



**T.C.
GEBZE TEKNİK ÜNİVERSİTESİ**

Bilgisayar Mühendisliği Bölümü

**REAL TIME WORD
RECOGNITION**

**Gözde DOĞAN
131044019**

**Danışman
Dr. Uraz Cengiz TÜRKER**

**Ocak, 2020
Gebze, KOCAELİ**



**T.C.
GEBZE TEKNİK ÜNİVERSİTESİ**

Bilgisayar Mühendisliği Bölümü

**REAL TIME WORD
RECOGNITION**

**Gözde DOĞAN
131044019**

**Danışman
Dr. Uraz Cengiz TÜRKER**

**Ocak, 2020
Gebze, KOCAELİ**

Bu çalışma 15/01/2020 tarihinde aşağıdaki jüri tarafından Bilgisayar Mühendisliği Bölümü'nde Lisans Bitirme Projesi olarak kabul edilmiştir.

Bitirme Projesi Jürisi

| | | |
|--------------|---------------------------|--|
| Danışman Adı | Dr. Uraz Cengiz TÜRKER | |
| Üniversite | Gebze Teknik Üniversitesi | |
| Fakülte | Mühendislik Fakültesi | |

| | | |
|------------|---------------------------|--|
| Jüri Adı | Doç. Dr. Didem GÖZÜPEK | |
| Üniversite | Gebze Teknik Üniversitesi | |
| Fakülte | Mühendislik Fakültesi | |

ÖNSÖZ

Bu kılavuzun ilk taslaklarının hazırlanmasında emeği geçenlere, kılavuzun son halini almasında yol gösterici olan Sayın Dr. Uraz Cengiz TÜRKER hocama ve bu çalışmayı destekleyen Gebze Teknik Üniversitesi'ne içten teşekkürlerimi sunarım.

Ayrıca eğitimim süresince bana her konuda tam destek veren aileme ve bana hayatlarıyla örnek olan tüm hocalarıma saygı ve sevgilerimi sunarım.

Ocak, 2020

Gözde DOĞAN

İÇİNDEKİLER

| | |
|---------------------------------------|------|
| ÖNSÖZ..... | iv |
| İÇİNDEKİLER | v |
| ŞEKİLLER TABLOSU | vi |
| KISALTMA LİSTESİ | vii |
| SEMBOL LİSTESİ | viii |
| ÖZET | ix |
| SUMMARY | x |
| 1. GİRİŞ | 1 |
| 1.1. PROJENİN TANIMI..... | 2 |
| 1.2. PROJENİN GEREKSİNİMLERİ..... | 3 |
| 2. TEMEL BİLGİLER VE ÇALIŞMALAR | 3 |
| 3. TASARIM..... | 10 |
| 3.1. PCA ile TASARIM | 11 |
| 4. UYGULAMA..... | 13 |
| 4.1. PCA ile GERÇEKLEME | 13 |
| 5. TARTIŞMA VE SONUÇ | 14 |
| KAYNAKLAR..... | 16 |
| EKLER..... | 18 |

ŞEKİLLER TABLOSU

| | |
|---|----|
| Şekil 1: Konuşma Tanıma (Speech Recognition) | 1 |
| Şekil 2: Problem Tanımı | 2 |
| Şekil 3: Genel Konuşma Tanıma Sistemi [7] | 4 |
| Şekil 4: Kovaryans matrisi | 8 |
| Şekil 5: x , y değerleri için kovaryans hesaplanması | 8 |
| Şekil 6: Varyasyon kovaryans eşitliği | 8 |
| Şekil 7: Öz değerlerin hesaplanması | 8 |
| Şekil 8: Öz vektörlerin hesaplanması | 9 |
| Şekil 9: Genel Tasarım Planı | 10 |
| Şekil 10: Konuşma tanıma işlemi | 10 |
| Şekil 11: Kelime Tanıma (Word Recognition) Tasarım Planı | 11 |
| Şekil 12: PCA ile gerçekleştirilen tasarım planı | 12 |
| Şekil 13: PCA Modülü | 12 |
| Şekil 14: PCA algoritması kullanılarak kelime tanıma işleminin gerçekleştirilmesi | 13 |

KISALTMA LİSTESİ

| | |
|--------------|---|
| G.T.Ü | : Gebze Teknik Üniversitesi |
| PCA | : Principal Component Analysis (Temel Bileşenler Analizi) |
| LDA | : Linear Discriminant Analysis (Lineer Ayırt Edici Analiz) |
| TDA | : Tekil Değer Ayrışımı |
| HMM | : Hidden Markov Models (Saklı Markov Modelleri) |
| DTW | : Dynamic Time Warping (Dinamik Zaman Atlama) |
| DNN | : Deep Neural Networks (Derin Sinir Ağları) |
| ASR | : Automatic Speech Recognition (Otomatik Konuşma Tanıma) |
| PDA | : Personal Digital Asistant (Taşınabilir Küçük Cep Bilgisayarları) |

SEMBOL LİSTESİ

| | |
|------------|--|
| S | : Standart Sapma |
| V | : Varyans |
| $cov(x,y)$ | : x ve y değerleri için kovaryans |
| Cov | : Kovaryans Matris |
| \bar{y} | : Y veri setinin değerlerinin ortalaması |
| λ | : Lambda (Öz Değer) |

ÖZET

Konuşma tanıma sistemi kullanılan kelimeleri veya cümleleri tanımak ve bunları ilgili metin biçimine çevirmek için kullanılan bilgisayar destekli bir sistemdir. Günümüzde mobil cihazların kullanımının artmasıyla beraber konuşma tanıma sisteminin kullanımı da yaygınlaşmıştır. Konuşma tanıma sistemi kullanıcı ile uygulama arasında kullanılan bir ara yüz olmuştur.

Konuşma sırasında geçen belirli kelimeleri yakalama işlemi ise kelime tanıma sistemi olarak adlandırılmaktadır. Konuşmada geçen belirli kelimeleri yakalama işlemi ses dosyası arama gibi çeşitli konularda kolaylık sağlayacaktır.

Bir gerçek zamanlı kelime tanıma sistemi öneriyoruz. Bu sistem, kullanıcı konuşurken elimizdeki listede yer alan kelimelerin yakalanması işlemini gerçekleştirecektir. Bu sistem makine öğrenmesi yöntemlerinden biri olan PCA (Principal Component Analysis – Temel Bileşenler Analizi) yöntemi ile uygulanacak gerçek zamanlı bir kelime tanıma sistemi olacaktır.

SUMMARY

The speech recognition system is a computer-aided system used to recognize words or phrases used and translate them into the corresponding text format. Today, with the increase in the use of mobile devices, the use of speech recognition system has become widespread. The speech recognition system is an interface between the user and the application.

The process of capturing certain words during speech is called the word recognition system. The process of capturing certain words in the conversation will provide convenience in various aspects such as audio file search.

We recommend a real-time word recognition system. This system will capture the words that we have in the list while the user is talking. This system will be a real-time word recognition system which will be implemented with PCA (Principal Component Analysis) method, which is one of the machine learning methods.

1. GİRİŞ

Günümüzde yaygınlaşan teknoloji ile elektronik cihaz kullanımı da artmıştır. Önceleri sadece kamu kuruluşlarında denk geldiğimiz bilgisayarlar artık bir günlük kullanım aracı olmuştur. Aynı şekilde mobil cihazlar yani tabletler ve cep telefonları da yaygınlaşan diğer teknoloji ürünleridir. Bu teknolojik araçlar ile kurulan etkileşim genellikle butonlar ve klavyeler yardımıyla olmaktadır. Yine son yıllarda bu arayüz etkileşimi yöntemine bir de dokunmatik ekranlar eklenmiştir. Bu şekilde kurulan etkileşim yani bu küçük cihazların küçük klavyelerine yazmak fazlasıyla zordur. Zor olduğu için de sıklıkla yanlışlıklar yapılmaktadır. Bu nedenle de elektronik cihazlar için alternatif arayüz arayışları başlamıştır. Son zamanlarda yaygınlaşan konuşma tanıma sistemleri bu arayışın bir sonucudur [1].

Konuşma tanıma elektronik cihaz üzerindeki mikrofon yardımıyla konuşmanın cihaz tarafından yazı olarak çevrilmesi olarak tanımlanmaktadır [2]. Kısaca sesin yazıya çevrilmesi de diyebiliriz. Kullanıcı komutunu ses ile verir ve bu ses cihaz tarafından algılanır, algılanan sesin yazıya dönüştürülmesi gerçekleşir. Bunun sonucunda da istenilen komut anlaşılır ve anlaşılan bu komut gerçekleştirilir.



Şekil 1: Konuşma Tanıma (Speech Recognition)

Konuşma tanıma sistemi bir çok alanda insanların günlük yaşantısına katkı sağlar. Konuşma tanımanın yanı sıra birde bu konuşma sırasında geçen belirli kelimelerin yakalanması durumu mevcuttur. Bu durumda kayıtlı kelimelere denk gelinirse

yazıya çevrilme işlemi gerçekleştirilir. Bu belirli bir komutun yakalanması ile sistemin aktif hale geleceği gibi durumlarda insanlar için faydalı olacaktır.

1.1. PROJENİN TANIMI

Kelime tanıma işlemi, elde kayıtlı olarak tutulan kelimelerin konuşma sırasında geçmesi halinde yakalanıp kayıt edilmesi işlemidir.



Şekil 2: Problem Tanımı

Yukarıdaki şekilde de görüldüğü gibi, mikrofon yardımıyla konuşmacının sesi toplanmalıdır. Bu sesin anlaşılıp incelenmesi gerekmektedir. İnceleme işlemini elimizdeki listede olan kelimelerden geçip geçmediğini anlamak için yapılacaktır. Bütün bu işlemler gerçek zamanlı olarak gerçekleştirilecektir.

Konuşmanın tanınmasını zorlaştıran bazı durumlar mevcuttur [2];

- Yüksek doğruluk oranının sağlanması için programın eğitilmesi gerekmektedir.
- Konuşmacı etrafındaki diğer sesler yani gürültüler sesi etkiler, mikrofon ile ortamdan alınan ses net olmayabilir.
- Programın öğrenmesi için hataların düzeltilmesi gerekir.

Bu gibi sorunlar ile sesin yazıya doğru bi şekilde çevrilmesi zorlaşmaktadır. Bu nedenle bu konu üzerine yani sesin yazıya doğru bir şekilde çevrilmesi üzerine bir çok çalışma yapılmıştır. Buna rağmen hala %100 başarı elde edilememektedir [3].

1.2. PROJENİN GEREKSİNİMLERİ

Kelime tanıma sistemini gerçekleştirmek için öncelikli olarak elimizde mikrofona bir cihaz olmalıdır. Bu cihaz programın çalıştırılacağı bilgisayar olacaktır.

Konuşma tanıma işlemi bir akustik model kullanılarak gerçekleştirilmektedir. Konuşma tanıma akustik modellerinden HMM (Hidden Markov Models – Saklı Markov Modelleri) kullanılacaktır.

Projenin kodlanması için C++ dili kullanılacaktır.

Kodlama IDE si olarak NetBeans 8.2 kullanılacaktır.

Makine öğrenmesi algoritmalarından aşağıdaki kullanılarak proje gerçekleştirilecektir;

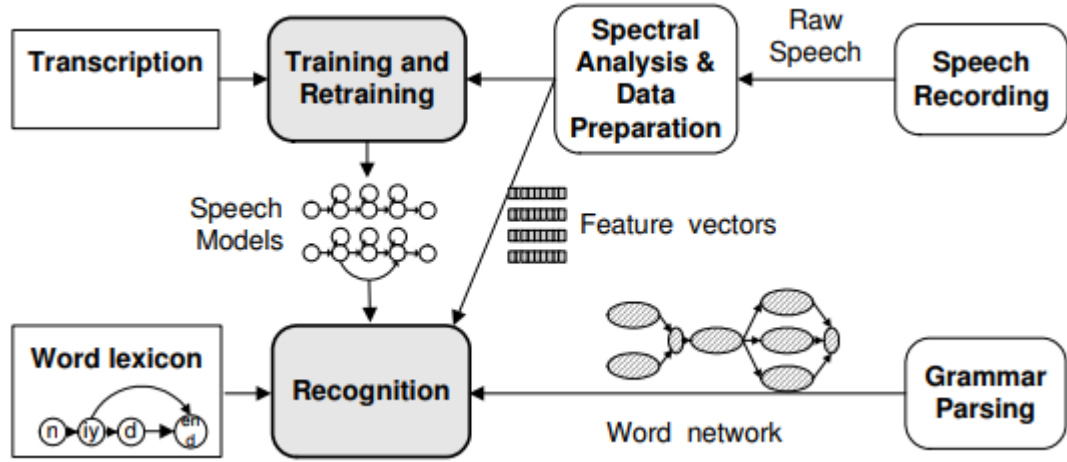
- PCA (Principal Component Analysis – Temel Bileşenler Analizi)

2. TEMEL BİLGİLER VE ÇALIŞMALAR

Kelime tanıma problemini çözebilmek için öncelikle konuşma tanıma problemini anlamamız gerekmektedir. Konuşma tanıma problemi konuşmanın mikrofona yardımıyla toplanarak ilgili cihazda ilgili dile çevrilmesidir [4]. Kısaca konuşmanın yazıya çevrilmesi olarak tanımlanabilir. Konuşma tanıma insan makine etkileşiminde çok fazla kolaylık sağlar. Bilgisayar ile konuşulduğunda ve bilgisayar cevap verdiğinde bilgisayarlara bir kişilik yansıtılır. Bilgisayar artık bir sese sahip olduğu için insan bilgisayar etkileşimi farklı bir boyut kazanır [2].

Konuşma tanıma çok aktif bir çalışma alanıdır. Çok aktif ve kullanışlı bir alan olduğu için de iki temel problemi vardır. Bu problemler doğruluk ve çalışma süresidir [1]. Gerçek zamanlı çalışması istenen bu sistemlerde anında geri dönüş almak önemlidir. Yani bir insan bu sistemin başında konuştuktan sonra bunun yazıya aktarılıp gerekli işlemin anında yapılmasını bekler. Doğruluk ise, söylenenlerin doğru yazıya çevrilmemesi halinde yanlış işlemlerin yapılma olasılığından dolayı temel bir problemidir. “Aç” denildiğinde “At” algılanması dosyanın silinmesi gibi kötü sonuçlar doğurabilir.

Konuşma tanıma problemini çözmek için ortaya atılan konuşma tanıma sistemlerinin ilk amacı teknolojinin bilgi ve iletişim teknolojilerini dünyanın az gelişmiş bölgelerine getirme çabasıdır. Bu nedenle de gerçek zamanlı konuşmaya dayalı kullanıcı ara yüzleri geliştirilmeye başlanmıştır [7].



Şekil 3: Genel Konuşma Tanıma Sistemi [7]

Konuşma Tanıma sistemlerindeki veri akışı şeması yukarıdaki gibidir. Konuşma ifadeleri alınır, konuşma modelleri çıkartılır, tanınacak ifadeleri önce spectral analiz safhasına tabi tutulur ve son olarak da kelime sözlüğü ve dil bilgisi kurallarının da yardımıyla ifade tanınır.

Konuşma tanıma sistemleri birçok alanda kullanılır;

- Araba içi sistemler (In-car systems)

- Saęlık hizmeti
- Askeri alanlarda
- Eęitim
- Gnlk hayatta telefon kullanımı
- Gnlk yařam
- Engelli insanlar iin iletiřim řekli

Konuřma tanıma iřlemini gerekleřtirmek akustik model ve dil modelleri kullanılır. Akustik model, ses sinyali ile dildeki fonetik birimler arasındaki iliřkiyi modeller [13]. Dil modeli, benzer grnen szckler ve ifadeleri arasında ayırım yapmak iin baęlam saęlar [14]. Bu iki model, belirli bir ses segmentine karřılık gelen en st sıradaki kelime dizilerini elde etmek iin bir araya getirilir.

Konuřma tanıma iřlemini gerekleřtirmek iin 4 akustik modelden bahsedilir. Bunlar;

- Hidden Markov Model (HMM – Saklı Markov Modelleri): İstatiksel bir modeldir. Otomatik olarak eęitilebilir ve basit kullanılması mmkndr.
- Dynamic Time Warping (DTW – Dinamik Zaman Atlama): Zamana ve hıza gre deęiřen iki sekans arasındaki benzerlięi len bir algoritmadır. Farklı konuřma hızlarıyla bař edebilmek iin kullanılır.
- Neural Networks (Sinir Aęları): HMM'lere gre istatiksel zelliklerin zellięi hakkında daha az aık varsayımlarda bulunur. Srekli tanıma grevinde nadiren bařarılı iken izole kelimelerin tanınmasında olduka bařarılı sonular vermiřtir.
- End-to-end Automatic Speech Recognition (Utan Uca Otomatik Konuřma Tanıma): Konuřma tanımlayıcının tm bileřenlerini birlikte ęrenirler. Bu eęitim ve daęıtım srecini basitleřtirdięi iin deęerlidir. HMM'ye gre

hafıza kullanımı daha azdır. Böylelikle bulut gibi ağ bağlantısı gerektiren şeylere ihtiyaç duymaz.

Konuşma tanıma sistemi gerçekleştirilirken makine öğrenmesinden yararlanılmaktadır. Makine öğrenmesi, matematiksel ve istatistiksel yöntemler kullanarak mevcut verilerden çıkarımlar yapan, bu çıkarımlarla bilinmeyene dair tahminlerde bulunan yöntem paradigmasıdır [8]. Makine öğrenmesi algoritmaları iki temel öğrenme yöntemine sahiptir. Bunlar; gözetimli (Supervised), gözetimsiz (Unsupervised) ve pekiştirmeli (Reinforcement) öğrenmedir.

Gözetimli öğrenmede, algoritmaya gönderilen eğitim verileri, etiket (label) olarak adlandırılan, istenilen çözümleri içerir [9]. Gözetimli öğrenmenin 2 türü vardır;

- Sınıflandırma (Classification): Her bir gözleme bir sınıf atması yapar [8].
- Eğri Uydurma (Regresyon): Eldeki verilerden sayısal bir değer tahmin edilmeye çalışılır [9].

Gözetimsiz öğrenmede, algoritmaya gönderilen eğitim verileri etiket içermez. Algoritma bir öğretmen olmadan öğrenmeye çalışır [9]. Gözetimsiz öğrenmenin 3 türü vardır;

- Öbekleme (Clustering): Gözlemleri homojen bölgelere ayırır [8].
- İlişkilendirme Kuralları (Association rule learning): Olayların birlikte gerçekleşme durumlarını çözümleyip verileri arası ilişkiler kurar [11].
- Boyut Azaltma (Dimensionality Reduction): Gözlemlerin mevcut özellik sayısını az ve öz hale indirir, bize en iyi öğrenme imkanı sunar [8].

Pekiştirmeli öğrenme ise hit ve deneme yöntemi olarak düşünülebilir. Her eylem için bir Ödül veya Ceza puanı verilir. Makine doğru bir seçenek seçerse, ödül noktasını kazanır ve bunun tersi de geçerlidir [10].

Burada gözetimsiz öğrenme yöntemlerinden Boyut Azaltma algoritmalarından olan PCA (Principal Component Analysis) algoritması kullanılmıştır. PCA veri bilimi için son derece kullanışlı ve istatistiksel bir algoritmadır. Boyut azaltmada çok faydalı bir yöntemdir. Verinin daha küçük alt kümelerle temsil ve analiz edilebilmesini sağlar [6].

Çok boyutlu verilere doğru açıdan bakılarak verilerdeki ilişkilerin açıklanabileceği esasına dayanan PCA, doğru açığı bulmayı amaçlar [6]. Açı dediğimiz şey burada koordinat sistemini ifade etmektedir. Uygun açıdan bakmak, koordinat sistemi kullanılarak verilerin incelenbilmesidir.

Orijinal veriler, “dik olan en büyük varyans yönleri [6]” bulunarak koordinat sisteminde gösterilir. Yani ilk olarak verilerin en büyük değişime sahip olduğu yönde bir eksen seçilir, ardından 2. eksen hem bu eksene dik olacak şekilde hem de yine kalan verilerin en büyük değişime sahip olduğu yönde seçilir. Bu şekilde eksenler seçilmeye devam edilir. Veri seti PCA ile koordinat sistemine yerleştirilerek görsel olarak gösterilmiş olur ve bu sayede incelenmesi de kolaylaşır.

Seçilen eksenler temel bileşenler olarak adlandırılır. Bu temel bileşenler kolerasyonsuzlardır, seçilen ilk temel bileşen toplam değişkenliği en çok açıklayan bileşendir, seçilen ikinci temel bileşen ise kalan değişkenliği en çok açıklayan bileşendir. PCA ile boyut azaltırken veri kaybı genellikle yaşanmamaktadır ama yine de bu garanti edilememektedir [6].

PCA’yı 4 adımda açıklamak mümkündür;

- Değişkenlerin ortalama değerleri bulunur ve verilerden bu ortalamalar çıkartılır. ($y_i - \bar{y}$)
- Kovaryans matrisi hesaplanır. (Cov)

$$\text{Cov} = \begin{bmatrix} \text{Var}(x) & \text{Cov}(x, y) & \text{Cov}(x, z) \\ \text{Cov}(y, x) & \text{Var}(y) & \text{Cov}(y, z) \\ \text{Cov}(z, x) & \text{Cov}(z, y) & \text{Var}(z) \end{bmatrix}$$

Şekil 4: Kovaryans matrisi

$$\text{cov}(x, y) = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{n - 1}$$

Şekil 5: x, y değerleri için kovaryans hesaplanması

$$\text{Var}(X) = \text{Cov}(X, X).$$

Şekil 6: Varyasyon kovaryans eşitliği

- Kovaryans matrisinden öz değerler (eigen values) ve öz vektörler (eigen vectors) bulunur.

Characteristic matrix

$$A - \lambda I = \begin{bmatrix} 1 & 2 \\ 3 & -4 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 - \lambda & 2 \\ 3 & -4 - \lambda \end{bmatrix}$$

Characteristic equation

$$|A - \lambda I| = (1 - \lambda)(-4 - \lambda) - (2)(3) = \lambda^2 + 3\lambda - 10 = 0$$

Eigenvalues: $\lambda_1 = -5, \lambda_2 = 2$

Şekil 7: Öz değerlerin hesaplanması

Eigenvalues

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & -4 \end{bmatrix} \quad \begin{matrix} \lambda_1 = -5 \\ \lambda_2 = 2 \end{matrix}$$

Determine eigenvectors: $\mathbf{Ax} = \lambda \mathbf{x}$

$$\begin{aligned} x_1 + 2x_2 &= \lambda x_1 & \Rightarrow & (1 - \lambda)x_1 + 2x_2 = 0 \\ 3x_1 - 4x_2 &= \lambda x_2 & \Rightarrow & 3x_1 - (4 + \lambda)x_2 = 0 \end{aligned}$$

Eigenvector for $\lambda_1 = -5$

$$\begin{aligned} 6x_1 + 2x_2 &= 0 \\ 3x_1 + x_2 &= 0 \end{aligned} \quad \Rightarrow \quad \mathbf{x}_1 = \begin{bmatrix} -0.3162 \\ 0.9487 \end{bmatrix} \text{ or } \mathbf{x}_1 = \begin{bmatrix} 1 \\ -3 \end{bmatrix}$$

Eigenvector for $\lambda_1 = 2$

$$\begin{aligned} -x_1 + 2x_2 &= 0 \\ 3x_1 - 6x_2 &= 0 \end{aligned} \quad \Rightarrow \quad \mathbf{x}_2 = \begin{bmatrix} 0.8944 \\ 0.4472 \end{bmatrix} \text{ or } \mathbf{x}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

Şekil 8: Öz vektörlerin hesaplanması

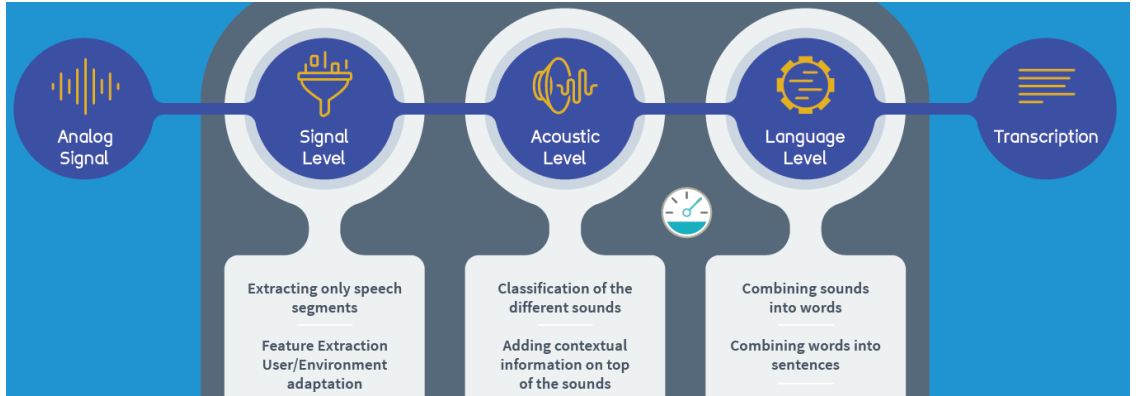
- Son olarak bulunan öz değerler kullanılarak n boyutlu veri n-1 boyuta indirgenecektir. Bunun için de kovaryans değerlerinden oluşan matris ile öz vektörler çarpılır ve indirgenmiş yeni veri seti elde edilir.

3. TASARIM



Şekil 9: Genel Tasarım Planı

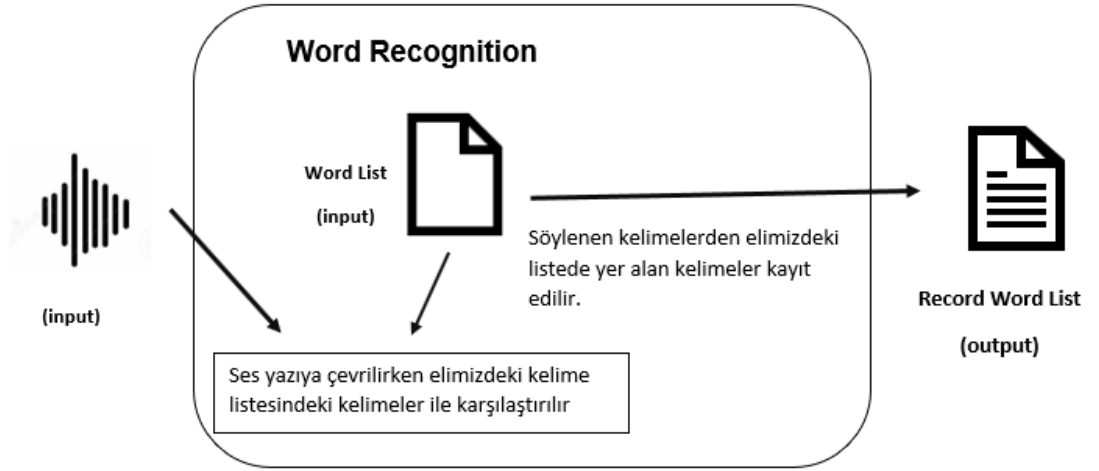
Kullanılan cihazın yani uygulamanın çalıştığı cihazın mikrofonu ile toplanan ses konuşma tanıma aşamalarından geçerek ilerler ve bir sonuç üretilir. Burada üretilen sonuç listede olan ve konuşmada geçen ortak kelimelerin bir listesi olmaktadır.



Şekil 10: Konuşma tanıma işlemi

Kelime tanıma işlemi de konuşma tanıma işlemi gibi ilerlemektedir. Öncelikle konuşma yakalanmaktadır. Yakalanan sinyal bir özellik çıkarma işleminden geçmektedir. Farklı sesler sınıflandırılmaktadır. Yani gürültü temizlenmektedir. Ve son olarak da sesler kelimelerle eşleştirilerek geçen kelimeler yakalanmaktadır. Burada konuşma tanımadan farklı olarak sadece listemizde var olan kelimeler ile eşleştirmeler yapılmaktadır. Böylelikle listemizde var olan ve konuşmada geçen kelimelerin yakalanması gerçekleştirilir.

Aşağıdaki şekilde yakalanan sesteki kelimelerin listedeki kelimeler ile karşılaştırılacağı belirtilmiştir.

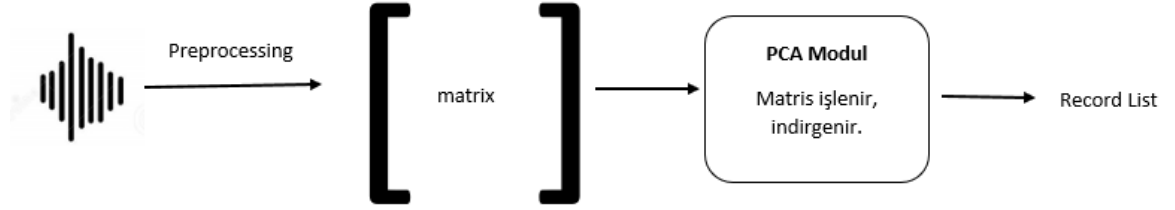


Şekil 11: Kelime Tanıma (Word Recognition) Tasarım Planı

Karşılaştırılan kelimelerin seste geçiyor olması durumunda bu kelimeler bir kayıt listesine eklenmektedir. Bu kayıt listesi ise çıktı olarak kullanıcıya gösterilmektedir.

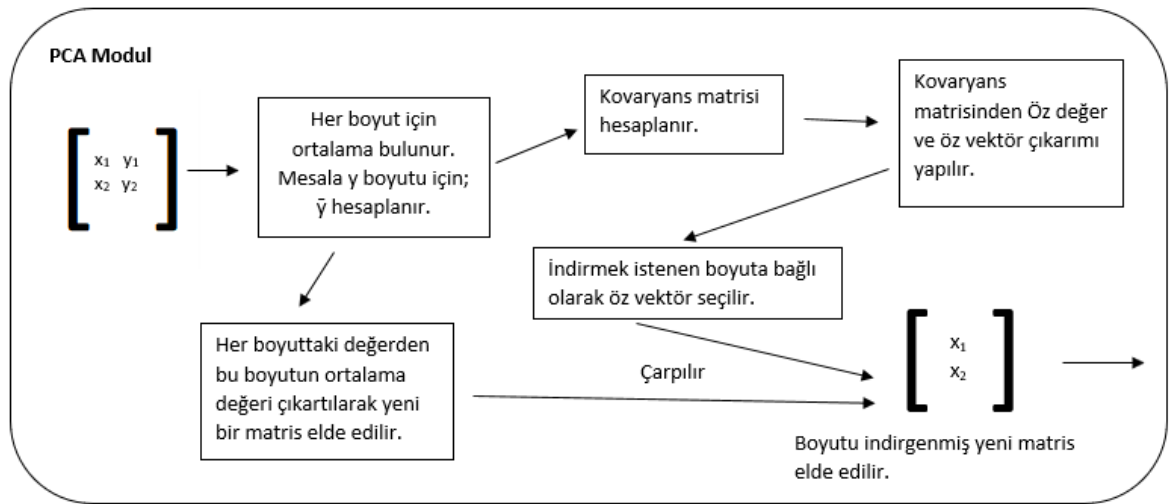
3.1. PCA ile TASARIM

PCA ile proje gerçekleştirilirken ses öncelikle bir matrise, vektöre de olabilir, çevrilmek için ön işleminden geçmektedir. Oluşturulan matris ise PCA modülüne gelir ve burada matrise boyut indirgeme işlemi uygulanır. Ardından kontrol edilerek kayıt listesinde yer alan kelimelerin geçip geçmediği bulunur.



Şekil 12: PCA ile gerçekleştirilen tasarım planı

PCA modülü, matrisi alarak bu matrisin boyutunu küçültür. Bunu öncelikle matrisin her sütunu yani boyutu için bir ortalama değer hesaplar. Ardından verilerden bu ortalama değerler çıkartılarak bir matris elde edilir. Başta PCA modülüne gelen matris ve hesaplanan ortalama değerler kullanılarak bir kovaryans matrisi oluşturulur. Kovaryans matrisi üzerinde işlemler yapılarak öz değerler ve öz vektörler bulunur. Öz vektörlerden seçim yapılır. Bu seçim indirgenmek istenen boyuta göre gerçekleştirilir. Son olarak da seçilen öz vektör ve kovaryans matrisinden önce oluşturulan matris çarpılarak boyut indirgenir.



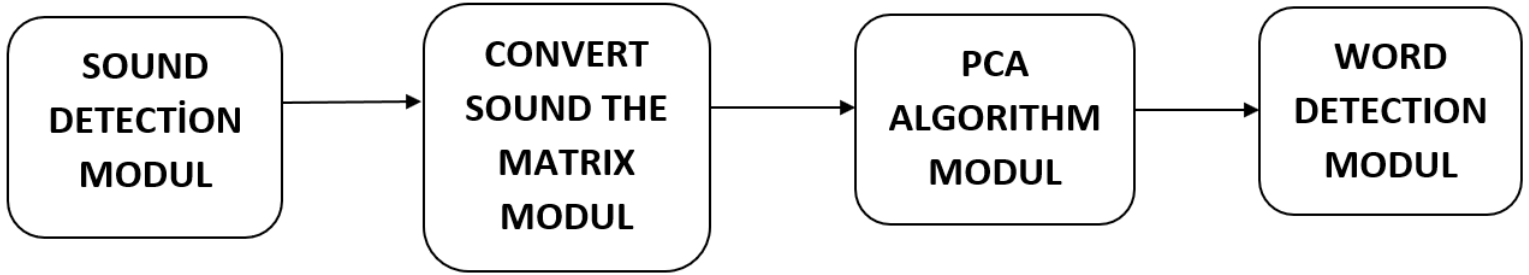
Şekil 13: PCA Modülü

PCA algoritmasının nasıl ilerlediği ayrıntılı olarak ‘2. TEMEL BİLGİLER ve ÇALIŞMALAR’ başlığı altında anlatılmıştır.

4. UYGULAMA

4.1. PCA ile GERÇEKLEME

Kelime tanıma işlemi PCA algoritması ile gerçekleştirilirken aşağıdaki şekilde gösterildiği gibi bir yol izlendi.



Şekil 14: PCA algoritması kullanılarak kelime tanıma işleminin gerçekleştirilmesi

Ses Yakalama Modülü (Sound Detection Modul) ‘nde mikrofona konuşan konuşmacının sesinin algılaması ve ardından konuşmanın kaydedilmesi işlemleri gerçekleştirildi.

Sesin Matrise Dönüştürülmesi Modülü (Convert Sound the Matrix Modul) ‘nde yakalanan sesin Matrise çevrilmesi işlemi gerçekleştirildi. PCA algoritması matris üzerinde işlem yaptığı için ses dalgası matrise çevrildi.

PCA Algoritması Modülü (PCA Algorithm Modul) ‘nde matrise çevrilen sesin boyutu indirildi. PCA Algoritması Modülü ‘3. TASARIM’ başlığı altında yer alan ‘3.1. PCA ile TASARIM’ başlığı altında ayrıntılı bir şekilde açıklanmıştır.

Kelime Yakalama Modülü (Word Detection Modul) ‘nde elimizdeki listede yer alan kelimeler ile yakalanan sesteki kelimeler karşılaştırılmıştır. Elimizdeki listede yer alan kelimelerden bu seste kullanılmış olanlar var ise kayıt edilmiştir. Kayıt işlemi bittikten sonra kullanıcıya elimizdeki listede yer alan kelimeler içinden kullandığı kelimelerin listesi gösterilmiştir.

Bu modüller gerçekleştirirken temel kütüphaneler dışında ekstra kütüphaneler kullanılmamıştır. Bunun yerine bütün kütüphaneler implement edilmiştir.

Ses Yakalama Modülü (Sound Detection Modul) ve Sesin Matrise Dönüştürülmesi Modülü (Convert Sound the Matrix Modul) 'lerinde sesin yakalanması için gerekli olan ve sese ait verilerinin tutulmasını sağlayan yapılar bir ses sınıfı oluşturularak gerçekleştirilmiştir.

PCA Algoritması Modülü (PCA Algorithm Modul) için gerekli olan matris kütüphanesi, bu matris için gerekli olan kovaryans, matris çıkarması, matris çarpması, öz değer ve öz vektör bulunması işlemleri bu matris kütüphanesi ile gerçekleştirilmiştir.

Kelime Yakalama Modülü (Word Detection Modul) için ise bir arama algoritması gerçekleştirilmiştir.

5. TARTIŞMA VE SONUÇ

Speech Recognition (Konuşma Tanıma) işlemi günümüzde çok yaygın kullanılan bir işlem olmuştur. Bütün akıllı cihazlarda insanların aradığı ilk özelliklerden olmuştur. Bu nedenle kelime tanıma işlemi de işleri kolaylaştıracağı için gerçekleştirilen bir sistemdir.

Bu sistemi gerçekleştirirken başta belirlenen başarı kriterlerinden 2 algoritma ile gerçekleştirilememiştir. En az 2 ve en çok 4 kişilik ortamlarda çalışıyor olması şartı için gerekli testler bitirilemediği için henüz bir şey söylenememektedir. Kelime tanıma başarısının %90 olması işlemi ise var olan ses dosyalarında sağlanırken gerçek zamanlı ses alma işleminde bir sorun çıktığı için henüz belirlenmemiştir.

Gelecekte bu sistemin geliştirilmesi ile video ve ses dosyalarında istenilen kelimelerin geçmesi durumuna göre bir arama işlemi gerçekleştirilebilir. Yani sesi dosyasında mesele “toplum” ve “değer” kelimelerinin geçip geçmemesine bakılarak

ses dosyaları bulunabilir. Bu da insanlar için birçok fayda sağlar. Bu sistem ile, adı aradığımız kritere yakın olmasına rağmen içeriği yakın olamayan video ve ses dosyalarının dinlenilerek gereksiz zaman kaybı sağlanması önlenabilir. Bunun gibi birçok konuda da fayda sağlayacağına inanmaktayım.

KAYNAKLAR

- [1] Andreas Hagen, Daniel A. Connors, Bryan L. Pellom, *The Analysis and Design of Architecture Systems for Speech Recognition on Modern Handheld-Computing Devices*, Newport Beach, California, USA, 2003.
- [2] Christine Flounders, *"Are you There Margaret? It's me, Margaret": Speech Recognition as a Mirror*, New York University, 2001.
- [3] Anuj Mohamed, K.N. Ramachandran Nair, *Continuous Malayalam speech recognition using Hidden Markov Models*, India, 2010.
- [4] Sai Sawant, Mangesh Deshpande, *Isolated Spoken Marathi Words Recognition using HMM*, Electronics & Telecommunication Engineering, Vishwakarma Institute of Technology, Pune, India, 2018.
- [5] Özbey Can, Bayar Salih, *Otomatik Ses Tanıma: Türkçe için Genel Dağarcıklı Akustik Model Oluşturulması ve Test Edilmesi*, İdea Teknoloji Çözümleri Ar-Ge Merkezi, Boğaziçi Üniversitesi, İstanbul, 2017.
- [6] Cömert Zafer, *Temel Bileşenler Analizine Genel Bir Bakış*, 2015
- [7] Sergiu Nedeveschi, Rabin K. Patra, Eric A. Brewer, *Hardware Speech Recognition for User Interfaces in Low Cost, Low Power Devices*, Department of Electrical Engineering and Computer Science University of California at Berkeley, Anaheim, California, USA, 2005.
- [8] Aladağ Emre, <https://www.emrealadag.com/makine-ogrenmesi-nedir.html> [Ziyaret Tarihi: 21 Ekim 2019].
- [9] Koptur Murat, <https://makineogrenimi.wordpress.com/2017/05/25/makine-ogrenmesi-teknikleri/> [Ziyaret Tarihi: 21 Ekim 2019].

- [10] Nar Kadir, <https://kadirnar.com/post/makine-ogrenmesi-nedir-/21> [Ziyaret Tarihi: 21 Ekim 2019]
- [11] Hatipoğlu Ekrem, <https://medium.com/@ekrem.hatipoglu/machine-learning-association-rule-mining-birliktelik-kural-%C3%A7%C4%B1kar%C4%B1m%C4%B1-apriori-algorithm-4326b8f224c3> [Ziyaret Tarihi: 21 Ekim 2019].
- [12] Tosun Gökhan, <http://www.gokhantosun.com/2019/02/14/pca-principal-component-analysis-algoritmasi-nedir/> [Ziyaret Tarihi: 16 Ekim 2019].
- [13] Wikipedia,
<https://www.wikizeroo.org/index.php?q=aHR0cHM6Ly9lbi53aWtpcGVkaWEub3JnL3dpa2kvQWNvdXN0aWNfbW9kZWw> [Ziyaret Tarihi: 10 Ekim 2019]
- [14] Wikipedia,
<https://www.wikizeroo.org/index.php?q=aHR0cHM6Ly9lbi53aWtpcGVkaWEub3JnL3dpa2kvTGFuZ3VhZ2VfbW9kZWw> [Ziyaret Tarihi: 10 Ekim 2019]

EKLER

A. Lisans Bitirme Projesi Konusu Bildirme Formu

**T.C.
GEBZE TEKNİK ÜNİVERSİTESİ
MÜHENDİSLİK FAKÜLTESİ
BİLGİSAYAR MÜHENDİSLİĞİ
LİSANS BİTİRME PROJESİ KONUSU BİLDİRME FORMU**

ÖĞRENCİ ADI SOYADI : Gözde DOĞAN

ÖĞRENCİ NO : 131044019

İMZA:

PROJE KONU BAŞLIĞI: REAL TIME WORD RECOGNITION

PROJENİN AMACI:

Projede konuşma sırasında geçen kelimeler elimizdeki kelimeler ile uyuyorsa yakalanacaktır. Amacımız konuşma tanıma sisteminin amaçları ile aynı amacı taşımaktadır. Yani amacımızın günlük hayatı kolaylaştırmak olduğunu söyleyebiliriz.

FAYDALANILACAK KAYNAKLAR:

[1] Zafer Cömert, *Temel Bileşenler Analizine Genel Bir Bakış*, 2015

[2] Christine Flounders, *"Are you There Margaret? It's me, Margaret": Speech Recognition as a Mirror*, New York University, 2001

PROJE DANIŞMANI: Dr. Uraz Cengiz TÜRKER

İMZA:

BÖLÜM BAŞKANI: Prof. Dr. Yusuf Sinan AKGÜL

İMZA: