

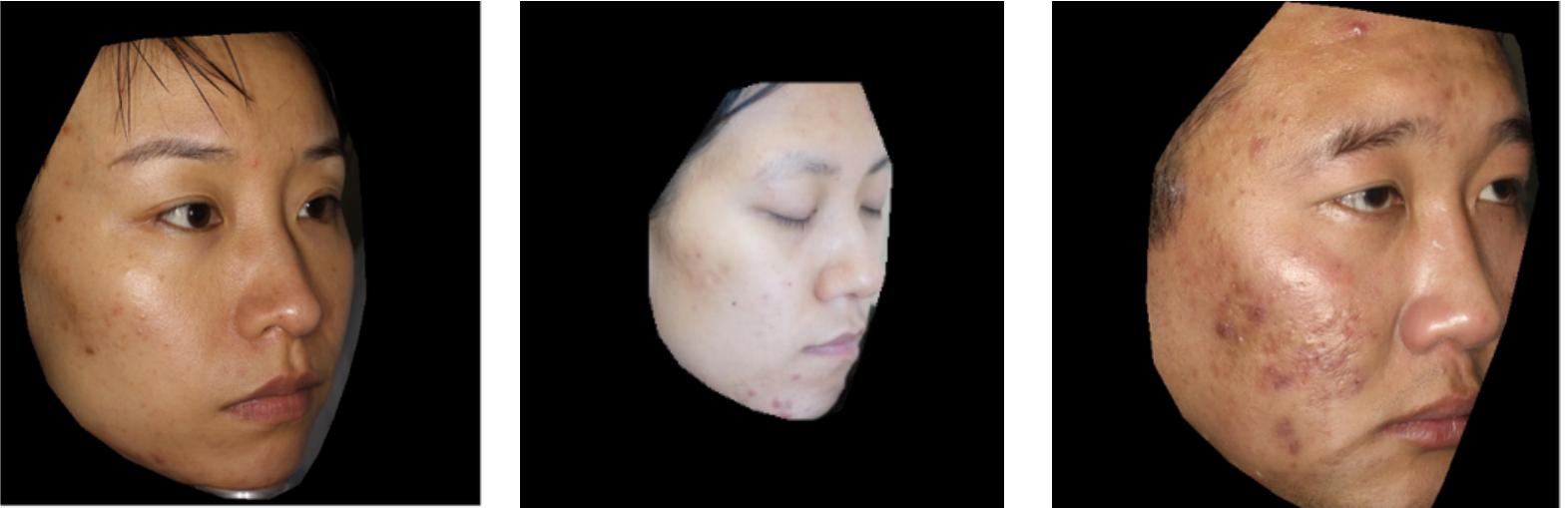
ACNE SEVERITY CLASSIFICATION WITH VISION TRANSFORMER (ViT)

SENA YILDIZ

DATASET TANIMI

Kaggle Acne Grading Classification Dataset

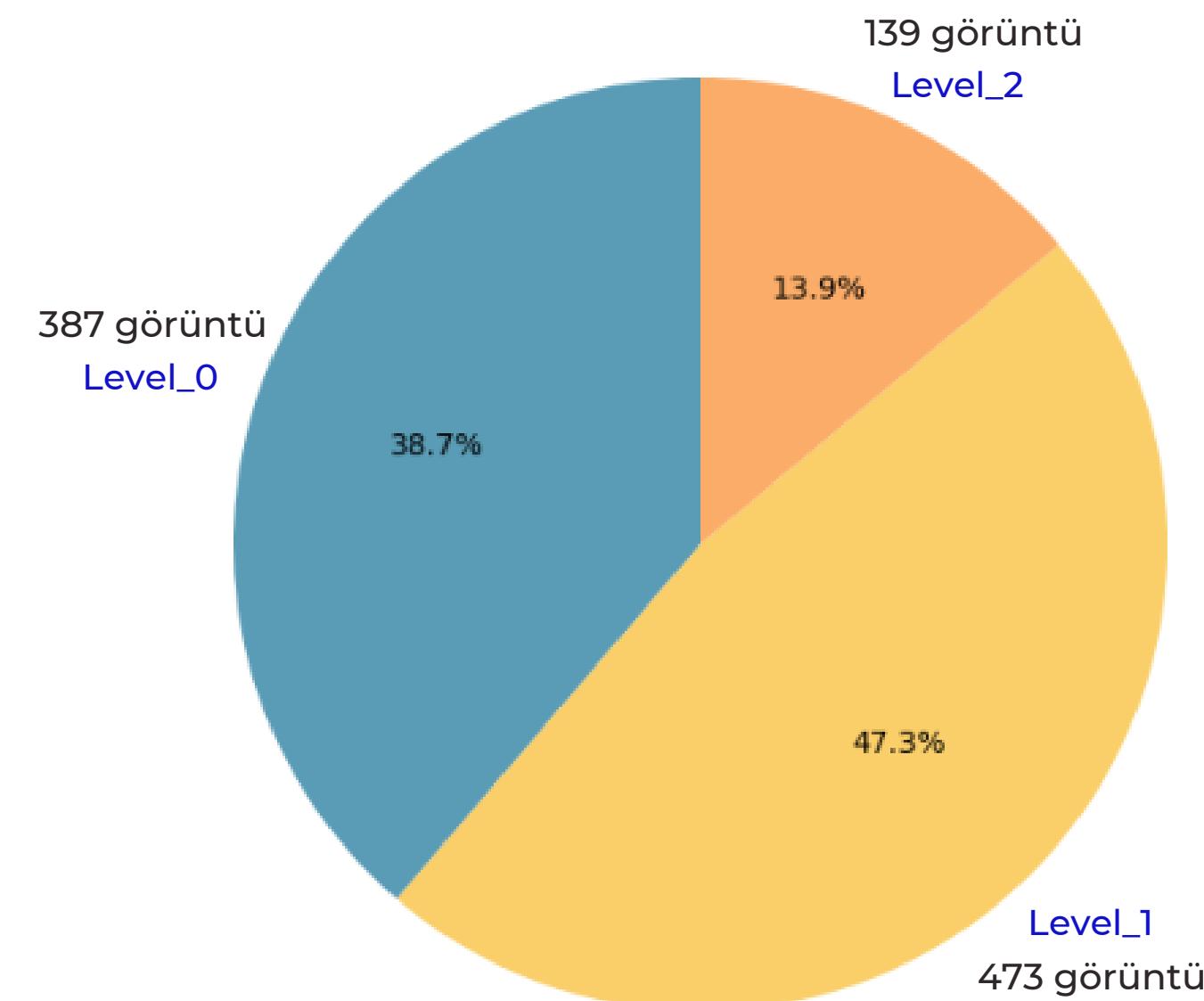
Veri seti, RGB frontal yüz görüntülerinden oluşan ve akne şiddetini Level_0, Level_1, Level_2 olmak üzere üç sınıfta ele alan çok sınıflı bir görüntü sınıflandırma problemidir. Toplam 999 görüntü bulunmaktadır. Level_2 sınıfının verinin yalnızca %13,9'unu oluşturması, nadir ve ağır vakaların öğrenilmesini zorlaştıran belirgin bir sınıf dengesizliği problemine işaret etmektedir.



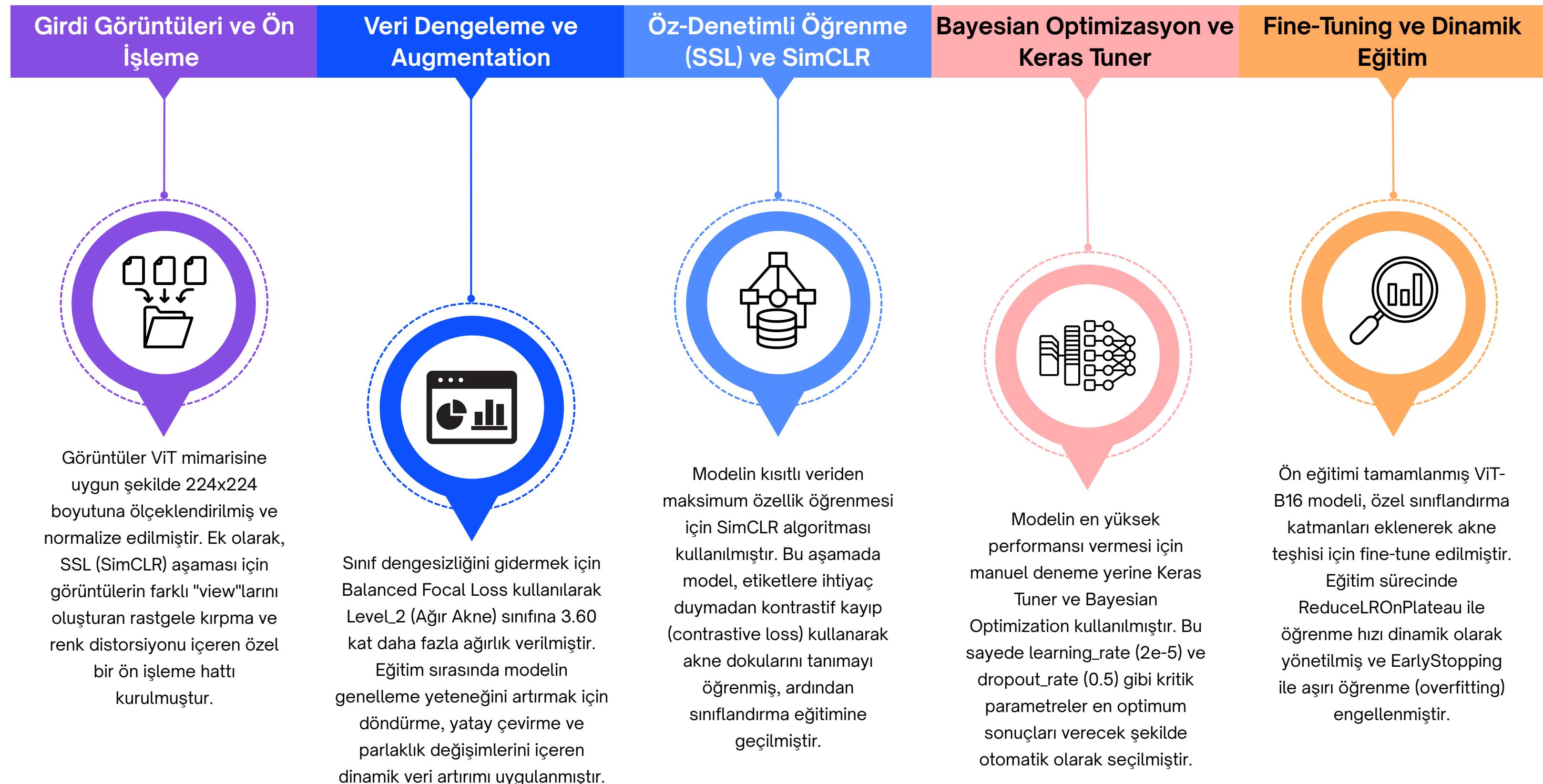
Level_0

Level_1

Level_2



UYGULANAN YÖNTEMLER



MODEL

Vision Transformer (ViT-B16)

Model olarak Google tarafından geliştirilen vit_b16 (Base 16) mimarisi temel alınmıştır.

SSL ve SimCLR Entegrasyonu

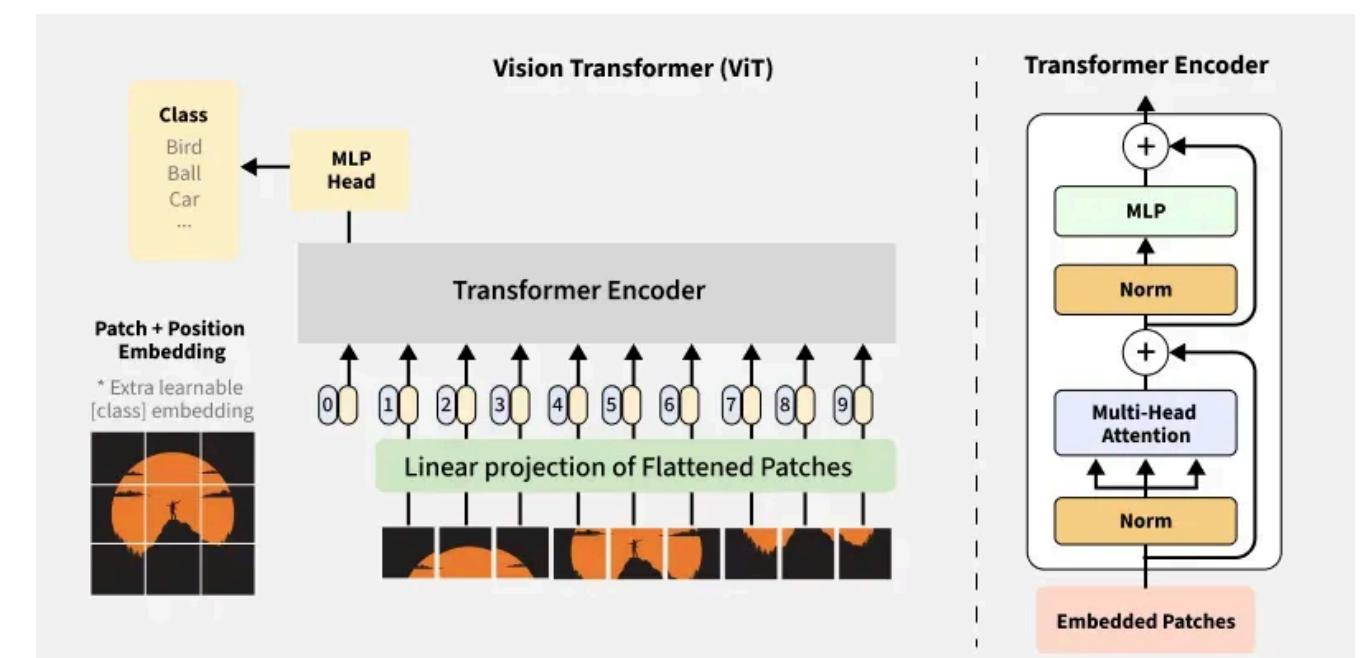
Self-Supervised Learning (SSL); sınırlı veriyi daha verimli kullanmak adına, etiketli eğitimden önce modelin dokuları kendi başına öğrenmesi sağlanmıştır. SimCLR Mimarisi, görüntülerin farklı varyasyonları (kırılmış, döndürülmüş) arasındaki benzerliği ölçen bir "Projection Head" katmanı eklenerek kontrastif öğrenme (Contrastive Learning) uygulanmıştır.

Classification Head

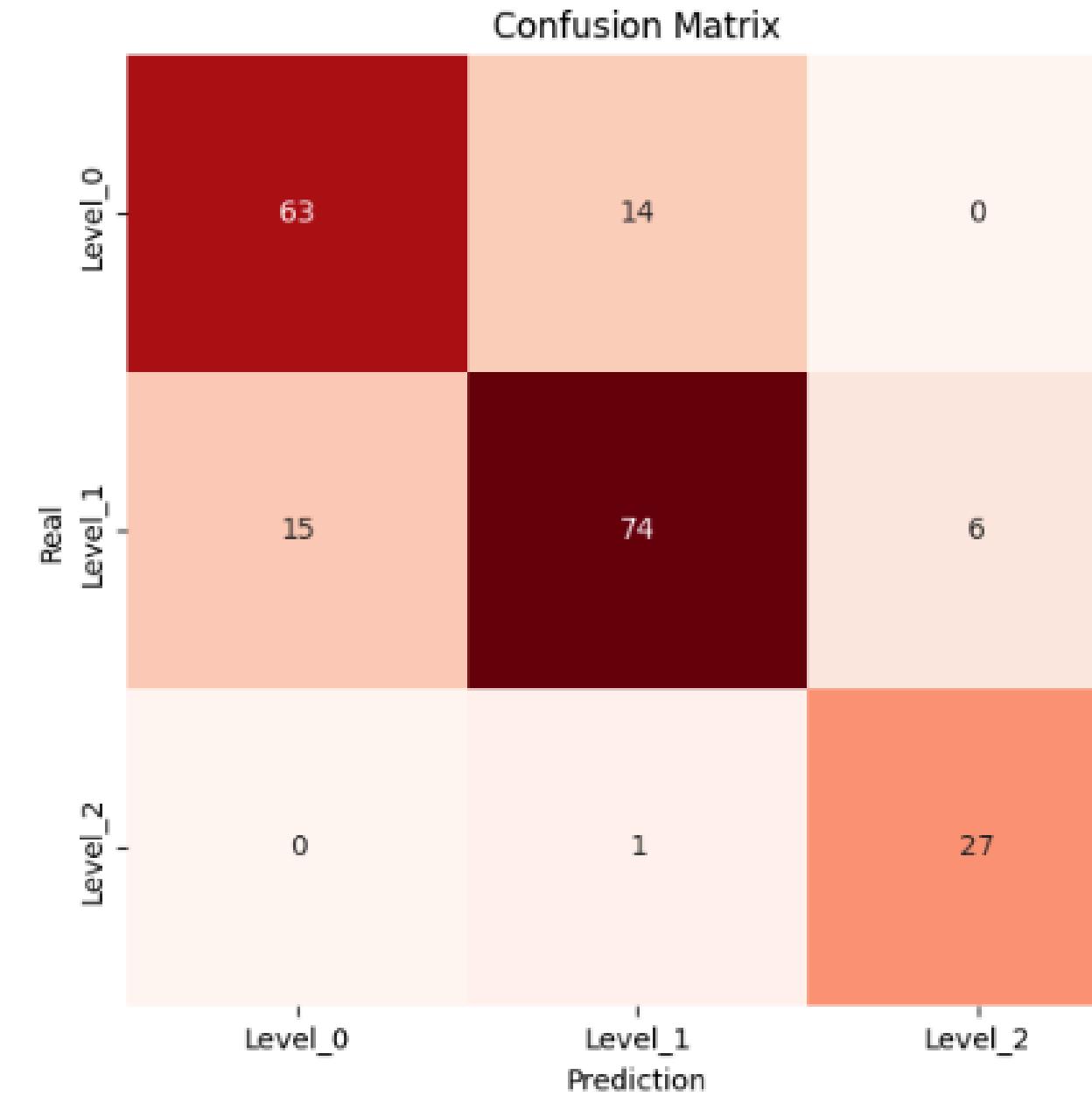
Kodda belirlenen en iyi parametrelerle şu katmanlar eklenmiştir: Global Average Pooling (1D), Dropout (0.5), Dense Layer (Level 0, 1, 2) ve Softmax katmanı.

Hiperparametre ve Optimizasyon

Optimizer olarak AdamW, Balanced Focal Loss, ve ReduceLROnPlateau kullanılmıştır.



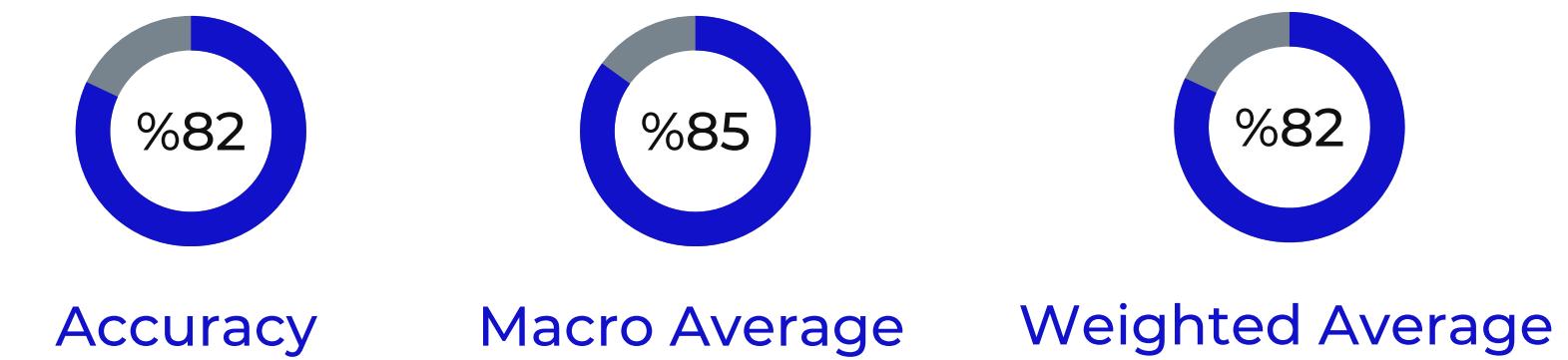
MODEL PERFORMANSI



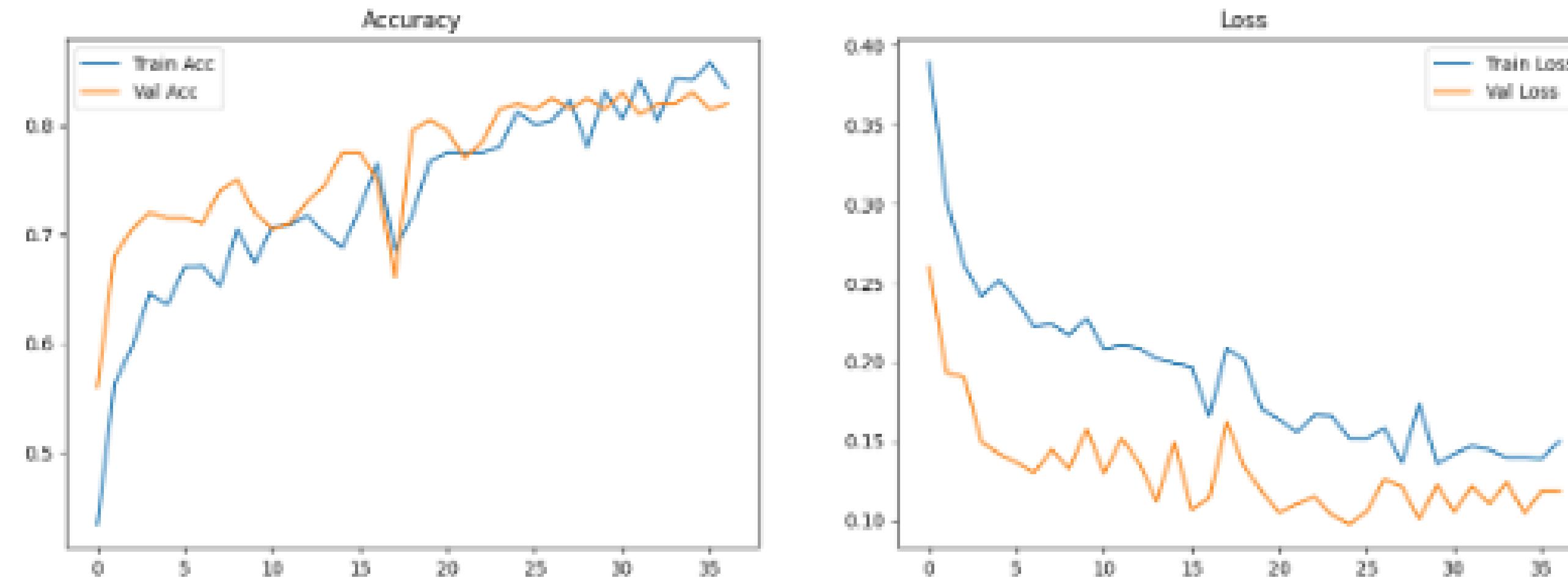
Sınıf Bazlı Performans

Sınıf	Precision	Recall	F1-score
Level_0	0.81	0.82	0.81
Level_1	0.83	0.78	0.80
Level_2	0.82	0.96	0.89

Genel Bazlı Performans



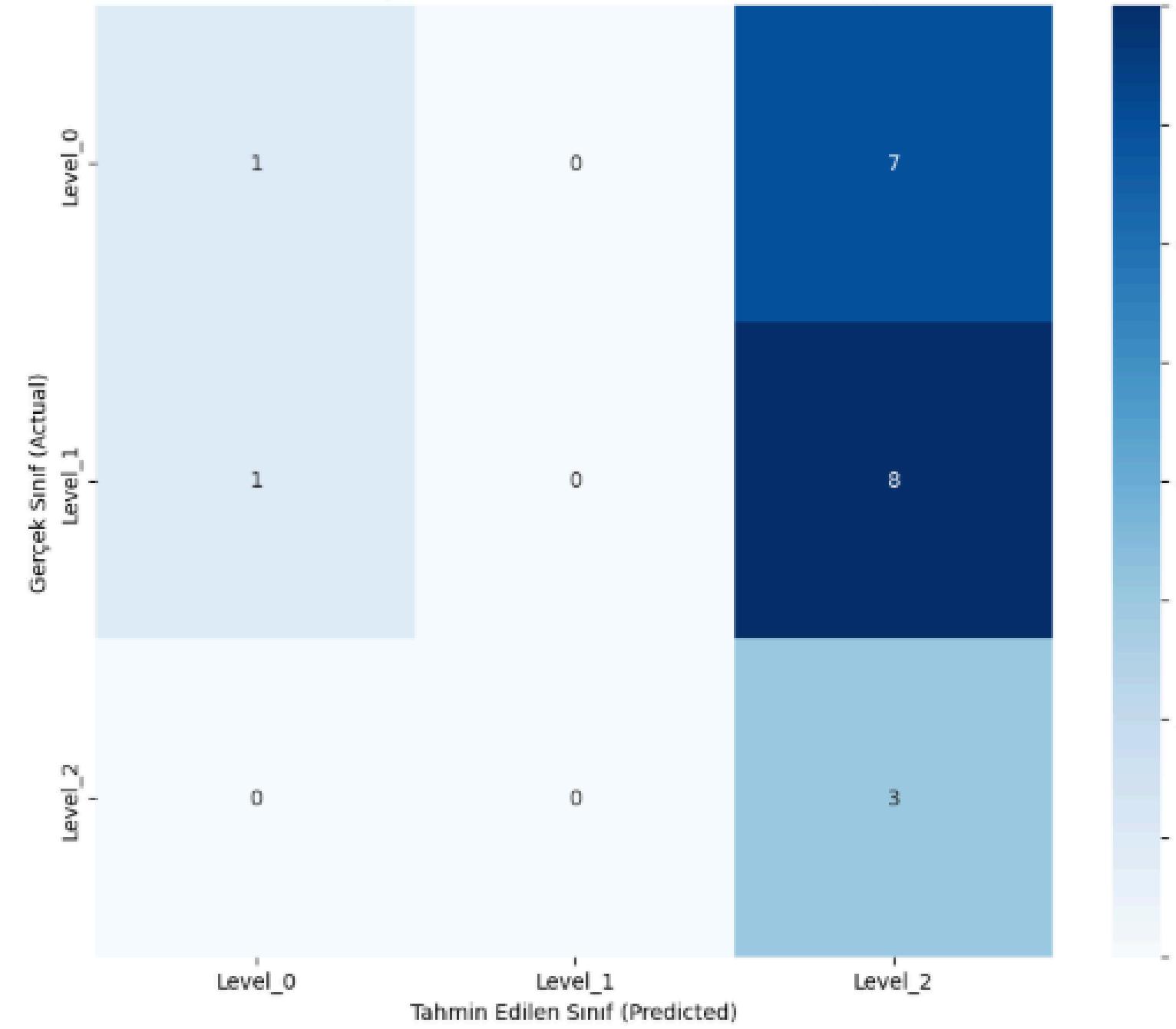
GENEL DEĞERLENDİRME



Vision Transformer (ViT) mimarisi, akne lezyonları arasındaki dokusal ilişkileri CNN tabanlı modellere kıyasla daha hassas yakalamıştır. Balanced Focal Loss ve Mixup kullanımı sayesinde en zor sınıf olan ağır vakalar (Level_2) %93 recall ile başarıyla tespit edilmiştir. Model, doğrulama setinde %81 genel doğruluk elde ederek klinik karar destek sistemleri için güçlü bir temel sunmuştur.

SSL MODEL PERFORMANSI

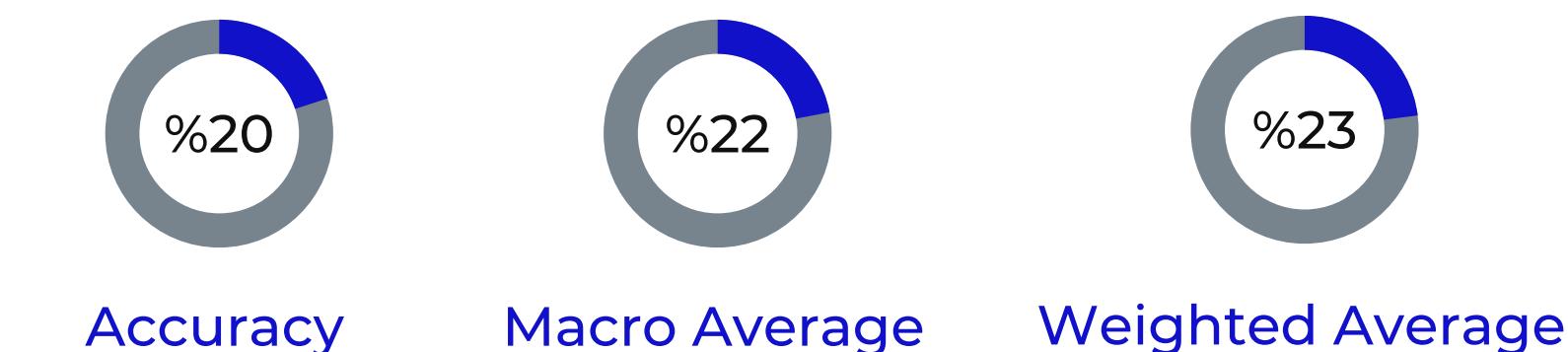
Karmaşık Matrisi (Confusion Matrix)



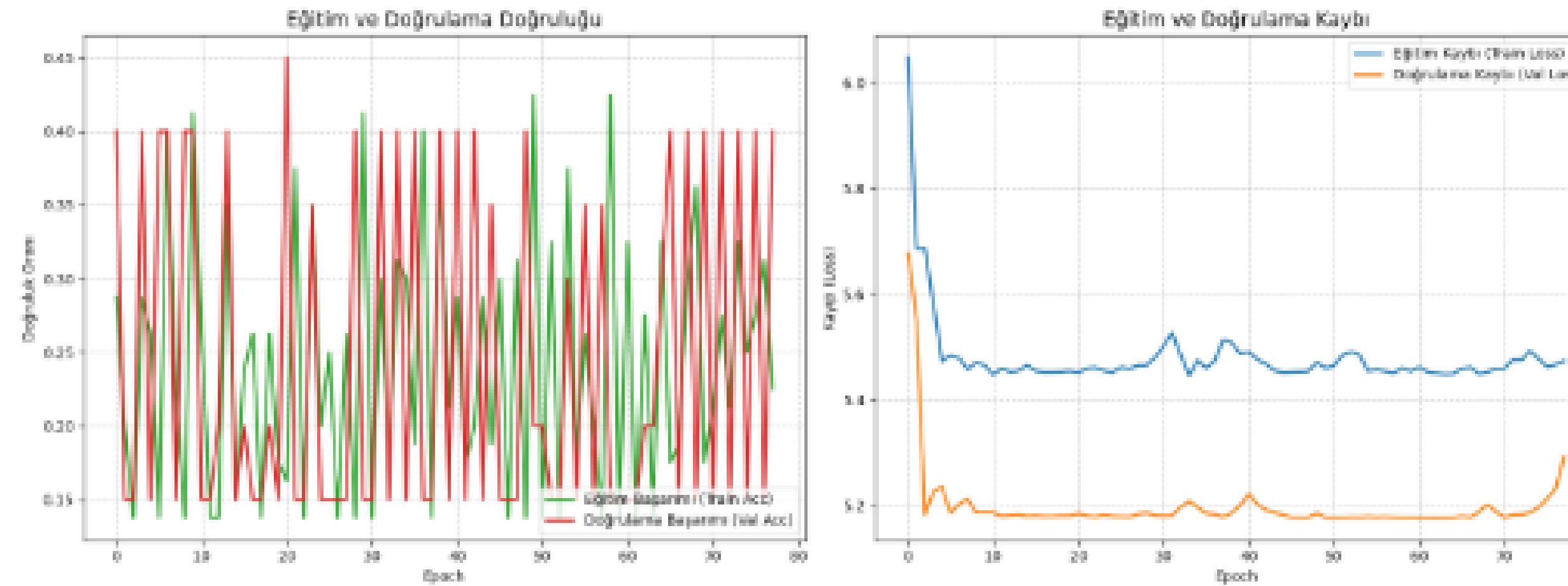
Sınıf Bazlı Performans

Sınıf	Precision	Recall	F1-score
Level_0	0.50	0.12	0.20
Level_1	0.80	0.00	0.00
Level_2	0.17	1.00	0.29

Genel Bazlı Performans



SSL GENEL DEĞERLENDİRME



SSL aşamasında SimCLR ile akne dokularına ait ayırt edici temsillerin öğrenilmesi amaçlanmış; ancak doğrulama başarısı %20 seviyesinde kalmıştır. Bunun başlıca nedenleri, SSL için gerekli geniş ölçekli etiketlenmemiş verinin sınırlı olması, kontrastif öğrenmenin yüksek epoch ve hesaplama maliyeti gerektirmesi ve şiddetli akne (Level_2) sınıfının dengesiz dağılımıdır.

SONUÇ

Bu çalışmada, akne şiddeti sınıflandırmasında Vision Transformer (ViT) mimarisi standart eğitim ve Öz-denetimli Öğrenme (SSL) yaklaşımılarıyla karşılaştırılmıştır. Bayesian optimizasyonu ve Balanced Focal Loss ile eğitilen ViT-B16 modeli, %82 genel doğruluk ve kritik Level_2 sınıfında %96 recall elde ederek klinik kullanım için güçlü bir performans göstermiştir. Buna karşılık, SimCLR tabanlı SSL yaklaşımı veri ve eğitim kısıtları nedeniyle %20 doğrulukta kalmış; ancak daha büyük etiketlenmemiş veri setleriyle genelleme potansiyeli sunduğu görülmüştür. Geliştirilen model, nadir ağır vakaları yüksek hassasiyetle tespit ederek dermatologlar için etkili bir karar destek aracı olma potansiyeli taşımaktadır.

Model	Accuracy	Recall
ViT	0.82	0.12
ViT + SSL	0.20	0.100



TEŞEKKÜRLER