

Pozlandırılmış Görüntüleri Sınıflandırarak Etiketleme

Tuğrul Güçlü

**YÜKSEK LİSANS TEZİ**

Elektrik-Elektronik Mühendisliği Anabilim Dalı

Şubat 2010

Labeling Posed Images By Classification

Tuğrul Güçlü

**MASTER OF SCIENCE THESIS**

Department of Electrical & Electronics Engineering

February 2010

Pozlandırılmış Görüntüleri Sınıflandırarak Etiketleme

Tuğrul Güçlü

Eskişehir Osmangazi Üniversitesi  
Fen Bilimleri Enstitüsü  
Lisansüstü Yönetmeliği Uyarınca  
Elektrik Elektronik Mühendisliği Anabilim Dalı  
Telekomünikasyon-Sinyal İşleme Bilim Dalında  
YÜKSEK LİSANS TEZİ  
Olarak Hazırlanmıştır

Danışman: Yrd. Doç. Dr. Nihat Adar

Şubat 2010

## ONAY

Elektrik-Elektronik Mühendisliği Anabilim Dalı Yüksek Lisans öğrencisi Tuğrul Güçlü'nün YÜKSEK LİSANS tezi olarak hazırladığı "Pozlanmış Görüntüleri Sınıflandırarak Etiketleme" başlıklı bu çalışma, jürimizce lisansüstü yönetmeliğin ilgili maddeleri uyarınca değerlendirilerek kabul edilmiştir.

**Danışman** : Yrd. Doç. Dr. Nihat Adar

**İkinci Danışman** : -

### **Yüksek Lisans Tez Savunma Jürisi:**

**Üye** : Yrd. Doç. Dr. Nihat Adar

**Üye** : Prof.Dr. İdris Dağ

**Üye** : Yrd. Doç. Dr. Erol Seke

**Üye** : Yrd. Doç. Dr. Kemal Özkan

**Üye** : Yrd. Doç. Dr. Selçuk Canbek

Fen Bilimleri Enstitüsü Yönetim Kurulu'nun ..... tarih ve ..... sayılı kararıyla onaylanmıştır.

Prof. Dr. Nimetullah BURNAK

Enstitü Müdürü

## ÖZET

Bu çalışmada insan vücut parçalarının açısız özelliklerine göre pozlandırıcıdan geçirilerek poz dizilimlerinin elde edilmesi ve bu dizilimler vasıtası ile el sallama, yürüme ve koşma hareketlerinin tanınmasında kullanılan örüntü algılama süreci sunulmaktadır. Çalışmanın ilk bölümlerinde resimsel yapılar modelinden yola çıkarak 10 parçadan oluşan bir insan vücudu modeli tasarlanmıştır. Bu model daha sonra iki farklı veritabanına ait görüntülerin elle işaretlenerek açısız özelliklerinin çıkarımı sürecinde kullanılmıştır. Bu sürecin sonunda elde edilen açısız özelliklerden bir açısız öz nitelik kütüphanesi oluşturulmuştur. Elle işaretleme yapılması vücut parçalarını otomatik bulabilecek görüntü işleme yöntemleri içinde bir soyutlama katmanı oluşturmuştur. Ayrıca farklı görüntü veritabanları ile çalışmakta açısız öz nitelik vektörlerinin farklı ölçek ve yönelimdeki görüntülerden etkilenmesini sınanabilir kılmıştır. Açısız öz niteliklerin elde edilmesinden sonra K-means algoritmasının kullanımı ile öz nitelik vektörlerinden poz kelimeleri oluşturulmuştur. Hareketin algılanması kısmında ise Yapay Sinir Ağları ve Destek Vektör Makineleri kullanılmıştır. Başarımlar vektör boyutlarını eşitlemede kullanılan tekrarlı ve ortak poz ekleme yöntemleri ile hesaplanmıştır. Buna ek olarak poz dizilimlerindeki sıralamanın hareket algılamaya olan etkisini görebilmek için poz dizilimleri karıştırılarak ve ters çevrilerek testler yapılmıştır. Böylece poz dizilimlerinin sıralamasının hareket algılamada ki önemi vurgulanmıştır.

Anahtar Kelimeler: İnsan hareketlerinin tanınması, poz tanıma, poz dizilimi, BOW

## SUMMARY

In this work, the process of recognition human actions such as handwaving, walking, running by using angular features of body parts is presented. In the first parts of the study a human body part model of 10 parts has been designed by using pictorial structures. This model is again used in the manual labeling process of images belonging to two different image database. At the end of this process, a library of angular features has been formed. Manual labeling also creates an abstraction layer for image processing methods which can automatically find body parts. Also working with different image databases makes it possible to test usage of angular features against images with different scales and orientations. After having angular features, k-means clustering is used to gather pose words from angular features of body parts. In order to recognize actions from pose words SVM and Neural Network classifiers are used. Results are then compared whether common pose or repetitive pose appending method is used to equalize input vector lengths. In addition to these tests, other tests are applied to understand the importance of pose arrangement order in action recognition. These test involve shuffling and reversing the pose arrangement order. Results also imply the importance of pose arrangement order in action recognition.

Keywords: Recognizing human actions, pose estimation , pose sentences, BOW

**TEŐEKKÖR**

Tez alıőmam boyunca beni yönlendiren, bilgisini, deęerli görüőlerini ve deneyimlerini bana aktarıp her türlü olanaęı bana saęlayan danıőmanım Sayın Yrd. Do.Dr. Nihat ADAR'a ok teőekkör ederim.

## İÇİNDEKİLER

<b>ÖZET.....</b>	<b>V</b>
<b>SUMMARY .....</b>	<b>VI</b>
<b>TEŞEKKÜR.....</b>	<b>vii</b>
<b>İÇİNDEKİLER .....</b>	<b>viii</b>
<b>ŞEKİLLER DİZİNİ .....</b>	<b>X</b>
<b>ÇİZELGELER DİZİNİ .....</b>	<b>xii</b>
<b>SİMGELER VE KISALTMALAR DİZİNİ .....</b>	<b>xiv</b>
<b>1 GİRİŞ .....</b>	<b>1</b>
1.1 Hareket Analizi Tabanlı Görü Uygulamaları.....	3
1.2 İnsan Hareketinin Analizi .....	7
1.3 İnsan vücut yapısının analizi.....	8
1.4 Poz Bilgisinin Çıkarımı.....	11
1.4.1 Vücut planları.....	12
1.4.2 Resimsel yapılar.....	13
1.5 Hareketin Algılanması .....	18
1.5.1 BoW modeli .....	19
1.5.2 BoW modelinin görüntü algılamada uygulanışı .....	20
1.6 Önerilen Yöntem.....	22
<b>2 KULLANILACAK KÜTÜPHANELERİNİN HAZIRLANMASI.....</b>	<b>25</b>
2.1 Kullanılan Veri Tabanları .....	25
2.1.1 ESOGU-IP LAB veritabanı .....	25
2.1.2 NADA veritabanı .....	26
2.1.3 Görüntü veritabanlarından derlenerek kütüphane oluşturulması.....	26
2.1.4 Görüntü veritabanlarının kullanım dağılımı .....	27
2.1.5 Açısal özelliklerden poz kelimeleri kütüphanesi oluşturulması .....	31
<b>3 ÖZ NİTELİKLERİN ÇIKARILMASI.....</b>	<b>33</b>
3.1.1 Kullanılan insan vücut modeli .....	33
3.1.2 Elle işaretleme.....	35
3.1.3 Açılarının hesaplanması.....	40



3.1.4	K-means ile poz kümelerinin oluşturulması .....	42
3.1.4.1	K-means algoritması adımları .....	43
3.1.4.2	Pozlandırıcı için küme sayısının çıkarılışı .....	44
3.1.4.3	Test gruplarının pozlandırıcıdan geçirilmesi .....	47
3.1.4.4	Poz dizilimlerinin oluşturulması .....	48
<b>4</b>	<b>DENEYSEL ÇALIŞMALAR.....</b>	<b>50</b>
4.1	Destek Vektör Makineleri .....	50
4.2	Yapay Sinir Ağları .....	53
4.3	Örüntü Algılamada Performans Ölçüm Kriterleri .....	56
4.4	Eğitim ve Test Verileri .....	59
4.5	Poz Dizilimlerinin Eğitimi ve Testi .....	59
4.6	Uygulama ve Test Detayları .....	60
4.6.1	1:3 yapılan testler .....	61
4.6.2	1:1 yapılan testler .....	62
4.6.3	3:3 yapılan testler .....	63
4.6.4	3:3 karışık ve ters çevrilmiş poz dizilimi testleri .....	64
4.6.5	Test başarımları .....	64
4.6.5.1	1:1 testleri başarımları .....	64
4.6.5.2	1:3 testleri başarımları .....	72
4.6.5.3	3:3 Testleri başarımları .....	75
4.6.5.4	3:3 Karıştırılmış poz dizilimi testi başarımları .....	76
4.6.5.5	3:3 Ters çevrilmiş poz dizilimi testi başarımları .....	77
<b>5</b>	<b>SONUÇ.....</b>	<b>79</b>
<b>6</b>	<b>KAYNAKLAR DİZİNİ .....</b>	<b>81</b>

## ŞEKİLLER DİZİNİ

Şekil 1-1: Piksel gösterimi .....	1
Şekil 1-2: Masa üstü kamera ile işaret dili tanıma .....	6
Şekil 1-3: Şapkaya takılı kamera ile işaret dili tanıma .....	6
Şekil 1-4: İnsan hareketinin genel analizi .....	7
Şekil 1-5: İnsan vücudu sembolik gösterimleri .....	8
Şekil 1-6: Model yapısına göre insan vücut parçaları .....	9
Şekil 1-7: At için hazırlanmış vücut planı .....	12
Şekil 1-8: İnsan yüzü için geliştirilen resimsel yapılar modeli .....	14
Şekil 1-9: 14 eklemlili ve 15 parçalı insan modeli .....	14
Şekil 1-10: Elle işaretleme yapılan eğitim imgesi .....	16
Şekil 1-11: Parça algılayıcılarının test imgelerinde buldukları parçalar .....	17
Şekil 1-12: BoW modeli ile oluşturulan sözlük .....	20
Şekil 1-13: BoW yöntemi ile oluşturulan kelime histogramları .....	20
Şekil 1-14: Çalışmaya ait genel akış diyagramı .....	23
Şekil 1-15: Sisteme ayrıntılı genel bakış .....	24
Şekil 2-1: Görüntülerden oluşturulan kütüphane .....	27
Şekil 2-2: ESOGU-IP-LAB veritabanı örnek resimler .....	29
Şekil 2-3: NADA veritabanı örnek resimler .....	30
Şekil 2-4: Veritabanlarının ortak kullanımı .....	31
Şekil 3-1: Elle işaretlenmiş insan vücudu modeli (a) El sallama (b) Koşma .....	35
Şekil 3-2: Elle işaretlenmiş resim .....	36
Şekil 3-3: Elle işaretleme uygulaması akış diyagramı .....	37
Şekil 3-4: Eksene göre simetrik parçalar .....	40
Şekil 3-5: İşaretlenmiş ve açıları hesaplanmış imgeler .....	41
Şekil 3-6: K-means ilk atanan kümeye göre farklı sonuçlar verir .....	43

Şekil 3-7: Kümeleme işleminin sembolik gösterimi.....	45
Şekil 3-8: Kümeleme sonrası elle düzeltme.....	46
Şekil 3-9: Test gruplarının pozlanmasına ait akış diyagramı.....	48
Şekil 4-1: Ayırıcı düzlem ve destek vektörleri .....	51
Şekil 4-2: Birden fazla aşırı düzlem olma durumu .....	52
Şekil 4-3: Destek Vektörleri. ....	53
Şekil 4-4: Sinir hücresi gösterimi .....	54
Şekil 4-5: Yapay sinir hücresi gösterimi.....	55
Şekil 4-6: Ortak poz ekleme .....	59
Şekil 4-7: Tekrarlı poz ekleme.....	60
Şekil 4-8: 1:3 Test genel yapısı.....	62
Şekil 4-9: 1:1 Test genel yapısı.....	63
Şekil 4-10: Çoklu sınıflandırıcı genel yapısı .....	63

## ÇİZELGELER DİZİNİ

Çizelge 1: Bilgisayarlı görü işleme uygulamaları ve alanları (Gavrila., 1999) .....	4
Çizelge 2: Modele göre parçaların dağılımı.....	15
Çizelge 3: Veri seti-hareket türüne göre görüntü sayısı dağılımı .....	28
Çizelge 4: Hareket türlerine göre veri seti-görüntü sayısı dağılımı .....	28
Çizelge 5: Vücut modeli-parça dağılımı .....	33
Çizelge 6: Elle işaretlemeye kullanılan veri yapısı.....	39
Çizelge 7: Kümeleme işleminde kullanılan veri yapısı .....	47
Çizelge 8: Poz dizilimi kütüphanesinin veri gruplarına göre dağılımı .....	49
Çizelge 9: Hata matrisi gösterimi .....	57
Çizelge 10: 1:1 testi T2 veri seti için NN ile ortak poz ekleme başarımları .....	65
Çizelge 11: 1:1 testi T2+T3 veri seti için NN ile ortak poz ekleme başarımları .....	65
Çizelge 12: 1:1 testi T3 veri seti için NN ile ortak poz ekleme başarımları .....	66
Çizelge 13: 1:1 testi T2 veri seti için SVM ile ortak poz ekleme başarımları .....	66
Çizelge 14: 1:1 testi T2+T3 veri seti için SVM ile ortak poz ekleme başarımları .....	67
Çizelge 15: 1:1 testi T3 veri seti için SVM ile ortak poz ekleme başarımları .....	67
Çizelge 16: 1:1 testi T2 veri seti için NN ile tekrarlı poz ekleme başarımları.....	68
Çizelge 17: 1:1 testi T2+T3 veri seti için NN ile tekrarlı poz ekleme başarımları .....	69
Çizelge 18: 1:1 testi T3 veri seti için NN ile tekrarlı poz ekleme başarımları.....	69
Çizelge 19: 1:1 testi T2 veri için SVM ile tekrarlı poz ekleme başarımları .....	70
Çizelge 20: 1:1 testi T2+T3 veri seti için SVM ile tekrarlı poz ekleme başarımları .....	70
Çizelge 21: 1:1 testi T3 veri seti için SVM ile tekrarlı poz ekleme başarımları.....	71
Çizelge 22: 1:3 testi T2 veri seti ortak poz ekleme başarımları.....	72
Çizelge 23: 1:3 testi T2+T3 veri seti ortak poz ekleme başarımları.....	72
Çizelge 24: 1:3 testi T3 veri seti ortak poz ekleme başarımları.....	73
Çizelge 25: 1:3 testi T2 veri seti tekrarlı poz ekleme başarımları .....	73

Çizelge 26: 1:3 testi T2+T3 veri seti tekrarlı poz ekleme başarımları.....	74
Çizelge 27: 1:3 testi T3 veri seti tekrarlı poz ekleme başarımları .....	74
Çizelge 28: 3:3 testi ortak poz ekleme ile başarımları .....	75
Çizelge 29: 3:3 testi tekrarlı poz ekleme ile başarımları.....	76
Çizelge 30: 3:3 testi ortak poz eklemeli karıştırılmış poz dizilimi başarımları .....	77
Çizelge 31: 3:3 testi ortak poz eklemeli ters çevrilmiş poz dizilimi başarımları.....	78

**SİMGELER VE KISALTMALAR DİZİNİ****Simgeler****Açıklamalar****Kısaltmalar****Açıklamalar**

Pixel

Picture Element

BoW

Bag Of Words

HOG

Histogram of Oriented Graphs

SVM

Support Vector Machines

RVM

Relevance Vector Machines

N-N

Neural Networks

NLP

Natural Language Processing

CGI

Computer-Generated Imagery

AVI

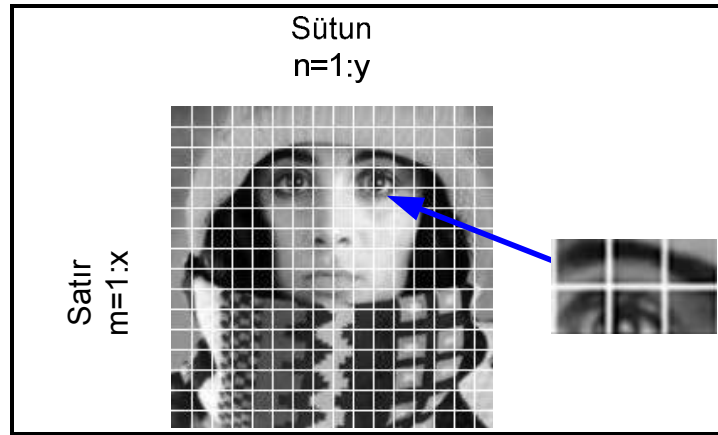
Audio Video Interleave

## 1 GİRİŞ

Günümüzde modern teknolojinin de gelişimi ile çevremizde ya da uzakta olup biten olaylara ait görüntüler sayısal olarak kayıt edilebilmekte ve üzerlerinde işlem yapılabilmektedir. Böylece görüntüdeki hareketli nesneyi diğerlerinden ayırmak, hareketin türünü anlamak ve yapılan hareketten anlam çıkarmak mümkün hale gelmiştir. İnternetin ve bilgisayar destekli görüntü yakalama sistemlerinin gelişimi bilgisayar tabanlı otomatik görü işleme, analiz ve hareket algılama sistemlerini gün geçtikçe daha da önemli hale getirmeye başlamıştır.

### *Sayısal Görüntünün Temelleri*

Sayısal görüntü  $\alpha(x,y)$  şeklinde ifade edilen ve 2 boyutlu sürekli bir sinyalin örnekleme yapılarak  $\alpha[m,n]$  şeklinde ve 2 boyutlu ayrık uzayda bulunan görüntüye çevrilmiş halidir (Young, et al., 1995). Sayısal görüntüler piksel adını verdiğimiz noktalardan oluşur. Pikseller üzerlerinde görüntüye ait renk bilgisini taşırlar.



Şekil 1-1: Piksel gösterimi

Bir sayısal görüntü  $m \times n$  boyutlu bir matris ile ifade edilir (bkz: **Şekil 1-1**). Matrisin satır ve sütunlarında bulunan değerler ise sayısal görüntü üzerindeki piksellerin renk değerlerini verir.

Pikseller görüntünün en küçük parçalarını oluşturup aynı zamanda renk bilgisi taşır ve bir araya gelerek görüntüdeki renk ve şekilleri oluştururlar. Renkler ise içerdikleri kanal sayılarına göre ayrılırlar. Örneğin bir RGB görüntü 3 kanala sahip iken bir gri seviye görüntü de bu sayı 1'dir. Bu nedenle gri seviyeli görüntüler siyah-beyaz algılanır. Bilgisayar ortamında sayısal görüntüler matrisler şeklinde işlem görürler. Matris halindeki sayısal görüntüler ise çeşitli işlemlere tabi tutularak istenen bilgiye ulaşılır. Girdi sinyalinin görüntü olduğu bu tarz sinyal işleme yöntemleri topluluğuna ise sayısal *görüntü işleme* denir.

İnsan görme sistemi için hareket algılama işlemi doğası gereği otomatik olarak yapılmaktadır. Ardışık 2 resim çerçeve arasında hareketli bir objenin hareketini bilgisayar ortamında algılamak ise bazı zorluklar içerir. Hareket yönündeki değişim, dönme ve eğilme etkileri, başka objelerin hareketin ilgi noktasını kapatması, arka plandaki karmaşıklık karşılaşılan güçlüklerden sadece birkaçıdır.

Bu yüksek lisans tezinin amacı insan hareketinin poz dizilimleri kullanılarak otomatik algılanmasıdır. Hecelerin kelimeleri, kelimelerin de cümleleri oluşturması gibi insan hareketleri de pozlar ve bu pozların dizilimlerinden oluşmaktadır.

*Tezin genel yapısı şu şekildedir:* Birinci bölümde hareket algılama, görüntü işleme konuları üzerine genel bilgi verilecektir.



2. bölümde çalışmada kullanmış olduğumuz hareketlere ait açısız özellik ve poz dizilimleri kütüphanelerinin hazırlanması üzerinde durulacaktır.

Bölüm 3’de çalışmanın 1. fazı detaylandırılmıştır. Bu bölümde imgelerin işaretlenmesi, açısız öznitelik vektörlerinin elde edilmesi ve kümeleme işleminde kullanılan k-means algoritması üzerinde durulacak ve bu işlemler için geliştirilen yazılım akış diyagramları üzerinden incelenecektir.

Bölüm 4’de ise çalışmanın 2. fazı üzerinde durulacaktır. Bu bölümde poz dizilimlerinin elde edilmesi ve bu dizilimlerin eğitim ve test kümelerindeki kullanımı açıklanacaktır. Daha sonra ise SVM, Yapay Sinir Ağları ve k-means yöntemlerinin kullanımları anlatılacaktır.

Test sonuçlarında kullanılan sınıflandırıcılar, başarımlarının karşılaştırılmasında kullanılan metrikler ve bu oranların karşılaştırılması Bölüm 4’de açıklanacaktır.

Sonuçların incelenmesi ise Bölüm 5’te anlatılacaktır.

## **1.1 Hareket Analizi Tabanlı Görü Uygulamaları**

Bilgisayar ortamındaki görüntü işleme ve analiz sistemlerinin ortak amacı görüntünün elde edildiği ortama dair geçmişe dönük ya da gerçek zamanlı bilgi elde etmektir. Görüntü işleme ve görüntü algılama çalışmalarının gelişmesindeki temel isteklendirme ise askeri uygulamalardan sosyal etkileşimli ağlara kadar birçok alan da uygulamalarının olmasıdır.

Çizelge 1: Bilgisayarlı görü işleme uygulamaları ve alanları (Gavrila., 1999)

Genel Çalışma Alanı	Özel Uygulama Alan
Sanal Gerçeklik	<ul style="list-style-type: none"> <li>— Etkileşimli sanal dünyalar</li> <li>— Oyunlar</li> <li>— Sanal stüdyolar</li> <li>— Karakter animasyonları</li> <li>— Uzaktan konferans uygulamaları</li> </ul>
Akıllı Güvenlik Sistemleri	<ul style="list-style-type: none"> <li>— Yetki kontrolü</li> <li>— Park alanları</li> <li>— Market, depo alanları</li> <li>— ATM cihazları</li> <li>Trafik</li> </ul>
Gelişmiş kullanıcı ara yüzleri	<ul style="list-style-type: none"> <li>— Sosyal ara yüzler</li> <li>— İşaret dili çevirileri</li> <li>— İfade kullanımı kontrol</li> <li>— Gürültülü ortamlarda haberleşme (havaalanları, fabrikalar)</li> </ul>

**Sanal Gerçeklik:** İnternet üzerinde özellikle çeşitli sanal sohbet odası uygulamalarında karşımıza çıkan bu özellik ile kişiler duygu ve düşüncelerini sadece metin tabanlı değil aynı zamanda 2B karakterlerin hareketleri olarak da anlatma olanağına sahiptirler. Bunu mümkün kılmak için gerçek dünyada ki 3B poz bilgisinin 2B ortama taşınması gerekmektedir.

Film ve animasyon endüstrisinde de görüntü tanıma uygulamaları sıkça karşımıza çıkmaktadır. Özellikle CGI (*Computer-Generated Imagery*) tabanlı filmlerde yaratılan karakterler gerçek karakterlerin hareketlerini çeşitli alıcı uygulamaları yardımı ile son derece gerçekçi olarak yansıtabilmektedir.

**Akıllı Güvenlik Sistemleri:** Buradaki *akıllı* kelimesi uygulamanın hareketleri ne kadar iyi algıladığının yanı sıra hatalı alarmlara karşı da ne duyarlı olduğunu göstermek için kullanılır. Sadece hareketin varlığı ile alarm üretecek bir sistem rüzgârın esmesi, çevrede hayvanların olması durumlarında da hatalı alarm üretebilir ki bu istenmeyen bir durumdur.

Günümüzde güvenlik sistemleri sadece insan varlığını saptamakla kalmayıp ek olarak yüz tanınması, retina taraması, parmak izi tanınması gibi uygulamaları da çok yaygın olarak kullanırlar.

Daha farklı güvenlik uygulamalarında da ise insanın varlığından çok ne yaptığı araştırılmıştır. Aggarwal'ın parmaklık *tırmanan insanları* algılamaya yönelik çalışması buna bir örnektir (Aggarwal and Yu., 2006).

**Gelişmiş Kullanıcı Ara yüzleri:** Bilgisayar uygulamalarının ve internetin de hızla gelişmesi ile birlikte gündeme gelen bir çalışma alanıdır. Sosyal iletişim ara yüzleri de çalışma konularından birisidir. Sosyal iletişim ara yüzleri bilgisayarların insana benzer hareketler sergileyerek çeşitli konularda kullanıcılarla etkileşimi üzerine çalışır. Özellikle robot-insan etkileşimi üzerine uygulamaları vardır. İnsanlarla beraber yürüyebilen robotlar, evde bir yabancıнын varlığını algılayabilen robotlar bu çalışmalara örnek olarak verilebilir.

Aynı zamanda gerçek zamanlı işaret dili algılama uygulamaları da insan-bilgisayar ara yüzüne örnek teşkil eder. Pentland çalışmasında Amerikan işaret dilini sürekli cümleler seviyesinde algılamak için masa üstü kamera ve şapkaya monte edilmiş kamera kullanmış ve gerçek zamanlı olarak karşıdaki kişinin kol hareketlerini takip ederek sırası ile %92 ve %98 başarımla elde etmiştir (Pentland, et al., 1998).



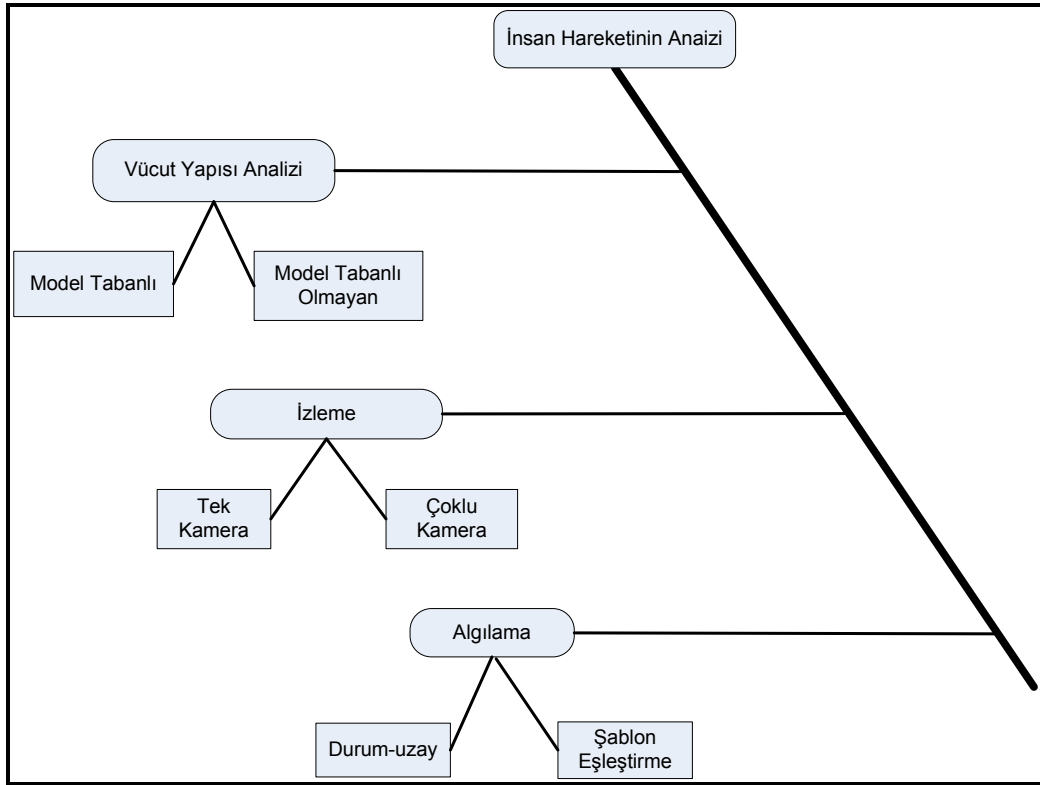
Şekil 1-2: Masa üstü kamera ile işaret dili tanırna



Şekil 1-3: Şapkaya takılı kamera ile işaret dili tanırna

## 1.2 İnsan Hareketinin Analizi

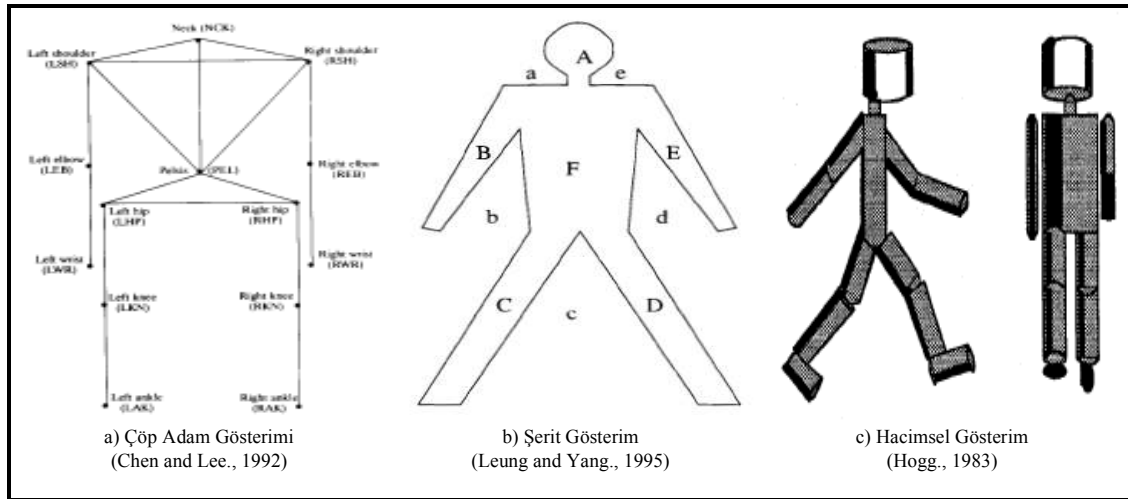
İnsan hareketinin analizi bilgisayar tabanlı görü uygulamalarının en çok üzerinde durdukları çalışma alanlarından birisidir. **Şekil 1-4**'de de belirtildiği gibi insan hareketinin analizini 3 ana başlıkta incelenir (Aggarwal and Cai., 1997).



Şekil 1-4: İnsan hareketinin genel analizi

### 1.3 İnsan vücut yapısının analizi

Bir görüntüde dizisinde insan görüntüsü aranırken öncelikle vücudun tamamı ya da vücut parçalarını bulmak gerekir. Bu anlamda vücut parçaları bulunurken çeşitli insan vücudu gösterimlerinden faydalanılır (bkz: **Şekil 1-5**)



Şekil 1-5: İnsan vücudu sembolik gösterimleri

Chen ve Lee yaptıkları çalışmada insan vücudunu kafa, gövde, kalça, kol ve bacaklar gibi bükülemeyen ve bazı özellik noktaları ile karakterize edilen alt parçalardan oluştuğunu varsayımlardır (Chen and Lee., 1992). Bunu da 14 adet eklem, 17 adet bölümden oluşacak şekilde modellemişlerdir (bkz: **Şekil 1-5** (a)).

Leung ve Yang ise ellerin kolların alt kısmının uzantısı olduğu, alt kolların ise üst kolların uzantısı olduğu fikrini öne sürmüştür (Leung and Yand, 1995). Bu şekilde vücut parçaları arasında hiyerarşik bir ilişkiyi modellemiştir. Çalışmalarında insan vücut parçalarını 2B olarak şeritler halinde göstermiştir (bkz:Şekil 1-5 (b)). Verilen örnekte

A,B,C,D,E,F => Vücuda ait parçaların şeritleri

a,b,c,d,e,f => Parçaların arasında kalan arka plana ait şeritlerdir.

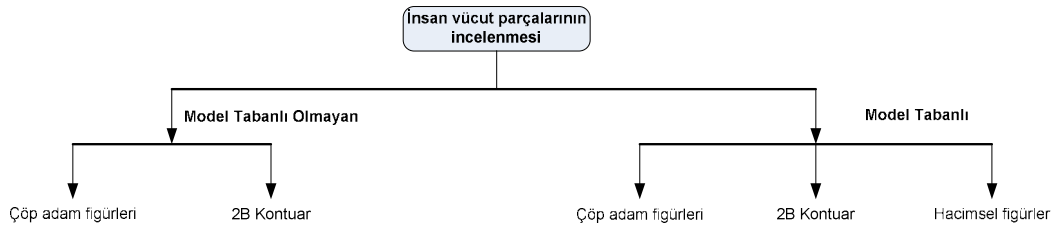
Hogg ise çalışmasında insan vücudunu hacimsel şekilde modellemiş ve eliptik 14 parçaya ayırmıştır (bkz:Şekil 1-5 (c)) (Hogg., 1983).

Bu çalışmada önerdiğimiz yöntemde de benzeri bir gösterim kullanılmıştır. Kullanılan gösterimde insan vücudu 10 parçadan oluşacak şekilde şerit parçaları halinde kabul edilmiştir.

Hareket analizinde 2 yaklaşım kullanılmaktadır (bkz: Şekil 1-6).

Bunlar;

- Model tabanlı yaklaşımlar
- Model tabanlı olmayan yaklaşımlar



Şekil 1-6: Model yapısına göre insan vücut parçaları

### Model tabanlı yaklaşım

Model tabanlı yaklaşımlarda bir nesnenin hareketin analizi yapılırken o nesneye ait daha önceden bilinene ya da tasarlanmış bir şekil yapısı kullanılır. Model tabanlı yaklaşımlar imge dizisinde model verisini eşleştirmeye çalışır. Bu başarılı olursa öz niteliklerin uyumlandırılması ise otomatik olarak yapılır.

### Model tabanlı olmayan yaklaşım

Bu yaklaşımda ise hareketi analiz edilmek istenen nesnenin şekli ile ilgili bir ön bilgi ya da model bilinmemektedir. Bu gibi durumlarda ardı ardına gelen imgelerde ki öz nitelik uygunluğu aranan öz niteliğe bağlıdır.

Örnek vermek gerekirse;

- Parçalarının mutlak ya da göreceli pozisyon bilgileri
- Parçaların hareket süratleri
- Parçalara ait şekil bilgileri
- Parçaların dokuları
- Parçalara ait renk bilgileri

Fakat her 2 yaklaşımda hareketin analizi sorusunu aynı çerçevede ele alır.

1. Öz niteliklerin çıkarımı
2. Öz niteliklerin uyumunun sorgulanması
3. Üst seviyede öz nitelik bilgilerini işlenmesi

Model tabanlı ve model tabanlı olmayan yaklaşım arasındaki en belirgin fark ardışık imgeler arasındaki öz nitelikler arasında bir uyum sağlanmasıdır.



Her 2 yaklaşımı ele aldığımızda ortaya şu sonuç çıkıyor. Model tabanlı yaklaşımlarda imge dizilerinde model verisini aramak algoritmik açıdan bakıldığında karmaşık bir iştir. Çoğu zaman kullanılan model bilgisini imgedeki piksel grupları içerisinde aramak yeterli olmaz. Vücut parçalarının değişik pozisyonlarını da düşünerek bu aramayı farklı döndürme açıları ve farklı büyüklük çarpanları ile denemek gerekir. Model verisi imge içinde bulunduğu zaman ise diğer imgelerde bulmak için harcanan çaba azalır. Hareket vektörleri ya da optik akış yöntemleri vasıtası ile hareketin açisal hızı incelenerek arama yapılan bölge azaltılır ve sonuca daha hızlı ulaşmak mümkün olur. Model tabanlı olmayan yaklaşımlarda ise imgelerde bulunan bir öz nitelik aranır. Bu her zaman şekil bilgisi olmak zorunda değildir. Kimi zaman renk kimi zaman ise pozisyon bilgisi olabilir. Yapılan çalışmada model tabanlı olmayan bir yaklaşım üzerinde durulmuştur. Öz nitelik bilgisi olarak ise vücut parçalarının birbirleri ile yaptıkları açılar kullanılmıştır. Bu sayede farklı vücut büyüklüklerinin hareket algılama sürecine edeceği olumsuz etkilerin önüne geçilmiştir.

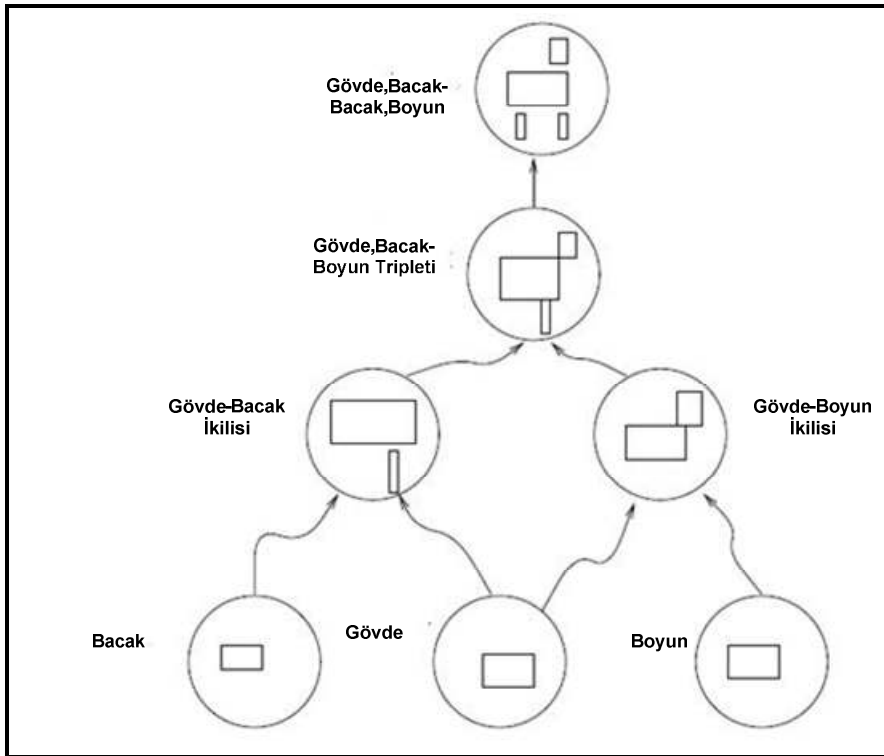
#### **1.4 Poz Bilgisinin Çıkarımı**

İmgelerdeki insanları bulmak çözümü güç bir problemdir. Bunun başlıca nedeni ise dışarıdan bakıldığında insan vücudunun birçok görünüş şekli alabilmesinden kaynaklanır. Vücut parçalarının farklı ölçeklerde olması, bakış açıları, aydınlatmadaki farklılıkların hepsi görüntüdeki objeyi bulmayı güçleştirir. Bu nedenle insanları günlük sıradan aktivitelerini yaparken algılayabilecek yöntemlere ihtiyaç doğmuştur (Ronfard, et all., 2002). Yapılan birçok çalışmada sıkça yapılan görüntünün kapanmaması, arka planın homojen olması gibi çeşitli kabullenmeler bu tip yöntemlerde yapılamaz.

Bu sorunun çözmek ve görüntüdeki insanı bulmak için genel bir yöntem bulmak adına *vücut planları* geliştirilmiştir (Forsyth and Fleck., 1997).

### 1.4.1 Vücut planları

Vücut planları insan ya da hayvan vücutlarının kesimlerini göstermek için oluşturulan gruplama safhalarıdır. Vücut planları insan ve hayvanları vücutlarını daire ya da silindir şekilleri olarak modeller (bkz: **Şekil 1-7**).



Şekil 1-7: At için hazırlanmış vücut planı

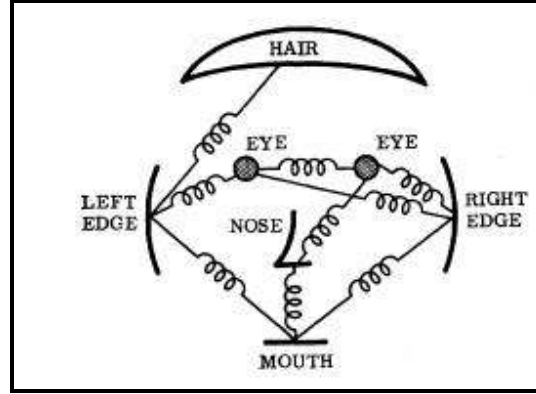
Bu modeli kullanan arama yöntemleri ise imge üzerindeki insan ya da hayvan şeklini ararken de çok safhalı bir arama yapar. Örnek vermek gerekirse, eğer resimde aranan bir at ise öncelikle gövde, ayak ve boyun parçaları aranır. Daha sonra ise gövde-ayak, gövde-boyun çiftleri aranır. Bu şekilde 3'lü ve 4'lü gruplar oluşturulmaya devam edilir. Bulunan vücut planları daha sonra eğitim ve test sınıflandırıcıları vasıta ile farklı imgelerde test edilir.

Vücut planları görüntüden insanı bulmak konusunda genel amaçlı yaklaşımlarda bir tanesidir. Fakat nesne üzerinde rahat-bol giysi olması durumunda başarımlar oranları düşebilmektedir.

Vücudun kesimlere ayrılmasını kullanılan bir diğer modelleme ise *Resimsel Yapılar*'dır.

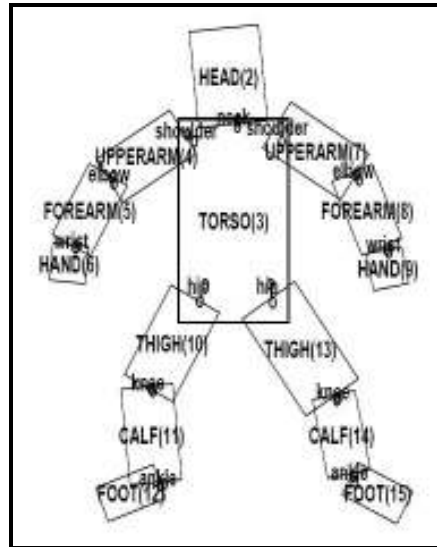
#### 1.4.2 Resimsel yapılar

Resimsel yapılar modelindeki temel fikir, bir nesneyi formu bozulabilir konfigürasyonlara göre gruplanmış parçalar topluluğu olarak ifade etmektir. Her parçanın görünüşü ayrı olarak modellenir ve parçaların çiftleri arasındaki formu bozulabilen konfigürasyon bağlantısı yay benzeri birleşimler ile gösterilir. Bu modeller görsel şekiller için nitelikli tanımlayıcılar olup genel algılama problemlerinin çözümleri için uygundur (bkz: **Şekil 1-8** (Felzenswalb and Huttenlocher., 1972)).



Şekil 1-8: İnsan yüzü için geliştirilen resimsel yapılar modeli

Resimsel yapılar imgelerdeki insanın bulunmasında sık başvurulan bir yöntemdir. Ronfard yaptığı çalışmada insanın vücudu için 14 parçadan oluşan eklemli bir model kullanmıştır (bkz: **Şekil 1-9**).



Şekil 1-9: 14 eklemli ve 15 parçalı insan modeli

Çizelge 2: Modele göre parçaların dağılımı

Parça ismi	Modeldeki yeri
Kafa	2
Gövde	3
Üst kol	4,7
Ön kol	5,8
El	6,9
Uyluk	10,13
Baldır	11,14
Ayak	12,15

**Şekil 1-9**'de verilen modelde her bir vücut parçası  $P_i$  olarak ifade edilip merkezi  $[x_i, y_i]$ , uzunluğu  $s_i$ , yönelimi  $\theta_i$  olarak gösterilir.

Ronfard'ın çalışmasında ilk etapta 100 adet imge de elle işaretleme yapılarak Çizelge 2'de belirtilen parçalar bulunup sınıflandırıcıların eğitimde kullanılmıştır. Elle işaretlenen parçalar Şekil 1-10' da gösterilmiştir.

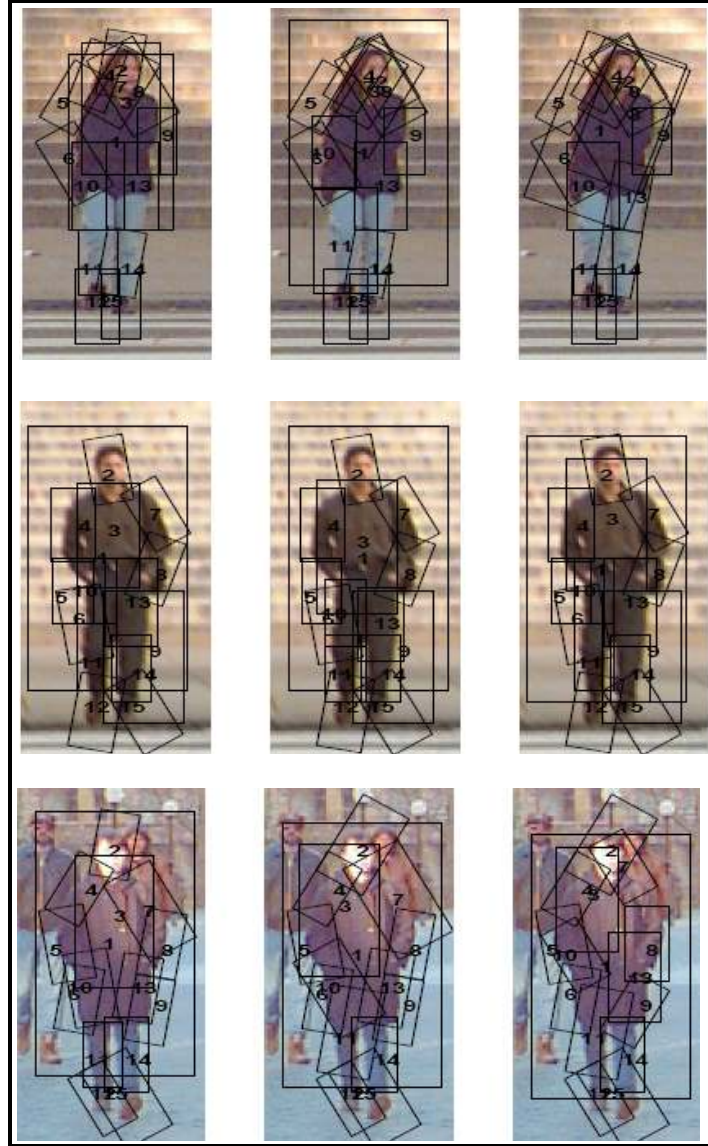
Kullanılan sınıflandırıcılar SVM ve RVM'dir. Elle işaretleme yapılan parçalara ise *parça algılayıcı* adını vermiştir. Dikkat edileceği üzere her insan için modeldeki vücut parçaları farklı büyüklük ve yönlendirmeye sahip olabilir. Bu problemin üstesinden gelebilmek için parça algılayıcıları test imgesinde ararken 8 farklı ölçekleme ve  $10^\circ$ 'lik adımlardan oluşan 36 farklı yönelim kullanılmıştır.

Tanımlayıcısı  $[x_i, y_i, s_i, \theta_i]$  olan resimdeki her bir dikdörtgen için ise 1. ve 2. türevlerini aldıktan sonra  $\sigma = 1$  için Gauss filtresi uygulamıştır. Böylece öz nitelik vektörleri  $\{G, \nabla_x G, \nabla_y G, \nabla_{xx} G, \nabla_{xy} G, \nabla_{yy} G\}$  şeklinde ifade edilmiştir.



Şekil 1-10: Elle işaretleme yapılan eğitim imgesi

Daha sonra öz nitelik vektörleri test imgeleri üzerinde farklı ölçek ve yönelimlerde gezdirilmiş, eşleşme sonuçları kayıt edilmiş ve sınıflandırıcılara verilerek başarımları ortaya çıkarılmıştır (bkz: **Şekil 1-11**).



Şekil 1-11: Parça algılayıcılarının test imgelerinde buldukları parçalar

Ronfard'ın çalışması insan algılamaya problemine yaklaşımı açısından bakıldığında çalışmamız ile benzeşim gösterir. Bahsi geçen çalışmanın ilk etabında parça algılayıcıları bulabilmek için elle işaretleme yöntemine gidilmiş ve elde edilen öz nitelik vektörleri sınıflandırıcının eğitiminde kullanılmıştır. Böylece pozları

algoritmik olarak bulan çeşitli görüntü işleme yöntemleri için bir soyutlama katmanı (*abstraction layer*) sağlanmıştır. Çalışmada aynı yöntem izlenerek öznelik vektörleri elle işaretlenen vücut parçalarından çıkarılmıştır. Bu sayede farklı öznelik vektörleri ile çalışan başka çalışmalar için bir ara yüz sağlanmıştır.

Çalışmamızın ayrıştığı noktalar ise görünmeyen parçaların, *occluded parts*, ele alınış şeklidir. Bu noktada Ronfard üstü kapanmış parçaları tahmin yoluyla işaretlemiş ve aynı zamanda bu parçaları “üstü kapalı” olarak etiketlemiştir. Bizim çalışmamızda ise üstü kapalı parçalar hiç işaretlenmemiştir fakat işaretlenmemiş olduğu uygulama esnasında kayıt altına alınmıştır. Bu anlamda daha az kapsamlı bir vücut modeli, 14 parçaya 10 parça, dolayısı ile de daha az sayıda öznelik vektörü kullanılması açısından çalışmamız farklılık göstermektedir.

## 1.5 Hareketin Algılanması

İmgelerdeki insan hareketinin algılanması için sık başvurulan bir yöntem poz bilgisini tanımlayan öznelik vektörlerinin sınıflandırıcılar vasıtasıyla eğitilip sonrasında ise test edilmesidir. Kullanılan öznelik vektörleri çok çeşitlilik gösterebilir.

Örneğin; İkizler çalışmalarında optik akış, çizgi ve akış histogramları kullanarak elde ettikleri hareket bilgilerini SVM ile test etmiştir ve yüksek tanıma sonuçları elde etmişlerdir (Ikizler vd., 2008)

Yine aynı çalışmada yerel özellikler yerine global özellikler kullanılarak genel bir hareket sınıflandırıcı üzerinde çalışılmıştır. Burada imgenin genelindeki global hız özelliklerine bakılarak bir hareketin “el sallama” ya da “koşma” olmasına göre sınıflandırılmıştır. Basit bir hareket sınıflandırıcı olması açısından başarılı sonuçlar veren



yaklaşım el sallayan bir insanın koşan bir insana göre hareket genelinde daha yavaş bir hız sergilemesinden yola çıkmıştır.

### 1.5.1 BoW modeli

Hareket algılamada kullanılan bir diğer yöntem ise poz cümlelerinin kullanımınıdır. Hareketlerin resimleri oluşturan pozların dizilimleri aracılığı ile algılanması BoW modellerinin örüntü algılama alanına uygulanması ile birlikte gündeme gelmiştir. Poz dizilimlerinin ve hareketin kendisi arasındaki bağıntının benzeri hece-kelime-cümle bağıntısından hareketle ortaya konulmuştur. Hecelerin birleşimi kelimeleri, kelimelerin birleşimi de cümleleri oluşturur. Bu noktadan hareketle bir harekete ait imgelerdeki pozlar çıkarıldıktan sonra bunların ardı ardına sıralanması hareketin kendisi tanımlar. Bu yaklaşım BoW (*bag of words*) modeli olarak da bilinmektedir. BoW yöntemi NLP ve belge sınıflandırma çalışmalarında sıkça başvurulan bir yöntemdir. Bu yöntemde kelimelerin sırası önemsizdir. Örneğin;

“*Ali okula gidiyor*” ve “*gidiyor Ali okula*” BoW modeli açısından aynı anlama gelmektedir.

BoW yönteminde metindeki kelimeler vektör elemanlarını temsil eder. Bu şekilde bir cümle içerdiği kelimeler kullanılarak vektörler şeklinde ifade edilebilir.

Örnek: Elimizde aşağıda cümleleri içeren 2 farklı belge olsun.

Cümle 1 = “Ali ve Ayşe film izlemeyi sever”

Cümle 2 = “Ali ayrıca maç izlemeyi sever”

Buradaki kelimeleri vektörler şeklinde ifade ederek bir sözlük oluşturacak olursak

Kod kitabı	1	2	3	4	5	6	7	8
	Ali	film	izlemeyi	sever	ayrıca	maç	ve	Ayşe

Şekil 1-12: BoW modeli ile oluşturulan sözlük

Kelimelerin cümlelerde ne kadar kullanıldığını vektör notasyonu olarak da gösterebiliriz.

Kod kitabı	1	2	3	4	5	6	7	8
	Ali	film	izlemeyi	sever	ayrıca	maç	ve	Ayşe
Cümle-1	1	1	1	1	0	0	1	1
Cümle-2	1	0	1	1	1	1	0	0

Şekil 1-13: BoW yöntemi ile oluşturulan kelime histogramları

Buradan da anlaşılacağı üzere her iki cümlede de kesişen bazı kelimeler vardır. Şekil 1-13'e bakılırsa kelimelerin tekrar sayısı gösterilmiştir. Hangi kelimenin tekrar edip etmediği ise sözlükteki sıra numaralarından çıkarılır.

### 1.5.2 BoW modelinin görüntü algılamada uygulanışı

BoW modelini hareket algılama problemlerinde kullanmak için şu adımlar izlenir;

- *Öz niteliklerin algılanması*: Verilen bir imgedeki parçaların çıkarılması işlemidir. Bu işlem sonucu elde edilen çıktılara “kelime” diyebiliriz
- *Öz niteliklerin tanımlarının çıkarılması*: Parçalar çıkarıldıktan sonra her bir parçanın sayısal bir vektör olarak ifade edilmesi işlemidir. Öz nitelik tanımlayıcılarının algılanan her bir öz niteliğe atı döndürme, büyüklük gibi değerleri ifade edebilen vektör dizilimleri olması gerekir.
- *Kod kitabı oluşturulması*: BoW modelinin oluşumundaki son adım kod kitabının oluşturulmasıdır. Kod kitabını öz nitelik tanımlayıcılarının hepsinin yer aldığı bir sözlük gibi düşünmek mümkündür. Kod kitabı oluştururken en çok başvurulan yöntem k-means yöntemidir. Bu yöntemle birbirine benzeyen öz nitelik tanımlayıcıları aynı etiketli kümelere yerleştirilir ve kod kitabı oluşmuş olur.

Kardelen ve Duygulu yaptıkları çalışmada test kümesindeki her hareketin diziliminin  $A_i$ , poz kelimelerinin sıralı diziliminden meydana geldiğini öne sürmüşlerdir. Poz kelimelerini bulabilmek için resimlerdeki bütün pozların  $f_{ij} \in A_i, j = 1 \dots |A|$ , HOG (*Histogram of Oriented Gradients*) özellikleri bulunmuştur Daha sonra bulunan bütün pozlar kümelere ayrılmış ve kümelerin merkezleri poz kelimeleri kabul edilmiştir,  $P = \{p_1 \dots p_k\}$ . Daha sonra ise her hareket kendisini oluşturan poz kelimelerinin sıralı dizilimi şeklinde ifade edilmiştir ve buna da poz cümlesi denmiştir.

$$A_i = a_1 a_2 \dots a_n \quad (a_n \in P)$$

Son kısımda ise katar karşılaştırma yöntemleri kullanılarak poz cümleleri birbirleri ile karşılaştırılmıştır (Kardelen ve Duygulu., 2008).

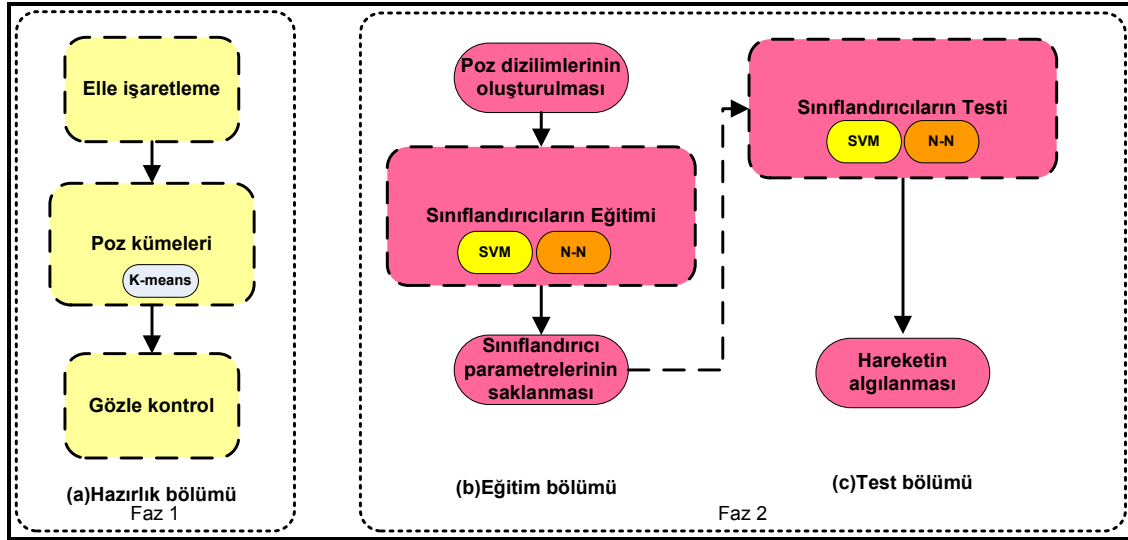
Kardelen ve Duygulu'nun çalışmalarında izledikleri yöntem incelendiğinde poz kelimelerinin kullanımı açısından çalışmamızla benzerlik gösterir. Öz nitelik vektörlerinin görüntü işleme yöntemleri ile çıkarımı bizim çalışmamızda soyutlanmıştır fakat bahsi geçen çalışmada HOG yöntemi kullanılmıştır.

Bir diğer ayrışım noktası ise çalışmamızda 2 farklı kütüphanenin görüntülerinin kullanılmış olmasıdır. Böylece hibrid bir görüntü kütüphanesi oluşturulmuştur. Ayrıca BoW modellerinde kullanılan katar karşılaştırma, *string matching*, yöntemi yerine SVM ve Yapay Sinir Ağları kullanılarak farklı sınıflandırıcıların başarımları karşılaştırılmıştır.

## 1.6 Önerilen Yöntem

Çalışma 3 kısımdan oluşmaktadır. Birinci kısımda öz nitelik vektörleri elde edilecektir. Bunun için resimlerdeki insan vücut parçalarının merkez noktaları elle işaretlenecektir. Daha sonra bu merkezlerin kartezyen koordinatları kullanılarak gövde ile diğer parçalar arasında açılar hesaplanacaktır. Bu açılar kullanacağımız öz nitelik vektörleridir. Bütün resimler işaretlenip öz nitelik vektörleri hesaplandıktan sonra k-means yöntemi ile açı vektörleri kümelere ayrılacaktır. Bu kümelerin her biri benzer açısal vektörler içeren görüntüler barındıracağı için poz kelimelerini oluşturacaktır.

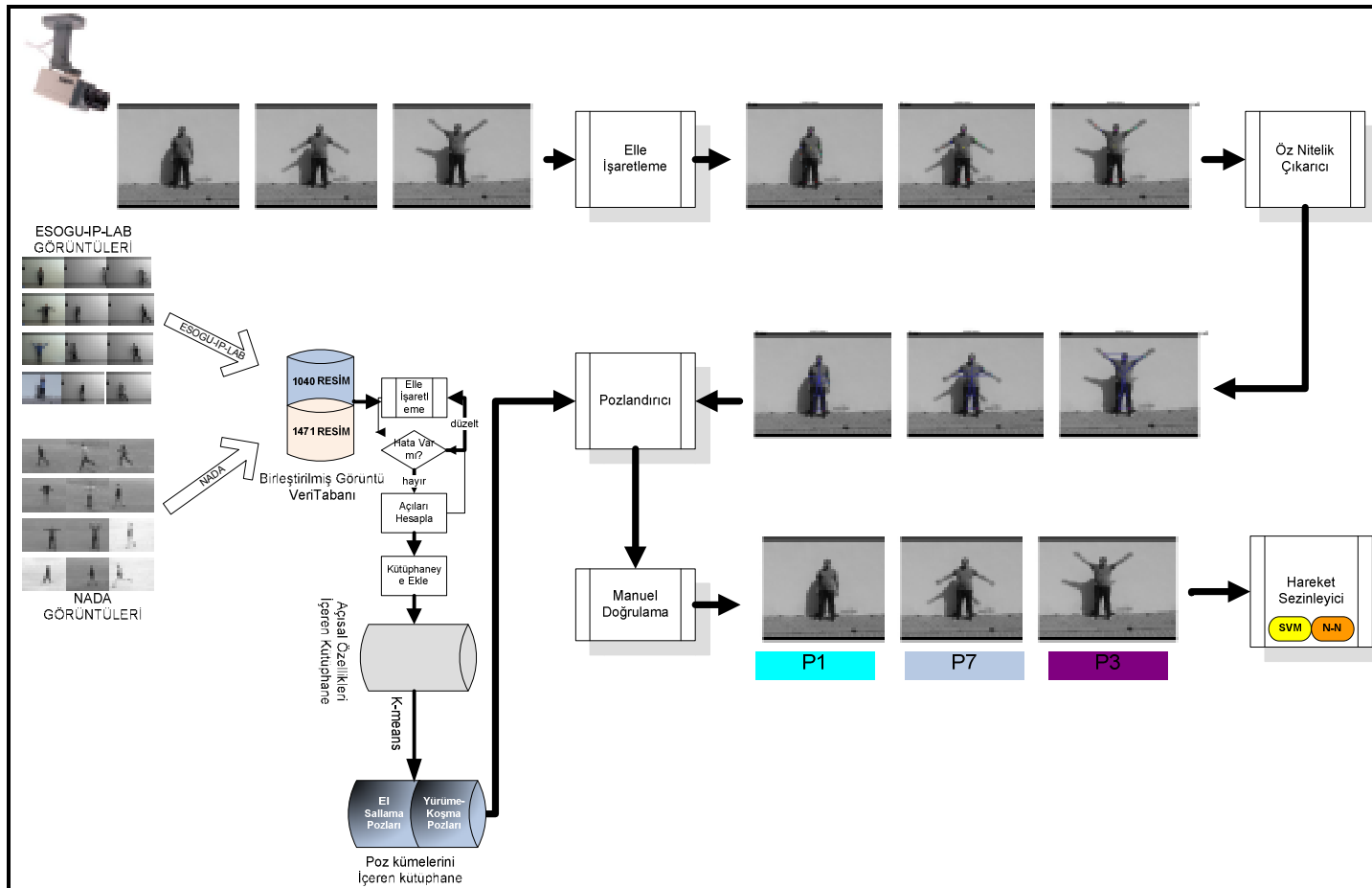
2. kısımda ise oluşturulan poz kelimelerinden poz dizilimleri elde edilecektir. Daha sonra ise poz dizilimleri eğitim ve test olarak gruplara ayrılarak farklı sınıflandırıcılar ile test edilecektir (bkz: **Şekil 1-14**).



Şekil 1-14: Çalışmaya ait genel akış diyagramı

Çalışmada öz nitelik vektörleri olarak vücut parçalarının gövde ile yaptıkları açıların kullanılması farklı vücut büyüklükleri ve yönelimlerin olumsuz etkisini en aza indirecektir.

3. bölümde ise sınıflandırıcılardan elde edilen sonuçlar ve başarımları değerlendirilecektir.



Şekil 1-15: Sisteme ayrıntılı genel bakış

## 2 KULLANILACAK KÜTÜPHANELERİNİN HAZIRLANMASI

Çalışma süresince 2 farklı görüntü veritabanı birleştirilmiş ve bir kütüphane haline getirilmiştir. Farklı veritabanlarının birleştirilmesindeki amaç çalışmamızı farklı ölçekleme, çözünürlük, yönelim karakterleri taşıyan görüntülerde sınamaktır. Öz nitelik vektörleri olarak açısız özellik kullanmamızı sebebi ise ölçekleme, yönelim gibi yerel etkilerden korunmaktır.

### 2.1 Kullanılan Veri Tabanları

Veritabanları sıkıştırılmış *avi* (*audio video interleave*) uzantılı videolardan oluşmaktadır. Çalışma kolaylığı açısından video görüntüleri VirtualDump yazılımı ile *jpg* uzantılı dosyalara çevrilmiştir. Bu tez çalışmasında veri tabanlarında bulunan el sallama, koşma ve yürüme hareketleri üzerinde durulmuş ve bu hareketlerin işaretlenmeleri ile bir kütüphane oluşturulmuştur. Diğer hareket türleri çalışma dışında bırakılmıştır.

#### 2.1.1 ESOGU-IP LAB veritabanı

Görüntü veritabanlarında ilki ESOGU-IP LAB ismi ile anılmakta olup Osmangazi Üniversitesi öğrencilerinin yer aldığı görüntülerden oluşmaktadır. Öğrenciler koşma, yürüme, jogging, el sallama, oturma-kalkma hareketlerini kapalı ve açık hava ortamlarında çeşitli kereler tekrarlamıştır. Görüntüler sıkıştırılmış *avi* formatında olup çözünürlük 720x576 pikseldir. Kamera çekim hızı 25fps'dir. Kameranın otomatik odaklanma özelliği çekimler esnasına açık kaldığı için kimi görüntülerde odak kaybı söz konusudur. **Şekil 2-2**'de bu veritabanına ait değişik görüntüler mevcuttur.

### 2.1.2 NADA veritabanı

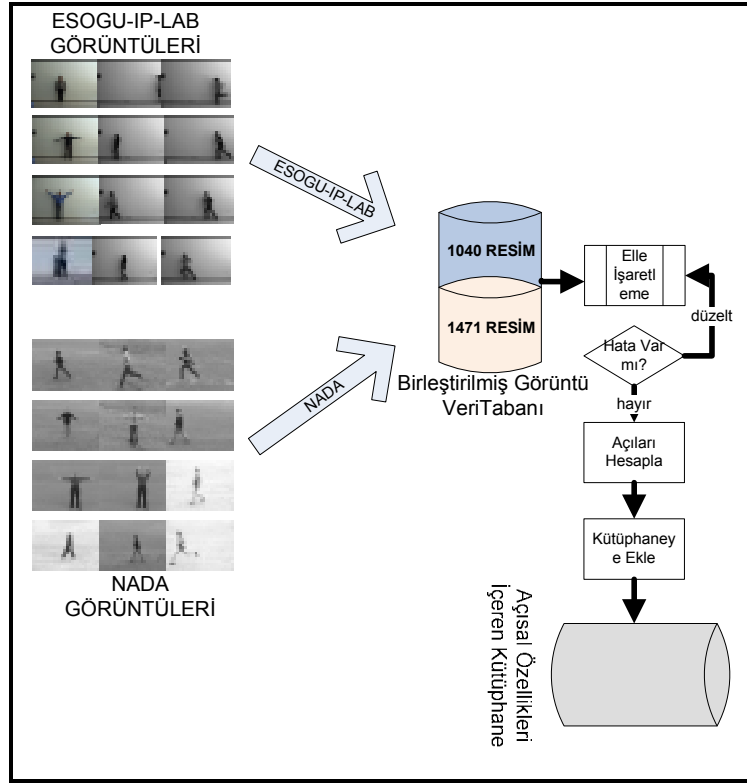
Kullanılan diğ er veritabanı ise eğitim ve test için, toplam 20 kişinin gerçekleştirdiği 5 farklı hareketin videolarının bulunduğu bir hareket veritabanıdır (Schüldt,et al., 2004). Bu veritabanı çalışma süresince NADA veritabanı adı ile anılmıştır. NADA veritabanı 6 değişik insan hareketi içermektedir: *yürüme, koşma, jogging, el sallama ve el çırpma*. Bu hareketler 25 kişi tarafından 4 farklı senaryoda tekrarlanmıştır. Bu senaryolar: *açık hava s1, açık hava ve farklı ölçeklerde s2, açık hava ve farklı kıyafetler s3 ve kapalı ortam s4* olarak tasarlanmıştır. Veritabanı 2391 resim içermektedir. Bütün resimlerde arka plan homojen olup kamera çekim hızı 25fps'dır ve görüntü çözünürlükleri 266x200 pikseldir. **Şekil 2-3**'de bu veritabanına ait değişik görüntüler mevcuttur.

Her iki veritabanındaki görüntülerde hareketi gerçekleştiren tek bir insan vardır ve hareketler tek yönde gerçekleşmektedir. Test ve eğitimlerde 2 veritabanının imgeleri ortak kullanılmış olup çeşitlilik sağlanmaya çalışılmıştır. Örnek görüntüler için **Şekil 2-4**'e bakınız.

### 2.1.3 Görüntü veritabanlarından derlenerek kütüphane oluşturulması

Veritabanlarındaki görüntülerden bir kütüphane elde etmek çalışmanın önemli bölümlerinin başında gelmektedir. Hazırlanan kütüphaneler daha sonraki çalışmalara da katkı sağlanmış olacaktır. Kütüphanenin derlenme süresi kullanılacak olan bütün görüntülerin elle işaretlenerek açıs al özelliklerin çıkarılması sürecini kapsar. **Şekil 2-1: Görüntülerden oluşturulan kütüphane** incelenebilir. Görüntülerin işaretlenerek açıs al özelliklerini çıkarıp kütüphane oluşturabilmek amacı ile Matlab ortamında bir uygulama geliştirilmiştir. Uygulama ilgili detaylar bölüm 3.1.2-Elle işaretleme kısmında detaylandırılacaktır.





Şekil 2-1: Görüntülerden oluşturulan kütüphane

#### 2.1.4 Görüntü veritabanlarının kullanım dağılımı

Çalışmamızda girdi olarak kullanılacak olan test ve eğitim setleri T1, T2, T3, T2+T3 olarak planlanmıştır. Bu setlerin her biri el sallama, yürüme, koşma hareketlerine ait.

El sallama hareketine ait bir resim 10x1'lik açısal özellik vektörü ile ifade edilmektedir. Buna karşın yürüme ve koşma hareketleri için açısal özellik vektörü 6x1'dir. Bu bilgiler ışığında **Çizelge 3**'e bakıldığında açısal özellikleri içeren kütüphane de;

$$\text{El sallama hareketi için} = 910 \times 10 = 9100$$

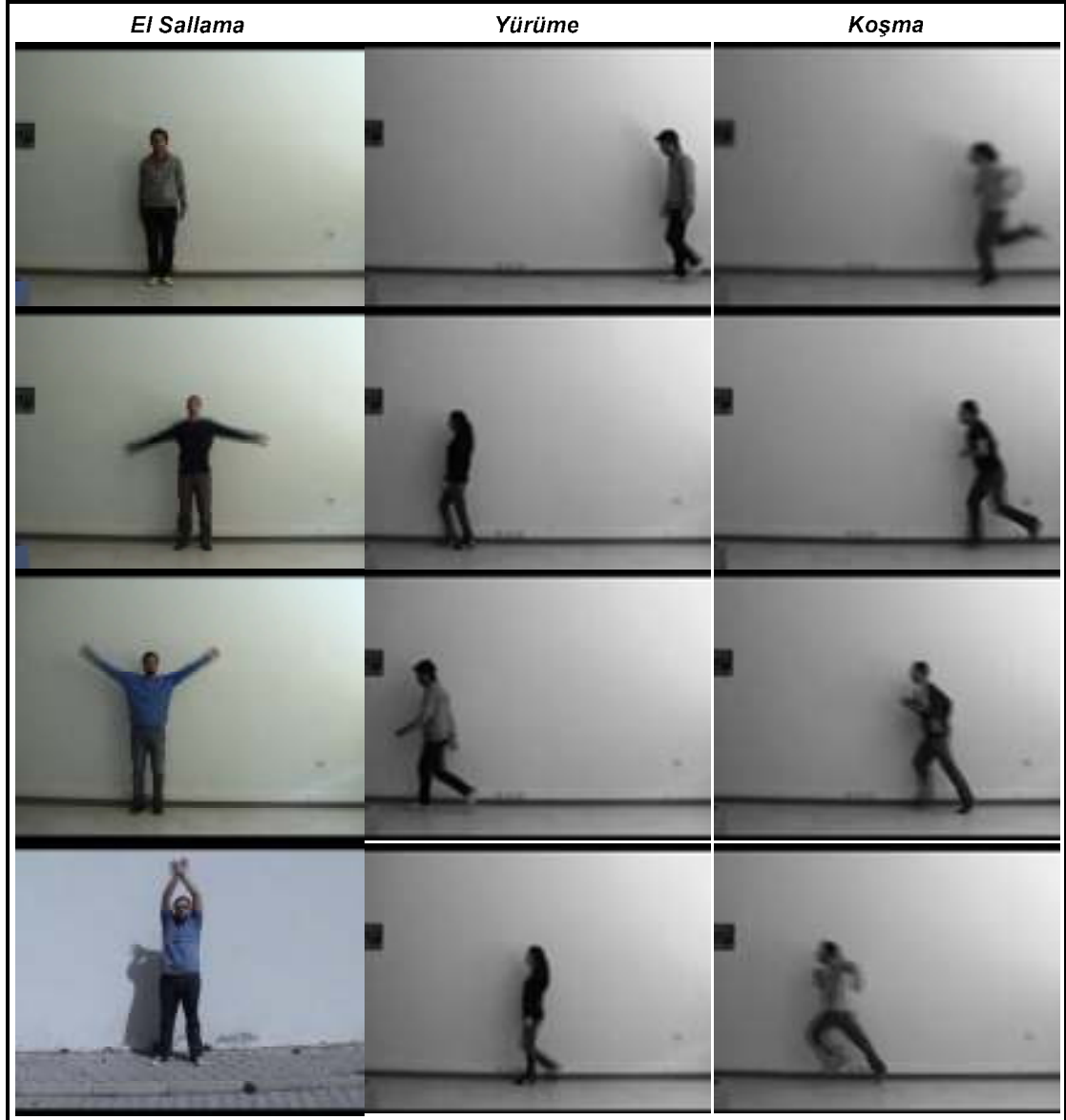
Yürüme ve Koşma için =  $1640 \times 6 = 9840$ , toplamda ise 18940 adet açısız özellik çıkarılmıştır.

Çizelge 3: Veri seti-hareket türüne göre görüntü sayısı dağılımı

Veri Seti	Hareket Türü			İmge Sayısı
	El Sallama	Yürüme	Koşma	
T1	415	460	182	1057
T2	415	448	169	1032
T3	80	267	114	461
Genel Toplam	910	1175	465	<b>2550</b>

Çizelge 4: Hareket türlerine göre veri seti-görüntü sayısı dağılımı

Veri Seti	Hareket Türü			
	El Sallama		Yürüme-Koşma	
	ESOGU-IP-LAB	NADA	ESOGU-IP-LAB	NADA
T1	350	65	15	627
T2	350	65	15	602
T3	0	80	349	32
Genel Toplam	700	210	379	1261



Şekil 2-2: ESOGU-IP-LAB veritabanı örnek resimler



Şekil 2-3: NADA veritabanı örnek resimler



Şekil 2-4: Veritabanlarının ortak kullanımı

### 2.1.5 Açısal özelliklerden poz kelimeleri kütüphanesi oluşturulması

Daha önce bahsedildiği üzere bu çalışma da hareket sezinleme yöntemi olarak poz dizilimlerinin sınıflandırıcılarda eğitilip test edilmesi yöntemi kullanılmıştır. Bu anlamda açısal özelliklerden oluşan kütüphane bu hali ile hareket sezinlemede kullanılamaz. Poz

kelimelerini oluşturmak için açısız özellik kütüphanesinde her girdi k-means yöntemi ile kümeler ayrılacak ve aynı kümelere düşen açısız özelliklerine sahip görüntüler aynı poz etiketini alacaktır. Daha sonra bu poz etiketleri hareketi yapan kişiler bazında başlangıçtan hareket sonuna kadar gruplandırıldığında poz dizilimleri kütüphanesi elde edilmiş olacaktır. T1, T2, T3 eğitim ve test gruplarının her biri için poz dizilimleri oluşturulmuştur. T1 eğitim grubu olduğu için bu grubun poz dizilimi SVM ve NN' de eğitiminde kullanılmıştır.

Kütüphaneden farklı çalışmalarda faydalanabilmek için açısız özelliklerin çıkarımının ve pozlandırma fazlarının elle ya da başka yöntemlerle tekrarlanması gerekmektedir. Bu noktadan sonra elde edilen poz dizilimleri farklı çalışmalarda da kullanılabilir. Poz dizilimlerini oluşturmak için Matlab ortamında bir uygulama geliştirilmiştir. Uygulamaya ait detaylar bölüm 3.1.4.4-Poz dizilimlerinin oluşturulmasında verilmiştir.

### 3 ÖZ NİTELİKLERİN ÇIKARILMASI

Çalışmada imgelerdeki hareketi algılayabilmek için öz nitelik olarak vücut parçalarının gövde ile yaptıkları göreceli açılar kullanılmıştır. Daha önce de açıklandığı üzere açısal özelliklerin kullanılmasının en önemli avantajı farklı vücut büyüklüklerine ve parça yönelimlerinin getireceği olumsuz etkilerden korunmaktır.

#### 3.1.1 Kullanılan insan vücut modeli

Öz nitelik vektörleri çıkarılırken insan vücudu için 10 parçalı bir model düşünülmüş ve parçalar dikdörtgen çerçevelerin içerisinde kalacak şekilde bir model oluşturulmuştur. Bu modeldeki parçalar gövde, kafa, sağ üst kol, sağ alt kol, sol üst kol, sol alt kol, sağ üst bacak, sağ alt bacak olacak şekilde gruplanmıştır.

Çizelge 5: Vücut modeli-parça dağılımı

Parça ismi	El Sallama	Koşma/Yürüme
	Parça No	
Gövde	1	1
Kafa	2	2
Sağ üst kol	3	- <sup>1</sup>
Sağ alt kol	4	-
Sol üst kol	5	5
Sol alt kol	6	6
Sağ üst ayak	7	-
Sağ alt ayak	8	-
Sol üst ayak	9	9
Sol alt ayak	10	10

<sup>1</sup> (-) ile işaretli parçalar bu hareket tiplerinde arkada kaldıkları için işaretlenmemiştir.

İşaretleme yapılırken parçaların merkez noktaları işaretlenecektir. İşaretleme bitiminde ise her parçanın merkezinin gövde ve kendi yatay eksen ile yaptığı açı bulunacak ve bir sonraki resim işaretlenecektir (bkz: Şekil 3-1(a))

Koşma ve yürüme hareketlerinde kol ve bacakların birbirlerini kapatması durumunda arkada kalan parçanın merkezi görülemediği için bu hareketlerde sadece önde kalan kol ve bacaklar işaretlenmiştir (bkz: Şekil 3-1(b)).

Bu da el sallama ve koşma ve yürüme hareketlerinin öz nitelik vektörlerinin uzunluklarının farklı olmasına neden olmuştur.

El sallama hareketi için  $Fe$  , Koşma ve Yürüme için ise  $Fkr$  açısız öz nitelik vektörü, gövde ve parçalar arasındaki açılarda şu şekilde olsun.

$\theta_g$  = Gövde-Gövde arası açı

$\theta_k$  =Gövde-Kafa arası açı

$\theta_{suk(1)}$  =Gövde-Sağ üst kol arası açı

$\theta_{sak(1)}$  =Gövde-sağ alt kol arası açı

$\theta_{suk(2)}$  =Gövde-sol üst kol arası açı

$\theta_{sak(2)}$  =Gövde-sol alt kol arası açı

$\theta_{sua(1)}$  =Gövde-Sağ üst ayak arası açı

$\theta_{saa(1)}$  =Gövde-sağ alt ayak arası açı

$\theta_{sua(2)}$  =Gövde-Sağ üst ayak arası açı

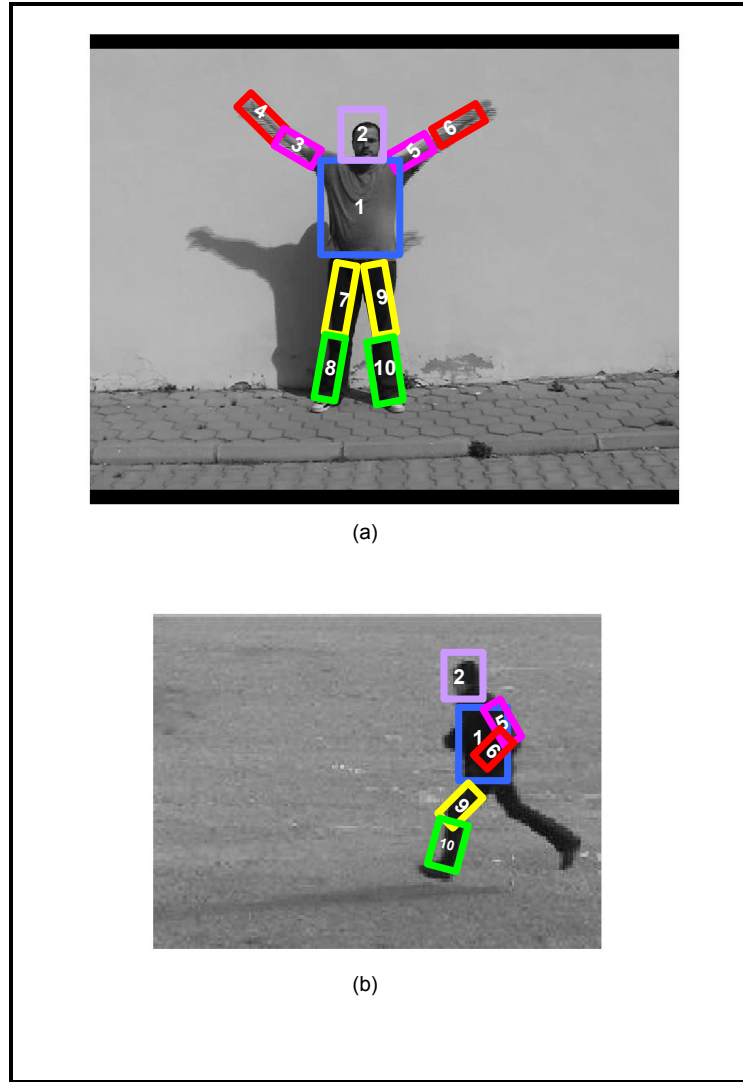
$\theta_{saa(2)}$  =Gövde-sağ alt ayak arası açı

Bu durumda;

$$Fe = \{\theta_g, \theta_k, \theta_{suk(1)}, \theta_{sak(1)}, \theta_{suk(2)}, \theta_{sak(2)}, \theta_{sua(1)}, \theta_{saa(1)}, \theta_{sua(2)}, \theta_{saa(2)}\}$$

$$Fkr = \{\theta_g, \theta_k, \theta_{suk(2)}, \theta_{sak(2)}, \theta_{sua(2)}, \theta_{saa(2)}\} \text{ şeklindedir}$$



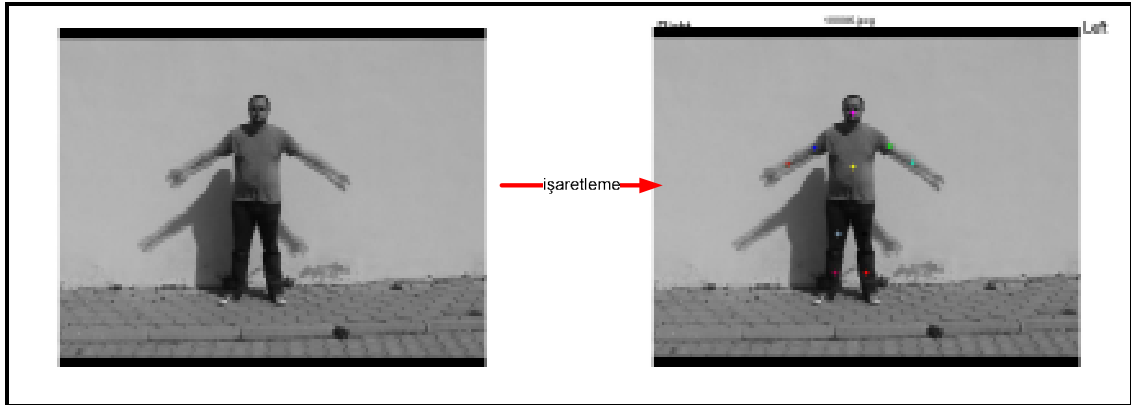


Şekil 3-1: Elle işaretlenmiş insan vücudu modeli (a) El sallama (b) Koşma

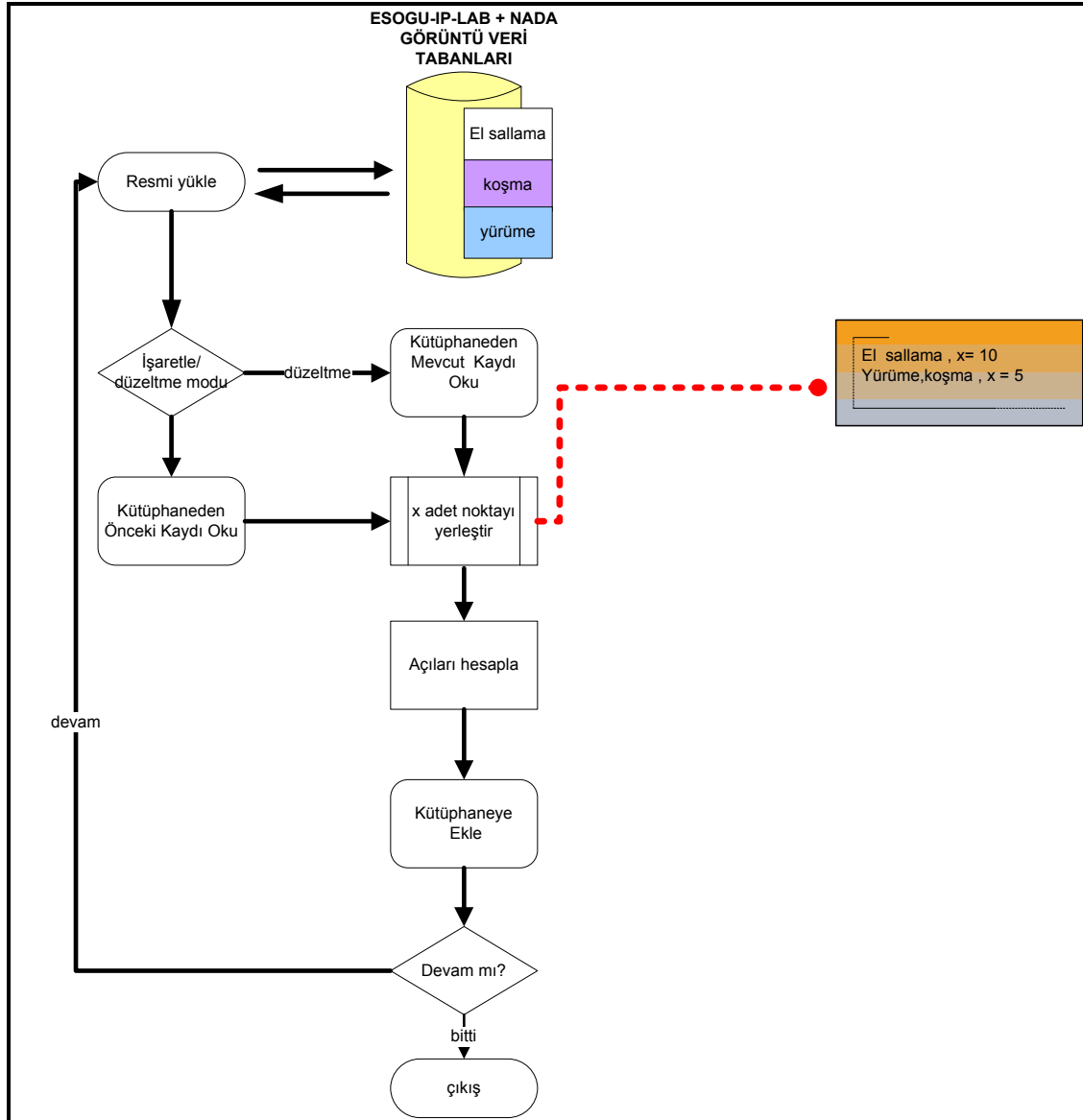
### 3.1.2 Elle işaretleme

MATLAB ortamında geliştirilen uygulama ile imgeler üzerinde işaretleme işlemi yapılmıştır. Elle işaretleme işlemi dikkat isteyen ve zorlu bir işlem olduğu için uygulama da çeşitli kolaylıklar geliştirilmiştir. Örneğin; bir imge ilk kez yüklenirken kendinden önceki işaretlenmiş imgenin parça koordinatlarını yüklenerek işaretleme işlemi

kolaylaştırılmıştır. El sallama hareketinde bacaklar hiç hareket etmediği için ilk imge işaretlendikten sonra diğer bütün el sallama hareketlerinde bacaklara ait parçalarda aynı merkez noktaları otomatik atanmıştır. İşaretlenen görüntüdeki merkez koordinatlar aynı isimdeki mat uzantılı dosyaya kayıt edildikten sonra bir sonraki görüntü uygulamaya yüklenmiştir. Bu şekilde işaretleme süreci hızlandırılmıştır. Eğer işaretlemede bir hata var ise gerçek zamanlı ya da geçmişe dönük olarak şekil üzerinde düzeltme yapılması mümkün hale getirilmiştir. Uygulamanın akış diyagramı **Şekil 3-3**' de gösterilmiştir.



Şekil 3-2: Elle işaretlenmiş resim



Şekil 3-3: Elle işaretleme uygulaması akış diyagramı

Uygulamada imgeleri işaretlemek için Matlab'in *impoint* fonksiyonu kullanılmıştır. Bu rutin imge üzerine görsel olarak seçilen bir koordinata istenilen renkte bir nokta çizer ve çağırıldığı uygulamaya bir işaretçi döndürür. Bu işaretçi vasıtası ile daha sonra da işaretlenen noktayı değiştirmek mümkündür.

Uygulamanın 2 çalışma modu vardır. *İşaretleme modunda* verilen dizin bilgisinde işaretlenecek imgeler olmalıdır. Uygulama imgeleri gri-seviye imge olarak okur. Burada Matlab'in *imread* rutini kullanılmıştır. İşaretleme modunda her seferinde baştan işaretleme yaparak hata oranını artırmamak adına o anda yüklü imgenin bir önceki kayıtları kontrol edilir. Şayet kayıt var ise kütüphanede okunan noktalar mevcut imgenin üzerine otomatik olarak işaretlenir. Bu sayede işaretleme işlemi hızlandırılmış olur ve olası işaretleme hatalarının önüne geçilmiş olunur.

*Düzeltilme modunda* ise kütüphanede mevcut imgenin kayıt dosyası okunur ve noktalar imge üzerine işaretlenir. Kullanıcı noktaları istediği yerlere sürükleyerek düzeltme işlemini tamamlar ve kütüphane kaydını güncellenir. Vücut parçaları işaretlenirken izlenen bir sıra vardır. Ayrıca hangi vücut parçası için ne renk nokta konulacağı bellidir. İşaretleme sırası şu şekildedir;

1. Gövde
2. Kafa
3. Sağ Üst kol
4. Sağ alt kol
5. Sol üst kol
6. Sol alt kol
7. Sağ alt bacak
8. Sağ üst bacak
9. Sol üst bacak
10. Sol alt bacak

*El sallama hareketinin işaretlenmesi:*

El sallama hareketinde bütün uzuvlar hep seçilebilir olduğu için 10 adet işaretleme yapılır.

*Koşma ve Yürüme hareketlerinin işaretlenmesi:*

Bu hareketleri içeren görüntüler özellikle ekranın solundan girip sağından çıkan kişilerin olduğu görüntülerden seçilmiştir. Bu sayede arkada kalan vücut parçaları her zaman sağ kol ve bacak olmuştur. Dolayısı ile bu parçalar hiç işaretlenmemiştir. Bu nedenle bu hareket tipinde toplam 6 işaretleme yapılmıştır.

Çizelge 6: Elle işaretlemede kullanılan veri yapısı

Değişken ismi	Açıklama	Büüklük
pointsMarked	İşaretlenen noktaların kartezyen koordinatları	[10x2]
angle	İşaretlenen noktaların gövde ve yatay eksenle yaptıkları açılar	[10x1]
poseID.actionType	Hareket türü	[1x1]
poseID.fullPath	Resmin bilgisayardaki yeri	[1x80]
poseID.frameNo	İşaretlene imgenin imge dizisindeki indeksi	[1x1]

### 3.1.3 Açılar hesaplanması

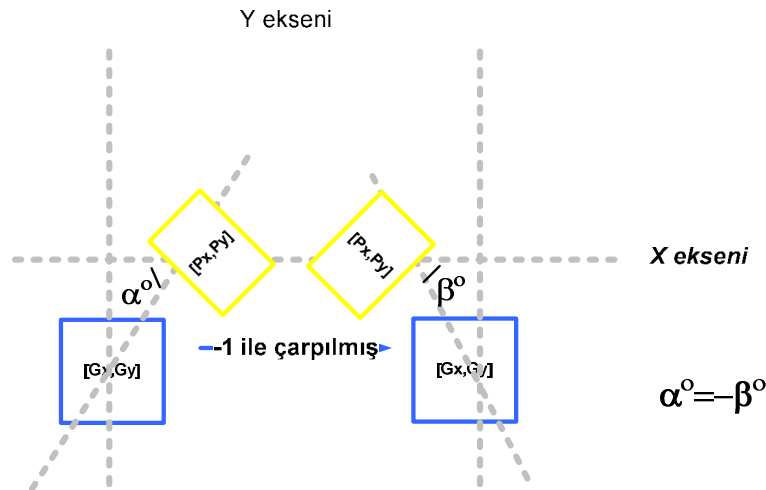
Açılar gövdenin merkezi-parçanın yatay eksenine-parçanın merkezi arasındaki yatay ve dikey eksenler arasındaki farkları gösteren şekilde hesaplanmıştır. Gövde merkezinin yatay ve dikey eksenin farklı kısmında kalan parçaların hesaplanan açılar -1 ile çarpılmıştır. Bunun nedeni eksene göre yansıması alınmış şekilde duran parçaların aynı pozisyona dâhil edilmesini önlemektir (bkz:Şekil 3-4).

$P(x, y)$  Gövdenin,  $P(m, n)$  diğer parçanın merkez noktalarını gösterecek olursa aralarındaki açı;

$$\alpha = \arctan\left(\frac{|Py - Gy|}{|Px - Gx|}\right) \text{ Şeklinde bulunur.}$$

Mutlak değer kullanılmasının sebebi  $\arctan$  fonksiyonunun kartezyen koordinat sisteminde bölge bilgisini vermemesinden kaynaklanmaktadır.

$$\text{Derece cinsinden açı bilgisi} = \frac{180 * \alpha}{\pi} \text{ elde edilir}$$



Şekil 3-4: Eksene göre simetrik parçalar



Şekil 3-5: İşaretlenmiş ve açıları hesaplanmış imgeler

### 3.1.4 K-means ile poz kümelerinin oluşturulması

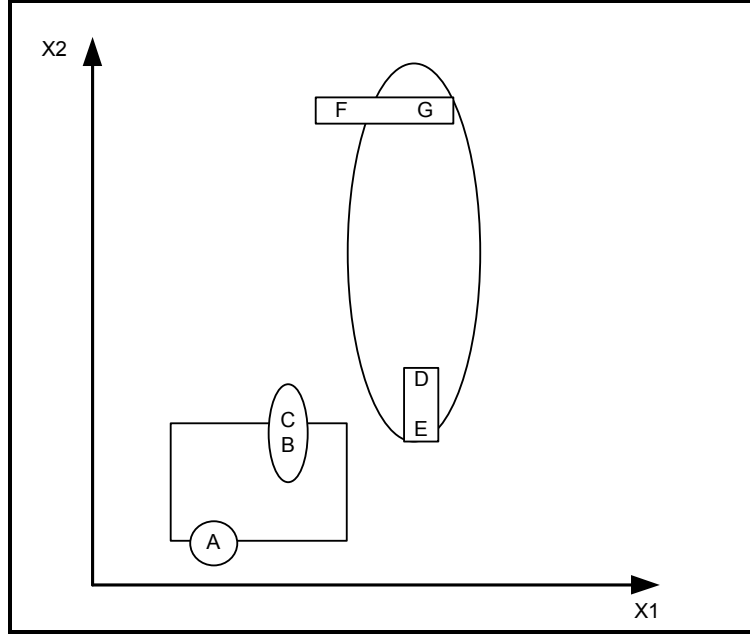
K-means verileri sınıflandırmaya yarayan bir yönsüz sınıflandırma algoritmasıdır. Verilen nesnelere özelliklerine göre  $k$  adet sınıfa ayırmak amacıyla kullanılır. Kümeleme yaparken her veriyi küme merkezlerine yakınlıklarına göre sıralar ve en yakın olduğu küme merkezinin sınırları içerisine yerleştirir. Küme merkezleri ve sınıflandırılmak istenen veri için aradaki uzaklık Öklid bağıntısı ile hesaplanır (MacQueen, 1967).

Algoritma önce rastsal olarak bir partiyon seçer ve bu partiyonu örnek ve küme merkezleri arasında bir örüntü oluşana dek tekrar tekrar eşleştirmeye devam eder. Bu işlem yakınsama koşulları sağlanan dek devam eder. Yakınsama koşulları;

- Kümeler arasında yeni bir eşleşme yapılamaması
- İterasyonlardan sonra hataların karesi sonucunun aşırı düşüş göstermesi

K-means algoritmasının popüler olmasını sebebi kolay uygulanabilir olması ve karmaşıklığının  $O(n)$  olmasında kaynaklanmaktadır. Algoritma ile ilgili önemli bir problem ise rastsal seçilen ilk kümenin atanmasından doğan farklılıklara aşırı duyarlı oluşudur. **Şekil 3-6**'ya bakılırsa eğer ilk noktalar olarak A,B,C seçilirse oluşan ilk küme bu noktaları çerçeveler ve sonuçta  $\{\{A\}, \{B,C\}, \{D,E,F,G\}\}$  elde edilir. Fakat A, D ve F ilk noktalar olarak alınırsa  $\{\{A,B,C\}, \{D,E\}, \{F,G\}\}$  kümelemesi yapılmış olur (Jain, et al., 1999). Yaptığımız çalışmalar da ardı ardına farklı kümeleme sonuçları almamak için yazılan uygulamalarda k-means algoritmasını hangi noktaları yerleştirmeye başlayacağı algoritmik olarak verilmiştir.





Şekil 3-6: K-means ilk atanan kümeye göre farklı sonuçlar verir

### 3.1.4.1 K-means algoritması adımları

K-means algoritmasının birçok farklı uygulaması mevcuttur. Başlangıç noktalarının seçimlerine göre değişik yaklaşımlar görülebilmektedir. Fakat genel mantık şu şekilde verilebilir (Jain, et al., 1999)

- (1) k adet küme merkezi rastgele ya da ön tanımlı olarak seçilir
- (2) Kümelenmesi istenen her nokta kendisini en yakın/benzer küme merkezi ile eşleştirilir
- (3) Yeni oluşan küme komşulukları göz önüne alınarak küme merkezler tekrar hesaplanır

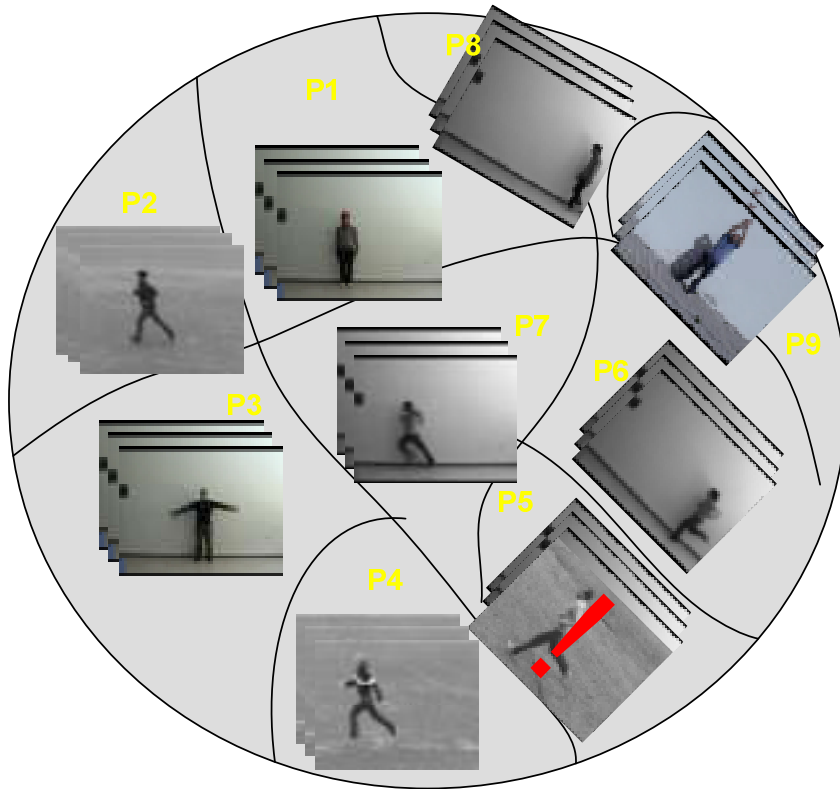
- (4) Yakınsama ölçütleri sağlanamadı ise 2.adıma geri dönülür. Tipik yakınsama koşulları
- a. Küme komşuluklarının değişmesi ya da çok az değişmesi
  - b. Hataların karesi sonucunda çok küçük sonuçların elde edilmesi

### 3.1.4.2 Pozlandırıcı için küme sayısının çıkarılışı

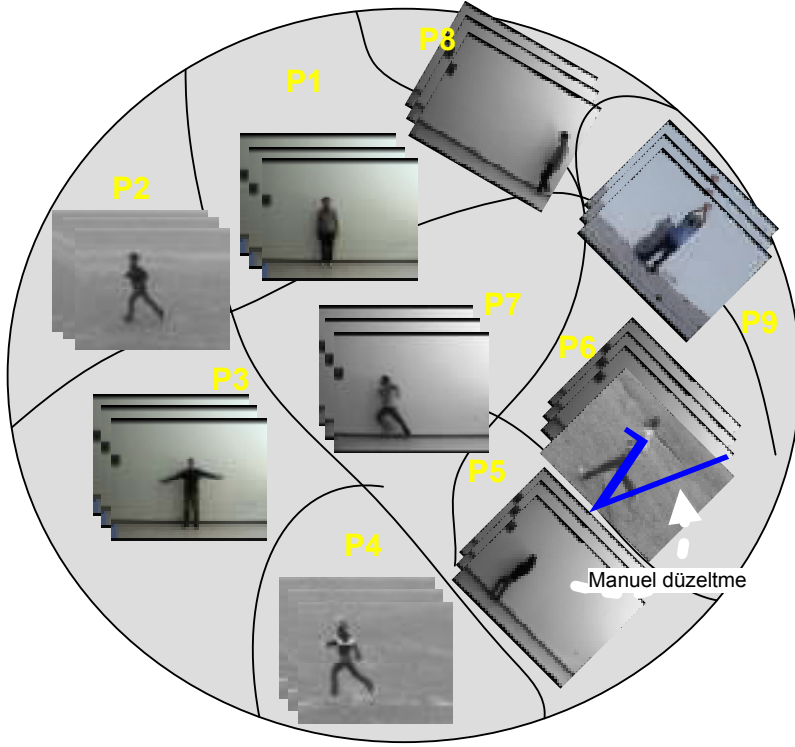
K-means ile işaretlenen imgelerin sınıflandırılması yapılmadan önce kaç adet poz kümesine ayrılması gerektiği ortaya konulmalıdır. Buradaki *k-küme sayısı* aynı zamanda hareketlerde kaç farklı poz olduğunu da gösterir. Hareketler için küme sayıları çıkarılırken bütün pozlar incelenmiş ve hareket türüne göre kaç farklı poz olabileceği belirlenmiştir. Buna göre el sallama hareketi için hareket aralığı ellerin yanda bitişik halde duruşundan yukarıda birleşmiş olmasına kadardır. Bu hareket aralığı içerisinde yaklaşık 10 farklı poz alabilmektedir. Koşma ve yürüme hareketlerinde ise hareket aralığı kişinin resim çerçevesine tamamen girip kol ya da ayağının çerçeve sınırına ilk kez değdiği aralık olarak belirlenmiştir. Burada el ve kolların hareket kombinasyonlarına bakıldığında 40'a yakın farklı poz olabileceği fark edilmiştir.

Pozlama işlemi bittiğinde gözle doğruluk sınaması yapılmıştır ve tek elemanlı kümelere düşen imgeler var ise bunlar benzeri başka imgelerin olduğu kümelere aktarılmıştır. Aynı zamanda birbiri ile hiç ilgisi olmayan imgeler aynı kümeye yerleştirilmiş ise bunlar benzeri imgelerin olduğu kümelere elle yerleştirilmiştir. El sallama hareketlerinin neredeyse tamamı kümelere uygun bir şekilde yerleşmesine rağmen, koşma ve yürüme hareketlerinde kimi zaman birbirleri ile ilgisi olmayan imgeler aynı kümelerde yerleştirilmiştir. Hatalı poz yerleşimlerinin bir kısmında kol ve ayak işaretlemelerinde yanlışlıkla arkadaki parçaların işaretlendiği bir kısmında ise işaretlerin merkez noktalarından uzağa konulduğu fark edilmiştir. Ayrıca ardı ardına gelen imgelerde noktaların bazen fazla kaydırılması da söz konusudur. Anlaşıldığı üzere

iřaretleme iřleminde ufak kaydırmalar dahi k-means kümelemesinde farklı kümelerin oluşmasına yol açabilmektedir. Şekil 3-7’de kümeleme iřlemi sembolik olarak gösterilmiştir.



Şekil 3-7: Kümeleme iřleminin sembolik gösterimi



Şekil 3-8: Kümeleme sonrası elle düzeltme

Test için kullanılan imgelerde işaretlendikten sonra aynı süreçten geçirilmiştir. Fakat bu imgeler için yukarıda anlatılan k-means kümelemesi yerine eğitim verileri için yapılan kümeleme işleminin sonucu olan kümelerin merkez noktaları verilmiştir. Bu şekilde eğitim ve test setleri üzerinde eşleme sağlanmıştır.

Çizelge 7: Kümeleme işleminde kullanılan veri yapısı

Değişken ismi	Açıklama	Büüklük
Angle	Gövde ile yapılan açılar	$[10 \times 1]^2$
personID	Hareketi yapan kişi	$[10 \times 20]$
actionType	Hareket türü	$[10 \times 1]$
frameNo	Kümelenen imgenin imge dizisindeki yeri	$[1 \times 1]$
clusterID	Yerleştirilen küme	$[1 \times 1]$
clusterNo	Maksimum küme sayısı	$[1 \times 1]$
pointsMarked	İşaretlenen noktaların kartezyen koordinatları	$[10 \times 2]$

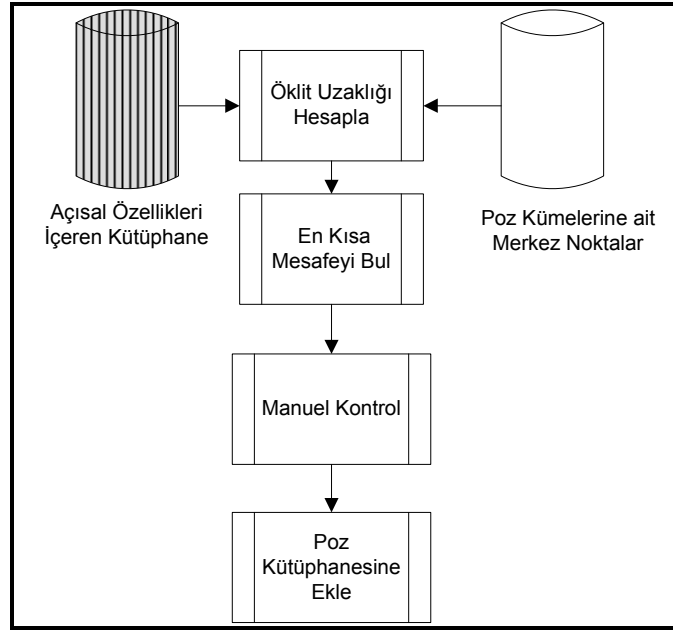
### 3.1.4.3 Test gruplarının pozlandırıcıdan geçirilmesi

**Şekil 1-15: Sisteme ayrıntılı genel bakış'** ta gösterildiği üzere test görüntülerinin pozlandırıcıda doğru pozlarla etiketlenmeleri için eğitim grubundaki k-means parametreleri ile işleme tabi tutulmaları gerekmektedir.

K-means ile T1 grubu kümeleme işlemine sokulduktan sonra her poz kümesine ait merkez noktaları saklanmıştır. Test grupları poz kümelerine yerleştirilecekleri vakit her bir açısal özellik vektörü ile poz kümesi merkezi arasındaki Öklid mesafesi ölçülür. Test grubundaki veri hangi poz kümesine en yakın ise o kümeye eklenir.

---

<sup>2</sup> Koşma ve yürüme hareketlerinde bu alan  $[6 \times 1]$ 'dir



Şekil 3-9: Test gruplarının pozlanmasına ait akış diyagramı

#### 3.1.4.4 Poz dizilimlerinin oluşturulması

Bütün test ve eğitim grupları pozlandırıcıdan geçirildikten sonra poz dizilimlerinin oluşturulmasına ihtiyaç vardır. Bu işlem için Matlab ortamında bir uygulama yazılmıştır. Uygulama her veri grubunu parametre olarak aldıktan sonra bir kişinin aynı hareketi yaparken çekilen her görüntüsüne ait poz bilgisini bir satır vektör yapısına ekler. Bu işlem her hareket grubu ve her kişi için tekrarlandığında poz dizilimleri elde edilmiş olur.

Sonuçta ortaya her veri grubundaki hareket tiplerine göre poz dizilim kütüphaneleri oluşmuş olur. **Çizelge 8: Poz dizilimi kütüphanesinin veri gruplarına göre dağılımı** tablosunda genel dağılım görülmektedir. Her insanın belli bir hareketi

yapma hızı diğerlerinde farklı olabileceği için kişi bazında poz dizilimi vektörlerinin uzunlukları arasında farklılar olabilmektedir. Bu durumlarda poz dizilim vektörü daha küçük olan ya da bir diğer deęişle hareketi hızlı yapan kişilerin poz dizilim vektörlerinin son kısmına en uzun vektörün boyuna eşitlemek için 0 ya da tekrarlanan poz dizisi eklenmiştir.

Çizelge 8: Poz dizilimi kütüphanesinin veri gruplarına göre dağılımı

Veri Grubu	Hareket Türüne Göre Poz Dizilim Kütüphanesi		
	El Sallama	Yürüme	Koşma
T1 <sup>3</sup>	20x32	20x32	21x32
T2 <sup>4</sup>	20x30	20x31	21x15
T3 <sup>5</sup>	5x23	5x60	5x30
T2+T3 <sup>6</sup>	76x60		

---

<sup>3</sup> T1 veri grubu eğitim grubu olup 20 kişiye ait hareketleri içerir

<sup>4</sup> T2 veri grubu T1 grubunda olup eğitimde kullanılmayan 20 kişiye ait imgelerden oluşmuştur

<sup>5</sup> T3 veri grubu eğitimde kullanılmamış olup 5 kişiye ait hareketi içerir

<sup>6</sup> T2+T3 veri grubu T2 ve T3'ün bütün hareketlerinin birleşiminden oluşmuştur

## 4 DENEYSEL ÇALIŞMALAR

Bilgisayar tabanlı örüntü tanıma uygulamalarında hareket türünü anlamak için sınıflandırıcıların kullanımı yaygın bir yöntemdir. Bu sınıflandırıcılardan bir kısmı ikili sınıflandırıcı, *binary classifier*, bir kısmı da çoklu sınıflayıcı, *multi class classifier*, olarak kullanılmaktadır. Makine tabanlı statiksel modellerinde gelişmesi ile birlikte birçok çalışmada kendine yer bulmuştur.

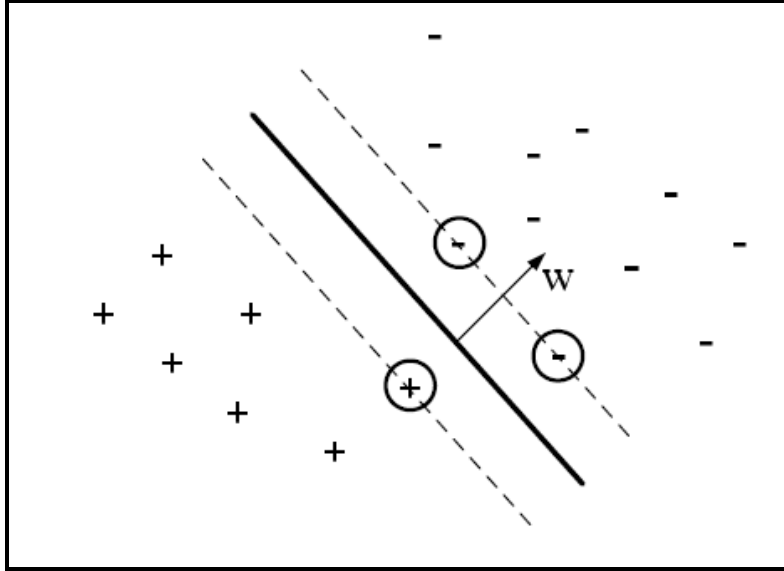
Çalışmanın hareketi tanıma bölümünde Destek Vektör Makineleri ve Yapay Sinir Ağları olmak üzere 2 farklı sınıflandırıcı kullanılmıştır.

### 4.1 Destek Vektör Makineleri

Destek vektör makineleri, ikili sınıflandırma ve gerçel değerli fonksiyon yaklaşılama (bağlanım kestirimi [regression estimation]) için kullanılan bir öğreticiyle öğrenme (supervised learning) yöntemleri kümesi olarak tanımlanabilir.

Diğer bir deyişle, etiketlenmiş eğitim verilerini kullanarak sonraki örneklemelerinin hangi sınıfa ait olduğunu söyler. Destek vektör makinesi ile sınıflandırma, eldeki veri kümesini iki ya da daha çok bölüme ayıran d-boyutlu bir aşırı düzlem (hyperplane) oluşturularak gerçekleştirilir. Temelde ikili sınıflandırıcı olan destek vektör makineleri ile çoklu sınıflandırma problemi ise veri kümesinin farklı biçimlerde ikiye bölünmesiyle çözülür. Örüntü tanıma problemlerinde aşırı düzlem, eğitim kümesindeki pozitif ve negatif örnekleri ayırma işlevini görür. Verilerin dağılımlarına göre birden fazla ayırıcı düzlem olabilir. Bu durumda, kendisine en yakın pozitif ve negatif verinin uzaklıkları en büyük olan ayırıcı düzlem seçilir. Ayırıcı düzleme en yakın pozitif ve negatif veriler "destek vektörleri" olarak adlandırılır ve karar sınırını belirler (Burges, 1998).





Şekil 4-1: Ayırıcı düzlem ve destek vektörleri

**Şekil 4-1: Ayırıcı düzlem ve destek vektörleri** gösterilmektedir. Kalın ve dik çizgi ayırıcı düzlemi, kesikli çizgiler ise destek vektörlerini verir. (+) ve (-) ise ayrılmak istenen sınıfları gösterir. Veri kümesindeki vektörleri ifade eden değişken  $x$ , bağıl konum (ofset) değeri  $b$  ve ayırıcı düzlemin normal vektörü  $w$  olmak üzere ayırıcı düzlem,  $w \cdot x + b = 0$  doğrusu ile ifade edilebilir.  $n$  adet örnek içeren eğitim kümesindeki sınıflarına (pozitif ya da negatif) göre işaretlenmiş veriler (vektörler) kullanılarak en iyi ayırıcı düzlem,  $x_i$   $d$ -boyutlu eğitim vektörü,  $y_i$  bu vektörün sınıfı ve  $b$  de ofset değeri olmak üzere, aşağıdaki ifadeyle elde edilebilir:

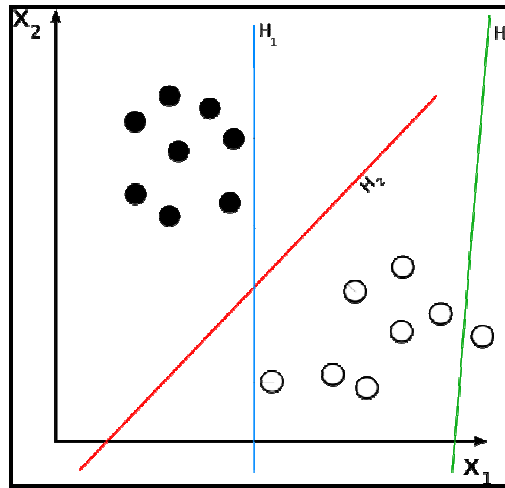
$$y_i(w \cdot x_i + b) - 1 \geq 0, \quad i = 1, 2, \dots, n, \quad x_i \in \mathbf{R}^d, y_i \in \{-1, 1\} \quad (4-1)$$

Burada, en yakın eğitim örneği ile arasındaki uzaklık (margin) değeri  $2 / \|w\|$  en büyük olan düzlem, ayırıcı düzlem olarak belirlenir. Ayırıcı düzlemi belirlenmiş (eğitilmiş) bir destek vektör makinesi için, verilen yeni bir verinin sınıflandırılması ( $y_i$ 'nin bulunması) ise şu denklemle gerçekleştirilebilir (Muller, et al., 2001).

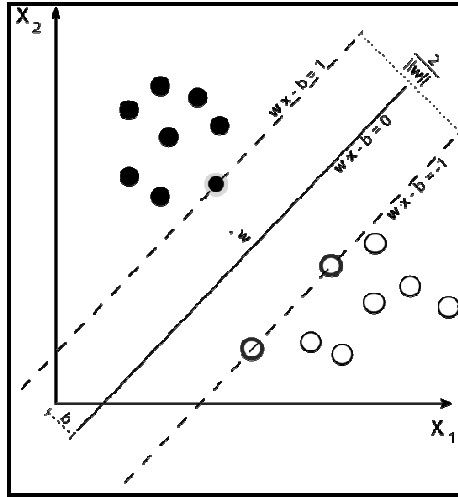
$$y_i = \text{sgn}(w \cdot x_i + b) \quad (4-2)$$

Doğrusal destek vektör makinelerinde en iyi ayırıcı düzlemi bulmak için en yakın eğitim örneğine olan uzaklığına bakılır. Doğrusal olarak sınıflandırılmayan veriler için ise, verilerin doğrusal olarak sınıflandırılabilmesi ve iç çarpımın tanımlı olduğu başka bir uzaya dönüşüm için çeşitli çekirdek fonksiyonları tanımlanmaktadır. Doğrusal çekirdek fonksiyonu aşağıdaki gibi ifade edilmektedir.

$$K(x_i, x_j) = (x_i \cdot x_j) \quad (4-3)$$



Şekil 4-2: Birden fazla aşırı düzlem olma durumu

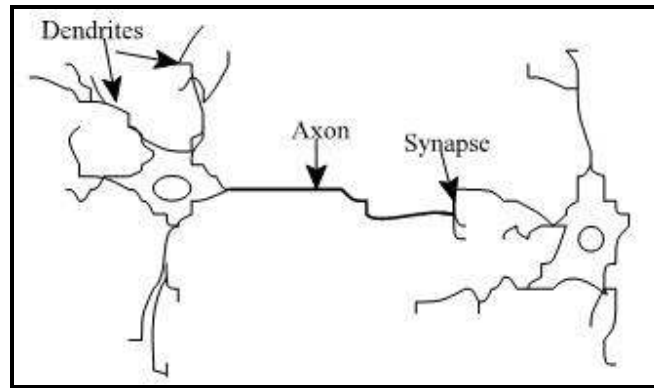


Şekil 4-3: Destek Vektörleri.

## 4.2 Yapay Sinir Ağları

Yapay sinir ağları (YSA), insan beyninin bilgi işleme teknolojisinden esinlenerek geliştirilmiş bir bilgi işlem teknolojisidir. YSA'nın temelleri karmaşık problemlerinin çözümünde kullanılan "parçala ve fethet" ilkesi üzerine kurulmuştur (Gershenson, 2003). YSA ile basit biyolojik sinir sisteminin çalışma şekli simüle edilir (benzetilir). Simüle edilen sinir hücreleri nöronlar içerirler ve bu nöronlar çeşitli şekillerde birbirlerine bağlanarak ağı oluştururlar. Bu ağlar öğrenme, hafızaya alma ve veriler arasındaki ilişkiyi ortaya çıkarma kapasitesine sahiptirler. Diğer bir ifadeyle, YSA'lar, normalde bir insanın düşünme ve gözlemlemeye yönelik doğal yeteneklerini gerektiren problemlere çözüm üretmektedir. Bir insanın, düşünme ve gözleme yeteneklerini gerektiren problemlere yönelik çözümler üretebilmesinin temel sebebi ise insan beyninin ve dolayısıyla insanın sahip olduğu yaşayarak veya deneyerek öğrenme yeteneğidir.

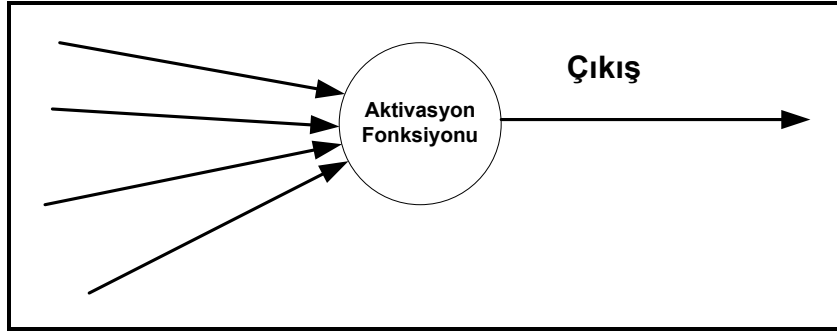
Biyolojik sistemlerde öğrenme, nöronlar arasındaki sinaptik(synaptic) bağlantıların ayarlanması ile olur. Yani, insanlar doğumlarından itibaren bir yaşayarak öğrenme süreci içerisine girerler. Bu süreç içinde beyin sürekli bir gelişme göstermektedir. Yaşayıp tecrübe ettikçe sinaptik bağlantılar ayarlanır ve hatta yeni bağlantılar oluşur. Bu sayede öğrenme gerçekleşir. Bu durum YSA için de geçerlidir. Öğrenme, eğitime yoluyla örnekler kullanarak olur; başka bir deyişle, gerçekleşme girdi/çıkı verilerinin işlenmesiyle, yani eğitime algoritmasının bu verileri kullanarak bağlantı ağırlıklarının (weights of the synapses) bir yakınsama sağlanana kadar, tekrar tekrar ayarlanmasıyla olur. **Şekil 4-4: Sinir hücresi gösterimi**'nde insana ait bir sinir hücresinin yapısı incelenebilir.



Şekil 4-4: Sinir hücresi gösterimi

YSA'lar, ağırlıklandırılmış şekilde birbirlerine bağlanmış birçok işlem biriminden (nöronlar) oluşan matematiksel sistemlerdir. Bir işlem birimi, aslında sık sık transfer fonksiyonu olarak anılan bir denklemdir. Bu işlem birimi, diğer nöronlardan sinyalleri alır; bunları birleştirir, dönüştürür ve sayısal bir sonuç ortaya çıkartır. Genelde, işlem birimleri kabaca gerçek nöronlara karşılık gelirler ve bir ağ içinde birbirlerine bağlanırlar;

bu yapı da sinir ağlarını oluşturmaktadır.Şekil 4-5: Yapay sinir hücresi gösterimi'nde bir yapay sinir hücresinin gösterimi incelenebilir.



Şekil 4-5: Yapay sinir hücresi gösterimi

Sinirsel (neural) hesaplamanın merkezinde dağıtılmış, adaptif ve doğrusal olmayan işlem kavramları vardır. YSA'lar, geleneksel işlemcilerden farklı şekilde işlem yapmaktadırlar. Geleneksel işlemcilerde, tek bir merkezi işlem birimi her hareketi sırasıyla gerçekleştirir. YSA'lar ise her biri büyük bir problemin bir parçası ile ilgilenen, çok sayıda basit işlem birimlerinden oluşmaktadır. En basit şekilde, bir işlem birimi, bir girdiyi bir ağırlık kümesi ile ağırlıklandırır, doğrusal olmayan bir şekilde dönüşümünü sağlar ve bir çıktı değeri oluşturur. İlk bakışta, işlem birimlerinin çalışma şekli yanıtıcı şekilde basittir. Sinirsel hesaplamanın gücü, toplam işlem yükünü paylaşan işlem birimlerinin birbirleri arasındaki yoğun bağlantı yapısından gelmektedir. Bu sistemlerde geri yayılım metoduyla daha sağlıklı öğrenme sağlanmaktadır.

Çoğu YSA'da, benzer karakteristiğe sahip nöronlar tabakalar halinde yapılandırılırlar ve transfer fonksiyonları eş zamanlı olarak çalıştırılırlar. Hemen hemen tüm ağlar, veri alan nöronlara ve çıktı üreten nöronlara sahiptirler.

YSA'nın ana ögesi olan matematiksel fonksiyon, ağın mimarisi tarafından şekillendirilir. Daha açık bir şekilde ifade etmek gerekirse, fonksiyonun temel yapısını ağırlıkların büyüklüğü ve işlem elemanlarının işlem şekli belirler. YSA'ların davranışları, yani girdi veriyi çıktı veriye nasıl ilişkilendirdikleri, ilk olarak nöronların transfer fonksiyonlarından, nasıl birbirlerine bağlandıklarından ve bu bağlantıların ağırlıklarından etkilenir.

Yapay sinir ağlarının üstünlüklerinin yanı sıra bazı sakıncaları da vardır. Bu sakıncalar şu şekilde listelenebilir:

- Sistem içerisinde ne olduğu bilinemez.
- Bazı ağlar hariç kararlılık analizleri yapılamaz.
- Farklı sistemlere uygulanması zor olabilir.

### 4.3 Örüntü Algılamada Performans Ölçüm Kriterleri

Örüntü tanıma uygulamalarında kullanılan sınıflandırıcıların performansları doğrulukları, *accuracy*, ile ölçülür. Doğruluk ise eğitim verilerinde doğru ya da pozitif olarak etiketlenen bir verinin test verilerinde doğru olarak bulunması ve eğitimde hatalı sınıfa ait olduğu belirtilen bir verinin test kümesinde hatalı sınıfa ait olarak şekilde sınıflandırılmasıdır. Sınıflandırıcı performanslarında hatalı sınıflandırmalardan

bahsederken 2 türlü hatalı sınıflandırmanın birbirlerinden ayrılması çok önemlidir. Örnek vermek gerekirse, kanser tümörü tetkikinde kullanılan bir MRI cihazında hatalı negatif değerlerden kaçınmak hayati önemlidir fakat az sayıdaki hatalı pozitif sınıflandırmalar personel tarafından ardı ardına izlenerek tolere edilebilir (Veropoulos, et all., 1999).

Yukarıda anlatıldığı üzere sınıflandırıcı performansları değerlendirirken hatalı negatif, hatalı pozitif, doğru pozitif, hatalı pozitif değerlerinin dağılımı çok önemlidir. Bu oranlar birleşerek hata matrislerini, confusion matrice, oluştururlar.

Hatalı pozitif (FP) : Bir değeri negatif iken pozitif etiketlenirse

Doğru pozitif(TP) : Bir değer pozitif iken pozitif etiketlenirse

Hatalı negatif(FN) : Bir değer pozitif iken negatif etiketlenirse

Doğru negatif(TN) : Bir değer negatif iken negatif etiketlenirse

Çizelge 9: Hata matrisi gösterimi

Hata Matrisi			
Eğitim	Test		
		P	N
	P	TP	FP
	N	FN	TN

**Çizelge 9'** daki notasyon aracılığı ile pozitif tahminleri TP + FN, yanlış tahminleri TN+FP, veri kümesindeki bütün örnekleme sayısını ise TP+FN+FN+TN olarak göstermek mümkündür.

Bu notasyonun yardımı ile duyarlılık, *sensitivity*, ve özgüllük, *specificity*, oranları tanımlanmıştır. Buna göre bir sınıflandırıcının duyarlılığı doğru pozitif tahminlerin test kümesindeki bütün doğru tahminlere oranıdır.

$$\text{Duyarlılık} = \frac{TP}{TP + FN}$$

Bir sınıflandırıcının özgüllüğü ise doğru negatif tahminlerin test kümesindeki negatif tahminlere oranıdır.

$$\text{Özgüllük} = \frac{TN}{TN + FP}$$

Doğruluk ise doğru sınıflandırılmış verilerin test kümesine oranıdır

$$\text{Doğruluk} = \frac{TP + TN}{TP + TN + FP + FN}$$

Bu çalışmada sınıflandırıcıların başarımları, burada verilmiş olan doğruluk, duyarlılık ve özgüllük oranları şeklinde verilecektir. Duyarlılık ve özgüllük oranları yüksek çıkan bir sınıflandırıcı yüksek başarımlı göstermiş anlamına gelecektir.



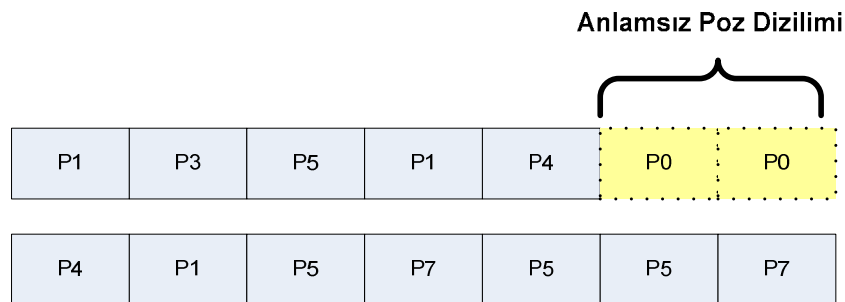
#### 4.4 Eğitim ve Test Verileri

Eğitim ve testler esnasında kullanılan kütüphane ve test ve eğitim gruplarının dağılımı Bölüm 2.1.5'e verilmiştir. Buna göre eğitim seti T1 grubu olarak isimlendirilmiştir. Test grubu ise T2, T3 ve T2 ile T3 bileşiminden oluşmaktadır.

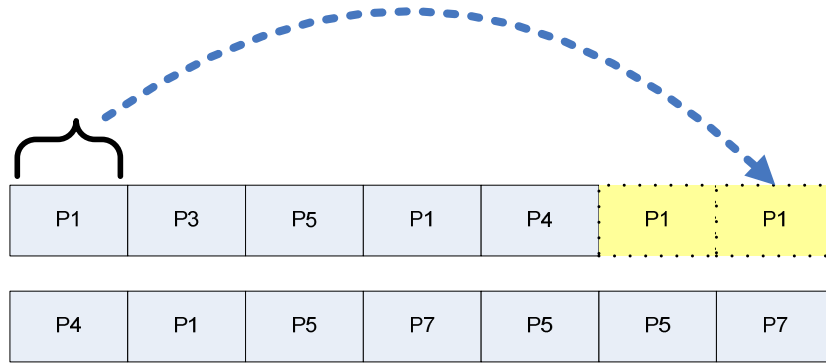
#### 4.5 Poz Dizilimlerinin Eğitimi ve Testi

Yapılan testlerde T1 grubu eğitim grubu, T2, T3, T2+T3 ise test grubu olarak kullanılmıştır. Eğitim ve test gruplarında eğer verilerin boyutları eşit değilse boyutları eşitlemek için 2 yaklaşım sergilenmiştir.

İlk yaklaşımda vektör boyutu küçük veri grubundaki dizilimlere poz dizilimlerinde olmayan bir ortak poz eklenmiştir (bkz: **Şekil 4-6: Ortak poz ekleme**) Diğer yaklaşımda ise sona ekleme yapmak için en baştaki poz diziliminden gerekli sayıda sona ekleme yapılmıştır. Bu yöntem ise tekrarlı poz ekleme ismi verilmiştir (bkz: **Şekil 4-7: Tekrarlı poz ekleme**).



Şekil 4-6: Ortak poz ekleme



Şekil 4-7: Tekrarlı poz ekleme

#### 4.6 Uygulama ve Test Detayları

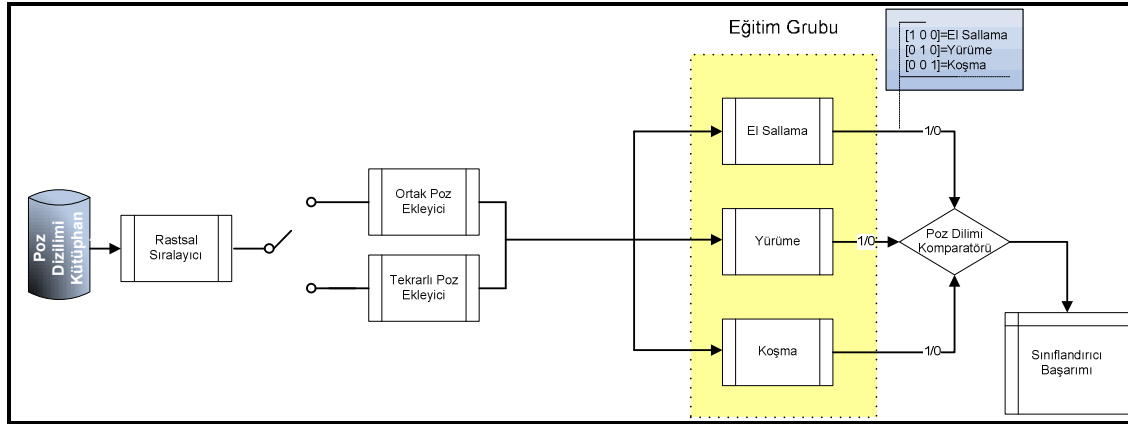
Sınıflandırıcı eğitimi ve testleri için Matlab ortamında farklı uygulamalar geliştirilmiştir. Bu uygulamalarda çeşitli eğitim ve test grubu kombinasyonlarını ele alınmıştır. Bunlar;

- 1:3 Sınıflandırma
  - Bir hareket türü için eğitilip 3 hareket türüne karşı teker teker test edilmesi
  - Kullanılan Yöntemler:SVM, NN
- 1:1 Sınıflandırma
  - Bir hareket türü için eğitilip bir hareket türüne karşı test edilmesi
  - Kullanılan Yöntemler:SVM, NN
- 3:3 Sınıflandırma
  - 3 hareket türü için eğitilip aynı anda 3 harekete göre test edilmesi

- Kullanılan Yöntem:NN
- 3:3 Sınıflandırma karıştırılmış poz dizilimi
  - T3 veri setindeki poz dizilimleri kişi bazında karıştırılarak poz dizilimlerinin hareket tanımadaki etkisinin doğrudan sınanması
  - Kullanılan Yöntem:NN
- 3:3 Sınıflandırma ters çevrilmiş poz dizilimi
  - T3 veri setindeki poz dizilimleri kişi bazında ters çevrilerek poz dizilimlerinin hareket tanımadaki etkisinin doğrudan sınanması
  - Kullanılan Yöntem:NN

#### 4.6.1 1:3 yapılan testler

Bu testlerde poz dizilimi kütüphanesinde ilk başta rastsal olarak sıralanmıştır. Buradaki amaç ardı ardına aynı hareket türüne ait örnekleri test etmektense karışık test etmektir. Rastsal sıralamadan sonra okunan hareket ve kişi bazındaki her bir örnek için vektör boyutu eğitim grubundaki vektör boyutuna ile eşitlenmiştir. Bu işlem için uygulama çalıştırılırken parametrik olarak verilen bir değişkenle ortak poz ya da tekrarlı poz seçiminin yapılması ile sağlanır. Önceden El Sallama, Koşma, Yürüme için eğitilmiş sınıflandırıcı ile test edilen test verisi bir karşılaştırıcıdan geçirilerek hangi hareketin sınıflandırıcısı ile daha yüksek eşleşme sağlamış ise bu hareket ile etiketlenir.

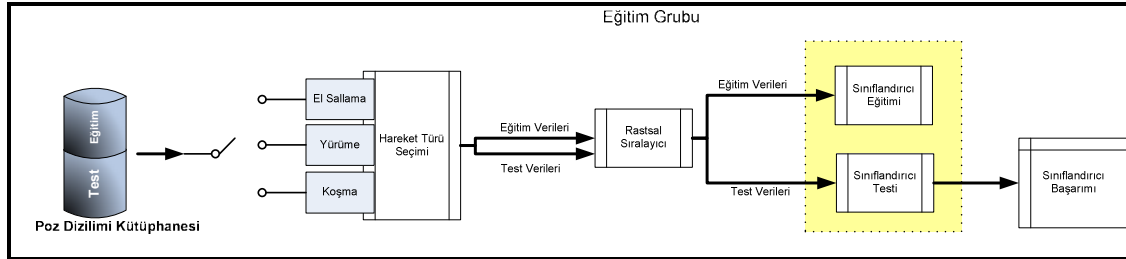


Şekil 4-8: 1:3 Test genel yapısı

Uygulamanın genel yapısı Şekil 4-8’de görülmektedir.

#### 4.6.2 1:1 yapılan testler

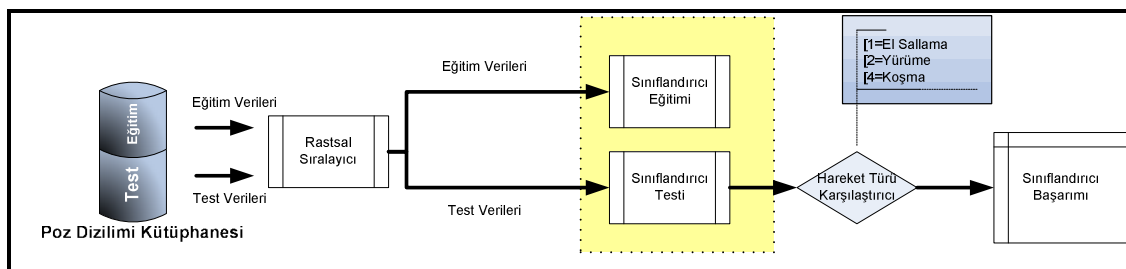
1:1 testlerde poz dizilimi kütüphanesinden bir hareket türü eğitim için, aynı hareket türündeki test verileri ise test amaçlı olarak seçilmiştir. Kütüphaneden seçilen veriler daha sonra rastsal olarak kendi içlerinde gözlem sıraları değiştirildikten sonra eğitim verileri sınıflandırıcıda eğitilmiş ve test verileri de aynı sınıflandırıcı ile test edilmiştir.



Şekil 4-9: 1:1 Test genel yapısı

#### 4.6.3 3:3 yapılan testler

Bu test aynı zamanda çok sınıflı sınıflandırıcı testidir. Yöntem olarak ise NN yöntemi kullanılmıştır. Poz dizilimi kütüphanesinden seçilen örnekler 3 hareket tipi için aynı zamanda eğitilirler. Eğitim esnasında El Sallama, Yürüme ve Koşma hareketleri ayrı etiketler ile işaretlenir, *yönlü eğitim*. Test grubundan rastsal olarak seçilen test verileri ise sınıflandırıcıya karşı test edilir. Test sonuçları her bir hareketin etiketi için verilen sayı limit alınacak şekilde bir karşılaştırıcıdan geçirilir ve başarı oranları kayıt altına alınır.



Şekil 4-10: Çoklu sınıflandırıcı genel yapısı

#### 4.6.4 3:3 karışık ve ters çevrilmiş poz dizilimi testleri

Bu testler 3:3 testleri ile aynı şekilde yapılan testlerdir. Fakat burada amaçlana poz dizilimlerindeki sıralamanın hareket algılama üzerindeki doğrudan etkisini sınamak amaçlıdır. Beklenen sonuçlar algılama oranının düşmesi yönündedir. Çünkü poz dizilimlerinin sırası algılamada önem arz eder. Sıralamanın önemsiz olması BoW modelinin uygulandığı durumlar için geçerlidir. Bütün hareketler için kişi bazındaki poz dizilimleri rastgele olarak karıştırılıp test edilmiştir. Daha sonra aynı test poz diziliminin tamamen ters çevrilmesi ile tekrar denenmiş ve sonuçlar kayıt altına alınmıştır.

#### 4.6.5 Test başarımlar oranları

Testlerdeki başarımlar oranları ortak poz ekleme ve tekrarlı poz ekleme yöntemleri için her sınıflandırıcıda ayrı test edilmiş ve sonuçları kayıt edilmiştir. Sınıflandırıcıların başarımları incelenirken doğruluk, duyarlılık ve özgüllük değerleri verilmiştir. Ayrıca hata matrisleri de çıkartılmıştır.

##### 4.6.5.1 1:1 testleri başarımlar oranları

Hata matrisindeki dizilim  $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$  şeklindedir

Ortak Poz Ekleme Yöntemi

Çizelge 10: 1:1 testi T2 veri seti için NN ile ortak poz ekleme başarımları

TEST SETİ	T2					
SINIFLANDIRICI	NN					
	El Sallama		Yürüme		Koşma	
<b>Doğruluk</b>	100%		96,7%		98,3%	
<b>Duyarlılık</b>	1		0,97		1	
<b>Özgüllük</b>	1		0,95		0,95	
	El Sallama		Yürüme		Koşma	
<b>Hata Matrisi</b>	41	0	40	1	40	0
	0	20	1	19	1	20

Çizelge 11: 1:1 testi T2+T3 veri seti için NN ile ortak poz ekleme başarımları

TEST SETİ	T2+T3					
SINIFLANDIRICI	NN					
	El Sallama		Yürüme		Koşma	
<b>Doğruluk</b>	100%		92,1%		96%	
<b>Duyarlılık</b>	1		0,98		0,96	
<b>Özgüllük</b>	1		0,8		0,96	
	El Sallama		Yürüme		Koşma	
<b>Hata Matrisi</b>	51	0	50	1	48	2
	0	25	5	20	1	25

Çizelge 12: 1:1 testi T3 veri seti için NN ile ortak poz ekleme başarımları

TEST SETİ	T3					
SINIFLANDIRICI	NN					
	El Sallama		Yürüme		Koşma	
Doğruluk	100%		100%		100%	
Duyarlılık	1		1		1	
Özgüllük	1		1		1	
	El Sallama		Yürüme		Koşma	
Hata Matrisi	10	0	10	0	10	0
	0	5	0	5	0	5

Çizelge 13: 1:1 testi T2 veri seti için SVM ile ortak poz ekleme başarımları

TEST SETİ	T2					
SINIFLANDIRICI	SVM					
	El Sallama		Yürüme		Koşma	
Doğruluk	100%		96%		98,3%	
Duyarlılık	1		0,97		1	
Özgüllük	1		0,95		0,97	
	El Sallama		Yürüme		Koşma	
Hata Matrisi	41	0	40	1	39	1
	0	20	1	19	0	21



Çizelge 14: 1:1 testi T2+T3 veri seti için SVM ile ortak poz ekleme başarımları

TEST SETİ	T2+T3					
SINIFLANDIRICI	SVM					
	El Sallama		Yürüme		Koşma	
<b>Doğruluk</b>	100%		90,7%		92,1%	
<b>Duyarlılık</b>	1		0,89		1	
<b>Özgüllük</b>	1		0,95		0,81	
	El Sallama		Yürüme		Koşma	
<b>Hata Matrisi</b>	51	0	50	1	44	6
	0	25	6	19	0	26

Çizelge 15: 1:1 testi T3 veri seti için SVM ile ortak poz ekleme başarımları

TEST SETİ	T3					
SINIFLANDIRICI	SVM					
	El Sallama		Yürüme		Koşma	
<b>Doğruluk</b>	100%		100%		100%	
<b>Duyarlılık</b>	1		1		1	
<b>Özgüllük</b>	1		1		1	
	El Sallama		Yürüme		Koşma	
<b>Hata Matrisi</b>	10	0	10	0	10	0
	0	5	0	5	0	5

Çizelge 10, Çizelge 11 ve Çizelge 12 sonuçlarına bakıldığında NN yöntemi ortak poz ekleme durumunda el sallama hareketinde tam başarımlar göstermektedir. Yürüme ve

koşma hareketlerinde başarımın daha düşük olması bu hareketlerde örtüşebilen pozlar olması ile sınıflandırıcının hatalı sınıflandırmasından kaynaklanmaktadır. T3 veri setinde tüm hareketlerin doğru algılanması ise bu veri setinde az sayıda poz dizilimi olmasından kaynaklanmaktadır.

**Çizelge 13, Çizelge 14 ve Çizelge 15** SVM başarım oranları açısından değerlendirildiğinde ise el sallama hareketi yine tam başarıyla ayırt edilmiştir. Yürüme ve koşma hareketleri açısından bakıldığında SVM, NN'e göre daha az başarılıdır. Burada ki fark ise SVM'in başarımının eğitim grubunun büyüklüğünden daha çok etkileniyor olmasındandır.

Tekrarlı Poz Ekleme Yöntemi

Çizelge 16: 1:1 testi T2 veri seti için NN ile tekrarlı poz ekleme başarımları

TEST SETİ	T2					
SINIFLANDIRICI	NN					
	El Sallama		Yürüme		Koşma	
<b>Doğruluk</b>	100%		98,3%		98,3%	
<b>Duyarlılık</b>	11		0,97		0,97	
<b>Özgüllük</b>	1		1		1	
	El Sallama		Yürüme		Koşma	
<b>Hata Matrisi</b>	41	0	40	1	40	0
	0	20	0	20	1	20

Çizelge 17: 1:1 testi T2+T3 veri seti için NN ile tekrarlı poz ekleme başarımları

TEST SETİ	T2+T3					
SINIFLANDIRICI	NN					
	El Sallama		Yürüme		Koşma	
<b>Doğruluk</b>	100%		98,6%		100%	
<b>Duyarlılık</b>	1		0,98		1	
<b>Özgüllük</b>	1		1		1	
	El Sallama		Yürüme		Koşma	
<b>Hata Matrisi</b>	51	0	50	1	50	0
	0	25	0	25	0	26

Çizelge 18: 1:1 testi T3 veri seti için NN ile tekrarlı poz ekleme başarımları

TEST SETİ	T3					
SINIFLANDIRICI	NN					
	El Sallama		Yürüme		Koşma	
<b>Doğruluk</b>	100%		100%		100%	
<b>Duyarlılık</b>	1		1		1	
<b>Özgüllük</b>	1		1		1	
	El Sallama		Yürüme		Koşma	
<b>Hata Matrisi</b>	10	0	10	0	10	0
	0	5	0	5	0	5

Çizelge 19: 1:1 testi T2 veri için SVM ile tekrarlı poz ekleme başarımları

TEST SETİ	T2					
SINIFLANDIRICI	SVM					
	El Sallama		Yürüme		Koşma	
<b>Doğruluk</b>	100%		75,4%		98,3%	
<b>Duyarlılık</b>	1		0,76		0,97	
<b>Özgüllük</b>	1		0,72		1	
	El Sallama		Yürüme		Koşma	
<b>Hata Matrisi</b>	41	0	38	3	40	0
	0	20	12	8	1	20

Çizelge 20: 1:1 testi T2+T3 veri seti için SVM ile tekrarlı poz ekleme başarımları

TEST SETİ	T2+T3					
SINIFLANDIRICI	SVM					
	El Sallama		Yürüme		Koşma	
<b>Doğruluk</b>	100%		76,3%		98,6%	
<b>Duyarlılık</b>	1		0,77		0,98	
<b>Özgüllük</b>	1		0,73		1	
	El Sallama		Yürüme		Koşma	
<b>Hata Matrisi</b>	51	0	47	4	50	0
	0	25	14	11	1	25

Çizelge 21: 1:1 testi T3 veri seti için SVM ile tekrarlı poz ekleme başarımları

TEST SETİ	T3					
SINIFLANDIRICI	SVM					
	El Sallama		Yürüme		Koşma	
<b>Doğruluk</b>	100%		80%		100%	
<b>Duyarlılık</b>	1		0,81		1	
<b>Özgüllük</b>	1		0,75		1	
	El Sallama		Yürüme		Koşma	
<b>Hata Matrisi</b>	10	0	9	1	10	0
	0	5	2	3	0	5

Tekrarlı poz eklemenin NN yönteminde etkisinin fazla olmadığı gözlenmektedir. Fakat bu etki SVM açısından bakıldığında dikkat çekicidir. **Çizelge 19**, **Çizelge 20**, **Çizelge 21**'e özellikle yürüme hareketi için bakılırsa ortak poz eklemeye kıyasla %18'e varan bir düşüş fark edilecektir. Bunu nedeni yine eğitim verisinin azlığına dayanmaktadır. Baştan alınıp sona eklenen poz test kümesindeki ilişim artırarak sınıflandırıcının başarımını düşürmektedir. Geçerli bir poz olmayan 0'ın poz olarak eklenmesi ise sınıflandırıcının işini kolaylaştırıp daha kolay ayırt edebilmesini mümkün kılmaktadır. Koşma hareketinin poz ekleme yönteminde yürüme hareketinde olduğu kadar çok etkilenmemesi ise bu hareket ait poz diziliminin daha kısa olmasıdır. Bu da gayet normaldir, çünkü koşma daha hızlı bir hareket olduğu için daha az sayıda görüntü ve poz oluşturmaktadır.

#### 4.6.5.2 1:3 testleri başarımları

Hata matrisindeki dizilim  $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$  şeklindedir

#### Ortak Poz Ekleme Yöntemi

Çizelge 22: 1:3 testi T2 veri seti ortak poz ekleme başarımları

TEST SETİ	T2			
SINIFLANDIRICI	NN		SVM	
<b>Doğruluk</b>	98,3%		98,3%	
<b>Duyarlılık</b>	0,99		0,98	
<b>Özgüllük</b>	0,96		0,98	
<b>Hata Matrisi</b>	121	1	120	2
	2	59	1	60

Çizelge 23: 1:3 testi T2+T3 veri seti ortak poz ekleme başarımları

TEST SETİ	T2+T3			
SINIFLANDIRICI	NN		SVM	
<b>Doğruluk</b>	94,3%		94,3%	
<b>Duyarlılık</b>	0,96		0,95	
<b>Özgüllük</b>	0,90		0,92	
<b>Hata Matrisi</b>	146	6	145	7
	7	69	6	70

Çizelge 24: 1:3 testi T3 veri seti ortak poz ekleme başarımları

TEST SETİ	T3			
SINIFLANDIRICI	NN		SVM	
Doğruluk	100%		100%	
Duyarlılık	1		1	
Özgüllük	1		1	
Hata Matrisi	30	0	30	0
	0	15	0	15

Tekrarlı Poz Ekleme Yöntemi

Çizelge 25: 1:3 testi T2 veri seti tekrarlı poz ekleme başarımları

TEST SETİ	T2			
SINIFLANDIRICI	NN		SVM	
Doğruluk	99,4%		91,2%	
Duyarlılık	1		0,97	
Özgüllük	1		0,78	
Hata Matrisi	121	1	119	3
	0	61	13	48

Çizelge 26: 1:3 testi T2+T3 veri seti tekrarlı poz ekleme başarımları

TEST SETİ	T2+T3			
SINIFLANDIRICI	NN		SVM	
Doğruluk	99,5%		91,6%	
Duyarlılık	1		0,97	
Özgüllük	1		0,80	
Hata Matrisi	151	1	148	4
	0	76	15	61

Çizelge 27: 1:3 testi T3 veri seti tekrarlı poz ekleme başarımları

TEST SETİ	T3			
SINIFLANDIRICI	NN		SVM	
Doğruluk	100%		93,3	
Duyarlılık	1		0,96	
Özgüllük	1		0,86	
Hata Matrisi	30	0	29	1
	0	15	2	13



**Çizelge 22, Çizelge 23 ve Çizelge 24** incelendiğinde 1:3 testlerinde ortak poz eklemenin SVM ve NN açısından aynı başarımlara yol açtığı görülmektedir.

**Çizelge 25, Çizelge 26 ve Çizelge 27** incelendiğinde ise tekrarlı poz eklemenin SVM sonuçlarını olumsuz etkilediği görülmektedir. T3 veri grubu bu etkiden T2'ye göre daha az etkilenmiştir. Bu da T3 deki poz dizilimlerinin daha az olmasından kaynaklanmaktadır. Bu açıdan bakıldığında 1:3 testleri sınıflandırıcı bazında 1:1 testleri ile paralellik sergiler.

#### 4.6.5.3 3:3 Testleri başarımları

Hata matrisindeki dizilim  $\begin{bmatrix} ElSallama = 1 \\ Yürüme = 2 \\ Kosma = 3 \end{bmatrix}$  şeklindedir.

#### Ortak Poz Ekleme Yöntemi

Çizelge 28: 3:3 testi ortak poz ekleme ile başarımları

SINIFLANDIRICI	NN								
	T2			T2+T3			T3		
TEST SETİ	T2			T2+T3			T3		
<b>Doğruluk</b>	100%			93,4%			100%		
<b>Duyarlılık</b>	1			1			1		
<b>Özgüllük</b>	1			1			1		
	T2			T2+T3			T3		
<b>Hata Matrisi</b>	20	0	0	25	0	0	5	0	0
	0	20	0	0	20	5	0	5	0
	0	0	21	0	0	26	0	0	5

Tekrarlı Poz Ekleme Yöntemi

Çizelge 29: 3:3 testi tekrarlı poz ekleme ile başarımları

SINIFLANDIRICI	NN								
	T2			T2+T3			T3		
TEST SETİ	T2			T2+T3			T3		
Doğruluk	100%			93,4%			100%		
Duyarlılık	1			1			1		
Özgüllük	1			1			1		
	T2			T2+T3			T3		
Hata Matrisi	20	0	0	25	0	0	5	0	0
	0	20	0	0	20	5	0	5	0
	0	0	21	0	0	26	0	0	5

Çizelge 28 ve Çizelge 29 incelendiğinde poz ekleme yönteminin NN üzerinde fark edilir bir etkisi olmadığı görülecektir. Burada farklılıkların NN ile tahmin edilen sınıfların etiketlenmesinde kullanılan komparatör için var olan eşik değerinin değiştirilmesi ile azaltılması mümkündür.

**4.6.5.4 3:3 Karıştırılmış poz dizilimi testi başarımları**

Hata matrisindeki dizilim  $\begin{bmatrix} ElSallama = 1 \\ Yürüme = 2 \\ Kosma = 3 \end{bmatrix}$  şeklindedir.

Ortak Poz Ekleme Yöntemi

Çizelge 30: 3:3 testi ortak poz eklemeli karıştırılmış poz dizilimi başarımları

SINIFLANDIRICI	NN								
	T2			T2+T3			T3		
TEST SETİ	T2			T2+T3			T3		
Doğruluk	55,74%			59,4%			66,6%		
Duyarlılık	1			1			1		
Özgüllük	0,53			0,49			0,8		
	T2			T2+T3			T3		
Hata Matrisi	20	0	4	25	0	0	5	0	0
	6	14	0	8	16	1	0	5	0
	13	8	0	18	4	4	2	3	0

Çizelge 30 incelendiğinde beklenildiği üzere poz dizilimlerinin karıştırılmasının algılamayı düşürmektedir. Bu düşüş ise öne sürdüğümüz poz dizilimlerdeki sıranın hareket algılamada doğrudan etkisini ortaya koymaktadır.

**4.6.5.5 3:3 Ters çevirilmiş poz dizilimi testi başarımları**

Hata matrisindeki dizilim  $\begin{bmatrix} ElSallama = 1 \\ Yürüme = 2 \\ Kosma = 3 \end{bmatrix}$  şeklindedir.

*Ortak Poz Ekleme Yöntemi*

Çizelge 31: 3:3 testi ortak poz eklemeli ters çevrilmiş poz dizilimi başarımları

SINIFLANDIRICI	NN								
	T2			T2+T3			T3		
TEST SETİ	T2			T2+T3			T3		
Doğruluk	65%			53,9%			66%		
Duyarlılık	1			0,96			1		
Özgüllük	0,51			0,76			0,9		
	T2			T2+T3			T3		
Hata Matrisi	20	0	0	24	1	0	5	0	0
	0	20	0	9	11	5	0	5	0
	20	1	0	3	17	6	1	4	0

Çizelge 31 incelendiğinde beklenildiği üzere poz dizilimlerinin tersine çevrilmesi hareket algılamayı düşürmektedir. Bu düşüş ise öne sürdüğümüz poz dizilimlerindeki sıranın hareket algılamada doğrudan etkisini ortaya koymaktadır.

## 5 SONUÇ

Yapılan çalışmada verilen görüntülerdeki hareketlerin poz kelimeleri ve dizilimleri kullanılarak tanınması üzerinde durulmuştur. Üç aşamalı çalışmanın birinci bölümünde bir insan vücut modeli oluşturulmuş olup ikinci bölümde ise oluşturulan bu model üzerindeki vücut parçalarının merkez noktaları elle işaretlenerek açısız özelliklerden oluşan bir kütüphane oluşturulmuştur. Bu kütüphanenin poz sayılarına göre kümelendirilmesi ile farklı hareket kategorilerine ait poz dizilimleri kütüphanesi elde edilmiştir. Son bölümde ise görüntülerde var olan hareketlerin bilgisayar ortamında otomatik olarak algılanması için farklı sınıflandırıcılar kullanılmıştır

Yapılan testlerde özellikle el sallama hareketinin her zaman eksiksiz algılanıyor olması , yürüme ve koşma hareketlerinde ise yine yüksek başarımlar elde edilmesi yöntemin başarımlarını açısından tatmin edicidir. Farklı poz dizilimi uzunluklarına göre girdi vektörlerine ortak ya da tekrarlı poz eklenmesinin de sınıflandırıcı başarımlarındaki etkisi incelenmiş ve girdi uzay büyüklüğü az olan veri gruplarında ortak poz eklemenin tekrarlı poz eklemekten daha iyi sonuçlar verdiği gözlemlenmiştir.

Eğitim ve testlerde kullanılan imge veritabanlarının farklı ölçekleme, çözünürlük ve yönelimler içermiş olmasına rağmen hareket algılamadaki yüksek başarımları açısız özelliklerin öz nitelik vektörü olarak kullanılmasının etkin bir yöntem olacağını göstermektedir.

Poz dizilimlerinin karıştırılarak ve ters çevrilerek test edilmesi ile algılamada düşüş görülmesi ise dizilim sıralamasının önemli olduğunu gösterir. Bu açıdan bakıldığında önerilen yöntemin BoW modeli kullanan algılama yöntemlerinden farklı olarak poz dizilimindeki sıralamaya bağlı olarak algılama yaptığı anlaşılmaktadır.

Elle işaretleme sürecinde üstü kapalı parçaların işaretlenmeyip pozlandırıcıya verilmemiş olması poz dizilimlerinde bir indirgemeye neden olmuş ve arka planda kalan parçaların poz dizilimlerine etkisi yok farz edilmiştir. Bu anlamda bakıldığında vücudun bütün parçalarının açılabilir özelliklerini çıkarılmış olması çalışmanın geliştirilebilecek yönlerinden birisidir. Bunun yanı sıra eğitim verilerinde kullanılan 60 kişiye ait hareket az bir örnek sayısı teşkil etmektedir ve daha geniş eğitim ve test imgeleri kullanımı çalışmayı daha kapsamlı kılacaktır.

Çalışmada yer alan elle işaretleme süreci bir soyutlama katmanı sağlamış olup aynı görevi otomatik olarak yapacak görüntü işleme yöntemleri için değişik kullanım olanakları sağlamaktadır. Bu yönde yapılacak bir çalışma, el sallama, yürüme, koşma gibi temel hareketlerin gerçek zamanlı olarak algılanmasında bir ilerleme sağlayabilir. Oluşturulan poz dizilimi kütüphanesi de bu tarz bir çalışma ile birleştirilmesiyle birlikte çalışma poz dizilimlerinden hareket algılama uygulamaları alanında daha verimli bir hale gelecektir.

## 6 KAYNAKLAR DİZİNİ

- Aggarwal, J.K. and Cai,Q.,1999,Human Motion Analysis: A Review,CVIU Journal,vol.73, no.3, 428-440
- Burges, C. J. C.,1998, A Tutorial on Support Vector Machines for Pattern Recognition, Data Mining and Knowledge Discovery, Volume 2 , Issue 2 (June 1998),121-167
- D.Hogg.,1983, Model-Based Vision: A Program to See a Walking Person, Image and Vision Computing, Vol. 1, No. 1. (1983), pp. 5-20.
- Fleck, M. M. and Forsyth,D.A.,1997,Body Plans, Proceedings of the 1997 Conference on Computer Vision and Pattern Recognition (CVPR '97), pp.678
- Felzenszwalb, P. F., Huttenlocher, D. P.,2005, Pictorial Structures for Object Recognition, International Journal of Computer Vision, Volume 61,Issue 1 (January 2005),55-79
- Fischler, M. A,Elschlager, R. A.,1973,The Representation and Matching of Pictorial Structures, IEEE Transactions on Computers, Volume 22 , Issue 1 (January 1973),67-92
- Gavrilla,D.M.,1999,"The Visual Analysis of Human Movement: A Survey",Computer Vision and Image Understanding, Vol. 73, No. 1, January, 82-98

**KAYNAKLAR DİZİNİ**(devam)

- Gershenson, Carlos, 2003, Artificial Neural Networks for Beginners, Neural and Evolutionary Computing, Volume 21, 2003, pp.8
- Hatun,K., Duygulu,P.,2008, A new representation for action recognition using sequence of pose words, 19th International Conference on Pattern Recognition (ICPR 2008)
- Jain, A.K., Murty, M. N. and Flynn, P. J., 1999, Data clustering: a review,ACM Computing Surveys (CSUR), Volume 31 , Issue 3 (September 1999),264-323
- Lee,H.-J., Chen, Z.,1992,Knowledge-guided visual perception of 3-D human gait from a single image sequence,Systems,Man and Cybernetics,IEEE Transactions on,Volume: 22, Issue: 2,336-342
- Leung,M. K.,Yang,Y-H.,1995, First Sight: A Human Body Outline Labeling System, IEEE Transactions on Pattern Analysis and Machine Intelligence,Volume 17 , Issue 4 (April 1995),359-377
- MacQueen,J.,1967, Some methods for classification and analysis of multivariate observations, Proc. Fifth Berkeley Symp. on Math. Statist. and Prob., Vol. 1 (Univ. of Calif. Press, 1967),281-297
- Ronfard,R., Schmid,C. and Triggs,B.,2002, Learning to Parse Pictures of People, Proceedings of the 7th European Conference on Computer Vision-Part IV, Lecture Notes In Computer Science; Vol. 2353,700-714



**KAYNAKLAR DİZİNİ**(devam)

- Schüldt, C., Laptev, I and Caputo, B., 2004, Recognizing human actions: a local SVM approach, Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on, Volume: 3,32-36
- Starner, T., Pentland A. and Weaver J.,1998,Real-Time American Sign Language Recognition Using Desk and Wearable Computer Based Video,IEEE Transactions on Pattern Analysis and Machine Intelligence,Volume 20, Issue 12 (December 1998),1371-1375
- Veropoulos, K. and Campbell, C. and Cristianini, N., (1999) Controlling the sensitivity of support vector machines. In Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence (IJCAI99),55-60
- Young, I.T., Gerbrands J.J and Van Vliet L.J., 1995,Fundamentals of Image Processing, Delft University of Technology,109p
- Yu, Elden., Aggarwal, J.K.,2006,Detection of Fence Climbing from Monocular Video,IEEE Internation conference on Pattern Recognition( ICPR'06), 375-378