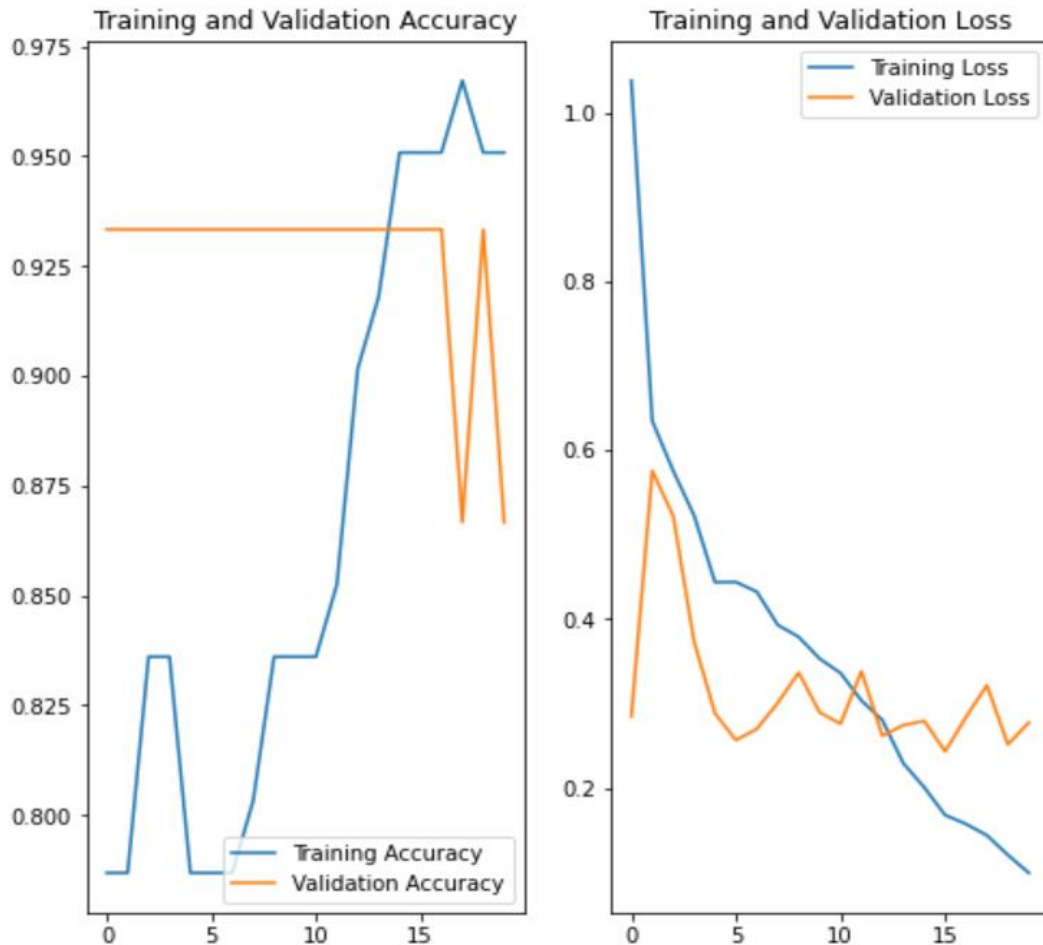


- 1) Tiroid Ultrason Görüntülerinin dijital veritabanı, açık kaynaklı CIM@LAB ve IDIME tarafından sağlanmıştır. Veritabanı 99 vaka ve 134 resim içermektedir. İyi huylu ve kötü huylu tümör görüntüleri benigns ve maligns olarak klasör altında bulunmaktadır. Veri etiketleri klasör isimlerinden alınarak kategorik olarak belirlenmiştir. Öznitelikler yaş, cinsiyet, tirad skoru, kireçlenme, kenarlar, ekojenite, tümör bileşeni (sertlik, olgunluk) ve tümör ultrason görüntüleridir.

```
<?xml version="1.0"?>
<case>
  <number>549</number>
  <age>31</age>
  <sex>F</sex>
  <composition>solid</composition>
  <echogenicity>hypoechogenicity</echogenicity>
  <margins>well defined smooth</margins>
  <calcifications>none</calcifications>
  <tirads>4c</tirads>
  <reportbackoff>
  <reportecol>
  <mark>
    <image>1</image>
    <svg>[{"points": [{"x": 292, "y": 123}, {"x": 293, "y": 129}, {"x": 298, "y": 138}, {"x": 300, "y": 139}, {"x": 306, "y": 144}, {"x": 308, "y": 145}, {"x": 311, "y": 148}, {"x": 314, "y": 149}, {"x": 315, "y": 150}, {"x": 319, "y": 152}, {"x": 323, "y": 154}, {"x": 332, "y": 155}, {"x": 336, "y": 157}, {"x": 337, "y": 157}, {"x": 345, "y": 157}, {"x": 350, "y": 156}, {"x": 355, "y": 154}, {"x": 361, "y": 152}, {"x": 364, "y": 150}, {"x": 367, "y": 148}, {"x": 369, "y": 142}, {"x": 369, "y": 142}, {"x": 370, "y": 137}, {"x": 370, "y": 131}, {"x": 371, "y": 128}, {"x": 371, "y": 126}, {"x": 368, "y": 127}, {"x": 365, "y": 114}, {"x": 361, "y": 111}, {"x": 361, "y": 110}, {"x": 356, "y": 106}, {"x": 352, "y": 104}, {"x": 350, "y": 104}, {"x": 349, "y": 103}, {"x": 346, "y": 103}, {"x": 340, "y": 103}, {"x": 338, "y": 101}, {"x": 335, "y": 101}, {"x": 334, "y": 102}, {"x": 328, "y": 105}, {"x": 322, "y": 105}, {"x": 320, "y": 107}, {"x": 319, "y": 110}, {"x": 318, "y": 112}, {"x": 318, "y": 113}, {"x": 316, "y": 118}, {"x": 314, "y": 121}, {"x": 312, "y": 123}, {"x": 307, "y": 126}, {"x": 304, "y": 128}, {"x": 300, "y": 129}, {"x": 293, "y": 129}], "annotation": {}, "regionType": "freeland"}]</svg>
    </mark>
    <image>2</image>
    <svg>[{"points": [{"x": 259, "y": 142}, {"x": 258, "y": 138}, {"x": 258, "y": 138}, {"x": 258, "y": 133}, {"x": 259, "y": 128}, {"x": 260, "y": 125}, {"x": 261, "y": 122}, {"x": 261, "y": 119}, {"x": 262, "y": 118}, {"x": 263, "y": 114}, {"x": 264, "y": 114}, {"x": 266, "y": 103}, {"x": 270, "y": 101}, {"x": 273, "y": 101}, {"x": 275, "y": 100}, {"x": 281, "y": 95}, {"x": 287, "y": 95}, {"x": 289, "y": 95}, {"x": 289, "y": 95}, {"x": 306, "y": 95}, {"x": 308, "y": 95}, {"x": 309, "y": 95}, {"x": 314, "y": 100}, {"x": 318, "y": 103}, {"x": 322, "y": 104}, {"x": 324, "y": 107}, {"x": 327, "y": 112}, {"x": 327, "y": 115}, {"x": 323, "y": 120}, {"x": 319, "y": 123}, {"x": 316, "y": 131}, {"x": 310, "y": 138}, {"x": 307, "y": 142}, {"x": 302, "y": 145}, {"x": 291, "y": 148}, {"x": 286, "y": 148}, {"x": 281, "y": 148}, {"x": 281, "y": 147}, {"x": 272, "y": 144}, {"x": 266, "y": 142}, {"x": 264, "y": 142}, {"x": 259, "y": 143}], "annotation": {}, "regionType": "freeland"}]</svg>
  </case>
```

- 2) Github'da Project.ipynb adı altında kodlar bulunmaktadır.
- 3) Daha büyük bir veriseti kullanarak eğitilen modelde daha başarılı sonuçların alınacağı aşikardır. Ayrıca her bir epochs için accuracy artışı dikkat çekicidir. Epochs sayısı artırmak model başarımını artıracaktır. Epochs sayıları ile validation/train accuracy kıyaslandığında 15.epoch tan sonra doğrulama başarımının dalgalandığı (negatif yönde) ve eğitim başarımının düşüşe geçtiği görünüyor. Bu durumda ideal epoch sayısını yaklaşık 15 civarı tutarak model başarımını zirvede tutabiliriz.



Training and Validation Accuracy grafiğine bakıldığında test ve eğitim başarımlarından bu modelde aşırı uyumluluk (overfitting) problemi olmadığı görünüyor. Bu sebeple çapraz doğrulama (cross-validation) yöntemine ihtiyaç duyulmamıştır.

- 4) Daha önce bilgisayarla görü-örüntü tanıma projelerinde çalışmamış olsam da, kütüphane dokümanlarından yararlanarak temel bir image recognition projesi gerçekleştirdim. Ancak tiroid test ve labaratuvar sonuçlarından yararlanarak (kan tahlilleri, kitle ölçüleri, hormon değerleri vs.) gibi sayısal veriler ile KNN, SVM, CNN gibi farklı algoritmalarından daha komplike sonuçlar alınabilirdi. K-En yakın komşu algoritmasının yüksek gürültü toleransı, Destek vektör makineleri algoritmasının hafıza verimi ve keskin marjlardaki yüksek doğruluk oranından yararlanarak sonuçlarımı kıyasladım.
- 5) Tırat Skorunun bu problemin çözümünde yetersiz kalması olasılıklardan bir tanesidir. Ancak ilk önce veri setinde çeşitliliği artırmaya giderdim. Sonrası veriler arasındaki netliği özellikle veri setinde yanlışlık olup olmadığını tespit ederdim. Ayrıca her bir görüntüde ki tümör sayısını doğrulayarak aykırı bir durum olup oluşmadığını kontrol ederim.

İrem Sönmez