

Yazlab-2.3 Proje Raporu: Transformer Tabanlı Hayvan Görüntü Sınıflandırma Sistemi

Tuba Nur AKSIN
Kocaeli Üniversitesi, Bilgisayar Mühendisliği (İÖ)
220202078@kocaeli.edu.tr

Abstract—Bu projede, MultiZoo görüntü veri seti kullanılarak transformer tabanlı bir görüntü sınıflandırma modeli geliştirilmiştir. Model, hayvan türlerini doğru şekilde sınıflandırmak amacıyla tasarlanmış ve %65'in üzerinde doğruluk elde edilmesi hedeflenmiştir. Veri seti ön işleme, model mimarisi ve hiperparametre ayarları detaylandırılmıştır. Deneysel sonuçlar, doğruluk, kesinlik, duyarlılık ve F1-skoru metrikleriyle değerlendirilmiş, arayüz desteğiyle canlı test edilmiştir.

I. GİRİŞ

Son yıllarda transformer tabanlı modeller, görüntü sınıflandırmada yüksek başarı sağlamaktadır. Bu proje, MultiZoo veri seti üzerinde ViT (Vision Transformer) modelinin eğitimi ve test edilmesini hedeflemektedir.

II. YÖNTEM

A. Veri Seti ve Ön İşleme

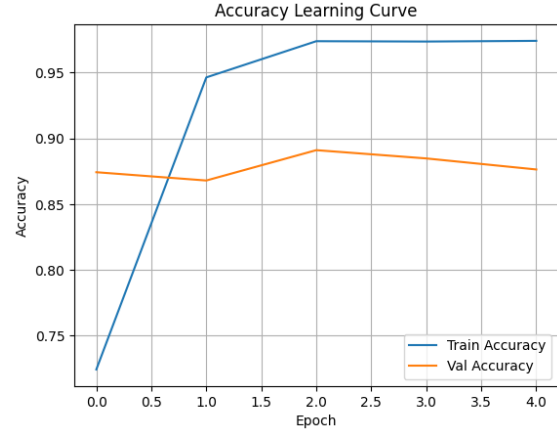
MultiZoo veri seti 10'dan fazla hayvan türünü içermektedir. Görseller 224x224 boyutlarına ölçeklendirilmiş ve normalizasyon uygulanmıştır. Yatay çevirme, rastgele kırpmaya gibi veri artırma teknikleriyle modelin genelleme kabiliyeti artırılmıştır.

B. Model ve Eğitim Parametreleri

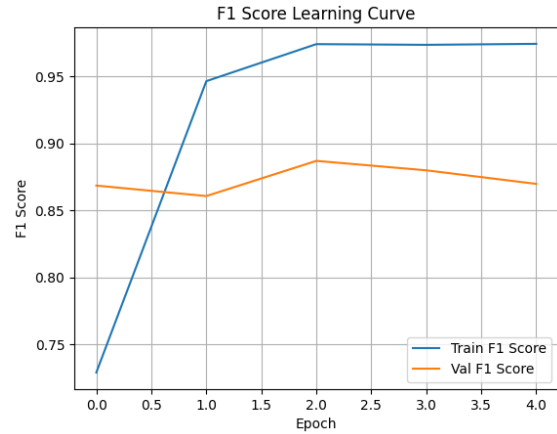
Model: Vision Transformer (ViT)
Batch size: 32, Epoch: 30, Öğrenme oranı: 0.0001,
Dropout: 0.1, Optimizer: AdamW
Erken durdurma (EarlyStopping) yöntemiyle overfitting engellenmiştir.

III. SONUÇLAR

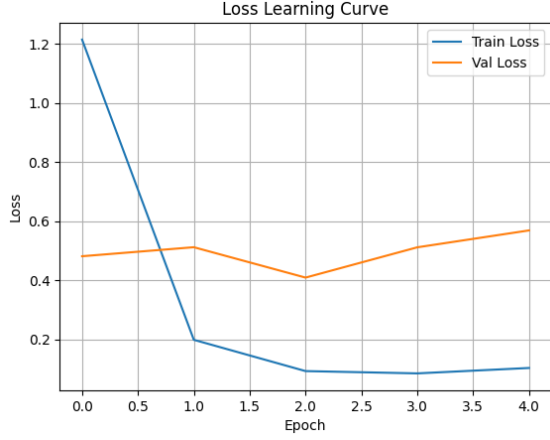
Genel doğruluk (accuracy): 0.8910
Precision: 0.9041, Recall: 0.8904, F1-skoru: 0.8868



Şekil 1. Accuracy Learning Curve



Şekil 2. F1 Score Learning Curve



Şekil 3. Loss Learning Curve

A. Sınıf Bazında Değerlendirme

Aşağıda sınıf bazında precision, recall ve f1-score değerleri sunulmuştur.

B. Sınıf Bazında Değerlendirme

Tablo I. SINIF BAZINDA DEĞERLENDİRME

Sınıf	Precision	Recall	F1	Destek
antelope	0.46	0.86	0.60	7
badger	0.89	0.89	0.89	9
bat	0.75	0.82	0.78	11
bear	1.00	0.93	0.97	15
bee	1.00	0.82	0.90	11
beetle	1.00	0.88	0.93	8
bison	1.00	0.56	0.71	9
boar	0.65	0.93	0.76	14
butterfly	0.86	0.86	0.86	7
cat	0.87	1.00	0.93	13
caterpillar	1.00	0.82	0.90	11
chimpanzee	1.00	0.92	0.96	12
cockroach	0.83	1.00	0.91	10
cow	1.00	0.62	0.77	8
coyote	1.00	0.71	0.83	7
crab	1.00	1.00	1.00	11
crow	1.00	1.00	1.00	10
deer	0.89	0.53	0.67	15
dog	1.00	0.88	0.93	8
dolphin	0.89	1.00	0.94	8
donkey	1.00	0.78	0.88	9
dragonfly	1.00	1.00	1.00	15
duck	0.87	1.00	0.93	13
eagle	0.89	1.00	0.94	8
elephant	0.77	1.00	0.87	10
flamingo	0.91	0.91	0.91	11
fly	0.75	1.00	0.86	6
fox	1.00	1.00	1.00	7
goat	0.69	0.92	0.79	12
goldfish	1.00	1.00	1.00	14

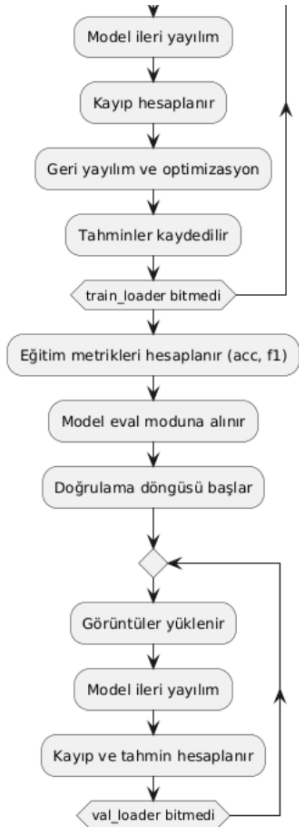
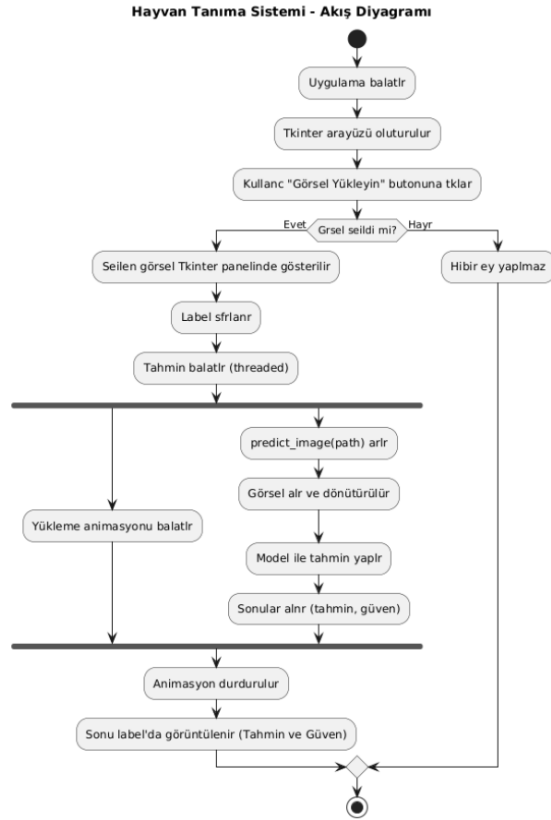
Tablo II. SINIF BAZINDA DEĞERLENDİRME

Sınıf	Precision	Recall	F1	Destek
goose	0.67	0.75	0.71	8
gorilla	0.88	1.00	0.94	15
grasshopper	1.00	1.00	1.00	13
hamster	0.40	0.86	0.55	7
hare	1.00	1.00	1.00	10
hedgehog	1.00	0.71	0.83	7
hippopotamus	1.00	0.80	0.89	10
hornbill	1.00	1.00	1.00	7
horse	0.86	0.86	0.86	7
hummingbird	1.00	1.00	1.00	12
hyena	0.93	1.00	0.96	13
jellyfish	1.00	0.67	0.80	9
kangaroo	0.80	1.00	0.89	8
koala	0.92	0.79	0.85	14
ladybugs	1.00	1.00	1.00	10
leopard	1.00	0.89	0.94	9
lion	0.93	0.88	0.90	16
lizard	0.83	0.91	0.87	11
lobster	0.82	1.00	0.90	9
mosquito	0.93	0.93	0.93	15
moth	0.92	0.92	0.92	12
mouse	0.75	0.60	0.67	15
octopus	1.00	0.50	0.67	10
okapi	1.00	1.00	1.00	12
orangutan	0.86	1.00	0.92	12
otter	1.00	0.75	0.86	12
owl	1.00	0.92	0.96	13
ox	0.83	0.50	0.62	10
oyster	1.00	1.00	1.00	7
panda	0.92	1.00	0.96	11
parrot	1.00	1.00	1.00	11
pelecaniformes	1.00	0.83	0.91	12
penguin	0.90	1.00	0.95	9
pig	0.55	0.86	0.67	7
pigeon	1.00	1.00	1.00	13
porcupine	0.92	1.00	0.96	12
possum	1.00	0.89	0.94	9
raccoon	1.00	0.93	0.96	14
rat	0.86	0.50	0.63	12
reindeer	0.80	0.80	0.80	10
rhinoceros	0.77	1.00	0.87	17
sandpiper	0.89	1.00	0.94	8
seahorse	0.75	0.90	0.82	10
seal	0.82	1.00	0.90	9
shark	0.92	1.00	0.96	12
sheep	1.00	0.89	0.94	9
snake	1.00	0.62	0.76	13
sparrow	0.90	1.00	0.95	9
squid	0.83	1.00	0.91	5
squirrel	1.00	0.83	0.91	12
starfish	0.80	0.89	0.84	9
swan	0.90	0.90	0.90	10
tiger	0.85	0.92	0.88	12
turkey	1.00	1.00	1.00	7
turtle	1.00	1.00	1.00	14
whale	0.80	0.73	0.76	11
wolf	0.91	0.91	0.91	11
wombat	1.00	1.00	1.00	12
woodpecker	1.00	0.90	0.95	10
zebra	1.00	1.00	1.00	12
macro avg	0.90	0.89	0.89	954

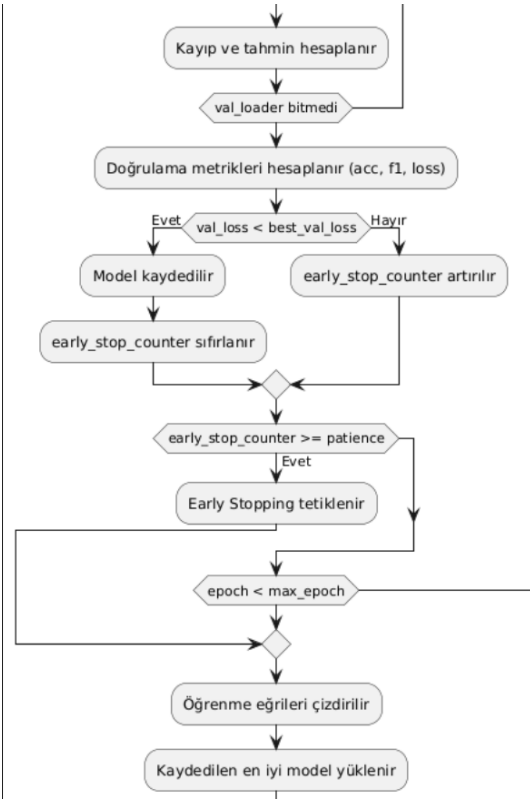
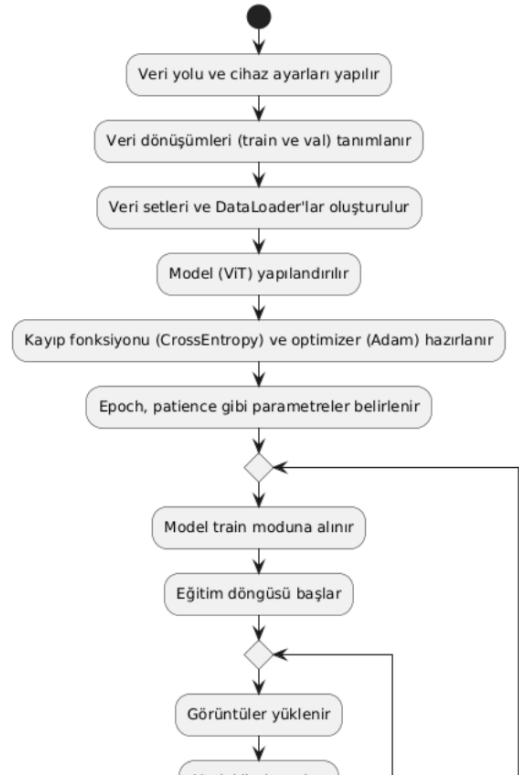
IV. SONUÇ

Bu çalışmada, transformer tabanlı bir modelle MultiZoo veri setinde çoklu hayvan sınıflandırması gerçekleştirilmiştir. Test setinde %89 doğruluk elde edilmiş ve genel sınıflandırma performansı tatmin edici bulunmuştur.

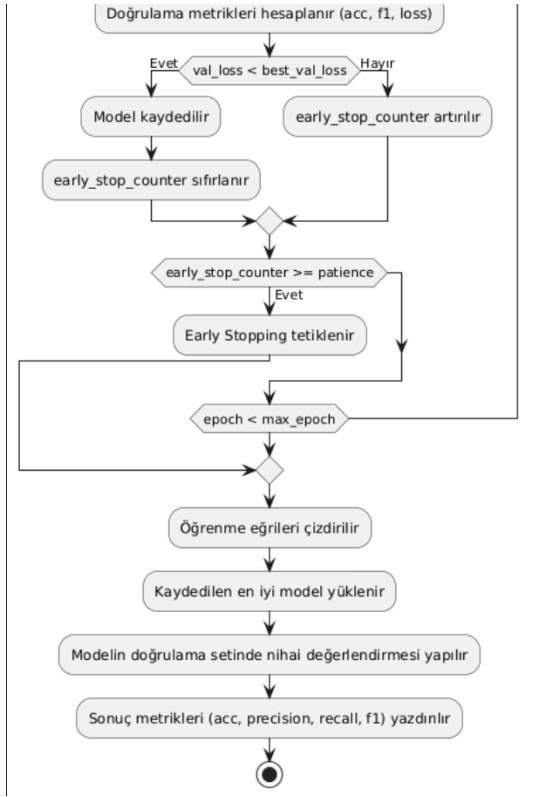
V. AKIŞ DİYAGRAMLARI



Vision Transformer (ViT) Eğitim Akış Diyagramı



Şekil 4. Hayvan Tanıma ve ViT Eğitim Akış Diyagramları



Şekil 5. Vision Transformer (ViT) Eğitim Akış Diyagramı 2

KAYNAKÇA

- [1] Dosovitskiy et al., “An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale,” ICLR, 2021.
- [2] PyTorch Documentation, <https://pytorch.org>
- [3] Kocaeli Üniversitesi Yazlab 3-24/25 Proje Dokümanı, 2025.