**DEEP LEARNİNG YÖNTEMLERİ İLE NESNE TANIMA**

Cansu KAPLAN, Merve ÇOBAN, İlyas ZİYAOĞLU, Harun KARAHAN

*Bilgisayar Mühendisliği*

*Kocaeli Üniversitesi*

*İzmit, Türkiye*

[cansukaplan97@gmail.com](mailto:cansukaplan97@gmail.com) [merveeecoban@gmail.com](mailto:merveeecoban@gmail.com) [ilyasziyaoglu@gmail.com](mailto:ilyasziyaoglu@gmail.com) karahan.harunn@gmail.com

***Özet*- Derin öğrenme teknolojileri, görüntü ve konuşmanın tanınmasından hesaplama ve sistem biyolojisine kadar her alana etki etmektedir. Otomatik Kodlayıcı, Seyrek Kodlama, Sınırlı Boltzmann Makinesi, Derin İnanç Ağları ve Konvolüsyonel sinir ağları, derin öğrenmede yaygın olarak kullanılan modellerdir. Farklı model tipleri arasında, Konvolüsyonel sinir ağlarının görüntü sınıflandırmada yüksek performans gösterdiği belirlenmiştir. Nesne tanıma, görüntüdeki bir nesneyi etkin bir şekilde tanımak için kullanılan klasik bir tekniktir. Özellikle bilgisayarlı görme alanındaki teknolojilerin, yerel özellikler algılama yöntemleri ile daha karmaşık görevleri algılaması ve tanıması beklenir. Bu makalede deep learning yöntemlerinden CNN ile nesne tanıma uygulaması araştırılmıştır.**

***Anahtar Kelimeler*- Derin Öğrenme, Konvolüsyonel Sinir Ağları, Nesne Tanıma**

***Abstract*- Deep learning technologies affect every area from recognition of image and speech to calculation and system biology. Auto Encoder, sparse coding, Restricted Boltzmann Machine, Deep Belief Networks and Convolutional neural networks is commonly used models in deep learning. Among different type of models, Convolutional neural networks has been demonstrated high performance on image classification. Object recognition is classic technique used to effectively recognize an object in the image. Technologies specifically in field of computer vision are expected to detect and recognize more complex tasks with help of local features detection methods. In this article, object recognition application of CNN with deep learning methods has been investigated.**

### *Keywords*- Deep Learning, [Convolutional Neural Network](http://scholar.google.com.tr/scholar?q=convolutional+neural+network&hl=tr&as_sdt=0&as_vis=1&oi=scholart), Object Recognition

I. GİRİŞ

Son yıllarda insanlar, günlük yaşamda olan tüm operasyonları dijital sistemlerde insan gücünü azaltarak ve otomatikleştirerek yapmaya çalışmaktadır. Bu otomasyon gereksinimi akıllı sistemlerin oluşturulmasını sağlamış, yapay zeka ve makine öğrenmesi olarak adlandırılan sistemlerin uygulanması için bir ortam sağlamıştır [1]. Derin öğrenme üzerine birçok çalışma yapılmıştır ve yapılmaya devam etmektedir[2,3]. Derin Öğrenme veya Derin Sinir Ağı terimi, çok katmanlı, Yapay Sinir Ağları (YSA) anlamına gelir. Görüntü sınıflandırma, nesne izleme, poz tahmini, metin tespiti ve tanıma, görsel özellik tespiti, eylem tanıma ve sahne etiketleme konularında derin öğrenme kullanılmıştır[4].

Derin öğrenme, insan beyninin yapısını simüle eden bir yöntemdir [5]. Bu yöntem, insan beyninin her zaman maruz kaldığı bir dizi duyusal verinin önemli bir bölümünü algılama biçimini simüle ederek girdi verilerinin hiyerarşik bir gösterimini bulmak için bir dizi algoritmadır [6].

Günlük yaşam aktivitelerinde çeşitli nesnelerin tanınması bilgisayar vizyonunun en önemli uygulamalarından biridir. Nesne algılama ve tanıma tanımlanabilir; özel algoritma tabanlı modellerin uygulanmasıyla hedeflenen tüm yerleri araştırmak ve her hedefi ait olabileceği belirli bir kategori ile temsil etmek için bir görüntü, metin veya video çerçevesi sağlamaktır. İnsanlar kaotik sahnelerdeki nesneleri günlük yaşamlarında kolayca tanıyabilir ve yerelleştirebilirler. Buna karşılık, bilgisayar bunu kolayca yapamamakta ve yaparken de iyi programlanması gerekmektedir [7].

Nesne tanıma, bilgisayarda görüntü bölümlendirme, video izleme ve sahne tanımlaması için temel rol oynamaktadır.

Makine öğrenmesi ve yapay zeka gelişimi ile birlikte robotikte otomatik nesne tanıma, akıllı video izleme ve diğer alanlar gittikçe daha popüler hale gelmektedir. Son yıllarda, nesne tanıma teknikleri önemli bir ilerleme elde etmiştir, ancak görüntünün arka plan karmaşıklığı, kategori çeşitliliği ve anlamsal çeşitliliğin varlığından dolayı zorluklar çıkmıştır [8].

Önerilen çözümlerden biri, ham verilerden gerçek zamanlı olarak nesneleri izlemek, tespit etmek ve sınıflandırmak için konvülasyonel bir sinir ağı algoritması kullanan derinlemesine öğrenmeye dayalı bir yazılımın uygulanmasıdır. Son birkaç yılda, derin konvolüsyonel sinir ağları göreceli olarak yüksek doğruluk ve hızlarından dolayı görüntü nesnesi tespiti ve sınıflandırma için güvenilir bir yaklaşım olduğunu göstermiştir[9].

Konvolüsyonel Sinir Ağı, , örüntü tanıma ile ilgili çeşitli alanlarda son on yılda çığır açan sonuçlar vermiştir. CNN'lerin en faydalı yönü, YSA'da parametre sayısını azaltmaktır. Bu başarı hem araştırmacıları hem de geliştiricileri, klasik YSA'larla mümkün olmayan karmaşık görevleri çözmek için daha büyük modellere yaklaşmaya yöneltmiştir [10].

Bu yazının ilk bölümü derin öğrenme ve konvolüsyonel sinir ağları hakkında kısa bir bilgi vermektedir. İkinci bölümde konu ile ilgili araştırmalar açıklanmıştır. Üçüncü bölüm ilgili çalışmalarda kullanılan metodolojiyi gözden geçirmektedir. Dördüncü bölüm sonunda, araştırma sonuçları verilmektedir.

II. LİTERATÜR ÇALIŞMALARI

Hubel ve Wiesel’in çığır açan bulgularından esinlenerek [11] [12], Fukushima [13] Neocognitron adlı basit bir birim katmanı ve karmaşık birim katmanı yığınlarından oluşan hiyerarşik bir model önermiştir.

LeCun ve diğ. [14] [15]. Temel olarak CNN, basit birimlerin alıcı alanlarını öğrenmek için geri yayılma (BP) algoritmasını kullanmıştır. Bu model Neocognitron’dan  farklıdır.

CNN, doğumundan bu yana yerel bağlantılar, ağırlık paylaşımı ve yerel havuzlama ile karakterize edilmiştir. İlk iki özellik, modelin MLP' den daha az ayarlanabilen parametrelerle yerel bilgilendirici görsel desenleri keşfetmesini sağlamaktadır.

Geçtiğimiz yıllarda, CNN'in performansını artırmak için birçok teknik geliştirilmiştir. Doğrultulmuş doğrusal fonksiyon [16](ReLU), BP algoritmasındaki  gradyan kayma etkisine karşı direnci nedeniyle en sık kullanılan aktivasyon işlevi haline gelmiştir. Rektifiye Doğrusal Birim(ReLU), derin öğrenme modellerinde en sık kullanılan aktivasyon işlevidir. Herhangi bir negatif giriş alırsa, işlev 0 değerini döndürmekte, ancak x için herhangi bir pozitif değer için bu değeri geri döndürmektedir.

Dropout[17], sinir ağlarının eğitimde fazla kullanılmasını önlemek için etkili bir tekniktir. Modelin ortalama  dropout yeteneğinden yararlanmak için, Goodfellow ve ark. [18], etkinleştirme işlevi olarak özellik kanalları üzerinden maksimum havuzlamayı kullandılar.

Evrişimsel birimlerin doğrusal olmayan halini güçlendirmek için, Lin ve ark. [19] ağın, evrişimin yerini girdi özellik haritaları üzerinde kayan yerel çok katmanlı algılayıcı (MLP) [20] ile değiştirdiği ağ (NIN) yapısında önermiştir. NIN'in aşırı öğrenmesini  önlemek için, tam bağlı katmanlar, genel ortalama havuzlama katmanı ile değiştirilmiştir.

CNN, hesaplama yoğun bir modeldir ve genellikle CPU üzerinde çalışması zordur. GPU kullanımı, CNN'nin büyük ölçekli veri setlerinde eğitimi ve test edilmesini büyük ölçüde kolaylaştırmıştır[21].

III. DEEP LEARNİNG YÖNTEMLERİ İLE NESNE TANIMA

Derin öğrenme, son zamanlarda görüntü tanıma alanında büyük etkisi olan bir makine öğrenme yöntemidir [22]. Son birkaç yılda, çok fazla veriyi kullanabildiği için en güçlü araçlardan biri olarak kabul edilmiş ve literatürde çok popüler olmuştur. Derin öğrenme, çoklu soyutlama seviyelerine sahip verilerin gösterimini öğrenmek için çoklu işleme katmanlarından oluşan hesaplama modellerine izin vermektedir. Bu yöntemler; konuşma tanıma, görsel nesne tanıma, nesne algılama, ilaç keşfi gibi diğer birçok alanda son teknolojiyi önemli ölçüde geliştirmiştir. Derin öğrenme, bir makinenin her katmandaki gösterimini önceki katmandaki gösterimden hesaplamak için kullanılan ağırlık parametrelerini nasıl değiştirmesi gerektiğini göstermek için geri yayılma algoritmasını kullanarak, karmaşık veri kümelerinde karmaşık yapıyı keşfeder.

Derin konvolüsyonel ağlar; görüntü, video, konuşma ve ses işlemede çığır açarken, tekrarlayan ağlar metin ve konuşma gibi ardışık veriler üzerinde etkilidir [23].

Derin Sinir Ağları (DNN'ler), problemlerin çeşitli doğasına uygulanmakta ve çok etkili sonuçlar almaktadır.

Spesifik olarak, Konvolüsyonel Sinir Ağları (CNN) modeli, doğal dil işleme, bilgisayar görüşü, görüntü segmentasyonu ve sınıflandırma ile ilgili problemleri çözmek için iyi sonuçlar vermektedir [7].

Derin öğrenme yöntemleri, veri setlerinin dağınık karakteristik ifadesini keşfetmek amacıyla düşük seviyeli özelliklerin bir kombinasyonundan başlayarak daha soyut bir özellik sunumuna kadar görüntüleri öğrenir.

Genel olarak, derin öğrenme için tipik ağ modelleri üç geniş kategoriye ayrılabilir. Bunlar; Konvolüsyonel Sinir Ağı (CNN), Kısıtlı Boltzmann Makineleri (RBM) ve Yığın Otomatik Kodlayıcı Ağı (SAE)’dır. En popüler derin sinir ağlarından biri Konvolüsyonel Sinir Ağı (CNN)’dır. Bu adı, konvolüsyon olarak adlandırılan matrisler arasındaki matematiksel doğrusal işlemden alır [8].

Derin öğrenme teknolojisi, düşük seviyeli verileri öğrenmek ve temsil etmek için birden fazla işleme katmanını toplayarak hesaplama modelleri oluşturur. İşlemci hızı, bellek ve grafik işlem birimleri gibi donanımlardaki gelişmeler, derin öğrenme teknolojisinin birçok alanda kullanılmasını ve büyük başarılar elde etmesini sağlar. Farklı derin öğrenme araçları arasında, konvolüsyonel sinir ağları (ConvNet), görüntü işleme sınıflaması gibi görsel işleme alanında iyi performans sağlayan başarılı bir derin sinir ağıdır. ConvNet, önceki bilgilerden bağımsız olarak çok az ön işleme ve el yapımı özellik çıkarımı gerektirir. Son zamanlarda, ConvNet ses veri setindeki uygulamaya geçmiştir [24].

Konvolüsyonel ve tamamen birbirine bağlı katmanlar parametrelere sahiptir ancak havuz katmanı ve doğrusal olmayan katmanlar parametrelere sahip değildir. CNN, makine öğrenme problemlerinde mükemmel bir performansa sahiptir. Özellikle en büyük resim sınıflandırma veri seti (Image Net), bilgisayar vizyonu ve doğal dil işleme (NLP) gibi görüntü verileriyle ilgili uygulamalarda iyi bir performans göstermektedir [10].

Konvolüsyonel sinir ağının ağırlık paylaşma performansı, onu biyolojik sinir ağlarına benzer hale getirmektedir. Bu da ağ modelinin karmaşıklığını ve parametre sayısını azaltmaktadır. CNN tarafından çözülen problemlerle ilgili en önemli varsayım, mekansal olarak bağımlı olan özelliklere sahip olmamasıdır. Başka bir deyişle, örneğin bir yüz tanıma uygulamasında, yüzlerin resimlerde bulunduğu yere dikkat etmemize gerek yoktur. Tek sorun, verilen resimlerdeki konumlarından bağımsız olarak onları tespit etmektir.

CNN'nin bir başka önemli yönü, giriş daha derin katmanlara doğru yayıldığında soyut özellikler elde etmektir.

Girdiler çok boyutlu görüntüler olduğunda, belirli bir yer değiştirme ve deformasyon değişmezliğini koruduğunda önemli avantajlar gösterir. Konvolüsyonel sinir ağı, girişler ve çıkışlar arasında herhangi bir kesin matematiksel ifadeye ihtiyaç duymadan, çok sayıda giriş ve çıkış arasındaki haritalama ilişkisini öğrenebilen özünde çıktının giriş haritasıdır. Öncelikle eğitim sürecinde BP algoritması uygulanmaktadır.

Eğitime başlamadan önce, tüm ağırlıklar farklı küçük rasgele sayılarla başlatılmalıdır. "Küçük rasgele sayı", ağın eğitimin başarısız olmasına neden olan çok büyük ağırlıklar için doygunluğa ulaşmamasını sağlar; "farklı", ağın normal şekilde öğrenmesini garanti eder. Aslında, ağırlık matrisini aynı sayı ile başlatırsak, ağ veri kümelerini öğrenme yeteneğine sahip değildir [8].

A. Konvolüsyonel Sinir Ağları

Yapay sinir ağı modelleri ya da basitçe "sinir ağları", bağlantı modelleri, paralel dağılmış işlem modelleri ve nöromorfik sistemler gibi birçok isime dayanır. İsmi ne olursa olsun, tüm bu modeller basit hesaplama elemanlarının yoğun şekilde birbirine bağlanmasıyla iyi performans elde etmeye çalışır. Bu bakımdan, yapay sinir ağı yapısı, biyolojik sinir sistemleri konusunda şu anki anlayışımıza dayanmaktadır.

Sinirsel modeller, birçok hipotezin paralel olarak sürdürüldüğü konuşma ve görüntü tanıma, yüksek hesaplama oranlarının gerekli olduğu ve mevcut en iyi sistemlerin insan performansının eşitlenmesinden uzak olduğu alanlarda en yüksek potansiyele sahiptir [25].

Yapay sinir ağları (YSA), insan beyninin çalışma yapısını taklit ederek geliştirilen yapay zeka tekniklerinden sadece biridir. İnsan beyni, nöron şeklindeki karmaşık çok katmanlı sinir hücrelerinden oluşur. YSA'lar bazı gözlemsel verilere dayanarak programlama yoluyla yönetilir ve bu nedenle bilgisayarların bazı öngörülebilir kararları öğrenmesine ve vermesine izin verir. Birçok model, YSA' lardan ilham almaktadır. Bunlar; Konvolüsyonel Sinir Ağları, Tekrarlayan Sinir Ağları, Derin İnanç Ağlarıdır [7].

Diğer derin öğrenme yöntemleriyle karşılaştırıldığında, konvolüsyonel sinir ağı, yerel bağlantı özellikleri ve ağırlık paylaşma özellikleri için görüntü işlemede benzersiz bir üstünlüğe sahiptir [8].

CNN, çoklu işleme katmanlarından oluşan bir makine öğrenme sınıfıdır. İleri beslemeli yapay sinir ağı veya geri yayılım algoritması kullanır. İlk konvolüsyonel ağ 1989'da sunuldu.

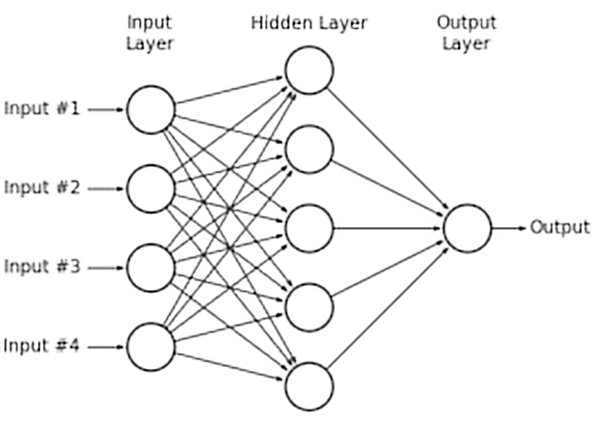
CNN'nin diğer sınıflandırıcılara göre birincil üstünlüğü, görüntü, video, konuşma ve ses işlemede daha fazla zenginleşme sağlayan diğerlerine kıyasla çok az ön işleme gerektirmesidir [26].

Konvolüsyonel sinir ağı (CNN), geleneksel sinir ağı yönteminden evrimleşmiş, aynı zamanda çok katmanlı algılayıcı olarak da bilinen çok katmanlı bir ağ yapısıdır [27]. Çok katmanlı algılayıcılar, giriş ve çıkış düğümleri arasında bir veya daha fazla düğüm katmanı bulunan ileri besleme ağlarıdır. Bu ek katmanlar, hem giriş hem de çıkış düğümlerine doğrudan bağlı olmayan gizli birimler veya düğümler içerir.

Çok Katmanlı algılayıcılar, tek katmanlı algılayıcıların sınırlamalarının çoğunun üstesinden gelir, ancak etkili eğitim algoritmaları bulunmadığından genellikle geçmişte kullanılmamışlardır. Bu durum son zamanlarda yeni eğitim algoritmalarının gelişmesiyle birlikte değişmiştir.

Ayrıca, bu algoritmaların tek katmanlı algılayıcılarda olduğu gibi birleştiği kanıtlanamamış olmakla birlikte, birçok sorun için başarılı oldukları kanıtlanmıştır.

Çok katmanlı algılayıcıların yetenekleri, düğümler içinde kullanılan doğrusal olmayanlardan kaynaklanmaktadır. Düğümler doğrusal elemanlar ise, uygun şekilde seçilen ağırlıklara sahip tek bir katman ağı, herhangi birçok katmanlı ağ tarafından gerçekleştirilen bu hesaplamaları tam olarak çoğaltabilir [25]. Şekil 1’de çok katmanlı algılayıcı gösterilmiştir.



Şekil 1. Multi-Layer Perceptron

Konvolüsyonel sinir ağı, nesne sınıflandırma görevlerini gerçekleştiren sentetik görüntü sistemleri için olarak birçok avantaja sahiptir.

Öncelikle, konvolüsyonlar gerçekleştirerek yerel alıcı alanlarla çalışırlar: konvolüsyon matrislerinde ağırlıkları paylaşırlar, böylece büyük görüntüler daha düşük ağırlıklarla işlenebilir. Ağdaki ağırlıkların sayısı, giriş görüntüsü ile orantılı olmadığı için bu önemlidir (yani, son işlem ağ boyutu belirli bir görev için sabitlenir.). İkinci olarak, uzaysal alt örnekleme / havuzlama, doğrusal olmayan hesaplamanın her adımında giriş veri boyutunu hiyerarşik olarak azaltmak için kullanılır. Küçük, yerel bir alıcı alanın kopyalanması, temel özellikleri büyük bir girişten çıkarırken sonucu alt örnekleme, bozulma ve ölçeğin etkisini azaltır. Bu özelliklerin birleştirilmesi, yüksek seviyeli görme sistemlerinin tipik bir özelliği olan çok iyi kayma, ölçek ve bozulma değişmezliği olan üst düzey özellikler üretir [28].

Konvolüsyon sinir ağı esas olarak girdi katmanı, konvolüsyon katmanı, havuz katmanı, tam bağlantı katmanı ve çıktı katmanı içerir.

1. Konvolüsyonel Katman

Konvolüsyon tabakası, yerel bağlantılara ve sert özelliklere sahip ağırlıklara sahip olan Konvolüsyonel sinir ağının çekirdek kısmıdır.

Konvolüsyon tabakasının amacı, girdilerin özellik gösterimlerini öğrenmektir. Katman, birkaç özellik haritasından oluşur. Aynı özellik haritasındaki her bir nöron, önceki katmandaki farklı konumların yerel özelliklerini çıkarmak için kullanılır, ancak tek nöronlar için çıkartılması, eski farklı özellik haritasındaki aynı konumların yerel özellikleridir. Yeni bir özellik elde etmek için, giriş özelliği haritaları ilk önce öğrenilmiş bir çekirdekle çevrilir ve ardından sonuçlar doğrusal olmayan bir etkinleştirme işlevine geçirilir.

2. Havuz Katmanı

Örnekleme işlemi bulanık filtrelemeye eşdeğerdir. Havuzlama katmanı, ikincil özellik çıkarımının etkisine sahiptir, özellik haritalarının boyutlarını azaltabilir ve özellik çıkarımının sağlamlığını artırabilir. Genellikle iki Konvülayonel katman arasına yerleştirilir.

Havuz katmanındaki özellik haritalarının boyutu, çekirdeklerin hareketli adımına göre belirlenir. Tipik havuzlama işlemleri ortalama havuzlama ve maksimum havuzlamadır. Birkaç konvolüsyonlu tabaka ve havuzlama tabakasını istifleyerek girdilerin yüksek seviyeli özelliklerini çıkarabiliriz.

3. Tamamen Bağlı Katman

Genel olarak, Konvolüsyonel sinir ağının sınıflandırıcısı bir veya daha fazla tamamen bağlı katmandır. Önceki katmandaki tüm nöronları alırlar ve onları mevcut katmandaki her bir nörona bağlarlar. Tam olarak bağlı katmanlarda korunan hiçbir uzamsal bilgi yoktur. En son tam olarak bağlı olan katmanı bir çıktı katmanı izler. Sınıflandırma görevleri için, softmax regresyonu, çıktıların iyi performans gösterilmiş bir olasılık dağılımını ürettiği için yaygın olarak kullanılır. Yaygın olarak kullanılan bir başka yöntem, farklı sınıflandırma görevlerini çözmek için CNN'lerle birleştirilebilen SVM'dir [29].

Hızlı eğitim ile özellik çıkarımı ve haritalama yapabilir ve yüksek tahmin doğruluğu vardır, bu nedenle genellikle sınıflandırma ve öngörmeye uygulanır [27].

Konvolüsyonel sinir ağları (CNN) derin öğrenme yöntemidir, son zamanlarda görüntü bölüşümü, nesne algılama, tanıma ve resim yazısı gibi bilgisayar görme alanında ileri ve canlı bir gelişme kaydetmiştir. Konvolüsyonel ağlar ve derin öğrenme kavramları altındaki diğer ilgili mimariler, insan beynindeki sinir ağlarıyla en iyi şekilde karşılaştırılabilir. Tıpkı hiyerarşik çok katmanlı ağların konumlandırılması yoluyla biyolojik nöronların yapısı ve işlemleri gibi, derin sinir ağının da özelliklerin çeşitli örneklem şemalarını öğrenmede çok etkili olduğunu ortaya koyması gerekir .

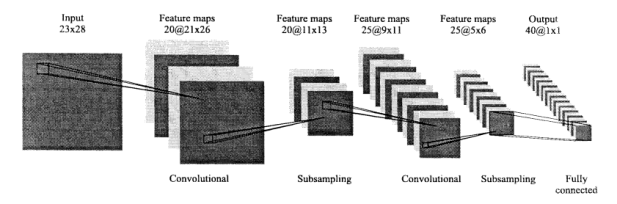
CNN, öğrenme sürecini sürdürmek için geri yayılım algoritmasını kullanır. Bununla birlikte, ağırlık filtresini ve bağlanma katsayısını güncellemek için stokastik gradyan kullanır. Bunların yanı sıra, optimize edilmiş özelliklerin tanınması için havuzlama (alt örnekleme) ve evrişimsel işlemler gibi doğrusal olmayan yöntemler kullanılır [7].

Konvolüsyonel sinir ağları resimlerdeki nesneleri, yüzleri ve sahneleri tanımak için yararlıdır. Doğrudan görüntü verilerinden ders alır, görüntüleri sınıflandırmak için desenleri kullanır ve manuel özellik çıkarımı ihtiyacını ortadan kaldırır. Bu model çok katmanlı bir yapıya sahiptir, her katman birden fazla iki boyutlu düzlemden ve her düzlem de birden fazla nörondan oluşur [9]. Bu katmanlar çoğunlukla girdi katmanları, gizli katmanlar ve çıktı katmanları olmak üzere üç kısımda incelenebilir. Öğrenme için gerekli olan tüm karmaşık işlemler gizli katmanlarda gerçekleşirken, verilerin sisteme girişi girdi katmanından, çıkışı ise çıktı katmanından elde edilir.

Birden fazla katmandan oluşan bu ağ, nöronların öğrenme eylemini paralel olarak gerçekleştirmelerini sağlar. Ek olarak, klasik makine öğreniminde cevap sadece 1 veya 0'dır ve bu ağ yapısını kullanan çalışmaların çıktısında, 0,2 ve 0,7 gibi 0 ile 1 arasındaki değerler elde edilebilir. Bu, problemin daha detaylı bir şekilde çözülmesini kolaylaştırır, öğrenmedeki başarısını arttırır ve daha iyi sonuçlar sağlar. Birçok katmandaki özellikleri öğrendikten sonra, konvolüsyonel sinir ağının bir sonraki kısmı sınıflandırmadır. Son katmanın yanında, x boyutunun bir vektörünü çıkaran, x'in ağın tahmin edebileceği sınıfların sayısı olduğu, tamamen bağlı bir katman bulunur. Bu vektör sınıflandırılan herhangi bir görüntünün her sınıfı için olasılıkları içerir. Konvolüsyonel sinir ağı mimarisinin son katmanı, sınıflandırma çıktısını sağlamak için bir sınıflandırma katmanı kullanır [4].

Konvolüsyonel Ağlar (CN), kısıtlamaları birleştirir ve üç fikir kullanarak bir dereceye kadar değişim ve deformasyon değişmezliği elde eder: yerel alıcı alanlar, paylaşılan ağırlıklar ve mekansal alt örneklemedir. Paylaşılan ağırlıkların kullanılması, sistemdeki genelleştirmeye yardımcı olacak parametre sayısını da azaltır. Değişken ağlar karakter tanıma işlemine başarıyla uygulanmıştır.

Tipik bir konvolüyonel ağı Şekil 2'de gösterilmiştir. Ağ, her biri bir veya daha fazla düzlem içeren bir dizi katmandan oluşur. Yaklaşık ortalanmış ve normalleştirilmiş görüntüler giriş katmanına girer. Bir düzlemdeki her birim, önceki katmanın düzlemlerindeki komşularından küçük bir girdi alır. Her katmanda genellikle birden çok düzlem kullanılır, böylece birden çok özellik algılanabilir. Bu katmanlara konvolüsyonel katmanlar denir [30].



Şekil 2. Tipik Konvolüsyonel Ağ

Görüntü işleme çok zorlu bir araştırma alanıdır ve günlük hayatımızın önemli bir parçası olmuştur. Konvolüsyonel sinir ağları (CNN), görüntü analizinin zorlu problemlerini çözmek için standart bir teknik olarak kabul edilir.

Bilgisayarla görme ve örüntü tanıma görevlerinde geleneksel yöntemlerden önemli ölçüde daha iyi performans gösterir. CNN yapay zekanın bir şeklidir ve makine öğrenme yöntemlerine aittir.

Bu ağlar günümüzde, derin mimarilerine göre sağlamlık ve genelleme yeteneklerinden dolayı görüntü işlemenin en zor görevleri için kullanılmaktadır .

Optimal CNN tasarımı ve eğitimi çok önemli bir iştir. Ağların çok sayıda olası mimari varyasyonları ve konfigürasyonları vardır ve mümkün olan her konfigürasyonu denemek hesaplamalı olarak mümkün değildir. Bir diğer önemli sorun, CNN'nin eğitim için büyük miktarda veri gerektirmesidir , kategori başına yaklaşık bin kişi önerilir . Bu çok yaygın bir sorundur ve veri eksikliği temelde fazladan donmaya ya da sinir ağının birleşmeye başlayamayacağı durumlara yol açmaktadır. Bazen, yapay olarak veri setini genişleten, veri büyütme adı verilen yaklaşıma yardımcı olabilir .

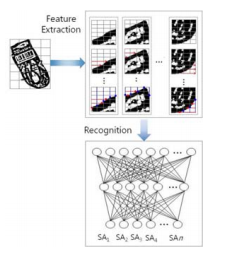
CNN, eğitim sırasında temsili özellikler aramaktadır. Sınıflandırma problemini göz önünde bulundurmak için görüntüde hangi nesnenin olduğunu öğrenmesi gerekir, böylece verilen nesneleri temsil eden ortak özellikleri araştırır . 2D konvolüsyon, göreceli olarak daha küçük konvolüsyonel çekirdeğe göre tüm görüntüyü işler; burada daha fazla katman daha karmaşık görüntü yapılarını işler ve daha yüksek soyutlama ile bilgi elde eder. Ancak, çekirdekler görüntünün belirli bir nesnenin değerli bilgisini içermeyen bölümlerini bile işlemektedir. Nesnenin görüntünün büyük bölümünü kapladığı sınıflandırma görevlerinde, eğitim her zaman çok zor bir problem değildir. Bu, büyük veri kümeleri olduğunda da mümkündür. Sorun, sınırlı miktarda eğitim verisi olduğunda ortaya çıkar, sınıflandırma için kullanılabilecek bilgiler görüntünün önemli bir bölümünü kapsamaz ve veriler eğitim sürecini zorlaştıran gereksiz bilgiler içerir. Sinyal / gürültü oranı birçok yöntem ve uygulamada genel olarak önemlidir. Görüntünün arka planı, CNN'nin eğitimi için gürültü olarak düşünülebilir [31].

B. CNN ile Nesne Algılama ve Nesne Tanıma

Gerçek dünyadaki bir uygulama alanı için nesne tanıma teknolojisi yıllardır aktif olarak gelişim göstermektedir. Görme tabanlı nesne tanıma, 2D veya 3D görme sensörleri kullanarak nesne şekli, kenar veya dönüm noktası özellikleri çıkarılarak işlenmektedir. Nesnenin uç bilgisi, nesneyi tanımak için en önemli özellikleri dikkate almıştır.

Kenar tespiti, esas olarak, bileşenlerin ürün hattındaki hatasını kontrol etmek için kullanılmıştır.

Nesneyi incelemek ve üretim sürecinde çeşitli nesneler tanımak için daha doğru kenar algılaması gerekmekteydi. Aydınlatma sistemi, nesnenin tanınmasının güvenilirliğini artırmak için dikkatlice seçilmektedir. Bununla birlikte, yüksek maliyet efekti ve düşük verimliliğe sahiptir çünkü optimize edilmiş görüş sensörü ve aydınlatma kurulumunu seçmek zaman alıcıdır. Parçalı nesneyi tanımak için geometrik ve istatistiksel özellikler çıkarılmıştır. Kafes, geçiş ve mesafe özellikleri, sırasıyla bölünmüş nesne görüntüsünün m × n alt pencerelerinde elde edilir. Sınıf, her nesnenin pozuna ve tarafına göre ayrıntılı türlere ayrılmıştır. Üç katmanlı algılayıcı, nesneyi sınıflandırmak için uygulanmıştır ve iyi bilinen bir değiştirilmiş geri yayılma algoritması ile eğitilmiştir [32]. Şekil 3'te gösterilmiştir.

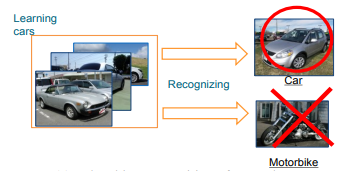


Şekil 3. Özellik Çıkarma ve Tanıma İşlemi

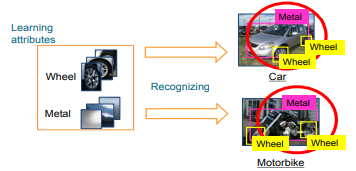
Son yıllarda, nesnelerin tanınması gibi işlemler gerekli olmuştur. Ancak, genel nesne tanıma yöntemleri çok yönlü değildir, çünkü bu yöntemler sadece öğrenilmiş nesneyi tanır. Örneğin, bu genel nesne tanıma yöntemlerinin bir araba öğrenmesi durumunda, aracı tanıyabilir ancak bir motosikleti tanıyamaz. Aksine, özelliklerine dayanan bir nesne tanıma yöntemi çok yönlü bir nesne tanıma yöntemi olarak önerilmiştir. Örneğin, bir otomobilin ve bir motosikletin tanınmasında, nitelik bazlı nesne tanıma yöntemi otomobili dört teker ve metal gövdeye sahip bir nesne olarak tanımlar ve motosikleti iki teker ve metal gövdeye sahip bir nesne olarak tanımlar. Daha sonra, niteliklere dayanan nesne tanıma yöntemi, otomobil ve motosikleti ayrı ayrı öğrenmeden teker teker ve metal bir malzeme öğrenerek, her iki nesnenin de aynı özelliklere sahip olması nedeniyle tanıyabilir. (Şekil 4 (b)) Ancak, nitelikleri kullanan geleneksel bir metodun iki sorunu vardır.

İlk olarak, geleneksel yöntem, niteliklerin bir bölümünün sorunlu olduğu bir hedef nesneyi tanıyamaz. Örneğin, bir yüz tanımada, geleneksel yöntemin bir yüzü, iki göze sahip bir nesne olarak tarif etmesi durumunda, bir burun ve bir ağız özelliklerinin bir sahnede sorun oluşturması durumunda, geleneksel yöntem bir yüzü tanıyamamaktadır; çünkü özniteliklerin bir kısmı eksiktir. (Şekil 4 (a)) İkincisi, geleneksel yöntem alakasız özelliklere sahip bir hedef nesneyi yanlış tanır.

Örneğin, bir otomobil tanımada, geleneksel yöntemin tekerlekleri, metal ve pencereleri olan bir nesneyi tanımlaması durumunda, geleneksel yöntem bir motosikletin metallerini ve bir evin camlarını bir otomobil olarak yanlış tanımaktadır. (Şekil 4 (b)). Bu nedenle, bu iki problemi çözmek için, nitelikler arasındaki ilişkiye dayanarak nesne tanıma önerilmektedir. İlk olarak önerilen yöntem, hedeflerin bir kısmının kapatıldığı bir hedef nesneyi tanıyabilir, çünkü önerilen yöntem, hedef nesnenin diğer görünür özellikleri arasındaki ilişkiyi yerel bir kısıtlama olarak tanımlar. İkinci olarak önerilen yöntem, ilgisiz niteliklere sahip olan bir hedef nesneyi hariç tutabilir çünkü önerilen yöntem, tüm özellikleri ilişkiyi global bir kısıtlama olarak kullanarak ilgili ilişkiyi bir nesne olarak değerlendirir [33].



(a) Genel Yöntem ile Nesne Tanıma



(b) Niteliklere Göre Nesne Tanıma

**Şekil 4. Nesnelerin nitelikler tarafından tanınmasının avantajı**

Otomatik nesne tanıma için iki durum göz önünde bulundurulmaktadır. Birincisi, bir nesnenin adı belirtildiğinde, sistem onu tanımaya çalışmaktadır. İkincisi ise bir kullanıcı bir alanda hangi nesnelerin olduğunu sorduğunda, sistem tüm nesneleri tanımaya çalışır [34]. Her iki durumda da, sistem aşağıdaki adımları izleyerek nesneleri tanır:

1.Nesneler için aday bölgeler çıkartılmalıdır.

2.Nesne türleri doğrulanmalı veya belirlenmelidir (kutu, şişe, PET şişe).

3.Çıkarılan bölgeler, nesne modelleriyle eşleştirilmelidir.

Nesne algılama, bilgisayarlı görüşte yaygın bir görevdir ve görüntü verilerinde belirli özelliklerin varlığının veya yokluğunun belirlenmesi anlamına gelir. Özellikler algılandıktan sonra, bir nesne önceden tanımlanmış bir sınıf kümesinden birine ait olarak da sınıflandırılabilir. Bu son işlem, nesne sınıflandırması olarak bilinir. Nesne tespiti ve sınıflandırma yapay zekanın temel yapı taşlarıdır. Yapay zeka ve makine öğreniminin görüntü analizine entegre edilmesinin önündeki en büyük zorluk, bu görevlerin gerçek zamanlı veya gerçek zamanlıya yakın gerçek zamanlı olarak uygulanabilir olmamasıdır [9].

Nesne algılama, otomatik sürüş ve güvenlik sisteminde yaygın olarak kullanılmaktadır [35]. Nesne algılama, bilgisayarların dünyayı anlamalarını ve tepki vermelerini sağlayan bilgisayar vizyonunda önemli bir iştir ve otomatik sürüş gibi yeni ortaya çıkan uygulamalar için büyük potansiyele sahiptir [36]. Son yıllarda, ön plan belirleme ve izleme için farklı araştırmacılar tarafından çeşitli yaklaşımlar önerilmiş ve gösterilmiştir. Ancak bu yaklaşımlar, radikal değişiklikler ve izleme sırasındaki hedef sapması gibi sorunları çözmede başarısız olmuştur. Hareketli nesne algılama ve izlemenin ana zorluğu, nesne pozisyonunu daha doğru tahmin etmektir. Hareketli nesne algılama, bilgisayarla görüş algoritmasının önemli bir adımıdır. Video analizi, tıbbi görüntüleme ve askeri uygulama gibi farklı uygulamalarda kullanılır.

Genel olarak, gözetleme sistemi hareketli nesne algılama ve hareket tahmini gibi iki ana adımdan oluşur. İlk adım, nesne tespitidir ve arka plandaki piksel bilgilerinden etkilenir. Bazı araştırmacılar, nesneleri bir video dizisinden tespit etmeye odaklanan birçok metodolojiyi önermiştir. Önerilenlerden biri olan arka plan çıkarma, ilgili hareketli nesneyi video karelerinden çeken yöntemdir .

Arka plan çıkarma, çoğunlukla durağan olmayan arka plan ve aydınlatma değişikliklerinden etkilenir. Uygulamada, bu dezavantaj optik akış algoritmasıyla ortadan kaldırılabilir, ancak karmaşık koşullar altında izleme algoritmaları için yanlış bilgi üretmektedir. Arka

plan çıkarma durumlarının çoğunda, nesne izleyicileri arka plan bilgisinden etkilenir, ancak yanlış sınıflandırmaya neden olur. Dahası, sağlam sınıflandırıcı seçimi, algoritmanın doğruluğunu arttırmak için önemlidir.

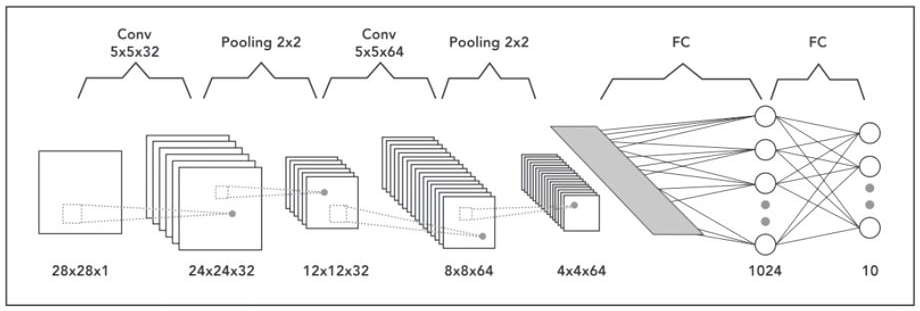
Derin öğrenmenin gelişmesiyle birlikte, CNN nesne tespitinde kullanılmıştır ve geleneksel yöntemlerden daha iyi performans gösterir.

Sınırlamaların üstesinden gelmek için, CNN tabanlı nesne algılama algoritması kullanılmaktadır. Bu yaklaşım, karmaşık sahnelerde ve karmaşık arka plan koşullarında sağlam bir şekilde algılanır ve izlenir.

CNN'ler, görüntülerin güçlü yerel iki boyutlu yapılara sahip olması gerçeğine dayanarak, gizli birimlerin alıcı alanlarını yerel olarak sınırlayarak yerel özelliklerin çıkarılmasını gerektirir [37].

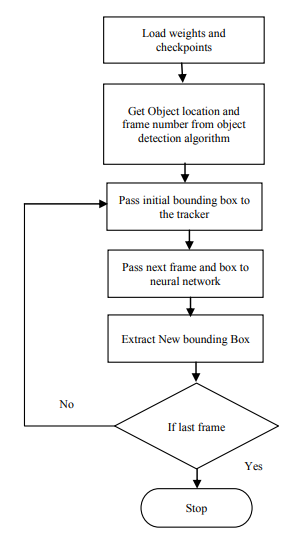
Özelliklerin manuel olarak seçildiği ve belirtilen algoritmalar yoluyla çıkarıldığı geleneksel makine öğrenme yöntemlerinin aksine, CNN veriden giderek daha yüksek seviye özellikleri otomatik olarak keşfeder [36]. Bununla birlikte, günümüzde çoğu nesne algılama algoritması, gerçek zamanlı algılamayı sağlamak için güçlü GPU'lara ihtiyaç duyar. Otomatik sürüş sistemi gibi birçok sistemde güçlü GPU'ları desteklemek mümkün değildir [35].

Konvolüsyonel sinir ağının önemli bir yönü, tüm parametrelerinin modellenecek verilerden öğrenilebilmesidir. Örneğin; Şekil 4'de sunulan ağda, filtre veri tabanlarından, havuzlama fonksiyonlarından ve aynı zamanda sınıflandırıcıdan gelen tüm ağırlıklar, etiketli veri setindeki stokastik gradyan inişi kullanılarak öğrenilir. El sıkışması olmayan özellik çıkarıcılarla karşılaştırıldığında, Konvolüsyonel sinir ağı daha kompakt ve genel amaçlı tanıma görevlerine uygundur [28].



Şekil 5.Nesne Tanıma İçin Tipik Bir Konvolüsyonel Ağın Mimarisi

CNN tabanlı nesne izleme algoritmasının Akışı, Şekil 6'te gösterildiği gibidir [37].



Şekil 6.Nesne Tespiti Akış Şeması

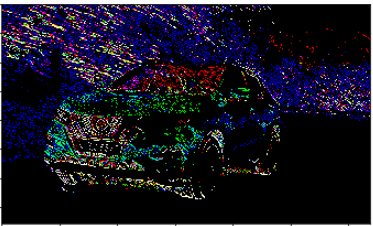
Akış şemasında ilk olarak ağırlıklar ve kontrol noktaları belirlenmektedir. Bu adım gerçekleştirildikten sonra nesnenin konumu ve nesne algılama algoritmasından çerçeve numarası alınır. Daha sonra kullanıcıya başlangıç durumu için sınırlayıcı kutu bilgileri verilir. Sonraki çerçeve ve sınırlayıcı kutu bilgileri sinir ağına verilmektedir. Sinir ağına iletim gerçekleştirildikten sonra yeni sınırlama kutusu çıkartılır. Eğer son kareye gelindi ise şema sonlandırılır, gelinmedi ise tekrar kullanıcıya başlangıç durumu için sınırlayıcı kutu bilgileri verilir.

IV.DENEYSEL SONUÇLAR

Projede derin öğrenme yöntemlerinden biri olan konvolüsyonel sinir ağı ile görüntü tanıma işlemi gerçekleştirilmektedir.

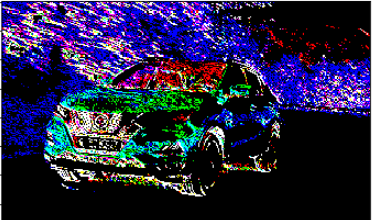
Konvolüsyonel sinir ağı temel olarak girdi katmanı, konvolüsyon katmanı, havuz katmanı, tam bağlantı katmanı ve çıktı katmanından oluşmaktadır.

Öncelikle Şekil 7’deki örnekteki gibi konvolüsyon katmanında girdi olarak alınan görüntülerin özelliklerini ön plana çıkarmak için çeşitli filtreler kullanılmaktadır. Bunlar; edgeDetection, grayScaleEdgeDetection,cornerDetection, diagonalLineDetection,inversion, optimalEdgeDetection filtreleridir. Projede Konvolüsyon işleminden sonra ReLU(Rectified Linear Unit) Fonksiyonu çalıştırılmaktadır. ReLU fonksiyonu yapay sinir ağlarında kullanılan bir aktivasyon fonksiyonudur. Diğer aktivasyon fonksiyonları çok işlem gerektirmektedir. ReLU fonksiyonu ise ağdaki bazı nöronların aktif olup, negatif eksende 0 değerlerini alması ağın daha hızlı çalışacağı anlamına gelir.



Şekil 7. Konvolüsyon İşleminden Geçirilen Resim

ReLU fonksiyonundan sonra Şekil 8’de yer alan örnekteki gibi max pooling işlemi gerçekleştirilmektedir. Bu katmanda konvolüsyon işlemi ile elde edilen matris üzerinden belirlenen filtre boyutunda maske geçirilerek en büyük değerleri bulunmaktadır. Bu değerler yeni oluşturulan matrise atanmaktadır.



Şekil 8.Havuzlama İşleminden Geçirilen Resim

Bu katmanlarla birlikte CNN işlemini gerçekleştirmiş olduk. Sonrasında elde edilen görüntüleri tanımladığımız yapay sinir ağına girdi olarak verdik.

Burada rastgele ağırlıklar tanımladık. Bu ağırlık değerleri kullanılarak girdiler ile bir çıktı üretilmiştir.

Yapay sinir ağından elde edilen çıktı değeri beklenen değer ile karşılaştırılarak hata değeri bulunur.

Hata değeri backpropagation algoritması kullanılarak azaltılmaya çalışılır. Backpropagation, Çok Katmanlı Perceptron (Yapay Sinir Ağları) eğitimi için denetimli bir öğrenme algoritmasıdır. Temel olarak hata miktarlarının gizli katmanda(Hidden Layer) bulunan sinir hücrelerine(neurons) geri dönmesi ve eğitimin başarısının arttırılmasıdır.

Hesaplanan hata değeri kullanılarak ağırlık parametreleri güncellenir. Hata değeri tekrar kontrol edilerek, minimum hataya ulaşıncaya kadar iterasyonlar devam eder.

Resim buraya gelecek

**Şekil 9.Hata Değeri**

V. SONUÇ

Bu makalede, derin öğrenme yöntemleri ile nesne tanıma yaklaşımı sunulmuştur. Konvolüsyonel sinir ağı kullanarak nesne algılama ve izleme için araştırmalar gerçekleştirilmiştir. Nesne tanıma modülü nesneyi sağlam bir şekilde algılar. Algılanan nesne, CNN algoritması kullanılarak izlenir.

KAYNAKÇA

[1] S. H. Tajmir, T. K. Alkasab, “Toward augmented radiologists: changes in radiology education in the era of machine learning and artificial intelligence,” Academic radiology, 2018, vol. 25, pp.747- 750.

[2] J. M. Valin, “A hybrid dsp/deep learning approach to real-time fullband speech enhancement,” 2017.

[3] Y. Zhou, O. Tuzel, “Voxelnet: End-to-end learning for point cloud based 3d object detection,” 2017.

[4] Aylin Sevik, Pakize Erdogmus, Erdi Yalcin,” Font and Turkish Letter Recognition in Images with Deep Learning”,2018.

[5] R. Wason, “Deep learning: Evolution and expansion,” Cognitive System Research, 2018, vol. 52, pp.701-708.

[6] E. Bati, “Deep convolutional neural networks with an application towards geospatial object recognition,” Diss. Middle East Technical University Ankara, 2014.

[7] Shaukat Hayat, She Kun, Zuo Tengtao, Yue Yu, Tianyi Tu, Yantong Du,” A Deep Learning Framework Using Convolutional Neural Network for Multi-class Object Recognition”,2018.

[8] Li Chuyi, Yu Weiyu,Hu yao, Hu Dan,” Gabor Feature based Convolutional Neural Network for Object Recognition in Natural Scene”,2016.

[9] V. V. Khryashchev , V. A. Pavlov , A. A. Ostrovskaya , A. S. Semenov, “Optimization Of Convolutional Neural Network For Object Recognition On Satellite Images”,2018.

[10] Saad ALBAWI , Tareq Abed MOHAMMED, Saad AL-ZAWI,” Understanding of a Convolutional Neural Network” 2017.

[11] D. H. Hubel and T. N. Wiesel. Receptive fields of single neurones in the cat’s striate cortex. The Journal of physiology, 148(3):574, 1959.

[12] D. H. Hubel and T. N. Wiesel. Receptive fields, binocular interaction and functional architecture in the cat’s visual cortex. The Journal of Physiology, 160(1):106, 1962.

[13] K. Fukushima. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. Biological cybernetics, 36(4):193–202, 1980.

[14] Y. LeCun, B. Boser, J. Denker, D. Henderson, R. Howard, W. Hubbard, and L. Jackel. Handwritten digit recognition with a back-propagation network. In Advances in Neural Information Processing Systems (NIPS), pages 396–404, 1990.

[15] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel. Backpropagation applied to handwritten zip code recognition. Neural Computation, 1(4):541–551, 1989.

[16] X. Glorot, A. Bordes, and Y. Bengio. Deep sparse rectifier networks. In Proceedings of the 14th International Conference on Artificial Intelligence and Statistics, volume 15, pages 315–323, 2011.

[17] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov. Dropout: A simple way to prevent neural networks from overfitting. Journal of Machine Learning Research, 15:1929–1958, 2014.

[18] I. J. Goodfellow, D. Warde-Farley, M. Mirza, A. C. Courville, and Y. Bengio. Maxout networks. In Proceedings of the 30th International Conference on Machine Learning (ICML), pages 1319–1327, 2013.

[19] M. Lin, Q. Chen, and S. Yan. Network in network. In International Conference on Learning Representations (ICLR), 2014.

[20] F. Rosenblatt. Principles of neurodynamics; perceptrons and the theory of brain mechanisms. Spartan Books Washington, 1962.

[21] Ming Liang Xiaolin Hu,” Recurrent Convolutional Neural Network for Object Recognition”,2015.

[22] Tomohiro Ishii1, Ryosuke Nakamura, Hidemoto Nakada, Yoshihiko Mochizuki, and Hiroshi Ishikawa,” Surface Object Recognition with CNN and SVM in Landsat 8 Images”,2015.

[23]URL- <https://www.nature.com/articles/nature14539>

[24] Haiyan Shu, Ying Song and Huan Zhou ,”Time-frequency Performance Study on Urban Sound Classification with Convolutional Neural Network “,2018.

[25] Richard P. Lippmann,” An Introduction' to Computing with Neural Nets”,1987.

[26] Rahul Nijhawan,Himanshu Sharma ,Harshita Sahni ,Ashita Batra,” A Deep Learning Hybrid CNN Framework Approach for Vegetation Cover Mapping using Deep Features”,2017.

[27] Weiguo Shen , Wei Wang ,”Node Identification in Wireless Network Based on Convolutional Neural Network”,2018.

[28] Cl´ement Farabet, Berin Martini, Polina Akselrod, Selcuk Talay, Yann LeCun and Eugenio Culurciello,” Hardware Accelerated Convolutional Neural Networks for Synthetic Vision Systems”,2010.

[29] Tianmei Guo, Jiwen Dong ,Henjian LiˈYunxing Gao,” Simple Convolutional Neural Network on Image Classification”,2017.

[30] Steve Lawrence, Member, C. Lee Giles, Senior Member, Ah Chung Tsoi, Senior Member, and Andrew D. Back, Member,” Face Recognition: A Convolutional Neural-Network Approach”,1997.

[31] Martin Rajnoha Radim Burget Lukas Povoda,” Image Background Noise Impact on Convolutional Neural Network Training”,2018.

[32] Kyekyung Kim , Sangseung Kang , Jaehong Kim , Jaeyeon Lee , Joongbae Kim1 and Jinho Kim,” Multiple Objects Recognition for Industrial Robot Applications”,2013.

[33] Hiroka Horiguchi, Kazuo Ikeshiro ,Hiroki Imamura,” Recognition for objects by relationship between attributes”,2016.

[34] Yasushi Makihara, Masao Takizawa, Yoshiaki Shirai, Jun Miura, and Nobutaka Shimada,” Object Recognition Supported by User Interaction for Service Robots”,2002.

[35] Liuchen Qian,Yuzhuo Fu.Ting Liu, “An Efficient Model Compression Method for CNN Based Object Detection”,2018.

[36] Wendi Cai1, Jiadie Li,hongzhao Xie1, Tao Zhao1, Kang Lu1,” Street Object Detection Based on Faster R-CNN”,2018.

[37] Shraddha Mane, Prof.Supriya Mangale, “Moving object detection and tracking Using Convolutional Neural Networks“,2018.