



DeepGraphNet: Grafiklerin Sınıflandırılmasında Derin Öğrenme Modelleri

Gökhan Altan^{1*}

¹ İskenderun Teknik Üniversitesi, Mühendislik ve Doğa Bilimleri Fakültesi, Bilgisayar Mühendisliği Bölümü, Hatay, Türkiye (ORCID: 0000-0001-7883-3131)

(This publication has been presented orally at HORA congress.)

(İlk Geliş Tarihi 1 Ağustos 2019 ve Kabul Tarihi 25 Ekim 2019)

(DOI: 10.31590/ejosat.638256)

ATIF/REFERENCE: Altan, G. (2019). DeepGraphNet: Grafiklerin Sınıflandırılmasında Derin Öğrenme Modelleri. *Avrupa Bilim ve Teknoloji Dergisi*, (Özel Sayı), 319-329.

Öz

Grafik sınıflandırma modeli henüz yeni bir araştırma alanı olarak ön plana çıkan bir görüntü işleme yaklaşımıdır. Özellikle verilerin görselleştirilmesi ve kolay okunabilirliğini sağlamak için tercih edilen grafikler, programlama dillerinin esnekliğini ve görselliğini her geçen gün artırmasına bağlı olarak hızla gelişen üçüncü parti ofis yazılımları, farklı görsel türlerde grafikler oluşturabilmektedir. Bu çalışmanın amacı, farklı giriş olarak verilen grafiğin hangi tür bir grafik olduğunu belirlemeyen bir derin öğrenme modeli oluşturmaktır. Analiz edilen görüntülerine düşük ve yüksek seviye öznitelik çıkarma algoritmaları uygulamak yerine, doğrudan derin öğrenme modellerine giriş olarak verilmiştir. Derin öğrenme algoritmalarının öznitelik öğrenme, paylaşılmış sınıflandırma ağırlıklarının transferi ve kendi içerisindeki ileri seviyeli görüntü işleme kabiliyetlerini kullanarak modellerin grafik sınıflandırmada ki etkinlikleri kıyaslanmıştır. Çalışmada konvolüsyonel sinir ağları ve derin inanç ağları modellerinin genel başarımları, hassasiyet, özgünlük, pozitif öngörü değeri ve negatif öngörü değeri gibi sınıflandırıcı performansları hesaplanmıştır. Analizlerde kullanılan veriler, çizgi grafiği, sütun grafiği, pasta grafiği ve dağılım grafiğini eşit sayıda içerecek şekilde toplam 1200 resimden meydana gelmektedir. Herbir resim dosyası 224x224 boyutta yeniden boyutlandırılarak, gri seviye resme dönüştürülmüştür. Analizlerin sınıflama süreçlerinde 5-katlı çapraz doğrulama yöntemi kullanılarak herbir verinin birbirinden bağımsız olarak test ve eğitim süreçlerine dâhil edilmesi sağlanmıştır. Deneysel çalışmalardan elde edilen sonuçlar göstermiştir ki önerilen konvolüsyonel sinir ağları düşük, orta ve yüksek seviyeli öznitelik çıkarma kapasitesiyle %92,92 genel başarımla dört farklı grafiği sınıflandırırken, derin inanç ağları %90,04 genel başarıma kadar ulaşabilmiştir. Görüntülerdeki verilerin istatistiksel ve enerji durumuna bağlı olarak yeniden oluşturulan yansımalar her ne kadar bu verilerin belirli filtrelerden konvolüsyon işlemi sonrası elde edilen yansımalarından daha düşük sınıflandırma başarımları elde etmiş olsa da denenen farklı katman sayısına sahip modeller için yüksek başarımlar elde edilmiştir. Tensörler haline getirilen katmanların en baskın özelliklerinin belirlenerek belirleyici piksel değerlerinin havuzlama ile bir sonraki katmana aktarıldığı konvolüsyonel sinir ağları modelleri, görüntü işleme yaklaşımları için esneklik ve etkin bir kullanım sağlamaktadır.

Anahtar Kelimeler: Derin Öğrenme, grafik sınıflandırma, Derin inanç ağları, Konvolüsyonel sinir ağı.

DeepGraphNet: Deep Learning Models in the Classification of Graphs

Abstract

The graph classification model is an image processing approach that has come into prominence as a new research area. Particularly preferred graphs to provide visualization and easy readability of data, third party office software which develops rapidly due to

* Sorumlu Yazar: İskenderun Teknik Üniversitesi, Mühendislik ve Doğa Bilimleri Fakültesi, Bilgisayar mühendisliği Bölümü, Hatay, Türkiye, ORCID: 0000-0001-7883-3131, gokhan.altan@iste.edu.tr

increasing the flexibility and visuality of programming languages, can create graphics of different visual types. The aim of this study is to propose a deep learning model that classifies the type of graph is given as an input. It is fed into deep learning models as the features, instead of applying low and high-level feature extraction algorithms to the analyzed images. The effectiveness of the models in graphical classification was compared using feature learning, transfer of shared classification weights and advanced image processing capabilities of deep learning algorithms. In this study, the performance of convolutional neural networks and deep belief networks models such as overall performance, sensitivity, specificity, positive predictive value, and negative predictive value were calculated. The dataset in the analyzes consists of a total number of 1200 images including an equal number of line graphs, column graphs, pie graphs and distribution graphs. Each graph was resized to 224x224 pixels and was converted to gray level image. In the classification process of the analysis, 5-fold cross validation algorithm was used to evaluate each image independently into test and training processes. The experimental results showed that the proposed convolutional neural networks model classifies four different graphs with an overall accuracy rate of 92.92% with low-, medium- and high-level feature extraction capacity, while deep belief networks reached to an overall accuracy rate of 90.04%. Although the generated new presentations of the input images depending on the statistical and energy status of the visible and hidden units have achieved lower classification performance than the representations of input data from definitive filters using convolution, high graph classification performances have been achieved using the proposed models for models with various hidden layers. Convolutional neural networks model, in which the predominant properties of the layers formed by tensors using the dominant pixels are transferred to the next layer by pooling, provides flexibility and effective use for image processing approaches.

Keywords: Deep Learning, graph classification, Deep belief networks, Convolutional Neural Networks.

1. Giriş

Büyük veri gibi kavramların hayatımıza girdiği günümüzde, uzun süreli verilerin toplanması analiz süreci için oldukça büyük önem taşımakla birlikte, bu verilerin görselleştirilmesi ve sınıflandırılması da verilerin takibi, analizi ve sonuçlandırılması için bilgisayar destekli modeller kullanımını mecburi hale getirmiştir. Belirli aralıkta kaydedilen verinin görselleştirilerek daha kolay yorumlanmasını sağlayan grafikler son yıllarda üçüncü parti yazılımların gelişmesiyle birlikte oldukça farklı ve genişletilmiş tasarımlara dönüşmüştür. Büyük veri analizi ve veri madenciliği yaklaşımları için önemli ve ayırt edici özniteliklerin çıkarılması, bu grafiklerin yorumlanması ve kolay anlaşılır duruma getirilmesi açısından daha anlamlıdır. Farklı üçüncü parti yazılımlarından kaynaklı değişik görselliklerdeki grafiklerin sınıflandırılması bu anlamda büyük önem taşımaktadır. Aynı verilerden elde edilen farklı grafik türleri bile görsellik ve anlam olarak farklı yorumlamalara sebep olabilmektedir. Bu durum grafiklerin basit öznitelik ve dağılım özellikleriyle doğru olarak sınıflandırılmasını neredeyse imkânsız kılmaktadır. Bu açıdan değerlendirildiğinde basit öznitelikler dışında düşük, orta ve yüksek seviyeli özniteliklerin bütünlük halinde kullanımına olanak sağlayan süreçler tasarlanması gerekmektedir.

Grafik ve resim sınıflandırma çalışmalarında özellikle elle çıkarılmış öznitelikler anlamlı sonuçlar elde edilmesini sağlarken, veri miktarının artması, resim işleme süreçlerinde sınıflandırma aşamasına gelinmeden henüz öznitelik çıkarma aşamasındayken oldukça zaman alan süreçlerle sonuçlanmaktadır. Bu durum göz önünde bulundurulduğunda bazı süreçlerin doğru modellenerek öznitelik çıkarma ve çıkarılan anlamlı özniteliklerin bir sonraki sürece transferinin otomatikleştirilmesi gerekliliği ortaya çıkmaktadır. Bazı çalışmalarda grafik tanıma, yeniden oluşturma, grafik türleri arasındaki dönüşüm için gerçekleştirilen analizlerde sistemler kullanıcı tarafında elle gerçekleştirilmesi gereken metinsel değerlerin girilmesi, grafiğin düzlemlerinin belirlenmesi, koordinatlar aralığının belirlenmesi ve grafiklerin odak bölgelerinin işaretlenmesi gibi ön işlemlere ihtiyaç duymaktadır (Jung ve ark., 2017; Savva ve ark., 2011; Shadish ve ark., 2009; Yang, Huang ve Tan, 2006). Poco ve Heer ise bu yaklaşımlardan biraz daha farklı olarak metinsel içeriğin görüntü işleme algoritmalarıyla tespiti ve elde edilen verilerin dışa aktarımı sağlanırken (Poco ve Heer, 2017), elde edilen bu verilerin yorumlanması ve kendi arasında farklı grafik modelleri arasında doğrudan dönüşüm gerçekleştirilmiştir (Dai, Wang, Niu ve Zhang, 2018).

Bilgisayar destekli görüntü analizleri özellikle görüntü işleme yaklaşımlarının gelişmesiyle paralel olarak üst seviyelere taşınmıştır. Düşük ve yüksek seviyeli öznitelik çıkarmayı sağlayan görüntü işleme algoritmaları resimdeki baskın ve anlamlı özellikleri belirleyip, belirlenen bu özelliklerin dağılımsal ve morfolojik özellikleri çeşitli makine öğrenmesi algoritmalarına bütünlük hâle getirilerek kullanılmıştır. Grafiklerden manuel olarak belirlenen basit özellikler üzerinde görüntü işleme yaklaşımlarının temel işlemleriyle değerler temsil edilmeye çalışılmıştır. Zhou vd. sütun grafiklerinin tanınması için Hough dönüşümünden elde ettiği öznitelikleri kullanmıştır. Sütun grafiğinin yanında farklı türden grafikleri de tanıma özelliğine sahip bir model önermişlerdir (Zhou ve Tan, 2000). Zhou vd. daha sonrasında doküman içeriğindeki grafiklerin tanınması için gradyan temelli öznitelikler kullanarak saklı Markov modeli ile grafiklerin sınıflandırılmasını gerçekleştirmiştir. Önerdikleri model her ne kadar gradyan temelli özellikleri kullansa da çok basit grafik türlerini tanıyabilmektedir (Zhou ve Tan, 2001). Bunun sebebi graydan yaklaşımın gürültüden fazla etkilenmesi olarak gösterilebilir. Butler vd. bir başka grafik türü olan akış diyagramlarını tanıyan bir model önermişlerdir. İlk aşamada grafik üzerindeki çizgi, ok, kutular, daire ve eliptik şekillerin tespitini yaparak sonrasında bu şekiller içerisindeki metinlerin çözümlemesini gerçekleştirmişlerdir. Sonraki aşamada elde edilen bu metin ve komut bildirimleri olan şekiller arasındaki ilişkileri belirleyerek gerçekleştirme yapmışlardır (Butler, Grogono, Shinghal ve Tjandra, 1995). Yu vd. grafik tanıma yaklaşımının uygulamasını yaparak adaptif bir hiper grafik öğrenme modeli önermişlerdir. Elde ettikleri bir görüntüde en yakın komşuların belirlenmesini gerçekleştirerek bu komşuluğun boyutlarına göre sütunlar arasında ilişki kurmuşlardır (Yu, Tao ve Wang, 2012). Poco ve Heer (Poco ve Heer, 2017) bu yaklaşımlardan biraz daha farklı olarak metinsel içeriğin görüntü işleme algoritmalarıyla tespiti ve verilerin dışa aktarımını sağlayan bir model gerçekleştirmişlerdir. Dai vd. (Dai ve ark., 2018) yapıyı biraz daha ilerleterek farklı grafik modelleri için tanıma ve içeriğin çözümlenmesi aşamalarını bütünlük hâle getirmiş ve yeni bir yaklaşım gerçekleştirmişlerdir. Tang vd. derin inanç ağları (Deep Belief Networks - DBN) ve konvolüsyonel sinir ağları (Convolutional Neural Networks – CNN) modellerini birleştirerek beş farklı türdeki

grafiklerin sınıflandırılmasını gerçekleştirmişlerdir. Bu yaklaşımda konvolüsyonel sinir ağları ile elde ettikleri farklı seviyedeki öznetelikleri derin inanç ağlarına giriş olarak vermiş ve basit öznetelikler olan elle elde edilmiş özneteliklere göre çok daha yüksek sınıflandırıcı performansı elde etmişlerdir (Tang ve ark., 2016).

Derin öğrenme, son yıllarda görüntü işleme olanaklarını da öznetelik öğrenme modeli olarak sınıflandırma sürecine eklemiş, daha detaylı analizlerle anlamlı sonuçlara yaklaşmayı hızlandırmış bir makine öğrenmesi alanıdır. Özellikle görüntü işleme, video işleme, görüntü oluşturma, ses analizleri, biyomedikal sinyal sınıflandırma ve doğal dil işleme gibi disiplinlerde oldukça popüler olan derin öğrenme algoritmalarının en önemli avantajı öznetelik çıkarma yöntemlerine ihtiyaç duymaksızın ön işlemden geçirilen ham veriyi yüksek performanslı olarak sınıflandırma, segmentasyon ve regresyon süreçlerine uyarlayabilmesidir. DBN, CNN, derin sinir ağları (Deep neural networks), çekişmeli üretici ağlar (Generative adversarial networks), uzun kısa süreli hafıza ağları (Long short term memory) ve derin aşırı öğrenme makineleri (Deep extreme learning machines) yüksek genelleme kapasiteleri sebebiyle derin öğrenme algoritmaları içerisinde en yaygın kullanıma sahip sınıflandırıcı modelleridir. Görüntü işleme üzerine en sık kullanılan modeller CNN yaklaşımıdır. Cireşan vd. grafik işlem biriminde eğitimini gerçekleştirdiği CNN modeli kullanarak nesne tespiti ve karakter tanıma üzerindeki etkileri üzerine çalışmalar yapmıştır. Farklı veritabanları üzerinde gerçekleştirdikleri nesne tanıma yaklaşımında %2,53 hataya kadar sınıflandırma performansına ulaşırken, yine CNN altyapısını kullanarak MNIST el yazısı veritabanıyla karakter tanıma üzerine %0,35 sınıflandırma hatası raporlamışlardır. Kullanılan CNN modelin esnekliği ve görüntü işleme yaklaşımlarına uyarlanabilirliği doğrultusunda olumlu görüşlerini paylaşmışlardır (Cireşan, Meier, Masci, Gambardella ve Schmidhuber, 2011). Cireşan vd. yine MNIST el yazısı veritabanı üzerinde karakter tanıma işlemini, CNN sınıflandırıcısı kullanarak farklı modeller üzerinde performansını artırmayı hedeflemiştir. Otomatik karakter tanıma başarımı olarak %0,27 hata raporlamış ve karakter tanıma üzerine en başarılı ve farklı durumlara karşı dayanıklı model olarak önermiştir (Cireşan, Meier, Gambardella ve Schmidhuber, 2011). Krizhevsky vd. geniş ve detaylı bir CNN modeli kullanarak düşük, orta ve yüksek seviyeli özneteliklerin belirlenmesine yönelik ImageNet isimli bir model üretmişlerdir. Bu model 1,2 milyon resim üzerinde resim analizi yaparak bin farklı sınıfta görüntünün tanınmasını sağlamaktadır. Önerilen modelin detaylı olması eğitim sırasında yaklaşık altmış milyon parametrenin optimizasyonunu gerektiren zaman alıcı bir süreç olsa da grafik işlem biriminin hesaplama kabiliyetini kullanarak hızlandırılan eğitim oldukça başarılı bir model geliştirmeyi Dropout algoritmasının esnekliğiyle yerine getirmiştir (Krizhevsky, Sutskever ve Hinton, 2012). Kalchbrenner vd. cümlelerin oluşturulması ve gerekli gramer yapılarının kurulması için bir CNN modeli önermişlerdir. Önerdikleri modelde diğer yapılardan farklı olarak dinamik havuzlama yapısı kullanarak genel bir yaklaşım elde etmişlerdir. Çok sınıflı fikir tahmini ve sosyal medyaya yönelik metinlerden duygu tanımlama gerçekleştirmişlerdir. Yaklaşımı farklı sınıflandırıcılar üzerinde modellemişler fakat en başarılı sınıflandırma performansını %86,8 başarıyla CNN sınıflandırıcısı ile elde etmişlerdir (Kalchbrenner, Grefenstette ve Blunsom, 2014). Kim, benzer şekilde CNN sınıflandırıcı kullanarak cümleden fikir tahmin çalışması yapmış ve soru kümesi sınıflandırma yaklaşımında başarılı sonuçlar elde etmiştir. Farklı veritabanları üzerinde yaptıkları deneylerle statik modelli, ön tanımlı ve çok kanallı kelime vektörlerinin farklı varyasyonlarla etkili ve dayanıklı bir şekilde sınıflandırılmasını sağlamıştır (Kim, 2014). Karpathy vd. CNN modelinin resimler üzerindeki başarılı sonuçlarına dayanarak yaklaşık 1 milyon video üzerinde 487 farklı sınıf içeren yüksek çözünürlüklü videoların sınıflandırılmasını gerçekleştirmişlerdir. Literatürdeki video sınıflandırıcıların başarımlarını %19,4 artırarak %63,3 seviyelerine getirmeyi başarmışlardır. Önerdikleri modellerde videoların hangi spor dalına ait olduğu, videodaki şahısların hangi aktivitede bulunduğunu tanıyabilme yetisine sahip kaynaştırma ve aşamalı kaynaştırmalı modeller önermişlerdir (Karpathy ve ark., 2014). Siddiqui vd. sualtında çekilmiş balık görüntülerinden balığın sınıflandırılmasını gerçekleştiren ve otomatik balık izleme modülü olarak kullanılan bir CNN modeli gerçekleştirmişlerdir. Sualtı ortamında gizleme özelliği olan balıkların tespitini gerçekleştirebilen sistem %94,3 genel başarımla belirli sayıda balık veritabanı üzerinde başarımla sağlamıştır (Siddiqui ve ark., 2018). DBN sınıflandırıcısı da benzer şekilde görüntü işleme aşamalarında hızlı eğitim sağlayan bir model olmasından kaynaklı sıklıkla tercih edilmiş, sürekli olarak denetimsiz modelinde ve geri yayılım sürecinde yeniliklerle ön plana çıkan bir algoritma olarak ön plana çıkmaktadır. Hinton vd. DBN modeli için hızlı ve dayanıklı öğrenme algoritmaları olan tamamlayıcı öncelik modellerinden ince ayar (Fine-tuning) algoritmasını önermişlerdir. Bu yaklaşımla katman sayısı ve her katmandaki nöron sayısı fazla modeller için geri yayılım algoritmalarının oldukça hızlanmasına katkı sağlamışlardır. MNIST karakter veri seti üzerinde yaptıkları detaylı analizler sonucunda sınıflandırıcı performansı olarak hata oranını çok düşük seviyelere taşımışlardır (Hinton, Osindero ve Teh, 2006). Mohamed vd. CNN modeline benzer şekilde DBN sınıflandırıcısını akustik konuşma tanıma işlemi için kullanmış Gauss karışımı modelleriyle telefon seslerinin modellemesini gerçekleştirmişlerdir (Mohamed, Dahl ve Hinton, 2012). Huang vd. görüntüden elle öznetelik çıkarmak yerine görüntülere uygulanmış konvolüsyon işleminin çıktılarını DBN modeline giriş olarak vererek evrimsel bir DBN modeli önermişlerdir. Analizlerini yüz tanıma uygulaması üzerine yaptıklarında farklı sınıflandırıcı ve modellere göre önerdikler DBN modelinin oldukça başarılı olduğunu raporlamışlardır (Huang, Lee ve Learned-Miller, 2012). Abdel-Zaher ve Eldeib, Wisconsin göğüs kanseri verisi üzerinde gerçekleştirdikleri göğüs röntgen filmlerinden elle çıkarılmış öznetelikleri kullanarak %99,68 genel başarımla %100 hassasiyet ve %99,47 seçicilikle sağlıklı göğüs kanseri verilerini, kanser verilerinden ayırmayı başarmışlardır. Burada DBN modelin dayanıklılığı ve performansını detaylı modeller için çok katmanlı ağlarda avantajlarını vurgulamıştır (Abdel-Zaher ve Eldeib, 2016). Görüntü analizlerinin yanı sıra zaman serilerinin analizinde de DBN model, sıklıkla kullanılmıştır. Altan vd. elektroensefalografi sinyallerine uygulanmış Hilbert-Huang dönüşümü ile elde ettikleri frekans modülasyon özneteliklerini DBN modeline giriş olarak vermişler ve felçli hastalar için olumlu ve olumsuz hareket belirleme işlemini gerçekleştirmişlerdir. Özellikle zaman serilerinin doğrudan giriş olarak değil de öznetelik çıkarma algoritmalarıyla işlenerek DBN modeline verilmesinin daha başarılı sınıflandırıcı modeller geliştirilmesine katkı sunmuşlardır (Altan, Kutlu ve Allahverdi, 2016b). Altan vd. yine farklı yapıdaki bir zaman serisi olan elektrokardiyogram sinyallerine benzer bir yaklaşım uygulayarak öncelikle sağlıklı ve artımili sinyallerden ikinci derece fark haritaları çıkarmış, bu haritaların nicelemelerini yaparak elde ettiği dağılım özneteliklerini kullanarak sınıflandırılmasını gerçekleştirmiştir. İkinci aşama olarak da hasta olarak etiketlenmiş verinin hangi türden aritmi olduğunun tespitini %96,10 genel başarımla raporlamışlardır (Altan, Kutlu ve Allahverdi, 2016a). Literatürdeki bu çalışmalar hem CNN hem de DBN sınıflandırıcılarının görüntü ve diğer süreçler için iyi birer makine öğrenmesi algoritması olduğunu gösteren parametrelerdir. Bu algoritmalarının gösterdikleri yüksek performanslar, derin öğrenme yaklaşımlarının popüler olmasındaki temel etkenler olma özelliği taşımaktadır.

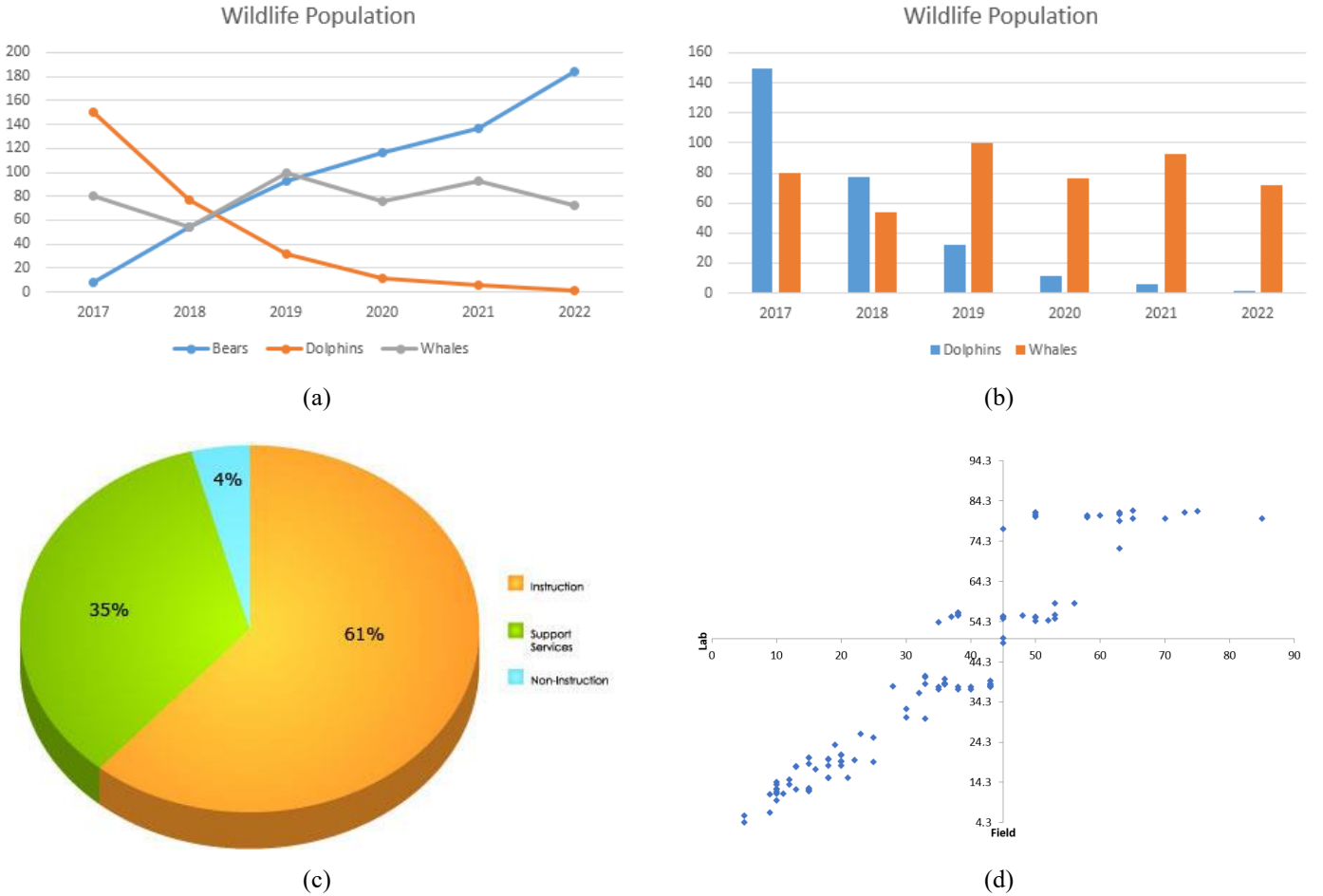
Bu çalışmada, DBN ve CNN modellerinin grafik sınıflandırmadaki başarımları analiz edilmiştir. Literatürdeki bazı ön tanımlı modellere nazaran daha basit modeller geliştirilerek sınıflandırıcı parametre sayıları en aza indirilmiş, bununla bağlantılı olarak eğitim süreci minimize edilmiştir. Optimizasyon sonrası elde edilen modeller CNN ve DBN sınıflandırıcıları için paylaşılacak, düşük, orta ve yüksek seviyeli öznelilik çıkarma modelleri ile farklı yansımalar oluşturularak istatistiksel bağlam modellemeleri sunulmuştur. Analizler, yine tasarlanan bir script tarafından elde edilmiş dört farklı grafik üzerinde gerçekleştirilmiştir. Modellerin etkinliği üzerinde yorumlar ve çıkarımlar yapılmıştır.

Çalışmanın geri kalan kısmında kullanılan veritabanı ve veri toplama süreci hakkında bilgi, elde edilen görüntüler üzerinde ön işlemler, DBN ve CNN hakkında detaylı bilgiler ve sınıflandırıcı parametrelerinin tanıtımı, sınıflandırma süreçlerine nasıl dâhil edildikleri ikinci bölümde; farklı DBN ve CNN modelleri ile gerçekleştirilen sınıflandırıcılar, deneysel kurulum ve parametre seçimleri, bu modeller için elde edilen deneysel sonuçlar üçüncü bölümde; modeller hakkında tartışma ve literatürdeki ön tanımlı CNN modellerine göre avantajları dördüncü bölümde anlatılmaktadır.

2. Materyal ve Metot

2.1. Veritabanı

Grafikler, özellikle sunum ve uzun süreli veri kümelerinin görselleştirilmesi için sıklıkla tercih edilen araçlardır. Belirli bir orandaki dağılımı nitelemek ve nicellemek için kullanılan bu görseller, aslında çok uzun süreçlerin yorumlanması için gerekli kontrol sürecini anlamsal bütünlere dönüştürebilirler. Grafiklerin veriyi anlamlandırmak üzerindeki bu önemli katkısı, büyük veri analizlerinde basit ama etkili yaklaşımlar oluşturulmasında temel teşkil etmektedir. Özellikle son yıllarda üç boyutlu grafikler, farklı dağılım mekanizmaları ile görselliği ön plana çıkaran üçüncü parti yazılımlarının çeşitlenmesi sebebiyle veriler grafikler tarafından çok farklı türlerde temsil edilmeye başlanmıştır. Bu yüzden verisetinin son günlerdeki yayınlanmış grafik türlerinden oluşmasına özen gösterilmiştir.



Şekil 1. Verisetinden rastgele seçilmiş çizgi grafiği (a), sütun grafiği (b), pasta grafiği (c) ve dağılım grafiği (d) görüntüleri

Veri toplamak için, Google resim arama motorunda son 1 yılda indekslenmiş, belirtilen anahtar kelimeye bağlı olarak resimleri tüm resim formatlarında indirebilen, farklı anahtar kelimeleri kendi içinde kategorize eden python temelli bir kütüphanesi oluşturulmuştur. Kütüphanenin parametreleri, kategorilenecek anahtar kelimedenden oluşmaktadır. Her bir anahtar kelimeyi arama motorunda post ederek, arama motorunun bu işleme karşılık bulduğu görselleri anahtar kelime ile aynı isimdeki bir klasöre kayderek kategorilemektedir. Burada dikkat edilmesi gereken durum anahtar kelimeye bağlı farklı içeriklerin de verisetine eklenmesi olasılığıdır. Bu durumun verisetinde

gürültü oluşmasına, grafikler üzerinde belirgin öznitelikler belirleyerek doğru bir sınıflandırma modelinin oluşturulamamasına, bu doğrultuda da CNN ve DBN modellerinin anlamsız sonuçlar vermesine sebep olacağından indirilen resimler tarafımdan kontrol edilmiştir. Her bir grafiğin belirtilen kategori içerisindeki bir örnek olması sağlanmıştır.

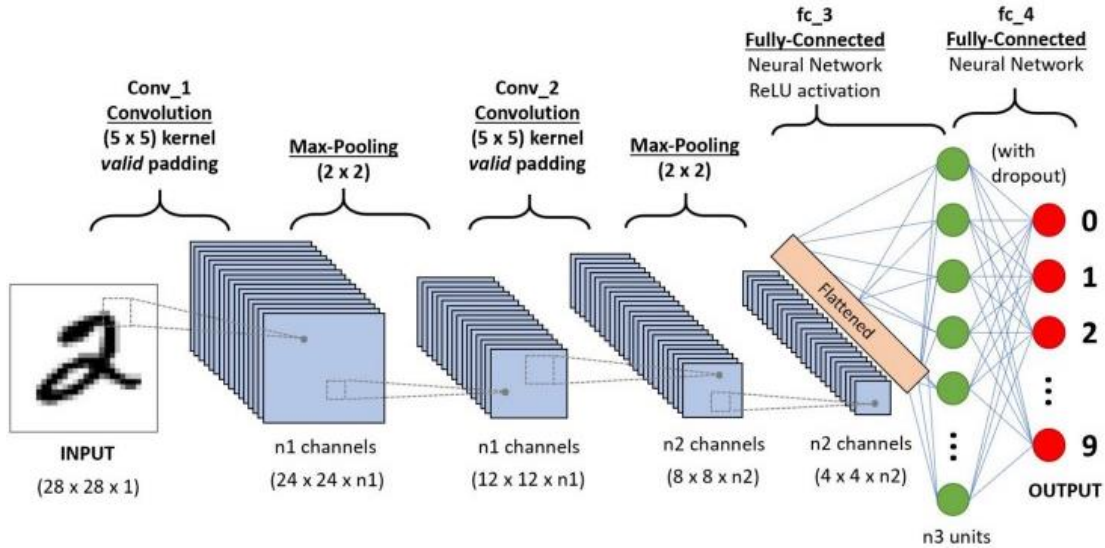
Grafik veriseti, çizgi grafiği, sütun grafiği, pasta grafiği ve dağılım grafiği olarak dört farklı türünü içermektedir. Her bir grafikten 300 adet olmak üzere toplamda 1200 grafik resmi bulunmaktadır. Verisetindeki grafikler farklı boyutlar, görsellik ve renk dağılımlarında oluşturulmuş görüntüler içermektedir. Şekil 1 verisetinden rastgele seçilmiş farklı grafik türlerini göstermektedir. Sınıflandırıcı modellerinde giriş sayısının aynı olma zorunluluğundan dolayı grafiklerin sabit bir boyut üzerinden CNN ve DBN modellerini beslemesi gerekir. Bunun için tüm grafikler 224x224 piksel büyüklüğüne yeniden boyutlandırılmıştır. Derin öğrenme modellerinin eğitimi oldukça zaman alan bir süreç olduğu için, görüntü analizlerinde ve modellemede harcanan zamanı azaltmak için resimler gri seviye resimlere dönüştürme tercih edilmektedir (Krizhevsky ve ark., 2012). Bu yüzden bu çalışmada da farklı resimlerin renk yoğunluklarını taramak için görüntünün belirli bantlarından (Kırmızı-R, Yeşil-G, Mavi-B) incelenmesi yerine tek banta düşürülerek yeniden boyutlandırmanın ardından tüm resimler gri seviye resme dönüştürülmüştür. Bu sayede grafik görüntüleri 0-255 arasında değişen piksel değerleri içeren bir matris oluşturmuştur.

2.2. Derin Öğrenme

Derin öğrenmenin diğer sınıflandırıcılardan en önemli farklı denetimli sınıflandırma modellerinin öncesine denetimsiz sınıflandırma modelleri ekleyerek her bir öznitelik için belirlenmiş rastgele ağırlık uzayını girişlere bağlı hale getirmesidir. Bu durum belirli bir uzaya toplanmış ağırlıkların geri yönlü besleme ile eğitilmesini kolaylaştıran en önemli özelliklerden biridir. Bu yapı önceden denetimsiz olarak hesaplanan nöron değerleri ve nöronlar arası bağlantıların ağırlıkların, denetimsiz olarak belirlenip bu ön tanımlı sınıflandırıcı parametrelerinin geri yayılım gibi denetimli öğrenme modellerine transferini yaparak optimizasyonunu gerçekleştirmektedir. Bu da eğitimin kısa zamanda minimum hata oranına yaklaşmasını dolayısıyla da hızlı ve dengeli öğrenme algoritmalarını üzerinde barındırmasını sağlamaktadır. Bu çalışmada giriş olarak verilen görüntünün farklı yansımalarını çıkararak bu yeni yansımalar üzerinde piksel değerlerinin baskınlığa göre anlamlı değerleri transfer eden CNN ve DBN algoritmalarının matematiksel altyapısı detaylandırılmıştır.

2.2.1. Konvolüsyonel Sinir Ağları (Convolutional Neural Networks)

CNN son yıllarda görüntü analizi süreçlerini üzerinde bulundurması sebebiyle bilgisayarlı görü çalışmalarında en yaygın kullanılan derin öğrenme modelidir. CNN, ilk aşamasında görüntünün piksel değerlerinden elde edilmiş matrise belirli katmanlarda yöntem ismini veren konvolüsyon işleminin uygulanması ile başlar. Konvolüsyon işlemi uygulandığı filtre türüne göre ilk aşamalarda düşük seviye, sonraki aşamalarda orta ve yüksek seviyeli özniteliklerin çıkarılmasını sağlar. Yani, ilk katmanlarda düşük seviyeli öznitelikleri öğrenen model sonraki katmanlara ilk katmanlardan elde ettiği düşük seviyeli özniteliklerden en baskın olan pikselleri sonraki katmanlar aktarır. Bu süreç öznitelik öğrenmesi (feature learning) ya da öğrenmenin transferi (transfer learning) isimleri verilir. CNN algoritması temelde konvolüsyon (Convolution-conv), havuzlama (pooling), doğrusallaştırılmış lineer birim (Rectified linear unit-RELU), düzleşme (Flattening) ve tam bağlı katmanlar (Fully connected layer-FC) katmanları, nöronları benzerliklerine göre saf dışı bırakan düşürme (Dropout) yöntemi ve genellikle yapay sinir ağları modelinin aktivasyon fonksiyonu olarak kullanılan softmax aşamasından meydana gelir. Genel olarak bir CNN modeli Şekil 2'de resmedilmiştir.



Şekil 2. Karakter tanıma üzerine modellenmiş bir CNN modeli (Kaggle, 2019)

Konvolüsyon işlemi, bir matrise belirli boyuttaki farklı filtre çekirdeklerinin belirli aralıklarla dolaştırılması işlemine denir. Her katmanda kullanılan filtre sayısı farklılık gösterebilir. Her filtre kendin özgü kenar tespit yapısı içerir. Filtreler genellikle kare matris olmasına karşın, bazı çalışmalarda satır sütun sayısı farklı filtreler de kullanılmıştır. Giriş resmine konvolüsyon işlemi uygulanan her filtre üst üste bildirilerek bir tensör elde edilir. Kullanılan filtre sayısı oluşacak olan tensörün derinliğini belirler. Konvolüsyon işlemi sonrasında elde edilecek yeni görüntünün boyutuna (W_{out}), giriş resminin boyutuna (W_{in}), filtre boyutuna (Filtre size- F), konvolüsyon

penceresinin adım büyüklüğüne (stride- S) ve resme çerçeve ekleme (padding- P) sayısına bağlı olarak değişir. $W_{out} = (W_{in} - F + 2P)/S + 1$ formülü kullanılarak hesaplanır.

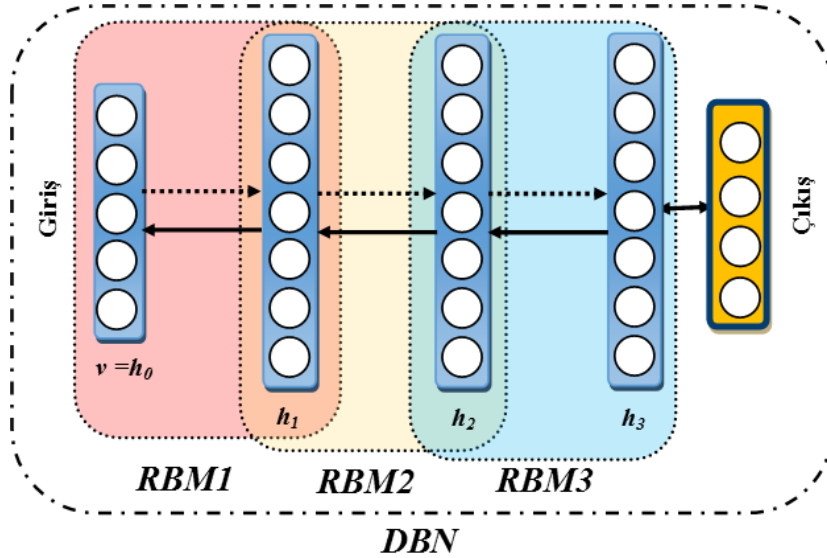
Havuzlama işleminde, konvolüsyon işleminden elde edilmiş her bir matrisi giriş resmine uygulanmış farklı bir filtrenin sonucu elde edilen tensöre belirlenen havuzlama filtresi büyüklüğünde ve adım aralığında tensör üzerinde hareket ettirilerek havuzlama filtresinin türüne göre minimum, maksimum, ortalama, orta değerlerden biri tüm derinlikler için hesaplanır. Havuzlama işlemi aslında bir örnek azaltma (down-sampling) işlemidir. Mantığında giriş olarak verilmiş resmin küçük olarak da görüntü içerisinde tespit edilmesi için girişin boyutunun azaltılarak sürece dâhil edilmesi vardır. Literatürde genellikle havuzlama filtre boyutu 2 olarak tercih edilerek piksel sayısı yani resmin boyutu yarıya indirilir.

RELU aşamasında, filtreleme sonrası elde edilen piksel değerlerinin negatif olanlarına 0 değeri verilirken, diğerlerine lineer bir yapıda kendi değeri sabit kalacak şekilde matrisi düzenler. Kısacası negatif değerler için sıfır, pozitif değerler için girişin kendi değerini çıkış olarak veren bir fonksiyondur. RELU kullanılma sebebi de CNN genellikle görüntü işleme süreçleri için kullanıldığından piksel değerinin 0-255 aralığında olma zorunluluğundan kaynaklanmaktadır.

Düzleştirme aşaması CNN modelinin çıkışında elde edilen son görüntü matrisinin vektöre dönüştürülmesini ifade etmektedir. Bu sayede girişe ait belirginleştirilmiş özellikler yapay sinir ağları modeline giriş olarak besleme yapılacak şekilde dönüştürülmektedir. Düzleştirme sonrası elde edilen vektörün her bir skaler büyüklük, yapay sinir ağları modeli için bir öznitelik olarak ele alınır. Tam bağlı yapay sinir ağları modelleri kullanılacak katman sayısını belirler. Her katman için CNN modelde bir FC görüntülenir. Her FC içerisinde bulunan tüm nöronların kendinden sonra gelen FC içerisindeki her bir nöronla arasında bağlantısının olması yapıya ismini veren temel etmendir. Nöronlar arasında tam bağın olması dolayısıyla eğitim ve öğrenme sürecinde geri yayılım algoritmaları için oldukça fazla parametrenin optimizasyonunu gerektirir. Bu aşamada dropout yöntemi ön plana çıkarak her bir FC içerisinde girilen argüman değeri oranında birbirine benzer nöronları modelden kaldırarak nöron sayısını dolayısıyla da optimize edilmesi gereken parametre sayısını azaltarak eğitimi hızlandırıcı bir yöntem olarak ön plana çıkar.

2.2.2. Derin İnanç Ağları (Deep Belief Networks)

DBN olasılık ve enerji temelli bir sınıflandırıcı modelidir. DBN iki aşamalı olarak eğitim sürecine dâhil olur. Birinci aşamada kısıtlanmış Boltzmann makineleri (Restricted Boltzmann Machines-RBM) kullanarak iki katman arasındaki koşullu olasılığa bağlı olarak denetimsiz öğrenme modelleriyle nöronlar arasındaki ağırlıkların değerleri için ön tanımlama yapar (Hinton ve ark., 2006). Bu aşama ağırlıklar için rastsallıktan çıkarıp belirli bir fonksiyona göre belirleme yaptığı için bu parametrelerin optimizasyonu oldukça kolay ve hızlı şekilde en uygun modelin tasarlanması sağlanabilmektedir. İkinci aşamada ise RBM tarafından ön tanımlaması yapılmış sınıflandırma parametrelerinin yapay sinir ağları modeliyle denetimli öğrenme dâhil edilerek ağırlıkların ince ayar (Fine-tuning) optimizasyonu yapılır (Altan ve ark., 2016b).



Şekil 3. Üç Gizli katmandan meydana gelen DBN modeli

DBN modeli her katmandaki nöron sayıları belirlenerek oluşturulur. Şekil 3 genel bir DBN modelinin yapısını göstermektedir. Her katman arasında sonraki katmandaki nöronlar arası tam bağlı durum vardır. Herbir katman içerisindeki nöronların birbirine bağlantısı bulunmamaktadır. Her iki katman bir araya gelerek bir RBM oluşturur. Giriş katmanı ile ilk gizli katman arasında hesaplanan bağlantı ağırlıkları kullanılarak sonraki katman için nöron ağırlıkları belirlenir. Gizli katman sayısı kadar RBM mevcuttur. Her bir RBM bağımsız olarak hesaplanarak elde edilen değerler girişten çıkışa doğru katman katman hırslı (Greedy layer-wise pre-training) şekilde hesaplanır (Hinton ve ark., 2006). Ağırlıkların denetimsiz RBM modeliyle belirlenmesi işlemine ön eğitimli model denir. RBM modelin bias ve ağırlık değerleri enerji fonksiyonunun olasılıksal dağılımına eşittir. DBN modelinin enerji fonksiyonu $E(v, h)$ ve olasılık fonksiyonu $P(v, h)$ aşağıdaki şekilde ifade edilir:

$$E(v, h; \theta) = - \sum_{i=1}^v \sum_{j=1}^h W_{ij} v_i h_j - \sum_{i=1}^v b_i v_i - \sum_{j=1}^h c_j h_j \quad (1)$$

formüllerdeki $\theta = \{W, b, c\}$ ve v giriş katmanına ait vektör, h gizli katmana ait vektörü ifade eder. Ayrıca b ve c DBN modelde sırasıyla giriş ve gizli katmanın biasleri olarak tanımlanır. DBN modelinin ön eğitim sürecindeki c_j ve b_i değerleri arasındaki ağırlıklar olarak W_{ij} olarak ifade edilir.

$$P(v, h) = \frac{e^{-E(v, h)}}{\sum_v \sum_h e^{-E(v, h)}} \quad (2)$$

RBM dağılımının giriş katmanı için tanımlanmış olasılık fonksiyonu $P(v)$ 'dir. W , RBM modeli için belirlenmiş olasılık fonksiyonunun kısmi türevi olarak formüle edilir.

$$P(v) = \frac{\sum_h e^{-E(v, h)}}{\sum_v \sum_h e^{-E(v, h)}} \quad (3)$$

$$\frac{\partial \log(P(v))}{\partial W_{ij}} = \langle v_i h_j \rangle_v - \langle v_i h_j \rangle_{model} \quad (4)$$

W ağırlığının her bir eğitim devirinde (epoch) güncellenmiş hali modelin öğrenme oranına (η) bağlı olarak güncellenir:

$$\nabla W_{ij} = \eta (\langle v_i h_j \rangle_{data} - \langle v_i h_j \rangle_{model}) \quad (5)$$

Gizli katmanlar ve giriş birimlerinin koşullu olasılık dağılımı δ , sigmoid fonksiyon olacak biçimde:

$$P(h_j = 1 | v) = \delta \left(\sum_i v_i W_{ij} + c_j \right) \quad (6)$$

$$P(v_i = 1 | h) = \delta \left(\sum_j h_j W_{ij} + b_i \right) \quad (7)$$

DBN eğitiminde iterasyonlar P fonksiyonunun değerini maksimize ederken, E fonksiyonunun değerini minimize eder. Geri kalan eşitlikler Gibbs sampling yaklaşımı temel alınarak eğitime devam edilir (Hinton, 2009; Hinton ve ark., 2006).

3. Araştırma Sonuçları

Özellikle uzun süreçlerin takibi ve büyük verilerin analizi söz konusu olduğunda bilgisayar destekli yöntemlerin kullanılabilirliği ve güvenilirliği oldukça önem kazanan parametrelere dönüşmektedir. Bilgisayar destekli analiz sistemleri özellikle süreçlerin yorumlanarak benzer durumlara uyarlanması ve yeni süreçlere aktarılmasında yeni fonksiyonelliklerle kullanılabilirliğini artırmıştır. Görüntü işleme yaklaşımlarıyla grafiklerin genel yapılarının farkına vararak çıkarımlarda bulunacak ve grafiklerin türünü belirleyecek nitelikte sistemler modellemeye yönelik çalışmalarda hala kat edilmesi gereken uzun bir yol mevcuttur. Grafik sınıflandırma modeli ileriki süreçler için genelleme kapasitesi yüksek olması gereken ön aşamaları teşkil etmektedir. Bir verisetinde sınıflandırılacak türlerin güvenilirlik ve geçerliliğinin belirlenmesinde kullanılan bağımsız test karakteristiklerinin hesaplanması ve yorumlanması öncelikli bir aşamadır. Bu çalışmada hassasiyet (recall, sensitivity), seçicilik (specificity, selectivity), kesinlik (precision, positive predictive value-PNV), tahmin edilen negatif değer (Negative predictive value-NPV), genel başarımlar (accuracy) ve F1 skoru (F1 score) değerleri sınıflandırıcı modelin değerlendirilmesi için kullanılan test karakteristikleridir (Duda, Hart ve Stork, 2000).

Dört farklı grafik türünden oluşturulan kütüphane ile elde edilen verilerin ön kontrolü gerçekleştirilerek analizlere uygun olup olmadığı, kodlanan python kütüphanesinin doğru bir kategorileme yapıldığı yapılmadığı kontrol edilmiştir. 224x224 piksel büyüklüğüne yeniden boyutlandırılıp gri seviyeye dönüştürülen resimler sınıflandırma sürecine hazır hale getirilmiştir. 1200 grafik görüntüsü 5 kat çapraz doğrulama kullanılarak eğitim ve testleri gerçekleştirilmiştir. Bu yöntemde eşit sayıda bulunan grafik türleri homojen olacak şekilde beş parçaya ayrılmıştır. Ayrılan her parça aynı grafik türünden eşit sayıda olacak biçimde görüntü içermektedir. Analizlerde bu 5 parçanın dördü sınıflandırıcı modelin eğitimi için kullanılırken, kalan parça eğitilen modelin test edilmesi için kullanılır. Dolayısıyla verisetinin %80'i eğitim, %20'si test için kullanılır. Bu işlem tüm parçalar birbirinden bağımsız olacak şekilde hem test hem de eğitimde kullanılabilecek kadar devam eder. Elde edilen test başarımlarının ortalaması sistemin genel başarımlarını belirler. Grafik sınıflandırma modelleri CNN ve DBN üzerinde denenmiştir. DBN ve CNN farklı sınıflandırıcı parametrelerine sahip olduğu için her ikisinin kesin bir kıyaslaması mümkün olmamakla birlikte, gizli katman sayıları, her gizli katmandaki nöron sayıları, konvolüsyon işleminde kullanılan filtre büyüklükleri ve havuzlama özellikleri deneysel olarak belirli aralıklarda denenecek modellenmiştir.

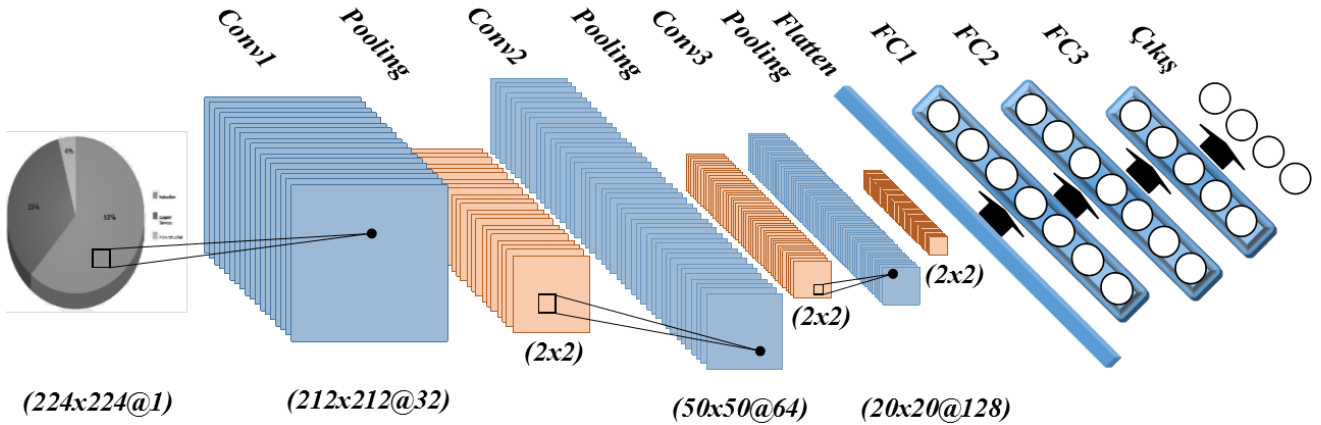
Modellenen DBN sınıflandırıcılar, grafikler üzerinde farklı yansımalarındaki analizini sağlamak için 3 gizli katmandan meydana gelecek şekilde oluşturulmuştur. Gizli katmalarda kullanılan nöron sayıları 50 ile 250 nöron arasında her deneyde her bir katmandaki nöron sayısı 10 artacak şekilde seçilmiştir. Yapı olarak belirlenen aralıklar brute force yapısı mantığında her durum için tekrarlanmıştır.

Belirlenen aralıkta oldukça fazla model gerçekleştirildiği için ulaşılan en yüksek başarımlar paylaşılmıştır. DBN için sigmoid aktivasyon fonksiyonu kullanılmış ve geriye yayılım yaklaşımında öğrenme oranı olarak 0,001 seçilmiştir. Öğrenme oranının düşük olması modelin global değere ulaşması için zorluk çıkarabileceği göz önünde bulundurularak, iterasyon sayısı 100 olarak belirlenmiştir. Sınıflandırıcılardan elde edilen test karakteristikleri farklı nöron sayısına sahip modeller için sistem performansları Tablo 1’de göstermiştir. CNN sınıflandırıcı için denenen parametreler ise farklı sayıda konvolüsyon işlemleri, her konvolüsyon için belirlenen farklı filtre boyutları, filtre sayıları, dropout benzerlik parametreleri, sinir ağları modeli için kullanılan FC sayısı ve her FC’de yer alan nöron sayısı olarak ayarlanmıştır. Bunlara ek olarak havuzlama katmanında uygulanan filtrenin boyutu sabitlenerek her seferinde görüntü boyutunun yarıya indirilmesi kararlaştırılmıştır. Kıyaslama yapılabilmesi için FC içerisinde denenen nöron aynı aralıkta değiştirilmiş ve üç gizli katman kullanılmıştır. Bu durumda kıyaslanan kısım aslında DBN ve CNN modellerinin farklı yansıma oluşturmadaki başarımlarının kıyaslanması sağlanmıştır. RBM ve konvolüsyon işlemlerinin grafik görüntülerinin yeni yansımaları için en anlamlı yeniden biçimlendirmenin hangi yöntem olduğunun belirlenmesi sağlanmıştır.

Tablo 1. Farklı derin öğrenme modellerinin grafik sınıflandırma performans karakteristikleri

Derin Öğrenme Modeli	Sınıflandırıcı	Hassasiyet	Seçicilik	Keskinlik	NPV	F1 değeri	Başarımlar
Conv1(9x9@32, S=1, P=0)	CNN	91,75	92,42	92,37	91,80	0,92	92,08
Conv2(9x9@64, S=2, P=0)							
Conv3(4x4@128, S=1, P=0)							
FC1=70, FC2=110, FC3=140							
Conv1(13x13@32, S=1, P=0)	CNN	94,50	91,33	91,60	94,32	0,93	92,92
Conv2(8x8@64, S=2, P=0)							
Conv3(6x6@128, S=1, P=0)							
FC1=120, FC2=90, FC3=190							
$h_1=80, h_2=110, h_3=190$	DBN	89,50	88,42	88,54	89,39	0,89	88,96
$h_1=60, h_2=180, h_3=180$		89,17	90,92	90,75	89,35	0,90	90,04

Derin öğrenme yaklaşımlarında belirli bir yapının olmaması, optimize edilmesi gereken parametrelerin fazla olması, katman sayısı ve katmanlardaki nöron sayılarının değişkenliği sebebiyle model geliştirmek zaman alan bir süreçtir. Literatürde görüntü üzerine AlexNet, VGGNet, GoogleNet gibi popüler modeller kullanılarak oldukça yüksek başarımlı sınıflamalar gerçekleştirilse de, bu modellerin daha az kompleks olarak oluşturulması ve modellenmesi günümüzde üzerine yoğunlaşmış en temel araştırma alanlarıdır. Önerilen CNN ve DBN modelleri bu manada literatürdeki çalışmalara nazaran daha basit modeller olarak yüksek başarımlar elde etmeyi temel alacak nitelikte gerçekleştirilmiştir. En yüksek performansa sahip derin öğrenme modelinin yapısına ilk gizli katmanında 130 nöron, ikinci gizli katmanında 140 nöron, üçüncü gizli katmanda 200 nöron bulunan DBN modeliyle gerçekleştirilmiştir. CNN modelinde en yüksek başarımlar 2 konvolüsyon katmanından oluşan modeller, yapay sinir ağları katmanında ilk FC’de 90, ikinci FC’de 170 ve üçüncü FC’de 190 nörona meydana geliştirilmiştir. Grafik sınıflandırıcıda gerçekleştirilen CNN modeli Şekil 4’te detaylandırılmıştır.



Şekil 4. En yüksek grafik sınıflandırma başarımına ulaşan CNN modelinin yapısı

En yüksek grafik sınıflandırma başarısı CNN model ile elde edilmiştir. Bu model üç konvolüsyon katmanı, her konvolüsyon sonrası maksimum havuzlama (max pooling) gerçekleştirilerek elde edilmiştir. Tüm maksimum havuzlama katmanları 2x2 filtre büyüklüğünde ele alınmıştır. Her pooling katmanından sonra RELU fonksiyonundan geçirilmiştir. İlk katmanda filtre boyutu 13x13 ve stride=1 alınarak 32 farklı filtre kullanılmıştır. İkinci konvolüsyon katmanında kullanılan filtre boyutu 8x8 ve stride=2 alınarak 64 farklı filtre kullanılmıştır. Son konvolüsyon katmanında filtre boyutu olarak 6x6 ve stride=1 alınarak 128 farklı filtreden geçirilen yansımalar sonrasında uygulanan maksimum havuzlama ile görüntü boyutu 10x10 piksele indirilmiştir. Grafik görüntülerinden farklı filtrelerle elde edilen 10x10 boyutundaki bu görüntü düzleştirme ile 100 öznitelik olacak şekilde yapay sinir ağları modeline verilerek eğitilmiştir. En yüksek başarımlar ilk FC katmanında 120 nöron, FC2’de 90 nöron ve FC3’te 190 nöron kullanılarak elde edilmiştir.

4. Sonuç ve Tartışma

Bilgisayar destekli grafik analizleri geçtiğimiz günlerde veri değerlendirme süreçler için ön plana gelmiştir. Grafik sınıflandırma çalışmaları grafiğin türünün belirlenmesi aşamasında önemli bir mihenk taşı olma özelliğine sahiptir. Çok aşamalı analiz sistemlerinde grafiklerin yorumlanmasından önce hangi grafik türünün olduğunu, bu grafik türüne ait farklı yapıların tanınması ve daha sonrasında nicelenen değerlerin yorumlanması gibi yaklaşımlar mevcuttur. Özellikle görüntü işleme tekniklerindeki hızlı ilerleyiş, bilgisayarlı görü, farklı görüntü türlerinden nesne tespiti, daha önceden öğrenilmiş örüntülerden yeni görüntüler meydana getirme yaklaşımlarına olumlu bir ivme kazandırmıştır. Bu çalışmada da grafik resimlerinin işlenmesi üzerine oldukça popüler CNN ve DBN modellerinin ayrıştırma kabiliyeti kıyaslanarak yetenekleri deneyimlenmiştir.

Tablo 1. Benzer Grafik Sınıflandırma Çalışmaları

Çalışmalar	Grafik Türleri	Veriseti	Sınıflandırıcı	Başarım (%)
Tang vd. (2016)	Sütun, Akış diagramı, Pasta, Çizgi, Dağılım	5000 grafik	DBN+ CNN	75,40
Karthikeyani vd. (2012)	Sütun, Donat, Pasta, Çizgi	155 grafik	k-NN	78,06
Savva vd. (2011)	Sütun, Pasta, Çizgi, Ven, Tablo, Radar, Dağılım, Pareto, Harita, Eğri	2601 grafik	Destek Vektör Makineleri	80,00
Huang vd. (2007)	Sütun, Pasta, Donat, Çizgi	210 grafik	Çoklu örnek öğrenme	94,00
Amara vd. (2017)	Sütun, Pasta, Çizgi, Ven, Tablo, Radar, Eğri, Dağılım, Pareto, Harita, Nod Bağlantı	3377 grafik	CNN	89,50
Önerilen Modeller	Sütun, Pasta, Çizgi, Dağılım	1200 grafik	DBN CNN	90,04 92,92

Tablo 2’de grafik sınıflandırma üzerine başarılı çalışmaların ayrıştırdığı grafik türleri, kullandığı grafik verisetinin büyüklüğü, sınıflandırıcı modelleri ve sınıflandırma başarımlarının kıyaslanması verilmiştir. Huang vd. (2007) şekil tespiti gerçekleştirilerek elde edilen şekillerden tanımlayıcı öznitelikler kullanarak korelasyon faktörleri belirlemişlerdir. Belirlenen bu korelasyon faktörleri çoklu örnek öğrenme modeli ile eğiterek yoğunluk ayırma algoritmasının (Diverse Density Algorithm) avantajlarıyla dört farklı grafik türünü %94 genel başarımla ayrıştırmışlardır. Savva vd. (2011) görüntü işleme ve makine öğrenmesinin esnekliklerini birleştirerek elde ettikleri sınıflandırıcı modelde on farklı grafik türünün ayrıştırılmasını sağlamıştır. Huang vd. göre çok daha geniş bir veriseti üzerinde analizlerini gerçekleştiren Savva ve arkadaşları düşük seviyeli görüntü öznitelikleri olan gradyan ve yerel alan dokularını kullanarak %80 genel başarımla ulaşılmıştır. Grafik sınıflandırmaya ek olarak sütun ve pasta grafiklerinin birbirine dönüşümü ve grafiklerden metinsel ifadelerin tespitini de gerçekleştirmişlerdir. Karthikeyani vd. (2012) çok sınırlı bir veriseti üzerinde gerçekleştirdikleri analizlerini pdf dökümanlarından elde edilen temelde dört, 3-boyutlu halleriyle onbir farklı grafik türünü, en yakın komşu sınıflandırıcısı ile %78,06 genel başarımla ayrıştırmayı başarmışlardır. Giriş olarak verilen grafiklerin gri seviye eş oluşum matrisi (Gray Level Co-Occurrence Matrix) öznitelikleri olan medyan, alan, minimum ve maksimum yoğunluk, kontrast, homojenlik, enerji, entropi, ortalama, varyans, standart sapma ve korelasyon özelliklerini kullanarak farklı makine öğrenmesi algoritmalarının etkinliğini test etmişlerdir. Tang vd. (2016) oldukça geniş bir grafik veriseti üzerindeki çalışmalarını DBN, CNN gibi derin öğrenme modelleriyle etkinliklerini kontrol etmişlerdir. Herhangi bir öznitelik çıkarma algoritmasına bağlı kalmaksızın resimleri sınıflandırıcı modellere giriş olarak besleyen araştırmacılar oldukça başarılı sonuçlar elde etmişlerdir. CNN modelinin son aşamasındaki denetimli öğrenme modeli olan yapay sinir ağıları modelinin yerine DBN sınıflandırıcısını bütünleştirerek yeni bir model öneren Tang ve arkadaşları bu modelle beş farklı grafik türünü %75,40 genel başarımla ayırmayı sağlayabilmişlerdir. Amara vd. (2017) farklı grafik verisetlerin birleştirilmesiyle bir araya getirdikleri detaylı bir grafik veritabanı üzerinde LeNET modeli ile gerçekleştirdikleri CNN sınıflandırıcısının etkisini kontrol etmişlerdir. Gerçekleştirdikleri sınıflandırma modeliyle onbir farklı türden grafiği %89,50 genel başarımla ayrıştırabilmişlerdir. Yalnızca sınıflandırma yaklaşımlarıyla yetinmeyip grafik anlamlandırmak için metin modelleme üzerine de çalışmış ve grafiklerden metinsel içerikleri ayrıştırma yoluna gitmişlerdir. Literatüdeki çalışmalar göstermektedir ki, klasik makine öğrenmesi modelleri sınıflandırma öncesinde belirli öznitelik çıkarma yaklaşımlarıyla anlamlı özniteliklerin belirlenmesine ihtiyaç duyarken, derin öğrenme yaklaşımları doğrudan grafikler üzerinde düşük, orta ve yüksek seviyeli öznitelik çıkarma kabiliyetleriyle sınıflandırma modellerine oldukça etkin ve değişikliklere karşı dayanıklı yaklaşımlar sağlamaktadır. Bu modellerin birbiriyle bütünleşik hale getirilmesi de farklı örüntülerin ve yansımaların üretilmesinde oldukça başarılıdır.

Deneyimlenen DBN ve CNN modellerinde literatürdeki popüler modellere göre oldukça az sınıflandırıcı parametre sayısıyla sırasıyla %90,04 ve %92,92 genel başarımla dört farklı grafiğin sınıflandırılması gerçekleştirilebilmiştir. Katman sayıları, her katmandaki nöron sayıları, konvolüsyon ve filtre özellikleri belirli aralıklarda denenmiş olduğu için daha detaylı ve derin modellerle daha yüksek başarımlar elde edilmesi mümkündür. İleri çalışmalarda bu aralıkların genişletilmesi, verisetindeki grafik sayısının artırılması derin öğrenme modelleri için eğitimde daha başarılı ayırt etme performanslara erişeceği öngörülmektedir. Ayrıca, başarılı bir grafik sınıflandırıcı modelin gerçekleşmesinin ardından verilerin nicellenmesi ve görsel araç-gerecin yorumlanması büyük verinin analizinde çok önemli bir adım olarak doğal dil işleme yaklaşımı gibi bir disiplin oluşturacaktır.

Kaynakça

- Abdel-Zaher, A. M., & Eldeib, A. M. (2016). Breast cancer classification using deep belief networks. *Expert Systems with Applications*. <https://doi.org/10.1016/j.eswa.2015.10.015>
- Altan, G., Kutlu, Y., & Allahverdi, N. (2016a). A Multistage Deep Belief Networks Application on Arrhythmia Classification. *International Journal of Intelligent Systems and Applications in Engineering*, 4(Special Issue-1), 222–228. <https://doi.org/10.18201/IJISAE.270367>
- Altan, G., Kutlu, Y., & Allahverdi, N. (2016b). Deep Belief Networks Based Brain Activity Classification Using EEG from Slow Cortical Potentials in Stroke. *International Journal of Applied Mathematics, Electronics and Computers*. <https://doi.org/10.18100/ijamec.270307>
- Amara, J., Kaur, P., Owonibi, M., & Bouaziz, B. (2017). Convolutional Neural Network Based Chart Image Classification. In *25th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision in co-operation with EUROGRAPHICS Association* (pp. 83–88).
- Butler, G., Grogono, P., Shinghal, R., & Tjandra, I. (1995). Analyzing the logical structure of data flow diagrams in software documents. In *Proceedings of the International Conference on Document Analysis and Recognition, ICDAR*. <https://doi.org/10.1109/ICDAR.1995.601962>
- Cireşan, Dan C., Meier, U., Masci, J., Gambardella, L. M., & Schmidhuber, J. (2011). Flexible, high performance convolutional neural networks for image classification. In *IJCAI International Joint Conference on Artificial Intelligence*. <https://doi.org/10.5591/978-1-57735-516-8/IJCAI11-210>
- Cireşan, Dan C., Meier, U., Gambardella, L. M., & Schmidhuber, J. (2011). Convolutional neural network committees for handwritten character classification. In *Proceedings of the International Conference on Document Analysis and Recognition, ICDAR*. <https://doi.org/10.1109/ICDAR.2011.229>
- Dai, W., Wang, M., Niu, Z., & Zhang, J. (2018). Chart decoder: Generating textual and numeric information from chart images automatically. *Journal of Visual Languages and Computing*. <https://doi.org/10.1016/j.jvlc.2018.08.005>
- Duda, R. O., Hart, P. E., & Stork, D. G. (2000). *Pattern Classification*. New York: John Wiley, Section. <https://doi.org/10.1038/npp.2011.9>
- Hinton, G. (2009). Deep belief networks. *Scholarpedia*, 4(5), 5947. <https://doi.org/10.4249/scholarpedia.5947>
- Hinton, G. E., Osindero, S., & Teh, Y.-W. (2006). A fast learning algorithm for deep belief nets. *Neural Computation*, 18(7), 1527–1554. <https://doi.org/10.1162/neco.2006.18.7.1527>
- Huang, G. B., Lee, H., & Learned-Miller, E. (2012). Learning hierarchical representations for face verification with convolutional deep belief networks. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. <https://doi.org/10.1109/CVPR.2012.6247968>
- Huang, W., Zong, S., & Tan, C. L. (2007). Chart image classification using multiple-instance learning. In *Proceedings - IEEE Workshop on Applications of Computer Vision, WACV 2007*. <https://doi.org/10.1109/WACV.2007.17>
- Jung, D., Kim, W., Song, H., Hwang, J., Lee, B., Kim, B., & Seo, J. (2017). ChartSense: Interactive Data Extraction from Chart Images. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI '17*. <https://doi.org/10.1145/3025453.3025957>
- Kaggle. (2019). CNN tutorial. Retrieved from <https://www.kaggle.com/nhlr21/deep-keras-cnn-tutorial>
- Kalchbrenner, N., Grefenstette, E., & Blunsom, P. (2014). A Convolutional Neural Network for Modelling Sentences. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. <https://doi.org/10.3115/v1/P14-1062>
- Karpathy, A., Toderici, G., Shetty, S., Leung, T., Sukthankar, R., & Li, F. F. (2014). Large-scale video classification with convolutional neural networks. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. <https://doi.org/10.1109/CVPR.2014.223>
- Karthikeyani, V., & Nagarajan, S. (2012). Machine Learning Classification Algorithms to Recognize Chart Types in Portable Document Format (PDF) Files. *International Journal of Computer Applications*. <https://doi.org/10.5120/4789-6997>
- Kim, Y. (2014). Convolutional Neural Networks for Sentence Classification. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. <https://doi.org/10.3115/v1/D14-1181>
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. *Advances In Neural Information Processing Systems*. <https://doi.org/http://dx.doi.org/10.1016/j.protcy.2014.09.007>
- Mohamed, A. R., Dahl, G. E., & Hinton, G. (2012). Acoustic modeling using deep belief networks. *IEEE Transactions on Audio, Speech and Language Processing*. <https://doi.org/10.1109/TASL.2011.2109382>
- Poco, J., & Heer, J. (2017). Reverse-Engineering Visualizations: Recovering Visual Encodings from Chart Images. *Computer Graphics Forum*. <https://doi.org/10.1111/cgf.13193>
- Savva, M., Kong, N., Chhajta, A., Li, F.-F., Agrawala, M., & Heer, J. (2011). ReVision: Automated classification, analysis and redesign of chart images. In *UIST'11 - Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology*. <https://doi.org/10.1145/2047196.2047247>
- Shadish, W. R., Brasil, I. C. C., Illingworth, D. A., White, K. D., Galindo, R., Nagler, E. D., & Rindskopf, D. M. (2009). Using UnGraph to extract data from image files: Verification of reliability and validity. *Behavior Research Methods*. <https://doi.org/10.3758/BRM.41.1.177>
- Siddiqui, S. A., Salman, A., Malik, M. I., Shafait, F., Mian, A., Shortis, M. R., & Harvey, E. S. (2018). Automatic fish species classification in underwater videos: Exploiting pre-trained deep neural network models to compensate for limited labelled data. *ICES Journal of Marine Science*. <https://doi.org/10.1093/icesjms/fsx109>

- Tang, B., Liu, X., Lei, J., Song, M., Tao, D., Sun, S., & Dong, F. (2016). DeepChart: Combining deep convolutional networks and deep belief networks in chart classification. *Signal Processing*. <https://doi.org/10.1016/j.sigpro.2015.09.027>
- Yan Ping Zhou, & Chew Lim Tan. (2002). Hough technique for bar charts detection and recognition in document images. <https://doi.org/10.1109/icip.2000.899506>
- Yang, L., Huang, W., & Tan, C. L. (2006). Semi-automatic Ground Truth Generation for Chart Image Recognition. https://doi.org/10.1007/11669487_29
- Yu, J., Tao, D., & Wang, M. (2012). Adaptive hypergraph learning and its application in image classification. *IEEE Transactions on Image Processing*. <https://doi.org/10.1109/TIP.2012.2190083>
- Zhou, Y., & Tan, C. L. (2001). Chart analysis and recognition in document images. In *12th Proceedings of the International Conference on Document Analysis and Recognition, ICDAR* (p. 1055). <https://doi.org/10.1109/ICDAR.2001.953947>