

**T.C.
FIRAT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

YÜZ GÖRÜNTÜLERİNDEN CİNSİYET TAHMİNİ

**YÜKSEK LİSANS TEZİ
Serdar ABUT
(092129107)**

**Anabilim Dalı: Bilgisayar Mühendisliği
Programı: Donanım**

Danışman: Yrd. Doç. Dr. Burhan ERGEN

Tezin Enstitüye Verildiği Tarih:

OCAK – 2015

T.C
FIRAT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

YÜZ GÖRÜNTÜLERİNDEN CİNSİYET TAHMİNİ

YÜKSEK LİSANS TEZİ

Serdar ABUT

(092129107)

Tezin Enstitüye Verildiği Tarih : 29 Ocak 2015

Tezin Savunulduğu Tarih : 22 Ocak 2015

Tez Danışmanı :

Yrd. Doç. Dr. Burhan ERGEN (F.Ü)

Diğer Jüri Üyeleri :

Yrd. Doç. Dr. Taner TUNCER (F.Ü)

Yrd. Doç. Dr. Turgay KAYA (F.Ü)

OCAK – 2015

ÖNSÖZ

Çalışmalarım boyunca benden yardımlarını esirgemeyen, beni yönlendiren değerli hocam Yrd. Doç. Dr. Burhan ERGEN'e, maddi ve manevi desteklerini çalışma boyunca hep hissettiğim Yrd. Doç. Dr. Sait İZGİ'ye teşekkür ederim.

Serdar ABUT

ELAZIĞ - 2015

İÇİNDEKİLER

ÖNSÖZ.....	II
İÇİNDEKİLER.....	III
ÖZET	V
ABSTRACT	VI
ŞEKİLLER LİSTESİ.....	VII
TABLolar LİSTESİ.....	IX
KISALTMALAR LİSTESİ	X
1. GİRİŞ.....	1
2. ÖNİŞLEME	4
2.1. Yumuşatma.....	5
2.2. Histogram Eşitleme	6
3. ÖZNİTELİK ÇIKARIM YÖNTEMLERİ	7
3.1. Gri Seviye Eş-Oluşum Matrisi	8
3.1.1. Enerji	9
3.1.2. Zıtlık.....	9
3.1.3. Homojenlik	10
3.2. Ayrık Dalgacık Dönüşümü.....	11
3.2.1. Haar Dalgacığı.....	17
3.2.2. Daubechies	17
3.2.3. Biortogonal	19
3.2.4. Coiflet.....	20
3.2.5. Symlets	21
3.2.6. Ters Biortogonal (ReversBiorthogonal)	22
4. SINIFLANDIRMA.....	24
4.1. Çapraz Doğrulama	24
4.2. Destek Vektör Makineleri	25
5. UYGULAMALAR.....	28
5.1. Materyal.....	28
5.2. GSEM Kullanarak Öznitelik Çıkarılması	29
5.2.1. GSEM Öznitelikleri ile Cinsiyet Sınıflandırması	31
5.3. ADD ve GSEM ile Öznitelik Çıkarımı.....	32
5.3.1. GSEM ve ADD Öznitelikler İle Cinsiyet Sınıflandırması	33

6.	SONUÇ	37
7.	KAYNAKÇA	38
	ÖZGEÇMİŞ	41

ÖZET

İnsanların yüz biyometrikleri üzerine yapılan çalışmalar, özellikle kamuya açık alanlarda uygulanmak istenen biyometrik sistemlerin uygulama alanlarının yaygınlaşması açısından önemlidir. İnsan yüzüne bakarak cinsiyeti doğru tahmin edebilme işini insanlar yüksek bir başarımla yapabilmektedir. Fakat, bunun bilgisayarlara yaptırılabilmesi, gerçek zamanlı ve geniş ölçekli uygulamalar açısından çok önemlidir.

Bu çalışmada, yüz görüntülerinden cinsiyet tanınması üzerine çalışılmıştır. Çalışmada materyal olarak insanların karcı cepheden alınmış yüz görüntüleri kullanılmıştır. Bu yüz görüntülerinden öznitelik çıkartmak amacı ile Gri Seviye Eş Oluşum Matrisi ve ayrık dalgacık dönüşümü kullanılmıştır. Dalgacık dönüşümü çalışmalarında cinsiyet belirlemesine yönelik öznitelik çıkarmak için 103 farklı dalgacık filtresinin başarımları belirlenmeye çalışılmıştır. Çıkarılan özniteliklerin sınıflandırılması için ise Destek Vektör Makineleri kullanılmıştır. Farklı öznitelik kombinasyonlarına göre kullanılan veri tabanı üzerinde %90'lara varan bir doğru sınıflandırma oranına erişilmiştir.

Anahtar Kelimeler: Cinsiyet tanıma, Ayrık Dalgacık Dönüşümü, Gri Seviye Eş Oluşum Matrisleri, Destek Vektör Makineleri

ABSTRACT

Gender Recognition With Facial Images

Studies on facial biometrics is vital as they are the driving force behind the development of applications in public spaces. By the help of these studies, researchers have been able to make gender recognition with higher accuracy rates. However, to perform these tasks with computers, real time and large scale applications are needed.

This study aims at which wavelet filters are more useful for extracting meaningful facial features by using frontal face images. In this study, an application using 103 wavelet filters with the help of frontal face images to perform better gender classification is developed. The accuracy rate of wavelet filters is tested using discrete wavelet transform and gray level co-occurrence matrix that produces 103 feature sets. As classifier, support vector machines are used.

Keywords: Gender Recognition, Discrete Wavelet Transform, Gray Level Co-Occurrence Matrix, Support Vector Machines.

ŞEKİLLER LİSTESİ

Sayfa No:

Şekil 1.1.Cinsiyeti tanımlamak için kullanılan geleneksel metotlarının şematik gösterimi ..2	
Şekil 1.2 Tez çalışması uygulama aşamalarının basit gösterimi.2	
Şekil 2.1 Önişleme aşamasının alt alanlarının şematik gösterimi4	
Şekil 2.2 Otomatik cinsiyet tanıma uygulamasının önişlem aşamaları5	
Şekil 2.3 a) Gri Seviyeye Dönüştürülmüş imge b) 3x3 ortalama alıcı filtre c) Gürültü indirgenmiş yumuşatılmış imge.....5	
Şekil 2.4 a) Gri seviye imge ve histogramı, b) Histogramı eşitlenmiş imge ve histogramı...6	
Şekil 3.1 Bir imge için GSEM'nin elde edilmesi. a) Orijinal İmge b) Griseviye Eş Oluşum Matrisi8	
Şekil 3.2 Bazı analiz yöntemleri arasındaki farklılıklar12	
Şekil 3.3 Tek boyutlu bir işaret için ayrık dalgacık dönüşümü.14	
Şekil 3.4 Üzerine gürültü eklenmiş bir sinyale ADD uygulanmasının şematik gösterimi[12]14	
Şekil 3.5 Dalgacık ayrıştırma ağacının genel gösterimi [12].15	
Şekil 3.6 Bir sinyale çoklu seviye ayrıştırması15	
Şekil 3.7 İşaret için ADD ayrıştırma ağacı ve farklı temsili16	
Şekil 3.8 2 Boyutlu bir imge işaretine uygulanan filtrelemenin şematik gösterimi16	
Şekil 3.9 Orjinal imgelere uygulanan filtreler ve oluşan yeni imgeler.17	
Şekil 3.10 a) Haar fonksiyonu, b) Orjinal İmge c) Yatay dalgacık sonucu d) Dikey dalgacık sonucu e) Diagonal dalgacık sonucu.17	
Şekil 3.11 Daubechies Dalgacık Örnekleri.18	
Şekil 3.12 a)Orjinal İmge b) Yatay db2 dalgacık sonucu c) Dikey db2 dalgacık sonucu d) Diagonal db2 dalgacık sonucu.....18	
Şekil 3.13 Kullanılan 14 adet biortogonal dalgacıklar.....19	
Şekil 3.14 a) Orjinal İmge b) Yatay bior1.5 dalgacık sonucu c) Dikey bior 1.5 dalgacık sonucu d) Diagonalbior 1.5 dalgacık sonucu20	
Şekil 3.15 Coiflet ailesindeki dalgacık tipleri20	
Şekil 3.16 a)Orjinal İmge b) Yatay Coif 1 dalgacık sonucu c) Dikey Coif 1 dalgacık sonucu d) DiagonalCoif 1 dalgacık sonucu.....20	
Şekil 3.17 Symlet ailesinden kullanılan dalgacıklar21	

Şekil 3.18 a) Orjinal İmge b) Yatay Sym5 dalgacık sonucu c) Dikey Sym5 dalgacık sonucu d) Diagonal Sym5 dalgacık sonucu	21
Şekil 3.19 Ters biortogonal dalgacıkların elde edilmesinin örnek gösterimi.....	22
Şekil 3.20 a) Orjinal İmge b) Yatay Rbior1.5 dalgacık sonucu c) Dikey Rbior 1.5 dalgacık sonucu d) DiagonalRbior 1.5 dalgacık sonucu	22
Şekil 4.1 Mesafeyi maksimize eden hiper düzlemin seçilmesi []......	27
Şekil 4.2 Verinin özellik uzayında ayrılması [25].	27
Şekil 5.1. Veritabanındaki erkek ve bayanlara ait örnek yüzler. (İlk satır: Erkekler, ikinci satır: Bayanlar).....	29
Şekil 5.2. 200 adet imgenin her birinin vektörle temsil edilerek kaynak VT'nin oluşturulmasının şematik gösterimi	29
Şekil 5.3. Tüm sınıflandırma denemelerinde kullanılan tek boyutlu indis değişkeni yapısının örnek gösterimi.	30
Şekil 5.4. 45 adet Daubechies dalgacık filtresi ile oluşturulan veri setinin sınıflandırma başarımları (db1 – db45)	33
Şekil 5.5. Coiflets dalgacık filtresi ile oluşturulan veri setlerinin sınıflandırma başarımları (coif1-coif5).....	33
Şekil 5.6. 23 adet Symlets dalgacık filtresi ile oluşturulan veri setinin sınıflandırma başarımları (sym2 – sym24)	34
Şekil 5.7. 15 adet Biorthogonal dalgacık filtresi ile oluşturulan veri setinin sınıflandırma başarımları	34
Şekil 5.8. Reverse Biorthogonal dalgacık filtresi ile oluşturulan veri setinin sınıflandırma başarımları	35
Şekil 5.9. Dalgacık filtre ailesi ile oluşturulan özelliklerin sınıflandırma başarımları ortalamaları.	35

TABLÖLAR LİSTESİ

Sayfa No

Tablo 3.1. Kullanılan dalgacık aileleri, tipleri ve sayıları.....	23
Tablo 5.1. Orijinal İmgelerden GSEM ile elde edilen özellik setinin sınıflandırma başarımı.	32

KISALTMALAR LİSTESİ

GSEM	: Gri Seviye Eş oluşum Matrisi
ADD	: Ayrık Dalgacık Dönüşümü
DVM	: Destek Vektör Makinaları
LANDSAT1	: Land Satellite 1
KSFD	: Kısa Süreli Fourier Dönüşümü
A	: Yaklaştırım
D	: Ayrıntı
CA	: Yaklaştırım katsayısı
CD	: Ayrıntı katsayısı
DB	: Daubechies dalgacığ
BİOR	: Biortogonal dalgacığ
COİF	: Coiflet dalgacığ
SYM	: Symlets dalgacığ
RBİOR	: Ters Biorthogonal dalgacığ
OKT	: Optik Karakter Tanıma
FEI	: Inaciana Eğitim Vakfı (<i>Fundação Educacional Inaciana</i>)

1. GİRİŞ

Sayısal görüntü işleme, sayısal imgenin bilgisayarlarla işlenmesi esasına dayanır. Görüntü zenginleştirme, uydu fotoğraflarının iyileştirilmesi ve savunma gibi pek çok alanlarda görüntü işleme teknikleri kullanılabilir. Sınıflandırma işlemi, sayısal görüntü işlemedeki aktif araştırma alanlarından biridir. Görüntülerin sınıflandırılmasında, bütün içerik, sınırlı sayıdaki sınıflara ayrılır [1].

Sunulan bu çalışmada, yüz imgelerinin iki farklı sınıfa ayrılmasında, sayısal görüntü işleme teknikleri kullanılmıştır. Birinci sınıf erkekler, ikinci sınıf ise bayanlar olarak işaretlenmiştir. Sınıflandırma başarımı ise, kaç kişinin cinsiyetinin doğru olarak işaretlendiğine bakılarak hesaplanmaya çalışılmıştır.

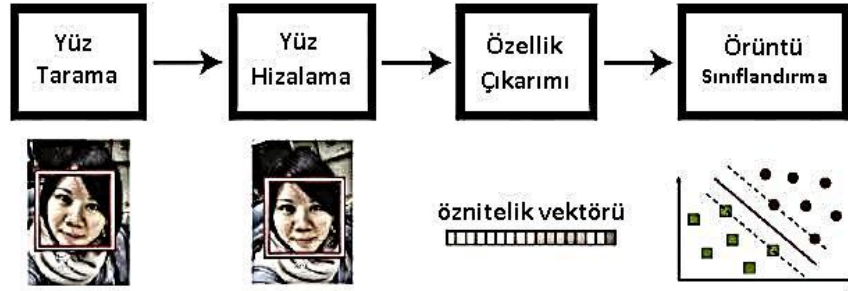
Sosyal aktiviteler ve iletişimde, cinsiyet sınıflandırmanın önemli bir rolü vardır. Günümüzde, cinsiyetin otomatik olarak sınıflandırılabilmesi çok faydalı ve büyük uygulama alanı potansiyeline sahiptir. Sinema, bina girişleri ve restoran gibi mekanlarda müşteri istatistikleri ve güvenlik taraması gibi alanlar otomatik cinsiyet tanıma işlevinin kullanım alanlarına örnek olarak verilebilir [2].

Cinsiyetin otomatik olarak tanınabilmesi bazı avantajlar sağlayabilir. Örneğin, binlerce insanın bulunduğu bir fotoğraf karesinde kaç erkek, kaç bayanın bulunabileceği hızlıca bulunmak isteniyorsa, cinsiyetin otomatik olarak sınıflandırılabilmesi bu işi kolaylaştırır.

Başka bir uygulama olarak, insanların yoğun olarak geçtiği kaldırımlar, metro girişleri gibi alanlara kameralar bırakılarak, günde kaç erkek kaç bayanın geçtiği saptanabilir. Hangi saat aralıklarında hangi cinsiyete sahip insanların daha fazla geçiş yaptığı gibi istatistiki bilgiler hızlıca tayin edilerek, o cinsiyetleri ilgilendiren reklamların bulunduğu değişebilen reklam afişleri hazırlanabilir.

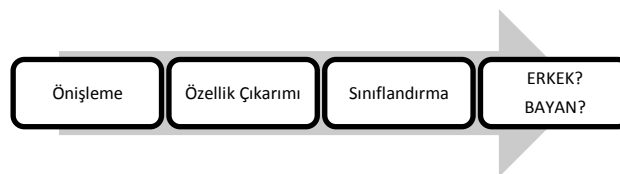
İnsan yüzündeki cinsiyete veya etnik kökene ait karakteristik özelliklerin hangileri olduğu konusunda çaba sarf edilmektedir. Sadece otomatik sistemler geliştirmek için değil, antropoloji açısından da bu konu önem arz etmektedir. Örneğin deri rengi, etnik köken

sınıflandırma çalışmalarında ayırıcı bir özellikken, cinsiyet sınıflandırma çalışmalarında etkisiz olabilmektedir. Sakal gibi özellikler ise bunun tersine örnektir [3]. Yüz imgelerinden cinsiyet sınıflandırma çalışmaları ile yüz tanıma çalışmalarının ortak özellikleri vardır. Bu ortak özellik Şekil 1.1’de de gösterilen çalışmalarda geleneksel olarak önileme, özellik çıkarımı ve sınıflandırma aşamalarının olmasıdır.



Şekil 1.1.Cinsiyeti tanımlamak için kullanılan geleneksel metodların şematik gösterimi [4].

Bu tez çalışması kapsamında karşı cepheden elde edilen 100 erkek ve 100 bayan fotoğrafından bilgisayar görmesi tekniklerini kullanarak cinsiyet tayini yapılmaya çalışılmıştır. Tez uygulamasında gerçekleştirilen işlemler kaba hatları ile Şekil 1.2’ de gösterilmiştir.



Şekil 1.2 Tez çalışması uygulama aşamalarının basit gösterimi.

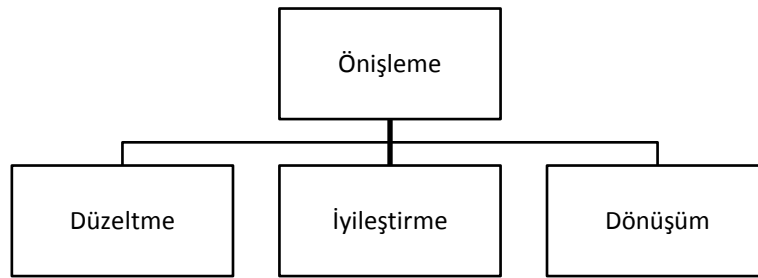
Karşı cephe imgelerin standart bir şekilde kullanımı için gereken Önileme aşaması Bölüm 2’de açıklanmıştır. Bölüm 3’de ise, özellik çıkarımında kullanılan yöntemlerden Gri Seviye Eş Oluşum Matrisi (GSEM), Ayırık Dalgacık Dönüşümü (ADD) ve bu dönüşümde kullanılan dalgacık aileleri konularına yer verilmiştir. Bölüm 4, özniteliklerin sınıflandırılmasında kullanılacak olan Destek Vektör Makineleri (DVM) anlatılmış, Bölüm 5’de ise çalışmada kullanılan materyal ve bu materyal üzerinden GSEM ile

öznitelik çıkarımı anlatılmıştır. Bölüm 6’da, imgelere ADD uygulandıktan sonra GSEM ile çıkarılan özeniteliklerin sınıflandırılması anlatılmış, Bölüm 7’de, uygulamalar sonrasında elde edilen sonuçlar sunulup ADD’nin cinsiyete ait öznitelikleri çıkarmadaki etkisi ve dalgacık ailelerinin, başarımdaki rolleri yorumlanmaya çalışılmıştır.

2. ÖNİŞLEME

Sayısal imge incelenip sayısal olarak yorum yapılmak istendiğinde imgenin özniteliklerinin çıkarılması için bazı önışlemlerden geçirilmesi gerekir. Eğer, tasarlanacak sistemde bir öğrenme süreci gerçekleştirecekse, kullanılacak yöntemeye dayalı olarak resim bazı önışlemlerden geçirilmelidir. Örneğin, bir harf tanıma sistemi düşünıldüğünde, eğitim verileri düzenli olarak yataya dik harflerden oluştuğu, harflerin eğimli olması gibi bir durumun olmadığı şartlarda sistem çalışacak şekilde tasarlanacaksa, eğitim kümesindeki harfler belli bir kurala göre hizalanıp düzenli bir veritabanı oluşturulması sağlanır.

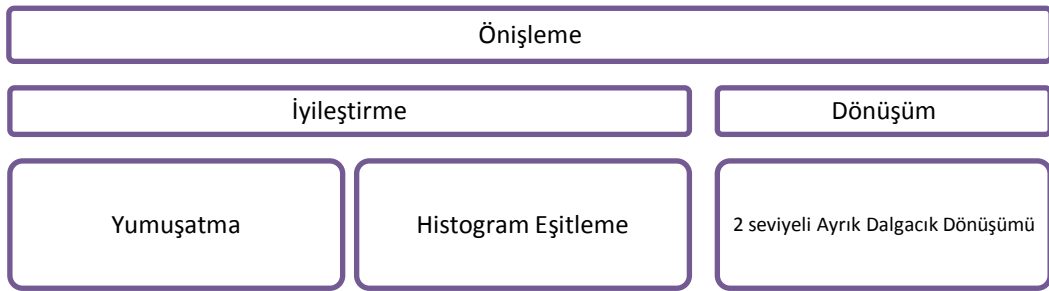
Sayısal görüntü önışleme aşaması; bilgisayar yardımı ile imgelerin amaca göre iyileştirilmesi ve yorumlanmasını içerir. Bir sayısal imgenin önışleme aşamasında imge, bazı matematiksel işlemlere girer. Bu matematiksel işlemler sonrasında yeni bir görüntü oluşur. Oluşan bu yeni görüntü depolanarak bundan sonraki işlemler için yeni görüntüler üzerinde çalışılır. Önışleme aşaması, kendi içinde farklı aşamalara ayrılabilir. Bu aşamalar Şekil 2.1’ deki gibi üç sınıfta değerlendirilebilir. Görüntü verisini “düzeltme” (*Restorasyon*), daha iyi görsel yorumlama için “iyileştirme”, görüntüden yeni görüntü elde edilebilmesini sağlayan “dönüşüm” işlemleri, önışleme sürecinin alt aşamalarıdır.



Şekil 2.1 Önışleme aşamasının alt alanlarının şematik gösterimi

Otomatik cinsiyet tanıma sisteminin oluşturulması için bu tez kapsamında gerçekleştirilen uygulamada, karşı cephe renkli imgeler kullanılmıştır. Bütün imgeler,

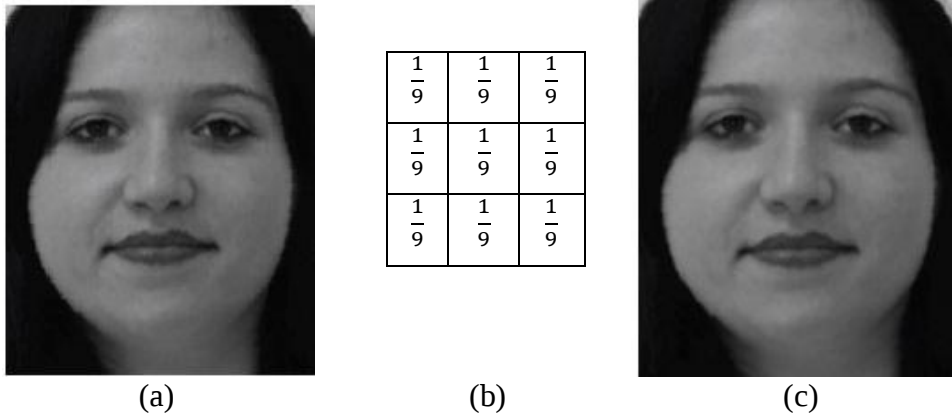
görüntünün anlamlı bir şekilde işlenebilmesi için gri seviyeye dönüştürme, yumuşatma ve histogram eşitleme işlemleri “iyileştirme” aşamasını kapsayacak şekilde gerçekleştirilmiştir. Ayrıca önileme aşamasında farklı oryantasyonda olan imgelere dönüşüm uygulanarak dikey pozisyona getirilmiştir. Daha sonra ise öznitelik çıkartım aşamasına geçilmiştir. Öznitelik çıkartım aşamasında ise ADD kullanılmıştır. Bu aşamaların şematik gösterimi Şekil 2.2’de gösterilmiştir.



Şekil 2.2 Otomatik cinsiyet tanıma uygulamasının önileme aşamaları

2.1. Yumuşatma

Renkli imgeler, gri seviyeye çevrilerek Şekil 2.3a’daki imge elde edilir. Resmin gri seviyeye dönüştürülmesinden sonra ortalama filtresi (Şekil 2.3b’de yapısı verilen ortalama alıcı filtresi) kullanılarak Şekil 2.3c’de verilen yumuşatılmış imge elde edilir. Böylelikle komşu pikseller arasındaki geçişler yumuşatılmış, resimdeki gürültü giderilmiş olur.



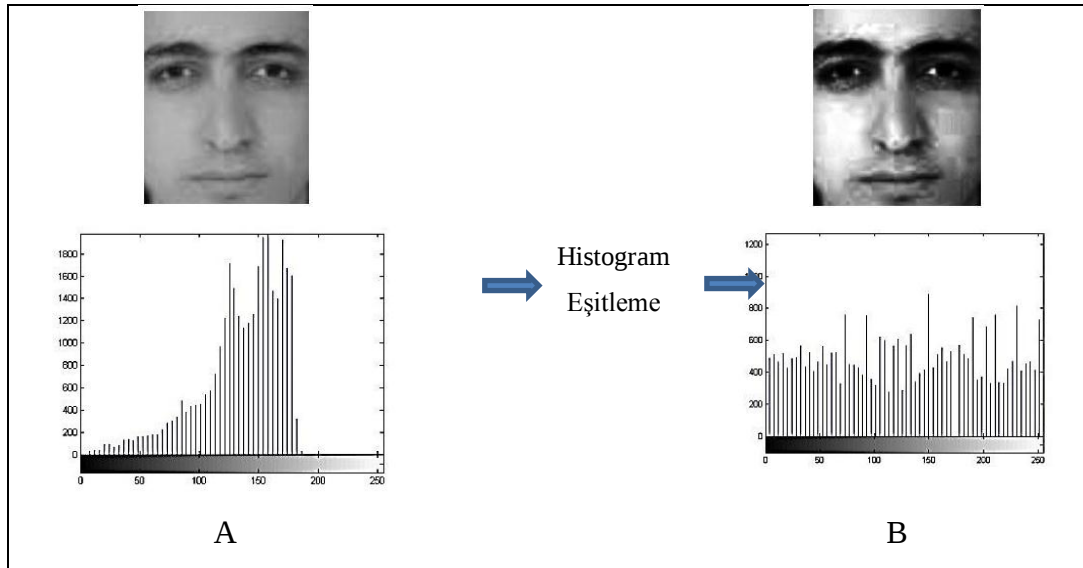
Şekil 2.3 a) Gri Seviyeye Dönüştürülmüş imge b) 3x3 ortalama alıcı filtre c) Gürültü indirgenmiş yumuşatılmış imge

2.2. Histogram Eşitleme

Histogram, sayısal bir imge içerisinde her gri seviye değerinden kaç adet olduğunu gösteren grafikdir. Bu grafiğe bakılarak resmin parlaklık durumu ya da tonları hakkında bilgi sahibi olunabilir.

Histogram eşitleme, kontrast zenginleştirme yöntemlerinden biridir. Histogram eşitlemesi ile görüntüde parlaklık en çok toplandığı alana en büyük kontrast zenginleştirilmesi uygulanmaktadır. Deneyimli bir görüntü analizcisi, genellikle görüntü histogramını inceleyerek ve yeterli sonuçlar elde edinceye kadar deneyerek, uygun kontrast zenginleştirme algoritmasını belirleyebilir [5].

Histogram eşitleme ile belirli bir ton etrafında toplanan histogram eğrisi 0-255 tonları arasına düzgün bir şekilde dağıtılır böylece resmin gri ton dağılımının homojen olarak yapılandırılması sağlanır. Histogram eşitleme işleminde, resmin kümülatif gri seviye dağılım skalası üzerinde normal dağılım uygulanmaktadır. Bu yeniden dağılım, gri seviye dağılımında dengeleme sağlamaktadır.



Şekil 2.4 a) Gri seviye imge ve histogramı, b) Histogramı eşitlenmiş imge ve histogramı

Açıklanan önışlem basamakları tüm veritabanına uygulanarak resimlerin standart bir hale gelmesi sağlanmıştır.

3. ÖZNETELİK ÇIKARIM YÖNTEMLERİ

Öznitelik çıkartımı, imge sınıflandırma sistemlerinde en önemli aşamadır. Bu aşamada imgeyi temsil edecek mümkün olduğunda az boyutlu öznitelikler belirlenmeye çalışılır. Özniteliklerin doğru belirlenmesi sınıflandırma başarısını doğrudan etkileyen en önemli çalışma alanıdır. Özniteliklerin belirlenmesi temelde iki şekilde yapılır. Birincisi yapısal yaklaşım ikincisi ise dokusal yaklaşımdır. Yapısal yaklaşım özellikle ikili imgeler (siyah-beyaz) için uygundur. Yapısal yaklaşımda imge üzerinde nesneyi tanımlayabilecek noktalar belirlenerek bu noktalar üzerinden öznitelik çıkartımı yapılır. İmge sınıflandırması için bu noktalardan elde edilen özniteliklerin karşılaştırılması yapılır. Dokusal yaklaşımda ise herhangi bir nokta belirlemesi yoktur. İmge içerisindeki bilgi bütünüyle çıkartılmaya çalışılır. Uzaysal bölgede çalışan (GSEM gibi) yöntemler olduğu gibi frekans incelemesi yapan (Ayrık Fourier Dönüşümü ve Ayrık Kosinüs Dönüşümü gibi) yöntemlerde literatürde önerilmiştir.

Belirlenen özniteliklerin uzunluğu da önemlidir. Öznitelik vektörünün çok uzun olması doğru tanımayı artıracığı anlamı taşımamaktadır. Ayrıca öznitelik vektörünün çok uzun olması çalışma zamanını artıracığından sistemin uygulanabilirliğini düşürecektir. Bu nedenle öznitelik vektörünü kısaltmak için bir boyut indirgemesi yapılabilir.

Veri seti içindeki bazı sütunlar, eğitim ve test kümelerinin oluşturulmasında verinin geneli hakkında anlamlı bilgi içermeyebilir. Hatta bazı veriler, verinin anlamlandırılmasını engelleyebilir [6].

Öznitelik çıkarımından önce, bir veri indirgemesi de yapılabilir. Çalışma kapsamındaki uygulamamızda resimler üzerinde sınıflandırma işlemi yapılması amaçlanmaktadır. Her resim 193x192 boyutunda bir matris olduğu düşünüldüğünde 200 resim 200x31266 yaklaşık 6 milyon kadar sayıdan oluşmaktadır. Bunun yerine 193x162 ebatlarındaki bir imge üzerinden anlamlı verilerin çekilmesi gerekmektedir. Bu çalışmamızda özellik çıkarım yöntemlerinden biri olan GSEM kullanılarak anlamlı öznitelikler bulunmaya çalışılmıştır.

3.1. Gri Seviye Eş-Oluşum Matrisi

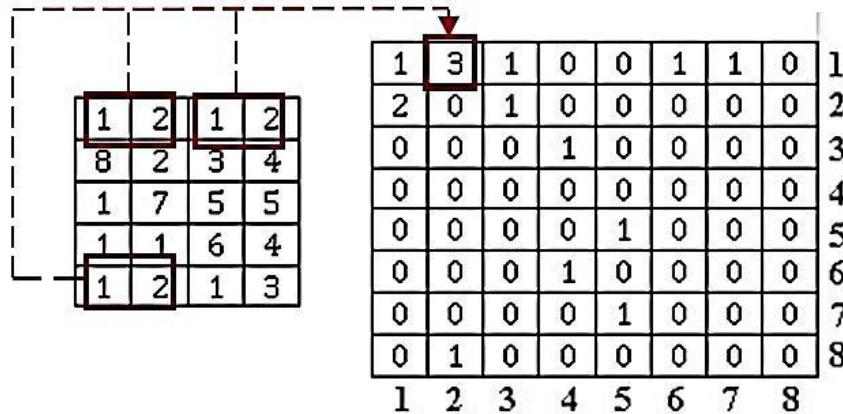
Gri seviye eş-oluşum matrisi (GSEM), Haralick ve arkadaşları tarafından 1973'te tanımlanmıştır. GSEM yardımı ile elde ettiği öznitelikleri, kumtaşı imgelerinin, pankromatik uzay fotoğraflarının ve LANDSAT1 uydusundan elde edilen görüntülerin bulunduğu 3 farklı veritabanı üzerinde sınıflandırma işlemi ile test etmiştir [7].

GSEM, ikinci derece istatistikleri kullanan en çok bilinen doku analiz yöntemlerinden biridir. GSEM, bir imge içerisindeki piksellerin gri seviyelerinin hangi gri seviyeden sonra oluştuğuna ilişkin bir matris oluşturmaktadır. Daha sonra bu matris üzerinde istatistiksel fonksiyonlar kullanılarak imgeyi ifade eden öznitelikler elde etmektedir. GSEM oluşturulması şu şekilde açıklanabilir: Oluşturulacak GSEM'deki (i, j) indisindeki değeri, imge üzerinde belirli bir pencere içindeki, θ yönündeki d komşu uzaklığına sahip değeri i ve j olan iki pikselin frekansını ifade eder. d genellikle 1 ya da 2 olarak seçilir. θ ise $0^\circ, 45^\circ, 90^\circ$ ve 135° olmak üzere 4 adet yön değeri alabilir.

GSEM(Gri Seviye Eş-oluşum Matrisi)'nin her elemanı şu şekilde hesaplanır;

$$P_{(i,j)} = \frac{P(i, j, d, \theta)}{\sum_{i=1} \sum_{j=1} P(i, j, d, \theta)} \quad (3.1)$$

Burada, $P(i, j, d, \theta)$, θ yönünde, aralarındaki komşuluk uzaklığı d olan iki pikselin (piksel değeri i olan ve piksel değeri j olan) frekansını belirtmektedir [8]. Şekil 3.1'de bir gri seviye imgeden GSEM oluşturulması gösterilmiştir.



Şekil 3.1 Bir imge için GSEM'nin elde edilmesi. a) Orijinal İmge b) Griseviye Eş Oluşum Matrisi

Şekil 3.1'deki örnekte, $d = 1$, $\theta = 0^\circ$ olduğu durumda imgede sıraya (1, 2) değerlerinin oluşması ve böylesi oluşumun tekrarlanma sıklığının 3 olduğu görülmektedir. GSEM, imgelerin gri seviye sayısını, yönlerini ve dağılım değişimleri gibi gri bilgileri sentezleyerek yansıtır [9]. GSEM'den elde edilen değerler öznitelik çıkarımı için kullanılabilir. GSEM tabanlı imge doku özniteliği üzerine Haralick ve arkadaşları, GSEM'nin nicel açıklama yöntemini teşkil eden 14 çeşit ayırt edici özellik önermişlerdir [7,10]. Sunulan bu çalışmada GSEM'den dört istatistiksel değer hesaplanarak öznitelik vektörleri oluşturulmuştur. Bu istatistiksel değerleri hesaplamak için enerji, homojenlik, zıtlık ve korelasyon fonksiyonları kullanılır.

3.1.1. Enerji

Bu parametre, açısız ikinci moment ve tekdüzellik kavramları ile de ifade edilmektedir. Enerji, piksel çiftlerinin tekrarları gibi dokusal tekdüzeliği ölçmektedir. Yüksek enerji değerleri, gri seviye dağılımının periyodik olduğunu gösterir. Enerji; Eşitlik 3.2.'de belirtildiği gibi ifade edilir.

$$ene = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} g^2(i, j) \quad (3.2)$$

Enerjinin; zıtlık ve varyans gibi birincil istatistiki değişkenler ile korelasyonu oldukça düşüktür. Enerji ya da açısız ikinci moment benzerliği, görüntü homojenliğinin bir ölçüsü olup görüntü ne kadar homojen bir yapıya sahipse değeri o kadar artar [11].

3.1.2. Zıtlık

Uzaysal frekanslar, bir ardışık piksel kümesinin en yüksek ve en düşük değerleri arasındaki farktır. Bu tanım GSEM, zıtlık (Contrast) kavramını iyi ifade etmektedir. GSEM zıtlığı, uzaysal frekansla sıkı bir korelasyon gösterir.

$\Delta_{con} = (i - j)^2$ olmak üzere;

$$con = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} (i - j)^2 \cdot g(i, j); = \overline{\Delta_{con}} \quad (3.3)$$

Zıtlık değeri, Denklem 3.3.'te belirtildiği gibi hesaplanır. GSEM zıtlık ve varyans değer çiftleri için, düşük zıtlığa sahip imgelerin varyanslarının da düşük olacağı söylenebilir. Fakat bunun tersi her zaman doğru değildir [10].

3.1.3. Homojenlik

Homojenlik değeri;

$$\sum_{i,j} \frac{p(i,j)}{(1+|i-j|)} \quad (3.4)$$

ile ifade edilebilir.

Görüntünün farklı bölgelerinde benzerlik varsa homojenlik değeri artar.

Korelasyon:

$\Delta_{cor} = (i - \mu) \cdot (j - \mu)$ olmak üzere;

$$cor = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} (i - \mu) \cdot (j - \mu) \cdot g(i, j) / \sigma^2; = \frac{\overline{\Delta_{cor}}}{\sigma^2} \quad (3.5)$$

GSEM korelasyonu, i ve j rastgele değişken arasındaki korelasyon katsayısıdır. Korelasyon; imge içerisindeki lineer bağımlılık olup, görüntünün doğrusal bir yapıya sahip olup olmadığını ölçmek için kullanılır [10].

GSEM üzerinden hesaplanan bu istatistiksel değerler kullanılarak bir imgeyi tanımlayan öznitelik vektörü oluşturulabilir. İmgelere ait öznitelik vektörleri arasındaki uzaklıklar ölçülerek benzerlik miktarı ortaya çıkartılabilir.

3.2. Ayrık Dalgacık Dönüşümü

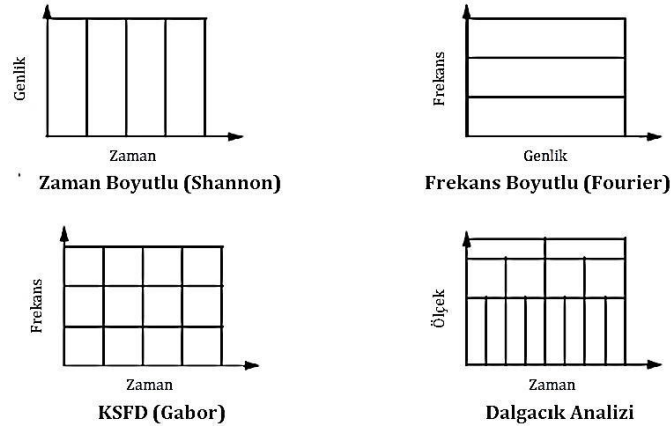
İşaret analizinde şimdiye kadar farklı yöntemler kullanılmış ve bu yöntemlerden en çok bilineni Fourier Dönüşümü'dür. Fourier Dönüşümü, basit olarak sinyali matematik formülasyonlar yardımıyla zaman bölgesinden frekans bölgesine dönüşümünü sağlamaktadır.

Birçok sinyal için Fourier Dönüşümü, frekans bileşenlerinin öneminden dolayı faydalıdır. Fakat, Fourier Dönüşümü'nün bazı dezavantajlarından dolayı, dalgacık analizi gibi diğer tekniklere ihtiyaç duyulmaktadır.

Frekans alanına çevrilme işleminde zaman bilgisi kaybedilir. Bir sinyalin Fourier dönüşümüne bakıldığında, belirli bir olayın ne zaman gerçekleştiğini bilmek imkânsızdır. Eğer bir işaret, zaman içinde çok değişmiyorsa (durağan işaret olarak adlandırılır) bu durum çok önemli değildir. Ancak, birçok işaret, durağan olmayan veya geçici karakteristikleri (kaymalar, ani değişiklikler, olay başlangıç ve sonları) barındırır.

Bu karakteristikler, genellikle işaretin en önemli parçalarıdır ve Fourier dönüşümü bunları tarayıp bulmak için uygun değildir. Bu eksikliği gidermek için, Dennis Gabor, 1946'da bir sinyalin sadece belirli bir zamandaki parçasını analiz ederek Fourier dönüşümünü adapte etmiştir. Bu teknik sinyali “pencereleme” diye adlandırılmıştır. Gabor'un Kısa Süreli Fourier Dönüşümü (Short Time Fourier Transform) diye adlandırılan uyarlamasında sinyal, zaman ve frekans olarak iki boyuttaki fonksiyonla temsil edilmiştir [12].

Dalgacık analizi, değişken boyuttaki bölgeler ile gerçekleştirilen bir pencereleme tekniği olarak Fourier ve Kısa Süreli Fourier analizlerinden bir sonraki mantıksal aşamayı gerçekleştirir. Dalgacık analizi ile, düşük frekans bilgilerine daha fazla dikkat edilmesi, istenilen yerlerde uzun zaman aralıklarının kullanılabilmesi ve yüksek frekans bilgilerine daha fazla dikkat edilmesi istendiğinde de kısa zaman aralıklarının kullanılabilmesi imkânı sağlanmaktadır.



Şekil 3.2 Bazı analiz yöntemleri arasındaki farklılıklar

Dalgacık, sınırlı bir süre içerisinde etki gösteren ortalama değeri sıfır olan bir dalga şeklindedir. Muhtemel tüm ölçeklerdeki dalgacık katsayılarının hesaplanması, dalgacık analizinin büyük bir kısmını oluşturur. Bu işlem sonrasında birçok anlamsız veri üretilmiş olur. Hesaplamalar yapılırken sadece belli bir ölçek veya pozisyon kümesinin seçilmesi de mümkündür.

Şema kullanarak filtrelemenin etkili bir yolu, 1988’de Mallat tarafından geliştirilmiştir. Mallat’ın algoritması, sinyal işleme çevreleri tarafından, “iki kanallı alt band kodlayıcı” diye bilinen klasik bir şemadır.

Bu pratik filtreleme algoritması, hızlı bir dalgacık dönüşümü sağlar. Bu algoritma, işaretin içinden geçtiği, dışından da dalgacık katsayılarının üretildiği bir kutu gibi düşünülebilir.

Birçok işaret için, düşük frekanslı bileşenler en önemli parçalardır. Düşük frekanslı bileşenler, işaretin kimliğini temsil eder. Diğer taraftan, yüksek frekanslı bileşenler ise, işaretin ince ayrıntılarıdır. İnsan sesi dikkate alındığında, yüksek frekanslı bileşenleri silindiğinde, konuşma sesi değişir, fakat ne söylenildiği halen anlaşılır. Bununla birlikte, düşük frekanslı bileşenlerin yeterince silinmesi durumunda, anlaşılamaz seslere dönüşür [12].

Dalgacık, sınırlı bir sürenin dalga formudur. Dönüşüm ise sinyalin farklı bir şekilde sunulmasıdır. Dalgacık dönüşümü, sinyallerin farklı çözünürlüklerde analiz edilmesi esasına dayanır (*MRA-Multi Resolution Analysis*). Ayırık dalgacık dönüşümü ve sürekli dalgacık dönüşümü, dalgacık dönüşümü metodlarından ikisidir. Bu dalgacık dönüşümleri,

bazı kaydırma ve genişleme operasyonlarına tutulan ana dalgacık fonksiyonundan elde edilmektedir [13].

$$\Psi_{a,b}(t) = \frac{1}{\sqrt{a}} \Psi[(t - b)/a] \quad (3.6)$$

a ölçek parametresine, b kaydırma parametresine sahip ana dalgacık fonksiyonu kabaca Denklem 3.6’da verilmiştir [13].

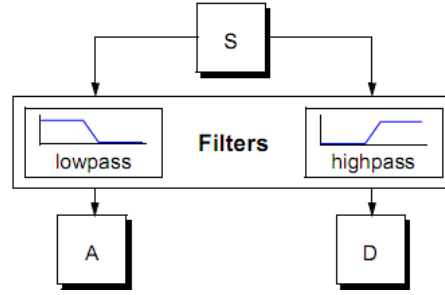
1 Boyutlu Ayırık Dalgacık Dönüşümü aşağıdaki eşitliklerle verilir [14].

$$W_{\varphi}(j, k) = \frac{1}{\sqrt{M}} \sum_n f(n) \varphi_{j,k}(n) \quad (3.7)$$

$$W_{\psi}(j, k) = \frac{1}{\sqrt{M}} \sum_n f(n) \psi_{j,k}(n) \quad (3.8)$$

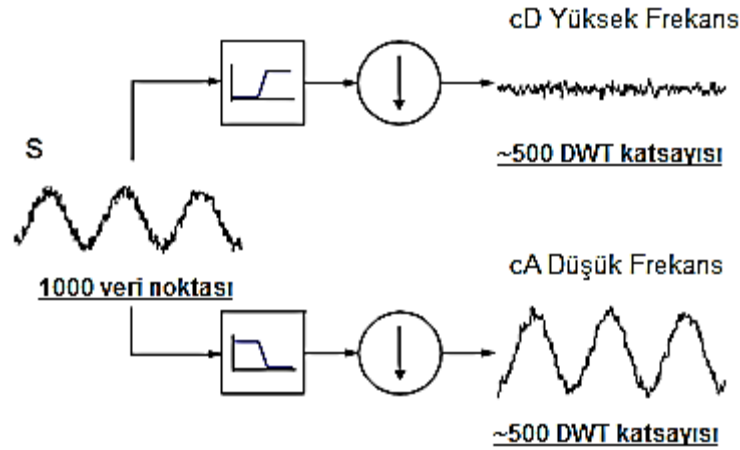
Denklem (3.1)’deki $W_{\varphi}(j, k)$ ifadesi dalgacık dönüşümündeki yaklaştırım (A) ifadesi ile eşdeğerdir. Denklem (3.2)’deki $W_{\psi}(j, k)$ ifadesi, ayrıntı (D) değeri ile eşdeğerdir. Denklem(3.1)’deki $\psi(n)$, dalgacık fonksiyonunu ifade etmektedir. Denklem (3.2)’deki $\varphi(n)$ ise ölçek fonksiyonunu (*scaling function*) ifade etmektedir [14].

Dalgacık dönüşümünde bu yüzden sıklıkla “*yaklaşım*” (Approximations) ve “*ayrıntı*” (Details) değerlerinden bahsedilir. Yaklaştırım değerleri; işaretin, yüksek ölçekli, düşük frekanslı bileşenleridir. Ayrıntılar ise işaretin, düşük ölçekli, yüksek frekanslı bileşenleridir. Filtreleme işleminin en basit gösterimi Şekil 3.3’te mevcuttur. Bu gösterimde, standart bir kabul gördüğü için *yaklaştırım* ve *ayrıntı* terimlerinin İngilizcedeki karşılıkları olan “Approximation” ve “Detail” terimlerinin baş harfleri kullanılmıştır. Bu tez çalışmasının bundan sonraki kısımlarındaki şematik gösterimlerde, Yaklaşım için “A”, Ayrıntı için “D” harfleri kullanılacaktır.



Şekil 3.3 Tek boyutlu bir işaret için ayrık dalgacık dönüşümü.

Örneğin, üzerine yüksek frekansta gürültü eklenmiş basit bir sinüs dalgası (S) üzerinde dalgacık dönüşümünün etkileri Şekil 3.4’da görülmektedir.

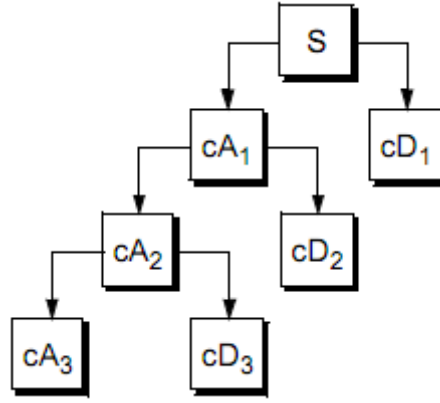


Şekil 3.4 Üzerine gürültü eklenmiş bir sinyale ADD uygulanmasının şematik gösterimi[12]

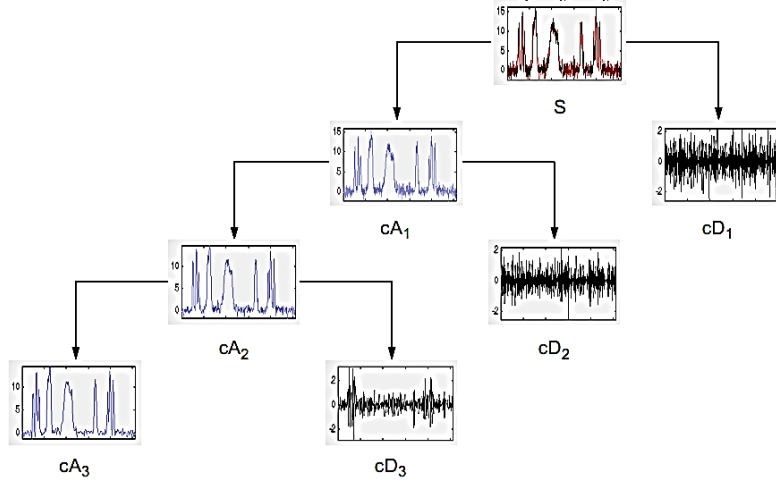
Şekil 3.4’te, ayrıntılar katsayısı olan cD, yüksek frekansta gürültü içerirken, yaklaşım değerleri katsayısı olan cA, orijinal sinyale göre daha az gürültü içerir.

Ayrık Dalgacık Dönüşümü’nde, yüksek frekans bantlarındaki katsayıların büyük olanları imgelerde genellikle kenarları temsil eder. Bu alanlar insan gözü tarafından zor farkedilir. Bundan dolayı bu alanlara gizli bilgiler gömülebilir (*watermarking*) [15].

Ayrıştırma işlemi, ardışıl yaklaşım değerleri üzerinden tekrarlanabilir. Böylelikle, bir sinyal düşük çözünürlüklü çokça bileşene ayrıştırılabilir. Buna “*dalgacık ayrıştırma ağacı*” denir [12].

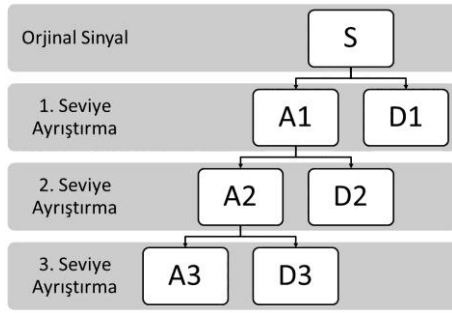


Şekil 3.5 Dalgacık ayrıştırma ağacının genel gösterimi [12].

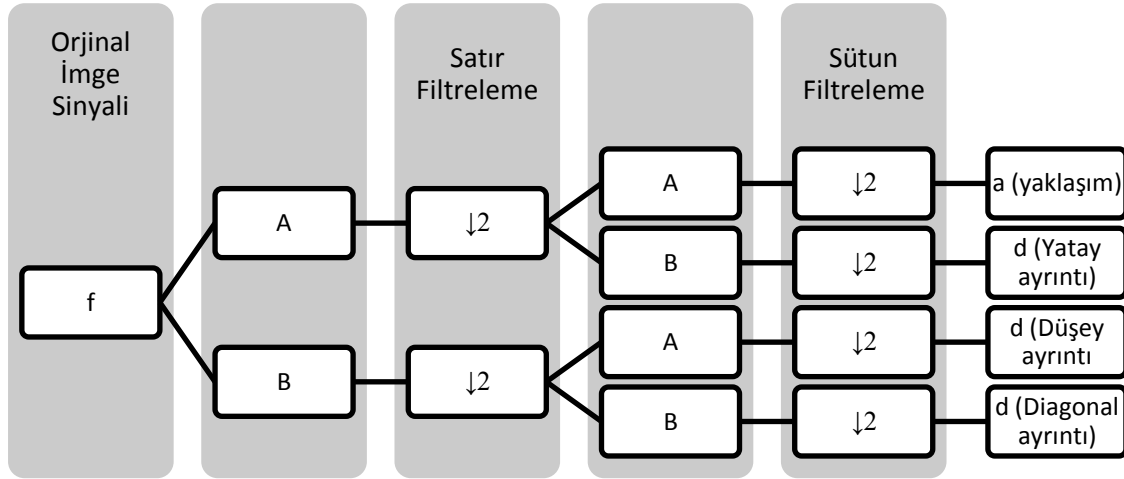


Şekil 3.6 Bir sinyale çoklu seviye ayrıştırması

Dalgacık analizinde, bir sinyal yaklaşım ve detay diye iki kısma ayrılır. Oluşan yaklaşıma aynı bölme işlemi tekrar yapıp ikinci seviye yaklaşım ve ikinci seviye detay elde edilir. Bu işlem sürekli olarak tekrar edilir. n - seviyeli bir ayrıştırma için, sinyali tekrar elde etmenin $n + 1$ çeşit yolu vardır [12].



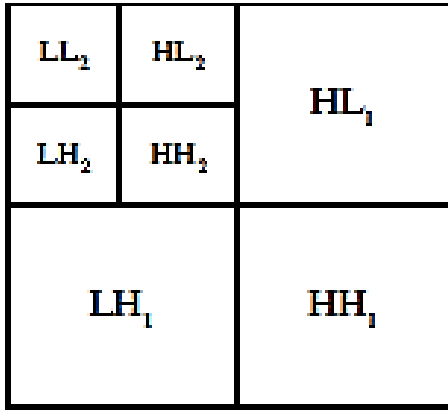
Şekil 3.7 İşaret için ADD ayrıştırma ağacı ve farklı temsili



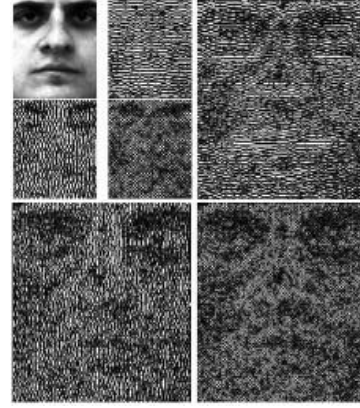
Şekil 3.8 2 Boyutlu bir imge işaretine uygulanan filtrelemenin şematik gösterimi

Şekil 3.8’de 2 boyutlu bir sinyale dalgacık dönüşümünün uygulamasının şematığı verilmiştir. İmgeler 2 boyutlu oldukları için bu çalışmada kullanılan dalgacık dönüşümü tekniğinin yukarıdaki gibi olduğu söylenebilir. A ve B harflerinin bulunduğu kutucuklar, ölçekleme ve dalgacık fonksiyonlarını göstermektedir. Örneğin, A ve B harfleri, sırasıyla alçak geçiren ve yüksek geçiren filtrelerdir [13]. Bu filtreler, sinyali Şekil 3.8’ de gösterilen yaklaşım (a) ve ayrıntı (d) katsayılarına dönüştürmektedir. Her bir dönüşüm uygulandıktan sonra oluşan imge, bir öncekinin yarısı kadar olur.

Bu tez çalışması kapsamında ADD, iki seviyede gerçekleştirilmiş, her bir ayrıntı imgeleri için yatay, dikey ve diagonal olmak üzere 3 adet imge elde edilmiştir. Şekil 3.9’ da gösterildiği gibi orijinal imgeden 7 adet imge elde edilmiştir.



(a)



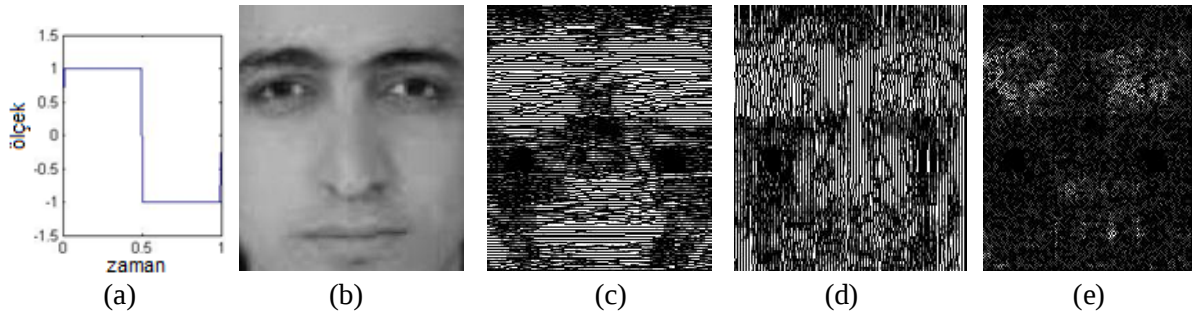
(b)

Şekil 3.9 Orjinal imgelere uygulanan filtreler ve oluşan yeni imgeler.

3.2.1. Haar Dalgacı

En fazla bilinen en temel dalgacık türü Haar dalgacıdır. Haar, kesikli adım fonksiyonuna benzer bir yapı sergiler. Daubechies dalgacık ailesinin en düşük dereceli db1 dalgacı ile aynı sonucu verir.

Haar dalgacı en eski ve en basit ortonormal dalgacıdır. Tamamen özyinelemeli(*recursive*) olarak çalışır. Haar iki komşu pikselin değişimini yansıtır [13].



Şekil 3.10 a) Haar fonksiyonu, b) Orjinal İmge c) Yatay dalgacık sonucu d) Dikey dalgacık sonucu e) Diagonal dalgacık sonucu.

3.2.2. Daubechies

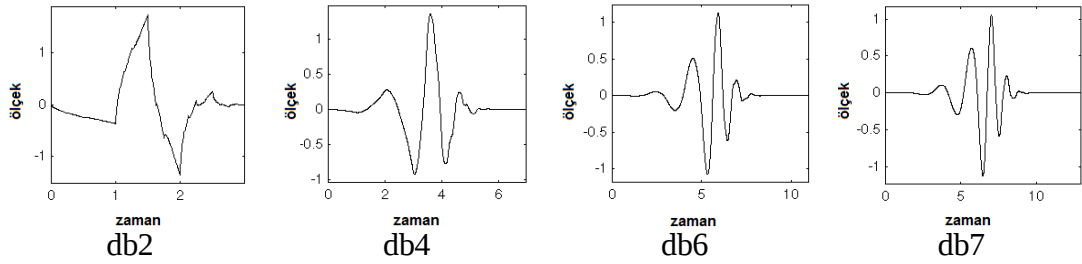
Ingrid Daubechies, dalgacık araştırmacıları çevrelerince önemsenen Belçikalı bilim kadınıdır. Daubechies, ayrık dalgacık analizini kolayca yapmaya yarayan tam destekli ortonormal dalgacıklar, keşfetmiştir [12].

Daubechies ailesinde, 2 ile 20 arasında deęerlerle numaralandırılmış ortogonal dalgacıklar mevcuttur [16].

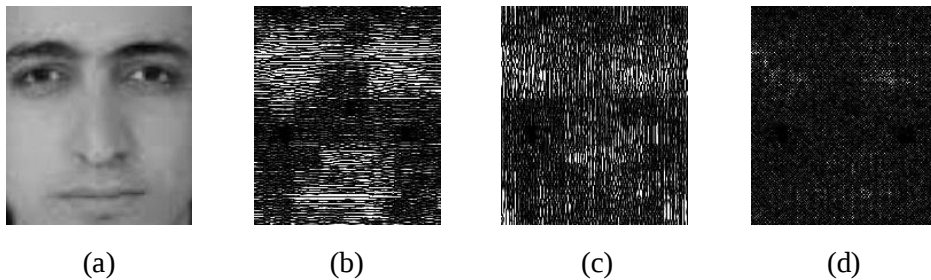
Daubechies ailesindeki dalgacıklar, sıralamalarına göre git gide karmaşıklıřır. 4 ve 6 numaralı Daubechies dalgacıkları en çok kullanılan dalgacıklardır. Bu dalgacıklarda en önemli anahtar kelime moment noktası (*vanishing moment*) kavramıdır. 2 nolu Daubechies dalgacığının moment noktası bir tanedir, 3 nolu Daubechies dalgacığının ise 2 adet moment noktası vardır, 4 nolu dalgacığının ise 3 adet moment noktası vardır. Bu böylece devam etmektedir. Yüksek sayıdaki moment noktaları, sinyalin düzenli parçalarını daha iyi sıkıştırır [17].

Daubechies ailesine mensup dalgacıklar dbN şeklinde gösterilir. N , dalgacığının sırasını, db ise dalgacığının soyadını belirtir. $db1$ bir önceki başlıkta belirtildięi gibi *Haar* ile aynıdır.

Şekil 3.11’de Daubechies dalgacık ailesine ait bazı örnekler verilmiştir.



Şekil 3.11 Daubechies Dalgacık Örnekleri.

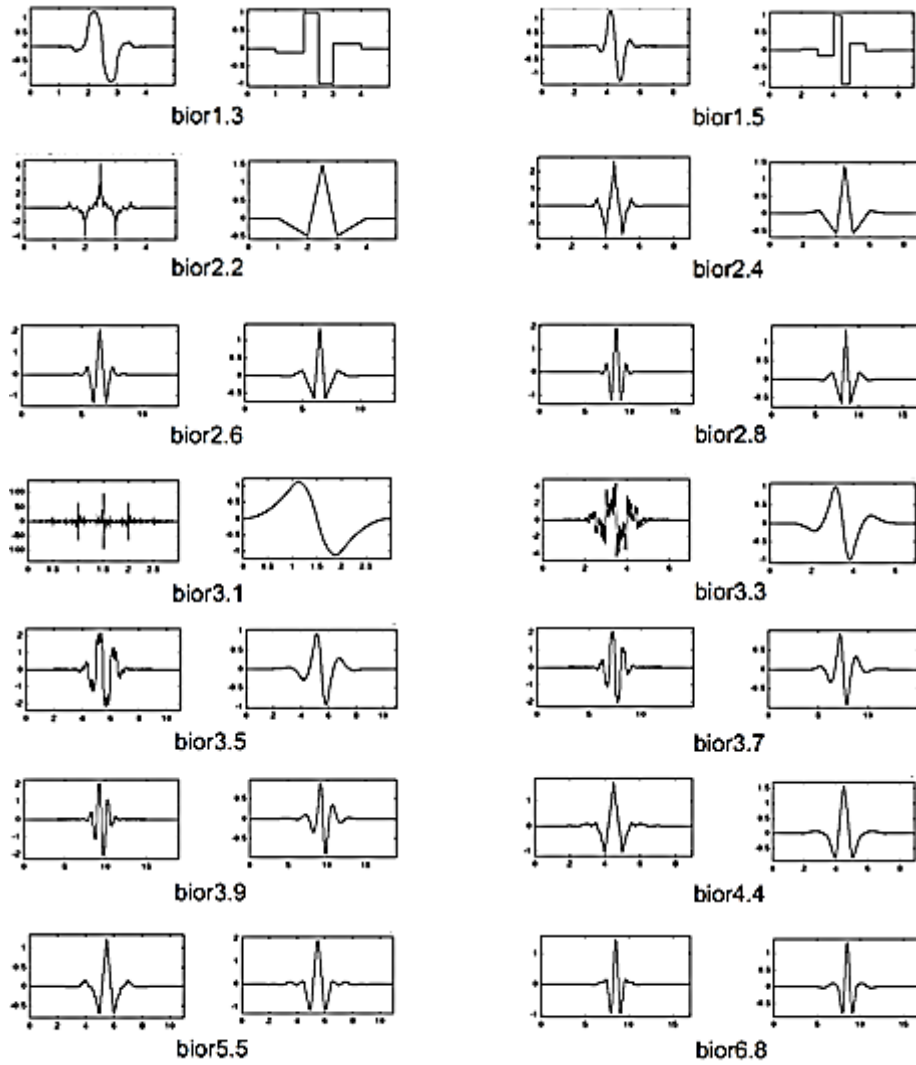


Şekil 3.12 a)Orjinal İmge b) Yatay db2 dalgacık sonucu c) Dikey db2 dalgacık sonucu d) Diagonal db2 dalgacık sonucu

Bu tez çalışması kapsamında, $db1$ (Haar) den $db45$ ’e kadar 45 adet Daubechies filtresi kullanılmıştır.

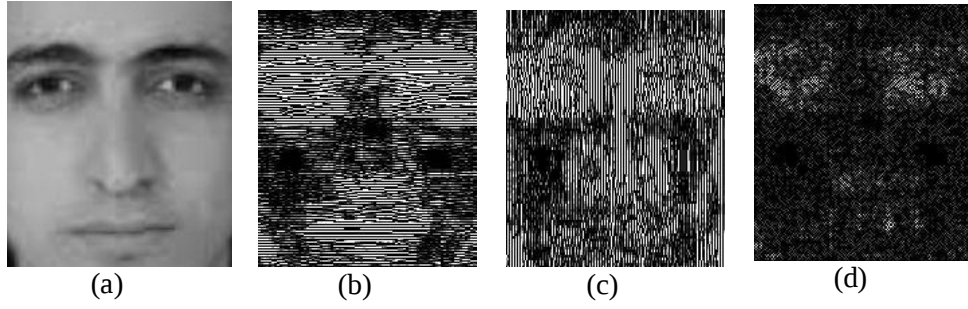
3.2.3. Biortogonal

Bu dalgacık ailesi, sinyalin ihtiyaç duyulan lineer faz özelliklerini açığa çıkarır. Bu dalgacık ailesinde iki dalgacık kullanılır. Birinci dalgacıkla sinyal parçalanıp, ikincisi ile tekrar toparlanarak, sinyalin ilginç özellikleri çıkarılır [12].



Şekil 3.13 Kullanılan 14 adet biortogonal dalgacıklar

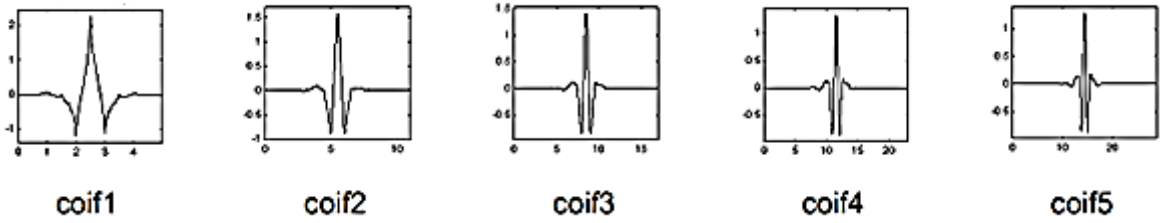
Bu tez çalışmasında, biortogonal ailesinden bior1.1 filtresi ve Şekil 3.13’de belirtilen filtreler kullanılarak 15 adet biortogonal dalgacık filtresi kullanılmıştır.



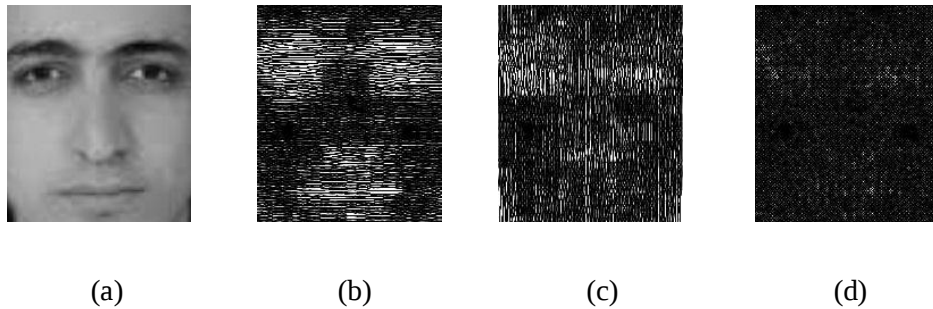
Şekil 3.14 a) Orjinal İmge b) Yatay bior1.5 dalgacık sonucu c) Dikey bior 1.5 dalgacık sonucu d) Diagonalbior 1.5 dalgacık sonucu

3.2.4. Coiflet

R. Coifman'ın ricası üzerine, I. Daubechies tarafından geliştirildi. Dalgacık fonksiyonu $2N$ momente, ölçek fonksiyonu ise $N-1$ momente sahiptir. Bu momentlerin toplamı 0'a eşittir.



Şekil 3.15 Coiflet ailesindeki dalgacık tipleri

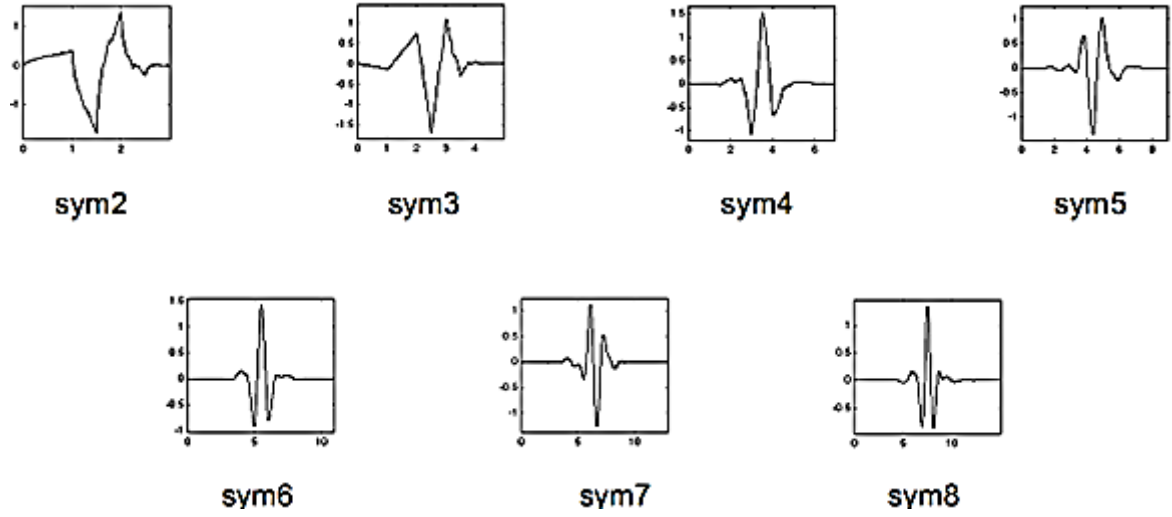


Şekil 3.16 a)Orjinal İmge b) Yatay Coif 1 dalgacık sonucu c) Dikey Coif 1 dalgacık sonucu d) DiagonalCoif 1 dalgacık sonucu

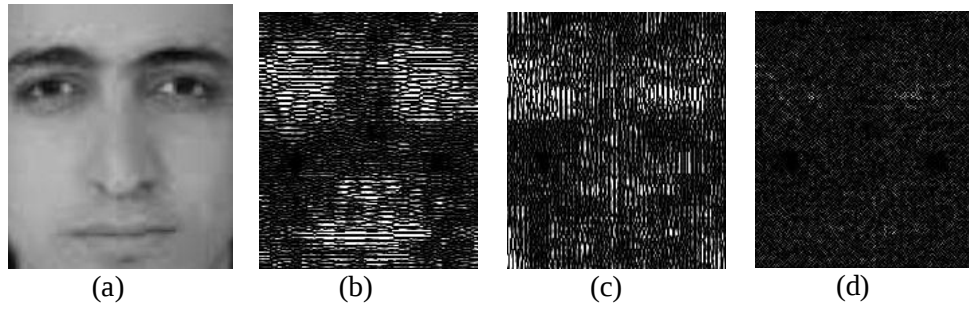
Bu tez çalışmasında, Coiflets ailesinin 5 adet dalgacığının tümü kullanıldı.

3.2.5. Symlets

Symlet ailesi, Daubechies tarafından geliştirilmiş, daubechies dalgacık ailesinin değiştirilip düzenlenmesi ile elde edilen simetriğe yakın dalgacıklardır.



Şekil 3.17 Symlet ailesinden kullanılan dalgacıklar

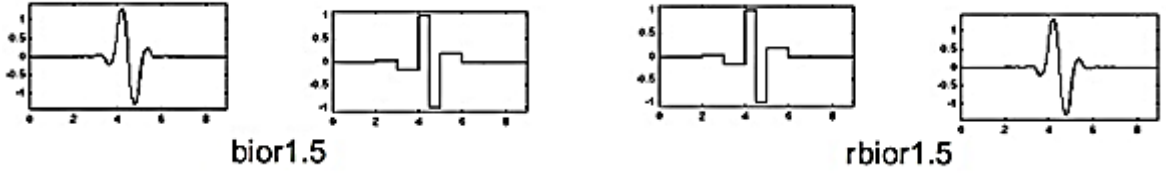


Şekil 3.18 a)Orjinal İmge b) Yatay Sym5 dalgacık sonucu c) Dikey Sym5 dalgacık sonucu d) Diagonal Sym5 dalgacık sonucu

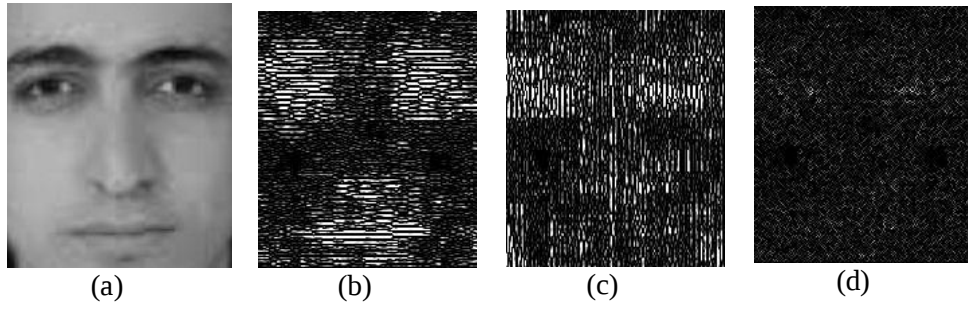
Symlet ailesinden, Sym2'den Sym24'e kadar toplam 23 adet dalgacık türü bu çalışmada incelenmiştir.

3.2.6. Ters Biortogonal

Biortogonal dalgacık çiftlerinin ters tanımlanması ile gerçekleşir. Şekil 3.19’de bior1.5 ve rbior1.5 dalgacıkları örnek olarak gösterilmiştir.



Şekil 3.19 Ters biortogonal dalgacıkların elde edilmesinin örnek gösterimi



Şekil 3.20 a) Orjinal İmge b) Yatay Rbior1.5 dalgacık sonucu c) Dikey Rbior 1.5 dalgacık sonucu d) DiagonalRbior 1.5 dalgacık sonucu

Tez çalışması kapsamında, Ters Biortogonal dalgacık ailesinden, biortogonal dalgacık ailesindeki tüm filtre çiftlerinin terslenmesi ile elde edilen 15 adet filtre kullanılmıştır.

Buraya kadar anlatılan dalgacık ailelerinden hangilerinin kullanıldığına dair bilgiler Tablo 3.1’de sunulmuştur. Çalışma kapsamında 103 adet dalgacıkların hepsi ile dalgacık dönüşümü uygulanmış, özellikler çıkarılmış ve başarımlar test edilmiştir.

Tablo 3.1. Kullanılan dalgacık aileleri, tipleri ve sayıları

Dalgacık Ailesi	Kullanılan Dalgacıklar	Toplam
<i>Daubechies</i>	db1, db2,..., db45	45 Adet
<i>Biortogonal</i>	bior 1.1, bior1.3, bior1.5, bior2.2., bior2.4, bior2.6, bior2.8, bior3.1, bior3.3, bior3.5, bior3.7, bior3.9, bior4.4, bior5.5, bior6.8	15 Adet
<i>Coiflets</i>	coif1, coif2,...,coif5	5 Adet
<i>Symlets</i>	sym2, sym3,...sym24	23 Adet
<i>ReversBiortogonal</i>	rbio 1.1, rbio1.3, rbio1.5, rbio2.2., rbio2.4, rbio2.6, rbio2.8, rbio3.1, rbio3.3, rbio3.5, rbio3.7, rbio3.9, rbio4.4, rbio5.5, rbio6.8	15 Adet
Toplam		103 Adet

4. SINIFLANDIRMA

Sınıflandırma, bilgisayar öğrenmesinde, bir veri kümesinin hedef kategoriler oluşturulacak şekilde işaretlenmesidir. Sınıflandırmadaki nihai amaç, veri seti içindeki her bir örneğin kategorisini doğru bir şekilde tahmin edebilmektir [18].

Bu çalışmada ise bir önceki bölümlerde anlatılan ADD ve GSEM ile oluşturulan değerler kullanılarak her bir örnek için bir öznitelik vektörü oluşturulmuş, bu vektörlere sınıflandırma uygulanarak hangi örneğin bayan, hangi örneğin erkek olduğu tahmin edilmeye çalışılmıştır.

Sınıflandırma probleminin en basit tipi hedef kategori çeşidinin iki olduğu durumlardır. Bu tür sınıflandırmalar ikili sınıflandırma olarak adlandırılır. Bu tez çalışmasında da hedef kategori çeşidi sayısı iki (erkek,bayan) olduğu için ikili sınıflandırma yapılmıştır.

Sınıflandırma işleminin öğrenme aşamasında, bilinen kategoriler ile tahmin edilen kategori sonuçları arasında bir ilişki bulunmaya çalışılır. Test aşamasında ise, öğrenme aşamasına hiç katılmayan örneklerin testi yapılır. Daha sonra gerçek değerler ile tahmin edilen değerler mukayese edilerek hata oranı hesaplanır.

Sınıflandırma işleminden önce eğitim ve test örneklerinin büyüklüğünün doğru belirlenmesi, çıkan sonuçların güvenilirliğini arttıracaktır. Eğitim ve test örneklerinin büyüklüğünü belirlemede farklı yöntemler mevcuttur. Örneğin, veri kümesinden rastgele olarak %70 i kadarı eğitim, geri kalan %30'u ise test için kullanılmak istenebilir. Fakat burada her deneme sonunda farklı sonuçlar elde edilebilecektir. Seçilen %30 test verisinin, örnekleri ne kadar temsil ettiğinin ise kesin bir cevabı yoktur. O yüzden bu çalışmada, sonuçların kararlı olarak elde edilebilmesi için 10 aşamalı Çapraz Doğrulama (*10 fold Cross Validation*) yöntemi ile sınıflandırma sonuçları test edilmiştir.

4.1. Çapraz Doğrulama

Eğitici öğrenme algoritmalarında, bir sınıflandırıcının başarımını doğru hesaplamak önemli bir işlemdir. Bu işlem sadece sınıflandırıcının ileriki uygulamalarda doğru

kestirimler yapabilmesi için değil, aynı zamanda model seçimi veya sınıflandırıcıların kombinasyonu için önemlidir [19].

Sınıflandırma işlemi sonucundaki hesaplanan başarımlar bazen aldatıcı olabilmektedir. Örneğin rastgele seçilen %10 luk test kümesi ile bir başarımlı hesaplanmaya çalışıldığında, seçilen kümenin ne hangi elemanlardan oluştuğu önemli olabilmektedir. Eğer rastgele alınan test kümesinde ekstrem örnekler var ise başarımlı aslında olduğundan düşük çıkacaktır. Tam tersi durumda ise, yani alınan %10'luk test verisi, şansa bağlı olarak eğitim verisine yakın elemanlardan oluşuyorsa, hesaplanan başarımlı, gerçek başarımlının üzerinde gözükür. Burada elde edilen sonuçların güvenilirliği ile ilgili bir problem vardır.

Bir sınıflandırıcının gerçek başarımlı hesaplanmaya çalışılırken, düşük varyasyon göstermesi beklenir [20]. Yani, sınıflandırıcı ile birkaç deneme yapılır. Bu denemelerin ortalamaları başarımlı belirlerken, oluşan varyasyon ise sınıflandırıcının güvenilirliğini kısmen gösterir.

Bir sınıflandırıcının, her zaman başarımlı yüksek bir güvenle kestirebileceğini, hesaplayabileceğini düşünmek yanlıştır [20].

Bu yöntemde, veri kümesi rastgele olarak 10 parçaya bölünür. Her bir parçaya bir indis numarası verilir. Birinci aşamada, birinci indise sahip olan veriler test verisi, geri kalanları eğitim verisi olarak sınıflandırıcıya gönderilir. Sınıflandırıcıdan elde edilen sonuçlar, birinci indise sahip verilerle test edilip başarımlı hesaplanır.

İkinci aşamaya geçildiğinde, iki numaralı indise sahip olan veriler test verisi, geri kalan 9 parça ise eğitim verisi olacak şekilde sınıflandırıcıya gönderilir. Böylelikle 10 kez sınıflandırma yapılmış olur. Her bir aşamada tüm verilerin %10'u test, %90'ı ise eğitim verisi olarak kullanılır. Bu 10 sınıflandırmanın başarımlı sonuçlarının ortalaması alınarak, genel başarımlı yüzdesi elde edilir. Bu sayede, veri setinin tamamı en az bir kere test, en az bir kere eğitim verisi olarak sınıflandırıcıya yönlendirilmiş olur.

4.2. Destek Vektör Makineleri

1936'da, R.A. Fisher tarafından, örüntü tanıma işlevini gerçekleştirmek için ilk algoritma geliştirildi. 1957 Frank Rosenblatt tarafından, ileri beslemeli sinir ağlarının basit bir modeli olan *perceptron*'u keşfedildi. Bundan 6 yıl sonra, Aizerman, Braverman ve Rozonoer tarafından çekirdeğin geometrik yorumuna giriş yapıldı. 1964'te daha iyi bir

algoritmanın gelişmesine öncülük eden Genelleştirilmiş Betimleme algoritması “Destek Vektör Makineleri (DVM)” Vladimir Vapnik ve Chervonenkis tarafından ortaya konuldu [21].

Vapnik ve arkadaşları, DVM’nin şimdiki formunu AT&T Bell Laboratuvarlarında geliştirmiştir. Bu endüstriyel başlangıcı nedeniyle, DVM ses oryantasyonundan diğer gerçek dünya uygulamalarına kadar kullanım alanı bulmuştur. Başlangıçtaki uygulamalar Optik Karakter Tanıma (OKT) üzerine olmuştur. Sonrasında DVM, kısa sürede OKT ve nesne tanıma uygulamalarında kullanılmak üzere en uygun sistem haline gelmiştir [22].

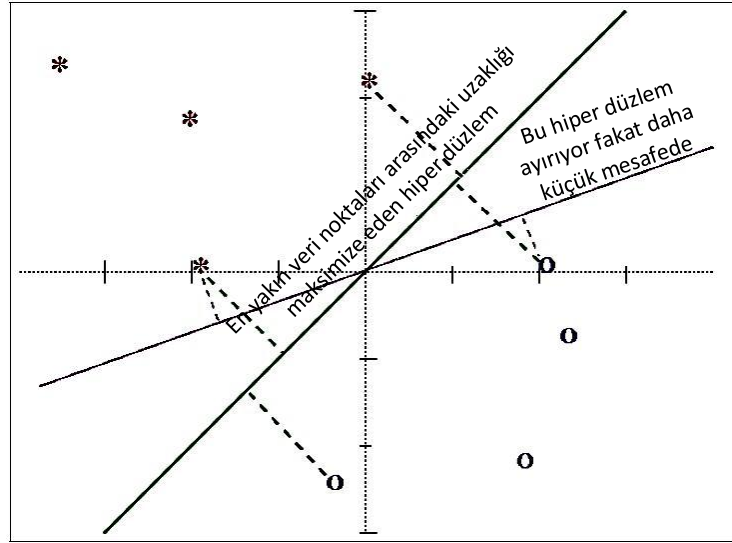
DVM’nin formüle edilebilmesi için sınıflandırmak için esas problem düşünülebilir. Örneğin, $y_i \in \{+1|-1\}$ ikili sınıflandırıcı düşünürse, c sayıda sınıfa ayrılması gereken N_c verikümesi c adet ikili sınıflandırıcılara ayrıştırılarak çözülebilir [23]. Yani örneğin, 5 hedef sınıfa ayırmak amacıyla düzenlenen bir sınıflandırıcı problemi, beş adet ikili sınıflandırıcıya ayrıştırma stratejileri ile çözülebilir.

DVM nin temel stratejisi sınıflar arasındaki farkı maksimum yapan düzlemleri bulmaktır. Öznitelik uzayı üzerindeki DVM sınıflandırıcının öğrenme eşitliği denklem 3.9’da verilmiştir [20].

$$f(x_i, w_f) = w_f^T \cdot x_i^f + b \quad (3.9)$$

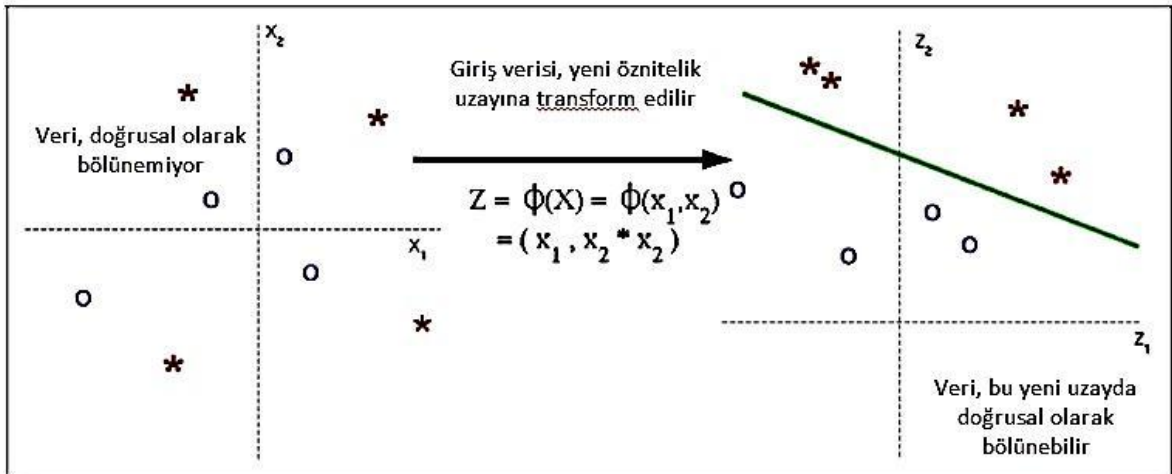
DVM, güçlü bir düzenlileştirme, tanzim etme kabiliyetine sahiptir. Bu özelliğini ise, modeli yeni verilere doğru genelleştirebilmesi ile kazanır. Kullanışlılık ve ölçeklenebilirlik DVM’nin en önemli tercih nedenleridir [24].

DVM, kategoriler arasında onları bölen hiper düzlemler bulmaya çalışır. Bu işlemi yaparken sinir ağları mantığı ile çalışarak öğrenir. En uygun hiper düzlem, birbirinden farklı kategorilerde yer alan en yakın komşu data noktaları ile hiper düzlemin arasındaki mesafenin maksimum olduğu düzlemdir. Şekil 4.1’de gösterildiği gibi iki hiper düzlem de veriyi uygun şekilde bölmeye karşın mesafenin maksimum olduğu yaklaşık 45 derece açıda olan hiper düzlem uygun olan düzlemdir.



Şekil 4.1 Mesafeyi maksimize eden hiper düzlemin seçilmesi [25].

Bazen veriler, bir hiper düzlem ile doğrusal olarak bölünebilecek bir şekilde bulunmuyor olabilirler. Bu tür durumlarda DVM, verilerin başka bir uzaya dönüşümünü sağlayarak oluşan yeni veri noktalarını doğrusal olarak bölmeye çalışır. Bu DVM'nin güçlü özelliklerindendir. Doğrusal olarak bölünemeyen kategorilerin farklı uzaya dönüştürülmesi işleminin basit bir modeli Şekil 4.2'de gösterilmektedir.



Şekil 4.2 Verinin özellik uzayında ayrılması [25].

5. UYGULAMALAR

Bu bölümde, uygulamada kullanılan veritabanı üzerinde hangi işlemler yapılarak özellik seti elde edilmeye çalışıldığı anlatıldı. Alınan imgeler, öncelikle orijinal hallerine GSEM uygulayarak özniteliklerin çıkarılması ve bu özniteliklere göre sınıflandırma yapmaya çalışılmıştır.

İlk başta GSEM, orijinal imgeler kullanılarak doğrudan öznitelik çıkardığı durumda cinsiyetlerin sınıflandırma başarımının yüksek olabileceği öngörülmüştü. Fakat sadece GSEM kullanılarak sınıflandırma başarımının yüksek olmadığı görülünce, imgelere önce ADD uygulayıp sonra GSEM ile öznitelikler elde edilmeye çalışılmıştır.

ADD uygulandığında farklı dalgacık ailelerinin farklı başarımları elde edildiği görülünce, mümkün mertebe tüm dalgacık aileleri ile başarımlar test edilmeye çalışılmıştır. Bazı dalgacık şekillerinin hesaplanması uzun sürdüğünden toplamda yalnız 103 dalgacık denenmiştir.

5.1. Materyal

Çalışmalarda kullanılan bazı veri tabanlarındaki bayanların başörtüsü, makyajlı vb. olmasından veya bazı veritabanlarında erkeklerin sakallı olanlarının sayısı fazla olmasından dolayı başarımlarının yüksek gözüktüğü görülmektedir. Tabi cinsiyete haiz özelliklerin farklılıklarının etnik köken, yaş gibi etkenlerle başarımları düşürüp yükseltebileceği açıktır. Veritabanında kullanılan resimlerden erkek ve bayanlarda ayırt edici özelliğin saç olduğu düşünülerek, tamamen yüz biyometrisine odaklanabilmek için, saçı kapsayan alanlar çıkarılan iç yüz alanları kullanılmıştır. Böylelikle elde edilen sonuçların güvenilirliği arttırılmıştır. Zira, karşı cephe fotoğrafların saçları da kapsayan tam bir fotoğrafın incelenmesi çalışmanın başarımını arttıracak, fakat elde edilen sonuçların yüze ait cinsiyete haiz özelliklerin bulunup bulunmadığını tam olarak belirlenemeyecektir. Çalışmamızda, bu yüzden özellikle iç yüz biyometrisinin cinsiyet tanımadaki başarımı hesaplanmaya çalışılmıştır.

Bu tez çalışmasında kullanılan yüz veritabanı, Brezilya'daki FEI [26] Üniversitesinin Yapay Zeka Laboratuvarlarında 2005 – 2006 yılları arasında toplanmıştır. 200 farklı bireyin her birinin 14 çeşit fotoğrafı çekilmiş, fakat bu çalışmada sadece bunlardan 1 pozu içeren altküme kullanılarak toplamda 200 imge üzerinde çalışılmıştır. Tüm resimler hizalanmış ve

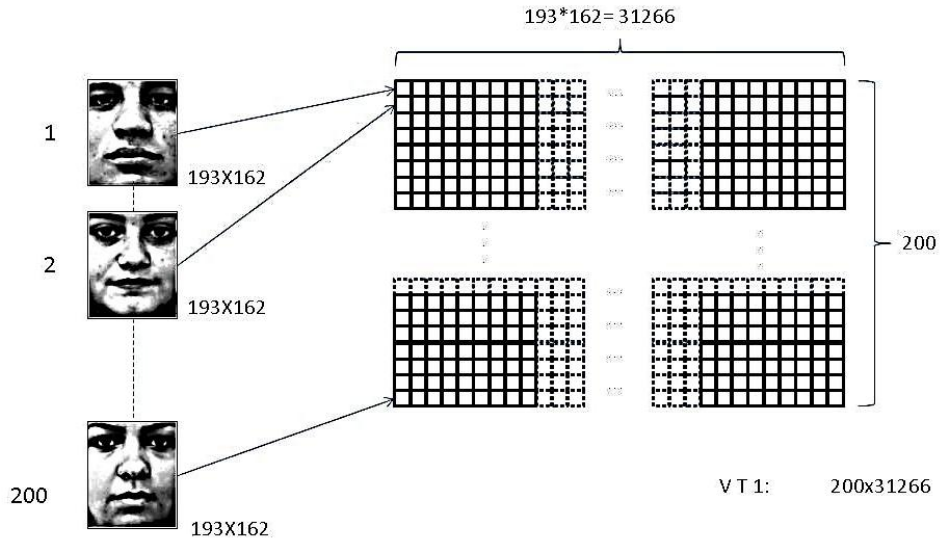
kırpılmış, 193x162 piksel ebatlarında olup, bazı fotoğraflar Şekil 5.1’de gösterilmiştir. Tüm yüzler yaşları 19 ile 40 arasında değişen FEI öğrencileri ve çalışanlarından elde edilmiştir. Erkek ve bayanların sayısı eşit olup, veritabanında 100 erkek ve 100 bayan kullanılmıştır [26].



Şekil 5.1. Veritabanındaki erkek ve bayanlara ait örnek yüzler. (İlk satır: Erkekler, ikinci satır: Bayanlar)

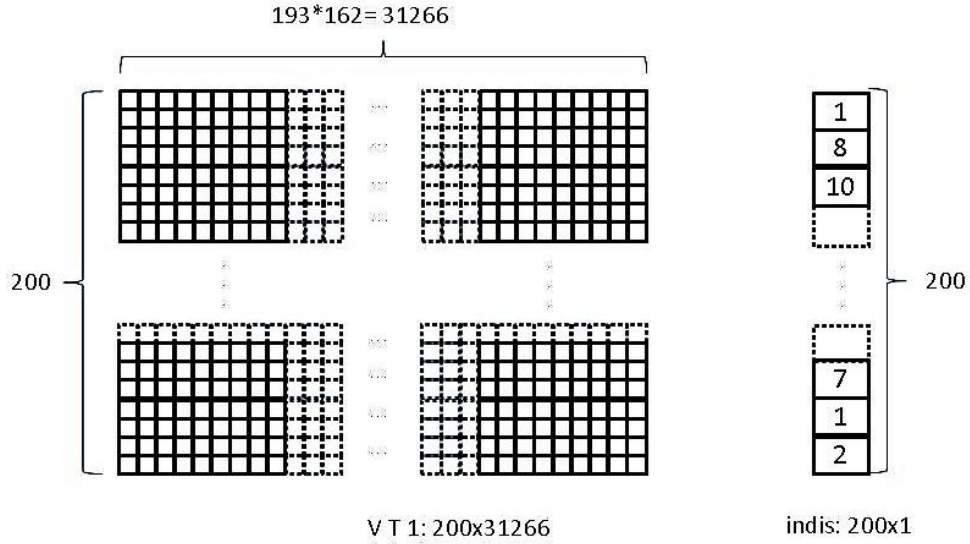
5.2. GSEM Kullanarak Öznitelik Çıkarılması

GSEM kullanarak öznitelik çıkarılmadan önce bazı değişkenler düzenlenip tanımlanmıştır. Öncelikle 193x162 boyutundaki 200 adet 2B imge, her biri 1B vektör şekline olacak şekilde 200X(193*162) boyutunda bir kaynak veritabanı oluşturulmuştur.



Şekil 5.2. 200 adet imgenin her birinin vektörle temsil edilerek kaynak VT'nin oluşturulmasının şematik gösterimi

Daha sonra her denemede ki sonucun, sınıflandırıcı değişikliklerinden etkilenmemesi için tüm denemelerdeki sınıflandırıcı olan 10 aşama çapraz doğrulamalı Destek Vektör Makine(10 fold CrossValidation SVM)'nin girdisi olacak şekilde 'indis' adında Şekil 5.3'deki gibi tek boyutlu matris oluşturuldu. İndis matrisinde, 4. Bölümde anlatılan 'Çapraz Doğrulama' (*Cross Validation*) işlemindeki, eğitim ve test elemanlarının hangi imge kümelerinden oluştuğunu belirten indisler bulunmaktadır.



Şekil 5.3. Tüm sınıflandırma denemelerinde kullanılan tek boyutlu indis değişkeni yapısının örnek gösterimi.

İndis değişkeni içinde bulunan sayılar, ona tekabül eden imgenin hangi aşamada test verisi olacağını belirler. Her bir aşama için 200 imgenin %10'u yani 20 adet imge rastgele belirlenmiştir. Örneğin 8 inci aşamada, indis değişkeninde 8 değeri bulunan 20 imge test verisi olarak kullanılmakta, diğer 180 imge eğitim verisi olarak kullanılmaktadır. Tüm bu aşamalarda elde edilen sınıflandırma sonuçların ortalaması ise tutarlı bir başarı oranını belirlenmekte, rastgeleliğin etkisi azaltılmaktadır.

'cinsiyetler' değişkeni, 200 imgenin tek tek incelenerek, 'erkek' ler için '1', 'bayan'lar için '0' değerini alacak şekilde 200x1'lik bir vektör şeklinde tanımlanmıştır.

100 Erkek ve 100 Bayan karşı cephe fotoğraflarına ilk etapta GSEM kullanılarak özellik çıkarımı yapıp DVM ile sınıflandırma yapılmaya çalışılmıştır. Bunun için öncelikle 200 imgenin GSEM ile elde edilen Bölüm 3.1'de belirtilen istatistiki değerleri kullanılarak bir öznitelik vektörü veritabanı oluşturulmuştur.

Her bir imge için piksel değerlerinin ortalaması, standart sapması ve entropi değerlerine ek olarak GSEM'in istatistiki değerleri olan zıtlık, korelasyon, homojenlik,

enerji deęerleri birleřtirilerek her imge iin 7 zellik belirlenmiřtir. Bu 7 deęerin birleřtirilmesi ile imgeler iin oluřturulan znitelik vektrleri, 200x7 boyutunda GSEM znitelik vektrleri oluřturulmuřtur.

5.2.1. GSEM znitelikleri ile Cinsiyet Sınıflandırması

znitelik kmesinin cinsiyete gre sınıflandırma iřlemi DVM ile gerekleřtirilmiřtir. Kullanılan sınıflandırıcı 10 ařamalı apraz doęrulama metoduna gre, 200 adet olan tm imgeler 10 adet gruba blnerek, her gruba 20 řer adet imge dřecek řekilde etiketlenir. Bu grupların her biri iin, o gruptaki 20 imge test verisi olarak, dięer 9 gruptaki imgeler ise eęitim verisi olarak sırayla kullanılır. Bylelikle her grup test verisi olacak řekilde 10 kez sınıflandırma yapılır. Bu 10 adet sınıflandırma sonucunun bařarım ortalaması, 10 ařamalı sınıflandırmanın sonucu olarak hesaplanır.

Orijinal imgelerden GSEM ile elde edilen öznitelik vektörü üzerinde Bölüm 4’te anlatılan DVM yardımı ile sınıflandırma yapıldığı zaman Tablo 5.1 elde edilmiştir.

Tablo 5.1. Orijinal İmgelerden GSEM ile elde edilen özellik setinin sınıflandırma başarımı.

		Gerçek	
		Erkek	Bayan
Tahmin	Erkek	59	40
Edilen	Bayan	41	60

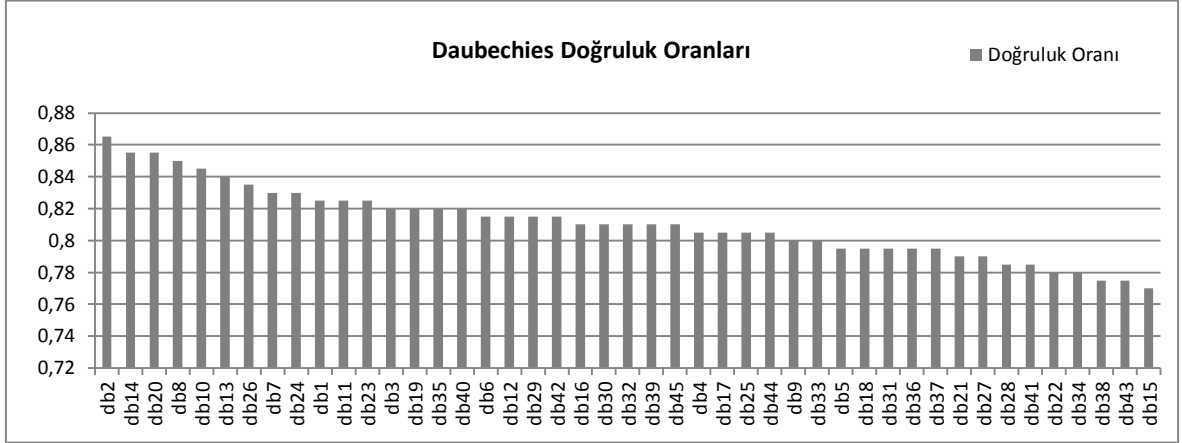
Tabloya göre 100 erkeğin doğru sınıflandırma oranı %59, 100 bayanın ise doğru sınıflandırma oranı %60 tır. 200 kişiden toplamda 119 kişinin doğru sınıflandırıldığı düşünülürse sınıflandırıcının başarı oranı $119/200$ oranı ile %59,5 olarak bulunur.

5.3. ADD ve GSEM ile Öznitelik Çıkarımı

İmgelerin orijinal hallerine sadece GSEM uygulanıp ardından sınıflandırma yapılarak elde edilen %59,5 luk başarı oranından sonra aynı imgelere Ayrık Dalgacık Dönüşümü ile elde edilen yeni öznitelik veritabanları oluşturuldu. Buradaki amaç, imgelere ADD uygulandıktan sonra GSEM uygulanmasıyla başarımın artıp artmayacağını belirlenmesi, başarımın artma durumunda ne kadar katkı sağlayacağını saptanmasıdır. Bu amacın gerçekleştirilebilmesi için Ayrık Dalgacık Dönüşümlerinde kullanılan Bölüm 3.2.1’de detaylandırılmış 5 adet dalgacık ailesinden toplam 103 adet dalgacık fonksiyonunu kullanılmıştır. Dolayısıyla 103 deneme yapılarak en yüksek başarımın ne kadar olduğu ve mevcut veritabanı için hangi dalgacık fonksiyonunun bu başarıyı gerçekleştirdiği bulunmaya çalışılmıştır. İmgelere her 2B ADD uygulandıktan sonra 7 adet imge Bölüm 3.2.1’de anlatıldığı gibi oluşmaktadır. Bu 7 adet imgenin her birine GSEM uygulanarak 7 istatistiki değer elde edildiği düşünüldüğünde her bir resim için 49 öznitelik çıkarılmış olur. Böylelikle ADD ile elde edilen özellik setinin boyutu 200×49 olur.

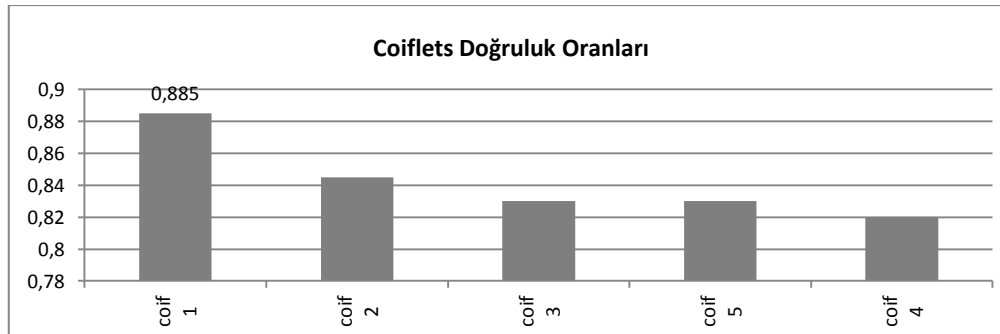
5.3.1. GSEM ve ADD Öznitelikleri ile Cinsiyet Sınıflandırılması

2B ADD'nin farklı filtre ailelerine göre denemeler yapılmıştır. Bu filtre aileleri kullanılarak oluşturulan özellik setleri ile ayrı ayrı sınıflandırmalar yapılmıştır. Daubechies ailesinin 45 adet fonksiyonu kullanılarak elde edilen başarımlar Şekil 5.4' de gösterilmiştir.



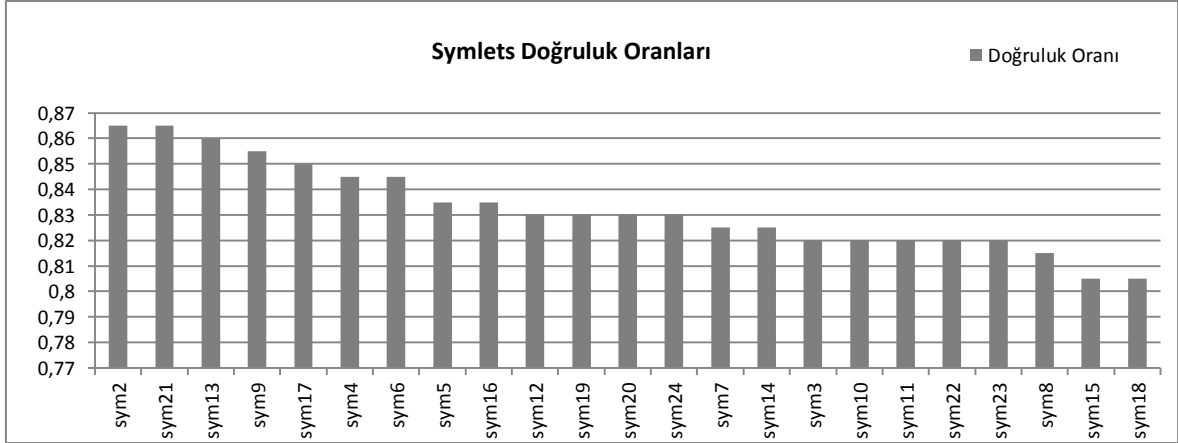
Şekil 5.4. 45 adet Daubechies dalgacık filtresi ile oluşturulan veri setinin sınıflandırma başarımları (db1 – db45)

Coiflets Dalgacık ailesinden 5 adet dalgacığın kullanılması ile elde edilen sınıflandırmalardaki başarımlar Şekil 5.5'de gösterilmiştir. Tüm denemelerin sonucunda Coiflets ailesinin 1. Dalgacığı ile çıkarılan özelliklerin en yüksek sınıflandırma oranına ulaştığı görülmüştür.



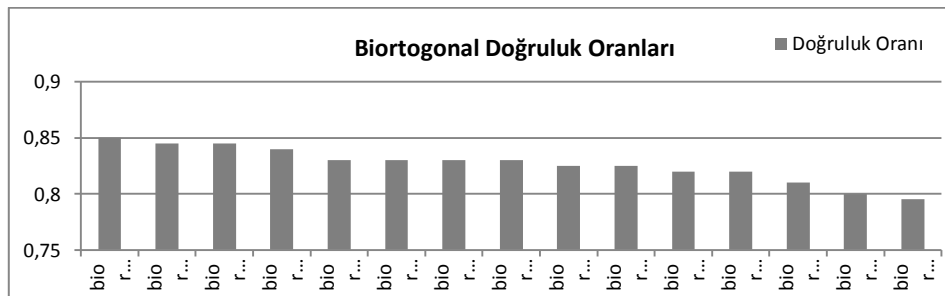
Şekil 5.5. Coiflets dalgacık filtresi ile oluşturulan veri setlerinin sınıflandırma başarımları (coif1-coif5)

Symlet dalgacık ailesinden ise 43 adet dalgacık bulunmasına rağmen 23 tanesiyle deneme yapılmıştır. Symlet ailesinin diğer dalgacıklarının hesaplanma süreleri uzun sürdüğünden bazı aşamalarda zorluklarla karşılaşıldı. Bu yüzden sadece 23 tanesi ile özellik setleri elde edilip sınıflandırma yapıldı. Coiflet ailesinden elde edilen özellik setlerinin sınıflandırılmasıyla ulaşılan başarımlar Şekil 5.6’de gösterilmiştir.



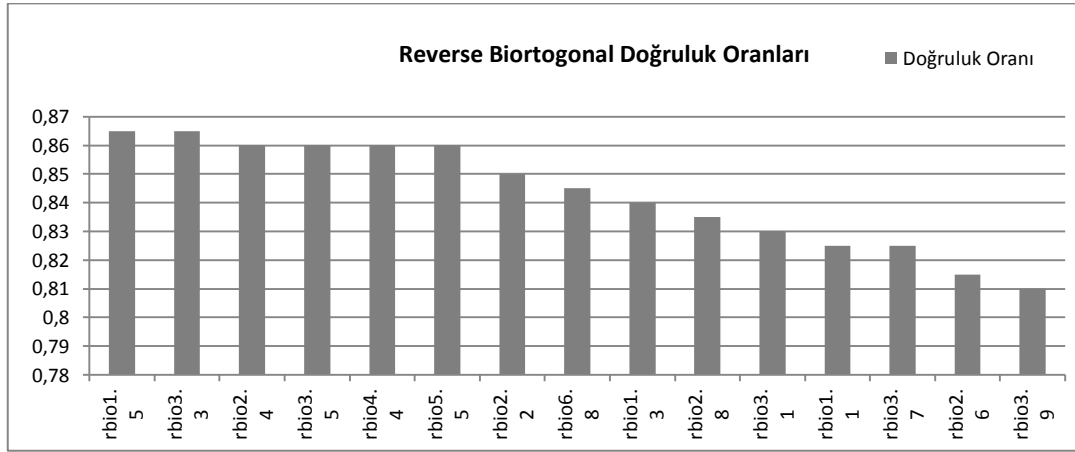
Şekil 5.6. 23 adet Symlets dalgacık filtresi ile oluşturulan veri setinin sınıflandırma başarımları (sym2 – sym24)

Dördüncü dalgacık ailesinden olan Biortogonal dalgacık ailesinden 15 dalgacık çeşidi kullanılarak, sonrasında GLCM ile öznelik vektörleri çıkarılmıştır. Ardından yapılan sınıflandırmada cinsiyetleri doğru tahmin edilenlerin oranları Şekil 5.7’de verilmiştir. Biortogonal dalgacık ailesinden en yüksek başarımlar *bior 5.5* adlı dalgacıkla elde edilmiştir.



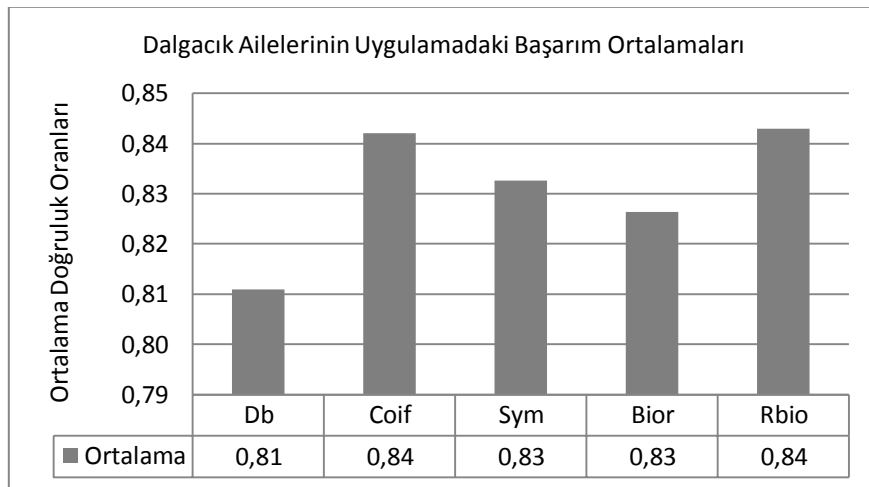
Şekil 5.7. 15 adet Biorthogonal dalgacık filtresi ile oluşturulan veri setinin sınıflandırma başarımları

Son olarak Reverse(Ters) Biortogonal dalgacık ailesi ile 15 farklı öznitelik veri seti elde edilmiştir. Bu özniteliklerin cinsiyetlere göre sınıflandırma durumunda Reverse Biortogonal dalgacık ailesi içindeki en yüksek başarımlar, rbio 1.5 dalgacığı %86-87 dolaylarında elde edilmiştir.



Şekil 5.8. Reverse Biorthogonal dalgacık filtresi ile oluşturulan veri setinin sınıflandırma başarımları

Sonuç olarak 5 farklı dalgacık ailesinden farklı sayılarda alınmış dalgacık örneklerine GSEM uygulanarak aynı öznitelikler çıkarılmış. Her biri ayrı ayrı sınıflandırıldıktan sonra başarımları hesaplanmış ve dalgacık ailelerinin başarımlar ortalamaları Şekil 5.9’da verilmiştir.



Şekil 5.9. Dalgacık filtre ailesi ile oluşturulan özelliklerin sınıflandırma başarımları ortalamaları.

Dalgacık ailelerinin kendi fonksiyonları kullanılarak elde edilen ortalama başarımlar Şekil 5.9’ de gösterildiği gibi bulunmuştur. Mevcut imge veritabanı için ‘reverse biorthogonal’ ve ‘coifflet’ dalgacıklarının yüksek oranda olduğu görülmektedir. 103 dalgacık fonksiyonundan en yüksek başarımlar Coif1 ile %88,5 olarak ölçülmüştür.

6. SONUÇ

Yapılan bu çalışmada, 2B dalgacık dönüşümünün cinsiyet tanıma üzerindeki etkisi araştırılmıştır. Orijinal imgelere uygulanan gri seviye eş oluşum matrisi ve 3 adet istatistiki değer ile tüm denemelerde 7 özellik üzerinde çalışılmıştır. Birinci yöntemde 200x7' lik bir özellik kümesi kullanılmışken, ikinci yöntemde Dalgacık dönüşümünün çıktısı olan 7 resimden dolayı 200x(7*7) yani 200x49'luk bir özellik kümesi kullanılmıştır. Birinci yöntemde özellik sayısının 7 olması da dikkate alınarak bakıldığında %60 lık bir başarı elde edilirken, ikinci yöntemdeki 49 özellik ile %88,5'a ulaşılmıştır. Bu sayede Coif1 dalgacık fonksiyonu kullanılarak %48 lik bir başarı artışı sağlanmıştır.

Gelecekteki çalışmalarda, ADD'nin farklı filtreleri ile oluşturulan bu 49 öz niteliğin, öz nitelik seçimi yöntemleri ile boyut azaltması yapılması planlanmaktadır. Bu sayede Ayrık Dalgacık dönüşümü uygulanan imgelerin cinsiyete ait özellikleri çıkarmadaki başarımı tam olarak ölçülecektir. Zira, burada başarımın artmasının iki nedeni vardır. Bunlardan biri özellik kümesinin genişlemesi, diğeri ise ADD'nin etkisidir. Özellik çıkarımı ve sınıflandırma işlemlerinde, en az özellikle en iyi performansı elde etmek amaçlandığından, özellik sayısı azaltılarak aynı performans elde edilmeye çalışılacaktır.

7. KAYNAKÇA

- [1] **Khan, S.A, Nazir, M., Akram, S. and Riaz, N.,** 2011. Gender classification using image processing techniques: A survey, *Multitopic Conference (INMIC), 2011 IEEE 14th International* , vol., no., pp.25,30, 22-24
- [2] **Khorsandi, R. and Abdel-Mottaleb, M.,** 2013. Gender classification using 2-D ear images and sparse representation, *Applications of Computer Vision (WACV), 2013 IEEE Workshop on* , vol., no., pp.461,466, 15-17.
- [3] **Farinella, G. and Dugelay, J.,** 2012, Demographic classification: Do gender and ethnicity affect each other?, *Informatics, Electronics & Vision (ICIEV), 2012 International Conference on* , vol., no., pp.383,390.
- [4] **Wen-S., C., Chun-R., H. and Chu-S., C.,** 2013. Gender classification from unaligned facial images using support subspaces, *Information Sciences*, Volume 221, Pages 98-109.
- [5] **Bayram, S.,** Histogram Eşitleme, <http://www.yildiz.edu.tr/~bayram/sgi/saygi.htm>, 2 Aralık 2014.
- [6] https://docs.oracle.com/cd/B28359_01/datamine.111/b28129/feature_extr.htm#DMCON010, Feature Selection. 2 Aralık 2014.
- [7] **Haralick, R.M., Shanmugam, K. and Dinstein, I.,** 1973. Textural Features for Image Classification, *Systems, Man and Cybernetics, IEEE Transactions on* , vol.SMC-3, no.6, pp.610,621.
- [8] **Zhu, L. and Zhang, Z.,** 2010. Auto-classification of insectimagesbased on colorhistogramand GLCM, *FuzzySystemsand Knowledge Discovery (FSKD), 2010 Seventh International Conference on* , vol.6, no., pp.2589,2593.
- [9] **Jia, L., Zhou, Z. and Li, B.,** 2012. Study of SAR Image Texture Feature Extraction Based on GLCM in Guizhou Karst Mountainous Region, *Remote Sensing, Environment and Transportation Engineering (RSETE), 2012 2nd International Conference on* , vol., no., pp.1,4.
- [10] **Anys, H., Bannari, A., He, D. C. and Morin, D.,** 1994. Texture Analysis For The Mapping Of Urban Areas Using Airborne MEIS-II Images, *Proceedings of*

- the First International Airborne Remote Sensing Conference and Exhibition, Pages 231-245.
- [11] **Baraldi, A. and Parmiggiani, F.**, 1995. An investigation of the textural characteristics associated with gray level cooccurrence matrix statistical parameters, *Geoscience and Remote Sensing, IEEE Transactions on* , vol.33, no.2, pp.293,304.
 - [12] **Michel, M., Yves, M., Georges, O. and Jean, M.**, 1996. Wavelet Toolbox User's Guide, Vol.1.
 - [13] **Sanjekar, P.S. and Dhabe, P.S.**, 2010. Fingerprint verification using haar wavelet. *Computer Engineering and Technology (ICCET), 2010 2nd International Conference on* , vol.3, no., pp.V3-361,V3-365.
 - [14] **Lalitha, N.V., Rao, S. and JayaSree, P.V.Y.**, 2013. DWT — Arnold Transform based audio watermarking. *Microelectronics and Electronics (PrimeAsia), 2013 IEEE Asia Pacific Conference on Postgraduate Research in* , vol., no., pp.196,199.
 - [15] **Jianwei Y., Chen, W. and Guosheng C.**, 2007. Two-Dimensional Wavelet Filters for Watermarking. *Computational Intelligence and Security, 2007 International Conference on* , vol., no., pp.935,939.
 - [16] **I. Daubechies, I.**, 1992, Ten Lectures on Wavelets, 1992, SIAM, Philadelphia.
 - [17] **Balakrishnan, P., Hasan, M.M. and Wahid, K.A.**, 2014. An Efficient Algorithm for Daubechies Lifting Wavelets Using Algebraic Integers. *Electrical and Computer Engineering, Canadian Journal of*, vol.37, no.3, pp.127,134.
 - [18] https://docs.oracle.com/cd/B28359_01/datamine.111/b28129/classify.htm#i1005746
Classification. 2 Aralık 2014.
 - [19] **Wolpert, D.H.**, 1992. Stacked Generalization. *Neural Networks* 5, 241-259.
 - [20] **Ron, K.**, 1995. A study of cross-validation and bootstrap for accuracy estimation and model selection. *In Proceedings of the 14th international joint conference on Artificial intelligence - Volume 2 (IJCAI'95)*, Vol. 2. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1137-1143.

- [21] **Negin, B.**, 2014. Post Operative Prognostic Prediction of Esophageal Cancer Cases Using Bayesian Networks And Support Vector Machines, *Master Thesis*, METU Department of Health Informatics, Ankara.
- [22] **Smola, A.J., Schölkopf, B.**, 2004. A tutorial on support vector regression, *Journal Statistics and Computing*, Vol. 14, Issue 3.
- [23] Chen, Q.; Song, Z.; Dong, J.; Huang, Z.; Hua, Y.; Yan, S., "Contextualizing Object Detection and Classification," *Pattern Analysis and Machine Intelligence*, IEEE Transactions on , vol.37, no.1, pp.13,27, Jan. 1 2015
- [24] https://docs.oracle.com/cd/B28359_01/datamine.111/b28129/algo_svm.htm#CHDDJFDJ, About Support Vector Machines. 2 Aralık 2014.
- [25] **Dustin, B.**, Introduction to Support Vector Machines , <http://dustwell.com/PastWork/IntroToSVM.pdf>. 2 Aralık 2014.
- [26] <http://fei.edu.br/~cet/facedatabase.html>, FEI Face Database, 14 Mart 2012.

ÖZGEÇMİŞ

Serdar ABUT, 1984 yılında Van’da doğdu. Ortaokul ve Liseyi Van Gölü Anadolu Lisesinde okudu. 2001-2008 yılları arasında Selçuk Üniversitesi Bilgisayar Mühendisliği Bölümünde Lisans eğitimini tamamladı. 2009 yılında Van Bölge İdare Mahkemesinde 5 ay Zabıt Katibi olarak çalıştı. Aralık 2009’da Siirt Üniversitesi Eğitim Fakültesi Bilgisayar ve Öğretim Teknolojileri Eğitimi Bölümü’nde göreve başlayıp halen aynı görevde bulunmaktadır. Şubat 2010’da Elazığ Fırat Üniversitesi Bilgisayar Mühendisliği Bölümü Donanım Anabilim dalında Yüksek Lisans eğitimine başladı.