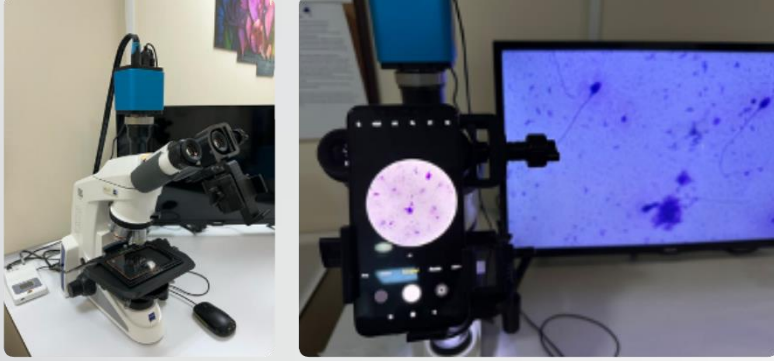


Hi-Lab SpermMorpho Veri Seti Bilgilendirmesi

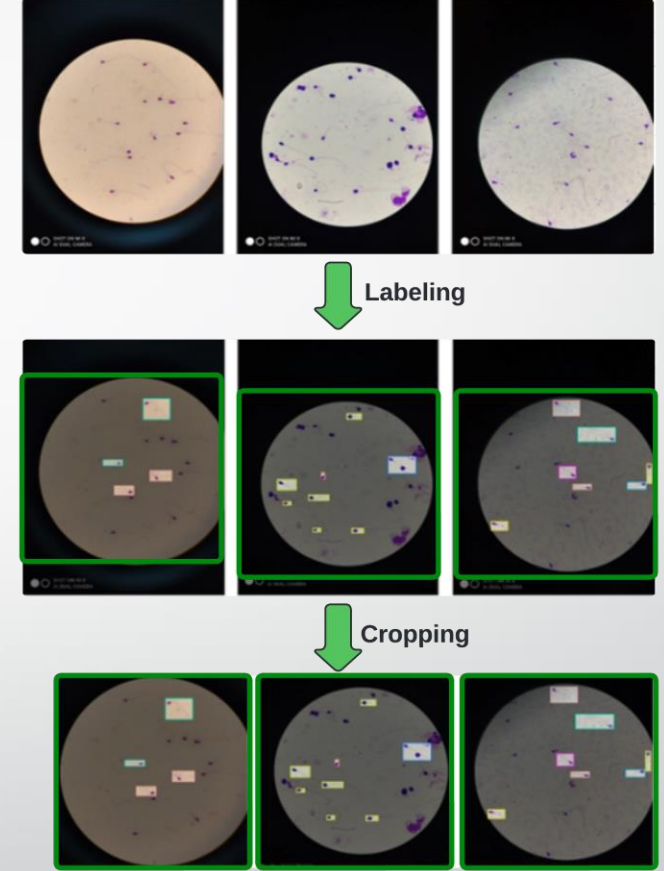
Notlandırma

- Projenin, genel notlandırmaya %30 etkisi olacaktır.
- Proje, uygulama ve sunum olmak üzere iki aşamada değerlendirilecektir.
- Veri üzerinde gerçekleştirilen uygulama kodunun sistem üzerinden gönderilerek, paylaşılmayan veriler üzerinde belli bir başarımla elde edilebilmesi proje puanının %80'ini oluşturacaktır.
- %80'in %50'si modelinizin F1 metriğinin performans sıralamasından, %50'si ise genel doğruluk (Accuracy) puanından performans sıralamasından hesaplanacaktır.
- Paylaşılan veri seti üzerinde yapılan çalışmalar ve elde edilen sonuçların sunumu ise proje puanının %20'lik kısmını oluşturacaktır.

Veri K mesinin Oluřturulması

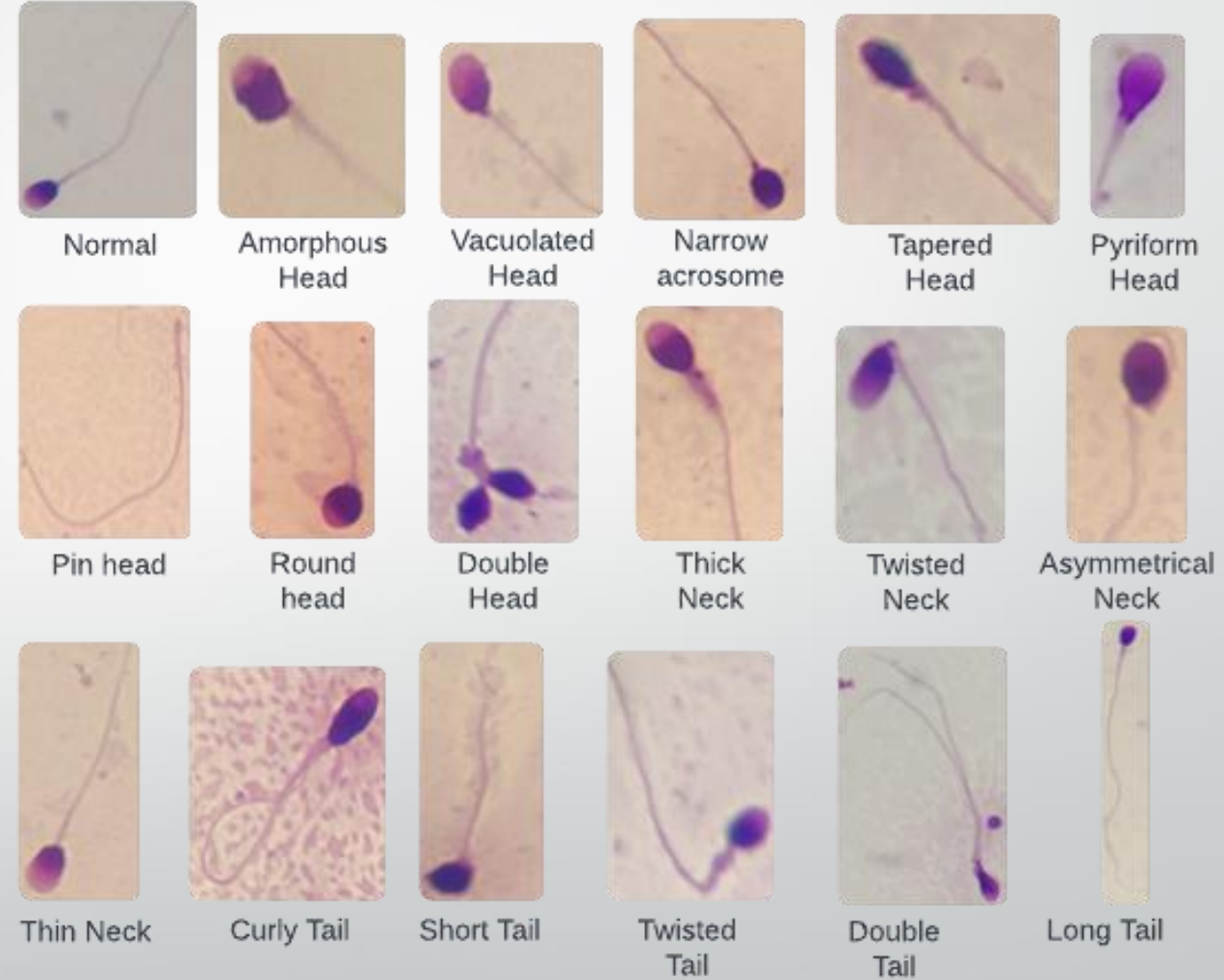


- G r nt ler Recep Tayyip ErdoĖan  niversitesi, EĖitim Arařtırma Hastanesi, Androloji laboratuvarında, gerekli etik izinler alınarak kaydedilmiřtir.
- G r nt lerin kaydı i in Cep Telefonu kamerası kullanılmıřtır.
- Semen  rneklerinin bio-iřaretleyici boyalar ile boyandıktan sonra mikroskop ok ler g r nt  kayıtları alınmıřtır.
- Mikroskop ok ler g r nt lerin laborant uzmanlar tarafından etiketlemesi ImageStudio ile ger ekleřtirilmiřtir.
- Manuel bir iřlem olarak etiket sınıfları ve koordinatlarına g re ana resimden morfolojik bozuntular ve normal sınıflı sperm g r nt leri koparılarak veri k mesi oluřturulmuřtur.

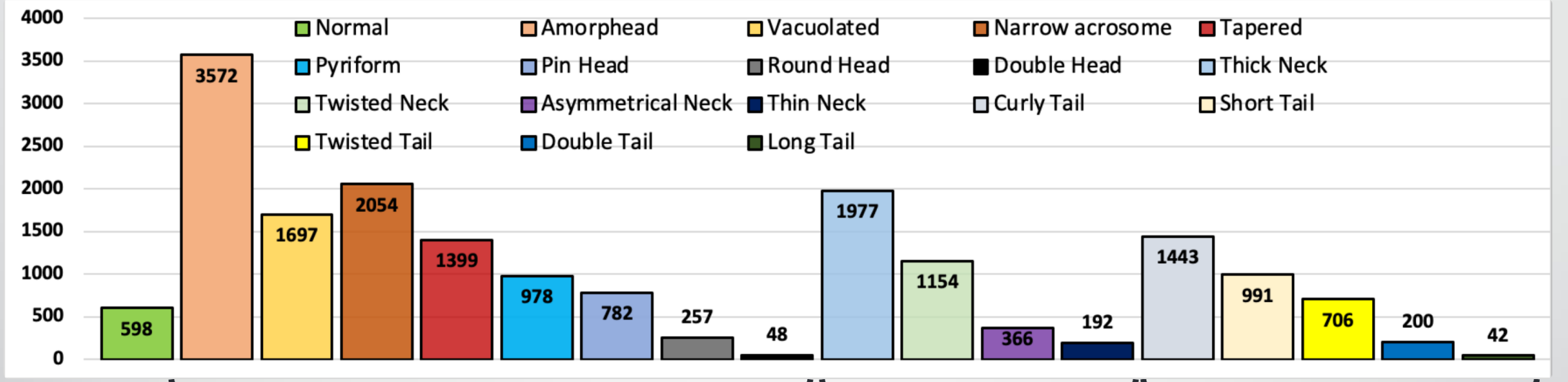


Veri Kümesi Detayları

- Hi-Lab Sperm Morpho veri kümesi toplamda 18 sınıftan oluşmaktadır. (1 Normal, 17 bozuntu sınıfı)
- Veri kümesinde yer alan sperm görüntüleri farklı yükseklik ve genişlikte olabilmektedir.
- Cep telefonu tabanlı görüntü kaydı alındığından çözünürlükleri düşük olabilmektedir.
- Veri Kümesinde toplamda 18462 adet sperm görüntüsü bulunmaktadır. Sizinle %30 oranında veri paylaşılmıştır. Toplamda 5530 adet görüntü sizinle paylaşılmıştır.
- Paylaşılan veri kümesindeki örnek – sınıf dağılımı, orijinal veri kümesinde olan dağılım ile birebir örtüşecek şekilde ayarlanmıştır.



Veri Kümesi Detayları



- Orijinal veri kümesine ait örnek – sınıf dağılımı görselde sunulmuştur.
- Sınıf İsimler:

AmorfHead	DoubleHead	NarrowAcrosome	PyriformHead	TaperedHead	TwistedNeck
AsymmetricNeck	DoubleTail	Normal	RoundHead	ThickNeck	TwistedTail
CurlyTail	LongTail	PinHead	ShortTail	ThinNeck	VacuolatedHead

Uygulama Değerlendirme

- Uygulamanızda «PyTorch» kütüphanesini kullanmanız gerekmektedir.
- Uygulamanızda önerdiğiniz yaklaşımın performansını scikit-learn kütüphanesinde yer alan «confusion_matrix» fonksiyonu ile karışıklık matrisini (confusion matrix) ve «classification_report» fonksiyonu ile raporlama sonuçlarını elde ederek değerlendiriniz.
- Karışıklık matrisinin görselleştirilmesi için seaborn veri görselleştirme kütüphanesi üzerinden aşağıda belirtilen formatta karışıklık matrisini elde ediniz.

```
import seaborn as sn
```

```
sn.heatmap(df_cm, annot=True, cmap="OrRd", fmt="d", annot_kws={"size": 10}, cbar=False)
```

- Proje dosyalarınızı tek bir «.zip» dosyası olacak şekilde online.yildiz.edu.tr platformu üzerinde tanımlanan proje kısmına yüklenmesi istenmektedir. Zip dosyası içerisinde bulunması gerekenler:
 - Uygulama dosyanız = **Sadece** test kısmını içeren kodunuz. Eğitim aşamasına ait kodların gönderilmesi **istenmemektedir**.
 - Model dosyanız = Test aşamasında çalışacak olan modelinizin dosyasını da sisteme yüklemeniz gerekmektedir. Bu hazır bir ağ bile olsa, sisteme yüklenmesi istenmektedir. Boyut ile alakalı engel oluşursa txt dosyası içerisine modelin indirilebileceği link paylaşılabilir.
 - Karışıklık Matrisi = Sizinle paylaşılan veri seti üzerinde modelinizin performansını gösteren karışıklık matrisi görüntüsü
 - Rapor = Sizinle paylaşılan veri seti üzerinde modelinizin detaylı performansını gösteren classification_report sonucunun görüntüsü

Uygulama Değerlendirme

- Proje teslim süresi sonrasında, test işlemi için açıklamalı bir şekilde hazırlanan kodunuzun sizinle paylaşılmayan veri kümesi üzerinde testi tarafımızdan gerçekleştirilerek performans sonuçları elde edilecektir.
- Tarafımızdan elde edilen performans sonuçları kapsamında F1 skoru ve Doğruluk puanı bazında genel sıralamalar oluşturulacaktır.
- Ayrıca sizinle paylaşılan veri kümesini kullanarak oluşturduğunuz ve performansını ölçeklendirdiğiniz modelin sizin tarafınızdan sisteme yüklenen sonuçları da toplu şekilde duyurulacak ve sıralanacaktır.
- Genel uygulama puanınız F1 skor ve Doğruluk puan sıralamanıza bağlı olarak oluşturulacaktır.

Sunum Değerlendirme

- Sunumda tasarladığınız model hakkında detaylı bilgi ve aşamalara yer vermeniz beklenmektedir.
- Model tasarımında baz aldığınız hazır çalışmalar var ise atıf vermeniz beklenmektedir.
- Etkin sunum hazırlama tekniklerine uygun şekilde sunumun hazırlanması beklenmektedir.
- Sizinle paylaşılan veri kümesi üzerinde gerçekleştirdiğiniz testler hakkında detaylı bilgi vermeniz beklenmektedir.
- Elde ettiğiniz sonuçlar nezdinde detaylı bulgulara ve yorumlara yer vermeniz beklenmektedir.
- Modelinizin ana veri kümesinde elde ettiği skoru yorumlamanız beklenmektedir.
- Gerçekleştireceğiniz sunum sonrası genel proje notunuzun %20'lik kısmı oluşacaktır.

BAŞARILAR...

Verileriniz temiz, algoritmalarınız hatasız, modelleriniz overfit olmasın! Bol şans ve düşük kayıp oranları dilerim. GPU'larınız asla ısınmasın, sonuçlarınız hep tatmin edici olsun!

Doç. Dr. Ali Can Karaca
ackaraca@yildiz.edu.tr

Doç. Dr. Hamza Osman İlhan
hoilhan@yildiz.edu.tr