STEGANALİZ YÖNTEMLERİ KULLANILARAK RESİM İÇERİSİNDEKİ SAKLI BİLGİLERİN TESPİT EDİLMESİ

DETECTING HIDDEN MESSAGES IN IMAGES USING STEGANALYSIS TECHNIQUES

FİRDES AKTAŞ

Hacettepe Üniversitesi

Lisansüstü Eğitim – Öğretim ve Sınav Yönetmeliğinin

ELEKTRİK ve ELEKTRONİK Mühendisliği Anabilim Dalı İçin Öngördüğü

YÜKSEK LİSANS TEZİ

olarak hazırlanmıştır.

2011

Fen Bilimleri Enstitüsü Müdürlüğü'ne,

Bu çalışma jürimiz tarafından **ELEKTRİK ve ELEKTRONİK MÜHENDİSLİĞİ ANABİLİM DALI'nda YÜKSEK LİSANS TEZİ** olarak kabul edilmiştir.

Başkan	:….. Prof. Dr. Abdullah ÇAVUŞOĞLU
Üye (Danışman)	:….. Yrd. Doç. Dr. Mehmet DEMİRER
Üye	:….. Yrd. Doç. Dr. Murat EFE
Üye	:….. Yrd. Doç. Dr. Derya ALTUNAY
Üye	:….. Yrd. Doç. Dr. Umut SEZEN

ONAY

Bu tez ...../...../2011 tarihinde Enstitü Yönetim Kurulunca kabul edilmiştir.

Prof. Dr. Adil DENİZLİ
Fen Bilimleri Enstitüsü Müdürü

STEGANALİZ YÖNTEMLERİ KULLANILARAK RESİM İÇERİSİNDEKİ SAKLI BİLGİLERİN TESPİT EDİLMESİ

Firdes Aktaş

ÖZ

Bu tez kapsamında teknolojinin gelişmesiyle birlikte sayıları gittikçe artan bilgi gizleme yöntemlerinden biri olan steganografinin tespit edilmesine yönelik bir uygulama geliştirilmiştir. Uygulamada yedi farklı steganaliz yöntemine yer verilmiştir.

Bunlardan dört tanesi LSB steganografinin tespit edilmesine yönelik olarak geliştirilen RS, Yakın Renk Çifti (RQP), Değişken Eşik Değerli Yakın Renk Çifti ve Resim Düzgünlüğü Steganaliz yöntemleridir. Diğer üç yöntem olarak kör steganaliz olarak adlandırılan İkili Benzerlik Ölçütleri, Yüksek Dereceli İstatistikler ve Koşu Uzunluğu Steganaliz yöntemleri kullanılmıştır.

Kör steganaliz yöntemlerinin eğitilmesinde DVM olarak açık kaynak kodlu “libsvm” kullanılmıştır. Farklı kaynaklardan elde edilen farklı tipteki resimler kullanılarak çeşitli steganografi programları vasıtasıyla eğitim için bir veritabanı oluşturulmuştur.

Yazılım dili olarak C# seçilmiş ve geliştirme ortamı olarak Visual Studio .NET kullanılmıştır. Yazılım aracılığı ile hem genel hem de steganaliz yöntemi bazlı resim içerinde gizli bilginin tespit edilmesine yönelik analizin yapılması sağlanmıştır.

Anahtar Kelimeler: Steganografi, Steganaliz, DVM, Visual Studio .NET, C#.

Danışman: Yrd. Doç. Dr. Mehmet DEMİRER, Hacettepe Üniversitesi, Elektrik ve Elektronik Mühendisliği Bölümü

DETECTING HIDDEN MESSAGES IN IMAGES USING STEGANALYSIS TECHNIQUES

Firdes Aktaş

ABSTRACT

In this thesis, an aplication has been developed for detecting steganography in images. Seven different steganalysis methods have been used in this application.

The first four in detecting LSB steganography: RS, RQP, RQP with Variable Threshold and Image Smoothness Steganalysis; and the others called blind steganalysis techniques: Binary Similarity Measure, High Order Statistics and Run Length Steganalysis have been used.

In training of blind steganalysis methods, open source tool libsvm has been used as Support Vector Machine. A database has been constructed by using different steganography programs and different types of images.

C# as a programming language and Visual Studio .NET as developing environment have been selected. In this application, a general steganalysis method to detect hidden messages has been implemented.

Keywords: Steganography, Steganalysis, SVM, Visual Studio .NET, C#.

Advisor: Asst. Prof. Mehmet DEMİRER, Hacettepe University, Department of Electrical and Electronics Engineering

TEŞEKKÜR

Yazar, bu çalışmanın gerçekleşmesinde katkılarından dolayı, aşağıda adı geçen kişi ve kuruluşlara içtenlikle teşekkür eder.

Sayın Yrd. Doç. Dr. Mehmet DEMİRER, tez çalışmasının gerçekleştirilmesi için gerekli ortamı hazırlamıştır ve karşılaşılan güçlüklerin aşılmasında yol gösterici olmuştur.

Yazar, ailesi ve arkadaşlarına, bu çalışmanın gerçekleşmesi esnasında gösterdikleri maddi, manevi destekler için müteşekkirdir.

İÇİNDEKİLER DİZİNİ

Sayfa

1. GİRİŞ 1

2. STEGANOGRAFİ ... 3

2.1. Steganografi Tarihçesi ... 3

2.2. Günümüzde Steganografi ... 5

2.3. Bir Steganografi Sisteminin Yapısı ... 5

2.4. Sayısal Bir Resmin Yapısı ... 6

2.5. Görüntü Steganografide Kullanılan Yaklaşımlar ... 7

2.5.1. Resim Uzayı Tabanlı ... 7

2.5.2. Maskeleme ve Filtreleme ... 10

2.5.3. Algoritma ve Dönüşümler ... 10

2.6. Steganografi Sistemlerinin Kriterleri ... 11

2.6.1. Taşıyıcı Resimdeki Değişim ... 11

2.6.2. Kapasite ... 11

2.6.3. Dayanıklılık ... 12

3. STEGANALİZ ... 13

3.1. Steganalizde Kullanılan Yaklaşımlar ... 15

3.1.1. Öğrenmeye Dayalı Steganaliz ... 16

3.1.2. Kör Tanımlamaya Dayalı Steganaliz ... 18

3.1.3. Parametrik İstatistiksel Tespite Dayalı Steganaliz ... 19

3.1.4. Hibrit Teknikler ... 20

4. STEGANALİZDE KULLANILAN BAZI SALDIRILAR ... 21

4.1. RS Steganaliz ... 21

4.2. RQP Steganaliz (Yakın Renk Çifti) ... 26

4.3. Değişken Eşik Değerli Yakın Renk Çifti Analizi ... 29

4.4. Resim Düzgünlüğüne Bağlı Steganaliz ... 30

4.4.1. Resim Düzgünlüğüne Dayanarak LSB Steganografinin Tespit Edilmesi ... 32

4.4.2. Stego-Resimlerin Resim Düzgünlüğünün Analizi ... 32

4.4.3. LSB Düzlemi Çevrildikten Sonra Stego-Resmin Resim Düzgünlüğünün Analizi ... 33

4.4.4. Veri Saklama Oranının Tahmin Edilmesi ... 34

4.5. Koşu Uzunluğuna Bağlı Steganaliz ... 35

4.5.1. Koşu Uzunluğu Nedir? ... 35
4.5.2. Steganaliz İçin Koşu-Uzunluğu Analizi ... 36
4.5.3. Parametize Edilmiş Koşu-Uzunluğu Gösterimleri ... 38
4.5.3.1. Nicemleme ile Koşu-Uzunluğu Gösterimi ... 38
4.5.3.2. Koşu Uzunluğu Fark Gösterimi ... 39
4.5.4. Steganaliz için Özellikler ... 40
4.6. İkili Benzerlik Ölçütlerine Dayalı Steganaliz ... 42
4.6.1. İkili Resimlerdeki Benzerlik Ölçütleri ... 42
4.7. Yüksek Dereceli İstatistiklere Dayalı Steganaliz ... 48
5. STEGANALİZ UYGULAMASI VE DEĞERLENDİRME ... 53
6. SONUÇ VE TARTIŞMA ... 77
KAYNAKLAR ... 79
EKLER ... 82
EK 1 ... 82
EK 2 ... 86
EK 3 ... 89
ÖZGEÇMİŞ ... 92

ŞEKİLLER DİZİNİ

Sayfa

Şekil 2.1. Bir Steganografi Sistemi 6

Şekil 2.2. Sayısal resmin temel yapısı 6

Şekil 3.1. Kör Tanımlamaya Dayalı Steganaliz..... 18

Şekil 4.1. RS diyagramı 23

Şekil 4.2. (a) Standart resim "Lena" (b) Fark değişkeni $\boldsymbol{XI}$'nin dağılımı 31

Şekil 4.3. Koşu-uzunluğu hesaplama yön gösterimleri (a) yatay $\boldsymbol{0^o}$, (b) dikey $\boldsymbol{90^o}$, (c) diyagonal $\boldsymbol{45^o}$, (d) diyagonal $\boldsymbol{135^o}$. 36

Şekil 4.4. Lena resminin normal ve steganografi uygulandıktan sonraki koşu-uzunluğu histogram örneği 38

Şekil 4.5. Orijinal ve steganografi uygulanmış Lena resmi için (a) $\boldsymbol{Q = 4}$ nicemleme koşu uzunluğu histogramı, (b) koşu uzunluğu fark histogramı 39

Şekil 4.6. Lena için koşu uzunluğu resim gösterimleri (a) normal Lena $\boldsymbol{Q = 4}$, (b) stego Lena $\boldsymbol{Q = 4}$, (c) normal Lena $\boldsymbol{\varepsilon = 2}$, (d) stego Lena $\boldsymbol{\varepsilon = 2}$..... 41

Şekil 4.7. Yelkenli resmi ve bu resmin kırmızı renk kanalına ait bit düzlem resimleri (1-8) 43

Şekil 4.8. Piksel komşuluğu 44

Şekil 4.9. (a) Ojala değer hesaplamasındaki ağırlıklı komşuluk. (b) Örnek olarak, eğer E, N, NE bitleri 1 ve diğer bitler 0 ise değer S=2+4+8=14'tür. 46

Şekil 4.10. Frekans uzayının çok ölçekli ve yönlü ayrıştırması. Üstten alta 0,1, 2 seviyeleri ve soldan sağa alçak geçiren, yatay, dikey ve diyagonal alt bantları. 49

Şekil 4.11. "disc" resminin üç yönlü ve üç seviyeli alt bant katsayılarının mutlak değeri. Artık olan alçak geçiren alt bant sol üst köşede verilmiştir. 50

Şekil 5.1. RS Steganaliz yöntemine ait akış diyagramı 55

Şekil 5.2. Yakın renk çifti ve değişken eşik değerli yakın renk çifti yöntemleri için ortak akış diyagramı 58

Şekil 5.3. Resim düzgünlüğüne bağlı steganaliz yöntemi için akış diyagramı..... 62

Şekil 5.4. Eğitim veritabanında kullanılan resimlerden bazıları..... 64

Şekil 5.5. Bit Düzlem resimlerini gösteren ekran görüntüsü (Yat.jpg resminin kırmızı renk kanalına ait bit düzlem resimleri, 1'den 8'e)..... 66

Şekil 5.6. İkili benzerlik ölçütüne dayalı steganaliz yöntemi için akış diyagramı 67

Şekil 5.7. Yüksek dereceli istatistiklere dayalı steganaliz yöntemi için akış diyagramı 70

Şekil 5.8. Resmin 4 seviyeli dalgacık dönüşümü katsayılar resmi 71

Şekil 5.9. Koşu uzunluğu histogramı arayüz formu ekran görüntüsü (Yat.jpg resminin kırmızı renk kanalındaki $\mathbf{90}\boldsymbol{o}$'lik fark koşu histogramı) 71
Şekil 5.10. Koşu uzunluğuna dayalı steganaliz yöntemine ait akış diyagramı..... 73
Şekil 5.11. Kör steganaliz yöntemleri için veritabanı eğitme arayüzü 74
Şekil 5.12. Resmin renk kanalı ve histogramını gösteren arayüze ait ekran görüntüsü 75

ÇİZELGELER DİZİNİ

Sayfa

Çizelge 3.1. Steganaliz Saldırı Çeşitleri 15
Çizelge 4.1. İkili benzerlik ölçütleri 46
Çizelge 5.1. Manmade.bmp resmi için RS steganaliz sonuçları 56
Çizelge 5.2. Foliage.bmp resmi için RS steganaliz sonuçları 56
Çizelge 5.3. Animal.bmp resmi için RS steganaliz sonuçları 56
Çizelge 5.4. Manmade.bmp resmi için yakın renk çifti analizi sonuçları 59
Çizelge 5.5. Foliage.bmp resmi için yakın renk çifti analizi sonuçları 59
Çizelge 5.6. Animal.bmp resmi için yakın renk çifti analizi sonuçları 60
Çizelge 5.7. Manmade.bmp resmi için Resim Düzgünlüğü mesaj oranı sonuçları 61
Çizelge 5.8. Foliage.bmp resmi için Resim Düzgünlüğü mesaj oranı sonuçları 61
Çizelge 5.9. Animal.bmp resmi için Resim Düzgünlüğü mesaj oranı sonuçları 61
Çizelge 5.10. İkili benzerlik ölçütüne dayalı steganaliz için sonuçlar 68
Çizelge 5.11. Yüksek dereceli istatistiklere dayalı steganaliz yöntemi için test sonuçları .. 69
Çizelge 5.12. Koşu Uzunluğuna Dayalı Steganaliz yöntemi için test sonuçları 72

SİMGELER ve KISALTMALAR DİZİNİ

AU	Ses Birimi (Audio Units)
DCT	Ayrık Kosinüs Dönüşümü (Discrete Cosine Transform)
DFT	Ayrık Fourier Dönüşümü (Discrete Fourier Transform)
DOS	Disk İşletim Sistemi (Disk Operating System)
DVM	Destek Vektör Makinesi
DWT	Ayrık Dalgacık Dönüşümü (Discrete Wavelet Transform)
EÖB	En Önemsiz Bit
GIF	Grafik Değiştirme Biçimi (Graphics Interchange Format)
GIMP	GNU Resim İşleme Programı (The GNU Image Manipulation Program)
HTML	Zengin Metin İşaretleme Dili (HyperText Markup Language)
IEC	Uluslararası Elektroteknik Komisyonu (International Electrotechnical Commission)
IP	İnternet Protokolü (Internet Protocol)
ISO	Uluslararası Standart Kuruluşu (International Organization for Standardization)
JPG/JPEG	Birleşik Fotoğraf Uzmanları Grubu (Joint Photographic Experts Group)
LSB	En Önemsiz Bit (Least Significant Bit)
MSE	Ortalama Karesel Hata (Mean Square Error)
PCX	Paintbrush Dosya Biçimi (PC Paintbrush File Format)
PDF	Adobe Taşınabilir Doküman Biçimi (Adobe Portable Document Format)
PNG	Taşınabilir Ağ Grafikleri (Portable Network Graphics)
POVs	Değer Çiftleri (Pairs of Values)
PSNR	Tepe Sinyal Gürültü Oranı (Peak Signal to Noise Ratio)
QMF	Dörtlü Ayna Filtreleri (Quadrature Mirror Filters)

RGB	Kırmızı-Yeşil-Mavi (Red-Green-Blue)
RLH	Koşu Uzunluğu Histogramı (Run Length Histogram)
RMSE	Ortalama Karekök Hata (Root Mean Square Error)
RQP	Çiğ Çabuk Çiftleri (Raw Quick Pairs)
RS	Düzenli-Tekil (Regular-Singular)
SVM	Destek Vektör Makinesi (Support Vector Machine)
TIFF	İşaretli Resim Dosya Biçimi (Tagged Image File Format)
WAV	Dalgaşekli Ses Dosya Biçimi (Waveform Audio File Format)

1. GİRİŞ

Gelişen teknolojiyle birlikte, bilgi gizleme ve iletişimin güvenliğinin sağlanması önem kazanmaktadır. Özellikle son yıllarda internet teknolojisinin hızlı ilerlemesi ile veri paylaşımı artmış ve bu paylaşılan verinin hem paylaşım sırasında hem de saklama aşamasında güvenliğinin sağlanması hususu öne çıkmıştır. Bu amaçla haberleşen iki kişi arasındaki iletişimi güvenli kılacak yöntemler geliştirilmiştir.

Haberleşmenin güvenli kılınması amaçlı kullanılan yöntemlerden biri şifrelemedir. Şifrelemede korunmak istenen bilgi çeşitli algoritmalar vasıtasıyla anlaşılmaz hale getirilir ve elde edilen bu karışık veri gönderilir. Şifreleme işlemi haberleşmenin içeriğinin güvenliğini sağlar. Ancak bazı durumlarda haberleşen iki tarafın bu olayı üçüncü şahıslardan okunmaya, dinlemeye ya da değiştirmeye karşı koruması gerekebilir. Ancak şifrelerin belli atak yöntemleri ile kırılabilmesi, şifrelemenin tek başına bilgi güvenliğini sağlamak açısından yetersiz kaldığını göstermiştir.

Bazı durumlarda verinin sadece karmaşık hale getirilmesinin yanı sıra arada gizli bir iletişimin varlığının saklanmasına ihtiyaç duyulmaktadır. Bunun nedeni şifreli bir iletişimin ortaya çıkması bu iletişimi saldırıların hedefi haline getirmektedir. İletişimin fark edilmesi durumunda araya giren kişi iletişimi değiştirebilir, kesebilir ya da engelleyebilir.

İletişimin gizli ve güvenli kılınması gerektiği durumlara bir örnek devletlerin diğer ülkelerde bulunan elemanlarının haberleşmesinin sadece güvenli değil aynı zamanda da gizli olması gereğidir. Bu örneğe ilave olarak bilgi gizleme tekniklerinin kullanıldığı bazı alanlar şöyle sıralanabilir:

i. Telif hakları bu uygulamaların başında gelmektedir. Teknolojinin gelişmesiyle video ve ses dosyaları kolaylıkla paylaşılabilmekte ve izinsiz şekilde çoğaltılabilmektedir. Bu da izinsiz kopyalamayı sağlamakta ve film, müzik ya da kitap sektöründeki firmaların zarar görmesine neden olmaktadır. Bu alanda damgalama ve parmak izi

uygulamaları, telif haklarını korumak ve olası izinsiz kopyalamaların takip edilmesini sağlamak amaçlı kullanılmaktadır.

ii. Kimi suç unsuru teşkil eden yapılanmalar dikkat çekmeyen haberleşme için bilgi saklama yöntemlerine yönelmişlerdir.

iii. Devletler kendi elemanları ile iletişimlerini sağlamak amaçlı kullanmaktadır.

iv. Bazı ülkeler vatandaşlarının şifreleme kullanmasına izin vermemekte ya da kısıtlama koymaktadır. Kişisel hayatın gizliliğine müdahale olarak gören kimi vatandaşlar bilgi gizleme yöntemlerine yönelmişlerdir.

Yukarıda ifade edilen konular kapsamında daha güvenli bir iletişim mekanizmasına ihtiyaç duyulmuştur ve haberleşmeyi üçüncü şahıslara görünmez kılan ve bilgiyi gizleyen bir haberleşme sistemi geliştirilmiştir. *Steganografi* belirtilen bu ihtiyaçları karşılayan, hem verinin içeriğinin güvenliğini hem de arada gizli bir haberleşmenin yapıldığının gizlenmesini gerçekleştiren bir daldır. Geçmiş zamanlarda görünmez mürekkeple mektuplar yazılarak gerçekleştirilen steganografi teknolojinin ilerlemesiyle dilin özelliklerini kullanmaya başlamıştır.

Bu tez kapsamında resim üzerinde uygulanan steganografi yöntemlerinin tespitine yönelik steganaliz yöntemleri incelenmiş ve bu yöntemleri kapsayan bir uygulama geliştirilmiştir. Geliştirilen bu uygulamada sadece belirli bir steganografi yöntemine yönelik tespit yerine, genel ve yeni tekniklere adapte edilebilecek bir uygulamanın geliştirilmesi amaçlanmıştır.

İkinci bölümde steganografi konusu genel olarak incelenmiştir. Steganografinin tarihçesi, geçmişte ve günümüzde yapılan uygulamaları, görüntü steganografide kullanılan yaklaşımlar açıklanmıştır. Üçüncü bölümde steganaliz ve steganalizde kullanılan yaklaşımlar incelenmiş, bu yaklaşımların avantaj ve dezavantajları açıklanmıştır. Dördüncü bölümde steganalizde bilginin varlığını tespit etmek için kullanılan bazı yöntemlere yer verilmiştir. Beşinci bölümde incelenen steganaliz yöntemlerini kapsayan bir uygulama geliştirilmiş ve elde edilen sonuçlar sunulmuştur.

2. STEGANOGRAFİ

Steganografi eski bir bilgi gizleme sanatıdır. Kelime kökeni Yunancadan gelmektedir ve tam olarak anlamı "*örtülmüş yazı*" (*covered writing*) demektir [25].

Steganografinin amacı gizli mesaj ya da bilginin varlığını saklamaktır. Taşınmak istenen mesaj masum görünüşlü başka bir ortamda saklanarak, üçüncü şahısların iletilen mesajın varlığından haberdar olması engellenir. Bu yaklaşımla ses, sayısal resim, video görüntüleri üzerine veri saklanabilmektedir. Görüntü dosyaları içerisine saklanacak olan veri metin dosyası olabileceği gibi, herhangi bir görüntü içerisine gizlenmiş başka bir görüntü dosyası da olabilir. Yine aynı şekilde bir ses dosyasının içine bir metin dosyası da saklanabilmektedir.

Steganografi şifrelemeye yakın bir kavram olmasına rağmen şifrelemeden amaç olarak faklıdır. Şifreleme mesajın içeriğinin korunması ile ilgilenirken steganografi mesajın varlığının gizlenmesi ile ilgilenmektedir. Dolayısıyla steganografi bir şifreleme yöntemi değil şifrelemeyi tamamlayıcı bir öğedir.

11 Eylül 2001 tarihinde Amerika Birleşik Devletleri'nde meydana gelen olayların ardından çeşitli haber sitelerinde teröristlerin iletişimlerini steganografi yöntemleri aracılığıyla gerçekleştirdikleri hakkında haberler çıkmıştır [3]. İletilecek olan mesajın internet üzerinde paylaşım sitelerinde bulunan resimlere saklandığı ve bu şekilde teröristlerin iletişim kurdukları söylenmiştir. Ayrıca E-Buy, Amazon ve pornografi içerikli bazı sitelerde bulunan resimler içerisinde de gizli bilgi bulunduğuna dair haberler yayınlanmıştır. Bu gelişme üzerine bir grup araştırmacı internet üzerinden servis sunan çeşitli siteler üzerinde bulunan iki milyondan fazla resim üzerinde incelemeler yapmış ancak inceleme sonucunda şüpheli bir bulguya rastlanmamıştır [3].

2.1. Steganografi Tarihçesi

Steganografinin tarihi milattan önce 440'lı yıllara kadar dayanmaktadır. Herodotus tarafından yazılan "Histories"de steganografi ile ilgili iki örnek verilmiştir [25]. Birincisi, Demeratus tarafından yapılacak olan istilayı haber vermek için tahta bir plakanın üzerine yazılan mesajın daha sonra üzerinin mum ile kaplanmasıdır. İlk görünüşte tablet boş gözükmekte ama alıcı üzerindeki mumu çıkarttıktan sonra

mesaj elde edilmektedir. İkinci örnek ise Histiaeus'a aittir. Bir kölenin saçı kazıtılır ve kafasının üzerine dövme olarak mesaj yazılır. Ardından kölenin saçları uzadıktan sonra mesaj görünmez olur ve haberci olarak köle yola çıkartılır. Saçın tekrar kazınması ile mesaj okunur hale gelir.

Teknoloji ilerledikçe steganografi teknikleri de ilerlemiştir. Önce süt, meyve suyu ya da ürin (idrar) gibi organik materyaller kullanılarak kâğıt üzerine yazılar yazılmış ve kuruduktan sonra görünmez olan mesaj iletilmiştir. Isı işlemi gibi çeşitli teknikler, bu mesaj içeren boş görünümlü kâğıda uygulandığında gizlenen mesaj tekrardan ortaya çıkarak okunur hale getirilmiştir.

Steganografi alanında diğer ilginç bir örnek ise 1600'lü yıllarda müzik notaları aracılığıyla yapılan saklama yöntemidir. Bu uygulamada her nota alfabede ayrı bir harfe karşılık gelmekte ve kullanılan bir anahtar aracılığıyla masum görünüşlü bir müzik parçasından gizli mesaj çıkartılmaktadır.

Fotoğrafçılığın ilerlemesiyle mikrofilmler yapılmaya başlanmış ve kıyafet ya da bir eşya üzerine kolaylıkla saklanabilen bu küçük filmler ile mesajlar taşınmıştır. Mikrofilm ile mesaj taşıma 1870 -1871 Franco-Prussian savaşı sırasında oldukça popüler olan bir yöntemdir. 1900'lü yıllarda ise mikrodot teknolojisinin geliştirilmesiyle harita gibi önem taşıyan dokümanlar fotoğrafik yöntemler ile nokta boyutlarına küçültülerek dikkat çekmeyen mektup, zarf gibi materyallerin üzerinde taşınması sağlanmıştır.

Yukarıda belirtilen bu örnekler dışında aşağıdaki örnekte İkinci Dünya Savaşı sırasında Alman ajanlarının birbirleriyle haberleşmek için kullandıkları bir steganografi örneği bulunmaktadır [29].

"Apparently neutrals protest is thoroughly discounted and ignored. Isman hard hit. Blockade issue affects pretex for embargo on by-products, ejecting suets and vegetable oils."

Yukarıda verilen paragrafta her kelimenin ikinci harfleri yan yana getirildiğinde "Pershing sails from NY June 1." mesajı ortaya çıkmaktadır.

2.2. Günümüzde Steganografi

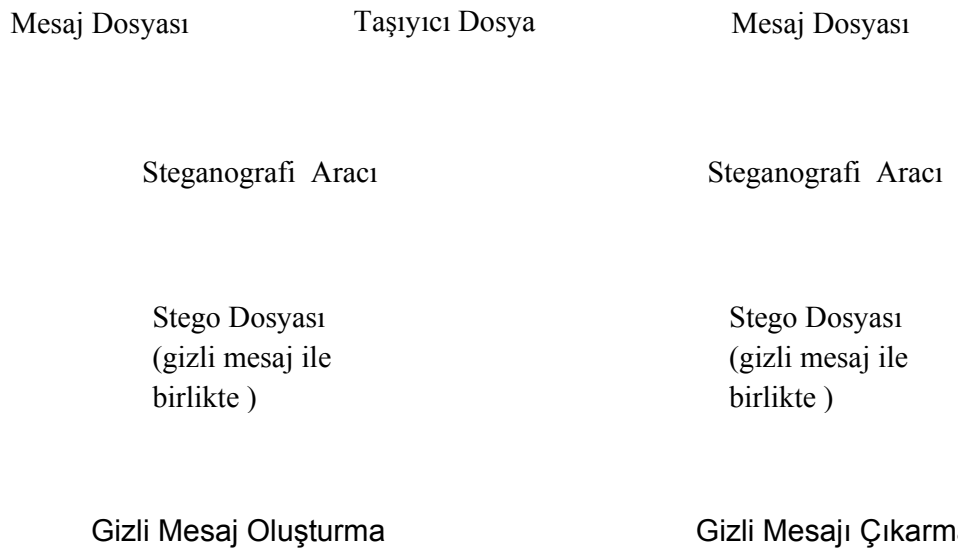
Gelişen teknolojiyle birlikte artık dijital nesneler üzerinde steganografi uygulamaları yapılmakta ve verilerimizi korumak amacıyla son yıllarda sıklıkla kullanılmaktadır. Gizli veri ilk önceleri yapılanlara benzer olarak yine masum görünüşlü bir dosyanın içerisine saklanmaktadır. Bunlardan en ilgi çekici ve çok kullanılanları dosya yapıları itibariyle daha fazla bilgi taşıyabildiklerinden dolayı resim, ses ve video dosyalarıdır. Benzer bir şekilde düz metin, html, exe gibi dijital ortamda bulunan dosyalar içerisine çeşitli yöntemler ile bilgi saklanabilmektedir [32]. Ayrıca sabit disklerde dosyalar için ayrılmış alanlarda ortaya çıkan artık kısımlar (slack space) ve IP paketlerinin ileride kullanılmak üzere ayrılmış bölümleri de günümüz teknolojisinde veri gizlemek için kullanılmaktadır.

Bilgisayar steganografisi iki temel prensip üzerine kurulmuştur. Bunlardan ilki sayısal hale getirilmiş resim veya ses dosyalarının, diğer türlerden farklı olarak, sahip oldukları fonksiyonlarını yitirmeden değiştirilebilmeleri ilkesidir. İkincisi ise, insanın, renk veya ses kalitesinde meydana gelen ufak değişiklikleri ayırt edememesine dayanır. Bunun mantığı da lüzumsuz bilgiler taşıyan nesnelerin içindeki bilgileri, başka bilgi parçacıklarıyla yer değiştirmektir.

2.3. Bir Steganografi Sisteminin Yapısı

Gizli bilgiyi bir resme saklama işleminde iki dosya söz konusudur [30]. Kapak resim, örtü verisi ya da taşıyıcı olarak adlandırılan ilk dosya, gizli bilgiyi saklayacak olan resim dosyasıdır. İkinci dosya ise gizlenecek bilgi olan mesajı oluşturacak dosyadır. Bu mesaj dosyası da stego olarak isimlendirilmektedir [26]. Mesaj; açık metin (plain text), şifreli metin (chipher text), herhangi bir dosya veya bit dizisi olabilir.

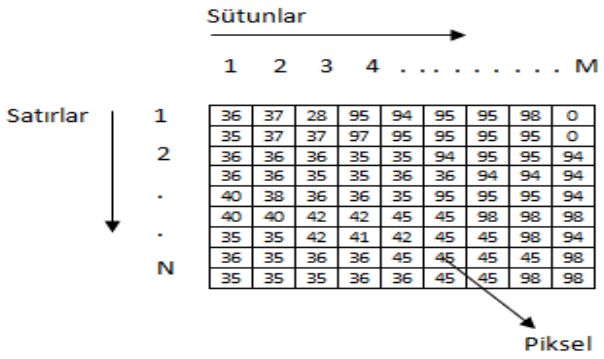
Bilgi saklama işleminde yerleştirilen verinin bir kayıp olmadan stego dosyasından çıkartılması önemlidir. Şekil 2.1'de bir steganografi sisteminin şematik gösterimi verilmiştir [23].



Şekil 2.1. Bir Steganografi Sistemi

2.4. Sayısal Bir Resmin Yapısı

Dijital resim satır ve sütunları x ve y ile temsil edilen N satır ve M sütunluk bir dizi ile gösterilmektedir. Bir resim dizisinde bulunan elemanlara piksel adı verilir. Bir resmin satır ve sütun olarak gösterimi Şekil 2.2'de verilmiştir. Resim pikselleri 0-255 arasında değerler alabilmektedir. Eğer pikseller sadece 0 ve 1 değerlerinden oluşuyor ise bu resimlere ikili (binary) resim denir [12]. Böyle bir resimde 0 siyah, 1 ise beyaz renk ile temsil edilir.



Şekil 2.2. Sayısal resmin temel yapısı

Sayısal resim dosyaları genellikle 8 ya da 24 bitlik; gri-seviye görüntüler ise 1-2-4-6 ya da 8 bitlik piksellerden oluşurlar. Bunlar dışında 16 ve 32 bitlik piksel değerlerine sahip resimler de bulunmaktadır. 32 bitlik resimlerde 8 bitlik kısımda resim hakkında farklı bilgiler yer almakta, resmin içerik bilgisi yine 24 bitlik kısımda tutulmaktadır.

24 bit resimlerde bir piksel başına 3 bayt kullanılmaktadır. Her pikselin rengi; Kırmızı (Red), Yeşil (Green), Mavi (Blue) olmak üzere üç ana renkten elde edilmektedir. Buna pikselin RGB değeri denmektedir. Her ana renk 8 bit ile temsil edilmektedir.

Resim dosyaları kayıplı ve kayıpsız olmak üzere iki formatta sayısal ortamlarda saklanmaktadır. Ek 1'de kayıpsız (bmp, gif, tiff) ve kayıplı (jpg) resim formatlarına ait kısa açıklamalar verilmiştir.

2.5. Görüntü Steganografide Kullanılan Yaklaşımlar

Görüntü dosyaları üzerinde bilgi gizlemek için geliştirilen steganografik yöntemler 3 başlık altında sınıflandırılabilir.

i. Resim Uzayı Tabanlı – En Önemsiz Bit
    a. EÖB Yer Değiştirme
    b. EÖB Eşleştirme
    c. EÖB Modül Fonksiyon
ii. Maskeleme ve filtreleme
iii. Algoritmalar ve dönüşümler

2.5.1. Resim Uzayı Tabanlı

Resim piksellerinin en önemsiz bitlerinin değişimine bağlı olarak bilgi saklanmasına dayanan yöntemlerdir. En yaygın kullanılan bilgi gizleme yöntemidir. Taşıyıcı ortamın en az önemli bitlerini insan gözünün fark edemeyeceği şekilde gizli veriyi saklamak amacıyla değiştirmeyi temel alır. Bilgi gizlemek için resmin renk değerleri ya da renk paleti değiştirilebilir.

Resim üzerinde yapılan değişim insan gözüyle algılanamaz. Eklenen mesaj resme gürültü olarak eklenir. Bu yöntem ile resim içerisine yüksek miktarda veri

yerleştirmek mümkündür ancak resim üzerinde yapılacak kesme, yapıştırma, döndürme sıkıştırma gibi değişimlere karşı yöntem hassasiyet göstermektedir.

EÖB yer değiştirme, EÖB eşleştirme ve EÖB modül fonksiyonu olmak üzere üç alt başlığa ayrılabilir.

2.5.1.1. EÖB Yer Değiştirme

En basit yaklaşımdır. Resmin EÖB'lerinin yer değiştirilmesi ile saklanacak olan veri gizlenmektedir. Yöntem, yaklaşık olarak saklanmak istenen mesaj verisinin yarısı kadar taşıyıcı resmin piksel değerinin değiştirilmesine sebep olmaktadır.

24 bitlik renkli resim kullandığımızı düşünelim. Her baytın son bitine bilgi saklanırsa piksel başına 3 bit saklanabilir. Aşağıda buna dair bir örnek verilmiştir.

$$10010101\ 00001101\ 11001001 \qquad (149, 13, 201)$$

$$10010110\ 00001111\ 11001010 \qquad (150, 15, 202)$$

$$10011111\ 00010000\ 11001011 \qquad (159, 16, 234)$$

Yukarıda verilen bu iki piksel değerine "X" verisi yani ikili karşılığı "01010111" saklanmak istenirse piksel değerleri,

$$10010100\ 00001101\ 1100100 \qquad (148, 13, 200)$$

$$10010111\ 00001110\ 11001011 \qquad (151, 14, 203)$$

$$10011111\ 00010001\ 11001011 \qquad (159, 17, 202)$$

haline gelir.

2.5.1.2. EÖB Eşleştirme

Bu yöntem, saklanmak istenen bilgiye göre resim içerisindeki piksellerin azaltılması ya da artırılmasına dayanır [7]. Eğer resim pikseliyle gizlenecek olan bilgi uyuşursa herhangi bir değişiklik yapılmaz. Eğer uyuşma olmaz ise piksel değeri azaltılarak ya da artırılarak veri saklanır. Herhangi bir pikselin değerinin

artırılması ya da azaltılması söz konusu olduğu için yer değiştirme yöntemindeki monotonluktan çıkılmış olunur. Bu da gizlenen bilginin tespitini zorlaştırmaktadır. [24][37]'de bu yöntem daha da dayanıklı hale getirilmiştir. Özel bir fonksiyon kullanılarak her bir piksel çiftinin değerinin artırılması sağlanarak her resim piksel çiftine iki bit veri gizlenebilmektedir.

En önemsiz bit eşleştirme steganografi ±1 yerleştirme olarak da bilinmektedir [11] ve EÖB yerleştirmenin daha karmaşık halidir. Matematiksel olarak Eş. 2.1.'deki gibi ifade edilebilir;

$$p_s = \begin{cases} p_c + 1, & \text{if } b \neq LSB(p_c) \text{ ve } (\kappa > 0 \; ya \; da \; p_c = 0) \\ p_c - 1, & \text{if } b \neq LSB(p_c) \text{ ve } (\kappa < 0 \; ya \; da \; p_c = 255) \\ p_c, & \text{if } b = LSB(p_c) \end{cases} \quad (2.1)$$

burada p_s bir stego resimdeki, p_c ise taşıyıcı resimdeki 0'dan 255'e kadar olan değerlerdeki piksel değerlerini simgeler. *b* gizlenecek mesaj biti ve $\kappa \{-1, +1\}$ aralığında tekdüze dağılıma sahip bir rastsal değişkendir. Saklanacak olan mesaj uzunluğuna bağlı olarak bütün EÖB düzlemi ya da sadece belli bir bölümü taşıyıcı olarak kullanılabilir. EÖB düzleminin boyutu ve mesajın uzunluğu arasındaki oran yerleştirme oranı olarak adlandırılır ve ρ ile simgelenir.

EÖB yer değiştirme göreceli olarak tespit edilmesi daha kolay olarak bilinmektedir. Bu yöntem ile her pikselin değişim ihtimali yüzde elliden daha azdır. Ancak yöntemin en büyük dezavantajı, saklanmak istenen verinin resim piksel sayısı kadar olmak zorunda olmasıdır.

2.5.1.3. EÖB Modül Fonksiyon

Saklanmak istenen verinin taşıyıcı resim içerisine bit-modül fonksiyonu kullanılarak yerleştirilmesini sağlar [7]. Yöntemde Eş. 2.2.'de verilen bit-modül fonksiyonu kullanılarak bütün olası değerler arasında, gerçek piksel değeri olan *y* değerine en yakın $\hat{y}$ değeri bulunmaya çalışılır. Burada *k* saklanmak istenen mesajın bit sayısını, *x* mesajın değerini ve $\hat{y}$ ise mesaj saklandıktan sonra pikselin alacağı değeri simgelemektedir.

$$x = \hat{y} \, mod 2^k \quad (2.2)$$

Bu yöntem ile saklanmak istenen her iki bitlik veri için taşıyıcı resimde en fazla 1 bitlik bir değişim olmaktadır. Bu da taşıyıcı resimde değişikliği azaltmakta ve tespit edilme oranını düşürmektedir.

2.5.2. Maskeleme ve Filtreleme

Bu tekniklerde genellikle 24 bitlik ve gri seviyeli resimler kullanılmaktadır. Burada mesaj kâğıt ortamlarda uygulanan damgalama uygulamalarına benzer şekilde saklanmaktadır. Steganografide gizli mesaj haberleşme objesidir. Damgalama uygulamalarında ise haberleşme objesi taşıyıcıdır. Gizlenecek olan bilgi sadece resmin gürültü taşıyabilen bölgelerine değil resimle bütünleşmiş şekilde gömülür. Genellikle ticari amaçlar için kullanılmaktadır. Televizyon kanallarının logoları, resim içerisindeki sayısal imzalar, telif hakları için resim içerisine yerleştirilen bazı görünür ya da görünmeyen yazılar bu tip uygulamalara örnek olarak verilebilir.

2.5.3. Algoritma ve Dönüşümler

Bu yöntemde saklanmak istenen bilgi, taşıyıcı resim başka bir uzaya dönüştürüldükten sonra resmin belirli alanlarına yerleştirilir. Dönüşüm tekniklerinde genellikle Ayrık Kosinüs Dönüşümü (DCT), Ayrık Fourier Dönüşümü (DFT) ve Ayrık Dalgacık Dönüşümü (DWT) kullanılmaktadır.

DCT ve DFT uygulamalarında resimler 8x8'lik bloklara ayrılır ve her bir bloğa bu dönüşümler uygulanır. Dönüşümler sonucunda elde edilen katsayılara veri gizleme işlemi gerçekleştirilir.

DWT uygulamalarında ise taşıyıcı resim yüksek ve alçak geçiren filtrelerden geçirilerek 4 parçaya ayrılır. Bu parçalar dikey, yatay, diyagonal ve alt bantlardır. Gizlenecek olan veri bu parçalardan resmin en az etkileneceği yatay, dikey ve diyagonal alt bantlarındaki katsayılara yerleştirilmektedir.

Dönüşüm tekniklerinin dezavantajı diğer yöntemlerle kıyaslandığında yüksek miktarda veri saklanamamasıdır. Avantajı ise resim üzerinde yapılan değişikliklere karşı diğer yöntemlere göre daha dayanıklıdır.

2.6. Steganografi Sistemlerinin Kriterleri

Bir steganografik sistemin güvenilirliği çeşitli açılardan değerlendirilmektedir. Bunlar bilgi gizlemenin taşıyıcı dosyasını ne kadar değiştirdiği, bilgi saklama kapasitesinin ve dayanıklılığının ne kadar olduğudur [2]. Aşağıda bu kavramlar kısaca açıklanmıştır.

2.6.1. Taşıyıcı Resimdeki Değişim

Steganografi uygulamalarında taşıyıcı dosyada meydana gelen değişimin miktarı önemli bir kriterdir. Değişimin ya da resimde ne kadar bozulmanın meydana geldiğinin ortaya konmasında birçok yöntem vardır. En çok bilinen ve kullanılanları Ortalama Karesel Hata (MSE), Ortalama Karekök Hata (RMSE) ve Tepe Sinyal Gürültü Oranı (PSNR)'dır.

$$MSE = \sigma^2 = \frac{1}{N} \sum_{n=1}^{N} (x_n - y_n)^2 \tag{2.3}$$

MSE Eş. 2.3'te verilen şekilde hesaplanmaktadır. Burada x_n orijinal resmin piksel değeri, y_n steganografi işlemi gerçekleştirildikten sonraki o pikselin değeri, *N* ise resimdeki toplam piksel sayısıdır.

RMSE hesaplaması ise MSE'nin karekökü alınmış halidir [18] ve matematiksel olarak Eş. 2.4.'te gösterilmiştir.

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{N} \sum_{n=1}^{N} (x_n - y_n)^2} \tag{2.4}$$

PSNR değeri, hatanın büyüklüğünün orijinal piksel değerinin en tepe değeri ile olan ilişkisini ölçer ve Eş. 2.5.'te ifade edildiği şekilde hesaplanır. Bu eşitlikte x_{tepe} resimde bulunan en büyük piksel değeri, σ_d^2 resmin MSE değeridir.

$$PSNR(dB) = 10 \log_{10} \frac{x_{tepe}^2}{\sigma_d^2} \tag{2.5}$$

2.6.2. Kapasite

Bir steganografi sisteminde kapasite, taşıyıcı resmin içerisine ne kadar veri saklanabileceğinin ölçüsüdür. Steganografi kapasitesi, resmin biçimi, resmin piksel

sayısı ve kullanılan yöntemle doğru orantılıdır. Aynı boyuttaki resimlere farklı steganografi algoritmaları ile farklı büyüklükte bilgi saklanabilmektedir. BMP ve GIF formatındaki resimler daha yüksek kapasiteye sahipken, JPEG formatındaki dosyalar ise 8x8'lik piksel bloklarına sadece 1 bayt saklanabildiğinden dolayı daha az kapasiteye sahiptirler.

2.6.3. Dayanıklılık

Steganografi sistemlerinde dayanıklılık, mesaj saklamak için kullanılan algoritmanın bilginin varlığının belirlenebilmesi için yapılan işlemlere karşı ne kadar karşı koyabildiğinin ölçüsüdür. Steganografi algoritmalarının dayanıklılığını ölçmek zor bir işlemdir. Resimlerde genellikle görsel olarak değişiklik meydana gelmediğinden gözle ayırt edilebilmesi mümkün olmamaktadır. Bundan dolayı algoritmaları test ederken dayanıklılığı ölçmek için resmin çeşitli özelliklerinden (renk dağılımı, dalgacık dönüşümü, vb.) faydalanılarak geliştirilmiş olan steganaliz yöntemleri kullanılmaktadır.

3. STEGANALİZ

Steganaliz, bir taşıyıcı verisi içerisinde, saklanmış bir bilgi olup olmadığını bulmayı, eğer var ise bu bilgiyi elde etmeyi amaçlayan ve steganografik sistemlere karşı yapılan saldırılara denmektedir.

Steganografi sonucunda oluşturulmuş olan stego resim orijinal haliyle kıyaslandığında görsel olarak herhangi bir değişim olmadığı görülür. Görsel olarak herhangi bir belirti olmadığı için steganografiyi belirlemek için daha detaylı analiz yöntemlerine ihtiyaç duyulmaktadır. Bazı steganografi araçları üzerinde yapılan incelemelerde [12][19] Johnson ve Jajodia birkaç elektronik iz keşfetmişlerdir. Bunlardan en belirgin olanı resmin boyutunda olan değişimdir. Belirlenen imzaların çoğu taşıyıcı resmin renk paletinde değişime neden olduğudur. Bu izler ya renk paletinde azalma ya da renk paletinde artmaya neden olmaktadır.

Stego resimlere yapılan saldırılar üç sınıf altında toplanabilir.

1. Gizli mesajın varlığının tespit edilmesi,
2. Gizli mesajın elde edilmesi,
3. Gizlenmiş olan mesajın yok edilmesi.

Bir resim içerisinde eğer gizli bilgi olduğu tespit edilmiş ise bu mesajı yok etmek ya da bozmak için birkaç tane yol vardır. Bunlardan bir tanesi resim bir resim aracı ile açılıp ve en basit olarak tekrardan JPEG formatında saklanmasıdır. Bu işlem özellikle LSB yöntemi ile saklanmış bilginin bozulması için işe yarayan bir yöntemdir. Ancak dönüşüm yöntemleri kullanılarak yaratılan stego resimler için daha farklı bir yöntem kullanılması gerekmektedir. Tek bir dönüşüm bu araçlar ile yaratılan resimlerdeki bilgilerin yok edilmesi için yeterli olmasa da birkaç kere farklı formatlarda resim dönüşümlerinin yapılması bu bilgilerin kaybolmasını sağlamaktadır [21]. Ayrıca kesme, bazı bölümlerini çıkartma, bulanıklaştırma, pikseller arasındaki kontrastı azaltma/arttırma, gürültü ekleme ya da çıkarma, tekrar örnekleme, bit yoğunluğunda değişiklik yapma (gri-seviye 8-bit, 24-bit vb.), gibi resim değiştirme işlemleri uygulanması durumunda eldeki resimde bulunan gizli bilginin yok edilmesi sağlanabilmektedir.

Resim içerisinden saklanmış olan bilginin çıkartılması işlemi incelenmesi gereken çok farklı bir konudur. Resim içerisinde bilgi olduğu tespit edilmesine rağmen bu bilginin çıkartılması mümkün olmayabilir. Kullanılan algoritma bilinmiyorsa, şifreleme kullanılıp kullanılmadığı bilinmiyorsa işlem oldukça zor hale gelmektedir. Bu konu detaylı olarak incelenmesi gereken bir konudur.

Steganografi sistemlerine saldırı yaparak bilginin varlığını ortaya çıkartmaya çalışan kişiye steganalist denmektedir. Genelde steganalistin kullanılan steganografi sistemini bildiği varsayılır. Eğer steganalist kullanılan sistemi bilmiyorsa, bu onun işini zorlaştıracaktır. Steganalistin bir steganografik sisteme saldırabilmesi için sahip olması gereken bazı veriler vardır. Bu sahip olduğu verilere göre saldırı modellerinden birini seçebilir. Steganalizde kullanılan bu modeller kriptanalizde kullanılan saldırılar kullanılarak açıklanabilir. Kriptanalizde kullanılan saldırılar, sadece şifreli metin, bilinen açık metin, seçilen açık metin ve seçilen şifreli metindir. Steganalizde ise sadece stego, bilinen taşıyıcı, bilinen mesaj, seçilen stego, seçilen mesaj ve bilinen stego saldırılardır [31].

1. Sadece stego saldırısı: Analiz için sadece stego-nesnesi (Stego-object) (Görüntü dosyası) bilinmektedir. Kriptanalizdeki sadece şifreli metin atağına benzemektedir. En zor durumdur. Elde sadece içerisinde bilgi olduğu şüphelenilen dosya bulunmaktadır.
2. Bilinen taşıyıcı saldırısı: Elde resmin hem mesaj gizlenmeden önceki hali hem de mesaj gizlendikten sonraki hali bulunmaktadır.
3. Bilinen mesaj saldırısı: Analist gizlenmiş olan mesajı biliyorsa stego medyayı analiz etmek için kullanabilir. Bu atakta oldukça zordur. Sadece stego atağıyla benzer olarak düşünülebilir.
4. Seçilen stego saldırısı: Kullanılan steganografi algoritması ve stego medyasının bilindiği durumdaki ataktır.
5. Seçilen mesaj saldırısı: Bu atakta steganalist elde bulunan stego resmini analiz edebilmek için seçilen mesajlardan çeşitli steganografi algoritmaları kullanarak stego medyalar elde eder. Amaç belirli steganografi araçlarını ya da algoritmalarını işaret eden stego resimdeki örüntüyü/modeli belirlemektir.

Eğer bunları bir tablo halinde göstermek gerekirse aşağıdaki gibi bir tablo oluşturulabilir.

Çizelge 3.1. Steganaliz saldırı çeşitleri

Saldırı Tipi	Saldırı Yapan kişinin elinde bulunması gerekenler			
Saldırı	Stego-Resmi	Taşıyıcı	Gizli Mesaj	Algoritma
Sadece stego saldırısı	✓			
Bilinen taşıyıcı saldırısı		✓		
Bilinen mesaj saldırısı			✓	
Seçilmiş stego saldırısı	✓			✓
Seçilmiş mesaj saldırısı			✓	✓

Çizelge 3.1'de görülen bu saldırı tipleri değerlendirildiğinde en önemli olanı sadece stego olanıdır. Çünkü karşımıza çıkması en olası olan ve gerçek uygulamalara en yakın olan durumdur.

3.1. Steganalizde Kullanılan Yaklaşımlar

İyi bir steganaliz yöntemi geliştirmek zordur. Çünkü belli bir steganografi yöntemi için başarılı olan bir analiz başka bir yönteme uygulandığı takdirde işe yaramayabilir. Steganalizde amaç daha çok tek bir yöntem yerine bütün yöntemlere uygulanabilir, geniş çaplı bir sistem geliştirmektir. Günümüzde kullanılan yöntemler göz önüne alındığında steganalizde iki yaklaşım olduğu söylenebilir [9].

- İncelenen resim hakkında çok az ya da hiçbir istatistiksel varsayım olmadan yapılan analiz; burada istatistikler çok geniş bir veritabanı üzerinden elde edilir.

- İncelenen resim için bir parametrik model önerilir ve kullanılacak olan istatistikler bu model üzerinden hesaplanarak steganaliz çalışmalarında kullanılır.

Steganaliz problemlerini çözmek geliştirilen yaklaşımları sınıflandırırsak;

- Öğrenmeye dayalı steganaliz,
- Kör tanımlamaya dayalı steganaliz,
- Parametrik istatistiksel steganaliz,
- Hibrit teknikler

olarak sıralanabilir.

[9]'da bu yaklaşımların avantajları ve dezavantajları incelenmiştir.

3.1.1. Öğrenmeye Dayalı Steganaliz

Öğrenmeye dayalı yöntemler, stego resimler ve stego olmayan resimler arasındaki farklılıkları ortaya çıkarmak için bir sınıflandırıcı oluştururlar. Bazı resim özellikleri ortaya çıkartılır ve bu özellikler eğiticiye giriş olarak öğrenici sisteme verilir. Bu verilen örnekler hem stego mesajları hem de stego olmayan mesajları içerir. Öğrenme sınıflandırıcısı iteratif olarak tablolarını yeniden düzenler. Bu işlem istenilen en son stego sınıflandırıcısına varılıncaya kadar devam eder.

Bu yöntemle ilgili olarak steganalizin daha başarılı olarak gerçekleştirilmesi için dikkat edilmesi gerekli bazı hususlar aşağıda sıralanmıştır.

- Eğer resim kalite metrikleri ya da dalgacık katsayı istatistikleri gibi stego resme ait özellikler kullanılırsa daha iyi performanslı bir steganaliz yapılabilmektedir.
- Eğer sınıflandırıcı özel belirlenmiş bir algoritma için eğitilirse tespit etme başarısı daha yüksek olur. Çünkü verilen örnekler doğrultusunda bir öğretme mekanizması çalışmakta ve bu doğrultuda analiz yapılmaktadır.
- Genel bir steganaliz tespit etme yöntemi tasarlanmak istenirse öğrenmeye dayalı teknikler daha uygundur.

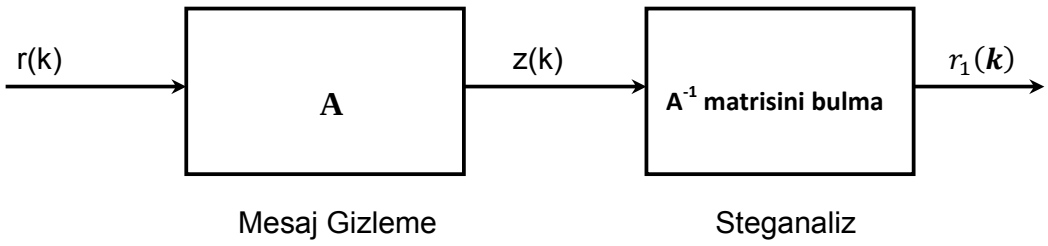
- Bu yöntem kullanıldığında standart özellikleri karşılamayan resimlerle karşılaşıldığında bile bir problem ortaya çıkmaz. Çünkü inceleme yapılırken ortalama değerler üzerinden yola çıkılır.
- Makine öğrenme son yıllarda ortaya çıkan bir teknoloji olduğundan ve bu konu üzerinde hala birçok araştırma yapıldığından, çok iyi bir teoriye ve metodolojiye sahiptir.
- İnternet üzerinden kolayca temin edilebilecek bazı ücretsiz yazılım paketleri ile kullanılacak olan dedektör eğitilebilir.

Bu tip steganaliz dedektörleri bazı faktörler ele alındığında çeşitli sınırlamalara sahip olabilirler. Bu sınırlamalar aşağıda belirtilmeye çalışılmıştır.

- İncelenecek her bir algoritma için ayrı bir sınıf eğitilmelidir. Bu da çok zaman alıcı bir durum olduğu için pratik değildir.
- Sınıflandırıcıyı eğitmek için özelliklerin seçimi çok önemlidir. Eğer seçilen özellikler uygulanacak algoritmaya uygun olarak seçilmedilerse detektör iyi çalışmayabilir hatta tamamıyla işe yaramaz olabilir. Bu özellikleri seçmek için kesin kurallar yoktur. Çoğunluğu sezgisel ya da deneme yanılma yöntemiyle bulunurlar.
- Bazı sınıflandırıcıların parametreleri steganalist tarafından seçilmelidir. Mesela, hangi çeşit çekirdek (kernel) seçilecek, doğrusal ya da doğrusal olmayan mı olacak, en son faza gelene kadar kaç adım ilerlemek gerekiyor ya da öğretmede kullanılacak örnek grup ne kadar genişlikte olacak. Bu değerler için de kesin seçme yöntemleri olmadığı için ancak deneme yanılma şeklinde seçilmek zorundadır.
- Eğer sınıflandırıcı çok yüksek doğruluklu olarak eğitilirse, genel özelliklerini kaybedebilir. Bu da bazı steganografi tekniklerinin tespit edilememesine neden olur.
- Bu yöntemlerde yanlış alarm ya da eksik olasılık steganalist tarafından kontrol edilememektedir. Yani steganalistin müdahale seçeneği yoktur.
- Resmin neresinde bir mesajın gizlendiğinin bulunması çok zordur hatta imkânsız denilebilir. Bu yüzden buradaki amaç sadece gizli bilgi vardır ya da yoktur yönünde bir karara varmaktır.

3.1.2. Kör Tanımlamaya Dayalı Steganaliz

$z(k)$ bir rastsal stego mesaj vektörü, A matris formundaki bir steganografi algoritmasını ve r de gizli mesajla birlikte taşıyıcı mesajı simgelesin [8]. Steganalist bu durumda $z(k)$'dan A^{-1}'i bulmak gibi bir problemle karşılaşır. Bu problem kör tanımlama problemi olarak tanımlanır ve Şekil 3.1'de gösterilmiştir.



Şekil 3.1. Kör Tanımlamaya Dayalı Steganaliz

Eğer burada A^{-1} bulunabilirse $r(k)$'nın bir tahminsel değeri olan $r_1(k)$ bulunabilir. Buradaki asıl steganaliz problemi $r(k)$'nın böyle bir parçasını elde edebileceğimiz doğrusal bir dönüşümün bulunabilmesidir.

Bu yaklaşım kullanılarak gerçekleştirilen steganaliz yöntemlerinin avantajları şunlardır:

- Bu steganaliz yaklaşımında herhangi bir eğitici veriye ihtiyaç yoktur. Her bir resim bireysel olarak sadece kendisine ait hesaplanan değerler kullanılarak analiz edilir. Bu durum resme ait istatistiklerin doğru olarak tahmin edilmesi durumunda iyi sonuç verir. Bu şekilde hesaplanan istatistikler resmin karakteristik özelliklerini daha iyi yansıttığından daha iyi sonuç alınır.
- Bu yaklaşımda gizli bilgi var ya da yok gibi bir sonuç elde etmekten çok gizli mesajın elde edilebilmesi mümkündür.
- Kör sistem tanımlama yapısı çok genel bir yapı olduğundan dolayı birçok farklı steganografi algoritması bu yapı içerinde modellenerek tespit edilmesi sağlanabilir.
- Başarılı bir steganalizin nasıl gerçekleştirilebileceği, orjinal ve gizli mesaj üzerinde bazı analitik sonuçlar ortaya konarak belirlenebilir.

Yukarıda avantajlarını belirttiğimiz bu yaklaşıma ait bazı dezavantajlar aşağıda görülmektedir.

- Dijital resimler istatistiksel olarak kararlı değillerdir. Kör tanımlama yöntemi verinin sabit olduğunu varsayarak uygulandığından, pratik olarak algoritmanın uygulanmasında sıkıntılar yaşanabilir.
- Sabit koşullar zorlandığında steganalizin başarılı olması için ekstra bir çaba gerekebilir.
- Eğer steganografi algoritması doğrusal değilse kör tanımlama probleminin çözümü daha zordur. Bu işlemi gerçekleştirebilmek için artı olarak daha yüksek seviyede birkaç istatistiğe ihtiyaç olabilir.
- Eğer stego resim ve gizli mesaj için yapılan istatistiksel modele dayalı varsayımlar doğru olmazsa çok ciddi performans kaybı yaşanabilir.

3.1.3. Parametrik İstatistiksel Tespite Dayalı Steganaliz

Bu yaklaşım birkaç farklı olgu olarak çalışılabilir. Özel olarak isimlendirmek gerekirse aşağıdaki olaylar ortaya çıkabilir [34].

- Tamamıyla Bilinen İstatistikler: Bu durumda steganografi algoritmasının bilindiği sadece kullanılan gizli anahtarın bilinmediği varsayılır. Bu sebeple resme ait bütün istatistikler steganaliz detektörü için elimizdedir.
- Kısmen Bilinen İstatistikler: Eğer steganografi algoritmasının sadece çalışan kodu elimizdeyse yani algoritmanın içyapısı bilinmiyorsa, çok büyük bir test seti kullanarak varsayımsal olarak istatistikler elde edilebilir.
- Tamamıyla Bilinmeyen İstatistikler: Bu durumda sadece stego resim bulunmaktadır. Steganografinin uygulanış amacını düşünürsek bu durum daha olası bir durumdur.

Eğer elimizde incelediğimiz sisteme ait bütün istatistikler varsa parametrik modele dayalı steganaliz ile gizli mesaj bulunabilir. Kısmen bilinen istatistikler için, parametrik model yine uygulanabilir fakat kullanılacak olan bütün parametreler bulunmaz. Bu parametreler tahmini olarak elde edilmeye çalışılır. Son olarak tamamıyla bilinmeyen istatistikler için değerlendirmede bulunursak ki bu durum

daha gerçekçi bir durumdur, bir model oluşturulur ve buna dayanılarak dedektör geliştirilir.

Bir parametrik modelin steganaliz dedektör olarak var olduğunu kabul edersek, aşağıda sayılan avantajlara sahiptir.

- Parametrik istatistiksel tespit etme teorisi çok gelişmiş bir alana dayanmaktadır. Bu yüzden bu alanda daha önce bilinen ve uygulanmış sonuçlar doğrudan steganaliz çalışmalarında kullanılabilir.
- Steganalistin analiz üzerinde kontrolü vardır.
- Bu yaklaşım ile gizli anahtarın mesajın geldiği yerin mesajın uzunluğu gibi bilgilerin tespit edilmesi mümkündür.

Yukarıda sayılan avantajlarına karşılık, bu yöntemin bazı sınırlamaları ve zorlukları bulunmaktadır. Bunlar maddeler halinde aşağıda sıralanmıştır.

- Yaklaşımın doğası gereği, parametrik steganaliz kullanılan parametrelerin varsayımlarının doğru olmamasına karşı duyarlıdır. Yani eğer tahmin edilen istatistikler doğru değilse bu doğrudan analizin performansını etkiler.
- Olasılıksal önceden bilme tartışmalı bir konudur. Genellikle objektif değildir ve bu yüzden kullanıcının doğrudan analizi etkilemesini sağlar. Yani yapılacak yanlış bir varsayım tespitin tamamıyla yanlış yönde gitmesine neden olur.
- Dijital resimlerin sabit olmayan istatistiksel özelliklerinden dolayı algoritmanın pratik olarak uygulanmasında bazı ciddi problemler ortaya çıkabilir.

3.1.4. Hibrit Teknikler

Hibrit tekniklerde, parametrik istatistiksel tespite dayalı steganaliz, kör tanımlamaya dayalı steganaliz ve öğrenmeye dayalı steganaliz yöntemlerinden birkaçı birlikte kullanılır. Analistin ihtiyaçlarına, elde bulunan veriye göre kullanılacak olan yöntemler seçilir ve uygulanır. Steganaliz yöntemlerinin zayıflık ve eksiklikleri düşünüldüğünde farklı yaklaşıma sahip bu yöntemlerin birkaçının bir arada kullanılmasıyla daha doğru ve güvenilir bir analiz gerçekleştirilmiş olur.

4. STEGANALİZDE KULLANILAN BAZI SALDIRILAR

Önceki bölümlerde incelendiği gibi birçok steganografi yöntemi ve uygulaması bulunmaktadır (Bkz. Ek 2). Bu yöntemlere karşılık çeşitli analiz yaklaşımları geliştirilmiştir. Bu yaklaşımlardan bazıları belirli steganografi yöntemlerine karşı geliştirilmiş olmakla beraber bazıları kör steganaliz olarak tanımlanan genel analiz yöntemleridir. Aşağıdaki bölümde EÖB saklama yöntemine karşı geliştirilmiş yöntemlerden ve kör steganaliz yöntemlerinden literatürde referans olarak gösterilen ve başarı oranları yüksek olan birkaç tanesi incelenmiştir.

4.1. RS Steganaliz

Çoğu resim için EÖB düzlemi rastsaldır ve kolayca anlaşılabilecek herhangi bir yapı/doku içermez. Rastsallaştırmanın derecesini ortaya çıkarmak için EÖB düzlemini kısıtlayan klasik istatistiksel nicelikleri kullanmak güvenli değildir. Rastsal görünmesine rağmen EÖB düzlemi diğer bit düzlemleriyle de ilişkilidir. Ancak bu ilişki doğrusal değildir.

Bu analizde, görüntülerde uzaysal ilintilerden üretilen duyarlı ikili istatistikleri kullanmaktadır. RS steganaliz 24 bit renkli ve 8 bit gri seviye görüntülerde kullanılmaktadır. Görüntü dosyaları üzerinde son bite ekleme yöntemine göre bilgi gizlenip gizlenmediğini anlamak için kullanılmaktadır. Yöntemde, bir resmin piksellerinin 3 bağımsız gruba: Düzenli (R-Regular), Tekil (S-Singular) ve Kullanılmayan (U-Unused) olarak ayrılması esastır [17].

Test edilen resmi *R* olsun. Bu *R* resmi *P* kümesinden değer alan $M \times N$ piksellerden oluşmaktadır. Örnek vermek gerekirse, 8-bit gri seviyeli bir resimde $P = \{0, \dots, 255\}$'dir. Yapılacak ilk işlem olarak *R*'nin, *n* komşu pikselden oluşan *G* ayrı gruba bölünmesidir.

$$G = \{x_1, x_2, \dots \dots . x_n\} \in R \tag{4.1}$$

Bu işlem yapılırken

$$f\,(G) = f\,(x_1, x_2\,, \dots . , x_n\,) \in R \tag{4.2}$$

ayırıcı fonksiyonu kullanılmaktadır. Ayırıcı fonksiyon *f* Eş. 4.3'de görüldüğü şekilde ifade edilmektedir.

$$f(x_1, x_2, \ldots\ldots, x_n) = \sum_{i=1}^{n-1} |x_{i+1} - x_i| \tag{4.3}$$

Yukarıda tanımlananlara örnek vermek gerekirse *n = 4* olarak seçildiğinde ayırma fonksiyonu Eş. 4.4.'teki şekilde tanımlanmaktadır.

$$G = (x_1, \ldots, x_4) \tag{4.4}$$

ve

$$f(x_1, x_2, x_3, x_4) = |x_2 - x_1| + |x_3 - x_2| + |x_4 - x_3| \tag{4.5}$$

şeklindedir. RS steganalizde üç adet tersine çevirme fonksiyonu *F* kullanılmaktadır. Bu tersine çevirme fonksiyonları Eş. 4.6'daki gibi tanımlanmaktadır.

- $F_1 : 2j \leftrightarrow (2j + 1) \qquad j = 0, 1, \ldots\ldots, 127$
  $F_1 : 0 \leftrightarrow 1, 2 \leftrightarrow 3, 4 \leftrightarrow 5, \ldots\ldots\ldots, 254 \leftrightarrow 255$

- $F_{-1} : (2j - 1) \leftrightarrow 2j \qquad j = 0, 1, \ldots\ldots\ldots, 128$
  $F_{-1} : -1 \leftrightarrow 0, 1 \leftrightarrow 2, 3 \leftrightarrow 4, \ldots\ldots\ldots, 255 \leftrightarrow 256$

- $F_0 : j \leftrightarrow j \qquad j = 0, 1, 2, \ldots\ldots\ldots 255$ (4.6)

$f(G)$ değerleri hesaplandıktan sonra bir maskeleme işlemi uygulanır. Maske *(M)*, $(-1,0,1)$ değerlerinden oluşmaktadır. Maske *G* grubuna uygulanır ve bu maskeye ait $F_M(G)$ değerleri hesaplanır. Maskenin değeri 1 ise F_1 fonksiyonu, maske değeri -1 ise F_{-1} fonksiyonu kullanılır. Daha sonra *−M* maskesi için de $F_{-M}(G)$ değerleri hesaplanır. Hesaplanan bu değerler aşağıdaki şartlara göre değerlendirilerek R_M , R_{-M} , S_M, S_{-M}, U_M ve U_{-M} sayıları hesaplanır.

− Eğer $f(F(G)) > f(G)$ ise *G* piksel grubu düzenlidir *(R)*.

− Eğer $f(F(G)) < f(G)$ ise *G* piksel grubu tekildir *(S)*.

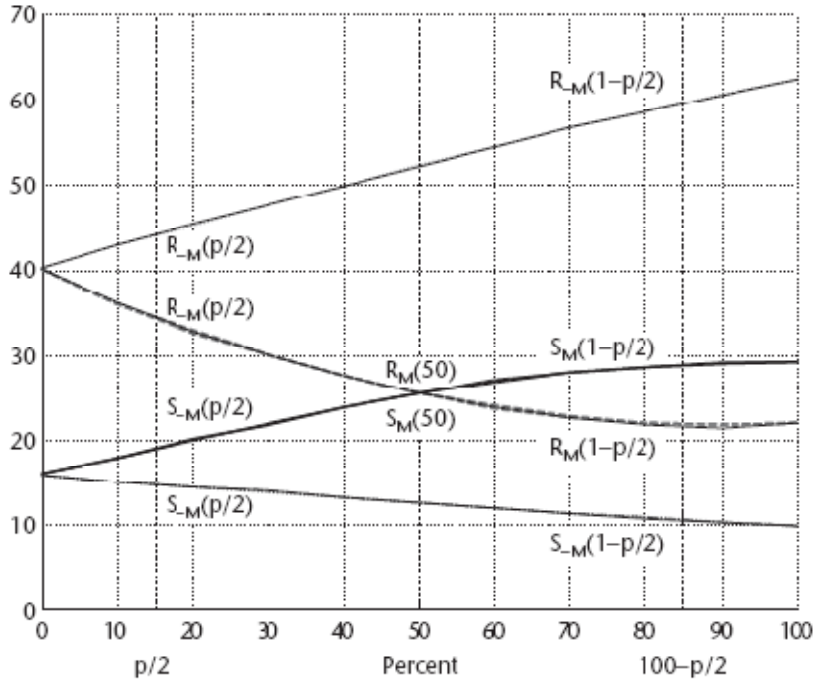
− Eğer $f\ (F(G))\ =\ f\ (G)$ ise *G* piksel grubu kullanılmayan *(U)* dır.

Tüm *G* grupları için pozitif ve negatif maskeler kullanılarak *R*, *S* ve *U* gruplarının sayısı belirlenir.

Daha sonra resmin tüm piksellerinin son bitleri değiştirilir ve yukarıdaki işlemler tekrar edilir. R_M , R_{-M} , S_M ve S_{-M} sayıları karşılaştırılarak bir sonuç elde edilir. RS steganaliz tekniğini geliştiren Fridrich [16], geniş bir veritabanı üzerinde yapmış olduğu testlerin ardından istatistiksel olarak yapılmış tahminleri deneysel olarak da ispatlamıştır.

$$R_M + S_M \leq 1 \qquad ve \qquad R_{-M} + S_{-M} \leq 1$$

$$R_M \cong R_{-M} \qquad ve \qquad S_M \cong S_{-M} \tag{4.7}$$



Şekil 4.1. RS diyagramı

Deneyler sonucunda Fridrich tarafından elde edilen *RS* diyagramı [16] Şekil 4.1'de sunulmuştur. Bu diyagrama dayanarak iki tane varsayımda bulunulmuştur.

1. R_M ve R_{-M} eğrilerinin kesim noktası ve S_M ve S_{-M} eğrilerinin kesim noktasıyla aynı *x* koordinatına sahiptir.
2. R_M ve S_M eğrileri $m = 50$'de kesişmektedir ya da başka bir ifadeyle $R_M\left(\frac{1}{2}\right) = S_M\left(\frac{1}{2}\right)$'dir.

p uzunlukta bir mesajın resim içerisine saklanmış olduğu kabul edilerek, F_1 ve F_{-1} tersine çevirme fonksiyonları kullanılarak Eş. 4.8.'de belirtilen 4 nokta hesaplanmıştır.

$$R_M\left(1-\frac{p}{2}\right), \quad S_M\left(1-\frac{p}{2}\right), \; R_{-M}\left(1-\frac{p}{2}\right) \; ve \; S_{-M}\left(1-\frac{p}{2}\right) \tag{4.8}$$

Bu analiz tekniğinde Eş. 4.7'de tanımlanan eşitlikler ve RS diyagramı kullanılarak Eş. 4.9.'da belirtilen eşitlik türetilir ve resim içerisine gizlenmiş olan mesaj uzunluğu *p* belirlenebilir.

$$2(d_1 + d_0)x^2 + (d_{-0} - d_{-1} - d_1 - 3d_0)x + d_0 - d_{-0} = 0$$

$$d_0 = R_M\left(\frac{p}{2}\right) - S_M\left(\frac{p}{2}\right)$$

$$d_1 = R_M\left(1-\frac{p}{2}\right) - S_M\left(1-\frac{p}{2}\right)$$

$$d_{-0} = R_{-M}\left(\frac{p}{2}\right) - S_{-M}\left(\frac{p}{2}\right)$$

$$d_{-1} = R_{-M}\left(1-\frac{p}{2}\right) - S_{-M}\left(1-\frac{p}{2}\right) \tag{4.9}$$

Eş. 4.9'da verilen denklemler *x* için çözülürse mesaj uzunluğu olan *p* değeri Eş. 4.10.'da belirtildiği şekilde bulunmaktadır.

$$p = \frac{x}{x-\frac{1}{2}} \tag{4.10}$$

Eğer belirtilen şartlar bir resim için sağlanıyorsa bilgi saklanmamış demektir. Değerlerin 0'a yakın çıkması resmin içinde bilgi olmadığını göstermektedir.

Bu analiz yönteminde ana ilke RS diyagramının dört eğrisini tahmin etmek ve çıkartımlar yardımıyla kesişimlerini hesaplamaktır.

Rastsal değişimlere bağlı olarak yapılan testler sonucunda orijinal içerisinde bilgi içermeyen resimlerde dahi sıfır olmayan mesaj uzunluklarının hesaplandığı görülmüştür. Bu başlangıç sapma değeri negatif ya da pozitif yönde olabilmektedir. Daha küçük boyutlu olan resimlerin başlangıç sapmaları bu resimlerin *S ve R* gruplarının sayıları daha küçük oldukları için daha yüksek olma eğilimi göstermektedir. Gürültülü resimler için *S* ve *R* grupları arasındaki fark daha küçük olmakta ve RS diyagramındaki eğrilerin kesişimi daha küçük bir açıyla gerçekleşmektedir. Bu da RS steganalizin doğruluğunun azalması anlamına gelmektedir.

Stego resim içerisinde mesajın rastsal olarak dağıtıldığı durumlarda RS Steganaliz sadece resmin küçük bir kısmına saklanmış olduğu durumdan daha başarılı sonuç vermektedir.

RS steganaliz yönteminin sözde programı (pseudo code) aşağıda verilmiştir [32].

Adım 1. Resmi seç

Adım 2. Maske değerlerini gir.

Adım 3. Her renk kanalı için ayrı ayrı uygulanmak üzere;

i. Resmi 4'lü *G* gruba böl.
ii. $f\ (G)$ ayırma fonksiyonu değerini hesapla.
iii. Maske *(M)* değerlerine göre uygun tersine çevirme fonksiyonlarını kullanarak $f\ (F(G))$ değerini hesapla.
iv. Ayırma ve tersine çevirme fonksiyonlarından elde edilen değerleri karşılaştırarak *R,S* ve *U* gruplarının sayılarını belirle.
v. $-M$ için de Adım *3i*, *3ii*, *3iii* ve *iv*'ü tekrarla.

Adım 4. Resmin tüm piksellerinin her baytının son bitlerini değiştir ve Adım 3'ü tekrarla.

Adım 5. Her renk kanalı için orijinal resim ve son bitleri değiştirilmiş resimden elde edilen R_M , S_M ve U_M sayıları arasındaki farkı hesapla.

Yukarıdaki kod sonucunda elde edilen farklar ne kadar 0'a yakınsa o derece incelenen resim içerisinde bilgi olma olasılığı azdır. İdeal durum bu değerin 0 çıkmasıdır.

4.2. RQP Steganaliz (Yakın Renk Çifti)

Bu yöntem de Fridrich [15] tarafından geliştirilmiştir. Önerilen bu yöntemin arkasındaki mantık, tipik taranmış resimler ya da dijital kameralardan elde edilmiş resimlerin EÖB düzlemleri rastsaldır ve bu EÖB şifrelenmiş mesajla değiştirmenin tespit edilebilir herhangi bir kalıntı bırakmayacak olmasıdır. Bu fikir eğer taşıyıcı resimdeki eşsiz renklerin sayısı resimdeki piksel sayısıyla karşılaştırılabilir seviyede ise doğrudur. Genel olarak 24 bit renkli resimler için eşsiz renklerin sayısının resimdeki piksel sayısından küçük olduğu gözlemlenmiştir. Normalde eşsiz renklerin piksel sayısına oranı 1:2, yüksek kalitedeki BMP resimler için 1:6 oranına kadar çıkmaktadır. Eşsiz renklerin sayısı JPEG sıkıştırmaya bağlı olarak bu tip resimlerde azalma eğilimi göstermektedir. Bu çıkartım 24 bit doğru renkli resimlerin göreceli olarak küçük renk paletine sahip olduklarına işaret etmektedir. Resim üzerine EÖB saklama işlemi gerçekleştirildikten sonra resmin renk paletinde ayırt edilir şekilde fazla sayıda yakın renk çiftleri oluşacaktır. Fazla sayıda yakın renk çiftine sahip olmak bir resim üzerinde EÖB saklama yöntemi kullanılarak bilgi saklandığına işaret edecektir. Bu yaklaşımdan yola çıkılarak stego resimlerdeki yakın renk çiftlerini analiz etmeye dayanan bir steganaliz yöntemi geliştirilmiştir.

Bir resimdeki eşsiz renklerin sayısı *U* ile ifade edilsin. *P*, resim paletindeki yakın renk çiftlerinin sayısı olsun. Eğer;

$$|R_1 - R_2| \leq 1, \quad |G_1 - G_2| \leq 1, \; ve \; |B_1 - B_2| \leq 1, \tag{4.11}$$

ise (R_1, G_1, B_1) ve (R_2, G_2, B_2) iki renk birbirine yakındır denir. Bu Eş. 4.12'deki şekilde daha iyi ifade edilebilir.

$$(R_1 - R_2)^2 + (G_1 - G_2)^2 + (B_1 - B_2)^2 \leq 3 \tag{4.12}$$

Bütün renk çiftlerinin sayısı;

$$\binom{U}{2} \geq P \tag{4.13}$$

dir. Resimdeki yakın renk çiftlerinin sayısı ile bütün renk çiftlerinin sayısı arasındaki *R* oranı, resimdeki göreceli yakın renk çiftlerinin sayısı hakkında bize bilgi verir.

$$R = \frac{P}{\binom{U}{2}} \tag{4.14}$$

İçerisinde saklanmış mesaja sahip olmayan bir resim için yakın renk çiftlerinin sayısının bütün olası renk çiftlerinin sayısına oranı, gizli mesaj içeren bir resmin değerlerinden daha küçük olacaktır. Bir resim zaten büyük bir mesaj içeriyorsa bunun içerisine başka bir mesaj daha yerleştirmek *R* oranında belirgin bir değişiklik yaratmayacaktır. Diğer bir yandan eğer resim gizli bir mesaj içermiyorsa *R* oranında fark edilir şekilde artma olacaktır. Buna dayanarak incelenmekte olan resim içerisine EÖB ekleme yöntemi kullanılarak mesaj saklanır. Mesaj saklama işleminin ardından resimdeki eşsiz renklerin sayısı *U'*, yakın renk çiftlerinin sayısı *P'* ve bu iki değer arasındaki *R'* oranı tekrardan hesaplanır. Daha sonra elde edilen bu iki değer birbirleriyle kıyaslanarak resim içerisinde gizlenmiş bir bilgi olup olmadığı hakkında karar verilir.

Karşılaştırma işlemi yapılırken bir eşik değerinin belirlenmesi gerekmektedir. Ancak eşsiz renkler resimlerde çok fazla değişiklik gösterdiklerinden *R* değeri için bir eşik değeri elde edebilmek oldukça zordur.

Analiz mekanizması aşağıdaki şekilde gerçekleşmektedir.

1. Bir resmin gizli mesaj içerip içermediğini tespit etmek için yakın renk çiftlerinin sayısı P ve bütün renk çiftlerinin sayısı arasındaki oran R hesaplanır.

$$R = \frac{P}{\binom{U}{2}}$$

2. Rastsal EÖB'e yerleştirme kullanılarak, MxN boyutlarındaki bir resme $\alpha 3MN$ bitlik bir test mesajı gizlenir.

3. Test mesajını yerleştirdikten sonra elde edilen yeni resim için karşılık gelen *U'* ve *P'* niceliklerini hesaplanır ve *R'* oranını elde edilir.

$$R' = \frac{P'}{\binom{U'}{2}}$$

Eğer daha önce resim içerisine büyük miktarda mesaj gizlendiyse iki oran yaklaşık olarak aynıdır $(R \cong R')$. Eğer resim gizlenmiş bir mesaja sahip değilse $R' > R$ olması beklenir. Bu yüzden ayırıcı bir özellik olarak $\frac{R'}{R}$ oranı alınır.

Yukarıda belirtilen analizde ayırıcı kıstas olan $\frac{R'}{R}$ oranının eşik değeri ve yerleştirilecek olan test mesajının büyüklüğünü belirleyen α değeri önem taşımaktadır. Fridrich [22] tarafından geniş bir veritabanı üzerinde yapılan testler sonucunda α değerinin %5 ve eşik değerinin $1{,}1$ olduğu durumun en iyi çözüm olduğu belirlenmiştir.

Aşağıda RQP Steganaliz yönteminin uygulanmasına dair bir programın sözde kodu verilmiştir. *R* yakın renk çiftlerinin tüm renk çiftlerine oranı, *R'* ise resim içerisine bilgi gizlendikten sonra hesaplanan orandır [33].

Adım 1. Resmi seç.

Adım 2. Yakın renk çiftlerinin sayısını hesapla (renk çiftleri arasındaki fark 3'ten küçük olanlar yakın renk çifti olarak seçilmiştir.)

Adım 3. Yakın renk çiftlerinin tüm renk çiftlerine oranını hesapla ve *R* olarak belirle.

Adım 4. Seçilen resmin içine bir test mesajı gizle ve oranı tekrar hesaplayıp *R'* olarak belirle.

Adım 5. *R* ile *R'* arasındaki farkı hesapla.

Adım 6. Eşik değerine göre karar ver.

4.3. Değişken Eşik Değerli Yakın Renk Çifti Analizi

Bu yöntemde daha önce açıklanan yakın renk çifti analiziyle benzerlikler taşımaktadır [27]. RQP yönteminde olduğu gibi yakın renk çiftlerinin sayısı tüm renk çiftlerinin sayısına oranına dayanan bir yöntemdir. Ancak karar verme açısından farklılık göstermektedir.

Analiz yönteminde öncelikli olarak incelenmekte olan resim için *U* eşsiz renklerin sayısı, *P* yakın renk çiftlerinin sayısı ve bunların oranın $R = P/U$ hesaplanır. Ardından EÖB saklama yöntemini kullanan bir steganografi algoritması ile incelenen resmin içerisine test mesajı saklanır. Daha sonra oluşturulan bu test resmi için *U', P' ve R'* değerleri hesaplanır. RQP steganaliz yöntemi ile arasındaki fark hesaplamalar sonucunda elde edilen verilerin değerlendirmesinde ortaya çıkmaktadır. Yöntemde karar değişken bir mekanizma ile sağlanmaktadır.

*R'*deki değişim oranı *R'*deki değişiklik yüzdesini veren *m* terimi ile ifade edilir.

$$m = \frac{(R - R')}{R * 100} \tag{4.15}$$

Farklı kategorideki resimlerden oluşan bir veritabanı üzerinde yapılan testler sonucunda *m* değeri için sabit bir eşik değeri atamanın bir kategori için doğru sonuç verdiği ancak farklı kategoriye ait performansının tatmin edici olmadığı belirtilmiştir [27]. Bu problemi aşmak için renk yoğunluğu, eşsiz renkler arasındaki fark gibi çeşitli resim istatistikleri kullanılarak değişken bir eşik değeri belirlenmiştir.

Eşik değeri,

$$t = \frac{(U' - P')}{d} \tag{4.16}$$

ya da,

$$t = \frac{\left(U' - \left(2 * P'\right)\right)}{d} \tag{4.17}$$

şeklinde ifade edilmiştir. Burada kullanılan *d* incelenen resmin ortalama renk yoğunluğu değeridir.

Genel olarak analiz aşağıdaki şekilde yapılmaktadır.

Adım 1. Standart bir EÖB saklama yöntemini kullanan steganografi aracı ile test resmi *i* için *s* stego objesini oluştur.

Adım 2. Gerekli değişkenleri tanımla ve başlat.

Adım 3. Test resmi *i*'nin piksel değerlerini oku.

Adım 4. Birbirini takip eden pikselleri karşılaştır. *U* ve *P* sayılarını yapılan tanımlamalara göre bul.

Adım 5. *R* değerini hesapla

Adım 6. Stego objesi *s* için Adım 2'den Adım 5'e kadar olan kısmı tekrar et. *U', P' ve R'* değerlerini hesapla.

Adım 7. *m* değerini hesapla.

Adım 8. *t*'yi hesapla.

Adım 9. *m ve t*'yi karşılaştır ve $m < t$ ise stego resim, $m \geq t$ ise stego olmayan resimdir.

4.4. Resim Düzgünlüğüne Bağlı Steganaliz

Bu yöntem EÖB steganografinin karakteristiğine dayanarak, resim içerisine herhangi bir mesaj yerleştirmenin resim kalitesine olan etkisini açığa çıkartmak için ortaya atılmıştır [36]. I resminin (i,j) noktasındaki yoğunluk değerini $I(i,j)$ simgelesin. İncelenmekte olan piksel, *KxK*'lık bir pencerede ele alınarak bölgesel komşuluk hesaplanır. Burada K tek bir sayıdır. Örnek vermek gerekirse *K=3* olduğunda 3x3'lük bloklar halinde inceleme yapılır ve pikselin kuzey, güney, doğu, batı, kuzeydoğu, kuzeybatı, güneydoğu ve güneydoğu komşu pikselleri hesaplamaya dâhil edilir. $X_I(i,j)$ fark değişkeni, incelenmekte olan piksel ve bu pikselin bölgesel komşuluğunda bulunan piksellerin ortalama değeri arasındaki farkı simgeler ve Eş. 4.18.'de görüldüğü gibi ifade edilir.

$$X_I(i,j) = I(i,j) - \frac{1}{K^2 - 1} \sum_{\substack{m,n=-(K-1)/2 \\ (m,n) \neq (0,0)}}^{(K-1)/2} I(i+m, j+n) \qquad (4.18)$$

X_I fark değişkeninin olasılık dağılımını ele alalım ve X_I'nin olasılık yoğunluk fonksiyonunu $f(x)$ olarak gösterelim. $f(x)$ olasılık yoğunluk fonksiyonu olduğundan aşağıdaki özelliklere sahip olacaktır.

1. Bütün domain üzerinde bir alan 1'e eşittir.

$$\int_{-\infty}^{+\infty} f(x)\, dx = 1 \qquad (4.19)$$

2. X_I rastsal değişkeninin beklenen değeri 0'a eşittir.

$$\int_{-\infty}^{+\infty} f(x)\, x\, dx = 0 \qquad (4.20)$$

Şekil 4.3'te standart resim "Lena " ve bu resme ait fark değişkeni X_I'nın istatistiksel dağılımı görülmektedir.

(a) (b)

Şekil 4.2. (a) Standart resim "Lena" (b) Fark değişkeni $\boldsymbol{X_I}$'nin dağılımı

X_I fark değişkeninin varyansı incelenmekte olan piksel bölgesel komşuluğunda bulunan piksellerin ortalama değerinden ne kadar farklılık

gösterdiğini yansıtır. Bu yüzden bu fark değişkeni resmin düzgünlüğünün bir ölçüsü olarak kullanılabilir. Buradan yola çıkılarak *I* doğal resminin düzgünlüğü Eş. 4.21.'de belirtildiği gibi ifade edilebilir.

$$SM_I = \sigma^2 = \int_{-\infty}^{+\infty} f(x).x^2\,dx \tag{4.21}$$

Resim düzgünlüğü değeri SM_I'nın değeri büyüdükçe incelenmekte olan piksel değeri bölgesel komşuluğunun ortalama değerinden daha yoğun şekilde farklılık gösterecektir. Bunun tam tersi de geçerli olacaktır.

4.4.1. Resim Düzgünlüğüne Dayanarak LSB Steganografinin Tespit Edilmesi

Yapılan incelemelerde EÖB veri yerleştirme işleminin resimlerin düzgünlüğünü azalttığı tespit edilmiştir [36]. Bu yüzden gizli bilgi içeren steganografi uygulanmış resimlerin ifade edilen şekilde resim düzgünlüğü hesaplandığında, orijinal resimlere göre daha büyük değere sahip olması beklenmektedir.

4.4.2. Stego-Resimlerin Resim Düzgünlüğünün Analizi

Orijinal resim yani gizli bilgi için taşıyıcı olacak resmi *C* ve gizli bilgi yerleştirdikten sonra ortaya çıkan stego-resmi *S* olarak tanımlayalım. *C* orijinal resmi için fark değişkeni X_C ve stego-resim *S* için ise X_S olarak tanımlayalım. X_C ve X_S fark değişkenlerinin olasılık yoğunluk fonksiyonları $f_C(x)$ ve $f_S(x)$ olarak tanımlansın. Taşıyıcı resim *C* için resim düzgünlüğü,

$$SM_C = \sigma^2 = \int_{-\infty}^{+\infty} f_C(x).x^2\,dx = \int_{-\infty}^{+\infty} f(x)\,.x^2\,dx \tag{4.22}$$

olarak ifade edilebilir. Stego resim *S*'nin *p* yerleştirme oranında gizli mesaj içerdiğini düşünelim. Orijinal taşıyıcı resim ile karşılaştırıldığında $(1-p/2)*M*N$ adet pikselin piksel değeri değişmeden kalır. Geri kalan değişen piksellerden $(p/4)*M*N$ kadar pikselin piksel değeri 1 arttırılmıştır ve $(p/4)*M*N$ pikselin piksel değeri ise 1 azaltılmıştır. Yapılan incelemede buna rağmen komşuluk ortalama piksel değerinin hemen hemen değişmeden kaldığı gözlemlenmiştir. Bunun nedeni bölgesel komşulukta değeri 1 artırılan piksel sayısının, değeri 1

azaltılan piksel sayısına yaklaşık olarak eşit olmasındandır. Değeri değişmeden kalan piksel için X_S 'in olasılık yoğunluk fonksiyonu yaklaşık olarak $f(x)$'e eşittir. Değeri 1 artan pikseller için X_S'in olasılık yoğunluk fonksiyonu yaklaşık olarak $f(x-1)$ ve değeri 1 azalan pikseller için bu $f(x+1)$'e eşittir. Bunlara dayanarak stego-resim *S* için X_S'in olasılık yoğunluk fonksiyonu Eş. 4.23.'te görüldüğü gibi ifade edilir.

$$f_S(x) \approx \left(1-\frac{p}{2}\right).f(x) + \frac{p}{4}.f(x-1) + \frac{p}{4}.f(x+1) \tag{4.23}$$

$f_S(x)$ açıkça görüldüğü gibi diğer üç dağılımın doğrusal kombinasyonu halindedir. Eş. 4.23.'teki denklem kullanılarak stego resim *S*'nin resim düzgünlüğü Eş. 4.24.'te verildiği gibi çıkartılabilir.

$$SM_S = \sigma^2 = \int_{-\infty}^{+\infty} f_S(x)\, x^2\, dx$$

$$\approx \int_{-\infty}^{+\infty} \left(\left(1-\frac{p}{2}\right)f(x) + \frac{p}{4}f(x-1) + \frac{p}{4}f(x+1)\right) x^2 dx$$

$$\approx \left(1-\frac{p}{2}\right) \int_{-\infty}^{+\infty} f(x)x^2\, dx + \frac{p}{4} \int_{-\infty}^{+\infty} f(x-1)x^2\, dx + \frac{p}{4} \int_{-\infty}^{+\infty} f(x+1)x^2\, dx$$

$$\approx \left(1-\frac{p}{2}\right) \sigma_C^2 + \frac{p}{4} \int_{-\infty}^{+\infty} f(y)(y+1)^2 dy + \frac{p}{4} \int_{-\infty}^{+\infty} f(y)(y-1)^2 dy$$

$$\approx \left(1-\frac{p}{2}\right)\sigma_C^2 + \frac{p}{4}\,(\sigma_C^2+1) + \frac{p}{4}\,(\sigma_C^2+1)$$

$$\approx \sigma_C^2 + \frac{p}{2} \tag{4.24}$$

4.4.3. LSB Düzlemi Çevrildikten Sonra Stego-Resmin Resim Düzgünlüğünün Analizi

Bir *S* stego resminin LSB düzleminin çevrilmesi, bütün resmin en değersiz bitlerine lojik NOT işleminin uygulanmasıdır. Bu işlem ile 0 olan bitler 1, 1 olan bitler ise 0 değerini alır. İşlem sonucunda elde edilen resim *T* ile simgelenir.

Orijinal taşıyıcı resim olan *C* ile karşılaştırıldığında *T* resminde $p/2 * M * N$ piksel değişmeden kalacaktır. Yine burada *p* saklanan mesaj oranını ifade etmektedir. *T* resmi içerisinde piksel değeri 1 artan $(1/2 - p/4) * M * N$ adet piksel ve değeri 1 azalan $(1/2 - p/4) * M * N$ piksel vardır. Yukarıda *S* resmi için yaptığımız işlemlerin benzerini *T* resmi için de yaparsak, fark değişkeni X_T'nin olasılık yoğunluk fonksiyonu Eş. 4.25.'teki gibi ifade edilebilir.

$$f_T(x) \approx \frac{p}{2} f(x) + \left(\frac{1}{2} - \frac{p}{4}\right) f(x-1) + \left(\frac{1}{2} - \frac{p}{4}\right) f(x+1) \tag{4.25}$$

Buradan yola çıkılarak değiştirilen resim düzgünlüğü hesaplanırsa,

$$SM_T = \sigma_T^2 = \int_{-\infty}^{+\infty} f_T(x) \, . \, x^2 \, dx$$

$$\approx \int_{-\infty}^{+\infty} \left(\frac{p}{2} f(x) + \left(\frac{1}{2} - \frac{p}{4}\right) f(x-1) + \left(\frac{1}{2} - \frac{p}{4}\right) f(x+1)\right) x^2$$

$$\approx \frac{p}{2}\sigma_C^2 + \left(\frac{1}{2} - \frac{p}{4}\right)(\sigma_C^2 + 1) + \left(\frac{1}{2} - \frac{p}{4}\right)(\sigma_C^2 + 1)$$

$$\approx \sigma_C^2 + 1 - \frac{p}{2} \tag{4.26}$$

elde edilir.

4.4.4. Veri Saklama Oranının Tahmin Edilmesi

SM_S ve SM_T eşitliklerinden verilen bir *S* stego resim için yerleştirilen mesajın uzunluğu tahmin edilebilir. Yerleştirme oranının tahmini,

$$\hat{p} = 1 + SM_S - SM_T = 1 + \sigma_S^2 - \sigma_T^2 \tag{4.27}$$

şeklindedir.

Gizli mesajın yerleştirme oranının doğru olarak tahmininde iki ana unsur önem taşır.

1. Bilgi saklama oranı *p,* 0 ya da 1'e çok yakın ise yukarıda verilen eşitlikler tam olarak sağlanamamaktadır.

2. Ortalama komşuluk piksel değerinin mesaj yerleştirme işleminden sonra yaklaşık olarak değişmeden kaldığı kabul edilmiştir. Fakat bu yerleştirilen bit miktarı çok az ise doğru olmamaktadır.
3. Resim verisinin farklılığı, gizli mesajın ve kullanılan yerleştirme işleminin rastsallığı yerleştirme oranının tahmin edilmesinde hatalara neden olmaktadır.

Bu steganaliz yönteminin çıkış noktası incelenen piksel ve bu pikselin komşularının ortalama değerinin farkının dağılımının istatistiksel modeline dayanmaktadır. Bu istatistiksel dağılımın varyansına dayanarak resim düzgünlüğü kavramı açıklanmıştır. Önerilen bu yöntem diğer yöntemler ile kıyaslandığında daha küçük hesaplama yükü, farklı fiziksel anlam ve düşük hesaplama kompleksliği getirmektedir ve gerçek zamanlı tespit etme için uygun bir yöntem olarak görülmektedir. Ancak analiz çeşitli varsayımlara dayandığı için her tip resimde doğru sonuçlar vermemektedir.

4.5. Koşu Uzunluğuna Bağlı Steganaliz

4.5.1. Koşu Uzunluğu Nedir?

Birbirini takip eden birçok veri elemanı arasında aynı değere sahip diziye “koşu” denmektedir [28]. Bu elemanların tekrar etme sayısına “koşu uzunluğu” denmektedir. Asıl kullanıldığı alan veri sıkıştırmadır ve genellikle koşu uzunluğu kodlama olarak bilinmektedir. Aşağıdaki bir örnekle bu açıklanmaya çalışılmıştır.

“aaaaaaaaaaabbbbbbbbbbbbaaaaaaaaaaaaaaaaaaabbbbbbbbbbcccccccacc”

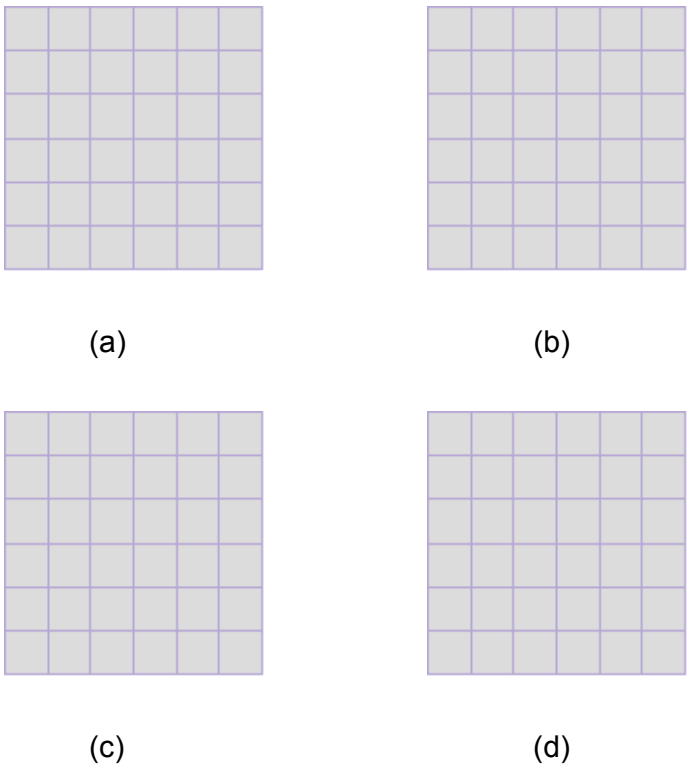
Eğer yukarıdaki dizide koşu uzunluğunu hesaplarsak gösterim şu şekilde olur.

11a12b18a10b6c1a2c

Bu şekilde 67 karakter uzunluğundaki bir dizi sadece 19 karakter ile simgelenebilmektedir. Bu işlem genellikle kayıpsız sıkıştırma işlemini sağlamak için kullanılmaktadır.

Koşu uzunluğunu resim için açıklarsak aynı değere sahip ardışık piksellerin sayısı olarak söylenebilir. Resim için koşu uzunluğunu hesaplarken yön önem

kazanmaktadır. 4 farklı yönde koşu uzunluğu hesaplaması yapılmaktadır. Bu yönlere ait gösterim aşağıda Şekil 4.3'te gösterilmektedir.



Şekil 4.3. Koşu-uzunluğu hesaplama yön gösterimleri (a) yatay $\mathbf{0}^{o}$, (b) dikey $\mathbf{90}^{o}$, (c) diyagonal $\mathbf{45}^{o}$, (d) diyagonal $\mathbf{135}^{o}$.

Koşu Uzunluğu Steganaliz yöntemi kör steganaliz olarak adlandırılan yöntemlerdendir [13]. Kör steganalizde iyi bir stego sınıflandırıcı ortaya çıkartmak için uygun özelliklerin seçimi hayati önem taşımaktadır. Resmin koşu-uzunluğu histogramının, karakteristik fonksiyonunun istatistiksel momentleri bu yöntemde özellik olarak seçilmiştir.

4.5.2. Steganaliz İçin Koşu-Uzunluğu Analizi

24 bitlik resimler için koşu kavramı düşünülürken tek bir renk kanalı üzerinden konuşmak gerekir. Koşu uzunluğu ise bu devam eden piksel dizisinin başka bir deyişle koşunun uzunluğu olarak ifade edilmektedir. Bunu bir resim için

düşünürsek, bir koşu uzunluğu matrisi $p(i,j)$ i gri seviye değerine sahip piksellerin j uzunluğundaki koşuların sayısını simgelemektedir.

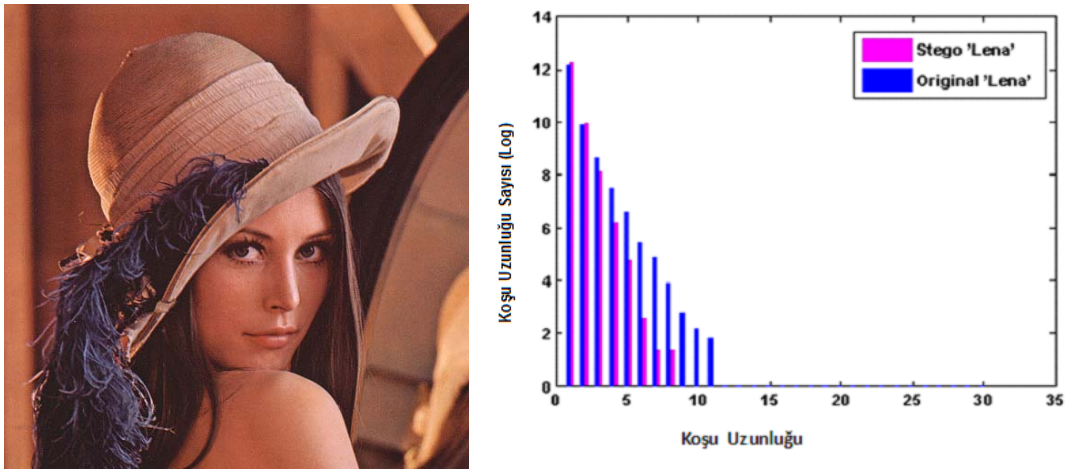
Örnek vermek gerekirse, 25 değerine sahip piksellerden ardı ardına 0 derece yönünde devam eden 7 tane varsa ve bu 7 uzunluklu 25 değerine sahip koşu tüm resim boyunca 3 kere tekrar ediyorsa $p(25,7) = 3$'tür ve bu şekilde ifade edilir. Daha düzgün şekilde bunu tam olarak resimlere göre açıklarsak, bir $p_\theta(i,j)$ koşu-uzunluğu matrisi için *M* gri seviyelerin sayısını (bir resim için maksimum değeri 255'tir.) ve *N* ise maksimum koşu sayısını ifade etsin. Bir resim için maksimum değer hangi yönde gidiliyorsa o yöndeki maksimum genişliktir. Yani eğer resmin genişliği yönünde, 0 derece, gidiliyorsa bu değer resmin enidir. Resmin koşu-uzunluğu histogramı (RLH) bir vektör olarak Eş. 4.28.'de gösterildiği gibi ifade edilebilir.

$$H_\theta(j) = \sum_{i=0}^{M} p_\theta(i,j) \qquad 1 < j < N \tag{4.28}$$

Günümüzde kullanılan steganografi algoritmalarının çoğu veri saklarken piksel başına işleme dayanan yöntemler kullanmaktadır. Saklanacak veriye ait bir bit saklanırken resmin karşılık gelen tek bir piksel değişir [13]. Bahsedilen bu işlem direkt olarak resmin bölgesel yoğunluk değişimiyle ilgilidir. Önerilen bu yöntemin çıkış noktası da bu teoriye dayanmaktadır.

Koşu uzunluğu istatistikleri belirli bir yönde yapının kalıbını yakalayabilmektedir. Koşunun uzunluğu resmin yapı elementlerinin detaylarını göstermektedir. Aynı zamanda koşu uzunluğu bir resmin bölgesel yoğunluk değişimini de yansıtabilmektedir. Bu teori koşu-uzunluğu histogramının neden steganaliz için özelliklerin tabanı seçildiğinin cevabıdır. Piksel başına veri yerleştirme işlemi yapıldıktan sonra, elde edilmiş olan bölgesel yoğunluğun sürekliliği bozulur ve karşılık gelen piksel koşu uzunlukları ve koşu değerleri değişmektedir. Örneğin EÖB steganografi yöntemlerinden biri uygulandıktan sonra bazı piksel değerleri ya azalacaktır ya da artacaktır. Doğal olarak da bu değişimler resmin koşu uzunluğu histogramını etkileyecektir. Bunun sonucunda, resimde uzun sayıda gerçekleşen koşular piksel değerlerinde olan değişimden dolayı daha küçük daha kısa koşulara bölünebilir ya da daha az sayıda uzun koşuya neden olarak kısa koşuların sayısı artabilir. Sonuç olarak resmin RLH'ının değeri

azalacaktır. Uzun koşuların bölünerek kısa koşulara dönüşümlerinin beklenmesinin yanı sıra kısa uzunlukta olan koşular da az bir ihtimalde olsa daha uzun koşular oluşturabilirler. Ancak doğal resimlerin uzaysal ilintilerinden dolayı uzun koşuların bölünmesi daha yüksek bir olasılıktır. Şekil 4.4'te "Lena" resmine ait normal RLH ve veri saklandıktan sonraki RLH gösterilmektedir. Burada açıkça uzunluklardaki küçülme görülmektedir.



Şekil 4.4. Lena resminin normal ve steganografi uygulandıktan sonraki koşu-uzunluğu histogram örneği

4.5.3. Parametize Edilmiş Koşu-Uzunluğu Gösterimleri

Doğal resimler için, bir resmin RLH'ında kısa koşuların sayısı gözle görülür şekilde uzun koşuların sayısından daha fazladır. Koşuların maksimum uzunluğu olası uzunluk değerlerinin aralığıyla kıyaslandığında genellikle daha kısıtlıdır. Resmin RLH'ının veri yerleştirme sonucunda kısalmasını daha açık ve duyarlı hale getirmek için iki tane yeni koşu-uzunluğu gösterimi açıklanmıştır [13]. Bu iki gösterim aynı mantıkta koşu uzunluğu hesaplamakta ancak farklı kurallara ve parametrelere dayanmaktadır.

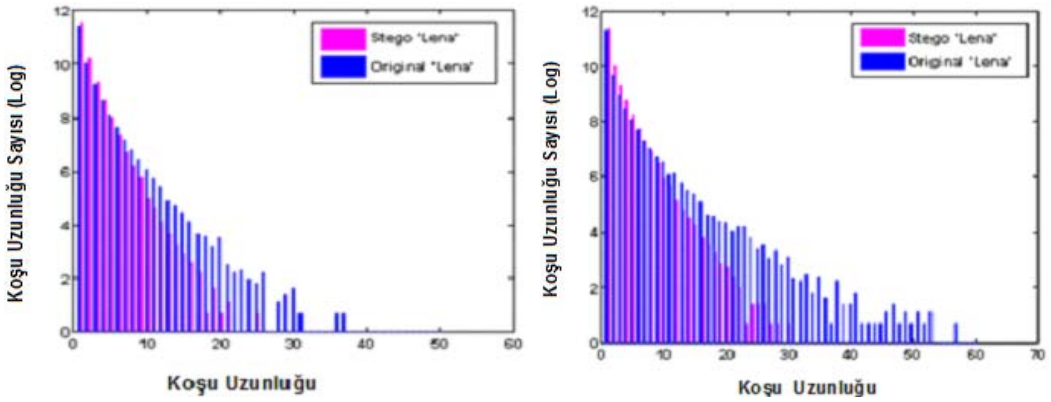
4.5.3.1. Nicemleme ile Koşu-Uzunluğu Gösterimi

Resmin piksellerinin koşu uzunluğu matrisini doğrudan bulmak yerine önce bir Q adım faktörü ile resim düzlemi nicemlenir. Daha sonra nicemlenmiş resmin RLH'ı hesaplanır. Örnek vermek gerekirse, 256 gri seviyeli bir resim Q=2 ile

nicemlenirse, yoğunluk değeri 0'dan 127 arasında değişen yeni bir resim matrisi elde edilmiş olur. Normal resmin RLH'ına kıyasla elde edilen yeni nicemlenmiş resmin RLH'ındaki uzun koşuların sayısı artacaktır. Çünkü her bir komşu yoğunluğa sahip piksel çiftlerinin nicemlemeden sonraki değeri aynı koşu değeri içerisine düşecektir. *Q* değeri daha büyük oldukça daha uzun koşular elde edilecektir. Geleneksel olarak uygulanan koşu uzunluğu hesaplama aslında *Q=1* olduğundaki nicemlemeye eşittir.

4.5.3.2. Koşu Uzunluğu Fark Gösterimi

Bu tip gösterimde koşu, bir yöndeki ortalama Є kadar birbirinden farklı değere sahip piksel dizisine denk gelmektedir. Ardı ardına gelen piksel değerlerinde çok küçük değişiklikler varsa bunlar tek bir koşu olarak simgelenir. Örneğin, 124, 125, 125, 126 değerlerine sahip 4 resim pikseli normal koşu uzunluğu hesaplaması ile $p(124,1),\ p(125,2)\ ve\ p(126,1)$ olarak simgelenir. Bunu fark koşu uzunluğu gösterimi ile simgelersek ve $\varepsilon = 1$ olarak alındığında gösterim $p(124,4)$ şeklini alır ve koşu uzunluğu artmış olur. Aynı şekilde ε değeri artarsa daha uzun koşular elde edilebilir. Normal koşu uzunluğu hesaplama, fark koşu uzunluğu hesaplamanın $\varepsilon = 0$ olarak alınmış özel halini temsil etmektedir.



(a) (b)

Şekil 4.5. Orijinal ve steganografi uygulanmış Lena resmi için (a) $\boldsymbol{Q = 4}$ nicemleme koşu uzunluğu histogramı, (b) koşu uzunluğu fark histogramı

Yukarıda açıklanan bu iki yeni koşu uzunluğu gösterimi geleneksel koşu uzunluğu gösterimine kıyasla daha çok ve daha uzun koşuların oluşmasını sağlamaktadır. Sonuç olarak, bu yeni iki yöntemin kullanılmasıyla veri saklama sonucunda oluşan RLH'taki kısalma daha da açık hale gelecek ve resim RLH'ı veri saklamaya daha duyarlı olacaktır. Lena resmi için bu iki koşu uzunluğu gösterimi Şekil 4.5'te gösterilmiştir [13]. Şekilden de görülebileceği gibi koşuların uzunlukları oldukça genişlemiş durumdadır. Grafiklere ilk bakıldığında yeni önerilen koşu uzunluğu yöntemlerinin steganaliz yöntemi için oluşturulan teorem ile uyuşmadığı düşünülebilir. Bakıldığında veri saklama işleminden sonra bu iki yöntem ile koşu uzunluğu hesaplaması yapıldığında değişimde bir azalma görülmektedir. Ancak yine de yapılan deneylerde Q ve ε için uygun değerler seçilerek resmin RLH'ı üzerinde yapılan genel bir analizin veri saklamanın etkilerini ifade ettiği gösterilmiştir [13]. Bu etki yukarıda açılanan iki gösterimden elde edilen koşuların resim halinde gösterilmesiyle daha iyi anlaşılabilir. Bu gösterim Şekil 4.6'da görülmektedir. Bu şekilde farklı koşu değerleri siyah ve beyaz olarak gösterilmiştir.

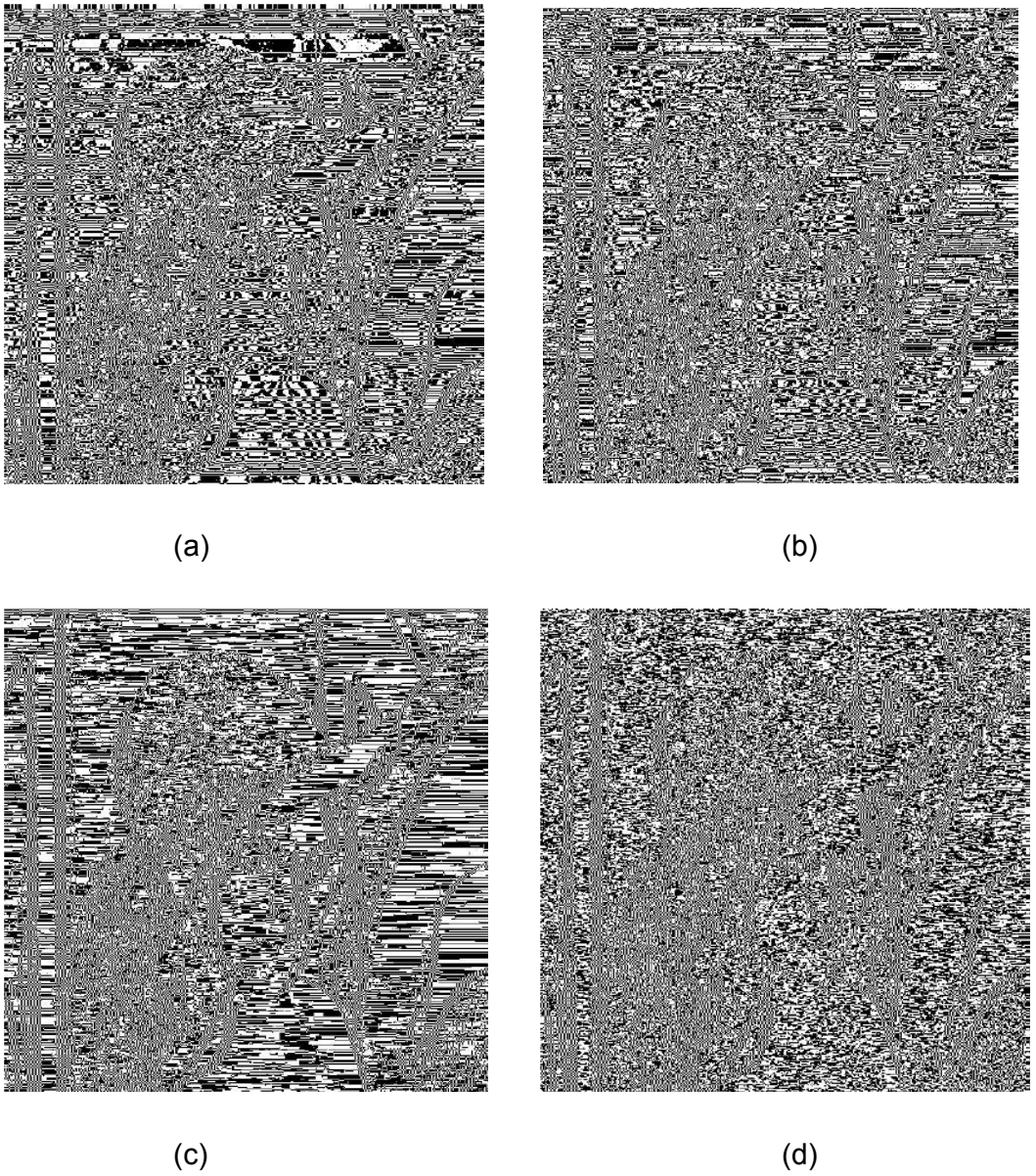
4.5.4. Steganaliz için Özellikler

Elde edilen üç tip resim RLH'ının karakteristik fonksiyonlarının çok dereceli momentleri ayırıcı özellik olarak kullanılmaktadır ve Eş. 4.29.'da ifade edildiği gibi gösterilmektedir.

$$M_n = \sum_{j=1}^{L/2} f_j^n \left|F_i(f_j)\right| / \sum_{j=1}^{L/2} \left|F_i(f_j)\right| \qquad i = 1,2,3 \tag{4.29}$$

Burada F_i, resmin RLH'ının karakteristik fonksiyonu (H_i'nin DFT'si), $F_i(f_j)$ F_i'nin f_j frekansındaki bileşeni ve *L* ise DFT dizi uzunluğudur. Özellik çıkartımında, 4 farklı yöndeki (0^0, 45^0, 90^0, 135^0) 3 farklı koşu uzunluğu yöntemi (H_1, H_2, H_3) ile hesaplanan RLH'ların her birinin karakteristik fonksiyonlarının ilk üç derece momentleri hesaplanır. Böylelikle resmin RLH'ının her bir tipi için 12-boyutlu özellik vektörü elde edilmiş olur. Karakteristik fonksiyonun dördüncü ve beşinci derece momentleri özellik olarak alınmamıştır çünkü bu iki moment diğer ilk üç momente göre çok daha etkisizdir. Bu elde edilen özellikler DVM kullanılarak bir model

oluşturulur ve modele dayanarak incelenen resmin hangi sınıfa ait olduğuna dair tahmin yapılmaya çalışılır.



(a) (b)

(c) (d)

Şekil 4.6. Lena için koşu uzunluğu resim gösterimleri (a) normal Lena $\boldsymbol{Q} = \mathbf{4}$, (b) stego Lena $\boldsymbol{Q} = \mathbf{4}$, (c) normal Lena $\boldsymbol{\varepsilon} = \mathbf{2}$, (d) stego Lena $\boldsymbol{\varepsilon} = \mathbf{2}$

4.6. İkili Benzerlik Ölçütlerine Dayalı Steganaliz

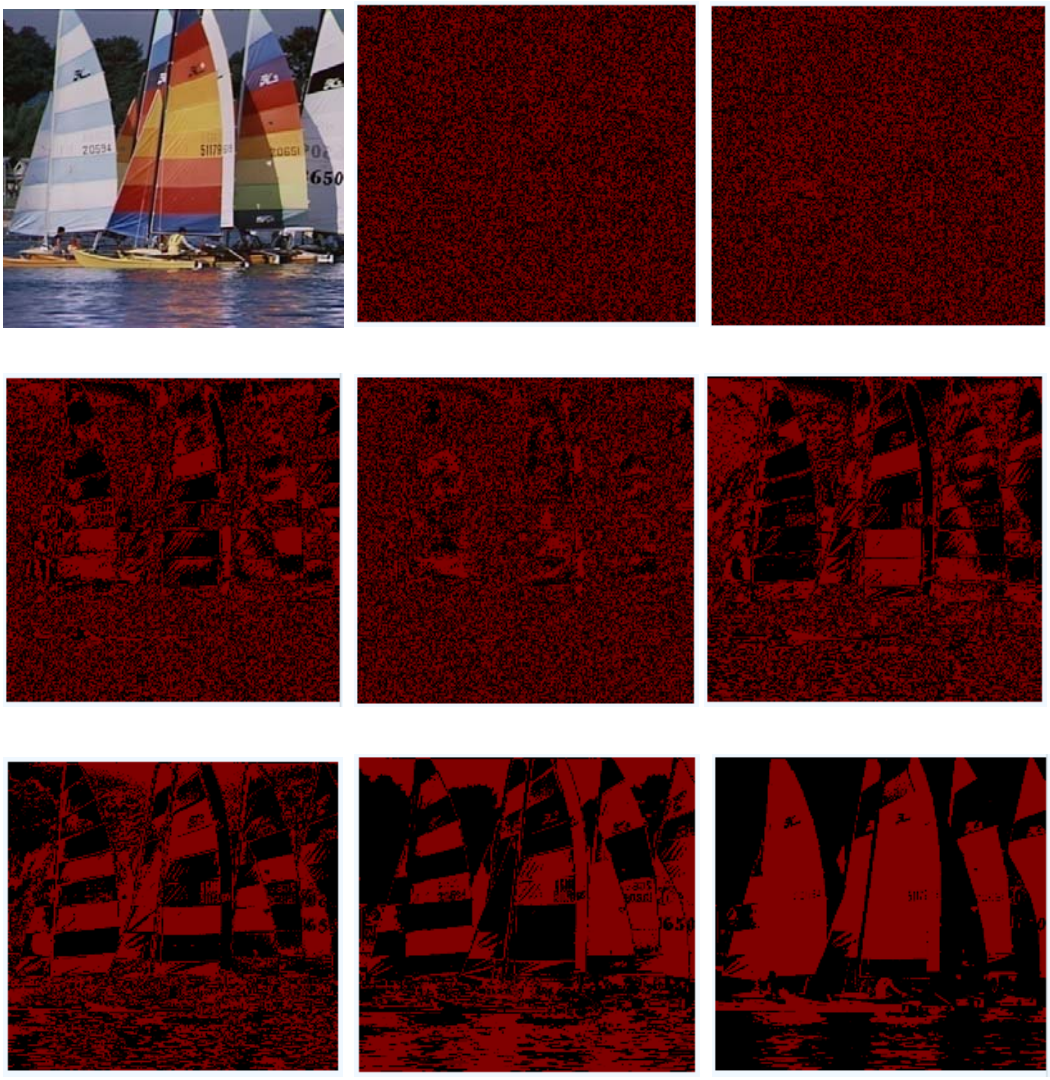
Bu analiz yöntemi, gizli mesajın varlığını tespit edebilmek için ardışık bit düzlemleri arasındaki ikili benzerlik ölçütlerine dayanmaktadır [10]. Yöntemdeki ana fikir, bit düzlemleri arasındaki karşılıklı ilinti ve bit düzlemlerindeki ikili doku karakteristiklerinin taşıyıcı ve stego resimler arasında farklılık göstermesidir [4]. Yöntemde yedinci ve sekizinci bit düzlemleri kullanılmakla beraber diğer bit düzlemleri de bu ikili benzerlik ölçütlerinin hesaplanmasında kullanılabilmektedir. Yöntem herhangi bir referans resim gerektirmemekle birlikte hem uzamsal/uzaysal hem de dönüşüm etki alanı steganografi kullanılarak bilgi saklama gerçekleştiren yöntemler için kullanılabilmektedir. Bu yöntem ile steganografik verinin/içeriğin varlığını ortaya çıkartmak için resimlerin kendine özgü istatistiksel özellikleri elde edilip incelemektedir.

4.6.1. İkili Resimlerdeki Benzerlik Ölçütleri

Analizde, gizli bilginin varlığını tespit etmek için, resimlerin düşük seviyeli bit düzlemlerinden çıkartılan istatistiksel özellikleri incelenmektedir. Renk kanalına ayrılmış olan bir resim 8 bitlik ayrı bir resim ifade eder. Bu 8 bitlik resmin her bir bit düzlemi yine 1 bitlik ayrı birer resim belirtmektedir [5]. Bu resimlere bir örnek Şekil 4.7'de verilmiştir. Resimlere herhangi bir steganografik içeriğin eklenmesi, incelenmekte olan bit değerinin yanı sıra bu bit değerinin bulunduğu bit düzleminin komşuluğunda bulunan yakın bit düzlemlerinde de değişikliğe neden olmaktadır. Yani bir başka deyişle bit düzlemleri arasındaki ilinti bozulmakta ve bit düzlem dokuları etkilenmektedir. Gerçekleşen bu etki bit değişimlerinin olasılıkları incelenerek bulunmaktadır. Resme veri eklemek bit düzlemleri arasındaki ilintiyi artırma ya da azaltma şeklinde ortaya çıkabilir.

Direkt olarak bit düzlemleri arasındaki bit ilintilerinin kullanılması resim içerisine saklanmış olan gizli bilginin tespit edilebilmesi için yeterli özellik kriterleri sağlayamamaktadır. Bu yüzden bit düzlemleri arasındaki direkt ilişki yerine ikili örüntü özelliklerinin istatistiklerine dayanan bir karşılaştırma kullanılmaktadır.

$x_i = \{x_{i,k} | k = 1, \dots\dots, K\}$ bir pikselin K komşuluğunu temsil eden bit dizisini temsil etsin. *K* değeri bir pikselin komşuluk sayısını temsil etmektedir. *K* değeri eğer 4 ise incelenmekte olan pikselin kuzey, güney, doğu ve batısında bulunan



Şekil 4.7. Yelkenli resmi ve bu resmin kırmızı renk kanalına ait bit düzlem resimleri (1-8)

piksel değerlerini içerir. Eğer bu değer 8 ise piksele ait tüm komşuluklar yani kuzey, doğu, batı, güney, kuzeydoğu, kuzeybatı, güneydoğu ve güneybatı değerlerini içerir. Bir pikselin komşuluk gösterimi Şekil 4.8'de görülmektedir. x_i değişkenindeki i indeks değeri bütün resim pikselleri boyunca değer almaktadır. İncelenmekte olan resmin boyutunun M x N olduğunu varsayarsak, 5 noktalı şablon fonksiyonu (uyuşma değişkeni) $X_{r,s}$ Eş. 4.30.'da görüldüğü gibi tanımlanmaktadır.

$$\chi_{r,s} = \begin{cases} 1\ , & eğer\ x_r = 0\ ve\ x_s = 0 \\ 2\ , & eğer\ x_r = 0\ ve\ x_s = 1 \\ 3\ , & eğer\ x_r = 1\ ve\ x_s = 0 \\ 4\ , & eğer\ x_r = 1\ ve\ x_s = 1 \end{cases} \quad (4.30)$$

İncelenen bir x_i pikseli için merkez-piksel ile komşu-piksel geçiş/değişim türünü belirten bir $\propto_i^j$ fonksiyonu tanımlanır. Bu $\propto_i^j$ fonksiyonu piksellerin komşuluk ilişkisini belirtmektedir. $\propto_i^j$ fonksiyonu,

$$\propto_i^j = \sum_{k=1}^{K} \delta(\chi_{i,k}\ , j), \qquad j = 1, \dots 4, \qquad K = 4 \quad (4.31)$$

şeklinde tanımlanmaktadır. Burada $\delta(m,n)$ değeri Kronecker delta fonksiyonudur ve

$$\delta(m,n) = \begin{cases} 1 & m = n \\ 0 & m \neq n \end{cases} \quad (4.32)$$

şeklinde ifade edilmektedir. Tüm resim üzerinden toplam uyuşma Eş. 4.33.'te ifade edildiği gibi hesaplanmaktadır.

$$a = \frac{1}{MN} \sum_i \propto_i^1,\ b = \frac{1}{MN} \sum_i \propto_i^2,\ \ c = \frac{1}{MN} \sum_i \propto_i^3\ ,\ \ D = \frac{1}{MN} \sum_i \propto_i^4 \quad (4.33)$$

Tanımlanan bu dört değişken $\{a, b, c, d\}$, bir ikili resimler için tek-adımlık eşoluşum (co-occurence) değerleri olarak ifade edilebilir. Yukarıdaki yapılan tanımlamalar kullanılarak birçok ikili resim benzerlik ölçütü Çizelge 4.1'de görüldüğü gibi tanımlanabilmektedir.

Kuzey Batı	Kuzey	Kuzey Doğu
Batı	Piksel	Doğu
Güney Batı	Güney	Güney Doğu

Şekil 4.8. Piksel komşuluğu

Bu tabloda dm_1'den dm_{10}'a kadar olan ölçütler resmin bir kanalının 7. ve 8. bit düzlemleri için elde edilmişlerdir. Elde edilen bu değerler ışığında üç tane benzerlik ölçüt kategorisi tanımlanmaktadır.

i. İlk grup, yedinci ve sekizinci bit düzlemleri için elde edilen benzerlik ölçütlerinin farkını içermektedir.

$$dm_i = m_i^{7th} - m_i^{8th} \ , \qquad i = 1, \dots\dots, 10 \tag{4.34}$$

ii. İkinci grup, histogramları ve entropik özellikleri içermektedir. Öncelikli olarak bit düzlemleri (yedinci ve sekizinci) uyuşmalarının histogramları normalize edilir.

$$p_j^b = \frac{\sum_i \alpha_i^j}{\sum_i \sum_j \alpha_i^j} \ , \qquad b = 7, 8 \tag{4.35}$$

b değeri bit düzlemlerini temsil eder. *Bu* normalize edilen 4-bin histogramlara dayanarak, minimum histogram farkı dm_{11}, mutlak histogram farkı dm_{12}, ikili karşılıklı entropi dm_{13} ve ikili Kullback Leibler uzaklığı dm_{14} tanımlanır. Bu ölçütlere ait eşitlikler Tablo 1'de verilmiştir.

iii. Üçüncü grup ise $dm_{15}, dm_{16}, dm_{17}\ ve\ dm_{18}$ ölçütlerinden oluşmaktadır ve diğer ölçütlere göre farklılık göstermektedir. Ojala tarafından önerilen komşuluk-ağırlık maskesi kullanılarak 512-seleli histogram hesaplanır. Sekiz yönlü ağırlıklı komşuluk Şekil 4.9'da gösterilmektedir. Hesaplama,

$$S = \sum_{i=0}^{7} x_i \, 2^i \tag{4.36}$$

kullanılarak yapılmaktadır. *N* histogramdaki sele sayısı olmak üzere, S_n^7 yedinci bit düzlemindeki, S_n^8 ise sekizinci bit düzlemindeki *n.* histogram sele sayısına karşılık gelmektedir. Bu 512-sele histogramları normalize

ettikten sonra Ojala minimum histogram farkı dm_{15}, Ojala mutlak histogram farkı ölçütü dm_{16}, Ojala ortak entropi dm_{17} ve Ojala Kullback-Leibler uzaklığı dm_{18} Çizelge 4.1'de [5] ifade edildiği gibi hesaplanmaktadır.

(a) (b)

Şekil 4.9. (a) Ojala değer hesaplamasındaki ağırlıklı komşuluk. (b) Örnek olarak, eğer E, N, NE bitleri 1 ve diğer bitler 0 ise değer S=2+4+8=14'tür.

Çizelge 4.1. İkili benzerlik ölçütleri

Benzerlik Ölçütü	**Açıklama**
Sokal & Sneath Benzerlik Ölçütü 1	$dm_1 = \frac{2(a+d)}{2(a+d)+b+c}$
Sokal & Sneath Benzerlik Ölçütü 2	$dm_2 = \frac{a}{a+2(b+c)}$
Kulczynski Benzerlik Ölçütü 1	$dm_3 = \frac{a}{b+c}$
Sokal & Sneath Benzerlik Ölçütü 3	$dm_4 = \frac{a+d}{b+c}$

Çizelge 4.1. devam ediyor

Benzerlik Ölçütü	**Açıklama**
Sokal & Sneath Benzerlik Ölçütü 4	$dm_5 = \frac{(a/(a+b))+(a/(a+c))+(d/(b+d))+(d/(c+d))}{4}$
Sokal & Sneath Benzerlik Ölçütü 5	$dm_6 \frac{ad}{\sqrt{(a+b)(a+c)(b+d)(c+d)}}$
Ochiai Benzerlik Ölçütü	$dm_7 = \sqrt{\left(\frac{a}{a+b}\right)\left(\frac{a}{a+c}\right)}$
Binary Lance & Williams Nonmetric Farklılık Ölçütü	$dm_8 = \frac{b+c}{2a+b+c}$
Örüntü Farkı (Pattern Difference)	$dm_9 = \frac{bc}{(a++c+d)^2}$
Varyans Farklılık Ölçütü	$dm_{10} = \frac{b+c}{4(a+b+c+d)}$
İkili Min Histogram Farkı	$dm_{11} = \sum_{n=1}^{4} min(p_n^7, p_n^8)$
İkili Mutlak Histogram Farkı	$dm_{12} = \sum_{n=1}^{4} \lvert p_n^7 - p_n^8 \rvert$
İkili Müşterek Entropi	$dm_{13} = -\sum_{n=1}^{4} p_n^7 \log p_n^8$

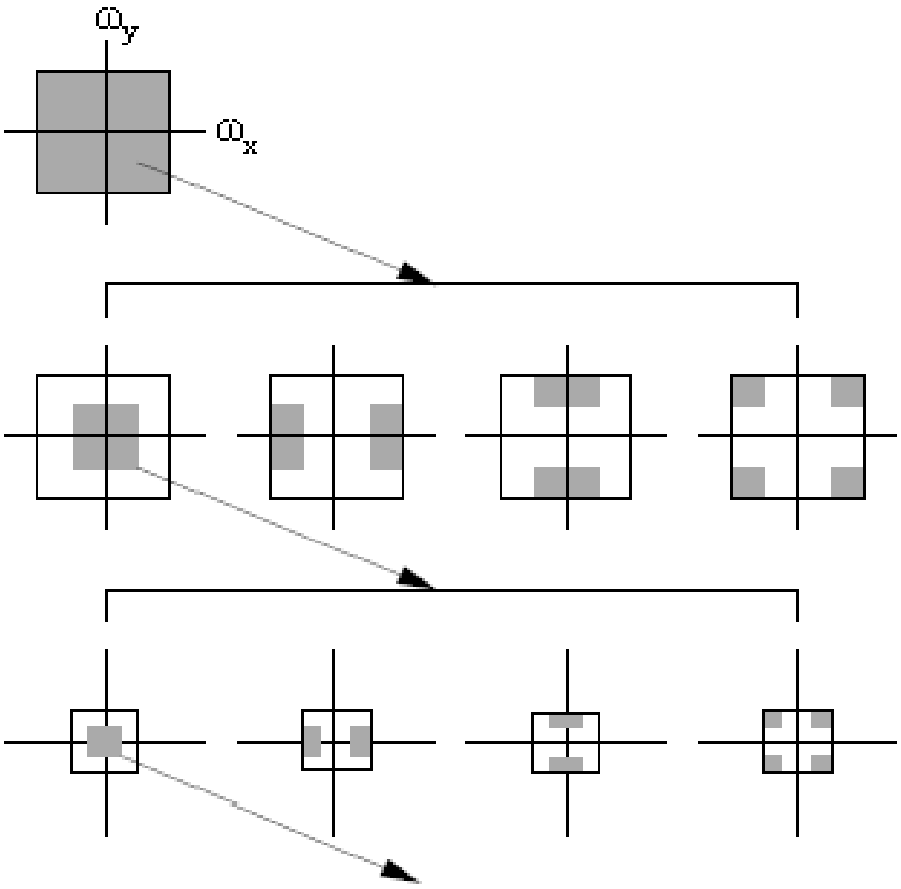
Çizelge 4.1. devam ediyor

Benzerlik Ölçütü	**Açıklama**
İkili Kullback-Leibler Farkı	$dm_{14} = -\sum_{n=1}^{4} p_n^7 \log \frac{p_n^7}{p_n^8}$
Ojala Min Histogram Farkı	$dm_{15} = \sum_{n=1}^{N} min(S_n^7, S_n^8)$
Ojala Mutlak Histogram Farkı	$dm_{16} = \sum_{n=1}^{N} \lvert S_n^7 - S_n^8 \rvert$
Ojala Müşterek Entropi	$dm_{17} = -\sum_{n=1}^{N} S_n^7 \log S_n^8$
Ojala Kullback-Leibler Farkı	$dm_{18} = -\sum_{n=1}^{N} S_n^7 \log \frac{S_n^7}{S_n^8}$

Bu ikili benzerlik ölçütleri resmin her renk kanalı için ayrı ayrı hesaplandıktan sonra toplamda 54 elemanlı bir özellik vektörü elde edilmiş olur. Bir DVM kullanılarak bir eğitici oluşturulur ve incelenen resmin sınıfı eğitici sonucuna göre karar verilir.

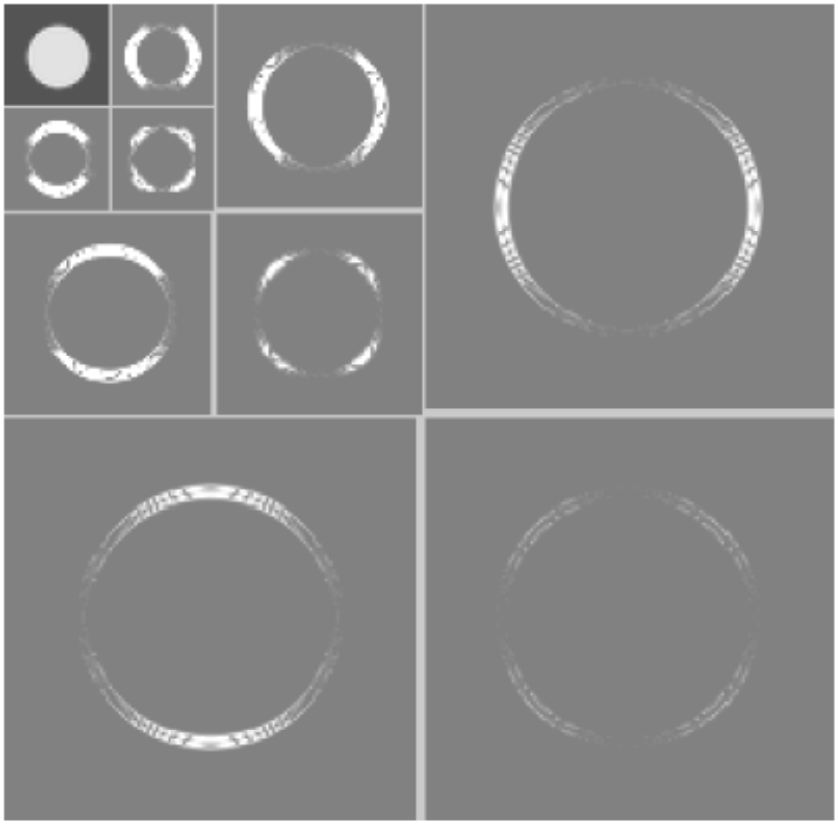
4.7. Yüksek Dereceli İstatistiklere Dayalı Steganaliz

Bu yaklaşımda resim içerisine bilgi gizlemenin resmin istatistiklerini bozduğu varsayılmıştır [22]. Uzamsal düzlemde, yönlü ve ölçeklenmiş dalgacık fonksiyonları kullanılarak resim ayrılması kullanımı, resim sıkıştırma, resim kodlama, gürültü temizleme gibi birçok alanda uygulama alanı bulmaktadır. Bu ayrıştırma işlemi resmin istatistiklerini iyi bir şekilde ortaya koymaktadır. Şekil 4.10'da böyle bir ayrıştırma işlemine ait şematik gösterim görülmektedir.



Şekil 4.10. Frekans uzayının çok ölçekli ve yönlü ayrıştırması. Üstten alta 0,1, 2 seviyeleri ve soldan sağa alçak geçiren, yatay, dikey ve diyagonal alt bantları.

Bu yöntemde uygulanan resim ayrıştırması dörtlü ayna filtrelerine (QMF) dayanmaktadır. Şekil 4.10'da gösterildiği gibi bu ayrıştırma frekans uzayını birçok ölçeğe ve yöne ayırır. Bu işlem resmin yatay dikey, diyagonal ve alçak geçiren alt bantları boyunca ayrı ayrı alçak ve yüksek geçiren filtreler uygulanmasıyla yapılır. Alt ölçekler alçak geçiren alt bandın tekrar tekrar uygulanması ile elde edilir. $i = 1, \ldots, n$'ye kadar olan yatay, dikey ve diyagonal alt bantları $V_i(x,y), H_i(x,y)$ ve $D_i(x,y)$ ile simgelenmektedir. Şekil 4.11'de "disc" resmine ait üç seviyeli ayrıştırma verilmiştir.



Şekil 4.11. “disc” resminin üç yönlü ve üç seviyeli alt bant katsayılarının mutlak değeri. Artık olan alçak geçiren alt bant sol üst köşede verilmiştir.

Bu ayrıştırma işlemi yapıldıktan sonra istatistiksel model $i = 1, \dots, n$ seviyelerinde ve her yöndeki alt bant katsayılarının ortalama, varyans, yamukluk ve savrukluk istatistiklerinin birleşimiyle oluşturulur. Bu istatistikler taban katsayı dağılımlarını karakterize eder. Toplanan ikinci istatistik seti katsayı büyüklüklerinin doğrusal tahmin edicilerinin hatalarına dayanmaktadır. Alt bant katsayıları bunların uzaysal/uzamsal, yönsel ve seviye komşularıyla ilintilidir. Bu ilintiyi gösterebilmek için, ilk olarak *i* yatay bandı $\mathrm{V_i(x,y)}$'yi ele alalım. Bütün olası komşulukların alt bantlarından elde edilen bu katsayıların büyüklüğü için tahmin edici Eş. 4.37'deki şekilde verilmektedir.

$$V_i(x,y) = w_1\, V_i(x-1,y) + w_2\, V_i(x+1,y) + w_3\, V_i(x,y-1) + w_4\, V_i(x,y+1) + w_5\, V_{i+1}(x/2\,,y/2) + w_6\, D_i\,(x,y) + w_7\, D_{i+1}(x/2\,,y/2) \quad (4.37)$$

Burada w_k değişkeni skaler ağırlıklandırma değerlerini belirtir. Bu doğrusal ilişki daha düzgün şekilde matris formunda,

$$\boldsymbol{V} = \boldsymbol{Q}\boldsymbol{w} \quad (4.38)$$

şeklinde gösterilmektedir. Burada $\boldsymbol{w} = (w_1, \ldots\ldots, w_7)^T$ sütun vektörüdür. $\boldsymbol{V}$ vektörü $V_i(x,y)$ katsayı büyüklüklerinin bir sütun vektörü şeklini, $\boldsymbol{Q}$ matrisinin sütunları ise yukarıdaki eşitlikte belirtilen komşu katsayıların büyüklük değerlerini içermektedir. Katsayılar karesel hata fonksiyonunun minimize edilmesiyle belirlenmiştir.

$$E(\boldsymbol{w}) = [\boldsymbol{V} - \boldsymbol{Q}\boldsymbol{w}]^2 \quad (4.39)$$

Bu hata fonksiyonu **w**'ya göre türev alınarak minimize edilir.

$$\frac{dE(\boldsymbol{w})}{d\boldsymbol{w}} = 2\boldsymbol{Q}^T[\boldsymbol{V} - \boldsymbol{Q}\boldsymbol{w}] \quad (4.40)$$

Yukarıdaki denklemde sonuç sıfıra eşitlenir ve **w** için çözülürse,

$$\boldsymbol{w} = (Q^T Q)^{-1} Q^T \boldsymbol{V} \quad (4.41)$$

elde edilir. Doğrusal tahmin edici oluşturulduktan sonra gerçek değerler ile tahmin edilen katsayılar arasındaki logaritmik hata şu şekilde hesaplanır:

$$\boldsymbol{E} = \log_2(\boldsymbol{V}) - \log_2(|\boldsymbol{Q}\boldsymbol{w}|) \quad (4.42)$$

Bu hata hesaplamasından ortalama, varyans, yamukluk ve savrukluk ekstra istatistikleri elde edilir. Bu işlem $i = 1, \ldots\ldots, n-1$ seviyelerindeki her bir dikey alt bantlar için tekrarlanır. Her seviye için yeni bir doğrusal tahmin edici oluşturulur. Benzer işlemler resmin yatay ve dikey alt bantları için de tekrar edilir. Yatay alt bantlar için doğrusal tahmin edici,

$$H_i(x,y) = w_1 H_i(x-1,y) + w_2 H_i(x+1,y) + w_3 H_i(x,y-1) + w_4 H_i(x,y+1) + w_5 H_{i+1}(x/2, y/2) + w_6 D_i(x,y) + w_7 D_{i+1}(x/2, y/2) \quad (4.43)$$

diyagonal alt bantlar için ise,

$$D_i(x,y) = w_1 D_i(x-1,y) + w_2 D_i(x+1,y) + w_3 D_i(x,y-1) + w_4 D_i(x,y+1) + w_5 D_{i+1}(x/2, y/2) + w_6 H_i(x,y) + w_7 V_i(x,y) \quad (4.44)$$

şeklindedir.

Dikey alt bant için ifade edilen aynı hata metrikleri ve hata istatistikleri yatay ve diyagonal alt bantlar için de hesaplanır. Sonuçta toplam olarak $12(n-1)$ hata istatistiği elde edilmiş olur. Burada *n* değeri kullanılmış olan seviyeyi ifade etmektedir. Bu elde edilen istatistikleri $12(n-1)$ katsayı istatistiği ile birleştirirsek toplam olarak $24(n-1)$adet özellik elde edilmiş olur. Bu elde edilen özellikler gizli bilgi içeren ve içermeyen resimlerin birbirlerinden ayırt edilmelerinde kullanılmaktadır.

Yukarıda bahsedilen bu özellik çıkartma işlemi resmin her renk kanalı için ayrı ayrı gerçekleştirilir. Eğitme aşamasında özellikleri çıkartırken 4 seviyeli ayrıştırma işlemi yapılmıştır. Her renk kanalı için $12 * (4-1)$ tane katsayılardan elde edilen özellik, $12 * (4-1)$ tane de hatalardan elde edilen özellik bulunur. 3 renk kanalı için toplamda $3 * 36 = 108$ tane katsayı özelliği ve $3 * 36 = 108$ tane hata özelliği yani toplamda renkli resimlerde kullanılacak 216 adet istatistiksel eğitim özelliği elde edilmiş olur. Elde edilen bu istatistikler Destek Vektör Makinesi-DVM kullanılarak bir model oluşturulur ve test edilen resimlerin sınıfının belirlenmesinde kullanılır.

5. STEGANALİZ UYGULAMASI VE DEĞERLENDİRME

Yapılan çalışmada bugüne kadar steganografinin tespit edilmesi için geliştirilmiş olan uygulamalar da incelenmiştir. Bu programlar çok fazla sayıda olmamakla beraber en bilindikleri, StegSpy[1], Stegdetect[2], StegBreak[3], Stego Suit[4] ve Stegoanalyzer[5]'dır. Bunlardan StegSpy ve Stegdetect programları açık kaynak kodlu uygulamalar olup sadece belirli programların üretmiş oldukları stego resimleri tespit edebilmektedir. StegSpy JP Hide and Seek, Invisible Secrets, Hiderman, JPegX, Masker steganografi programlarını, Stegdetect ise Jsteg, jpHide, InvisibleSecrets, Outguess, F5, Camouflage adlı steganografi programlarını tespit edebilmektedir. StegBreak yazılımı ise steganografinin tespit edilmesine yönelik bir program değildir. Daha çok steganografi uygulanmış ise içerisindeki mesajın çıkartılmasına yönelik bir uygulamadır. Hedef aldığı steganografi programları ise JstegShell, JPHide ve Outguess'dir.

Stego Suit Wetstone firmasına ait bir yazılımdır. Program 4 modülden oluşmaktadır. Bunlar

- Stego Hunter - steganografi programı tespiti,
- Stego Watch - steganografi tespit aracı,
- Stego Analyst - görüntü ve ses dosyası analizcisi,
- Stego Break - Steganografi şifresi kırmadır.

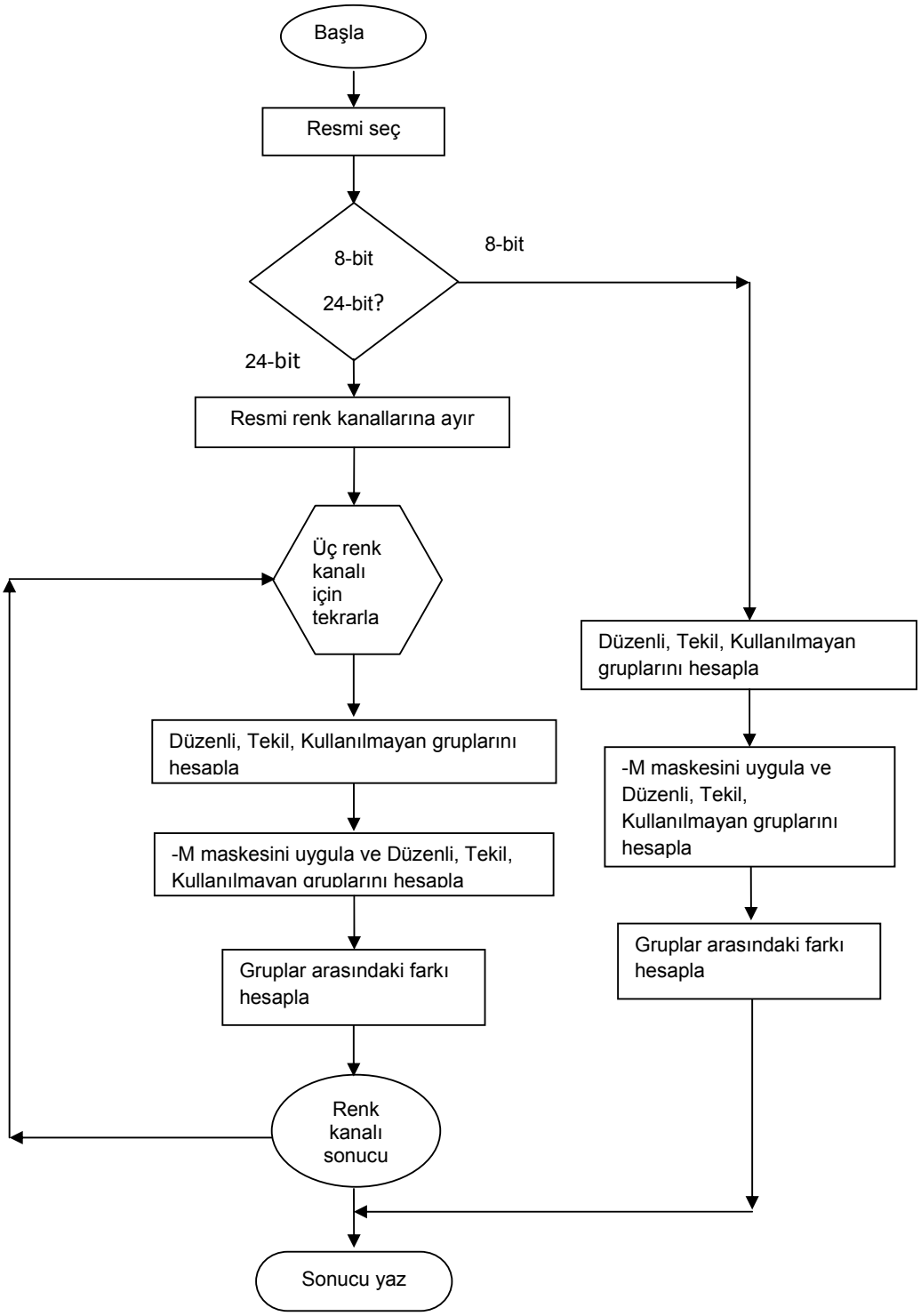
Steganalyzer ise Backbone Security firmasına ait bir yazılımdır. Bu iki üründe lisanslı olarak satılan uygulamalardır. İnternet siteleri üzerinden demo uygulamaları bulunmamaktadır. Programların çok yüksek doğrulukla gizli bilgiyi tespit edebildiği ifade edilmektedir. Ancak kullanmış olduğu yöntem ve yaklaşımlar hakkında bilgi bulunamamıştır. Bu programların daha çok ülkelerin güvenlik güçlerinin kullanımına yönelik olarak geliştirildiği ifade edilmektedir.

[1] http://www.spy-hunter.com/stegspydownload.htm
[2] http://www.outguess.org/detection.php
[3] http://www.digipedia.pl/man/doc/view/stegbreak.1/
[4] http://www.wetstonetech.com
[5] http://www.sarc-wv.com/

Bu çalışmada daha önceki bölümlerde açıklanmaya çalışılan steganaliz yöntemleri C# programlama dilinde Visual Studio.NET geliştirme ortamı kullanılarak bir uygulama geliştirilmiştir. İnternet üzerinde yaygın olarak kullanılan steganografi programları incelendiğinde büyük çoğunluğunun LSB steganografi kullandığı tespit edilmiştir. Bu yüzden öncelikle LSB steganografinin tespit edilmesinde başarılı olarak değerlendirilen ve farklı yaklaşımları kullanan daha önceki kısımda da açıklanan RS Steganaliz, RQP Steganaliz ve Resim Düzgünlüğüne Dayanan Steganaliz yöntemlerini içeren bir uygulama geliştirilmiştir. Bunların yanı sıra değişken eşik değerine sahip yakın renk çifti analizi de bu uygulamaya eklenmiştir.

Geliştirilen uygulamayı test etmek üzere farklı kategorilerde test resimleri seçilmiştir. Resimlerin maksimum mesaj saklama kapasitelerinin %10, %30, %50, %70 ve %100 oranlarında test mesajları oluşturularak LSB yöntemini kullanan Hermetic Stego programı aracılığıyla bu resimlere ait stego resimler elde edilmiştir. Mesajların resim içerisinde rastsal olarak dağılmasını sağlamak için oluşturulan test mesajları şifreledikten sonra saklama işlemi gerçekleştirilmiştir.

Uygulamada gerçekleştirilen ilk yöntem olan RS steganaliz yöntemine ait akış diyagramı Şekil 5.1'de verilmiştir. Yöntemde resmin Düzenli (R) ve Tekil (S) grup sayıları arasındaki fark dikkate alınır ve karar vermede rol oynar. R ve S değerleri arasındaki farkın 0'a yakın çıkması resim içerisinde saklanmış bilgi olmadığına işaret etmektedir. Geliştirilen uygulamanın 3 test resminin 5 farklı boyutta mesaj içeren stego versiyonları için elde edilen sonuçlar Çizelge 5.1, 5.2 ve 5.3'te verilmiştir. Çizelgelerde her renk kanalı için Düzenli-Tekil (R-S) farkı ayrı ayrı gösterilmiştir. Sonuçlar genel olarak değerlendirildiğinde R-S farkının yüzler hatta binler seviyesine çıkmakta olduğu görülmüştür. Farkın yüzden büyük olması stego resim olarak sınıflandırmada eşik değeri olarak kullanılmıştır. Çizelge 5.1'deki %70 ve % 50 mesaj oranlarında kırmızı renk kanalındaki değerler onlar seviyesinde bulunmasına rağmen elde edilen sonuçlar [16] ve [32]'deki sonuçlar ile karşılaştırıldığında benzer değerler elde edildiği görülmüştür.



Şekil 5.1. RS Steganaliz yöntemine ait akış diyagramı

Çizelge 5.1. Manmade.bmp resmi için RS steganaliz sonuçları

		Mesaj Oranı				
Renk Kanalı		**%10**	**%30**	**%50**	**%70**	**%100**
Kırmızı	**RS Farkı**	282	249	77	42	273
	Sonuç	Stego	Stego	Normal	Normal	Stego
Mavi	**RS Farkı**	780	257	152	145	1478
	Sonuç	Stego	Stego	Stego	Stego	Stego
Yeşil	**RS Farkı**	688	254	313	304	970
	Sonuç	Stego	Stego	Stego	Stego	Stego

Çizelge 5.2. Foliage.bmp resmi için RS steganaliz sonuçları

		Mesaj Oranı				
Renk Kanalı		**%10**	**%30**	**%50**	**%70**	**%100**
Kırmızı	**RS Farkı**	19	401	124	503	75
	Sonuç	Normal	Stego	Stego	Stego	Normal
Mavi	**RS Farkı**	672	729	132	150	298
	Sonuç	Stego	Stego	Stego	Stego	Stego
Yeşil	**RS Farkı**	115	114	489	201	829
	Sonuç	Stego	Stego	Stego	Stego	Stego

Çizelge 5.3. Animal.bmp resmi için RS steganaliz sonuçları

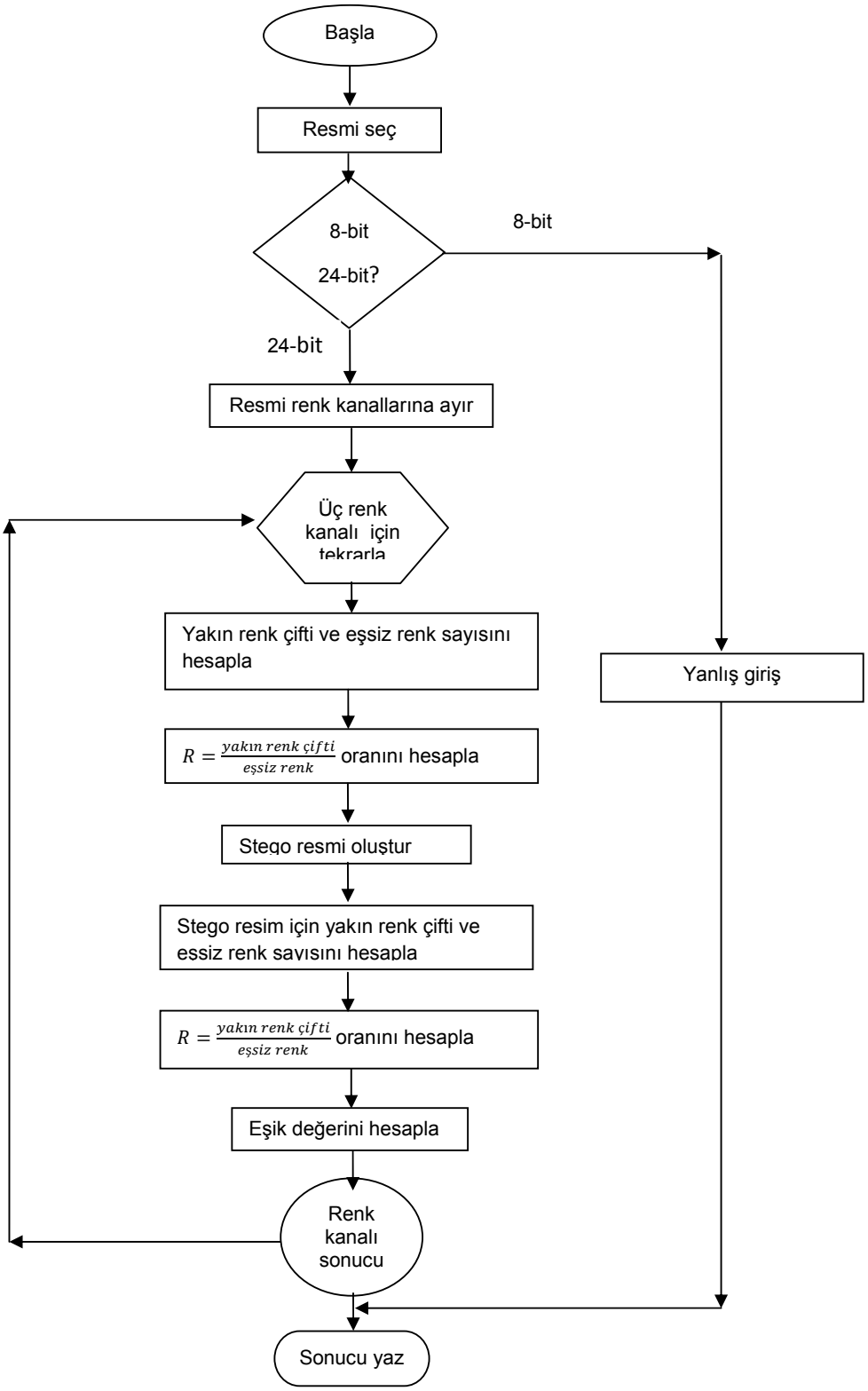
		Mesaj Oranı				
Renk Kanalı		**%10**	**%30**	**%50**	**%70**	**%100**
Kırmızı	**RS Farkı**	373	602	279	820	795
	Sonuç	Stego	Stego	Stego	Stego	Stego
Mavi	**RS Farkı**	173	36	437	1151	1326
	Sonuç	Stego	Normal	Stego	Stego	Stego
Yeşil	**RS Farkı**	468	650	1004	913	1150
	Sonuç	Stego	Stego	Stego	Stego	Stego

İkinci yöntem olarak uygulanan Yakın Renk Çifti ve Değişken Eşik Değerli Yakın Renk Çifti yöntemleri 24 bit renkli resimler üzerinde çalışmaktadır. Yöntemler dayandıkları mantık olarak birbirleriyle aynıdır sadece karar verme aşamasında farklılık göstermektedirler. Bu yüzden yöntemlerin uygulamada kullanılan akış diyagramları birleştirilerek Şekil 5.2'de verilmiştir. Yöntemde seçilen resmin yakın renk çiftlerinin sayısının tüm renk çiftlerinin sayısına oranı (R) bulunur. Resim içerisine bir test mesajı gizlenir ve bu oran (R') tekrar hesaplanır. Bu hesaplanan iki oran arasındaki farkın büyük olması resim içerisinde bilgi olmadığına işaret etmektedir. Oranların birbirine yakın olması ise resmin saklanmış mesaj içerdiğini göstermektedir. Yakın Renk Çifti yöntemi sabit bir eşik değeri kullanmaktadır. [15]'te R'/R oranın için eşik değeri olarak 1,1 değeri alınmıştır. Geliştirilen uygulamada bu değer kullanılmakla birlikte Değişken Eşik Değerli Yakın Renk Çifti yönteminde kullanılan eşik değeri hesaplaması da kullanılmıştır. Değişken Eşik Değerli yöntemde,

$$m = \frac{(R-R')}{R*100} \quad \text{ve} \quad t = \frac{(U'-P')}{d}$$

değerleri hesaplanarak karar verilmektedir. *m* değerinin *t*'den küçük olması resim içerisinde bilgi olduğu anlamına gelmektedir.

Çizelge 5.4, 5.5 ve 5.6'da bu iki yöntem için geliştirilen uygulamadan elde edilen sonuçlar sunulmuştur. R ve R' için elde edilen sonuçlar [15] ve [33]'teki sonuçlar ile karşılaştırıldığında yakın sonuçlara ulaşıldığı görülmüştür. Çizelgelerden de görüldüğü gibi 1,1 sabit eşik değerini kullanmak, yerleştirilen her mesaj oranı için doğru sonuçlara neden olmamıştır. Bu değerin 0,9 ya da 1,0 olarak kullanılmasının daha sağlıklı olacağı düşünülmektedir. Değişken eşik değeri hesaplaması ile sabit eşik değeri kullanılması ile elde edilen bazı sonuçların önüne geçilebileceği görülmüştür. Çizelge 5.5'te böyle bir örnek görülmektedir. Ancak sadece değişken eşik değeri kullanmanın da yeterli olmadığı Çizelge 5.6'da %10 mesaj oranında elde edilen sonuç incelendiğinde görülmüştür. Burada sabit eşik değeri ile içerisinde bilgi içeriyor olarak nitelendirilen resim, değişken eşik değeri ile normal resim olarak sınıflandırılmıştır.



Şekil 5.2. Yakın renk çifti ve değişken eşik değerli yakın renk çifti yöntemleri için ortak akış diyagramı

Çizelge 5.4. Manmade.bmp resmi için yakın renk çifti analizi sonuçları

		Yerleştirilen Mesaj Oranı				
	Orijinal	**%10**	**%30**	**%50**	**%70**	**%100**
R	0,04438	0,03323	0,03437	0,03399	0,03433	0,03732
R'	0,037335	0,03328	0,0338	0,03373	0,03394	0,03518
R'/R	0,8412	1,0015	1,1172	0,9923	0,9886	0,9426
Sabit Eşik Değeri Sonuç	Normal	Stego	Stego	Stego	Stego	Stego
m	15,870	15,046	1,0156	0,764	1,136	5,734
t	34,646	14,578	3,575	3,456	8,539	12,582
Değişken Eşik Değeri Sonuç	Stego	Normal	Stego	Stego	Stego	Stego

Çizelge 5.5. Foliage.bmp resmi için yakın renk çifti analizi sonuçları

		Yerleştirilen Mesaj Oranı				
	Orijinal	**%10**	**%30**	**%50**	**%70**	**%100**
R	0,00478	0,00411	0,00427	0,00431	0,00413	0,00458
R'	0,00383	0,00366	0,00369	0,00368	0,00354	0,00378
R'/R	0,802	0,890	0,864	0,858	0,857	0,825
Sabit Eşik Değeri Sonuç	Normal	Normal	Normal	Normal	Normal	Normal
m	19,782	10,795	13,556	14,585	14,323	17,444
t	31,272	29,333	13,142	13,608	21,252	19,056
Değişken Eşik Değeri Sonuç	Stego	Stego	Normal	Normal	Stego	Stego

Çizelge 5.6. Animal.bmp resmi için yakın renk çifti analizi sonuçları

		Yerleştirilen Mesaj Oranı				
	Orijinal	**%10**	**%30**	**%50**	**%70**	**%100**
R	0,26860	0,17084	0,18352	0,18701	0,19280	0,20818
R'	0,23362	0,19053	0,19716	0,19901	0,20007	0,20866
R'/R	0,8697	1,1152	1,0743	1,0641	1,0377	1,0024
Sabit Eşik Değeri Sonuç	Normal	Stego	Stego	Stego	Stego	Stego
m	13,023	11,525	7,432	6,416	3,770	0,230
t	47,015	8,324	7,890	7,869	1,357	3,584
Değişken Eşik Değeri Sonuç	Stego	Normal	Stego	Stego	Normal	Stego

Geliştirilen uygulamaya son LSB yöntemi olarak Resim Düzgünlüğüne Dayalı Steganaliz yöntemi eklenmiştir. Yönteme ait akış diyagramı Şekil 5.3'te gösterilmektedir. Resim önce renk kanallarına ayrılır ve ardından her renk kanalı için sırasıyla resmin tüm pikselleri taranarak piksel komşulukları elde edilerek düzgünlük değeri hesaplanır. Resmin LSB düzleminde 1'ler 0, 0'lar 1 yapılır. Tekrar elde edilen değiştirilmiş resim için resim düzgünlüğü hesaplanır ve elde edilen bu iki düzgünlük değerinin farkı alınarak resimde saklı mesaj uzunluğunun tahmini değeri bulunmaya çalışılır.

Bu yönteme ait test resimleri ve bu resimlerin farklı büyüklükte mesaj içeren stego halleri için geliştirilen uygulamadan elde edilen sonuçlar Çizelge 5.7, 5.8 ve 5.9'da verilmiştir. Çizelgelerde her renk kanalına ait resim içerisine saklanmış olan mesaj uzunluğunun, steganaliz yöntemi ile tahmin edilen değerleri görülmektedir. Orijinal resim için mesaj uzunlukları negatif değer olarak elde edilmiştir. Bu sonuçlar [27] ile karşılaştırıldığında normal olarak gözükmektedir. Mesaj oranlarının sonuçları incelendiğinde maksimum mesaj kapasitesinin yarısı civarında mesaj saklandığı durumlarda yöntemin gerçek mesaj oranına yakın tahminlerde bulunduğu gözlemlenmiştir. Ancak saklanmış olan mesaj oranı %30 ya da daha az olduğu durumlarda ve maksimum mesaj saklama kapasitesine yakın oranlarda doğru olmayan sonuçlar elde edilmiştir. Sonuçlar incelendiğinde

mesaj oranlarının yüksek doğrulukla tahmin edilemediği görülmüştür. Ancak orijinal resim her zaman negatif değerler elde edildiğinden kriter olarak bu sonucun kullanılıp, negatif ise normal değilse stego diye sınıflandırma yapılmasının faydalı olacağı düşünülmektedir.

Çizelge 5.7. Manmade.bmp resmi için resim düzgünlüğü yöntemi mesaj oranı sonuçları

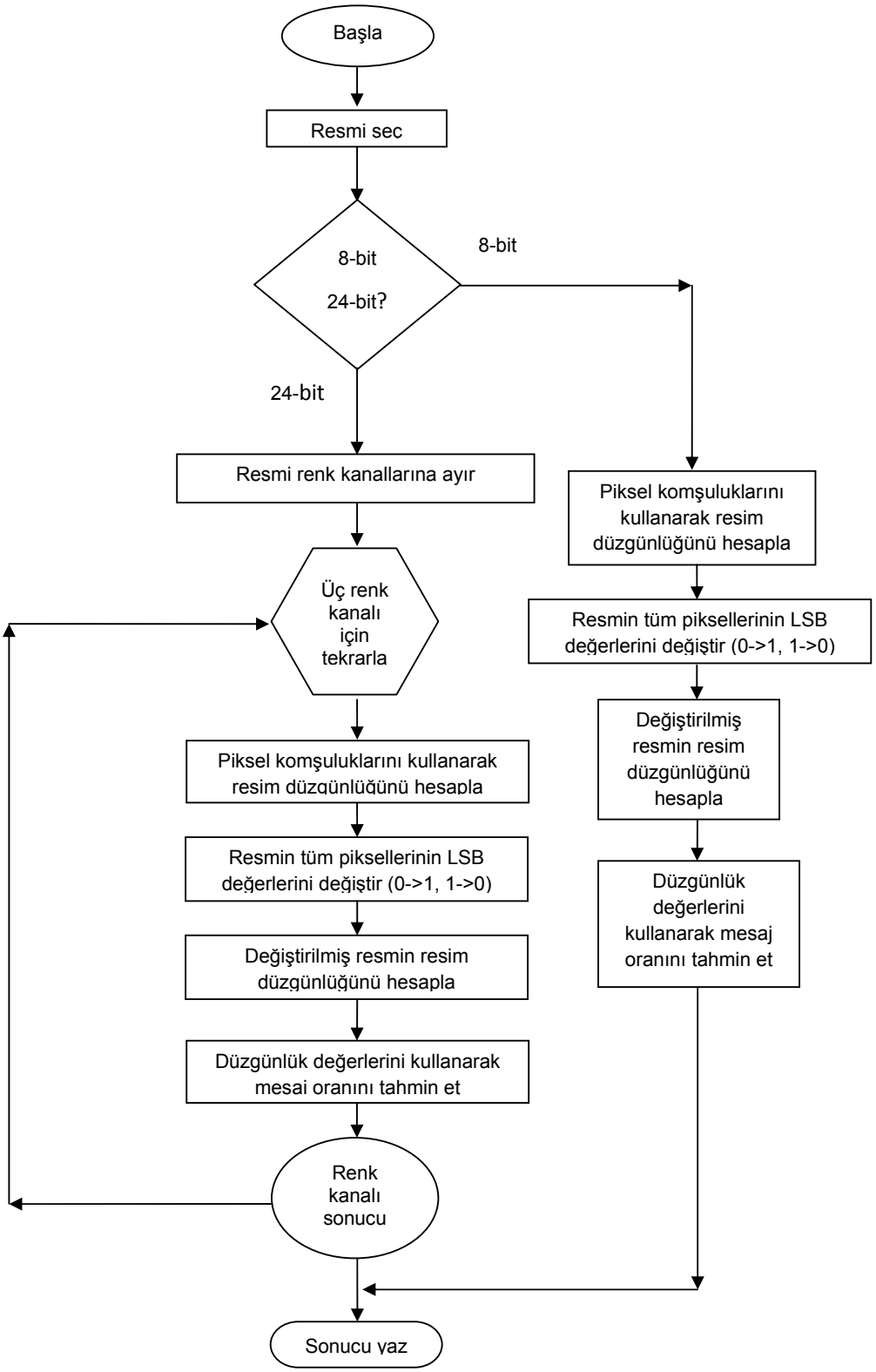
		Mesaj Oranı				
Renk Kanalı	**Orijinal**	**%10**	**%30**	**%50**	**%70**	**%100**
Kırmızı	-12,481	0,25	18,93	52,65	81,35	68,50
Mavi	-11,918	1,35	9,16	48,51	72,75	42,02
Yeşil	-12,450	1,34	1,19	31,21	78,01	70,82

Çizelge 5.8. Foliage.bmp resmi için Resim Düzgünlüğü mesaj oranı sonuçları

		Mesaj Oranı				
Renk Kanalı	**Orijinal**	**%10**	**%30**	**%50**	**%70**	**%100**
Kırmızı	-15,870	1,97	3,08	50,02	62,17	57,03
Mavi	-11,493	1,31	5,18	48,18	89,51	45,35
Yeşil	-10,716	1,23	1,09	44,59	68,49	57,44

Çizelge 5.9. Animal.bmp resmi için Resim Düzgünlüğü mesaj oranı sonuçları

		Mesaj Oranı				
Renk Kanalı	**Orijinal**	**%10**	**%30**	**%50**	**%70**	**%100**
Kırmızı	-17,339	11,70	2,25	43,79	79,30	52,52
Mavi	-6,510	3,75	9,99	44,56	70,46	69,22
Yeşil	-6,700	1,43	16,83	41,38	59,31	69,52



Şekil 5.3. Resim düzgünlüğüne bağlı steganaliz yöntemi için akış diyagramı

Kullanılmış olan bu yöntemler LSB steganografi uygulamalarının tespit edilmesinde başarılı olarak değerlendirilmektedir. Ancak yapılan incelemelerde LSB steganografinin yanı sıra dönüşüm tabanlı tekniklerinde kullanıldığı görülmüştür. İlaveten teknolojinin gelişmesiyle birlikte yeni steganografi teknikleri ortaya çıkmaktadır. Her çıkan yeni yaklaşıma karşı bir analiz uygulamasının yorucu ve yetersiz kalacağı düşünüldüğünde kör steganaliz olarak değerlendirilen ve yöntemden bağımsız, resim üzerindeki geleneksel istatistiklere dayanan analiz uygulamalarının kullanılmasının faydalı olacağı düşünülmüştür. Bu kapsamda önceki kısımda açılanan İkili Benzerlik Ölçütü Steganaliz, Yüksek Dereceli İstatistik Steganaliz ve Run-Uzunluğu Steganaliz yöntemleri de uygulamaya dahil edilmiştir.

Kör steganaliz yöntemlerinde karar verme mekanizması olarak Destek Vektör Makinesi (DVM) kullanılmaktadır. Geliştirilen uygulamada DVM olarak açık kaynak kodlu libsvm[1] aracı kullanılmıştır. Bu aracın seçilmesinin nedeni açık kaynak kodlu olması ve pratik uygulamalarda kolay uygulanabilir olmasıdır. DVM'nin kullanılmadan, uygun özellikler çıkartılarak eğitilmesi gerekmektedir. DVM'nin eğitilmesinin ardından üç algoritma için üç model dosyası oluşturulmuş ve bu model dosyalarına dayanarak resmin hangi sınıfa ait olduğu belirlenmiştir.

Geliştirilen uygulamada DVM, her analiz yöntemi için ayrı ayrı eğitilmiştir. Eğitim için yaklaşık 2500 adet resim çeşitli internet sitelerinden[2] indirilmiştir. Oluşturulan resim veritabanı farklı tür ve tipteki resim dosyalarından meydana gelmektedir. Kullanılan resimler özellikle yüz, nesne, manzara gibi farklı tiplerden seçilerek çeşitlilik arttırılmış ve DVM'nin daha kapsamlı olarak eğitilmesi amaçlanmıştır. Veritabanında kullanılan resimlerden birkaçı Şekil 5.4'te görülmektedir. Bu orijinal resimlerden günümüzde kullanımı daha fazla olan 5 adet steganografi uygulaması ile farklı boyutlarda gizli mesajlar saklanarak stego resimler elde edilmiştir.

[1] Açık kaynak kodlu DVM kütüphanesi, http://www.csie.ntu. edu.tw/~cjlin/libsvm/

[2] http://www.vision.caltech.edu/,
http://www.cs.columbia.edu/,
http://www.cs.cmu.edu,
www.kodak.com
www.cipr.rpi.edu

Farklı boyutlarda seçilen JPG ve BMP biçimlerindeki resimleri için her steganografi programı ve her resim boyutu için maksimum mesaj saklama kapasitesi hesaplanmış, bu kapasitenin %20, %50, %80 ve %100 oranlarında test mesajları hazırlanmıştır. Eğitim sırasında kullanılan Outguess, F5, Stegano, StegHide ve JStegShell programları ile hazırlanan bu test mesajları resimlerin içerisine yerleştirilerek stego resimler elde edilmiştir. Toplamda $2500\ x\ 4\ mesaj\ boyutu\ x\ 5\ steganografi\ programı = 50.000$ adet resimden oluşan bu veritabanı ile DVM eğitilmiş ve karar vermede kullanılacak olan model dosyaları oluşturulmuştur.

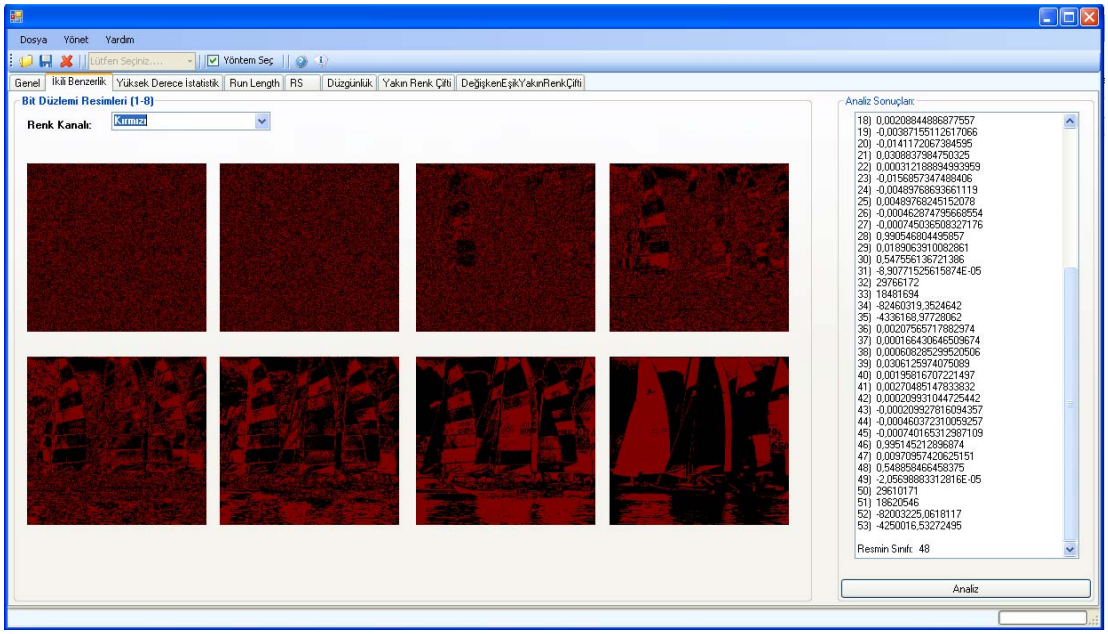
Şekil 5.4. Eğitim veritabanında kullanılan resimlerden bazıları

Uygulanan kör steganaliz yöntemlerini test etmek için 10 adet test resmi seçilmiştir. Bu test resimleri internet sitelerinden indirilmiş olan resimler arasından seçilmiş ve bu resimler eğitim veritabanının oluşturulmasında kullanılmamıştır.

Test resimlerine, yöntemlerin farklı boyutlardaki mesajları tespit edip edemediğini belirlemek amaçlı maksimum mesaj saklama kapasitelerinin %10, %30, %50, %70 ve %100 oranlarında mesaj saklanmıştır. Test mesajlarını resimlerin içerisine saklarken stego veritabanı oluşturulurken kullanılan Outguess, F5, Stegano ve StegHide steganografi programları kullanılmıştır. Bu 4 steganografi programının yanı sıra eğitim sırasında kullanılmayan Invisible Secrets programından da test mesajlarının oluşturulması sırasında faydalanılmıştır. Bu program ile yine aynı oranlarda test mesajları seçilen resimlerin içerisine yerleştirilmiştir. Toplamda $10\ x\ 5\ mesaj\ boyutu\ x\ 5\ steganografi\ programı = 250$ resimden oluşan bir test veritabanı oluşturulmuştur.

Kör steganaliz yöntemlerinde eğitim ve test veritabanları oluşturulurken sadece 24 bitlik renkli resimler kullanılmıştır. Günümüzde 8 bitlik gri seviyeli resimlerin sadece özel uygulamalarda kullanılmasından ve gizli haberleşmede gri seviyeli resim kullanımının dikkat çekici olacağından böyle bir seçim yapılmıştır. Ayrıca bu seçimde, ayrı bir veritabanı oluşturulup algoritmaların eğitilmesinin işlem gücü ve zaman açısından yüksek maliyetli oluşu etkili olmuştur.

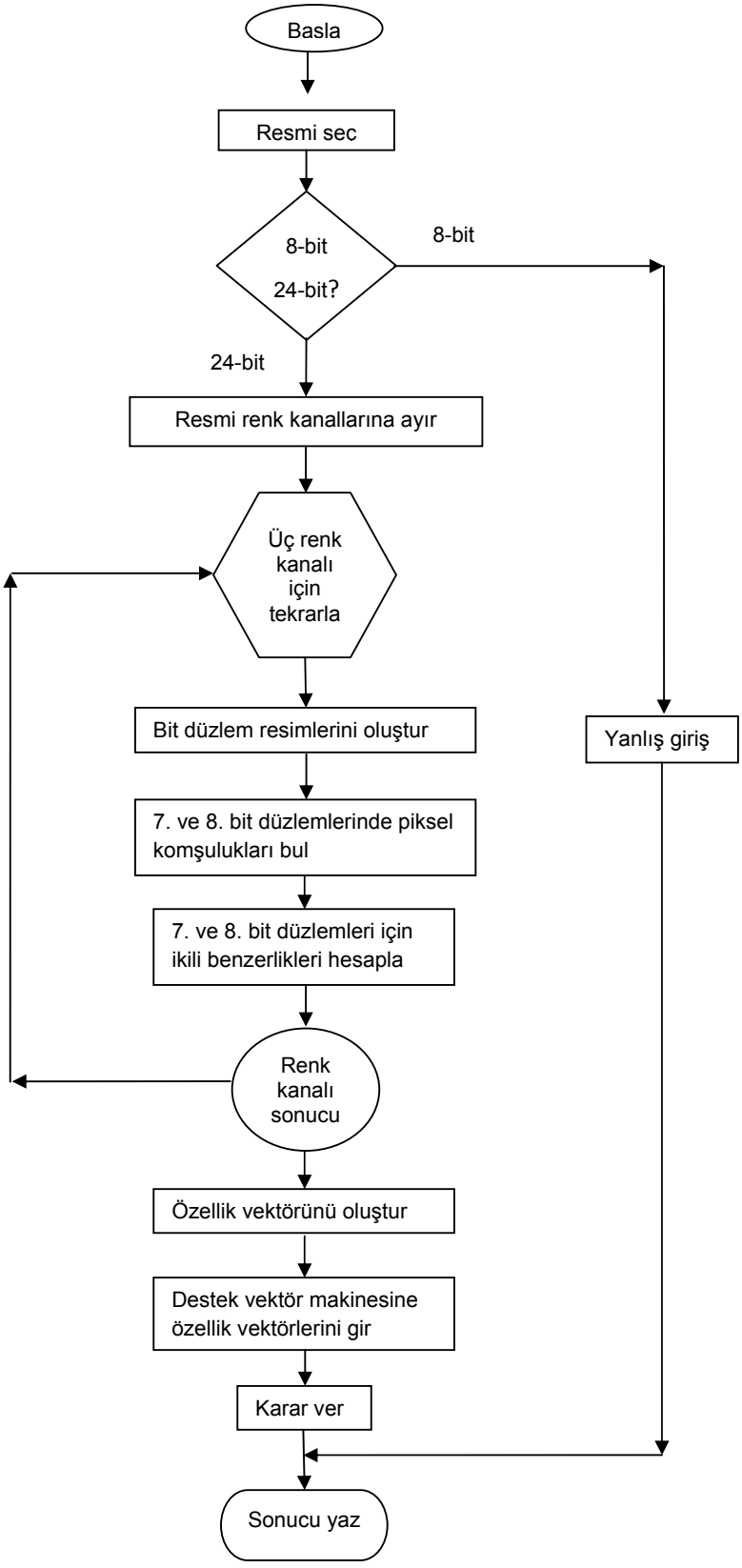
Kör steganaliz yöntemlerinden ilk uygulanan İkili Benzerlik Ölçütlerine Dayalı Steganaliz yöntemidir. Yöntemin akış diyagramı Şekil 5.6'da verilmiştir. Yöntem 8 bitlik resmin bit düzlemlerine ayrılıp birbiriyle ardışık gelen iki bit düzlemi arasındaki ilişkiye dayanmaktadır. 24 bitlik renkli resim kullanılması durumunda resim önce renk kanallarına ayrılarak 3 tane 8 bitlik resim elde edilerek analize devam edilir. Ardışık iki bit düzlemi olarak yedinci ve sekizinci bit düzlemleri seçilmiştir. Geliştirilen uygulamaya eklenen bit düzlem resimlerini gösteren form ile düzlemler arasındaki ilişki daha iyi görülmektedir. Bu forma ait bir ekran görüntüsü Şekil 5.5'te verilmiştir. Bu şekilde resmin kırmızı renk kanalı ait 1'den 8'e kadar olan düzlem resimleri görülmektedir.



Şekil 5.5. Bit Düzlem resimlerini gösteren ekran görüntüsü (Yat.jpg resminin kırmızı renk kanalına ait bit düzlem resimleri, 1'den 8'e)

Seçilen bit düzlemleri üzerinden hesaplanan ikili benzerlik ölçütleri ile üç renk kanalı için toplam 54 değerden oluşan bir özellik seti elde edilmiştir. Yaklaşık 50.000 resimden meydana gelen normal ve stego resimlerden oluşan eğitim veritabanı kullanılarak bütün resimlerin özellikleri çıkartılmıştır. Elde edilen bu özellik setleri DVM'yi eğitmek için kullanılmış ve sınıflandırıcıyı meydana getiren model dosyası oluşturulmuştur.

İkili benzerlik ölçütüne dayalı steganaliz yöntemi için test mesajları için elde edilen sonuçlar Çizelge 5.10'da verilmiştir. Çizelgede toplamda kullanılan 250 test resmi için elde edilen sonuçlar verilmiştir. Buradaki yüzdeler doğru tespit edilme oranını göstermektedir. Elde edilen sonuçlar [10]'da verilen sonuçlar ile karşılaştırıldığında daha yüksek oranlarda başarı elde edildiği görülmüştür. Bu başarıda kullanılan eğitim veri tabanının çok sayıda ve farklı boyutlarda mesajlar içeren stego resimlerden meydana gelmesinin etkili olduğu düşünülmektedir.



Şekil 5.6. İkili benzerlik ölçütüne dayalı steganaliz yöntemi için akış diyagramı

Çizelge 5.10. İkili benzerlik ölçütüne dayalı steganaliz yöntemi için sonuçlar

		Mesaj Yüzdesi				
Steganografi Programı	**Kullandığı Yöntem**	**%10**	**%30**	**%50**	**%70**	**%100**
Invisible Secrets	LSB	%90	%100	%100	%100	%100
Outguess	Dönüşüm	%100	%100	%100	%90	%100
F5	Dönüşüm	%90	%100	%100	%100	%100
Stegano	LSB	%100	%80	%100	%90	%100
StegHide	Dönüşüm	%100	%100	%100	%100	%100

İkinci kör steganaliz yöntemi olarak Yüksek Dereceli İstatistiklere Dayalı Steganaliz yöntemi geliştirilen uygulamaya eklenmiştir. Bu yöntemin akış diyagramı Şekil 5.7'de görülmektedir. Yöntemde resim renk kanallarına ayrılarak her renk kanalı için 4 seviyeli dalgacık dönüşümü uygulanır ve 3 yönde (yatay, dikey, diyagonal) dalgacık katsayıları elde edilir. Daha sonra bu dalgacık katsayıları kullanılarak yine her seviye ve her yönde hata katsayıları elde edilir. Elde edilen bu dalgacık ve hata katsayılarının dördüncü seviye istatistikleri özellik olarak kullanılmaktadır. Uygulamada dalgacık dönüşümünde Bior Dalgacık filtresi kullanılmıştır.

Geliştirilen uygulamaya resmin dalgacık katsayılarını gösteren bir form eklenmiştir. Bu forma ait ekran görüntüsü Şekil 5.8'de verilmiştir. Yöntemde toplamda her renk kanalı için 72 özellik, tüm resim için 216 özellik elde edilmektedir. Oluşturulan eğitim veritabanındaki resimler için bu 216 özellik seti hesaplanarak steganaliz yöntemi için sınıflandırıcı oluşturulmuştur. Uygulama hazırlanan veritabanı üzerinde test edilmiş ve elde edilen sonuçlar Çizelge 5.11'de sunulmuştur.

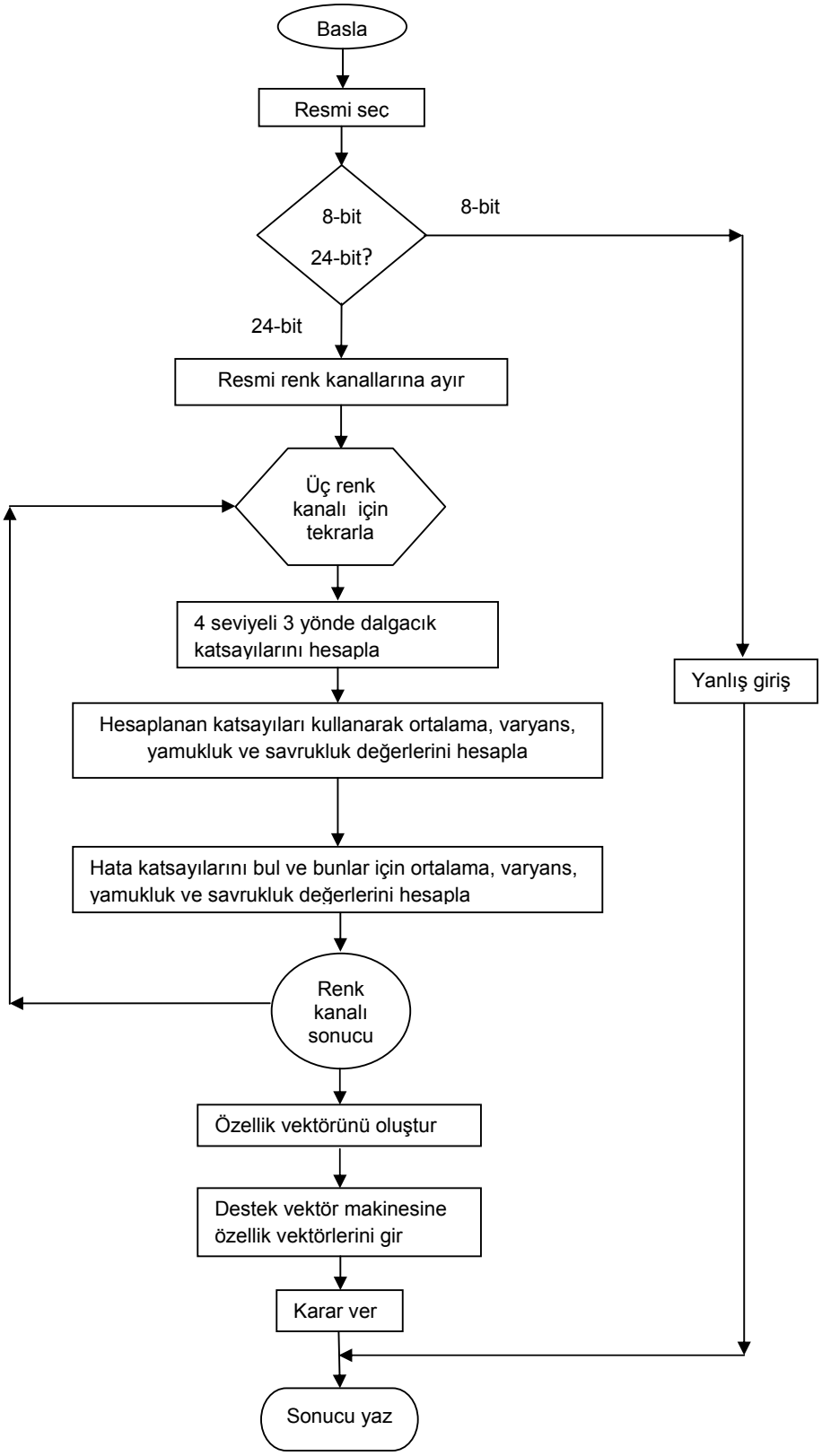
Test sonuçları incelendiğinde eğitim veritabanı oluşturulurken kullanılmış olan F5, Outguess, Stegano ve StegHide programları kullanılarak elde edilmiş resimlerde yöntemin %90 ve %100 seviyelerinde başarılı olduğu görülmektedir. Ancak eğitimde kullanılmayan Invisible Secret programıyla oluşturulan stego resimler için

aynı durum söz konusu değildir. Bu resimlerde mesaj oranı %30 ve %50 oranlarda olan stego resimlerde başarılı (%80) olarak nitelendirebileceğimiz sonuçlar elde edilmiştir. Fakat mesaj oranı %80 ve %100 olduğu durumlarda tespit etme oranı düşmekte ve hatta mesaj oranı %10 olduğu durum için bu değer 0 olmaktadır.

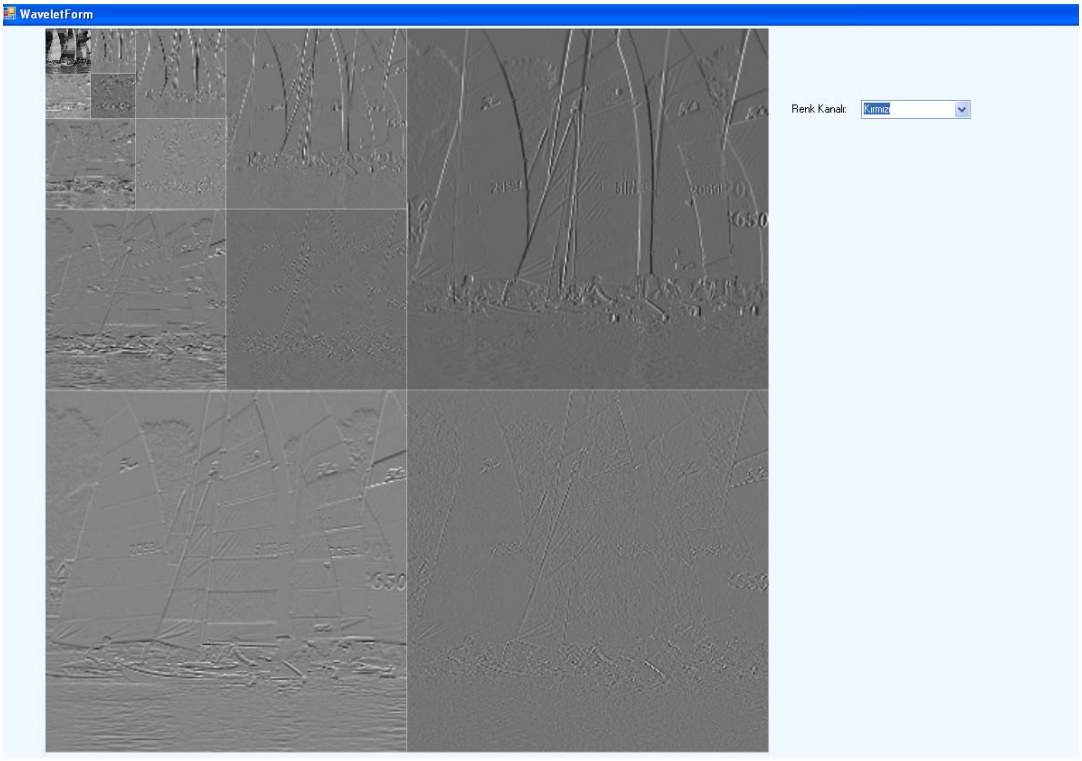
Çizelge 5.11. Yüksek dereceli istatistiklere dayalı steganaliz yöntemi için test sonuçları

		Mesaj Yüzdesi				
Steganografi Programı	**Kullandığı Yöntem**	**%10**	**%30**	**%50**	**%70**	**%100**
Invisible Secrets	LSB	%0	%100	%100	%10	%10
Outguess	Dönüşüm	%90	%80	%100	%100	%100
F5	Dönüşüm	%80	%100	%100	%100	%100
Stegano	LSB	%100	%90	%100	%100	%100
StegHide	Dönüşüm	%100	%90	%100	%100	%100

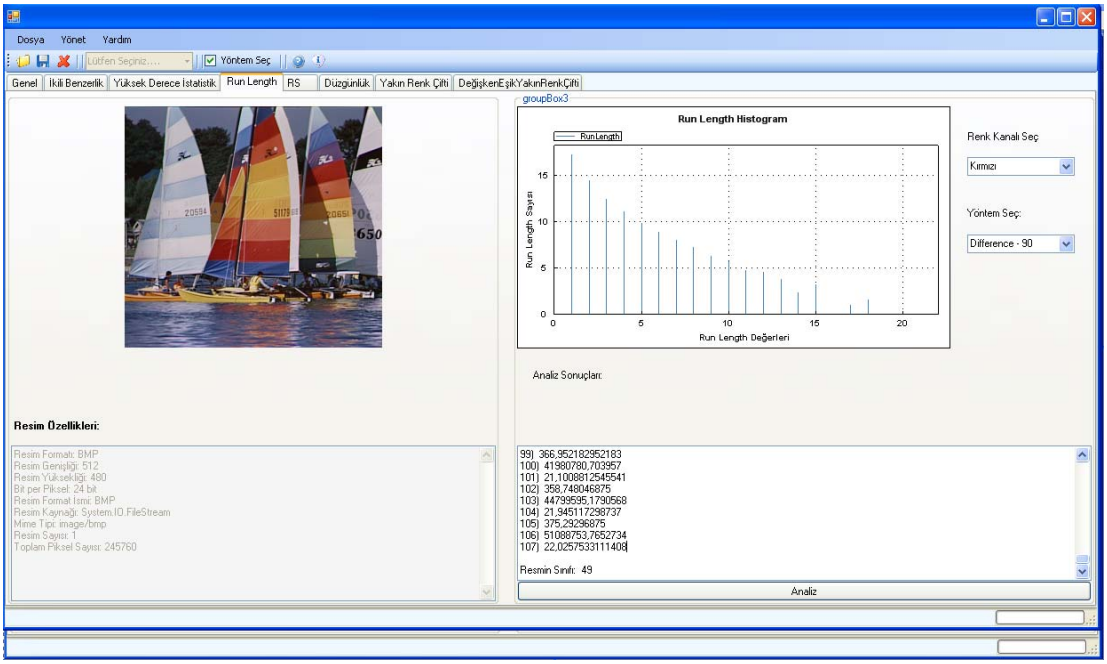
Son kör steganaliz yöntemi olarak geliştirilen uygulamaya, farklı bir yaklaşıma sahip Koşu Uzunluğuna Dayalı Steganaliz yöntemi eklenmiştir. Bu analiz yönteminde resmin renk kanallarına ayrılarak kaç tane aynı değere sahip pikselin birbirini takip ettiğinin sayısı bulunmaya çalışılır. Bu sayı bulunurken 4 farklı yönde (0^0, 45^0, 90^0 ve 135^0) hesaplama yapılır. Ayrıca [13]'te önerilen iki farklı koşu uzunluğu hesapla yöntemi ile analiz resim üzerinde yapılabilecek değişikliklere karşı daha hassas hale getirilmiştir. Geliştirilen uygulamaya resmin koşu uzunluğu histogramını gösteren bir arayüz eklenmiştir. Bu arayüz ile orijinal resminde elde olduğu durumda iki resim arasındaki farkın belirlenebilmesi için bir karşılaştırma imkânı sunulmuştur. Bu arayüze ait bir ekran görüntüsü Şekil 5.9'da görülmektedir. Bu şekilde "Yat.jpg" resmine ait kırmızı renk kanalının dikey (90^0) yöndeki koşu uzunluğu histogramı gösterilmiştir. Form üzerinde bulunan bileşim kutuları ile istenen renk kanlı ve istenen yöndeki koşu uzunluğu histogramları görüntülenebilmektedir.



Şekil 5.7. Yüksek dereceli istatistiklere dayalı steganaliz yöntemi için akış diyagramı



Şekil 5.8. Resmin 4 seviyeli dalgacık dönüşümü katsayılar resmi



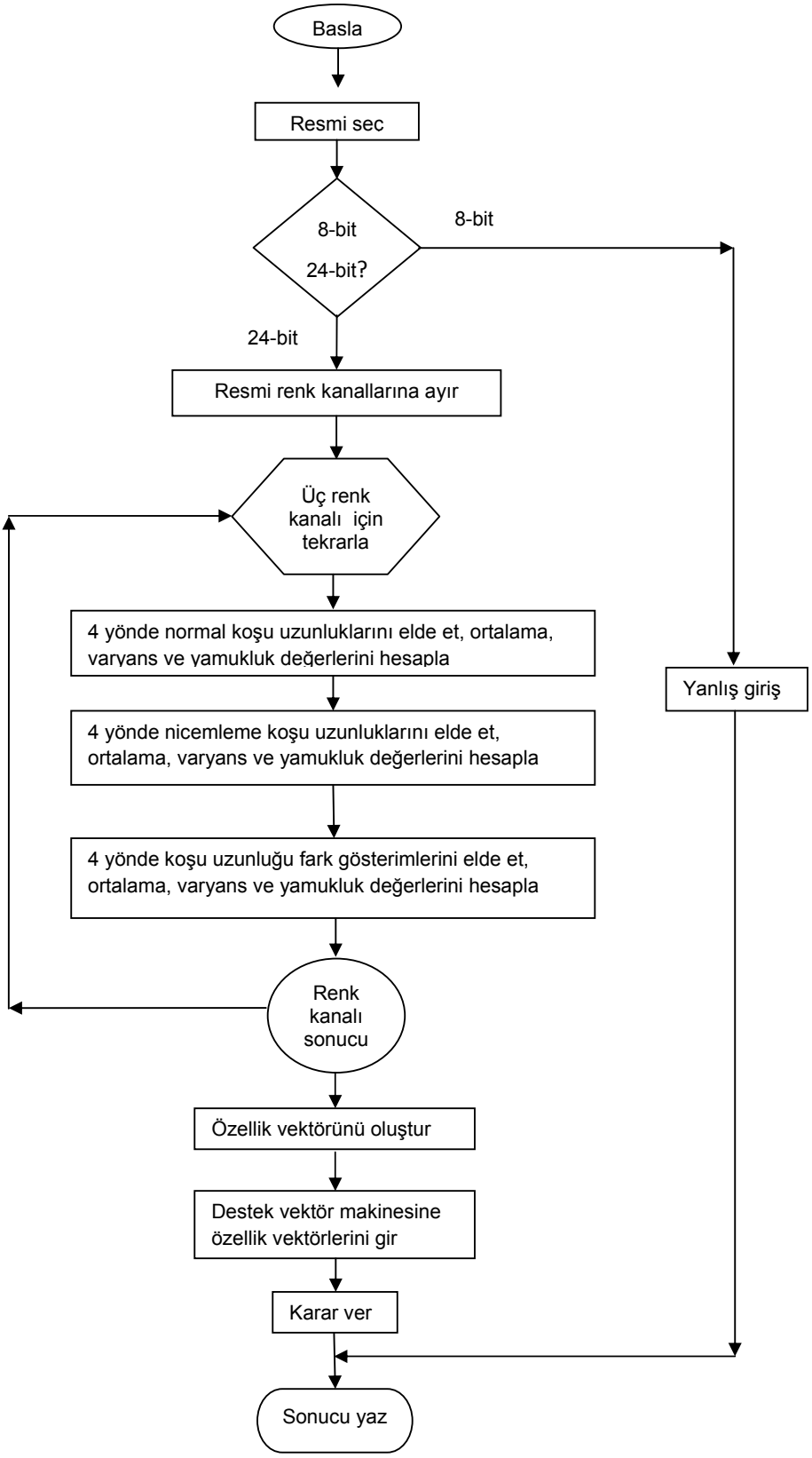
Şekil 5.9. Koşu uzunluğu histogramı arayüz formu ekran görüntüsü (Yat.jpg resminin kırmızı renk kanalındaki $\mathbf{90^o}$'lik fark koşu histogramı)

Yöntemde hesaplanan 3 farklı tip ve 4 farklı yöndeki koşu uzunluğu hesaplamalarının üçüncü derece istatistikleri kullanılmaktadır. Her renk kanalı için 36, toplamda da 108 istatistik değeri özellik olarak sınıflandırıcıya verilmiştir. Koşu uzunluğuna dayalı steganaliz yöntemi için akış diyagramı Şekil 5.10'da sunulmuştur.

Geliştirilen uygulamanın test veritabanı üzerinde elde edilen sonuçları Çizelge 5.12'de gösterilmiştir. Sonuçlar incelendiğinde yüksek dereceli istatistiklere dayalı steganaliz yöntemiyle benzer sonuçlar elde edildiği görülmüştür. Eğitim veritabanında kullanılan Outguess, F5, Stegano ve StegHide programları ile üretilmiş resimler üzerindeki tespit oranı %90 ve %100 seviyelerindedir. Elde edilen sonuç [13]'te elde edilen sonuçlar ile karşılaştırıldığında daha yüksek yüzdelerin elde edildiği görülmüştür. Ancak eğitim veritabanında kullanılmayan Invisible Secret programında bu başarı oranı azalmıştır. Mesaj uzunluğunun maksimum kapasitenin yarısı seviyelerinde tespit etme oranı kabul edilebilir seviyelerdedir (%70). Ancak %10 ve %100 mesaj oranlarında tespit etme oranı 0 olarak elde edilmiştir.

Çizelge 5.12. Koşu Uzunluğuna Dayalı Steganaliz yöntemi için test sonuçları

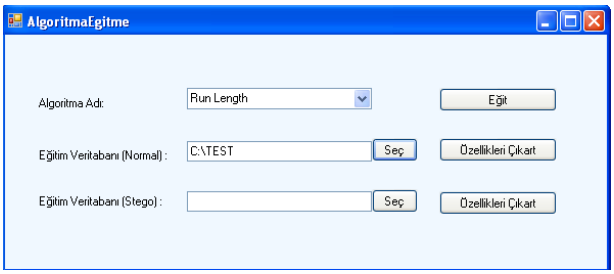
		Mesaj Yüzdesi				
Steganografi Programı	**Kullandığı Yöntem**	**%10**	**%30**	**%50**	**%70**	**%100**
Invisible Secrets	LSB	%0	%100	%70	%0	%0
Outguess	Dönüşüm	%100	%90	%100	%100	%100
F5	Dönüşüm	%80	%100	%100	%100	%100
Stegano	LSB	%100	%90	%100	%80	%100
StegHide	Dönüşüm	%100	%100	%100	%100	%100



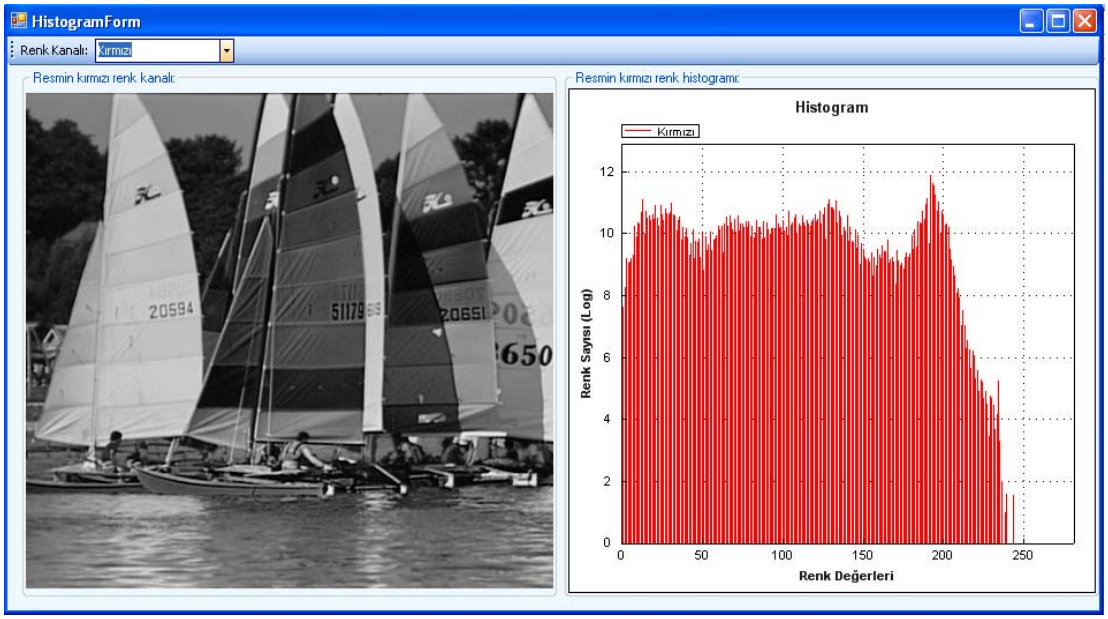
Şekil 5.10. Koşu uzunluğuna dayalı steganaliz yöntemine ait akış diyagramı

Uygulanan bu üç kör steganaliz yöntemine ait sonuçlar incelendiğinde eğitim veritabanı oluşturulurken kullanılan yöntemlerde üç yönteminde yüksek doğruluğa sahip olduğu görülmüştür. Bu başarı oranında eğitim veritabanının doğru şekilde oluşturulmuş olması etkili olmuştur. Eğitim veritabanında kullanılmayan Invisible Secrets programının sonuçları incelendiğinde İkili Benzerlik Ölçütüne Dayalı Steganaliz yöntemi için yüksek başarı oranı yakalanmıştır. Ancak Yüksek Dereceli İstatistiklere Dayalı Steganaliz ve Koşu Uzunluğuna Dayalı Steganaliz yöntemleri için aynı başarı oranı elde edilememiştir. Bu yöntemlerde saklanan mesaj oranı %30'un altında olduğunda ya da mesaj oranı %70'in üzerine çıktığında tespit etme oranı %10'ların altına düşmekte hatta 0 olmaktadır.

Kör steganaliz uygulamalarında elde edilen bu başarının kullanılan eğitim veritabanının doğru ve yüksek sayıda resimden oluşmasından kaynaklandığı düşünülmektedir. Bu nedenle geliştirilen uygulamaya eğitim veritabanı güncelleştirilmesine imkân sağlayan bir arayüz geliştirilmiştir. Bu arayüze ait ekran görüntüsü Şekil 5.11'de gösterilmiştir. Eklenmek istenen resimlerin bulunduğu dizin ve steganaliz yöntemi seçilerek yeni resimlerin özellikleri çıkartılarak daha önce oluşturulmuş olan model dosyasının güncellenmesine imkân tanınmıştır. Ayrıca geliştirilen uygulamaya daha detaylı inceleme sağlayabilmek için resmin renk kanallarını ve bu renk kanalına ait histogramını gösteren bir arayüz eklenmiştir. Bu arayüze ait ekran görüntüsü Şekil 5.12'de görülmektedir.



Şekil 5.11. Kör steganaliz yöntemleri için veritabanı eğitme arayüzü



Şekil 5.12. Resmin renk kanalı ve histogramını gösteren arayüze ait ekran görüntüsü

Kör steganaliz yöntemleri genel analize sahip olmaları ve steganografi algoritmalarında yapılacak değişikliklere kolaylıkla adapte olabilmeleri nedeniyle daha kullanışlı olarak düşünülmektedir. Bu yöntemlerde sınıflandırıcı ne kadar doğru eğitilirse yapılan algoritma da o doğrulukla tespit etmektedir. Elde edilen sonuçlar doğrultusunda sınıflandırıcı eğitilirken farklı tiplerde ve farklı kaynaklardan elde edilen resimlerin seçilmesinin gerekli olduğu sonucuna varılmıştır. Ayrıca eğitim veritabanına eklenecek stego resimler oluşturulurken kullanılacak steganografi algoritmalarının da çeşitlendirilmesi gerektiği ortaya çıkmıştır. Eğitim sırasında kullanılmayan Invisible Secret programının çıktılarında sadece İkili Benzerlik Ölçütlerine Dayalı Steganaliz yöntemi başarılı sonuçlar vermiştir. Diğer iki yöntem %50'ye yakın mesaj oranlarında doğru sonuçlara ulaşmıştır. Bu yüzden farklı yaklaşımlara sahip steganografi programlarının stego resimleri oluştururken kullanılmasının daha başarılı sonuçlara ulaşılmasında faydalı olacağı düşünülmektedir.

Yapılan çalışmada sınıflandırıcıları eğitmek için yaklaşık 50.000 adet normal ve stego resim kullanılmıştır. Bu büyüklükte bir veritabanı başarılı sonuçların elde edilmesine neden olmuştur. Ancak bu büyüklükte bir veritabanının

kullanılması resimler test edilirken programda yavaşlamaya neden olmuştur. Bu veritabanına eklenecek her resim programın biraz daha uzun sürede karar vermesine neden olacaktır.

Belirtilen bu iki noktanın yanı sıra eğitim veritabanında kullanılan stego resimlerde farklı mesaj büyüklüklerinin kullanılması %10 gibi az oranda mesaj saklanmış durumlarda bile başarılı (%90) sonuçlar elde edilmesine neden olmuştur.

6. SONUÇ VE TARTIŞMA

Bu tezde resim steganografi yöntemlerinin tespitine yönelik steganaliz yöntemleri incelenmiş ve bu yöntemleri kapsayan bir uygulama geliştirilmiştir. Geliştirilen bu uygulamada sadece belirli bir steganografi yöntemine yönelik tespit yerine, genel ve yeni tekniklere adapte edilebilecek bir uygulamanın geliştirilmesi amaçlanmıştır.

Uygulamada steganografi uygulamalarında en çok kullanılan yaklaşım olan LSB veri saklamayı hedef alan RS, Yakın Renk Çifti, Resim Düzgünlüğüne Dayalı ve Değişken Eşik Değerli Yakın Renk çifti steganaliz yöntemleri gerçekleştirilmiştir. Bu analiz teknikleri LSB steganografi için başarılı örnekler olmakla birlikte karar vermede önemli bir kriter olan eşik değeri uygun seçilmez ise içerisinde bilgi olan resimler normal resim, bilgi olmayan resimler de stego resim olarak sınıflandırılabilmektedir. Ancak analistin resimler üzerinde elde edilen sonuçlara bakarak müdahale etme imkanı bulunmaktadır. Edinilecek tecrübe ile eşik değeri kullanılmadan resmin sınıflandırılma olanağı bulunmaktadır.

Uygulanan bu yöntemler sadece LSB steganografi tekniklerinin tespit edilmesinde kullanılan yöntemlerdir. Daha genel bir uygulama yapabilmek için geliştirilen uygulamaya kör steganaliz teknikleri olarak adlandırılan yaklaşımlardan İkili Benzerlik Ölçütü, Yüksek Dereceli İstatistik ve Koşu Uzunluğuna Dayalı Steganaliz yöntemleri eklenmiştir. Bu analiz yöntemleri matematiksel işlem olarak daha karmaşıklardır. Uygulamada yüksek matematiğinden dolayı işlem gücüne bağlı olarak analiz süresi daha uzun sürebilmektedir. Ayrıca bu steganaliz yöntemlerinde karar aşamasında analistin müdahale şansı bulunmamaktadır.

Kör steganaliz yöntemlerinde sınıflandırıcı olarak kullanılan destek vektör makinelerinin doğru eğitilmesinin tespit etmede algoritmanın başarısını doğrudan etkilediği belirlenmiştir. Kör steganaliz yöntemlerinin karar verme mekanizması olan DVM eğitilirken farklı tipteki resimlerin kullanılmasının ve farklı steganografi algoritmaları kullanılarak stego resimlerin oluşturulması gerekliliği ortaya çıkartılmıştır. Ayrıca seçilen resimlerin tipi, ve farklı formatlarda olmaları analiz yöntemini etkilemektedir. Bunların yanı sıra eğitim seti hazırlanırken yaratılan

stego resimlere veri saklarken kullanılan gizli mesajın miktarının da önem taşıdığı anlaşılmıştır.

Çok yüksek sayıda resimden oluşan eğitim veritabanı kullanıldığında kör steganaliz yöntemleri için programın çalışma süresinin uzadığı görülmüştür. Ancak çok sayıda resim veritabanında kullanılarak algoritmaların başarı yüzdeleri arttırılmıştır. Bu nedenle hem veritabanını daha da zenginleştirebilmek ve hem de çalışma süresinin bundan daha az etkilenmesini sağlamak için ayrı kategorilerde veritabanının oluşturulmasının faydalı olacağı düşünülmektedir. Steganografi programı bazlı ya da resimlerin içerdikleri mesaj oranlarına bağlı olarak veritabanlarını ayırıp, ayrı ayrı model dosyaları oluşturarak hem veritabanının zenginleştirilebileceği hem de değerlendirme süresinin artmayacağı düşünülmektedir.

İleriye yönelik olarak yapılacak çalışmalarda geliştirilmeye çalışılan steganaliz aracının kapsamını ve başarım oranını artırabilmek için daha fazla sayıda steganaliz yönteminin eklenmesi gerekecektir. Ayrıca kör steganaliz yöntemlerinin eğitilmesinde kullanılan resimlerin daha çeşitli hale getirilmesi ve stego resimlerin oluşturulmasında kullanılan steganografi programlarının sayılarının artırılmasına ihtiyaç duyulacaktır. Bunların yanı sıra her tipteki resim için ayrı bir veritabanı oluşturularak ayrı DVM'lerin eğitilmesi kullanılan steganaliz programlarının daha başarılı olmasını sağlayabilecektir.

KAYNAKLAR

[1] Amasyalı M.F., 2006, Makine Öğrenmesine Giriş, Yıldız Teknik Üniversitesi, Bilgisayar Mühendisliği Bölümü, http://www.ce.yildiz.edu.tr/mygetfile.php?id=868, Rapor, 22s.

[2] Anderson R.J., Petitcolas F.A.P., 1998, On The Limits of Steganography, IEEE Journal of Selected Areas in Communications, 16(4):474-481.

[3] Anderson R.J., Petitcolas F.A.P., 1999, Information Hiding-A Survey, Proceedings of the IEEE, Special Issue on Protection of Multimedia Content, 87(7):1062-1078.

[4] Avcıbaş İ., Kharrazi M., Memon N., Sankur B., 2005, Image Steganalysis with Binary Similarity Measures, EURASIP Journal on Applied Signal Processing, 17, 2749–2757.

[5] Bayram S., Avcıbaş İ., Sankur B., Memon N., 2006, Image Manipulation Detection With Binary Similarity Measures, Journal of Electronic Imaging, vol. 15(4), 041102, pp. 1-17.

[6] Cancelli G., Do'err G., Barni M., and Cox I.J., 2008, A Comparative Study of ±1 Steganalyzers, IEEE,MMSP, pp. 791-796.

[7] Chan S.C., Chang C.C., 2007, A Survey of Information Hiding Schemes for Digital Images, IJCSES International Journal of Computer Sciences and Engineering System, Vol.1, No:3, 187-200.

[8] Chandramouli R., 2003, A mathematical framework for active steganalysis, ACM Multimedia Systems, vol. 9, no. 3, pp. 303–311.

[9] Chandramouli R., Subbalakshmi K.P., 2004, Current Trends in Steganalysis: A Critical Survey, Control, Automation, Robotics and Vision Conference,USA, Vol.2, 964-967.

[10] Choi S.S., Cha S.H., Tappert C.C., 2010, A Survey of Binary Similarity and Distance Measures , Journal of Systemics, Cybernetics and Informatics, vol.8 no.1, pp. 43-48.

[11] Cortes C., Vapnik V., 1995, Support-Vector Networks, Machine Learning, 20,pp. 273-297.

[12] Cox J., Miller L., Bloom J., Fridrich J., Kalker T., 2008, Digital Watermarking and Steganography, Morgan Kaufmann, İkinci Basım.

[13] Dong J., Tan T., 2008, Blind Image Steganalysis Based On Run-Length Histogram Analysis, The International Conference on Image Processing-ICIP, Berkeley, 4s.

[14] Engle S., 2003, Current State of Steganography: Uses, Limits, & Implications, University of California, Davis College of Engineering Department of Computer Science, Project Report, 12s.

[15] Fridrich J., Du R., Long M.,2000, Steganalysis of LSB Encoding in Color Images, IEEE, Multimedia and Expo, ICME, 1279-1282.

[16] Fridrich J., Goljan M., Du R., 2001, Detecting LSB Steganography in Color and Gray- Scale Images, State University of New York, Binghamton, IEEE Multimedia and Security, pp 22-28.

[17] Fridrich J., Goljan M., 2002, Practical Steganalysis of Digital Images – State of the Art", In Proceedings of SPIE, Security and Watermarking Multimedia Contents IV, 21–24.

[18] Hmood A.K., Kasirun Z.M., Jalab H.A., Alam G.M., Zaidan A.A., Zaidan B.B., 2010, On The Accuracy of Hiding Information Metrics: Counterfeit Protection for Education and Important Certificates, International Journal of the Physical Sciences, Vol.5(7), pp.1054-1062.

[19] Johnson F., Jajodia S., 1998, Steganalysis of Images Created Using Current Steganography Software, Information Hiding, LNCS 1525, pp. 273-289,

[20] Johnson F., Jajodia S., 1999, Steganalysis of Images Created Using Current Steganography Software, Arizona University.

[21] Katzenbeisser, S., Petitcolas, F.A.P., 2000, Information Hiding Techniques for Steganography, Artech House, INC.

[22] Lyu S., Farid H., 2002, Detecting Hidden Messages Using Higher-Order Statistics and Support Vector Machines, 5th International Workshop on Information Hiding, Noordwijkerhout, The Netherlands, 15s.

[23] Manoharan S., 2008, An Empirical Analysis of RS Steganalysis, Department of Computer Science University of Auckland, New Zealand, IEEE Computer Society, 172-177.

[24] Mielikainen J., 2006, LSB Matching Revisited, Signal Processing Letters, IEEE,Signal Processing Letters, Vol. 13, No:5, 285-287.

[25] Neil F.Johnson,1998, Sushil Jajodia, Exploring Steganography: Seeng the Unseen, IEEE Computing Practices, 0018-9162/98, 26-34.

[26] Rabah K., 2004, Steganography-The Art of Hiding Data, Information Technology Journal 3 : 245-269.

[27] Raja K.B., Shankara N., Venugopal K.R., Patnaik L.M., 2006, Steganalysis of LSB Embedded Images Using Variable Threshold Color Pair Analysis, Intelligent Sensing and Information Processing, ICISIP, Fourth International Conference, pp. 11-16.

[28] Dipperstein M., 2010, Run Length Encoding (RLE) Discussion and Implementation, http://michael.dipperstein.com/rle/index.html.

[29] Sabu M Thampi, 2004, Information Hiding Techniques: A Tutorial Review, Potentials, Department of Computer Science & Engineering LBS College of Engineering, India, ISTE-STTP on Network Security & Cryptography.

[30] Sellars D., 2007, An Introduction to Steganography, http://www.zoklet.net/totse/en/privacy/encryption/163947.html.

[31] Stallings W., 2006, Cryptography and Network Security Principles and Practices, Pearson Prentice Hall, Dördüncü Basım.

[32] Şahin A., Buluş E., Buluş H.N., Sakallı M.T., 2006, 24-Bit renkli Resimler Üzerinde Uygulanan RS Steganalizde Maske Seçimlerinin Etkileri, Elektrik-Elektronik-Bilgisayar Mühendisliği Sempozyumu, Rapor, 4s.

[33] Şahin A., Buluş E., Sakallı M.T., Buluş H.N., 2007, Resim İçerisindeki Gizli Bilginin RQP Steganaliz Yöntemiyle Sezilmesi, Akademik Bilişim 2007, Dumlupınar Üniversitesi, Kütahya, 5s.

[34] Westfeld A., Pfitzmann A., 1999, Attacks on Steganographic Systems, Information Hiding. Third International Workshop, IH'99, Dresden, Germany.

[35] Westfeld, A., 2001, High Capacity despite Better Steganalysis: F5–A Steganographic Algorithm, In: Moskowitz, I.S. (eds.): Information Hiding. 4th InternationalWorkshop. Lecture Notes in Computer Science, Vol.2137. Springer-Verlag, Berlin Heidelberg New York, 289–302.

[36] Zhang T., Zhang Y., Ping X., Song M., 2006, Detection Of Lsb Steganography Based On Image Smoothness, Multimedia and Expo, IEEE, 1377-1380.

[37] Zhang J., Cox J., Doerr G., 2007, Steganalysis for LSB Matching in Images with High-frequency Noise, IEEE, MMSP, 385-387.

EKLER

EK 1

RESİM FORMATLARI

Resim dosyaları kayıplı ve kayıpsız olmak üzere iki formatta sayısal ortamlarda saklanmaktadır. Aşağıda kayıpsız (bmp, gif, tiff) ve kayıplı (jpg) resim formatlarına ait kısa açıklamalar yapılmıştır.

BMP

BMP, herhangi bir sıkıştırma yapmadan resmin özelliklerini tutan, Microsoft firmasına ait bir resim dosyası biçimidir. Sıkıştırma yapmadığı için PNG, JPEG gibi dosya biçimlerine göre çok daha fazla yer kaplar.

En temel resim formatı BMP[1] (Bitmap)'dir. BMP'nin birbirinden farklı bir kaç türü bulunmaktadır. X-Windows ile MS-Windows ya da OS/2 ortamındaki BMP dosyaları için arada farklar vardır. X-Windows üzerindeki BMP formatı sadece 2 rengi desteklemektedir. MS-Windows üzerinde bulunan BMP dosyaları 16 ya da daha çok renk kaydedebilecek, herhangi bir sıkıştırma yapmayan oldukça hızlı bir formattır. Bu formatta resmin içindeki renk sayısı değil, resmin büyüklüğü önemlidir.

16 renk, 800x600 çözünürlüğünde bir BMP dosyası, 800x600x1/2=240000 baytlık bir yer kaplamaktadır. Resmin içinde 1, 2 ya da 12 renk olması hiç önemli değildir. 256 renk olarak kaydedilen bir dosya ise, 800x600x1=480000 baytlık yer tutmaktadır.

JPG

JPEG[2], Birleşik Fotoğraf Uzmanları Grubu (Joint Photographic Experts Group) tarafından standartlaştırılmış bir sayısal görüntü kodlama biçimidir. Bu biçim, 1994 yılında ISO 10918-1 adıyla standartlaşmıştır.

[1] http://www. po.metu.edu.tr/links/inf/css25/bolum13.html
[2] http://www.jpeg.org/JPEG

JPEG standardında görüntü saklayan dosya biçimi genellikle JPEG olarak adlandırılır. Bu dosyalar .jpg, .jpe ya da .jfif uzantılıdır, ancak çoğunlukla .jpg uzantısı kullanılır.

JPEG standardı sadece görüntünün nasıl kodlanacağını tanımlar, görüntünün herhangi bir saklama ortamında depolanma biçimini belirtmez. JPEG olarak bildiğimiz dosya biçimi, Independent JPEG Group adlı başka bir grubun JFIF (JPEG File Interchange Format - JPEG Dosya Alışveriş Biçimi) adlı standardı tarafından tanımlanmıştır.

Bu dosya biçimi, internet üzerinden görüntü iletmek ve fotografik görüntü saklamak için en çok kullanılan dosya biçimidir. JPEG/JFIF formatı, oldukça başarılı bir depolama ve veri transfer yapısına sahiptir.

JPEG formatında, ayarlanabilir kayıplı sıkıştırma kullanılmaktadır. Dolayısıyla JPEG verisinden okunan görüntü ile veriyi yaratmak için kullanılan görüntü aynı değildir. Ancak, kayıplar insan görme sisteminin daha az önem verdiği detaylarda gerçekleştiği için çoğu zaman fark edilmezdir.

Standart JPG formatında, resmin kalitesinden bir miktar ödün verilerek sıkıştırma uygulanır. Böylece dosya boyutu bir hayli düşer. Özellikle 24 bit renkli uygulamalarda resim kalitesinin düştüğünü anlamak mümkün değildir. Bu yüzden bu tip uygulamalarda JPG tercih edilir. JPEG'den ne kadar sıkıştırma istendiği (0-100 arası bir faktör) seçilebilmektedir ama genellikle 5-95 arası kullanılır. 95'den fazlası resimde detay kaybına yol açar ve 5'ten küçük seçilmesi durumunda da dosya boyutunda fazla küçülme olmaz.

JPEG formatında her renk bileşeni, 8x8 bloklar halinde ayrık kosinüs dönüşümü ile dönüştürülür. Bu sayede resmin enerjisi az sayıda pikselde yoğunlaştırılır. Dönüştürülen blokların nicemlenmesi sonrasında da sıfırdan farklı az sayıda değer ile bloğu ifade etmek mümkün olur. Dönüşüm uzayındaki yüksek frekans pikselleri, resmin görsel kalitesinde daha az rol oynarlar, dolayısıyla yüksek frekans pikselleri daha az sayıda değere nicemlenmiş olur.

GIF

GIF, İngilizce Grafik Değiştirme Biçimi anlamına gelen Graphics Interchange Formatin kısaltmasıdır ve bir sayısal resim saklama biçimidir. Kayıpsız sıkıştırma kullanır. 8-bit renge (yani 256 renk) kadar destek verir ve tek renk için1-bit'lik saydamlık sunar. JPEG ile birlikte bilgisayar dünyasında kullanılan en yaygın resim saklama biçimlerinden biridir. Az renk içermeleri dolayısıyla genellikle grafiklerin saklanması için kullanılır[1].

GIF, COMPUSERVE'in geliştirdiği bir resim formatıdır. İyi bir sıkıştırma algoritması vardır ve görüntüleme de oldukça hızlı bir şekilde gerçekleştirilir. 256 renk dışında (8 bit) herhangi önemli bir sorunu bulunmamaktadır.

GIF formatının iki farklı versiyonu bulunmaktadır. Bunlar 87a ve 89a formatlarıdır. 89a versiyonu, tek bir GIF dosya içinde birden çok GIF formatlı resim yerleştirilmesine ve anime edilmesine olanak tanır. Ayrıca, GIF89a versiyonu, katmanlı görüntü saklama özelliğine de sahiptir. Bu, özellikle internet üzerindeki resimlerde kullanılır.

PNG

PNG[2], "Taşınabilir Ağ Grafiği" anlamındaki Portable Network Graphics 'in kısaltmasıdır ve kayıpsız sıkıştırarak görüntü saklamak için kullanılan bir saklama biçimidir. PNG biçiminde paletli ya da gerçek renkte görüntüler seçimlik bir saydamlık kanalıyla saklanabilir.

Hâlihazırda GIF gibi kabul edilebilir başarımda ve yaygın bir kayıpsız sıkıştırma algoritması varken PNG'nin geliştirilmesini motive eden şey, Unisys'in GIF'de kullanılan algoritması üstündeki patent hakkının ihlallerini takip edeceğini duyurması olmuştur. Gelişen ve yaygınlaşan donanım teknolojisiyle beraber GIF biçimi yetersiz kalmaya da başlamış olduğundan, PNG 1.0 sürümüyle 1 Temmuz 1996'da yayımlanmıştır. 1.1 ve 1.2 sürümleriyle yeni genişletmeler tanımlanmış ve 1.2 sürümü küçük değişikliklerle ISO/IEC 15948:2003 adıyla bir ISO standardı olmuştur.

[1] http://www.image-formats.com/pcx-pcpaint brush-file-format
[2] http://www.w3.org/TR/ PNG

TIFF

TIFF grafik, fotoğraf gibi dosyalar için kullanılan bir biçimdir. Aldus isimli şirket tarafından üretilip 1986 yılında ilk sürümü duyurulmuştur. 1994 yılında Aldus Corp ile Adobe Systems'in birleşmesinden sonra TIFF 6.0 geliştirilmiş ve bir çok yeni özellikler eklenmiştir[1]. JPEG ve PNG gibi TIFF de yüksek renk derinliği olan görüntülerde kullanılır. Photoshop, GIMP gibi görüntü işleme programları TIFF biçimini desteklemektedir.

TIFF esnek ve uyarlanabilir bir dosya biçimidir ve dosya başlığında etiketler kullanarak tek bir dosyada birden fazla görüntüyü ve veriyi barındırabilir. Etiketler görüntünün boyutları gibi temel geometrisini veya hangi sıkıştırma tercihinin kullanıldığını belirtebilir.

TIFF biçimi birden fazla sayfayı desteklediği için, çok sayfalı dokümanlar ayrı dosyalar yerine tek bir TIFF dosyası olarak kaydedilebilir. TIFF dosyaları, kayıpsız görüntü depolama özellikleri sayesinde iyi bir görüntü arşivi olarak kullanılabilirler. Standart JPEG dosyalarının aksine, kayıpsız sıkıştırılmış ya da hiç sıkıştırılmamış bir TIFF dosyası, görüntü kalitesinde bir kayıp olmaksızın düzenlenip yeniden kaydedilebilir. Bu özellik TIFF dosyası içinde yer alan JPEG biçimindeki görüntüler için geçerli değildir.

PCX (Personal Computer eXchange)

PCX formatı Zsoft firması tarafından PC Paintbrush yazılımı için geliştirilmiştir[2]. PC Paintbrush'ın resim formatı olarak geliştirilmiş ve DOS resim standardı olarak kabul edilmiştir. Bugünkü yazılımların çoğu PCX formatının 5.inci sürümünü desteklemektedir. 3.sürüm özel renk paletini desteklemez. Bu nedenle 3. sürüm bir PCX dosyası açılırken standart bir renk paleti kullanılır.

[1] http://partners.adobe.com/public/developer/en/tiff/ TIFF6.pdf
[2] http://courses.engr.illinois.edu/ece390/books/labman uals/graphics-pcx.html

EK 2

ÇEŞİTLİ STEGANOGRAFİK YÖNTEMLER VE BUNLARA AİT UYGULAMALAR

İnternet ortamında ücretli ve ücretsiz olarak sunulan steganografi ile ilgili birçok araç bulunmaktadır. Yapılan araştırmada bu uygulamaların büyük çoğunluğunun LSB yöntemini kullandığı görülmüştür. Bu araçlar resmin pikselinin renk değerini, palet tipi resimlerin indislerini ya da JPEG resimlerin DCT katsayılarını değiştirerek LSB değiştirme yapmaktadır [19]. Ayrıca yapım açısından yani uygulamanın geliştirilmesi açısında LSB daha kolay bir yaklaşım olduğu için bu yöntem seçilmiştir. Aşağıda bu araçlardan birkaç tanesi incelenmiştir.

WhiteNoiseStorm

WhiteNoiseStorm[1] çok yönlü DOS ortamında çalışan bir steganografi uygulamasıdır. LSB yaklaşımını kullanarak PCX dosyaları içerisine bilgi saklamaktadır. Resim içerisine bilgi saklandıktan sonra resim kalitesinde bir azalma olmamaktadır. Ayrıca resim içerisine saklanan bitlerin rastsal olmasını sağlamak için şifreleme kullanmaktadır.

S-Tools

S-Tools[2] uygulaması da LSB steganografi tekniğini kullanmaktadır Saklanan bitlerin rastsallığını sağlamak için şifreleme algoritması kullanılmaktadır. S-Tools uygulaması BMP, GIF ve WAV dosyalarının içerisine bilgi gizleyebilmektedir.

StegoDos

Uygulama DOS ortamında çalışan ve GIF ya da PCX resimlerinin içerisine bilgi saklamayı sağlayan bir araçtır[3]. Program 320x200x256 boyutlarındaki resimlerinin içerisine bilgi saklamayı gerçekleştirir. Bu program da LSB yöntemini kullanarak mesaj gizlemektedir.

---

[1] ftp://ftp.funet.fi/pub/crypt/steganography/wnsto rm-stegography-wns210.zip

[2] ftp://ftp.funet.fi/pub/crypt/mirrors/idea.sec.dsi.unimi. it/code/

[3] http://www.filewatcher.com/m/stegodos.zip.21958. 0.0.html

JStegShell

Resimlerin içerisine LSB yöntemini kullanarak veri saklayan bir programdır[1]. JPEG formatlı resimlerin içine veri saklamakta ve DOS tabanlı olarak çalışmaktadır.

JPEG Hide And Seek

Bu program iki farklı araçtan oluşmaktadır[2]. JPHide bir JPEG resim dosyasının içerisine bilgi saklar ve JPSeek ise saklanan bilginin geri elde edilmesini sağlar. Hem görsel etkileri hem de istatistiksel etkileri azaltmak için gizlenecek olan bilgi JPEG dosyası içerisine dağıtılarak gizlenmektedir. Verinin rastsallığını saklamak için şifreleme tekniklerini kullanmaktadır.

OutGuess

OutGuess[3] herhangi bir taşıyıcı dosya içerisine saklama işlemi gerçekleştirmektedir. Linux, Mac, Windows gibi çeşitli işletim sistemleri üzerinde çalışabilmektedir. Saklanan verinin rastsallığını sağlamak için şifreleme teknikleri kullanılmaktadır. Dönüşüm steganografi tekniklerine dayanan bir uygulamadır.

GifShuffle

GIF resimleri içerisine veri saklayan bu uygulama, saklama işlemi öncesinde sıkıştırma ve şifreleme seçenekleri de sunmaktadır[4]. DOS ortamında çalışan bir programdır.

Steghide

Steghide[5] çeşitli resim ve ses dosyaları içerisine veri saklayabilen bir steganografi uygulamasıdır. Sıkıştırma ve şifreleme imkânı da sunmaktadır. JPG, BMP, WAV ve AU dosyalarını taşıyıcı olarak desteklemektedir.

[1] http://www.securityfocus.com/tools/1434
[2] http://www.infosyssec.com/infosyssec/Steganography/programs.htm
[3] http://www.outguess.org/
[4] http://www.darkside.com.au/gifshuffle/
[5] http://steghide.sourceforge.net/

wbStego4open

Windows ve Linux işletim sistemleri üzerinde çalışan açık kaynak kodlu bir steganografi uygulamasıdır. BMP, HTML, TXT ve PDF formatlarındaki dosyalar içerisine her türlü sayısal veriyi saklayabilmektedir[1]. Verinin rastsal hale getirilmesi için şifreleme uygulamaktadır.

Hermetic Stego

Kullanılışı kolay olan bu yazılım BMP resim dosyaları içerisine veri saklamaktadır. Yeterli miktarda taşıyıcı resim dosyası sağlandığı müddetçe saklanacak olan verinin boyutunda herhangi bir sınırlama yoktur. Hem rastgele saklama hem de şifreleme için bir anahtar kullanmaktadır[2].

EzStego

GIF dosyaları içerisine bilgi saklamayı sağlayan LSB tabanlı bir uygulamadır. Tespit edilebilmeyi azaltmak için resmin renk paletini tekrar sıralama işlemi yapmaktadır [34].

J-Steg

Sadece JPEG dosyaları içerisine veri gizlemek için geliştirilmiş ve LSB yöntemini kullanan bir uygulamadır[3]. Aynı zamanda şifreleme kullanarak gizlenen verinin karmaşık hale getirilmesi sağlanır.

F5

Yöntem JPEG dosyalarına saklanan verinin boyutunu artırmak için tasarlanmıştır. Uygulama resmin DCT katsayılarının EÖB'lerinin değerini değiştirmek yerine katsayıların mutlak değerinin bir azaltılması mantığında dayanarak geliştirilmiştir [35].

[1] http://wbstego.wbailer.com
[2] http://www.hermetic.ch/hst/hst.htm
[3] http://islab.oregonstate.edu/documents/ftpsites/berkeley/ jsteg

EK 3

DESTEK VEKTÖR MAKİNELERİ-DVM

Destek Vektör Makineleri (Support Vector Machine-SVM) sınıflandırma tekniği, istatistiksel öğrenme teorisine dayalı bir eğitimli sınıflandırma tekniğidir. Temelleri Vapnik ve Chervonenkis tarafından ortaya atılmıştır [11]. Sinir ağları, bulanık modeller ve sinir-bulanık ortak sistemleri gibi geleneksel öğrenme ve sistem modelleme yöntemleriyle karşılaştırıldığında, DVM yüksek genelleme başarımı ve yüksek boyutlu küçük sayıda veri üzerinde dahi çalışabilme gibi özelliklerinden dolayı son yıllarda oldukça fazla uygulama alanı bulmuştur. DVM'ler günümüzde veri madenciliği, finans sektörü, görüntü işleme, tıp, gıda ve çeşitli mühendislik problemleri gibi birçok alanda uygulanmaktadır [1].

Bir destek vektör makinesi girilen verileri optimum olarak iki kategoriye ayıran n boyutlu bir hiperdüzlem oluşturur. DVM'ler sinir ağları ile yakından ilişkilidir. Aslında sigmoid bir kernel fonksiyonu kullanan bir DVM, iki katmanlı, ileri beslemeli bir sinir ağına denk gelmektedir. DVM'ler saf istatistiksel kriter yerine marj tabanlı geometrik bir kriter kullanan bir sınıflandırma yöntemi kullanmaktadır[1]. DVM'ler:

i. Geleneksel sınıflandırıcılara kıyasla yüksek sınıflandırma doğruluğuyla ve çok iyi genelleme yetkinlikleriyle sonuçlanan, öz etkinlikleri
ii. Mimari tasarım için gerekli olan sınırlı efor
iii. Öğrenme problemini doğrusal olarak kısıtlanmış kuadratik programlama yöntemleriyle çözme olasılığı

nedeniyle DVM'lere ilgi gittikçe artmaktadır.

DVM'lerde temel mantık doğrusal olarak ayrıştırılabilen veri yapıları için en iyi ayırıcı düzlemin belirlenmesidir. Doğrusal olarak ayrıştırılamayan veri yapıları ise dönüşüm tekniği ile farklı bir boyuta taşınarak çözülür. DVM'ler öğrenme, basit fikirler üzerine kurulma ve pratik uygulamalarda yüksek performans göstermesi bakımından oldukça kullanışlıdır.

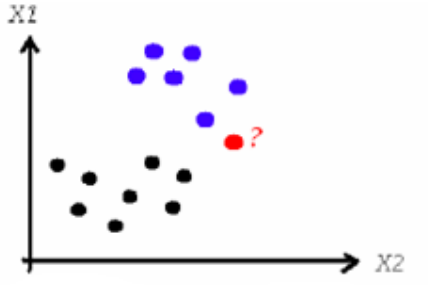
[1] http://www.support-vectormachines .org

DVM'lerde kullanılacak örnek sayısı önemli değildir. Eğitim esnasında görülmemiş verileri de sorunsuz olarak sınıflandırır ki bu da DVM'nin genelleştirebilme yeteneğinden kaynaklanmaktadır.

Destek Vektör Makine Algoritması

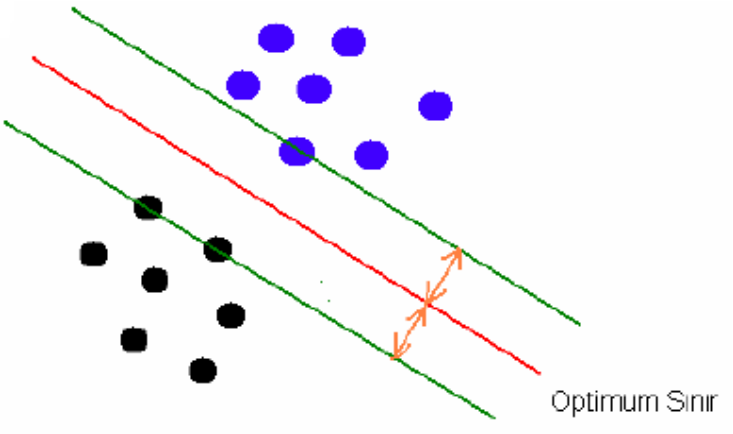
DVM temelinde, öncelik değişkenini bir özellikle çağırıp, çok boyutlu düzlemde kullanılan niteliğe dönüştürür. En uygun temsili seçmenin amacı, özellik seçimi olarak bilinir. Özellikler doğru seçilirse iyi temsiller elde edilerek doğru sonuçlara ulaşılabilir. Bir olayı tanımlayan özellik takımı bir vektör ile çağırılır. Bundan dolayı DVM modelinin amacı, hedef değişkeninin bir kategorisiyle olayların vektör kümelerini ayıran optimal hiperdüzlemi bulmaktır [36].

DVM'lerde genel amaç, sınıfları birbirinden ayıran özel düz bir çizginin bulunmasıdır. Şekil 3.1'de buna dair bir örnek verilmeye çalışılmıştır. Şekilde kırmızı ile belirtilen noktanın hangi sınıfa ait olacağının bulunabilmesi için mavi ve kırmızı noktalar arasına optimum bir çizginin çizilmesi gerekmektedir.



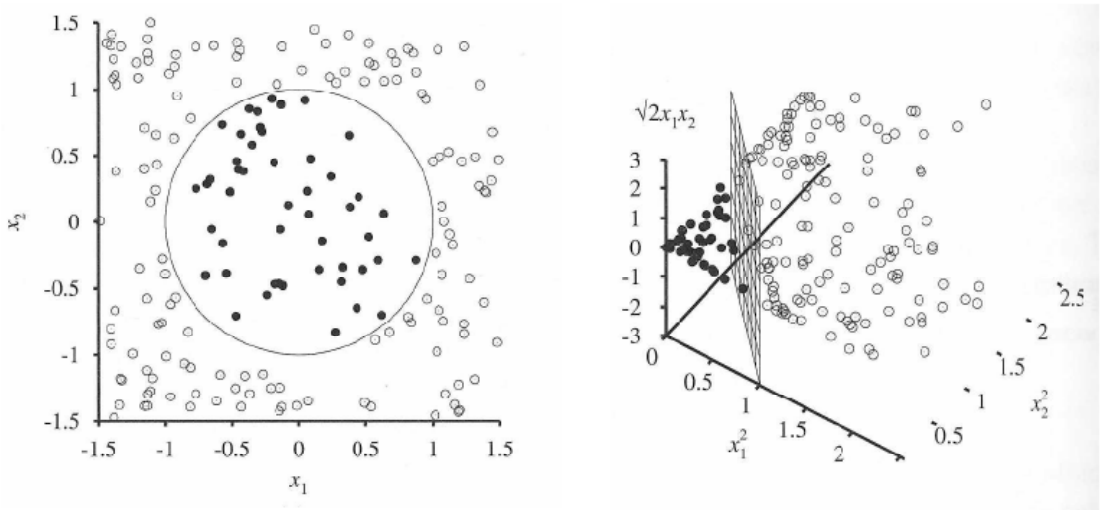
Şekil 3.1. Sınıflandırma problemi (Kırmızı hangi sınıftandır?)

Şekil 3.2'de gösterilen durumda bu çizginin birçok farklı şekilde çizilebilme ihtimali vardır. DVM'ler bu çizgilerden her iki sınıfa en uzak olanını bulmaya çalışarak hatayı en aza indirmeyi amaçlar. Eğitim verileriyle sınır çizgisi bulunduktan sonra test verileri sınırın hangi tarafında kaldıklarına göre sınıflandırılırlar.



Şekil 3.2. DVM ile sınıflandırma

Doğrusal olarak ayrılamayan örnekler için örnekler daha yüksek boyutlu başka bir uzaya taşınır ve sınıflandırma o uzayda yapılır. Örnek olarak Şekil 3.3a'da iki boyutlu uzayda (x_1, x_2) doğrusal olarak ayrılamazken, Şekil 3.3b'de üç boyutlu uzaya $\left(x_1^2, x_2^2, sqrt(2x_1x_2)\right)$ taşınarak ayrılabilmektedirler.



(a) (b)

Şekil 3.3. (a) 2 boyutta doğrusal ayrılamayan veriler, (b) 3 boyutta doğrusal ayrılabilen veriler

ÖZGEÇMİŞ

Adı Soyadı : Firdes AKTAŞ

Doğum Yeri : Çankırı

Doğum Yılı : 1980

Eğitim ve Akademik Durumu:

Lise : 1994 - 1997 İstanbul İhsan Mermerci Lisesi

Lisans : 1997 - 1999 Ankara Üniversitesi

Elektronik Mühendisliği Bölümü

1999 - 2003 İstanbul Üniversitesi

Elektronik Mühendisliği Bölümü

Yabancı Dil: İngilizce

İş Tecrübesi:

2004 - Başbakanlık

Proje Mühendisi