

# **Ham Piksel Tabanlı Yüz İfadesi Tanıma CNN ile 8 Sınıflı Duygu Sınıflandırması**

**Öğrenci:** YasirÖmer ALPARSLAN

**Danışman:** Öğretim Üyesi

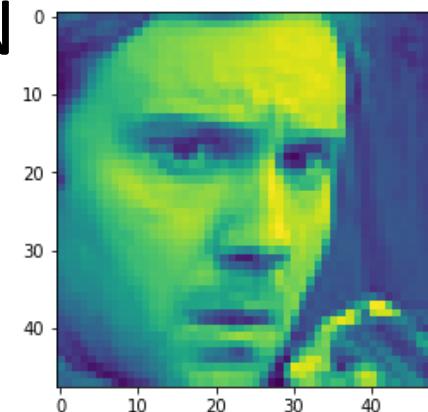
Dr. SELİM YILMAZ

**MUĞLA ÜNİVERSİTESİ/ FEN BİLİMLERİ ENSTİTÜSÜ / Bölüm:YZ**

**Tarih:** 21-12-2025

# Problem Tanımı ve Motivasyon

- Yüz ifadesi tanımanın İnsan–bilgisayar etkileşimi HCI, güvenlik, sağlık ve eğitim gibi alanlarda yüz ifadesi tanımanın önemi
- Klasik yöntemler: HOG, LBP, SIFT + SVM (manuel öznitelik çıkarımı)
- Bu çalışmada yalnızca ham pikseller ve Tamamen derin öğrenme tabanlı CNN kullanılmıştır



# Çalışmanın Amacı

- 48×48 gri seviye yüz görüntülerinden duygusal tanıma
- 8 sınıf **8 duygusal sınıfını** tahmin etmek:  
Anger, Contempt, Disgust, Fear, Happiness, Neutral, Sadness, Surprise  
(Öfke, Küçümseme, Tiksinti, Korku, Mutluluk, Tarafsızlık, Üzüntü, Şaşkınlık)



- Uçtan uca (end-to-end), hafif ve tekrar üretilen CNN modeli.

Modelin:

- Yeterli doğrulukta (%55 ≈ test doğruluğu)
- Hafif ve tekrar üretilen olması (Keras kodu + kayıtlı ağırlıklar)

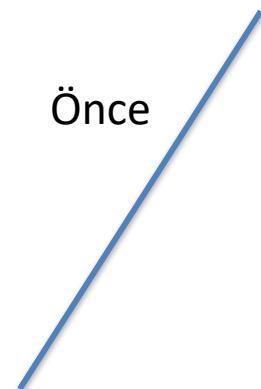
# Veri Seti

- $48 \times 48$  boyutunda gri seviye yüz görüntüleri
- Klasör yapısı: train/ (0–7) $\rightarrow$ 8 , test/ (0–7) $\rightarrow$  8
- Test seti: toplam 400 görüntü



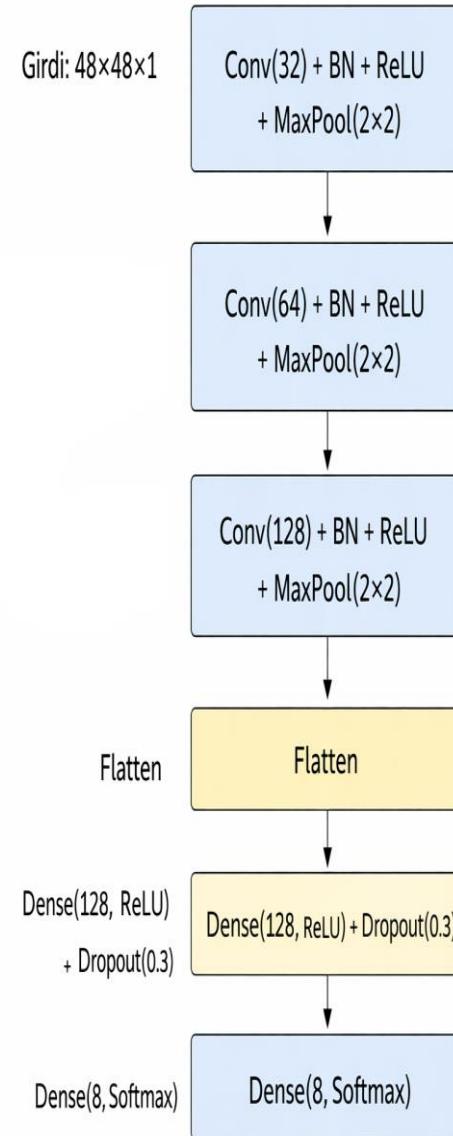
# Ön İşleme ve Veri Artırma

- Normalizasyon:  $0-255 \rightarrow 0-1$
- One-hot label encoding (8 sınıf)
- %80 eğitim – %20 doğrulama
- Data augmentation(Veri büyütme):  
döndürme, kaydırma, horizontal flip
- Çok küçük ve gri tonlamalı görüntüler, bu da  
görevi zorlaştırır ancak modelin boyutunu  
küçültür.



# CNN Mimarisi

- Girdi:  $48 \times 48 \times 1$
- Conv(32) Dönüşüm → Conv(64) Dönüşüm → Conv(128) Dönüşüm + MaxPool
- Dense(128) MaksHavuzYoğun + Dropout(0.3) Eksik
- Çıkış: Dense\* Yoğun\*(8, Softmax)

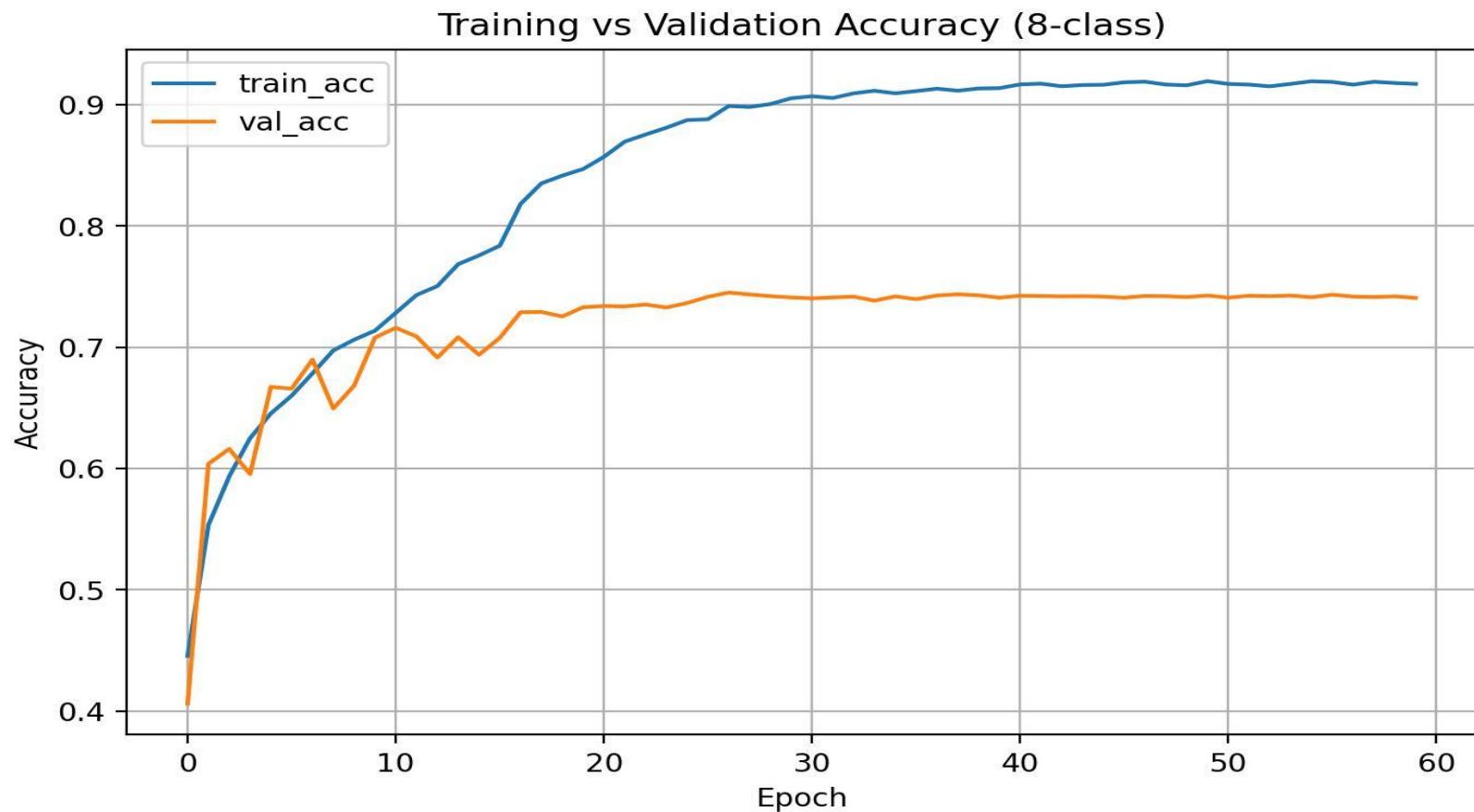


# Eğitim Kurulumu

- Batch size: 64, Epoch: 60,45,360. Optimizer: Adam ( $\text{lr} = 1\text{e-}3$ )  
**(EarlyStopping yok, tüm epoch'lar tamamlandı)**
- ModelCheckpoint      **best\_emotion\_cnn\_8class.keras**
- ve Reduce LROn Plateau    **val\_loss** göre öğrenme oranını düşürdü bunları ikisini kullanıldı.
- **ResNet veya transfer** öğrenme yöntemlerini kullanmadı; bunun yerine, sıfırdan oluşturulmuş hafif bir model kullanıldı.

**Bu adımlar, modelin sadece tren sayısını azaltmakla kalmayıp genelleme yapmasına da yardımcı olur.**

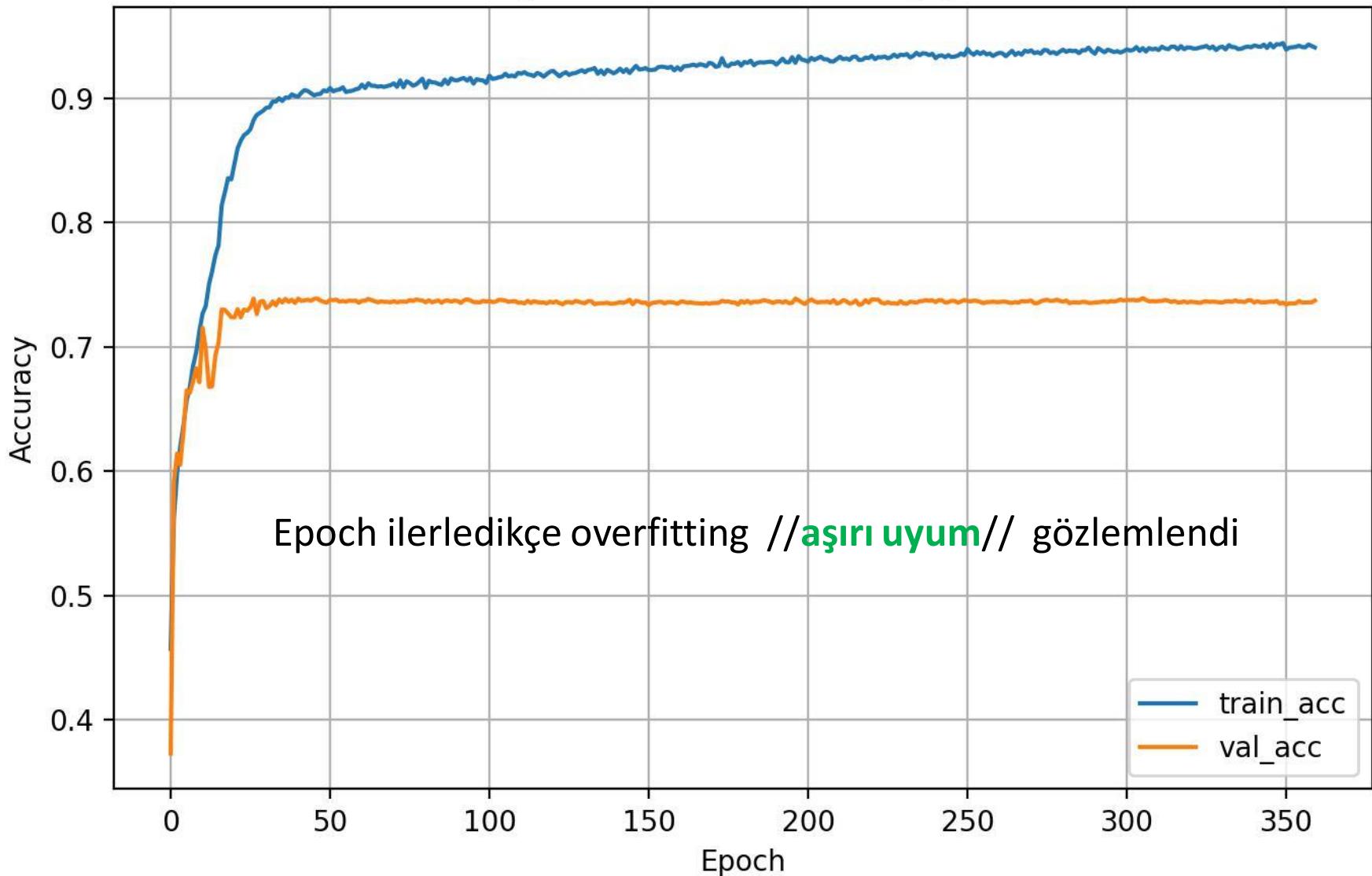
# Eğitim ve Doğrulama Doğruluğu



- Training accuracy: %43 → %90+
- Validation accuracy //Doğrulama doğruluğu//: %40 → %75
- Epoch ilerledikçe overfitting// aşırı uyum // gözlemlendi

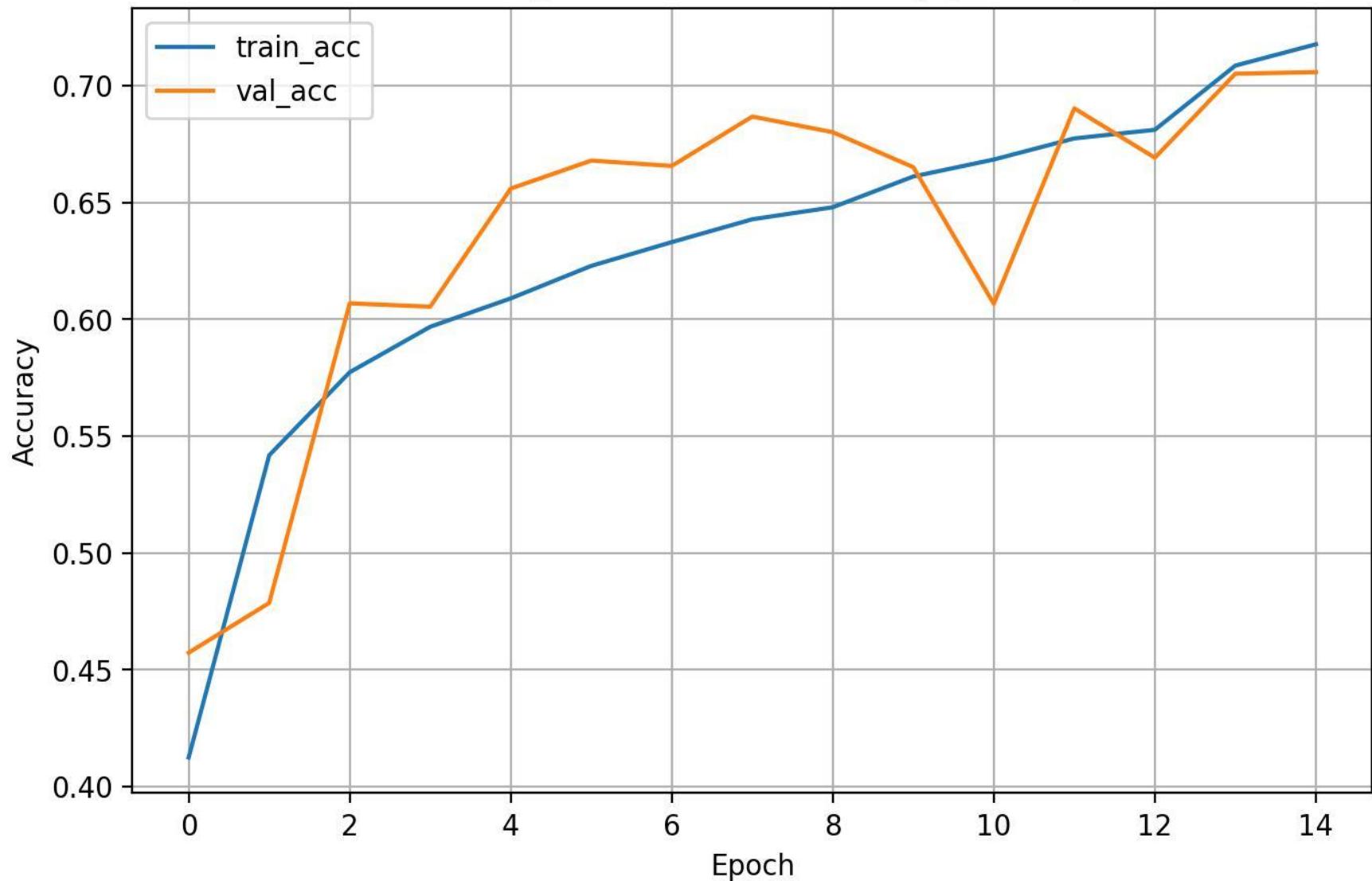
# Eğitim ve Doğrulama Doğruluğu

Training vs Validation Accuracy (8-class)



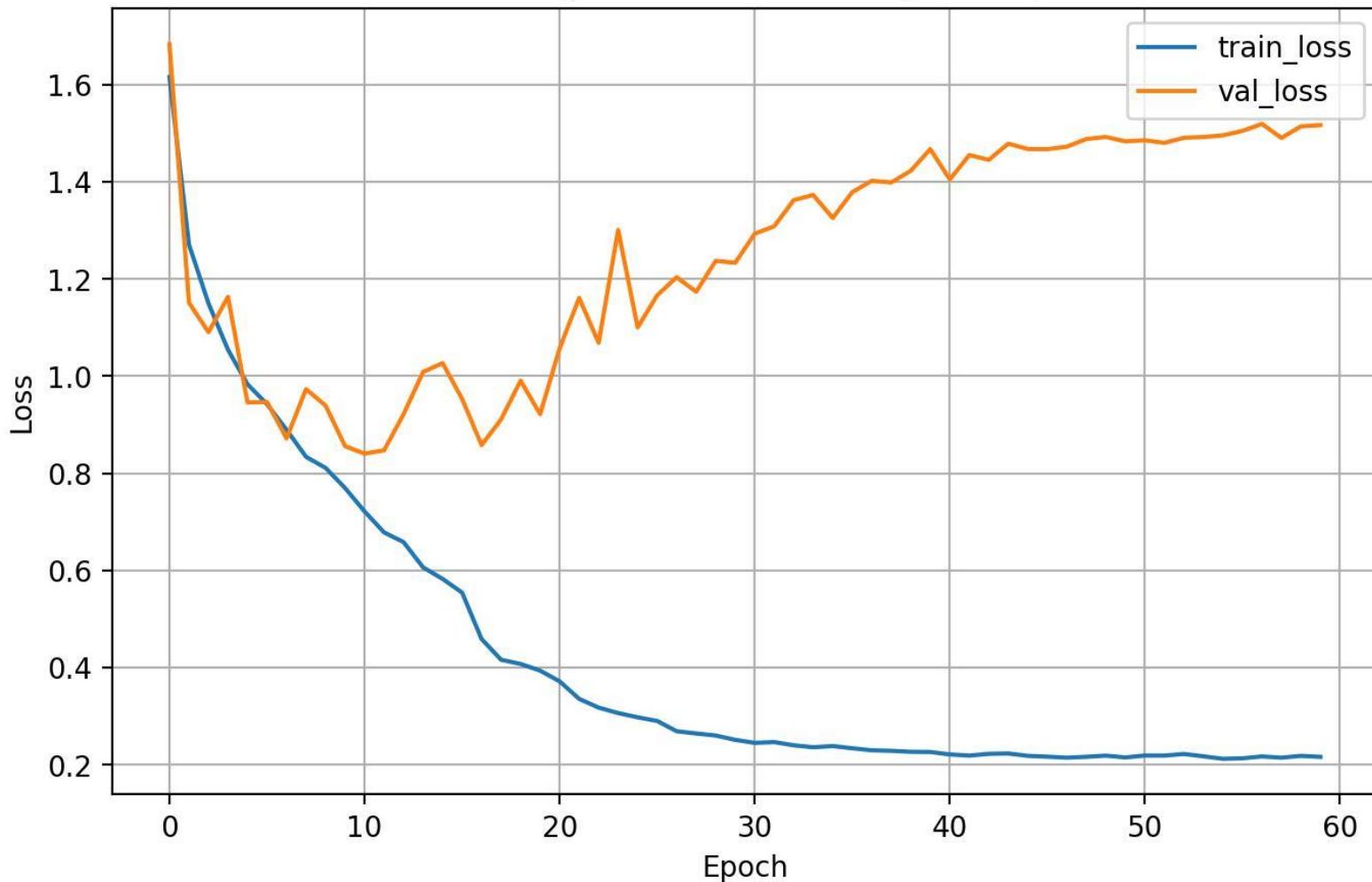
# Eğitim ve Doğrulama Doğruluğu

Training vs Validation Accuracy (8-class)



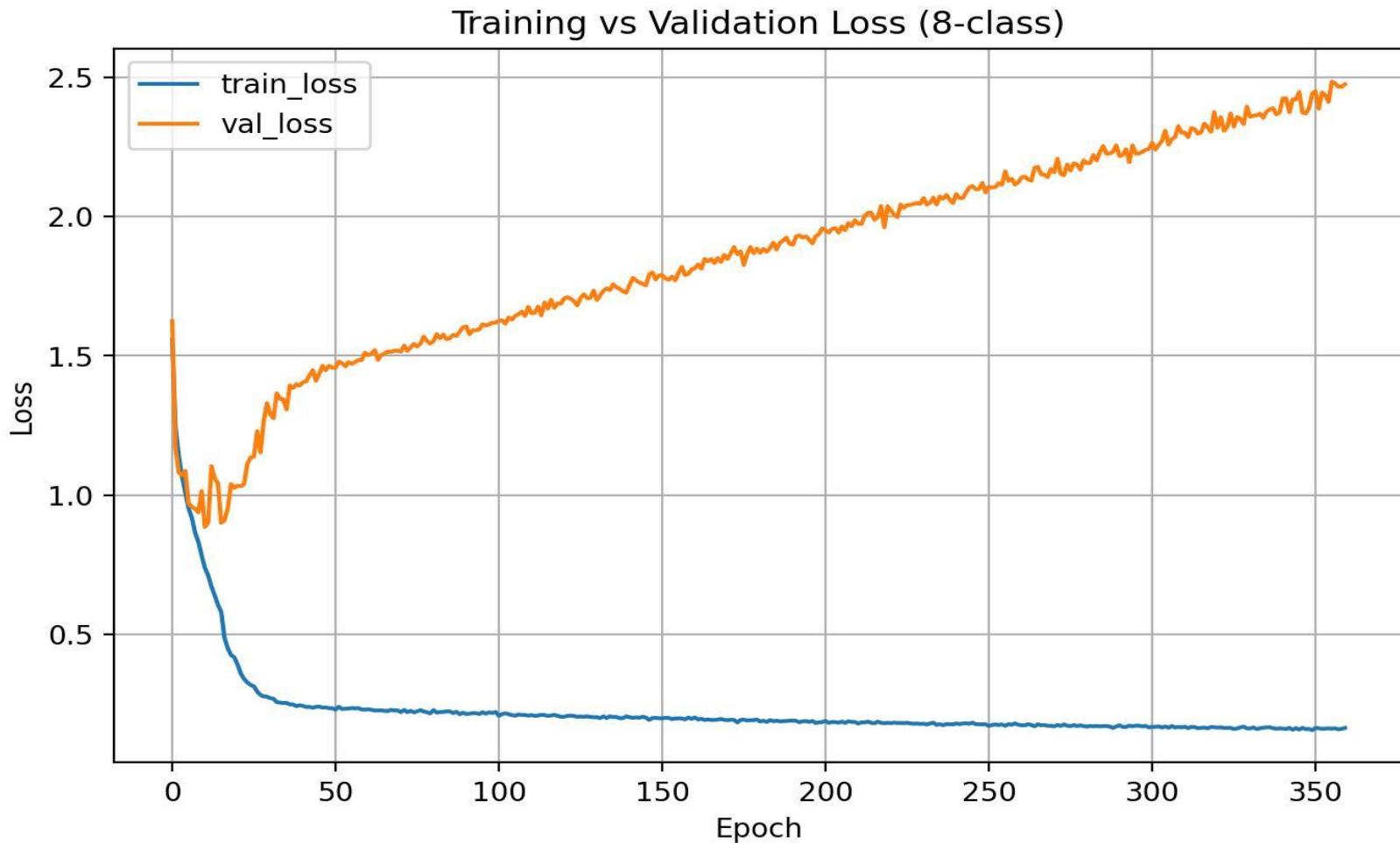
# Eğitim ve Doğrulama Kaybı

Training vs Validation Loss (8-class)



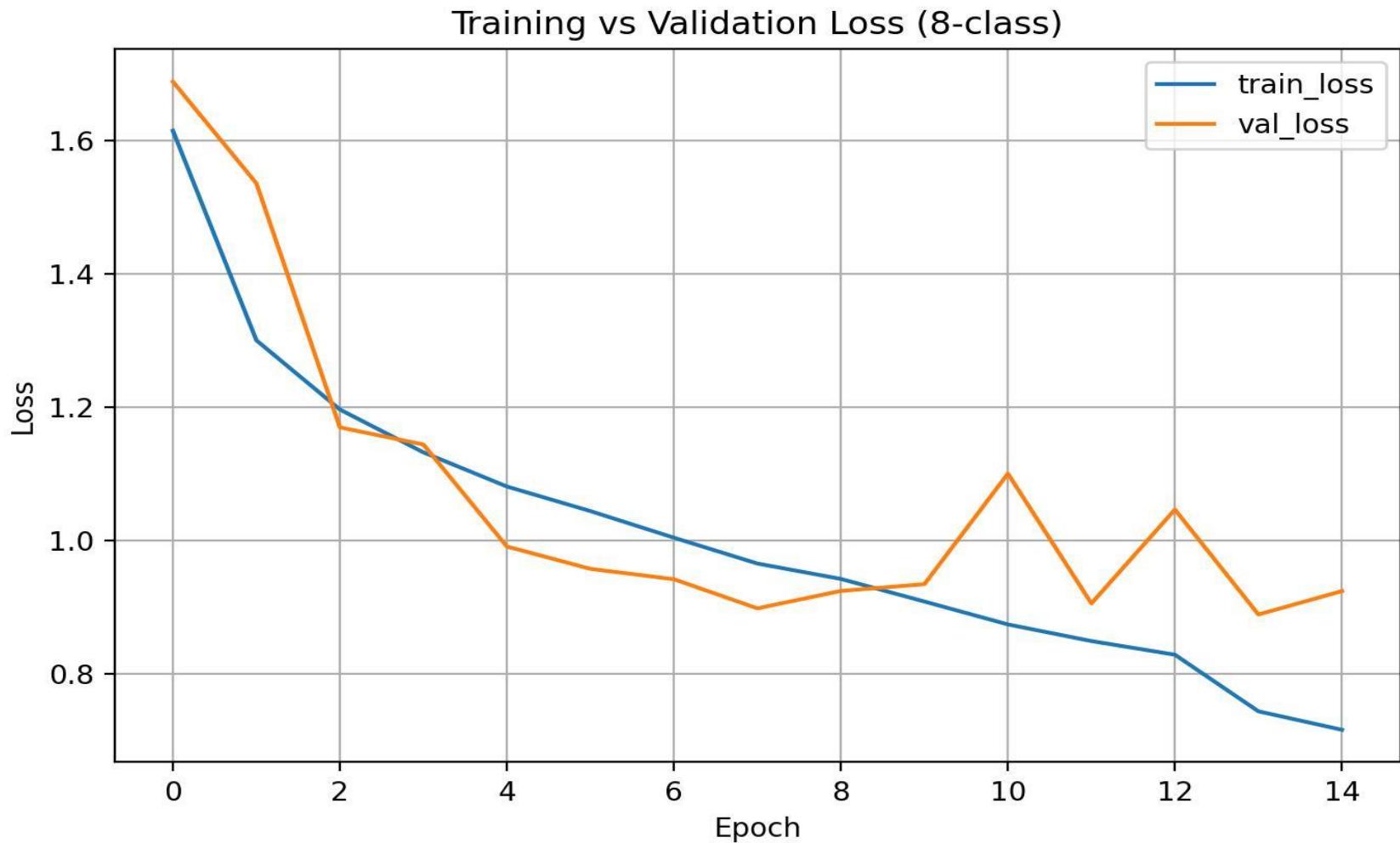
- Training loss:  $\sim 1.6 \rightarrow 0.16$
- Validation loss: en iyi nokta epoch 10–15
- Sonrasında validation loss artışı (overfitting //**aşırı uyum**// )

# Eğitim ve Doğrulama Kaybı



- Training loss:  $\sim 1.6 \rightarrow 0.16$
- Validation loss: en iyi nokta epoch 10–15
- Sonrasında validation loss artışı (overfitting //**aşırı uyum**// )

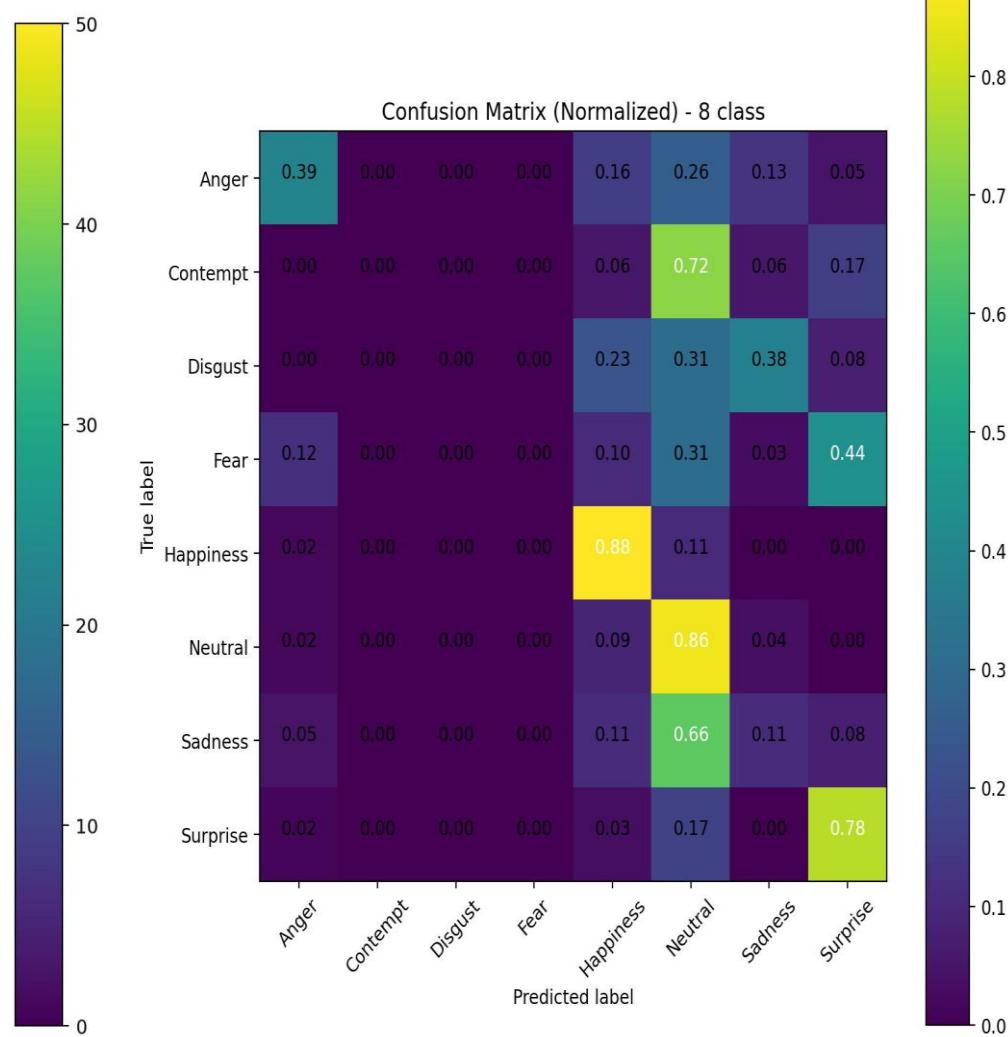
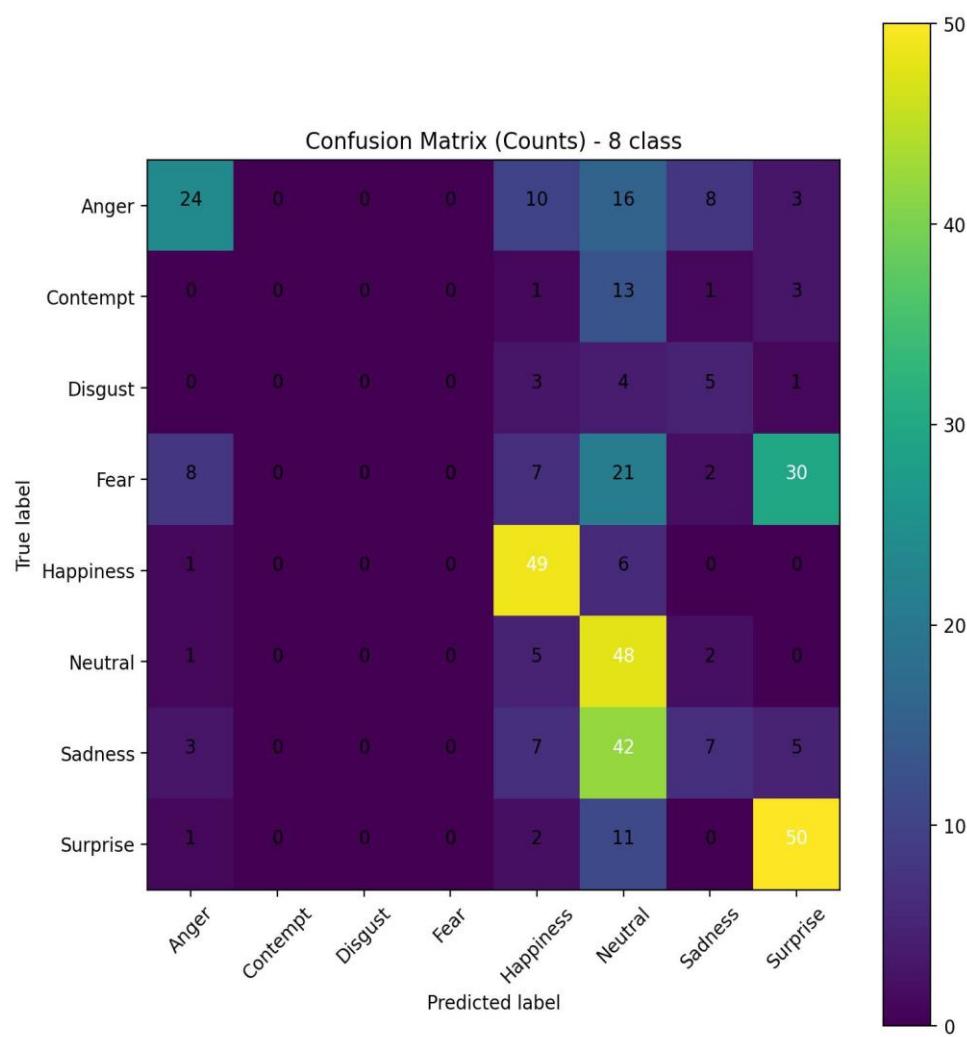
# Eğitim ve Doğrulama Kaybı



- Training loss:  $\sim 0.16 \rightarrow 1.6$
- Validation loss: en iyi nokta epoch 15–10

# Karmaşıklık Matrisi

- Test doğruluğu: %54.5 for 60 epoch , Test doğruluğu: 45% for 10-15 epoch
- En iyi sınıflar: Happiness, Neutral, Surprise, Anger
- Zayıf sınıflar: Contempt, Disgust



# Sınıf Bazlı Performans

- Macro F1  $\approx 0.41$
- Weighted F1  $\approx 0.50$
- Dengesiz sınıflar performansı düşürmektedir

Destek (Support)	F1-Score	Recall	Precision	Sınıf (Class)
61	0.3670	0.3279	0.4167	Anger
18	0.0000	0.0000	0.0000	Contempt
13	0.0000	0.0000	0.0000	Disgust
68	0.0000	0.0000	0.0000	Fear
56	0.7234	0.9107	0.6000	Happiness
56	0.4653	0.8393	0.3219	Neutral
64	0.3125	0.2344	0.4688	Sadness
64	0.6144	0.7344	0.5281	Surprise

# Sınıf Bazlı Ölçütler

Değer	Metrik
3.2419	Test Loss
0.4500	Test Accuracy
0.2919	Macro Ortalama Precision
0.3808	Macro Ortalama Recall
0.3103	Macro Ortalama F1-Score
0.3521	Ağırlıklı Ortalama Precision
0.4500	Ağırlıklı Ortalama Recall

# Örnek Tahminler

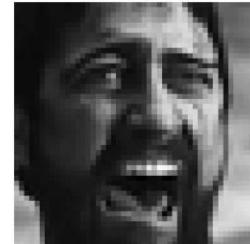
T:Anger  
P:Happiness



T:Anger  
P:Sadness



T:Anger  
P:Anger



T:Anger  
P:Neutral



T:Anger  
P:Anger



T:Anger  
P:Anger



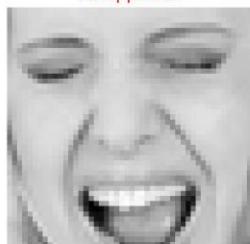
T:Anger  
P:Surprise



T:Anger  
P:Neutral



T:Anger  
P:Happiness



T:Anger  
P:Neutral



T:Anger  
P:Neutral



T:Anger  
P:Surprise



T:Anger  
P:Sadness



T:Anger  
P:Anger



T:Anger  
P:Anger



T:Anger  
P:Anger



Doğru ve yanlış sınıflandırma örnekleri

Görsel benzerlikten kaynaklanan hatalar

Anger – Neutral – Happiness karışmaları

# Sonuçlar

- Ham piksel tabanlı CNN ile 8 sınıfı duyu tanıma
- Test doğruluğu  $\approx$  %54.5 tüm ham veri epoch 15 için Basit ama etkili bir mimari
- Bu çalışmada geliştirilen CNN tabanlı duyu sınıflandırma modeli, **Mutluluk**, **Şaşkınlık** ve **Tarafsızlık** gibi duygularda görece yüksek başarı göstermesine rağmen, **Korku (Fear)**, **Tıksıntı (Disgust)** ve **Küçümseme (Contempt)** sınıflarında başarısız olmuştur. (sınıf dengesizliği ve görsel benzerlik).

# Sonuçlar

48×48 gri seviye yüz görüntülerinden, sadece ham piksel verisi kullanarak 8 sınıfı duygusal tanıma gerçekleştirildi.

**Sonuç olarak, Fear, Disgust ve Contempt sınıflarındaki başarısızlık, veri dengesizliği, düşük çözünürlük, duygular arası görsel benzerlik ve model kapasite sınırlamalarının birleşik etkisinden kaynaklanmaktadır.** Bu sınıflar için daha gelişmiş mimariler ve veri stratejileri gereklidir.

Bu sonuçlar, ham piksel tabanlı basit bir CNN ile bile duyguları makul doğrulukta tanıyalabileceğimizi gösteriyor

# Gelecek Çalışmalar

- Class imbalance için class\_weight, focal loss
- Overfitting için EarlyStopping
- Transfer learning (MobileNet, ResNet)

## Önerilen iyileştirme Stratejileri:

\* Bu sınıflardaki başarısızlığı azaltmak için aşağıdaki yöntemler önerilmektedir:

• **Sınıf dengesizliği için:** Class-weight veya Focal Loss kullanımı

• **Overfitting'i azaltmak için:**

EarlyStopping, daha fazla Dropout, daha az epoch.

• **Veri artırma (augmentation) Daha güçlü modeller:**

özellikle Fear ve Disgust sınıfları için hedefli artırma Daha derin CNN veya transfer learning (ResNet, EfficientNet)

• Yüksek çözünürlüklü ve RGB veri kullanımı

• **Ek veri setleri:** RAF-DB, AffectNet -----gerçek hayatı daha yakın yüz ifadeleri

# Teşekkürler

Dinlediğiniz için teşekkür ederim.