



Improvement algorithm application with image processing interface software for the improvement of mini gel electrophoresis images

Ahmet Reşit Kavsaoglu^{1*}, Kerim Özkara²

¹Department of Biomedical Engineering, Faculty of Engineering, Karabuk University, 78050, Karabuk, Turkey

²Department of Electronics and Automation, Sinop University, 57400, Sinop, Turkey

Highlights:

- Applying gel electrophoresis image processing function
- Creating the best BP value calculation algorithm function based on RF value
- User interface design showing UV gel image and BP counts on the image

Keywords:

- Gel electrophoresis
- Image processing
- Embedded system
- Base pair (BP) counts
- User interface

Article Info:

Research Article
Received: 15.05.2021
Accepted: 29.08.2021

DOI:

10.17341/gazimmfd.937650

Correspondence:

Author: Ahmet Reşit Kavsaoglu
e-mail: kavsaoglu@karabuk.edu.tr
phone: +90 370 418 7050

Graphical/Tabular Abstract

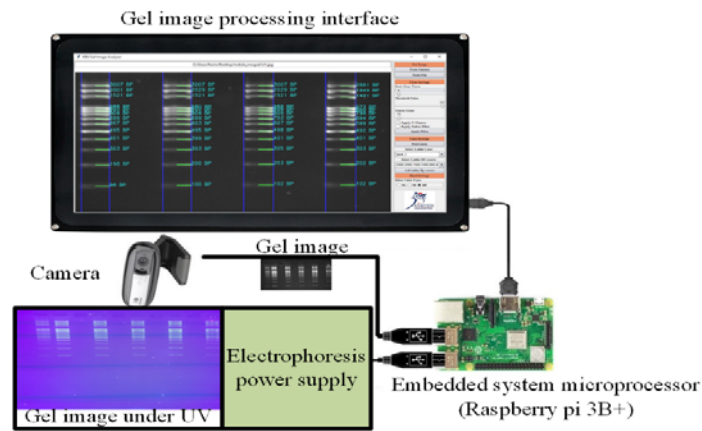


Figure A. Gel image acquisition and image processing system

Purpose: By applying image processing techniques on the image of the electrophoresis applied gel, the base pair (BP) counts can be found very close to the reference BP numbers. In this study, image processing techniques that enable the designed system to find BP numbers with the lowest error compared to previously made software are explained.

Theory and Methods:

To find the BP counts with the lowest error, filtering into the gel image, scanning the image horizontally and vertically, converting the image into a signal, correcting the baseline and finding the peak points by derivation, finding the lanes and bands and calculating the BP values according to the RF values were performed. For the best-filtered image, different combinations of Background Clear (BG Clear), Adaptive Histogram Equalization (Clahe), Gabor filter and K-mean methods were applied. As a result of these analyzes, it was seen that the most successful method was BG Clear + Gabor + K-mean combination. To find the lanes in the filtered image, the vertical and horizontal sums of the image pixels were taken, respectively, and the image was converted into a signal, and then the peaks were obtained by derivation process and the lanes and bands were found. Regression analysis was performed by creating a 2-term power function to calculate BP numbers according to the RF values of the bands found.

Results:

As a result of the analysis, $R_2 = 0.9999533$ and error rates for the lanes were found in the range of 0.55-0.86% and the closest values to the actual reference BP values were calculated. The calculated values are presented to the user on the image with the user interface software.

Conclusion:

This study presents a gel image processing application controlled by the embedded system to estimate the BP counts of the bands in the images. With the application, it is revealed that the estimated average error rates of BP count measurements decreased to 0.55 - 0.86% by using the software algorithm function.



Mini jel elektroforez görüntülerinin iyileştirilmesine yönelik görüntü işleme arayüz yazılımı ile iyileştirme algoritması uygulaması

Ahmet Reşit Kavsaoglu^{1*}, Kerim Özkara²

¹Karabük Üniversitesi, Mühendislik Fakültesi, Biyomedikal Mühendisliği Bölümü, 78050, Karabük, Türkiye

²Sinop Üniversitesi, Ayancık Meslek Yüksekokulu, Elektronik ve Otomasyon Bölümü, 57400, Sinop, Türkiye

Ö N E Ç I K A N L A R

- Jel elektroforez görüntü işleme fonksiyonunun uygulanması
- RF değerine dayalı en iyi BP değeri hesaplama algoritması fonksiyonunu oluşturma
- UV jel görüntüsünü ve görüntü üzerinde BP sayılarını gösteren kullanıcı arayüzü tasarımı

Makale Bilgileri

Araştırma Makalesi

Geliş: 15.05.2021

Kabul: 29.08.2021

DOI:

10.17341/gazimmfd.937650

Anahtar Kelimeler:

Jel elektroforez,
görüntü işleme,
gömülü sistem,
python kullanıcı arayüzü

ÖZ

Python programlama dili tabanlı, mini jel elektroforez sistemi görüntü işleme arayüzü yardımıyla jel elektroforez bant görüntülerinin iyileştirilmesini sağlayan algoritma fonksiyonu oluşturularak deneysel bir çalışmanın yapılması amaçlanmıştır. Bu sistem ile jel görüntülerinin UV (Ultraviyole) ışık altındaki analizi, tasarlanan arayüz yazılımı ile yapılmıştır. Görüntüler kamera aracılığıyla arayüz yazılımına aktarılıp şerit ve bantları belirginleştiren filtreler uygulanmıştır. BP (Baz Çifti) sayıların en doğru sonuçla tahmin eden iki terimli kuvvet fonksiyonu kullanılmıştır. Yazılım, Raspberry Pi 3B+’da OpenCV kütüphanesi Python programlama dili ile mini jel elektroforez sistemine entegre edilerek kullanılabilir. Bu çalışmada, gömülü sistem tarafından kontrol edilen mini jel elektroforez sisteminden elde edilen jel görüntüleri, görüntülerdeki bantların BP sayılarını tahmin etmek için kullanılmıştır. Arayüz yazılım algoritması, jel görüntülerini sistemdeki kamera üzerinden veya dosyadan içe aktararak el ile yapılan ölçümlerdeki BP sayılarının tahminini ve ortalama hata oranlarını % 0,55 -% 0,86 aralığına düşürülmesini sağlamıştır. Jel görüntülerinde şeritlerin ve BP sayılarının en düşük hata oranıyla hesaplanabilmesini sağlayan arayüz yazılımında, üstel fonksiyonların kuvvet fonksiyonu olarak da tanımlanabilmesi ilkesinden yola çıkarak iki terimli üstel fonksiyon yerine iki terimli kuvvet fonksiyonu uygulanmıştır. Bu iki terimli kuvvet fonksiyonu ile hesaplanan BP değerlerin, gerçek BP değerlerine yakınlığını gösteren $R^2=0,9999533$ değeri ile amacına uygun olarak çalıştığı ortaya konulmuştur.

Improvement algorithm application with image processing interface software for the improvement of mini gel electrophoresis images

H I G H L I G H T S

- Applying gel electrophoresis image processing function
- Creating the best BP value calculation algorithm function based on RF value
- User interface design showing UV gel image and BP counts on the image

Article Info

Research Article

Received: 15.05.2021

Accepted: 29.08.2021

DOI:

10.17341/gazimmfd.937650

Keywords:

Gel Electrophoresis,
image processing,
embedded systems,
python user interface

ABSTRACT

It is aimed to carry out an experimental study by creating an algorithm function that enables the improvement of gel electrophoresis band images with the help of the Python programming language-based, mini gel electrophoresis system image processing interface. With this system, the analysis of gel images under UV (Ultraviolet) light was made with the designed interface software. The images were transferred to the interface software via the camera and filters were applied to highlight the lanes and bands. The two-term power function, which estimates BP (Base Pair) numbers with the most accurate result, was used. The software can be used on Raspberry Pi 3B+ by being integrated into the mini gel electrophoresis system via Python programming language with OpenCV library. In this study, gel images obtained from the mini gel electrophoresis system controlled by the embedded system were used to estimate the BP numbers of the bands in the images. By importing the gel images from the camera or from the file in the system, the interface software algorithm enabled the estimation of BP numbers in the manual measurements and reducing the average error rates to the range of 0.55% - 0.86%. In the interface software, which enables the calculation of lanes and BP numbers in gel images with the lowest error rate, two-term power function has been applied to the image processing algorithm instead of the two-term exponential function, based on the principle that exponential functions can also be defined as power functions. It has been revealed that the BP values calculated with this two-term power function work in accordance with its purpose with the value of $R^2=0.9999533$, which shows the closeness of the actual BP values.

*Sorumlu Yazar/Yazarlar / Corresponding Author/Authors : *kavsaoglu@karabuk.edu.tr, kerim.ozkara@sinop.edu.tr /

Tel: +90 370 418 7050

1004

1. GİRİŞ (INTRODUCTION)

Makromoleküllerin elektrik alanda yürütülmesi ile oluşan BP (Base Pair-Baz Çifti) sayılarının gerçeğe en yakın değerde bulunabilmesi; mikrobiyoloji, kimya ve tıp alanları için oldukça önemlidir. Makromoleküllerin yürütülmesi bir yerden başka bir yere gitmek anlamına gelen "phoresis" kelimesinden türetilen elektroforez işlemi ile yapılmaktadır [1]. Elektroforez, bir elektrik alanının etkisi altında yüklü parçacıkların göçüne dayanır. Bu makromoleküller bakteri, virüs, protein molekülleri, nükleik asitler veya sentetik partiküller olabilir. Elektroforez, etkin ayırma gücü, uygulama kolaylığı ve ucuzluğu nedeniyle sürekli geliştirilen ve tercih edilen bir elektrokinetik yöntemdir [1, 2].

DNA jel elektroforezi, farklı boyutlardaki DNA parçalarını ayıran bir moleküler biyoloji tekniğidir. DNA jel elektroforezinin uygulamaları arasında DNA parmak izi (genetik tanı), DNA boyutu tahmini için DNA ayrımı yer alır. Elektroforetik görüntülerden DNA bandlama modellerinin doğru yorumlanması zahmetli ve elle belirlenen birçok band olduğundan hataya açık olabilmektedir [3].

Green ve Sambrook'un [4] araştırmasına göre, DNA'nın çoğu rutin elektroforetik analizi, agaroz mini jeller kullanılarak çok hızlı bir şekilde gerçekleştirilebilir. Mini jeller, klonlama protokolünde bir sonraki adım atılmadan önce hızlı bir yanıt gerektiğinde özellikle yararlıdır. Kuyucuklar daha küçük ve jeller daha ince olduğundan, görselleştirme için normalden daha az DNA gereklidir. Ayrıca, jeller hızlı bir şekilde hazırlanıp çalıştırılabildiğinden ve daha az reaktif gerektirdiğinden, zaman ve maliyet açısından önemli tasarruf sağlar. Mini jeller, 3 kilo bazdan (3000 baz) küçük DNA fragmanlarının analizi için en uygun boyutlardır. Daha büyük parçacıklar, daha büyük voltajlar kullanılması ve jelin nispeten kısa uzunluğu nedeniyle zayıf bir şekilde ayrılır. Her jel kuyusu, jelin kalınlığına ve tarak dişlerinin genişliğine bağlı olarak 3-12 µl numune alır. Genel olarak, tercih edilen jel doldurma tamponundaki 10-100 ng DNA bir kuyuya uygulanır. Kevin M. Bourzac vd. [5] etidyum bromürü kanserojen bir boya olarak belirterek, jelin boyanmasında kullandıkları daha güvenilir ve geliştirilmiş boyama özelliklerine sahip olan SyBr Safe yeşil boya ile boyanan jel uygun mesafeye getirilene kadar jelin boyutuna bağlı olarak cm başına 5 ila 20 Volt, 30 ila 60 dakika süreyle uygulanır. Buna göre, 5 cm'lik bir agaroz jelle 25-100V'luk bir voltaj uygulanmalıdır. Uygulama tamamlandığında, agaroz jelin görüntüsü UV ışığı altında bir CCD kamera ile alınabilir [6]. Pek çok farklı alanlarda kullanılan görüntü işleme yöntem ve teknikleri (örneğin: elektrotermal mikro-aktüatörün karakterizasyonu [7], ekmek doku analizi [8], hızlı üst eksremite değerlendirme [9] gibi) kamera ile alınan bu görüntüler üzerinde de uygulanabilir.

Kamera ile jel görüntüsünün istenilen doğrulukta alınabilmesi ve bu görüntünün güncel teknolojiler ile

işlenebilmesini sağlayan bir çalışmada; Gömülü sistem üzerinde oluşturulan bir kontrol yazılımının, bilinmeyen DNA'ların boyutlarını belirlemek için jel görüntülerinin tespit süresini iyileştirebileceği varsayılarak oluşturulan Mini Jel Elektroforez Sisteminde; SyBr yeşil boya ile boyanmış DNA görüntüsünün gömülü sistem üzerindeki analizinde, en az hata ile en hızlı ve en hassas şekilde kullanılan yöntemler geliştirilmiştir [10].

Kavsaoğlu ve Mersinkaya'nın [11] çalışması Python programlama dili ile oluşturulan kontrol yazılımı ve bu yazılım ile kontrol edilen güç kaynağı devresi, elektroforez süresince bant görüntülerinde etkilerin oluşmasına neden olan gerilim ve zamana bağlı değerlerin işlem süresince sabit tutulmasını sağlamıştır ve bant görüntülerindeki şeritlerin, görüntü işleme yazılımı ile bütünleştirildiğinde düzgün ve ölçülebilir bir şekilde ayrılabilceği belirtilmiştir.

Jel elektroforez işlemi sonrası jel görüntüsünde oluşan bantların belirlenmesinde görüntü işleme için çeşitli programlar kullanılmaktadır. S. Khakabimamaghani [12] ve arkadaşlarının çalışmasında jel elektroforez görüntü analizi için birkaç araç ve yazılımlar karşılaştırılmıştır, ancak çoğu ticari yazılımlardır. Ayrıca bu yazılımlardan bazıları kullanıcıların tüm gereksinimlerini karşılamamaktadır. Örneğin, QIAxcel System [12, 13], görüntüleri nicelleştirebilir ancak örnekleri kümeleyemez. BioDocAnalyze [14] ve Gel Doc EZ [15] düşük kaliteli görüntüleri işlerken bu yazılımlarla çalışmak kolay değildir. Bahsedilen yazılımlara ek olarak, kümeleme yeteneğine sahip başka araçlar da vardır. Örneğin, ClusterVis [16], GelQuest tarafından sağlanan bilgilere dayanarak örneklerin filogenetik ağacını çizebilir. GelQuest ile çalışmak kolay olsa da, görevlerin çoğu manuel olarak yapılmaktadır. Bu da görüntülerin analizi için geçen süreyi yavaşlatır. Ayrıca, GelScan [17], Phoretix 1D [18] ve Quantity One [19], bahsedilen araçların tüm özelliklerini paylaşan üç ticari yazılımdır. Ayrıca örnekleri fizyolojik olarak kümeleyebilirler. Bu araçlar, kullanıcıların gereksinimlerinin çoğunu karşılarken, kullanıcıların öğrenebilmek için çok uygulama yapmasını gerektiren karmaşık kullanımlarının olması ve yüksek fiyatlarla kullanıcılara sunulması dezavantajlarıdır [20]. Ayrıca, Rementeria ve ark. [21] ve Cardinali ve ark. [22], jel görüntülerine dayalı olarak bantların saptanmasında ve genotiplerin sayısının belirlenmesinde bazı ticari paketlerin dezavantajlarını ve aralarındaki farklılıkları tartışmıştır. Bazı araştırmacılar bu amaç için kendi ticari olmayan ve basit araçlarını geliştirmeye çalışmışlardır. Örnek olarak PyElph [23], ticari yazılımların tüm yeteneklerini sağlayan, açık kaynaklı, kullanıcı dostu, eğitim amaçlı uygulamalar için tasarlanmış, ancak sütunları ve bantları tespit etmede çok başarılı olmayan bir yazılımdır. Diğer yazılım, java tabanlı ancak açık kaynaklı olmayan ve şeritleri manuel tespit eden bir ürün olan GelAnalyzer [24]'dir. Sistemimizde bulunan özellikler ile diğer yazılımların bazı özellikleri Tablo 1'de gösterilmiştir. Tabloda verilen programlar gömülü sistemlere uygun olarak tasarlanmamıştır. Sadece Python

Tablo 1. Bazı band görüntüsü işleme programlarının özellikleri (feature of some of band image processing programs)

| Yazılım | Ticari | Otomatik şerit ve bant bulma | BP tahmin R ² değeri | UV görüntüleme sistemi | Gömülü sistem uyumluluğu |
|-----------------------|--------|------------------------------|---------------------------------|------------------------|--------------------------|
| QIAxcel Sistem | Evet | Evet | Bilinmiyor | Hayır | Hayır |
| BioDocAnalyze | Evet | Evet | Bilinmiyor | Hayır | Hayır |
| Gel Doc EZ System | Evet | Evet | Bilinmiyor | Hayır | Hayır |
| Gelquest & ClusterVis | Evet | Manuel | Bilinmiyor | Hayır | Hayır |
| GelScan | Evet | Evet | Bilinmiyor | Hayır | Hayır |
| Phoretix 1D | Evet | Evet | Bilinmiyor | Hayır | Hayır |
| Quantity One | Evet | Evet | Bilinmiyor | Hayır | Hayır |
| GelQuest | Hayır | Manuel | Bilinmiyor | Hayır | Hayır |
| PyElph | Hayır | Manuel | Bilinmiyor | Hayır | Hayır |
| GelAnalyzer | Hayır | Manuel | 0,9850 | Hayır | Hayır |
| GelJ | Hayır | Manuel | 0,9952 | Hayır | Hayır |

tabanlı, pyElph gömülü sistemde çalışmaya en uygun yazılımdır. Bu yazılım Python 2.7 sürümü ile tasarlanmıştır ancak günümüzde Python 2.7 sürümü artık desteklenmemektedir. Bu çalışmada tasarlanan yazılımın kullanımı diğer arayüzlere göre daha basit olup, BP tahmin değerlerini yakınlığını gösteren R² değerinde 1'e en yakın (0,9999) değere ulaşmıştır. Ayrıca ortalama hata oranlarını, geliştirilen algoritma ile % 0,55 -% 0,86 aralığına düşürülmesini sağlayan yenilikçi bir sistem oluşturulmuştur. Diğer yazılımlardan farklı olarak kendisine ait bir UV görüntüleme sistemiyle tümleşik olarak tasarlanmıştır. Gömülü sistem ile tasarlandığı için bir bilgisayardan ve ayrı olarak oluşturulan görüntüleme sisteminden daha az yer kaplayan, daha az enerji tüketen, daha düşük maliyetle işlemleri yapabilen bir sistem geliştirilmiştir.

Tasarlanan sistem, UV ışık kaynağını arayüz yazılımı üzerinden kontrol ederek UV görüntüleme sistemi ile bütünsel olarak çalışmaktadır. Bu özelliği kamera üzerinden anlık jel görüntülerinin arayüz üzerine alınabilmesi sağlanmaktadır. Jel elektroforez görüntülerindeki BP bandlarını belirginleştiren filtre kombinasyonu önerilmektedir. Önerilen bu kombinasyonun bandların konumlarının tespitinin kolaylaştırdığı bir arayüz tasarımı sunulmaktadır. Ayrıca makromoleküllerin en az hata oranı ile hesaplanabilmesi için arayüz yazılımı ile bütünsel iyileştirme algoritması geliştirilmektedir. Arayüz yazılımı kullanıcının jel görüntülerini yüklemesi sonrasında görüntü iyileştirme algoritma özelliklerini ayarlayabilmesini de içermektedir.

2. DENEYSEL METOT (EXPERIMENTAL METHOD)

Jel elektroforez deneyleri dijital görüntülerde gösterilebilir olduklarından önemli bilgileri çıkarmak için moleküler biyoloji uygulamalarında yaygın olarak kullanılmaktadır [25]. Mini jel elektroforez işleminde görüntü elde etme yöntemi, yaygın olarak bir dijital kamera ile UV ışığı altında SyBr yeşil boya ile boyanmış DNA görüntüsünün elde edilmesini içerir. Daha sonra bu görüntü bilgisayar ortamına aktarılıp açılabilir bozukluğu varsa düzeltilip, görüntü analizi

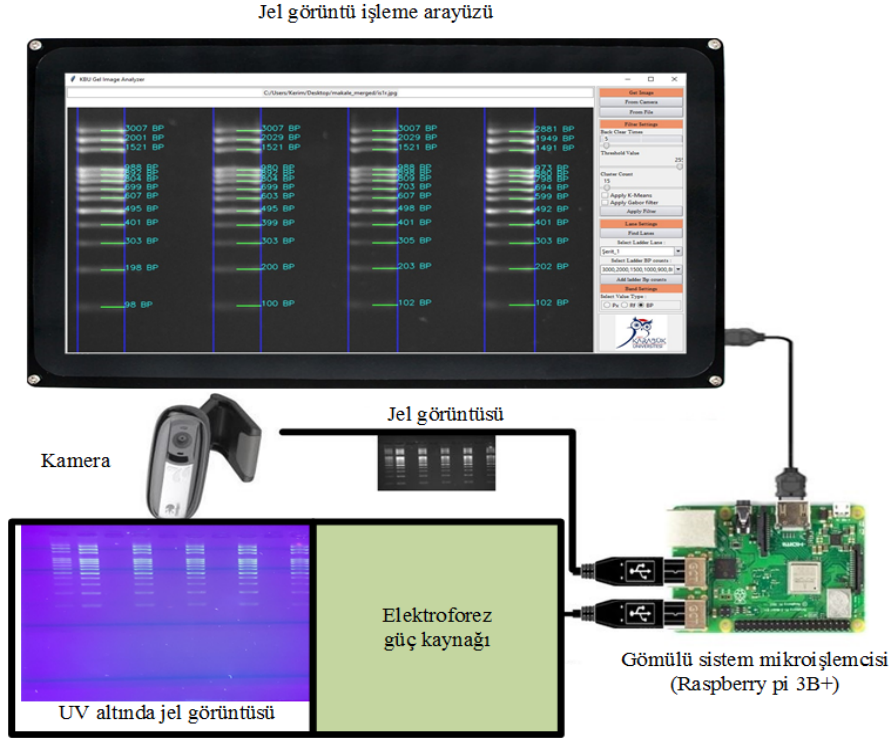
yaptırılmak üzere bir ölçüm programına aktarılmaktadır. Bu programda bantların yürüdükleri mesafeleri, belli bir referansı baz alan sanal bir cetvel ile ölçülebilmektedir. Bu yöntemde kamera açısını sabitleyerek bozulma önlenmesi de ışık yansımaları ya da gönderilen ışığın yüzeyde ya da ortamda patlamasıyla oluşan kontrast bozunmalarını önlemek pek mümkün değildir.

Bu çalışmada oluşturulan sistem, Şekil 1'de verilen bir UV görüntüleme elemanı, tümleşik bir kamera ve gömülü sisteme sahip bir görüntü işleme arayüz sistemidir. Gömülü sistem, UV ışığını kontrol ederek yazılıma bağlı kamera ile UV altındaki agaroz jel görüntüsünün ekrana aktararak işlendiği ve görüntü ayarlarının kullanıcı tarafından yapılabildiği uygulama yazılımı ile görüntünün analizini sağlayan bir sistemin oluşturulması sağlanmıştır.

Makale çalışmasında jel görüntülerinin dijital ortalama alınmasından sonra görüntü üzerinde sırasıyla filtreleme, görüntünün yatay ve dikey olarak taranıp sinyale çevrilmesi, taban çizgisi düzeltme ve türev olarak tepe noktaları bulma işlemleri ile şerit ve bantların bulunması sonrasında RF değerlerine göre BP değerlerini hesaplayan fonksiyonlar oluşturulmuştur. Görüntü işleme teknikleri, OpenCV kütüphanesi ile Python programlama üzerinde ve analiz yöntemleri kullanılarak, görüntünün kullanıcı arayüzünde otomatik olarak işlenmesi üzerinde durulmaktadır.

2.1. Arayüz (Interface)

Çalışmada, program arayüzü, programın kontrolünü sadeleştiren bir yapı olarak tasarlanmıştır. Şekil 2 programın ana ekranını göstermektedir. Kullanıcılara dosyadan veya kameradan görüntü alma özelliği sunmaktadır. Arayüz yazılımı ile jel görüntülerine uygulanan görüntü işleme teknikleri görüntü filtreleme, şeritleri ve bantları bulma, bant konumlarını belirleme ve belirlenen bant konumlarını bulurken RF (Relative Front-Görelî ön taraf) ile BP'yi hesaplamak için üstel fonksiyondan uyarlanarak oluşturulan kuvvet fonksiyonunun çalışması ve uygulama üzerinde yapılan çalışmalar açıklanmaktadır.



Şekil 1. Tasarlanan görüntü elde etme ve görüntü işleme sistemi arayüzü blok şeması (Designed image acquisition and image processing system interface block diagram)

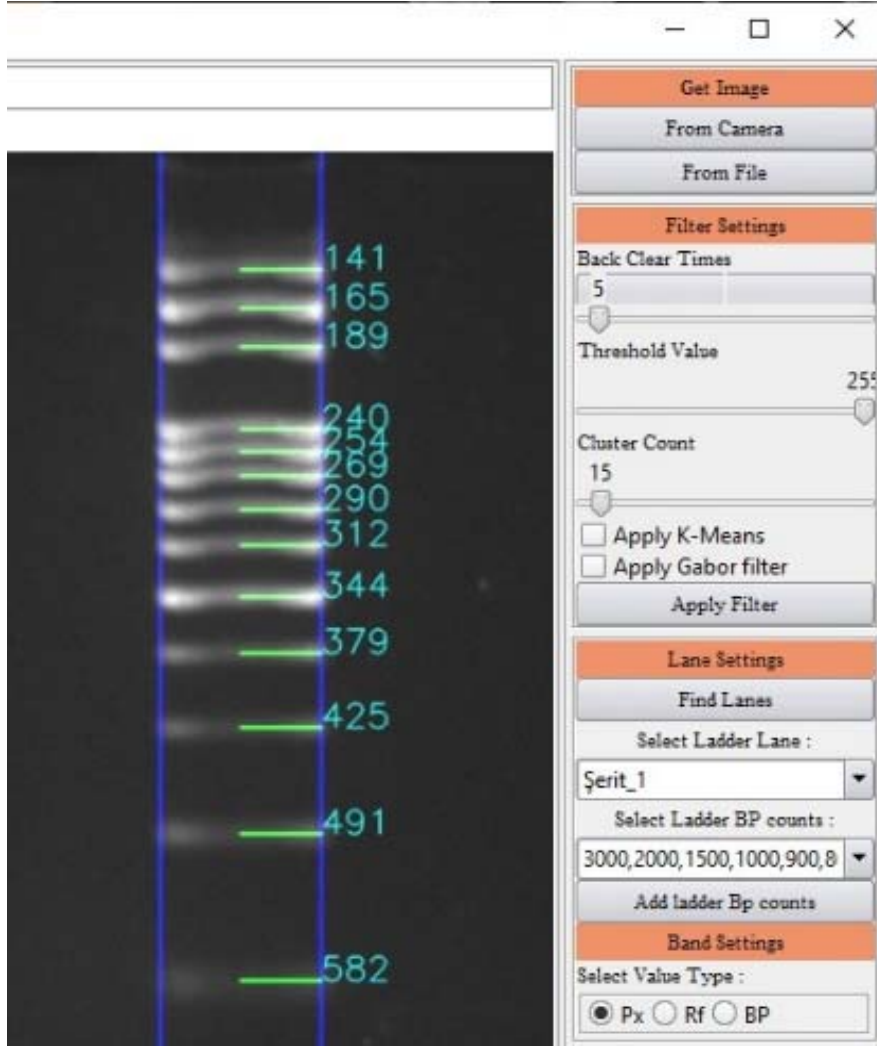
2.2. Jel Görüntüsü Filtreleme (Gel Image Filtering)

Agaroz jel görüntüleri üzerinde elektroforetik hareketi sağlanmış moleküllerin şerit ve bantlarını bulmak için kullanılan birçok yöntem vardır. Hemen hemen her yöntem, görüntü iyileştirme ile elde edilen sinyalden bilgi çıkarmaya dayanır. Ayrıca band görüntülerini iyileştirmede kullanılan birçok yöntem ve teknik mevcuttur. Uygulanan en basit yöntemlerden biri, jel görüntülerinde parlaklık değerlerini bir eşik değerine göre ayırarak ve arka plan ile bantlar arasındaki ton farkını açarak bantları belirlemektir [27]. Diğer yöntemlerden bazıları bulanıklaştırma [28], kenar algılama, uyarlamalı eşikleme, histogram eşitleme [29, 30], Morfolojik Geri Çatma (Morphological Reconstruction - MR) [31-33], Gabor filtre (GF) uygulaması [34, 35], K-means kümeleme [36, 37] gibi bir veya daha fazla görüntü filtreleme ve görüntü iyileştirme işlemi uygulandıktan sonra bantların konumunu netleştirerek bantların tespit edilmesidir.

Belirtilen yöntemlerden Morfolojik Geri Çatma (Morphological Reconstruction - MR) yöntemi, matematiksel morfoloji bağlamında, görüntüdeki sınırlar, iskeletler ve dışbükey zarf gibi bölgesel şekillerin tanım ve temsilinde görüntü bileşenlerini elde edebilmek için kullanılacak teknik ya da teknikleri niteler. Söz konusu morfolojik teknikler, budama (ayıklama), inceltme ve morfolojik filtreleme gibi ön veya son işlemlerde kullanılabilir [17]. “Opening” yöntemi aşındırma (Erozyon - erode) ve genişletme (Dilation - dilatasyon) yöntemlerinin artarda tekrarlarıyla elde edilmektedir. Bir görüntüye önce

aşındırma işlemi uygulayıp ardından genişletme yöntemi uygulanırsa beyaz alanlar önce küçülür, küçük beyaz alanlar yok olur, sonra ise küçülen beyaz alanlar tekrar genişletilerek eski haline getirilir. Böylelikle gürültüler yok edilmektedir [20]. MR işlemi, görüntünün farklı bölümlerindeki arka plan seviyelerini tahmin eder ve bunları orijinal görüntüden çıkarır. Uzamsal bir filtreleme tekniği olan Genişletme/aşındırma (Opening) işlemi görüntüdeki her pikseli kendisinin ve en yakın sekiz komşusunun değerini maksimum/minimum gri değeri ile değiştirmektedir. Genel olarak, bir resmi n kez genişletmek ve ardından genişletilmiş resmi n kez küçültmek, alanları $(2n+1)$ pikselden daha az olan nesnelere, yani, açılış (Opening) hareketleri gibi etkili bir şekilde alçak geçiren bir filtre gibi filtrelemektedir. Görüntüden kaldırılacak nesnelere boyutu, görüntüye uygulanan genişletme ve aşındırma işlemlerinin sayısı ile belirlenmektedir [19]. Bir diğer yöntem Gabor Filtresi (GF), bir görüntüyü farklı bölgelere ayırma işlemi olan görüntü bölütlemeye yaygın olarak uygulanan bir yaklaşımdır. Bu yaklaşım, gerçek dünya koşullarında görsel yorumlama için kullanılan insan görsel sisteminin çok kanallı çalışmasından esinlenmiştir. GF, Gauss zarfları tarafından modüle edilen karmaşık sinüzoidal dalgalar olarak tanımlanabilen iki boyutlu Gabor filtrelerinden oluşur. Bu filtre, iki boyutlu bir sinyalin yerleştirilmiş ve yönlendirilmiş bir frekans analizini gerçekleştirir. GF, Eş. 1’de zaman düzlemi, Eş. 2’de frekans düzlemi olarak verilmektedir [38].

$$h(x, y) = e^{\left\{ -\frac{1}{2} \left[\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} \right] \right\}} \cos(2\pi u_0 x) \quad (1)$$



Şekil 2. Arayüz yazılımı ana ekranı (Home screen of interface software) [26]

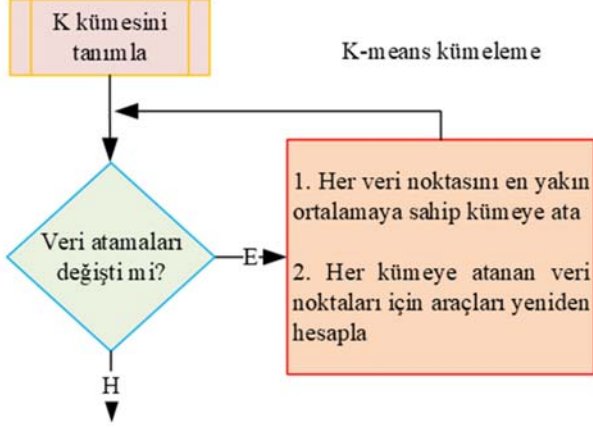
$$H(u, v) = A \left(e^{-\frac{1}{2} \left[\frac{(u-u_0)^2}{\sigma_u^2} \right] + \frac{v^2}{\sigma_v^2}} + e^{-\frac{1}{2} \left[\frac{(u-u_0)^2}{\sigma_u^2} \right] + \frac{v^2}{\sigma_v^2}} \right) \quad (2)$$

Burada; $\sigma_u = 1/2\pi\sigma_x$, $\sigma_v = 1/2\pi\sigma_y$, $A = 2\pi\sigma_x\sigma_y$ 'dir.

Gabor filtresinin çekirdeği “cv2.getGaborKernel(ksize, sigma, theta, lambda, gamma, psi, ktype)” kodundaki parametreler incelendiğinde; ksize, Gabor çekirdeğinin büyüklüğüdür. Eğer ksize = (a,b) ise, o zaman Gabor çekirdeği boyutunda axb piksellere sahip olunmaktadır. Sigma, Gabor filtresinde kullanılan Gauss fonksiyonunun standart sapmasıdır. Teta, Gabor fonksiyonunun normalin paralel çizgisine açısal olarak yönlendirmesidir. Teta haricindeki parametrelerin sabit olması koşuluyla açı değeri arttıkça Gabor filtresi saat yönünde dönmektedir. Lambda, Gabor filtresindeki sinüzoidal faktörün dalga boyudur. Lambda değeri arttıkça Gabor filtresi kalınlaşmaktadır. Gama, uzaysal en boy oranıdır. Diğer parametreler sabit tutulması kaydıyla gamma parametresinin değeri azaldıkça gabor filtresinin boyu uzamaktadır. Psi, faz kaymasıdır.

Ktype, Gabor çekirdeğindeki her pikselin alabileceği değerlerin türünü ve aralığını gösterir [21]. K-means kümeleme (K-means Clustering) yöntemi ise, denetimsiz makine öğrenimi yöntemi, otomatik veri gruplaması için basit ama güçlü bir algoritmadır. Bu yöntemeye dayalı olarak, benzer özelliklere sahip görüntü pikselleri tek bir kümede gruplandırılabilir. Buna göre, bir kümeye ait veri örnekleri en küçük varyasyon derecesine sahiptir [38]. Şekil 3'te K-means kümeleme algoritması gösterilmektedir. Algoritmanın isminde yer alan “k” harfi, aslında küme sayısını belirtir: Algoritma, hata hesaplamada yaygın olarak kullanılan karesel hata fonksiyonunu en aza indirgeyecek “k” küme sayısını da aramaktadır. Verilen “n” sayıdaki veri seti “k” tane kümeye bu hata fonksiyonunu en aza indirgeyecek şekilde yerleştirilmektedir. Bu nedenle küme benzerliği kümedeki değerlerin ortalamaya yakınlıkları ile ölçülmektedir. Bu da kümenin ağırlık merkezi olmaktadır. Kümenin merkezinde yer alan değer kümenin temsilci değeridir ve medoid olarak adlandırılmaktadır [24]. Tasarlanan arayüz yazılım algoritmasında kullanılan K-means kümeleme filtresi arayüz üzerinde seçilebilir.

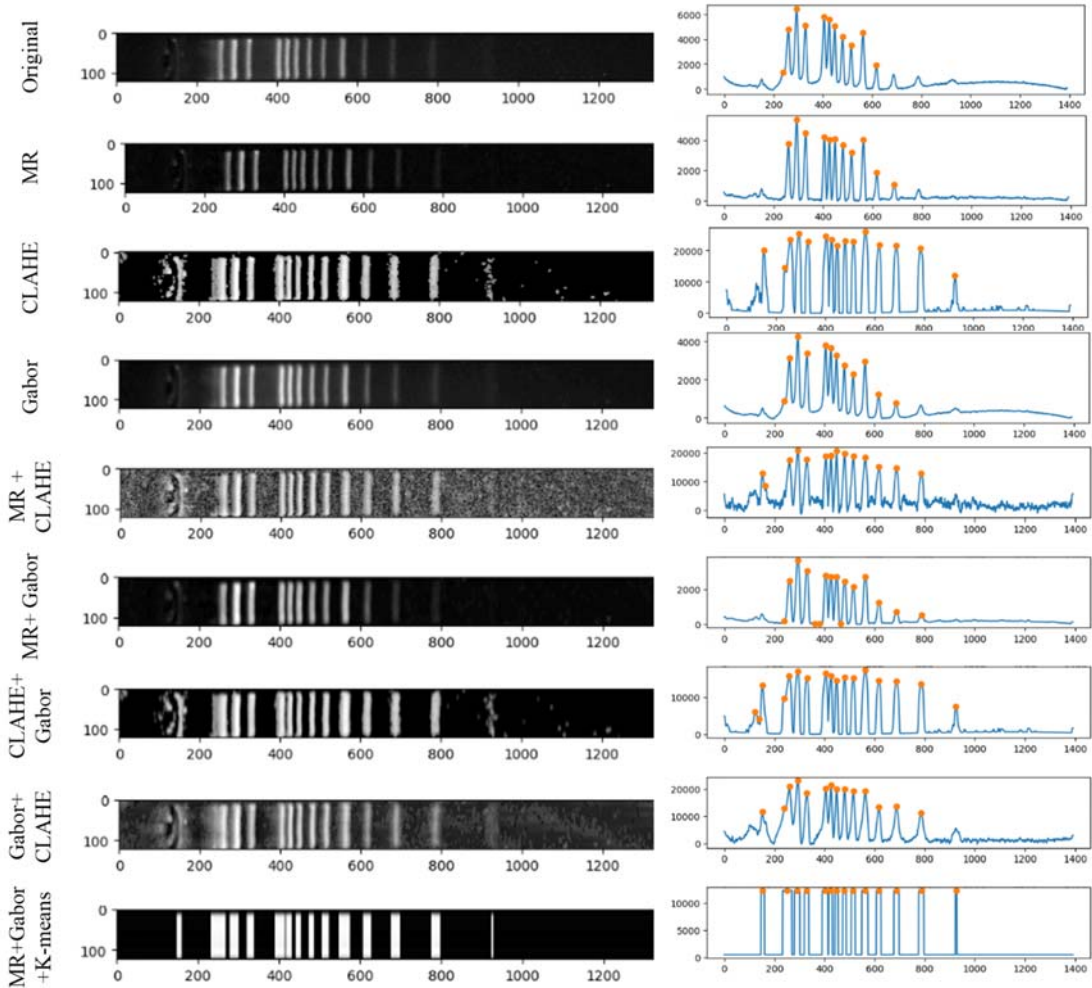
özelliğine sahiptir. Kullanıcı yüklediği görüntü için en uygun k değerini arayüz üzerinden seçebilmektedir. Şekil 4’te kullanılan K-means filtreleme yönteminde $k=2$ olarak alınmıştır.



Şekil 3. K-means kümeleme algoritması
(K-means clustering algorithm) [38]

Çalışmada jel görüntüsüne Şekil 4’te verilen filtre yöntemleri ve bu yöntemlerin birleşimleri uygulanmış, algoritma fonksiyonu ile bulunan tepeler verilmiştir. Grafiklerde son bantların bazı birleşimlerde tamamen kaybolduğu, bazı birleşimlerde oldukça belirgin hale geldiği ve fonksiyon tarafından kolaylıkla tespit edilebildiği görülmektedir. Belirsiz bantları en belirgin hale getiren yöntem birleşimi MR + Gabor + K-means olarak bulunmuş ve bu yöntem birleşimi görüntü işleme fonksiyonunda kullanılmıştır. Filtreleme sonucunda ladder DNA katalog görüntüsünde belirtilen 13 bantın tamamı geliştirilen bu yöntem ile bulunmuştur. Clahe (Contrast Limited Adaptive Histogram Equalization - Kontrast sınırlamalı adaptif histogram eşitleme) 15 tepe ve Clahe + Gabor 17 tepe bulunurken, MR + Gabor + K-means görüntüsünde 13 tepe bulunmuştur. Bu analizler sonucunda en başarılı yöntemin MR + Gabor + K-means birleşimi olduğu görülmektedir.

MR yönteminde görüntülerin arka alan gürültülerine göre tekrar miktarı değişen genişletme/aşındırma uygulanarak bantlar arka plandan ayrıştırılmıştır. K-means yöntemiyle birbirine yakın piksel değerleri her görüntüye göre farklı küme sayısına kümelenebilir görüntüde keskin farklar



Şekil 4. Filtre yöntemlerinin birleşim halinde karşılaştırılması (Combination comparison of filter methods) [26]

oluşturulmuştur. Bu sayede bantların konumları daha fazla belirginleştirilmiştir.

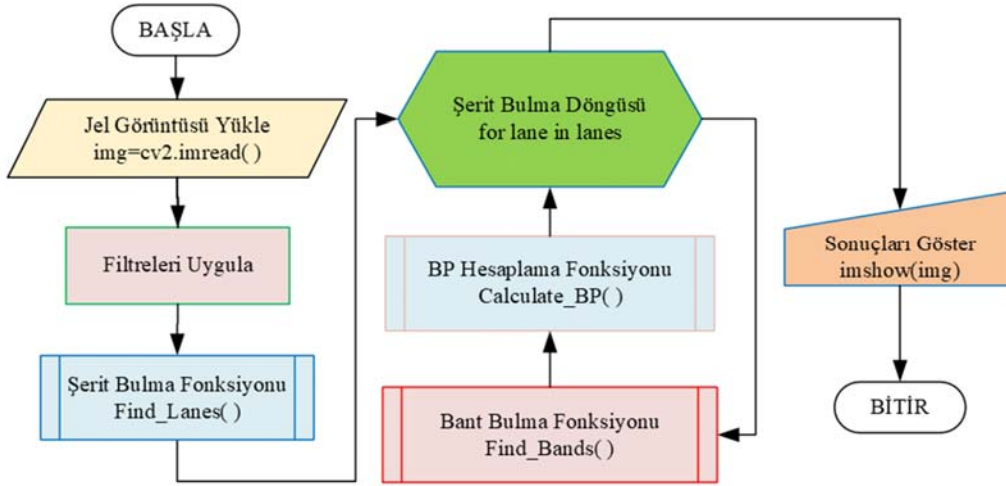
2.3. Şeritleri ve Bantları Bulma (Finding Lanes and Bands)

Agaroz jel görüntülerinde şerit ve bant bulma işlemi için öncelikle filtrelenmiş görüntüdeki pikseller dikey olarak toplanır. Daha sonra findLane() fonksiyonu ile sinyalin taban çizgisi düzeltilir ve şeritlerin kenarları tespit edilir. Oluşturulan findBands() fonksiyonu ile görüntüde şeritlerin bulunduğu bölgelerdeki pikseller yatay olarak toplanarak elde edilen sinyalden bantların bulunduğu tepe noktalarının indisleri kat edilen piksel mesafesi olarak bir diziyeye aktarılır. Bu dizideki değerler ile Ladder DNA'daki BP arasında bir fonksiyon oluşturmak için calculateBP() fonksiyonu çağırılır. Şerit ve bant bulma fonksiyonu için arayüz yazılımı akış şeması Şekil 5'te gösterilmektedir. Şerit ve bant bulma işlemi için Şekil 6a'daki sinyal, görüntüdeki piksellerin dikey toplamını (mavi) ve sinyalin taban çizgisini (turuncu) vermektedir. Görüldüğü üzere sinyalde arka ışık

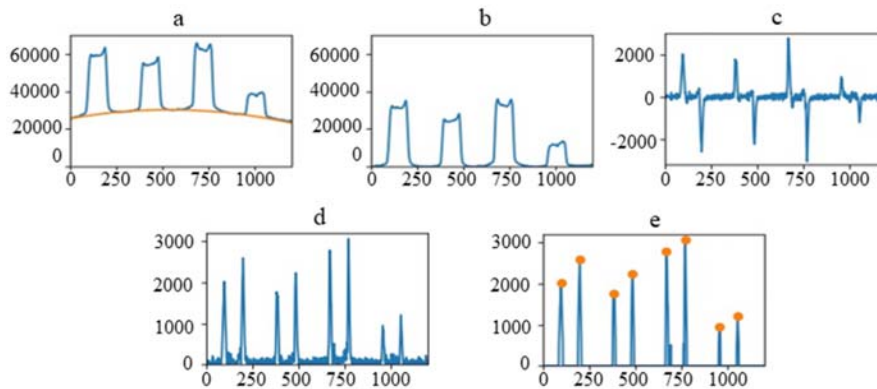
parlamaları ve kontrast farklılıkları nedeniyle bozulmalar vardır. Bu bozulmalar, gürültülü sinyalden taban sinyali (turuncu) çıkarılarak, Şekil 6b'deki düzeltilmiş sinyal elde edilir. Sinyaldeki büyük değişimlerin yerleri, görüntüdeki şeritlerin kenarlarını temsil eder. Kenar noktalarının doğru belirlenmesi için sinyalin türevi alınarak Şekil 6c'deki sinyal elde edilir. Sinyalin türevine pozitif ortalama ve negatif ortalama filtreleri uygulanarak Şekil 6d'de belirtilen sinyaldeki küçük parazitlerin yok edildiği sinyal elde edilir. Ardından, Şekil 6e'de elde edilen sinyalin tepe noktaları bulunur. Gösterilen tepeler, şeritlerin [39] kenar noktalarının konumlarını vermektedir. MR filtresi uygulanmış jel görüntüsü Şekil 7a üzerinde şerit bulma fonksiyonu çalıştırılarak, Şekil 7b'deki şeritlerin mavi çizgi ile belirtildiği görüntü elde edilmektedir.

2.4. Bant Konumlarının Tespiti (Detection of Band Positions)

Dikey olarak toplanan pikseller her şerit için bir de yatay olarak toplanır ve Şekil 8'deki sinyal elde edilir. Buradaki

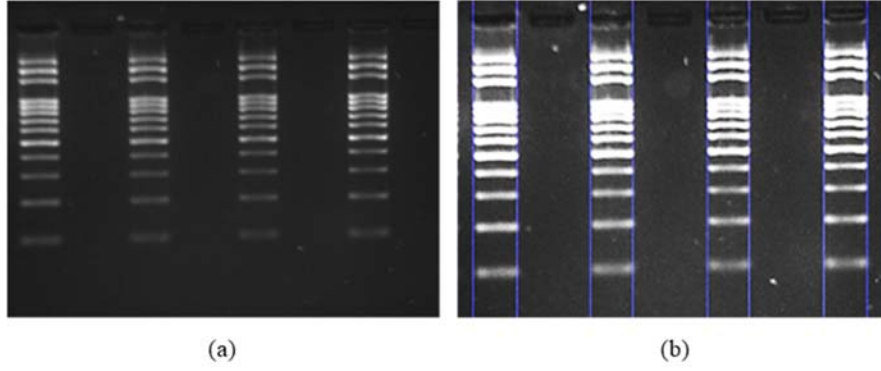


Şekil 5. Arayüz yazılımı akış şeması (Interface software flowchart)

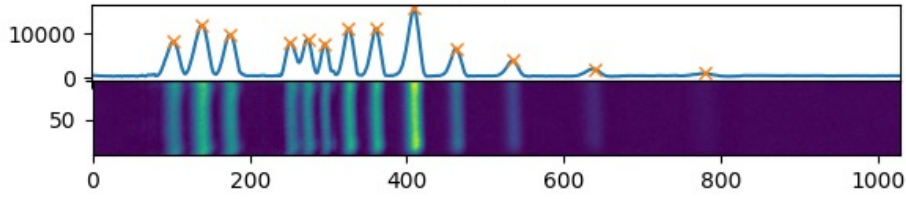


Şekil 6. (a) Jel görüntüsünde piksellerinin dikey toplamı (mavi) ve sinyalin taban çizgisi (turuncu), (b) Düzeltilmiş sinyal, (c) Düzeltilmiş sinyalin türevi, (d) "c" deki sinyalin mutlak değeri, (e) "d" deki sinyale ortalama filtre uygulandığında, sinyalin tepe noktaları (turuncu) şeritlerin kenarlarıdır

((a) Vertical sum of its pixels in the gel image (blue) and baseline of the signal (orange), (b) Smoothed signal, (c) Derivative of the smoothed signal, (d) Absolute value of the signal in "c", (e) When the average filter is applied to the signal in "d", the peaks of the signal (orange) are the edges of the strips.)



Şekil 7. (a) MR filtresi uygulanmış jel görüntüsü, (b) Şerit bulma fonksiyonu uygulanmış jel görüntüsü
(a) Gel image with MR filter applied, (b) Gel image with lane finding function applied

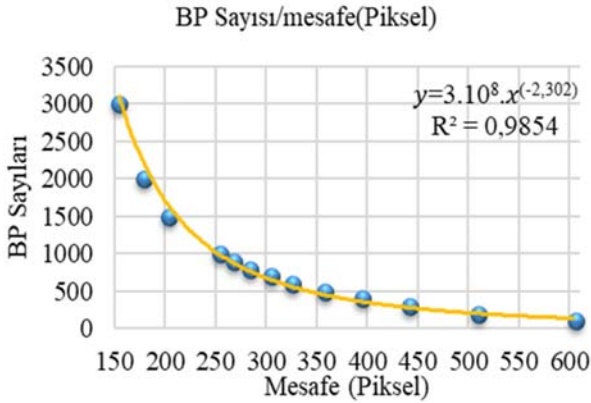


Şekil 8. Şeritteki piksellerin yatay toplanması sonucu bulunan sinyal (mavi) ile sinyalin tepe noktaları (turuncu)
(The signal (blue) and the peaks (orange) of the signal found by the horizontal collection of pixels in the strip)

sinyal incelendiğinde tepelerin tespit edilmesi ile bantların piksel olarak konularının tespit edilmesi mümkündür. Bu sinyal aynı zamanda taban çizgisi sinyalini de içerdiğinden düzeltilmiş bir sinyal, gürültülü sinyalden çıkartılarak elde edilir. Daha sonra elde edilen bu sinyalin tepeleri bulunur ve bantların yerleri piksel olarak ifade edilir. Şekil 8 filtrelenmiş ve tepe noktası tespit edilmiş sinyali göstermektedir.

Piksel mesafeleri ile BP sayıları arasındaki ilişkiyi sağlayan üstel fonksiyon Eş. 3'te verilmektedir. Şekil 9, BP sayıları ile pikseller arasındaki ilişkinin grafiğini ve eğri uydurmanın tutarlılığı-R² değerini göstermektedir.

$$y = 3.10^8 x^{-2,302} \quad (3)$$

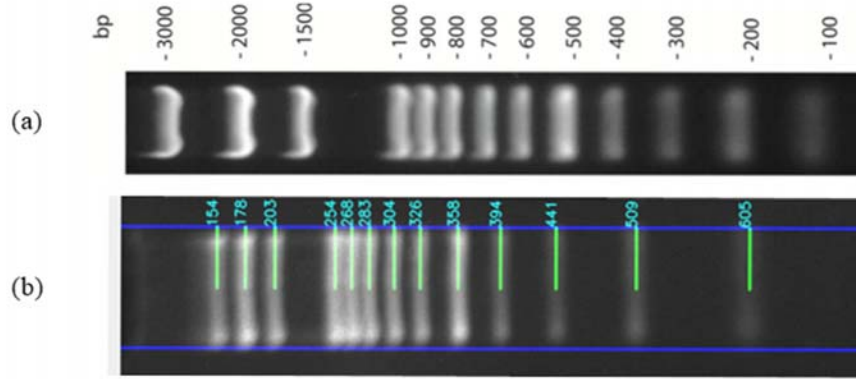


Şekil 9. BP Sayıları ve Mesafesi Grafiği (piksel) (mavi), eğriye uydurma grafiği (Turuncu)
(BP Numbers and Distance Graph (pixel) (blue), curve fit graph (Orange))

Ladder DNA'yı üreten firmadan alınan Şekil 10a'daki bilgi sayfasından DNA örneğinde hangi bandında kaç baz çifti bulunduğunu öğrenilip, baz çifti sayısı ile bantların yürüdüğü mesafe arasında bir fonksiyon oluşturulmaktadır. Oluşturulan bu fonksiyon sayesinde diğer DNA örneklerindeki hareket eden baz çiftlerinin kat ettikleri mesafeler hesaplanarak barındırdıkları baz çift sayıları tahmin edilmektedir. Bu işlemin doğruluğu ölçülen baz çift mesafelerinin gerçeğe yakınlığı ve R² ölçüsü, doğrusal modeller için en genel olarak kullanılan ve rapor edilen hata ve uyumluluk ölçüsüdür. R² tahmini, 0,0 ve 1,0 civarında bir yerde bulunan uyumunu ölçer. R², 1'e ulaştığında, eğri uydurma sürecini durdurmak daha iyidir. R², SSE (sum of squares of errors of regression - regresyon hatalarının karelerinin toplamı) ve SST (sum of squares of total - toplamın karelerinin toplamı) olmak üzere iki terim kullanılarak belirlenir. SSE, gözlemlenen değer ile tahmin edilen değerlerin karelerinin toplamından hesaplanır. Buna regresyonun karelerinin toplamı da denir. SST, gözlemlenen değer ile ortalama değerlerin karelerinin toplamından hesaplanır. Eğri uyumu iyiye, normalde SSE, SST'den daha küçüktür. Daha sonra, Eş. 4'deki denklem kullanılarak R² hesaplanır [40].

$$R^2 = \left[1 - \left(\frac{SSE}{SST} \right) \right] \quad (4)$$

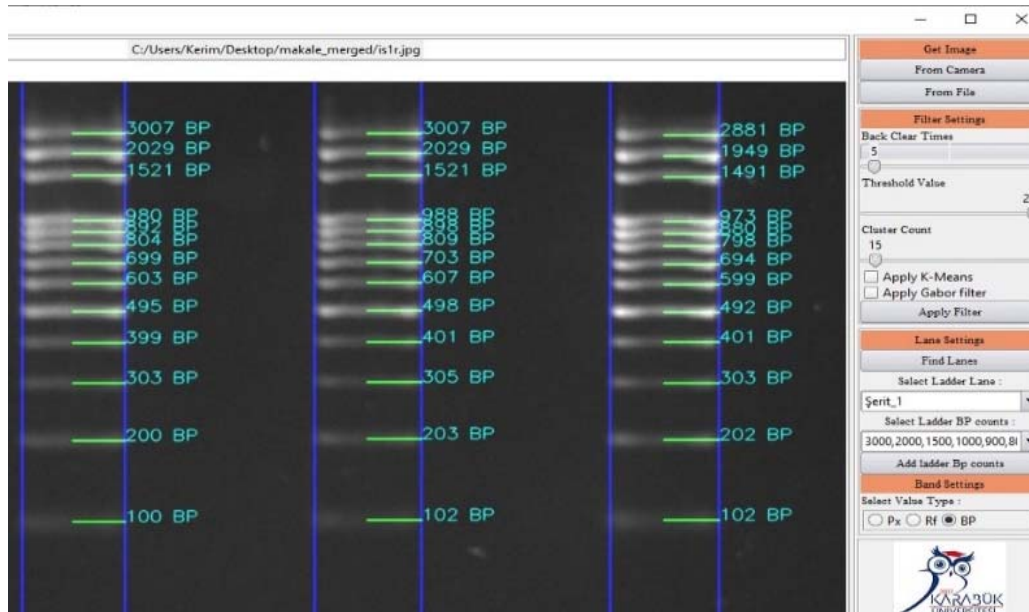
Bu değer 1'e ne kadar yakın ise formül gerçek değerlere o kadar yakın değerleri tahmin ettiği söylenebilmektedir. Yüzde hata değerinin hesaplanmasında her band için gerçek BP sayısından tahmin edilen BP sayısının çıkarılarak hata farkı bulunmaktadır. Bu hata farkı bandın gerçek değerine bölünüp 100 ile çarpılması ile yüzde hata (hata oranı) hesaplanmaktadır. Bulunan piksel tepelerinin mesafeleri Tablo 2'de verilmektedir.



Şekil 10. (a) Ladder DNA bilgi sayfasındaki BP sayıları, (b) Bantların hesaplanan piksel konumları
(a) BP counts of Ladder DNA in datasheet, (b) Calculated pixel locations of the bands)

Tablo 2. Tahmini BP sayılarının ve Gerçek BP sayılarının tasarlanan arayüz yazılımı ile karşılaştırılması
(Comparison of estimated BP numbers and Real BP numbers with the designed interface software.)

| Bant No | Mesafe | Gerçek BP | Tahmini BP | Fark | Hata (%) |
|--------------------------|--------|-----------|------------|------|-------------|
| 1 | 154 | 3000 | 2894 | 106 | 3.53 |
| 2 | 178 | 2000 | 2150 | -150 | 7.50 |
| 3 | 203 | 1500 | 1627 | -127 | 8.47 |
| 4 | 254 | 1000 | 993 | 7 | 0.70 |
| 5 | 268 | 900 | 879 | 21 | 2.33 |
| 6 | 283 | 800 | 776 | 24 | 3.00 |
| 7 | 304 | 700 | 658 | 42 | 6.00 |
| 8 | 326 | 600 | 559 | 41 | 6.83 |
| 9 | 358 | 500 | 445 | 55 | 11.00 |
| 10 | 394 | 400 | 356 | 44 | 11.00 |
| 11 | 441 | 300 | 270 | 30 | 10.00 |
| 12 | 509 | 200 | 191 | 9 | 4.50 |
| 13 | 605 | 100 | 124 | -24 | 24.00 |
| <i>Ortalama Hata (%)</i> | | | | | <i>7.60</i> |



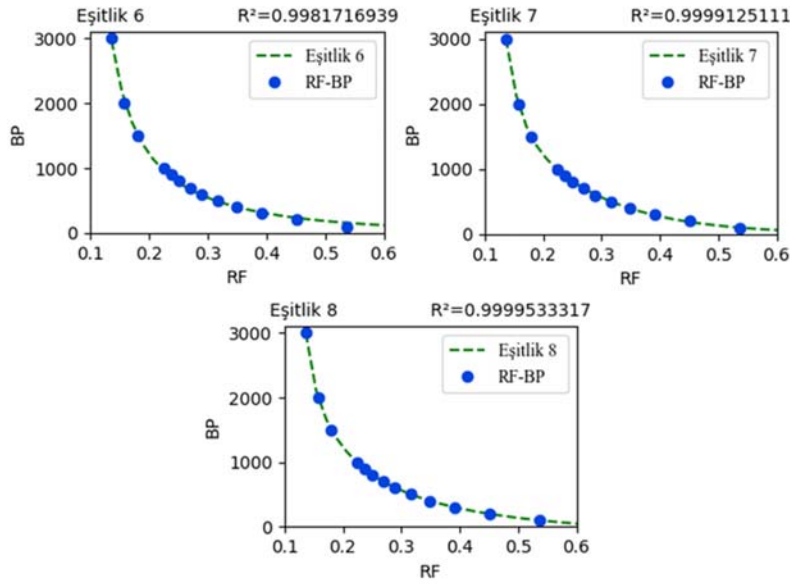
Şekil 11. Arayüz yazılım fonksiyonu tarafından bulunan bantların BP sayıları
(BP numbers of bands found by the interface software function)

Bant konum sinyallerinin tepe deęerinin bulunduęu piksel deęerleri görüntü üzerinde belirtilerek, bu deęerler kullanıcıya arayüz üzerinde Şekil 10b'deki gibi gösterilmektedir. Şekil 10a'da verilen Ladder DNA bilgi sayfasındaki BP sayıları referans alınarak, Şekil 10b'deki baz çiftlerinin ne kadar piksel yol kat ettikleri tespit edilmiştir. Örneğin elde edilen bu deęerler incelendiğinde 3000 BP içeren ilk bantın 154 piksellik bir mesafede, 100 BP içeren 13. bantın ise 605 piksellik bir mesafe olduęu görülmektedir. Bant bulma işleminin tüm şeritlere uygulanması sonucunda Şekil 11'deki görüntü elde edilir. Arayüz fonksiyonu ile görüntüdeki tüm bantların mesafeleri piksel olarak belirlenir. Şeritler ve bantlar belirlendiğinde görüntü üzerinde çizilerek kullanıcıya sunulmaktadır.

2.5. RF (Göreceli yüzey) deęeri ile BP sayısını hesaplayan fonksiyon oluşturma (Creating a Function That Calculates BP Counts with RF (Relative front) Value)

Bantların konumları piksel olarak belirlendikten sonra bu deęerler son mesafeye oranlanarak RF (Relative Front-Göreceli yüzey) deęerlerine dönüştürülür. Ladder DNA BP sayıları, bantların yer deęiştirme sırasına göre regresyon uygulanarak belirlenir.

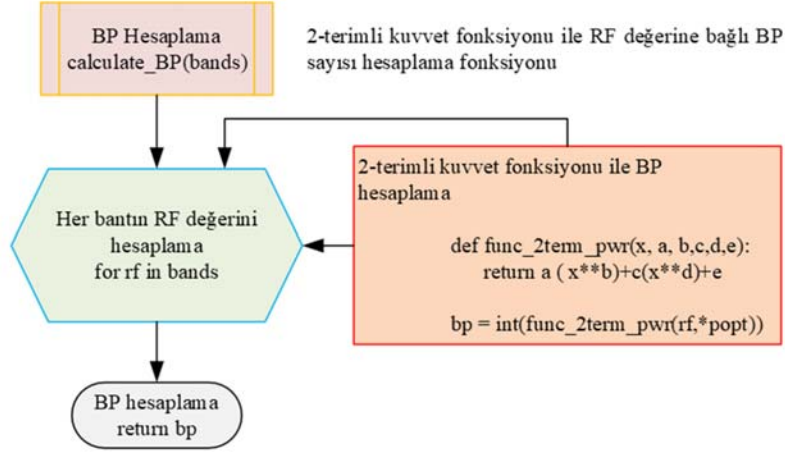
BP sayıları azaldıkça bantların ilerleme mesafesinin arttığı, Tablo 3 incelendiğinde görülmektedir. Yüksek BP sayılarına sahip bantlar arasındaki mesafe kısa iken, BP sayıları azaldıkça iki bant arasındaki mesafe genişir.



Şekil 12. Oluşturulan fonksiyonların grafikleri (Graphs of created functions)

Tablo 3. Arayüz yazılım fonksiyonu tarafından bulunan bantların piksel konumları ve RF deęerleri (Pixel locations and RF values of the bands found by the interface software function)

| Bant No | BP sayıları | Şerit 1 | RF 1 | Şerit 2 | RF 2 | Şerit 3 | RF 3 | Şerit 4 | RF 4 |
|---------|-------------|---------|-------|---------|-------|---------|-------|---------|-------|
| 1 | 3000 | 154 | 0,137 | 149 | 0,132 | 144 | 0,128 | 141 | 0,125 |
| 2 | 2000 | 178 | 0,158 | 173 | 0,153 | 168 | 0,149 | 165 | 0,146 |
| 3 | 1500 | 203 | 0,180 | 198 | 0,176 | 193 | 0,171 | 189 | 0,168 |
| 4 | 1000 | 254 | 0,225 | 250 | 0,222 | 244 | 0,216 | 240 | 0,213 |
| 5 | 900 | 268 | 0,238 | 263 | 0,233 | 257 | 0,228 | 254 | 0,225 |
| 6 | 800 | 283 | 0,251 | 277 | 0,246 | 272 | 0,241 | 269 | 0,238 |
| 7 | 700 | 304 | 0,270 | 299 | 0,265 | 293 | 0,260 | 290 | 0,257 |
| 8 | 600 | 326 | 0,289 | 322 | 0,285 | 316 | 0,280 | 312 | 0,276 |
| 9 | 500 | 358 | 0,317 | 353 | 0,313 | 348 | 0,309 | 344 | 0,305 |
| 10 | 400 | 394 | 0,349 | 389 | 0,345 | 383 | 0,340 | 379 | 0,336 |
| 11 | 300 | 441 | 0,391 | 436 | 0,387 | 429 | 0,380 | 425 | 0,377 |
| 12 | 200 | 509 | 0,451 | 501 | 0,444 | 494 | 0,437 | 491 | 0,435 |
| 13 | 100 | 605 | 0,536 | 596 | 0,528 | 587 | 0,520 | 582 | 0,516 |



Şekil 13. İki terimli kuvvet fonksiyonu akış şeması (Binomial force function flow chart)

Tablo 3'te belirtilen RF değerleri, gidilecek son mesafeye göre makromoleküllerin ne kadar göç ettiklerini belirleyerek hesaplanır. Fonksiyonların parametreleri, RF değerleri ile BP sayıları arasında oluşturulan regresyon ile bulunur. Regresyon kullanılarak elde edilen parametreler ile, BP sayılarını RF değerlerine göre tahmin eden bir fonksiyon oluşturulabilir. Üstel fonksiyonların kuvvet fonksiyon serisi olarak da belirtilebilmesi ile Eş. 5'te verilen fonksiyon, görüntü işleme fonksiyonu olarak yazılım üzerinde jel görüntülerine uygulanmıştır. Eş. 5'te belirtilen "x" değişkeni RF değeri, "y" değişkeni ise BP değeri olarak kullanılmaktadır. BP yi hesaplamak için Python programlama dilinde scipy kütüphanesi kullanılarak regresyon fonksiyonunun a, b, c, d ve e katsayıları hesaplanmıştır. Regresyon fonksiyonu tanımlandıktan sonra Eş. 8 elde edilmiş ve en düşük tahmin hatası ile sonuçlara ulaşılmıştır.

$$y = a(x^b) + c(x^d) + e \quad (5)$$

Makale çalışmasında, en iyi fonksiyona ulaşmak için 3 farklı fonksiyon parametreleri ile birlikte hesaplanarak; Üstel fonksiyon [41] (Eş. 6), iki terimli üstel fonksiyon [42] (Eş. 7) ve iki terimli kuvvet fonksiyonu (Eş. 8) oluşturulmuştur.

$$y = -0,00245832 + 0,00100707(1 + e^{-x/(-0,30)}) \quad (6)$$

$$y = 899542,94. e^{-37,99.x} + 5293,50. e^{-6,24x} \quad (7)$$

$$y = 1390319982. x^{-1,46} + 4,33. 10^{-6}. x^{-9,49} - 247,25 \quad (8)$$

Bu fonksiyonlar ile oluşturulan RF-BP grafikleri Şekil 12'de gösterilmektedir. R^2 değerleri karşılaştırıldığında en düşük değer Eş. 6'daki $R^2=0,99817216$, biraz daha büyük değer olan Eş. 7'deki $R^2=0,9999125$ ve en yüksek değer olarak Eş. 8'deki $R^2=0,9999533$ olarak bulunmuştur.

R^2 değerinin daha yüksek bir değer olarak hesaplanabilmesi için, üstel fonksiyonun kuvvet fonksiyonu biçimine uyarlanması sonucunda iki terimli kuvvet fonksiyonu

oluşturularak 1'e en yakın $R^2=0,9999533$ değeri elde edilmiştir. Bu değer ile Şekil 13'te verilen blok şemada tanımlanan tahmin algoritma fonksiyonu ortalama hata değerini %2'den %0,55-%0,86 arasında değişen değerlere düşürülebilmektedir. Hata oranlarının jel görüntülerinin en düşük değer ile iyileştirilmesinde kullanılabilecek yeni bir görüntü işleme algoritma fonksiyonu, bu çalışma ile önerilmektedir. Fonksiyonlar arasındaki farkların daha iyi görülebilmesi için, bantların RF değerleri fonksiyonlara uygulandığında tahmini BP sayıları ile gerçek BP sayıları ve ortalama tahmin hata oranları Tablo 4'te detaylı olarak karşılaştırılmaktadır. Tablodaki değerlere göre; R^2 değerleri 1'e en yakın olan Eş. 7 ve Eş. 8 fonksiyonlarının tahmin ettiği BP sayılarında, ortalama tahmin hata oranı en düşük Eş. 8 olmasına rağmen, Eş. 7'nin bazı bantlarda daha yakın tahminler yaptığı söylenebilir.

4. SONUÇLAR VE TARTIŞMALAR (RESULTS AND CONCLUSIONS)

Bu çalışmada tasarlanan sistem ile gömülü sistem üzerinde elektroforez jel görüntüleri üzerinde görüntü işleme analizi yapmak için görüntü filtreleme yöntemleri ve farklı birleşimleri de uygulanarak arka plan gürültüsünü gidermek için en etkili yöntem olarak MR + Gabor + K-means filtreleme yöntem birleşimi geliştirilmiştir.

Bantların konumlarını bulma işlemi ile piksel üzerinde RF değeri olarak hesaplanmış ve bu RF değerleri ile gerçek BP sayıları arasında bir fonksiyon geliştirilmiştir. Geliştirilen fonksiyon, matematiksel olarak üstel fonksiyonların aynı zamanda kuvvet fonksiyonu olarak da ifade edilebilmesi özelliğine dayanılarak iki terimli üstel fonksiyonun iki terimli kuvvet fonksiyonu biçiminde kullanılması ile elde edilmiştir. Elde edilen fonksiyon jel görüntüsüne uygulandığında BP sayılarının tahmininde R^2 değerinin 0,99995 değeri ile 1'e en yakın değer olmasıyla %0,55-%0,86 değerleri arasında oldukça düşük bir ortalama hata değeri elde edilmiştir. Bu sonuç ile jel görüntü analizi için yüksek R^2 değerine sahip en düşük ortalama hata ile kullanılabilecek bir görüntü işleme algoritma fonksiyonu önerilmektedir.

Tablo 4. RF değerleri fonksiyonlara uygulandığında tahmini BP sayıları ve gerçek BP sayıları ile ortalama tahmin hata oranları (The predicted BP counts and real BP counts with average predictive error rates when RF values are applied to functions)

| Şerit 1 Bant No | Gerçek BP | Eşitlik 6 | | | Eşitlik 7 | | | Eşitlik 8 | | |
|--------------------|--------------|------------------|------|---------|-----------|------|---------|-----------|------|---------|
| | | Tahmin | Hata | Hata(%) | Tahmin | Hata | Hata(%) | Tahmin | Hata | Hata(%) |
| 1 | 3000 | 2972 | 28 | 0,93 | 2999 | 1 | 0,03 | 3000 | 0 | 0,00 |
| 2 | 2000 | 2062 | 62 | 3,10 | 2005 | 5 | 0,25 | 1996 | 4 | 0,20 |
| 3 | 1500 | 1533 | 33 | 2,20 | 1495 | 5 | 0,33 | 1510 | 10 | 0,67 |
| 4 | 1000 | 963 | 37 | 3,70 | 987 | 13 | 1,30 | 989 | 11 | 1,10 |
| 5 | 900 | 865 | 35 | 3,89 | 896 | 4 | 0,44 | 894 | 6 | 0,67 |
| 6 | 800 | 776 | 24 | 3,00 | 811 | 11 | 1,38 | 805 | 5 | 0,63 |
| 7 | 700 | 674 | 26 | 3,71 | 707 | 7 | 1,00 | 700 | 0 | 0,00 |
| 8 | 600 | 587 | 13 | 2,17 | 613 | 13 | 2,17 | 608 | 8 | 1,33 |
| 9 | 500 | 487 | 13 | 2,60 | 499 | 1 | 0,20 | 498 | 2 | 0,40 |
| 10 | 400 | 401 | 1 | 0,25 | 397 | 3 | 0,75 | 401 | 1 | 0,25 |
| 11 | 300 | 318 | 18 | 6,00 | 294 | 6 | 2,00 | 302 | 2 | 0,67 |
| 12 | 200 | 233 | 33 | 16,50 | 190 | 10 | 5,00 | 198 | 2 | 1,00 |
| 13 | 100 | 156 | 56 | 56,00 | 103 | 3 | 3,00 | 99 | 1 | 1,00 |
| | | Ortalama Hata(%) | | | | | | 0,61 | | |
| | | R ² | | | 0,99817 | | | 0,99991 | | |
| Şerit 2 Bant No | Gerçek BP | Eşitlik 6 | | | Eşitlik 7 | | | Eşitlik 8 | | |
| | | Tahmin | Hata | Hata(%) | Tahmin | Hata | Hata(%) | Tahmin | Hata | Hata(%) |
| 1 | 3000 | 2971 | 29 | 0,97 | 2999 | 1 | 0,03 | 3000 | 0 | 0,00 |
| 2 | 2000 | 2063 | 63 | 3,15 | 2004 | 4 | 0,20 | 1996 | 4 | 0,20 |
| 3 | 1500 | 1534 | 34 | 2,27 | 1496 | 4 | 0,27 | 1511 | 11 | 0,73 |
| 4 | 1000 | 956 | 44 | 4,40 | 982 | 18 | 1,80 | 983 | 17 | 1,70 |
| 5 | 900 | 865 | 35 | 3,89 | 898 | 2 | 0,22 | 895 | 5 | 0,56 |
| 6 | 800 | 782 | 18 | 2,25 | 817 | 17 | 2,13 | 812 | 12 | 1,50 |
| 7 | 700 | 674 | 26 | 3,71 | 707 | 7 | 1,00 | 701 | 1 | 0,14 |
| 8 | 600 | 583 | 17 | 2,83 | 609 | 9 | 1,50 | 604 | 4 | 0,67 |
| 9 | 500 | 487 | 13 | 2,60 | 499 | 1 | 0,20 | 498 | 2 | 0,40 |
| 10 | 400 | 401 | 1 | 0,25 | 396 | 4 | 1,00 | 400 | 0 | 0,00 |
| 11 | 300 | 317 | 17 | 5,67 | 293 | 7 | 2,33 | 301 | 1 | 0,33 |
| 12 | 200 | 235 | 35 | 17,50 | 193 | 7 | 3,50 | 200 | 0 | 0,00 |
| 13 | 100 | 158 | 58 | 58,00 | 105 | 5 | 5,00 | 98 | 2 | 2,00 |
| | | Ortalama Hata(%) | | | | | | 0,63 | | |
| | | R ² | | | 0,99804 | | | 0,99988 | | |
| Şerit 3 Bant No | Gerçek BP | Eşitlik 6 | | | Eşitlik 7 | | | Eşitlik 8 | | |
| | | Tahmin | Hata | Hata(%) | Tahmin | Hata | Hata(%) | Tahmin | Hata | Hata(%) |
| 1 | 3000 | 2972 | 28 | 0,93 | 2999 | 1 | 0,03 | 3000 | 0 | 0,00 |
| 2 | 2000 | 2062 | 62 | 3,10 | 2005 | 5 | 0,25 | 1997 | 3 | 0,15 |
| 3 | 1500 | 1532 | 32 | 2,13 | 1495 | 5 | 0,33 | 1508 | 8 | 0,53 |
| 4 | 1000 | 961 | 39 | 3,90 | 985 | 15 | 1,50 | 986 | 14 | 1,40 |
| 5 | 900 | 869 | 31 | 3,44 | 900 | 0 | 0,00 | 897 | 3 | 0,33 |
| 6 | 800 | 779 | 21 | 2,63 | 813 | 13 | 1,63 | 808 | 8 | 1,00 |
| 7 | 700 | 675 | 25 | 3,57 | 708 | 8 | 1,14 | 702 | 2 | 0,29 |
| 8 | 600 | 584 | 16 | 2,67 | 609 | 9 | 1,50 | 605 | 5 | 0,83 |
| 9 | 500 | 484 | 16 | 3,20 | 495 | 5 | 1,00 | 495 | 5 | 1,00 |
| 10 | 400 | 400 | 0 | 0,00 | 395 | 5 | 1,25 | 400 | 0 | 0,00 |
| 11 | 300 | 317 | 17 | 5,67 | 294 | 6 | 2,00 | 302 | 2 | 0,67 |
| 12 | 200 | 235 | 35 | 17,50 | 193 | 7 | 3,50 | 200 | 0 | 0,00 |
| 13 | 100 | 158 | 58 | 58,00 | 106 | 6 | 6,00 | 99 | 1 | 1,00 |
| | | Ortalama Hata(%) | | | | | | 0,55 | | |
| | | R ² | | | 0,99815 | | | 0,99990 | | |
| Şerit 4 Bant No | Gerçek BP | Eşitlik 6 | | | Eşitlik 7 | | | Eşitlik 8 | | |
| | | Tahmin | Hata | Hata(%) | Tahmin | Hata | Hata(%) | Tahmin | Hata | Hata(%) |
| 1 | 3000 | 2972 | 28 | 0,93 | 3000 | 0 | 0,00 | 3001 | 1 | 0,03 |
| 2 | 2000 | 2056 | 56 | 2,80 | 2000 | 0 | 0,00 | 1993 | 7 | 0,35 |
| 3 | 1500 | 1542 | 42 | 2,80 | 1503 | 3 | 0,20 | 1515 | 15 | 1,00 |
| 4 | 1000 | 963 | 37 | 3,70 | 987 | 13 | 1,30 | 987 | 13 | 1,30 |
| 5 | 900 | 864 | 36 | 4,00 | 896 | 4 | 0,44 | 892 | 8 | 0,89 |
| 6 | 800 | 775 | 25 | 3,13 | 809 | 9 | 1,13 | 804 | 4 | 0,50 |
| 7 | 700 | 672 | 28 | 4,00 | 705 | 5 | 0,71 | 700 | 0 | 0,00 |
| 8 | 600 | 584 | 16 | 2,67 | 611 | 11 | 1,83 | 608 | 8 | 1,33 |
| 9 | 500 | 484 | 16 | 3,20 | 497 | 3 | 0,60 | 498 | 2 | 0,40 |
| 10 | 400 | 401 | 1 | 0,25 | 397 | 3 | 0,75 | 402 | 2 | 0,50 |
| 11 | 300 | 318 | 18 | 6,00 | 295 | 5 | 1,67 | 304 | 4 | 1,33 |
| 12 | 200 | 235 | 35 | 17,50 | 193 | 7 | 3,50 | 199 | 1 | 0,50 |
| 13 | 100 | 160 | 60 | 60,00 | 108 | 8 | 8,00 | 97 | 3 | 3,00 |
| | | Ortalama Hata% | | | | | | 0,86 | | |
| | | R ² | | | 0,99807 | | | 0,99992 | | |

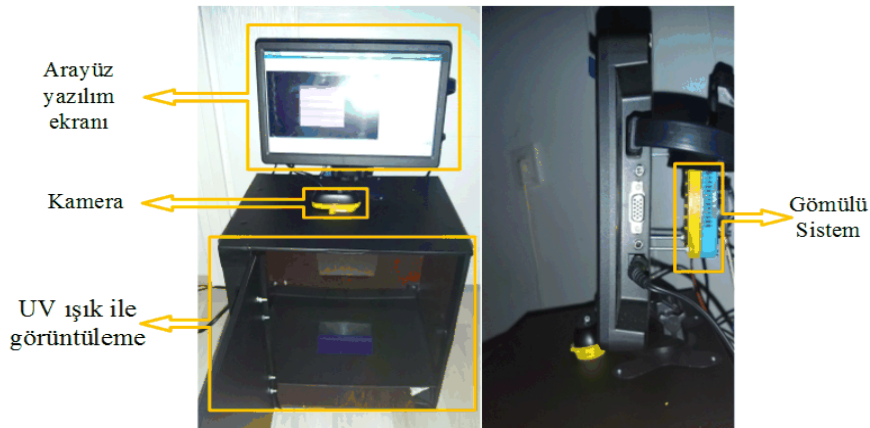
Tablo 5. Tasarlanan arayüz yazılımı ile mevcut jel görüntü analiz programlarının karşılaştırılması
(Comparison of the designed interface software and existing gel image analysis programs)

| | GelQuest | PyElph | GelAnalyzer | GelJ | Tasarlanan Arayüz Yazılımı |
|-----------------------------------|-----------|--------|-------------|--------|----------------------------|
| R ² | - | - | 0,985 | 0,9952 | 0,9999 |
| Otomatik Şerit Bulma | - | + | + | - | + |
| Otomatik Bant Bulma | - | + | + | + | + |
| Filtreler | K-Means | - | - | - | + |
| | Gabor | - | - | - | + |
| | MR | - | - | - | + |
| | Diğer | - | - | - | + |
| Görüntü Alma | Dosyadan | + | + | + | + |
| | Kameradan | - | - | - | + |
| UV Görüntüleme Sistemi | - | - | - | - | + |
| Arayüzde UV Işık Kontrol İşlevi | - | - | - | - | + |
| Gömülü Sisteme Uyumluluk | - | - | - | - | + |
| Bulunan Bant Sayısına (13 olmalı) | 13 | 14 | 12 | 13 | 13 |
| Ortalama Hata Oranı (%) | - | 9,813 | 8,422 | 14,02 | 1,174 |

Tablo 6. Programların hesapladığı BP sayılarında ortalama hatalarının karşılaştırılması
(Comparison of mean errors in BP numbers calculated by programs)

| GBP* | Pyelph | | | GelAnalyzer | | | GelJ | | | Tasarlanan Arayüz Yazılımı | | |
|----------------|--------|------|-------|-------------|------|-------|--------|------|-------|----------------------------|------|-------|
| | Tahmin | Hata | Hata% | Tahmin | Hata | Hata% | Tahmin | Hata | Hata% | Tahmin | Hata | Hata% |
| 3000 | 2243 | 757 | 25,23 | 2815 | 185 | 6,167 | 3,016 | 16 | 0,533 | 2993 | 7 | 0,233 |
| 2000 | 1898 | 102 | 5,1 | 2173 | 173 | 8,65 | 1,902 | 98 | 4,9 | 2000 | 0 | 0 |
| 1500 | 1606 | 106 | 7,067 | 1666 | 166 | 11,07 | 1,508 | 8 | 0,533 | 1542 | 42 | 2,8 |
| 1000 | 1113 | 113 | 11,3 | 1007 | 7 | 0,7 | 1,039 | 39 | 3,9 | 999 | 1 | 0,1 |
| 900 | 1021 | 121 | 13,44 | 882 | 18 | 2 | 950 | 50 | 5,556 | 902 | 2 | 0,222 |
| 800 | 924 | 124 | 15,5 | 767 | 33 | 4,125 | 866 | 66 | 8,25 | 817 | 17 | 2,125 |
| 700 | 793 | 93 | 13,29 | 651 | 49 | 7 | 773 | 73 | 10,43 | 708 | 8 | 1,143 |
| 600 | 674 | 74 | 12,33 | 546 | 54 | 9 | 669 | 69 | 11,5 | 610 | 10 | 1,667 |
| 500 | 539 | 39 | 7,8 | 442 | 58 | 11,6 | 550 | 50 | 10 | 496 | 4 | 0,8 |
| 400 | 418 | 18 | 4,5 | 367 | 33 | 8,25 | 442 | 42 | 10,5 | 400 | 0 | 0 |
| 300 | 300 | 0 | 0 | 306 | 6 | 2 | 317 | 17 | 5,667 | 305 | 5 | 1,667 |
| 200 | 186 | 14 | 7 | 261 | 61 | 30,5 | 165 | 35 | 17,5 | 203 | 3 | 1,5 |
| 100 | 95 | 5 | 5 | - | 100 | - | 7 | 93 | 93 | 103 | 3 | 3 |
| Ortalama Hata% | | | 9,813 | | | 8,422 | | | 14,02 | | | 1,174 |

*GBP: Gerçek BP değerleri

**Şekil 14.** Tasarlanan görüntü elde etme ve görüntü işleme sistemi

Makale çalışması ve geliştirilen jel görüntü işleme arayüz yazılımı, jel görüntülerinden BP sayılarının hesaplanabilmesi için gerekli özellikler karşılaştırıldığında,

mevcut olarak kullanılan programlara göre üstünlükleri mevcuttur. Bu özellikler bir tablo halinde kullanımda olan bazı açık kaynak programlar ile karşılaştırıldığında Tablo

5'te belirtilen üstünlükler ortaya çıkmıştır. Bu tabloda “+”, “-” işaretleri sırasıyla tasarlanan band görüntüsü işleme arayüz sistemi ile diğer yazılımların belirtilen özelliklere sahip olup/olmama durumunu göstermektedir. Tasarlanan görüntü işleme arayüz yazılımı R^2 değeri 1'e yakınlığı başta olmak üzere birçok özellikte diğer jel görüntüsü işleme yazılımlarından daha üstün olduğu tabloda belirtilmiştir. Geliştirilen algoritmanın gömülü sistem üzerinde çalışma zamanı *process_time_ns* kodu ile hesaplandığında 2,5 sn gibi kısa bir sürede görüntü işleme analizini gerçekleştirdiği tespit edilmiştir. Bu süre jel görüntüsü yüklendikten sonra band bulma işleminin başlatılması ile görüntü üzerinde bandların çizilmesi ve BP sayılarının görüntülenmesine kadar geçen süredir. Ortalama yüzde hata değeri olan Tablo 5'teki dört program için BP sayılarının tahminindeki ortalama hata hesaplaması Tablo 6'da verilen gerçek BP değerlerine göre yapılarak karşılaştırılmış ve tasarlanan arayüz yazılım programının en düşük hataya sahip olduğu gösterilmiştir. PyQuest programında her band için hesaplanan BP değeri ve R^2 değeri programdan elde edilemediği için karşılaştırmaya dahil edilmemiştir.

Tasarlanan sistem Şekil 14'te gösterildiği gibi karanlık alt bölgede bir UV ışık kaynağı üzerinde, jel görüntülerine dik açıda yerleştirilmiş bir kameradan gömülü sistem üzerindeki arayüz yazılımına görüntülerin anlık aktararak işlenmesi sağlanmaktadır. Jel görüntüleri üzerinden BP sayılarının en düşük yüzde hata ile hesaplanmasını ve bu sayıların jel görüntüsü üzerinde kullanıcıya sunulmasını sağlayan bu sistem, Gömülü Sistem Kontrollü Mini Jel Elektroferez sistemi [11] ile birleştirilerek kullanılabilir. Bu iki sistem, teknoloji temelinde özgün yazılımlar ile elektroferez işlem sürecini ve sonrasında jel görüntülerini gömülü sisteme aktaran ve görüntüleri diğer muadillerine göre az hata oranıyla analiz eden daha kompakt bir tasarım oluşturulmuştur. Bu sayede daha az yer kaplayan ve uygun maliyetli bir sistem olarak mikrobiyoloji, kimya ve tıbbi laboratuvarlarda kullanılabilir olduğu bu çalışmalar ile ortaya konulmuştur.

KAYNAKLAR (REFERENCES)

1. Yalcintepe L., New Approaches in Electrophoresis, Journal of Engineering and Naturel Sciences, Sigma 27, 151–160, Mar. 2009.
2. Francisco da Silva Neto G., Luíza de Andrade Rodrigues M., and Fonseca A., A New Quantitative Gel Electrophoresis Method With Image-Based Detection for The Determination of Food Dyes and Metallic Ions, Talanta, 221, 121602, Jan. 2021.
3. Intarapanich A., Kaewkamnerd S., Shaw P., Ukosakit K., Tragoonrunng S., and Tongshima S., Automatic DNA Diagnosis for 1D Gel Electrophoresis Images Using Bio-Image Processing Technique, BMC Genomics, 16, S15, Dec. 2015.
4. Green M. R. and Sambrook J., Agarose Gel Electrophoresis, Cold Spring Harbor Protocols, 2019 (1), 87–94, 2019.
5. Bourzac K. M., LaVine L. J., and Rice M. S., Analysis of DAPI and SYBR Green I as Alternatives to Ethidium Bromide for Nucleic Acid Staining in Agarose Gel Electrophoresis, Journal of Chemical Education, 80 (11), 1292–1296, Nov. 2003.
6. Yılmaz M., Öziç C., and Gök İ., Principles of Nucleic Acid Separation by Agarose Gel Electrophoresis, Gel Electrophoresis - Principles and Basics, 33-40, 2012.
7. Ülker O., Ertuğrul İ., Akkuş N., Characterization Of Electrothermal Micro-Actuator Using Image Processing Algorithm, Journal of the Faculty of Engineering and Architecture of Gazi University, 36 (2), 1160-1170, 2021.
8. Alkan A., Selcuk T., Çolakoğlu A.S., Bread Texture Analysis and Development of User Interface Using Image Processing Techniques, Journal of the Faculty of Engineering and Architecture of Gazi University, 33 (1), 31-41, 2018.
9. Can G.F., Fiğlalı N., Image Processing Based Rapid Upper Limb Assessment Method, Journal of the Faculty of Engineering and Architecture of Gazi University, 32 (3), 719-731, 2017.
10. Ausubel F. M., Current Protocols in Molecular Biology, Molecular Reproduction and Development, 1(2), 146, Jan. 1989.
11. Mersinkaya İ. and Kavsaoglu A. R., Python ile Mini Jel Elektroferez Kontrol Yazılımı ve Sistem Tasarımı, Gazi Üniversitesi Fen Bilimleri Dergisi Part C: Tasarım ve Teknoloji, 7 (4), 969–984, 2019.
12. Khakabimamaghani S., Najafi A., Ranjbar R., and Raam M., GelClust: A Software Tool for Gel Electrophoresis Images Analysis and Dendrogram Generation, Computer Methods and Programs in Biomedicine, 111 (2), 512–518, Aug. 2013.
13. Abu-Asab M. S. et al., Biomarkers in the Age of Omics: Time for a Systems Biology Approach, OMICS A Journal of Integrative Biology, 15 (3). 105–112, Mar. 2011.
14. Adamczyk M., van Eunen K., Bakker B. M., and Westerhoff H. v., Enzyme Kinetics for Systems Biology: When, Why and How, in Methods in Enzymology, 500, 233–257, 2011.
15. Gel Doc EZ, Bio-Rad. Gel Doc EZ Gel Documentation System. <https://www.bio-rad.com/en-tr/product/gel-doc-ez-gel-documentation-system?ID=O4950KKG4>, Erişim tarihi Mayıs 15, 2021.
16. Adams L. G. et al., Multi-Comparative Systems Biology Analysis Reveals Time-Course Biosignatures of in Vivo Bovine Pathway Responses to B. Melitensis, S. Enterica Typhimurium and M. Avium Paratuberculosis, BMC Proceedings, 5 (S4), Dec. 2011.
17. Alan A. et al., A Systems Biology Approach to Infectious Disease Research: Innovating the Pathogen-Host Research Paradigm, mBio, 2(1), e00325-10, Jun. 2021.
18. Afacan N. J., Fjell C. D., and Hancock R. E. W., A Systems Biology Approach to Nutritional Immunology - Focus on Innate Immunity, Molecular Aspects of Medicine, 33 (1), 14–25, Feb. 2012.
19. Agustí A., Sobradillo P., and Celli B., Addressing the Complexity of Chronic Obstructive Pulmonary Disease: From Phenotypes and Biomarkers to Scale-Free

- Networks, Systems Biology, and P4 Medicine, *American Journal of Respiratory and Critical Care Medicine*, 183 (9), 1129–1137, 2011.
20. Gerner-Smidt P., Graves L. M., Hunter S., and Swaminathan B., Computerized Analysis of Restriction Fragment Length Polymorphism Patterns: Comparative Evaluation of Two Commercial Software Packages, *Journal of Clinical Microbiology*, 36 (5), 1318–1323, 1998.
 21. Rementeria A., Gallego L., Quindós G., and Garaizar J., Comparative Evaluation of Three Commercial Software Packages for Analysis of DNA Polymorphism Patterns, *Clinical Microbiology and Infection*, 7 (6), 331–336, 2001.
 22. Cardinali G., Martini A., Preziosi R., Bistoni F., and Baldelli F., Multicenter Comparison of Three Different Analytical Systems for Evaluation of DNA Banding Patterns from *Cryptococcus Neoformans*, *Journal of Clinical Microbiology*, 40 (6), 2095–2100, 2002.
 23. Pavel A. B. and Vasile C. I., PyElph - A Software Tool for Gel Images Analysis and Phylogenetics, *BMC Bioinformatics*, 13 (1), 2012.
 24. Lazar I., GelAnalyzer 19.1. <http://www.gelanalyzer.com>, Erişim tarihi Mayıs 15, 2021.
 25. Horgan G. W. and Glasbey C. A., Uses of Digital Image Analysis in Electrophoresis, *Electrophoresis*, 16 (1), 298–305, 1995.
 26. Özkara K., Elektroforez Görüntü İşleme Arayüzü ve Yazılımı Geliştirme, Yüksek Lisans Tezi, Karabük Üniversitesi, Lisansüstü Enstitüsü, Karabük, 2019.
 27. Kusim A. S., Abdullah N. E., and Ab Rahim A. A., An Image Processing Enhancement Approach to Extract Information in Gel Electrophoresis Image, *Proceedings of 2013 International Conference on Technology, Informatics, Management, Engineering and Environment, TIME-E 2013*, 9–16, 2013.
 28. Meşe A. K., Erdamar A., and İşeri Ö. D., Image Analysis for Single Cell Gel Electrophoresis, *25th Signal Processing and Communications Applications Conference (SIU)*, 1–4, 2017.
 29. Gonzalez R. C., Woods R. E., and Masters B. R., *Digital Image Processing*, Third Edition, *Journal of Biomedical Optics*, 14 (2), 029901, 2009.
 30. Chityala R., and Pudipeddi S., *Image Processing and Acquisition Using Python 2nd Edition*, Chapman and Hall/CRC, 2020.
 31. Goutsias J., Chapter 4 Morphological Methods for Biomedical Image Analysis, *Handbook of Medical Imaging*, 175–263, 2009.
 32. Walker J. M., Chapter 20 Computer Analysis of 2-D Electrophoresis Gels, *New Protein Techniques*, 1988.
 33. Bradsky G. and Kaehler A., *Learning OpenCV- Computer Vision with the OpenCV Library*, 16 (3), 2009.
 34. Fogel I. and Sagi D., Gabor Filters as Texture Discriminator, *Biological Cybernetics*, 61 (2), 103–113, 1989.
 35. Chen C. C. and Chen D. C., Multi-Resolutional Gabor Filter in Texture Analysis, *Pattern Recognition Letters*, 17 (10), 1069–1076, Sep. 1996.
 36. Minichino J. and Howse J., *Learning OpenCV 3 Computer Vision with Python Second Edition*, Birmingham, 152, 2015.
 37. Wiley J., *Unsupervised Learning-Clustering Using K-Means*, *Python® Machine Learning*, First Edition, 221–242, 2019.
 38. Hoang N. D., Image Processing-Based Spall Object Detection Using Gabor Filter, Texture Analysis, and Adaptive Moment Estimation (ADAM) Optimized Logistic Regression Models, *Advances in Civil Engineering*, 2020, 8829715-16, 2020.
 39. Turan M. K., Sehirli E., Elen A., and Karas I. R., Identification of Column Edges of DNA Fragments by Using K-Means Clustering and Mean Algorithm on Lane Histograms of DNA Agarose Gel Electrophoresis Images, *Proc.SPIE*, 9631, Jul. 2015.
 40. Vidyullatha P. and Rajeswara Rao D., Machine Learning Techniques on Multidimensional Curve Fitting Data Based on R-Square and Chi-Square Methods, *International Journal of Electrical and Computer Engineering*, 6 (3), 974–979, Jun. 2016.
 41. Rill R. L., Beheshti A., and van Winkle D. H., DNA Electrophoresis in Agarose Gels: Effects of Field and Gel Concentration on The Exponential Dependence of Reciprocal Mobility on DNA Length, *Electrophoresis*, 23 (16), 2710–2019, 2002.
 42. M. L. Marshall, Fitting The Two-Term Mixed Exponential and Two-Parameter Lognormal Distributions to Grouped and Censored Data, *Journal of Applied Statistics*, 23 (3), 313, 1974.