Image Processing-based Turkish Sign Language Education Application – Mimica

GIZEM GÜNYOLU[1], İPEK KURAL[2], SENA YAREN EBCİN[3],
SUDE ŞAHİN[4], KEMAL GOKHAN NALBANT[5]

Istanbul Beykent University, Faculty of Engineering and Architecture, Department of Software Engineering, Ayazağa, Hadım Koruyolu St. No:19, 34398 Sarıyer/Istanbul
TURKEY

Abstract: Communication involves information exchange between a sender and receiver through various means. Sign language is a form of nonverbal communication that relies on facial expressions and hand movements. Sign language has long been used to help people with hearing impairments communicate and participate in society - from the earliest human history records.This study presents a mobile application designed to support Turkish Sign Language education and public awareness. It is an application for those with hearing loss or who want to communicate with them. With a set of level-based video lessons and an integrated image processing system, the application gives immediate feedback to encourage accurate gesture execution and support effective learning.The project presents a more accessible design by analyzing similar solutions, their shortcomings, and addressing gaps. Main objectives are to spread Turkish Sign Language and increase social inclusion for the hearing impaired.

Keywords: Turkish Sign Language, educational app, image processing, public awareness

Received: July 11, 2025. Revised: August 13, 2025. Accepted: September 9, 2025. Published: November 5, 2025.

1 Introduction

People with hearing and speech impairments rely on sign language, however its low use frequency often creates additional challenges for people with disabilities in their everyday lives. Just as every nation has a spoken language, sign languages also vary from country to country (Öztürk, 2024). Such culturally influenced differences can cause communication barriers even between deaf people from different regions.

A new international sign language has been proposed, yet its use remains largely restricted to multilingual settings like global conferences or professional interpretation (Karaca, 2018).

Although many digital tools are available worldwide for hearing impaired people and those who use them, many international platforms lack Turkish language support. This renders them less accessible to Turkish users. Additionally, almost all local resources in Turkey are dictionary-based, which can make the learning seem mechanical and hard to follow (Vişne & Yıldırım, 2020).

The limited availability of software-supported education options in Turkey makes the Mimica application described herein significant. It is a structured process based on two types of assessments starting with instructional videos from Mimica. The first type involves users choosing from multiple-choice options what signs are supposed to mean in a video. First, they are asked to demonstrate the sign for a given word.

These assessments use 15 randomly selected questions from a dataset to reinforce learning. The practical test component applies image processing to dynamic hand and finger movements recorded by the device's camera. The system then evaluates whether the sign was performed correctly and gives feedback. Technical details are provided in Section 2.1 "Dataset."

An improved design based on a review of similar applications and their limitations has been implemented. A major drawback of existing platforms is the absence of feedback mechanisms to enable users to check the sign accuracy. This causes improper learning. Mimica instead gives immediate, informative feedback to the user.

Additionally, the application includes a dictionary section. Image processing needs camera access to track gestures and motion in real time. Such technology organizes inconsistent or complex movements via classification algorithms (Çalık, 2024).

As digital transformation continues to shape educational landscapes, learning environments have to be adapted for online access (Taşkıran, 2017). The World Health Organization's 2011 World Disability Report said that 31 of the 93 countries surveyed provided no official sign language services and 30 had fewer than 20 interpreters (WHO, 2011). Up-to-date WHO data show that hearing-impaired people constitute about 5% of the population worldwide and projections state that 1 in 10 people will have hearing loss by 2050 (Öztürk, 2024). According to 2023 data from the Ministry of Family and Social Services in Turkey, approximately 2% of the population is hearing impaired.

These figures indicate the need for accessible sign language education. Platforms like Mimica are becoming more important for bridging communication gaps as internet tools are increasingly embedded into education.

1.1 Digital Support for Sign Language Education

Table 1. Degrees of Hearing Loss (Uğur, n.d.).

	CHILDREN	**ADULTS**
Very Mild	(16-25 dBHL) 64-100	(21-35 dBHL) 84-140
Mild	(26-40 dBHL) 104-160	(36-45 dBHL) 144-180
Moderate	(41-55 dBHL) 164-220	(46-55 dBHL) 184-220
Moderate-Severe	(56-70 dBHL) 224-280	(56-70 dBHL) 224-280
Severe	(71-90 dBHL) 284-360	(71-90 dBHL) 284-360
Profound	(91 dBHL and above) 364	(91 dBHL and above)

Table 1 illustrates how everyday activities such as making phone calls, watching television or socializing that are routine for most people can be challenging for those with hearing loss. These difficulties depend on the degree of hearing loss. Studies addressing these challenges have revealed that incorporating mobile technologies into the learning process helps students with hearing loss to improve their access to information (Trezek & Wang, 2006).

1.2 Related Works

Table 2. Literature Review

Authors	Year	Title	Methodologies	Findings	Limitations
Elliott et al.	2000	The development of language processing support for the ViSiCAST project	NLP, Motion Capture, Virtual Avatar	Avatars were used to improve accessibility for the hearing impaired	Lack of language support
Dibeklioğlu et al.	2007	Sign language motion tracking and generating 3D motion pieces using 2D features	Marker-based dataset, 2D to 3D conversion	Enabled 3D motion generation from 2D features	Issues with real-time performance
Özkul	2015	Türk işaret dili için insansı robotlar üzerinde vücutlandırma çalışmaları	Robotics and AI Technologies	Real-time simulation using humanoid robots	Lacked emotional expression
Demircioğlu et al.	2016	Leap Motion ile Türk İşaret Dili Tanıma	Leap Motion Sensor	High recognition accuracy achieved	Focused on only 12 signs

Yalçınkaya et al.	2016	Hareket Geçmişi Görüntüsü Yöntemi ile Türkçe İşaret Dilini Tanıma	MHI & KNN Algorithm	Converted signs detected via camera into text	Decreased performance with complex gestures
Ebling et al.	2018	Smile Swiss German Sign Language dataset	Automated SL-based system	Swiss-German sign language education	Difficult-to-use interface
Bansal et al.	2021	CopyCat: Using sign language recognition to help deaf children acquire language skills	Sign language recognition, Game Design	Educational support tailored for children	Limited accessibility
More et al.	2021	Sign language recognition using image processing	Region filling, object selection	Dataset with 50 variations per sign	Sensitive to lighting conditions
Öztürk et al.	2021	Recognition of Sign Language Letters Using Image Processing and Deep Learning Methods	Image processing, CNN	Achieved 87% true positive rate	13% false negatives due to data limitations
Papadimitriou et al.	2022	Greek Sign Language recognition for the SL-ReDu learning platform	Signer-independent SL tech, low-res camera analysis	High accuracy in recognizing fingerspelled sequences	Requires a laptop, struggles with continuous fingerspelling

According to Table 2, Mimica is unique in having a user-friendly interface and content developed for Turkish Sign Language learners. Like many applications, Mimica lets users track their own progress at three levels - beginner, intermediate and advanced - and score themselves on interactive testing. Also, through gamified elements, Mimica makes learning fun for all ages.

2. Methodology

The system architecture was defined using use case, component, deployment and network diagrams generated during application development process. Both functional and non-functional requirements were precisely defined. A prototyping UI mock-up was designed with user experience in mind and refined based on feedback.

This app was built with Kotlin in Android Studio with the user interface coded and structured. Training of artificial neural networks (ANN) in Python environments started the integration of gesture recognition systems. The trained model was converted to TensorFlow Lite (TFLite) format for native execution on mobile devices. This setup allows capturing hand landmarks from the device camera with the MediaPipe library and receiving real-time classification feedback from the integrated TFLite model.

2.1 Dataset

Instructional content for the TSL alphabet is currently presented as static images/photographs of gestures for visual reference by users. Even though this photo-based version teaches the alphabet now, future updates will add video-based content, in collaboration with certified TSL instructors, for a more engaging and effective learning experience.

There are 29 hand gestures representing the TSL alphabet and essential vocabulary at beginner, intermediate, and advanced levels. Each sign was captured from different angles and poses for model training to ensure the robustness of the dataset; on average, 75 data points per sign were collected. The instructional video content comprises approximately 65 minutes of learning material. The dataset includes examples from the Turkish Sign Language alphabet, introduced at the beginner level of the application, as shown in Figure 1. These gestures are analyzed and compared with user inputs via feature extraction and feedback is given accordingly during practice tests.

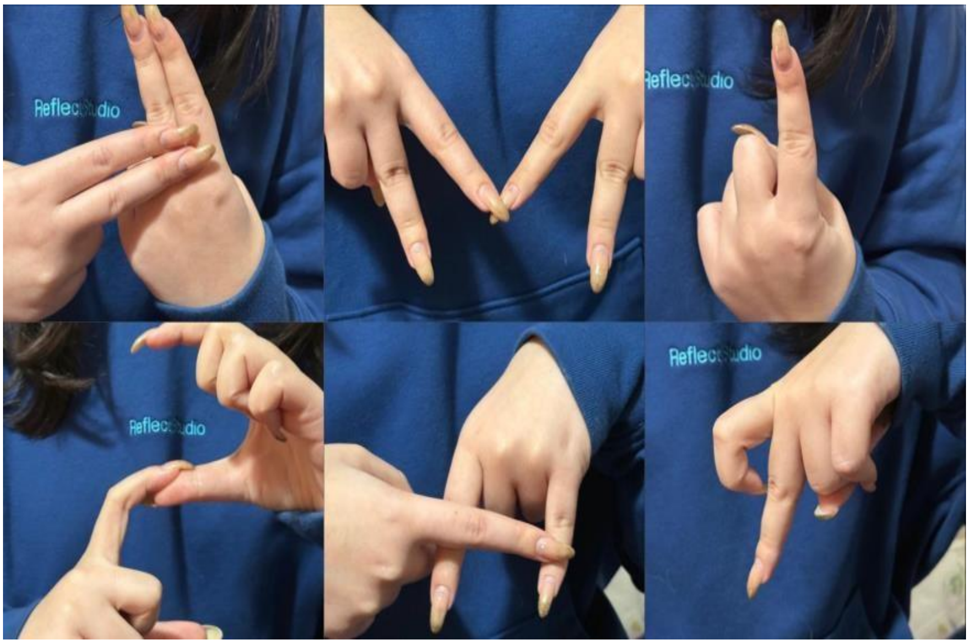

Figure 1 Sample Images from the Turkish Sign Language Alphabet

2.2 Image Processing Technology

At the end of each level - beginner, intermediate, advanced – users are asked to make specific signs during test sessions. These gestures are detected via image processing and feedback is given depending on accuracy.

Real-time hand tracking and a classification model are at the core of the Mimica application recognition system using the Google MediaPipe library. It uses MediaPipe to detect 21 key hand landmarks on the hand -finger joints, fingertips, and the wrist- instead of processing complex raw images directly. This process converts raw image data into a mathematical representation of the hand skeleton. Such a methodology is illustrated in Figure 2, where a hand posture is detected in real time and mapped onto a skeleton during the data collection phase.

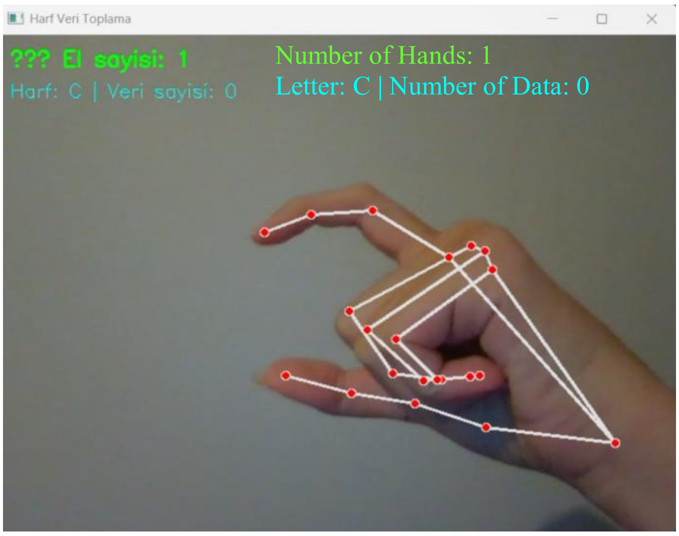

Figure 2. Real-time Landmark Detection using MediaPipe during Data Collection

Image processing workflow includes image capture followed by object clarity enhancement via filtering. This helps the system process hand shapes and gestures faster. Figure 3 summarizes the image processing pipeline in general (Solak & Altınışık, 2018).

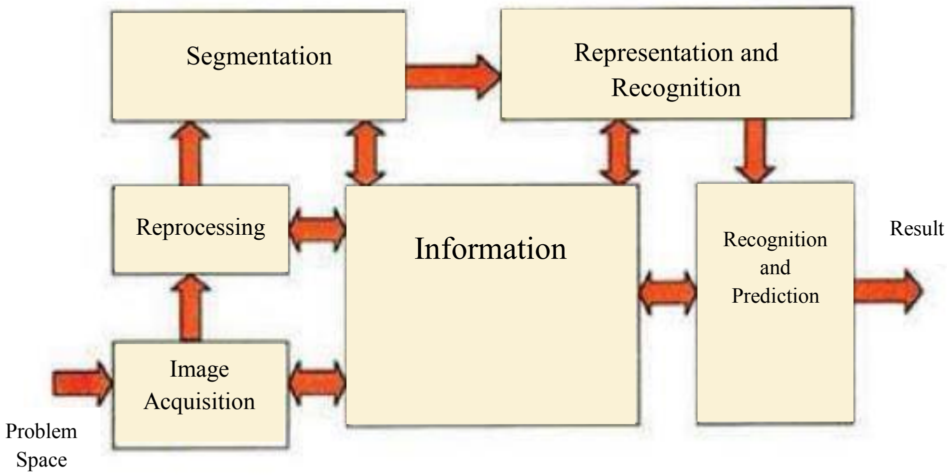

Figure 3. Stages of Image Processing (Smith, 1997).

2.3 Gesture Recognition Methodology

Each detected hand position returns 63 numeric values (21 landmarks x 3 coordinates) along the X, Y, and Z axes. Then these raw coordinates are normalized. Normalization ensures that the model is focused only on the geometric shape of the sign and independent of individual hand size or camera distance. This is achieved by:

1. Reference Point Translation: All landmark coordinates are re-positioned relative to the wrist point, which is the origin.

2. Scale Invariance: Those coordinates are then proportionally scaled by some measure of hand size.

After normalization the 63-value input vector is fed into a trained Artificial Neural Network (ANN) classification model. The model training taught the unique spatial configuration of the finger structure and hand posture for each of the 29 TSL signs. The model then predicts the sign to be performed based on geometric features extracted from the user's hand.

Segmentation is the most challenging step among the overall image processing pipeline steps (Figure 3). It's mainly because visual data is complex and irregular. The MediaPipe landmark detection approach avoids traditional segmentation problems by focusing analysis on critical anatomical points instead of broad image regions.

In digital images, pixels are analyzed as row-and column matrix elements. Each matrix element corresponds to a pixel in Figure 4 (Perihanoğlu, 2015).

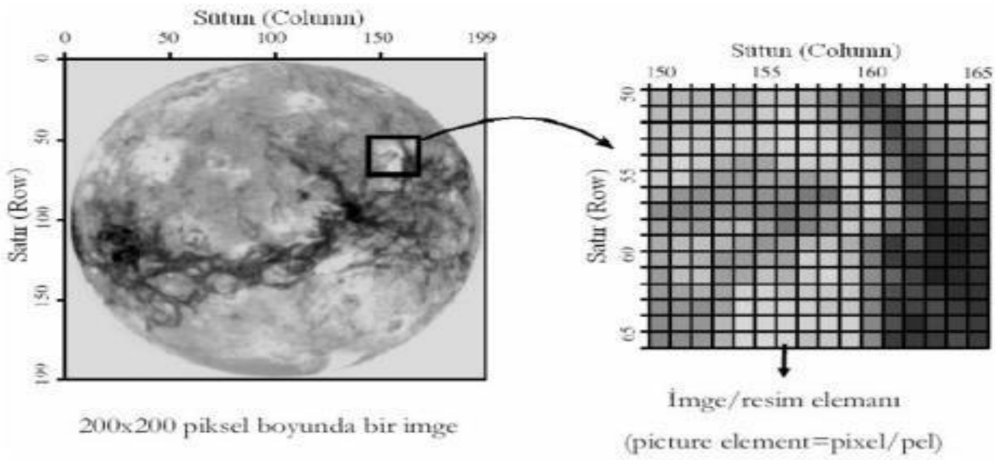

Figure 4. Pixel Representation in a Matrix Structure (Gonzalez & Woods, 1993).

2.4 Motion Analysis Through Image Processing

Motion analysis involves four main phases in image processing:

1. Detecting the presence of motion

2. Locating the moving object

3. Tracking the movement

4. Identifying and labelling the gesture

Segmentation techniques used for motion analysis include:

- Background Subtraction Methods
- Statistical Methods
- Optical Flow Methods

(Peker, 2009)

Gesture-specific motion analysis can be used also in sign language. Such tools include the Microsoft Kinect sensor. Kinect has infrared sensors for 3D depth sensing and real time motion tracking. The device was first used in games and later on in communications and engineering. Figure 5 shows that it can detect skeletal movements and audio input in real time (Zhang, 2012).

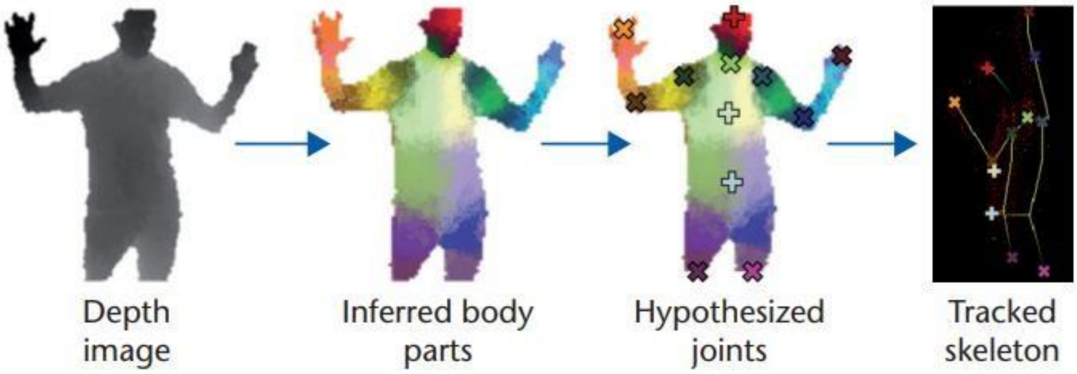

Figure 5. Skeletal Tracking with Kinect Sensor (Microsoft, 2012).

The Kinect is ideal for analyzing sign language gestures in real time - useful for the deaf and hard of hearing. A second is the Leap Motion Controller which analyses finger joints and movements finely enough to capture detailed gestures (Elons et al., 2014).

3. Conclusion and Evaluation

In this work a sign language learning application for hearing or speech impaired persons is developed. The correct usage of gestures and facial expressions in sign language is emphasized, and it is concluded that image processing technologies enhance educational value.

The dataset includes video demonstrations of Turkish Sign Language vocabulary and phrases at beginner, intermediate and advanced levels by a professional instructor. While currently designed for mobile devices, the application is also adaptable to desktop devices, which may expand its reach and user base further.

Its technological design aside, the study demonstrates the social value of inclusive communication environments. The image processing implemented in the system further improves gesture recognition accuracy and provides real-time feedback to users for continuous learning and self-correction. By combining educational design principles with artificial intelligence and computer vision, Mimica provides a scalable and adaptive learning solution.

Future improvements might include enlarging the dataset to include more complex gestures, facial expression detection for semantic accuracy, and gamification to keep user engagement. This is where Mimica offers a new vision of inclusive education that supports digital transformation of learning and makes people with hearing impairments socially active participants.

References

[1]. Bansal, D., Ravi, P., So, M., Agrawal, P., Chadha, I., Murugappan, G., & Duke, C. (2021). CopyCat: Using sign language recognition to help deaf children acquire language skills. In Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems (CHI EA '21), 1–10. https://doi.org/10.1145/3411763.3451523

[2]. Çalık, N. (2024). Ön-eğitilmiş derin sinir ağları kullanılarak ultra geniş bant radar tabanlı el hareketi tanıma. Mühendislik

Bilimleri ve Tasarım Dergisi, 12(1), 205-216. https://doi.org/10.21923/jesd.1418355

[3]. Davis, J. W., & Bobick, A. E. (1997). The representation and recognition of human movement using temporal templates. In Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference on (pp. 928–934). IEEE. https://doi.org/10.1109/34.910878

[4]. Demircioğlu, B., Bülbül, G., & Köse, H. (2016). Leap Motion ile Türk İşaret Dili tanıma. In 2016 24th Signal Processing and Communication Application Conference, SIU 2016 - Proceedings (pp. 589-592). Makale 7495809 (2016 24th Signal Processing and Communication Application Conference, SIU 2016 - Proceedings). Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/SIU.2016.7495809

[5]. Dibeklioğlu, H., Dikici, E., Santemiz, P., Balcı, K., & Akarun, L. (2007). Sign language motion tracking and generating 3D motion pieces using 2D features. 15th IEEE Signal Processing and Communications Applications, Eskişehir, Türkiye, 1–4. https://doi.org/10.1109/SIU.2007.4298843

[6]. Ebling, S., Camgöz, N. C., Boyes Braem, P., Tissi, K., Sidler-Miserez, S., Stoll, S.,

[7]. Hadfield, S., Haug, T., Bowden, R., Tornay, S., Razavi, M., & Magimai-Doss,

[8]. M. (2018). Smile Swiss German Sign Language dataset. In Proceedings of the 11th Language Resources and Evaluation Conference (LREC 2018) (pp. 4221–4229).

[9]. Elons, A. S., Ahmed, M., Shedid, H., & Tolba, M. F. (2014). Arabic Sign Language recognition using Leap Motion Sensor. 9th International Conference on Computer Engineering & Systems (ICCES). https://doi.org/10.1109/ICCES.2014.7030987

[10]. Elliott, R., Glauert, J. R., Kennaway, J. R., & Marshall, I. (2000). The development of language processing support for the ViSiCAST project. Proceedings of the Fourth International ACM Conference on Assistive Technologies, 101–108. https://doi.org/10.1145/354324.354349
Engelli ve Yaşlı İstatistikleri Bülteni. (2024). https://aile.gov.tr/media/135432/eyhgm_istatistik_bulteni_nisan_23.pdf

[11]. Gonzalez, R. C., & Woods, R. E. (1993). Digital image processing. Addison–Wesley Publishing.

[12]. Karaca, M. F. (2018). Üç boyutlu sanal model ile Türk işaret dili simülasyonu [Yayımlanmamış doktora tezi]. Karabük Üniversitesi.

[13]. Microsoft. (2012). Kinect skeletal tracking team honored with the Outstanding Technical Achievement Award. https://www.microsoft.com/about/technicalrecognition/Kinect-SkeletalTracking.aspx

[14]. More, V., Sangamnerkar, S., Thakare, V., Mane, D., & Dolas, R. (2021). Sign language recognition using image processing. JournalNX - A Multidisciplinary Peer Reviewed Journal, 85–87. Retrieved from https://repo.journalnx.com/index.php/nx/article/view/2117

[15]. Özkul, A. (2015). Türk işaret dili için insansı robotlar üzerinde vücutlandırma çalışmaları (Tez No. 504121501) [Yüksek lisans tezi]. İstanbul Teknik Üniversitesi. http://hdl.handle.net/11527/12948

[16]. Öztürk, A., Karatekin, M., Saylar, İ. A., Bardakcı, N. B. (2021). Recognition of Sign Language Letters Using Image Processing and Deep Learning Methods.

[17]. Journal of Intelligent Systems: Theory and Applications, 4(1), 17-23. https://doi.org/10.38016/jista.808458

[18]. Öztürk, Ş. (2024). Yeni eğitimsel Türk işaret dili veri kümesi (E-TSL) kullanarak nöral makine dönüşümü yöntemleri ile sürekli işaret dili çevirisi (Tez No. 896403) [Yüksek lisans tezi, Hacettepe Üniversitesi]. Ulusal Tez Merkezi. Erişim adresi: https://tez.yok.gov.tr/UlusalTezMerkezi/TezGoster?key=UjlM15wKZGQW6T LC0pvCt6X2HnZf6dAX92jcd6D_8v48LAxVAbHL60l0QjSU5YIP

[19]. Papadimitriou, K., Potamianos, G., Sapountzaki, G., Goulas, T., Efthimiou, E., Fotinea, S.-E., & Maragos, P. (2022). Greek Sign Language recognition for the SL-ReDu learning platform. In Proceedings of the 7th International Workshop on Sign Language

Translation and Avatar Technology: The Junction of the Visual and the Textual: Challenges and Perspectives (pp. 79–84). European Language Resources Association.
[20]. Peker, M. (2009). Görüntü işleme tekniği kullanılarak gerçek zamanlı hareketli görüntü tanıma. (Yayınlanmamış Yüksek Lisans Tezi). Sakarya Üniversitesi Fen Bilimleri Enstitüsü; Sakarya.
[21]. Perihanoğlu, G. M. (2015). Dijital görüntü işleme teknikleri kullanılarak görüntülerden detay çıkarımı (Doctoral dissertation, Fen Bilimleri Enstitüsü).
[22]. Smith, S. W. (1997). The scientist and engineer's guide to digital signal processing. https://www.dspguide.com/
[23]. Solak, S., & Altınışık, U. (2018). Görüntü işleme teknikleri ve kümeleme yöntemleri kullanılarak fındık meyvesinin tespit ve sınıflandırılması. Sakarya University Journal of Science, 22(1), 56–65. https://doi.org/10.16984/saufenbilder.303850
[24]. Taşkıran, A. (2017). Dijital çağda yükseköğretim. Açıköğretim Uygulamaları ve Araştırmaları Dergisi, 3(1), 96-109.
[25]. Tonguç, G., (2023). Çokluortam veri işleme ve analiz kütüphaneleri: Mediapipe görüntü işleme araçları. Sosyal, Beşerî ve İdari Bilimler Alanında Gelişmeler 4 (pp.463-488), Ankara: Platanus Publishing.
[26]. Trezek, B. J., & Wang, Y. (2006). Implications of utilizing a phonics-based reading curriculum with children who are deaf or hard of hearing. Journal of Deaf Studies and Deaf Education, 11, 202-213. https://doi.org/10.1093/deafed/enj031
[27]. Uğur, E. (n.d.). Hearing. Doktor Odyolog Emel Uğur. https://emelugur.com/en/hearing/
[28]. Vişne, T., & Yıldırım, S. (2020). İşitme yetersizliği yaşayanların iletişimlerine yönelik bir işaret dili çeviri sisteminin geliştirilmesi ve uzman değerlendirmesi. Cumhuriyet International Journal of Education, 9(3), 679–707. http://dx.doi.org/10.30703/cije.639035
[29]. World Health Organization (WHO), (2011). Dünya engellilik raporu. (Çevrimiçi). https://www.engelliler.biz/forum/image2/Dunya_Saglik_Orgutu_Engellilik_Raporu_Turkce_Ozeti.pdf
[30]. Yalçınkaya, Ö., Atvar, A., & Duygulu, P. (2016). Hareket geçmişi görüntüsü yöntemi ile Türkçe işaret dilini tanıma uygulaması. http://dx.doi.org/10.1109/SIU.2016.7495861