

# Seslerin Görüntüsü

## Image of Sounds

Merve Karagöz

Bilişim Sistemleri Mühendisliği

Kocaeli Üniversitesi

karagoz.merve15@gmail.com

### Özet

Görüntü işleme teknikleri kullanarak yüklemesi yapılan görüntü üzerindeki metni okuyan ve seslendiren bir program hazırlanmıştır. Bu programın amacı görme engelli bireylerin gün içerisinde karşılaştıkları zorlukları minimum dereceye indirmek ve yaşadıkları zorlukların günlük yaşantılarına bir engel olmamasını sağlamaktır. Programda bulunan canlı görüntü üzerindeki metni okuma özelliği görme engelli bireylerin gidecekleri yerleri, yol üzerinde bulunan metinlerin seslendirmesini sağlamaktadır. Program yalnızca görme engelli bireylere değil görüntü üzerinde bulunan herhangi bir yazıyı metni çevirmek isteyen kişilere de hitap edecek şekilde hazırlanmıştır.

Anahtar kelimeler: Görüntü işleme, metin okuma, seslendirme.

### Abstract

A program has been prepared that reads text and voices text when image processing is displayed. The aim of this program is to minimize the difficulties faced by visually impaired individuals during the day and to ensure that the difficulties they experience are not an obstacle to their daily lives. The reading under live view feature in the program enables visually impaired individuals to vocalize the texts outside. The program is designed to appeal not only to visually impaired people, but also to people who want to translate any text in the image.

Keywords: Image processing, text reading, vocalization.

### 1. Giriş

Birleşmiş Milletler Engelli Hakları Sözleşmesine (BMEHS) göre engelli birey, çeşitli engellerle karşılaşması halinde diğerleriyle eşit bir şekilde topluma tam ve etkili şekilde katılmalarını engelleyen uzun süreli fiziksel, zihinsel, ruhsal ve duyuşsal rahatsızlığı olan kişiler şeklinde tanımlanmaktadır [1]. Kısacası herhangi bir engeli bulunmayan insanlar, engelli bireyleri kısıtlayıcı yapılar, söylemler ve hareketlerde bulunmadığı sürece yeti kaybı yaşayan bireyler için hayat çok daha kolay ve engelsiz bir hale getirilebilir. Bu düşüncelere bağlı olarak Türkiye’de yasal mevzuatta bulunan özürü ve

sakat gibi kavramlar kaldırılmış ve engelli kavramının kullanılmasına ilişkin düzenleme yapılmıştır [2].

Türkiye’de Aile, Çalışma ve Sosyal Hizmetler Bakanlığı Ulusal Engelli Veri Sistemi’nde kayıtlı 281 bin 439 görme engelli birey bulunmaktadır [3]. Gözleri görmeyen bir bireyi günlük yaşantısında, sosyal hayatında, iş hayatında veya eğitim-öğretim hayatında herhangi bir engelle karşılaştırılmadığı sürece yaşadığı yeti kaybı onun engeli olmaktan çıkacaktır. Yapılan bu proje de bu ideolojiden yola çıkılarak seçilmiş ve gerçekleştirilmek istenmiştir.

Yapılan projede öncelikli olarak görme engelli bireylerin temel özgürlüklere erişim konusunda karşılaşılabilecekleri sorunlar gözlemlenerek çözümler üretilmiştir. Başlıca sorunlardan birinin günlük hayatlarında ulaşım ihtiyaçları doğrultusunda ve yazılı belgelerin okunmasında yaşanabilecek zorluklar olduğu tespit edilmiştir. Bu sorunlara çözüm oluşturmak amacıyla bir resim üzerinden veya canlı olarak sağlanan telefon kamerası bağlantısıyla görüntü üzerinden metni algılayan ve seslendirmeye yarayan bir program yapılmaya çalışıldı.

### 2. Proje Planı

#### 2.1. Projenin Amacı

Projenin amacı, görme engelli bireylerin günlük yaşantısını kolaylaştırmak, metinlerin kamera aracılığıyla ortamdan dijital ortama aktararak seslendirilmesini sağlamaktır.

#### 2.2. Kullanım Alanları

Proje görme engelli bireylere ve görüntüyü dijital ortama aktarmayı isteyen insanlara hizmet etmek amacıyla oluşturulmuştur. Özellikle görme engelli bireylerin eğitim, iş ve sosyal hayatlarında çevrelerinde bulunan yazılı uyarıların, belgelerin veya kitapların dijital ortama aktarılıp anlamlandırılarak kolayca seslendirmelerine yarayacaktır.

### 3. Projede Kullanılan Yöntem ve Metotlar

Proje görüntü işleme yöntemiyle gerçekleştirilmiştir. Öncelikle bir fotoğrafın sisteme yüklenmesi sağlanarak fotoğraf üzerindeki metni algılayıp bir çıktı üretmesi ve sonrasında bu çıktıyı bilgisayara kaydederek seslendirmesi sağlanmıştır. Daha sonra canlı olarak görüntü alabilmek amacıyla telefon uygulaması olan IP Webcam yüklenerek bilgisayarla iletişimi sağlanmıştır. Canlı olarak alınabilen görüntü üzerindeki metinler çıktı haline getirilerek yazdırılmış ve seslendirme işlemi yapılmaya çalışılmıştır.

#### 3.1. Görüntü İşleme

Projenin ortaya çıkmasındaki temel yöntem görüntü işleme teknikleridir. Görüntü işleme, bir görüntüyü dijital forma dönüştürmeyi amaçlayarak görüntü sinyallerini yeniden yapılandırır. Elde edilen görüntüden anlamlı ifadeler çıkarılmasını sağlar. Bu işlemlerin gerçekleştirilebilmesi için görüntü üzerinden piksellere matematiksel işlemler uygulanır. Görüntü işleme temel olarak aşağıdaki üç adımı içerir.

- Görüntünün optik tarayıcı veya dijital fotoğraflarla alınması.
- Veri sıkıştırma, görüntü iyileştirme ve uydu fotoğrafları gibi insan gözü olmayan lekelenme kalıplarını içeren görüntüyü analiz etme-kullanma.
- Çıktı, sonuçları görüntü analizine dayalı olarak değiştirilerek, kullanıma hazır hale getirilmesi.

Görüntü işlemenin amaçları şu şekilde açıklanabilir.

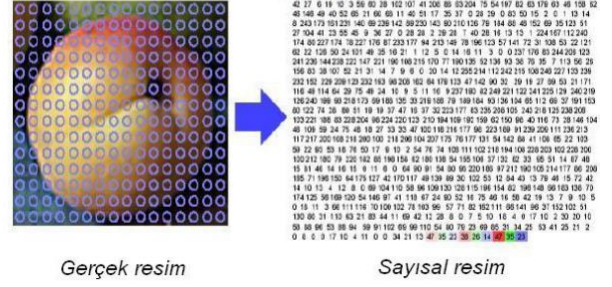
- Görselleştirme ile bir kısmı görünmeyen, deforme olmuş ve görünmesi zor görüntülerin görünür hale getirilmesi,
- Görüntü iyileştirme ve keskinleştirme ile lekeleri ortadan kaldırılarak görüntülerin iyileştirilmesi,
- Görüntünün elde edilmesi ile yüksek çözünürlüklü bir görüntü ortaya çıkarılması,
- Şekil, desen veya istenilen herhangi bir şeyin tanımlanması ve ayırt edilmesidir.

Görüntü işleme uygulama alanları konuları şu şekildedir [4]:

- Yapay Sinir Ağları
- Dalgacık Dönüşümü
- Markov Rasgele Alan Süzgeçleri
- İteratif Hücresel Resim İşleme Algoritması
- Yönlendirme Süzgeçleri
- Genetik Algoritma
- Görüntü İletimi
- Bulanık Mantı

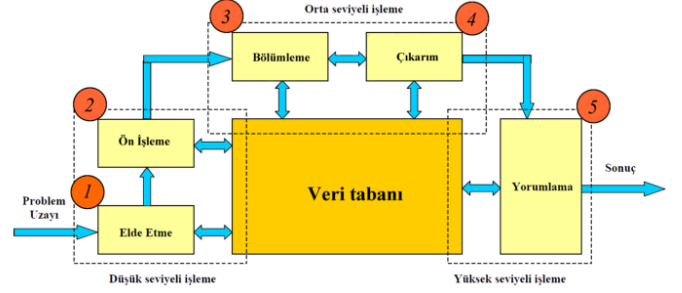
##### 3.1.1. Görüntü işleme yöntemleri

Görüntü İşleme için iki farklı yöntem kullanılmaktadır. Bu yöntemler, dijital (sayısal) ve analog görüntü işleme yöntemleridir [5]. Resim 3.1.1 'de analog bir görüntünün sayısal karşılığına yer verilmiştir.



Resim 3.1.1: Analog görüntü ve sayısal karşılığı

Dijital (sayısal) görüntü işleme yöntemi, analog bir görüntünün sayısal bir biçime dönüştürülmesi ve sonrasında onarma, iyileştirme, keskinleştirme ve sınıflandırma gibi amaçlar doğrultusunda sayısal bilgisayarlar aracılığıyla işlenmesi işlemleriyle ilgilenen yöntemdir. Sayısal görüntüleme işlemi temel basamaklarına Şekil 1 'de yer verilmektedir.

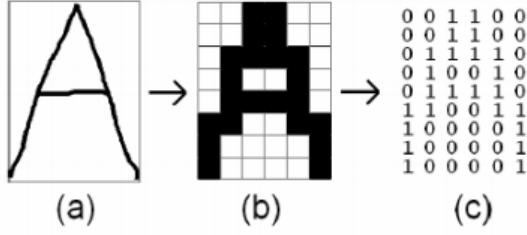


Şekil 1: Sayısal görüntüleme işlemi temel basamakları

Şekil 1 'de görülen terimlerin açıklamaları şu şekildedir:

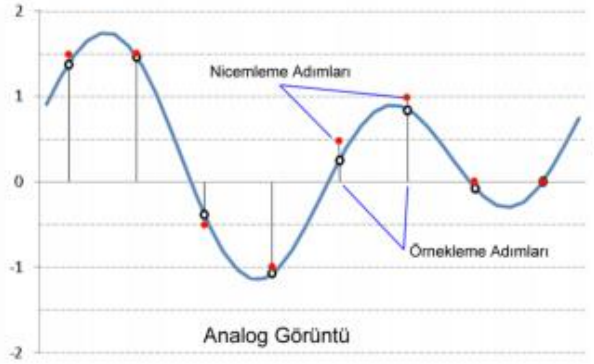
- Elde etme işlemi, sayısal kamera ile sayısal görüntüye ulaşılması işlemidir.
- Ön işleme, sayısal görünümün başarılı sonuçlar elde etme amacıyla görüntünün iyileştirme, onarma ve sıkıştırma işlemlerinden geçmesidir.
- Bölümleme işlemi, sayısal görüntü işleme basamaklarının en zorudur. Bir görüntü içerisindeki ilgilenilen alan harici diğer özelliklere sahip alanlardan ayrılmasıdır.
- Çıkarım işlemi, görüntüden elde edilen istenilen bilginin içerisindeki ayrıntıların ön plana çıkarılmasını sağlayan işlemidir.
- Yorumlama işlemi, yüksek seviyeli görüntü işleme grubunda yer almaktadır. Elde edilen ham görüntünün sınıflandırılmasında rol alan işlemidir.

Sayısal bir görüntünün en temel parçası piksel olarak adlandırılmaktadır. Resim 3.1.2 'de görüldüğü üzere sayısal görüntü işleme görsel karakterlerin sayısallaştırılmasında piksellerden yararlanılmaktadır.

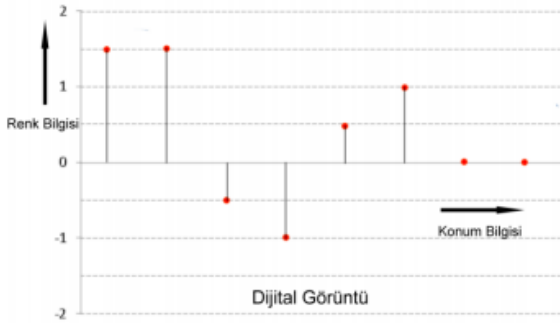


Resim 3.1.2: Görsel karakterlerin sayısallaştırılması örneği

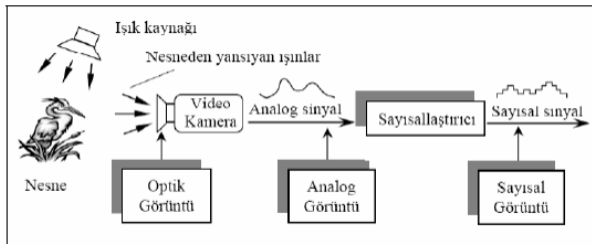
Analog görüntünün işlenebilmesi için sayısal görüntüye çevrilmesi gerekmektedir. Bu işlemde görüntü üzerinde konuma yönelik örnekleme ve renk derinliğine yönelik niceleme yapılmaktadır. Şekil 2 ve Şekil 3’de görüldüğü üzere x ekseninde bulunan her değer örnekleme, y ekseninde ise alınabilecek en yakın değer ise niceleme olarak adlandırılmaktadır. Resim 3.1.3 ‘de detaylı bir sayısal görüntüye çevrilme işlemine yer verilmiştir.



Şekil 2



Şekil 3



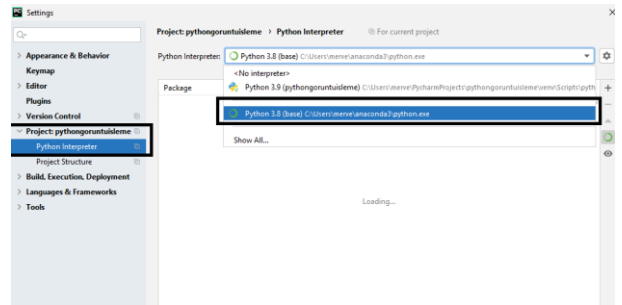
Resim 3.1.3: Görüntü sayısallaştırma aşamaları

## 4. Proje Yapımında Kullanılan Araçlar

Proje yapımında PyCharm geliştirme ortamı ve canlı görüntü bağlantısını sağlayabilmek amacıyla IP Webcam uygulaması kullanılmıştır. Kullanılan bu araçlara aşağıda yer verilmiştir.

### 4.1. PyCharm

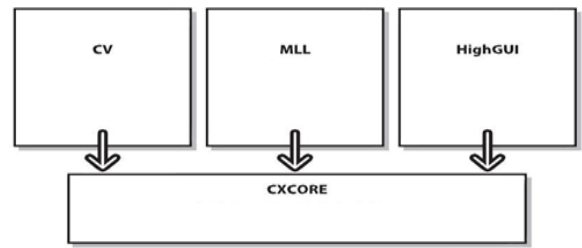
PyCharm, Python geliştirme ortamıdır. Çapraz platform işletim sistemi kullanmaktadır. Java ve Python dillerinde yazılarak Temmuz 2010 yılında ilk kez yayınlanmıştır. PyCharm bütün işletim sistemlerinde çalışılabilir biçimde geliştirilmiştir. Professional ve Community olarak 2 sürümü mevcuttur. Bu projede Community sürümü ile çalışılmıştır. PyCharm projesine başlamadan önce Anaconda Navigator üzerinden Python 3.8 sürümü kullanılacak şekilde ayarlanmıştır (Resim 4.1.1).



Resim 4.1.1: Anaconda Navigator Python 3.8 sürümü ayarlama

#### 4.1.1. OpenCV kütüphanesi

OpenCV (Open Source Computer Vision Library, anlamı Açık Kaynak Bilgisayar Görüşü Kütüphanesi), genellikle tüm görüntü işleme işlemlerinde kullanılır. Görüntü işleme alanında öncü bir açık kaynak kodlu kütüphanedir. İlk olarak 1999 yılında Intel tarafından geliştirilmeye başlanılmıştır. C++ diliyle yazılmıştır. Birincil arayüzü C++ içindir. Python, Java, MATLAB için kullanımı geliştirilmiştir. OpenCV ‘nin 4 temel bileşeni Şekil 4 ‘de gösterilmektedir.



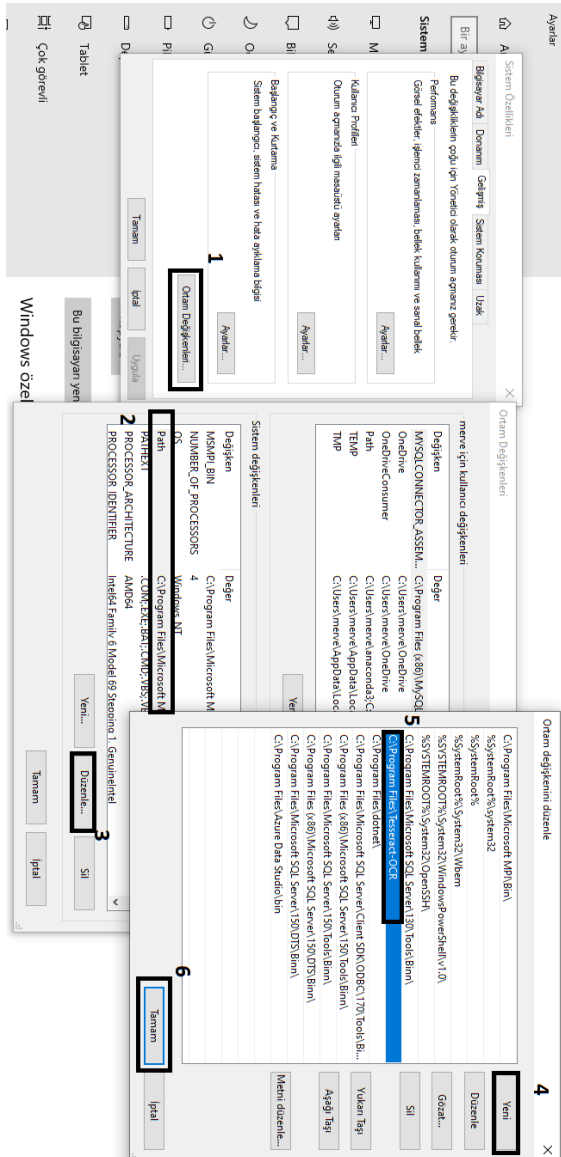
Şekil 4: OpenCv ‘nin 4 temel bileşeni

OpenCV kütüphanesi proje dahil edilmek için terminal kısmına “py -3.8 -m pip install opencv-python” kodları yazılması gerekmektedir “python -m pip install --upgrade pip” şeklinde bir öneriyle alınan hata karşısında bu kodlar terminalde çalıştırılarak openCV kütüphanesi kurulmuş olur.

#### 4.1.2. Tesseract kütüphanesi

Tesseract, OCR (Optical Character Recognition, Optik Karakter Tanıma) işlemi yapan bir kütüphanedir. Bir görüntüde yer alan yazıların bilgisayar ortamına aktarılması için kullanılmaktadır. Yapılan proje, görüntü üzerindeki yazıları sanal ortama aktararak seslendirme gayesi taşıdığından dolayı bu kütüphane uygun görülmüştür.

Kütüphane yüklemesi “https://github.com/UB-Mannheim/tesseract/wiki” adresinden gerçekleştirilmiştir. Bilgisayara kurulumu gerçekleştirilen kütüphanenin kullanılması için Resim 4.1.2.1 ‘deki işlemler sırasıyla gerçekleştirilmelidir. Bu işlemler sonrası proje dosyası üzerinde “import pytesseract” komutu yazılarak Tesseract kütüphanesinin projeye yüklenmesi gerçekleştirilmiştir.



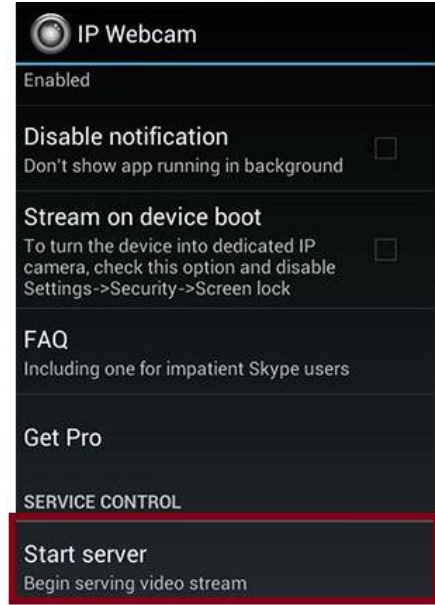
Resim 4.1.2.1: Tesseract Kurulumu

#### 4.1.3. gTTS kütüphanesi

Proje içerisinde okunan metnin seslendirilmesi istenildiğinde gTTs (Google Text-to-Speech, anlamı Google Metin-Konuşma) kütüphanesi kullanılmıştır. Google Translate'in metin okuma API'si (API, uygulama programlama ara yüzü) ile ara yüz oluşturmak için bir Python kütüphanesidir.

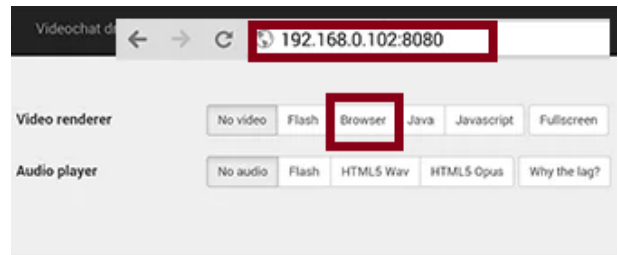
#### 4.2. IP Webcam

IP Webcam, Google Play Store Aracılığıyla android telefona yüklenerek bilgisayar ve telefon arasında kamera bağlantısı yapmaya yaran bir uygulamadır. Resim 4.2.1 'de görüldüğü üzere telefon üzerinden kurulan programa Start Server denilerek kamera başlatılır ve ip adresi alınır.



Resim 4.2.1: IP Webcam uygulaması

Start Server sonrası elde edilen ip adresi tarayıcı üzerinden aratılır (Resim 4.2.2) ve Browser seçeneğiyle telefondaki görüntüyü bilgisayarda görme işlemi gerçekleştirilir.

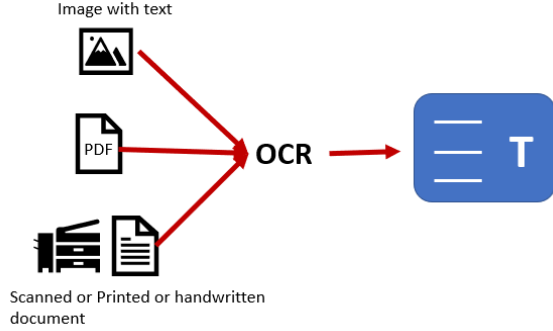


Resim 4.2.2: Tarayıcı üzerinden android kameraya erişim sağlama

Tarayıcı adres çubuğundan url kopyalanarak proje içerisinde belirtilir. Gerekli kodlar yazılarak proje üzerinden bağlantı sağlanmış olur.

## 5. Metin Tanıma Sistemi

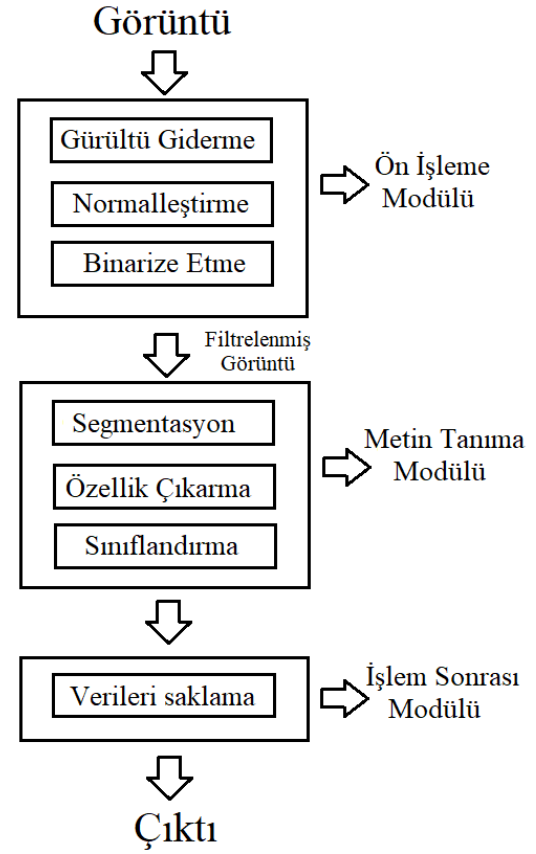
Metin tanıma sistemi, tarayıcıdan elde edilen bir çıktının veya el yazısı içeren bir belgenin, pdf formatında programa yüklenmiş olan belgelerin veya üzerinde metin bulunan bir fotoğrafın sisteme yüklenip bu metinlerin bulunarak dijital ortamda gösterilmesini sağlar, Resim 5.1 'de metin tanımlama sisteminin genel hali gösterilmiştir.



Resim 5.1

Proje içerisinde kullanılan metin tanıma sisteminin genel mimarisi Şekil 5 'de gösterildiği gibi açıklanabilir. Bir metin tanıma sistemi, bazı metin bilgilerini içeren görüntü biçiminde bir girdi alır. Bu sistemin çıktısı elektronik formattadır, yani görüntüdeki metin bilgileri bilgisayar tarafından okunabilir biçimde saklanır.

Metin tanıma sistemi aşağıdaki modüllere ayrılabilir: Ön işleme modülü, metin tanıma modülü ve son işlem modülü. Ön işlem modülü (pre-processing module); gürültü giderme (noise removal), normalleştirme (normalization) ve binarize etme (binarization) işlemlerinden oluşmaktadır. Metin tanıma modülü (text recognition module); segmentasyon (segmentation), özellik çıkarma (feature extraction) ve sınıflandırma (classification) işlemlerinden oluşmaktadır. Son olarak işlem sonrası modülü (post-processing module) verileri saklama işleminden oluşmaktadır. Her modül aşağıda ayrıntılı olarak açıklanmıştır.



Şekil 5: Metin tanıma mimarisi

### 5.1. Ön İşlem Modülü

Belgeler genellikle optik tarayıcı tarafından taranır ve bir resim formuna dönüştürülür. Bir resim, piksel olarak da bilinen resim öğelerinin birleşiminden meydana gelmektedir. Proje de kullanılan görüntüler ön işlem modülü sayesinde daha fazla analiz edilebilir, böylece önemli bilgiler alınabilir. Bu nedenle, giriş görüntüsünün kalitesini iyileştirmek için, gürültü giderme, normalleştirme, binarize etme gibi görüntünün iyileştirilmesi için işlemler gerçekleştirilir.

#### 5.1.1. Gürültü Giderme

Gürültü giderme en önemli ön işlem modülü işlemlerinden biridir. Görüntünün kalitesi ön işlem nedeniyle artacak ve görüntülerde daha iyi metin tanıma için tanıma sürecini etkileyecektir. Ve metin tanıma işleminin sonunda daha doğru çıktı üretilmesine neden olur. Resim 5.1.1.1 gürültü giderme işlemine bir örnektir.





Resim 5.1.1.1



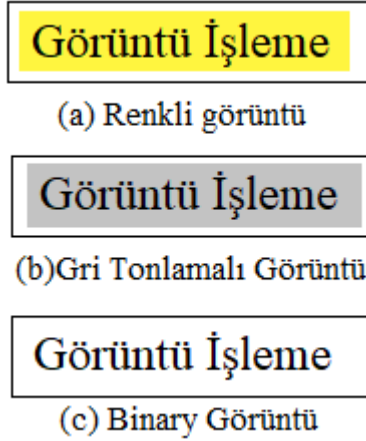
Resim 5.1.3.1

### 5.1.2. Normalleştirme

Normalleştirme, metin tanıma için önemli ön işleme işlemlerinden biridir. Normalleştirme, görüntü içerisinde tek tip boyut, eğim ve dönüş karakterleri elde etmek için uygulanır.

### 5.1.3. Binarize Etme

Binarize etme, metin tanıma için önemli ön işleme modülü işlemlerinden biridir. Yazdırılan bir belge önce taranır ve gri tonlamalı bir görüntüye dönüştürülür. Binarize etme yöntemi, gri tonlamalı görüntülerin ikili görüntülere dönüştürüldüğü bir tekniktir. Metnin arka plandan bu şekilde ayrılması bölümlene gibi bazı işlemler için gereklidir. Şekil 6 'da bir metin görüntüsünün renkli bir görüntüsü (a), gri görüntüsü (b) ve ikili görüntüsü (c) gösterilmektedir. Resim 5.1.3.1 'de gri tonlamalı ve renkli bir resim örneğine yer verilmiştir.



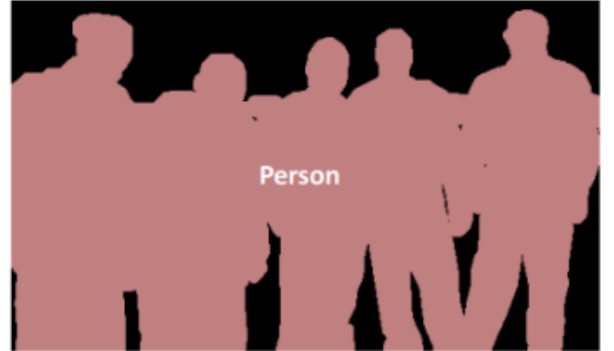
Şekil 6: Renklendirilmiş metin görüntüleri

## 5.2. Metin Tanıma Modülü

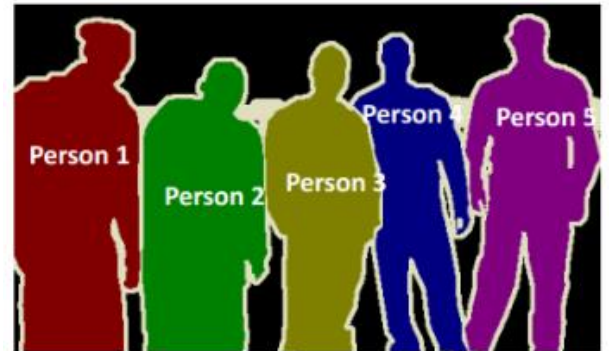
Bu modül, ön işleme modelinin çıktı görüntüsünde yer alan metni tanıma için kullanılır ve bilgisayardan anlaşılır biçimde çıktı verilerinin elde edilmesini sağlar. Dolayısıyla bu modülde aşağıdaki teknikler kullanılmaktadır.

### 5.2.1. Segmentasyon

Metin tanıma modülünde segmentasyon en önemli süreçtir. Bir görüntünün ayrı ayrı karakterleri arasında ayırım yapmak için bölümlene yapılmasını sağlamaktadır. Resim 5.2.1.1 'de görülen görüntüde insan sayısı bulunmak istenilmiştir buna bağlı olarak segmentasyon işlemi ile Resim 5.2.1.2 'de bulunan görüntü elde edilmiştir.



Resim 5.2.1.1



Resim 5.2.1.2

### 5.2.2. Özellik Çıkarma

Özellik çıkarma, ham verilerden istenilen verileri alma işlemidir. Bir karakterin farklı özelliklerini saklamak için farklı sınıflar yapılır. Prensip Bileşen Analizi (PCA), Doğrusal Ayrım Analizi (LDA), Bağımsız Bileşen Analizi (ICA), Zincir Kodu (CC), bölgeleme, Gradyan Tabanlı özellikler, Histogram vb. gibi özellik çıkarmayı için kullanılan birçok teknik vardır.

### 5.2.3. Sınıflandırma

Sınıflandırma, her bir karakterin tanımlanması ve ona doğru karakter sınıfının atanması sürecidir, böylece görüntülerdeki metinler bilgisayar tarafından anlaşılır hale getirilir. Sınıflandırıcılar, giriş özelliğini depolanan modelle karşılaştırır ve giriş için en uygun sınıfı bulur. Yapay Sinir Ağı, Şablon Eşleştirme, Destek Vektör Eşleştirme (SVM) gibi sınıflandırma için kullanılan birçok teknik vardır.

## 5.3. İşlem Sonrası Modülü

Metin tanıma modülünün çıktısı, bilgisayar tarafından anlaşılan metin verisi formundadır, bu nedenle, bu verileri düzenlemek veya aramak gibi daha uzak kullanım için bazı uygun formatlarda saklanmasını sağlamaktadır [6].

## 5.4. Metin Tanıma İşlemi Faydaları

### 5.4.1. Aranabilirlik

Taranması yapılan bir dosya veya görüntü üzerindeki metinler makine tarafından okunabilir metne dönüştürülerek .doc, .rtf, .txt, .pdf vb. formatlarda kaydedilme işlemi yapılabilmektedir. Bu dosya formatlarında aranılan bilgiye kolaylıkla erişim sağlanabilmektedir. Görüntü üzerinden elde edilmiş olan dosya formatlarını çeşitli dosya yükleme sitelerinin uygun veri tabanlarına yüklenerek, bu belgelere dünya çapında erişilebilir hale getirilmesinin sağlanması metin tanıma işleminin faydalarından biridir.

### 5.4.2. Düzenlenebilirlik

Tarayıcıdan geçirilen veya görüntü üzerinde bulunan bir metnin düzenlenme işlemini yapabilmek metin tanıma işleminin sağladığı faydalar arasındadır. Yükleme yapılan belgeler veya görüntüler metin tanıma işlemi ile dijitalleştirildikten sonra, istenilen düzenleme işleminin yapılması sağlanmaktadır.

### 5.4.3. Ulaşılabilirlik

Bir belge veya görüntü metin tanıma işlemi ile dijital ortama geçirildiğinde ve ortak bir veri tabanında kullanıma sunulduğunda, bu veri tabanına erişimi olan herkes tarafından kullanılabilirliği sağlanmaktadır. Bu özellik eski belgelerin dijital ortama geçirilmesi istenildiğinde ve devlet bünyesi altında kullanılması için büyük fayda sağlamaktadır.

### 5.4.4. Depolanabilirlik

Belgelerin dijitalleştirilmesi, aynı bilgiler için gereken depolama alanını düşürerek, daha verimli kullanım için potansiyel olarak yer açar. Ayrıca, kağıt israfı önlenerek geri dönüştürülebilir böylece metin tanıma işlemi çevre sağlığını iyileştirmede ve masrafi azaltmada fayda sağlamaktadır.

### 5.4.5. Yedeklemeler

Dijital yedeklemeler ucuza ve potansiyel olarak sınırsız sayıda yapılabilmektedir. Önemli belgeleri kağıt biçiminde evraklarla saklamak yerine metin tanıma işlemi ile dijitalleştirme, evrak işlerini çok daha sürdürülebilir hale getirmektedir.

### 5.4.6. Çevrilebilirlik

Metin tanıma işlemi için kullanılan algoritmalar, Arapçadan Hint alfabesine ve Japonca 'ya kadar çok çeşitli komut dosyalarını işleyebilmektedir. Unicode standardı ve makine çeviri yazılımı (Google çevirisi gibi) ile birleştğinde bu çeşitli dillerde yazılmış bir belgenin taranabileceği, dijitalleştirilebileceği ve başka bir dile çevrilebileceği anlamına gelmektedir ve çevirmenlerin basılı belgeleri titizlikle gözden geçirme ihtiyacını ortadan kaldırarak iş süresini ve uğraşını azaltmaktadır. [7].

## 6. Uygulama Alanları

Metin tanıma teknolojisinin tüm endüstrilerde uygulanabilirliği mevcuttur. Bu teknoloji, taranan belgelerin görüntü dosyalarının ötesine geçmesini ve bilgisayarlar tarafından tanınan metin içeriğiyle tamamen aranabilir belgeler haline gelmesini sağlamaktadır. Bu teknolojinin yardımıyla, insanlar artık elektronik veri tabanlarına girerken önemli belgeleri manuel olarak yeniden yazmak zorunda kalmayarak bunun yerine metin tanıma sistemi ilgili bilgileri elde ederek otomatik olarak işlem yapmayı sağlamaktadır. Böylece, daha kısa sürede doğru ve verimli bilgi işleme gerçekleştirilmektedir. Aşağıda, metin tanıma sisteminin bazı uygulamalarına yer verilmiştir.

### 6.1. Bankacılık

Görüntü üzerinden metni tanımanın kullanım alanları birçok alanda farklılık göstermektedir. Yaygın olarak bilinen bir metin okuma uygulaması da bankacılıktır, insan müdahalesi olmadan kontrolleri işlemek için kullanılır. Bir makineye çek iliştirilerek, üzerindeki yazı anında taranır, dijital ortama aktırılır ve doğru miktarda para aktarma işlemi kolay bir şekilde gerçekleştirilebilmektedir.

### 6.2. Yasal

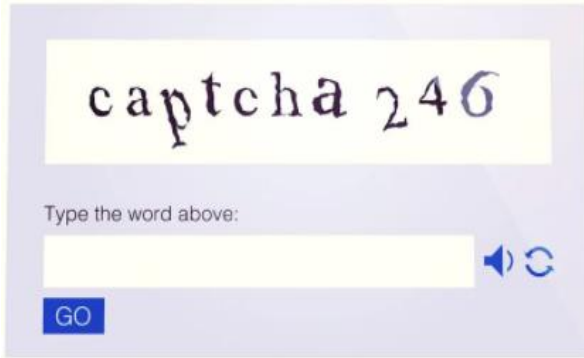
Hukuk endüstrisinde, eski belgelerin, tutanakların ve birçok kullanımda olan yazılı metinlerin dijital ortama aktarılması için de metin okuma uygulaması önemli bir fayda sağlamaktadır. Yerden tasarruf etmek, kağıt israfını azaltmak ve dosya kutularını eleme ihtiyacını ortadan kaldırmak için belgeler taranarak dijital ortama geçirilerek bilgisayar veri tabanlarına girilmektedir. Görüntü metin tanıma, belgeleri metin içinde aranabilir hale getirerek dosyalara erişim sürecinin basitleştirilmesini sağlamaktadır. Hukuk alanında çalışan insanlar artık, yalnızca birkaç anahtar sözcük yazarak bulabilecekleri, elektronik biçimdeki devasa bir belge kitaplığına hızlı ve kolay erişime metin okuma teknolojisi sayesinde ulaşabilmektedirler.

### 6.3. Sağlık

Sağlık hizmetleri alanında da evrak işlerini dijital ortama kolay ve güvenilir bir şekilde aktarabilmek için görüntüden metin tanıma teknolojisi kullanılmaktadır. Sağlık alanında, her hasta için sigorta formları ve genel sağlık formları gibi birçok evrak yükü oluşmaktadır. Bu bilgilere kolay ulaşım sağlayabilmek için, ilgili verileri gerektiğinde erişilebilen elektronik bir veri tabanına kaydetmek gerekmektedir. Görüntüden metni tanıma teknolojisini kullanarak formlardan istenilen bilgiler elde edilebilir ve her hastanın verileri veri tabanlarına kaydedilerek bu bilgilere kolay erişim sağlanabilmektedir. Sonuç olarak, sağlık hizmeti alanında hizmet veren çalışanların her hastaya mümkün olan en iyi hizmeti sunması ve iş yüklerinin azalması sağlanabilmektedir [8].

### 6.4. Captcha

CAPTCHA, insanların geçebileceği ancak mevcut bilgisayar programcılarının geçemediği testleri oluşturabilen ve derecelendirebilen bir programdır (Resim 6.4.1). Bilgisayar korsanlığı, internet kullanımı için ciddi bir tehdittir. Özellikle korunması gereken kişisel verilerin güvenliği için çözümler gerekmektedir. CAPTCHA sayesinde, görüntü bozulma teknikleri, boyut ve yazı tipi çeşitliliği, dikkat dağıtıcı arka planlar, rastgele bölümler, görüntüdeki vurgular ve gürültü ile karartılmış sayı harflerinden oluşan bir görüntü oluşturulur. Bu sistem, bu gürültüyü gidermek ve görüntüyü OCR (Optik Karakter Tanıma) sistemleri için izlenebilir hale getirmek için görüntüyü bölümlere ayırmak için kullanılabilir.



Resim 6.4.1: Captcha uygulamasına bir örnek

### 6.5. Kurumsal Depolar ve Dijital Kitaplıklar

Kurumsal depolar, bir üniversite veya araştırma kurumu bünyesinde oluşturulan çıktıların dijital koleksiyonlarıdır. Bir kurumun, özellikle de toplandığı, korunduğu ve yayınlandığı bir araştırma kurumunun entelektüel verilerinin çevrimiçi bir yerel araradır. Bir kurumun çıktılarını açmaya ve ona dünya çapında görünürlük ve daha fazla etki sağlamaya yardımcı olmayı hedeflemektedirler. Araştırma için disiplinler arası yaklaşımları etkinleştirerek dijital öğretim materyalleri ve yardımcıların geliştirilmesini sağlamaktadırlar. Temelde hakemli dergi makaleleri, konferans bildirileri, araştırma verileri, monografiler, kitaplar, tezler ve sunumlardan oluşan bir koleksiyondur. İlk görevleri, Açık Erişim literatürünü

sağlamaktır. Bunun pratik uygulaması, belgeleri tarayan tarayıcıdan oluşan bir sistemin kurulmasını içermektedir. Taranan bu belge daha sonra görüntüden metni okuma teknolojisi sayesinde dijital ortama aktarılarak ulaşımın kolay ve güvenilir bir hale gelmesini sağlamaktadır.

### 6.6. Otomatik Plaka Tanıma

Otomatik plaka tanıma, araç plakalarını tanımlamak için görüntüler üzerinde metin okuma teknolojisinden faydalanarak bir araç gözetim tekniği olarak kullanılmaktadır. Otomatik plaka tanıma ayrıca, lisans plakasından alınan numaralar dahil olmak üzere kameralar tarafından yakalanan görüntüleri sistemde kayıtlı tutarak saklama amacıyla yapılmıştır. Otomatik plaka tanıma teknolojisi, bölgeye özgü bir teknoloji olduğu için, ülkelere veya şehirlere göre farklılıklar göstermektedir. Çeşitli polis kuvvetleri tarafından ve kullanım başına ödemeli yollarda elektronik geçiş ücreti toplama yöntemi olarak ve trafik veya bireylerin hareketlerini kontrol ederek insan sağlığına faydalı olmak için kullanılmaktadır [9].



Resim 6.6.1: Otomatik plaka tanıma uygulamasına bir örnek

### 6.7. Diğer Sektörlerde Görüntüden Metin Tanıma

Görüntüden metin tanıma teknolojisi, eğitim, finans ve devlet kurumları dahil olmak üzere diğer birçok alanda yaygın olarak kullanılmaktadır. Bu teknoloji sayısız metni çevrimiçi olarak erişilebilir hale getirerek öğrenciler için para tasarrufu sağlayarak bilginin paylaşılmasına ve ulaşılmasının kolaylaştırılmasını sağlamaktadır. Fatura görüntüleme uygulamaları, birçok işletmede finansal kayıtları takip etmek ve birikmiş ödemelerin birikmesini önlemek için de kullanım alanı bulmaktadır. Devlet kurumlarında ve bağımsız kuruluşlarda, görüntüden metin tanıma teknolojisi, diğer süreçlerin yanı sıra veri toplama ve analizinin kolaylaştırılmasında kullanılmaktadır.

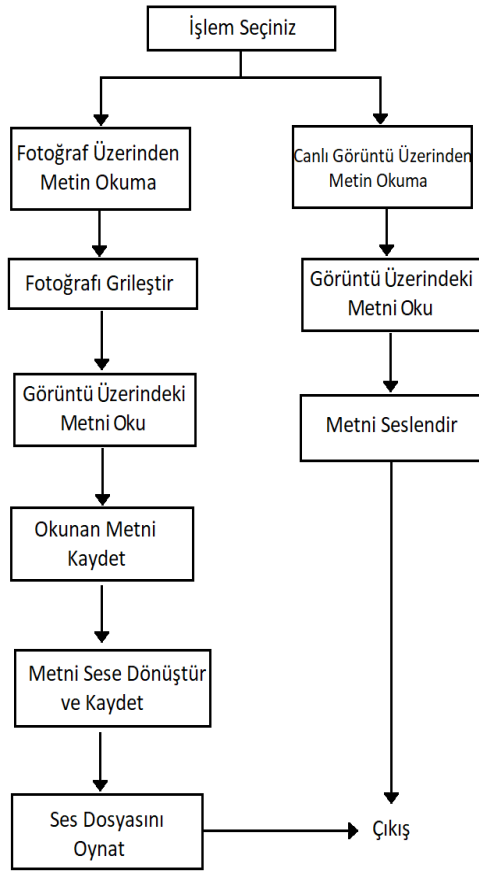
Günümüzde teknolojinin hızlı gelişmeye devam etmesi nedeniyle görüntüden metni tanıma teknolojisi gittikçe daha fazla uygulama alanı bularak hayatı kolaylaştırmaktadır [8].



## 7. Proje Akış Şeması

Yapılan proje 2 farklı işlem üzerinden gerçekleştirilmektedir. 1.işlem fotoğraf üzerinden metnin algılanmasını sağlamaktadır. Bu işlemde öncelikli olarak fotoğrafa grileştirme işlemi uygulanarak metin belirlenir ve dijital ortama aktarılması sağlanır. Aktarımı yapılan metin bir .txt dosyasına kaydedilerek seslendirme işlemine gönderilir. Sese dönüştürülen metin öncelikli olarak kaydedilir ve kayıt oynatılarak işlem tamamlanır.

2.işlemde ise IP Webcam uygulamasına bağlanarak ortaya çıkarılan canlı görüntüye ulaşılır. Görüntü üzerindeki metinler tek tek belirlenir ve seslendirme işlemi uygulanarak program sonlandırılır.



Resim 6.1: Proje akış şeması

## 8. Sonuçlar

Sonuçlar bölümünde projenin çalışma stili ve çalışması esnasında alınan sonuçlar yer verilmiştir.

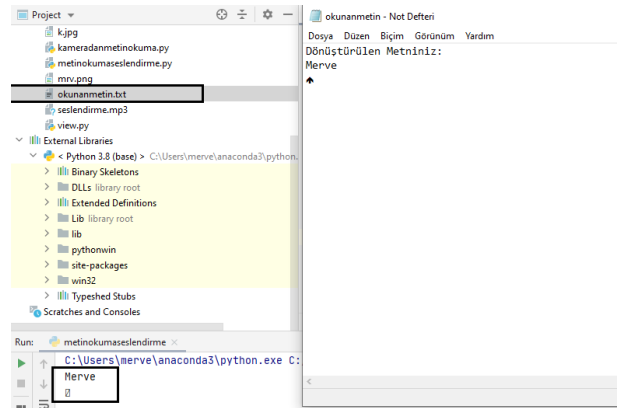
Görüntü işleme projesi olarak yola çıkılan görüntü üzerinden metin okuma konusunda istenilen işlemler gerçekleştirilmeye çalışılmıştır. Bu projenin görme engelli bireylerin kullanabileceği görüntü üzerinden metinleri algılayarak seslendirmeye yarayacak bir uygulama olarak yola çıkmıştır. Proje yapım esnasında geniş kitlelere hitap edebilmesi adına projeye çeşitli özellikler eklenmiştir.

Proje yapımına öncelikli olarak fotoğraf üzerinden metin okuma özelliği ile başlanmıştır. Fotoğraf üzerinden metin okuma işlemi için Tesseract ve OpenCv kütüphanelerinden yararlanılmıştır. Öncelikli olarak kodlar üzerinden girilen resmi kullanıcıya gösteren bir pencere tasarlanmıştır (Resim 8.1).



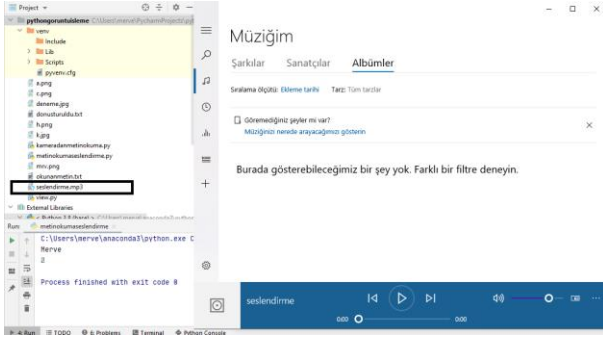
Resim 8.1: Okunacak resmin pencere içerisinde gösterilmesi

Okunması istenilen resmi kullanıcıya gösteren pencerenin manuel olarak kapanması sonrasında dönüştürülen metin Resim 8.2 'de görüldüğü üzere terminal üzerinden bir çıktı olarak gösterilir ve okunan metnin bir .txt dosyasına kaydedilir.



Resim 8.2: Okunan resimden bir çıktı üretilmesi ve dosyaya kaydedilmesi

Resmin üzerinden metni okuma ve dosyaya kaydetme işlemleri tamamlandıktan sonra sese dönüştürülen metin seslendirme.mp3 olarak kaydedilir ve seslendirme otomatik olarak açılır (Resim 8.3).

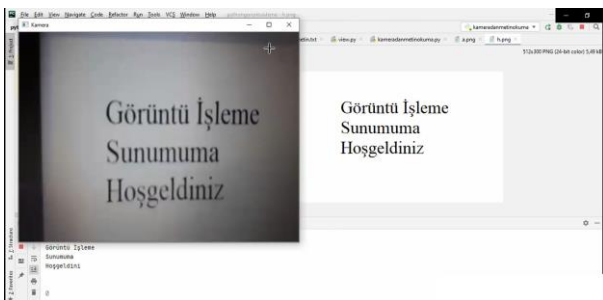


Resim 8.3: Seslendirilen dosyanın kaydedilmesi ve açılması

Canlı olarak görüntü alabilmek için telefona IP Webcam uygulaması yüklenerek IP adresi kopyalanmış ve projeye dahil edilmiştir. Böylece telefon kamerasıyla program arasındaki bağlantı sağlanmıştır (Resim 8.4). Canlı olarak metinleri okuyarak aynı anda seslendirme işlemi gerçekleştirilmiştir (Resim 8.5).

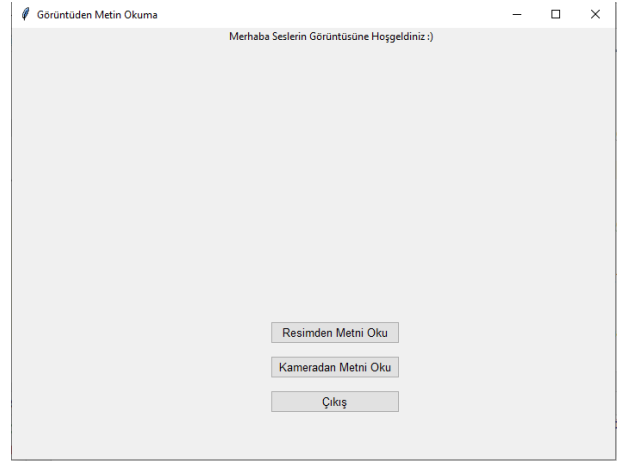


Resim 8.4: IP Webcam telefon kamerası görüntüsü



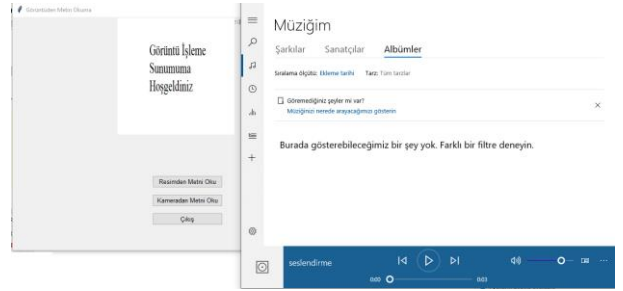
Resim 8.5: Canlı olarak görüntüyü metni çevirme

Proje içerisinde gerçekleştirilmek istenilen iki işlem yapıldıktan sonra bir ara yüz yapılmak istenerek Tkinter ara yüz aracı kullanılmıştır. Öncelikli olarak basit bir pencere oluşturularak başlığa ve istenilen 3 işlevi yapmaya yarayacak butonlara (resimden metni okuma, kameradan metni okuma ve programdan çıkış yapma) yer verilmiştir (Resim 8.6).



Resim 8.6: Tkinter ile oluşturulan ara yüz

Daha önce yapılan iki işlemle tkinter ile tasarlanan bu yapının tek farkı resmi bilgisayarda var olan dosyalardan seçmeye yaramasıdır. Seçilen resmi daha sonra Resim 8.7 'de olduğu gibi pencere içerisinde göstererek seslendirme işlemini gerçekleştirmeyi sağlamaktadır.



Resim 8.7: Dosyadan seçilen resmin görüntülenmesi ve seslendirilmesi

## 9. Kaynaklar

- [1] Yıldız, Sıtkı ve Gürler, Seda, “GÖRME ENGELLİ BİREYLERİN ENGELLİ HAKLARINA DAİR BİLGİ DÜZEYLERİNİN ÖLÇÜLMESİ”, *Kırıkkale Üniversitesi Sosyal Bilimler Dergisi (KÜSBD)*, Ocak 2018, Sayfa 241-268.
- [2] 3 05/04/2013 Tarih ve 6462 Sayılı “Kanun ve Kanun Hükmünde Kararnamelerde Yer Alan Engelli Bireylere Yönelik İbarelerin Değiştirilmesi Amacıyla Bazı Kanun ve Kanun Hükmünde Kararnamelerde Değişiklik Yapılmasına Dair Kanun”, *03/05/2013 Tarih ve 28636 Sayılı Resmi Gazete*.
- [3] Hürriyet Gazetesi yazarı, “Kamuda yaklaşık 11 bin görme engelli istihdam edildi”, *Hürriyet Gazetesi*, Ocak 2020
- [4] Karakoç, Mehmet, “GÖRÜNTÜ İŞLEME, TEKNOLOJİLER VE UYGULAMALARI”, *Akademik Bilişim*, Şubat 2012
- [5] Çayıroğlu, İbrahim, ”Görüntü İşleme”, *Karabük Üniversitesi Mühendislik Fakültesi Ders Notları*, Sayfa 7-9
- [6] Manwatkar, Pratik Madhukar ve Singh, Dr. Kavita R.,” A Technical Review on Text Recognition from Images”, *IEEE Sponsored 9th International Conference on Intelligent Systems and Control (ISCO) 2015*, 2015, Sayfa 721-723
- [7] Raghavan, Raamesh Gowri, Kakde, Prashant Manoharrao ve Khurana, Sukant, “Applications of OCR You Haven’t Thought Of”, <https://medium.com> internet sitesi, 2018
- [8] Application of OCR, <http://www.cvisiontech.com> internet sitesi
- [9] Asif, Ali Mir Arif Mir, Hannan, Shaikh Abdul, Perwej, Yusuf ve Vithalrao, Mane Arjun, “AN OVERVIEW AND APPLICATIONS OF OPTICAL CHARACTER RECOGNITION”, *International Journal of Advance Research In Science And Engineering IJARSE*, Vol. No.3, Issue No.7, Temmuz 2014, Sayfa 265-268