

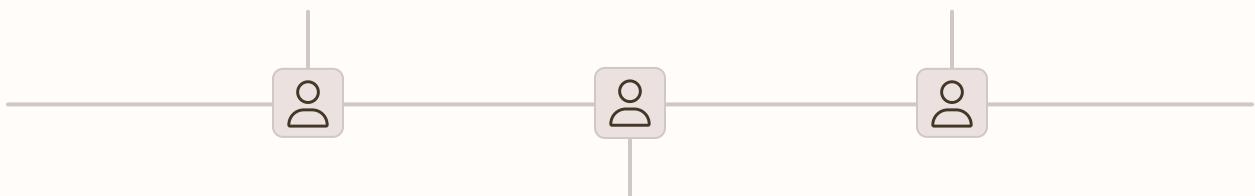
# Türk İşaret Dili Tanıma ve Metne Dönüştürme Uygulaması

Ayça Nilay ÖZOĞUL

170421012

Serra K.TALASLI

170421050



Merve HASASLARI

170421922

# Giriş

İşaret dili, işitme ve konuşma engelliler için hayatı bir iletişim aracıdır.

Bu dilin anlaşılması, toplumsal katılım için büyük önem taşır.

Teknolojimiz, gerçek zamanlı işaret tanıma ve çeviriyle günlük yaşamda erişilebilirlik sağlayarak işitme engelli bireylerin Türkçe ve İngilizce konuşanlarla iletişimini kolaylaştırmayı amaçlar.



# Giriş: İşaret Dilinin Önemi ve Zorluklar

## İletişim Aracı

İşaret dili, duyma ve konuşma engelli bireyler için temel bir iletişim aracıdır. El hareketleri, yüz ifadeleri ve vücut duruşu gibi birçok bileşeni içerir.



## Erişim Engelleri

İşaret dili bilen kişi sayısının azlığı, işitme engelli bireylerin eğitim, sağlık ve sosyal etkileşim gibi alanlarda dezavantajlımasına yol açmaktadır.



# Türk İşaret Dili (TİD) ve Küresel İletişim



## Benzersiz Yapı

Her işaret dilinin kendine özgü fonolojik, morfolojik ve sentaktik özellikleri vardır. TİD, Türkçe'den ve diğer işaret dillerinden farklı bir dilbilgisi yapısına sahiptir.



## Küresel Zorluklar

Küresel bir işaret dili standardının olmaması, farklı ülkelerden gelen işitme engelli bireyler arasında iletişimini zorlaştırmaktadır.



## Teknolojik Çözümler

Yapay zeka ve derin öğrenme teknolojileri, işaret dili tanıma sistemlerinin geliştirilmesinde önemli ilerlemeler kaydetmiştir. Bu sistemler, iletişim engellerini ortadan kaldırma potansiyeli taşır.

# Literatür Taraması: Türk İşaret Dili Tanıma Yaklaşımları

## 1 Transfer öğrenme

ile %98.8 doğruluk (SL-GCN, DANN, MCC)

## 2 CNN+LSTM

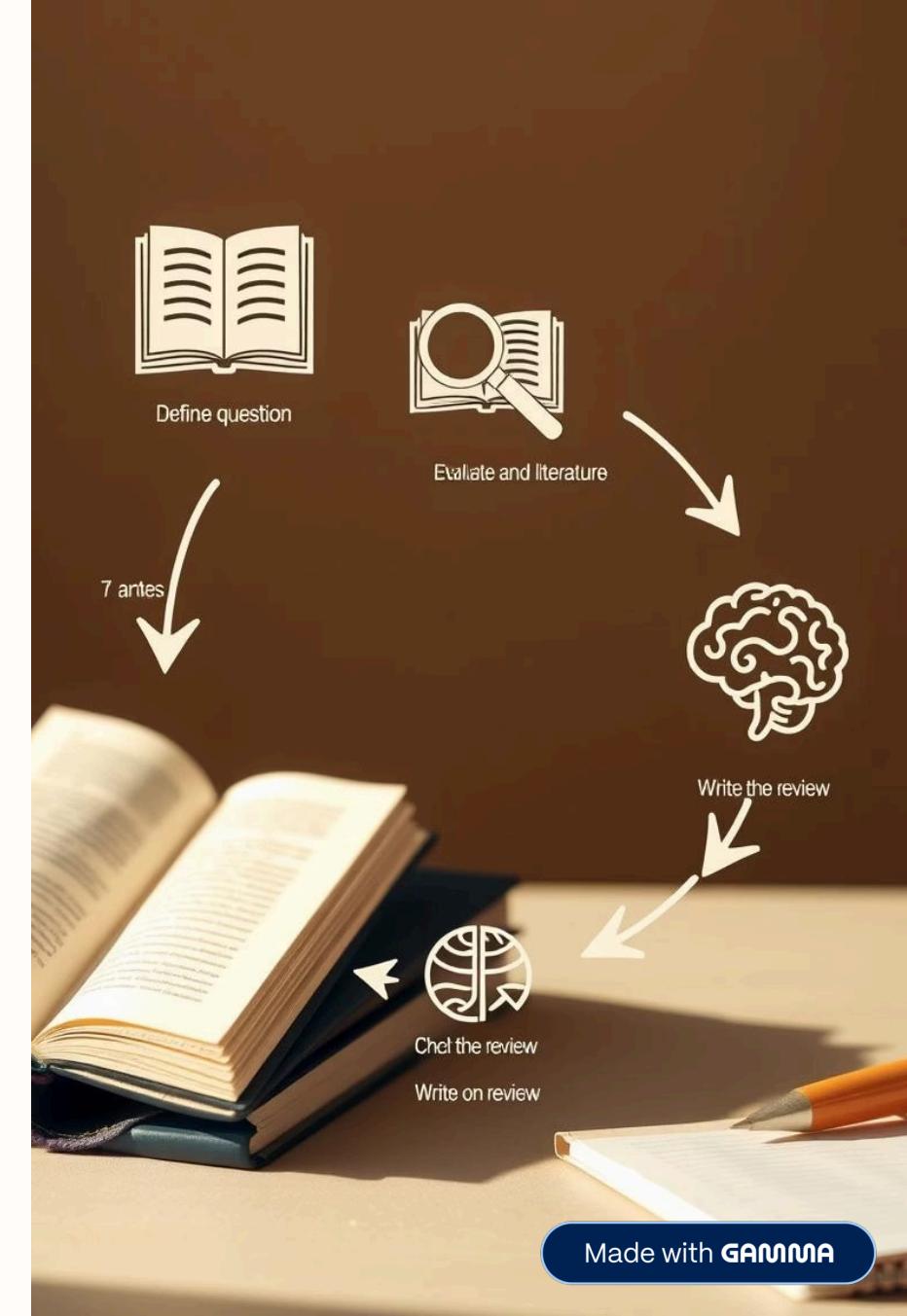
ile %95.95 doğruluk, kullanıcılardan bağımsız testte %62.02

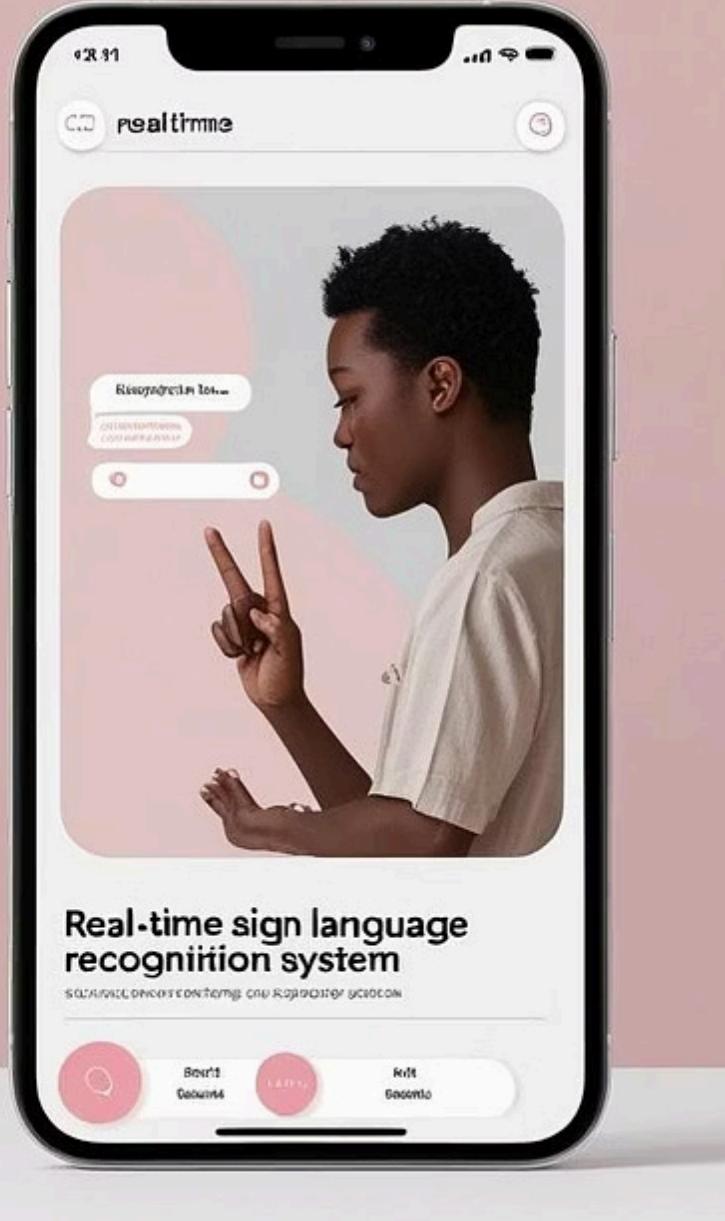
## 3 GCN, I3D ve BLSTM

ile RGB verisinde %98.53 doğruluk

## 4 Video Transformer Network (VTN)

ile %92.92 doğruluk





# Literatür Taraması: Türk İşaret Dili Tanıma Yaklaşımları

Literatürde en yüksek doğruluk, çok modlu veri ve transfer öğrenme ile elde edilmiştir (%98.8).

Çalışmamızda en yüksek doğruluk, iki katmanlı çift yönlü LSTM kullanılarak %74 olarak elde edilmiştir.

# Proje Hedefleri ve Beklenen Sonuçlar



## Yüksek Doğruluk

TİD'in otomatik tanınması için son teknoloji derin öğrenme yaklaşımlarıyla yüksek doğruluk oranına sahip bir model geliştirmek.



## Gerçek Zamanlı Performans

İşitme engelli bireylerin dijital platformlarla etkin etkileşimiini sağlamak için gerçek zamanlı performans sunmak.



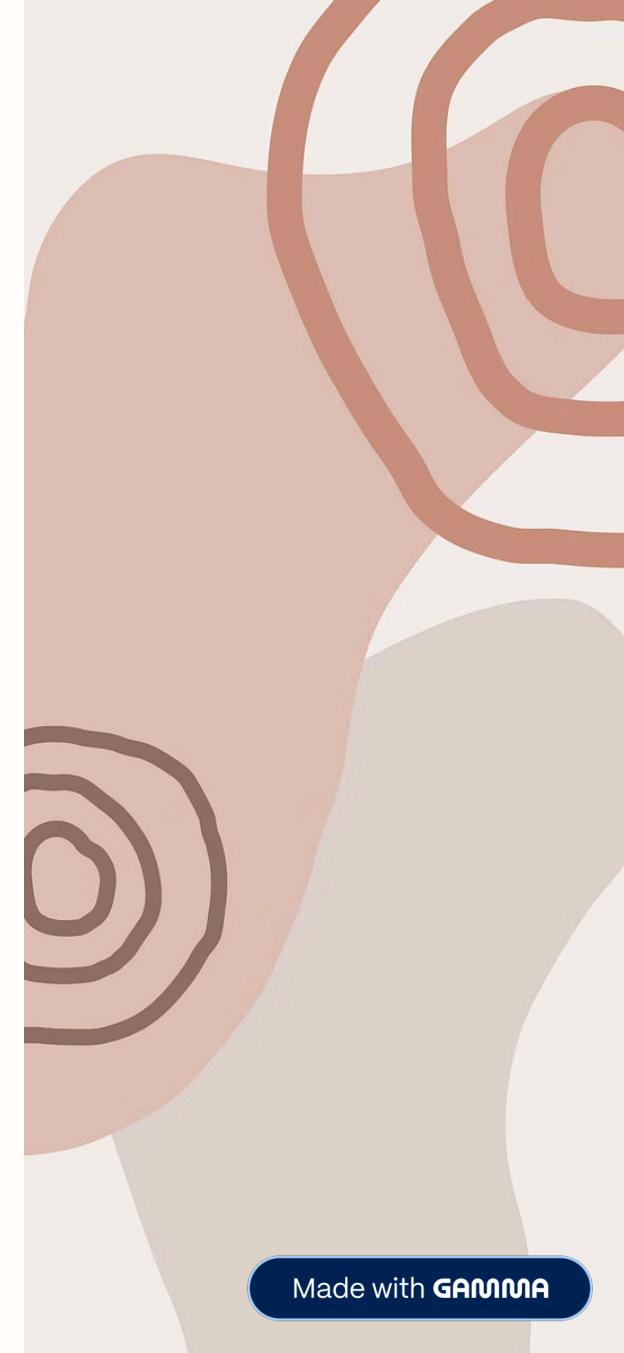
## İngilizce Çeviri

Tanınan TİD'i İngilizce metne çevirerek, TİD kullanıcıları ile İngilizce konuşanlar arasındaki iletişim köprülemek.



## Sosyal Kapsayıcılık

İnsan tercüman ihtiyacını en aza indirerek, işitme engelli bireylerin daha özgür ve bağımsız iletişim kurmasını sağlamak.



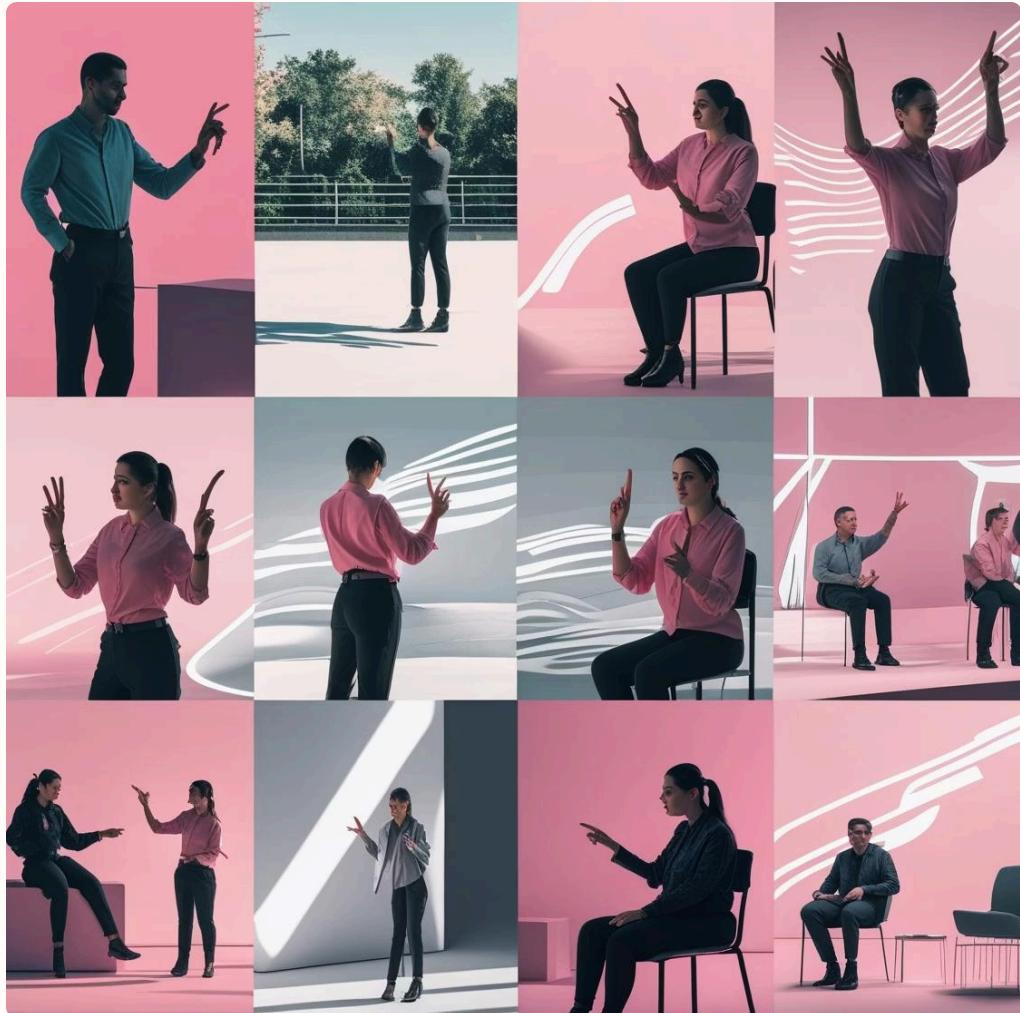
# Veri Seti: AUTSL Özellikleri

Bu çalışmada Türk İşaret Dili (TİD) tanıma için kapsamlı AUTSL veri seti kullanılmıştır. Bu veri seti, modelin gerçek dünya koşullarına uyum sağlamasını hedeflemektedir.

İşaret Sayısı	226
İşaretleyici Sayısı	43
Toplam Örnek	38,336
Farklı Arka Plan	20
Ortalama Örnek/İşaret	169.6



# Veri Seti: AUTSL Özellikleri



Veri seti, el hareketlerinin çeşitliliği ve bileşik işaretler içerir.

Dinamik ortamlarda (rüzgâr, değişken ışık) çekimler yapılarak modelin genelleme yeteneği artırılmıştır.

# Veri Ön İşleme Adımları

## Boş Kareleri Kaldırma

Videoların başlangıç ve bitişindeki boş veya durağan kareler tespit edilip kaldırılmıştır.

## Çözünürlük Ölçekleme

Tüm kareler 224x224 piksel boyutlarına yeniden ölçeklendirilmiştir.

## Bozuk Videoları Ayıklama

Eksik veya bozuk içerikli videolar veri kümesinden çıkarılmıştır.

## Kısa Videoları Filtreleme

Belirlenen eşik sürenin altındaki videolar veri setinden çıkarılmıştır.

## Düşük Kaliteli Görüntülerini Ayıklama

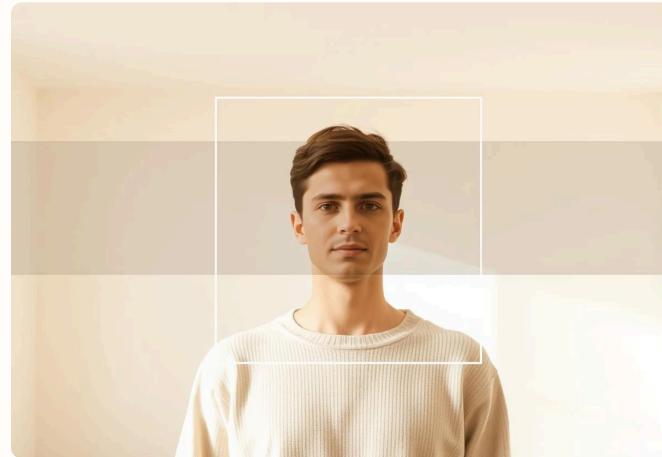
Düşük çözünürlüklü veya yetersiz görüntü kalitesine sahip videolar ayıklanmıştır.

# İskelet Çıkarımı Yöntemleri

İşaret dili tanıma için iskelet çıkarma işlemi, videoların karelere ayrılmasıyla daha doğru sonuçlar verir. İskelet çıkarımı için üç farklı yöntem denenmiştir:

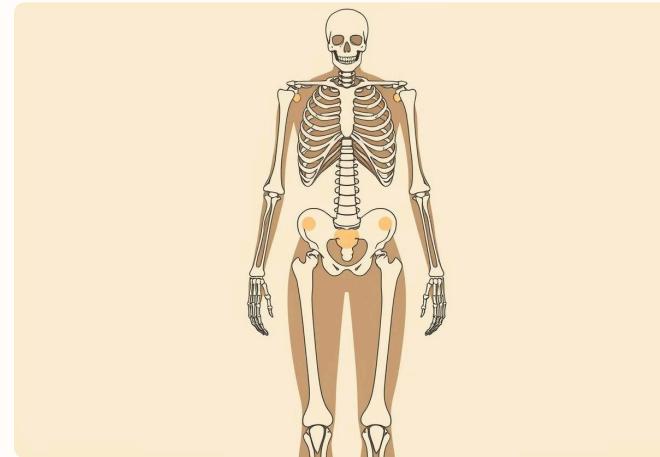
## YOLO

Nesne tespiti için kullanılır ancak el ve parmak detayları için yeterli değildir.



## OpenPose

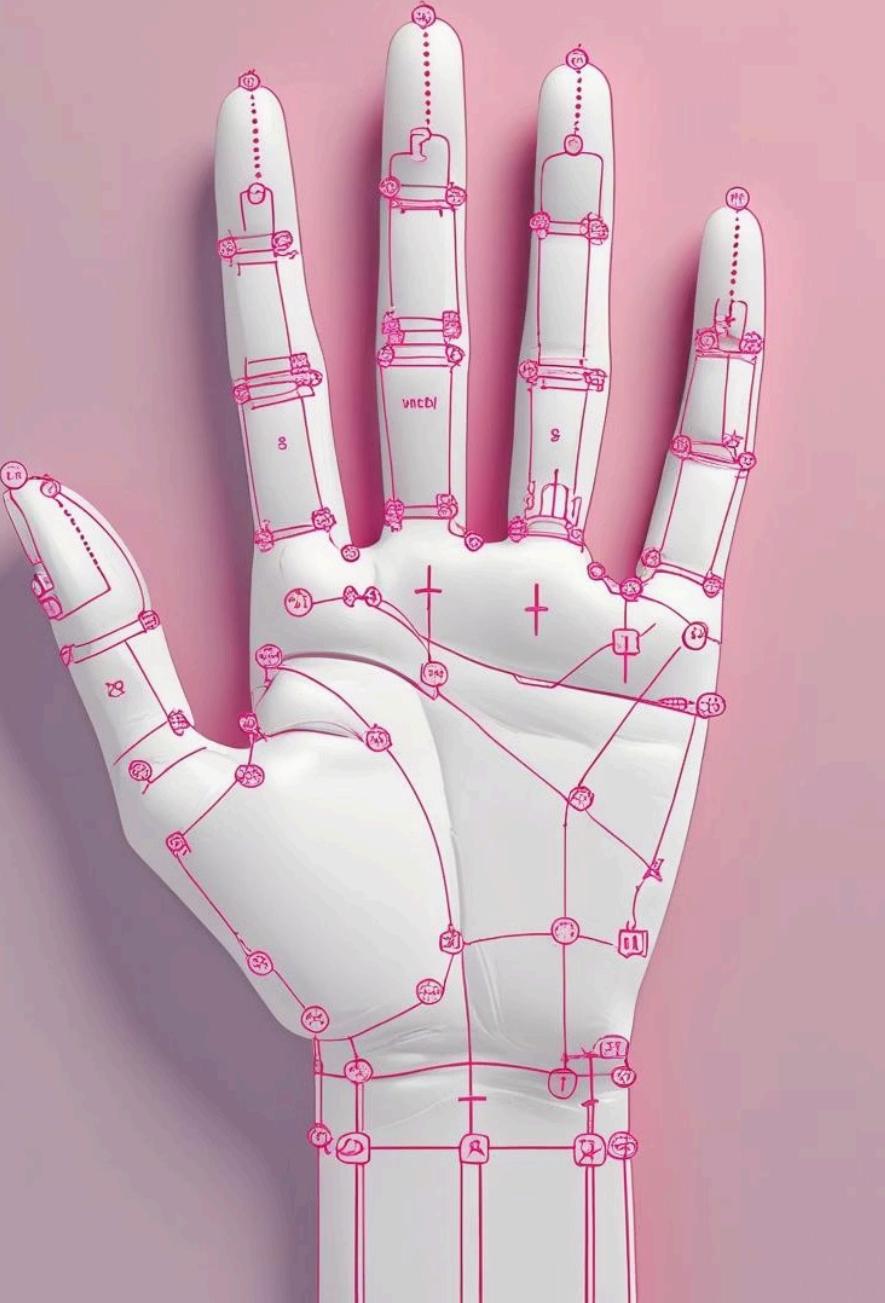
Yüz, vücut ve el iskelet noktalarını çıkarır ancak kurulumu karmaşık ve zaman alıcıdır.



## MediaPipe

Google tarafından geliştirilen, hafif ve GPU destekli bir kütüphanedir. Kurulumu kolay, gerçek zamanlı ve detaylı el takibi sunar.





# MediaPipe

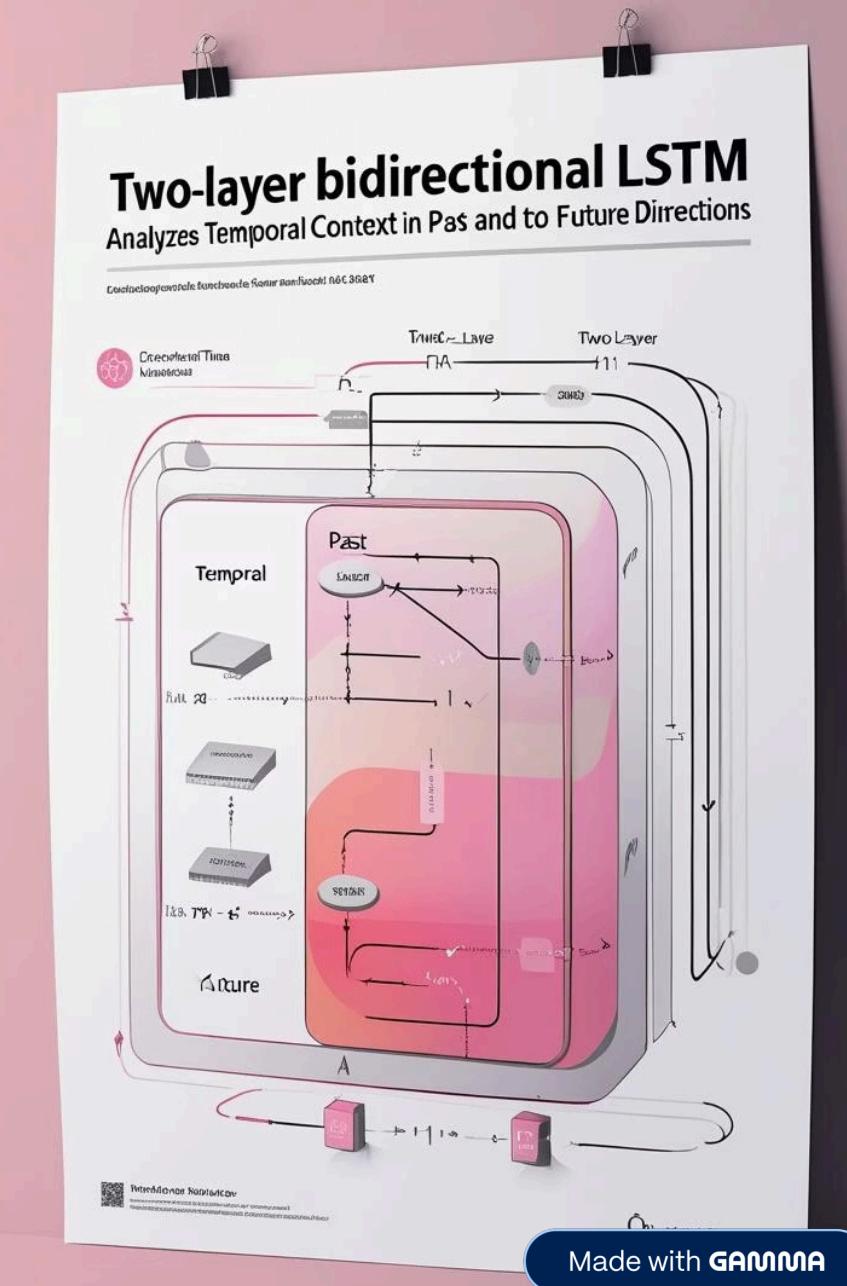
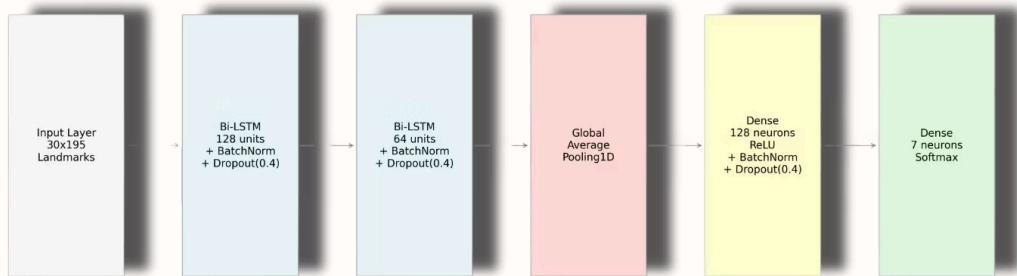
Denenen yöntemler arasında MediaPipe en uygunu olarak seçilmiştir.

MediaPipe, el ve vücut için x, y ve z koordinatları içeren toplam 195 landmark çıkarmaktadır.

Koordinatlar normalizasyonla standartlaştırılmıştır.

# Model Mimarisi: Bidirectional LSTM

Model, el hareketlerinin zamansal analizine odaklanan iki katmanlı çift yönlü LSTM yapısı üzerine kurulmuştur. BiLSTM, zamansal bağlamı hem geçmiş hem gelecek yönünde analiz eder. Önceden çıkarılmış görsel özellikler doğrudan BiLSTM'e beslenir.



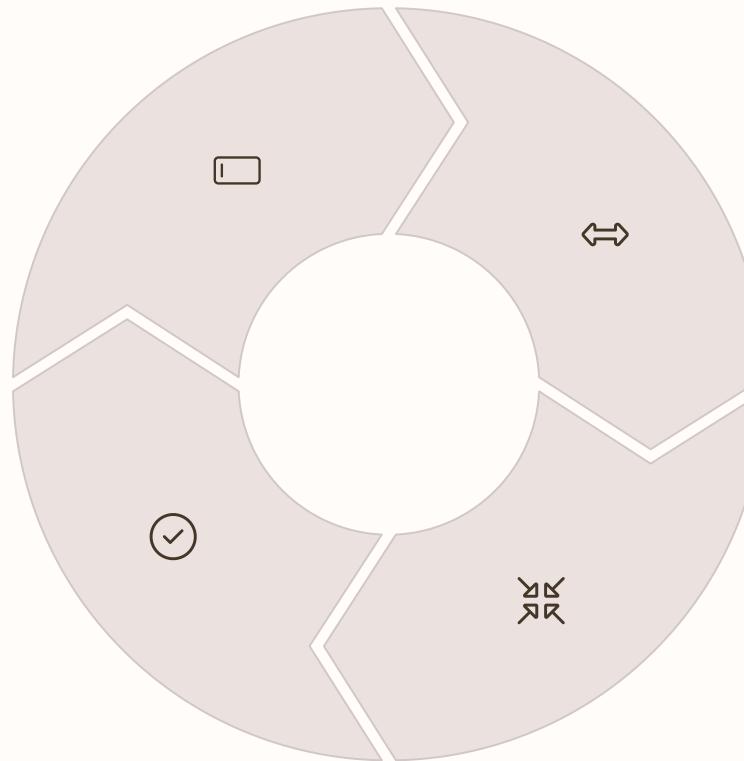
# Model Mimarisi: Bidirectional LSTM

## Girdi Özellikleri

Modelin girişi, her biri görsel bir kareyi temsil eden ve belirli sayıda özelliğe sahip sıralı veri olarak tanımlanmıştır.

## Çıkış Katmanı

Son katman, sınıf sayısı kadar nörona sahip olup Softmax aktivasyon fonksiyonu ile sınıflandırma gerçekleştirir.



## Bidirectional LSTM Katmanları

İki katmanlı LSTM, hareketlerin zamansal örüntülerini her iki yönde öğrenir. Aşırı öğrenmeyi önlemek için L2 regularizasyonu, Batch Normalization ve Dropout uygulanır.

## Havuzlama ve Tam Bağlantılı Katman

GlobalAveragePooling1D ile boyut düşürülür. 128 nöronlu, ReLU aktivasyon fonksiyonlu tam bağlantılı katman, öğrenilen zamansal özelliklerini işler.

# Model Eğitimi ve Değerlendirme Süreci

**Optimizasyon:**

Adam

**Loss Function:**

Categorical Crossentropy

**Grid Search ile Test Edilenler:**

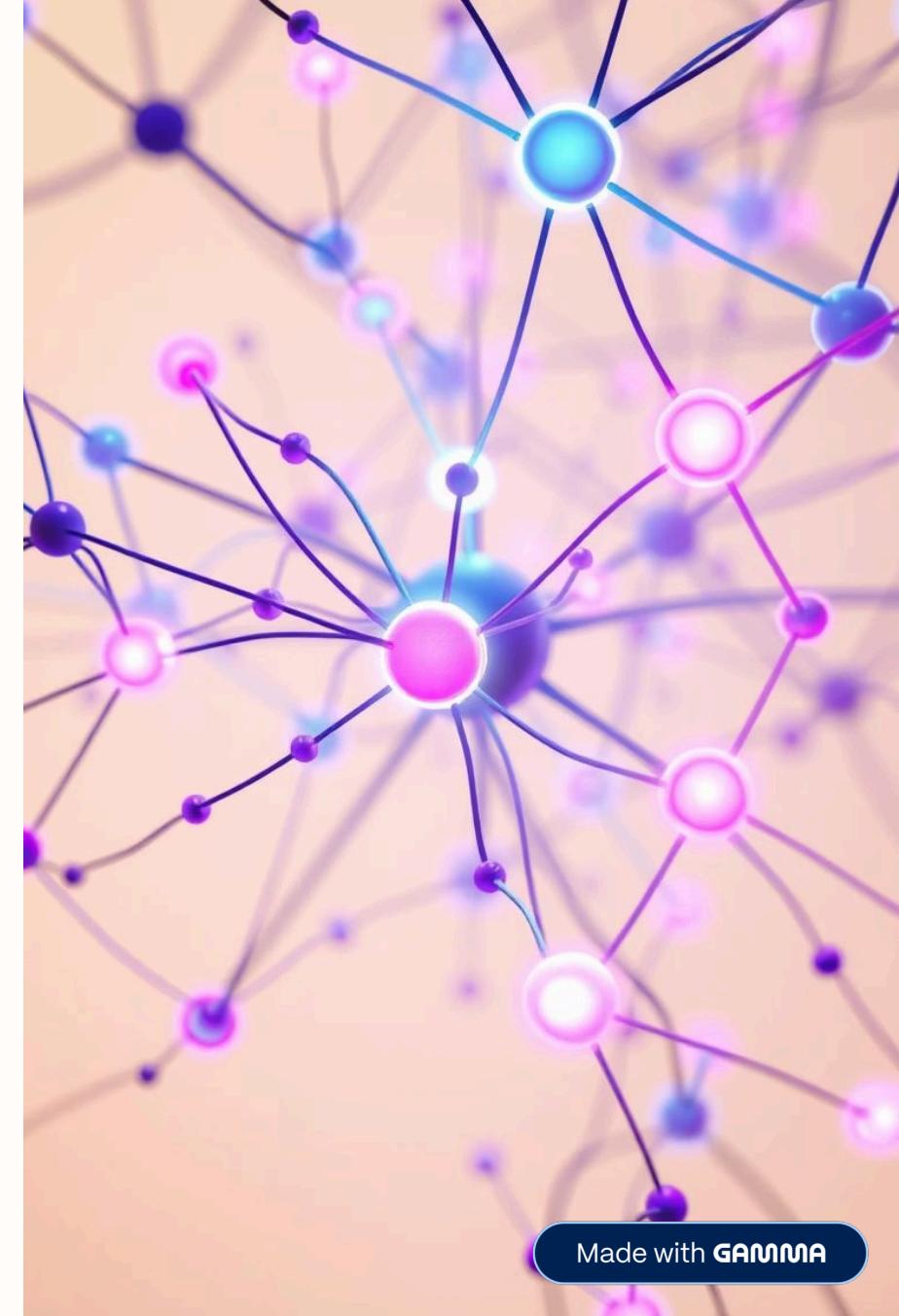
Learning Rate: 0.001,  
0.0001, 0.0003

Dropout: 0.4

Batch Size: 16, 32, 64

**Değerlendirme:**

Accuracy, F1 Skoru, Confusion Matrix



# Eğitim ve Test Aşamaları

## Veri Bölme

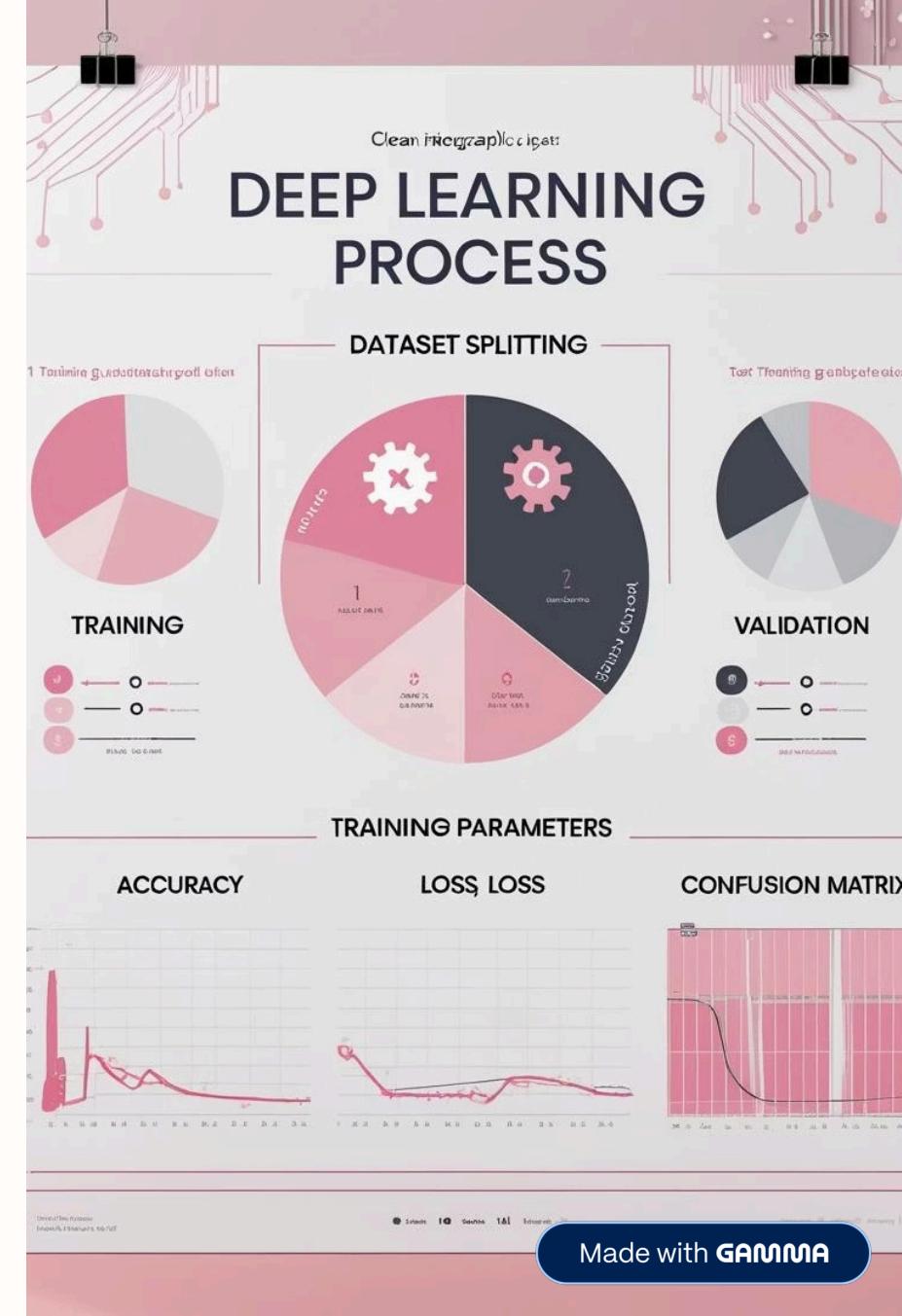
Veri seti %78 eğitim, %12 doğrulama ve %10 test olarak bölünmüştür. Eğitim sırasında çapraz doğrulama (k-fold cross-validation) uygulanmıştır.

## Eğitim Parametreleri

Adam optimizasyon algoritması, kategorik çapraz entropi kaybı kullanılmıştır. Öğrenme oranı (0.001, 0.0001, 0.0003), batch boyutu (16, 32, 64) ve 150 epoch boyunca eğitim yapılmıştır.

## Model Değerlendirme

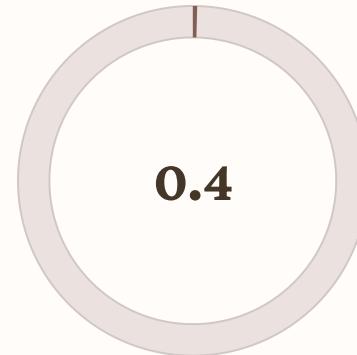
Modelin sınıflandırma başarısı doğruluk (Accuracy), F1 Skoru ve Karışıklık Matrisi (Confusion Matrix) ile değerlendirilmiştir.



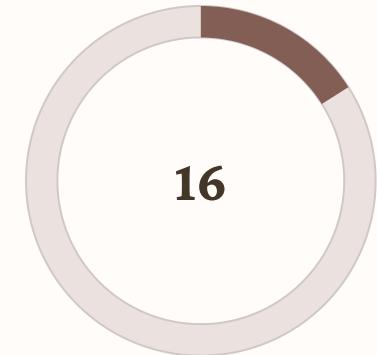
# Seçilen Hiperparametre

Overfitting'i önlemek için Dropout ve L2 regularizasyon kullanılmıştır

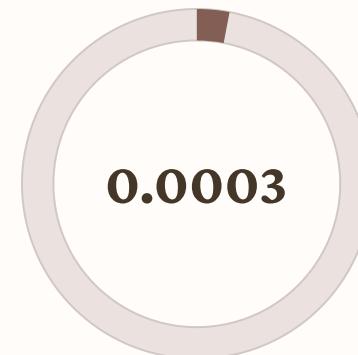
En yüksek doğruluk oranına 54. epoch sonunda ulaşarak model eğitimi durdurulmuştur.



**Dropout**



**Batch Size**



**Learning Rate**

## Model Performans Tablosu

Metrik	Final Epoch	Epoch 54 (En İyi Val Accuracy)
Training Loss	1.4183	1.4571
Training Accuracy	77.03%	75.09%
Validation Loss	2.3190	2.2951
Validation Accuracy	63.81%	64.17%

### Yorum:

1 Modelin en dengeli genel performansı, overfitting'in en az seviyeye indiği 54. epoch'ta gözlemlenmiştir. Bu epoch'ta doğrulama doğruluğu %64.17, eğitim doğruluğu ise %75 civarındadır.

2 Validation kaybı (loss), training'e göre daha yüksek olsa da, overfitting sınırlı düzeydedir.

3 Bu performans, gerçek zamanlı işaret dili tanıma uygulamaları için kullanılabilir düzeyde olup, gelecekte yapılacak hiperparametre ve veri çeşitliliği iyileştirmeleriyle daha da geliştirilebilir.

# Kullanılan Teknolojiler



## Frontend

HTML, CSS, JavaScript



## Backend

Python, Flask, WebSocket



## ML/Görüntü İşleme

MediaPipe, OpenCV, TensorFlow



## Servisler

EmailJS, Gemini API

# Şifre Yenileme Sistemi

1 Kullanıcı "Şifremi Unuttum" →



E-posta Adresi

Kod Gönder Geri Dön

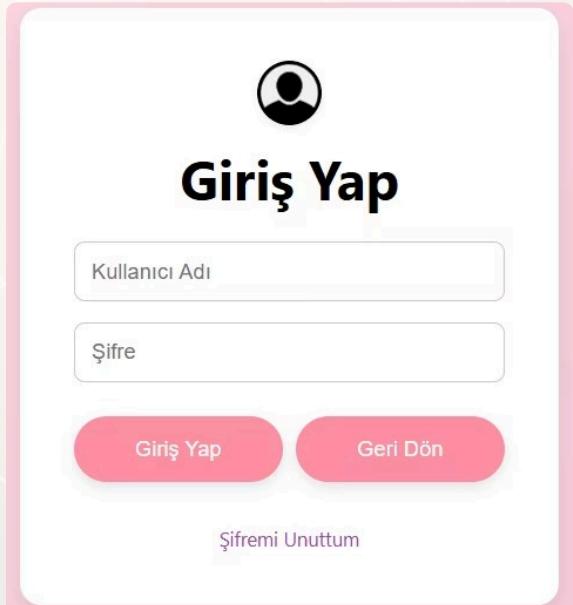
Şekil 1. Şifre Yenileme

2 EmailJS ile 6 haneli kod gönderimi →



Şekil 2. Şifre Yenileme Maili

3 Kod girişi sonrası yeni şifre belirleme



Kullanıcı Adı

Şifre

Giriş Yap Geri Dön

Şifremi Unuttum

Şekil 3. Giriş Ekranı

# Kullanıcı Profil Yönetimi



**Kullanıcı Dostu Tasarım:** Profil bilgileri (kullanıcı adı, e-posta, şifre) tek bir sayfada düzenlenebilir.



**Güvenlik Odaklı:** Yeni şifre belirleme ve mevcut şifre doğrulama alanlarıyla yetkisiz erişim engellenir.



**Basit İş Akışı:** "Kaydet" butonuyla değişiklikler anında uygulanır.



**Teknik Altyapı:** Bilgiler, tarayıcı tabanlı `localStorage` ile saklanır. Şifre güncellemeleri, EmailJS entegrasyonuyla doğrulanabilir.

## Profil Bilgileri

**Kullanıcı Adı**

**E-posta Adresi**

**Yeni Şifre**

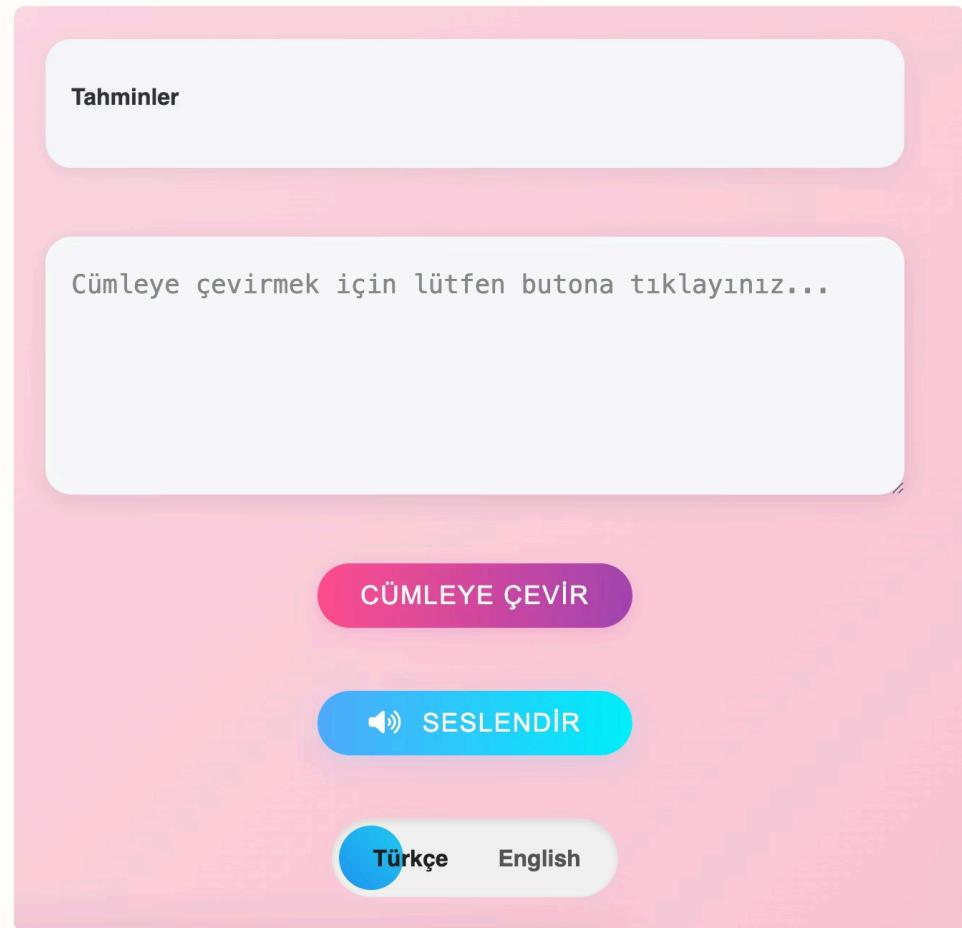
**Güncel Şifre**

**KAYDET**

Şekil 4. Kullanıcı Bilgileri Değiştirme

# Anlamlı Sonuç Üretme

- Cümleye Çevirme**  
1  
Kullanıcılar, yalnızca işaretleri tanıtmakla kalmaz, aynı zamanda bu tahminleri anlamlı diyaloglara dönüştürebilir. Gemini API'nin dil modeli, dağınık tahminleri akıcı bir cümle haline getirir ve çift yönlü iletişim sağlar.
- Seslendirme**  
2  
Seslendirme butonu, metni konuşmaya dönüştürerek işitme engelli olmayan bireylerle iletişimini kolaylaştırır.
- İngilizce'ye Çevirme**  
3  
Türkçe ve İngilizce seçenekleriyle kullanıcılar, çeviriyi istedikleri dilde alabilir.

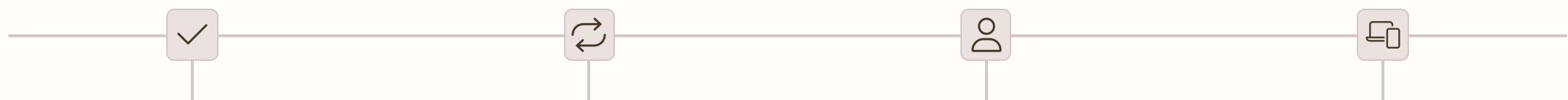


Şekil 5. Kelimelerin Anlamı Cümleye Çevrilmesi, İngilizce'ye Çeviri ve Seslendirme



# Sonuç

Geliştirilen sistem, Türk İşaret Dili tanıma ve çeviri alanında önemli bir ilerleme kaydetmiş, yapılan testlerde literatürdeki benzer çalışmalara kıyasla oldukça rekabetçi doğruluk oranları sunmuştur.



## Doğruluk Oranları

Model, **%75 doğruluk** oranıyla literatürdeki benzer RGB tabanlı çalışmalarla rekabet edebilir düzeydedir.

## Gerçek Zamanlı Performans

**MediaPipe** ve **Bidirectional LSTM** entegrasyonu sayesinde, kameradan alınan işaretler **anında** metne ve İngilizce'ye çevrilebilmektedir.

## Kullanıcı Deneyimi

Basit ve **etkileşimli arayüz** sayesinde kullanıcılar, işaret dilini kolayca metne ve sese dönüştürebilir.

## Teknolojik Katkı

**AUTSL veri seti** üzerinde yapılan derin öğrenme modeli, Türk İşaret Dili için **özgün bir yaklaşım** sunmuştur.

Dinlediğiniz için teşekkür ederiz...

