırlarını bölütlemektedir. 78.902 bağlı bileşen içeren 200'den fazla el yazımı imge üzerinde test edilen yöntem metin satırlarının % 97.65'ini doğru bir şekilde tespit edebilmiştir.

Su vd. (2007), el yazımı ile yazılmış Çince dokümanların eğikliğini tespit etmek için yatay vuruş histogramını kullanmışlardır. Vuruş eğikliği düzeltme algoritması yatay vuruş histogramından satırların eğiklik açısını bulur ve eğikliği düzelttikten sonra izdüşüm profilleri kullanarak metin satırlarını bölütler. Ayrıca bu çalışma ile Çince el yazımı tanıma alanındaki ilk veri seti olan HIT-MW veri seti oluşturulmuştur. 780'den fazla katılımcı tarafından yazılan 853 adet doküman imgesi bulunmaktadır. Veri setinde kesin referans olarak 8673 adet metin satırı bulunmaktadır. Basit bir hesapla her bir doküman imgesinde ortalama 10.16 metin satırı bulunmaktadır.

Li vd. (2008), satırların tespit edilmesi işlemini bir görüntü bölütleme problemi olarak ele almışlardır. Yoğunluk kestirimine ve seviye kümesi bölütleme yöntemine dayanan bir yaklaşım önerilmiştir. İkili imge anizotropik (yönbağımlı) bir Gauss çekirdeği kullanılarak gri seviyeli imgeye dönüştürülmektedir. Doküman imgesi üzerinden her bir elemanın bir metin satırına ait olma ihtimalini temsil ettiği olasılık haritası hesaplanmaktadır. Daha sonra, başlangıç tahmini geliştirilerek komşu metin satırlarının sınırlarını belirlemek için seviye kümesi yöntemi kullanılmaktadır. Yöntem Arapça, Korece, Çince ve Hintçe gibi çeşitli harfler içeren serbest el yazımı dokümanlar üzerindeki test edildiğinde iyi bir performans göstermektedir. Yapılan farklı deneyler önerilen algoritmanın ölçek değişikliği, dönüş ve gürültüye karşı dayanıklı olduğunu göstermektedir. Önerilen yöntem toplam 100 adet Arapça el yazımı doküman imgesi içeren bir veri seti üzerinde test edildiğinde % 85.6 başarı oranıyla 2691 satırın 2303'ü tespit edilmiştir. 100 Çince el yazımı doküman imgesinden oluşan başka bir veri seti üzerinde uygulandığında ise önerilen yöntem 1672 satırın 1532'sini bularak % 92 oranında bir başarı elde etmiştir. Çalışmada iki satırın birbirine önemli ölçüde temas etmesinden dolayı büyük hataların olduğu belirtilmektedir.

Louloudis vd. (2008), üç farklı adımdan oluşan bir yöntem önermişlerdir. İlk adım imge ikilileştirmesi ve iyileştirmesini, bağlı bileşen çıkarımını, bağlı bileşenlerin üç ayrı alt kümeye ayrıştırılmasını ve ortalama karakter yüksekliğinin hesaplanmasını içermektedir. İkinci adımda potansiyel metin satırlarının tespiti için blok tabanlı Hough dönüşüm yöntemi kullanılmaktadır. Üçüncü adım ise bir önceki adımda ortaya çıkarılamayan metin satırlarını tespit etmek için, olası bölünmeleri düzeltmek ve son olarak dikey olarak bağlı karakterleri ayırmak ve onları ilgili metin satırlarına atamak için kullanılmaktadır.

Önerilen yöntem Atina Üniversitesi'nin tarihi arşivlerinden, Dallas kütüphanesinden ve ICDAR2007 el yazımı bölütleme yarışmasından alınan İngilizce, Yunanca, Fransızca ve Almanca el yazımından oluşan 152 adet doküman üzerinde test edilmiştir. Tüm doküman imgeleri için manuel olarak metin satırı tespiti için kesin referans oluşturulmuştur. Oluşturulan veri seti toplam 3382 metin satırı içermektedir. Önerilen yöntem % 95.8 tespit oranı ve % 93.8 tanıma doğruluğu başarıları elde etmiştir.

Du vd. (2009), Mumford-Shah modeline dayanan bir yöntem tanımlamışlardır. El yazımı metin satırlarını bölütlemek için Mumford-Shah modeli ve düzey kümesi yöntemi kullanılmıştır. Metin satırlarını bölütlemek için öncelikle pencere boyutu yatay yönde geniş olan Gauss filtresi kullanılarak doküman imgesi bulanıklaştırılmaktadır. Düzey kümesi fonksiyonu iç, dış ve metin satırı bölgelerinde farklı değerlere sahiptir. Yöntemde Mumford-Shah modelinin enerji fonksiyonunun minimize edilmesi, böylece imgenin iyi bir şekilde bölgelere bölütlenmesi gereklidir. Son olarak metin satırlarını ayırtmak ve bazı gürültüleri gidermek amacıyla aşındırma ve genişleme gibi bazı morfolojik işlemler yapılmaktadır. Algoritma dilden bağımsızdır. Komşu metin satırları arasındaki örtüşmeleri kaldırmak ve kırık olanları bağlamak için işlem sonrası adım olarak morfolojik işlemler kullanılmaktadır. Yöntem farklı eğiklik açıları durumunda metin satırlarını doğru bir şekilde elde edememektedir.

Yin ve Liu (2009), kümeleme tabanlı bir yöntem olan asgari tarama ağacı yöntemini kullanmışlardır. Bu yöntem çoklu eğiklik durumu içeren satırları bölütleyebilirken, temas eden metin satırlarında başarısızdır. Bağlı bileşenler ilk olarak birkaç geometrik kısıtlamaya dayalı olarak bölünür, daha sonra gözetimli öğrenme tarafından tasarlanan uzaklık metriği ile asgari tarama ağacı aracılığıyla bir ağaç yapısına gruplanır. Seçilen kenarlar dinamik olarak ağaçtan kesilerek metin satırları elde edilir. Yapay parametrelerden kaçınmak ve bölütleme doğruluğunu artırmak için, gözetimli öğrenme aracılığıyla uzaklık metriği tasarlanmıştır. Çince el yazımı dokümanlar içeren HIT-MW veri setinden rastgele seçilen ve 508 metin satırı içeren 50 adet doküman imgesi kullanılarak uzaklık metriği öğrenimi ve parametre ayarlaması yapılmıştır. Daha sonra önerilen yöntem veri setindeki 8169 metin satırı içeren 803 adet doküman imgesi üzerinde test edilmiş ve % 98.02 tespit oranı ve % 2.47 hata oranı elde edilmiştir.

Ziaratban ve Faez (2010), doküman imgesini uyarlanabilir bloklara bölen aşağıdan yukarıya gruplama algoritması kullanmışlardır. Her bir bloğun eğikliği hesaplanmıştır. Yöntemi farklı yazarlara uyarlamak için üç parametre tanımlanmıştır. HIT-MW veri seti üzerinde uygulanan yöntem % 98.34 oranında tespit oranı başarıları elde etmiştir. 282 kişi

tarafından yazılan 7186 yazar içeren 1129 adet Farsça doküman imgesi içeren FHT veri seti üzerinde test edildiğinde ise % 91.22 tespit oranı elde edilmiştir. FHT veri setinde bulunan zor metin satırlarını içeren 371 imgelik bir grup imge üzerindeki tespit oranı ise % 86.26 olarak elde edilmiştir.

Alaei vd. (2011), doküman imgesinin ön plan bölümünü lekelemek için yeni bir boyama tekniği kullanmışlardır. Öncelikle imgedeki bileşenlerin ortalama genişlikleri hesaba katılarak imge dikey hatlara bölünmektedir. Boyama tekniği, metin satırlarının kolayca algılanmasını sağlayan ön ve arka plan bölümlerin ayrılabilmesini kolaylaştırmaktadır. Her bir metin satırı için tek bir bileşen elde etmek için boyanmış imgenin ön plan bölümünde bir genişleme işlemi uygulanmıştır. Aday çizgi ayırıcıları adı verilen bir dizi ayırıcı çizgi elde etmek için genişletilmiş imgenin arka bölümünün inceltilmesi ve daha sonra bazı düzeltme işlemleri gerçekleştirilmiştir. Aday çizgi ayırıcılarının başlangıç ve bitiş noktaları kullanılarak ve aralarındaki uzaklıklar analiz edilerek bölütlenmiş metin satırlarının elde edilmesi için ilgili aday çizgi ayırıcıları birleştirilir. Önerilen satır bölütleme yöntemi üç farklı veri setine uygulanmıştır. İlk olarak 152 adet İngilizce, Yunanca, Fransızca ve Almanca el yazımı dokümanlardan oluşup toplam 3382 satır içeren veri seti kullanılmıştır. Dokümanlar Atina Üniversitesi, Dallas kütüphanesi ve ICDAR2007 el yazımı bölütleme yarışmasından alınmıştır. Önerilen yöntem daha sonra toplam 4034 satır içeren 200 el yazımı doküman imgesinden oluşan ICDAR2009 el yazımı bölütleme yarışmasında kullanılan veri setine uygulanmıştır. Yöntem son olarak 78 el yazımı metin sayfasından oluşan ve toplam 1044 satır içeren Bangla, Oriya dili ve Farsça yazılardan oluşan gri seviyede taranmış doküman imgelerine uygulanmıştır. İlk veri setinde ortalama doğru tespit edilen satır oranı % 95.17 iken, ikinci veri setinde bu oran % 98.35 olarak elde edilmiştir. Son veri setinde ise ortalama doğru tespit edilen satır oranı % 95.87'dir. Önerilen yöntem -10 ve +10 eğiklik aralığında iyi çalışmaktayken -20 ve +20 arası eğiklik durumunda hatalı sonuçlar vermektedir. Ayrıca karmaşık ve düzensiz arka plan olduğu durumlarda yöntem hatalı ikileme nedeniyle başarısızdır.

Adıgüzel vd. (2012), satır bölütleme işleminde karşılaşılan problemlerin üstesinden gelebilmek için bağlı bileşen ve projeksiyon tabanlı bilgileri birleştiren bir yaklaşım sunmaktadır. Önerilen yöntem, her bir bağlı bileşenin taban çizgilerini bulmaktadır. Satırlar, her bir satıra ait olan bağlı bileşenlerin taban çizgilerinin projeksiyon bilgisi kullanılarak gruplanmasıyla tespit edilmektedir. Bileşenler, boyutlarıyla ilgili olarak farklı mesafe ölçümlerine göre satırlara atanmaktadır. Literatürde sıklıkla kullanılan yöntemlerden biri olan dikey izdüşüm yöntemi geliştirilerek tarihi dokümanlarda karşılaşılan sorunları çözebilecek şekilde değiştirilmiştir. Ayrıca dikey izdüşüm profillerindeki

tepe noktalarının tespiti için yeni bir yöntem önerilmiştir. Çalışmada bir projeksiyon profilindeki tepeleri saptamak için yeni bir yöntem olan Fourier eğri uydurma yöntemi önerilmektedir. Küçük bileşenler çıkarılıp, kırık karakterler birbirine bağlandıktan sonra yatay projeksiyon profiliyle Fourier eğri uydurma yöntemi kullanılarak satırları ayırma noktası bulunmaktadır. Farklı komşu satırlar arasındaki kesme noktasının yerini belirleyen bağlı bileşenin taban çizgisini çıkarmak için kontur bilgisi kullanılır. Bileşenler, taban hatlarındaki piksellere uyan bir polinom eğrisi ile yaklaşık olarak en yakın hatta göre atanır. Başlangıçta kaldırılan küçük boyutlu bileşenler dört yöndeki en yakın komşu bağlı bileşene yeniden atanır. Algoritma, farklı basılı ve el yazımı Osmanlı veri kümeleri üzerinde test edilmiştir. Bu çalışmada bir bölgenin birden fazla metin satırı içermesine rağmen tek bir metin satırı olarak algılanması durumu bulunmaktadır. Ayrıca metin satırları arasında bulunan küçük boyutlu bileşenlerin ayrı bir metin satırı olarak algılanması durumu da bulunmaktadır. Çalışma 6 farklı el yazımı kitap üzerinde test edildiğinde yaklaşık %93 doğruluk oranıyla satırlar bölütlenmiştir. Ayrıca yöntem matbu belgeler üzerinde test edildiğinde % 96.56 oranında F1-skoru elde etmiştir. Hataların büyük bir kısmı küçük boyutlu bileşenlerin ait oldukları satırlara atanmamasından kaynaklanmaktadır. Sonuçlar, yöntemin hem matbu hem de el yazımı belgelerinden, farklı yazma koşullarında en az % 92 doğrulukta satırları ayırmayı başardığını göstermektedir.

Koo ve Cho (2012), düzensiz karakter boyutları, uzamsal değişiklik gösteren yazı yönelimi ve satırlar arasındaki girişimden dolayı satır bölütlemenin serbest el yazımı dokümanlar için hâlâ zorlu bir problem olduğunu belirtmektedirler. Satır çıkarımı işlemi bağlı bileşenlerin gruplanması problemi olarak düşünülmüştür. Satır bölütlemeye karşılaşılan problemlerin üstesinden gelinmesi için satırlar arasındaki etkileşimi ve her bir satırın eğriselliğini göz önünde bulunduran bir maliyet fonksiyonu önerilmiştir. Maliyet fonksiyonunun minimize edilmesi sonucunda metin satırları elde edilmektedir. Maliyet fonksiyonunun minimizeasyonu iki aşamadan oluşmaktadır. İlk olarak aşağıdan yukarı gruplama ile kaba bir çözüm elde edilir. Daha sonra kaba çözüm iteratif bir şekilde küçük değişiklikler yapılarak iyileştirilir. Bu değişiklikler dört şekilde yapılmaktadır: birleştir, ayır, birleştir-ayır ve birleştir-birleştir-ayır. Birleştirme işleminde iki komşu satır bir satır olacak şekilde birleştirilir. Ayırma işleminde tek bir satır iki ayrı satır olacak şekilde ayrılır. Diğer işlemler de benzer şekilde tanımlanmıştır. Bitişik el yazımı ve Hintçe yazılarında çok yaygın bir durum olan az sayıda bağlı bileşen olduğunda yöntem başarısız olmaktadır. Bir satırda çok az sayıda karakterin bulunması durumunda iki komşu satır tek bir satır olarak algılanmakta ve hatalı bölütlenmektedir. Önerilen yöntem Çince el yazımı doküman imgeleri içeren HIT-MW veri seti üzerinde

test edilmiştir. Buna göre tespit oranı % 99.52 iken hata oranı % 0.32 olarak elde edilmiştir.

Ryu vd. (2014), el yazımı dokümanlar için dile bağımlı olmayan bir metin satırı çıkarımı algoritması önermişlerdir. Yöntem bağlı bileşen tabanlıdır ancak geleneksel yöntemlerin aksine vuruşlar ve eksik bölütlenmiş bağlı bileşenlerin normalize olanlara parçalanmış halleri analiz edilmektedir. Yöntem kısaca bağlı bileşenlerin çıkarılması ve onların gruplanarak satırların oluşturulması şeklinde özetlenebilir. Minimizasyonu sonucu satırların elde edildiği bir maliyet fonksiyonu oluşturulmaktadır. Önceki çalışmalarındaki yaklaşım uzamsal olarak değişen durumları (satırlar arası boşluklar ve yönelim) bağlı bileşenlerin dağılımından hesaplamaktadır ve Latin tabanlı el yazımı içeren dokümanlarda başarılı sonuçlar vermemektedir. Ayrıca bölütleme başarısı çoğu karakterin birbirine bağlı olduğu Hintçe yazılarda daha kötüdür. Çince'nin yanı sıra diğer dillerde de çalışabilen bir yaklaşım için önceden önerdikleri yöntemlerini geliştirmişlerdir. Eksik bölütlenen bağlı bileşenler normalize olanlara parçalanarak etkili bir bağlı bileşen bölütleme yöntemi geliştirilmiştir. Ancak geliştirilen bu yöntem enerji minimizasyonunda sorunlara yol açmıştır, çünkü parçalamadan dolayı bazen bağlantı bilgisi kaybolmaktadır. Bunu önlemek için enerji minimizasyon yöntemi de geliştirilmiştir. Yöntem bir satırın çok az sayıda bağlı bileşen içermesi ve/veya komşu satırların birbirine çok yakın olması durumunda başarısız olmaktadır. Önerilen yöntem ICDAR2013 el yazımı bölütleme yarışmasına katılan 10 yöntem arasında %98.64 tespit oranı ve %98.68 tanıma doğruluğu ile en iyi satır çıkarım performansını göstermiştir. Ayrıca yöntem HIT-MW veri seti üzerinde denendiğinde %99.78 tespit oranı ve % 99.88 tanıma doğruluğu elde edilirken ICDAR2009 veri seti üzerindeki test edildiğinde %99.60 tespit oranı ve %99.63 tanıma doğruluğu elde edilmiştir.

Saabni vd. (2014), otomatik metin satırı çıkarımı için dilden bağımsız bir yöntem sunmuşlardır. Önerilen yöntemle metin imgesinin enerji haritası hesaplanmakta, satırların üzerinden ve arasından geçen seam adı verilen çizgiler belirlenmektedir. Çalışmada ikili ve gri seviyeli imgeler için iki farklı algoritma geliştirilmiştir. İkili imgeler üzerinde çalışan ilk algoritma metin satırı boyunca bileşenlerin çıkarılabileceğini varsaymaktadır. Ön işleme adımında yüksekliği büyük (yüksekliği ortalama yüksekliğin önemli ölçüde üzerinde olan) dikey bileşenler, temas eden bileşenler olarak sınıflandırılır ve ortadan ayrılır. Seam ortadan ve metin satırı boyunca geçmekte ve o satırın kelime ve harflerini işaretlemektedir. Daha sonra işaretlenmemiş bileşenleri en yakın metin satırına atamaktadır. İkinci algoritma doğrudan gri seviyeli imgeler üzerinde çalışabilmektedir. Gri seviyeli imgeden uzaklık dönüşümünü hesaplar ve ortadaki seam ve ayırma seam'i olmak üzere iki çeşit

seam üretir. Ortadaki seam metin satırlarını tanımlamaktadır. Ayırma seam'i ise ortadaki seam'e göre belirlenir ve metin satırlarının alt ve üst sınırlarını tanımlamaktadır. İkili imgeler için önerilen algoritma ICDAR2009 el yazımı bölütleme yarışmasının veri seti ve 100 adet imgeden oluşan özel bir koleksiyon veri seti üzerinde test edilmiştir. Önerilen yöntem ICDAR 2009 veri setinde %98.5, özel koleksiyon veri setinde ise %98.85 başarı elde etmiştir. Hatalı sınıflandırmalar genellikle metin satırındaki kelimeler arasında büyük boşlukların olmasından kaynaklanmaktadır. Bu durum seam'in bir satırdan komşu satıra sıçramasına neden olmaktadır. Gri seviyeli imgeler için önerilen algoritmayı test etmek için farklı kaynaklardan toplanarak 215 tane el yazımı doküman imgesinden oluşan bir veri seti oluşturulmuştur. Yapılan testlerde % 95.21 ve % 98.35 arasında bölütleme doğruluğu elde edilmiştir.

Han vd. (2017), spektral kümelemeye dayalı bir satır bölütleme algoritması önermişlerdir. Doküman imgesi ilk olarak ikili hale dönüştürülmekte, daha sonra piksel noktalarının bitişik matrisi hesaplanmaktadır. Bu benzerlik metriği üzerinde spektral kümeleme uygulanmaktadır. Metin satırlarını gruplandırmak için dikey k-ortalamlar kümeleme algoritması kullanılmıştır. Bu yöntemin kısıtlaması parametrelerin seçilmesi konusunda hassas olmasıdır. Ayrıca büyük veri setleri için hesaplama açısından karmaşıktır. Önerilen yöntem Çince el yazımı doküman imgeleri içeren HIT-MW veri seti üzerinde test edildiğinde, tespit oranı %95.42, tanıma doğruluğu %95.61 ve FM performans metriği %95.51 olarak elde edilmiştir.

Makalenin takip eden bölümlerinde çalışmada sırasıyla kullanılan yöntem izah edilmiş, veri seti ve performans metrikleri sunulmuş ve elde edilen sonuçlar verilerek yöntemin bölütleme başarısı tartışılmıştır.

3. Materyal ve Yöntem

Bu çalışmada, el yazımı dokümanlardaki farklı davranışlara sahip satırları bölütleyebilen adaptif bir yöntem geliştirilmiştir. Yöntemin en temel amacı el yazımı satırlar arasında elde edilebilecek en güvenli sınırın bulunarak her bir satırın kayıpsız bir şekilde elde edilmesidir. Satırları oluşturan karakterlere ait pikseller süperpiksel yapıları olarak kabul edilerek piksel kümeleri arasındaki en uygun sınırı bulabilen optimizasyon tabanlı adaptif bir yöntem geliştirilmiştir. Çalışmada önerilen yöntem dört alt aşamadan oluşmaktadır. Bu aşamalar Şekil 2'de verilmiştir. İlerleyen bölümlerde bu aşamalar ayrıntılı olarak anlatılmaktadır.



Şekil 2. Önerilen yöntemin alt aşamaları

3.1. Yatay İzdüşüm Yöntemi

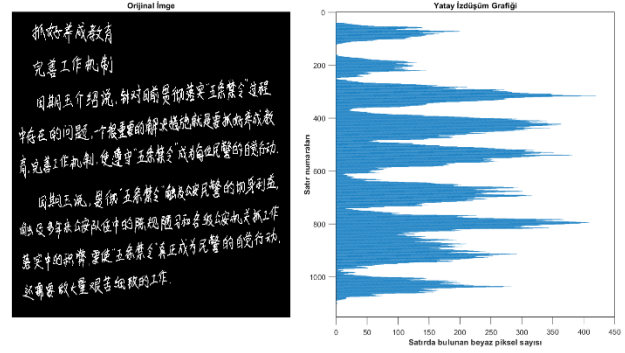
Bu aşamada temel amaç, her ardışık satır çifti için bölütleme işleminin başlatılacağı en uygun başlangıç konumlarının belirlenmesidir. Bu amaçla yatay izdüşüm yöntemi kullanılmıştır. Bu yöntemin seçilmesinin nedeni çalışmada kullanılan veri setlerinin ikili imge türünde olmasıdır. İkili imgede arka plan piksellerinin "0" ve metin içeren piksellerin "1" ile temsil edilmesi ile her satırdaki piksel değerleri toplanarak doküman imgesindeki arka plan ve metin içeren bölgeler tespit edilebilmektedir.

B, boyutları $M \times N$ olan ikili bir imge olmak üzere imgenin yatay izdüşümü $P_h[i]$ Eşitlik (1) ile hesaplanmaktadır. Burada i satırı, j ise sütunu temsil etmektedir (Hull, 1998).

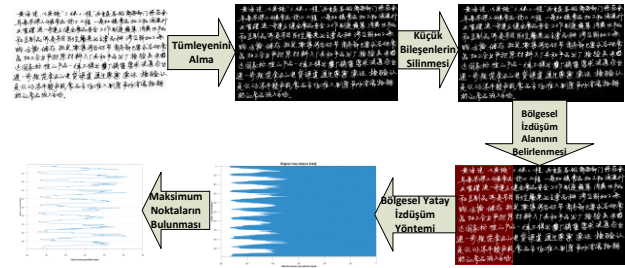
$$P_h[i] = \sum_{j=1}^N B(i, j) \quad (1)$$

Şekil 3'te bir el yazımı doküman imgesi için elde edilen yatay izdüşüm grafiği verilmiştir. Metin içeren piksellerin değeri "1" olacak şekilde imgenin tümleyeni alınmış ve sonrasında yatay izdüşüm grafiği elde edilmiştir. Grafikten anlaşıldığı üzere arka plan bölgesine ait satırlarda bulunan piksellerin toplamı metnin bulunduğu satırlardakinden daha düşük değerdedir. Ayrıca eğik satırların olduğu durumlarda -matbu imgelerin aksine- satırların yata izdüşümde örtüşmesi nedeniyle güvenli bölgelerin tespiti kolay olmamaktadır. Dolayısıyla izdüşüm grafiğinde bölgesel minimum noktalarının bulunması gerekmektedir. Bu amaçla, çalışmada yatay izdüşüm yöntemi imgenin bütününe değil başlangıç noktasının istendiği tarafta belirlenen bir bölgeye uygulanmıştır. Bu sayede, bölgesel minimum noktalar daha etkili bir şekilde elde edilmiştir.

Komşu satırların birbirine temas etmesi durumunda satırlar arasındaki boşluk bölgesinin bulunması daha zordur. Ayrıca iki komşu satır arasındaki sınır çizgisinin bulunması için birbirine temas eden karakterlerin veya kelimelerin birbirinden ayrıştırılması gerekmektedir. Metin satırları arasındaki sınır çizgisinin bulunması için öncelikle sınır çizgilerinin başlangıç noktaları yatay izdüşüm yöntemiyle belirlenmiştir. Metin satırlarını ayıran sınır çizgilerinin başlangıç noktalarının bulunma aşamalarını anlatan akış diyagramı Şekil 4'te verilmiştir.

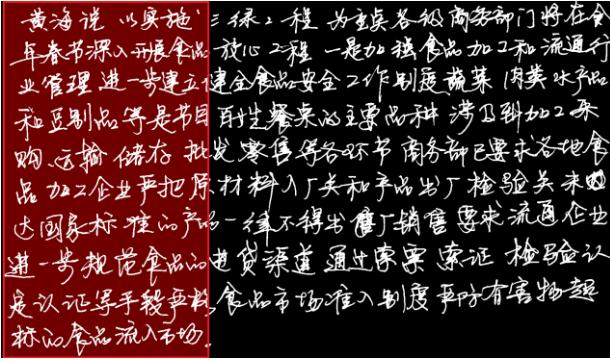


Şekil 3. Doküman imgesi ve yatay izdüşüm grafiği

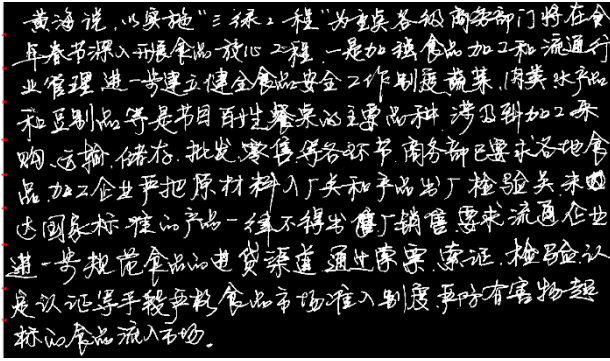


Şekil 4. Satırları ayıran çizgilerin başlangıç noktalarının bulunması

Doküman imgesindeki metinlerin yoğun olduğu bölgelerin tespit edilebilmesi için öncelikle doküman imgesinin tümleyeni elde edilmiştir. Yatay izdüşüm yöntemiyle belirlenen başlangıç noktalarının daha yüksek doğrulukta bulunabilmesi için imge içerisindeki gürültüler, noktalamalar ve vurgular gibi piksel sayısı 50'den daha az olan bağılı bileşenler geçici olarak silinmektedir. Böylece izdüşüm yöntemi daha etkili bir şekilde kullanılabilir. Doküman imgelerinde satırlar arasındaki eğiklik durumu, örtüşen satırlar ve birbirine temas eden satırlar bulunduğundan dolayı başlangıç noktalarının yüksek doğrulukta bulunması için yatay izdüşüm yöntemi bölgesel olarak uygulanmıştır. İzdüşüm yönteminin imgedeki hangi bölgeye uygulanacağını ayarlamak için imgedeki sütunlar taranarak yazıların bulunduğu ilk sütun bulunur. Bulunan bu sütun ile imgenin genişliğinin 1/3'ü kadar uzaklık arasında yatay izdüşüm yöntemi uygulanmıştır. Şekil 5'te doküman imgesi üzerinde yatay izdüşüm yönteminin uygulandığı bölge kırmızı dikdörtgen içerisinde belirtilmiştir. İmgenin bölgesel yatay izdüşümü elde edildikten sonra izdüşüm grafiğindeki tepe noktaları bulunur. Tepe noktalarının olduğu kısımlar en çok metin pikselinin olduğu satırları ifade etmektedir. Metin satırlarını ayıran bölge ise iki tepe noktasının arasındadır. Tepe noktalarının araları satırları ayıran çizginin başlangıç noktaları olarak seçilmiştir. Seçilen başlangıç noktaları doküman imgesi üzerinde Şekil 6'da en soldaki sütunda kırmızı renkli noktalarla belirtilmiştir.



Şekil 5. Yatay izdüşüm yöntemi uygulanan bölge

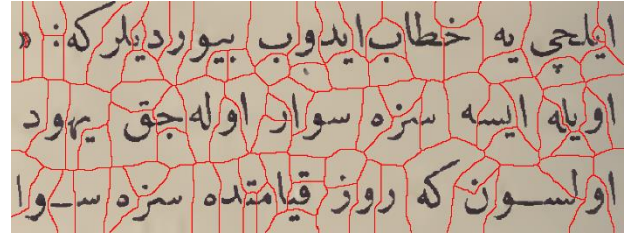


Şekil 6. Metin satırlarını ayıran çizgilerin başlangıç noktaları

3.2. Havza Bölütlemesi

İmgede bulunan pikseller renk, doku veya yoğunluk gibi özelliklerinin benzerliklerine göre gruplandırıldığında elde edilen yapılar süperpiksel olarak ifade edilmektedir. Süperpiksel algoritmaları, pikselleri anlamlı küçük bölgelere gruplamaktadır. Süperpiksel tabanlı yöntemler derinlik tahmini (Hoiem vd., 2005), imge bölütleme (Li vd., 2005; He vd., 2006; Zhang vd., 2014), vücut modeli kestirimi (Mori, 2005), nesne yerini belirleme (Fulkerson vd., 2009) gibi birçok uygulamada kullanılmaktadır. İmgedeki pikseller benzer özellikleri taşıyan piksellere gruplanarak anlamlı süperpikseller oluşturulmaktadır. Böylece daha sonraki aşamalarda hesaplama karmaşıklığı azaltılarak yöntemin daha hızlı çalışması sağlanmaktadır.

Metin içerisindeki harfler doğru bir şekilde gruplandırıldığında sözcükler ve uygun sözcükler gruplandırıldığında ise metin satırları elde edilebilmektedir. Örnek bir doküman imgesinin süperpikselle bölütlenmesi sonucu elde edilen imge Şekil 7'de verilmiştir. Burada kırmızı renkli çizgiler her bir süperpiksel bölgesinin sınırlarını belirtmektedir.



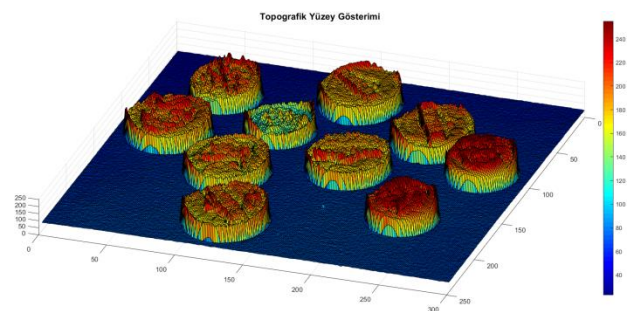
Şekil 7. Süperpikseller şeklinde bölütlenmiş doküman imgesi

Gri seviyeli herhangi bir imge topoğrafik bir yüzey olarak düşünülebilir. Topoğrafik gösterimde imgedeki madeni paraların olduğu açık renkli bölgeler daha yüksek piksel değerine sahip olup yükseltmeleri, koyu renkli bölgeler ise daha düşük seviyeli piksel değerine sahip olup havzaları temsil etmektedir. Böyle bir topoğrafik yüzeyde noktalar üç farklı durumda olabilirler. Bunlardan ilki bölgesel minimuma ait olan noktalardır. Diğerleri bu noktalara bir su damlası konulduğunda su damlası bir minimuma düşecektir. Belirli bir bölgesel minimum için, bu durumdaki nokta kümesine o minimumun havzası denilir. Son olarak ise bu noktalarda konulan suyun birden fazla minimuma eşit derecede düşebileceği noktalardır. Bu durumdaki noktalar topoğrafik yüzeyin tepe çizgilerini oluşturur ve havza çizgileri olarak adlandırılırlar.

Şekil 8'de verilen madeni paraların bulunduğu gri seviyeli imgenin topoğrafik yüzey olarak gösterimi Şekil 9'da verilmiştir.

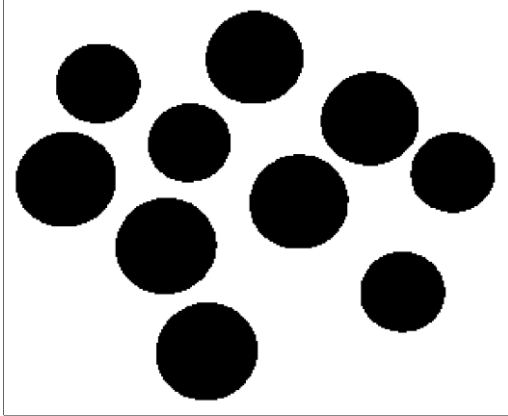


Şekil 8. Gri seviyeli imge

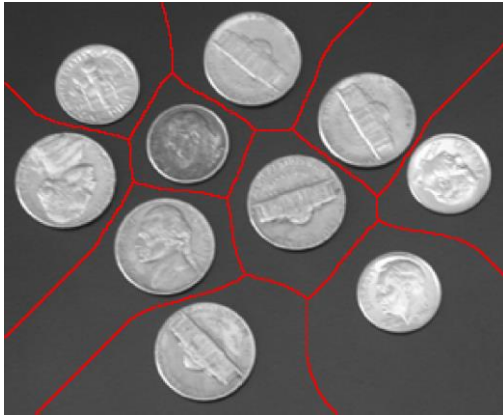


Şekil 9. Topoğrafik yüzey gösterimi

Şekil 8'deki imge ikili hale getirilip açma işlemi sonucunda Şekil 10'daki imge elde edilmiştir. Havza bölütlemesi sonucunda elde edilen bölütlenmiş imge ise Şekil 11'de verilmiştir. Burada kırmızı renkli çizgiler havza çizgilerini, bir başka deyişle süperpiksel sınırlarını belirtmektedir.



Şekil 10. İkili hale getirilen imge



Şekil 11. Havza bölütlemesi sonucu

Bu çalışmada süperpiksel algoritmalarından biri olan havza bölütleme algoritması kullanılmıştır. Havza bölütleme yönteminde bir imge yükseklik ve alçaklıklardan oluşan coğrafi bir yüzey gibi düşünülmektedir. Coğrafi yüzeyin yerel minimum noktalarından delinerek suya batırıldığı varsayılırsa, su bu yüzeyin en düşük seviyedeki noktasından itibaren dolmaya başlayarak farklı havzalara dolmaya devam edecektir. Farklı kaynaklardan gelen suların birleşmesini önlemek amacıyla, suların iki farklı havzadan gelerek birleşmek üzere olduğu bölgelere barajlar kurularsa, suya daldırma işleminin sonucunda minimumların her biri, bu minimumla ilişkili olan havzayı kapsayacak şekilde barajlarla çevrilmiş olacaktır. Su düzeyi bu coğrafi yüzeyin en yüksek noktasına ulaştığında bölütleme işlemi tamamlanmakta ve işlem sonlandırılmaktadır. İşlem sonucunda coğrafi yüzeyi farklı bölgelere ya da başka bir deyişle havzalara ayıran barajlara havza çizgileri adı verilmektedir. Bu havza çizgileri imgedeki bölgelerin sınır çizgilerini temsil etmektedir. Böylece

imge, havza çizgileri ve havzalar olmak üzere iki farklı kümeye ayrılmaktadır. Vincent ve Soille (1991), tarafından havza dönüşümünün suya daldırma benzetimine dayalı bir tanımı geliştirilmiştir.

İgri seviyeli bir imge olsun. I imgesinin en küçük ve en büyük yoğunluk değerleri sırasıyla h_{\min} ve h_{\max} olarak tanımlansın. h_{\min} değerinden h_{\max} değerine kadar değişen h gri seviye değeri ile bir özyineleme tanımlanmaktadır. Özyinelemenin başlangıcında X_h havza kümesi, h_{\min} değerindeki noktalar kümesine ($T_{h_{\min}}$) eşit olacak şekilde atanmaktadır. Daha sonra X_h havza kümesinin T_{h+1} eşik kümesi içerisindeki etki alanı ($IZ_{T_{h+1}(I)}(X_h)$) ardışık olarak genişletilir.

$$X_{h+1} = \text{MIN}_{h+1} \cup IZ_{T_{h+1}(f)}(X_h), \forall h \in [h_{\min}, h_{\max} - 1] \quad (2)$$

Eşitlik 2'de verilen MIN_h ifadesi, h yüksekliğindeki minimuma ait noktalar kümesini temsil etmektedir. Özyineleme işlemi sonucunda elde edilen $X_{h_{\max}}$, I imgesine ait havzalar kümesini vermektedir. Havza dönüşümü ise $X_{h_{\max}}$ kümesinin imge içerisindeki tümleyeni olarak bulunmaktadır.

Önerilen yöntemin uygulandığı veri setinde iki komşu satırın birbirine temas etmesi, örtüşmesi gibi durumlar olduğundan dolayı bu bölgelerde de satırlarını ayıran çizgilerin bulunabilmesi için havza bölütlemesi aşamasına geçilmeden önce bu bileşenler bölütlenmektedir. Doküman imgesindeki tüm karakterlerin yükseklik değerleri ve ortalama yükseklik değeri hesaplanmaktadır. Daha sonra ortalama yükseklik değerinin iki katından daha yüksek olan bileşenler tespit edilip havza bölütlemesi yöntemiyle daha küçük parçalara bölünmektedir. Böylece metin satırlarını ayıran çizgiler, satırlar arasında temas etme durumu olsa bile temas eden bileşenler bölütlendiği için bu bölümden geçebilmektedir.

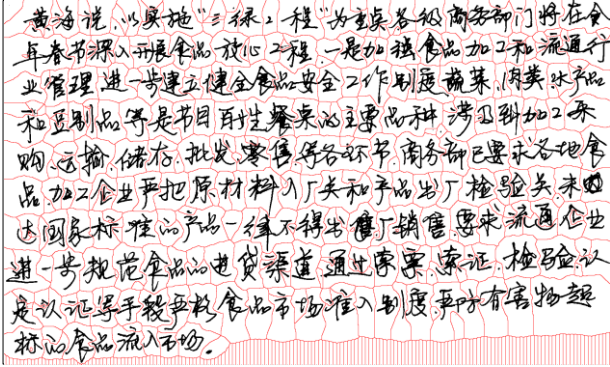
İkili seviyede bulunan doküman imgelerinde havza bölütlemesi yapabilmek için bir uzaklık dönüşüm matrisi hesaplanmaktadır. Uzaklık dönüşümü, siyah beyaz imgedeki her bir piksel için, bu piksel ile siyah beyaz imgedeki en yakın sıfır (beyaz) olmayan piksel arasındaki mesafe olan bir sayı atamaktadır. Uzaklık dönüşümünün hesaplanması ile ilgili bir örnek gösterim Şekil 12'de verilmiştir. Doküman imgelerinde en alt metin satırının da düzgün bir şekilde elde edilebilmesi için imgede en alt satıra geçici olarak yapay noktalar eklenmiştir. Doküman imgesinin uzaklık dönüşüm matrisi elde edilmiştir.

Elde edilen dönüşüm matrisine havza bölütlemesi uygulanmıştır ve elde edilen imge Şekil 13'te verilmiştir. Burada her bir bağlı bileşen farklı bir bölgede olacak şekilde bölge sınırları belirlenmiştir.

Her bir bölge piksellerden oluşan birer süperpiksel olarak adlandırmaktadır.

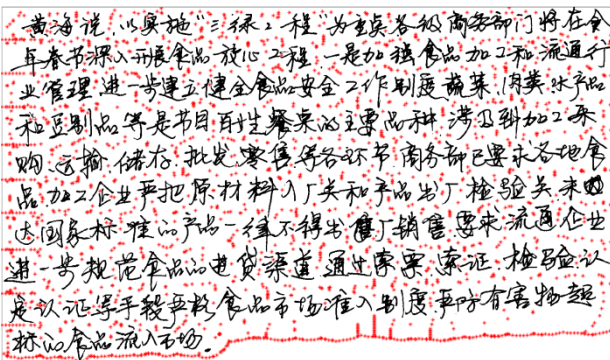
0	0	0	0	0	Uzaklık Dönüşümü	1.41	1.00	1.41	2.24	3.16
0	1	0	0	0		1.00	0	1.00	2.00	2.24
0	0	0	0	0		1.41	1.00	1.41	1.00	1.41
0	0	0	1	0		2.24	2.00	1.00	0	1.00
0	0	0	0	0		3.16	2.24	1.41	1.00	1.41
0	0	0	0	0						

Şekil 12. Uzaklık dönüşüm matrisinin elde edilmesi



Şekil 13. Havza bölütlemesi uygulanan imge

Havza bölütleme sonucunda elde edilen Şekil 13'teki görüntüde bulunan sınırlardaki üç veya daha fazla sınır çizgisinin birleştiği düğüm noktaları hesaplanmıştır. Bu noktalara metindeki düğüm noktalarını artırabilmek için sınır çizgilerinden örnek noktalarda eklenerek elde edilen düğüm noktaları imgesi Şekil 14'te verilmiştir. Başlangıçta sapma olmaması için başlangıç noktalarından itibaren her 10 sütunda bir olacak şekilde 9 adet düğüm noktası daha eklenmiştir. Ayrıca en alttaki metin satırını düzgün bir şekilde bulabilmek için havza bölütlemesinden önce en alttaki satıra yapay noktalar eklenmiştir. Böylece oluşan düğüm noktaları daha kullanışlı bir hal almıştır.



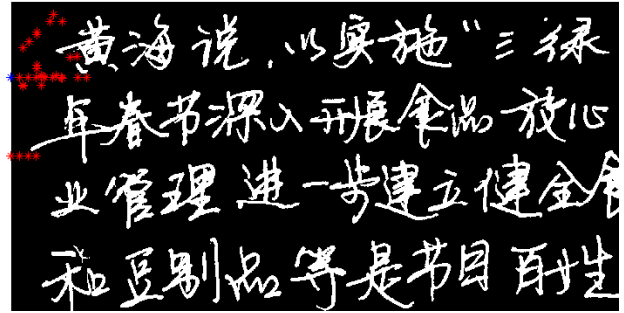
Şekil 14. Düğüm noktaları imgesi

3.3. Düğümler Arasındaki İlişkilerin Belirlenmesi

İmgedeki beyaz piksellerin kendine en yakın siyah pikselle olan uzaklıkları uzaklık dönüşümü ile elde

edilmiştir. Düğüm noktalarının uzaklık değerleri bu uzaklık dönüşüm matrisinden elde edilmiştir. Elde edilen bu uzaklık değerleri imgedeki piksellerin güvenlik değerleri olarak ifade edilebilir. Yani bir pikselin uzaklık değeri ne kadar yüksek ise kendisine en yakın yazı pikselinden o kadar uzaktır.

Her bir düğümün diğer düğümlere olan uzaklık değerleri öklid uzaklığı olarak hesaplanmıştır. Daha sonra hesaplanan bu uzaklık değerlerine göre her bir düğümün kendisinden daha ileride(sağ tarafta) ve kendisine en yakın olan 30 adet düğüm noktası bulunmuştur. Şekil 15'te mavi nokta ile gösterilen başlangıç noktasına en yakın 30 düğüm kırmızı noktalarla gösterilmiştir.



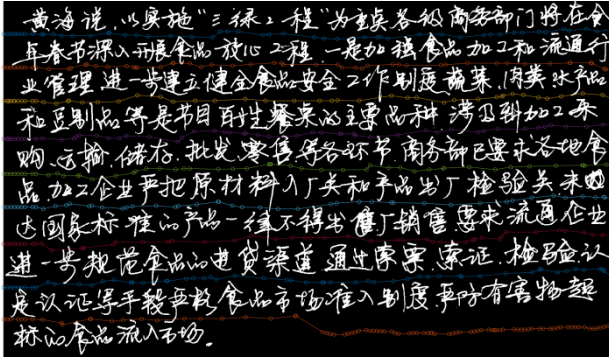
Şekil 15. Kendisine en yakın düğüm noktalarının elde edilmesi

Her bir başlangıç noktasından başlanarak daha sonra hangi noktanın seçileceği ile ilgili bazı parametreler hesaplanmaktadır. İlgili düğüm noktası ile en yakın komşu düğüm noktaları arasındaki açı değeri, dikey uzaklıklar ve yatay uzaklıklar hesaplanmaktadır. Daha sonra bu hesaplanan değerler ve daha önce elde edilen güvenlik değerlerine göre ilgili düğüm noktasından sonra hangi noktanın seçileceğine dair bir seçim fonksiyonu oluşturulmuştur. Oluşturulan fonksiyon Eşitlik (3) ile verilmiştir.

$$f(\alpha, Y, X, \beta) = w_1 \times (\alpha) + w_2 \times (Y) + w_3 \times (X) - w_4 \times (\beta) \quad (3)$$

Burada α değeri iki düğüm noktası arasındaki açı bilgisini, Y değeri düğümler arasındaki dikey uzaklığı, X değeri düğümler arasındaki yatay uzaklığı, β değeri düğüm noktasının güvenlik değerini belirtmektedir w_i değerleri ise ağırlık katsayıları olup her imge için boyut ve satırlar arası ortalama genişlik değerlerine göre optimize edilmektedir. Eşitlik (3) ile verilen seçim fonksiyonuna göre ilgili düğüm noktasıyla kendisine en yakın olan 30 nokta arasında açı değerleri, dikey uzaklık değerleri ve yatay uzaklık değerleri düşük, güvenlik değeri ise yüksek bir nokta seçilmesi hedeflenmiştir. Burada Seçim fonksiyonundan elde edilen değerlerden en küçük olan değer belirlenir ve bu değere göre düğüm noktası seçilir ve bu işlemler doküman imgesindeki en sağda bulunan sütuna kadar

devam etmektedir. Şekil 16'da her bir başlangıç noktasından satır sonuna kadar Eşitlik (3) ile verilen seçim fonksiyonuna göre seçilen düğüm noktaları gösterilmiştir.



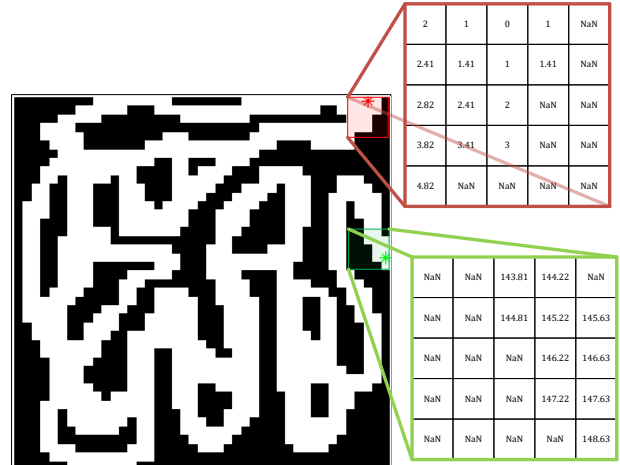
Şekil 16. Seçilen düğüm noktaları

3.4. Düğümler Arası En Kısa Yolun Bulunması

Başlangıç noktasından hedef noktasına en kısa yoldan gitmek için Şekil 14'te verilen ara düğümlerin tamamını kullanmak işlem sayısını artırmakta ve sistemin çalışma süresini uzatmaktadır. Bu yüzden başlangıç ve hedef noktaları arasındaki ara düğümlerin tamamı değil de aradaki 6 adet düğüm seçilmiştir. Başlangıç noktasından hedef noktasına giden en kısa yolu bulurken 6 adet durak noktasından geçecek şekilde en kısa yol bulunmuştur. İki nokta arasındaki en kısa yolu bulurken jeodezik uzaklık bilgisi kullanılmıştır (Soille, 2014):

$$d = \begin{cases} |x_1 - x_2| + (\sqrt{2} - 1)|y_1 - y_2|, & \text{eğer } |x_1 - x_2| > |y_1 - y_2| \\ (\sqrt{2} - 1)|x_1 - x_2| + |y_1 - y_2|, & \text{diğer} \end{cases} \quad (4)$$

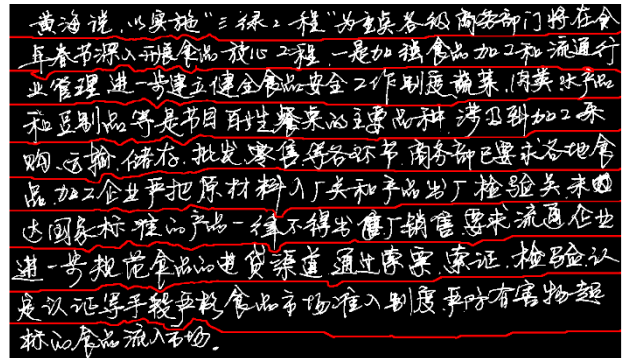
Doküman imgesi üzerinde başlangıç noktası için bir uzaklık matrisi hesaplanmıştır. Şekil 17'de verilen imgede başlangıç noktası kırmızı renkli yıldız, bitiş noktası ise yeşil renkli bir yıldız ile gösterilmiştir. Bu noktalar için hesaplanan uzaklık matrisinin küçük bir kısmı şekil üzerinde belirtilmiştir. Başlangıç noktası koordinatlarının olduğu matris elemanı 0 değeri alınırken diğer beyaz pikseller bu noktadaki piksele göre mesafe değeri alırlar, matristeki yazı piksellerine karşılık gelen yerlere herhangi bir değer atanmaz. Aynı şekilde hedef noktası için de benzer bir uzaklık matrisi hesaplanmakta ve bu iki matris toplanarak toplam uzaklık matrisi elde edilmektedir. Daha sonra başlangıç noktasından hedef noktasına giden en kısa yol bu toplam uzaklık matrisindeki değerlerden en küçük olanlar seçilerek bulunur. Şekil 17'deki imge için hesaplanan en kısa yol imge üzerinde Şekil 18'de verilmiştir. Şekil 16'da verilen Çince doküman imgesi için oluşturulan yollar Şekil 19'da kırmızı renkli çizgilerle gösterilmiştir. Böylece metin satırlarını ayıran sınır çizgileri bulunmuştur.



Şekil 17. Uzaklık matrisi gösterimi

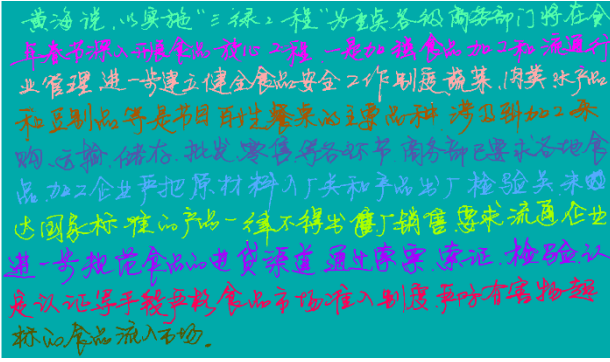


Şekil 18. Hesaplanan en kısa yol



Şekil 19. Metin satırları arasındaki sınır çizgileri

Satır çizgileri bulunduğundan sonra her bir sınır içerisindeki metin içeren piksellere farklı etiketler atanarak piksel seviyesinde etiketleme gerçekleştirilmiştir. Şekil 16'daki doküman imgesi satırlarına bölümlendikten sonra piksel seviyesinde etiketlenip daha sonra ise renklendirilerek Şekil 20'de verilmiştir.



Şekil 20. Piksel seviyesinde etiketlenmiş doküman imgesi

4. Araştırma Bulguları

Bu çalışmada doküman imgelerinin satır bölütlemesi için geliştirilen algoritma Çince el yazımı dokümanları içeren büyük bir veri seti üzerinde uygulanmış ve literatürde bulunan diğer yöntemlerle karşılaştırılmıştır. İlerleyen kısımda kullanılan veri seti ve yöntemin başarısının test edilmesi için kullanılan performans metrikleri açıklanmıştır. Daha sonra bu çalışmada elde edilen deneysel sonuçlar açıklanarak diğer yöntemlerle karşılaştırılmıştır.

4.1. Veri Seti

Önerilen yöntem Harbin Teknoloji Enstitüsü tarafından toplanarak halka açık olarak kullanıma sunulan, çeşitli yazarlar tarafından yazılmış Çince el yazımı doküman imgelerinden oluşan geniş bir veri seti olan, HIT-MW veri seti (Su vd., 2007) üzerinde uygulanmıştır. Veri setindeki imgeler el yazımı karakter tanıma ve metin satırı bölütleme amacıyla kullanılan ikili doküman imgelerinden oluşmaktadır. Veri setinde satırlar arası boşluğun dar olduğu, eğimli metin satırlarının, farklı yazarlardan dolayı çeşitli yazı stillerinin bulunduğu, satırlar arası temas eden ve örtüşen bileşenlerin bulunduğu doküman imgeleri mevcuttur. Veri seti (<https://sites.google.com/site/hitmwd/>) 780'den fazla yazar tarafından yazılmış 853 adet taranmış Çince doküman imgesini içermektedir. Doküman imgelerinde toplam 8673 adet metin satırı bulunmaktadır. Veri seti metin satırlarının kırılmış imge halinde kesin referanslarını da içermektedir. Önerilen yöntemin başarısının test edilebilmesi için piksel seviyesinde etiketlemeye dayalı performans metrikleri kullanıldığı için veri setindeki metin satırlarının da piksel seviyesinde etiketlenmiş verileri elde edilmiştir.

4.2. Performans Metrikleri

Önerilen yöntemin başarısının test edilebilmesi için ICDAR2013 El Yazımı Bölütleme Yarışması'nda (Stamatopoulos vd., 2013) da kullanılan performans metrikleri kullanılmıştır. Satır bölütlemesi için

önerilen yöntemin performansının değerlendirilmesi, algoritma tarafından bölütlenen metin satırları ve kesin referanstaki metin satırlarının piksel seviyesinde eşleşme miktarının sayılmasına dayanmaktadır. Eşleşmelerin tespiti için değerleri sonuç ve kesin referanstaki piksellerin kesişimini belirten bir Eşleşme Skoru tablosu kullanılmıştır. Buna göre I imgedeki tüm pikselleri belirten bir küme, G_j j. kesin referans bölgesindeki tüm pikselleri içeren bir küme, R_i ise i. sonuç bölgesindeki tüm pikselleri içeren bir küme olsun. $T(s)$ ise s kümesindeki pikselleri sayan bir fonksiyon olsun. EşleşmeSkoru(i,j) ifadesi i. sonuç bölgesi ve j. kesin referans arasındaki eşleşmeyi belirtir ve Eşitlik(5)'teki gibi ifade edilebilir.

$$EşleşmeSkoru(i, j) = \frac{T(G_j \cap R_i \cap I)}{T((G_j \cup R_i) \cap I)} \quad (5)$$

Bir bölge çifti, sadece değerlendiricinin kabul eşiği T_a 'ya eşit veya bundan daha fazla ise, bire bir eşleşme olarak kabul edilmektedir. Çalışmada T_a değeri 0.95 olarak seçilmiştir (Stamatopoulos vd., 2013). N kesin referans satırlarının sayısı, M sonuç satırlarının sayısı ve o2o ise bire bir eşleşen satır sayısı olsun, bu durumda tespit oranı (DR) ve tanıma doğruluğu (RA) aşağıdaki gibi tanımlanır:

$$DR = \frac{o2o}{N} \quad (6)$$

$$RA = \frac{o2o}{M} \quad (7)$$

Tespit oranı, kesin referansta verilen satırlardan ne kadarının doğru bir şekilde tespit edildiğini ifade etmektedir. Benzer şekilde tanıma doğruluğu ise doğru tespit edilen satırların sayısının bölütleme sonucunda elde edilen toplam satır sayısına oranı olarak tanımlanmaktadır. Tespit oranı (DR) ve tanıma doğruluğu (RA) değerlerinin harmonik ortalaması olarak tanımlanan FM (F-Measure) performans metriği elde edilir:

$$FM = \frac{2 \cdot DR \cdot RA}{DR + RA} \quad (8)$$

FM ise kesin referans ve elde edilen sonuçlara göre yöntemin genel test başarısını ölçmektedir. Bu çalışma kapsamında geliştirilen metin satırı bölütleme algoritması için eşleşme skoru tabloları elde edilmiş ve eşleşme skoru değeri, 0.95 olarak belirlenen T_a eşik değerinden büyük olan metin satırları bire bir eşleşme olarak kabul edilmiştir. Bire bir eşleşme sayıları kullanılarak önerilen yöntemin performansının test edilebilmesi için tanıma doğruluğu, tespit oranı ve FM performans metriği değerleri elde edilmiştir.

4.3. Deneysel Sonuçlar

Çalışma kapsamında geliştirilen satır bölütleme algoritmasının HIT-MW veri seti üzerinde

uygulanarak performans metrikleri elde edilmiştir. Veri setindeki 853 doküman imgesi için kesin referansta 8673 adet metin satırı bulunmaktadır. Önerilen yöntemle tespit edilen satır sayısı 8706 adettir. Bir metin satırının bire bir eşleşme kabul edilebilmesi için eşik değeri 0.95 olarak alınmıştır. Buna göre bire bir eşleşen metin satırı sayısı 8502 olarak elde edilmiştir. Bu sonuçlara göre yöntemin tespit oranı % 98.03, tanıma doğruluğu % 97.66 ve FM metriği ise % 97.84 olarak elde edilmiştir. Önerilen yöntem ve karşılaştırılan diğer yaklaşımların sonuçları Tablo 1’de verilmiştir. Karşılaştırma yapılan diğer yöntemlerin aksine önerilen yöntemde hatalı çıkan satırların bölütlenmesi için ayrıca bir işlem yapılmamaktadır.

Tabloda yer alan ilk altı çalışmada uzaklık metriği öğrenimi ve parametre ayarlaması için rastgele 50 imge seçilmiştir, bu imgeler 508 metin satırı içermektedir. Geriye kalan 803 imgede 8169 metin satırı bulunmaktadır. Doğru tespit edilen satır sayısının kesin referanstaki satır sayısına oranı tespit oranıdır. Tespit edilen satırlardan hatalı olanlarının sayısının tespit edilen satırlara oranı ise hata oranını

vermektedir.

Karşılaştırma yapılan yöntemler arasında en başarılı olan yöntemlerde bölütleme işlemi sonrası başarı oranını artıran ilave işlemler yapılmaktadır. Bu çalışmada son aşamada herhangi bir iyileştirme çalışması yapılmamış olup, bölütleme işlemi sonrası elde edilen sonuçlar sunulmuştur.

Tanıma doğruluğu sonuçlarının tespit oranı sonuçlarına kıyasla biraz daha düşük olması önerilen yöntemin bazı satırları bölütlerken ilave satırlar üretmesi ile açıklanabilir. Bu durumun iki temel nedeni vardır. Birincisi temas eden veya örtüşen satırların varlığıdır. İkinci neden ise aynı satıra ait karakterler arasındaki boşluğun ardışık satırlar arasındaki boşluktan daha fazla olduğu durumlardır. Bu durum yöntemin hatalı ilave satırlar üretmesine neden olmaktadır. Ancak tabloda verilen sonuçların ışığında önerilen yöntemin tespit oranı ve tanıma doğruluğu başarısının oldukça iyi olduğu görülmektedir.

Tablo 1. Önerilen yöntemin ve literatürdeki diğer yöntemlerin bölütleme başarıları

	Tespit Oranı (%)	Hata Oranı (%)	Tanıma Doğruluğu (%)	FM (%)
Yin ve Liu (2009) (işlem sonrası ile) (803 imge)	98.02	2.47	97.53	97.78
Yin ve Liu (2009) (işlem sonrası olmadan) (803 imge)	95.75	7.37	-	-
Nagy vd., (1992) (803 imge)	45.07	57.21	-	-
O’Gorman (1993) (803 imge)	65.38	44.38	55.62	60.11
Su vd., (2007) (803 imge)	55.34	23.33	-	-
Arivazhagan vd., (2007) (803 imge)	92.07	7.72	92.28	92.17
Ziaratban ve Faez, (2010) (853 imge)	98.34	-	-	-
Gao vd., (2011) (Ta = 0.9 853 imge)	98.68	-	98.76	-
Du vd., (2009) (853 imge)	95.92	3.14	96.86	96.39
Koo ve Cho, (2012) (853 imge)	99.52	0.32	-	-
Ryu vd., (2014) (853 imge)	99.78	-	99.88	99.83
Han vd., (2017) (853 imge)	95.42	-	95.61	95.51
Önerilen Yöntem (853 imge)	98.03	2.34	97.66	97.84

5. Sonuç ve Tartışma

Bu çalışmada el yazımı doküman imgeleri için süperpiksel tabanlı bir satır bölütleme algoritması geliştirilmiştir. Önerilen yöntem Çince el yazımı doküman imgelerinden oluşan büyük bir veri seti üzerinde uygulanmıştır. Farklı eğiklik açılara sahip, örtüşen veya birbirine temas eden bileşenlerin olduğu el yazımı dokümanlarda metin satırlarının bölütlenmesi problemini yüksek doğruluk oranı ile çözebilmektedir. Önerilen yöntem, temas eden ve/veya örtüşen karakterlerin olmadığı durumlarda kayıpsız bir şekilde bölütleme işlemini gerçekleştirmektedir. Ancak, temas eden ve/veya örtüşen karakterlerin varlığı, hatalı ilave satırların üretilmesine ya da satırların tespit edilememesine neden olabilmektedir ve bu durum da başarı oranını düşürmektedir ki bu durum imge analizi uygulamaları için genel bir problemdir. Gelecek çalışmalarda, önerilen yöntemin öncelikle temas eden ve örtüşen satırları kayıpsız bir şekilde bölütlemesini sağlayan bir algoritmanın geliştirilmesi ve çalışmanın etkili bir iyileştirme adımı ile desteklenmesi hedeflenmektedir.

Teşekkür

Bu çalışma, Ali Alper DEMİR tarafından Süleyman Demirel Üniversitesi, Fen Bilimleri Enstitüsü'nde Dr. Öğr. Üyesi Ufuk ÖZKAYA danışmanlığında yürütülen "Süperpiksel Tabanlı Satır Bölütleme" başlıklı yüksek lisans tezinden üretilmiştir. Bu çalışma Süleyman Demirel Üniversitesi ÖYP Koordinasyon Birimi tarafından ÖYP06748-YL-16 nolu proje kapsamında desteklenmiştir.

Conflict of Interest / Çıkar Çatışması

Yazarlar tarafından herhangi bir çıkar çatışması beyan edilmemiştir.

No conflict of interest was declared by the authors.

Kaynaklar

Adiguzel, H., Sahin, E., Duygulu, P., 2012. A Hybrid Approach for Line Segmentation in Handwritten Documents. In *Frontiers in Handwriting Recognition (ICFHR), 2012 International Conference on* (pp. 503-508). IEEE.

Alaei, A., Pal, U., Nagabhushan, P., 2011. A New Scheme for Unconstrained Handwritten Text-Line Segmentation. *Pattern Recognition*, 44(4), 917-928.

Arivazhagan, M., Srinivasan, H., Srihari, S. (2007, January). A statistical approach to line segmentation in handwritten documents. In *Document Recognition and Retrieval XIV* (Vol. 6500, p. 65000T). International Society for Optics and Photonics.

Ataer, E., Duygulu, P., 2006. Retrieval of Ottoman Documents. In *Proceedings of the 8th ACM International Workshop on Multimedia Information Retrieval (MIR '06)*. ACM, New York, NY, USA, 155-162.

Du, X., Pan, W., Bui, T. D. (2009). Text line segmentation in handwritten documents using Mumford-Shah model. *Pattern Recognition*, 42(12), 3136-3145.

Fulkerson, B., Vedaldi, A., Soatto, S. (2009, September). Class segmentation and object localization with superpixel neighborhoods. In *Computer Vision, 2009 IEEE 12th International Conference on* (pp. 670-677). IEEE.

Han, X., Yao, H., Zhong, G. (2017, February). Handwritten text line segmentation by spectral clustering. In *Eighth International Conference on Graphic and Image Processing (ICGIP 2016)* (Vol. 10225, p. 102251A). International Society for Optics and Photonics.

He, X., Zemel, R. S., Ray, D. (2006, May). Learning and incorporating top-down cues in image segmentation. In *European conference on computer vision* (pp. 338-351). Springer, Berlin, Heidelberg.

Hoiem, D., Efros, A. A., Hebert, M. (2005, July). Automatic photo pop-up. In *ACM transactions on graphics (TOG)* (Vol. 24, No. 3, pp. 577-584). ACM.

Hull, J. J. (1998). Document Image Skew Detection: Survey and Annotated Bibliography. *Document Analysis Systems II*, 40-64.

Koo, H. I., Cho, N. I. (2012). Text-line Extraction in Handwritten Chinese Documents Based on An Energy Minimization Framework. *IEEE Transactions on Image Processing*, 21(3), 1169-1175.

Li, Y., Sun, J., Tang, C. K., Shum, H. Y. (2004, August). Lazy snapping. In *ACM Transactions on Graphics (ToG)* (Vol. 23, No. 3, pp. 303-308). ACM.

Li, Y., Zheng, Y., Doermann, D., Jaeger, S. (2008). Script-Independent Text Line Segmentation in Freestyle Handwritten Documents. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(8), 1313-1329.

Likforman-Sulem, L., Hanimyan, A., Faure, C. (1995, August). A Hough Based Algorithm for Extracting Text Lines in Handwritten Documents. In *Document Analysis and Recognition, 1995., Proceedings of the Third International Conference on* (Vol. 2, pp. 774-777). IEEE.

Louloudis, G., Gatos, B., Pratikakis, I., Halatsis, K. (2006, October). A Block-Based Hough Transform

- Mapping for Text Line Detection in Handwritten Documents. In Tenth International Workshop on Frontiers in Handwriting Recognition. Suvisoft.
- Louloudis, G., Gatos, B., Pratikakis, I., Halatsis, C. (2008). Text Line Detection in Handwritten Documents. *Pattern Recognition*, 41(12), 3758-3772.
- Nagy, G., Seth, S., Viswanathan, M. (1992). A prototype document image analysis system for technical journals. *Computer*, 25(7), 10-22.
- O'Gorman, L. (1993). The document spectrum for page layout analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(11), 1162-1173.
- Ryu, J., Koo, H. I., Cho, N. I. (2014). Language-Independent Text-Line Extraction Algorithm for Handwritten Documents. *IEEE Signal Processing Letters*, 21(9), 1115-1119.
- Saabni, R., Asi, A., El-Sana, J. (2014). Text Line Extraction for Historical Document Images. *Pattern Recognition Letters*, 35, 23-33.
- Soille, P. (2013). *Morphological Image Analysis: Principles and Applications*. Springer Science & Business Media.
- Stamatopoulos, N., Gatos, B., Louloudis, G., Pal, U., Alaei, A. (2013, August). ICDAR 2013 handwriting segmentation contest. In *Document Analysis and Recognition (ICDAR), 2013 12th International Conference on* (pp. 1402-1406). IEEE.
- Su, T., Zhang, T., Guan, D. (2007). Corpus-based HIT-MW database for offline recognition of general-purpose Chinese handwritten text. *International Journal of Document Analysis and Recognition (IJ DAR)*, 10(1), 27.
- Vincent, L., Soille, P. (1991). Watersheds in digital spaces: an efficient algorithm based on immersion simulations. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (6), 583-598.
- Yin, F., Liu, C. L. (2009). Handwritten Chinese text line segmentation by clustering with distance metric learning. *Pattern Recognition*, 42(12), 3146-3157.
- Zhang, L., Gao, Y., Xia, Y., Lu, K., Shen, J., Ji, R. (2014). Representative Discovery of Structure Cues for Weakly-Supervised Image Segmentation. *IEEE Trans. Multimedia*, 16(2), 470-479.
- Ziaratban, M., Faez, K. (2010, August). An adaptive script-independent block-based text line extraction. In *Pattern Recognition (ICPR), 2010 20th International Conference on* (pp. 249-252). IEEE.