**Göğüs Kanseri’nin CNN ile Tespiti**

İsa Kulaksız

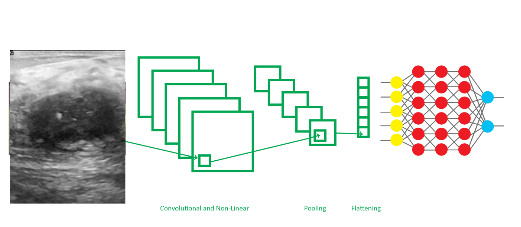
**1** Atatürk Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, Erzurum, Türkiye (ORCID: 0000-0000-0000-0000)

**Öz**

Göğüs kanseri tüm dünyada insanlar için ikinci ölüm nedenidir. Hastalığn sebebi bilinmediği için erken teşhis ve göğü kanserinin kontrolünün anahtarıdır. Erken teşhis, hayat kurtarabilir ve özellikle her geçen gün ilerleyen teknolojinin yardımıyla (Ultrason) maiyeti azaltılabilir hale gelmiştir. Ultrsaon görüntüleme, göğüs anormalliklerini tespit etmek ve sınıflandırmak için kullanılan tanı araçlarından biridir. Yazılım dünyasında var olan bu problemin çözümünü, operatör bağımlılığını ortadan kaldırıp teşhis doğruluğunu artırmak için bir karar destek sistemi (DSS) tasarlandı. Genel olarak, bir karar destek sistemi (DSS) Derin öğrenme’de CNN yapısı üzerinden bir model oluşturmaktadır.

1. Giriş

Üzerinde işlem yapılacak olan veri seti kötü huylu (malignant), iyi hutlu (benign) ve normal verilerden oluşmaktadır.



Kullanılacak öğrenme tipinin (Derin Öğrenme) Makine öğrenimi arasındaki temel fark, verinin sisteme nasıl sunulduğuna dayanır. Makine öğrenimi algoritması, çoğunlukla her zaman yapılandırılmış veriye ihtiyaç duyarken, Derin öğrenme ağları yapay sinir ağlarının katmanlarına dayanır.[1]

Derin Öğrenme algoritmaları, veri setine göre değişiklik gösterebilir. Metin üzerinde işlem yapmak için RNN’i (Yinelemeli Sinir Ağı) kullanmak mantıklıyken görüntü işlemede CNN (Evrişimli Sinir Ağları) kullanılır. Veri setinde hastaların göğüs röntgenleri olduğu için CNN modeli tercih edilmiştir. Derin derken kastedilen Derin öğrenmenin birtakım derin bilgiler elde etmesi değil birbirini takip eden gösterim katmanları ile ifade edilmesidir.

Bir CNN modeli, temelde yapay sinir ağlarına benzer şekilde neuron’lardan oluşur ve öğrenmek için ağırlıklar ile bias değerine sahiptir. Her neuron girişleri alır, birleştirir ve çoğunlukla non-linear bir fonksiyon (sigmoid vb.) ile çıkış üretir.[2]

CNN 3B tensörler üzerinde işlem yapar ve bu tensörler nitelik haritası adını alır. Genişlik, yükseklik ve derinlik vardır. Burada yer alan derinlik herhangi bir şey olabilir. Çünkü derinlik evrişim katmanının bir parametresidir ve derinlik eksenindeki farklı kanallar artık RGB gibi resmin kanallarını değil filtre olarak tanımlanan yapıları oluşturur.

2. Materyal ve Metot

**2.1. Veri Ön İşleme**

Sklearn kütüphanesi yardımıyla veri seti test ve eğitim olarak iki sınıfa ayrıldı. Bölme işlemi gerçekleştikten sonra eğitim ve test boyutları Tablo 1’de gösterilmiştir.

|  |  |
| --- | --- |
| X | Boyut (Shape) |
| X\_train | 1183, 128, 128, 1 |
| X\_test | 395, 128, 128, 1 |

*Tablo 1 X*

***2.1.1. Model’in Hazırlanması***

Diagram

Description automatically generated

Kayıp fonksiyonu, hedefin gerçek ve tahmin edilen değeri arasındaki farkı ölçer. Hesaplanan değer üzerinden modelin istenen değere ne kadar yaklaştığı sonucuna ulaşılılabilir. Oluşturulan modelde categorical\_crossentropy fonksiyonu kullanılmıştır. Categorical\_crossentropy etiketler ve tahminler arasındaki çapraz entropi kaybını hesaplar.

En iyileci için optimizasyon algoritmalarından Adam optimizer kullanıldı. Adam optimizer, Rmsprop ve momentum yöntemlerinin avantajlı yönlerinin birleştirilmesi ile önerilen gradient descent algoritmasıdır.[3]

*Tablo 2. Çalışmada Kullanılan Parametreler*

|  |  |
| --- | --- |
| Parametre Optimizasyonu | DEĞER |
| Adam Optimizasyon Algoritması (Adam Optimizer(lr=0.001)) | 0.001 |
| Kayıp (Loss) | Categorical\_crossentropy |
| Aktivasyon Fonk | RELU |
| EPOCHS | 20 |

3. Araştırma Sonuçları ve Tartışma

**3.1. Araştırma Sonuçları**

Tablo 2’de yer alan parametreler kullanılarak model eğitildikten sonra matplotlib kütüphanesi kullanılarak doğruluk (accuracy) değeri Görsel 2’de yer almaktadır.

Diagram, table

Description automatically generatedA picture containing text

Description automatically generated

*Confusion Matrix*

Doğruluk (accuracy) bir modelin başarısını ölçmek için kullanılan önemli bir metriktir.[4]

Chart, line chart

Description automatically generated

*Görsel 2 Eğitim ve Test Doğruluk Değeri*

*Chart, line chart

Description automatically generated*

*Görsel 3 Eğitim ve Test Kayıp Değeri*

Model üzerinde test verisi tahmin edildiğinde doğruluk (accuracy) değeri 0,80 olarak hesaplandı. Gerçek ve tahmin edilen değerler “seaborn” kütüphanesi yardımıyla Correlation Matrix Görsel 4’ de yer almaktadır.

Chart, treemap chart

Description automatically generated

Görsel 4 Correlation Matrix

Elde edilen sonuçlar, göğüs üzerinde CNN ile kitle tespitinin %80 oranda başarılı olduğunu kanıtladı. Bu yöntem sayesinde hastaneler ve doktorlar üzerinde iş yükü miktarı ve maliyetler azaltılabilir hale gelmiş oldu. Yapılan bu araştırmayla ortaya konan Karar Destek Sistemi (DSS) sayesinde tüm bu işlemler otomatize edilmiş oldu.

4. Tartışma

Daha önceki benzer çalışmalarda kötü huylu meme tümörlerinin çevredeki normal dokulardan daha düşük elektriksel empedansa sahip olduğuna dair bazı kanıtlar saptanmıştır.[5]

Proje üzerinde en iyileyici algoritması olarak Adam optimizer yerine SDG kullanılması başarı oranında değişikliklere yol açabilir.

CNN ile tespit edilen kitlelerin iyi huylu olanları, kanserli hücreler olmasa bile kötü huylu olma potansiyelleri vardır. Bu yüzden oluşturulan karar destek sistemi (DSS) ile hastaya doktoruyla iletişime geçmesi için mesaj verilmelidir. Kötü huylu tümörler ise kanserli hücre olabilir ve bu hücreler büyüyebilir, vücudun diğer bölgelerine yayılabilir.

Karar Destek Sistemi (DSS) nin kullanıcının daha rahat kullanılabilmesi için Web üzerinde proje çalıştırılabilir.

Kaynakça

[1] B. I. Academy, “DEEP LEARNING & MACHINE LEARNING,” *Bilginç IT Academy*. https://bilginc.com/tr/blog/5427/deep-learning-and-machine-learning (accessed May 28, 2022).

[2] M. A. Akcayol and G. Üniversitesi, “Derin Öğrenme Deep Learning,” p. 17.

[3] E. Seyyarer, F. Ayata, T. Uçkan, and A. Karci, “Derin Öğrenmede Kullanilan Optimizasyon Algoritmalarinin Uygulanmasi Ve Kiyaslanmasi,” p. 9.

[4] G. Öğündür, “Doğruluk (Accuracy) , Kesinlik(Precision) , Duyarlılık(Recall) ya da F1 Score ?,” *Medium*, Apr. 22, 2020. https://medium.com/@gulcanogundur/do%C4%9Fruluk-accuracy-kesinlik-precision-duyarl%C4%B1l%C4%B1k-recall-ya-da-f1-score-300c925feb38 (accessed May 28, 2022).

[5] Y. Zou and Z. Guo, “A review of electrical impedance techniques for breast cancer detection,” *Med. Eng. Phys.*, vol. 25, no. 2, pp. 79–90, Mar. 2003, doi: 10.1016/S1350-4533(02)00194-7.