



Recognition of human action in motion detected images with GMACA

Serkan Peldek^{1*}, Yaşar Becerikli²

¹Kastamonu University, Computer Engineering, 37150, Kastamonu, Turkey

²Kocaeli University, Computer Engineering, 41380, Kocaeli, Turkey

Highlights:

- The ability of the generalized multiple attractor cellular automata (GMACA) to attract the patterns was used in image processing.
- Motion detection has been performed on video images with GMACA.
- Motion detection images detected with the GMACA have been used to recognize human actions is a complex computer vision problem.

Graphical/Tabular Abstract

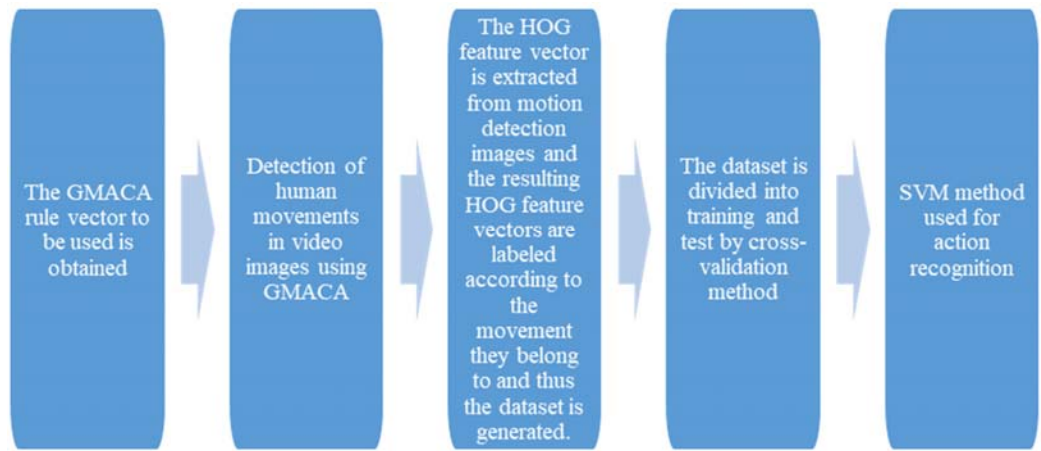


Figure A. Recognition of human action in motion detected image with GMACA

Keywords:

- Generalized multiple attractor cellular automata
- Motion detection
- Action recognition
- Cellular automata

Purpose: Generalized Multiple Attractor Cellular Automata (GMACA) is the type of Cellular Automata applied to more than one cell using the rule vector. Obtaining the rule vector in the application of the GMACA has a critical precaution. The purpose of the present study is to recognize human action in motion detected images obtained with GMACA.

Theory and Methods:

In this study, the GMACA rule vector, which is required to produce a single-length cycle attractor, was obtained using reachability tree-based methods. Detection of human motion in video images has been accomplished using attractors generated by the rule vector. In the developed action recognition method, the video images are first converted to the gray color space. Then, the GMACA rule vector to be used for motion detection is created. Motion detection is performed using GMACA. The HOG feature vector is extracted from the motion detection images and the resulting HOG feature vectors are labeled according to their motion. The dataset is created in this way. The generated dataset is decomposed into training and test data sets by cross-validation. Recognition of human action is performed by the SVM method. Experimental results are shown by the confusion matrix

Results:

The overall accuracy rate on the KTH dataset is 72.7%. The accuracy rates of boxing and hand clapping actions are 83% and 82%, respectively. The overall accuracy rate on the WEIZMANN data set is 77.3%. Accuracy rates for on-the-spot actions in the WEIZMANN dataset up to 98%. These ratios indicate that the motion detection method using GMACA can be used in action recognition applications.

Conclusion:

In this study, action recognition using motion detection images obtained with GMACA shows that the developed method can be used in complex computer vision applications. The operating logic, which allows for discrete and simultaneous calculation, makes the use of cellular automata in image processing and computer vision applications advantageous.

Article Info:

Research Article
Received: 26.09.2017
Accepted: 25.04.2018

DOI:

10.17341/gazimmfd.460500

Correspondence:

Author: Serkan Peldek
e-mail: speldek@kastamonu.edu.tr
phone: +90 366 280 2935



GMACA ile hareket tespiti yapılan video görüntülerde insan hareketlerinin tanınması

Serkan Peldek^{1*}, Yaşar Becerikli²

¹Kastamonu Üniversitesi, Bilgisayar Mühendisliği Bölümü, 37150, Kastamonu, Türkiye

²Kocaeli Üniversitesi, Bilgisayar Mühendisliği Bölümü, 41380, Kocaeli, Türkiye

Ö N E Ç İ K A N L A R

- Genelleştirilmiş çoklu cezbedici hücresel otomatlar (generalized multiple attractor cellular automata(gmaca))'ın örüntü çekim gücü görüntü işlemede kullanılmıştır
- GMACA ile video görüntülerde hareket tespiti yapılmıştır
- GMACA ile yapılan hareket tespit görüntüleri karmaşık bir bilgisayarlı görü problemi olan insan hareketlerinin tanınmasında kullanılmıştır

Makale Bilgileri

Araştırma Makalesi
Geliş: 26.09.2017
Kabul: 25.04.2018

DOI:

10.17341/gazimmfd.460500

Anahtar Kelimeler:

Genelleştirilmiş çoklu
cezbedici hücresel otomatlar,
hareket tespiti,
hareket tanıma,
hücresel otomatlar

ÖZET

Bu makale kapsamında insan hareketlerinin tanınması, Genelleştirilmiş Çoklu Cezbedici Hücresel Otomatlar (Generalized Multiple Attractor Cellular Automata(GMACA)) ile elde edilen hareket tespit görüntüleri kullanılarak gerçekleştirilmiştir. GMACA, Hücresel Otomatların birden fazla hücreye kural vektörü kullanılarak uygulanmasıdır. GMACA'nın uygulanmasında kural vektörünün elde edilmesi kritik öneme sahiptir. Bu çalışmada tek uzunluk döngülü cezbedici üretmesi istenen GMACA kural vektörü, erişilebilirlik ağacı tabanlı yöntemler kullanılarak elde edilmiştir. Video görüntülerdeki insan hareketlerinin tespiti, kural vektörünün ürettiği cezbediciler kullanılarak gerçekleştirilmiştir. Geliştirilen hareket tanıma yönteminde ilk önce görüntüler gri renk uzayına dönüştürülür. Daha sonra hareket tespiti için kullanılacak GMACA kural vektörü oluşturulur. GMACA kullanılarak hareket tespiti yapılır. Hareket tespit görüntülerinden HOG özellik vektörü çıkarılır ve elde edilen HOG özellik vektörleri ait oldukları harekete göre etiketlenir. Bu şekilde veri seti oluşturulur. Oluşturulan veri seti çapraz-doğrulama yöntemi ile eğitim ve test veri setlerine ayrıştırılır. İnsan hareketlerinin tanınması SVM yöntemi ile gerçekleştirilir. Deneysel sonuçlar karışıklık matrisi ile gösterilmiştir. Karışıklık matrisi kullanılarak geliştirilen tanıma yönteminin sınıflandırma performansı ortaya konmuştur. GMACA ile elde edilen hareket tespit görüntüleriyle yapılan hareket tanıma uygulaması mevcut arka plan çıkarma çalışmaları kadar iyi sonuç vermiştir. Elde edilen sonuçlar GMACA ile elde edilen hareket tespit görüntülerinin hareket tanıma çalışmalarında kullanılabileceğini göstermektedir. GMACA'nın zayıf yanı ikili örüntüler üzerinde uygulanabiliyor olmasıdır.

Recognition of human action in motion detected images with GMACA

H I G H L I G H T S

- The ability of the generalized multiple attractor cellular automata (GMACA) to attract the patterns was used in image processing.
- Motion detection has been performed on video images with GMACA.
- Motion detection images detected with the GMACA have been used to recognize human actions is a complex computer vision problem.

Article Info

Research Article
Received: 26.09.2017
Accepted: 25.04.2018

DOI:

10.17341/gazimmfd.460500

Keywords:

Generalized multiple
attractor cellular automata,
motion detection,
action recognition,
cellular automata

ABSTRACT

In this article, recognition of human action was performed using motion detection images obtained with Generalized Multiple Attractor Cellular Automata (GMACA). GMACA is the type of Cellular Automata applied to more than one cell using the rule vector. Obtaining the rule vector in the application of the GMACA has a critical precaution. In this study, the GMACA rule vector, which is required to produce a single-length cycle attractor, was obtained using reachability tree-based methods. Detection of human motion in video images has been accomplished using attractors generated by the rule vector. In the developed action recognition method, the images are first converted to the gray color space. Then, the GMACA rule vector to be used for motion detection is created. Motion detection is performed using GMACA. The HOG feature vector is extracted from the motion detection images and the resulting HOG feature vectors are labeled according to their motion. The dataset is created in this way. The generated dataset is decomposed into training and test data sets by cross-validation. Recognition of human action is performed by the SVM method. Experimental results are shown by the confusion matrix. The classification performance of the recognition method developed using the confusion matrix has been demonstrated. Motion recognition with motion detection images obtained with GMACA is as good as existing background subtraction studies. The obtained results show that motion detection images obtained by GMACA can be used for action recognition. The weakness of GMACA is that it can be applied on binary patterns.

*Sorumlu Yazar/Corresponding Author: speldek@kastamonu.edu.tr, ybecerikli@kocaeli.edu.tr, / Tel: +90 366 280 2935

1. GİRİŞ (INTRODUCTION)

Video görüntülerindeki insan hareketlerinin tanınması, üzerinde uzun yıllar çalışılan ancak olgunlaşmamış bir bilgisayarlı görü problemidir. Video görüntülerdeki insan hareketlerinin otomatik olarak yorumlanabilmesi; yaşlıların ve hastaların uzaktan bakımı ve tedavisi; görme engelli kişiler için yapay gözlüklerin üretilmesi, bilgisayar sistemlerinin insanlar tarafından uzaktan kontrolü, güvenlik amaçlı ortam gözetleme, sporcu hareketlerinin analizi, akıllı insan bilgisayar etkileşimi vb. birçok iş için çözüm sunma potansiyeline sahiptir. Amacı insan hareketlerini otomatik olarak ayırtmak, elde etmek ve tanımak olan hareket analizi bilgisayarlı görme alanında üzerinde yoğunlukla çalışılmaktadır [1]. İnsan hareketlerini tanımının amacı bilinmeyen videolardan(video ardışık görüntülerde oluşur) devam eden hareket, eylem ve etkinliklerin otomatik analiz edilmesidir. Daha basit bir durum için düşünülürse; sadece bir insan hareketini içeren bir video parçası için amaç videoyu içerdiği eyleme göre doğru eylem kategorisine göre sınıflandırmaktır. Daha genel ve karmaşık bir durum için; video görüntülerde süregiden insan etkinliklerinin tanınması insanlar hareketlerinin başlangıç ve sonunun tespit edilmesiyle gerçekleştirilir [2]. Hareket analizi ve tanıma çalışmaları başlangıç tarihi itibarıyla eski bir çalışma alanıdır. Ancak gerçek hayat uygulamaları ve performans açısından olgunlaşmamıştır [3-5].

Hareket tanımının ilk adımı ardışık resim çerçevelerindeki farklılıklardan veya insan ait ten rengi ve şekil gibi özelliklerden yararlanarak hareketli insan nesnesinin tespitidir. Nesne tespitini zorlaştıran ışık yetersizliği, nesnenin gölgesi ve nesnelerin çakışması gibi zorluklarla uğraşılması gerekir. Tespit edilen nesnenin takip edilmesi gerekir. Tespit ve takip edilen nesneden faydalı ve kararlı özellikler çıkartılır Son olarak hareketli nesneden çıkartılan özellikler makine öğrenmesi modelleri kullanılarak tanınır.

Hareketlerin zaman akışı içerisinde nerede başlayıp nerede bittiğini belirlemek hareket tanıma kadar önemlidir. Çünkü hareket için yapılacak yanlış sınırlar hareketlerin tanınmamasına veya yanlış tanınmasına doğrudan etki edecektir. Bu konuda yapılan çalışmalar, Weinland vd. [5] tarafından üçe ayrılmıştır; sınır tespiti, kayan pencereler ve gramer birleştirme.

Bu çalışmada video görüntülerdeki insana ait hareketlerin tanınması GMACA ile yapılan hareket tespit görüntüleri kullanılarak gerçekleştirilmiştir. Ayrık ve eş zamanlı hesaplamaya imkân sunan Hücreli Otomata türü olan GMACA ile yapılan hareket tespit görüntülerinin hareket tanıma gibi zor bir bilgisayarlı görü probleminde kullanılabilirliğini göstermesi, bu çalışmanın ana katkısıdır. Video görüntülerdeki hareket tespiti, GMACA kural vektörünün ürettiği tek uzunluk döngülü cezbediciler ile belirlenmiştir. Yapılan literatür değerlendirmesinde hareket tanıma kullanılan İngilizce kelimelerin (motion, action, gesture, activity vb.) hangi hareket yapıları için kullanıldığı

açıklanmıştır. Ayrıca hareket tanımının aşamaları, hareket tanıma yaklaşımları, durağan görüntülerdeki hareket tanıma, semantik yaklaşımlar ve farklı video görüntüleme teknolojilerine değinilmiştir.

Yöntem bölümünde video görüntülerde hareket tespiti için kullanılan hücreli otomatlar ve çoklu cezbedici hücreli otomatlar açıklanmıştır. Özellik çıkarımı için kullanılan Histogram eğilim büyüklükleri(Histogram Oriented Gradient (HOG)) ve hareketlerin sınıflandırılması için kullanılan destek vektör makineleri konuları başlıklar halinde ele alınmıştır.

DeneySEL sonuçlar bölümünde geliştirilen yöntem KTH ve WEIZMANN veri tabanlarında yer alan videolar üzerinde test edilmiştir. Elde edilen sonuçların karışıklık matrisleri, sınıflandırma raporları ve literatürde aynı veri tabanları üzerinde yapılan çalışmaların sonuçları ile karşılaştırılması tablolar halinde verilmiştir. Son olarak çalışmanın zayıf yanları ve geliştirilmesi için yapılabilecek olası çalışmalara değinilmiştir.

2. LİTERATÜRDE VIDEO GÖRÜNTÜLERDE İNSAN HAREKETLERİNİN TANINMASI (RECOGNITION OF HUMAN ACTION IN VIDEO IMAGES IN THE LITERATURE)

2.1. İnsan Hareketleri İçin Kullanılan Farklı Terimler (Different Terms Used for Human Movements)

Literatürde video görüntülerindeki insan hareketlerinin kavramsal anlamlarının farklı seviyelere göre adlandırılması vardır. Nagel bu seviyeleri değişim(change), olay(event), fiil(verb), hadise(episode) ve geçmiş(history) olarak adlandırmaktadır. Bu adlandırmalara karşılık gelen anlamlar şöyle tarif edilmiştir: Değişim gözlenebilir hareketler, olay kompleks bir davranışın ilk değişimi, fiil eyleme geçme hali, hadise eylemler zinciri, geçmiş ilişkili eylemlerin sıralı uzantısı olarak tanımlanmıştır [6]. Gonzalez vd. Nagel'in çalışmasını esas alarak yeni bir adlandırma yapmışlardır; hareket(motion), eylem(action), etkinlik(activity) ve durum(situation). Gonzalez vd. yaptıkları adlandırmalara karşılık gelen anlamlar şöyle tarif edilmiştir: Hareket sıralı resim çerçeveleri içerisindeki insan pozunun veya yerinin değişimi, eylem insanın gerçekleştirdiği anlamı olan bir iş, etkinlik insanın yaptığı bir veya daha fazla eylemler dizisi, durum gerçekleştiği şartlara göre farklı anlamları olan etkinlikler şeklinde tanımlanmıştır [7]. Fanti eylem-etkinliklerin şematik seviye gösterimini yapmıştır. Bu seviyeler hareket(movement), eylem(action) ve etkinliktir(activity) [8].

Fanti'nin yaptığı adlandırmalara karşılık gelen anlamlar şöyle tarif edilmiştir: Hareket eğitilmiş veri kullanılmadan elde edilen atomik hareketlerdir. Birkaç resimlik süre alır ve bazı hareketler üst üste gelebilir. Eylem hareketlerden meydana gelmiş kısa olaylardır ve bunlar bir araya gelince bir etkinliği netice verme ihtimali vardır. Tipik olarak bir eylem bir kaç saniye içerisinde sonlanır. Yürümek, koşmak

gibi bazı tekrarlı olaylar yine eylem olarak kabul edilir. Etkinlik en üst seviyede yer alır ve geniş bir zaman dilimine yayılır [8]. Diğer bir isimlendirme ilkel eylem(primitive action), eylem(action) ve etkinlik(activity) şeklindedir. Bu adlandırmalara karşılık gelen anlamlar şöyle tarif edilmiştir: İlkel eylem, eylemleri oluşturan atomik hareketlerdir. Eylem bir görevi tamamlamak için gerekli sıralı ilkel eylemler serisidir. Etkinlik çevrenin, insanın etkileşiminin veya nesnenin mahiyetine bağlı olan kapsamlı olaylardır. Örneğin tenis sporunu düşünürsek ilkel eylemler sağ vuruş, sol vuruş, sola koşmak, sağa koşmak vb. hareketlerdir. Eylem top tutmak için yapılan bir dizi ilkel eylemlerden meydana gelir. İlkel eylemlere bağlı olarak topu tutmanın farklı yolları olabilir. Nihayetinde tenis oynamak bir etkinliktir [9].

Tablo 1. Hareketlerin seviyeleri (Level of Movements) [8]

Süre	Adlandırma	Gerçek dünya karşılığı
1 Gün 1 Saat	Etkinlik	Spor, yemek pişirme, antrenman, yüzme
1 Dakika	Eylem	Yürüme, koşma, içme, açma/kapama, işaret etme
1 kaç Saniye ve Frame	Hareket	Çekme, itme, sol/sağa adım atma, fırlatma

Eylemler; sade ve insanlar tarafından gerçekleştirilen basit vücut hareketlerinin birleşimi olarak değerlendirilebilir. Bir eylem bir hareketle yapılabildiği gibi birden fazla ilkel eylemden meydana da gelebilir. İkisi arasında belirgin bir çizgiden bahsetmek zordur [10].

Etkinlik birçok basit eylemin bir araya gelmesinden oluşur. Bir veya daha fazla özne tarafından gerçekleştirilebilir. Basit eylemler olduğu gibi kompleks eylemler vardır. Kompleks eylemler etkinlik olarak da kabul edilebilir. İnsanın veya bir grup insanın davranışları bir etkinlik olabilir veya etkinliklerin bir araya gelmesinden meydana gelebilir.

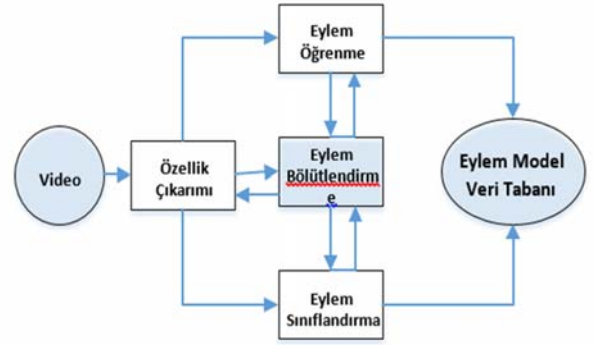
Eylem ve etkinliklerin farklı oluşumları vardır. Bazı eylemler çok yavaştır. Araba sürme gibi bazı eylemler durağandır. Koşmak, zıplamak gibi eylemler tüm vücudun hareket etmesini gerektirir. Alkışlamak, boks yapmak, ateş etme eylemlerinde vücudun üst kısmı, top vururken alt kısmı kullanılır. Kahve içmek veya at sürmek vücudun kullanılan bölgeleri ve hareket bilgileri açısından birbirinden çok farklıdır. Bu nedenle eylem veya etkinliğin birçok yönü vardır ve tek bir yöntem altında ele alınması çok zordur. Bir hareket tanıma çalışmasının kavramsal bilgileri ve nesnenin tespiti konularını dikkate alması gerekir. Örneğin bardak kahve içme eylemi için, futbol topu şut çekme eylemi için kullanılan nesnelere [3].

2.2. İnsan Hareketlerinin Tanınması (Recognition of Human Action)

Bilgisayarlı görüde tanıma(Recognition) eldeki sayısal görüntü verisi ile özel bir nesneyi, özelliği veya hareketi tanımadır [11]. Hareket tanıma, video görüntüler içerisinde hareketi, eylemi veya etkinliği tanıma işlemidir. İnsan hareketlerinin tanınması birçok zorluğu içerisinde

barındıran bir problemdir. Hareketler dinamik yapıdadır ve zaman içerisindeki bir noktaya yönelmekle tanınmazlar. Hareketlerin insanlar tarafından farklı yapılması veya aynı kişinin aynı hareketi farklı yapması hareket tanımayı zorlaştırmaktadır. Başarılı bir hareket tanıma için arka plan çıkarımının başarılı bir şekilde yapılması gerekir.

Eylem tanıma eylemlere isim verme işlemidir, genelde bu bir eylem fiilinin basit formudur. Teknik olarak eylem insanın bir görevi yaparken ortaya koyduğu hareketler dizisidir. Şekil 1. Tipik bir eylem tanıma sisteminin veri akış diyagramı gösterilmektedir.



Şekil 1. Eylem tanıma sisteminin veri akış diyagramı (Data flow diagram of the action recognition system) [5]

İnsan eylemlerine göre ayırt edici olan vücut duruşunun ve hareket işaretlerinin video görüntülerden çıkartılması özellik çıkarma olarak adlandırılmaktadır ve hareket tanımadaki temel işlemdir. Kompleks vücut modellerinden basit karartı resimlerine geniş bir yelpazede farklı temsiller kullanılabilir. Her durumda kişinin konumu, bölgesel çakışmalara karşı dayanıklılık, arka plan karışıklığı, gölgeler ve farklı ışıklandırmalar gibi etkenler göz önünde bulundurulmalıdır. Gelişmiş temsiller farklı elbise ve vücut yapılarından az etkilenirler. Eylem öğrenme ve sınıflandırma çıkartılan özelliklerden istatistiksel öğrenme adımlarıdır ve gözlenen yeni özellikleri sınıflandırmak için bu özellikler kullanılır. Temel sorun gerçekleştirilen bir eylemin farklı cinsiyette ve ölçülerde kişiler tarafından farklı hız ve stillerde yapılmasıdır. İtme, vurma veya el sallama gibi eylemlerin sınıflandırılması bize kolay gelebilir ancak, pratikte bunların uygulanması çok farklılık arz edebilir. Bu nedenle eylem tanımda özel bir sorunda bir hareketi temsil eden model tanımlarken aynı zamanda bu modelin eylemin değişik formlarına da uyarlanabilir olması gerekir. Eylem bölümlendirme tek bir eylem örneğinin hareket akışının belli bir yerde kesilmesi veri tabanında modellenen eylemle tutarlı olması açısından önemlidir [5].

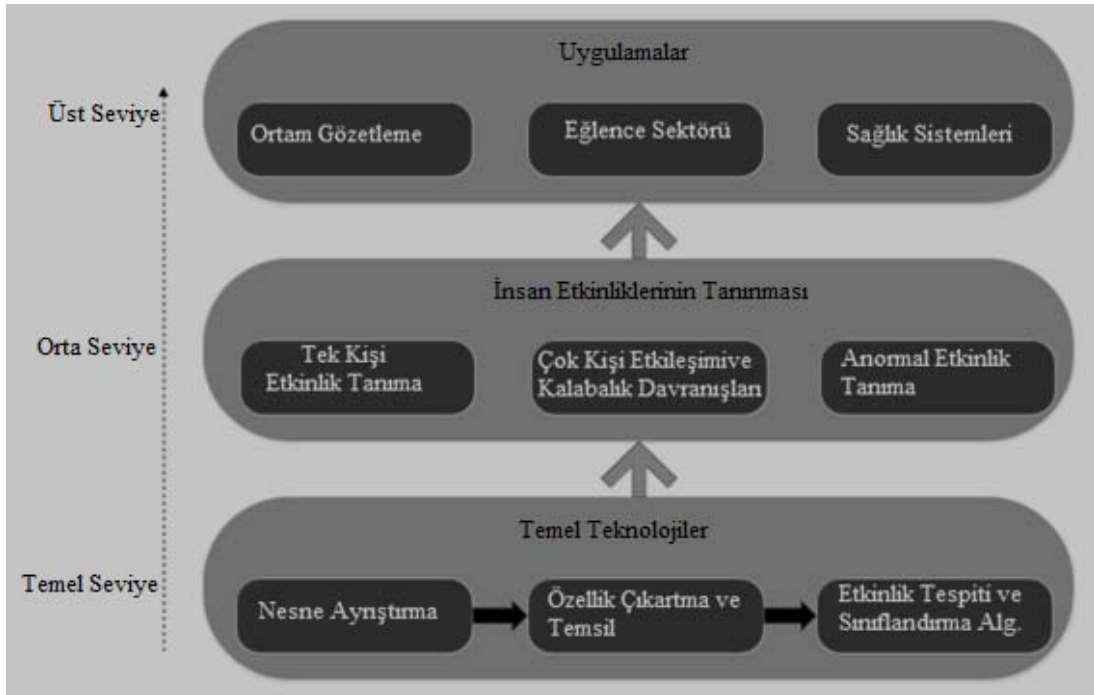
Ke vd. yaptıkları literatür araştırmalarına göre insan etkinliklerinin tanınması üzerine yapılan çalışmaları üç ayrı seviyeye ayırmışlardır [12]. Birinci seviyede insan hareketlerinin tanınması için kullanılan temel yöntemler ve teknolojiler, orta seviyede insan etkinliklerinin tanınması ve üçüncü seviyede uygulamaların gerçekleştirilmesi yer alır. Bu seviyeler Şekil 2.'de görülmektedir.

Temel seviyede üç ana işlem aşaması vardır; 1) Nesne ayrıştırma, 2) Özellik çıkarma ve temsil, 3) Eylem tespiti ve sınıflandırma. İlk önce insan nesnesi video içerisinden ayrıştırılır. Şekil, silüet, renkler, duruş ve vücut hareketleri gibi insan nesnesinin karakteristik özellikleri çıkartılır ve temsil edilir. Ardından değişik insan eylemelerinin tanınması için çıkartılan özelliklere tespit ve sınıflandırma algoritmaları uygulanır. İnsan hareketlerinin tanınma sisteminin orta seviyesinde üç farklı tanıma sistemi ele alınmıştır; 1) Tek kişi etkinliklerinin tanınması, 2) Birden fazla kişinin etkileşimi ve kalabalık davranışları, 3) Anormal etkinlik tanıma. Son olarak üst seviyede insan hareketlerinin tanınması ortam gözetleme, eğlence endüstrisi ve sağlık sistemleri gibi pratiklere uygulanması yer almaktadır.

Hareket tanıma yöntemlerinde görüntülerden çıkartılan temel özelliklerin yanında eylemlerin semantik özelliklerini çıkartan yöntemler de kullanılmıştır [13]. Semantik özellikler "Bu hareketi yapmanın anlamı nedir?" veya "Bir eylemi nasıl anlamalıyız?" gibi soruların cevaplarını bulmayı amaçlar. Semantik terimi anlam çalışmalarını ifade

eder. Örneğin araba ve yolun aynı resimde görünmesi anlamlıyken, zürafa ve mutfağın aynı resimde görülmesi anlamlı değildir. Semantik özellikler aynı sınıftaki farklı eylem oluşumlarının tespit edilmesinde kullanışlıdır. Burada aynı sınıftan maksat aynı eylemin farklı yapılarıdır ve aynı eylemin farklı oluşumlarının birbirlerine nasıl benzediğidir. Şekil 3.'de görüldüğü gibi aynı eylemi farklı şekilde gerçekleştirebilirler veya aynı kişi aynı eylemi farklı durumlarda farklı gerçekleştirebilir. İnsanlar giydikleri elbiseler, vücut şekilleri ve görünme açılarının çeşitliliğine göre çok farklı şekilde görünebilirler. Semantik özellikler ortak semantiğe sahip ancak görüntüde farklılık arz eden benzer eylemlerin ayırt edilmesine yardımcı olur.

Her ne kadar hareket tabanlı/video tabanlı insan hareket tanıma bilgisayarlı görme ve örüntü tanımada yoğunlukla çalışılmakta ise de son zamanlarda yapılan çalışmalarda durağan görüntüler içerisinde hareket tanıma araştırılmaya başlanmıştır. Guo ve Lai "Birçok eylem tek bir görüntü(hareket ve video sinyali içermeyen) ile anlatılabilir ve kişinin görünümüne göre yaptığı hareket anlaşılabilir"



Şekil 2. İnsan hareketleri tanıma sistemine genel bakış (Overview of human movement recognition system) [12]



Şekil 3. "Tekmeleme" eylemi. Aynı eylem kamera açısı, elbise, vücut şekilleri gibi farklılıklardan dolayı farklı görünmektedir (The "kicking" action. The same action seems different because of differences in camera angle, dress, body shape, etc.) [13].

[14] düşüncesinin bu araştırmaların temel motivasyonu olduğunu belirtmiştir. Bu düşünce durağan görüntülerde hareketlerin analizini ve tanınması için önerilen hesaplama algoritmalarının gelişimine katkı sağlamaktadır. İnternet üzerindeki büyük sayıdaki tek görüntü formatları düşünüldüğünde insan hareketlerinin bu görüntüler üzerinden analiz edilmesi değerli hale gelmektedir.

Görüş açısı insan hareketleri tanımda çok önemli dar boğazlardan biridir [15]. Şekil 4.'de görüldüğü gibi ortam izleme kameraları genelde rasgele açıyla görüntü alır bu nedenle yapılacak uygulamaların farklı görüş açılarına göre uygulanabilir olması gerekir. İnsan hareketlerini analiz eden yöntemlerin çoğu tek bir görüş açısına göre uyarlanır. Görüntüdeki insanların çoğu ya yüzü kameraya dönüktür ya da kameraya paralel bir şekilde hareket eder. Gerçek dünyada böyle bir kısıtla görüntü almak imkânsızdır.

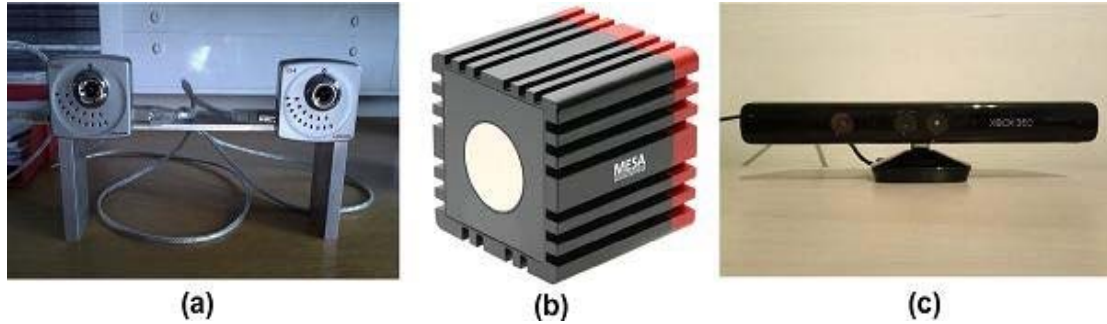
Görüş açısı kısıtlamalarından dolayı çok sayıda insan hareketlerini tanıma yöntemi geniş kapsamlı olarak kullanılmamaktadır. Son yıllarda görüş açısında bağımsız poz kestirimi, eylem temsili ve tanıma gibi görüş açısı etkilerini azaltan çalışmalar yapılmıştır. Ji ve Liu yapılan bu çalışmaları ikiye ayırmışlardır; görüş açısından bağımsız poz temsili ve kestirimi; görüş açısından bağımsız eylem temsili ve tanıma. Bu ikisi arasındaki fark; birincinin görüntü dizilerinden 3D modeller oluşturmaya çalışırken ikinci insan eylemlerini çıkarmak ve tanımaya yoğunlaşır [15].

Şimdiye kadar yapılan çalışmaların çoğu yoğunluk görüntüsü veya yoğunluk videosu verilerini girdi olarak kullanan çalışmalardır. Yoğunluk görüntüleri zengin bilgi içermesine karşın ışık değişimine duyarlıdır. Çıkartılacak özellikler nesnenin genel şeklinden çok yüzeysel doku yapılarına göre elde edilmesi durumlarında yoğunluk görüntüleri sınırlı bilgiye sahiptir. Arka plandan nesne çıkarımı karmaşık arka planlarda zordur. Derinlik görüntülerin yoğunluk görüntülere göre avantajı vardır. Öncelikli olarak ışık değişiminden etkilenmezler. İkinci olarak görünümün üç boyutlu yapısını sağladıklarından dolayı arka plan çıkarımı, nesne ayrıştırma ve hareket kestirimi gibi işlemleri önemli ölçüde kolaylaştırırlar.

Derinlik görüntüleri derinlik algılayıcıları tarafından elde edilirler. Bilgisayarlı görmede uygulanan üç ana derinlik algılayıcı teknoloji vardır; stereo kameralar, time-of-flight (ToF) kameralar ve kodlanmış ışık (coded light) kameralar. Şekil 5'te üç farklı kameranın görüntüleri vardır. Derinlik görüntüler yoğunluk görüntüler gibi ön işleme ihtiyaç duyabilirler. Stereo yapay görme, insanın görme sisteminden esinlenerek yapılmıştır. İki veya daha fazla görüş açısından alınmış resimlerden 3D görüntüler oluştururlar. Stereo geometri hesaplamalarının karışıklığından dolayı stereo görüntülerden derin haritaların inşa edilmesi hala zorluğunu devam ettirmektedir. Derinlik haritalarının oluşturulması için kullanılan yoğunluk görüntüleri ışık değişimine duyarlı olduklarından dolayı üçgenleme eşleştirmesini



Şekil 4. UCM veri setinde ortam gözetleme sahnesi (Surveillance scene in the UCM data set) [16]



Şekil 5. Derinlik kamera örnekleri. a) Stereo kamera b) Time-of-flight c) Kodlanmış ışık (Microsoft Kinect) (Depth camera samples. (a) Stereo camera (b) Time-of-flight (c) Coded light (Microsoft Kinect)) [1]

zorlaştırmaktadır. Bu nedenle stereo görüntülerden derinlik haritalarının inşa edilmesi çoğu gerçek dünya uygulamaları için uygun değildir. Time-of-flight (ToF) kameralar nesne yüzeyine gönderdiği darbeli ışıkların geri dönüş zamanı ve ışık hızına göre nesnenin uzaklığını elde eder. ToF kameralar yüksek görüntü karesi oranlarına ulaştıklarından dolayı gerçek zamanlı uygulamalarda elverişlidir. ToF kameralarının en önemli eksiği yüksek fiyatlı ve düşük çözünürlükte olmasıdır.

2010 Kasım ayında Microsoft, kişisel tüketicilerin alabileceği fiyata sahip Kinect adını verdiği görüntüleme cihazını piyasaya sürdü. Microsoft Kinect Guinness rekorlarına en hızlı satılan elektronik eşya olarak girdi. Kinect RGB kamerasına ve derinlik sensörüne sahiptir. Derinlik sensörü 640X480 çözünürlükte saniyede 30 resim karesi sağlayabilmektedir. Kameranın algılayabildiği uzaklık skalası 0,8 ile 3,5 metre aralığındadır. Tablo 2.'de derinlik kameralarını özellik karşılaştırılması verilmiştir [1].

3. ÖNERİLEN YÖNTEM (PROPOSED METHOD)

Bu çalışmada insan hareketlerinin tanınması KTH ve WEIZMANN veri tabanlarında yer alan insan hareketlerini içeren videolar kullanılarak gerçekleştirilmiştir. KTH veri tabanında altı çeşit insan eylemi(yürümek, tempolu koşu, koşu, boks, elle sallama ve elle el çırpma) içeren videolar yer almaktadır. Eylemler 25 farklı kişi tarafından dört farklı senaryo içerisinde gerçekleştirilmiştir. Eylemlerin gerçekleştirildiği senaryolar şöyledir; açık alanda, açık alanda ve ölçek değişikliği, farklı giysilerle açık alanda ve kapalı alanlarda. Veri tabanı 600 video içermektedir. Tüm videolar, statik bir kamera ile kaydedilmiştir. Videolardaki arka planlar sabittir. Videolar, 25fps kare hızı ve 160x120 piksel çözünürlüğe sahiptir [17]. WEIZMANN veri tabanında on çeşit insan eylemi(eğilme, tek el sallama, iki el sallama, jack(iki el ve iki bacağın yerinde zıplayarak beraber açılıp kapanması), yerinde zıplama, ileri doğru zıplama, yan yürüme, yürüme, koşma ve tek ayak üzerinde sekerek yürüme) içeren videolar yer almaktadır. Eylemler 9 farklı

kişi tarafından tek bir senaryo olan açık alanda gerçekleştirilmiştir. Veri tabanı statik kamera ile çekilen, 25 fps kare hızı ve 180x144 çözünürlüğe sahip 90 video içermektedir [18].

Gerçekleştirilen hareket tanıma yönteminin adımları şöyle özetlenebilir;

- Görüntüler gri renk uzayına dönüştürülür
- Hareket tespiti için kullanılacak GMACA kural vektörü oluşturulur
- GMACA kullanılarak hareket tespiti yapılır
- Hareket tespit görüntülerinden HOG özellik vektörü çıkarılır ve elde edilen HOG özellik vektörleri ait oldukları harekete göre etiketlenir. Bu şekilde veri seti oluşturulur
- Oluşturulan veri seti çapraz-doğrulama yöntemi ile eğitim ve test veri setlerine ayrıştırılır
- Hareketlerin tanınması SVM yöntemi ile gerçekleştirilir
- Yukarıda verilen hareket tanıma yönteminin adımları aşağıdaki bölümlerde başlıklar altında detaylıca incelenmiştir.

3.1. Renkli Görüntülerin Griye Dönüştürülmesi (Conversion of Color Images to Gray)

Görüntü tanıma ve analiz uygulamalarını çoğu gri görüntüler kullanır [19]. Kanan ve Cottrell yaptıkları çalışmada renkli görüntülerin gri görüntüye dönüştürülmesinde kullanılan yöntemin görüntü tanıma başarısına etki ettiğini göstermektedir.

Görüntüleri griye dönüştürmek için kullanılacak yöntemler Eş. 1 ve Eş. 2'de verilmiştir;

$$gri_{\text{yoğunluk}} = \frac{R+G+B}{3} \quad (1)$$

$$gri_{\text{Luminance}} = 0,299 * R + 0,587 * G + 0,114 * B \quad (2)$$

Tablo 2. Derinlik kameralarının karşılaştırılması (Comparison of depth cameras) [1].

Algılayıcı tipi	Stereo kamera	Time-of-flight (ToF)	Kodlanmış ışık
Çözünürlük	Yüksek: 640X480 veya daha fazla	Düşük: 64X48 ile 200X200 aralığı	Yüksek: 640X480
Hız	Yavaş	Hızlı	Hızlı
Kapsam alanı	Kamera özelliğine bağlı	5 ile 10 metre arası (iç ve dış mekân)	0,8 ile 3,5 metre arası (iç mekân)
Derinlik çözünürlüğü	Kamera özelliğine çözünürlüğüne bağlı ve	5 mm den az	1 cm den az
Görüş alanı	Kamera lensine bağlı	Yaklaşık yatay 43 dikey 69 derece açı	Yaklaşık yatay 43 dikey 57 derece açı
Derinlik haritasında boşluklar	Evet	Hayır	Evet
Fiyat	Ucuz	Pahalı	Ucuz
Işık duyarlılığı	Evet	Hayır	Hayır

Bu çalışmada yukardaki griye dönüştürme yöntemlerinden Eş. 2’de gösterilen *grLuminance* yöntemi kullanılmıştır.

3.2. GMACA İle Hareket Tespiti (Motion Detection with GMACA)

CA zaman ve uzay ayrık olarak tanımlanmış hücresel matris üzerinde yerel etkileşimlere göre işleyen dinamik bir sistemdir. CA üç temel karakteristiği sahiptir [20]; 1)Basitlik: CA’nın temel elemanı bilinen basit hücredir, 2) Geniş paralellik: CA paralel bilgisayarlardan bile daha çaplı paralellik gerçekleştirir, 3) Yerellik: CA’nın işleyişi her hücrenin komşularıyla olan etkileşimiyle gerçekleşir.

CA’nın çalışma mantığı basit olduğundan farklı disiplinlerdeki araştırmacıların dikkatini çekmiştir. Şekil 6 incelenerek CA’nın çalışma mantığı anlaşılabilir. İlk adım, başlangıç durumu kabul edilir. İkinci adımda hücrelerin alacağı değerler kendi, sağ ve solundaki hücrelerin değerlerine göre belirlenir. Şekil 6’da her bir hücre için uygulanan kural 254 olduğu için tekrarlar sonucu elde edilen görüntü beşinci adımdaki gibi olacaktır. Hücreler Kural 254’e göre neden aşağıdaki değerleri almaktadır? Bunun açıklaması 254 sayısının ikilik tabandaki karşılığında gizlidir. 254’ün ikili karşılığı 11111110’dır. Her bir hanenin yer numarası 7 ile 0 arasındadır. Mesela 7’inci hanede 1 var iken 0’inci hanede 0 değeri vardır.

Şimdi tekrardan hücre, sağdaki ve soldaki komşuları düşünülün: Üç hücrenin (hanenin) ikilik tabanda alabileceği kaç farklı değer vardır? Elbette sekiz, 0’dan 7’ye kadar. Örneğin hücrelerin durumu Şekil 6.’daki gibi ilk durumda beyaz-siyah-beyaz olsun bunun ikili karşılı 010’dır. 010’ın onluk karşılığı 2’dir. 254 sayısının ikilik tabandaki karşılığı olan 11111110’ın 2’inci hanesinde 1 değeri olduğundan dolayı siyah hücrenin bir altındaki hücrenin değeri 1 ve rengi siyahtır. Matristeki tüm hücrelerin değeri her bir tekrarda bu şekilde güncellenmektedir. Hücrenin bir sonraki durumunun belirlenmesi dijital mantıkta kullanılan doğruluk tablolarını hatırlatır. Siyah kutuların 1, beyaz kutuların 0 olduğu varsayılarak, Tablo 3’teki gibi bir doğruluk tablosu oluşturulabilir. Tablo 3, hücrelerin bir sonraki durumunun(Şekil 6’ya göre alttaki hücre için) Kural 254’e göre nasıl elde edildiğini göstermektedir.

Burada ilk 3 kutunun $2^3=8$ farklı durumuna göre ortadaki hücrenin bir sonraki durumu belirlenmiştir. Tablo 3’ün en sağdaki sütununda elde edilen 11111110 dizisi aynı zamanda

kullanılan kuralın numarasını verir. Tablo 3’de ortadaki hücrenin bir sonraki durumu Kural254’e göre belirlenmiştir. Bunun nedeni, ikili sayı sistemine dayanmaktadır; $11111110 = 254$.

İki duruma(0 ve 1) ve üç komşuya(sol, kendi ve sağ) sahip tek boyutlu hücresel otomatlar Tablo 3’te görüldüğü gibi 8 farklı dizilişe sahip olabilirler. Dolayısıyla, 8 basamaklı bu dizilime göre $2^8=256$ farklı kural elde edilebilir. Durum sayısı k ve komşu sayısı m olan bir hücresel otomatın sahip olabileceği kural sayısının elde edilmesi şöyle genelleştirilebilir; 2^{km} . İki durum(k=2) ve üç komşuya(m=3) sahip hücresel otomatlar için kural sayısı şöyle olacaktır; $2^{23} = 256$

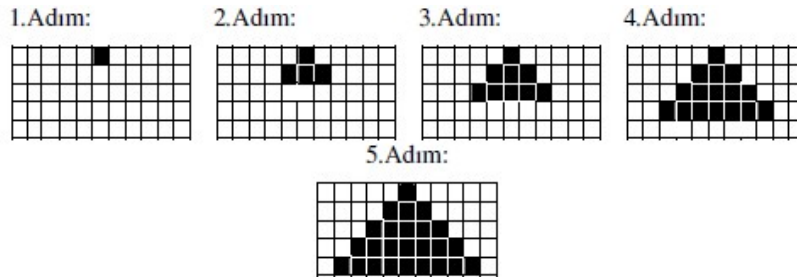
Tablo 3. Kural254 için doğruluk tablosu (The truth table for Rule254)

Sol komşu	Hücre	Sağ komşu	Ortadaki hücrenin bir sonraki durumu
0	0	0	0
0	0	1	1
0	1	0	1
0	1	1	1
1	0	0	1
1	0	1	1
1	1	0	1
1	1	1	1

Kurallar, kural vektörü kullanarak birden fazla hücreye uygulanabilir. Birden fazla hücre için kural vektörünün uygulanması üzerine birçok araştırma bulunmaktadır [21-24]. Kural vektörünün tanımı aşağıdaki gibi yapılabilir;

Kural Vektörü: Birden fazla hücrenin bir sonraki durumunu belirlemek için kurallardan oluşan vektöre denir; $R = \langle R_1, R_2, \dots, R_i, \dots, R_n \rangle$.

Birden fazla hücre üzerine uygulanan CA’nın ortaya koyduğu sonuçlar birçok araştırmacının dikkatini çekmiştir. Uzunluğu n olan bir kural vektörünün n uzunluklu ikili örüntüler üzerine uygulandığında bazı durumların diğer durumları çektiği gözlemlenmiştir. Bu durumlar çekim merkezi gibi davranmaktadır. Durumların bu özelliğinden dolayı “cezbedici(attractor)” veya “cezbedici kap(attractor basin)” olarak adlandırılmıştır. “Cezbediciler” kendine dönen durumlar olarak da adlandırılmıştır. Kendine dönen bu durumlar doğrudan kendine dönebildiği gibi birkaç adım



Şekil 6. Kural254’ün beş adımı (Five steps to Rule254)

sonrada kendine dönebilmektedir. Doğrudan kendine dönen “cezbediciler” tek uzunluk döngülü cezbediciler; birkaç adım sonra kendine dönen cezbediciler çok uzunluk döngülü cezbediciler olarak adlandırılmıştır [21].

N tane hücre üzerine uygulanan hücresele otomatlar kural vektörünün yapısına göre iki farklı isimle adlandırılmıştır; Çoklu Cezbedicili Hücresele Otomatlar(MACA(Multiple Attractor Cellular Automata)) ve Genelleştirilmiş Çoklu Cezbedicili Hücresele Otomatlar (GMACA(Generalized Multiple Attractor Cellular Automata)). Kural vektörü sadece lineer ve toplamsal kurallardan oluşuyorsa; MACA olarak adlandırılır. Kural vektöründe lineer ve toplamsal olmayan kurallar varsa; GMACA olarak adlandırılır. Farklı tipte kurallar vardır. Bunlar aşağıdaki gibi tanımlanabilir.

Lineer Kurallar: Bir sonraki durumun XOR mantıksal operatörüne göre yapıldığı kurallardır. İki durum ve üç komşu sahip hücresele otomatların 256 kuralından sadece yedi tanesi lineer kuraldır. Bu kurallar şöyledir; 60, 90, 102, 150, 170, 204, 240

Toplamsal Kurallar: Bir sonraki durumun XOR veya XNOR mantıksal operatörüne göre yapıldığı kurallardır. İki durum ve üç komşu sahip hücresele otomatların 256 kuralından sadece on dört tanesi toplamsal kuraldır. Bu kurallar şöyledir; lineer kurallar(60, 90, 102, 150, 170, 204, 240) ve bu lineer kuralların tamlayıcı(ikilik tabandaki karşılığının) olan kurallar; 195, 165, 153, 105, 85, 51, 15.

Lineer Olmayan Kurallar: Bir sonraki durumun AND veya OR mantıksal operatörüne göre yapıldığı kurallardır. İki durum ve üç komşu sahip hücresele otomatların 256 kuralından on dört tanesi hariç diğerleri lineer olmayan kurallardır.

Durum geçiş diyagramları, kural vektörünün ikili örüntüleri uygulanmasıyla bir durumun bir sonraki duruma geçişini gösterir. Durum geçiş diyagramları kural vektöründeki kuralların yapısına ve dizilimine göre şekillenir. GMACA ile elde edilen durum geçiş diyagramlarının yapısı, GMACA'nın kullanılacağı uygulama açısından çok önemlidir. GMACA'nın örüntü tanıma [23] ve sınıflandırma [21] uygulamaları çoğu kez tek uzunluk döngülü cezbediciler kullanılarak gerçekleştirilir. Bu çalışmada

GMACA'nın meydana getirdiği “cezbediciler” video görüntülerindeki hareket tespiti için kullanılmıştır.

3.2.1. Genelleştirilmiş çoklu cezbedicili hücresele otomatlar (Generalized multiple attractor cellular automata)

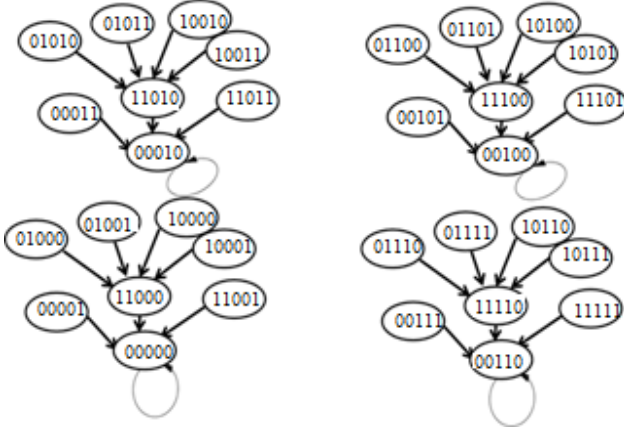
Öncelikle tek uzunluk döngülü cezbedicilerin nasıl elde edildiğini anlaşılması gerekir. Bunun için ilk önce hücrenin bir sonraki durumunu belirlemede kilit öneme sahip Kural Min Terimi(Rule Min Term) kavranmalıdır. Kural Min Terimi(KMT), sıralı bir şekilde yan yana olan üç hücrenin bir sonraki durumunu belirleyen fonksiyon olarak düşünülebilir. T zamanındaki bir hücrenin değerini S_i^t ile gösterilirse hücrenin bir sonraki durumunun belirlenmesini fonksiyon biçiminde tanımlanması şöyle olacaktır; $S_i^{t+1}=f(S_{i-1}^t+S_i^t+S_{i+1}^t)$ Bu fonksiyonun aldığı üç değer kısaca KMT olarak adlandırılmaktadır. Burada fonksiyon, bir sonraki durumu belirlemek için kullanılan kuraldır. Üç komşu ve iki durumlu hücresele otomatların sekiz farklı KMT'si vardır; KMT0, KMT1, ... , KMT7. Kurallar ve KMT'ler arasındaki ilişki aşağıdaki Tablo 4'e bakarak daha iyi anlaşılabilir. Örneğin t zamanındaki 101 durumunun ortasındaki değer için bir sonraki durum kuralların KMT5'i ne göre belirlenmektedir. 101 durumunun ortasındaki hücrenin bir sonraki durumu Kural6, Kural60, Kural204 ve Kural160'a göre sırasıyla şöyledir; 0, 1, 0 ve 1. Tablo 4.'te bu sütün gri arka planla gösterilmektedir. Araştırmacılar, kuralın "cezbedici" oluşumuna katkısını açıklayan bir tanım yapmışlardır. Bu tanıma "Özellik 1" ismi verilmiştir [21].

Özellik 1: Bir kural; KMT0, KMT1, KMT4 veya KMT5'ten en az biri 0 ise veya KMT2, KMT3, KMT6 veya

KMT7'den en az bir 1 olduğunda tek uzunluk döngülü cezbedici oluşumuna katkıda bulunabilir [21]. Kurallar tek uzunluk döngülü cezbedici oluşumuna yaptıkları katkıya göre dokuz gruba ayrılmıştır. Gruplar Grup0, Grup1, . . , Grup8 biçiminde adlandırılmıştır. Kurallar “özellik 1” karşılayan KMT sayılarına göre gruplarda yer alır. Örneğin; Kural6 özellik 1'i dört KMT'sinde karşıladığı için Grup4'de yer alır. Kural204 özellik 1'i tüm KMT'lerinde karşıladığı için Grup8'de yer alır. CA<6,60,204,204,160> kural vektörü tarafında üretilen durum geçiş diyagramı Şekil 7'de yer almaktadır.

Tablo 4. Ortadaki hücrenin Kural6, Kural60, Kural204 ve Kural160 kurallarına göre bir sonraki durumu (Next state of the middle cell according to Rule6, Rule60, Rule204 and Rule160)

	111	110	101	100	011	010	001	000
Kural Min Terimi(KMT):	7	6	5	4	3	2	1	0
Ortadaki hücrenin Kural6'ya göre t+1 zamanındaki durumu	0	0	0	0	0	1	1	0
Ortadaki hücrenin Kural0'a göre t+1 zamanındaki durumu	0	0	1	1	1	1	0	0
Ortadaki hücrenin Kural204'e göre t+1 zamanındaki durumu	1	1	0	0	1	1	0	0
Ortadaki hücrenin Kural160'a göre t+1 zamanındaki durumu	1	0	1	0	0	0	0	0



Şekil 7. CA<6,60,204,204,160> kural vektörüne ait durum geçiş diyagramı. Kural vektörü dört tane tek uzunluk döngülü cezbedici üretmiştir (State transition diagram of the rule vector CA <6,60,204,204,160>. The rule vector produced four singlelength cycle attractors.)

Hareket tespiti, video görüntülerindeki ardışık iki nokta veya bölge arasındaki renk değişikliğine odaklanır. Değişimin tespit edildiği yerlerin hareket içerdiği kabul edilir [25]. K cezbedici durumuna sahip GMACA, videolardaki nesnenin K sayıda farklı durumu olarak değerlendirilebilir. Bu bağlamda nesne, t zamanındaki tüm video görüntüsünü temsil eder. Nesnenin ardışık çerçevelerdeki değişimi cezbedicilerin değişimi olarak ele alınabilir. Bir pikselin ardışık çerçevelerde bir durumdan diğerine geçmesi durumunda hareket algılaması yapılır. Ardışık çerçevelerde piksel durumunda herhangi bir değişiklik olmazsa, hareket algılaması gerçekleştirilmez. GMACA ile hareket algılamasının önemli noktası, görünümün uygun sayıda cezbediciyle temsil edilmesini sağlamaktır. Uygun cezbedici sayısının elde edilmesi kural vektörünün tasarımı ile ilgilidir. Yukarıdaki bölümde, kural vektöründeki kuralların cezbedici oluşumunu nasıl etkilediğini belirtilmiştir. Bu çalışmada, [21] tarafından geliştirilen algoritmalara göre kural vektörü elde edilmiştir.

Yukarıdaki bölümde belirtildiği gibi iki farklı uzunluğa sahip cezbediciler vardır. Hareket tespit amacına uygun olarak bu çalışmada tek uzunluk döngülü cezbediciler kullanılmıştır. Das vd. geliştirdikleri tek uzunluk döngülü cezbediciler üreten kural vektörü elde etme yöntemi $O(n)$ zaman karmaşıklığına sahiptir [21]. Geliştirdikleri yöntem ikili bir ağaç olan erişilebilirlik ağacı tabanlıdır. Erişilebilirlik ağacındaki yapraklara kadar olan her bir seviye kural vektöründeki kuralların soldan sağa taranması ile oluşturulur. Ayrıca her bir seviye ikili örüntüdeki bir biti temsil eder. Eğer ilgili seviyede yer alan düğüm kapsayıcı bit ise, n tane cezbediciyi daha az bitle temsil etme kapasitesine sahiptir. Eğer kural vektörü m tane kapsayıcı bite sahip ise, bu bitler 2^m tane cezbedici durumu temsil edebilir. Matematiksel bir denklemlerle gösterilecek olursa kapsayıcı bitlerin temsil edebileceği bit sayısı şöyledir; $2^m = n$ dir. Bu m tane kapsayıcı bite sahip kural vektörünün ürettiği cezbedici bitlerin $O(\log_2 n)$ hafıza karmaşıklığına sahip olduğunu gösterir. Bu konuda detaylı bilgi Das vd. [21]

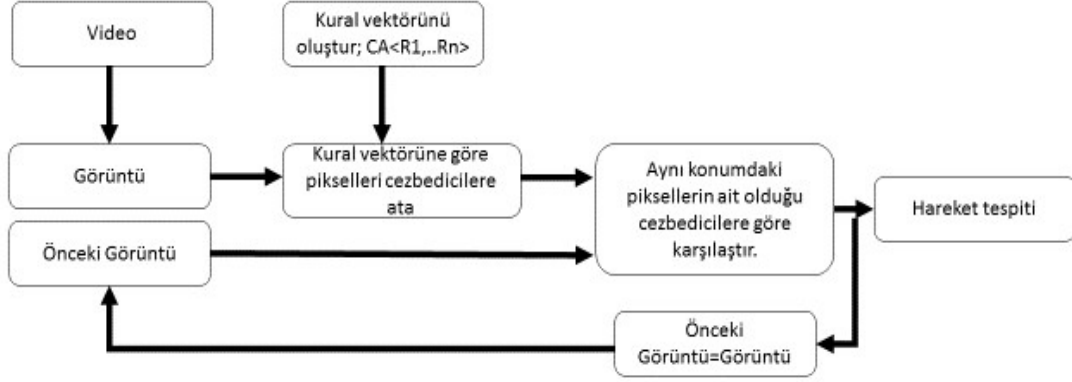
yaptıkları çalışmada bulunabilir. Önceki araştırmalarda GMACA örüntü tanıma [26], örüntü üretme [27], yüz tanıma [28] ve boyut azaltma [29] gibi bilgisayarlı görüntü uygulamalarında kullanılmıştır. Önceki çalışmalarda belirtildiği gibi kullanılacak kural vektörünün uzunluğu sekizden az olmamalıdır. Bunun nedeni, yoğunluk görüntülerindeki piksel değerlerinin 0 ile 255 arasında değerleri almasıdır. Sekiz uzunlukta bir kural vektörü, ondalık değerlerin sekiz uzunluğa sahip ikili değerlerine uygulanabilir. Sekiz uzunlukta ikili değerler 0 ile 255 arasında bir ondalık değer içerebilir. Bu nedenle, kural vektörünün uzunluğu en az sekiz olmalıdır. Kural vektörünün sekizden fazla uzunluğu olabilir. Bununla birlikte, şunu belirtmek gerekir ki: Kural vektörü tarafından üretilen cezbedici durumların 0 ile 255 arasında değerleri olmalıdır. Hareket tespit yöntemi, ardışık görüntülerin karşılaştırmasına dayanmaktadır. Bu açıdan ardışık görüntülerin farkına dayalı hareket tespit etme yöntemleri altında değerlendirilebilir. Kullanılan yöntemin daha önceki yapılan çalışmalardan en önemli farkı GMACA ile cezbediciler kullanılarak yapılmasıdır. Yöntemde ardışık görüntüler arasındaki farka dayalı hareket tespiti yapılmıyor, bunun yerine ardışık iki görüntü arasında karşılık gelen piksel konumlarının aynı cezbediciye ait olup olmamasına göre yapılıyor. Daha açık bir ifade ile ardışık iki görüntünün karşılık gelen pikselleri aynı cezbediciye ait ise hareket olmadığı, farklı ise hareket olduğu sonucuna varılır.

Hareket tespit yöntemin adımları şöyle sıralanabilir;

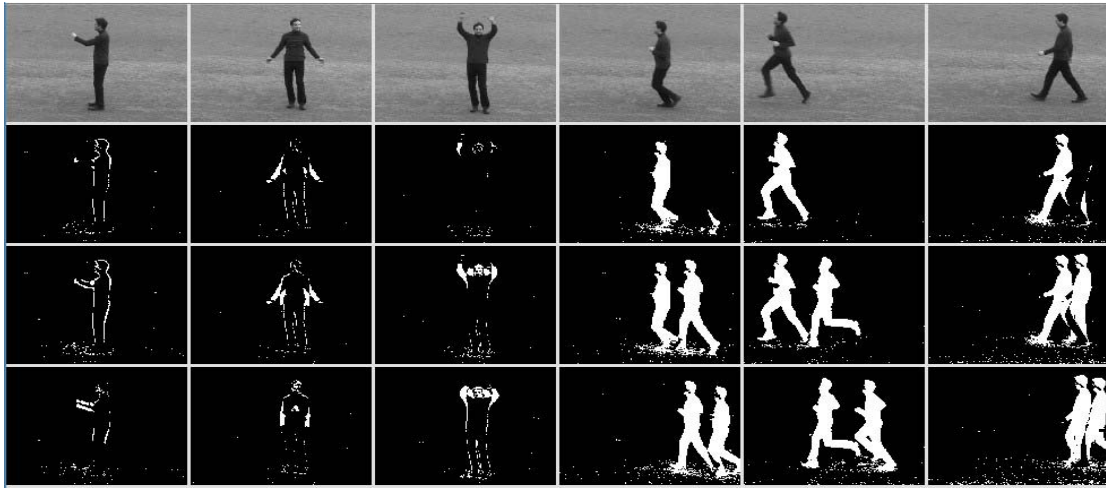
- İstenen sayıda cezbedici üreten kural vektörü belirlenir. Kural vektörünün nasıl belirlendiğini yukarıdaki bölümlerde detaylıca anlatılmıştır.
- O anki görüntünün piksel değerlerini seçilen kural vektörüne göre cezbedici durumlara atanır.
- Bir önceki görüntü ile o anki görüntü karşılaştırılır. Eğer görüntülerin aynı konuma karşılık gelen pikselleri aynı cezbediciye sahip ise hareket olmadığı, aynı cezbediciye sahip değil ise hareket olduğu sonucuna varılır. Karşılaştırma işleminde aynı konumdaki piksellerin ait oldukları cezbedici duruma bakılır. Eğer piksellerin cezbedici durumunda bir değişim var ise bu hareket tespiti olduğunu gösterir.

Eğer piksellerin cezbedici durumunda bir değişim yok ise bu hareket tespiti olmadığını gösterir. Kullanılan yöntemin blok diyagramı Şekil 8.'de gösterilmektedir;

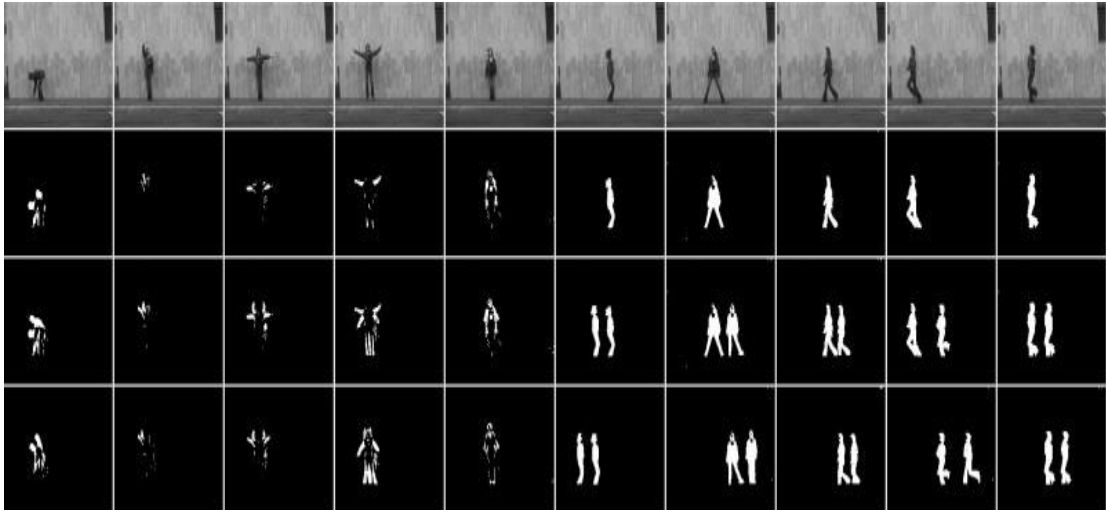
Şekil 9. ve Şekil 10.'da GMACA kullanılarak elde edilen, KTH ve WEIZMANN veri tabanlarına ait hareket tespit görüntüleri yer almaktadır. Şekillerde yer alan resimler yatay ve dikey bölümlere ayrılmıştır. Dikey bölümlerin her birisi ayrı bir hareketi göstermektedir. Yatay bölümler orijinal hareket görüntülerini ve farklı hareket tespit görüntülerini göstermektedir. Yatay ve dikey bölümlendirmeyi sırasıyla satır ve sütun olarak da düşünebilir. Sütunlarda KTH veri tabanı için sırasıyla boks, alkış, el sallama, tempolu koşu, normal koşu ve yürüme hareketlerine ait görüntüler yer almaktadır.



Şekil 8. Kullanılan hareket tespit yönteminin blok diyagramı (Block diagram of used motion detection method)



Şekil 9. KTH veri tabanında yer alan altı farklı hareket için hareket tespit görüntüleri (Motion detection images for six different movements in the KTH database)



Şekil 10. WEIZMANN veri tabanında yer alan on farklı hareket için hareket tespit görüntüleri (Motion detection images for ten different movements in the WEIZMANN database)

WEIZMANN veri tabanı için sırasıyla eğilme, tek el sallama, iki el sallama, jack(iki el ve iki bacağın yerinde zıplayarak beraber açılıp kapanması), yerinde zıplama, ileri

doğru zıplama, yan yürüme, yürüme, koşma ve tek ayak üzerinde sekerek yürüme hareketlerine ait görüntüler yer almaktadır. Satırlarda sırasıyla orijinal görüntü $I(x, y, t)$, t

zamanındaki gerçek hareket tespit görüntüsü; $D_g(x, y, t)$, t zamanındaki hareket tespit görüntüsü $D(x, y, t)$ ve $t-1$ zamanındaki hareket tespit görüntüsü; $D(x, y, t - 1)$ yer almaktadır.

Hareket tespit görüntülerinin elde edilmesini aşağıda verilen Eş. 3, Eş. 4, Eş. 5 ve Eş. 6 eşitliklerinde görülmektedir.

$$D_{(x,y,t)} = \begin{cases} 0, & \text{GMACA}(I(x,y,t)) = \text{GMACA}(I(x,y,t-1)) \\ 255, & \text{GMACA}(I(x,y,t)) \neq \text{GMACA}(I(x,y,t-1)) \end{cases} \quad (3)$$

$$D_{(x,y,t-1)} = \begin{cases} 0, & \text{GMACA}(I(x,y,t-1)) = \text{GMACA}(I(x,y,t-2)) \\ 255, & \text{GMACA}(I(x,y,t-1)) \neq \text{GMACA}(I(x,y,t-2)) \end{cases} \quad (4)$$

$$D_g(x,y,t-1) = D_{(x,y,t)} \cap D_{(x,y,t-1)} \quad (5)$$

$$D_g(x,y,t) = D_{(x,y,t)} \cap D_g(x,y,t-1) \quad (6)$$

Eş. 3 ve Eş. 4 ile GMACA kullanılarak t zamanına ve $t-1$ zamanına ait hareket tespit görüntüleri elde edilir. Eş. 5 ile $t-1$ zamanına ait gerçek hareket tespit görüntüsü elde edilir. Eş. 6 ile t zamanına ait gerçek hareket tespit görüntüsü elde edilir.

3.3. HOG Özelliklerinin Çıkarılması (HOG Feature Extraction)

HOG ilk defa, Fransız Ulusal Bilgisayar Bilimi ve Kontrol Enstitüsü (INRIA) araştırmacılarından Navneet Dalal ve Bill Triggs tarafından geliştirilmiştir [30]. HOG, görüntülerdeki piksel değerlerinin yerel eğilimlerine veya kenarların yönüne göre görüntülerden öznitelik elde eder. HOG tanımlayıcılarının arkasındaki temel fikir, bir görüntü içindeki yerel nesne görünümünün ve şeklinin yoğunluk eğilimlerinin veya kenar yönlerinin dağılımı ile tanımlanabilmesidir. HOG bilgisayarla görü ve görüntü işlemede sıkça kullanılmaktadır [31, 32].

HOG yönteminin görüntülerden öznitelik çıkarımı şöyle özetlenebilir: Görüntü hücre (cell) adı verilen birbirine bağlantılı küçük bölümlere ayrıştırılır. Blok olarak adlandırılan hücre bölümlerini içine alan daha geniş yapılar oluşturulur. Her bir blok için piksel değerlerinin yönelimi hesaplanır. Bu yönelimlerden eğim açıları elde edilir. Elde edilen iki farklı eğilimden eğilim histogramı oluşturulur. Yöntemin ana fikrini bu eğilim histogramının elde edilmesi oluşturur. Bu nedenle yöntem adını bu işlemde alır. Görüntüdeki ışık farklılıklarını en aza indirmek için elde edilen histogram normalleştirme işlemine tabi tutulur. Her bir blok için elde edilen normalleştirilmiş histogramlar birbirine eklenerek görüntüye ait öznitelik vektörü elde edilir [33]. Bu çalışmada HOG öz niteliklerini hareket tespiti yapılan görüntülerden elde edilmiştir. Özniteliklerin çıkarma işlemini dört saniyede bir gerçekleştirilmiştir. KTH veri tabanını oluşturan araştırmacıların belirttiği gibi veri setindeki hareketler dört saniyede bir gerçekleşmektedir. Makine öğrenme modelleri veriler üzerinde tanıma işlemi yapmak için iki şeye ihtiyaç duyar; birincisi özellik vektörüne, ikincisi özellik vektörünün ait olduğu sınıfın

etiketine. Elde edilen HOG özellik vektörüne ilgili hareketin etiketi atanır. Etiket atama işlemi doğrudan sınıf adının atanması yerine sınıflara verilen numaraların atanmasıyla gerçekleştirilir. Örneğin yapılan uygulamada var olan altı harekete 0'da 5'e kadar numaralar atanır.

3.4. Çapraz Doğrulama (Cross Validation)

Veri biliminde birçok modelleme tekniği vardır. Asıl soru hangisinin daha iyi olduğudur. İyi bir modelin iki özelliği vardır: İyi tahmin gücüne ve görmediği veriler üzerinde genelleme yeteneğine sahiptir. Bunu başarmak için bir hata ölçütü (modelin ne kadar yanlış olduğunu) ve bir doğrulama stratejisi tanımlanması gerekir [34, 35].

Çapraz doğrulama, eldeki veriyi eğitim ve test setine ayırmadan genelleme performansını daha istikrarlı ve kapsamlı bir şekilde artırmak için geliştirilmiş istatistiksel bir yöntemdir. Çapraz doğrulamada, veriyi eğitim ve test setine bölmek yerine, veriler tekrar tekrar bölünür ve model birden çok defa test edilir [36]. Çapraz doğrulama için çeşitli yöntemler önerilmesine rağmen, K-katlı(K-fold) çapraz doğrulama yaygın şekilde kullanılmaktadır. Bu yöntemde, eğitim seti K eşit parçaya bölünür ve bunlardan biri doğrulama seti olarak kullanılır. Kalan K-1 kısımları eğitim setini oluşturmaktadır. Her seferinde doğrulama seti olarak farklı bir bölüm kullanılır. Bu işlem K kez tekrar edilir ve doğrulama seti üzerindeki başarısı her defasında hesaplanır. Böylece, K tane başarı değeri elde edilir. Elde edilen K tane başarı değerinin ortalaması modelin başarısı kabul edilir [37]. Doğrulama son derece önemlidir çünkü modelinizin gerçek yaşam koşullarında çalışıp çalışmadığını belirler. Daha açık bir ifade ile önerilen modelin bir değerinin olup olmadığını gösterir [34]. Eldeki veri setini bir eğitim ve test setine bölme yerine çapraz doğrulamayı kullanmanın birçok avantajı vardır. Verimizi eğitimi ve test setlerine ayırdığımızda modellemesi zor verilerin eğitimde modellenmesi kolay verilerin test setinde ya da tam tersi gelme ihtimali vardır. Bu önerilen modelin gerçek performansının çok altında veya üstünde sonuçlar üretmesine yol açacaktır.

Çapraz doğrulama, veri setindeki her örneğin eğitim setinde ve test setinde kullanılmasını garanti edecektir. Dolayısıyla, modelin tüm çapraz doğrulama skorlarının (ve ortalamalarının) yüksek olması için veri kümesindeki tüm örnekler üzerinde iyi sonuç vermesi gerekir. Ayrıca çapraz doğrulama önerilen modelin vereceği başarı sonucunun aralığını belirler [38]. Örneğin geliştirilen öğrenme modeli 5 farklı eğitim ve test setlerini kullanarak en kötü %85 en iyi %95 başarı oranı ortaya koydu. Bu model için kullanılan veriye göre %10'luk bir başarı değişikliği olduğu varsayımında bulunabilir. Çapraz doğrulama yaklaşımları, hesaplama açısından maliyetli olabilir, çünkü eğitim verilerinin farklı alt kümelerini kullanarak aynı makine öğrenmesi yöntemi birden çok kez test edilir. Bununla birlikte, hesaplama gücünün geldiği nokta, yeniden örnekleme yöntemlerinin hesaplama gereksinimlerini genelde karşılayabilecek seviyededir [39].

3.5. Destek Vektör Makineleri (Support Vector Machine (SVM))

SVM yönteminde amaç iki sınıfı birbirinden ayıran doğruyu bulmaktır [40]. Bu doğru çok boyutlu sınıfları birbirinden ayırdığı için bir hiperdüzlem olmalıdır. Bu hiperdüzlemin doğrusal ayrılabilen örnekler üzerinde tanımlanması N tane örneğe sahip bir eğitim kümesi için Eş. 7’de olduğu gibi tanımlansın;

$$X = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\} \\ | x_i \in \mathbb{R}^d, y \in \{+1, -1\} \quad (7)$$

Denklemden d, x vektörünün özellik sayısı yani vektörün uzunluğunu, y ise özellik vektörünün ait olduğu sınıfı göstermektedir. SVM iki sınıflı bir sınıflandırma yöntemi olduğu için y’nin sınıf etiketleri +1 ve -1’dir.

Bu iki sınıfı birbirinden ayıran hiperdüzleme karar fonksiyonu olarak kabul edilirse, tanımlaması Eş. 8’deki gibi şöyle olacaktır;

$$f(x) = w^T x_i + b \quad (8)$$

Denklemden w^T ağırlık vektörüdür ve b bias değeridir. Sınıflandırılması istenen x vektörü f(x) fonksiyonuna sokulur ve fonksiyon ağırlık vektörü w^T ile x vektörünün çarpımının sonucu bias değeri ile toplanır ve bir değer döner. Bu değer 0’dan büyük ise x vektörünün + sınıfına ait olduğu, sıfırdan küçük ise eksi sınıfına ait olduğu kabul edilir. Bu sınıflandırma işlemi Eş. 9’daki gibi şöyle gösterebilir;

$$\text{sign}(f(x)) = \begin{cases} +, & \text{Eğer } f(x) > 0 \\ -, & \text{Eğer } f(x) < 0 \end{cases} \quad (9)$$

Bu çalışmada SVM için kullanılan çoklu sınıf stratejisi one-to-rest’dir.

4. DENEYSEL SONUÇLAR (EXPERIMENTAL RESULT)

Bu bölümde yapılan hareket tanıma çalışmasının sonuçlarına yer verilecektir. Sonuçlar sınıflandırma performansını detaylı gösteren karışıklık matrisi ve bu matris kullanılarak elde edilen istatistiksel metrikler ile beraber verilmiştir. İlk önce, sınıflandırma modelinin her bir sınıf için başarısını gösteren karışıklık matrisini anlamak faydalı olacaktır.

4.1. Karışıklık Matrisi (Confusion Matrix)

Karışıklık matrisi, sınıflandırma algoritmasının performansını özetlemek için kullanılan bir tablolama tekniğidir [41]. Karışıklık matrisi neden ihtiyaç vardır?

Bunun nedeni; sınıflandırma sonuçlarının sadece başarı oranına göre değerlendirilmesinin yanıltıcı olmasıdır. Başarı oranı doğru yapılan sınıflandırma sayısının toplam sınıflandırma sayısına bölümünden elde edilir. Sınıflandırma başarısı metriğindeki asıl eksiklik, sınıflandırma modelinin performansını daha iyi anlamak için gereken ayrıntıları gizlemesidir. Elimizdeki verilerde ikiden fazla sınıf varsa başarı oranı hangi sınıfın ne oranda başarı oranına sahip olduğu bilgisini veremez. Ayrıca veri setindeki sınıflar eşit sayıda örneğe sahip değilse başarı oranı sınıflandırma modelinin başarısını tam yansıtamaz. Sınıflandırma başarı oranı, modelin performansını belirlemek için ihtiyaç duyulan ayrıntıları gizleyebilir. Karışıklık matrisi kullanarak bu ayrıntıyı ayırt edebilir. Karışıklık matrisi, sınıflandırma modelinin haklı çıktığı durumları ve hangi hataların oluştuğu hakkında daha iyi bir fikir verir.

Karışıklık matrisi geliştirilen sınıflandırma modelinin tahmin sonuçlarını özetleyen bir tablodur [42]. Doğru ve yanlış tahminler her bir sınıf için gösterilir. Sınıflandırma modelinin yaptığı yanlış tahminlerin detayını verir. Sadece, sınıflandırma modelinin hata yaptığını söylemek yerine hatayı nasıl yaptığını söyler [42-44]. Böylece “başarı oranı sonucunda” gizli kalan detaylar karışıklık matrisi kullanılarak ortaya çıkar. Karışıklık matrisinde yer alan hücrelerin genel bir adlandırması vardır. Bu adlandırma her bir sınıf için ayrı bir anlam ifade eder. Bu adlandırma, her problem için aynı terimleri kullanma imkânı veriyor. Tablo 5’te iki sınıf için karışıklık matrisindeki hücrelerin adlandırması verilmiştir.

Tablo 6.’ye göre bazı istatistiksel metriklerin nasıl hesaplandığı Eş. 10, Eş. 11, Eş. 12 ve Eş. 13’de gösterilmektedir;

$$\text{Başarı Oranı (Accuracy Rate)} = \frac{TP+TN}{TP+TN+FP+FN} \quad (10)$$

$$\text{Doğru Pozitif Oranı (Recall)} = \frac{TP}{TP+FN} \quad (11)$$

$$\text{Kesinlik (Precision)} = \frac{TP}{TP+FP} \quad (12)$$

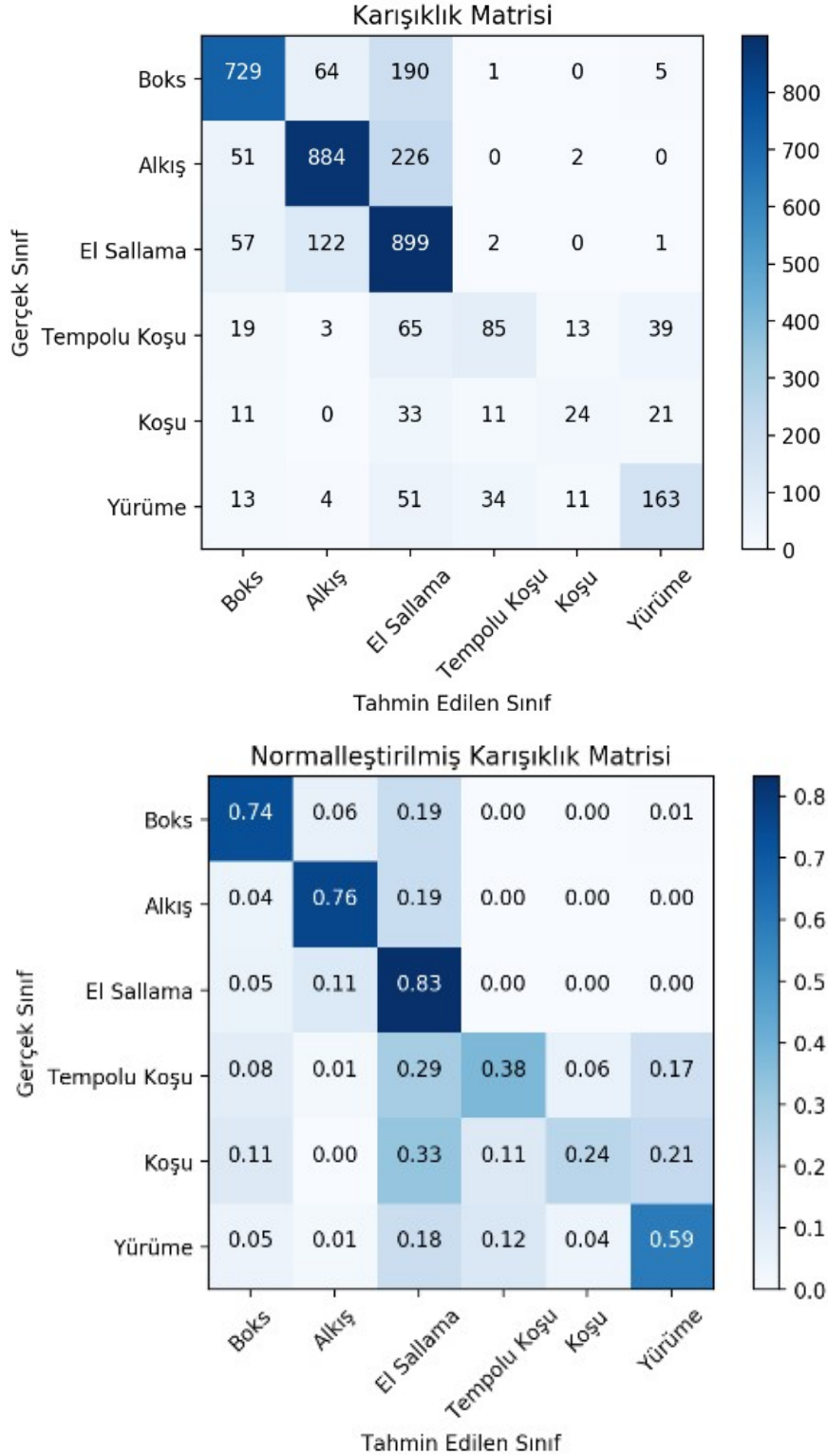
$$\text{F1 skor} = \frac{2TP}{2TP+FP+FN} \quad (13)$$

4.2. Deneysel Sonuçlar (Experimental Result)

Bu kısımda KTH ve WEIZMANN hareket tanıma veri tabanları üzerinde elde edilen hareket tanıma sonuçlarına yer verilecektir. GMACA hareket tespit görüntüleri kullanılarak

Tablo 5. İki sınıf için karışıklık matrisi (Confusion matrix for two classes)

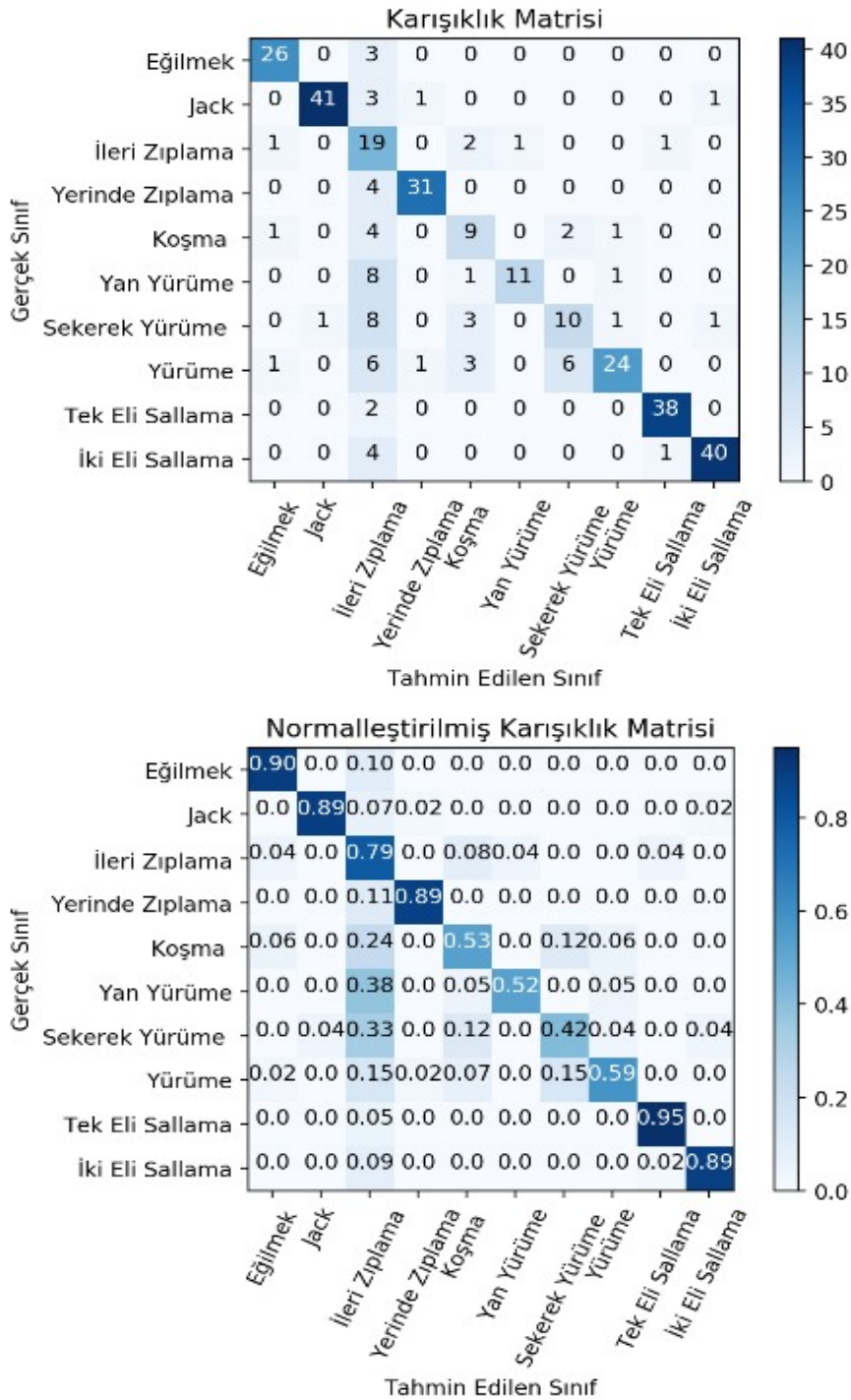
		Tahmin Edilen	
		Sınıf-A	Sınıf-B
Gerçek	Sınıf-A	True Positive (TP)	False Positive (FP)
	Sınıf-B	False Negative (FN)	True Negative (TN)



Şekil 10. KTH veri tabanı için önerilen GMACA hareket tespit yöntemine göre karışıklık matrisleri
(Confusion matrixes based on GMACA motion detection method we proposed for KTH database)

elde edilen tanıma sonuçlarının karışıklık matrisleri, normalleştirilmiş karışıklık matrisleri ve sınıflandırma raporları verilmiştir. Şekil 10'da KTH ve Şekil 11'de

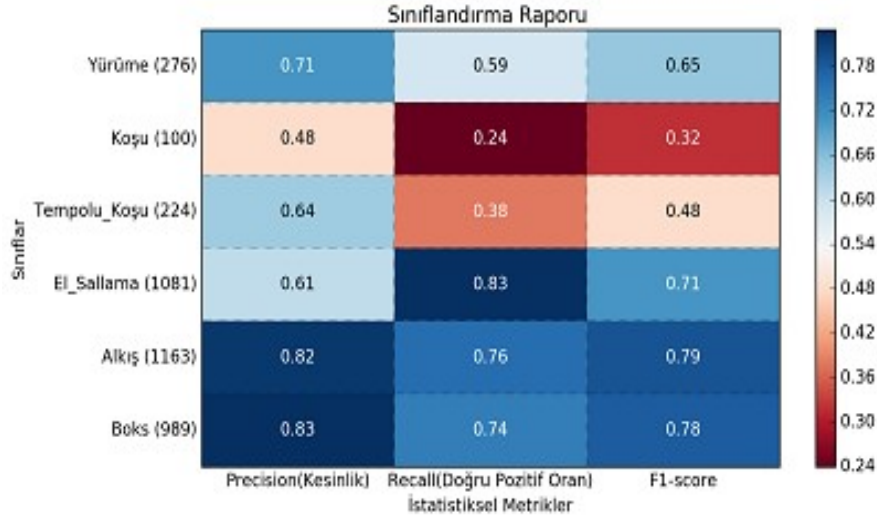
WEIZMANN veri tabanları için elde edilen karışıklık matrisi ve karışıklık matrisinin normalleştirilmiş hali verilmiştir. Karışıklık matrisinde sayısal değerler daha iyi anlaşılırken,



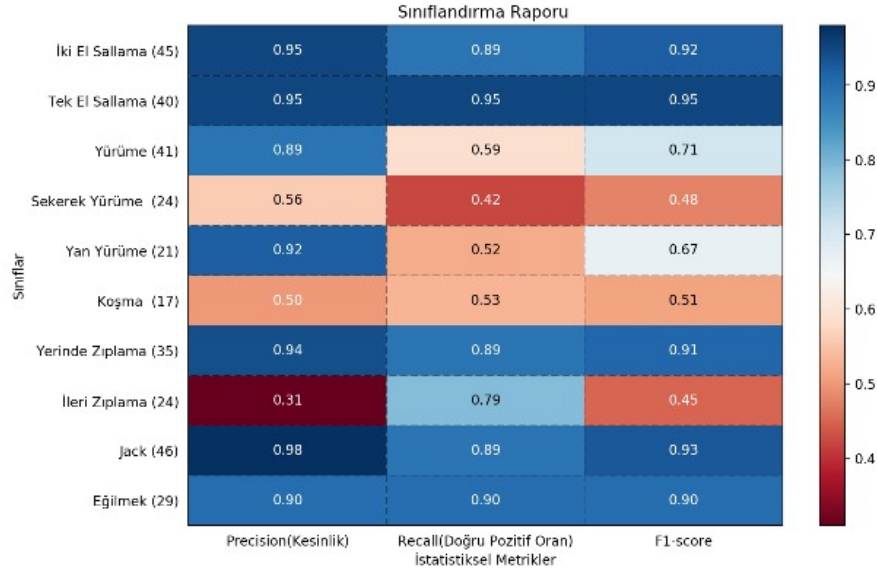
Şekil 11. WEIZMANN veri tabanı için önerilen GMACA hareket tespit yöntemine göre normalleştirilmiş karışıklık matrisleri (Normalized confusion matrixes based on GMACA motion detection method proposed for WEIZMANN database)

normalleştirilmiş karışıklık matrisinde başarı oranları daha iyi anlaşılmaktadır. Şekil 12’de KTH veri tabanı için ve Şekil 13’de WEIZMANN veri tabanı için karışıklık matrisi kullanılarak elde edilen sınıflandırma raporları yer almaktadır. Sınıflandırma raporlarında Doğru Pozitif Oranı(Recall), Kesinlik(Precision) ve F1 skor sonuçları verilmiştir. Sonuçlar söz konusu çalışma ile elde edilmiş olup Eş. 11, Eş. 12 ve Eş. 13 eşitlikleri ile hesaplanmıştır.

Şekil 10 ve 12’de KTH veri tabanındaki videolarda yer alan altı farklı insan hareketi için karışıklık matrisleri ve sınıflandırma raporları verilmiştir. Videolarda insan tespiti yapıldıktan sonra veri seti üretilmiştir. Veri seti; özellik vektörünün elde edilmesi ve elde edilen özellik vektörüne ilgili hareketin atanmasıyla oluşturulmuştur. Yukarıdaki tablolarda görüldüğü gibi her bir harekete ait farklı sayıda veri seti oluşturulmuştur. Bunun nedeni özellik çıkarımının



Şekil 12. KTH veri tabanı için GMACA hareket tespitine göre Sınıflandırma Raporu
(Classification Report according to GMACA motion detection for KTH database)



Şekil 13. WEIZMANN veri tabanı için GMACA hareket tespitine göre Sınıflandırma Raporu
(Classification Report according to GMACA motion detection for WEIZMANN database)

video görüntülerindeki insan nesnesinin tespitinden sonra gerçekleştirilmesinden kaynaklanmaktadır. En az veri seti “koşma” eylemine aittir. Bunun iki nedeni vardır: Birincisi “koşma” eyleminin hızlı gerçekleşmesi, ikincisi ise koşan bir kişinin tespit edilmesinin zorluğudur. KTH veri tabanındaki koşma eyleminde kişiler kameraya paralel yürüdükleri için tespit edilmesi zor olmuştur. En çok veri seti “alkış” eylemine aittir. “Alkış” eylemi sabit duruşta gerçekleştirildiği için insan tespiti kolay olmuştur. Ayrıca alkış yapan kişi daima görüntü içerisinde olduğu için daha çok veri setinin oluşmasına neden olmuştur. Diğer hareketlerde “koşma” ve “alkış” eylemleri için verilen bilgiler dikkate alınarak değerlendirilebilir. Şekil 11 ve 13’te WEIZMANN veri tabanındaki videolarda yer alan on farklı insan hareketi için karışıklık matrisleri ve sınıflandırma

raporları verilmiştir. Veri setinin oluşturulması KTH veri tabanı için veri seti oluşturulmasından daha kolay olmuştur. Bunun nedeni WEIZMANN veri tabanındaki videoların tümünün aynı şartlarda çekilmiş olması ve hareket yapan insanın daima video içerisinde olmasıdır. Hareketler için oluşturulan veri setinde KTH veri tabanına göre daha dengeli bir dağılım söz konusudur. En az veri seti hızlı gerçekleşen koşma, sekerek yürüme ve yan yürüme hareketlerine aittir.

Şekil 10 ve 11’deki karışıklık matrisleri incelenerek sınıflandırma performansları daha iyi anlaşılabilir. En düşük başarı oranları her iki veri tabanı için yürüme ve koşma benzeri eylemlerde görünüyor. Bunun nedeni bu eylemlerin birbirine çok benzemesidir. Bu eylemlerin tümünde, vücut organları aynı şekilde çalışmaktadır. Eylemler arasındaki

fark eylemin gerçekleştirilme hızında ortaya çıkmaktadır. Bu eylemlere ait yanlış tanımanın genelde birbirleri arasında olduğu görülmektedir. Geliştirilen hareket tanıma yönteminin zayıf yanını göstermektedir. İleriki çalışmalarda hareketlerin zamansal özelliklerini dikkate alarak başarı oranı yükseltilmeye çalışılacaktır.

El sallama, alkış ve boks eylemleri benzer özellikler taşımaktadır. Bu eylemlerin hepsinde vücudun el ve kol uzuvları aktif kullanılmaktadır. Eylemler arasındaki fark kolların kullanımında görülmektedir. Bu eylemler arasındaki yanlış tanıma birbirleri arasında olmaktadır. Kol hareketlerinin daha iyi tespiti tanıma performansını artıracaktır. Bunun için vücut uzuvlarının kısmi hareket tespiti faydalı olabilir. İleriki çalışmalarda vücudun kısmi hareket tespitleri dikkate alınacaktır.

KTH veri seti üzerindeki genel başarı oranı %72,7'dir. Boks ve alkış eylemlerinin kesinlik oranları sırasıyla %83 ve %82'dir. WEIZMANN veri seti üzerindeki genel başarı oranı %77,3'tür. WEIZMANN veri setindeki yerinde yapılan eylemler için kesinlik oranları Şekil 13'teki sınıflandırma raporunda görüldüğü %98'e kadar çıkmaktadır. Bu oranlar GMACA kullanılarak yapılan hareket tespit yönteminin hareket tanıma uygulamalarında kullanılabilirliğini göstermektedir. Ayrıca GMACA eş zamanlı hesaplamaya imkân sunması bilgisayarlı görü üzerine çalışan araştırmacıların dikkatini çekeceği düşünülmektedir.

KTH ve WEIZMANN veri tabanları üzerinde yapılan daha önceki hareket tanıma çalışmalarının sonuçları Tablo 7 ve 8'de yer almaktadır.

Tablo 7. GMACA hareket tespiti kullanılarak geliştirilen yöntemin ve diğer yöntemlerin KTH veri tabanı üzerindeki başarı oranları

(The accuracy rate of the method developed by using GMACA motion detection and other methods on the KTH database)

Yöntem	Başarı Oranı (%)
Schüldt vd. [17]	71,63
Schindler ve Gool [45]	87,98
Baumann [46]	94,31
Laptev vd. [47]	91,8
O'Hara and Draper [48]	97,9
Geliştirilen Yöntem	72,70

Tablo 8. GMACA hareket tespiti kullanılarak geliştirilen yöntemin ve diğer yöntemlerin WEIZMANN veri tabanı üzerindeki başarı oranları

(The accuracy rate of the method developed by using GMACA motion detection and other methods on the WEIZMANN database)

Yöntem	Başarı Oranı (%)
Niebles ve Fei-Fei [49]	72,80
Dollar vd. [50]	86,70
Ali vd. [51]	92,60
Wang ve Suter [52]	97,80
Geliştirilen Yöntem	77,3

5. SİMGELER (SYMBOLS)

b	: Bias
$gri_{yoğunluk}$: Üç renk kanalının ortalamasına göre gri renge dönüştürme
$gri_{luminance}$: Renk kanallarını belirli ağırlıklara göre hesaplayıp gri renge dönüştürme
I	: Görüntü
$I(x, y, t)$: Görüntüdeki x ve y konumundaki pikselin t zamanındaki değeri
D	: Hareket tespit görüntüsü
$D(x, y, t)$: Hareket tespit görüntüsündeki x ve y konumundaki pikselin t zamanındaki değeri
D_g	: Gerçek hareket tespit görüntüsü
$D_g(x, y, t)$: Gerçek hareket tespit görüntüsündeki x ve y konumundaki pikselin t zamanındaki değeri
S_i^t	: i'ninci hücrenin t zamanındaki durumu
w	: Ağırlık vektörü
wT	: Ağırlık vektörünün transpozunu
X	

5.1. Kısaltmalar :

CA	: Cellular Automata
GMACA	: Generalized Multiple Attractor Cellular Automata
FN	: False Negative
FP	: False Positive
HOG	: Histogram of Gradient
KMT	: Kural Min Terimi
SVM	: Support Vector Machine
TN	: True Negative
TP	: True Positive

6. SONUÇLAR (CONCLUSIONS)

Bu çalışmada modern bir hesaplama modeli olan CA tabanlı hareket tespit yöntemi KTH ve WEIZMANN veri tabanlarındaki videolarda yer alan insan hareketlerinin tanınmasında kullanılmıştır. Hareket tespitinde CA'nın birden çok hücreye farklı kurallar kullanarak uygulanan GMACA modeli kullanılmıştır. GMACA tabanlı harekete tespit yönteminde ilk önce görüntülerin durum uzayında ele alınmasını sağlayacak kural vektörü belirlenir. Kural vektörü görüntülere uygulandıktan sonra görüntüler kural vektörünün ürettiği cezbedici durumlarla temsil edilir. Video görüntü çerçevelerindeki her bir görüntü cezbedici durumlarla temsil edilir. Hareket tespiti cezbedici durumlarla temsil edilen görüntülerin karşılaştırılmasıyla elde edilir. Ardışık görüntülerde aynı konumdaki hücrelerin durumları karşılaştırılır. Eğer piksel konumunda durum değişimi varsa hareket tespiti yapılır. Eğer piksel konumunda durum değişimi yoksa hareket olmadığı varsayılır.

GMACA ile elde edilen hareket tespit görüntüleri ile yapılan hareket tanıma uygulaması mevcut arka plan çıkarma çalışmaları kadar iyi sonuç vermiştir. Elde edilen sonuçlar,

GMACA ile yapılan hareket algılamanın hareket tanıma çalışmalarında ve gerçek zamanlı uygulamalarda kullanılabilirliğini göstermektedir. Bu çalışmada GMACA ile elde edilen hareket tespit görüntüleri kullanarak yapılan hareket tanıma uygulaması, geliştirilen yöntemin karmaşık bilgisayarlı görü uygulamalarında kullanılabilirliğini göstermektedir. Ayırık ve eş zamanlı hesaplamaya imkân sunan çalışma mantığı, hücreli otomatların görüntü işleme ve bilgisayarlı görü uygulamalarında kullanılmasını avantajlı hale getirmektedir.

GMACA'nın zayıf yanı ikili örüntüler üzerinde uygulanabiliyor olmasıdır. Kullanılan hareket tespit yöntemi onluk tabandaki piksel değerlerinin ikilik tabana dönüştürülmesinden sonra uygulanabilir.

KAYNAKLAR (REFERENCES)

1. Chen L., Wei H., Ferryman J., A Survey of Human Motion Analysis Using Depth Imagery, *Pattern Recognition Letters*, 34 (15), 1995-2006, 2013.
2. Aggarwal J., Ryoo M., Human activity analysis: A review, *ACM Computing Surveys (CSUR)*, 43 (3), 16-65, 2011.
3. Ahad M., *Motion History Images for Action Recognition and Understanding*, 1st ed., Springer, London, 2013.
4. Ahad M., *Computer Vision and Action Recognition*. 1st ed., Atlantis Press, 2011.
5. Weinland D., Ronfard R., Boyer E., A Survey of Vision-Based Methods for Action Representation, Segmentation and Recognition, *Computer Vision and Image Understanding*, 115 (2), 224-241, 2011.
6. Nagel H., From Image Sequences Towards Conceptual Descriptions, *Image and Vision Computing*, 6 (2), 59-74, 1988.
7. González J., Varona J., Roca F., Villanueva J., aSpaces: Action Spaces for Recognition and Synthesis of Human Actions, Editors: Perales F., Hancock E., *Articulated Motion and Deformable Objects*, Springer, Berlin, 189-200, 2002.
8. Fanti C., *Towards Automatic Discovery of Human Movemes*, Doktora Tezi, California Institute of Technology, California, 2008.
9. Krüger, V., Kragic D., Ude A., Geib C., The Meaning of Action: A Review on Action Recognition and Mapping, *Advanced Robotics*, 21 (13), 1473-1501, 2007.
10. Gaidon A., Harchaoui Z., Schmid C., *Action Sequence Models for Efficient Action Detection*, *Computer Vision and Pattern Recognition (CVPR)*, Providence, Amerika, 20-25 Haziran 2011.
11. Demirci M. F., Osmanhoğlu Y., Shokoufandeh A. and Dickinson S., Efficient many-to-many feature matching under the l1 norm, *Computer Vision and Image Understanding*, 115 (7), 976-983, 2011.
12. Ke S.-R., Thuc H. L., Lee Y.-J., Hwang J.-N., Yoo J.-H., Choi K.-H., A Review on Video-Based Human Activity Recognition, *Computers*, 2 (2), 88-131, 2013.
13. Ziaefard M. T., Bergevin R., *Semantic Human Activity Recognition: A Literature Review*, *Pattern Recognition*, 48 (8), 2329-2345, 2015.
14. Guo G., Lai A., A Survey on Still Image Based Human Action Recognition, *Pattern Recognition*, 47 (10), 3343-3361, 2014.
15. Ji X., Liu H., *Advances in View-Invariant Human Motion Analysis: A Review*, *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 40 (1), 13 - 24, 2010.
16. Gross R., Shi J., *The CMU Motion of Body (MoBo) Database*, CiteSeer, <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.16.5015> Yayın tarihi: 2001, Erişim tarihi 10 Ekim 2017).
17. Schuldt C., Laptev I., Caputo B., *Recognizing Human Actions: A Local SVM Approach*, *ICPR'04*, Cambridge, İngiltere, 23-26 Ağustos 2004.
18. Blank M., Gorelick L., Shechtman E., Irani M., Basri R., *Actions as Space-Time Shapes*, Tenth IEEE International Conference on Computer Vision (ICCV'05), Beijing, Çin, 17 Ekim 2005.
19. Kanan C., Cottrell G., *Color-to-Grayscale: Does the Method Matter in Image Recognition?*, *PLoS ONE*, 7 (1), e29740-29750, 2012.
20. Wolfram S., *A New Kind of Science*, Wolfram Media, Champaign, 2002.
21. Das S., Mukherjee S., Naskar N., Sikdar B., *Characterization of Single Cycle CA and its Application in Pattern Classification*, *Electronic Notes in Theoretical Computer Science*, 252 (1), 181-203, 2009.
22. Das S., *Theory and Applications of Nonlinear Cellular Automata In VLSI Design*, Doktora Tezi, The Bengal Engineering and Science University, Bengal, 2006.
23. Ganguly N., Maji P., Sikdar B.K., Chaudhuri P.P., *Generalized Multiple Attractor Cellular Automata (Gmaca) Model For Associative Memory*, *International Journal of Pattern Recognition and Artificial Intelligence*, 16 (7), 781-793, 2002.
24. Maji P., Shaw C., Ganguly N., Sikdar B.K. ve Chaudhuri P.P., *Theory and Application of Cellular Automata For Pattern Classification*, *Fundamenta Informaticae*, 58(1), 321-354, 2003.
25. Yu Z., Chen Y., *A Real-Time Motion Detection Algorithm for Traffic Monitoring Systems Based on Consecutive Temporal Difference*, *Asian Control Conference (ASCC)*, Hong Kong, Çin, 27-29 Ağustos 2009.
26. Peldek S. ve Becerikli Y., *Hücreli Otomata Tabanlı İkili Sınıflandırma*, *Bilişim Teknolojileri Dergisi*, 12 (1), 43-58, 2019.
27. Peldek S. ve Becerikli Y., *Using Multiple Attractor Cellular Automata as Pattern Generator in Images*, 8. *War Technologies Kongres*, Ankara, 2016.
28. Peldek S., *Hücreli Hareketli Tekniği İle Yüz Tanıma*, Yüksek Lisans Tezi, Karabük Üniversitesi, Fen Bilimleri Enstitüsü, Karabük, 2012.
29. Peldek S., Becerikli Y., *Çoklu Cezbedici Hücreli Otomatların Boyut Azaltmada Kullanımı*, *International*

- Conference on Computer Science and Engineering (UBMK), Antalya, Türkiye, 5-8 Ekim 2017.
30. Dalal N., ve Triggs, B., Histograms of Oriented Gradients for Human Detection, Computer Vision and Pattern Recognition (CVPR'05), San Diago, Amerika, 20-26 Haziran 2005.
 31. Akyol, K., Şen, B., Bayır, Ş., Cakmak, H. B., Assessing the importance of features for detection of hard exudates in retinal, Turkish Journal of Electrical Engineering & Computer Sciences, 25 (2), 1223 – 1237, 2017.
 32. Peker M., Altun H. Karakaya F., Hardware implementation of a scale and rotation invariant object detection algorithm on FPGA for real-time applications, Turkish Journal of Electrical Engineering & Computer Sciences, 24 (1), 4368 – 4382, 2016.
 33. Cerit B., Bölük S. A., Demirci M. F., Analysis of the effect of image resolution on automatic face gender and age classification, Signal Processing and Communication Application Conference (SIU), Zonguldak, 2016.
 34. Cielen D., Meysman A., Ali M., Introducing Data Science, Manning Publications, Shelter Island, 2016.
 35. Kaya Y., Ertuğrul Ö. F., A novel feature extraction approach for text-based language identification: Binary patterns, Journal of the Faculty of Engineering and Architecture of Gazi University, 31 (4), 1085-1094, 2016.
 36. Vehtari, A., Gelman, A., Gabry, J. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC," Statistics and Computing, 27 (5), 1413–1432, 2017.
 37. Murty M., Devi V., Introduction to Pattern Recognition and Machine Learning, World Scientific Publishing, Singapore, 2015.
 38. Mueller A., Guido S., Introduction to Machine Learning with Python, O'Reilly Media Inc, Sebastopol, 2016.
 39. James G., Witten D., Hastie T., Tibshirani R., An Introduction to Statistical Learning with Applications in R, Springer, New York, 2013.
 40. Akıncı G., Polat E., Koçak O. M., A video-based eye pupil detection system for diagnosing bipolar disorder, Turkish Journal of Electrical Engineering & Computer Sciences, 21 (2), 2367 – 2377, 2013.
 41. Bulut F., A new clinical decision support system with instance based ensemble classifiers, Journal of the Faculty of Engineering and Architecture of Gazi University, 32 (1), 65-76, 2017.
 42. López V., Fernández A., Herrera F., On the importance of the validation technique for classification with imbalanced datasets: Addressing covariate shift when data is skewed, Information Sciences, 257, 1-13, 2014.
 43. Kuyumcu B., OpenCV Görüntü İşleme ve Yapay Öğrenme, Level, Istanbul, 2016
 44. Utku A., Doğru İ.A., Permission based detection system for android malware, Journal of the Faculty of Engineering and Architecture of Gazi University, 32 (4), 1015-1024, 2017.
 45. Schindler K., Gool L.V., Action Snippets: How Many Frames Does Human Action Recognition Require?, IEEE Computer Vision and Pattern Recognition (CVPR), Anchorage, Amerika, 23-28 Haziran 2008.
 46. Baumann F., Action Recognition with HOG-OF Features, German Conference on Pattern Recognition, Berlin, Almanya, 13-14 Eylül 2013.
 47. Laptev I., Marszalek M., Schmid C., Rozenfeld B., Learning Realistic Human Actions From Movies, Computer Vision and Pattern Recognition (CVPR), Anchorage, Amerika, 23-28 Haziran 2008.
 48. O'Hara S., Draper B. A., Scalable Action Recognition with A Subspace Forest, Computer Vision and Pattern Recognition (CVPR), Providence, Amerika, 16-21 Haziran 2012.
 49. Nieves J. C., Fei-Fei, A Hierarchical Model of Shape and Appearance for Human Action Classification, IEEE Conference on Computer Vision and Pattern Recognition, Minneapolis, Amerika, 17-22 Haziran 2007.
 50. Dollar P., Rabaud V., Cottrell G., Belongie S., Behavior Recognition Via Sparse Spatiotemporal Feature, IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance, Beijing, Çin, 15-16 Ekim 2005.
 51. Ali S., Basharat A., Shah M., Chaotic Invariants for Human Action Recognition, IEEE 11th International Conference on Computer Vision, Rio de Janeiro, Brezilya, 14-21 Ekim 2007.
 52. Wang L., Suter D., Recognizing Human Activities from Silhouettes: Motion Subspace and Factorial Discriminative Graphical Model, IEEE Computer Vision and Pattern Recognition, Minneapolis, Amerika, 17-22 Haziran 2007.

