

Yapılandırılmamış Dijital Belgelerin Düzeninin Çözümlemesi İçin Bir Yaklaşım

An Approach to Analyzing the Layout of Unstructured Digital Documents

Mert ATAY^a, Meltem KALAYCI^b,
Hasan APİK^c, Vedat AYBAR^d
Mobildev

İstanbul, Türkiye
{^amert.atay, ^bmeltem.kalayci, ^chasan.apik,
^dvedat.aybar}@mobildev.com

Faruk SERİN^e, A. Oğuz AKYÜZ^f
^{e,f}Bilgisayar Mühendisliği Bölümü
^eMersin Üniversitesi, Mersin, Türkiye,
fserin@mersin.edu.tr
^fODTÜ, Ankara, Türkiye,
akyuz@ceng.metu.edu.tr

Özetçe— Kişisel verilerin korunabilmesi için bu verileri içeren belgelerin içerisinden kişisel bilgilerin çekilebilmesi oldukça önemlidir. Makine öğrenmesi, doğal dil işleme, görüntü işleme, optik karakter tanıma vb. alanlarda uygulanan yöntemler kullanılarak, standart bir yapıya sahip bir form veya belgenin içerisindeki kişisel bilgiler belirlenebilmektedir. Ancak belge standart bir yapıya sahip değilse, bu belgenin düzeninin çözülmesi ve belgenin içerisindeki farklı belge veya form yapılarının ayrı ayrı işlenmesi gerekebilmektedir. Literatürde belge düzeni çözümü için kullanılan farklı yöntemler bulunmaktadır. Ancak bu yöntemlerin her birinin kendine özgü parametreleri bulunmaktadır. Bu çalışmada, ortak parametre ayarlaması için yeni bir yaklaşım önerilmektedir. Önerilen yaklaşım ile gerekli parametre sayısı azaltılmış ve parametrelerin ayarlanma aralığı düşürülmüştür. Önerilen yaklaşım deneysel sonuçlarda %94 duyarlılık başarıları göstermiştir.

Anahtar Kelimeler — kişisel veri; kişisel veri tespiti; belge düzeni çözümü; görüntü tespiti; otomatik parametre tespiti.

Abstract—Extracting personal information from the documents is important to protect personal data. Personal information in a form or document with a standard structure can be determined by using the methods applied in the fields of machine learning, natural language processing, image processing, optical character recognition, etc. However, if the document does not have a standard structure, analyzing the layout of this document and processing the different document or form structures in the document separately may be necessary. There are different methods used for document layout analysis in the literature. However, each of these methods has its own parameters. In this study, a new approach for common parameter tuning is proposed. With the proposed approach, the number of required parameters is reduced and the adjustment range of the parameters is decreased. The proposed approach showed a recall of 94% in the experimental results.

Keywords — personal data; personal data detection; document layout analysis; image detection; automatic parameter estimation.

I. GİRİŞ

Belge düzeni çözümü (BDÇ), bilişsel bilgi işlem ve belge görüntülerinden bilgi çıkaran süreçlerde çok önemli bir adımdır. BDÇ belge anlama sistemlerinin ön işleme adımıdır. BDÇ belge alma, içerik sınıflandırma, metin tanıma ve benzeri gibi birkaç önemli uygulamaya sahiptir. BDÇ'nin amacı, belge homojen bloklarını tanımlayarak ve ilişkilerini belirleyerek sonraki analiz/tanıma aşamalarını kolaylaştırmaktır.

Kişisel veri, Kişisel Verilerin Korunması Kanunu (KVKK) kapsamında, bir kişiyi doğrudan veya dolaylı olarak tanımlayabilen, kişi hakkında özel veya genel bilgi içeren her türlü veri kişisel veri olarak ifade edilmektedir. 2016 yılından itibaren Türkiye'de yürürlükte olan 6698 sayılı KVKK ve Avrupa Birliği kapsamında yayınlanan GDPR (General Data Protection Regulation - Genel Veri Koruma Kanunu) ile birlikte kişisel verilerin güvenliğinin sağlanabilmesi için gerekli olan prensipler ve yaptırımlar düzenlenmiştir.

Kişisel verilerin korunabilmesi için bu verileri içeren belgelerin içerisinden kişisel bilgilerin çıkarılabilmesi ve/veya bilginin türünün belirlenip yerinin tespit edilmesi gerekmektedir. Makine öğrenmesi, doğal dil işleme, görüntü işleme, optik karakter tanıma vb. alanlarda uygulanan yöntemler kullanılarak, standart bir yapıya sahip bir form veya belgenin içerisindeki kişisel bilgiler belirlenebilmektedir. Ancak standart bir yapısı olmayan belgenin düzeninin çözülmesi ve belgenin içerisindeki farklı belge yapılarının ayrı ayrı işlenmesi gerekebilmektedir. Özellikle birden fazla belgenin beraber taranarak tek bir belge haline getirilmesi durumlarında, belgenin düzeninin çözülmesi ve her bir belgenin tespiti zorunluluk arz etmektedir. Bunu gerektiren bir diğer sebep de taranan belgelerin metin yönleri veya taranma düzeninin aynı yönde olmama durumlarıdır. Örnek olarak, beraber taranmış bir kimlik ve ikametgâh belgesini ele aldığımızda, kimlik yukarıdan aşağıya okunacak düzende iken ikametgâh belgesi de ters yerleştirilmiş olup, sağdan sola okuma düzeninde olabilmektedir. Bunların yanında, birden

fazla ve daha farklı açılardan taranmış belgeleri içeren daha karmaşık belgeler de olabilmektedir. Tüm bu durumlar, daha etkili ve etkin bir kişisel bilgi belirlemesi için belge düzeni çözümlenmesini zaruri bir işlem haline getirmektedir.

II. LİTERATÜR

Konuyla ilgili [1] çalışmasında, önceden belirlenmiş belgelerin sınıflandırması yapılmıştır. Burada, bir belge kendi içinde alt belgeler içermiyorsa belge tipleri başarıyla sınıflandırılabilir. Ancak belge alt belgeler içeriyorsa belgenin çoklu belge içerdiği tespiti yapılamamaktadır. Bu problemin çözülmesi için belgelerin içerisindeki alt belgelerin ayrıştırılması gerekmektedir.

Literatürde belge düzeni çözümü için kullanılan farklı yöntemler bulunmaktadır. Ancak bu yöntemlerin her birinin kendine özgü parametreleri bulunmaktadır. Bu parametrelerin bazıları şunlardır: yatay tarama eşik değeri, dikey tarama eşik değeri, morfolojik genişleme için yapısal elemanın boyut, genişleme uygulama sayısı, bağlı bileşenler için en küçük alan, bağlı bileşenler için en büyük alan. Bu parametrelerin, farklı düzene sahip her bir belge için yeniden ayarlanması gerekmektedir. Veri setindeki belgelerin yapısı homojen ise bu parametrelerin ayarlanması görece daha kolay olabilmektedir. Bu durumda birkaç deneme ile uygun parametreler tespit edilebilir. Ancak veri setindeki belgeler çok farklı yapılar içeren heterojen bir yapıya sahipse bu durumda belgeler için ortak parametre belirleyebilme durumu zorlaşmaktadır.

BDÇ ardışık düzeni, belgelerin düzenlerine ve son analiz hedeflerine bağlı olarak BDÇ yöntemleri arasında değişebilen birkaç aşamadan oluşur. Bu bağlamda, her türlü belge düzenine uyan veya tüm analiz hedeflerini karşılayan evrensel bir BDÇ algoritması henüz geliştirilmemiştir. Binmakhshen ve ark. yaptıkları bir çalışmada, farklı belge düzeni analiz tekniklerinin eleştirel bir çalışmasını sunmuşlar. Çalışma, BDÇ'yi takip etmenin motivasyon nedenlerini vurgulamaktadır ve BDÇ algoritmalarının farklı aşamalarını, alandaki araştırmaların gözden geçirilmesinin bir sonucu olarak oluşturulan genel bir çerçeveye dayanarak kapsamlı bir şekilde tartışmaktadır. BDÇ çerçevesi ön işleme, düzen analizi stratejileri, işlem sonrası ve performans değerlendirme aşamalarından oluşur. Makale, genel olarak belge düzeni analizinde daha fazla araştırma yapmak için temel bir yapı sunmaktadır [2].

Belge gösterimi, eşik oluşturma, eğritme algılama, geometrik düzen analizi ve mantıksal düzen çözümlemesi tekniklerini içerir. Belge anlama sisteminin performansı büyük ölçüde sayfa bölütleme doğruluğuna ve metin, tablo, resim, çizim ve cetvel gibi farklı bölgelerin etiketlenmesine bağlıdır. Jain ve ark. sayfa bölütleme ve bölge tanımlamasını verimli bir şekilde uygulamak için bağlı bileşen ayıklamaya dayalı geleneksel aşağıdan yukarıya yaklaşımını kullanmışlar [3].

Antonacopoulos ve ark. çalışmalarında, taranan geçmiş belgeler için düzen analizi yöntemlerinin objektif bir karşılaştırmalı değerlendirmesini sunmuşlar. Bu çalışmada, yöntemlerden biri bölgeleri doğru bir şekilde bölütlere ayırabilme yeteneğini, diğeri ise bölütleme ve bölge sınıflandırmasının tüm işlem hattını (metin ayıklama hedefiyle) değerlendiren iki senaryo raporlanmaktadır.

Sonuçlar, yaklaşımdaki bazı varyasyonlarla belirli bir metodolojiye yakınsama olduğunu göstermektedir. Bununla birlikte, tarihsel belgelerin özellikleriyle ilgilenen sağlam yöntemler geliştirmeye hala önemli bir ihtiyaç duyulduğu belirtilmektedir [4].

Bilgisayar görüşü için geliştirilen derin sinir ağlarının, belge görüntülerinin düzenini analiz etmek için etkili bir yöntem olduğu kanıtlanmıştır. Zhong ve ark. PubLayNet üzerinde eğitilen derin sinir ağlarının bilimsel makalelerin düzenini doğru bir şekilde tanıdığını ifade etmişler [5].

Oliveira ve ark. konvolüsyonel sinir ağlarına (CNN) dayalı metin, şekil ve tabloları göz önünde bulundurarak otomatik belge düzeni analizi için hızlı bir tek boyutlu yaklaşım önermişler. Çalışmada, boyut analizini iki boyutlu belge görüntülerinden 1B'ye indirmek için metin ve tablo bloklarında gözlenen tek boyutlu desenden yararlanarak genel performansın önemli ölçüde iyileştirildiği, klasik iki boyutlu CNN yaklaşımıyla karşılaştırıldığında genel doğrulukta kayıp olmadan önemli ölçüde daha hızlı yürütme süreleri ve daha kompakt veri kullanımı sunulduğu ifade edilmektedir [6].

Minouei ve ark. çalışmalarında, doğal sahne nesne dedektörlerinden ilham alan derin bir sinir ağı sunmuşlar. Ağ, büyük bir genel veri kümesinden etiketli örnekler kullanılarak eğitilmiş ve sınanmıştır. Sonuçlar, düzen analizi için nesne dedektörlerini kullanmanın potansiyelinin olduğunu göstermektedir [7].

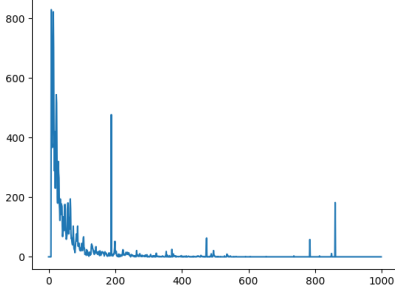
Li ve ark. farklı kategori bloklarını yerelleştirmek ve tanımlamak için belgelerin derin görsel, siyah görsel ve metin özelliklerini birleştiren bir VTLayOut modeli önermişler. Model esas olarak iki aşama içermektedir ve üç özellik çıkarıcı ikinci aşamada inşa edilmektedir. İlk aşamada, Cascade Mask R-CNN modeli doğrudan belgelerin tüm kategori bloklarını yerelleştirmek için uygulanmaktadır. İkinci aşamada, belgelerin kategori bloklarını tanımlamak için derin görsel, siyah görsel ve metin özellikleri ayıklanmaktadır. Çalışmada, deneysel sonuçlarda üstün bir başarının elde edildiği ifade edilmektedir [8].

Düzen analizi yöntemlerini değerlendirmek ve performanslarını ayrıntılı olarak incelemek için gerçekçi bir veri kümesine önemli bir ihtiyaç vardır. Antonacopoulos ve ark. makalelerinde, çok çeşitli çağdaş belgelere dayalı yeni bir veri kümesi ve onu oluşturmak için kullanılan metodoloji sunmaktadır [9].

III. YÖNTEM

Bu çalışmada, Wong ve ark. tarafından önerilen yatay tarama, dikey tarama ve genişleme tabanlı [10] yöntemde el ile tanımlanan yatay tarama ve dikey tarama eşik değerlerinin otomatik olarak belirlenmesini sağlayacak bir yaklaşım önerilmiştir. Önerilen yaklaşımda, öncelikle görüntüler ikili görüntü formatına dönüştürülür. Beyaz piksellerin 0'larla ve siyah piksellerin 1'lerle temsil edildiği ikili görüntü dizileri kullanılarak yatay ve dikey pikseller için iki ayrı histogram oluşturulur. Histogram hesaplamasında, görüntünün yatay ve dikey piksellerinde kaç adet ardışık 0 rakamının yer aldığı bilgisinden yararlanılmıştır. Örneğin 0 0 1 0 0 0 0 1 0 0 1 0 0 1 1 0 0 1 0 0 0 1 şeklindeki bir dizilimde 2 ardışık "0" 4 defa tekrarlanmıştır.

Şekil 1’de, veri setindeki örnek bir ardışık “0” sayısı dağılımı gösterilmektedir. Şekildeki sıfır noktasındaki yükseliş programda tanımlanan değişken dizisinin başlangıç değerlerinin 0 olmasından kaynaklanmaktadır. Bu yüzden göz ardı edilmelidir. Şekilde görüldüğü üzere, ardışık “0” sayısı bazı değerlerde zirve yapmakta ama genel olarak azalan bir eğilim izlemektedir.



Şekil 1. Veri setindeki örnek bir ardışık “0” sayısı dağılımı

Yatay tarama ve dikey tarama eşik değerlerinin hesaplanmasında satırlar ve sütunlar için oluşturulan histogramlar kullanılır ve bir ön işlem olarak histogramlar üzerinde aykırı değer analizi yapılır. Histogramların ortalama ve standart sapma değerleri hesaplanır ve bu değerler yardımıyla Denklem (1) ve Denklem (2) kullanılarak aykırı değer tespitinde kullanılan üst sınır değerleri hesaplanır. Histogramlarda, otomatik olarak hesaplanan üst sınırı aşan değerler sıfırlanır ve sıfır değerleri göz ardı edilerek ortalama ve standart sapma değerleri yeniden hesaplanır. Aykırı değerler göz ardı edilerek tekrar hesaplanan ortalama ve standart sapma değerlerinin yardımıyla Denklem (3) ve Denklem (4) kullanılarak yatay tarama ve dikey tarama eşik değerleri hesaplanır.

Burada; $l_y, l_d, t_y, t_d, \mu_y, \mu_d, \mu'_y, \mu'_d, \sigma_y, \sigma_d, \sigma'_y, \sigma'_d, \alpha, \beta, \theta$ sembolleri sırasıyla yatay üst sınır, dikey üst sınır, yatay eşik değeri, dikey eşik değeri, yatay ortalama, dikey ortalama, aykırı değerler göz ardı edilerek hesaplanan yatay ortalama, aykırı değerler göz ardı edilerek hesaplanan dikey ortalama, yatay standart sapma, dikey standart sapma, aykırı değerler göz ardı edilerek hesaplanan yatay standart sapma, aykırı değerler göz ardı edilerek hesaplanan dikey standart sapma, yatay ağırlıklandırma katsayısı, dikey ağırlıklandırma katsayısı, üst limit ağırlıklandırma katsayısını ifade etmektedirler.

$$l_y = \mu_y + \theta * \sigma_y \quad (1)$$

$$l_d = \mu_d + \theta * \sigma_d \quad (2)$$

$$t_y = \mu_y + \alpha * \sigma'_y \quad (3)$$

$$t_d = \mu_d + \beta * \sigma'_d \quad (4)$$

Bu işlemleri uygulamadaki motivasyon kaynağı, belirli bir belgeyi oluşturan parçaların arasındaki yatay ve dikey boşluklar ile belgeyi oluşturan içerikler arasındaki tutarlı bir ilişkinin olması gerektiği düşüncesidir. Bir diğer motivasyon

kaynağı ise, bir belgenin içerisine yerleştirilmiş olsa da farklı belgeleri oluşturan parçaların arasındaki boşluklar ile belirli bir belgeyi oluşturan içerikler arasındaki boşlukların birbirinden farklı bir yapıya sahip olması gerektiği düşüncesidir. Bu bilgilerden yola çıkarak, farklı belgeleri içerisinde barındıran bir belgenin düzen çözümlemesi için belgenin kendi öz niteliklerinden yararlanarak ayarlanması gereken parametrelerin azaltılabileceği veya parametre ayarlama aralığının düşürülebileceği fikri bu çalışmayı yapmamızı sağlamıştır. Çalışmanın sonucunda, gerekli parametre sayısı azaltılmış ve parametrelerin ayarlanma aralığı düşürülmüştür.

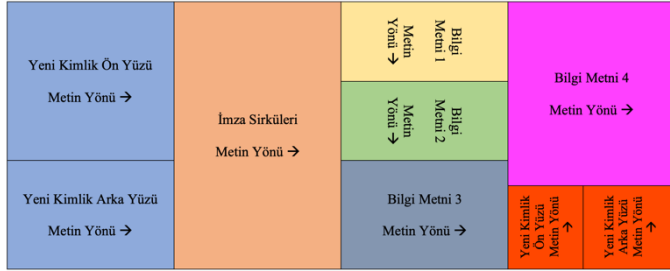
IV. DENEYSEL SONUÇLAR

Bu çalışmada, genellikle kimlik fotokopisi, ikametgâh belgesi, pasaport fotokopisi gibi kişisel veriler bulunduran dokümanları içeren belgeler üzerine odaklanılmıştır. Burada belgeler içerisindeki alt belgelerin parçalanmadan tespit edilebilmesi önemli bir kriter olacaktır. Yapılan çalışmalarda genellikle farklı amaçlar doğrultusunda farklı tablo, resim, metin yapıları içeren veri setleri üzerinde deneyler yapılmış ve veri setleri kısmen paylaşılmıştır. Ancak, kişisel verilerin korunması hassasiyeti nedeniyle, kişisel veriler içeren dokümanları bulunduran belgeleri paylaşan geniş kapsamlı bir veri setine ulaşılamadığı için elde edilen birtakım belgeler ve sentetik veriler üzerinde çalışılmıştır. Gizlilik gereğiyle, üzerinde çalışılan belgelerin içerisindeki bilgiler yerine belgelerin düzen yapısı sunulmuştur.

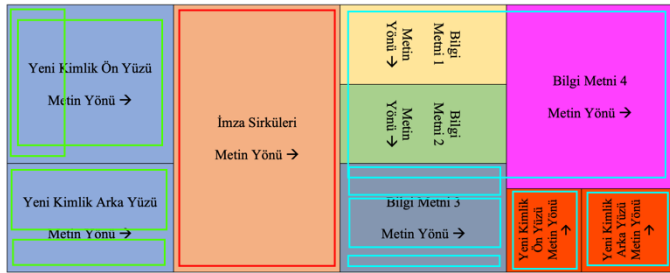
Şekil 2’de, 7 farklı yapıya ve 3 farklı metin okuma yönüne sahip veri setindeki en karmaşık belgelerden birinin temsili gösterimi sunulmaktadır. Şekil 3’te de Şekil 2’deki düzene sahip belgenin önerilen yöntemin uygulanması sonucu ayrıştırılmış düzeni gösterilmektedir. Burada görüldüğü üzere, imza sirküleri, yeni kimlik ön ve arka yüzleri ile bilgi metni 3 herhangi bir bilgi kaybı olmadan ayrıştırılmıştır. Bilgi metni 3’te bulunan küçük kutucuklar bilgi içermeyen gürültüden dolayı yapılmış ayrıştırmalardır. Bunlar göz ardı edilmiştir. Yeni kimlik arka yüzü iki parça olarak ayrıştırılmıştır. Ancak burada kişisel bilgiyi içeren metin kısmı büyük parça içerisinde ayrıştırılmıştır. Yeni kimlik ön yüzü de iki parça olarak ayrıştırılmıştır. Burada her ne kadar bilginin çoğunluğu büyük parça içerisinde olsa da bilgi bütünlüğü kısmi olarak bozulmuştur. Bilgi metni 1, 2 ve 4 tek parça halinde tespit edilmiştir. Ancak bu metinlerin okunma yönleri farklı olduğu için bilgi bütünlüğü bozulmuştur. Şekildeki dikdörtgenlerin çakışması görsel olarak durumu izah edebilmek için yapılan çizimden dolayıdır. Göz ardı edilebilir.

Tablo 1’de görüldüğü üzere, $\alpha = 2, \beta = 2, \theta = 3$ ve genişleme sayısı 5 alındığında belgeler yaklaşık %94 duyarlılıkla ayrıştırılabilmektedir. Bu analizler kişisel veriler içeren birden fazla alt belgeyi barındıran ve karmaşık düzene sahip 15 adet belge üzerinde yapılmıştır. Duyarlılık hesaplaması için bir belgenin içerisindeki alt belgelerin nasıl ayrıştırılması gerektiği konusunda uzman kişilerin belirlediği çıktılar ile algoritmanın ürettiği sonuçlar değerlendirilmiştir. Burada doğrudan belge düzeninden ziyade bilgi bütünlüğünün korunarak belgenin ayrıştırılması üzerine odaklanılmıştır. Bir kimlik veya formatı tanımlanmış bir belgenin hedeflenen bilgileri birden fazla parçaya ayrılmışsa burada bilgi

bütünlüğü bozulmuştur. Benzer bir şekilde bir metin okuma yönünün dikine parçalanmışsa bilgi bütünlüğü bozulmuştur. Ancak bir metnin bazı bölümleri okuma yönünde parçalanmışsa esas bilgi bütünlüğünün bozulmadığı ifade edilebilir. Bu değerlendirme sonuçları önerilen yaklaşımın uygulanabilirliğini göstermektedir.



Şekil 2. Çoklu düzene sahip örnek gerçek bir belgenin gösterimi



Şekil 3. Şekil 2'deki düzene sahip belgenin ayrıştırılmış düzeni.

TABLO I. FARKLI PARAMETRE DEĞERLERİYLE ÜRETİLEN SONUÇLAR

α	β	θ	#genişleme	Duyarlılık
1	2	3	3	0.67
1	1	3	2	0.71
1	1	3	5	0.81
1	1	3	7	0.84
1	1	5	3	0.73
2	2	3	3	0.80
2	2	3	5	0.94
3	3	3	3	0.74
3	3	3	5	0.83
3	3	5	3	0.80
3	3	5	5	0.73

V. SONUÇLAR

Kişisel bilgiler içeren belgelerin tespiti bu bilgilerin korunması açısından oldukça önemlidir. Çoğu durumda bu belgeler çoklu bir şekilde beraber taranırlar. Başarılı kişisel veri tespiti için bu belgelerin yeniden ayrıştırılması gerekebilmektedir. Bunun için belge düzen çözümü yöntemleri kullanılabilir. Bu çalışmada, bu amaçla kullanılan bir yöntemin görüntü çözünürlüğünden bağımsız olarak parametrelerinin nasıl hesaplanabileceği üzerine bir

yaklaşım önerilmiştir. Böylece parametre ayarlama aralığı düşürülmüş ve bu ayarlamalar yaygın istatistiksel temellere dayandırılmıştır. Önerilen yaklaşımın başarılı olduğu deneysel olarak gösterilmiştir.

Gelecek çalışmalarda, bazı belgeler üzerindeki başarıyı arttırmak için belge görüntüleri üzerinde belgenin öz bilgilerinden yararlanılarak ön işlem yaklaşımları geliştirilebilir ve daha büyük veri setleri üzerinde çalışılabilir.

BİLGİLENDİRME

Bu çalışma, 1501-TÜBİTAK Sanayi AR-GE Projeleri Destekleme Programı tarafından desteklenen 3210172 numaralı projenin bir parçasıdır.

KAYNAKLAR

- [1] B. Bakar, F. Aksoy, A. Yayık, S. İçöz, V. Aybar, and T. Güngör, "Turkish Rule-Based Official Document Type Detection," in *2020 28th Signal Processing and Communications Applications Conference (SIU)*, Oct. 2020, pp. 1–4. doi: 10.1109/SIU49456.2020.9302159.
- [2] G. M. Binmakhshen and S. A. Mahmoud, "Document Layout Analysis: A Comprehensive Survey," *ACM Comput. Surv.*, vol. 52, no. 6, p. 109:1-109:36, Oct. 2019, doi: 10.1145/3355610.
- [3] A. K. Jain and B. Yu, "Document representation and its application to page decomposition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 3, pp. 294–308, Mar. 1998, doi: 10.1109/34.667886.
- [4] A. Antonacopoulos, C. Clausner, C. Papadopoulos, and S. Pletschacher, "Historical Document Layout Analysis Competition," in *2011 International Conference on Document Analysis and Recognition*, Sep. 2011, pp. 1516–1520. doi: 10.1109/ICDAR.2011.301.
- [5] X. Zhong, J. Tang, and A. Jimeno Yepes, "PubLayNet: Largest Dataset Ever for Document Layout Analysis," in *2019 International Conference on Document Analysis and Recognition (ICDAR)*, Sep. 2019, pp. 1015–1022. doi: 10.1109/ICDAR.2019.00166.
- [6] D. Augusto Borges Oliveira and M. Palhares Viana, "Fast CNN-Based Document Layout Analysis," 2017, pp. 1173–1180. Accessed: Feb. 16, 2022. [Online]. Available: https://openaccess.thecvf.com/content_ICCV_2017_workshops/w18/html/Oliveira_Fast_CNN-Based_Document_ICCV_2017_paper.html
- [7] M. Minouei, M. R. Soheili, and D. Stricker, "Document Layout Analysis with an Enhanced Object Detector," in *2021 5th International Conference on Pattern Recognition and Image Analysis (IPRIA)*, Apr. 2021, pp. 1–5. doi: 10.1109/IPRIA53572.2021.9483509.
- [8] S. Li, X. Ma, S. Pan, J. Hu, L. Shi, and Q. Wang, "VTLayout: Fusion of Visual and Text Features for Document Layout Analysis," in *PRICAI 2021: Trends in Artificial Intelligence*, Cham, 2021, pp. 308–322. doi: 10.1007/978-3-030-89188-6_23.
- [9] A. Antonacopoulos, D. Bridson, C. Papadopoulos, and S. Pletschacher, "A Realistic Dataset for Performance Evaluation of Document Layout Analysis," in *2009 10th International Conference on Document Analysis and Recognition*, Jul. 2009, pp. 296–300. doi: 10.1109/ICDAR.2009.271.
- [10] K. Y. Wong, R. G. Casey, and F. M. Wahl, "Document Analysis System," *IBM J. Res. Dev.*, vol. 26, no. 6, pp. 647–656, Nov. 1982, doi: 10.1147/rd.266.0647.