

**HİSTOPATOLOJİK GÖRÜNTÜLERDEN YUMURTALIK KANSERİ SINIFLANDIRMASI
İÇİN VGG-19 TABANLI ÇOKLU SINIF MODELİ**
**VGG-19 BASED MULTICLASS MODEL FOR OVARIAN CANCER CLASSIFICATION
FROM HISTOPATHOLOGIC IMAGES**

Hazal PARLAK

*Bilecik Şeyh Edebali Üniversitesi, Lisansüstü Eğitim Enstitüsü,
Bilgisayar Mühendisliği ABD, Merkez, Bilecik.*

Dr. Öğr. Üyesi Burakhan ÇUBUKÇU

*Bilecik Şeyh Edebali Üniversitesi, Mühendislik Fakültesi,
Bilgisayar Mühendisliği Bölümü, Merkez, Bilecik.*

ÖZET

Bu çalışma da kadınlarda oldukça yaygın görülen yumurtalık kanserini histopatolojik görüntüler üzerinden beş alt türünü tahmin etmek ve sınıflandırmak için VGG-19 modeli önerilmektedir. Yumurtalık kanseri jinekolojik kanser türleri arasında en sık görülen, en agresif ve sağkalım oranı oldukça düşük olan kanser türlerinden biridir. Teşhis süreci genellikle zaman alıcı ve maliyetlidir. Bu sebeple yumurtalık kanserinin erken teşhisi ve alt türlerinin sınıflandırılması tedavi için oldukça önemlidir. Bu çalışmada histopatolojik görüntülerinin otomatik sınıflandırması ve patolojlara ikinci bir görüş sağlaması amacıyla derin öğrenme tabanlı transfer öğrenme yöntemi olan VGG-19 önerilmektedir. Yapılan çalışmada VGG-19 ön-eğitilmiş ağı kullanılarak görüntüler eğitilmiş ve test edilmiştir. Test veri seti için elde edilen sonuçlara göre doğruluk oranı %85, hassasiyet %87 ve duyarlılık %85 olarak belirlenmiştir. Eğitim veri seti için ise doğruluk %87, hassasiyet %88 ve duyarlılık %87 olarak belirlenmiştir. Başarım grafikleri ve karmaşıklık matrisi kullanılarak, modelin eğitim süreci ve performansı detaylı bir şekilde incelenmiştir. Sonuçlar incelendiğinde görüntü sınıflandırmada konusunda başarılı sonuçlar elde eden modelin, patolojlara güçlü bir destek sunma potansiyeli olduğu görülmüştür. Bu noktada, elde edilen sonuçların iyileştirilmesi için veri kümesinin çeşitliliğinin artırılması ve daha fazla örnek eklenmesi modelin genelleme yeteneğini geliştirebileceği düşünülmektedir. Ayrıca modeli farklı veri kümeleri veya kanser türleri üzerinde de genelleme yeteneğini değerlendirmesi gerekmektedir. Bu sayede modelin farklı klinik senaryolarda kullanılabilirliği sağlanabilir.

Anahtar Kelimeler: Yumurtalık Kanseri, Transfer öğrenme, Sınıflandırma, Histopatolojik görüntü, VGG-19

ABSTRACT

In this study, transfer learning techniques are used to predict and classify five subtypes of ovarian cancer, which is very common in women, based on histopathological images. Ovarian cancer is one of the most common, aggressive and low survival rates among gynaecological cancers. The diagnostic process is usually time-consuming and costly. Therefore, early diagnosis of ovarian cancer and classification of its subtypes are very important for treatment. In this study, VGG-19, a deep learning-based transfer learning method, is proposed for automatic classification of histopathology images and to provide a second opinion to pathologists. In this study, images were trained and tested using VGG-19 pre-trained network. According to the results obtained for the test data set, the accuracy rate was 85%, precision 87% and sensitivity 85%. For the training data set, the accuracy was 87%, precision was 88% and sensitivity was 87%. Using the performance graphs and complexity matrix, the training process and performance of the model were analysed in detail. When the results are analysed, it is seen that the model, which has achieved successful results in image classification, has the potential to provide strong support to pathologists. At this point, it is thought that increasing the diversity of the

dataset and adding more samples to improve the results obtained can improve the generalisation ability of the model. It is also necessary to evaluate the generalisation ability of the model on different datasets or cancer types. In this way, the usability of the model in different clinical scenarios can be ensured.

Keywords: Ovarian Cancer, Transfer learning, Classification, Histopathological image, VGG-19

GİRİŞ

Yumurtalık kanseri, dünya genelinde kadınlarda oldukça sık rastlanan sekizinci tür maligniteye örnek gösterilmektedir (Sung vd., 2021). Yetersiz tarama ve menopozun neden olduğu semptomlara benzeyen spesifik olmayan bulgular, kanserin tespit ve teşhis edilmesini oldukça zor kılmaktadır (Menon vd., 2021).

Ultrason görüntülerinden yumurtalık kanserinin teşhisi için KHO-CNN algoritmasının geliştirilmesi amaçlanmıştır. Ultrason görüntülerindeki istenmeyen etkiler ve gürültülerin ortadan kaldırılması amacıyla Birleşik Dönüşüm ve Yerel İkili Desen yöntemiyle özellik çıkarma uygulanmıştır. Uygulanan bu yöntemler %99.18'lik yüksek bir sınıflandırma doğruluğu sunmuştur (Kongara vd., 2024).

Büyük ölçekli doku histopatoloji görüntülerinde segmentasyon, görselleştirme ve sınıflandırma amacıyla derin evrişimli sinir ağı (CNN) aktivasyon özelliklerinin kullanılması önerilmektedir. Çalışmada son gizli katmandaki bireysel nöronların görselleştirilmesi CNN özelliklerini daha iyi anlamak için kullanılmıştır. Yapılan deneylerin sonucuna göre histopatoloji görüntülerinin otomatik analizi için oldukça verimli ve etkili görülmektedir (Xu vd., 2017).

Beyin tümörünün sınıflandırılması amacıyla MRI görüntüleri kullanılarak bir derin öğrenme modeli geliştirilmiştir. Residual Network (ResNet) kullanılarak ve çeşitli veri artırma teknikleri ile başarılı sonuçlar elde edilmiştir. 3064 MRI görüntüsü ile test edilmiş görüntü düzeyinde %99 hasta düzeyinde ise %97 doğruluk göstermiştir (Abdelaziz Ismael vd., 2020).

BreakHis veri seti kullanılarak Inception-V3Net, ShuffleNet ve Resnet 18 gibi derin sinir ağları kullanılarak meme kanseri görüntülerinin benign ve malign olarak ikili sınıflandırılması aynı zamanda çok sınıflı sınıflandırma yapılarak kanserli hücrelerin etkili bir şekilde edilip sınıflandırılması gerçekleştirilmiştir (Aljuaid vd., 2022).

Meme kanseri patolojik görüntüleri için önerilen AlexNet-BC modelinde aşırı uyum problemini hafifletmek amacıyla yeni bir düşük entropi çıkış cezası önerilmiştir. Sonuçlar incelendiğinde önerilen yöntemin mevcut yöntemlerden daha iyi sonuçlar elde ettiği gözlemlenmiştir (Tang vd., 2022).

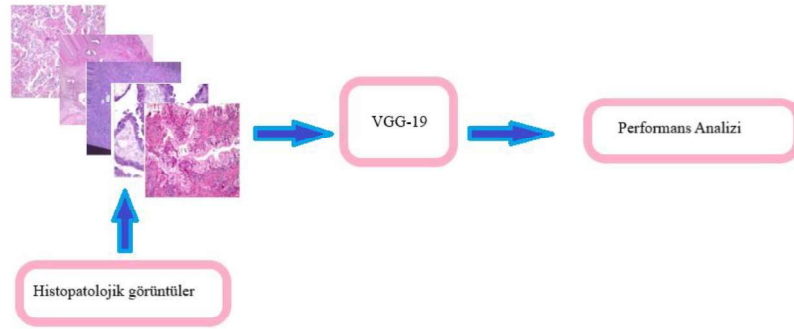
Yapılan başka bir çalışmada histopatolojik görüntüler üzerinden IDC (İnvasive Ductal Carcinoma) tespiti için derin transfer öğrenme tekniği olarak ResNet ve DenseNet kullanılmıştır. Modellerin sadece son katmanları eğitime dahil edilmiştir. DenseNet-161 ile en yüksek denge hassasiyeti %91,57 iken ResNet-50 modelinde en yüksek F puanı %94,11 olarak bulunmuştur (Celik vd., 2020).

Literatürde yapılan çalışmalar incelendiğinde histopatolojik görüntüler üzerinde yumurtalık kanseri ile ilgili yapılmış çalışmaların azlığı gözlemlenmektedir. Bu çalışma, literatüre bu konuda katkıda sunmak amacıyla yumurtalık kanseri alt tiplerinin histopatolojik görüntüler üzerinden tahmin edilmesi ve sınıflandırılmasında VGG-19 mimarisi kullanılarak transfer öğrenme modelinin geliştirilmesi amaçlanmıştır. Özgün bir sınıflandırıcı katman eklenerek yeni bir model oluşturulmuştur. Veri setindeki çeşitlilikte artırılarak eğitim tamamlanmış sonuçlar başarımlar grafikleri ve karmaşıklık matrisi ile analiz edilmiştir.

Çalışmanın devamında materyal ve metotlar bölümünde kullanılan veri seti, model yapılandırılması, eğitim ve test süreçleri ile ilgili detaylar sunularak diğer bölümlerde sırasıyla bulgular ve tartışmalar ve sonuç bölümlerine yer verilmektedir.

MATERYAL VE METOT

Bu bölümde kullanılan ve önerilen transfer öğrenme yöntemleri ve veri kümesinin hazırlanması ele alınmaktadır. Başlangıçta önceden eğitilmiş VGG-19 modelinin seçimini ve yapılandırılması incelenmiştir. Ardından modelin eğitimi için kullanılan veri seti ve eğitim süreci ele alınmıştır. Modelin eğitim sürecinin detayları ve performans değerlendirilmesi, çalışmanın metodolojisi ve temel adımları sunulmuştur. Çalışmada histopatolojik görüntüler kullanılarak yumurtalık kanserinin tespiti için derin transfer öğrenmeye dayalı bir yaklaşım önerilmektedir. Çalışmanın akış şeması Şekil 1'de verilmiştir.



Şekil 1. Yumurtalık Kanseri Sınıflandır Akış Şeması

Çalışmada kullanılan modelin yapılandırılması, öncelikle transfer öğrenme yöntemi kullanılarak önceden eğitilmiş bir VGG-19 modelinin seçilmesiyle başlamıştır. Bu model, geniş bir veri kümesi olan Imagenet üzerinde eğitilmiş ve genel özellikleri öğrenilmiştir. VGG-19 modeli yüklendikten sonra özgün sınıflandırma katmanı çıkarılmış ve temel özelliklerin korunması için tüm katmanlar doldurulmuştur. Çıkarılan sınıflandırma katmanları yerine Global Average Pooling 2D katmanı eklendi ve önceki katmanların çıktıları global bir ortalama değerine dönüştürüldü. Ardından aktivasyon fonksiyonu olarak Relu kullanılan bir tam bağlantılı (dense) katman eklendi. Son olarak model beş sınıf için sınıflandırma yapacağı için çıkış sınıf sayısı beş olan ve softmax aktivasyon fonksiyonunu kullanan bir çıkış katmanı eklendi. Model sınıfı kullanılarak modelin giriş çıkış katmanları belirlenmiş ve bu şekilde yapılandırılan model, önceden öğrenilmiş temel özelliklerini korurken, son katmanlarda özelleştirilmiş sınıflandırma görevi için uygun hale getirilmiştir. Sonuç olarak önceden eğitilmiş derin öğrenme modeli alınarak, modelin alt katmanları donduruldu ve bu model üzerinde özel bir sınıflandırıcı katman eklenerek yeni bir model oluşturuldu. Bu yapılandırma sayesinde, yeni görev için daha az veri ve hesaplama gücü ile etkin bir eğitim yapılmıştır.

Çalışmada kullanılan veri setinin %70'i eğitim, %20'si test ve %10'u ise doğrulama için kullanılmıştır. Veri setinde artırma işleminden sonra eğitim için 7309 görüntü, test için 2091 ve doğrulama için 1046 görüntü kullanılmıştır. VGG-19 mimarisi kullanarak eğitim 25 epoch olarak gerçekleştirilmiştir.

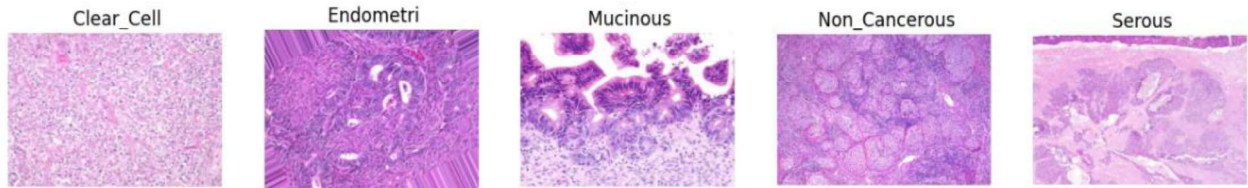
Eğitim sürecinin yönetiminde çeşitli önemli parametreler ve işlemler bulunmaktadır. Öncelikle 'ModelCheckpoint' ve 'EarlyStopping' kullanılarak modelin eğitim sürecinde takip edilmiştir ve en iyi performansa sahip model kaydedilmiştir. Doğrulama seti doğruluğu baz alınarak 'ModelCheckpoint' geri çağırısı, eğitim sırasında modelin performansının izlenmesini ve en iyi modelin kaydedilmesini sağlamaktadır. Kullanılan sabır (patience) değeri ise belirlenen değere göre eğitimi otomatik durdurarak aşırı uyumun önüne geçmektedir. Sabır (patience) değeri 10 olarak belirlendi. Bu sayede, 10 epoch boyunca doğrulama setindeki doğrulukta herhangi bir gelişme olmaması durumunda eğitim otomatik olarak durduruldu. Bu noktada 'model.compile' işlevi, derleme aşamasında modelin optimize ediciyi, kayıp fonksiyonunu ve metrikleri belirlemekte kullanıldı. Çalışmada optimizer parametresi olarak Adam optimize edici kullanılarak kategorik çapraz entropi kaybı(categorical_crossentropy) ve doğruluk(accuracy) metriği kullanılmıştır. Eğitim aşamasında 'fit' fonksiyonu çağılarak model eğitilmiş ve 'class_weights_dict' kullanılarak sınıf dengesizliği ele alınmıştır. Bu bilgiler ile eğitim süreci takip ayrıntılı bir şekilde takip edildi ve model istenilen

performansa ulaştığında eğitime son verildi ve modelin performansını değerlendirme aşamasına geçildi.

Yumurtalık Kanseri Veri Seti

The Cancer Genome Atlas – Yumurtalık Kanseri (TCGA-OV) deposundan elde edilen veri seti Pune Hindistan'daki Smt. Kashibai Navale Tıp Koleji ve Genel Hastanesi'nde tedavi gören 42 hasta üzerinde yapılan araştırmada, kansersiz hücrelere sahip 85 görüntü toplama süreci değerlendirilmiştir. Veri seti oluşturulurken yaşanan zorluğun nedeni Seröz karsinom dışındaki sınıflar için tedavi görmüş hasta sayısının azlığından kaynaklanmıştır. Uzman patoloğlar tarafından toplanan örnekler Leica ICC50 mikroskop kamerası kullanılarak yakalanmıştır. TCGA, Ulusal Kanser Enstitüsü (NCI), Ulusal Sağlık Enstitüleri (NIH), Ulusal İnsan Genom Araştırma Enstitüsü (NHGRI) ve Sağlık ve İnsan Hizmetleri Departmanı arasında yapılan bir iş birliğidir. Deposu, web sitelerine kaydılduktan sonra genel kullanıma açıktır. Elde edilen RGB görüntüleri, 224x224x3, JPG formatına yeniden boyutlandırılmıştır(K. R. Kasture vd., 2022)

Çalışmada, toplamda 175 Serous, 100 Mucinous, 60 Endometri, 80 Clear_Cell, ve 85 Non_Cancerous olmak üzere 500 etiketli histopatolojik görüntü, Ulusal Kanser Enstitüsü Genom Veri merkezi portalı, TCGA-OV deposundan toplanmıştır. Bu platform, kanser araştırmacıları ve biyobilimcilere kanser verilerini arayıp indirme imkanı sunan bir platformdur. Örnekleri Şekil-2'de gösterilen veri seti Mendeley veri web sitesinde yayınlanmıştır(K. Kasture, 2021) Veri seti tüm alt tiplere göre kategorize edilmiş ve etiketlenmiş olduğundan, gereksiz görüntü bulunmaması sağlanmıştır(K. R. Kasture, Shah, vd., 2021)



Şekil 2. Veri Setinden Örnek Görüntüler

Veri Arttırma (Data Augmentation)

Çalışmada TensorFlow kütüphanesinde yer alan ImageGenerator sınıfının içerdiği Şekil 3'de gösterilen döndürme, yakınlaştırma yatay/dikey çevirme genişlik/yükseklik kaydırma ve parlaklık ayarlama teknikleri kullanılmıştır. Her sınıfa ait orijinal görüntüler üzerinde bu teknikler uygulanarak verileri arttırıldıktan sonra oluşan yeni örneklerin her biri ayrı klasörlerde saklanmıştır. Modelin performansı arttırıp, genelleme yeteneğinin güçlenmesi için yapılan bu işlem sonucunda orijinal veri setinden çok daha büyük boyutta bir veri seti elde edilmiştir. Bu çeşitlilik sayesinde modelin ezberleme (over-fitting) durumunun engellenmesi amaçlanmıştır.

Performans değerlendirme metriği olarak Doğruluk (Accuracy), Hatırlama (Recall), Hassasiyet (Precision), F1-Puanı(F1-Score) kullanılmıştır. Model performansının değerlendirilmesi için yaygın olarak kullanılan başarı metrikleri doğruluk tabanlı bir performans değerlendirmesine dayanmaktadır. Modelin doğruluğu Doğru Negatif (TN), Doğru Pozitif(TP), Yanlış Negatif(FN)ve Yanlış Pozitif(FP) ile ölçülmüştür(Sokolova vd., 2006).

Burada DN doğru şekilde kanserli olarak sınıflandırılan görüntüleri, YP yanlış şekilde kanserli olarak tanımlanan kanserli olmayan görüntüleri, YP kanserli sınıfa ait olan görüntülerin kanserli değil olarak sınıflandırıldığı anlamına gelirken, DN doğru sınıflandırılan kanserli olmayan görüntüleri temsil eder. Doğruluk (Accuracy) toplam tahminler içinde doğru tahminlerin yüzdesini gösteren orandır. Doğru tahminlerin toplam tahminlere oranını belirtir(K. R. Kasture, Sayankar, vd., 2021).Şu şekilde hesaplanır:

$$\text{Doğruluk(Accuracy)} = \frac{(TP + TN)}{(TP + TN + FP + FN)}$$

$$\text{Hatırlama (Recall)} = \frac{TP}{(TP + FN)}$$

$$\text{Hassasiyet (Precision)} = \frac{TP}{TP + FP}$$

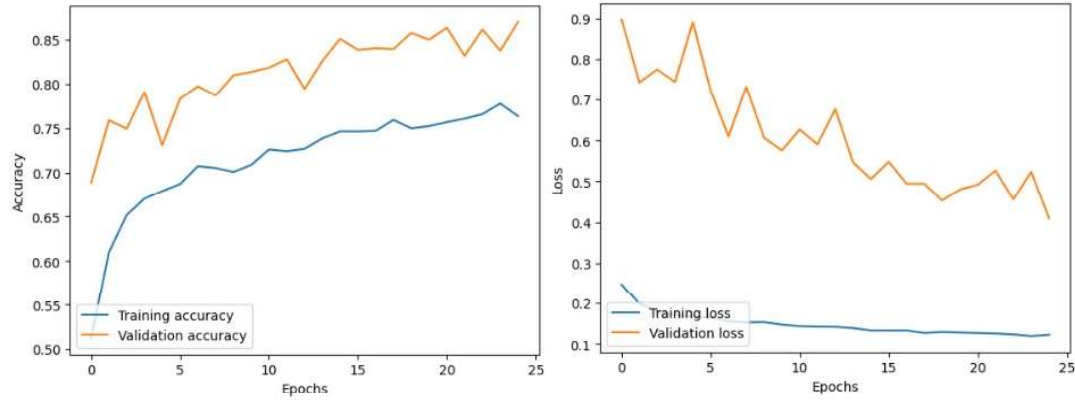
$$\text{F1 Puanı (F1 Score)} = 2 * \frac{(\text{Hassasiyet} * \text{Hatırlama})}{(\text{Hassasiyet} + \text{Hatırlama})}$$

BULGULAR VE TARTIŞMA

Bu bölümde, modelin eğitim boyunca elde ettiği performansın detaylı bir analizi sunulmaktadır.Başarım grafikleri ve karmaşıklık matrisi sonuçları, modelin eğitim sürecinin ve performansının değerlendirilmesinde önemli bir role sahiptir. Başarım grafikleri modelin eğitim sürecindeki performansını görsel olarak ifade ederken, karmaşıklık matrisleri ise modelin her sınıf için ne kadar doğru ve yanlış sınıflandırma yaptığını göstermektedir. Yapılan çalışmada başarım grafikleri ve karmaşıklık matrisi bir arada kullanılarak modelin performansı analiz edilmiştir. Başarım grafiği Şekil 5’te, karmaşıklık matrisi ise Şekil 6’da bulunmaktadır.

Şekil 5’te verilen başarım grafiklerinde “Training Accuracy” ve “Validation Accuracy” grafiği modelin doğruluğunu göstermektedir. Eğitim başlangıcında düşük olan eğitim ve doğrulama doğruluğu eğitim süreci ilerledikçe artmaktadır. Modelin eğitim sürecindeki performansı iyileşirken, model eğitim ve doğrulama verilerine de uyum sağlamaktadır. Modelin eğitim doğruluğu ve doğrulama doğruluğu arasındaki farkın zamanla azalması modelin aşırı uyuma dirençli olduğunu ve genelleme yeteneğinin arttığını göstermektedir.

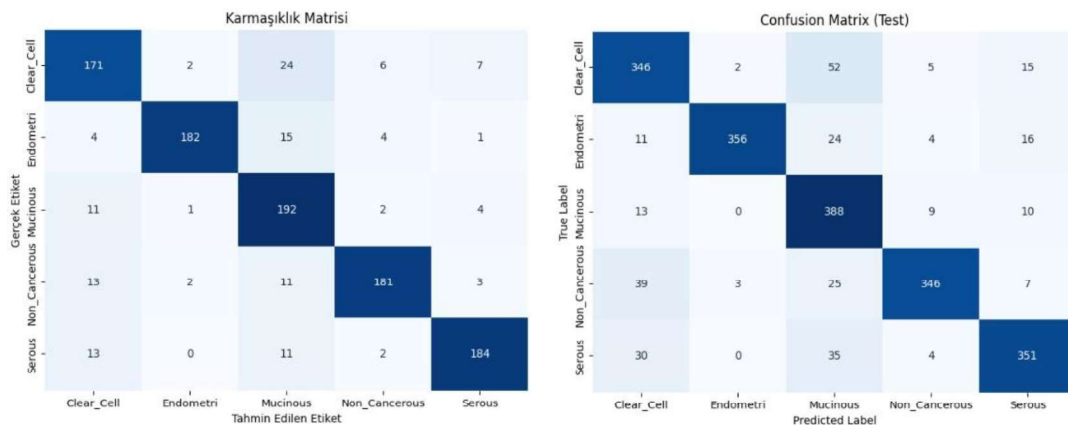
“Training Loss” ve “Validation Loss” grafiğinde ise değerleri arasındaki ilişki ve modelin eğitim sürecindeki başarımı gözlemlenmektedir. Eğitim sürecinin başlangıcında eğitim kaybı (Training Loss) ve doğrulama kaybı (Validation Loss) yüksektir. Fakat eğitiminin devam etmesiyle kayıpların azaldığı, değerlerin birbirine yaklaştığı görülmektedir. Bu durum modelin aşırı uyumdan kaçınarak verilere daha iyi uyum sağladığını göstermektedir. Eğitim süresince devam eden 25 epoch boyunca eğitim kaybı ve doğrulama kaybı arasındaki farkın sürekli olarak azaldığı bu sayede modelin eğitim boyunca istikrarlı bir şekilde devam ettiği gözlemlenmektedir. Sonuç olarak eğitim süreci boyunca eğitim doğruluğu ve doğrulama doğruluğu artmış, eğitim kaybı ve doğrulama kaybı azalmıştır. Bu durum modelin eğitim süresince verilere uyum sağladığını ve başarılı bir şekilde öğrenme sağladığını göstermektedir.



Şekil 5. VGG-19 Başarım Grafikleri

Karmaşıklık matris analizi, doğrulama ve test setlerinde modelin performansını değerlendirmek için önemli bir araçtır. Çalışma kapsamında, modelin performansını değerlendirmek için oluşturulan karmaşıklık matrisinin görselleştirilmiş hali Şekil 6'da sunulmuştur. Bu analizde Endometri, Non_Cancerous ve Serous hastalıklarının yüksek geri çağırma oranları ile oldukça iyi performans gösterdiği görülmektedir. Mucinous ve Clear_Cell hastalıklarının daha düşük geri çağırma oranları ile daha zayıf performans gösterdiği belirlenmiştir. Doğrulama seti ile en iyi performans gösteren hastalıklar arasında Non_Cancerous 192 doğru tahmin ile ilk sırayı alırken bunu 184 doğru tahmin ile Serous ve 182 doğru tahmin ile Non_Cancerous izlemektedir. Test verileri ile yapılan sınıflandırma da ise en iyi performansı 356 doğru tahmin ile Endometri, 351 doğru tahmin ile Serous ve 346 doğru tahmin ile Non_Cancerous göstermektedir. En iyi performans gösteren hastalıklar genellikle Endometri ve Non_Cancerous olarak belirlenmiştir. Endometri ve Non_Cancerous hastalıkları için precision, recall ve F1-score değerleri hem doğrulama hem de test setlerinde yüksek olmuştur. Endometri için test setinde %99 precision ve %87 recall elde edilirken, Non_Cancerous için ise %94 precision ve %82 recall elde edilmiştir. Bu yüksek değerler, bu hastalıkların model tarafından başarılı bir şekilde tanındığını göstermektedir.

En kötü performans gösteren hastalıklar ise Clear_Cell ve Mucinous olarak belirlenmiştir. Clear_Cell hastalığı için precision, recall ve F1-score değerleri hem doğrulama hem de test setlerinde düşük olmuştur. Mucinous hastalığı ise test setinde precision açısından daha düşük değer gösterirken, recall değeri daha yüksektir. Bu durum, test setinde Mucinous hastalığının daha fazla yanlış pozitif tahminlere sahip olduğunu, ancak daha az yanlış negatif tahminlere sahip olduğunu gösterebilir. Dolayısıyla, modelin özellikle Mucinous hastalığını tanıma konusunda iyileştirmeye ihtiyacı vardır.



Şekil 6. Doğrulama ve Test Sonuçlarının Karmaşıklık Matrisi

Tablo 1’de, VGG-19 modelinin performansını değerlendirmek için kullanılan sınıflandırma raporlarını göstermektedir. Test ve doğrulama verileri ile iki farklı veri seti üzerinden elde edilen sonuçlar raporlanmıştır. İlk olarak %70 eğitim, %20 test verileri kullanılarak eğitilen modelin test veri seti üzerinde elde ettiği sonuçlar sunulmuştur. Her bir hastalık sınıfı için Precision, Recall ve F1-Score metrikleri raporlanmıştır. Son olarak, tüm sınıflar için Accuracy (doğruluk oranı) %85 olarak raporlanmıştır. Aynı model %70 eğitim, %10 doğrulama veri kullanılarak veri seti üzerindeki performansı raporlanmıştır. Aynı şekilde her sınıf için Precision, Recall ve F1-Score değerleri verilmiştir. Tüm sınıfların doğruluk oranı %87 olarak raporlanmıştır. Tablo 1 sonuç olarak modelin her sınıf için ne kadar doğru tahmin yaptığını, yanlış sınıflandırmaların ne kadar olduğunu ve doğruluk oranlarının test ve doğrulama veri setlerine göre nasıl değiştiğini göstermektedir.

Tablo 1’deki sonuçlar incelendiğinde modelin sınıfları iyi tahmin ettiği görülmektedir. Fakat sınıflar arasındaki başarı oranlarında bir miktar değişimler olduğu görülmektedir, bazı sınıfların diğerlerine göre daha iyi performans gösterdiği gözlemlenmektedir. Precision (Hassasiyet) değeri incelendiğinde modelin çoğu sınıfı yüksek hassasiyet ile tahmin ettiği fakat Mucinous sınıfında iyileştirmeye ihtiyacı olduğu belirlenmiştir. Recall (Duyarlılık) değerlerinde ise modelin özellikle Mucinous sınıfında oldukça başarılı, Non Canceous sınıfında diğer sınıflara göre zayıf kaldığı görülmektedir. F1- Puanı (F1-Score) değerlendirildiğinde modelin çoğu sınıfı dengeli bir şekilde tahmin ettiği görülmekte, Mucinous sınıfının hassasiyetinin biraz düşük olması sebebi ile bu alanda iyileştirme yapılması gerektiği belirlenmiştir.

Tablo 1. VGG-19 Performans Değerlendirme

VGG-19 Performans Değerlendirme					
	Sınıflar	Precision	Recall	F1 Score	Accuracy
Eğitim (%70) – Test (%20)	Clear Cell	0,79	0,82	0,81	
	Endometroid	0,99	0,87	0,92	
	Mucinous	0,74	0,92	0,82	
	Non Cancerous	0,94	0,82	0,88	%85
	Serous	0,88	0,84	0,86	
Eğitim (%70) – Doğrulama (%20) - Test (%10)	Clear Cell	0,81	0,81	0,81	
	Endometroid	0,97	0,88	0,93	
	Mucinous	0,76	0,91	0,83	
	Non Cancerous	0,93	0,86	0,89	
	Serous	0,92	0,88	0,90	%87

Sonuçlara bakıldığında modelin oldukça etkin bir şekilde çalıştığı görülmektedir. Test verileri üzerinden elde edilen sonuçta “Clear Cell” sınıfı için %81 hassasiyet, %81 duyarlılık, %81 F1 puanı elde edilirken aynı şekilde diğer sınıflar içinde yüksek performans elde edilmiştir. Bu sonuç modelin sınıflar arasında dengeli bir şekilde çalıştığını göstermektedir. Doğrulama verileri üzerinde yapılan değerlendirmede benzer başarı göstermektedir. Genel olarak sonuçlara bakıldığında kullanılan derin öğrenme modelimizin yumurtalık kanseri alt türlerini başarı ile sınıflandırdığı ve dengeli bir performans sergilediğini göstermektedir.

SONUÇ

Çalışmada, yumurtalık kanserinin sınıflandırılması için derin transfer öğrenmeye dayalı bir yaklaşım sunulmuş ve önceden eğitilmiş VGG-19 modeli temel alınarak yeni bir model yapılandırılmıştır. Yumurtalık kanseri sınıflandırılması için geliştirilen bu derin öğrenme modeli, öncelikle veri setini uygun oranlarda eğitim, test ve doğrulama olarak üçe bölünmesi ile başlamaktadır. Eğitim verisi %70 test verisi %20 ve doğrulama verisi %10 olarak belirlenmiştir. Bölünen veri seti, her bir sınıfa ait örneklerin dengeli bir şekilde dağılmasını sağlayarak modelin genelleme yeteneğini arttırmıştır. Ardından eğitim verileri kullanılarak veri artırma işlemi yapılmıştır. Bu artırılmış görüntüler ile modelin genel ve sağlıklı bir öğrenme sağlamasını hedeflenmiştir. Görüntü artırma işlemleri, rotasyon, kayma parlaklık ayarları gibi çeşitli teknikler uygulanmıştır. Eğitim 25 epoch gerçekleştirilmiştir ve 'EarlyStoopping', 'ModelCheckpoint' gibi önemli parametreler de kullanılmıştır. Eğitim süresince kullanılan erken durdurma ve kontrol noktası sayesinde modelin aşırı uyuma direnci artırılmış, daha iyi sonuçlar alınmasına yardımcı olmuştur. Bu aşama modelin eğitim ve değerlendirme aşamalarını önemli ölçüde etkilemekte ve sürecin güvenilir bir şekilde gerçekleşmesini sağlamak adına önemlidir.

Model, transfer öğrenme yoluyla önceden eğitilmiş ağırlıkları kullanarak yumurtalık kanseri sınıflandırma görevini başarılı bir şekilde gerçekleştirmiştir. Çalışmada kullanılan derin öğrenme projesi Colab ortamında yüksek RAM ve TPU kullanarak gerçekleştirildi. Fakat buna rağmen eğitim sürelerinin oldukça uzun sürdüğünü gözlemledik. Gelecekte yapılacak çalışmalar için öncelikle daha fazla hesaplama gücüne sahip donanım kaynaklarına erişim sağlamak için yüksek kapasiteli GPU veya daha gelişmiş TPU kullanımı düşünülebilir. Bunlara ilaveten optimizasyon ve hiperparametre ayarlamalarında farklı kombinasyonlar denenebilir.

Elde edilen sonuçlar umut vaat etmektedir, gelecekte çalışmada kullanılan derin öğrenme modelinin iyileştirilmesiyle, veri setinin artırılması ve aynı veriler ile MobileNet-V2, InceptionV3, Resnet50, Xception gibi daha fazla transfer öğrenme yönteminin kullanılması ile daha başarılı sonuçlar elde edilebilir. Sonuç olarak algoritma ve donanım ile ilgi iyileştirmeler eğitim sürelerini kısalmasına ve daha etkili ve hızlı sonuçlar elde edilmesine imkan sağlayacaktır.

KAYNAKLAR

- Abdelaziz Ismael, S. A., Mohammed, A., & Hefny, H. (2020). An enhanced deep learning approach for brain cancer MRI images classification using residual networks. *Artificial Intelligence in Medicine*, 102. <https://doi.org/10.1016/j.artmed.2019.101779>
- Aljuaid, H., Alturki, N., Alsubaie, N., Cavallaro, L., & Liotta, A. (2022). Computer-aided diagnosis for breast cancer classification using deep neural networks and transfer learning. *Computer Methods and Programs in Biomedicine*, 223. <https://doi.org/10.1016/j.cmpb.2022.106951>
- Celik, Y., Talo, M., Yildirim, O., Karabatak, M., & Acharya, U. R. (2020). Automated invasive ductal carcinoma detection based using deep transfer learning with whole-slide images. *Pattern Recognition Letters*, 133, 232–239. <https://doi.org/10.1016/J.PATREC.2020.03.011>
- Kasture, K. (2021). *OvarianCancer&SubtypesDatasetHistopathology*. 1. <https://doi.org/10.17632/KZTYMSRJX9.1> <https://data.mendeley.com/datasets/kztymsrjx9/1>
- Kasture, K. R., Choudhari, D., & Matte, P. N. (2022). Prediction and Classification of Ovarian Cancer using Enhanced Deep Convolutional Neural Network. *International Journal of Engineering Trends and Technology*, 70(3), 310–318. <https://doi.org/10.14445/22315381/IJETT-V70I3P235>
- Kasture, K. R., Sayankar, B. B., & Matte, P. N. (2021). Multi-class Classification of Ovarian Cancer from Histopathological Images using Deep Learning - VGG-16. *2021 2nd Global Conference for Advancement in Technology, GCAT 2021*. <https://doi.org/10.1109/GCAT52182.2021.9587760>
- Kasture, K. R., Shah, D. D., & Matte, P. N. (2021). A new deep learning method for automatic ovarian cancer prediction & subtype classification. *turcomat.org*, 12(12), 1233–1242. <https://doi.org/10.17632/w39zgzksp6n.1>

- Kavitha, S., Dhanapriya, B., Vignesh, G. N., & Baskaran, K. R. (2021). Neural Style Transfer Using VGG19 and Alexnet. *2021 International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation, ICAECA 2021*. <https://doi.org/10.1109/ICAECA52838.2021.9675723>
- Kongara, S. R., Prakasha, S., Brindha, A., Pathak, S. K., Miya, J., Taqui, S. N., Almoallim, H. S., Alharbi, S. A., & Raghavan, S. S. (2024). Performance evaluation of optimized convolutional neural network mechanism in the detection and classification of ovarian cancer. *Multimedia Tools and Applications*. <https://doi.org/10.1007/S11042-024-18115-0>
- Menon, U., Gentry-Maharaj, A., Burnell, M., Singh, N., Ryan, A., Karpinskyj, C., Carlino, G., Taylor, J., Massingham, S. K., Raikou, M., Kalsi, J. K., Woolas, R., Manchanda, R., Arora, R., Casey, L., Dawnay, A., Dobbs, S., Leeson, S., Mould, T., ... Parmar, M. (2021). Articles Ovarian cancer population screening and mortality after long-term follow-up in the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS): a randomised controlled trial. *www.thelancet.com*, 397. [https://doi.org/10.1016/S0140-6736\(21\)00731-5](https://doi.org/10.1016/S0140-6736(21)00731-5)
- Sokolova, M., Japkowicz, N., & Szpakowicz, S. (2006). Beyond accuracy, F-score and ROC: A family of discriminant measures for performance evaluation. *AAAI Workshop - Technical Report, WS-06-06*, 24–29. https://doi.org/10.1007/11941439_114
- Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F. (2021). Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: A Cancer Journal for Clinicians*, 71(3), 209–249. <https://doi.org/10.3322/CAAC.21660>
- Tang, Y., Liu, M., Hu, L., Wang, C., He, Y., Zeng, C., Lin, K., He, Z., & Huo, W. (2022). A deep learning method for breast cancer classification in the pathology images. *ieeexplore.ieee.org* M Liu, L Hu, Y Tang, C Wang, Y He, C Zeng, K Lin, Z He, W Huo *IEEE Journal of Biomedical and Health Informatics*, 2022•*ieeexplore.ieee.org*, 26(10), 5025. <https://doi.org/10.1109/JBHI.2022.3187765>
- Xu, Y., Jia, Z., Wang, L. B., Ai, Y., Zhang, F., Lai, M., & Chang, E. I. C. (2017). Large scale tissue histopathology image classification, segmentation, and visualization via deep convolutional activation features. *BMC Bioinformatics*, 18(1). <https://doi.org/10.1186/S12859-017-1685-X>
- Zheng, Y., Yang, C., & Merkulov, A. (2018). Breast cancer screening using convolutional neural network and follow-up digital mammography. <https://doi.org/10.1117/12.2304564>, 10669, 1066905. <https://doi.org/10.1117/12.2304564>