

1. Proje Amacı ve Kapsamı

Bu proje, derin öğrenme teknolojilerini kullanarak insan yüz ifadelerinden duygu tanınması yapan bir yapay zeka sistemi içermektedir. Sistem, VGG16 transfer öğrenme modelini kullanarak dört temel duyguyu (kızgın, mutlu, üzgün, korku) tanıma yeteneğine sahiptir. Proje, veri ön işleme, model eğitimi ve web tabanlı bir kullanıcı arayüzünü içeren kapsamlı bir yapı sunmaktadır.

2. Kullanılan Veri Setleri

Projede, yüz ifadelerini içeren farklı veri setleri kullanılmıştır:

- **İlk Aşama:**
İlk aşamada, CK+ veri seti kullanılmıştır. Bu veri setinden sadece 4 duygu sınıfı seçilmiştir. CK+ veri seti, orijinalinde 920 görüntü içermektedir ve bu görüntüler 48x48 piksel boyutunda, gri tonlamalı ve yüzler Haar Cascade sınıflandırıcısı kullanılarak kırılmıştır.
- **Genişletme Aşaması:**
Modeli zenginleştirmek ve daha fazla veriyle eğitmek amacıyla, FER2013 veri setinden bir kısım veri alınarak CK+ veri setiyle birleştirilmiştir. FER2013 veri seti, 48x48 piksel boyutunda gri tonlamalı yüz görüntülerinden oluşmaktadır. Yüzler otomatik olarak hizalanmış ve her görüntüde yüz benzer bir alanı kaplayacak şekilde işlenmiştir. Bu veri setinin eğitim seti 28.709 örnek, test seti ise 3.589 örnek içermektedir. Ancak bu aşamada, FER2013 veri setinin tamamı değil, belirli bir kısmı kullanılmıştır.
- **Son Aşama:**
Birleştirilen veri setiyle yapılan denemelerde model performansı düşük çıktığı için, FER2013 veri setinin tamamı kullanılmıştır. Bu süreçte, veri setlerinin boyutları ve çeşitliliği artırılarak modelin genelleme yeteneği güçlendirilmeye çalışılmıştır.

3. Kullanılan Modeller ve Elde Edilen Sonuçlar

Projede farklı model mimarileri denenmiş ve aşağıdaki test doğruluk oranları elde edilmiştir:

- **VGG16:** %72.55 (En iyi sonuç)
- **ViT (Vision Transformer):** %52.70
- **Diğer denemeler:** %32.35, %43.14, %50.83, %56.44, %57.48, %64.97

Görüldüğü üzere, **VGG16 modeli en iyi performansı göstermiştir.**

4. Model Eğitimi ve Kullanılan Teknikler

Veri Ön işleme:

- Veri setlerinin birleştirilmesi ve uygun formata dönüştürülmesi
- Görüntülerin 48x48 piksel boyutuna yeniden boyutlandırılması
- FER2013 görüntüleri için gri tonlama ve kontrast iyileştirme
- CK+48 görüntüleri için renkli formatın korunması
- Veri artırma teknikleri (rotasyon, kaydırma, yakınlaştırma, yatay çevirme, parlaklık ayarı)

Ön işleme aşamasında, veri setinin kaynağına göre farklı stratejiler uygulanmıştır:

- FER2013 görüntüleri genellikle düşük kaliteli ve gürültülü olduğundan gri tonlamaya çevrilip Gaussian bulanıklaştırma ve CLAHE (Contrast Limited Adaptive Histogram Equalization) uygulanmıştır.
- CK+48 görüntüleri daha kaliteli olduğundan renkli formatları korunmuştur.

Yaklaşım:

- **Sıfırdan CNN Modeli:** Basit bir konvolüsyonel sinir ağı
- **Transfer Öğrenme Modeli:** Imagenet veri seti üzerinde önceden eğitilmiş VGG16 mimarisi

Uygulama, transfer öğrenme yaklaşımını benimsemektedir:

- VGG16 ağıının son 12 katmanı fine-tuning için eğitilebilir bırakılmış
- Diğer katmanlar dondurularak önceden öğrenilmiş özellikler korunmuş
- GlobalAveragePooling2D, Dense, BatchNormalization ve Dropout katmanları eklenmiş
- Adam optimizier ve categorical_crossentropy kayıp fonksiyonu kullanılmış
- Early stopping ve learning rate reduction callback'leri ile eğitim optimizasyonu
- VGG16 mimarisinde uygulanan fine-tuning stratejisi, modelin ağırlıklarını tamamen sıfırdan eğitmek yerine, önceden eğitilmiş ağırlıkları kullanarak daha hızlı ve daha iyi sonuçlar elde etmeyi sağlamıştır. Son 12 katmanın eğitilebilir bırakılması, modelin duygu tanıma görevine özel özellikleri öğrenmesine olanak tanırken, erken katmanların dondurulması temel görsel özellikleri korumuştur.

Hiperparametre Optimizasyonu:

Model eğitiminde kullanılan hiperparametreler, ağın performansını önemli ölçüde etkilemektedir:

- Düşük öğrenme oranı (0.00005), modelin aşırı öğrenme riskini azaltmış ve daha stabil bir eğitim sağlamıştır.
- 0.4 oranındaki dropout, overfitting'i önleyerek modelin genelleme yeteneğini artırmıştır.
- 128 birimlik yoğun katman, model karmaşıklığı ve performans arasında optimal bir denge sağlamıştır.

Veri Artırma Teknikleri:

Veri artırma teknikleri, modelin gerçek dünya koşullarında daha dayanıklı olmasını sağlamıştır:

- Görüntü rotasyonları (± 40 derece) farklı baş pozisyonlarına adaptasyon sağlar
- Görüntü kaydırma ve yakınlaştırma, yüzün çerçevedeki konumuna karşı dayanıklılık kazandırır
- Yatay çevirme, simetrik yüz özelliklerinin öğrenilmesine yardımcı olur
- Parlaklık ve kesme parametreleri, farklı aydınlatma koşullarına ve perspektiflere adaptasyonu geliştirir

Eğitim Süreci:

- **Optimizer:** Adam, AdamW, SGD denendi. En iyi sonucu Adam optimizer verdi.
- **Loss Function:** Sparse Categorical Crossentropy ve Focal Loss denendi.
- **Epoch Sayısı:** 50 Epoch olarak belirlendi.
- **Batch Size:** 32 olarak seçildi.
- **Learning Rate Scheduler:** Cosine Annealing kullanıldı.

5. Modelin Hata Analizi ve Geliştirme Çalışmaları

Hata Analizi:

- **Sınıflandırma raporu** incelendi.
- "Fear" ve "Sad" sınıflarının en düşük recall değerine sahip olduğu görüldü.
- Yanlış sınıflandırmalarda veri seti çarpıklığı gözlemlendi.

Geliştirme Çalışmaları:

- Daha düzgün bir veri dengesi sağlamak için veri seti ön işlendi.
- Modelin ağırlıkları optimize edildi.

- ViT modeli denenmesine rağmen beklenen sonuç alınamadı.

6. Modelin Gerçek Zamanlı Tahmin Entegrasyonu

Streamlit kütüphanesi kullanılarak geliştirilen web tabanlı bir arayüz:

- Görüntü yükleme ve analiz etme yeteneği
- Güven oranlarını gösteren görselleştirilmiş sonuçlar
- Duygu sınıflarını emoji ve renkli grafiklerle görselleştirme
- Profesyonel ve modern bir kullanıcı deneyimi

Arayüz özellikleri:

- Modern ve temiz bir tasarım
- Duygu durumlarını temsil eden emoji görselleştirmesi
- Model güvenini gösteren renkli çubuk grafikleri
- Yardımcı ipuçları ve bilgi kutuları
- Duygu sınıfları için görsel renk kodlaması (kızgın: kırmızı, mutlu: yeşil, vb.)

7. Sonuç

Bu proje, modern derin öğrenme tekniklerini kullanarak yüz ifadelerinden duygu tanıma alanında kapsamlı bir çözüm sunmaktadır. VGG16 mimarisi üzerine inşa edilen transfer öğrenme yaklaşımı, sınırlı veri ile bile yüksek performans elde edilmesini sağlamıştır. Streamlit ile geliştirilen web arayüzü, teknik olmayan kullanıcıların da sistemi kolayca kullanabilmesine olanak tanımaktadır.

Projenin modüler yapısı, gelecekte yeni duygu sınıflarının eklenmesi, farklı model mimarilerinin denenmesi veya gerçek zamanlı video analizi gibi iyileştirmelerin kolayca entegre edilebilmesini sağlamaktadır. Bu çalışma, duygu tanıma teknolojisinin kullanıcı deneyimi, pazarlama araştırmaları, eğitim ve sağlık hizmetleri gibi çeşitli alanlardaki potansiyel uygulamalarına temel oluşturmaktadır.