Osmanlı Nüfus Kayıtlarının Otomatik Yerleşim Analizi ile İncelenmesi

Developing an Automatic Layout Analysis System for Ottoman Population Registers

Yekta Said CAN and M. Erdem Kabadayı

Koc Üniversitesi

Rumelifeneri Yolu 34450

Sarıyer, İstanbul, Türkiye

Özetçe —Tarihsel belgelerden bilgi çıkarmak için sayısallaştırma çabaları son yıllarda önemli ölçüde artmıştır. Doğru yerleşim analizi, araştırmacıların bu belgelerden anlamlı bilgiler elde edecek daha sağlam HTR ve OCR teknikleri geliştirmelerine yardımcı olacaktır. Değişken yerleşimler, düşük kaliteli ve bozuk tarihi belgeler, modern belge işlemeye kıyasla farklı sorunlar yaratmaktadır. Arapça yazılar kendine has bazı özelliklerinden dolayı otomatik işlem sistemleri için daha da fazla sorun yaratmaktadır. Bu çalışmada, ilk Osmanlı nüfus kayıtlarının Arap harfleriyle yazılmış yerleşimlerini otomatik olarak analiz etmek için bir araç geliştirdik. 1840'lar ve 1860'lar arasında Osmanlı İmparatorluğu'nun ilk nüfus kayıtlarından seçilen sistemimizin performansını test etmek için bir veri seti oluşturduk. Bu belgelerde iki farklı nesne türünü başarıyla sınıflandırdık.

Anahtar Kelimeler—sayfa bölütleme, tarihi döküman analizi, evrişimsel sinir ağları, arapça yazı yerleşim analizi

Abstract—For extracting information from the historical documents, digitization efforts have increased dramatically in the recent decades. Accurate layout analysis will help researchers for developing more robust HTR and OCR techniques which will extract meaningful information from these documents. Variable layouts, low quality and distorted images of historical documents create different problems to deal with when compared to modern document processing. Arabic script features have even more problems for these automatic processing systems. In this study, we have developed a tool for automatically analyzing the layouts of the first Ottoman population registers which are written in Arabic script form. We built a dataset for testing the performance of our system which are chosen from the first population records of the Ottoman Empire between the 1840s and 1860s. We successfully classified two different object types in those documents.

Keywords—page segmentation, historical document analysis, convolutional neural networks, Arabic layout analysis

I. GİRİŞ

Tarihi belgeler, geçmişin tarihi, sosyal ve ekonomik yönlerinin incelenmesini sağlayan değerli kültürel kaynaklardır.

Bunların sayısallaştırılması, araştırmacılara ve halka bu arşivlere anında erişim sağlar. Ancak, bakımlar nedeniyle, bunlara erişim mümkün olmayabilir veya sınırlandırılabilir. Bahsedilen dijitalleşme süreçlerinden sonra bu belgelerden yeni bilgiler analiz edilebilecek ve çıkarılabilecektir. Geçmiş belgelerin dijitalleştirilmesi için farklı alanların sayfa bölümlenmesi, belge analizi ve optik karakter tanıma (OCR) [1] için kritik bir süreçtir. Sayfa bölütleme teknikleri, görüntüyü arka planlar, metinler, grafikler, süslemeler vb. farklı bölgelere bölerek belgeyi analiz eder [2]. Tarihsel belge bölütlemesi, belge görüntülerinin bozulması, sayısallaştırma hataları ve değişken yerleşim türleri nedeniyle daha zordur. Bu nedenle, projeksiyon tabanlı veya kural tabanlı yöntemler uygulayarak bunları segmentlere ayırmak zordur [2].

Sayfa bölütleme hatalarının, el yazısı veya basılı metni sayısallaştırılmış karakterlere dönüştüren OCR çıktısı üzerinde doğrudan etkisi vardır [1]. Bu nedenle, tarihsel dokümanlar için sayfa bölütleme teknikleri doğru sayısallaştırma için önem kazanmaktadır. Özellik çıkarma ve sınıflandırıcı algoritma tasarımı, sayfa bölütleme yöntemlerinin performansı için çok önemlidir. Her ne kadar belge görüntü analizi daha geleneksel makine öğrenimi sınıflandırıcıları ile başlamış olsa da, Evrişimsel Sinir Ağlarının (CNN) ortaya çıkmasıyla birlikte, çoğu çalışma daha iyi performansları nedeniyle bunları kullanmaktadır. Arapça yazılar, örneğin Osmanlı, Arapça, Urduca, Kürtçe, Farsça [3] sayfa bölütleme prosedürünü zorlaştıran farklı şekillerde yazılabilmektedir. Arapça yazılar ayrıca sayfa bölütleme işlemlerinde daha da fazla zorluğa neden olan bağlı harflerin bitişik harfler oluşturduğu el yazısı betikleri, noktalar ve hareke işaretleri içerebilir [3].

Bu çalışmada, köylerde kayıtlı Osmanlı nüfusunu saymak için sayfaları otomatik olarak bölümlere ayıran ve nesneleri tanıyan bir yazılım geliştirdik. Verilerimiz, Osmanlı İmparatorluğu'nun 1840'larda gerçekleşen ilk nüfus kayıtlarından gelmektedir. Bu kayıtlar, imparatorluğun her bir erkek vatandaşının yaşı, etnik veya din, askeri ve mali durumunu kaydetmeyi amaçlayan ilk idari operasyonun sonuçlarıdır. Erkek nüfusu için evrensel bir kapsama sahip olmak hedeflenmesinden dolayı bu kayıtlara (proto-) sayımlar denilebilir. Bu kayıtların coğrafi kapsamı, 19. yüzyılın ortalarında, bugün Güneydoğu Avrupa ve Orta Doğu'da yaklaşık iki düzine halef devle-

tinin topraklarını kapsayan tüm Osmanlı İmparatorluğu'dur. Bu çalışma için iki konuma odaklanılmıştır: Türkiye'de Batı Anadolu'da bulunan İznik ve Bulgaristan'da bir Tuna kıyı kasabası olan Zıstovi. Bu sayımlarda memurlar el yazması veya basılı tablo kullanmadan dökümanları el yazısı ile hazırladılar. Ayrıca, önceden belirlenmiş bir sayfa yapısı yoktu. Sayfa düzenleri farklı bölgelerde farklılık gösterebilmektedir. Memura bağlı olarak yapısal değişiklikler de görülebilmektedir. Denetimli öğrenme algoritmalarına girdi olarak vermek üzere etiketli bir veri kümesi oluşturduk. Bu veri kümesinde, farklı bölgeler ve nesneler farklı renklerle işaretlendi. Daha sonra tüm pikseller sınıflandırıldı ve aynı piksel türlerinden oluşan bölgeleri bağlandı. Bu sayfalarda köyün başlangıç noktalarını ve kişi nesnelerini sınıflandırarak tüm köy ve sayfalardaki insan sayısını saydık. Sistemimiz farklı köylerdeki nüfusu başarıyla saymıştır.

Makalenin kalan kısımlarının yapısı aşağıdaki gibidir. İkinci bölümde, tarihsel doküman analizindeki ilgili çalışmalar gözden geçirilecektir. Oluşturulan veritabanının Bölüm 3'te yapısını tarif ettik. Sayfa bölütlemesi ve nesne tanıma yöntemimiz Bölüm 4'te açıklanmıştır. Deney sonuçları ve tartışma Bölüm 5'de sunulmaktadır. Çalışmanın sonuç ve gelecekteki çalışmalarını ise Bölüm 6'da sunuyoruz.

II. İLGİLİ ÇALIŞMALAR

Belgelerde görüntü analizi çalışmaları 2000'li yılların başında başlamıştır. Laven ve diğ. [4] istatistiki öğrenme tabanlı sayfa bölütleme sistemi geliştirdi. Akademik dergilerin 932 sayfalık görüntülerini içeren bir veri kümesi oluşturdular ve fiziksel yerleşim bilgilerini elle etiketlediler. Bir lojistik regresyon sınıflandırıcısı kullanarak, 25 etiketle yaklaşık %99 doğruluk elde ettiler. Bölütleme algoritması, XY kesim algoritmasının [5] bir varyasyonuydu. Arapça belge yerleşim analizi de literatürdeki geleneksel algoritmalar ile incelenmiştir. Hesham ve diğ. [3] Arapça belgeler için otomatik bir düzen algılama sistemi geliştirdi. Sauvola ikileştirme [6], gürültü filtreleme ve çarpıklık düzeltme algoritmaları uyguladıktan sonra, metin ve metin olmayan bölgeleri Destek Vektör Makinesi (SVM) algoritmasıyla sınıflandırdılar.

Sinir ağlarının ortaya çıkmasından sonra Arapça belge analiz sistemleri üzerinde de test edilmiştir. Buhari ve diğ. [7] bir otomatik yerleşim algılama sistemi geliştirdi. MultiLayer Perceptron (MLP) algoritmasını kullanarak ana gövdeyi ve yan metni sınıflandırdılar. Kudüs'ün eski kentindeki özel bir kütüphaneden 38 tarihi belge görüntüsünü içeren bir veri kümesi oluşturdular. % 95 sınıflandırma doğruluğu elde ettiler. Evrimsel Sinir Ağı aynı zamanda görüntü işleme uygulamalarının çoğunda kullanılabilen bir tür derin sinir ağıdır [8]. CNN ve Uzun Kısa Süreli Bellek (LSTM), İngilizce olarak yazılmış bilimsel dergi makalelerinin belge düzeni analizi için kullanılmıştır [9] ve [10]. Amer ve diğ. [11] Arapça gazeteler ve Arapça basılı metinler için CNN tabanlı bir belge düzeni analiz sistemi önerdi. Metin ve metin olmayan bölgeleri bulmada yaklaşık % 90 doğruluk elde ettiler. CNN'ler ayrıca tarihi belgeleri bölümlere ayırmak için de kullanılmıştır [12], [2] ve [1]. Fakat bilginiz dahilinde bu çalışma CNN tabanlı tarihi Arapça belge yerleşim analizi (bölütleme ve nesne tanıma) yapan literatürdeki ilk çalışmadır.

Şekil 1: Kayıtların üç örnek sayfası üç farklı bölgeye aittir. Sayfaların yerleşimi ilçeler arasında değişebilmektedir.

III. KAYITLARIN YAPISI

Vaka çalışmamızda kullandığımız İznik ve Zıstovi ilçe kayıtları NFS.d. 1411, 1452 ve NFS.d. 6314, sırasıyla Türkiye Cumhuriyeti Cumhurbaşkanlığı Devlet Arşivi - Osmanlı Arşivi Bölümü'nde jpeg formatında mevcuttur. 1840'lar ile 1860'lar arasında hazırlanan imparatorluğun çeşitli bölgelerinden benzer kayıtların da etkin bir şekilde okunması için uygulanacak bir metodoloji geliştirmeyi hedeflemekteyiz. Yukarıda belirtildiği gibi bu kayıtlar, hane halkı mensupları erkekler hakkında aile ilişkileri, yaş ve meslek bilgilerini içermektedir. Kadınlar hanelerde kayıtlı değildi. Toplam sayıları 11.000 civarında olan Osmanlı devletine ait kayıtlar, 2011 yılında araştırmaya açılmıştır. Şimdiye kadar herhangi bir sistematik çalışmaya tabi tutulmamışlardır. Sadece bireysel kayıtlar parça parça bir şekilde yazıya çevrilmiştir. Kayıtların dijital imgeleri genellikle 2100x3000 piksel boyutu civarındadır. Daha önce

Şekil 2: Bir örnek kayıt sayfası ve etiketli sürümü gösterilmiştir. Farklı renkler farklı nesne türlerini temsil etmektedir.

de belirtildiği gibi, bu kayıtların yerleşimi bölgeden bölgeye değişebilir (bkz. Şekil 1), bu da görevimizi daha karmaşık hale getirmektedir. Bu çalışmada, bu belgelerin genel özellikleri üzerinde çalışıyoruz. İlk nesne, köy başlangıç sembolüdür. Bu sembol birçok ilçede kullanılır ve bir önceki köyün sonunu ve yeni köyün başlangıcını işaretleyebilir. Kayıtlarda kalan kümeler nüfus sayımında sayılan kişilerdir ve bunlar kişiler hakkında demografik bilgiler içerir. Bu kayıtlarda, askerlik hizmetine gidenleri ya da ölen kişileri belirten güncellemeler de vardır. Memurlar genellikle güncellenen bireyin kaydının üzerine bir çizgi çizer ve bazen bireyi, bölütleme algoritmasında bazı hatalara neden olabilecek bitişik bir çizgiyle komşu bireye bağlayabilir.

IV. OSMANLI NÜFUS KAYITLARININ SAYILMASI İÇİN GELİŞTİRİLEN OTOMATİK SAYFA BÖLÜTLEME VE NESNE TANIMA SİSTEMİ

A. Veri Kümesi Oluşturma

DhSegment yazılımını [12] kullanabilmek için etiketli bir veri kümesi oluşturduk. 4 farklı sınıf oluşturuldu. Birincisi, sayfa kenarlıkları ile belge kenarlıkları arasındaki bölge olan arka plandır. Bu bölgeyi siyah olarak işaretledik. İkinci sınıf sayfa bölgesidir ve mavi ile işaretlenmiştir. Üçüncüsü köy başlangıcıdır ve onu yeşille boyadık. Sonuncusu bireysel kayıtlardır ve biz onları kırmızı ile işaretledik. 173 sayfayı açıklanan etiketlerle işaretledik. Bunların 51'i Zıştovi bölgesine, 122'si İznik bölgesine aittir. Örnek bir orijinal görüntü ve etiketli versiyon Şekil 2'de gösterilmiştir.

B. CNN Mimarisi Eğitimi

Sistemimiz için bir CNN eğitmek için dhSegment [12] yazılımını kullandık. Bu yazılım, önceden hazırlanmış Resnet-50 mimarisini [13] kullanarak bir sistemi eğitmiştir. Yazılım hem bir örnek seyreltme katları (Resnet-50 [13] içindeki derin artık ağı takip eder) hem de düşük çözünürlüklü özellikleri orijinal yüksek çözünürlüklü özelliklerle eşleyen örnek çoğaltma katlarına sahiptir (genişleyen ve daralan yollar için terminolojiye bakılabilir [14]) [12]. Genişleyen yol beş bloktan ve piksel sınıflandırması için bir evrişimsel katmandan oluşur ve her bir evrişimsiz adım bir görüntünün büyütülmesinden, özellik haritasının bir büzülme, 3x3 evrişimsel ve bir Relu tabaka bloğuyla birleştirilmesinden oluşur.

Modeli eğitmek için, yazılım 10–6 weight decay [12] ile L2 normalleştirmesi kullandı. Xavier başlatma [15] ve Adam optimizasyon [16] algoritmaları uygulanmıştır. Çeşitlilik eksikliği sorunundan kaçınmak için toplu renormalizasyon [17] kullanıldı. Yazılım, resimleri daha da küçülterek belleğe daha iyi oturması ve toplu eğitim için destek sağlaması için 300x300 yamalara bölmüştür. Kenar boşluklarının eklenmesi ile sınır efektleri önlenir. Ağda önceden ölçülen ağırlıkların kullanılması nedeniyle, eğitim süresi önemli ölçüde azalmıştır [12]. Eğitim süreci, döndürme, ölçekleme ve yansıtma gibi çeşitli anında veri büyütme tekniklerinden yararlanır. Sistem, her pikselin eğitilmiş nesne türlerinden birine ait olma olasılıklarını vermektedir.

C. Veri Kümesini Değerlendirmeye Hazırlama

Sistemimizin performansını değerlendirmek için beş farklı model geliştirdik. İlk iki model bir bölgenin kaydı ile eğitildi ve tamamen farklı bir bölgenin kaydı ile test edildi. Son model için iki bölge kayıtları birleştirildi ve birleşik bir model üzerinde eğitim yapıldı. Bu model 10 kat çapraz doğrulama ile test edildi. Son iki model de aynı bölgenin kaydı ile eğitilip, aynı bölgenin kaydı ile 10 kat çapraz doğrulama ile test edildi.

D. Son İşleme

Bizim sorumuzda dört farklı sınıfımız bulunmaktadır: arka plan, sayfa, birey ve köy başlangıcı. Bu nedenle, sınıflardan birine ait piksellerin olasılıklarını değerlendirdik. Her sınıf için, bir pikselin kendilerine ait olma olasılığını gösteren ikili bir matris vardır. Bu matrisleri kullanarak pikseller bağlanmalı ve bileşenler oluşturulmalıdır. Bağlı bileşen analizi aracı [12] nesne oluşturmak için kullanılmıştır. Nesneler tüm sınıflar için oluşturulduktan sonra, sistemimizin performansı ölçülmüştür.

V. DENEYSEL SONUÇLAR VE TARTIŞMA

Bu bölümde önce sistemimizi değerlendirmek için kullanılan metrikler tanımlanmıştır. Sonra sonuçlar sunulup tartışılmıştır.

A. Metrikler

Sistem performansımızı değerlendirmek için dört farklı metrik kullandık. İlk iki metrik, düşük düzey değerlendiricilerdir ve nesne algılama sorunlarında yaygın olarak kullanılmaktadır. Sistemimizin doğruluğunu değerlendirmek için üçüncü ve dördüncü üst düzey metrikler tanımlanmıştır. Bu metrikler, nüfus kayıtlarındaki kişileri saymaya yönelik uygulamamıza özeldir. İlk metrik, piksel doğruluk oranıdır. İkinci, doğru sınıflandırılmış pikselleri tüm piksel sayısına bölerek hesaplanabilir. Üçüncüsü, Kesişim - Birleşim oranı (IoU) metriğidir. Bu metrik için temel doğruluk bileşenleri ve modelimizden tahmin edilen bileşenler vardır. Bu metrik, bu iki bileşenin bölgelelerinin kesişiminin bu iki bileşenin bölgelerinin birleşmesine bölünmesiyle hesaplanabilir. Bireyleri saymak için, yüksek düzeyli metrik, öngörülen obje sayısı bölü gerçek obje sayısı olarak tanımlanabilir. Bu metrik kişi sayımında bireysel sayma doğruluğu (ICA) olarak adlandırılabilir. Ayrıca, köy başlangıç objeleri için benzer bir yüksek seviye metriği tanımlanıp köy başlangıç sayımı doğruluğu (VSCA) olarak adlandırıldı.

B. Sonuçlar ve Tartışma

İznik bölgesinden iki, Zıştovi bölgesinden bir kayıt defterimiz bulunmaktadır. Model 1'de İznik kayıtları ile eğitim yapıldı ve Zıştovi kayıtları ile test edildi. Model 2'de Zıştovi bölge kayıtlarına sahip bir model eğittik ve İznik nüfus kayıtları ile test ettik. Aynı bölgedeki kayıtlarda 10 kat çapraz doğrulama kullanarak validasyon yapılmıştır. Model 3'te, Zıştovi kayıtları üzerinde eğitim ve test gerçekleştirilmiş, Model 4'te ise İznik kayıtları üzerinde eğitim ve test yapılmıştır. Model 5'te, tüm veri kümesini birleştirdik ve modeli 10 kat çapraz doğrulama ile değerlendirdik. Piksel doğruluğu, IoU ve sayım hassasiyeti sonuçları Tablo 1'de verilmiştir. İlk iki metrik 4 sınıflı bir sınıflandırma için sonuçlar vermiştir. Son iki metrik ise bireylerin sayısını ve köy başlangıç nesnesini bulma doğruluğudur. En iyi ICA sonuçları eğitim için Zıştovi kayıtları kullanıldığında elde edilmiştir. En kötü doğruluk oranları ise, sistem İznik kayıtları ile eğitildiğinde ve Zıştovi kaydı ile test edildiğinde elde edilmiştir. Ayrıca, köy başlangıç sayma doğruluğu tüm modeller için % 100'dür.

Daha önce de belirtildiği gibi, kayıtların yerleşimi ilçelere ve memura bağlı olarak değişebilmektedir. Kayıtlarımız için, İznik'teki bireylere geniş yer ayrılırken, kayıtlar arasındaki mesafe Zıştovi kayıtlarında daha azdır. Bir İznik kayıt sayfasındaki kayıtlı kişi sayısı yaklaşık 40, bir Zıştovi kaydında ise yaklaşık 80'dir. Bu nedenle, sistem gevşekçe yerleştirilmiş İznik kayıtları ile eğitildiğinde ve Zıştovi'deki yakın yazılı kümelerde test edildiğinde, sayım doğruluğu azalır ve birden fazla kaydı saymaya yönelik bir hata oluşmaya başlar. Oysa, eğitim ve test parçalarını değiştirirsek, kümeleri saymak için sistem doğruluğu beklediğimiz gibi % 100'e yaklaşır. Veri kümesini karıştırıp 10 kat çapraz doğrulama uygularsak, bu iki sonuç arasında sayım doğrulukları elde ettik. Bizim amacımız için, üst düzey metrikler daha önemli olmasına rağmen, düşük

TABLO I: Farklı ölçütlere ait sonuçlar beş farklı model için sunulmuştur.

Eğitim	Test	Piksel Doğ. Oranı(%)	IoU(%)	ICA(%)	VSCA(%)
İznik	İznik	94.25	80.91	98.35	100
İznik	Ziştovi	92.72	73.6	88.43	100
Ziştovi	Ziştovi	93.04	72.29	99.24	100
Ziştovi	İznik	85.36	47.95	99.73	100
Karışık	Karışık	87.26	48.54	97.74	100

düzeyleyici metrikler sistemimizin genel performansını göstermekte ve farklı modellerle performans karşılaştırması için de faydalıdır.

Şekil 3: Sistemimiz tarafından yapılan örnek bir tahmin şekilde gösterilmiştir. Solda, bireyleri saymak için ikili bir tahmin görüntüsü, ortada köy başlangıcını saymak için ikili bir görüntü ve sağda bulunan nesneler dikdörtgen kutularla çevrilidir. Bireysel kayıt sayımı için yeşil kutular ve köy başlangıç nesnesini saymak için kırmızı kutular kullanılmıştır.

VI. SONUÇ VE GELECEK ÇALIŞMALAR

Bu çalışmada, Osmanlı İmparatorluğu'nun ilk nüfus sayımında 1840-1860 yılları arasındaki kayıtlar için otomatik bireysel sayma sistemi geliştirdik. Kayıtlar Arapça harflerle yazılmıştır ve yerleşimleri büyük ölçüde bölgeye ve yetkili memura bağlıdır. Üç kayıt için etiketlenmiş bir veri kümesi oluşturduk ve sistemimizi bu veri kümesinde değerlendirdik. Köy başlangıç sembollerini sayan sistemimiz % 100 doğruluğa ulaştı. Ayrıca, bireysel sayım doğruluk oranında maksimum % 99,73 sınıflandırma oranı elde ettik. Bu sonuçlardan modellerin yakın yerleştirilmiş ve gürültülü kayıtlarla eğitilmesi gerektiğini çıkardık (Vaka çalışmamızda Ziştovi kaydı). Bu modeller temiz ve gevşek bir şekilde yerleştirilmiş bir modelle test edildiğinde (bu vaka çalışmasında İznik kaydı), sistem bireyleri doğru olarak saymaktadır. Bununla birlikte, bir model gevşek yerleştirilmiş kayıtlarla eğitilmiş ve yakın yerleştirilmiş bir modelle test edilmişse, sayım hatalarının sayısı artmaktadır. Amacımız, 1840'lar ve 1860'lar arasında hazırlanan tüm kayıtların verimli sayımı ve uzaktan okunması için uygulanabilecek bir sistem geliştirmektir. Tüm kayıtları etiketlemek imkansız olduğundan, böyle bir sistemi geliştirmek için stratejik olarak yakın yerleştirilmiş ve gürültülü olanları etiketleyeceğiz. Gelecekteki çalışmalar olarak, bölümlere ayrılmış bireysel kayıt nesneleri için otomatik bir el yazısı tanıma sistemi geliştirmeyi planlamaktayız.

TEŞEKKÜR

This work has been supported by European Research Council (ERC) Project: "Industrialisation and Urban Growth

from the mid-nineteenth century Ottoman Empire to Contemporary Turkey in a Comparative Perspective, 1850-2000" under the European Union's Horizon 2020 research and innovation programme grant agreement No. 679097.

KAYNAKLAR

- [1] C. Wick and F. Puppe, "Fully convolutional neural networks for page segmentation of historical document images," in *2018 13th IAPR International Workshop on Document Analysis Systems (DAS)*. IEEE, 2018, pp. 287–292.
- [2] Y. Xu, W. He, F. Yin, and C.-L. Liu, "Page segmentation for historical handwritten documents using fully convolutional networks," in *2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR)*, vol. 1. IEEE, 2017, pp. 541–546.
- [3] A. M. Hesham, M. A. Rashwan, H. M. Al-Barhamtoshy, S. M. Abdou, A. A. Badr, and I. Farag, "Arabic document layout analysis," *Pattern Anal. Appl.*, vol. 20, no. 4, pp. 1275–1287, Nov. 2017.
- [4] K. Laven, S. Leishman, and S. Roweis, "A statistical learning approach to document image analysis," in *Eighth International Conference on Document Analysis and Recognition (ICDAR'05)*. IEEE, 2005, pp. 357–361.
- [5] Jaekyu Ha, R. M. Haralick, and I. T. Phillips, "Recursive x-y cut using bounding boxes of connected components," in *Proceedings of 3rd International Conference on Document Analysis and Recognition*, vol. 2, Aug 1995, pp. 952–955 vol.2.
- [6] J. Sauvola, T. Seppanen, S. Haapakoski, and M. Pietikainen, "Adaptive document binarization," in *Proceedings of the Fourth International Conference on Document Analysis and Recognition*, vol. 1. IEEE, 1997, pp. 147–152.
- [7] S. S. Bukhari, T. M. Breuel, A. Asi, and J. El-Sana, "Layout analysis for arabic historical document images using machine learning," in *2012 International Conference on Frontiers in Handwriting Recognition*. IEEE, 2012, pp. 639–644.
- [8] T. Matsumoto, T. Yokohama, H. Suzuki, R. Furukawa, A. Oshimoto, T. Shimmi, Y. Matsushita, T. Seo, and L. O. Chua, "Several image processing examples by cnn," in *IEEE International Workshop on Cellular Neural Networks and their Applications*, Dec 1990, pp. 100–111.
- [9] T. M. Breuel, "Robust, simple page segmentation using hybrid convolutional mlstm networks," in *2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR)*, vol. 01, Nov 2017, pp. 733–740.
- [10] D. Augusto Borges Oliveira and M. Palhares Viana, "Fast cnn-based document layout analysis," in *Proceedings of the IEEE International Conference on Computer Vision*, 2017, pp. 1173–1180.
- [11] I. M. Amer, S. Hamdy, and M. G. M. Mostafa, "Deep arabic document layout analysis," in *2017 Eighth International Conference on Intelligent Computing and Information Systems (ICICIS)*, Dec 2017, pp. 224–231.
- [12] S. Ares Oliveira, B. Seguin, and F. Kaplan, "dhsegment: A generic deep-learning approach for document segmentation," in *2018 16th International Conference on Frontiers in Handwriting Recognition (ICFHR)*, Aug 2018, pp. 7–12.
- [13] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.
- [14] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015, pp. 234–241.
- [15] X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," in *Proceedings of the thirteenth international conference on artificial intelligence and statistics*, 2010, pp. 249–256.
- [16] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.
- [17] S. Ioffe, "Batch renormalization: Towards reducing minibatch dependence in batch-normalized models," in *Advances in neural information processing systems*, 2017, pp. 1945–1953.