

Türkçe Manzara Metni Veri Kümesi

Scene Text Dataset in Turkish

Nesli Erdoğan

Bilgisayar Mühendisliği Bölümü
İzmir Yüksek Teknoloji Enstitüsü
İzmir, Türkiye
neslierdoğmus@iyte.edu.tr

Özetçe —Manzara metni tespit etme ve tanıma gerçek dünya imgelerinden anlamlı içerik çıkarmada ve bu imgelerden oluşan büyük veri kümelerinde aranan metni bulup getirmede büyük kolaylık sağladığından araştırmacılardan artan bir ilgi görmektedir. Buna rağmen, bu alanda yaygın olarak kullanılan imge kümelerinin büyük kısmı İngilizce metinler içerdiklerinden yapılan çalışmalar tek dil ile kısıtlı kalmıştır. Buradan yola çıkarak, geliştirilen manzara metni tespit etme ve tanıma teknolojilerinin Türkçe metinlere uygulanması, performanslarının analiz edilmesi ve Türkçe'ye özgün algoritmaların geliştirilmesi amacıyla literatürde ilk kez bir Türkçe manzara metni veri kümesi oluşturulmuştur. Bu bildiriye, kısaca STRIT (Scene Text Recognition In Turkish) olarak adlandırılan veri kümesinin içeriği anlatılmaktadır. Buna ek olarak, Türkçe manzara tespiti ve tanınması için iki taban çizgisi yöntemi denenmekte ve ön sonuçlar takdim edilmektedir.

Anahtar Kelimeler—Türkçe, veri kümesi, manzara metni, tespit, tanıma.

Abstract—Scene text localization and recognition keeps attracting an increasing interest from researchers due to its valuable advantage in extracting content from real world images and in image retrieval via text search. Nevertheless, due to the fact that the majority of the image datasets that are commonly used in this field is comprised of text in English, the related studies have mostly been limited to a single language. On that account, in order to apply the technologies developed for scene text detection and recognition to Turkish scene text, analyze their performances and to develop Turkish language specific algorithms, a Turkish scene text database is collected for the first time in the literature. In this paper, the contents of this database, shortly called STRIT (Scene Text Recognition In Turkish), are detailed. Additionally, two baseline methods are tested to detect and recognize scene text in Turkish and the preliminary results are presented.

Keywords—Turkish, dataset, scene text, detection, recognition.

I. GİRİŞ

Mobil iletişim ve dijital görüntüleme olanaklarının günümüzde ulaştığı seviye sayesinde imge üretme ve paylaşma hızı her geçen gün artmaktadır. Bu gelişmeler ile akıl almaz boyutlara ulaşan görsel verinin idaresi ve kullanımı için görüntü anlamlandırma ve imgelerden doğru ve kesin içerik çıkarma ihtiyacı büyük önem kazanmıştır.

Renkleri, doku ve şekilleri kapsayan algısal içeriklere kıyasla, olaylar ve bunlar arasındaki ilişkileri içeren anlambilimsel içerikleri elde etmek daha zor ancak kullanmak daha

kolaydır. Bütün anlambilimsel içerik çeşitleri arasında ise metin bilgisi bünyesinde barındırdığı bilgi zenginliği ve buna rağmen sahip olduğu çıkarım kolaylığı ile diğerlerinden bir adım öne çıkmaktadır.

Farklı tipte imgelerden elde edilebilen metin bilgileri artan karmaşıklıklarına göre şu şekilde sıralanabilirler: Homojen bir arkaplan üzerinde metin içeren belge resimleri, kitap ya da CD kapağı gibi çok renkli belge resimleri ve ilk iki gruba dahil olmayan doğal manzara resimleri. Son grup daha detaylı iki alt-kategoriye ayrılır:

- Bir manzara resmine sonradan eklenen yapay metinler (örneğin altyazı)
- Aslen manzara resminin içinde bulunan ve manzara metni olarak adlandırılıp bu çalışmada ele alınan metinler (örneğin bir sokak resminde yer alan tabela yazıları)

Manzara metinleri diğer metin gruplarıyla kıyaslandığında açı değişikliklerine, perspektife ve örtülmeye bağlı bozulmalara daha fazla maruz kalabilmektedir. Bu sebepten dolayı, günümüzde yaygın olarak kullanılan ve özellikle belge resimlerinde yüksek başarımlarıyla sahip OCR (optical character recognition) yöntemleri manzara metinlerini okumada yetersiz kalmaktadırlar [2]. Bu da, manzara metinlerine özel yaklaşımların geliştirilmesi ihtiyacını doğurmaktadır.

Tüm yapay öğrenme alanlarında olduğu gibi, bilgisayarla görü ve imge anlamlandırma alanında da veri kümeleri, yeni tekniklerin geliştirilebilmesi ve denenmesi için anahtar bir role sahiptirler. Önerilen metotların çeşitli senaryolarda denenmesi, farklı yaklaşımların karşılaştırmalı değerlendirilmesi ve tekrarlanabilir araştırma sonuçlarının sunulması ancak araştırmacıların kullanımına açık umumi veri kümelerinin temini ile mümkün olur. Bu durum manzara metni tanıma konusu için de farklı değildir. Ancak, maalesef, manzara resimlerini tespit etme ve tanıma araştırmaları için yaygın olarak kullanılan imge kümelerinin büyük kısmı yalnızca İngilizce metinler içermektedirler. Bu durumda, Türkçe için bir manzara metni uygulaması istendiği takdirde, var olan yöntemlerin Türkçe'ye uyarlanması ve bu uyarlama sonucu bir performans analizi yapılması gerekecektir.

Bu ihtiyaç doğrultusunda, halihazırda önerilmiş olan manzara metni tespit etme ve tanıma teknolojilerinin Türkçe'ye uygulanması, performanslarının analiz edilmesi ve Türkçe'ye özgün yeni algoritmaların geliştirilmesi amacıyla literatürde ilk kez bir Türkçe manzara metni veri kümesi oluşturulmuştur. Ek

TABLEO I: MEVCUT MANZARA METNİ VERİ KÜMELERİ

Veri kümesi	Yıl	Büyükük	Amaç	Dil
ICDAR 2003 [16]	2003	181 eğitim + 251 test imgesi	tespit etme, tanıma	İngilizce
ICDAR 2005 [17]	2005	1001 eğitim +489 test imgesi	tespit etme, tanıma	İngilizce
Chars74k [13]	2009	74k+ doğal/sentetik imge	tanıma	İngilizce
KAIST [18]	2010	3000 imge	tespit etme, bölütleme, tanıma	Korece/İngilizce
Street View Text (SVT) [19]	2010	350 eğitim + 250 test imgesi	tespit etme	İngilizce
ICDAR 2011 [20]	2011	229 eğitim + 255 test imgesi	tespit etme, tanıma	İngilizce
NEOCR [21]	2011	659 imge	tespit etme, tanıma	Çoğunlukla Almanca ve İngilizce
MSRA Text Detection 500 Database (MSRA-TD500) [22]	2012	500 imge	tespit etme	Çince/İngilizce
StanfordSynth [23]	2012	62 karakter imgesi	tanıma	-
IIIT 5K-Words [24]	2012	2000 eğitim + 3000 test imgesi	tanıma	İngilizce
ICDAR 2013 [25]	2011	229 eğitim + 233 test imgesi	tespit etme, bölütleme, tanıma	İngilizce
Synthetic Word Dataset [26]	2014	90000 kelimenin 9 milyon sentetik imgesi	bölütleme, tanıma	İngilizce
ICDAR 2015 [27]	2015	1000 eğitim + 500 test imgesi	tespit etme, bölütleme, tanıma	İngilizce
COCO-Text [28]	2016	63686 imge	tespit etme, tanıma	İngilizce

olarak, bu veri kümesi kullanılarak bir metin tespit etme, bir de metin tanıma yöntemi denenmiş ve ön sonuçları taban çizgisi olarak sunulmuştur.

II. İLGİLİ ÇALIŞMALAR

Bu bölümde başlıca manzara metni tespit etme ve tanıma yöntemlerine değinilecek ve buna ek olarak bugün itibarı ile araştırmacıların kullanımına açık olan diğer manzara metni veri kümeleri ile ilgili bilgi verilecektir.

A. Mevcut Yöntemler

Manzara metni tespiti (yer tespiti ya da sezimi) manzara resimlerinde yer alan metin bölgelerini bulmayı amaçlar. Bu görev için önerilen mevcut teknikler ana hatlarıyla üç grupta incelenebilir: örüntü tabanlı metotlar, bileşen tabanlı metotlar ve melez metotlar [1].

Resimdeki tüm konum ve ölçeklerin incelendiği ve bu sebeple hesapsal olarak maliyetli olabilen örüntü tabanlı metotlar, metin görüntülerinin örüntüsel özelliklerini kullanarak imgedeki metin alanlarını bulmayı hedefler. [2], [3] Kayar bir pencerenin tüm resim üzerinde dolaştırılıp metin varlığının test edildiği doku temelli yaklaşımlar da bu sınıfa dahil edilebilirler. [4], [5]

Bileşen tabanlı metotlar için öncelikle resimdeki her piksel için belli yerel öznitelikleri hesaplanır. Ardından benzer özniteliklere sahip pikseller bir araya getirilir ve oluşturulan bağlantılı bileşenlerin bir metne ait olup olmadığına karar verilir. [6]–[9] Ölçek değişimsiz bu yöntemler ile karakter kesimlemesi de yapılabilir.

Son olarak melez metotlar, hem örüntü hem de bileşen tabanlı bir karışım olarak karşımıza çıkarlar. [10], [11] Bu yaklaşımlar iki türün de avantajlı yanlarına sahiptirler.

İmgede bir metin kutusu tespit edildikten sonra atılması gereken ikinci adım bu metni okumak ya da bir diğer deyişle tanımadır. Metindeki karakterler ayrı ayrı okunabileceği gibi [12], [13], tüm metin resmi için bir kelime de bulunabilir [14], [15]. Ancak ikinci yaklaşım için metin kutucuklarının eşleştirilebileceği bir kelimeler listesine (sözlük) ihtiyaç vardır.

B. Mevcut Veri Kümeleri

Yapay öğrenme alanındaki tüm araştırmalar için veri kümeleri büyük önem arz etmektedir. Manzara metinlerini otomatik

tespit etme ve tanıma yöntemleri geliştirilebilmesi için de birçok veri kümesi toplanmış ve araştırmacıların hizmetine sunulmuştur. Bunlardan yaygın olarak kullanılanlar hakkındaki bilgiler Tablo I’de özetlenmiştir.

Görüldüğü üzere son yıllarda manzara metni araştırmalarına paralel olarak toplanan ve yayımlanan veri kümelerinde ve bu kümelerin büyüklüklerinde dikkate değer bir artış olmuştur. Ancak istisnalar olmakla birlikte, bu veri kümelerindeki imgeler büyük oranda yalnızca İngilizce metinler içermektedirler. Bu tek dillilik geliştirilen yöntemlerin Türkçe’ye uygulanması ya da uyarlanması için bir engel teşkil etmektedir.

Öncelikle her ne kadar büyük oranda bir örtüşme varsa da Türkçe alfabede olan ve İngilizce alfabede olmayan ya da tam tersi İngilizce alfabede olan ve Türkçe alfabede bulunmayan harfler mevcuttur.

İkinci olarak, tespit edilen bir manzara metnini tanıran karşılaşılan belirsizlikleri ya da bölütleme hatalarını gidermede dil modellerinden sıkça yararlanılmaktadır. Bu nedenle Türkçe manzara metinleri için özellikle sözlük temelli bir yaklaşım geliştirilmesi için bu hedefe özgü bir veri kümesine ihtiyaç duyulmaktadır.

III. STRIT VERİ KÜMESİ

STRIT veri kümesinde bugün itibarı ile Canon EOS 550D fotoğraf makinesi ile çekilmiş 2592x1728 boyutlarında 30 manzara imgesi bulunmaktadır. Bu 30 imgede toplam 256 Türkçe kelime dikkate alınmış ve her bir kelime için metin kutucukları elle işaretlenmiştir.

Manzara metni konumları için yapılan elle işaretleme ICDAR denektaşı protokollerinden iki açıdan farklıdır. İlk olarak ICDAR veri kümelerinde metin kutucukları için onları sınırlayan dörtgenin yalnızca iki köşesi verilmektedir. Bu işaretleme, yatay olmayan metinlerin konumlarında belirsizliklere yol açmaktadır. STRIT veri kümesinde bulunan imgelerdeki manzara metinleri için sınırlayıcı kutucuklar herhangi bir yataylık kısıtı olmadan, dörtgenlerin tüm köşe koordinatları kaydedilerek işaretlenmişlerdir. Önemli bir nokta, istenildiği takdirde bu koordinatların ICDAR protokollerine uygun hale getirilebilmesidir. Bu uygulama nedeniyle, kırılmış kelime imgelerinin elde edilmesi için metin kutucuklarına kırpmaya ek olarak perspektif dönüşüm de uygulanmıştır. Fakat tabi ki istenildiği takdirde bu kırpmaya işlemi için de ICDAR protokollerine paralellik sağlanması için bir engel yoktur. (Şekil 1)



Şekil 1: STRIT veri kümesinden örnek imge ve veriler. Üstte: Orjinal bir manzara resmi ve üzerine metin kutucuğu işaretlenmiş hali. Ortada: Manzara imgesinden kırılmış metin imgeleri ve metin konum ve çevriyazı gerçek referans değerlerinin kaydedildiği doküman. Alttta: Veri kümesinden iki örnek daha.

İkinci fark olarak, sınırlayıcı dörtgenler işaretlenirken tiyografi elemanları dikkate alınmıştır. ICDAR veri kümelerinde olduğu gibi bu işlem için çoğunlukla benimsenen yaklaşım bir kelimenin tüm harflerinin tüm parçalarını içine alan en küçük dikdörtgeni seçmektir. Ancak bu yaklaşım beraberinde bir dezavantaj getirmektedir: Kelimelerin içinde yer alan harflere göre her harfin tek başına sınırlayıcı bir kutucuk içindeki konumu kelimedenden kelimeye değişebilmektedir. Örneğin, bir kelimenin işaretlenmesi içinde alt çıkıntıya taşan parçaları olan "y", "p"; ya da üst çıkıntıya taşan parçaları olan "S", "k" gibi karakterler barındırıp barındırmamasına göre farklılık gösterecektir. Bu belirsizliği ortadan kaldırma amacıyla, STRIT veri tabanında kelimedede yer alan harflerden bağımsız olarak sınırlayıcı kutucukların alt kenarı için taban çizgisi, üst kenarı içinse görünen ya da tahmini üst çıkıntı yüksekliği temel alınmıştır. (Şekil 2)

IV. STRIT İÇİN TESPİT VE TANIMA TABAN ÇİZGİSİ

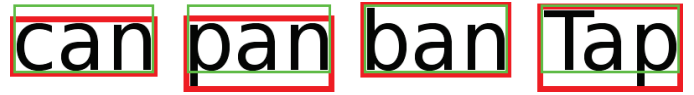
Bu bölümde, oluşturulan veri tabanında ön performans testleri yapmak için biri tespit biri tanıma amaçlı olmak üzere iki yöntem test edilmektedir.

A. Tespit

Manzara menti tespiti için imgeden Ölçek Değişimsiz Öznitelik Dönüşümü (ÖDÖD/SIFT) [29] anahtar noktaları çıkarılmış, bu anahtar noktalar yaklaşık en yakın komşularının oylamasıyla metin ve metin değil şeklinde sınıflandırılmışlardır. Metne ait olarak sınıflandırılan anahtar noktaları kullanarak yazı çizgilerinin geçtiği yerler ile ilgili hipotezler oluşturulmuş ve böylece metin bölgelerinin tayini sağlanmıştır. [30]

B. Tanıma

Manzara metni tanıma için sözlük temelli, daha açık olmak gerekirse dilde kullanılan sözcüklerin modellendiği bir yakla-



Şekil 2: Örnek sınırlayıcı kutucuklar. Kırmızı: ICDAR veri kümelerinde de kullanılan yaygın yaklaşım. Yeşil: STRIT için önerilen yaklaşım. Örnekte görüldüğü üzere, yaygın yaklaşımla farklı kelimelerde aynı "a" harfi kutucuk içerisinde 4 farklı konuma denk gelebiliyor. Önerilen yaklaşımda ile böyle bir değişkenlik söz konusu değil.

şım benimsenmiştir. Öncelikle, tespit edilen kelime kutucukları içinde ÖDÖD noktaları bulunmuş, çıkarılan noktalar yine en yakın komşuluk oylaması ile bu kez Türkçe alfabedeki harflerin, rakamların ve bazı noktalama işaretlerinin oluşturduğu semboller kümesinin bir elemanı olarak sınıflandırılmışlardır.

Bu sınıflandırma sonucu kutucuk içinde kayan bir pencere içinde her karakter için bir var olma olasılığı hesaplanmıştır. Bu olasılıklar görsel veriden gelen ipuçlarından türetilmiştir. Ancak, metin tanımda en az imge kadar önemli bir diğer veri kaynağı ise dilin yapısıdır. Bu sebeple ODTÜ Türkçe Bütünce'si [31] kullanılarak Türkçe'de kullanılan sözcükler içerisinde her sembolün hangi tek ve iki harfli alt sözcükleri takiben belirlediği incelenmiş ve istatistiksel olasılıkları hesaplanmıştır.

Son olarak, elde edilen Türkçe sözcük modelinin harf sınıflandırma modülüne bütünleştirilmesi için Saklı Markov Modeli (SMM) ile Viterbi algoritması kullanılmıştır. Model için sembol kümesindeki her karakter bir duruma karşılık gelmektedir. Bu durumlar direkt görülebilir değildir, ancak bu durumların tetiklediği gözlemlenebilir fakat gürültülü sinyaller mevcuttur. Bu problem özelinde gözlemler görsel karakter olasılıklarıdır. Durumlar arası geçiş olasılıkları içinse çıkarılan Türkçe sözcük modeli kullanılmıştır. Tespit edilen ve kırılan her metin resmi için durum dizisi Viterbi algoritması kullanılarak bulunmuştur.

V. DENEYLER VE SONUÇ

Seçilen taban çizgisi yöntemleri STRIT veri kümesi üzerinde test edilmiştir. STRIT yalnızca test için kullanılmış, eğitim için iki kaynaktan yararlanılmıştır: ICDAR 2011 eğitim veri kümesindeki, Türkçe'de bulunmayan karakterleri içermeyen kelime imgeleri ve sentetik olarak üretilen Türkçe metin imgeleri.

STRIT imge sayısı iyi bir eğitim için yeterli olmadığından ve Türkçe'ye özgü karakterleri içeren başka büyük bir veri kümesi bulunmadığından sentetik veri oluşturma yoluna gidilmiştir. Bu amaçla, ODTÜ Türkçe Bütünce'sinden rastgele seçilen kelimeler sonlarına bazı rakamlar ve noktalama işaretleri eklenerek farklı yeğinlikte arka planlar üzerinde ve farklı yeğinlikte renkler ile sentezlenmişlerdir. (Şekil 3)

Sentetik ve gerçek imgelerden oluşan eğitim kümesi oluşturulduktan sonra, eğitim imgelerinden ÖDÖD anahtar noktaları çıkarılmıştır. ICDAR veri kümesi elle işaretlendiğinden ve sentetik imgelerde harflerin basıldığı koordinatlar tam olarak bilindiğinden, çıkarılan bu anahtar noktalarının metne ait olup olmadığı, metne ait ise hangi karaktere ait olduğu bilgisi



Şekil 3: Üretilen sentetik Türkçe metin imgelerinden örnekler.

mevcuttur. Bu sayede, test imgelerinden çıkarılan ÖDÖD noktaları için eğitim kümesindeki yaklaşık en yakın komşuluklar bulunduğu ve noktalar ile ilgili çıkarım yapılabildiği. Elde edilen sonuçlar Tablo II'de verilmiştir.

Bu bildiri literatürde ilk kez toplanıp işaretlenen Türkçe manzara metni veri kümesi sunulmuş ve detayları açıklanmıştır. Buna ek olarak, Türkçe manzara tespiti ve tanınması için iki taban çizgisi yöntemi test edilmiş ve ön sonuçlar takdim edilmiştir. Toplanan veri kümesi tüm araştırmacı ve öğrencilerin kullanımına açıktır.

TABLO II: STRIT İÇİN METİN BULMA (1), KIRPILMIŞ KELİME TANIMA (2) VE BAŞTAN SONA MANZARA RESMİ TANIMA (3) SONUÇLARI

(1)	R, P ve F-1 skorları	%51.90	%50.28	%51.08
(2)	Ortalama düzenleme uzaklığı	2,7		
(3)	R, P ve F-1 skorları	%33.16	%81.67	%47.17

BİLGİLENDİRME

Bu çalışma BİDEB-114C025 nolu "Scene Text Recognition and Its Application in Turkish" başlıklı proje kapsamında TÜBİTAK tarafından desteklenmiştir.

KAYNAKLAR

- [1] Y. Zhu, C. Yao, X. Bai, "Scene Text Detection and Recognition: Recent Advances and Future Trends", *Frontiers of Computer Science* 10, no.1, 19-36, 2016.
- [2] X. Chen, A.L. Yuille, "Detecting and Reading Text in Natural Scenes", *IEEE Conference on Computer Vision and Pattern Recognition*, Vol. 2, pII-II, 2004.
- [3] K.I. Kim, K. Jung, J.H. Kim, "Texture-Based Approach for Text Detection in Images Using Support Vector Machines and Continuously Adaptive Mean Shift Algorithm", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(12), 1631-1639, 2003.
- [4] J.J. Lee, P.H. Lee, S.W. Lee, A. Yuille, C. Koch, "Adaboost for Text Detection in Natural Scene", *IEEE Conference on Document Analysis and Recognition*, 429-434, 2011.
- [5] C. Shi, C. Wang, B. Xiao, Y. Zhang, S. Gao, Z. Zhang, "Scene Text Recognition Using Part-Based Tree-Structured Character Detection", *IEEE Conference on Computer Vision and Pattern Recognition*, 2961-2968, 2013.
- [6] A. Mosleh, N. Bouguila, A.B. Hamza, "Image Text Detection Using a Bandlet-Based Edge Detector and Stroke Width Transform", *British Machine Vision Conference*, 1-12, 2012.
- [7] L. Neumann, J. Matas, "Real-Time Scene Text Localization and Recognition", *IEEE Conference on Computer Vision and Pattern Recognition*, 3538-3545, 2012.
- [8] C. Yao, X. Bai, W. Liu, Y. Ma, Z. Tu, "Detecting Texts of Arbitrary Orientations in Natural Images", *IEEE Conference on Computer Vision and Pattern Recognition*, 1083-1090, 2012.

- [9] W. Huang, Z. Lin, J. Yang, J. Wang, "Text Localization in Natural Images Using Stroke Feature Transform and Text Covariance Descriptors", *IEEE International Conference on Computer Vision*, 1241-1248, 2013.
- [10] L.I.U. Yangxing, T. Ikenaga, "A Contour-Based Robust Algorithm for Text Detection in Color Images", *IEICE Transactions on Information and Systems*, 89(3), 1221-1230, 2006.
- [11] Y.F. Pan, X. Hou, C.L. Liu, "A Hybrid Approach to Detect and Localize Texts in Natural Scene Images", *IEEE Transactions on Image Processing*, 20(3), 800-813, 2011.
- [12] T.E. de Campos, B.R. Babu, M. Varma, "Character Recognition in Natural Images", *International Conference on Computer Vision Theory and Applications*, 273-280, 2009.
- [13] L. Neumann, J. Matas, "Real-Time Lexicon-Free Scene Text Localization and Recognition", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38(9), 1872-1885, 2016.
- [14] T. Wang, D.J. Wu, A. Coates, A.Y. Ng, "End-to-End Text Recognition with Convolutional Neural Networks", *IEEE International Conference on Pattern Recognition*, 3304-3308, 2012.
- [15] V. Goel, A. Mishra, K. Alahari, C.V. Jawahar, "Whole is Greater than Sum of Parts: Recognizing Scene Text Words", *IEEE International Conference on Document Analysis and Recognition*, 398-402, 2013.
- [16] S.M. Lucas, A. Panaretos, L. Sosa, A. Tang, S. Wong, R. Young, "ICDAR 2003 Robust Reading Competitions", *IEEE Conference on Document Analysis and Recognition*, 682, 2003.
- [17] S.M. Lucas, "ICDAR 2005 Text Locating Competition Results", *IEEE Conference on Document Analysis and Recognition*, 80-84, 2005.
- [18] S. Lee, M.S. Cho, K. Jung, J.H. Kim, "Scene Text Extraction With Edge Constraint and Text Collinearity", *IEEE International Conference on Pattern Recognition*, 3983-3986, 2010.
- [19] K. Wang, S. Belongie, "Word Spotting in the Wild", *European Conference on Computer Vision*, 591-604, 2010.
- [20] A. Shahab, F. Shafait, A. Dengel, "ICDAR 2011 Robust Reading Competition Challenge 2: Reading Text in Scene Images", *IEEE Conference on Document Analysis and Recognition*, 1491-1496, 2011.
- [21] R. Nagy, A. Dicker, K. Meyer-Wegener, "NEOCR: A Configurable Dataset for Natural Image Text Recognition", *CBDAR Workshop at IEEE Conference on Document Analysis and Recognition*, 53-58, 2011.
- [22] C. Yao, X. Bai, W. Liu, Y. Ma, Z. Tu, "Detecting Texts of Arbitrary Orientations in Natural Images", *IEEE Conference on Computer Vision and Pattern Recognition*, 1083-1090, 2012.
- [23] Y. Netzer, T. Wang, A. Coates, A. Bissacco, B. Wu, A.Y. Ng, "Reading Digits in Natural Images with Unsupervised Feature Learning", *NIPS Workshop on Deep Learning and Unsupervised Feature Learning*, 2011.
- [24] A. Mishra, K. Alahari, C.V. Jawahar, "Scene Text Recognition using Higher Order Language Priors", *British Machine Vision Conference*, 2012.
- [25] D. Karatzas, F. Shafait, S. Uchida, M. Iwamura, L.G. i Bigorda, S.R. Mestre, J. Mas, D.F. Mota, J.A. Almazan, L.P. de las Heras, "ICDAR 2013 Robust Reading Competition", *IEEE Conference on Document Analysis and Recognition*, 1484-1493, 2013.
- [26] M. Jaderberg, K. Simonyan, A. Vedaldi, A. Zisserman, "Synthetic Data and Artificial Neural Networks for Natural Scene Text Recognition", *arXiv preprint arXiv:1406.2227*, 2014.
- [27] D. Karatzas, L. Gomez-Bigorda, A. Nicolaou, S. Ghosh, A. Bagdanov, M. Iwamura, J. Matas, L. Neumann, V.R. Chandrasekhar, S. Lu, F. Shafait, "ICDAR 2015 Competition on Robust Reading", *IEEE Conference on Document Analysis and Recognition*, 1156-1160, 2015.
- [28] A. Veit, T. Matera, L. Neumann, J. Matas, S. Belongie, "Coco-text: Dataset and benchmark for text detection and recognition in natural images", *arXiv preprint arXiv:1601.07140*, 2016.
- [29] D. G. Lowe, "Distinctive image features from scale-invariant keypoints", *International Journal of Computer Vision*, 60:2, 91-110, 2004.
- [30] N. Erdoğan, M. Özuysal, "Scene text localization using keypoints," *Signal Processing and Communications Applications Conference (SIU)*, 1917-1920, 2015.
- [31] K. Oflazer, B. Say, D.Z. Hakkani-Tür, G. Tür, "Building a Turkish Treebank", *Invited chapter in Building and Exploiting Syntactically-annotated Corpora*, Anne Abeille Editor, Kluwer Academic Publishers, 2003.