

**TC.
İSTANBUL TOPKAPI ÜNİVERSİTESİ**



SWE303 PROGRAMLAMA DİLLERİ PROJESİ

Alzheimer Erken Tanı ve Analiz Sistemi

Mehmet POYRAZ - 23040301020

1. Özет

Bu proje, Alzheimer hastalığının erken teşhisindeki zorlukları aşmak amacıyla geliştirilmiş yapay zeka tabanlı bir Klinik Karar Destek Sistemi'dir. Çalışma kapsamında, dünyaca ünlü OASIS-2 boylamsal veri seti kullanılmış; 3 boyutlu MR görüntülerini işleyen derin öğrenme algoritmaları (3D CNN) ile klinik verileri analiz eden makine öğrenmesi modelleri (Random Forest) hibrit bir yapıda birleştirilmiştir. Geliştirilen sistem, tek modlu analizlere kıyasla daha yüksek doğruluk oranı sunarak radyolog ve nörologlara objektif bir "ikinci görüş" imkanı sağlamaktadır.

2. Problem Tanımı ve Mevcut Durum

Alzheimer hastalığı, modern tıbbın en karmaşık nörodejeneratif problemlerinden biridir. Hastalıkta teşhisinde yaşanan temel sorunlar şunlardır:

- **Subjektif Değerlendirme:** Radyologların MR görüntülerini incelerken göz yorgunuğu veya kişisel deneyim farklılıklarını nedeniyle hipokampal atrofi gibi ince detayları gözden kaçırma riski.
- **Veri Karmaşıklığı:** Bir hastanın yaşı, eğitim durumu, genetik geçmişi ve radyolojik bulgularının (beyin hacmi vb.) zihinsel olarak bütünlüğünü analiz edilmesinin zorluğu.
- **Geç Teşhis:** Semptomların (unutkanlık vb.) yaşılılığın doğal bir sonucu olarak görülmesi ve klinik başvurunun gecikmesi.

Bu proje, bu süreçleri otomatize ederek, veri odaklı ve ölçülebilir bir risk analizi sunmayı hedefler.

3. Proje Amacı ve Çözüm Yaklaşımı

Projenin temel amacı, sadece görüntüye veya sadece teste dayalı olmayan, Multimodal (Çok Modlu) bir analiz sistemi geliştirmektir.

Temel Hedefler:

1. **Hacimsel Analiz (Voxel-Based):** MR görüntülerini 2 boyutlu kesitler (slice) olarak değil, beynin derinliğini ve hacmini koruyan 3 boyutlu bloklar halinde işlemek.
2. **Hibrit Karar Mekanizması:** Görüntüden elde edilen yapısal bozulma verisi ile hastanın klinik tablosunu (MMSE skoru, Yaş, nWBV) birləştererek teşhis doğruluğunu artırmak.
3. **Klinik Görselleştirme:** Doktorların hastayı genel popülasyonla (Sağlıklı vs. Demanslı) kıyaslayabileceği grafiksel arayüzler sunmak.

4. Kullanılan Veri Seti: OASIS-2

Projede, Washington Üniversitesi tarafından sağlanan OASIS (Open Access Series of Imaging Studies) veri seti kullanılmıştır. Veri seti 150+ özgün hastanın çoklu ziyaretlerini (longitudinal) kapsar.

Modelin başarısında kilit rol oynayan parametreler şunlardır:

- **nWBV (Normalized Whole Brain Volume):** Beyin parankim hacminin kafa içi toplam hacme oranıdır. Alzheimer hastalarında bu oranın hızla düşüğü (atrofi) gözlemlenmiştir.
- **MMSE (Mini-Mental State Exam):** 30 üzerinden değerlendirilen bilişsel testtir. 24 ve altı puanlar demans şüphesini kuvvetlendirir.
- **eTIV (Estimated Total Intracranial Volume):** Kafa taşı hacmidir. Hastalar arası normalizasyon (kıyaslama) yapabilmek için referans değer olarak kullanılır.
- **MRI Verisi:** T1 ağırlıklı, yüksek çözünürlülüklü 3 boyutlu anatomik taramalar.

5. Metodoloji ve Teknik Mimari

Proje, ham veriden teşhise giden süreçte 4 aşamalı bir "Pipeline" (Borusu Hattı) mimarisini kullanır:

Tıbbi görüntülerin ve klinik verilerin modele uygun hale getirilmesi sürecidir.

- **Klinik Veri: Eksik veriler (missing values)** medyan değerlerle doldurulmuş, kategorik veriler (Cinsiyet vb.) sayısal formata (Label Encoding) dönüştürülmüştür.
- **Görüntü İşleme:**
 - **Boyut İndirgeme:** İşlem yükünü azaltmak için görüntüler 64x64x32 voksel boyutuna yeniden örneklenmiştir.
 - **Gürültü Filtreleme:** MR cihazından kaynaklı parazitleri temizlemek için Gaussian Filter ($\Sigma=0.5$) uygulanmıştır.
 - **Normalizasyon:** Görüntü piksel yoğunlukları 0-1 aralığına çekilerek modelin

daha hızlı yakınsaması (convergence) sağlanmıştır.
Sistem, iki farklı yapay zeka modelinin güçlerini birleştirir:

A. 3D CNN (Derin Öğrenme Modeli):

- Beynin 3 boyutlu uzaysal yapısını analiz eder.
- Yapı: 3 adet Conv3D (Özellik çıkarıcı) katmanı ve bunları takip eden MaxPooling3D (Boyut azaltıcı) katmanlarından oluşur.
- Çıktı: Görüntü üzerindeki yapısal bozulmaya dayalı 0-1 arası bir Alzheimer olasılık skoru üretir.

B. Random Forest (Makine Öğrenmesi Modeli):

- Klinik ve demografik veriler arasındaki karmaşık ve doğrusal olmayan ilişkileri analiz eder.
- Yaş, Eğitim, MMSE ve nWBV gibi 7 farklı özniteliği kullanarak karar ağaçları oluşturur.

C. Füzyon (Karar Birleştirme):

- İki modelden gelen olasılıklar "Ağırlıklı Ortalama" (Weighted Average) yöntemiyle birleştirilir.
- Nihai Karar Formülü: Skor = $(0.4 \times \text{CNN Olasılığı}) + (0.6 \times \text{Random Forest Olasılığı})$
- Not: Klinik verilerin (özellikle MMSE ve nWBV) tanısal gücü yüksek olduğu için Random Forest modeline daha yüksek ağırlık verilmiştir.

6. Bulgular, Analiz ve Tartışma

6.1. İstatistiksel Korelasyonlar

Veri seti üzerinde yapılan analizler şu tıbbi gerçekleri doğrulamıştır:

- Güçlü Negatif İlişki (-0.42): nWBV (Beyin Hacmi) azaldıkça Alzheimer riski belirgin şekilde artmaktadır.
- Bilişsel İlişki (-0.38): MMSE test puanı düştükçe hastalık evresi ilerlemektedir.
- Yaş Faktörü (+0.25): İleri yaş, en büyük risk faktörlerinden biri olarak teyit edilmiştir.

6.2. Model Performansları

Test veri seti üzerinde yapılan denemelerde elde edilen doğruluk (accuracy) oranları:

Model	Doğruluk	Açıklama
Sadece 3D CNN	%68	Sadece görüntü ile sınırlı başarı.
Sadece Random Forest	%76	Klinik testlerin gücü sayesinde iyi başarı.
Hibrit Model	%82	En yüksek başarı.

Sonuç: Görüntü ve klinik verinin birleşimi, tek başına kullanılan yöntemlerden daha üstün bir performans sergilemiştir.

7. Sonuç ve Gelecek Çalışmalar

Geliştirilen sistem, Alzheimer teşhisinde yapay zekanın "karar verici" değil, "karar destekleyici" olarak kullanılmasının potansiyelini ortaya koymuştur.

Gelecek Çalışmalar İçin Öneriler:

- Skull Stripping (Kafatası Soyma): Görüntü işleme aşamasında kafatası dokusunun tamamen temizlenmesi, modelin sadece beyin dokusuna odaklanması sağlayarak başarayı artırabilir.
- Explainable AI (XAI): Grad-CAM gibi tekniklerle, yapay zekanın beyin hangi bölgelere bakarak karar verdiği (örneğin Hipokampus bölgesi) görselleştirilmesi, doktorların sisteme olan güvenini artıracaktır.
- Çok Sınıflı Sınıflandırma: Sistemin sadece "Hasta/Sağlıklı" değil, "Erken Evre", "Orta Evre" gibi detaylı sınıflandırma yapacak şekilde geliştirilmesi.

8. Teknik Altyapı (Tech Stack)

- Dil: Python 3.10+
- Arayüz: Streamlit (Web UI)
- Yapay Zeka: TensorFlow, Keras, Scikit-learn
- Veri Analizi: Pandas, NumPy
- Görüntü İşleme: Nibabel, SciPy
- Görselleştirme: Matplotlib, Seaborn