Lessons from Deploying the First Bilingual Peruvian Sign Language - Spanish Online Dictionary

Joe Huamani-Malca[*], Carlos Vasquez[*], Sabina Oporto[*], Francisco Cerna[*], César Ramos[*], Miguel Rodriguez-Mondoñedo[*], Gissella Bejarano[†]

[*]Pontificia Universidad Católica del Perú
San Miguel, Lima-Perú
{huamani.jn, c.vasquezr, sabina.oporto, francisco.cerna, cesar.ramos, mmondon}@pucp.edu.pe

[†]Marist College
Poughkeepsie, NY-USA
gissella.bejarano@marist.edu

Abstract

Bilingual dictionaries present several challenges, especially for sign languages and oral languages, where multimodality plays a role. We deployed and tested the first bilingual Peruvian Sign Language (LSP) - Spanish Online Dictionary. The first feature allows the user to introduce a text and receive as a result a list of videos whose glosses are related to the input text or Spanish word. The second feature allows the user to sign in front of the camera and shows the five most probable Spanish translations based on the similarity between the input sign and gloss-labeled sign videos used to train a machine learning model. These features are constructed in a design and architecture that differentiates among the coincidence for the Spanish text searched, the sign gloss, and Spanish translation. We explain in depth how these concepts or database columns impact the search. Similarly, we share the challenges of deploying a real-world machine learning model for isolated sign language recognition through Amazon Web Services (AWS).

Keywords: bilingual dictionary, bidirectional dictionary, sign language, Peruvian Sign Language (LSP), Spanish

1. Introduction

Bilingual dictionaries are key tools for learners when expanding their vocabulary or translating terms. For sign languages and oral languages, the bidirectional translation is especially more challenging due to their multimodal (visual and text) representations. For example, before online settings and up to now, the printed dictionary built by Stokoe et al. (1976) is considered revolutionary due to the signs listed by their handshapes instead of their English translations (Hochgesang and Miller, 2016). In an online setting, the oral language - sign language direction can be performed by querying a word related to the video of a sign in a database. In the sign language - oral language direction, the user performs a sign in front of a camera and receives the oral language word back. This is a perfect application of video recognition, given that signs can be processed as sequences of frames. Besides, advances in machine learning for action (Asadi-Aghbolaghi et al., 2017), movement (Cust et al., 2019; Li et al., 2021), and gesture recognition (Aggarwal et al., 2023; Adaloglou et al., 2021) have contributed to building better models for isolated sign language recognition (Rastgoo et al., 2021).

Despite the advances in machine learning to develop more sign language recognition technology, the creation of dictionaries could still be limited by the creation of large and varied corpora. For example, more observations of the same sign help train more robust video recognition models and study of signs variations can help on more accurate and standardized annotations. Some sign language apps or websites function as an oral language to sign language word bank through a text search query or letter index vocabulary. The videos on this kind of platforms are usually not accessible as a corpora but as individual videos, i.e. SignASL.com for American Sign language (ASL) [1], signbsl.com for British Sign Language (BSL) [2] or the Auslan Signbank Dictionary for Australian Sign Language (Auslan) [3]. Alternately, more accessible corpora have been created by sign language researchers and linguists (Prillwitz et al., 2008; Crasborn and Zwitserlood, 2008). However, some seemed to have been built mostly for linguistics analysis rather than computational research.

This work presents lessons learned from our attempt to create the first bilingual Peruvian Sign Language - Spanish Online Dictionary (DLSP)[4] highlighting Dos and Dont's. More specifically, we address challenges, learned insights, and share the discussions of our interdisciplinary work in three stages: 1) an information pipeline to preprocess and organize the collection of videos in a database,

[1]www.signasl.org
[2]www.signbsl.com
[3]https://auslan.org.au/about/dictionary/
[4]https://diccionariolsp.pucp.edu.pe/

LREC-COLING 2024, pages 10316–10323

and to feed the training for the machine learning models, 2) the deployment of the machine learning model in AWS SageMaker, and 3) the interface and setting of the video camera input for the search by text and sign. Our team consisted of linguists, a Peruvian Deaf consultant, an Peruvian Sign Language (LSP) interpreter, and machine learning researchers. The videos were all performed by members of the Peruvian Deaf community, which currently has minimal access to education and information (León Pacheco and León Pacheco, 2019).

2. Related work

Bilingual sign language-oral language dictionaries are still very limited, especially for end users. The work of Moryossef (2023) presents bi-directional continuous translation between sign languages and spoken languages [5] based on sign writing datasets. However, this might not be a finalized platform for end users and does not provide instructions in sign language. There are two works closest to ours in the sign language - oral language direction that has just been published. They incorporate a more user-centric approach and have coincided in providing sign language instructions and videos of the possible results so the user can verify the translation. The first is the French Belgian Sign Language - French dictionary [6], a work by Fink et al. (2023), where the users are guided to locate their face and hands so they can produce a video of a sign that will be translated into French. This website allows recognition of 700 signs with a Top-10 accuracy of 83%. The second one (Boháček and Hassan, 2023) presents the initial design of a sign language dictionary system (ASL-English), and user perception results regarding the use of the dictionary and system performance. Additionally to allow the user signing in front of the camera in real time, they allow users to upload their pre-recorded videos. Their underlying sign recognition model is SPOTER (Boháček and Hrúz, 2022), a work that present state of the art for WLASL100 AND WLASL300 where the numbers represent the number of classes or signs addressed. Their final platform was trained on a subset of WLASL2000, another subset of the dataset released by Li et al. (2020). Our work provides explicit lessons for more bilingual dictionary development and paves the way for the design of the database and architecture to support it. We will share our dataset under research request.

[5]https://research.sign.mt/

[6]https://dico.corpus-lsfb.be/

3. Information Pipeline and Database

We describe the available recorded videos, how we annotate them by sign and utterances, and how we use a pipeline to process videos and prepare them for different machine learning, databases, and application tasks (Bejarano et al., 2022).

3.1. Video Recording and Annotation

From the original corpora creation (Rodriguez-Mondoñedo et al., 2023)[7], isolated signs from 24 previous stories were selected and shown to four Deaf participants. Then, they were requested to provide related signs, making a total of 144 signs. In addition, the participants were asked to elaborate a sentence using each of the 144 signs. Over that new dataset, we asked Deaf consultants, CODAs, an interpreter, and linguists to translate and annotate the sign language sentences and isolated signs using ELAN (Wittenburg et al., 2006), a software for video annotation. After a revision to identify unique glosses, we generated up to 517 total items for the dictionary, even when they were not individually elicited from the Deaf participants. The new segmented videos added to the number of samples needed to train the model for the search by sign.

Lesson 1 Request meaning: as in many other languages' translation, a Spanish word can be related to several different signs, and different Spanish words can be associated with the same sign. For that reason, although the elicitation of sentences should be spontaneous, we suggest having additional time to clarify terms and meanings with the participants.

gloss	spanish word(s)	web sign name
ABRIR-CAJON	abrir	Abrir (cajón)
ABRIR-CORTINA	abrir	Abrir (cortina)
ABRIR-PUERTA	abrir	Abrir (puerta)
ABRIR-VENTANA	abrir	Abrir (ventana)

Table 1: Three fields: "gloss", "Spanish word(s), "web sign name", which describe our database signs

3.2. Pipeline and Database

The search by text and by sign are features that share a common database where each entry represents a LSP sign with three fields or columns, as shown in Table 1: a) gloss to name a unique sign defined by their hand shape, location, movement, orientation, and other latent features, b) the

[7]https://hdl.handle.net/20.500.12534/JU4OLG

Figure 1: Showcasing the web sign names as entries of the LSP signs that have their Spanish words internally matching exactly the input text ***abrir***

Spanish word(s), to which the input text will be compared in the search by text and c) the web sign name, which is the text to show once the input text (in the search by text) or the gloss (in the search by sign) have matched to, as shown in Figure 1. For example, if a user searches for the input text or word "abrir" (to open), the platform will match it with three different signs that refer to the action of opening different objects and whose translations or web sign names are:

- abrir (cajón): to open the drawer.
- abrir (cortina): to open a curtain.
- abrir (puerta): to open a door.
- abrir (ventana): to open a window.

Besides these three fields, our database includes the path to a video of the sign to be shown as a result of the search queries. Given the additional number of samples generated by the annotation of the sentences, as explained in 3.1, one same sign had several video examples. The video selected to represent one sign is selected randomly from all these samples.

In Appendix A, we include a diagram of the pipeline that processes the data and stores it in AWS. To streamline our text search service deployment, we adopted the Serverless Framework, which simplified the process of setting up serverless applications on AWS. This approach enabled us to define AWS Lambda functions, DynamoDB tables, and API endpoints within an AWS CloudFormation template, thereby facilitating the organized deployment and effective management of AWS resources required for the functionality of this service. Our implementation was guided by examples found within the Serverless Framework's documentation[8].

Lesson 2 Select what to match: as for now, the match between the input text and the Spanish word(s) field in the database is exact. Making the match more flexible by allowing queries of more than one word match at least one of the words in the Spanish word(s) field could provide another set of results.

Lesson 3 Select video to show as search result entry: while selecting a random video of the sign samples makes the overall process more efficient and standardized for all the signs, this video might not be the best representation of the sign. In case the sign was originally elicited, the corresponding video will be shown as a result in the searches instead of any other entry or video sample generated by the rest of the annotations.

4. Deployment of ML model in the cloud

We utilized the SPOTER SRL architectures model (Boháček and Hrúz, 2022) that receives a skeleton or set of keypoint landmarks as input data. It modifies a transformer model removing the self-attentional module from the decoder for efficiency. Additionally, the decoder is fed by a query class that enables the last layer to perform the classification tasks of sign language recognition. We reduced its number of parameters to work with our dataset of 38 classes which was splited in 80% for training and 20% for validation. Although our validation results obtained 66.19% for Top-1 and 89.52% for Top-5, at the production setting, we obtained close to 22% for Top-1 and 50% for Top-5. Where Top-1 and Top-5 represent that the sign is the most prob-

[8]https://github.com/serverless/examples

able or five of the most probable predicted signs matches the real sign.

Initially, we tried setting up the Inference model service on AWS SageMaker with a custom configuration, which involved preparing the computing environment and building the service from scratch. However, it became time-consuming due to technical complexities, particularly in managing library and program dependencies. To simplify, we used Python PIP directly for dependency installation, Instead of through a conda environment, which allowed us to utilize AWS SageMaker's automatic configuration. Besides, it reduced the "cold start" time, the duration needed for the service to initialize and start serving. This is key in a serverless approach, which used the dynamic service manager of AWS that adjusts resources as needed.

To enhance service security, we opted for using an AWS Lambda as an intermediary service positioned between the website and the translation service. This allows us to modify internal settings without exposing them on the website, and it also facilitate the testing of the core service.

Finally, we developed a deployment script using the AWS SDK for Python, facilitating model deployment and service activation online. Appendix A contains the interaction of the AWS components used in our architecture.

Lesson 4 Keep architecture simple: Navigating the learning curve of AWS services can be challenging at the beginning. However, the abundance of resources in AWS documentation and the active AWS community online facilitate the learning process for the project's specific needs.

Lesson 5 Enhance Security: Security to safeguard against potential cyber-attacks and ensure service capacity when it reaches its limits is vital in developing a website. We use AWS Sagemaker to configure uninterrupted service through the autoscaling functionality and AWS Lambda intermediary service to ensure continuity during updates and provide protection against attacks.

5. Interface and Video Settings

In the search by test functionality, the user inputs a text, which is then compared exactly with the column Spanish Word(s) in the database. If matches with more than one entry, all of their web names and respective videos and sentence examples, are listed as the result. Figure 1 shows the searched text and the web names along with the result of videos and example sentences of each sign related to the searched word.

When searching by a sign, the results to show are the web names of the list of glosses that represent the classes or signs more probable to match the sign in terms of video representation. As this process could involve more steps, we explain it in the following subsections.

5.1. Instructions in Peruvian Sign Language

The written instructions in Spanish were translated into LSP by a Deaf consultant and shown in the video. The instructions contained all the steps since when the user is located in front of the camera, waits for the start of the recording to finally list the results or video entries. We found two main challenges in this stage. The first one is that the signed version of the instructions seemed longer than the written version of it. This might have been due to the fact that the Deaf consultant needed to explain much more background about the web page and the technology it involves, and not only focus on the start and end of time to perform the sign. The second challenge was the Peruvian Deaf community's unfamiliarity with the data and artificial intelligence concepts, making it challenging to convey the idea of Top-5, explained in section 5.3, which was often misunderstood to include the correct sign.

Lesson 6 Retranslation and additional edition: Ideally, Deaf researchers and programmers can validate the signed instructions. If a platform is developed by hearing personnel, the signed instructions should be interpreted back to an oral language. With this measure, the correctness of the translation can be verified. Alternatively, testing instructions with deaf participants, interpreters, or main stakeholders is recommended. Enhancing the video to show interface sections highlighted and click locations can also be beneficial.

Lesson 7 Disseminate about the technical solution: Meet with a group of deaf users to talk about how this technology works inside. This challenge made us think about the need to explain AI concepts, data, probability, and privacy in sign language technology to the Peruvian Deaf community.

5.2. Live Camera

For a sign to be recognized, users must be properly located. More specifically, their locations should be similar to the ones in the videos used for training the AI model. For this reason, we show a set of template key landmarks of a face and a pair of hands to suggest the location and distance from the camera, as shown in Figure 4. The current location of a user is calculated through the user pose estimated by the MediaPipe javascript library from Google Research embedded in the website. Verifying a correct location involves simulating an exact match of the user pose and the template of landmarks. We ensure that the position of the users' eyes is at least within the area of small invisible rectangles surrounding the midpoint between the eyes in the template. Besides, we verify that the users' palms are within the area of small invisible rectangles surrounding the template's palms.

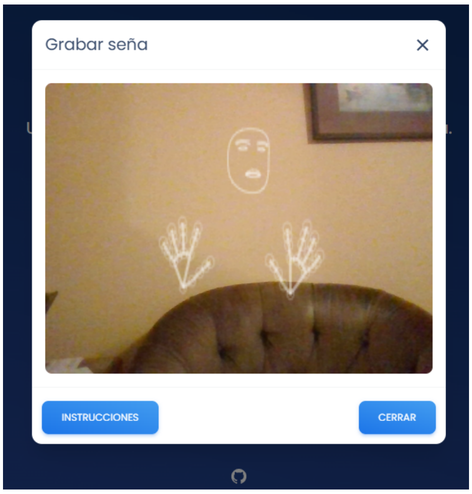

Figure 2: Pose template that guides the user to locate in front of the camera

After a correct position is verified, our platform shows a regressive counter to notify the user when the recording starts. Then, the continuous pose estimations of the movements of the sign are obtained again from the MediaPipe library. However, the pose estimation is a model that consumes significant time to generate real-time results on the website. Therefore, some frames are transmitted without pose estimation and ignored from the input sequence for the inference model, resulting in smaller frames per second than the videos used in the training stage of the model.

Our tests reunite a group of deaf potential users in a computer lab. They were driven to use the search by text and the search by sign. When performing their signs, each of the computer cameras recorded not only the person in front of it but the rest of the persons behind them. This may cause discrepancies between training videos and participants' streams in the real world.

Lesson 8 Ensure Similar Stream Frequency: To the best of our knowledge, not all the streamed frames will be processed by MediaPipe, given that some of them will be lost while waiting for the results or previous frames. We are trying to implement a buffer or option that allows us to process all of them. We want to highlight the importance of comparing the distribution of the video's length due to the streaming process rate or the camera frame. In that way, the sequences of frames produced on the website are similar to the videos used for training.

Lesson 9 Ensure Similar Video distribution: It could be possible that the MediaPipe model had combined the body structure of more than one person in one video streaming. Besides, the noisy backgrounds are something to consider when training machine learning models. In our case, the machine learning model was trained on video signs with white backgrounds only. In general, when training the machine learning models, the video samples should be modified or simulated as much as possible videos coming from the real platform interface.

5.3. Top-5 Results

Besides, when presenting the results of the search by sign, our tool links each of the five most probable classes or gloss predictions to their respective sign video. This is similar to the search by text, but the difference is that the search is performed by the gloss names or classes found matching the input sign. A unique result or top-1 has a lower probability of showing the correct result and, therefore a lower performance.

Lesson 10 Include videos of the possible results: Even when the users are warned that the search by sign does not provide the correct result 100% of the times, they can verify the result by watching the video corresponding to each sign.

6. Conclusions

We have presented lessons learned during our work on building the first bilingual Peruvian Sign Language - Spanish Dictionary. We explain the two main functionalities of this dictionary: the search by text, where the input is a text and the results are a list of videos and their respective web names, and the search by sign, where the input is a sign performed in front of a camera and the results are the top-5 most probable glosses of the signs similar to the input sign. We have highlighted lessons learned about the creation and processing of the dataset, the deployment of the machine learning model as a cloud service, and the interface and camera settings.

7. Acknowledgements

We express our gratitude to Amazon Web Services (AWS) for their research grant supporting this project. Our machine learning models were trained in Chameleon, supported by the National Science Foundation (Keahey et al., 2020). Finally, to Jenny Vega and Daniel Carbajal for their advice on AWS services, Juan Villamonte for the LSP consultancy, Alexandra Arnaiz for interpreting LSP on the sessions with the deaf users, and signing the instructions video, and Julio Mendoza and Rossy Latorraca for the initial web and cloud design.

8. Ethical considerations

The research involves the work of Deaf participants, who have been thoroughly informed about the use of their data for annotations in the dataset and have provided consent for testing procedures. It is important to note that the Deaf persons involved in the annotation process may not be the same individuals participating in the testing procedures. On the other hand, our website does not store videos of users used for LSP to Spanish translation. As part of our ethical practices, we plan to organize a talk to communicate the results of this research to the Deaf community in Peru.

9. Bibliographical References

Nikolas Adaloglou, Theocharis Chatzis, Ilias Papastratis, Andreas Stergioulas, Georgios Th Papadopoulos, Vassia Zacharopoulou, George J Xydopoulos, Klimnis Atzakas, Dimitris Papazachariou, and Petros Daras. 2021. A comprehensive study on deep learning-based methods for sign language recognition. *IEEE Transactions on Multimedia*, 24:1750–1762.

Apeksha Aggarwal, Nikhil Bhutani, Ritvik Kapur, Geetika Dhand, and Kavita Sheoran. 2023. Real-time hand gesture recognition using multiple deep learning architectures. *Signal, Image and Video Processing*, 17(8):3963–3971.

Maryam Asadi-Aghbolaghi, Albert Clapés, Marco Bellantonio, Hugo Jair Escalante, Víctor Ponce-López, Xavier Baró, Isabelle Guyon, Shohreh Kasaei, and Sergio Escalera. 2017. Deep learning for action and gesture recognition in image sequences: A survey. *Gesture Recognition*, pages 539–578.

Gissella Bejarano, Joe Huamani-Malca, Francisco Cerna-Herrera, Fernando Alva-Manchego, and Pablo Rivas. 2022. PeruSIL: A framework to build a continuous Peruvian Sign Language interpretation dataset. In *Proceedings of the LREC2022 10th Workshop on the Representation and Processing of Sign Languages: Multilingual Sign Language Resources*, pages 1–8, Marseille, France. European Language Resources Association.

Matyáš Boháček and Saad Hassan. 2023. Sign spoter: Design and initial evaluation of an automatic video-based american sign language dictionary system. In *Proceedings of the 25th International ACM SIGACCESS Conference on Computers and Accessibility*.

Matyáš Boháček and Marek Hrúz. 2022. Sign pose-based transformer for word-level sign language recognition. In *Proceedings of the IEEE/CVF winter conference on applications of computer vision*, pages 182–191.

Onno Crasborn and Inge Zwitserlood. 2008. The corpus NGT: an online corpus for professionals and laymen. In *Proceedings of the 3rd Workshop on the Representation and Processing of Signed Languages: Construction and Exploitation of Sign Language Corpora*, pages 44–49, Marrakech, Morocco. European Language Resources Association.

Emily E Cust, Alice J Sweeting, Kevin Ball, and Sam Robertson. 2019. Machine and deep learning for sport-specific movement recognition: A systematic review of model development and performance. *Journal of sports sciences*, 37(5):568–600.

Paula Escudeiro, Nuno Escudeiro, Rosa Reis, Jorge Lopes, Marcelo Norberto, Ana Bela Baltasar, Maciel Barbosa, and José Bidarra. 2015. Virtual sign – a real time bidirectional translator of portuguese sign language. *Procedia Computer Science*, 67:252–262. Proceedings of the 6th International Conference on Software Development and Technologies for Enhancing Accessibility and Fighting Info-exclusion.

Jérôme Fink, Pierre Poitier, Maxime André, Loup Meurice, Benoît Frénay, Anthony Cleve, Bruno Dumas, and Laurence Meurant. 2023. Sign language-to-text dictionary with lightweight transformer models. pages 5968–5976.

Ye Gu, Xiaofeng Ye, Weihua Sheng, Yongsheng Ou, and Yongqiang Li. 2020. Multiple stream deep learning model for human action recognition. *Image and Vision Computing*, 93:103818.

Julie A. Hochgesang and Marvin T. Miller. 2016. Special feature: A celebration of the dictionary of american sign language on linguistic principles: Fifty years later. *Sign Language Studies*, 16(4):563–591.

Kate Keahey, Jason Anderson, Zhuo Zhen, Pierre Riteau, Paul Ruth, Dan Stanzione, Mert Cevik, Jacob Colleran, Haryadi S. Gunawi, Cody Hammock, Joe Mambretti, Alexander Barnes, François Halbah, Alex Rocha, and Joe Stubbs.

2020. Lessons learned from the chameleon testbed. In *2020 USENIX Annual Technical Conference (USENIX ATC 20)*, pages 219–233. USENIX Association.

Lucía León Pacheco and Olenka León Pacheco. 2019. Todo niño tiene derecho a aprender en su lengua materna: el caso de la comunidad sorda en el perú. In *II Congreso Iberoamericano Interdisciplinario de la Familia e Infancia: Violencia contra Niños, Niñas y Adolescentes en Colectivos Vulnerables. Concurso Internacional de Ponencias*, pages 91—117, Lima, Perú. Pontificia Universidad Católica del Perú. Departamento Académico de Derecho. Centro de Investigación, Capacitación y Asesoría Jurídica.

Dongxu Li, Cristian Rodriguez, Xin Yu, and Hongdong Li. 2020. Word-level deep sign language recognition from video: A new large-scale dataset and methods comparison. In *Proceedings of the IEEE/CVF winter conference on applications of computer vision*, pages 1459–1469.

Xiaoli Li, Peilin Zhao, Min Wu, Zhenghua Chen, and Le Zhang. 2021. Deep learning for human activity recognition.

Amit Moryossef. 2023. sign.mt: Real-time multilingual sign language translation application.

Siegmund Prillwitz, Thomas Hanke, Susanne König, Reiner Konrad, Gabriele Langer, and Arvid Schwarz. 2008. DGS: Corpus project – development of a corpus based electronic dictionary German Sign Language / german. In *Proceedings of the 3rd Workshop on the Representation and Processing of Signed Languages: Construction and Exploitation of Sign Language Corpora*, pages 159–164, Marrakech, Morocco. European Language Resources Association.

Razieh Rastgoo, Kourosh Kiani, and Sergio Escalera. 2021. Sign language recognition: A deep survey. *Expert Systems with Applications*, 164:113794.

Yukai Song, William Taylor, Yao Ge, Muhammad Usman, Muhammad Ali Imran, and Qammer H Abbasi. 2022. Evaluation of deep learning models in contactless human motion detection system for next generation healthcare. *Scientific Reports*, 12(1):21592.

William C. Stokoe, Dorothy C. Casterline, and Carl G. Croneberg. 1976. *A dictionary of American Sign Language on linguistic principles*, 1st edition. Linstok Pr.

Guilherme Augusto Silva Surek, Laio Oriel Seman, Stefano Frizzo Stefenon, Viviana Cocco Mariani, and Leandro dos Santos Coelho. 2023. Video-based human activity recognition using deep learning approaches. *Sensors*, 23(14):6384.

Haijing Wang, Alexandra Stefan, Sajjad Moradi, Vassilis Athitsos, Carol Neidle, and Farhad Kamangar. 2012. A system for large vocabulary sign search. In *Trends and Topics in Computer Vision*, pages 342–353, Berlin, Heidelberg. Springer Berlin Heidelberg.

Peter Wittenburg, Hennie Brugman, Albert Russel, Alex Klassmann, and Han Sloetjes. 2006. ELAN: a professional framework for multimodality research. In *Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC'06)*, Genoa, Italy. European Language Resources Association (ELRA).

Jimin Yu, Maowei Qin, and Shangbo Zhou. 2022. Dynamic gesture recognition based on 2d convolutional neural network and feature fusion. *Scientific Reports*, 12(1):4345.

10. Language Resource References

Rodriguez-Mondoñedo, Miguel and Arnaiz, Alexandra and Pérez-Silva, Jorge and Velázquez, Héctor and Carlín Mariana and Raico, Giovanna and Marino, Mateo and Ramírez, Abraham and Zúñiga, Joselin and Moreno, Ana Rosa and Villamonte, Juan and Oporto, Sabina and Cerna, Francisco and Ramos, César and Vasquez and Bejarano, Gissella. 2023. *Materiales para el Diccionario Anotado de Lengua de Señas Peruana. Proyecto PUCP DGI-305*.

Appendices

A. Appendix: Diagrams for Pipeline Process and AWS components

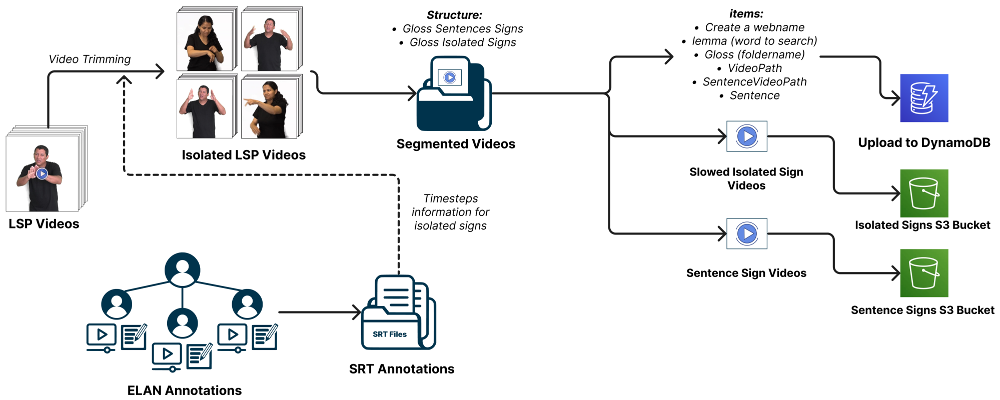

Figure 3: Video segmentation and gloss processing for storage in a database in AWS

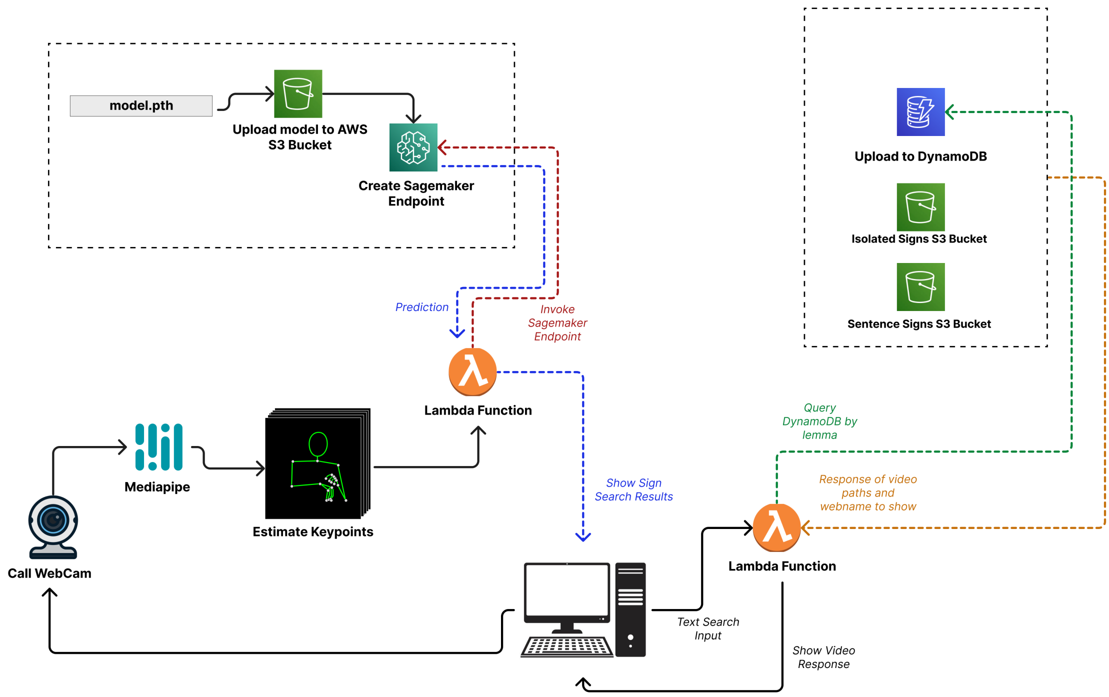

Figure 4: Platform components interaction: website, text-based search and sign-based search