İSTANBUL TEKNİK ÜNİVERSİTESİ ★ FEN BİLİMLERİ ENSTİTÜSÜ

YÜZDEKİ NİRENGİ NOKTALARININ KOŞULLU REGRESYON ORMANLARI İLE SAPTANMASI

YÜKSEK LİSANS TEZİ

Gencer VURAL

Bilgisayar Mühendisliği Anabilim Dalı

Bilgisayar Mühendisliği Programı

HAZİRAN 2015

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ FEN BİLİMLERİ ENSTİTÜSÜ

YÜZDEKİ NİRENGİ NOKTALARININ KOŞULLU REGRESYON ORMANLARI İLE SAPTANMASI

YÜKSEK LİSANS TEZİ

Gencer VURAL
(504121520)

Bilgisayar Mühendisliği Anabilim Dalı

Bilgisayar Mühendisliği Programı

Tez Danışmanı: Prof. Dr. Muhittin GÖKMEN

HAZİRAN 2015

İTÜ, Fen Bilimleri Enstitüsü'nün 504121520 numaralı Yüksek Lisans Öğrencisi **Gencer VURAL**, ilgili yönetmeliklerin belirlediği gerekli tüm şartları yerine getirdikten sonra hazırladığı **"YÜZDEKİ NİRENGİ NOKTALARININ KOŞULLU REGRESYON ORMANLARI İLE SAPTANMASI"** başlıklı tezini aşağıdaki imzaları olan jüri önünde başarı ile sunmuştur.

Tez Danışmanı :	**Prof. Dr. Muhittin GÖKMEN** İstanbul Teknik Üniversitesi	.....
Jüri Üyeleri :	**Doç. Dr. Hazım Kemal Ekenel** İstanbul Teknik Üniversitesi	.....
	Yrd. Doç. Dr. Berk Gökberk MEF Üniversitesi	.....
		.....

Teslim Tarihi : **4 Mayıs 2015**
Savunma Tarihi : **2 Haziran 2015**

Sevgili aileme,

ÖNSÖZ

Çalışmalarım süresince bilgi ve tecrübelerini benden esirgemeyerek bana yardımcı olan ve yol gösteren çok değerli danışman hocam Sayın Prof. Dr. Muhittin Gökmen'e teşekkür ederim.

Bugüne kadar atmış olduğum her adımda bir an olsun desteğini esirgemeyen aileme teşekkür ederim.

Yüksek lisans öğrenimime vermiş olduğu burs ile finansal katkı sağlayan TÜBİTAK'a teşekkürlerimi sunarım.

Haziran 2015 Gencer VURAL

İÇİNDEKİLER

Sayfa

ÖNSÖZ ..... **vii**
İÇİNDEKİLER ..... **ix**
KISALTMALAR ..... **xi**
ÇİZELGE LİSTESİ ..... **xiii**
ŞEKİL LİSTESİ ..... **xv**
ÖZET ..... **xvii**
SUMMARY ..... **xix**
1. GİRİŞ **1**
1.1 Literatür Çalışması 2
1.2 Tezin Organizasyonu 3
2. KULLANILAN YÖNTEMLER **5**
2.1 Yerel Zernike Momentleri 5
2.1.1 Yerel Zernike momentlerinin hesaplanması 5
2.1.2 YZM öznitelik vektörlerinin oluşturulması 9
2.2 Rastgele Ormanlar 9
2.2.1 Koşullu rastgele regresyon ormanları 12
2.2.1.1 Yüzün bakış açısının saptanması 13
2.2.1.2 Rastgele regresyon ormanlarının eğitimi 13
2.2.1.3 Yüzdeki nirengi noktalarının saptanması 14
3. DENEYSEL ÇALIŞMALAR **17**
3.1 Kullanılan Veri Seti 17
3.2 Yapılan Çalışmalar 17
4. SONUÇ VE ÖNERİLER **29**
KAYNAKLAR **31**
ÖZGEÇMİŞ **33**

KISALTMALAR

YZM	**:**	Yerel Zernike Momentleri
LFW	**:**	Labeled Faces in the Wild
LZM	**:**	Local Zernike Moments
FGH	**:**	Faz-Genlik Histogramları
RO	**:**	Rastgele Ormanlar
RRO	**:**	Rastgele Regresyon Ormanları
KRRO	**:**	Koşullu Rastgele Regresyon Ormanları
IG	**:**	Information Gain

ÇİZELGE LİSTESİ

Sayfa

Çizelge 3.1: [1] çalışmasında bildirilen başarım sonuçları. 18
Çizelge 3.2: [1] çalışmasına en yakın rastgele ormanlarla elde edilen başarım sonuçları. 18
Çizelge 3.3: Çift katmanlı YZM ile normalizasyon uygulanmayıp integral bilgisi kullanılarak, 40 bileşenle elde edilen başarım sonuçları. 19
Çizelge 3.4: YZM'nin 1. katmanı ile normalizasyon uygulanıp, integral bilgisi kullanılarak 40 bileşenle elde edilen başarım sonuçları. 19
Çizelge 3.5: Çift katmanlı YZM ile normalizasyon uygulanmadan ve integral bilgisi kullanılmadan, 40 bileşenle elde edilen başarım sonuçları..... 20
Çizelge 3.6: Çift katmanlı YZM ile normalizasyon uygulanıp, integral bilgisi kullanılmadan, 40 bileşenle elde edilen başarım sonuçları..... 20
Çizelge 3.7: YZM'nin 1. katmanı ile normalizasyon uygulanıp, integral bilgisi kullanılmadan 40 bileşenle elde edilen başarım sonuçları..... 21
Çizelge 3.8: YZM'nin 2. katmanı ile normalizasyon uygulanıp, integral bilgisi kullanılmadan 40 bileşenle elde edilen başarım sonuçları..... 21
Çizelge 3.9: $n_1 = 3$, $n_2 = 3$ parametreleriyle çift katmanlı YZM kullanılarak elde edilen başarım sonuçları. 24
Çizelge 3.10: Çift katmanlı YZM ve Gabor bileşenlerinin birlikte kullanımıyla elde edilen başarım sonuçları. 24
Çizelge 3.11: Çift katmanlı YZM kullanılarak FERET veri setinde elde edilen başarım sonuçları. 25

ŞEKİL LİSTESİ

Sayfa

Şekil 1.1 : Yüzde saptanması amaçlanan nirengi noktaları [1]. 2
Şekil 2.1 : YZM çekirdekleri ile yapılan konvolusyon işlemi. 6
Şekil 2.2 : $k = 9$ için elde edilen filtreleme çekirdeklerinden ilk 8 tanesinin gerçel bileşenleri [2]. 7
Şekil 2.3 : $k = 9$ için elde edilen filtreleme çekirdeklerinden ilk 8 tanesinin sanal bileşenleri [2]. 8
Şekil 2.4 : Yüz tanımada kullanılan YZM yönteminin genel gösterimi [3]. 10
Şekil 2.5 : Çimlerin ıslaklığının tespiti için örnek karar ağacı. 11
Şekil 2.6 : Örnek ikili karar ağacı. 11
Şekil 2.7 : Örnek karar ormanı. 12
Şekil 3.1 : %5 Tolerans için n_1 ve n_2 parametrelerinin ve bileşen sayısının başarıma etkisi. 22
Şekil 3.2 : %10 Tolerans için n_1 ve n_2 parametrelerinin ve bileşen sayısının başarıma etkisi. 22
Şekil 3.3 : %15 Tolerans için n_1 ve n_2 parametrelerinin ve bileşen sayısının başarıma etkisi. 23
Şekil 3.4 : %20 Tolerans için n_1 ve n_2 parametrelerinin ve bileşen sayısının başarıma etkisi. 23
Şekil 3.5 : Çalışmadaki yöntemle nirengi noktalarının isabetli saptandığı örnek çıktılar. 26
Şekil 3.6 : Çalışmadaki yöntemle nirengi noktalarının hatalı saptandığı örnek çıktılar-1. 27
Şekil 3.7 : Çalışmadaki yöntemle nirengi noktalarının hatalı saptandığı örnek çıktılar-2. 28
Şekil 3.8 : Çalışmadaki yöntemle nirengi noktalarının hatalı saptandığı örnek çıktılar-3. 28

YÜZDEKİ NİRENGİ NOKTALARININ
KOŞULLU REGRESYON ORMANLARI İLE SAPTANMASI

ÖZET

Teknolojinin gelişimine paralel olarak bilgisayarlar da gelişmektedir. Bilgisayarların gelişmesiyle beraber fiyatları da daha ucuz ve daha makul seviyelere gelmektedir. Bu sayede yaygınlaşan bilgisayarları kullanarak insan hayatını kolaylaştıracak sistemler oluşturma gayesi, bilgisayar bilimlerinde bir çok yeni araştırma alanının oluşumuna ve var olan bir çok alanın da hızla gelişmesine yol açmıştır. Bu alanlardan bir tanesi "Bilgisayarla Görü"dür.

Bilgisayarla görü, eğlence ve endüstri gibi alanlardan, günümüzde insanlar için büyük önem arz eden sağlık ve güvenlik alanlarına kadar çok geniş bir kullanım alanı yelpazesine sahiptir. Bilgisayarla görü, bu alanlarda kullanılan, nesne tanıma, insan bilgisayar etkileşimi, yüz tanıma, yüz analizi gibi pek çok araştırma alanını bünyesinde barındırmaktadır. İnsan bilgisayar etkileşimiyle ve yüz analiziyle olan alakasının yanı sıra yüz tanımada ön adım olarak yüzdeki nirengi noktalarının saptanması da bilgisayarla görünün aktif alanlarından bir tanesidir.

Orta ve yüksek kaliteli iki boyutlu durağan imgelerde yüzdeki nirengi noktalarının saptanması üzerine çok sayıda çalışma yapılmış olsa da, yüzdeki nirengi noktalarının düşük kaliteli imgelerde saptanması ve çoğu uygulama alanı için bir gereksinim olan gerçek zamanlı olarak saptanması bu alanın güncel problemlerindendir.

Bu tez çalışmasında, bu güncel problemlerle ilgili olarak, yüzdeki 10 nirengi noktasının saptanması amaçlanmıştır. Bu noktalar; sol gözün solu, sol gözün sağı, sağ gözün solu, sağ gözün sağı, burnun sağı, burnun solu, ağzın sağı, ağzın solu, ağzın üstü ve ağzın altıdır. Bu çalışmada, bahsedilen nirengi noktalarının etkili ve başarılı bir şekilde saptanması için son yıllarda bilgisayarla görü problemlerinin çözümünde etkili ve çok yönlü bir araç olduğu görülmüş olan Regresyon Ormanları ile moment tabanlı bir öznitelik çıkarma yöntemi olan Yerel Zernike Momentlerini (YZM) içeren bir yapı kullanılmaktadır.

Bu çalışmanın amaçlarından birisi de yüz tanımada etkili ve başarılı bir şekilde kullanılan YZM'nin yüzdeki nirengi noktalarının saptanmasında da başarılı bir şekilde kullanılabileceğinin gösterilmesidir. YZM'nin uygulandığı yamalar hem regresyon ormanlarının eğitiminde hem de yüzdeki nirengi noktalarının saptanması aşamasında kullanılmaktadır. Yüzdeki nirengi noktaları saptanırken, yamalar YZM uygulandıktan sonra karar ormanlarında değerlendirilmektedir. Regresyon ormanları ise görüntü yamaları ile yüzdeki nirengi noktaları arasındaki uzamsal ilişkiyi öğrenmektedir. Ayrıca, bu çalışmada farklı olarak regresyon ormanları, genelde olduğu gibi bütün eğitim kümesinden rastgele bir alt küme seçilerek eğitilmek yerine yüzün bakış açısına göre koşullu olarak eğitilmektedir. Bu sayede, yüzdeki bütün görünüm ve şekil değişimleri ile ilgilenmek yerine sadece ilgili bakış açısına ait olanları bilmeleri yeterli olduğu için ağaçların öğrenmesi daha kolay olmaktadır.

Bu çalışmada kullanılan yöntem, temel olarak yüzün bakış açısının saptanmasında kullanılacak karar ormanının eğitimi, yüzdeki nirengi noktalarının saptanması için kullanılacak karar ormanlarının yüzlerin bakış açısına göre koşullu olarak eğitilmesi ve test aşamasında nirengi noktalarının yüzün bakış açısına uygun koşullu karar ağaçları kullanılarak saptanmasından oluşmaktadır.

Bu tez çalışmasında kullanılan yöntemi test etmek amacıyla, 5749 kişinin 13233 yüz imgesinin bulunduğu ve çok çeşitli görüntüleme koşullarına sahip imgelerden oluşan Labeled Faces in the Wild (LFW) veri seti kullanılmıştır. Bu veri setinin seçilmiş olmasının sebebi, imgelerin farklı pozları, değişik ışıklandırma koşulları, çözünürlük ve imge kalitelerini, faklı yüz ifadelerini ve cinsiyetleri ve yüz bölümlerinin kapanması gibi saptama işlemini zorlaştırabilecek etkenleri barındırdığı zorlayıcı bir veritabanı olmasıdır.

Bu çalışmada kullanılan yöntemin başarımı LFW veri setinde çeşitli hata toleransları belirlenerek değerlendirilmiştir. Tolerans, hata payının iki göz arası uzaklığa oranının alabileceği maksimum değerdir. YZM ile koşullu regresyon ormanlarının birlikte kullanıldığı bu yöntem kullanılarak %15 tolerans için yüzdeki nirengi noktaları ortalama %89.33 başarımla saptanmaktadır. Artan tolerans değerlerinde daha yüksek başarımlar elde edilmektedir. Aynı yöntemin %20 tolerans için yüzdeki nirengi noktalarını saptama başarımı ortalama %94.48 olmaktadır.

FACIAL FEATURE DETECTION
USING CONDITIONAL REGRESSION FORESTS

SUMMARY

In parallel with development of technology, computers are developing either. Computer prices are getting cheaper and more reasonable with the development of computers. The aim of creating systems which make people's lives easier by using computers becoming widespread thanks to reduction in their prices leads up to generation of many new research areas and rapid development of many existing areas. One of these areas is "Computer Vision".

Computer vision has wide range of areas of usage from entertainment and industry to medical and security which are very important for people at the present time. Computer vision contains many fields of study such as object recognition, human computer interaction, face recognition and face analysis which are used in these areas of usage. Besides relationship with human computer interaction and face analysis, facial feature detection, as a pre-step of face recognition, is one of fresh fields of computer vision.

Even though there are many studies on facial feature detection from two dimensional, high and medium quality, still images, facial feature detection from low quality still images and real-time facial feature detection are two of current problems in this field of study.

Studies on facial feature detection can be classified into two categories depending on usage of holistic or local features. Holistic methods constructs a model by using entire face region, and images are tested by using this model. However, holistic methods are not able to deal with lighting changes and low resolution images accurately. In recent years, local methods, which are preferred and used mostly, uses image patches around facial landmarks while constructing a model. But these methods are not able to cope with global variations in the face due to being constructed by using limited face regions. Therefore global information about face like head pose is also used in local methods commonly in order to improve accuracy.

In this thesis study, related to mentioned current problems, detection of 10 facial features is aimed. These facial features are left of left eye, right of left eye, left of right eye, right of right eye, right of nose, left of nose, right of mouth, left of mouth, upper of mouth and lower of mouth. In this study, a structure containing Regression Forests, which is proven as an effective and versatile tool for solving computer vision problems in recent years, and Local Zernike Moments (LZM), which is moment-based feature extraction method, is used in order to detect mentioned facial features effectively and successfully.

One of aims of this thesis study is showing that LZM, is used for face recognition effectively and successfully, can also be used to detect facial features successfully. Patches on which LZM is applied are used for both training of regression forests and detection of facial features. After LZM is applied on patches, they are evaluated by

decision forests during detection of facial features. Regression forests learn spatial relationship between facial features and image patches. Additionally, in this study, regression forests are trained conditionally to head pose instead of being trained by using a random subset of entire training set dissimilarly. Regression forests have to deal with variations in appearance and shape of only related head pose instead of entire set, and also trees are able to learn more easily thanks to this difference.

Basically, the method which is used in this study consists of training decision forest for head pose detection, training decision forests conditionally to head pose for facial feature detection and detecting facial features by using suitable conditional decision trees depending on head pose.

In this thesis study, regression trees which are used to detect facial features are chosen from entire trained regression trees depending on head pose. Therefore, head pose must be detected primarily. For this purpose, regression forest which detects the head pose among defined 5 head poses, 'left profile', 'left', 'front', 'right', 'right profile', is trained by using LFW dataset. Regression forest which is trained for head pose detection consist of 15 trees whose maximum depth are 10.

During training of regression forests which are used to detect facial features, 13233 images, in which coordinates of 10 facial features are located, are used. 1500 images are randomly selected for training each regression tree. Faces in these images are detected. After face detection, images are rescaled in order to obtain 100x100 pixels face bounding boxes. Then, face bounding boxes are enlarged by %30 to make sure face bounding boxes contain all facial features. After defining face bounding boxes, 200 20x20 pixels patches, 150 from inside of boxes and 50 from rest area of the image, are randomly collected from each image. Finally, a regression tree whose maximum depth is 20 is constructed by applying training procedure mentioned in this study on collected patches.

Similarly to training process of regression forests used for facial feature detection, face in the image is detected firstly during facial feature detection process. After face detection, image is rescaled in order to obtain 100x100 pixels face bounding box and then, face bounding box is enlarged by %30 to make sure face bounding box contains all facial features. 20x20 pixels patches, mostly from inside of box, are randomly collected from image. These patches are provided as input to regression trees chosen for facial feature detection depending on head pose of face. Patches are evaluated at each node and reach a leaf node at each tree. As a result of this evaluation, a set of leaf nodes is obtained for each patch. Facial features are tried to be predicted by using Gaussian kernel and density estimator on this set. Finally, facial features are detected by applying Mean-Shift on each one of predicted facial feature points.

Labeled Faces in the Wild (LFW) dataset, which contains 13233 face image of 5749 people and large variations in imaging conditions, is used for testing the method used in this thesis study. This challenging dataset has been chosen because it includes variations in pose, lighting conditions, resolution, quality, facial expressions and gender, and challenging factors such as occlusion of facial area.

Success of the method used in this study is evaluated in LFW dataset depending on various tolerance values. Tolerance is the maximum value of ratio of error margin to inter-ocular distance. When tolerance is defined as %15, %89.33 average success rate is obtained by using this method consisting a combined structure of LZM and

conditional regression forests. If greater tolerance values are selected, greater success rates are observed. %94.48 average success rate is measured by using same method when tolerance is defined as %20.

Speed of the method used in this thesis study is also tested. If location of face is provided to the system as input, average duration of facial feature detection is measured as 56 milliseconds. If the system in this study also detects face firstly, average duration of facial feature detection is measured as 72 milliseconds.

If the results obtained in this study are examined, it is seen that these results which are obtained from a challenging dataset including variations in pose, lightning conditions, resolution, quality, facial expressions and gender are promising results.

1. GİRİŞ

Teknolojinin gelişimiyle birlikte bilgisayarlar da gelişmektedir. Bilgisayarların gelişip ucuzlaması ve dolayısıyla yaygınlaşması sonucu bilgisayarları kullanarak insan hayatını kolaylaştıracak sistemler oluşturma gayesi, bilgisayar bilimlerinde bir çok yeni araştırma alanının oluşumuna ve var olan bir çok alanın da hızla gelişmesine yol açmıştır. Bu alanlardan bir tanesi "Bilgisayarla Görü"dür.

Eğlence ve endüstri gibi alanlardan, günümüzde insanlar için büyük önem arz eden sağlık ve güvenlik alanlarına kadar birçok alanda kullanılan bilgisayarla görü, nesne tanıma, insan bilgisayar etkileşimi, yüz tanıma, yüz analizi gibi pek çok araştırma alanını bünyesinde barındırmaktadır. İnsan bilgisayar etkileşimiyle ve yüz analiziyle olan alakasının yanı sıra yüz tanımada ön adım olarak yüzdeki nirengi noktalarının saptanması da bilgisayarla görünün aktif alanlarından bir tanesidir.

Orta ve yüksek kaliteli iki boyutlu durağan imgelerde yüzdeki nirengi noktalarının saptanması üzerine çok sayıda çalışma yapılmış olsa da, yüzdeki nirengi noktalarının düşük kaliteli imgelerde saptanması ve çoğu uygulama alanı için bir gereksinim olan gerçek zamanlı olarak saptanması bu alanın güncel problemlerindendir.

Bu tez çalışmasında, bu güncel problemlerle ilgili olarak, yüzdeki 10 nirengi noktasının saptanması amaçlanmıştır. Şekil 1.1'de görüldüğü gibi bu noktalar; sol gözün solu, sol gözün sağı, sağ gözün solu, sağ gözün sağı, burnun sağı, burnun solu, ağzın sağı, ağzın solu, ağzın üstü ve ağzın altıdır. Bu çalışmada, bahsedilen nirengi noktalarının etkili ve başarılı bir şekilde saptanması için son yıllarda bilgisayarla görü problemlerinin çözümünde etkili ve çok yönlü bir araç olduğu görülmüş olan Regresyon Ormanları ile moment tabanlı bir öznitelik çıkarma yöntemi olan Yerel Zernike Momentlerini (YZM) [3] içeren bir yapı kullanılmaktadır.

Şekil 1.1: Yüzde saptanması amaçlanan nirengi noktaları [1].

1.1 Literatür Çalışması

Yüzdeki nirengi noktalarının saptanması alanında önceden yapılmış çalışmalar bütünsel ve yerel öznitelikleri kullanımına göre iki ana başlıkta toplanabilir.

Bütünsel yöntemlerde, bütün yüz alanı kullanılarak bir model oluşturulur ve oluşturulan bu modelle görüntüler test edilir. Fakat bu yöntemler ışıklandırma değişimleriyle ve düşük çözünürlüklü görüntülerle başa çıkma konusunda zayıf kalmaktadırlar.

Son yıllarda belirgin olarak daha çok başvurulan yerel yöntemler ise model oluştururken yüzün belirli noktalarının etrafında alınan görüntü yamalarını kullanmaktadır. Bu yöntemlerle elde edilen modeller yüzde sınırlı bir bölge kullanılarak oluşturuldukları için yüzün görünümünde meydana gelen geniş kapsamlı değişimlerin üstesinden gelme konusunda sorun yaşamaktadırlar. Bu yüzden, bu yöntemlerde yüzün görünümü ile ilgili daha geniş çaplı (global) bilgiler de genellikle kullanılmaktadır.

Daha önceki [4–7] çalışmalarında da görüleceği gibi, yüzdeki nirengi noktalarını oldukça isabetli bir şekilde saptayan yöntemler bulunmaktadır. Fakat bu nirengi noktalarının gerçek zamanlı bir performansla saptanması konusunda bir yöntem eksikliği göze çarpmaktadır.

[1] çalışmasında, son yıllarda bilgisayarla görü problemlerinin çözümünde etkili ve çok yönlü bir araç olduğu görülmüş olan regresyon ormanlarına [8] dayalı bir yöntem kullanılarak yüzdeki nirengi noktaları gerçek zamanlı olarak saptanmaktadır. Regresyon ormanlarının görüntü yamaları ile yüzdeki nirengi noktaları arasındaki uzamsal ilişkiyi öğrendiği [1] çalışmasında, gerçek zamanlı performansın yanı sıra, farklı olarak regresyon ormanları, genelde olduğu gibi bütün eğitim kümesinden rastgele bir alt küme seçilerek eğitilmek yerine yüzün bakış açısına göre koşullu olarak eğitilmektedir. Bu sayede, yüzdeki bütün görünüm ve şekil değişimleri ile ilgilenmek yerine sadece ilgili bakış açısına ait olanları bilmeleri yeterli olduğu için ağaçların öğrenmesi daha kolay olmaktadır.

Bu çalışmada kullanılan yöntem, [1] çalışmasını ve [3] çalışmasındaki YZM yöntemini bir araya getirmektedir. Bu yöntem, temel olarak [1] çalışmasındaki yönteme benzemekle beraber, bu yöntemin YZM [3] yöntemiyle çeşitli şekillerde entegrasyonunun sağlanması ve yüz tanımada etkili ve başarılı bir şekilde kullanılan YZM'nin yüzdeki nirengi noktalarının saptanmasında da başarılı bir şekilde kullanılabileceğinin gösterilmesi amaçlanmıştır.

1.2 Tezin Organizasyonu

Bölüm 2'de, bu tez çalışmasında kullanılan Yerel Zernike Momentleri, Rastgele Ormanlar ve Koşullu Regresyon Ormanları yöntemlerinden bahsedilmiştir. 3. Bölüm olan Deneysel Çalışmalar'da bu tez çalışması kapsamında yapılan testlerden bahsedilmiş ve elde edilen sonuçlar gösterilmiştir.

2. KULLANILAN YÖNTEMLER

Bu bölümde, bu çalışmada kullanılan yöntemlerden bahsedilmektedir. Bu yöntemler Yerel Zernike Momentleri, Rastgele Ormanlar ve Koşullu Regresyon Ormanları'dır.

2.1 Yerel Zernike Momentleri

Global görüntü momentleri, karakter ya da parmak izi gibi belirgin şekilleri içeren imgelerle yapılan sınıflandırma ve tanıma işlemlerinde başarılı sonuçlar vermektedir [3]. Zernike momentleri de bu momentlerden biridir. Zernike momentleri, her biri imgenin farklı bir karakteristiğine karşılık düşen farklı derecelerde radyal çokterimlilerini kullanarak imgeyi bütünsel olarak özgün bir biçimde tanımlamaktadır [9].

Yüz imgeleri ile gerçekleştirilen uygulamalarda, imgelerin içerdiği yerel bilgiler, bütünsel bilgilerle kıyaslandığında daha fazla önem taşımaktadır. Bu bilgiden hareket edilerek [3] çalışmasında, Zernike momentlerinin, her gözek etrafında yeni bir moment değerinin hesaplanmasıyla yerelleştirildiği Yerel Zernike Momentleri (YZM) yöntemi geliştirilmiştir. Geliştirilen bu yöntem ile yüz tanımada oldukça başarılı sonuçlar elde edilmiştir.

Yüzdeki nirengi noktalarının saptanmasında da imgelerin içerdiği yerel bilgiler önemli bir yer tutmaktadır. Bu sebeple, bu tez çalışmasında YZM yöntemi farklı koşullarla denenmiş ve kullanılan yöntemde imgelerden toplanan yamalara iki katmanlı olarak uygulanmıştır.

2.1.1 Yerel Zernike momentlerinin hesaplanması

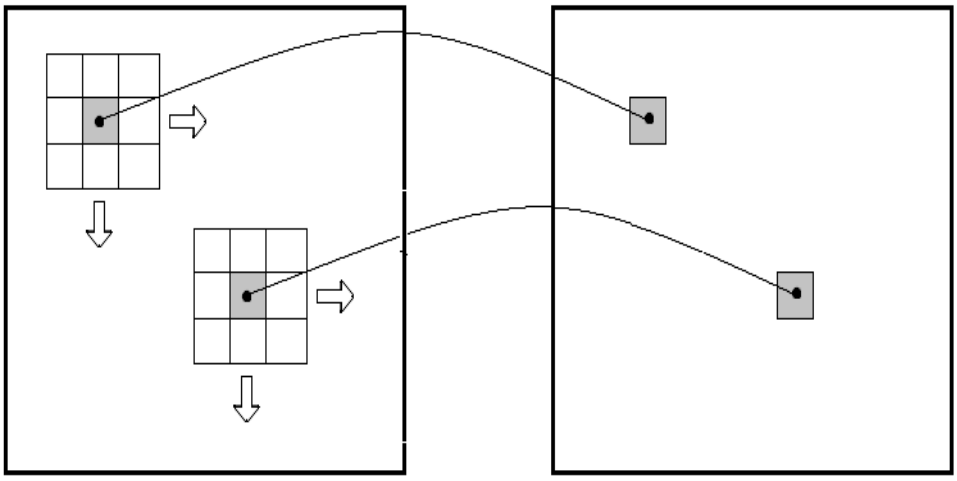
Moment değerlerinin tüm imgenin gözek değerleri kullanılarak hesaplanması yerine her gözeğin etrafında yerel bir şekilde hesaplandığı YZM dönüşümü, eşitlik 2.1 kullanılarak türetilen moment tabanlı V_{mn}^{k} çekirdeklerini kullanmaktadır.

$$\boldsymbol{V}_{nm}^{k}(i,j) = V_{nm}(\rho_{ij}, \theta_{ij}) \tag{2.1}$$

kxk boyutlu olan bu çekirdekler, imge filtrelemede sıklıkla kullanılan konvolusyon çekirdekleriyle benzerlik göstermektedirler. YZM dönüşümünün tanımı türetilen bu çekirdekler kullanılarak şu şekilde yapılmaktadır:

$$\boldsymbol{Z}_{nm}^{k}(i,j) = \sum_{p,q=-\frac{k-1}{2}}^{\frac{k-1}{2}} f(i-p, j-q)\boldsymbol{V}_{nm}^{k}(p,q) \tag{2.2}$$

YZM yönteminde moment değerleri *kxk* boyutundaki V_{mn}^{k} çekirdeklerinin imgenin üzerinde Şekil 2.1'de gösterildiği gibi kaydırılması ile hesaplanmaktadır. Bunun sonucunda her bir gözeğin çevresindeki yoğunluk değişimi ortaya çıkarılıp zengin bir imge gösterimi elde edilmektedir [3].



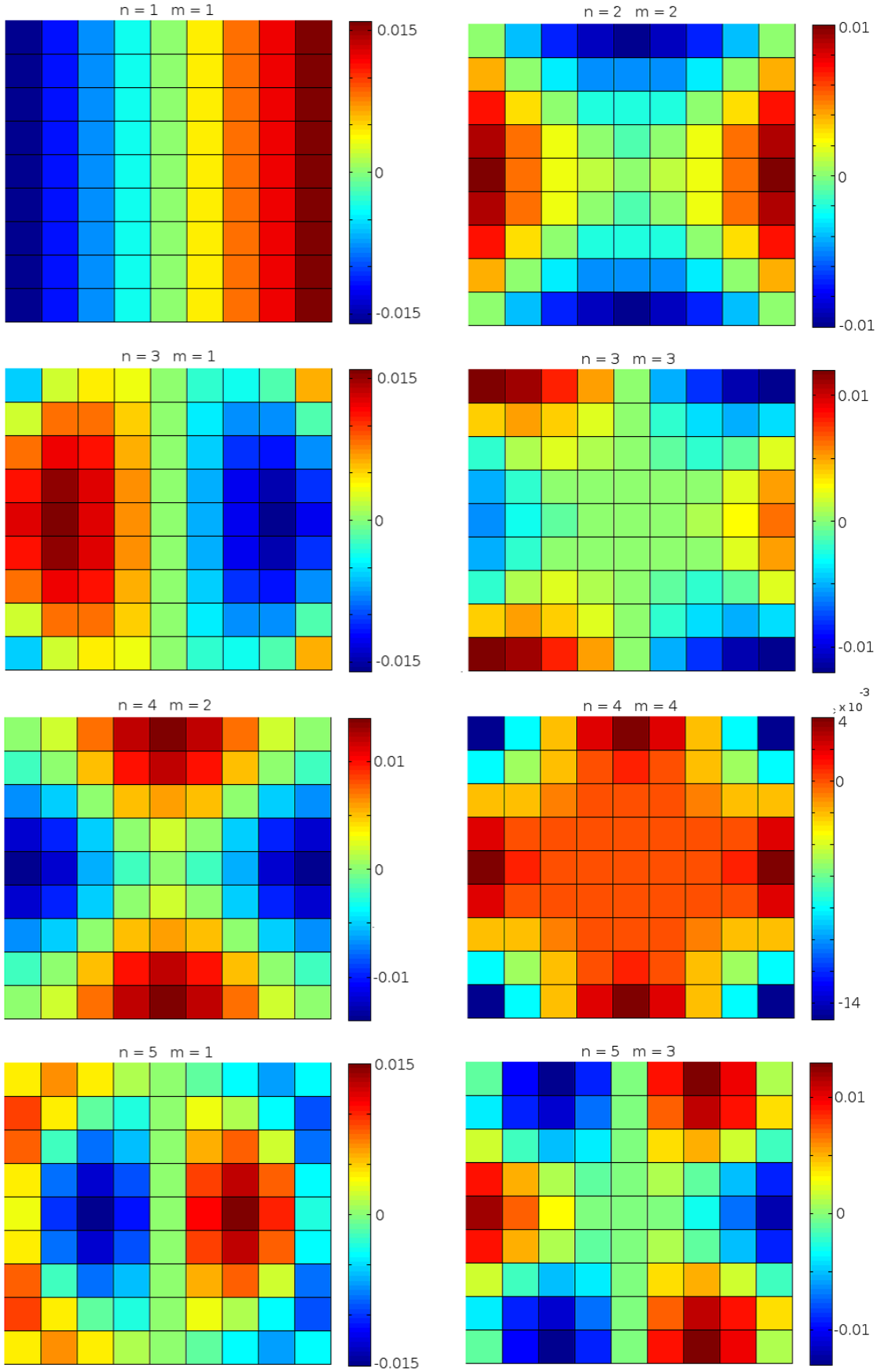
Şekil 2.1: YZM çekirdekleri ile yapılan konvolusyon işlemi.

Moment derecesi parametresi olan n, YZM yönteminde kullanılan en önemli parametrelerden biridir. Moment derecesi n'in alacağı değere bağlı olarak farklı sayıda gerçel ve sanal V_{mn}^{k} çekirdekleri elde edilmektedir. $k = 9$ için elde edilen çekirdeklerden ilk 8 tanesinin gerçel ve sanal bileşenleri sırasıyla Şekil 2.2 ve Şekil 2.3'te gösterilmektedir.

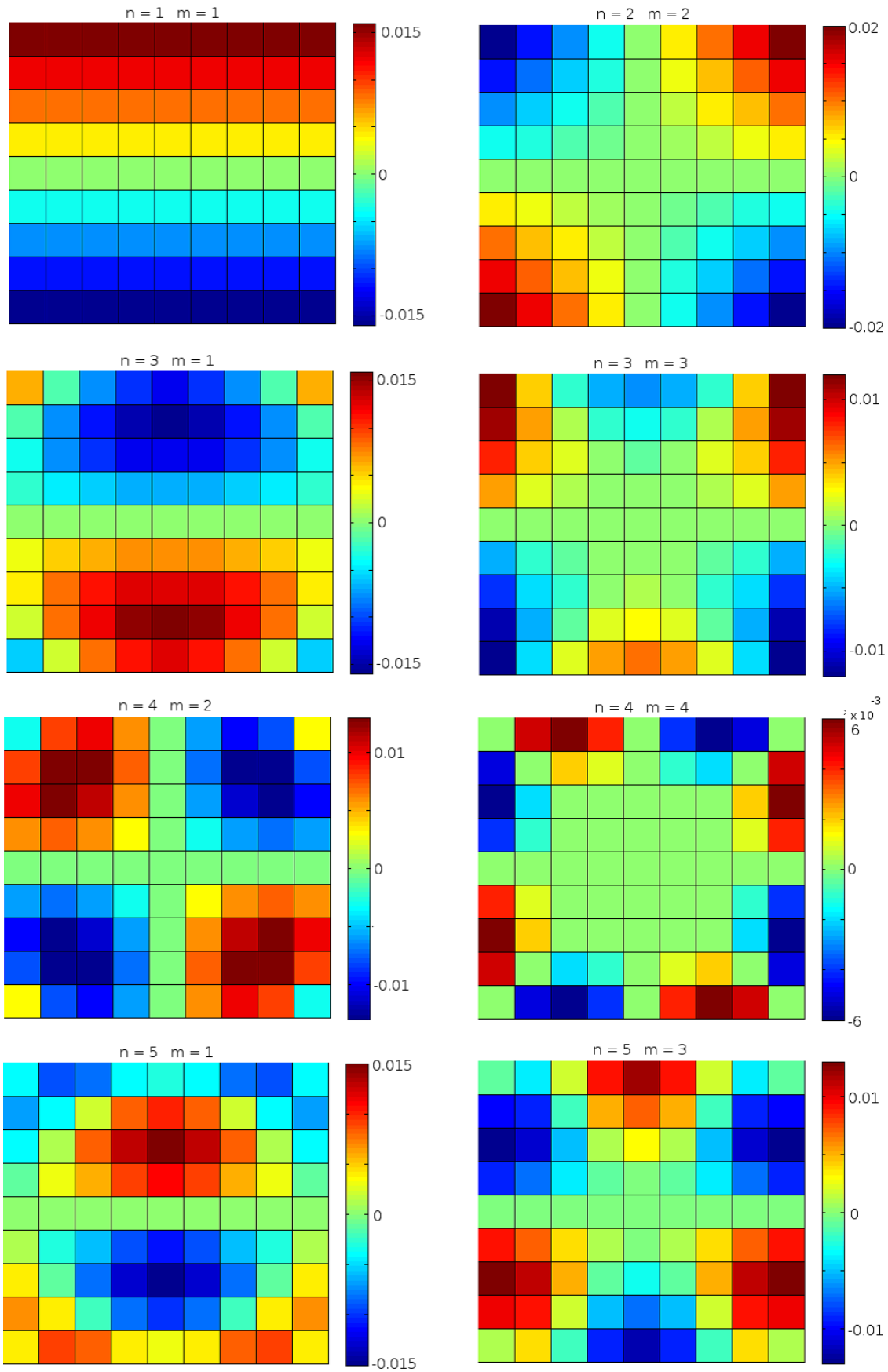
Eşitlik 2.2'den de anlaşılacağı gibi, YZM dönüşümü ile kullanılan gerçel ve sanal V_{mn}^{k} çekirdeği sayısı kadar gerçel ve sanal imgeler üretilmektedir. $m = 0$ için elde edilen filtreler, V_{mn}^{k} çekirdeklerinin sanal kısımlarının değeri 0 olduğundan dolayı YZM dönüşümünde kullanılmamaktadır.

YZM dönüşümü ile elde edilen karmaşık değerli imge sayısı şu şekilde hesaplanabilir:

$$K(n) = \begin{cases} \frac{n(n+2)}{4} & \text{n çift ise,} \\ \frac{(n+1)^2}{4} & \text{n tek ise.} \end{cases} \tag{2.3}$$



Şekil 2.2: $k = 9$ için elde edilen filtreleme çekirdeklerinden ilk 8 tanesinin gerçel bileşenleri [2].



Şekil 2.3: $k = 9$ için elde edilen filtreleme çekirdeklerinden ilk 8 tanesinin sanal bileşenleri [2].

2.1.2 YZM öznitelik vektörlerinin oluşturulması

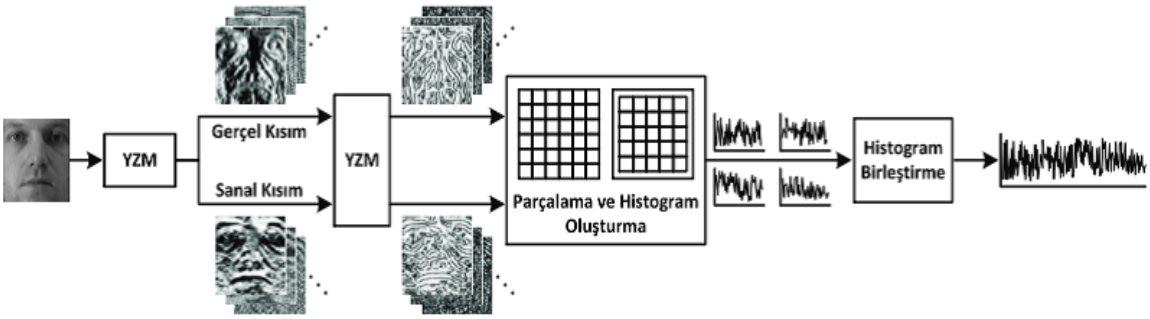
Yüz tanıma uygulamalarının birçoğunda, imgeleri alt bölgelere ayırarak işlem yapmanın avantajlarından bahsedilmiştir. Bundan dolayı, [3] çalışmasında da YZM imgeleri alt bölgelere ayrılmıştır ve bu işlem iki aşamalı olarak gerçekleştirilmiştir. İmgeler ilk aşamada NxN eşit boyutta alt bölgelere ayrılmaktadır. İkinci aşamada ise, imgeler sol üst köşeden yarım hücre kaydırılarak oluşturulan ızgara ile $(N-1)x(N-1)$ alt bölgeye ayrılmaktadır ve toplamda her imge için $N^2x(N-1)^2$ alt bölge elde edilmektedir.

YZM ile üretilen imgelerden öznitelik vektörü oluşturmak için tüm imgelerin alt alanlarından elde edilen bilgiler, histogram yöntemi kullanılarak birleştirilmektedir. Bu histogramlar, imgeler üzerinde farklı şekillerde oluşturulabilmektedir. [3] çalışmasında yapılan testler sonucunda, en iyi sonuçlar faz-genlik histogramları (FGH) kullanıldığında elde edilmiştir. Bu histogramları oluşturmak için öncelikle, $[0, 2\pi]$ aralığı b seleye ayrılmaktadır. Daha sonra her pikselin genlik değeri ($|\boldsymbol{Z}_{nm}^{k}(i,j)|$), aynı pikselin faz değerine ($\angle\boldsymbol{Z}_{nm}^{k}(i,j)$) karşılık gelen seleye eklenmektedir. Bu işlemler YZM imgelerinin tüm alt bölgelerinde uygulanarak yerel FGH'ler elde edilmektedir. Son işlem olarak ise, elde edilen FGH'ler birim boylu olacak şekilde düzgelendikten sonra art arda birleştirilerek sonuç öznitelik vektörü oluşturulmaktadır.

[3] çalışmasında , YZM dönüşümünün iki kere art arda uygulanmasıyla elde edilen imgeler de test edilmiş ve tek katmanlı YZM dönüşümünden daha başarılı sonuçlar elde edilmiştir. Çift katmanlı YZM dönüşümünde, birinci YZM dönüşümü sonucunda ışıktan bağımsız yerel şekil özellikleri elde edilirken, ikinci YZM dönüşümü sonucunda ise elde edilen şekil özelliklerinin istatistikleri belirlenmektedir. YZM'nin çift katlı olarak uygulandığı yöntemin genel gösterimi Şekil 2.4'te sunulmaktadır.

2.2 Rastgele Ormanlar

Rastgele Ormanlar (RO), temel olarak birden fazla karar ağacının kombinasyonundan oluşan bir yapı olarak tanımlanabilirler. Uygun ağaç yapılarıyla, sınıflandırma, kümeleme ve regresyon problemlerinde hızlı sonuçlar veren, çok sınıflı durumlar için



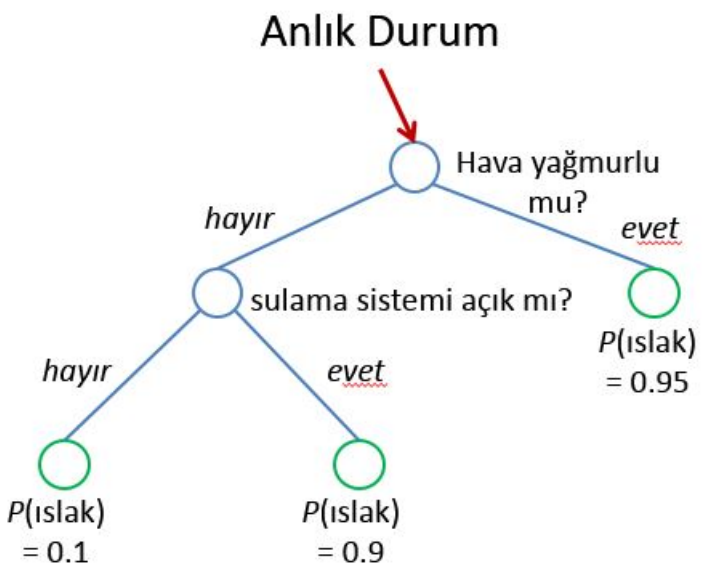
Şekil 2.4: Yüz tanımada kullanılan YZM yönteminin genel gösterimi [3].

uygun bir yöntem olan RO yöntemi, bilgisayarla görüde key point tanıma, nesne bölümleme, poz tahmini ve organların tespiti gibi uygulamalarda kullanılmaktadır.

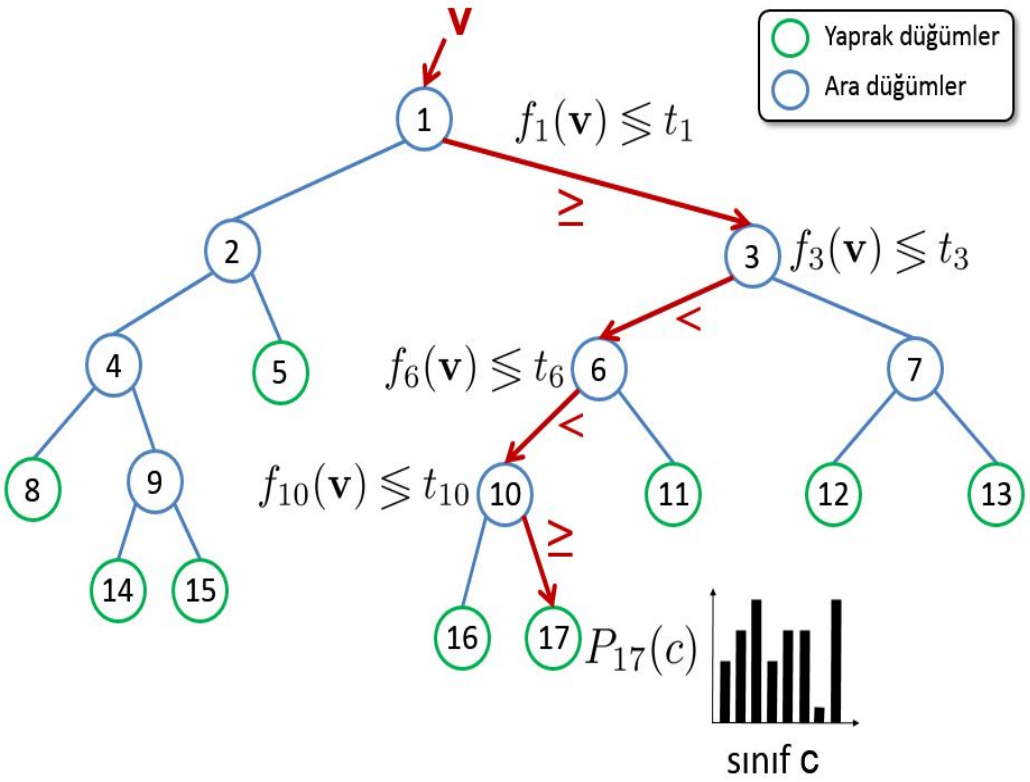
Karar ağaçları temel olarak, bir girdinin dallanma düğümlerindeki sorgulara göre sağa veya sola dallanarak yaprak düğüme ulaştığı ve ulaşılan yaprak düğümde belli bir olasılıkla bir sonucun üretildiği yapılardır. Şekil 2.5'te çimlerin ıslaklığına karar vermek için hazırlanmış basit bir karar ağacı sunulmuştur. Anlık durumu girdi olarak alan bu ağaç, ilk dallanma düğümünde "Hava yağmurlu mu?" sorgusuna alacağı cevaba göre dallanmanın yönünü tayin etmektedir. Bu sorguya alınacak "Evet" cevabıyla girdi sağa dallanacak ve yaprak düğüme ulaşacaktır. Bu yaprak düğümle anlık durum girdisi için bu karar ağacının vereceği sonuç %95 olasılıkla çimlerin ıslak olduğu olacaktır. İlk dallanma sorgusuna alınacak "Hayır" cevabı ise girdinin soldaki başka bir ara düğüme dallanmasını sağlayacak ve burada "Sulama sistemi açık mı?" sorgusuna alınacak cevaba bağlı olarak karar ağacının vereceği sonuç %10 ya da %90 olasılıkla çimlerin ıslak olduğu yönünde olacaktır.

Şekil 2.6'da sunulan ikili karar ağacı örneğinde ise Şekil 2.5'te gösterilen karar ağacına benzer şekilde, girdi olarak ağaca verilen öznitelik vektörü ara düğümlerde dallanma fonksiyonlarıyla değerlendirilip, belli bir eşik değerinden büyük veya küçük oluşuna göre sağa ya da sola dallanarak yaprak düğüme ulaşmakta ve yaprak düğümde olasılıksal sonuçlar elde edilmektedir.

Aynı öznitelik vektörünün birden fazla karar ağacında aynı anda Şekil 2.6'da olduğu gibi değerlendirildiği ve yaprak düğümlerde olasılıksal sonuçların elde edildiği karar ormanı örneği Şekil 2.7'de gösterilmektedir. Karar ormanlarında ağaçlarda elde edilen olasılıksal sonuçlar toplu olarak değerlendirilerek ortak bir karar alınmaktadır. Ortak karar, basitçe olasılıksal sonuçların ortalaması veya bu çalışmadaki gibi sonuçların

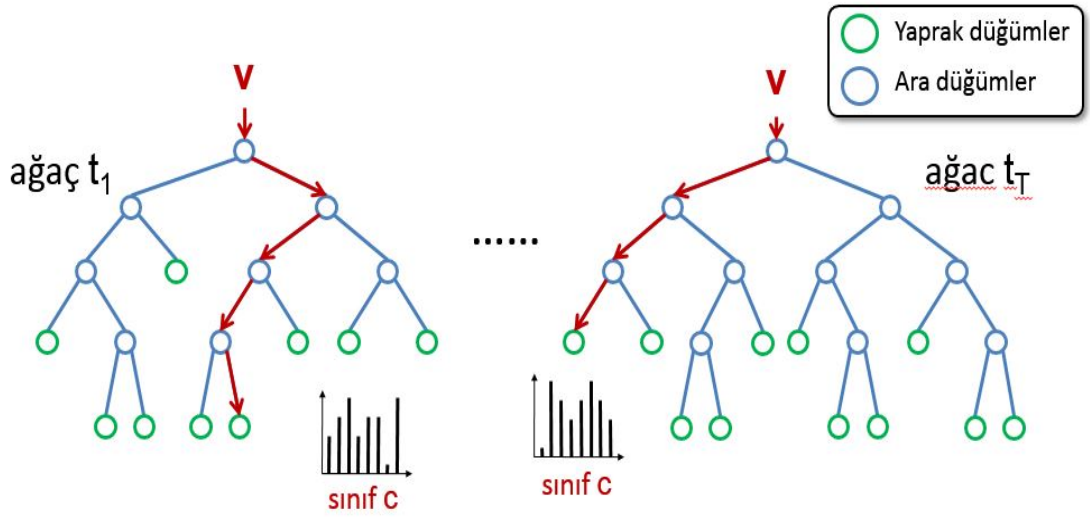


Şekil 2.5: Çimlerin ıslaklığının tespiti için örnek karar ağacı.



Şekil 2.6: Örnek ikili karar ağacı.

Gaussian kernel ve yoğunluk kestirici kullanılarak değerlendirilmesi gibi çeşitli şekillerde elde edilebilmektedir.



Şekil 2.7: Örnek karar ormanı.

RO yapısı ise Şekil 2.7'de gösterilen karar ağaçlarının eğitiminin rastgele eğitim kümeleriyle yapılması prensibine dayanmaktadır.

Bu çalışmada Rastgele Regresyon Ormanları(RRO), RO yönteminde olduğu gibi bütün eğitim kümesinden rastgele bir alt küme seçilerek eğitilmek yerine yüzün bakış açısına göre koşullu olarak eğitilmektedir. Bu sebeple kullanılan yöntem Koşullu Rastgele Regresyon Ormanları (KRRO) olarak adlandırılmaktadır.

2.2.1 Koşullu rastgele regresyon ormanları

Bu tez çalışmasında Koşullu Rastgele Regresyon Ormanları (KRRO) imgelerden rastgele olarak alınan görüntü yamaları ile yüzdeki nirengi noktaları arasındaki uzamsal ilişkiyi öğrenmektedir. Ayrıca yüzün bakış açısına göre koşullu olarak eğitilmesi sayesinde, yüzdeki bütün görünüm ve şekil değişimleri ile ilgilenmek yerine sadece ilgili bakış açısına ait olanları bilmeleri yeterli olduğu için bu yöntemde ağaçların öğrenmesi daha kolay olmaktadır.

Kullanılan KRRO yöntemi, temel olarak yüzün bakış açısının saptanmasında kullanılacak karar ormanının eğitimi, yüzdeki nirengi noktalarının saptanması için kullanılacak karar ormanlarının yüzlerin bakış açısına göre koşullu olarak eğitilmesi ve test aşamasında nirengi noktalarının yüzün bakış açısına uygun koşullu karar ağaçları kullanılarak saptanmasından oluşmaktadır.

2.2.1.1 Yüzün bakış açısının saptanması

Bu çalışmada kullanılan yöntemde yüzdeki nirengi noktalarının saptanmasında kullanılacak olan ağaçlar yüzün bakış açısına göre belirlenmektedir. Bu yüzden, öncelikli olarak yüzün bakış açısının tahmin edilmesi gerekmektedir. Bu amaçla LFW [10] veri seti kullanılarak, 'sol profil', 'sol', 'ön', 'sağ' ve 'sağ profil' olmak üzere 5 bakış açısı arasından bir saptama yapacak regresyon ormanı, [11, 12] çalışmalarına benzer şekilde eğitilmiştir. Eğitilen orman maksimum derinliği 10 olan 15 ağaçtan oluşmakta olup, %72 başarımla yüzün bakış açısını tahmin edebilmektedir.

2.2.1.2 Rastgele regresyon ormanlarının eğitimi

$T = T_t$ ormanındaki her T ağacı, birbirinden farklı ve rastgele seçilmiş eğitim imgeleri kümesiyle eğitilmektedir. Her imgeden rastgele olarak, I_i görünümü, D_i yüzdeki nirengi noktalarından kayıklığı temsil etmek üzere, $Pi = (I_i, D_i)$ kare yamalar kümesi elde edilir.

[1] çalışmasında, yamanın I_i bileşeni, 1 adet yamanın ham imge halinin gri değerlerini içeren kanal, 1 adet aydınlatma değişimlerini dengelemek için normalize edilmiş gri değerleri içeren kanal ve 8 farklı rotasyon, 4 farklı faz ile Gabor filtrelemesi sonucu elde edilen 32 kanal olmak üzere toplam 34 kanaldan oluşmaktadır. Bu çalışmada ise YZM [3] yöntemiyle elde edilen bileşenler, yüzdeki nirengi noktalarının saptanmasında Bölüm 3'de bahsedileceği üzere çeşitli şekillerde [1] çalışmasında kullanılan Gabor kanalları yerine kullanılmıştır. Yamanın $D_i = (d_i^1, d_i^2, d_i^3, \ldots, d_i^{10})$ bileşeni ise yamanın merkezinin her bir yüz özniteliğinden dikey ve yatay yönlü kayıklığını belirten 2 boyutlu vektörleri içermektedir.

Yama karşılaştırma özniteliği, [1] çalışmasında olduğu gibi, R_1 ve R_2 yama içerisindeki iki dikdörtgeni, α ise seçilen görünüm kanalını temsil etmek üzere $\theta = (R_1, R_2, \alpha)$ parametreleriyle 2.4'te gösterildiği şekilde belirlenmiştir.

$$f_\theta(P) = \frac{1}{|R_1|} \sum_{q \varepsilon R_1} I^\alpha(q) - \frac{1}{|R_2|} \sum_{q \varepsilon R_2} I^\alpha(q) \tag{2.4}$$

Yöntemin eğitim adımları Breiman [8] tarafından ortaya atılmış olan rastgele orman çatısını (framework) takip etmektedir:

- Dallanma adaylarından $\phi = (\theta, \tau)$ bir havuz oluşturulur.

- Her bir ϕ için yamalar kümesi P_L ve P_R olmak üzere 2.5 ve 2.6'da ifade edildiği gibi iki alt kümeye bölünür.

$$P_L(\phi) = \{P | f_\theta(P) < \tau\} \tag{2.5}$$

$$P_R(\phi) = P \setminus P_L(\phi) \tag{2.6}$$

- $H(P)$ sınıf belirsizlik ölçütü olmak üzere; bilgi kazancı (Information Gain - IG) değerlendirme fonksiyonunu maksimum yapan dallanma adayı 2.7 ve 2.8 eşitlikleri ile gösterilen şekilde seçilir.

$$\phi^* = \arg\max_{\phi} IG(\phi) \tag{2.7}$$

$$IG(\phi) = H(P) - \sum_{S\varepsilon\{L,R\}} \frac{|P_S(\phi)|}{|P|} H(P_S(\phi)) \tag{2.8}$$

- Ağaç için maksimum derinliğe ulaşıldıysa ya da bilgi kazancı (IG) önceden belirlenmiş bir eşik değerinin altına düştüyse yaprak düğümü oluşturulur. Aksi durumda eğitim P_L ve P_R alt kümeleri için birinci adımdan özyinelemeli olarak devam eder.

Rastgele ormanların eğitiminde LFW veri setindeki [10] 10 nirengi noktasının koordinatları belirlenmiş 13.233 imge kullanılmaktadır. Her bir ağacın eğitimi sırasında LFW veri setinden 1500 imge rastgele olarak seçilmekte ve Viola-Jones [13] yüz bulma algoritmasıyla yüzler bulunarak $100x100$ gözek boyutlarında olacak şekilde yeniden ölçeklenmektedir. Bulunan ve yeniden ölçeklenen yüz bölgesi %30 genişletilerek yüzdeki bütün nirengi noktalarını içermesi kesinleştirilmeye çalışılmaktadır. Bu aşamadan sonra, her imge için yüz bölgesinden 150 adet, imgenin kalan kısmından 50 adet olmak üzere toplam 200 adet $20x20$ gözek boyutunda rastgele yama alınarak Breiman'ın [8] rastgele eğitim prosedürüyle, her bir bakış açısı için maksimum derinliği 20 olan 10 adet rastgele orman ağacı oluşturulmaktadır.

2.2.1.3 Yüzdeki nirengi noktalarının saptanması

Yüzdeki nirengi noktaları bulunurken ilk adım olarak Viola-Jones [13] yüz bulma algoritmasıyla yüzün yeri ve büyüklüğü bulunmakta ve rastgele ormanların eğitimi aşamasında olduğu gibi yeniden ölçeklenip %30 oranında genişletilmektedir. Yine eğitim aşamasına benzer şekilde, çoğunlukla yüz bölgesinden yamalar alınmaktadır.

Bu yamalar, yüzün bakış açısına göre koşullu olarak yüzdeki nirengi noktalarını bulmakta kullanılacak ağaçlara verilmektedir. Yamalar düğümlerde değerlendirilmekte ve bu değerlendirme sonucu ağaçta sağa veya sola ilerleyerek yaprak düğüme ulaşmaktadır. Bu işlemin çok sayıda ağaçta gerçekleşmesi sonucu, her yama için yaprak düğümler kümesi L_i elde edilmektedir. Elde edilen L_i yaprak düğümleri kümesi, Gaussian çekirdeği ile birlikte yoğunluk kestiricide kullanılarak yüzdeki nirengi noktalarının yeri kestirilmeye çalışılmaktadır. Yüzdeki nirengi noktaları bu işlem sonucu kestirilen noktaların her biri için Mean-Shift uygulanması sonucu elde edilmektedir.

3. DENEYSEL ÇALIŞMALAR

Bu bölümde, kullanılan veri seti ile ilgili genel bilgiler verilmektedir. Ayrıca bu tez çalışması kapsamında yapılan deneysel çalışmalar ve elde edilen sonuçlardan da söz edilmektedir.

3.1 Kullanılan Veri Seti

Bu tez çalışmasında kullanılan yöntemi test etmek amacıyla, 5749 kişinin 13233 yüz imgesinin bulunduğu ve çok çeşitli görüntüleme koşullarına sahip imgelerden oluşan Labeled Faces in the Wild (LFW) [10] veri seti kullanılmıştır. Bu veri setinin seçilmiş olmasının sebebi, farklı pozları, değişik ışıklandırma koşulları, çözünürlük ve imge kalitelerini, faklı yüz ifadelerini ve cinsiyetleri ve yüz bölümlerinin kapanması gibi saptama işlemini zorlaştırabilecek etkenleri barındırdan imgelerin bulunduğu, zorlayıcı bir veri seti olmasıdır.

Bu çalışmada kullanılan yöntemin başarımı, LFW veri setinde çeşitli hata toleransları belirlenerek değerlendirilmiştir. Tolerans, hata payının iki göz arası uzaklığa oranının alabileceği maksimum değerdir. LFW veritabanı için iki göz arası uzaklık ortalama 50.247 gözektir.

3.2 Yapılan Çalışmalar

Bu tez çalışması kapsamında ilk olarak yüzlerin bakış açılarının ve nirengi noktalarının saptanmasında kullanılacak rastgele ormanlar eğitilmiş, parametre değişimleriyle, [1] çalışmasında Çizelge 3.1'de sunulduğu gibi elde edildiği belirtilen başarımlara en yakın sonuçları veren ormanlar oluşturulmaya çalışılmıştır. Bu kapsamda elde edilebilen, [1] çalışmasına en yakın başarımlar farklı toleranslarla Çizelge 3.2'de görüldüğü gibi olmuştur. Bu sonuçlar YZM ile elde edilen başarımlarla karşılaştırmada baz olarak kullanılmıştır.

Çizelge 3.1: [1] çalışmasında bildirilen başarım sonuçları.

Nirengi Noktası	Tolerans (%)
	10
Sol Göz (Sol)	87,7%
Sol Göz (Sağ)	93,5%
Ağız (Sol)	81,9%
Ağız (Sağ)	80,8%
Ağız (Alt)	71,5%
Ağız (Üst)	86,7%
Sağ Göz (Sol)	92,9%
Sağ Göz (Sağ)	86,2%
Burun (Sol)	90,4%
Burun (Sağ)	88,2%
Ortalama:	**85,98%**

Çizelge 3.2: [1] çalışmasına en yakın rastgele ormanlarla elde edilen başarım sonuçları.

Nirengi Noktası	Tolerans (%)			
	5	**10**	**15**	**20**
Sol Göz (Sol)	39,68%	78,54%	91,05%	93,78%
Sol Göz (Sağ)	50,17%	91,35%	98,26%	98,82%
Ağız (Sol)	27,44%	65,97%	88,19%	95,59%
Ağız (Sağ)	30,22%	74,44%	92,39%	96,74%
Ağız (Alt)	20,67%	57,29%	81,85%	92,22%
Ağız (Üst)	27,15%	71,55%	91,80%	97,46%
Sağ Göz (Sol)	44,99%	87,50%	97,78%	98,99%
Sağ Göz (Sağ)	37,11%	80,37%	91,96%	94,14%
Burun (Sol)	34,64%	78,57%	94,96%	98,38%
Burun (Sağ)	36,27%	79,63%	95,28%	98,55%
Ortalama:	**34,83%**	**76,52%**	**92,35%**	**96,47%**

Koşullu rastgele regresyon ormanlarının YZM ile entegrasyonunu konu edinen bu çalışmada, ilk aşamada, çift katmanlı YZM [3] uygulanması sonucu elde edilen 40 bileşenin, [1] çalışmasında kullanılan yöntemdeki Gabor kanalları yerine kullanıldığı bir test ortamında, yüz görüntüsünün normalizasyonu ile ağaçların eğitiminde kullanılan kanalların oluşturulmasında integral bilgisinin kullanılmasının başarıma etkisi incelenmiştir. Bu inceleme sonucunda integral bilgisinin kullanıldığı durumlarda hem çift katmanlı hem de tek katmanlı YZM ile Çizelge 3.3 ve 3.4'te örnek olarak sunulduğu gibi düşük başarımlı sonuçlar elde edilmiştir. İntegral bilgisinin kullanılmadığı durumlarda elde edilen sonuçlar ise Çizelge 3.5 ve 3.6'da görüldüğü gibi olmuştur.

Çizelge 3.3: Çift katmanlı YZM ile normalizasyon uygulanmayıp integral bilgisi kullanılarak, 40 bileşenle elde edilen başarım sonuçları.

Nirengi Noktası	Tolerans (%)			
	5	**10**	**15**	**20**
Sol Göz (Sol)	0,10%	0,61%	1,59%	3,11%
Sol Göz (Sağ)	3,14%	10,85%	22,23%	34,25%
Ağız (Sol)	0,02%	0,17%	0,72%	2,21%
Ağız (Sağ)	0,03%	0,13%	0,42%	1,42%
Ağız (Alt)	0,00%	0,05%	0,23%	0,74%
Ağız (Üst)	1,53%	6,41%	15,68%	27,99%
Sağ Göz (Sol)	3,81%	13,48%	27,90%	42,59%
Sağ Göz (Sağ)	0,01%	0,06%	0,23%	0,62%
Burun (Sol)	2,36%	9,18%	20,64%	36,01%
Burun (Sağ)	1,90%	8,05%	19,05%	35,11%
Ortalama:	**1,29%**	**4,90%**	**10,87%**	**18,41%**

Çizelge 3.4: YZM'nin 1. katmanı ile normalizasyon uygulanıp, integral bilgisi kullanılarak 40 bileşenle elde edilen başarım sonuçları.

Nirengi Noktası	Tolerans (%)			
	5	**10**	**15**	**20**
Sol Göz (Sol)	0,34%	1,01%	2,40%	4,50%
Sol Göz (Sağ)	5,18%	17,34%	32,03%	45,97%
Ağız (Sol)	4,47%	15,89%	32,92%	50,55%
Ağız (Sağ)	1,84%	9,49%	23,93%	42,02%
Ağız (Alt)	0,50%	2,92%	8,88%	18,77%
Ağız (Üst)	7,59%	27,75%	52,46%	71,48%
Sağ Göz (Sol)	9,90%	30,82%	51,58%	67,18%
Sağ Göz (Sağ)	0,20%	1,10%	2,86%	5,38%
Burun (Sol)	23,81%	60,47%	83,26%	92,61%
Burun (Sağ)	13,77%	42,09%	68,70%	85,15%
Ortalama:	**6,76%**	**20,89%**	**35,90%**	**48,36%**

Çizelge 3.5 ve 3.6'da görüleceği üzere çift katmanlı YZM ile en iyi başarım sonuçlarına normalizasyon uygulanıp, integral bilgisi kullanılmadan elde edilen 40 bileşenle ulaşılmıştır.

Bu aşamadan hareketle, tek katmanlı YZM ile normalizasyon uygulanıp integral bilgisi kullanılmadan oluşturulacak 40 bileşenle elde edilecek başarımlar incelenmiştir. Sadece YZM'nin 1. katmanı kullanılarak elde edilen başarım sonuçları Çizelge 3.7'de, sadece YZM'nin 2. katmanı kullanılarak elde edilen başarım sonuçları ise Çizelge 3.8'de gösterildiği gibi olmuştur.

Çizelge 3.5: Çift katmanlı YZM ile normalizasyon uygulanmadan ve integral bilgisi kullanılmadan, 40 bileşenle elde edilen başarım sonuçları.

Nirengi Noktası	**Tolerans (%)**			
	5	**10**	**15**	**20**
Sol Göz (Sol)	28,34%	66,77%	83,86%	89,43%
Sol Göz (Sağ)	38,26%	84,55%	96,58%	98,22%
Ağız (Sol)	24,36%	59,95%	82,26%	92,12%
Ağız (Sağ)	29,69%	69,74%	88,49%	94,44%
Ağız (Alt)	22,44%	56,98%	79,37%	89,58%
Ağız (Üst)	29,44%	71,61%	91,09%	96,81%
Sağ Göz (Sol)	39,60%	84,14%	96,39%	98,50%
Sağ Göz (Sağ)	28,44%	70,86%	87,24%	91,23%
Burun (Sol)	31,60%	74,27%	92,47%	97,10%
Burun (Sağ)	34,16%	77,28%	93,92%	97,64%
Ortalama:	**30,63%**	**71,62%**	**89,16%**	**94,51%**

Çizelge 3.6: Çift katmanlı YZM ile normalizasyon uygulanıp, integral bilgisi kullanılmadan, 40 bileşenle elde edilen başarım sonuçları.

Nirengi Noktası	**Tolerans (%)**			
	5	**10**	**15**	**20**
Sol Göz (Sol)	30,49%	68,77%	85,31%	90,26%
Sol Göz (Sağ)	39,45%	85,20%	96,77%	98,38%
Ağız (Sol)	24,31%	60,41%	82,41%	92,04%
Ağız (Sağ)	30,12%	70,16%	88,66%	94,97%
Ağız (Alt)	22,23%	57,13%	79,60%	89,71%
Ağız (Üst)	30,02%	71,72%	90,97%	96,88%
Sağ Göz (Sol)	39,90%	84,21%	96,48%	98,51%
Sağ Göz (Sağ)	29,50%	71,80%	87,57%	91,25%
Burun (Sol)	33,09%	75,49%	92,69%	97,13%
Burun (Sağ)	34,86%	78,06%	93,91%	97,60%
Ortalama:	**31,40%**	**72,30%**	**89,44%**	**94,67%**

Çizelge 3.6, 3.7 ve 3.8 karşılaştırıldığında en iyi sonuçların çift katmanlı YZM ile elde edildiği görülmüştür. Bunun üzerine, yöntemin başarımı, çift katmanlı YZM ile aynı koşullarda değişik sayıda bileşenler oluşturularak test edilmiştir. Bu testlerde, YZM yönteminin 1. katman ve 2. katman için moment derecelerini temsil eden n_1 ve n_2 parametrelerine değişik değerler verilip, değişik sayıda bileşenler oluşturulmuştur. n_1 ve n_2 parametrelerinin değişik değerleriyle elde edilen farklı sayılarda bileşenler ile %5, %10, %15 ve %20 toleranslarda ortalama başarımdaki değişim, Şekil 3.1, 3.2, 3.3 ve 3.4'te gösterildiği gibi olmuştur.

Çizelge 3.7: YZM'nin 1. katmanı ile normalizasyon uygulanıp, integral bilgisi kullanılmadan 40 bileşenle elde edilen başarım sonuçları.

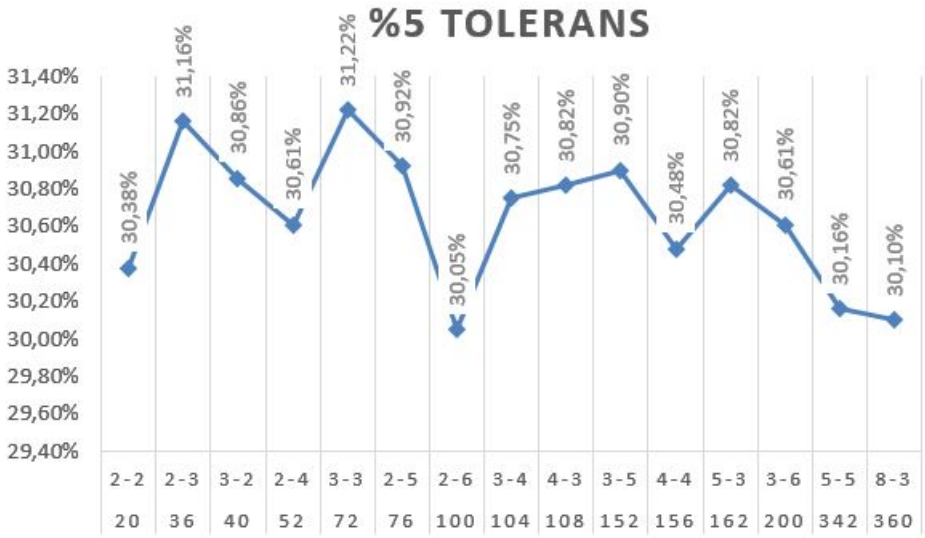
Nirengi Noktası	Tolerans (%)			
	5	10	15	20
Sol Göz (Sol)	29,56%	68,24%	84,90%	90,06%
Sol Göz (Sağ)	37,63%	83,91%	96,39%	98,25%
Ağız (Sol)	22,05%	55,83%	78,35%	89,32%
Ağız (Sağ)	27,37%	66,60%	86,96%	93,85%
Ağız (Alt)	22,39%	56,49%	77,96%	88,68%
Ağız (Üst)	27,75%	69,24%	89,35%	95,87%
Sağ Göz (Sol)	38,94%	83,58%	96,29%	98,35%
Sağ Göz (Sağ)	29,74%	71,91%	87,56%	91,66%
Burun (Sol)	30,22%	71,93%	90,90%	96,41%
Burun (Sağ)	33,59%	75,88%	93,13%	97,38%
Ortalama:	**29,92%**	**70,36%**	**88,18%**	**93,98%**

Çizelge 3.8: YZM'nin 2. katmanı ile normalizasyon uygulanıp, integral bilgisi kullanılmadan 40 bileşenle elde edilen başarım sonuçları.

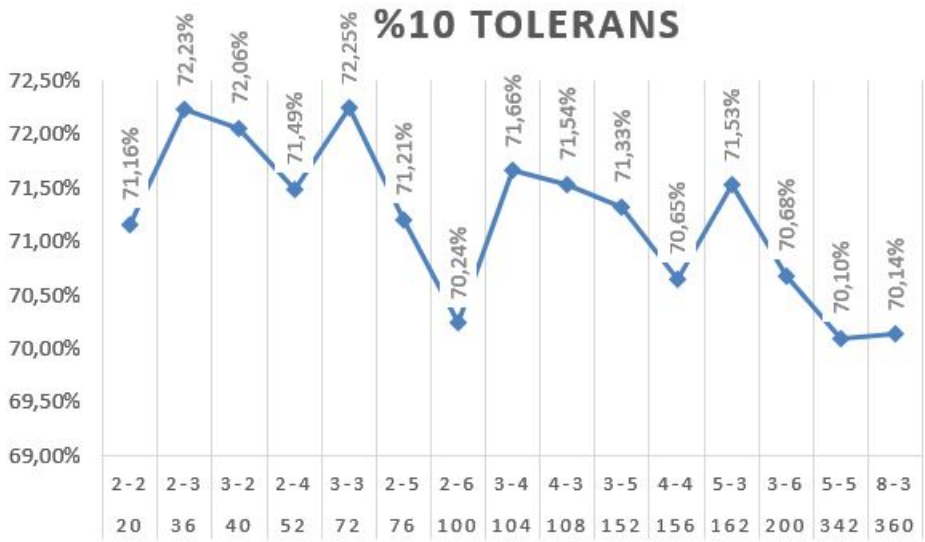
Nirengi Noktası	Tolerans (%)			
	5	10	15	20
Sol Göz (Sol)	29,97%	66,70%	82,02%	86,65%
Sol Göz (Sağ)	45,11%	88,00%	96,49%	97,63%
Ağız (Sol)	23,76%	60,21%	83,00%	92,27%
Ağız (Sağ)	25,76%	68,52%	88,44%	94,75%
Ağız (Alt)	19,46%	52,99%	77,05%	88,06%
Ağız (Üst)	26,71%	69,89%	90,43%	96,63%
Sağ Göz (Sol)	40,19%	84,00%	95,81%	97,88%
Sağ Göz (Sağ)	30,09%	70,32%	84,45%	87,82%
Burun (Sol)	33,99%	76,92%	93,31%	97,32%
Burun (Sağ)	32,08%	75,89%	93,24%	97,42%
Ortalama:	**30,71%**	**71,34%**	**88,43%**	**93,64%**

Bu sonuçlar incelendiğinde bileşen sayısının başarım ile doğru orantılı bir ilişkiye sahip olmadığı ve belirli bir aralıkta daha iyi sonuçlar verdiği gözlemlenmiştir. Elde edilen bu sonuçlara bakıldığında, ortalamalara göre en yüksek başarım, $n_1 = 3$ ve $n_2 = 3$ parametreleriyle elde edilmiştir. $n_1 = 3$ ve $n_2 = 3$ parametreliyle elde edilen başarımlar Çizelge 3.9'da sunulduğu gibidir.

YZM ile elde edilen bu başarımların yanı sıra, YZM ve Gabor bileşenlerinin birlikte kullanımının başarıma etkisi de incelenmiştir. Bu doğrultuda, Çizelge 3.9'daki başarım sonuçlarını veren $n_1 = 3$ ve $n_2 = 3$ parametreleri ile oluşturulan 72 adet çift katmanlı YZM bileşeniyle, [1] çalışmasında olduğu gibi 8 rotasyon ve 4 fazdan oluşan 32 gabor



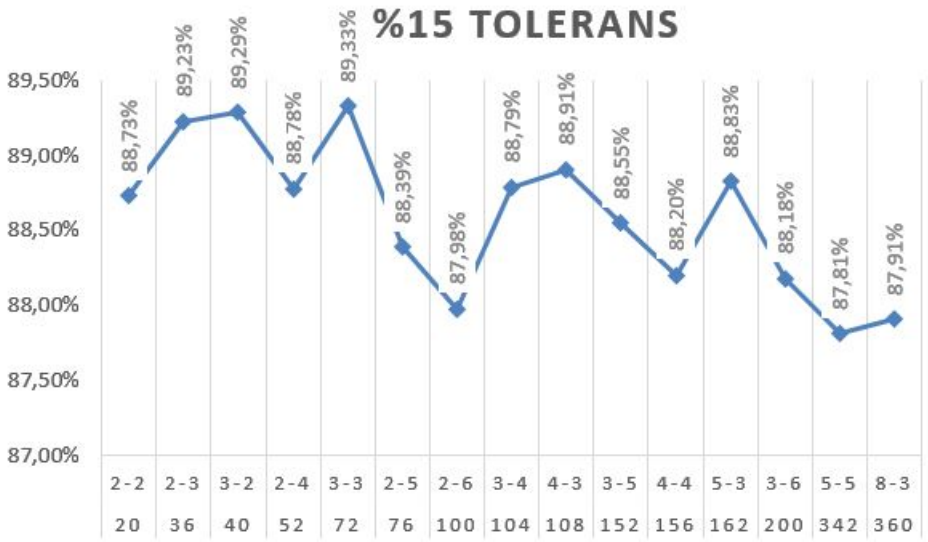
Şekil 3.1: %5 Tolerans için n_1 ve n_2 parametrelerinin ve bileşen sayısının başarıma etkisi.



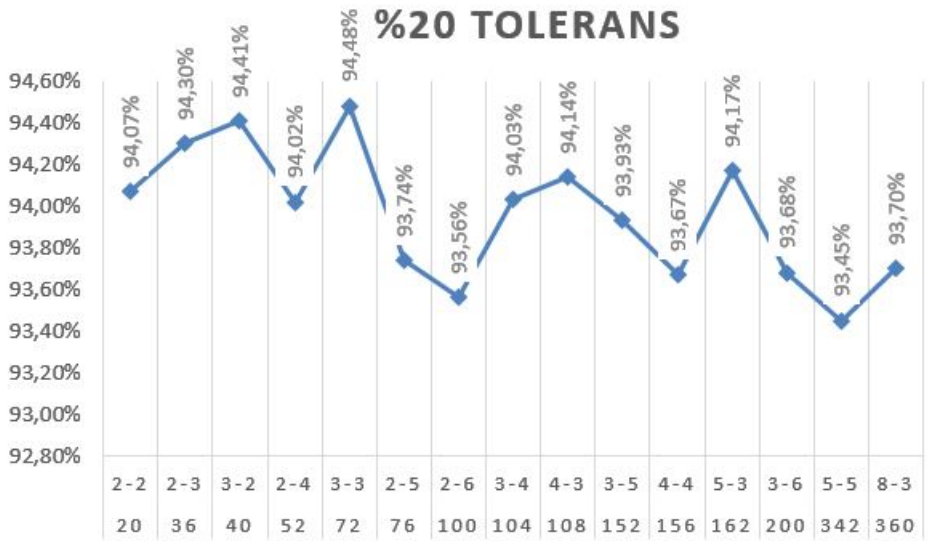
Şekil 3.2: %10 Tolerans için n_1 ve n_2 parametrelerinin ve bileşen sayısının başarıma etkisi.

bileşeni birlikte kullanılarak test edilmiştir. YZM ve Gabor bileşenlerinin bu şekilde kullanılmasıyla elde edilen başarımlar Çizelge 3.10'da gösterildiği şekilde olmuştur.

Çizelge 3.10'da elde edilen sonuçlar incelendiğinde Gabor bileşenlerinin kullanımının, çift katmanlı YZM bileşenleriyle elde edilen başarım sonuçlarını yükselttiği gözlemlenmiştir. Değişik Gabor ve YZM bileşen sayılarıyla bu test birkaç kere tekrarlanmış ama daha yüksek başarım sonuçlarına ulaşılamamıştır.



Şekil 3.3: %15 Tolerans için n_1 ve n_2 parametrelerinin ve bileşen sayısının başarıma etkisi.



Şekil 3.4: %20 Tolerans için n_1 ve n_2 parametrelerinin ve bileşen sayısının başarıma etkisi.

Bunun yanı sıra, sunulan çizelgelerdeki sonuçlar incelendiğinde, yüz bölgesinin daha iç kısmında kalan "Sol Göz (Sağ)" ve "Sağ Göz (Sol)" noktaları daha yüksek başarımlarla saptanırken, yüzün bölgesinin dışına yakın bölgelerde yer alan "Sol Göz (Sol)" ve "Sağ Göz (Sağ)" noktaları nispeten daha düşük başarımlarla saptanabilmiştir. Aynı sebepten dolayı, yüz bölgesinin dış kısmına yakın bölgelerde yer alan bir başka nirengi noktası olan "Ağız (Alt)" noktasının da düşük başarımlarla saptanabildiği gözlemlenmiştir.

Çizelge 3.9: $n_1 = 3$, $n_2 = 3$ parametreleriyle çift katmanlı YZM kullanılarak elde edilen başarım sonuçları.

Nirengi Noktası	Tolerans (%)			
	5	10	15	20
Sol Göz (Sol)	29,75%	67,71%	84,30%	89,13%
Sol Göz (Sağ)	41,44%	86,78%	96,80%	98,19%
Ağız (Sol)	24,65%	61,00%	83,15%	92,29%
Ağız (Sağ)	28,92%	70,40%	89,20%	95,00%
Ağız (Alt)	21,31%	55,32%	78,42%	89,34%
Ağız (Üst)	28,73%	71,01%	90,77%	96,71%
Sağ Göz (Sol)	40,33%	84,84%	96,46%	98,44%
Sağ Göz (Sağ)	29,39%	71,50%	87,18%	90,76%
Burun (Sol)	33,40%	76,17%	93,08%	97,40%
Burun (Sağ)	34,28%	77,77%	93,90%	97,51%
Ortalama:	**31,22%**	**72,25%**	**89,33%**	**94,48%**

Çizelge 3.10: Çift katmanlı YZM ve Gabor bileşenlerinin birlikte kullanımıyla elde edilen başarım sonuçları.

Nirengi Noktası	Tolerans (%)			
	5	10	15	20
Sol Göz (Sol)	38,59%	77,67%	90,40%	93,46%
Sol Göz (Sağ)	48,58%	90,95%	98,07%	98,81%
Ağız (Sol)	27,04%	65,39%	87,43%	95,24%
Ağız (Sağ)	29,81%	73,88%	91,77%	96,55%
Ağız (Alt)	20,84%	57,70%	81,41%	91,88%
Ağız (Üst)	26,66%	70,03%	91,17%	97,17%
Sağ Göz (Sol)	44,11%	87,21%	97,63%	99,01%
Sağ Göz (Sağ)	36,44%	79,67%	91,72%	93,95%
Burun (Sol)	33,86%	77,93%	94,63%	98,26%
Burun (Sağ)	36,15%	79,86%	95,17%	98,62%
Ortalama:	**34,21%**	**76,03%**	**91,94%**	**96,29%**

Ayrıca, bu tez çalışması sırasında çift katmanlı YZM kullanılarak LFW veri seti ile eğitilen KRRO, iki göz arası uzaklığın ortalama 132.147 gözek olduğu 512x768 gözek boyutunda imgeler içeren FERET veritabanında da test edilmiş ve Çizelge 3.11'de gösterilen başarımlar elde edilmiştir. FERET veri setinde "Sağ Göz", "Sol Göz", "Burun" ve "Ağız" olmak üzere 4 adet nirengi noktası işaretli olarak bulunmaktadır. Bu yüzden, başarım sonuçları hesaplanırken "Sol Göz" nirengi noktası için saptanan "Sol Göz (Sol)" ve "Sol Göz (Sağ)" nirengi noktalarının, "Sağ Göz" nirengi noktası için saptanan "Sağ Göz (Sol)" ve "Sağ Göz (Sağ)" nirengi noktalarının, "Burun" için saptanan "Burun(Sol)" ve "Burun(Sağ)" nirengi noktalarının, "Ağız" için ise "Ağız

(Sol)", "Ağız(Sağ)", "Ağız (Alt)" ve "Ağız(Üst)" nirengi noktalarının orta noktası hesaplanarak karşılaştırma yapılmıştır.

Çizelge 3.11: Çift katmanlı YZM kullanılarak FERET veri setinde elde edilen başarım sonuçları.

Nirengi Noktası	**Tolerans (%)**			
	5	**10**	**15**	**20**
Sağ Göz	1,28%	37,90%	88,32%	94,03%
Sol Göz	16,68%	70,70%	92,04%	93,35%
Burun	0,71%	9,05%	41,32%	76,15%
Ağız	2,86%	33,55%	74,19%	90,53%
Ortalama:	**5,38%**	**37,80%**	**73,97%**	**88,51%**

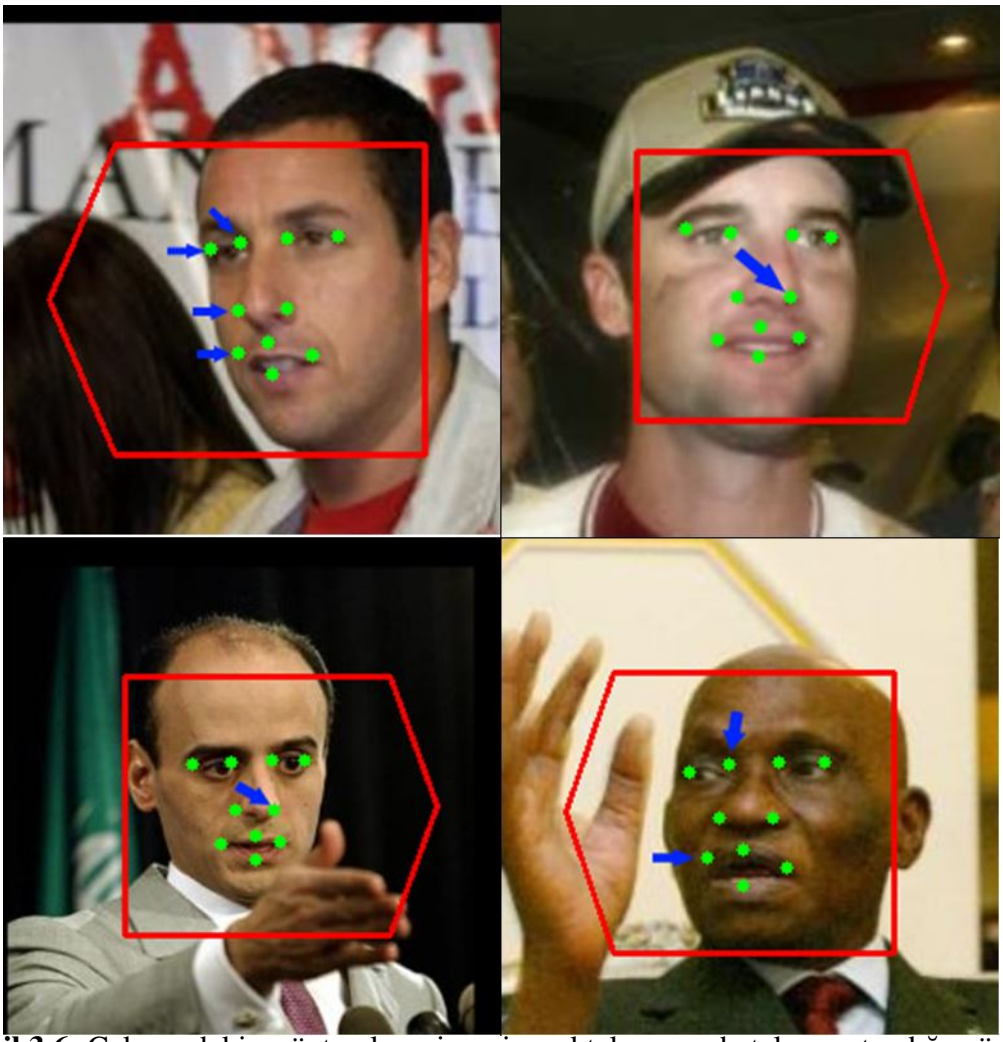
Çizelge 3.11 ile sunulan sonuçlarda görülebileceği üzere, FERET veri setinde LFW veri setine kıyasla daha düşük başarımlar elde edilmiştir. Bunun sebebi, kullanılan KRRO'nın LFW veri seti kullanılarak 10 adet nirengi noktası saptamak üzere eğitilmiş olmasıdır.

Çizelge 3.9'da sonuçları verilen yöntem, bu tez çalışması kapsamında yüzdeki nirengi noktalarının saptanması konusunda sadece YZM kullanarak yaptığımız çalışmalar arasında en yüksek başarıma sahiptir. Bu yöntemle elde edilen çıktılardan nirengi noktalarının isabetli saptandığı birkaç örnek Şekil 3.5'te sunulmuştur. Bu örnek çıktılarda da görülebileceği üzere, hatalı saptanan nirengi noktalarının bir kısmı yüzün bakış açısından dolayı görüş açısında bulunmayan ya da çok dar bir imge alanında görülebilen noktalardır. Bunun yanı sıra gözlük, yüzün duruş açısı veya yüzün bir bölümünün kapanması gibi etkenler de nirengi noktalarının saptanmasında sapmalara yol açmaktadır.

Bunun yanı sıra, bu çalışmada elde edilen çıktılarda nirengi noktalarının hatalı saptandığı durumlar da meydana gelmiştir. Nirengi noktalarının hatalı saptandığı örnek şekiller Şekil 3.6, 3.7 ve 3.8'de sunulmuştur.

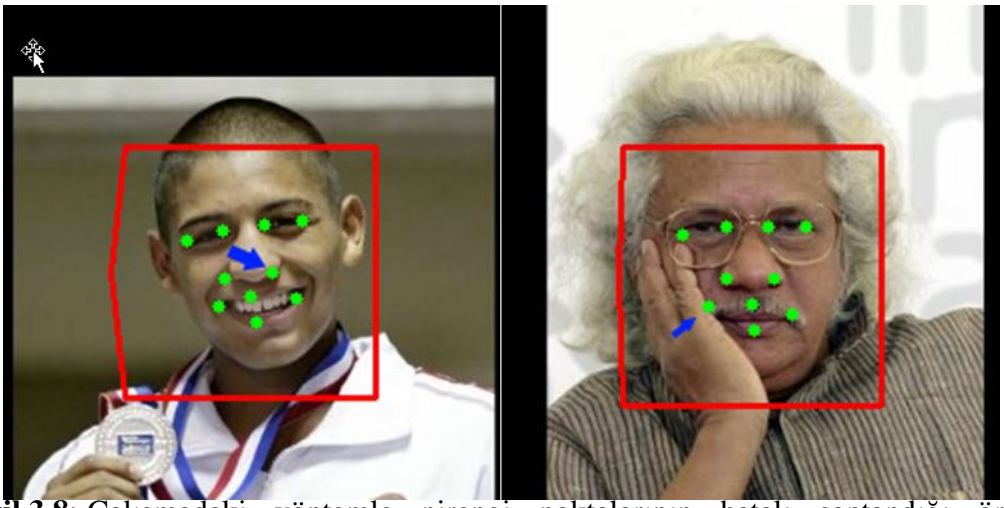
Bu çalışmada kullanılan yöntem başarımın yanı sıra hız testine de tabi tutulmuştur. Çizelge 3.9'da gösterilen başarımı sağlayan YZM sisteminde nirengi noktalarının bulunma süresi yüzün konumu dışarıdan girdi olarak verildiğinde ortalama 56 ms, yüzün saptanması işlemi de sisteme bırakıldığında ise ortalama 72 ms olarak ölçülmüştür.

Şekil 3.5: Çalışmadaki yöntemle nirengi noktalarının isabetli saptandığı örnek çıktılar.



Şekil 3.6: Çalışmadaki yöntemle nirengi noktalarının hatalı saptandığı örnek çıktılar-1.

Şekil 3.7: Çalışmadaki yöntemle nirengi noktalarının hatalı saptandığı örnek çıktılar-2.



Şekil 3.8: Çalışmadaki yöntemle nirengi noktalarının hatalı saptandığı örnek çıktılar-3.

4. SONUÇ VE ÖNERİLER

Bu tez çalışmasında, yüz tanımada etkili ve başarılı bir yöntem olduğu görülmüş olan YZM'nin yüzdeki nirengi noktalarının saptanmasındaki kullanılabilirliği ve başarısı araştırılmıştır.

Koşullu Rastgele Regresyon Ormanları ile Yerel Zernike Momentleri'ni birlikte kullanılan yöntemin başarımı LFW veri setinde çeşitli hata toleransları belirlenerek değerlendirilmiştir. Çalışmada kullanılan bu yöntem ile %15 tolerans için yüzdeki nirengi noktaları ortalama %89.33 başarımla saptanmaktadır. Artan tolerans değerlerinde daha yüksek başarımlar elde edilmektedir. Aynı yöntemin %20 tolerans için yüzdeki nirengi noktalarını saptama başarımı ortalama %94.48 olmaktadır.

Bu çalışmada anlatılan yöntem başarımın yanı sıra hız testine de tabi tutulmuştur. Bu yöntemle nirengi noktalarının bulunma süresi, yüzün konumu dışarıdan girdi olarak verildiğinde ortalama 56 ms, yüzün saptanması işlemi de sisteme bırakıldığında ise ortalama 72 ms olarak ölçülmüştür.

Bu tez çalışmasında elde edilen sonuçlar incelendiğinde, farklı pozları, değişik ışıklandırma koşulları, çözünürlük ve imge kalitelerini, faklı yüz ifadelerini ve cinsiyetleri ve yüz bölümlerinin kapanması gibi saptama işlemini zorlaştırabilecek etkenleri barındırdan imgelerin bulunduğu, zorlayıcı bir veri seti olan LFW ile elde edilen sonuçların umut verici olduğu gözükmektedir.

KAYNAKLAR

[1] **Dantone, M.**, **Gall, J.**, **Fanelli, G. ve Gool, L.V.** (2012). Real-time Facial Feature Detection using Conditional Regression Forests, *CVPR*.

[2] **Başaran, E.** (2013). Nicemlenmiş yerel zernike momentlerle trafik işaretlerinin sınıflandırılması / Traffic sign classification with quantized local zernike moments.

[3] **Sariyanidi, E.**, **Dagli, V.**, **Tek, S.**, **Tunç, B. ve Gökmen, M.** (2012). Local Zernike Moments: A new representation for face recognition, *ICIP*, s.585–588.

[4] **Everingham, M.**, **Sivic, J. ve Zisserman, A.** (2006). Hello! My name is... Buffy – Automatic Naming of Characters in TV Video, *British Machine Vision Conference*, s.899–908.

[5] **Valstar, M.F.**, **Martínez, B.**, **Binefa, X. ve Pantic, M.** (2010). Facial point detection using boosted regression and graph models, *Computer Vision and Pattern Recognition*, s.2729–2736.

[6] **Amberg, B. ve Vetter, T.** (2011). Optimal landmark detection using shape models and branch and bound, *International Conference on Computer Vision*, s.455–462.

[7] **Belhumeur, P.N.**, **Jacobs, D.W.**, **Kriegman, D.J. ve Kumar, N.** (2011). Localizing parts of faces using a consensus of exemplars, *Computer Vision and Pattern Recognition*, s.545–552.

[8] **Statistics, L.B. ve Breiman, L.** (2001). Random Forests, *Machine Learning*, s.5–32.

[9] **Chong, C.W.**, **Raveendran, P. ve Mukundan, R.** (2003). A comparative analysis of algorithms for fast computation of Zernike moments, *Pattern Recognition*, ***36***(3), 731 – 742.

[10] **Huang, G.B.**, **Ramesh, M.**, **Berg, T. ve Learned-Miller, E.** (October, 2007). Labeled Faces in the Wild: A Database for Studying Face Recognition in Unconstrained Environments., **Teknik Rapor**, University of Massachusetts, Amherst, Technical Report 07-49.

[11] **Fanelli, G.**, **Weise, T.**, **Gall, J. ve Gool, L.V.** (2011). Real Time Head Pose Estimation from Consumer Depth Cameras.

[12] **Huang, C.**, **Ding, X. ve Fang, C.** (2010). Head Pose Estimation Based on Random Forests for Multiclass Classification, *International Conference on Pattern Recognition*, s.934–937.

[13] **Viola, P.A. ve Jones, M.J.** (2004). Robust Real-Time Face Detection, *International Journal of Computer Vision*, ***57***, 137–154.

ÖZGEÇMİŞ

Ad Soyad: Gencer VURAL

Doğum Yeri ve Tarihi: Konak - 26.09.1989

Adres: Yalıköy Mahallesi Hüsrebey Sokak No: 4 Beykoz/İSTANBUL

E-Posta: vuralge@itu.edu.tr

Lisans: İstanbul Teknik Üniversitesi - Bilgisayar Mühendisliği

Y. Lisans: İstanbul Teknik Üniversitesi - Bilgisayar Mühendisliği

Mesleki Deneyim ve Ödüller:

Yayın ve Patent Listesi:

TEZDEN TÜRETİLEN YAYINLAR/SUNUMLAR

- **Vural, G.**, Gökmen, M., Yüzdeki Nirengi Noktalarının Koşullu Regresyon Ormanları ile Saptanması, *23. Sinyal İşleme ve Uygulamaları Kurultayı(SIU), 2015.*