



Robot-Nesne Etkileşiminde Çok Kipli Hata Sezme

Arda İnceoğlu¹ , Gökhan İnce² , Yusuf Yaslan³ , Sanem Sariel^{4*} 

^{1,2,3,4} İstanbul Teknik Üniversitesi, Yapay Zeka ve Robotik Laboratuvarı, Bilgisayar ve Bilişim Fakültesi, İstanbul, Türkiye

sariel@itu.edu.tr, gokhan.ince@itu.edu.tr, yyaslan@itu.edu.tr, sariel@itu.edu.tr

Öz

Hizmet robotları için gündelik etkileşimlerinde yürütme, sensör veya ortamsal faktörlerden dolayı emniyetsiz durumlar oluşabilir. Bu tür durumlarda emniyetin sağlanması kritik öneme sahiptir. Bu durumların sezilebilmesi için sürekli eylem gözetimi ve hata sezme bileşenlerine ihtiyaç duyulmaktadır. Bu amaçla, bu çalışmada nesne etkileşim hatalarının sezilmesi için çok kipli bir eylem gözetimi ve hata sezme sistemi sunulmuştur. Tek bir sensör kipine bağımlı kalmak yerine, farklı tipte sensörlerden alınan gözlemler tümleştirilerek farklı hata senaryoları için hata sezme başarımı artırılmıştır. Önerilen sistemde iç algı, işitsel algı ve görsel algı kipleri birbirlerinden bağımsız olarak işlenerek semantik yüklemeler elde edilmiş ve bu yüklemeler hata sezme için birleştirilmiştir. İnsansı robotumuz ile masa üstünde yapılan deney sonuçlarına göre sensör verilerinin hata sezme katkılarının tamamlayıcı olduğu gözlenmiştir. Çok kipli sensör füzyonuyla hata sezme, tutma eylemi için %86, bırakma ve itme eylemleri için %95 oranında hata sezme başarımı ile tek kipli hata sezmeden daha başarılı sonuçlar üretmiştir.

Anahtar kelimeler: Bilişsel robotlar, hata sezme, çok kipli algılama.

Multimodal Failure Detection in Robot-Object Interaction

Abstract

Unsafe situations might arise for service robots in everyday manipulation settings due to operational, sensory or environmental factors. Ensuring safety is crucial for these settings. In order to detect these situations, onboard continuous execution monitoring and failure detection procedures are needed. To address these issues, we present a multimodal failure monitoring and detection system to detect manipulation failures. Rather than relying only on a single sensor modality, we consider integration of different modalities to get better detection performance in different failure cases. In our system, high level proprioceptive, auditory and visual predicates are extracted by processing each modality separately. Then, the extracted predicates are fused altogether. Experiments on our humanoid robot for tabletop manipulation scenarios indicate that the contributions of modalities are complementary of each other. Multimodal fusion-based failure detection outperforms the unimodal detection with 86% success rate for pick and 95% success rates for place and push actions.

Keywords: Cognitive robots, failure detection, multi-modal perception.

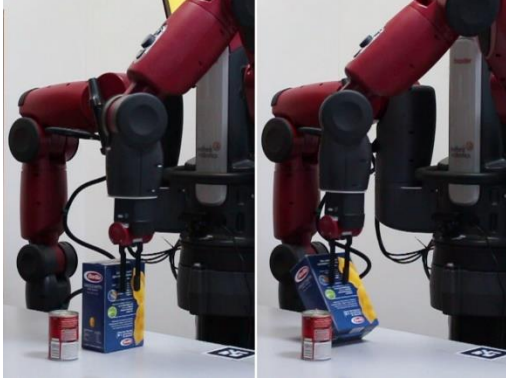
1. Giriş (Introduction)

Robotların kendilerine verilen görevleri insanlara, çevrelere ve kendilerine zarar vermeden tamamlamaları gerekmektedir. Fakat robotlar ev ve ofis gibi kendileri için tasarlanmayan düzensiz (unstructured) ortamlarda çalışırken algılamadaki ve ortamdaki belirsizlikler sebebiyle beklenmedik durumlar meydana gelebilir. Nesne etkileşiminde dengesiz tutuş sebebiyle nesnenin düşürülmesi, nesnenin taşınması sırasında ortamdaki diğer nesnelere ile çarpışması, sensör motor arayüzündeki farklar sonucu oluşan hatalar beklenmedik durumlara örnek

olarak verilebilir. Örnek bir senaryo Şekil 1' de görülmektedir. Bu senaryoda robot masa üzerindeki makarna kutusunu dengesiz bir noktadan itmeye çalışması sebebiyle kutu devrilmektedir. Robotun eylemi yürütmesi sırasında oluşan beklenmedik durumların tespiti için hem robotun ve hem de ortamın durumu sürekli olarak gözetilmelidir. Bunun için gözlemlenen durumdan hataları ve beklenmedik durumları tespit eden bir eylem gözetim sistemine ihtiyaç vardır. Endüstriyel ortamlarda çalışan robotların güvenliği için belli başlı standartlar (ISO 10218-2:2011, ISO 10218:2011) bulunmasına karşın mutfak gibi

* Sorumlu yazar.
E-posta adresi: sariel@itu.edu.tr

düzensiz ortamlarda çalışan otonom robotlar için bu durum henüz çözüme ulaşmamıştır.



Şekil 1. Hizmet robotunun kutuyu devirmesi (The service robot knocked over the object)

Bu çalışmada, çok kipli bir hata sezme sistemi önerilmiştir. Görsel, işitsel ve iç algı kipleri zamansal olarak birleştirilerek Saklı Markov Model yöntemi ile robotun yürüttüğü eylemler *başarılı* ve *hatalı* olarak sınıflandırılmıştır. Bu çalışma, daha önce ICRA-MRP18 çalıştayında (Inceoglu vd., 2018-1) ve IROS 2018 konferansında (Inceoglu vd., 2018-2) sunulan çalışmalardan derlenmiş ve TORK 2019 konferansında sunulmuştur. Çalışma, TÜBİTAK tarafından 115E368 no'lu proje kapsamında desteklenmiştir.

2. Literatür Özeti (Related Work)

Literatürde, eylem yürütme hatalarını sezme için modele dayalı ve modelden bağımsız yöntemler sunulmuştur. En yaygın olarak kullanılan modele dayalı yöntemde, eylemlerin yürütülmesi sonucunda ortamda oluşması beklenen etkilerin, eylemlerin gerçek sonuçları ile karşılaştırılması sağlanır (Pettersson, 2005). Modelden bağımsız yöntemlerde ise sensör bilgisi herhangi bir modelden bağımsız olarak işlenir (Pettersson vd., 2007). Hata sezme üzerine sunulan kapsamlı bir çalışmada (Fritz, 2005), hata sezme üç kategori altında incelenmiştir: dış gözlemlene, beklentiye dayalı gözlemlene ve modele dayalı gözlemlene.

Modele dayalı çalışmalar arasında (Mendoza vd., 2012) normal çalışma durumlarını modelleme yoluyla eyleyici (actuator) hatalarının tespiti (Fourlas vd., 2014) ve sensörlerden gelen bilginin odometri modeliyle karşılaştırılması yoluyla hata tespitinin yapılması sağlanmıştır (Stavrou vd., 2015). Başka bir çalışmada hata durumları, varlık bileşen sistemi (Entity-Component-System) ile saptanmıştır (Orendt vd., 2015). RoboCup yarışması için eylemlerin oluşturması beklenen çıktılar ile sensör bilgileri karşılaştırılmıştır (Mendoza vd., 2015). Başka bir çalışmada Focused Anomalous Region Optimization (FARO) algoritması (Mendoza vd., 2014) kullanılarak çapraz-entropi yöntemi (Rubinstein, 1999) ile belirgin bir şekilde sapmaların olduğu durumlar listelenmiştir. Ayrıca

bölütleme algoritmaları (Schleyer ve Russell, 2011) ve planlama algoritmaları (Micalizio, 2013) kullanılarak da hata tespiti ve tanısı yapılmaktadır. Self Organizing Maps (SOPs) (Häussermann vd., 2015) ve Probabilistic Graphical Models (PGMs) gibi olasılıksal yöntemler de hataların uzam-zamansal olarak sezilmesi için analiz edilmiştir. Bir çalışmada (Hovland vd., 1998), hatalar Saklı Markov Modelleri (SMM) (Hidden Markov Models, HMMs) kullanılarak tork ve güç sinyalleri göz önüne alınarak sezilmektedir. Bir başka çalışmada (Bouguerra vd., 2007), robot ortamlarındaki belirsizlik durumları karşısında robotun beklediği çıktılar için bir olasılık dağılımı kullanılarak ortamda oluşan hatalar sezilmiştir. Yapay Sinir Ağları (Artificial Neural Networks, ANN) da eylem gözetimi için kullanılmaktadır. Yapay Sinir Ağları ve Kalman Filtreleri kullanılarak gezgin bir robot sisteminde sensör ve mekanik hataların sezilmesi konusunda bir çalışma da yapılmıştır (Goel vd., 2000). Hata sezme ve tanıma için Dinamik Bayes Ağları (DBA, Dynamic Bayesian Networks) da kullanılmaktadır (Kirchner ve Geighs, 2014). Başka bir çalışmada ise çok katmanlı Bayes Ağları (Multi-layered Bayesian Networks) kullanılarak hata sezme ve tanıma yapılmıştır (Sassi vd., 2016). Anormal durumlar kontrol kuralları kullanılarak da sezilebilir. Bir çalışmada (Adam vd., 2014), kontrol kuralları uygulama alanına özgü bir dilde temsil edilip donanım hataları sezilmiştir. Başka bir çalışmada (Kapotoglu vd., 2014), Metric Temporal Logic (MTL) (Kvarnström vd., 2008) kullanılarak eylem kontrol formleri temsil edilmiştir. El ile kodlanmış bu formler hata sezme için ileri yayılım (progression) algoritması ile eylem yürütülmesi süresince kontrol edilmiştir. İnsansız hava araçları için de hata sezme yöntemleri önerilmiştir (Doherty vd., 2009).

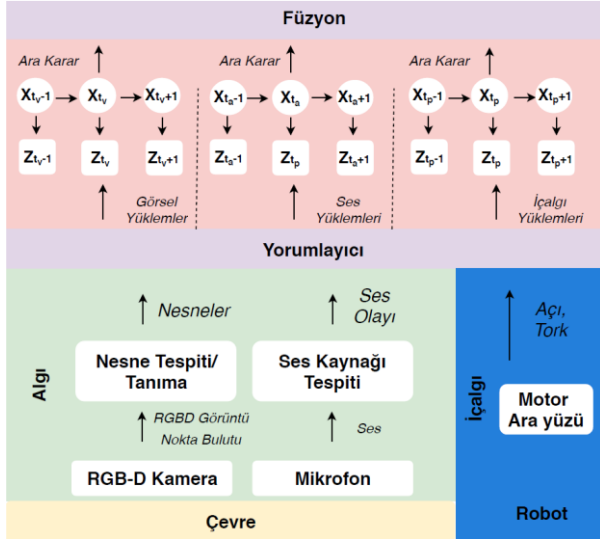
Hataların tümleşik olarak sezilmesi için bir hata sezme, tanıma ve hata kotarma mimarisi sunulmuştur (Saied vd., 2015). Bu çalışmada, lineer olmayan bir gözlemlene birimi kullanılarak kalıntı (residual) üretimi yapılmıştır. Daha sonra ise, koşul-sonuç kuralları kullanılarak hata hakkında bir çıkarım yapılmıştır.

Önceki çalışmalar incelendiğinde, robotlarda hata sezme için görsel, işitsel ve içalgı sensör bilgilerini birarada değerlendiren bir çalışma bulunmamaktadır. Bu makale, bu konudaki öncü çalışmayı sunmaktadır.

3. Önerilen Yöntem (Proposed Method)

Bu çalışmada ara seviye (intermediate-level) sensör füzyonu tabanlı bir hata sezme sistemi (Şekil 2) önerilmiştir. Bu sistem sensörlerden (RGB-D kamera ve mikrofon) alınan gözlemlerin işlendiği algılama iş hattı, motor arayüzü, semantik yüklemelerin oluşturulduğu yorumlayıcı katmanı ve bu yüklemelerin zamansal olarak işlendiği füzyon katmanından oluşmaktadır. Sensörlerden elde edilen gözlemler hata sezme için sembolik yüklemelere (görsel, işitsel ve içalgı) dönüştürülmektedir. Bu yüklemeler hata sezme ile ilgili anlamsal bilgiler içermektedir. Oluşturulan yüklemeler

füzyon katmanında birleştirilerek, yürütülen eylem için *başarılı* veya *hatalı* kararı olasılıksal olarak oluşturulur.



Şekil 2. Hata sezme mimarisi (Failure Detection Architecture)

3.1. İçalgı (Proprioception)

İçalgı (Proprioception), robotun motor arayüzünden elde edilen ölçümlerin değerlendirilmesi ile elde edilir. Motor arayüzü, motorların konumlarına, hareketlerine ve motorlara uygulanan yük miktarına ilişkin ölçümleri içerir. Hata sezme için robot koluna bağlı iki parmaklı paralel tutucunun anlık durumunu ifade eden yüklemeler oluşturulmaktadır. Tutucu parmakları arasında bulunan kodlayıcılar (encoder) ile ölçülen konum bilgisine göre oluşturulan parmaklar *açık*, *kapalı* ve *hareket* halinde yüklemelerinin yanı sıra parmaklarda ölçülen yük miktarına göre tutucu dolu yüklemi oluşturulmaktadır.

Paralel Tutucu İçalgı Yüklemeleri: Robotun paralel tutucusundaki parmakları arasındaki uzaklık (D) ve parmaklarda ölçülen kuvvet F değeri aşağıdaki yüklemelere dönüştürülmektedir (τ_D ve τ_F uzaklık ve kuvvet eşik değerlerini temsil etmektedir):

- *Açık:* Tutucunun parmakları açık konumdadır: $D > \tau_D$.
- *Kapalı:* Tutucunun parmakları kapalı konumdadır: $D < \tau_D \wedge F < \tau_F$.
- *Hareketli:* Tutucunun parmakları hareket halindedir; açılmakta ya da kapanmaktadır.
- *Dolu:* Tutucuda ölçülen kuvvet değeri eşik değerinden üzerindedir, $F > \tau_F$.

3.2. İşitsel İş Hattı (Audio Pipeline)

Yürütme esnasında robota ses bilgisinin sağlanabilmesi için robot üzerine bir mikrofon yerleştirilmiştir. İşitsel iş hattı 3 bileşenden oluşmaktadır: ön işleme, öznelik çıkarma ve

sınıflandırma. Ses sinyalleri 4 kanallı bir mikrofon aracılığıyla 16 KHz örnekleme yapılarak elde edilir. Ön işleme sırasında, elde edilen ses sinyali 32 milisaniyelik çerçevelere bölünür ve çerçevelere Hızlı Fourier Dönüşümü (FFT) (Brigham vd., 1998) uygulanarak zaman alanından frekans alanına geçirilir. O andaki çerçevenin güç spektrumu hesaplanarak daha önceden belirlenmiş başlangıç eşik değeri (onset) ile karşılaştırılır. Eşik değerinin altında kalan çerçeveler gürültü olarak değerlendirilir. Başlangıç eşik değerini aşan ses sinyali, bitiş eşik değerinin (offset) altına düşene kadar dinlenmeye devam edilir.

Öznelik çıkarma adımı, eşik değerinin üzerindeki çerçeveler üzerinde Mel-Frekanslı kepsstral katsayıları (MFCC) (Logan 2000) hesaplanarak 12 boyutlu bir öznelik vektörü oluşturulur ve ilk 10 çerçevenin ortalaması alınır. Ayrıca, ses olayının süresi hesaplanarak öznelik vektörüne eklenir. Sınıflandırma adımı ise, elde edilen öznelik vektörleri Lineer Destek Vektör Makinesi (SVM) aracılığıyla düşme, çarpma ya da robot gürültüsü olarak sınıflandırılır.

İşitsel Algı Yüklemeleri: İşitsel algı iş hattının çıktıları aşağıdaki ses olaylarına ($E = e_j$) dönüştürülür:

- *Ses yok:* Herhangi bir ses olayının olmaması durumu. İşlenmekte olan ses sinyali çerçevesinin toplam enerjisi önceden belirlenmiş başlangıç eşik (onset) değerinin altında kalması durumunda ses olayı yok olarak kabul edilmektedir.
- *Düşürme:* Robotun bir nesneyi düşürmesi ile oluşan ses olayı.
- *Çarpma:* Robotun ortamdaki bir nesneye çarpması ile oluşan ses olayı. Bu ses olayı yalnızca *tutma* eyleminin yürütülmesi sırasında gözlemlenmektedir. Bu nedenle *birakma* ve *itme* eylemlerinin modellerinde dahil edilmemiştir.
- *Robot gürültüsü:* Robot motorlarının hareketi sırasında meydana gelen gürültü.

3.3. Görsel İş Hattı (Visual Pipeline)

Görsel iş hattı, bir RGB-D kameradan elde edilen veriler üzerinde paralel olarak çalıştırılan nesne tanıma algoritmalarından ve bu algoritmaların sonuçlarının filtrelenerek tutarlı bir dünya modeli oluşturmak için birleştirilmesinden oluşmaktadır.

Ortamda bulunan nesnelere tanımlanması için şablon tabanlı 3 boyutlu bir nesne tanıma algoritması, LINEMOD (Hinterstoisser vd., 2012), kullanılmaktadır. Robotun çalışırken karşılaşacağı nesnelere şablonları eğitim aşamasında kaydedilir. Çalışma sırasında robot, karşılaştığı nesnelere kayıtlı şablonlar ile eşleştirilerek tanımlanır. Nesne tanımlanmaya ek olarak, ortamdaki nesnelere ait nokta bulutları, 3B Bölütleme (Aldoma

vd., 2012) algoritması kullanılarak masa düzleminden ayrıştırılır.

Nesne tanıma algoritmalarının çıktılarını sahne yorumlama birimine iletilir. Tanınan nesnelere kategori, konum, boyut, renk gibi fiziksel özellikleri ile beraber bir bilgi tabanına kaydedilir. Bu özelliklere ek olarak, nesnenin sahnedeki varlığı ile ilgili bir güven değeri hesaplanır. Bu değer, nesne ile ilgili gözlemler alındıkça olasılıksal olarak artırılır veya belirli bir süre gözlem alınmazsa azaltılır. Güven değeri belirli bir değerin altına düştüğünde nesne bilgisi bilgi tabanından kaldırılır (Inceoglu vd., 2018-3).

Görsel Algı Yüklemeleri: Robotun eyleme başlamadan önceki ve eylemi bitirdikten sonraki sahneler Violet sisteminin çıktısı kullanılarak temsil edilir. Bu temsil üzerinden daha sonra aşağıdaki yüklemeler oluşturulur:

- ΔA : Sahnenin nokta bulutu (point cloud) temsili üzerinde hesaplanan yüzey alanı (A) değişimini ifade etmektedir.
- ΔL : Robotun etkileşime geçtiği nesnenin etkileşim öncesi ve sonrası 3B konumu arasındaki farkı ifade etmektedir. Her bir eksen için ayrı ayrı hesaplanmaktadır. (o robotun etkileşim halinde olduğu nesneyi temsil etmektedir.)

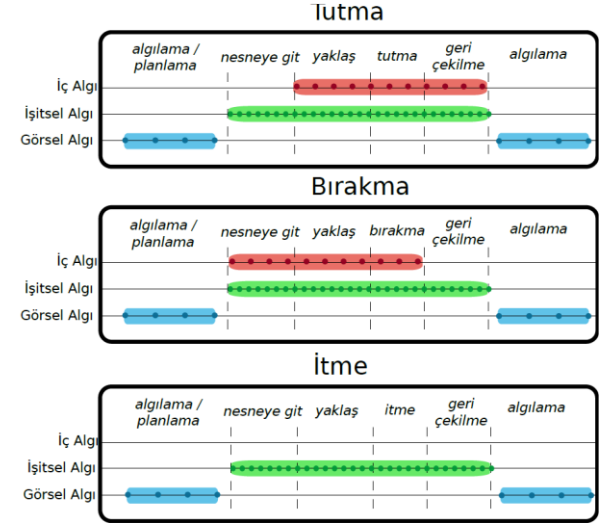
4. Bulgular (Results)

Deneylerde Baxter insansı robotu, 2 parmaklı tutucu, robotun baş kısmına yerleştirilmiş Asus Xtion Pro RGBD kamera ve gövde kısmına yerleştirilmiş PSEye mikrofon yerleştirilmiştir. Sistemin yazılım bileşenleri ROS (2022) platformu üzerinde gerçekleştirilmiş, işitsel algı için HARK (2022) kütüphanesi kullanılmıştır. İnsansı robotun karar mekanizması için Intel İ7 İşlemci ve 8 GB belleğe sahip bir dizüstü bilgisayar kullanılmıştır.

4.1. Algı Kiplerinin Niteliksel Değerlendirmesi (Qualitative Evaluation of Modalities)

Şekil 3'te tutma, bırakma ve itme eylemleri için içalgı, işitsel ve görsel algı kiplerinin, eylemlerin adımları esnasında (algılama/planlama, nesneye gitme, yaklaşma, tutma, geri çekilme ve algılama) gözlemlenebilirliğine ilişkin grafik sunulmuştur. Şekillerdeki noktalar gözlemlere ve gözlemler arası uzaklıklar gözlem frekanslarına karşılık gelmektedir. Mavi renkle işaretlenen gözlemler görsel algı kipine, yeşil renkle işaretlenen gözlemler işitsel algı kipine, kırmızı renkle işaretlenen gözlemler iç algı kipine aittir. Tutma eylemi için nesneye-git (move-to) adımında ve bırakma eylemi için geri çekilme (retreat) adımında herhangi bir içalgı gözlemi içermemektedir. Benzer şekilde, itme eylemi sırasında içalgı gözlemleri sabit kalmaktadır. Diğer yandan işitsel algı kullanılarak

eylemin herhangi bir adımında oluşan hatalar tespit edilebilmektedir. Diğer yandan, robotun eylemi yürütmesi esnasında robot kolunun görüntüyü engellemesi sebebiyle görsel algı ile sahnenin tamamı gözlenememektedir. Bu yüzden görsel algı ile yalnızca eylemin öncesinde ve sonrasında sahne gözlemlenerek durum tespiti yapılmaktadır.



Şekil 3. Tutma, bırakma ve itme eylemleri için algı kiplerinin gözlemlenebilirliği (Visualization of observable action phases for pick, place and push actions)

4.2. Hata Sezme Deneyleri (Failure Detection Experiments)

Bu bölümde güvenli görev yürütme mimarisinde hata sezme yöntemlerinin başarımlarını analiz edilmektedir. *Tutma*, *bırakma* ve *itme* veri kümeleri kendi içlerinde rastgele olarak %50 eğitim ve %50 test kümelerine bölünmüştür. Bu işlem sırasında hata ve başarı durumlarının dağılımları korunmuştur. İşlem 10 kez tekrarlanarak elde edilen sonuçların ortalaması alınmıştır.

Hata sezme için çerçeve ve seri olmak üzere iki değerlendirme ölçütü tanımlanmıştır. Çerçeve bazlı değerlendirmede her t anındaki gözlem ile bir tahmin yapılarak gerçek-değer ile karşılaştırılmaktadır. Seri bazlı değerlendirmede ise gözlem serisinde bulunan tüm gözlemler kullanılarak tek bir karar üretilir ve gerçek değer ile karşılaştırılır.

4.2.1. Hata Sezme için Önerilen Yöntemler (Proposed Methods for Failure Detection)

Hata sezme için ele alınan yöntemler aşağıda sıralanmaktadır:

- *İçalgı-SM*: Hata ve başarı durumları için birer adet tek kipli SMM modeli eğitilir.
- *İşitsel-SMM*: İşitsel yüklemeler ile tek kipli SMM eğitilir.

- *İşitsel-SMM (MFCC)*: MFCC öznitelikleri kullanılarak SMM modeli eğitilir.
- *Görsel (ΔA)*: Görsel algı ile hesaplanan alan değişimi ΔA özneliği kullanılarak hata tahmini yapılır. Alan değişiminin artması hata göstergesi olarak kabul edilmiştir (örn., kübik bloklardan oluşan yapının devrilerek etrafa saçılması).
- *Görsel (ΔL)*: Nesnenin konumu baz alınarak hata tahmini yapılır. 3 boyutta her eksen için beklenen ve gözlemlenen konum arasındaki fark hesaplanarak ikili bir yüklem oluşturulur. Sonrasında Karar Ağacı algoritması (Decision Tree (DT)) kullanılarak hata tahmini yapılır.
- *Görsel-SMM (HOG)*: Robotun kamerasından elde edilen 2B resimler üzerinde Yönlü Gradyan Histogramı (Histogram of Oriented Gradients (HOG) (Freeman vd., 1994) hesaplanarak hata tahmini yapılır. Resimler 3 saniyelik aralıklarla örneklenmiş ve her bir resim üzerinde ilgili alan kırılmıştır. Hesaplanan HOG öznitelikleri Temel Bileşen Analizi (Principal Component Analysis (PCA) (Pearson, 1901) kullanılarak 100 boyuta indirgenmiştir. Bu öznitelikler kullanılarak hata ve başarı durumları için ikişer saklı durum içeren SMM modelleri eğitilmiştir.
- *Görsel-SVM (HOG)*: Her bir eylem için gözlem sekansındaki son görselden elde edilen

HOG öznitelikleri kullanılarak bir Destek Vektör Makinesi (Support Vector Machine (SVM)) eğitilmiştir.

- *Füzyon-SMM*: Farklı kiplerinden alınan gözlemler, zaman bilgisi kullanılarak senkronize edilir ve tek bir vektör haline getirilir. Elde edilen yeni öznitelik ile SMM modelleri eğitilir.
- *Füzyon-DT*: Görsel ve işitsel gözlemler tek bir öznitelik vektörü haline getirilir ve Karar Ağacı (DT) eğitilerek hata sezme gerçekleştirilir.

4.2.2. Çerçeve Bazlı Hata Sezme (Frame level Failure Detection)

İçalgıdan elde edilen ardışık gözlemler birbirinden bağımsız olmasına karşın, bir ses olayı birden fazla çerçeveye etki etmektedir. Ses sinyallerinin işlenmesi sırasında oluşturulan çerçeveler birbirinden bağımsız veya bütün olarak değerlendirilebilir. Bu bölümde sunulan sonuçlar elde edilirken, ses sinyallerine ait çerçeveler birbirinden bağımsız olarak değerlendirilmiştir. Hata sezme modellerinin eğitimi için, gözlem serileri her t anı için etiketlenmiştir.

Tablo 1 ve Tablo 3'te tutma ve bırakma eylemleri için çerçeve bazlı hata sezme sonuçları sunulmuştur. Tablo 2 ve Tablo 4'te ise son kararlar gerçek değerle karşılaştırılarak ölçülmüştür.

Tablo 1. Çerçeve bazlı tek kipli tutma hatası sezme sonuçları (Experimental results for frame level unimodal pick action failure detection)

Yöntem	<i>F1 Ölçüsü</i>	<i>Precision</i>	<i>Hatırlama (Recall)</i>
İçalgı-SMM	0.65 ± 0.09	0.81 ± 0.03	0.71 ± 0.11
İşitsel-SMM	0.55 ± 0.16	0.63 ± 0.06	0.56 ± 0.13
İşitsel-SMM (MFCC)	0.73 ± 0.07	0.76 ± 0.05	0.72 ± 0.08

Tablo 2. Çerçeve bazlı tek kipli tutma hatası sezme sonuçları (son çerçeve değerlendirme) (Experimental results for frame level unimodal pick action failure detection, considering the last frame)

Yöntem	<i>F1 Ölçüsü</i>	<i>Precision</i>	<i>Hatırlama (Recall)</i>
İçalgı-SMM	1.00 ± 0.00	1.0 ± 0.00	1.0 ± 0.00
İşitsel-SMM	0.62 ± 0.20	0.66 ± 0.25	0.64 ± 0.15
İşitsel-SMM (MFCC)	0.85 ± 0.09	0.89 ± 0.06	0.85 ± 0.09
Görsel-SMM (HoG)	0.51 ± 0.10	0.63 ± 0.20	0.60 ± 0.05
Görsel-SVM (HoG)	0.55 ± 0.12	0.58 ± 0.12	0.54 ± 0.13

Tablo 3. Çerçeve bazlı tek kipli bırakma hatası sezme sonuçları (Experimental results for frame level unimodal place action failure detection)

Yöntem	<i>F1 Ölçüsü</i>	<i>Precision</i>	<i>Hatırlama (Recall)</i>
İçalgı-SMM	0.39 ± 0.03	0.51 ± 0.03	0.50 ± 0.01
İşitsel-SMM	0.70 ± 0.20	0.83 ± 0.06	0.73 ± 0.17
İşitsel-SMM (MFCC)	0.62 ± 0.11	0.81 ± 0.02	0.64 ± 0.09

Tablo 4. Çerçeve bazlı tek kipli bırakma hatası sezme sonuçları (son çerçeve değerlendirme) (Experimental results for frame level unimodal place action failure detection, considering the last frame)

Yöntem	<i>F1 Ölçüsü</i>	<i>Precision</i>	<i>Hatırlama (Recall)</i>
İçalgı-SMM	0.48 ± 0.15	0.52 ± 0.29	0.58 ± 0.09
İşitsel-SMM	0.75 ± 0.20	0.81 ± 0.21	0.78 ± 0.16
İşitsel-SMM (MFCC)	0.98 ± 0.05	0.99 ± 0.04	0.98 ± 0.05
Görsel-SMM (HoG)	0.54 ± 0.05	0.48 ± 0.06	0.63 ± 0.07
Görsel-SVM (HoG)	0.79 ± 0.14	0.84 ± 0.13	0.80 ± 0.13

4.2.3. Sekans Bazında Hata Sezme (Temporal Failure Detection)

Bu bölümde sunulan sonuçlar elde edilirken bir ses olayına ait çerçeveler toplu olarak değerlendirilip tek bir sınıflandırma sonucu üretilmiştir. Eylemlere ait gözlemler her sekansa bir etiket karşılık gelecek şekilde etiketlenmiştir. Tablo 5, Tablo 6 ve Tablo 7'de sırasıyla *tutma*, *birakma* ve *itme* eylemlerine ait hata sezme sonuçları verilmiştir. Şekil 3'te yapılan analize göre çok kipli hata sezme yöntemleri incelenmiştir.

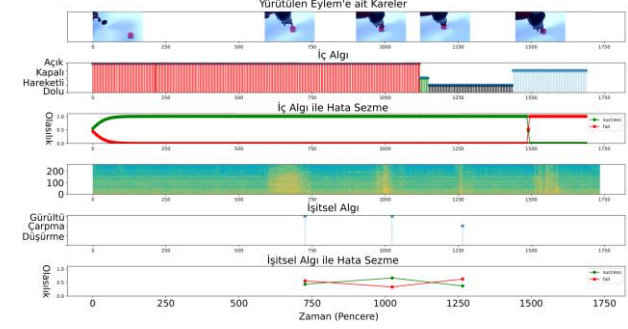
Tablo 5'te sunulan sonuçlardan görülebileceği üzere, *tutma* eylemi hatası sezme için ana bilgi kaynağı içalgıdır.

Birakma eylemi için ise (Tablo 6), içalgı tutucunun parmakları açıldıktan sonra eylemin sonucuna ait herhangi bir ek bilgi sunamamaktadır. Görsel algı incelendiğinde, toplam alan değişimi özneliğinin sahnedeki önemli değişiklikleri tespit edebildiği gözlemlenmiştir.

İtme eyleminin yürütülmesi sırasında (Tablo 7), içalgıdan elde edilen gözlemler değişmemektedir. Hedef nesnenin konumuna bağlı olarak eylem öncesi ve eylem sonrası sahne karşılaştırılması yapıldığında (ΔL), *itme* eylemi için daha yüksek başarımlar elde edilmiştir.

Şekil 4'te örnek bir nesne tutma hatası için içalgı ve işitsel algıdan elde edilen gözlemler (içalgı ölçümü ve ses güç spektrumu), bu gözlemlerden elde edilen yüklem (iç algı için *açık*, *kapalı*, *hareketli*, *dolu* ve işitsel algı için *gürültü*, *çarpma* ve *düşürme* yüklemelerinin oluşma durumu) ve SMM tabanlı tek kipli hata sezme

sonuçları başarı ve hata sınıfları için olasılıksal olarak sunulmaktadır.



Şekil 4. İç algı ve işitsel algı gözlemlerinin ve sistemin hata sezme çıktısı kararının görselleştirilmesi (Visualization of proprioceptive and auditory observations and failure detection)

Hata sezme sonuçlarına bakıldığında, içalgının *tutma* hatalarını sezme için önemi ortaya çıkmaktadır. Bazı durumlarda işitsel algı kullanılarak hata sezme başarımları artırılabilir. *Birakma* ve *itme* eylemleri için ise *içalgı* hata sezimi için bir bilgi sunamamaktadır. Bu nedenle görsel ve işitsel algı gibi destekleyici bilgilere ihtiyaç duyulmaktadır. Görsel ve işitsel algı karşılaştırıldığında, görsel algı ile daha yüksek başarımlar ile hatalar sezilebilmektedir. Fakat sahnenin tam olarak gözlemlenebilmesi için robotun eylemi yürütmeyi bitirmesi gerekmektedir. Diğer yandan, işitsel algı ile düşme gibi olaylar anlık olarak tespit edilebilmektedir.

Tablo 5. Tutma hatası sezme sonuçları (Experimental results for pick action)

Yöntem	F1 Ölçüsü	Precision	Hatırlama (Recall)
İçalgı	0.85 ± 0.00	0.80 ± 0.00	0.80 ± 0.00
İşitsel	0.76 ± 0.05	0.76 ± 0.05	0.76 ± 0.05
Görsel (ΔA)	N/A	N/A	N/A
Görsel (ΔL)	N/A	N/A	N/A
Füzyon-SMM (İçalgı & İşitsel)	0.86 ± 0.02	0.82 ± 0.03	0.80 ± 0.04

Tablo 6. Birakma hatası sezme sonuçları (Experimental results for place action)

Yöntem	F1 Ölçüsü	Precision	Hatırlama (Recall)
İçalgı	0.26 ± 0.13	0.44 ± 0.33	0.39 ± 0.10
İşitsel	0.87 ± 0.04	0.90 ± 0.02	0.87 ± 0.04
Görsel (ΔA)	0.93 ± 0.05	0.95 ± 0.03	0.93 ± 0.05
Görsel (ΔL)	N/A	N/A	N/A
Füzyon-DT (Görsel - ΔA & İşitsel)	0.95 ± 0.06	0.96 ± 0.04	0.95 ± 0.06

Tablo 7. İtme hatası sezme sonuçları (Experimental results for push action)

Yöntem	F1 Ölçüsü	Precision	Hatırlama (Recall)
İçalgı	0.48 ± 0.00	0.39 ± 0.00	0.62 ± 0.00
İşitsel	0.64 ± 0.07	0.70 ± 0.10	0.64 ± 0.07
Görsel (ΔA)	0.74 ± 0.09	0.74 ± 0.09	0.74 ± 0.09
Görsel (ΔL)	0.95 ± 0.03	0.96 ± 0.02	0.95 ± 0.03
Görsel-SMM (HoG)	0.60 ± 0.13	0.66 ± 0.21	0.68 ± 0.10
Görsel-SVM (HoG)	0.69 ± 0.08	0.71 ± 0.09	0.71 ± 0.07
Füzyon-DT (Görsel - ΔA & İşitsel)	0.70 ± 0.12	0.73 ± 0.13	0.71 ± 0.10
Füzyon-DT (Görsel - ΔL & İşitsel)	0.95 ± 0.03	0.96 ± 0.02	0.95 ± 0.03

5. Sonuçlar (Conclusion)

Bu makalede içalgı, görsel ve işitsel algı tabanlı çok kipli bir eylem gözetimi ve hata sezme sistemi sunulmuştur. *Tutma, bırakma* ve *itme* eylemleri için algılama kiplerinin hata sezme için katkısı analiz edilmiştir. Bu analize göre, çok kipli eylem gözetiminin hataları sezmede tekil sensör sistemlerine göre daha başarılı sonuçlar ürettiği ve sensör kiplerinin birbirlerini tamamlayıcı olduğu gözlenmiştir. İleriki çalışmalarda iç algının dokunma sensörleri ile zenginleştirilmesi planlanmaktadır. Ayrıca nesne ve eylem kümesi çeşitliliği artırılarak daha uzun nesne etkileşimi gerektiren senaryolar üzerinde çalışılması planlanmıştır.

Teşekkür (Acknowledgment)

Bu çalışma, TÜBİTAK tarafından 115E368 no'lu proje kapsamında desteklenmiştir.

Kaynaklar (References)

Adam S., Larsen M., Jensen K., and Schultz U. P., 2014. Towards Rule-based Dynamic Safety Monitoring for Mobile Robots. *Simulation, Modeling, and Programming for Autonomous Robots*. Springer, 207–218.

Aldoma A., Marton Z.C., Tombari F., Wohlkinger W., Potthast C., Zeisl B., Rusu R. B., Gedikli S., and Vincze M., 2012. Point cloud library. *IEEE Robotics & Automation Magazine*, 1070(12).

Bouguerra A., Karlsson L., and Saffiotti A., 2007. Handling uncertainty in semantic-knowledge based execution monitoring. *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 437–443.

Brigham E. O., Pastor J. R., Apostol T. M., Rodríguez M., Martínez M. R., Edwards C. H., Edwards D. E. H., 1998. *The fast Fourier transform and its applications*, Prentice Hall.

Doherty P., Kvarnström J., and Heintz F., 2009. A temporal logic-based planning and execution monitoring framework for unmanned aircraft systems. *Autonomous Agents and Multi-Agent Systems*, 19(3), 332–377.

Freeman W. T. and Roth M., 1994. Orientation histograms for hand gesture recognition. *Mitsubishi Electric Research Laboratories Teknik Raporu*.

Fritz C., 2005. Execution monitoring – a survey. *University of Toronto Teknik Raporu*.

Fourlas G. K., Karkanis S., Karras G. C., and Kyriakopoulos K. J., 2014. Model based actuator fault diagnosis for a mobile robot. *IEEE International Conference on Industrial Technology (ICIT)*, 79–84.

Goel P., Dedeoglu G., Roumeliotis S. I., and Sukhatme G. S., 2000. Fault detection and identification in a mobile robot using multiple model estimation and neural network. *IEEE International Conference on Robotics and Automation (ICRA)*, 2302–2309.

HARK: <http://www.hark.jp/>, 12.09.2022.

Hüsserle K., Zweigle O., and Levi P., 2015. A novel framework for anomaly detection of robot behaviors. *Journal of Intelligent & Robotic Systems*, 77(2), 361–375.

Hinterstoisser S., Cagniard C., Ilic S., Sturm P., Navab N., Fua P., and Lepetit V., 2012. Gradient response maps for real-time detection of textureless objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(5), 876–888.

Hovland G. E., and McCarragher, B. J., 1998. Hidden markov models as a process monitor in robotic assembly. *The International Journal of Robotics Research*, 17(2), 153–168.

Inceoglu A., Ince G., Yaslan Y., Sariel S., 2018-1. Comparative Assessment of Sensing Modalities on Manipulation Failure Detection. *IEEE International Conference on Robotics and Automation (ICRA) Workshop on Multimodal Robot Perception*.

Inceoglu A., Ince G., Yaslan Y., Sariel S., 2018-2. Failure Detection Using Proprioceptive, Auditory and Visual Modalities. *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*.

Inceoglu A., Koc C., Kanat B. O., Ersen M., and Sariel S., 2018-3. Continuous visual world modeling for autonomous robot manipulation. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 49(1), 192 – 205.

ISO 10218-2:2011, 2011. *Robots and Robotic Devices – Safety Requirements for Industrial Robots – Part 2: Robot Systems and Integration*. ISO, Geneva, Switzerland.

ISO 10218:2011, 2011. *Robots and Robotic Devices – Safety Requirements for Industrial Robots – Part 1: Robots*. ISO, Geneva, Switzerland.

Kapotoglu M., Koc C., Sariel S., and Ince G., 2014. Action monitoring in cognitive robots. *The 22nd Signal Processing and Communications Applications Conference (SIU)*, 2154–2157.

Kirchner D., and Geihs K., 2014. Qualitative bayesian failure diagnosis for robot systems. *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*.

Kvarnström J., Heintz F., and Doherty P., 2008. A temporal logic-based planning and execution monitoring system. *ICAPS*, 198–205.

Logan B., 2000. Mel frequency cepstral coefficients for music modeling. *International Symposium on Music Information Retrieval (ISMIR)*.

Mendoza J. P., Veloso M., and Simmons R., 2012. Mobile robot fault detection based on redundant information statistics. *IEEE/RSJ International Conference on Intelligent Robots and Systems*.

Mendoza J. P., Veloso M., and Simmons R., 2015. Plan execution monitoring through detection of unmet expectations about action outcomes. *IEEE International Conference on Robotics and Automation (ICRA)*, 3247–3252.

Mendoza J. P., Veloso M., Simmons R., 2014. Focused optimization for online detection of anomalous regions. *IEEE International Conference on Robotics and Automation (ICRA)*, 3358–3363.

Micalizio R., 2013. Action failure recovery via model-based diagnosis and conformant planning. *Computational Intelligence*, 29(2), 233–280.

Orendt E. M. and Henrich D., 2015. Design of robust robot programs: Deviation detection and classification using

- entity-based resources. IEEE International Conference on Robotics and Biomimetics (ROBIO), 1704–1710.
- Pearson K., 1901. On lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 2(11), 559–572.
- Pettersson O., 2005. Execution monitoring in robotics: A survey. *Robotics and Autonomous Systems*, 53(2), 73–88.
- Pettersson O., Karlsson L., and Saffiotti A., 2007. Model-free execution monitoring in behavior-based robotics. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 37(4), 890–901.
- ROS, 2022: <http://www.ros.org/>, 12.09.2022.
- Rubinstein R., 1999. The cross-entropy method for combinatorial and continuous optimization. *Methodology and Computing in Applied Probability*, 1(2), 127–190.
- Saied M., Lussier B., Fantoni I., Francis C., Shraim H., and Sanahuja G., 2015. Fault diagnosis and fault-tolerant control strategy for rotor failure in an octorotor. *IEEE International Conference on Robotics and Automation (ICRA)*.
- Sassi I., Gouin A., and Thiriet J., 2016. A Bayesian Network for Diagnosis of Networked Mobile Robots. *Risk, Reliability and Safety: Innovating Theory and Practice (ESREL)*.
- Schleyer G. and Russell R. A., 2011. Disturbance and failure classification in walking robots. *Australasian Conference on Robotics and Automation*, 1–8.
- Stavrou D., Eliades D. G., Panayiotou C. G., and Polycarpou M., 2015. Fault detection for service mobile robots using model-based method. *Autonomous Robots* 40(1), 383–394.