



Anlam Çözümleme Modelleri ile Geliştirilen Türkçe İşaret Dili Çeviri Sisteminin Nesnelerin İnterneti Üzerinden Web Ortamında Uygulaması

Yelda FIRAT^{*}, Taşkın UĞURLU^{*}

Çanakkale Onsekiz Mart Üniversitesi, Çanakkale Teknik Bilimler Meslek Yüksekokulu, Bilgisayar
Teknolojileri Bölümü, Çanakkale, Türkiye

Geliş / Received: 26.09.2019, Kabul / Accepted: 19.12.2019

Öz

Günlük hayatta insanlar; duygularını, düşüncelerini ve yaşantılarını dili kullanarak ya da kullanmaksızın çevrelerindeki insanlara aktararak iletişime geçerler. Dilsel ve işitsel yetisi olmayan kişiler ise çevreleriyle bu iletişimi işaret dilini kullanarak gerçekleştirirler. İşaret dili ise, engelli kişilerin hem kendi aralarında hem de bu dili bilen engelsiz kişilerle jest, mimik ve vücut hareketleriyle anlaşmalarını sağlayan bir dil olarak tanımlanır. Bu anlamda Nesnelerin İnterneti üzerinden anlam çözümleme modelleri ile web (ağ) ortamında çalışmasına yönelik geliştirilen bu çalışma, duyma ve konuşma bozukluğu olan kişilerle işaret dilini bilmeyenler arasındaki iletişimi sağlayan bir Türkçe İşaret Dili Çeviri Sistemidir. Gerçekleştirilen bu sistemle; işaret dilini gösteren el hareketleri, derinlik kamerasıyla yakalanmış, MQTT (Message Queuing Telemetry Transport /Sıralı Telemetry Mesaj İletimi) sunucusuna gönderilmiştir. Gönderilen bu görüntüler Üç Boyutlu Modelleme ve Hareket Analizinde kullanılan yapay sinir ağları algoritmaları ve makine öğrenmesi yöntemleri ile analiz edilmiştir. Analiz edilen görüntülerden elde edilen sözcüklerin gerçek anlamları ise, Biçimsel Kavram Analizi Teorisini hazırlanan tematik rollerden oluşan gerçeklik modelleriyle bulunmuştur. Gerçek anlamları bulunan bu kelimeler, içinde geçtiği cümle ile birlikte lokal ağ ortamında yayınlanmıştır. Böylece, engelli bir kişi aynı veya farklı mekanlarda yer alan kişi veya kişilerle iletişim kurabilmiştir.

Anahtar Kelimeler: İşaret dili tanımlama, derinlik kamerası, biçimsel kavram analizi, tematik roller, MQTT

Application of the Turkish Sign Language Translation System Developed by the Semantic Analysis Models over the Internet of Things at the Web Medium

Abstract

In daily life, people communicate by conveying their emotions, thoughts, and experiences to the people around them with or without language. However, persons who do not have speech and hearing skills make this communication with their environment by using sign language. As for that the sign language is defined as a language enabling handicapped persons communicate through gesture, mimic and body movements between themselves and with other persons who know this language. In this sense, this work developed for its functioning at the web medium through the semantic analysis models over the Internet of Things is a Turkish Sign Language Translation System which provides a communication opportunity between hearing impaired and speech handicapped persons and those who do not know the sign language. Through this system, movements of hands reflecting the sign language were captured by a depth camera and sent to the MQTT (Message Queuing Telemetry Transport) server. These images were analysed by means of the methods of three dimensional modelling, artificial neural networks algorithms used at movement analysis and machine learning. Real meanings of the words obtained from the analysed images were found through the reality models consisting of thematic roles prepared in accordance with the Formal Concept Analysis Theory. These words the real meanings of which had been identified were published at the local network medium together with the sentence that they had been mentioned. Thus, a handicapped person could communicate with other person or persons being at the same place or elsewhere.

Keywords: Sign language recognition, depth camera, formal concept analysis, thematic roles, MQTT

1. Giriş

İşaret dili; dilsel ve işitsel yetisini kullanamayan kişilerin kendi aralarında anlaşabilmeleri için el hareketlerini ve yüz mimiklerini kullanarak oluşturdukları sessiz, görsel bir dildir.

Her ülkenin farklı dili olduğu gibi işaret dili de ülkeden ülkeye değişim göstermektedir. Meksika, Kore, Japon, İspanya, Norveç, Finlandiya ve Türk İşaret Dili bunlardan bazılarıdır. İş hayatında, sosyal yaşamda ya da aile içerisinde işitsel-dilsel anlamda engeli olmayan bir kişinin, işitme kaybı ve konuşma bozukluğu olan bir kişiyle anlaşabilmesi ancak o kişinin kendi ülkesine ait işaret dilini bilmesi ile olur. Normal insanlar tarafından bu dilin bilinmemesi ise, işitme ve konuşma engelli kişilerle iletişimde karmaşıklığa neden olur. Buna bağlı olarak; esnek, düşük maliyetli ve senkronize çalışarak işaret dilini yazılı veya sesli ortamlara aktarabilecek bir çeviri sisteminin geliştirilmesi, işitme ve konuşma engelli kişilerin engeli olmayan kişilerle sağlıklı bir şekilde iletişime geçmelerini sağlayabilir (Adithya vd., 2013; Rajaganapathy vd., 2015; Nikam ve Ambekar, 2016; Fırat ve Uğurlu, 2018).

Bu bağlamda Nesnelerin İnterneti üzerinden, el işaretlerine karşılık gelen kelimelerin gerçek anlamlarının çözümlenmesi ile web (ağ) ortamında çalışan bir Türkçe İşaret Dili Çeviri Sistemi geliştirilebilir. Özellikle MQTT (Message Queuing Telemetry Transport /Sıralı Telemetri Mesaj İletimi) protokolü sayesinde dünyanın farklı yerlerindeki insanlar bu Türkçe İşaret Dili çeviri sistemini rahatlıkla kullanılabilir.

Sözlü olmayan iletişimde, bir işaret dili çeviri sistemi geliştirmek için el hareketini

tanımlama önemli bir araştırma konusu olmaktadır. Günümüzde bunun için çeşitli özelliklere sahip kameralar, eldiven tabanlı sistemler, görme algoritmaları ve renk belirleyici yaklaşımlar kullanılmaktadır. Özellikle insan-bilgisayar etkileşimi açısından etkili bir yöntem olarak görülen görüntü işleme tekniğinde, kameralarla yakalan görüntüler ya da el hareketleri işlenmek üzere bilgisayarlara aktarılır. Bilgisayarlara aktarılan bu görüntülerin analizinde ise görüntülerin renk, doku gibi özellikleri kullanılır. Basit görünmesine rağmen bu teknik, görüntünün arka planıyla, aydınlatma koşullarıyla, deri renginin tanınmasıyla, görüntünün tanınma süresiyle ilgili bazı problemleri de bünyesinde barındırır (Ibraheem ve Khan, 2012; Pandey ve Jain, 2015).

Bu anlamda Rajam ve Balakrishnan (2011), beş parmağın ikili pozisyonlarını temsil eden 32 işareten oluşan bir veri seti tanımlamıştır. Görüntüler sağ elin avuç içi tarafındadır. Önerilen bu yöntem, hem eğitim hem de test aşamasında tek bir kullanıcı için geliştirilmiştir. Statik görüntüler, özellik noktası çıkarma yöntemi kullanılarak önceden işlenerek her bir işaret için 10 adet görüntü ile eğitilmiştir. Ayrıca elde edilen görüntüler, parmak ucu konum tanımlanmaları üzerinden metne dönüştürülmüştür. Test görüntüleri ile sunulan sonuçlar, önerilen İşaret Dili Tanıma Sisteminin, 320 görüntü ile eğitildiğinde ve 160 görüntü ile test edildiğinde görüntüleri % 98.125 hassasiyetle tanıyabildiğini göstermiştir. Tewari ve Srivastava ise (2012), Hint İşaret Dili için geliştirdikleri el hareketlerini tanımlama sistemlerinde, sınıflandırma amacıyla kendi kendini organize eden harita (The Self-Organizing Map-SOM) kullanmışlardır. Yaklaşımlarında, öncelikle el görüntüsünü pixel değerleri üzerinden cilt ve arka plan olmak üzere

bölümlendirerek arka planı siyah renk ile temsil etmişlerdir. El görüntüsünün bu şekilde parçalara ayrılmasından sonra, el bölgesi için iki boyutlu ayırık kosinüs dönüşümünü (discrete cosine transform- 2-D DCT) hesaplamışlardır. Bu hesaplama iki basamaktan oluşmuştur: (i) Görüntünün her satırı için tek boyutlu ayırık kosinüs dönüşümünü 'nün (1-D DCT) hesaplanması. (ii) Görüntünün her sütunu için tek boyutlu ayırık kosinüs dönüşümünü 'nün (1-D DCT) hesaplanması. Ayırık kosinüs dönüşümünün hesaplanmasından sonra el hareketlerini ve dolayısıyla işaret dilini sınıflandırmak için kendi kendini düzenleyen harita'yı uygulamışlardır. Bu haritayı, verileri sınıflandırmak için kullanılan bir tür sinir ağı olarak ele almışlardır.

İşaret dilini oluşturan hareketlerin eş zamanlı olarak hızlı ve kararlı bir şekilde tanımlanması için son yıllarda Kinect cihazı kullanılmaya başlanmıştır. Derinlik kamerası olarak ta kullanılan Microsoft Kinect; el, kol ve bacak hareketlerini algılama prensibi ile çalışan bir tür algılayıcıdır. Başlangıçta Microsoft Xbox 360 oyun sistemi için geliştirilen Kinect; kullanıcıların oyunları, filmleri ve müzikleri fiziksel hareket veya sesli komutlarla joystick veya klavye gibi ayrı bir giriş kontrolüne gerek kalmadan kontrol etmelerini sağlayan eklenti bir cihazdır. Son yıllarda ise görüntü işleme ve derinlik-uzaklık algılama özelliklerinden dolayı bilimsel çalışmalarda da kullanılmaktadır. Murata ve Shin (2014) yaptıkları çalışma ile havada el ile yazılmış sayısal ve alfabetik karakterleri Kinect ile yakalayarak tanımlamaya çalışmışlardır. Giriş rakamları ve alfabetik karakterler, her bir karakter için kaydedilen vuruş sayıları bilgisine dayanan dinamik programlama eşleşmesi (dynamic programming matching) ile çözümlenmiştir. Sayısal ve alfabetik karakterler için sırasıyla %95 ve %98.9

tanıma oranları elde edilmiştir. Jiang ve arkadaşları (2015) Kinect cihazı ile çok katmanlı bir hareket tanımlama sistemi önermişlerdir. İlk katmanda Kinect cihazı ile yakalanan görüntülerin hareketleri modellenmiştir. İkinci katmanda konum bileşeni sınıflandırması için parçacık bazlı bir tanımlayıcı geliştirmişlerdir. Son katmanda ise şekil bileşenleri sınıflandırılmıştır. Ma ve Peng (2018) ise el hareketlerini tanımlamada derinlik bilgisine sahip Kinect sensörüne dayanan bir eşik segmentasyonu yöntemi (the threshold segmentation method) geliştirmişlerdir. Bu yöntemde ilk olarak, RGB (renk) görüntü verileri ve Kinect sensörü tarafından elde edilen derinlik görüntü verileri, medyan filtreleme (median filtering) ile ön işleme tabi tutulmuştur. İkinci olarak, derinlik bilgisi ve cilt rengi eşığı ile birlikte hareketleri çıkarmak için geliştirilmiş bir uzamsal tabakalaşma yöntemi (the spatial stratification method) önerilmiştir. Hareketler böylece karmaşık ortamlarda geniş bağlamlarda tanımlanabilmiştir. Son olarak, yerel komşu yöntemiyle (the local neighbor method), insan elinin hareketleri parçalara bölünmüştür. Önerilen yöntemin etkinliğini doğrulamada ise, parmak uçlarını tespit etmek için k-kosinüs eğrilik yöntemi kullanılmıştır.

Bu makalede gösterilen çalışmada ise; Kinect sensörü kullanılarak web ortamında çalışan, Türkçe İşaret Dili Çeviri Sistemi geliştirilmiştir. Gerçekleştirilen bu sistemde önce Türkçe İşaret Dili sözlüğünden popüler hareketler seçilmiştir. Daha sonra bu el hareketlerine ait ham veriler derinlik kamerası sensörü yani Kinect aracılığı ile algılanarak MQTT protokolü ile sunucu tarafına gönderilmiştir. Bu gönderilen görüntüler (ham veriler) yapay sinir ağları algoritmaları ve makine öğrenmesi yöntemleri ile çözümlenmiştir. Analiz edilen bu görüntülerden elde edilen kelimelerin cümle

içerisinde kazandığı temel anlamları, Biçimsel Kavram Analizi Metodu ile hazırlanan anlam çözümleme modelleriyle bulunmuştur. Anlamları bulunan bu sözcükler, içinde geçtikleri cümlelerle birlikte web ortamında yayınlanmıştır. Ayrıca geliştirilen bu ara-yüzle işaret dilini öğretmek için bir eğitim ortamı da yaratılmıştır.

Bu bağlamda çalışmanın ikinci bölümünde materyal olarak kullanılan Kinect sensörü, anlam çözümleme modelleri, MQTT protokolü, sürücü- yazılım geliştirme kitleri (Software Development Kit-SDK) ve çalışma çerçevelerine (frameworks) yer verilmiştir. Ayrıca bu bölümde, geliştirilen sistemin çalışma evreleri de anlatılarak önerilen Türkçe İşaret Dili Çeviri Sistemi tanıtılmıştır. Üçüncü bölümde bulgular gösterilmiştir. Dördüncü bölümde ise sonuçlara ulaşılmıştır.

2. Materyal ve Metot

2.1. Microsoft kinect sensörü

Microsoft tarafından geliştirilen bir hareket sensörü eklentisi olan Kinect, insan uzuvlarını algılama prensibi ile çalışan bir aygıttır. Xbox One, Xbox 360 ve Windows bilgisayarlar için oyun oynarken veya çeşitli günlük aktiviteler yaparken kullanıcılara ekstra bir deneyim yaşatmak amacıyla piyasaya sürülmüştür. Xbox'ta herhangi bir ek bileşen kullanmadan oyun oynamayı sağlayan bu cihaz, sahip olduğu bir kızılötesi (infrared) sensör ile saniyede 2 gigabit veri işleyebilme özelliğine sahiptir. Ayrıca Kinect bir oyuncunun kalp atış hızını, yüz ifadesini, mimiklerini, 25 ayrı eklemine pozisyonunu ve oryantasyonunu algılayabilen bir cihazdır. Son yıllarda insan hareketlerini görüntüsel algılaması ve bu görüntüleri çözümleme işlevinden dolayı işaret dilini konuşulan dile çevirecek teknolojik çalışmalarda da sıkça

kullanılmaktadır. Microsoft Kinect sensörü ve onu oluşturan öğeler Şekil 1'de gösterilmiştir.



Şekil 1. Microsoft kinect sensörü

Şekil 1'de de görüldüğü üzere basit bir konsol aksesuarı olarak görülmemesi gereken bu teknolojinin insan iskeletini algılayabilmesi için en az 2,5 metrelik mesafeye ihtiyacı vardır. Bu işlemi cihazın üzerindeki lazer projektörü ve CMOS (Complementary Metal Oxide Semiconductor/Tamamlayıcı Metal Oksit Yarı İletken) algılayıcısı gerçekleştirir. İşlem sırasında projektör, görüş açısına sürekli lazer ışını gönderirken CMOS algılayıcı ise, ışığın yansıma hızını hesaplayarak bir derinlik görüntüsü oluşturur. Bu görüntünün cihazın içerisindeki görüntü işleme algoritmaları üzerinden işlenmesiyle de insan iskeleti meydana getirilir.

2.2. Anlam çözümleme modelleri

Son yıllarda, doğal dilin anlamsal çözümlenmesine yönelik olarak istatistiksel veya kural tabanlı modelleme yöntemleri üzerine çalışmalar hız kazanmaktadır. Oluşturulan modeller, insan zihni ile dil arasında matematiksel bir köprü oluşturarak kavramların gerçek çıkarımlarının elde edilmesini sağlar. Yaratılan bu modeller, uygulamalarından doğabilecek güçlüklerden dolayı bilgisayar ortamında

hazırlanmalıdırlar. Ayrıca Türkçe için hazırlanan modellerle de, cümleleri oluşturan öğelerin anlamsal çıkarımları elde edilerek Türkçe İşaret Dili Çeviri Sistemi için önemli bir kaynak elde edilir (Fırat, 2017).

Bu amaçla, Biçimsel Kavram Analizi Metodu üzerinden tematik rollerle anlam çözümleme modelleri yaratılabilir. Özellikle cümleleri oluşturan öğelerin anlamsal analizlerinde tematik roller önemli argumanlardır. Dilbilimsel bir kavram olan tematik rol terimi, bir ad öbeğinin sözdizimsel-biçimbilimsel olarak cümlenin eylemi ile girdiği anlamsal ilişkisi olarak ifade edilir. En önemli tematik roller Tablo. 1’de verilmiştir (Fırat vd., 2014; Fırat, 2017; Wittenberg vd., 2017; Fırat ve Uğurlu, 2018).

Tablo 1. Tematik roller

Roller	Anlamı
TEMA	Eylem boyunca varlığını gösteren
ETKİLEYEN	Bir eylemi farkındalıkla, bilinçli bir şekilde başlatan
ETKİLENEN	Bir eylemden farkındalıkla, bilinçli bir şekilde etkilenen
KAYNAK	Eylemin başlangıç noktası
HEDEF	Eylemin yöneldiği nokta
YOL	Bir varlığın eylem boyunca hareket ettiği yol
ARAÇ	Bir eylemi gerçekleştirmede gücünden yararlanan nesne

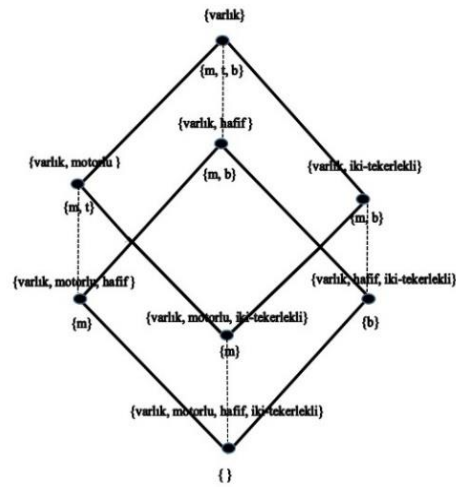
Biçimsel Kavram Analizi Metodu ise biçimsel bağlamları ve bu bağlamlara karşılık gelen latis diyagramlarını kullanarak veri analizi yapan matematiksel bir teoridir. Bir başka deyişle dili oluşturan kavramları özellik-nesne ikilisi içerisinde analiz eden bir disiplindir. Bu anlamda Biçimsel Kavram

Analizi Metodu, nesnelere taşıdıkları özelliklerine göre gruplandırılarak hiyerarşik bir yapı içerisinde gösterir. Bu hiyerarşik yapıda ilk olarak kavramların her biri bir düğüme yerleşir. Daha sonra üst düğümler alt düğümlere kalıtım verecek şekilde çizgilerle birleştirilir. Oluşturulan bu sıralı küme, kavram latisi olarak isimlendirilir (Ganter ve Wille, 1999; Fırat ve Uğurlu, 2018).

Tablo 2’de basit bir bağlam örneği, Şekil 2’de ise bu bağlama ait Biçimsel Kavram Analizi Teorisi üzerinden geliştirilen kavram latisi görülmektedir (Fırat vd., 2014).

Tablo 2. Bağlam örneği

özellik nesne	motorlu	motorsuz	hafif	ağır	iki- tekerlekli	çok- tekerlekli	tekerleksiz
tekne	x			x			x
motorsiklet	x			x	x		
bisiklet		x	x		x		



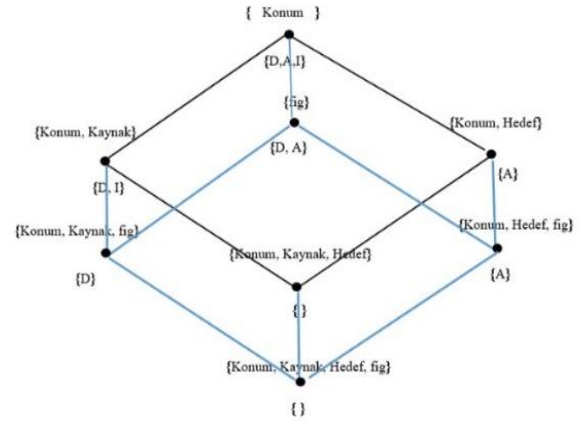
Şekil 2. Örnek kavram latisi

Tablo 2’de görülen bağlam örneğinde ve Şekil 2’de gösterilen örnek kavram latisinde nesnelere; *motorsiklet (m)*, *tekne (t)* ve *bisiklet (b)* özellikler ise; *varlık*, *motorlu*, *hafif* ve *iki-tekerlekli* olarak ifade edilir. Ayrıca düğümlerin üst kısmında özellikler, alt kısmında ise nesnelere yer alır.

Amerikan dilbilimci Jackendoff’un teorisinde ise kavramlaştırma, cümleyi oluşturan öğelerin karakterizasyonunu çıkarmaktır. Bunu da ayrıştırma yöntemiyle gerçekleştirmiştir. Ayrıştırma yöntemi: bir cümleyi oluşturan kelimelerin ve bu kelimelerin anlamlarını temel alan kavramların, en küçük öğelerine ayrılmasıdır. Bir yapı içerisinde ayrıştırılan bu sözcüksel kavramlar ontolojik kategorilere ya da katmanlara ayrılır. Tematik roller ise bu katmanların vazgeçilmez elemanlarıdır. (Jackendoff, 2006).

Jackendoff’un teorisinde ve anlam çözümleme üzerine yapılan diğer çalışmalarda, tematik rollerin soyut ilkeler temelinde kalmasından ve doğrusal bir hiyerarşi ile gösterilmesinden dolayı oldukça yetersiz kaldığı görülmüştür. Bu sebeple matematikselleştirilmiş bir yapının doğal dile bütünleşik hale getirilmesi gerekir. Bu çalışmada geliştirilen üç boyutlu modellerde Jackendoff’un geliştirdiği kategorilerden *alansal katman* ve *hareket katmanları* kullanılmıştır. Konumsal rolleri içeren *alansal katmanda* KONUM, KAYNAK, HEDEF ve YOL gibi lokasyon rolleri mevcuttur. Herhangi bir yere konumlanan yani biçimsel (figüral) rolleri içeren *hareket katmanı* ise ETKİLEYEN, ETKİLENEN ve TEMA rollerini bünyesinde barındırır. Bu iki katman (1) nolu cümle üzerinden latis diyagramları ile tanımlanabilir. *Hareket katmanında* birincil giriş *figüral (fig)* olmalıdır.

(1) Deniz Ayşe-ye İzmir-den kazak gönderdi. cümlesinde; *kazak tüm* eylem boyunca varlığını gösterdiğinden merkezi rol olarak TEMA rolünü alırken, *İzmir alansal katmanda* KAYNAK tematik rolünü gösterir. Ayrıca *Deniz* ve *Ayşe* kelimelerinin *alansal katman* üzerinde KAYNAK ve HEDEF rollerini üstlenirken, aynı zamanda *hareket katmanında* ETKİLEYEN ve ETKİLENEN rollerini aldığı görülür. Anlaşılacağı üzere bir kelimenin iki farklı rol alması anlam belirsizliğine neden olmaktadır. Bu durum Şekil 3’teki *alansal* ve *hareket katmanları* arasındaki bağlantı noktaları ile konumsal ve biçimsel rollerin hiyerarşik dizilişleriyle giderilebilir (Fırat ve Uğurlu, 2018; Fırat vd., 2014).



Şekil 3. Alansal katman ile hareket katmanı arasındaki bağlantı noktaları

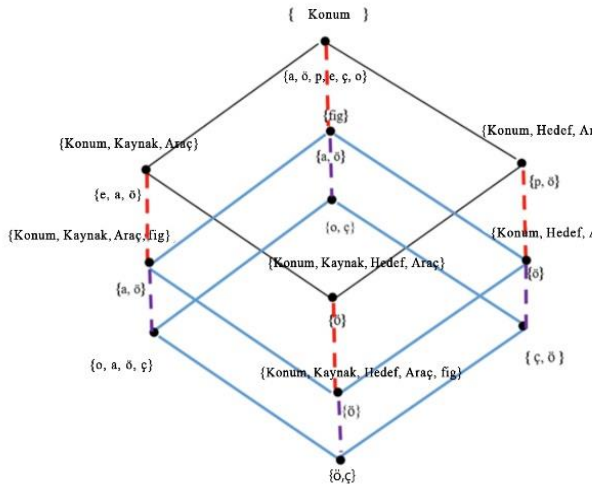
Şekil 3’deki *D, A, I* sırasıyla *Deniz, Ayşe* ve *İzmir* için ilgili sembollerdir. Üstteki dörtgen *alansal katmanı* işaret ederken, alttaki ise *hareket katmanını* gösterir.

Tematik roller üzerinden oluşturulan bu kavram latisleri ile söylem düzeyinde de Türkçe cümlelerin anlamsal analizi yapılabilir. Örneğin Türkçe cümlelerdeki

zamirlerin hangi varlığı temsil ettiği rahatlıkla bulunabilir. Örneğin (2) nolu Türkçe cümle dizisindeki *O* zamirine ait belirsizliğinin, çözümlendiği görülebilir.

(2) Ali otobüsle parka gitti. *O* çantasını evde unuttu.

cümlesinde *O* zamirinin Ali'yi mi, otobüsü mü yoksa parkı mı temsil ettiği belirsizliği Şekil 4'de gösterilen latis diagramında çözüme kavuşmuştur. Şekil 4'deki *a*, *ö*, *p*, *o*, *ç*, *e* sembolleri sırasıyla *Ali*, *otobüs*, *park*, *O zamiri*, *çanta*, *ev* kelimelerini ifade etmektedir.



Şekil 4. (2) nolu cümle dizisindeki *O* zamirinin analizi

Şekil 4'deki düğümlerin temsil ettiği nesnelerin özellikleri yani tematik rolleri dikkate alınarak düğümler üzerinden kavramların hiyerarşisine dikkat edildiğinde *O* zamirinin *Ali* kelimesi ile eşleştiği yani temsil ettiği görülür. Ayrıca (2) nolu cümle dizisinin ikinci eyleminde çanta kelimesi olayın tüm evresinde kendisini gösterdiğinden merkezi rol olarak TEMA tematik rolünü üstlenir. Böylece tematik roller üzerinden önerilen bu gerçeklik modeliyle bir başka ifadeyle tematik rol latisleriyle Türkçe'deki

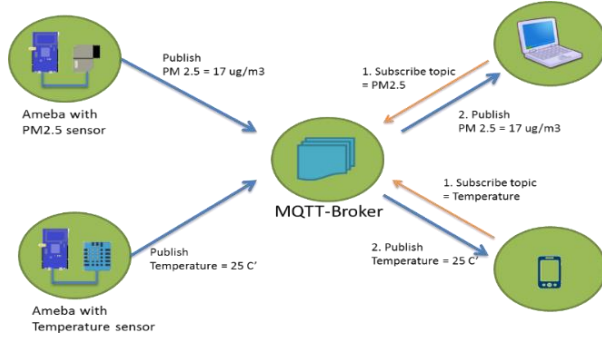
zamirlerin temsil ettiği varlıkların belirsizliği de ortadan kalkmış olur.

Kavram latisleriyle oluşturulan bu gerçeklik modelleri anlam belirsizliklerini ortadan kaldırırken, İşaret Dili Çeviri Sistemine de önemli bir referans olur.

2.3. MQTT protokolü

MQTT, TCP/IP (Transmission Control Protocol/İnternet Protocol-İletim Denetimi Protokolü/İnternet Protokolü) üzerinden çalışan mesajlaşma protokolüdür. Bu protokol, *yayın* (*publish*)-*abone* (*subscriber*) yapısında TCP/IP bağlantısı ile Linux, Windows, Android gibi işletim sistemlerinde çalışır. Aynı zamanda network (ağ) cihazlarının bir *aracıya* (*broker*) yayın yapmasına izin veren bir *yayın/abone* protokolüdür. İstemciler bu *aracıya* bağlanır ve *aracı* istemciler arasındaki iletişimi düzenler. Her cihaz belirlenen *konularda* (*topic*) abone olabilir veya abonelikten çıkabilir. Bir cihaz; abonesi olmuş olduğu *konuyla* ilgili bir mesaj yayınladığında, *aracı* tarafından bu mesaj kendisine iletilir. Ayrıca MQTT iki yönlü çalışan ve durum bilgisini koruyan bir protokoldür. Bir cihaz bağlantıyı kaybederse, abone olan tüm istemciler sunucu tarafından bilgilendirilir. Böylece sistemdeki cihazlardan biri tarafından gönderilen mesaj er ya da geç aynı abonelikteki tüm cihazlara iletilir. MQTT'nin hafifliği ve verimliliği, izlenen veya kontrol edilen verilerin miktarını önemli ölçüde artırmayı mümkün kılar. (Graff, 2018).

MQTT mimarisi Şekil 5'de gösterilmiştir (Kömeçoğlu, 2017).



Şekil 5. MQTT mimarisi

Şekil 5'teki mimaride de görüldüğü üzere, günümüzde Nesnelerin İnterneti sayesinde daha da akıllı hale gelen elektronik cihazlar birbirleri ile haberleşip belirlenen kurallara göre insan hayatını daha da kolaylaştırabilmektedirler. Bu noktada da MQTT, Nesnelerin İnterneti üzerinden geliştirilen projeler için vazgeçilmez bir haberleşme protokolüdür.

2.4. Görüntü işlemede kullanılan sürücü, yazılım geliştirme kütüphaneleri ve çalışma çerçeveleri

Bu çalışmada görüntülerin analizi için gerekli yazılım geliştirme kütüphaneleri ve çalışma çerçeveleri ile Kinect cihazının kullandığı sürücüler Tablo 3'de gösterilmiştir.

Tablo 3. Microsoft Kinect'in sürücüleri ve yazılım geliştirme kütüphaneleri

İsim	Programlama Dili/Sürücü
Microsoft Kinect için USB Sürücüsü	LibusbK-3.0.7.0
SimpleOpenNI Kütüphanesi	Processing
OpenCV (Open Source Computer Vision)	Python
Tensorflow (framework)	Python
Paho MQTT	Python, Processing
Open Kinect	Processing

2.5. Sistemin çalışma evreleri

2.5.1. Sistemin eğitilmesi

Yapılan bu çalışmada Üç Boyutlu Modelleme ve Hareket Analizi olmak üzere iki yöntem kullanılır. Kinect sensörü ile yakalanan

görüntülerden elde edilen veriler ile görüntünün Üç Boyutlu Modellemesi yapılarak her kelime örneği için 100'er adet görüntü kaydedilir. Ayrıca Kinect sensörü ile yakalanan görüntülerden elde edilen veriler ile görüntünün hareketini ifade eden;

periyodik aralıklarla x-y-z pozisyonlarından oluşan çok boyutlu matrisler tek boyutlu dizide toplanır, linear grafiği oluşturulur ve her kelime için 50 şer grafik görüntüsü kaydedilir. Böylece Üç Boyutlu Modelleme ve Hareket Analizi için gerekli veri setleri oluşturulur.

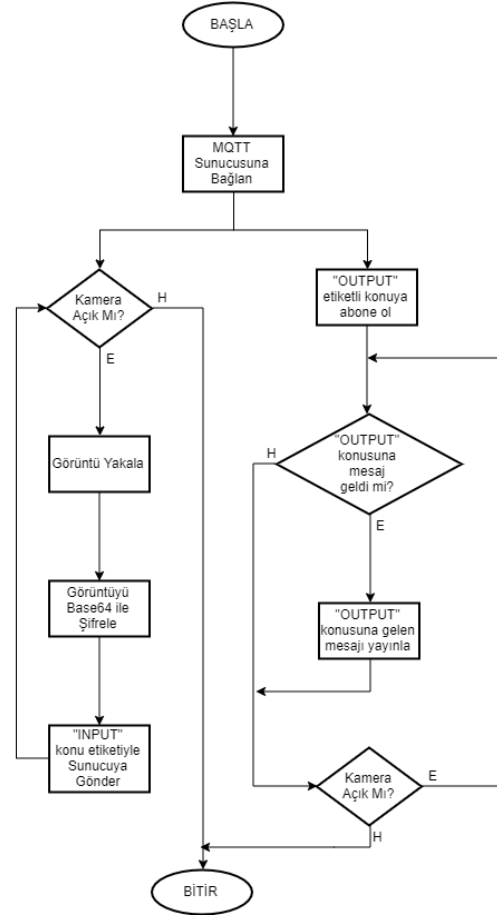
Veri setlerindeki görseller karşılıklarıyla birlikte etiketlenir. 1 adet nörondan oluşan 1 giriş katmanlı, 5 adet nörondan oluşan 1 çıkış katmanlı ve 6 gizli katmanlı bir yapay sinir ağı modeli oluşturularak bu model kaydedilir. Veri setlerindeki görsellerin bu modele giriş olarak sırasıyla verilmesiyle model eğitilir. Eğitilen modelin edindiği tecrübeler kaydedilir.

2.5.2. Sistemin test edilmesi

Kaydedilmiş model edindiği tecrübelerle birlikte oluşturulan program tarafından çağrılır. Veri setlerinde bulunan görüntü dosyalarına benzerlik gösteren rastgele seçilmiş görüntüler, tecrübe sahibi modele giriş olarak verilir. Ve çıktılar kontrol edilir. Bilgisayar tarafından yapılmış tahminlerin, kabul edilebilir oranda doğru olması beklenir.

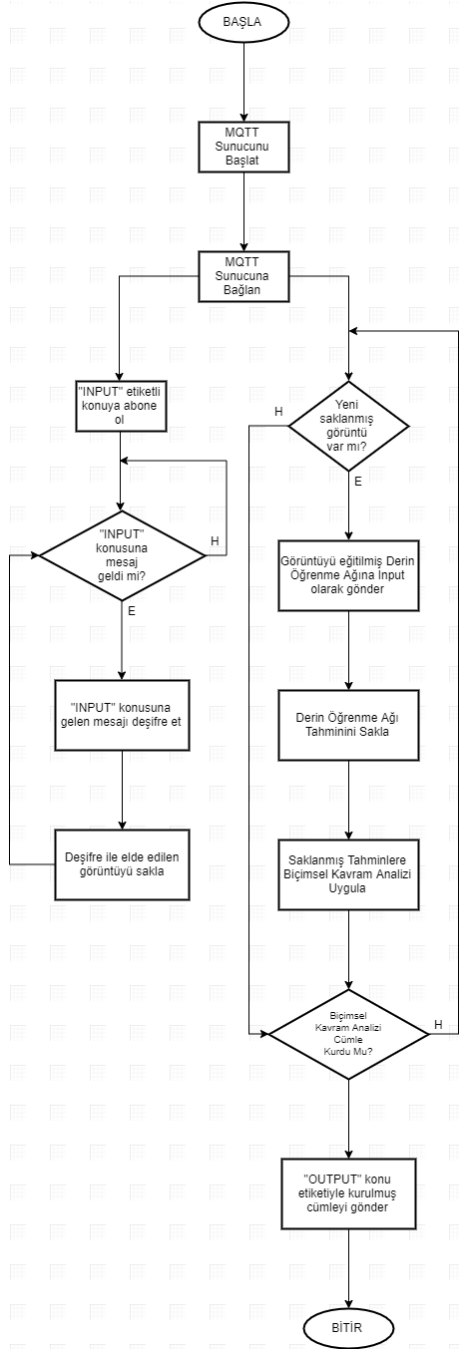
2.5.3. Sistemin çalıştırılması

Gerçekleştirilen sistemin çalışması, lokal ağ üzerinde kurulan MQTT sunucusu ile bu sunucuya bağlanan cihaz(lar) arasında olmaktadır. Şekil 6'da ve Şekil 7'de sırasıyla sistemin istemci ve sunucu tarafındaki çalışmalarının akış şemaları verilmiştir.



Şekil 6. Sistemin istemci tarafındaki çalışmasının akış şeması

Şekil 6'da istemci tarafında sistemin çalışma aşamaları gösterilmiştir. Bu şemada program ilk olarak, MQTT sunucusuna bağlanarak OUTPUT (ÇIKIŞ) konusuna abone olur. Kinect sensörü ile yakaladığı her bir kareyi base64 ile şifreleyerek INPUT (GİRİŞ) konu altında post eder. Sunucunun, istemci tarafından gönderilen bilgilerden yarattığı cümleyi OUTPUT konu altında yayınlaması ile birlikte istek gönderen istemci çeviriyi yakalar ve çıktıyı oluşturur.



Şekil 7. Sistemin sunucu tarafındaki çalışmasının akış şeması

Şekil 7’de ise sunucu tarafında sistemin çalışma aşamaları verilmiştir. Bu şemada da görüldüğü üzere, model önce tecrübeleriyle birlikte program tarafından çağrılır. Daha sonra sistem MQTT sunucusuna bağlanır ve INPUT konusuna abone olur. INPUT konusuna base64 ile şifrelenmiş katar (string) geldiğinde yakalanır, deşifre edilir ve görüntü olarak yorumlanır. Bu yorumlanmış sıralı görüntüler yapay sinir ağı modeline giriş olarak verilerek Hareket ve Üç Boyutlu Modelleme Analizlerinden geçirilerek en büyük olasılıklı birleşik tahmin seçilir. Bu seçim aşağıda verilen (1) nolu matematiksel formülün her bir kelimeye uygulanması ve en büyük birleşik oranın (2 nolu formül) seçilmesi ile yapılır.

$$\text{Birleşik_Oran}_n = \text{nesne}_n_{\text{olasılık}}(\text{üçboyutlumodel}) * \text{nesne}_n_{\text{olasılık}}(\text{hareket}) \quad (1)$$

$$\text{En Büyük Oran} = \text{Birleşik_Oran}_{(\text{max})} \quad (2)$$

Bir sonraki aşamada ise seçilen kelime hafızada tutulur. Bu şekilde base64 ile şifrelenmiş katarların ard arda gelmesi, anlamlı bir cümle oluşturulana kadar devam eder. Sunucuya gelen kelimeler kodlamada oluşturulan metotlar ile tematik rollerine ayrılarak kelimelerin ekleri ve doğru sıralaması elde edilir. Oluşturulan cümle

MQTT sunucusuna OUTPUT konusu altında gönderilir.

3. Bulgular

Veri Seti

Materyal ve metot bölümünde anlatıldığı gibi oluşturulan veri setinden bir kısım Şekil 8’de gösterilmiştir.



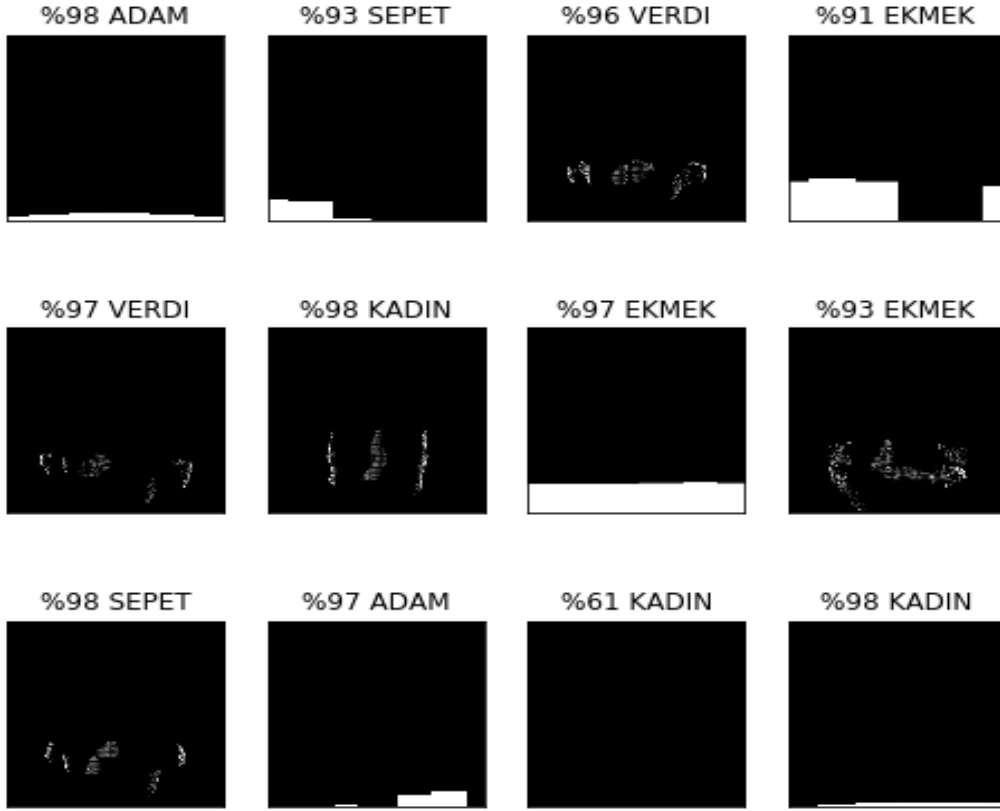
Şekil 8. Oluşturulmuş veri setinden bir örnek

Eğitim

Aynı anlama karşılık gelen benzer görüntülerden oluşan veri seti sistemin eğitilmesinde kullanılmış ve 50 kez pekiştirilmiştir. Eğitilmiş bu sistemi test etmek için girilen 40 görüntü üzerinden çoğunlukla %95 in üzerinde doğruluk oranıyla hareketlerin tanımlanabildiği görülmüştür.

Test

Materyal ve metot bölümünde anlatıldığı üzere test veri setiyle gerçekleştirilen test işleminin çıktıları Şekil 9’da gösterilmiştir.



Şekil 9. Test çıktıları

Çalıştırma

Şekil 10 ve Şekil 11’de de kamera tarafından yakalanan görüntü kareleri eş zamanlı olarak üç boyutlu modellenmiş kareler ve hareket ifade eden grafikler şeklinde gösterilmiştir.



(a)

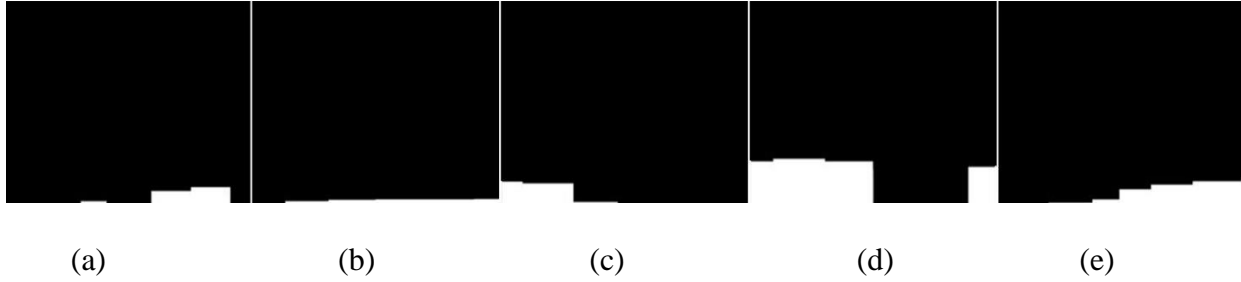
(b)

(c)

(d)

(e)

Şekil 10. Üç boyutlu modellenmiş kareler (a: Adam, b: Kadın, c: Sepet, d: Ekmek, e: Verdi)



Şekil 11. Hareketi ifade eden grafikler (a: Adam, b: Kadın, c: Sepet, d: Ekmek, e: Verdi)

Materyal ve metot bölümünde anlatıldığı ve Şekil 12’de de görüldüğü üzere kelimelerin tahminleri yapılmış, anlam çözümleme modelleriyle rolleri belirlenmiş ve çeviri çıktısı elde edilmiştir. Çeviri çıktısı MQTT protokolüyle OUTPUT *konusu* altında yayınlanmış ve *konunun* abonesi olan istemciler tarafından yakalanmıştır.

```
ADAM : 0.95043653
KADIN : 0.9605443
SEPET : 0.9188244
EKMEK : 0.8538642
VERDI : 0.016100073
{'AGENT': 'ADAM', 'PATIENT': 'KADIN', 'PATH': None, 'LOC': 'SEPET', 'SOURCE': 'SEPET',
'THEME': 'EKMEK', 'GOAL': 'KADIN', 'VERB': 'VERDI', 'PATHOBJ': None, 'PATHRAW': None}
Path is not constructed.
PathObject is not constructed.
Translated Sentence: ADAM KADINA SEPETTEN EKMEK VERDI
```

Şekil 12. Tahminler, roller ve çeviri çıktısı

MQTT'nin çalışma performansı

Geliştirilen bu sistemde istemci-sunucu arasındaki mesajlaşmada, MQTT protokolü 10 saniyede yaklaşık 60 mesajı gönderip 730 mesajı alarak cihazlar arasındaki iletişimi hızlı bir şekilde sağlamıştır.

4. Sonuç ve Tartışma

Yapılan çalışmada; Türkçe İşaret Dilini bilmeyen kişiler için, işaret dilini konuşulan dile yazılı olarak çeviren bir sistem geliştirilmeye çalışılmıştır. Çalışma üç kısımdan oluşmuştur. Birinci kısımda, Türkçe İşaret Dilinden seçilen belirli el hareketleri Kinect sensörü tarafından algılanarak

yakalanmıştır. Elde edilen bu görüntüler Processing ile yazılmış bir istemci program ile MQTT protokolü üzerinden sunucuya gönderilmiştir. Çalışmanın ikinci kısmında, sunucu tarafından işlenen görüntüler yapay öğrenme metotları ile analiz edilerek sembollerin karşılıkları tespit edilmiştir. Çalışmanın üçüncü kısmında ise bulunan bu karşılıkların gerçek anlamlarına Biçimsel Kavram Analizi Teorisi ile hazırlanan tematik rol latisleri üzerinden erişilmiştir. Gerçek anlamları belirlenmiş olan kelimeler daha sonra anlamlı bir biçimde sıralanarak bir cümle haline getirilmiştir. Sunucu tarafından oluşturulan ve yayınlanan bu cümle(ler) istemci tarafından yakalanmıştır.

Bu bağlamda test görüntüleri ile sunulan sonuçlar, önerilen Türkçe İşaret Dili Çeviri Sisteminin, 750 görüntü ile eğitildiğinde ve 40 görüntü ile test edildiğinde el hareketlerini %95’in üzerinde bir doğruluk oranıyla tanıyabildiğini göstermiştir. Ayrıca MQTT protokolünün farklı ortamlarda bulunan cihazları web ortamına bağladığı, sistemle cihazlar arasındaki entegrasyonu kolay ve hızlı bir şekilde gerçekleştirdiği görülmüştür.

Teşekkür

Bu çalışma, Çanakkale Onsekiz Mart Üniversitesi Bilimsel Araştırma Projeleri Koordinasyon Birimince desteklenmiştir.

Proje Numarası: FBA-2018-2663. Birime sonsuz teşekkürlerimizi sunarız.

5. Kaynaklar

Adithya, V., Vinod, P. R. ve Gopalakrishnan, U. 3013. "Artificial Neural Network Based Method for Indian Sign Language Recognition", *IEEE Conference on Information and Communication Technologies (ICT)*, 1080- 1085.

Fırat, Y., Uçar, Ö. ve Kılıçaslan, Y. 2014. "Semantic Analysis with a LatticeBased FrameNet", *Journal of International Scientific Publications: Language, Individual & Society*, 8, 512-518.

Fırat, Y. 2017. "The Semantic Inferences And Mappings Realized In Computer Through The Formal Concept Analysis", *Journal of the International Scientific Researches*, 2 (1), 86-107.

Fırat Y. ve Uğurlu T. 2018. "Latis Tabanlı Anlam Çözümlemesi İle Türkçe İşaret Dili Tercüme Sistemi", *Ömer Halisdemir Üniversitesi Mühendislik Bilimleri Dergisi*, 7 (2), 490-503.

Ganter, B. ve Wille, R.1999. "Formal Concept Analysis Mathematical Foundation", Berlin: Springer, Verlag, 5-23.

Graff, D., (2018), <https://gist.github.com/davidgraeff/6c00d2ad5e4bc553f5f209118307d8fd> Son Erişim Tarihi: 23.09.2019

Ibraheem, N. A. ve Khan, R. 2012. "Survey on Various Gesture Recognition Technologies and Techniques", *International Journal of Computer Applications*, 50 (7), 38-44.

Jackendoff, R. 2006. "On conceptual Semantics", *Intercultural Pragmatics*, 3 (3), 353-358.

Jiang, F., Zhang, S., Wu, S., Gao, U. ve Zhao, D. 2015. "Multi-layered Gesture Recognition with Kinect", *Journal of Machine Learning Research*, 16, 227-254.

Kömeçoğlu, Y., (2017), <http://devnot.com/2017/mqtt-nedir-nasil-bir->

mimaride-calisir/ Son Erişim Tarihi: 23.09.2019

Ma, X. ve Peng, J. 2018. "Kinect Sensor-Based Long-Distance Hand Gesture Recognition and Fingertip Detection with Depth Information", *Journal of Sensors*, 2018, 1-9.

Murata, T. ve Shin, J. 2014. "Hand Gesture and Character Recognition Based on Kinect Sensor", *International Journal of Distributed Sensor Network*, 10 (7), 1-6.

Nikam, A. S. ve Ambekar, A.G. 2016. "Sign language recognition using image based hand gesture recognition techniques", *IEEE 2016 Online International Conference on Green Engineering and Technologies (IC-GET)*, Coimbatore, India.

Pandey, P. ve Jain, V. 2015. "Hand Gesture Recognition for Sign Language Recognition: A Review", *International Journal of Science, Engineering and Technology Research (IJSETR)*, 4 (3), 464-470.

Rajaganapathy, S., Aravind, B., Keerthana, B. ve Sivagami, M. 2015. "Conversation of Sign Language to Speech with Human Gestures", *Procedia Computer Science*, 50, 10-15.

Rajam, P. S. ve Balakrishnan, G. 2011. "Real Time Indian Sign Language Recognition System to aid Deaf and Dumb people", *IEEE 13th International Conference on Communication Technology(ICCT)* , Jinan, China, 737-742.

Tewari, D. ve Srivastava, S. K. 2012. "A Visual Recognition of Static Hand Gestures in Indian Sign Language based on Kohonen Self- Organizing Map Algorithm", *International Journal of Engineering and Advanced Technology (IJEAT)*, 2, 165- 170.

Wittenberg, E., Khan, M. ve Snedeker, J. 2017, "Investigating Thematic Roles through Implicit Learning: Evidence from Light Verb Constructions", *Front Psychol*, 8 (1089).