

Fatura İçerik Tablolarının Öznitelik Kaynaştırması Yöntemiyle Analizi

Invoice Content Table Analysis with Feature Fusion

Ummu Seleme Unal^{1,2}

Ethem Unver²

Tugrul Karakaya^{1,2}

Yusuf Sinan Akgul¹

¹ GIT Vision Lab, <http://vision.gyte.edu.tr/>,

Bilgisayar Mühendisliği Bölümü

Gebze Teknik Üniversitesi, Kocaeli, 41400, Türkiye
ummuselemee@gmail.com, akgul@gtu.edu.tr

²Kuveyt Türk Katılım Bankası

Kocaeli, 41400, Türkiye

{ethem.unver, tugrul.karakaya}@kuveytturk.com.tr

Özetçe— Fatura işleme ve içerisinden finansal veriler elde etme son yıllarda araştırmalara sıkça konu olmaktadır. Katılım bankaları gibi firmalarda günlük binlerce fatura el ile işlenmektedir. Özellikle tablolarda yer alan ürün, miktar, fiyat gibi önemli bilgilerin sistemlere otomatik olarak aktarılması öncelik arz etmektedir. Bu çalışmada faturalar üzerindeki tabloların tespiti ve analizi için özgün yöntemler tanımlanarak, deneylerle etkinlikleri ölçülmüştür. Kullanılan faturalar genellikle müşteriler tarafından taranarak veya faks yolu ile kurumlara gönderilen ve görüntü kaliteleri oldukça düşük olan faturalardır. Yüksek gürültüye sahip ve yazı kalitesi düşük olan bu faturaları işleyebilmek için geliştirilen yöntemler farklı parametrelerle bir kaç defa çalıştırılmıştır. Elde edilen sonuçlar kaynaştırılarak aday tablolar elde edilmiştir. Bu çalışmada kullanılan yöntemler sayesinde, düşük Optik Karakter Tanıma (OKT) performansı, yüksek gürültü ve düşük çözünürlük gibi sorunlara rağmen etkin sonuçlar elde edilmiştir. Ayrıca sunulan yöntem sınırları belirli ve belirsiz tablolar üzerinde başarılı bir şekilde çalışmaktadır. Geliştirilen metotlar bankacılıkta kullanılan gerçek faturalar üzerinde test edilmiş ve sistem başarısı nicel olarak gösterilmiştir.

Anahtar Kelimeler — *tablo algılama, tablo analizi, fatura işleme.*

Abstract—Invoice processing and financial information extraction have been popular topics among researchers for decades. Corporations like participation banks process invoices manually. Extraction of important information from invoices like product, price, amount etc. is a prerequisite for these banks. In this paper we propose a novel technique for processing invoice image tables automatically. Invoice images we process are mostly sent by customers as scanned or fax images which have low image quality. In order to process these invoices we run different methods several times with different parameters. Results from each method are fused to get candidate tables. The proposed methods are robust to the character set used in a document, the image resolution and the noise ratio of the document image, and can perform detection operations in a highly effective manner. In addition to success in low quality images, this method can be applied both on tables with and without borders. The quantitative results obtained by applying this method on real business invoices have very favorable results.

Keywords— *table detection; table analysis; invoice processing.*

I. GİRİŞ

Faturaların işlenmesi özellikle katılım bankaları için önemli bir süreçtir. Bu tür bankalarda, fon ve finansman işlemlerinin belgelendirilmesi [1] için faturalar sıklıkla kullanılmaktadır. Örneğin Kuveyt Türk Katılım Bankasında bir günde yaklaşık 1000 fatura işlenmekte ve her biri yaklaşık olarak 6 dakika sürmektedir. İçerik tabloları ise faturaların en kritik parçalarından birisidir. Katılım bankacılığı kuralları gereği bir ticari işlemin kredilendirilebilmesi için kontrolü gerekli olan mal veya hizmetlerin ayrıntılı bilgisi tablo kısmında yer alır. Bu sebeple faturalardaki tabloların algılanması ve analiz edilmesi bankalar için hayati önem taşımaktadır.

Literatürde tablo algılama ve analizi için birçok çalışma bulunmaktadır. Bu çalışmaları 3 ana gruba ayırabiliriz. İlk grup tabloların içerikleri üzerine yoğunlaşarak, satır ve sütun analizi yapmıştır. Bart vd. [2] satır, sütun, ayırıcı gibi algısal ipuçları kullanarak tablonun farklı özelliklerini kategorize etmişler ve belge üzerinde tablonun algılandığını varsayımlarıdır. Faturalar üzerinde tabloların algılanması ile ilgili çalışmada [3] ise kelime morfolojisi yaklaşımı kullanılmıştır. Bu çalışmanın temel varsayımı ise, her satırda olması gereken fiyat etiketinin reel sayıyla ifade edilmesi ve reel sayıların satır ayırıcı olarak kullanılmasıdır.

Tablolarla ilgili çalışma yapan diğer grup ise OKT'den gelen verileri kullanmıştır. Li vd. [6] ise OKT sonuçlarına göre genel tablo konumuna karar vermiş ve sonuçları projeksiyon yöntemi ile güçlendirmiştir.

Diğer bir grup ise tabloların geometrik özelliklerinden yararlanmışlardır. [7] belgelerdeki yatay ve dikey çizgileri bulduktan sonra, bu doğruların kesişimlerinden yola çıkarak tablo algılama işlemini gerçekleştirmiştir.

Literatürde fazla sayıda tablo algılama ve analizi çalışması [4] olmasına rağmen, faturalar üzerinde yer alan tablolarla ilgili çalışmalar oldukça azdır. Faturaların tek sayfalık belgeler olması, OKT motorlarının belgenin font büyüklüğüne ve stiline karar vererek, daha kaliteli okuma yapmasını engeller. Diğer yandan, faturalar kurumlara müşteriler tarafından faks yoluyla ya da taranarak gönderildiği için, faturaların kaliteleri kontrol edilememektedir. Diğer belgelerde yapılan tablo algılama ve

analiz çalışmalarını, faturalar üzerinde uygulamak tek başına yeterli değildir.

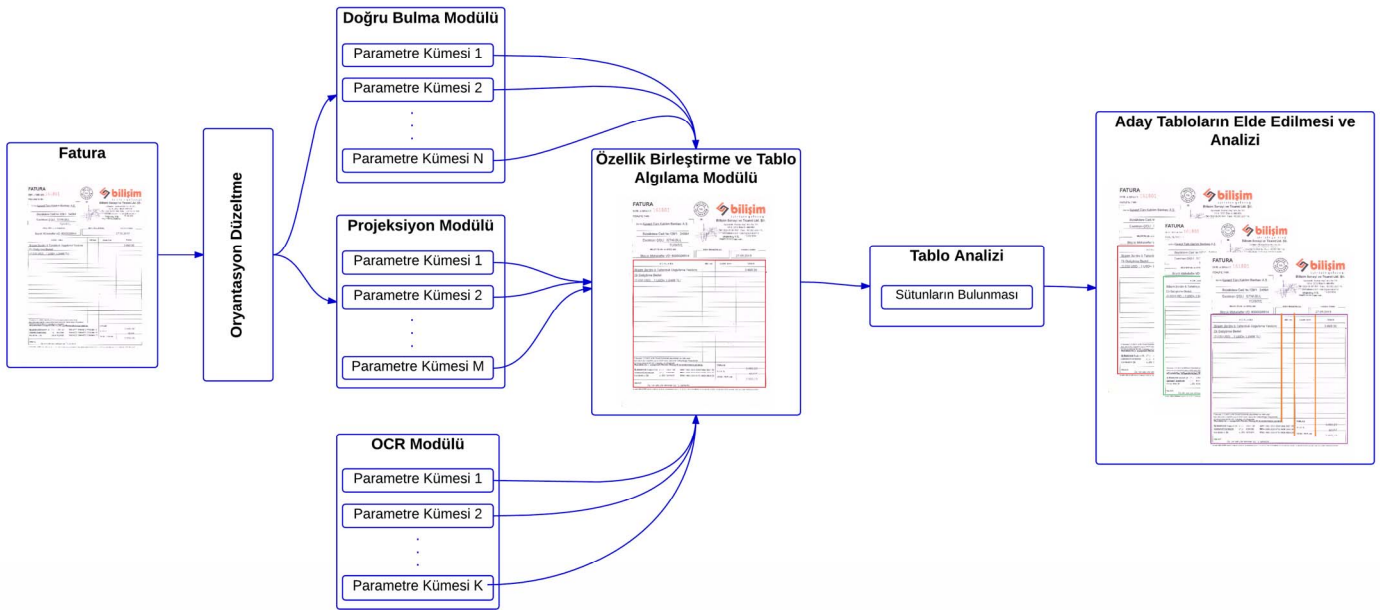
Sunduğumuz yöntemin diğer çalışmalardan en önemli farkı, tablo algılama ve analiz için geliştirilen algoritmaların aynı anda ve birden fazla kez çalıştırılmasıdır. Bu yöntemlerden her biri farklı tipteki faturaların farklı problemlerini giderir. Örneğin projeksiyon yöntemi satır ve sütunları olmayan tablolar için faydalıyken, doğru bulma yöntemi ise satır ve sütunları olan tablolar için kullanışlıdır. OKT ise tabloların satır ve sütun analizinde kullanılır. Farklı yöntemlerin aynı anda çalıştırılması faydalı olsa da, her algoritma için doğru parametrelerin bulunması oldukça zor bir problemdir. Önerilen sistemde tablo algılama ve analizi için geliştirilen yöntemleri birden fazla kez, farklı parametre kümeleri ile çalıştırıp, sonuçlar kaynaştırılarak bu probleme bir çözüm sunulmaktadır. (Şekil 1’de gösterilmektedir.) Bu

çalışmada, kullanılan yöntemlerden hangisinin daha etkili olduğu ile ilgili analizler yapılmakta ve bu yöntemlerden hangisinin gürültüye karşı daha fazla dayanıklı olduğu gösterilmektedir. Bu analizlerin sonucu, yöntemlerde kullanılan parametre kümelerinin seçimine ve sayısına karar verme konusunda fikir verecektir.

Bu bildirinin artı kalan kısmında sunulan yöntemin detaylarından bahsedilecektir. II. Bölümde; tablo tespiti ve analizi için geliştirilen yöntemler detaylı bir şekilde anlatılmıştır. Son olarak sistem üzerinde yapılan deneyler ve sonuçlara yer verilmiştir.

II. YÖNTEM

Tablo algılama ve analizi için sunulan yöntem birden fazla alt işlemlerden oluşmaktadır. Sunulan yöntemin temel adımları aşağıda açıklanmıştır.



Şekil 1. Önerilen Sistemin Akış Şeması

A. Oryantasyon Düzeltme

Müşteriler tarafından faks ya da tarayıcı yardımıyla kurumlara gönderilen faturaların kaliteleri kontrol edilemediğinden, birtakım açığı hatası problemleri olabilir. Bu faturalar üzerinde işlemler yapmak için bu tip problemlerin ortadan kaldırılması gerekmektedir. Faturanın açığı hatası, yatay ve dikey doğruların bulunması ile tespit edilir ve elde edilen değere göre fatura üzerinde çevirme işlemi yapılarak fatura düzeltilir. Dokümanların verev açığı[5] problemlerinin de giderilmesi gerekmektedir ancak sonraki işlem basamaklarında bu problem ihmal edilebilmektedir.

B. Doğru Bulma Modülü

Tablolar genel olarak yatay ve dikey doğrular ile sınırlanmıştır. Olasılıksal Hough Doğru Dönüşümü[9] yöntemi temel alınarak geliştirilen doğru bulma yöntemi, tablo algılama ve analizinin önemli bir parçası olup, adımları aşağıda açıklanmaktadır:

1. Faturalar çoğunlukla gürültü içermektedir. Fatura görüntüsü üzerinde kenarların, yatay ve dikey doğruların

düzgün bir şekilde elde edilebilmesi için görüntü üzerindeki gürültünün azaltılması gerekmektedir. Bunun için fatura görüntüsü üzerinde yatay ve dikey olarak Gauss Yumuşatma Filtresi uygulanarak görüntü yumuşatılmıştır.

2. Canny Kenar Bulma algoritması ile yatay ve dikey kenarlar elde edilmiştir.

3. Non-maximum-suppression temel alınarak geliştirilen yöntem ile kenarların parçası olmayan pikseller silinmiş ve görüntü üzerinde sadece aday kenarlar kalmıştır.

4. Olasılıksal Hough Doğru Dönüşümü metodu bulunan kenarlar üzerinde çalıştırılarak yatay ve dikey doğrular elde edilmiştir.

5. Bu yöntemle elde edilen doğrular arasında kesinti olmamalıdır. Aksi takdirde, satır ve sütun doğrularını tespit etmek zorlaşacaktır. Elde edilen tüm yatay doğruların birbirlerine yakınlığı ve dikey pozisyonları kontrol edilerek birleştirilir. Birleştirme işleminin ardından takip edilemeyen ve tablonun bir parçası olamayacak kadar kısa olan doğrular

temizlenir. (Şekil 2’de gösterilmiştir.) Aynı işlem tüm dikey doğrulara da uygulanmıştır.



Şekil 2. (a) Doğru bulma yöntemi ile elde edilen kesik doğrular. (b) Doğruların düzgünleştirilmiş hali

Farklı tipteki faturalar üzerinde bu yöntemin başarılı bir şekilde çalışması için tek parametre kümesi yerine, birden fazla parametre kümesi kullanılmaktadır. Doğru bulma adımlarından 1., 2. ve 4. adımlar farklı parametre kümeleri ile kullanılmayı gerektirir. Şekil 3’te gösterildiği gibi, bu parametre kümelerinden bazıları tabloların ana hatlarını elde etmek için kullanılırken, bazıları ise yatay ve dikey doğrular ile ilgili daha detaylı bilgiler elde etmek için kullanılır.



Şekil 3. Doğru Bulma Yöntemi Sonucu (a) Parametre kümesi 1 (b) Parametre kümesi 2

C. Projeksiyon Modülü

Yatay ve dikey sınırları belirli olan tablolar kadar, sınırları belirli olmayan tablolar da sıklıkla kullanılmaktadır. Bu tip tabloları algılamak için görüntü projeksiyonları [8] yöntemi kullanılmıştır. Görüntü üzerinde yatay ve dikey olarak projeksiyon alındığında, projeksiyon değerlerinin hızla yükseldiği noktalar fatura görüntüsü üzerindeki satır ve sütunların başlangıç noktalarına denk gelmekte ve değerlerin hızlı düşüş gösterdiği noktalar ise ayırıcıları temsil etmektedir.

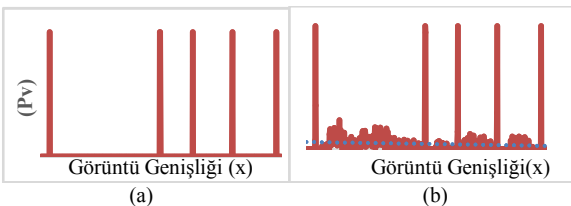
Bu yöntem uygulanmadan önce fatura görüntüsü üzerinde Gauss Yumuşatma Filtresi uygulanarak görüntü yumuşatılmış ve otomatik eşikleme yöntemi kullanılarak fatura görüntüsü, ikili belge görüntüsüne dönüştürülmüştür. I , ikili fatura görüntüsünü ifade etsin. Dikey projeksiyon şu formülle hesaplanabilir:

$$P_d(x) = \sum_{j=0}^y I(x, j), \quad (1)$$

Formülde $x = 0, \dots, g$ olacak şekilde, g ve y ikili fatura görüntüsünün genişlik ve yüksekliğini ifade etmektedir. Benzer şekilde yatay projeksiyon hesaplamak için kullanılacak formül ise şu şekildedir:

$$P_y(y) = \sum_{i=0}^g I(i, y). \quad (2)$$

Projeksiyon modülünde, Gauss Yumuşatma Filtresi ve eşikleme gibi parametre gerektiren yöntemler, birden fazla parametre kümesi ile birlikte kullanılır. Şekil 4’te fatura görüntüsü üzerinde dikey projeksiyonların farklı parametre kümeleri ile uygulanmış örnek çıktıları gösterilmektedir.



Şekil 4. Farklı parametre kümeleri ile kullanılan Projeksiyon Yöntemi (a) Parametre kümesi 1 (b) Parametre Kümesi 2

D. OKT Modülü

Faturalarda yer alan içerik tablolarının başlık ve alt bilgi kısımlarında olabilecek, “açıklama, ürün, fiyat, toplam” gibi anahtar kelimeler bellidir ve bu kelimelerin sayısı sınırlıdır. Fatura görüntüleri üzerinde OKT motorları çalıştırılarak ve bu kelimelerin konumları göz önünde bulundurularak tablonun başlık ve alt bilgisi ile ilgili bilgiler elde edilir. Sınırları belirsiz tablolar için bu kelimeler, tablonun varlığını göstermektedir ve tablo bu kelimelerin pozisyonlarına göre konumlandırılmıştır. Tabloların başlık ve alt bilgi kısımlarında yer alabilecek kelimelerin sayısı sınırlı olsa da, OKT motorları, faturaların yazı tipleri ve boyutlarının değişmesi nedeniyle, tanıma sırasında hatalar yapmaktadır. Bu hataları ortadan kaldırmak için fatura görüntüleri üzerinde 2 farklı parametre kümesi kullanılarak OKT motoru çalıştırıldı ve elde edilen sonuçlar kaynaştırıldı.(Şekil 5’te gösterilmiştir.)

TOPLAM	USD	3.240,00
KDV (%18)	USD	3.240,00
GENEL TOPLAM	USD	21.240,00

TOPLAM	USD	18.000,00
KDV (%18) <td>USD <th>3.240,00</th> </td>	USD <th>3.240,00</th>	3.240,00
GENEL TOPLAM <td>USD <th>21.240,00</th> </td>	USD <th>21.240,00</th>	21.240,00

Şekil 5. Farklı parametrelerle çalıştırılmış OKT sonuçları (a) Parametre kümesi 1 (b) Parametre kümesi 2

E. Özellik Birleştirme ve Tablo Bulma Modülü

Tablo algılama problemi için; önceki 3 modülde anlatılan yöntemlerden elde edilen doğrular, geliştirilen kural tabanlı mekanizma ile birleştirilir. Modüllerin her biri kendi içerisinde 2 farklı parametre kümesi ile çalıştırılmıştır (Şekil 1’deki $N=2, M=2$ ve $K=2$ olacak şekilde). Birleştirme işleminin ana kuralları şu şekildedir: OKT modülünden elde edilen anahtar kelimeler, tablonun dikey sınırlarını oluşturmaktadır. Doğru bulma modülünden elde edilen doğrular ve projeksiyon modülünden elde edilen zirve noktaları birbirleriyle eşleştiği takdirde, bunlar güçlü doğrular olarak kabul edilir. Farklı yöntemlerden elde edilen ve birbirleriyle eşleşmeyen doğrular da göz önünde bulundurularak, tüm doğrular kaynaştırılıp, tablo algılama işlemi için kullanılır. Nihai yatay ve dikey doğruların kesişimleri ile aday tablolar elde edilir. Her bir aday tabloya bir skor değeri verilerek, en yüksek skorlu tablo elde edilir.

F. Tablo Analizi Modülü

Tablonun algılanmasının ardından, tablonun sütunlarının belirlenmesi gerekmektedir. Sütunların belirlenmesi için öncelikle OKT modülünden elde edilen anahtar kelimelere bakılır. Ancak OKT motorları her zaman fatura üzerindeki tüm kelimeleri okuyamayabilir. Bu durumda iki doğru ya da iki ayırıcı birbirine yakınsa, bunlar arasında OKT sonucu olmasa da bir anahtar kelime olabilir. Bundan emin olmak için başlık kısmında, dikey projeksiyon alınarak balık bilgisi olup olmadığına bakılır ve varsa sütun olarak işaretlenir.

III. DENEYLER

Tablo algılama ve analizi için geliştirilen sistemin performansını ölçmek için, bulunan tablo alanı ile gerçek tablo alanı karşılaştırıldı ve aşağıdaki formül kullanıldı.

$$Performans(\%) = \frac{100 * kA * (bA + gA)}{2 * gA * bA} \quad (3)$$

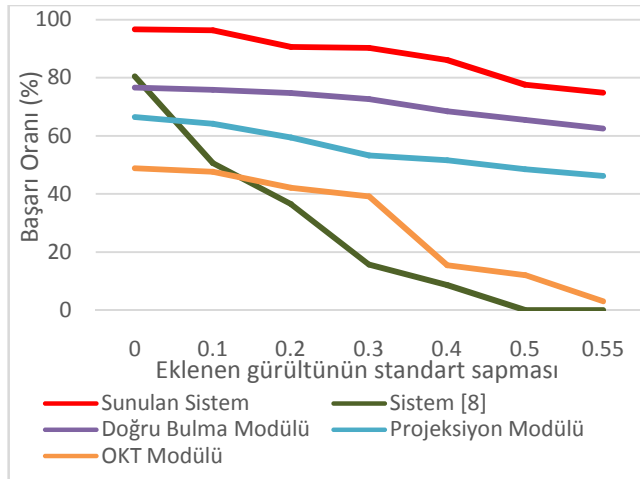
Burada, kA bulunan tablo ve gerçek tablo alanlarının kesişimleri olmak üzere, bA bulunan tablo alanını, gA ise gerçek tablo alanını ifade etmektedir. Sistem, 120 gerçek fatura üzerinde test edildi ve başarı oranı % 96.7 olarak

hesaplandı. Geliştirilen sistemin alt bileşenleri olan yöntemlerin, bu başarı oranını ne ölçüde etkilediği ile ilgili analizler yapıldı. Farklı parametre kümeleri ile elde edilen test sonuçları Tablo 1'de gösterilmektedir. Bu tabloda yer alan O1 ve O2 OKT Modülünün, P1 ve P2 Projeksiyon Modülünün, L1 ve L2 ise Doğru Bulma Modülünün iki farklı parametre kümesini ifade etmektedir. Tablolarda yer alan 1 sayısı, o yöntemin kullanıldığını; 0 ise o yöntemin kullanılmadığını göstermektedir. Her yöntemin birbiriyle beraber veya tek tek çalıştırılmasıyla elde edilen 64 farklı kombinasyonun, en iyi 16 sonucu Tablo 1.a'da gösterilirken, en kötü 16 sonuç ise Tablo 1.b'de gösterilmektedir.

L1	L2	P1	P2	O1	O2	Başarı Oranı(%)	L1	L2	P1	P2	O1	O2	Başarı Oranı(%)
1	1	1	1	1	1	96.7	0	0	1	1	1	1	65.7
1	1	1	1	1	0	96.4	1	0	0	0	1	0	65.4
1	1	1	0	1	1	94.8	0	0	0	1	1	1	59.8
1	1	1	0	1	0	94.5	0	1	0	0	1	0	59.1
1	0	1	1	1	1	94.2	1	0	0	0	0	1	52.8
1	1	0	1	1	1	93.9	1	0	0	0	0	0	52.6
1	0	1	1	1	0	93.9	0	0	1	1	0	1	49.9
1	1	0	1	1	0	93.6	0	0	0	0	1	1	49.6
1	0	1	0	1	1	92.3	0	0	1	1	0	0	49.6
1	0	0	1	1	1	91.4	0	0	0	0	1	0	48.8
0	1	1	1	1	1	91.3	0	1	0	0	0	1	29.7
1	0	0	0	1	1	89.5	0	0	1	0	0	1	28.5
0	1	1	0	1	1	89.4	0	0	1	0	0	0	28.2
0	1	0	1	1	1	88.5	0	1	0	0	0	0	25.4
1	1	1	1	0	1	88.3	0	0	0	1	0	1	19.7
1	1	1	1	0	0	88.0	0	0	0	1	0	0	19.4
1	0	0	1	1	0	86.6	0	0	0	0	0	1	3.4

Tablo 1. Yöntemlerin farklı kombinasyonlar ile çalıştırılması (a) en iyi 16 test sonucu (b) en kötü 16 test sonucu

Yapılan analizler sonucunda sistemin yüksek başarısı, tüm yöntemlerin bir arada ve 2 farklı parametre kümesi ile beraber çalıştırıldığı durumda elde edilmiştir. Yöntemlerin ayrı ayrı çalıştırılması ile elde edilen başarı oranlarına bakıldığında; sistemin genel başarısına en büyük katkının beklenildiği üzere, Doğru Bulma modülünün 1. parametre kümesi (L1 = 52.6) olduğu görüldü. Ardından en yüksek katkının OKT Modülünün 1. Parametre kümesi (O1= 48.8) ve onu takip eden Projeksiyon Modülünün 1. Parametre kümesi (P1= 28.2) olduğu analiz edildi.



Şekil 6. Sunulan yöntemlerin gürültüye karşı dayanıklılığının başka bir sistem ile karşılaştırılarak gösterilmesi

Geliştirilen sistemin gürültüye karşı ne kadar dayanıklı olduğu, faturalar üzerine belirli miktarlarda yapay Gauss gürültüsü eklenerek ve tablo algılama, analizi ile ilgili çalışma yapan başka bir sistem[8] ile karşılaştırılarak gösterildi. Aynı zamanda, sunulan sistem içerisinde kullanılan yöntemlerin gürültüye karşı dayanıklılıkları ayrı ayrı test edildi. Şekil 6'da, geliştirilen sistemin yüksek gürültü oranlarında da başarılı bir şekilde çalıştığı görülmektedir. [8], gürültüsü çok az olan faturalar üzerinde iyi sonuçlar verse de gürültüye karşı dayanıklı değildir. Diğer yandan geliştirilen sistem içerisinde kullanılan yöntemlerden gürültüye en dayanıklı olanın Doğru Bulma yöntemi ve ardından projeksiyon yöntemi olduğu görüldü. Yöntemlerden gürültüye karşı en hassas olanın OKT olduğu anlaşıldı.

IV. SONUÇLAR

Bu çalışmada fatura görüntüleri üzerinde tabloların bulunması ve analiz edilmesi için özgün bir metot sunulmaktadır. Metot farklı yöntemlerin, farklı parametre kümeleri ile birkaç defa çalıştırılıp, elde edilen sonuçların kaynaştırılması esasına dayanmaktadır. Geliştirilen sistemin başarısının yüksek düzeyde olduğu, gürültüye karşı ne kadar dayanıklı olduğu ve birden fazla yöntemin birkaç defa çalıştırılmasının yüksek başarı sağladığı, gerçek fatura örnekleri üzerinde yapılan deneylerle gösterilmiştir. Kullanılan yöntemlerin, sistemin başarısını hangi oranda etkilediği ve gürültüye karşı dayanıklılıkları ayrı ayrı test edilerek, analiz edilmiştir. Elde edilen %96.7 başarı oranı, pratik olarak bankacılık uygulamalarında kullanılması için yeterlidir. Ancak ileride yapacağımız çalışmalarda, analiz sonuçlarından edinilen bilgilere göre, yöntemlerde kullanılan parametre kümeleri sayısı artırılarak, daha güçlü bir sistem geliştirilecektir.

KAYNAKÇA

- [1] L. Hardy, "The Evolution of Participation Banking in Turkey." Al Nakhlah Online Journal of Southwest Asia and Islamic Civilization (2012).
- [2] Evgeniy Bart, "Parsing Tables by probabilistic modeling of perceptual cues", IAPR International Workshop on Document Analysis Systems, 2012
- [3] Y. Belaid and A. Belaid, "Morphological tagging approach in document analysis of invoices," in ICPR, 2004
- [4] D. W. Embley, M. Hurst, D. Lopresti, and G. Nagy, "Table processing paradigms: a research survey," IJDAR, vol. 8, no. 2, pp. 66-86, 2006
- [5] Yin, P.Y.: Skew detection and block classification of printed documents. Image and Vision Computing 19, (2001) 567-5
- [6] Li, J., Wang, K., Hao, S., & Wang, Q. (2012). Location and Recognition of Free Tables in Form. In Software Engineering and Knowledge Engineering: Theory and Practice (pp. 685-692). Springer Berlin Heidelberg
- [7] B. Gatos, D. Danatsas, I. Pratikakis and S. J. Perantonis, "Automatic table detection in document images", Proc. Intl. Conf. Advances Patt. Recog. pp. 609-618, 2005
- [8] KALYON, L. İlham; AKGUL, Yusuf Sinan. Document table detection and analysis using projection scale space. In: Signal Processing and Communications Applications Conference (SIU), 2014 22nd. IEEE, 2014. p. 1319-1322.
- [9] Matas, J. and Galambos, C. and Kittler, J.V., Robust Detection of Lines Using the Progressive Probabilistic Hough Transform. CVIU 78 1, pp 119-137 (2000)