

Healtyntelligence

AMAÇ:

Healtyntelligence, tarımda insan gücünü azaltmaya ve verimi arttırmaya yönelik olarak insan hayatına girmesi muhtemel otomasyon sistemlerinin, ürünün kalitesini saptamak için kullanabileceği bir Image Classification çalışmasıdır. Proje, basitçe domates içeren bir fotoğraftan, domatesin çürük ya da sağlam olduğunu saptamak üzere tasarlanmıştır. Proje sayesinde, domates hasadı esnasında çiftçiye ürün seçimi yapmaktan kurtararak vakit ve iş gücü tasarrufu hedeflenmektedir.

VERİ:

Çalışma için çeşitli dijital kaynaklardan ve marketlerden veri toplanmıştır. Toplam 338 domates fotoğrafının 247'si training, 91'i validation için ayrılmıştır. Training verisinin 147'si sağlam, 100 tanesi çürük domateslerden oluşmaktadır. Validation verisi için de 49 sağlam, 42 çürük domates fotoğrafı ayrılmıştır.

TEKNİKLER:

Proje için Pytorch kullanılarak ImageNet üzerinde Transfer Learning uygulanmıştır.

Training verileri üzerinde TorchVision ile data augmentation uygulanmıştır. Kullanılan yöntemler Random Resized Crop ve Random Horizontal Flip'tir. Yapılan testler sonucunda performansı çok kötü etkilemediği göz önünde bulundurularak Random Rotation (90, 270) da eklenmiştir. Ayrıca hem training, hem de validation için ayrılan veriler Normalize edilmiştir.

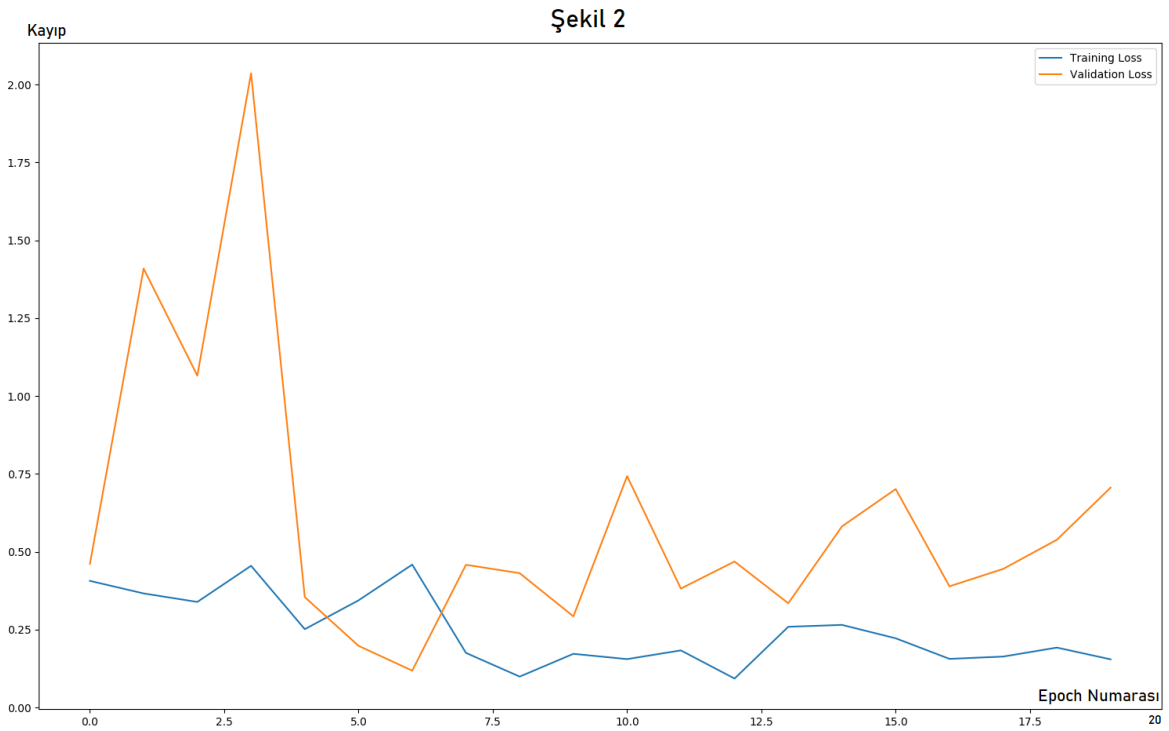
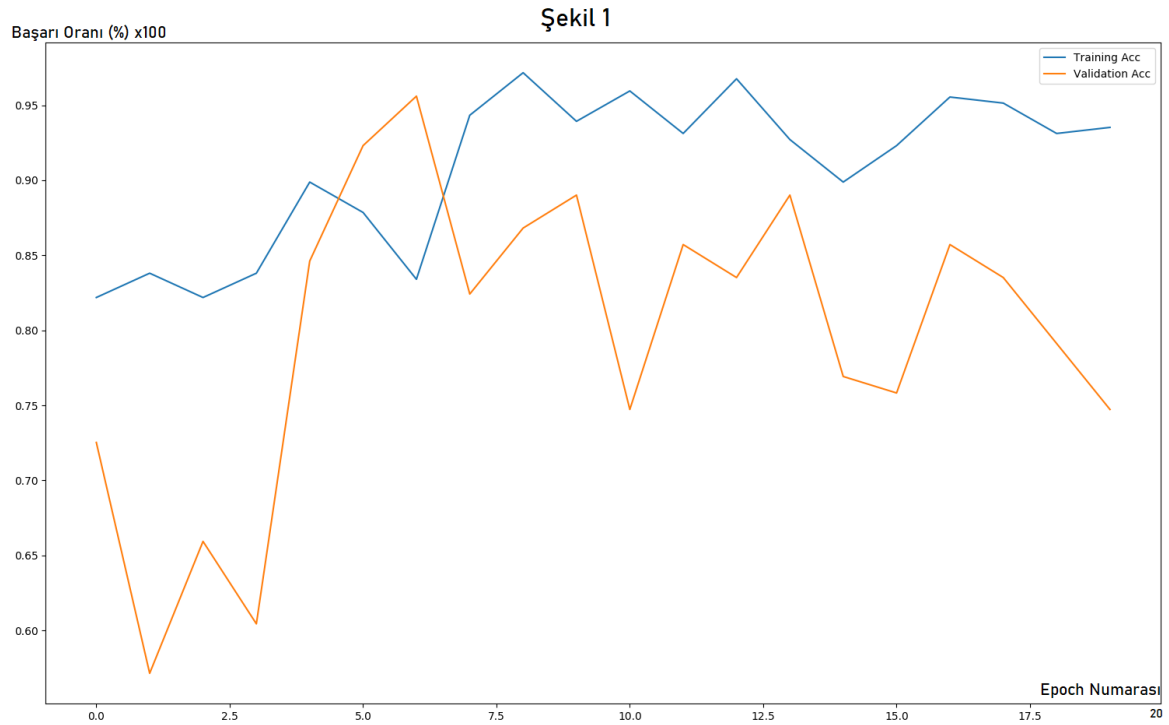
Eğitim için Learning Rate 0.001 olarak belirlenmiştir. Scheduler da 7 adımda bir 10% oranında ayarlanmıştır. 20 epoch ile eğitim yapılmıştır.

SONUÇLAR:

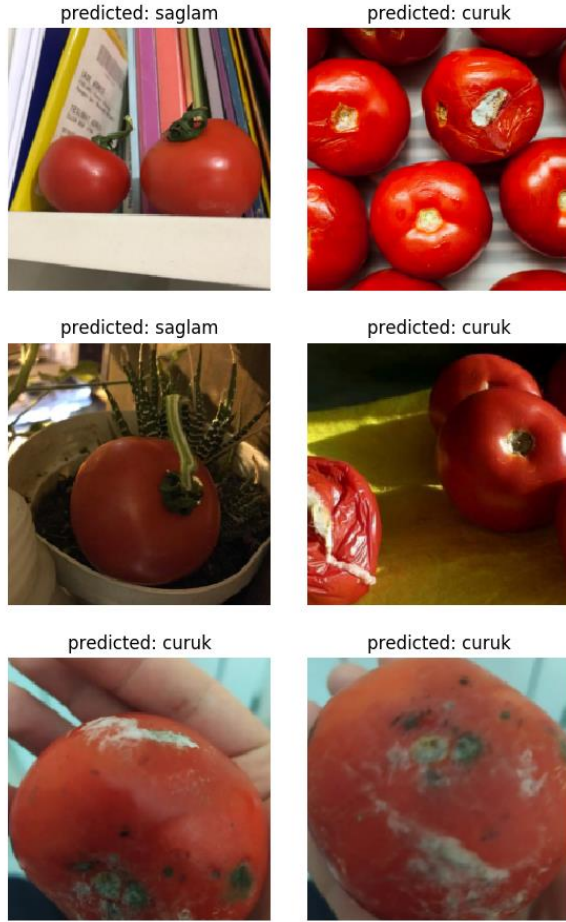
Eğitim 26 dakika 48 saniye sürmüştür, 95.6% oranında validation accuracy sağlanmıştır. Model "bestmodel.pth" ismiyle proje dizinine kaydedilmiştir. Epochlara ait doğruluk ve loss değerleri şöyle olmuştur:

Epoch 0/19 ----- train Loss: 0.4068 Acc: 0.8219 val Loss: 0.4613 Acc: 0.7253	Epoch 5/19 ----- train Loss: 0.3442 Acc: 0.8785 val Loss: 0.1983 Acc: 0.9231	Epoch 10/19 ----- train Loss: 0.1557 Acc: 0.9595 val Loss: 0.7429 Acc: 0.7473	Epoch 15/19 ----- train Loss: 0.2225 Acc: 0.9231 val Loss: 0.7015 Acc: 0.7582
Epoch 1/19 ----- train Loss: 0.3665 Acc: 0.8381 val Loss: 1.4091 Acc: 0.5714	Epoch 6/19 ----- train Loss: 0.4590 Acc: 0.8340 val Loss: 0.1187 Acc: 0.9560	Epoch 11/19 ----- train Loss: 0.1837 Acc: 0.9312 val Loss: 0.3821 Acc: 0.8571	Epoch 16/19 ----- train Loss: 0.1565 Acc: 0.9555 val Loss: 0.3893 Acc: 0.8571
Epoch 2/19 ----- train Loss: 0.3392 Acc: 0.8219 val Loss: 1.0656 Acc: 0.6593	Epoch 7/19 ----- train Loss: 0.1757 Acc: 0.9433 val Loss: 0.4584 Acc: 0.8242	Epoch 12/19 ----- train Loss: 0.0933 Acc: 0.9676 val Loss: 0.4689 Acc: 0.8352	Epoch 17/19 ----- train Loss: 0.1639 Acc: 0.9514 val Loss: 0.4451 Acc: 0.8352
Epoch 3/19 ----- train Loss: 0.4551 Acc: 0.8381 val Loss: 2.0358 Acc: 0.6044	Epoch 8/19 ----- train Loss: 0.0996 Acc: 0.9717 val Loss: 0.4316 Acc: 0.8681	Epoch 13/19 ----- train Loss: 0.2595 Acc: 0.9271 val Loss: 0.3349 Acc: 0.8901	Epoch 18/19 ----- train Loss: 0.1927 Acc: 0.9312 val Loss: 0.5388 Acc: 0.7912
Epoch 4/19 ----- train Loss: 0.2518 Acc: 0.8988 val Loss: 0.3546 Acc: 0.8462	Epoch 9/19 ----- train Loss: 0.1726 Acc: 0.9393 val Loss: 0.2928 Acc: 0.8901	Epoch 14/19 ----- train Loss: 0.2655 Acc: 0.8988 val Loss: 0.5817 Acc: 0.7692	Epoch 19/19 ----- train Loss: 0.1550 Acc: 0.9352 val Loss: 0.7061 Acc: 0.7473
Training complete in 26m 48s Best val Acc: 0.956044			

Training ve validation doğruluk oranlarının epoch bazında değişimi Şekil1’de, kayıpların epoch bazında değişimi Şekil 2’de gösterilmiştir:



Son olarak, modelin bazı tahminlerine yer verilmiştir:



YORUMLAR:

Epoch bazında doğruluk oranlarının değişimine bakıldığında; 4, 5 ve 6. epochlarda validation doğruluğunun bir sıçramasını gözlemliyoruz. Daha sonra tekrar eğitim doğruluğunun altına düşmüş ve 75% ile 90% arasında dalgalanma yaşamış. Aynı süreçte eğitim doğruluğu ile arasındaki ortalama 10%'luk fark sürekli olarak korunmuş. Aynı şekilde kayıplar arasında da korunan 0.05 puanlık bir fark var.

Eğitimin bazı noktalarında grafiklerin birbirine yaklaştığı durumlar da mevcut. Fakat akabinde büyük açılmalar gözlemleniyor.

Dolayısıyla seçilen adım boyutu anlamsız olmamakla birlikte, daha küçük bir adım boyutu belirlense model daha iyi eğitilebilir. Epoch sayısı arttırılırsa da model iyileşme eğilimi gösterecektir. Veri setinin de arttırılması aynı şekilde faydalı olacaktır.

Modelin rastgele seçip kendini denediği 6 fotoğrafta hata yapmadığı gözlemlenmektedir. Fakat 4. fotoğraf, kullanıcı inisiyatifine bağlı şekilde “sağlam” olarak labellanabilirdi. Bu durumda model yanlış yapmış olacaktı. Diğer “çürük” tahminlerine bakılacak olursa, modelin “ortasında beyaz leke olan çürüktür” gibi bir öğrenme eğiliminde bulunduğu söylenebilir. Böyle bir durumda veri seti, üstünde beyaz leke olmayan ve çürük olan domateslerle zenginleştirilebilir.

Mert Ünal