







Real-Time Recognition of Peruvian Sign Language Using Convolutional Neural Networks (CNNs)

Sonia J. León-Jimenez¹ , Claudia León-Chavarri² , Rafael Chavez-Ugaz³ , Lucia B. Suni-Chavez⁴ ,
Fabricio H. Paredes-Larroca⁵ , Ezilda M. Cabrera-Gil⁶ 

¹(20201141@aloe.ulima.edu.pe), ²(cleon@ulima.edu.pe), ³(rchavezu@ulima.edu.pe), ⁴(20194642@aloe.ulima.edu.pe),
⁵(fparedes@ulima.edu.pe), ⁶(ecabrera@ulima.edu.pe)

¹²³⁴⁵⁶ Carrera de Ingeniería Industrial, Universidad de Lima, Perú

Received: 15 August, 2025 / Accepted 5 September, 2025 / Published: 5 June, 2026
<https://doi.org/10.26439/ciii2025.8660>

ABSTRACT—Inclusive education for people with hearing impairments in many countries still lacks accessible technological tools. This work introduces a prototype for automatic translation of the Peruvian Sign Language (PSL) finger alphabet based on convolutional neural networks (CNNs) combined with support vector machines (SVMs). The system recognizes letters in real time without requiring additional sensors or wearable devices. A proprietary dataset containing up to 50 images per class was used for training under controlled conditions. The prototype achieved an average accuracy of 97%, a word error rate (WER) of 15%, and a response time of 1.8–2.0 s and a processing speed of up to 125 frames per second (fps). These results demonstrate the viability of the system as an inclusive educational tool in both controlled environments and real-life school settings.

Index Terms—Assistive technology, convolutional neural networks (CNNs), inclusive education, Peruvian Sign Language (PSL), sign language recognition.

I. INTRODUCTION

The inclusion of people with hearing disabilities in the education system continues to be a considerable challenge worldwide, especially in developing countries such as Peru [1]. Despite various institutional efforts, deaf students continue to face persistent barriers, including a shortage of qualified interpreters, limited teacher training in the use of Peruvian Sign Language (PSL), and the low availability of inclusive technological tools [1], [2], [3].

According to the World Health Organization (WHO), approximately 5% of the world's population—about 430 million people—live with significant hearing loss, a figure that is projected to increase to one in ten people by 2050 [7]. In Peru, children living in poverty are at increased risk

of hearing impairment, partly due to untreated middle ear disease and limited access to pediatric ear care, which is associated with worse academic performance [22]. In low- and middle-income countries, people with hearing loss frequently experience restricted access to education, employment and health services, leading to substantial psychosocial burden [8]. Limited access to appropriate language in deaf children can lead to delays in language development, which can cause academic problems and social exclusion. These challenges can also influence emotional development, particularly in contexts where there are communication barriers that restrict interaction with colleagues or educators. [9]. In the long term, the educational barriers can negatively impact the academic outcomes, social participation and inclusion in society. [10].

In response to this problem, deep learning-based technologies, such as CNNs, have demonstrated remarkable effectiveness in visual recognition tasks, including automatic sign interpretation [4], [13], [14]. CNNs are particularly suitable for real-time applications as they enable efficient processing and faster inference compared to traditional computer vision approaches [14], [16]. Recent studies indicate that the integration of CNNs with techniques such as support vector machines enhances classification accuracy for highly similar gestures, thereby increasing their applicability in educational contexts [14], [17].

Currently, several prototypes aimed at improving communication for people with hearing disabilities have been proposed, using a variety of technologies, as described and illustrated in Fig. 1. However, many of these approaches exhibit adaptability limitations and rely on wearable devices, such as sensorized gloves, whose accuracy typically ranges from 93.4% to 96% under standard conditions [5], [6]. Although some studies using specialized sensors have achieved accuracies of up to 96% in controlled scenarios [5], [6], these solutions remain intrusive, expensive, and

impractical for deployment in conventional classroom environments. In contrast, international studies based on computer vision and CNN models report more consistent accuracies, typically >90% accuracy, in some controlled scenarios above 95% [4], [13], Using polynomial classifiers, recognition accuracy can range from 93.41% to 98.4%, depending on the number of training patterns [25]. These findings indicate that non-invasive and context-aware approaches represent a more viable and relevant alternative, thereby reinforcing the proposed approach of this study.

This project proposes an autonomous, portable, and non-intrusive tool capable of recognizing PSL gestures in real time and automatically translating them into text and voice, without requiring gloves or an internet connection. Using an HD camera and local processing on a computer with an integrated GPU, the proposed system is positioned as an accessible alternative for public educational institutions. Unlike models based on international datasets (ASL, ISL) [11], [17], this prototype will be trained using data specific to the Peruvian context, thereby increasing its relevance and accuracy in national scenarios. It is expected to achieve an average accuracy in the range of 95% to 99%, a word error rate (WER) of approximately 15%, and a response time of no more than 2.5 s, so that the tool can effectively contribute to the educational inclusion of people with hearing disabilities through a scalable, accessible, and adaptable solution to the Peruvian school system.

A. Prototypes Based on Virtual Dictionaries and Educational Platforms

One of the most relevant developments in the national context is the PSL (PUCP) Virtual Bilingual Dictionary, which enables the translation of approximately 750 words from Spanish into sign language through manual on-screen selection [2]. Another example is *Yapaykuy*, a mobile application that converts speech to text and text to speech, with plans to incorporate predefined gesture recognition [3]. These initiatives have favored basic access to sign language; however, their operation remains static and does not support real-time visual recognition. Furthermore, these solutions rely on limited databases and do not facilitate natural interaction in dynamic educational contexts, reflecting the gaps that still exist in inclusive education in Peru [1], as well as the international challenges reported in terms of accessibility and educational outcomes for people with hearing disabilities [7]–[10].

- Comparative technical attributes are summarized as follows:
- Input: Manual selection; no automatic visual recognition [2] [3].
- Output: Text and speech; gestural fluidity is not considered [2].
- Limitation: They depend on closed databases without real-time scalability [1] [7].

B. Prototypes Based on Physical Devices (Gloves, Sensors and Wearables)

At the international level, some developments have used physical devices such as sensorized gloves to capture hand movements. Promising recognition performance in controlled settings has been approached by wereable instrumented gloves and inertial sensors; however, they require calibration and may reduce the comfort and portability in an educational environment. Although these solutions have achieved accuracies of 96.3% in segmented samples and 91.2% in real time—and in some cases up to 96% when using specialized sensors— [5], [6], they present limitations in terms of cost, user comfort, and the need for constant calibration. Furthermore, several studies have indicated that deaf people prefer non-intrusive technologies that respect their natural body language [12]. In this sense, although technically effective, gloves interfere with natural gestures and are impractical for prolonged use in educational environments.

- Comparative technical attributes are summarized as follows:
- Accuracy: >90%, but dependent on calibration, hardware used and constant maintenance [5] [6].
- Cost: high because it requires specialized hardware [5] [6].
- Limitation: They interfere with natural gestures and reduce comfort in a classroom [12].

C. Prototypes Based on Visual Recognition With Machine Learning

The latest systems use deep learning models, such as CNNs, trained on gesture images. Examples include SignNet and BranTNet—the latter based on transfer learning—which have been designed for sign languages such as ASL or ISL, using high-quality standardized datasets [4], [11], [13], [14]. These models have reported accuracies ranging from 93% to 98% under controlled conditions [4], [13], [25]. However, these models typically require large volumes of data, connectivity to external servers, and training in cultural contexts different from that of Peru. Furthermore, in many cases, they focus exclusively on isolated letters or words, without considering sentence-level structure or the particularities of the PSL [11], [17].

Comparative technical attributes are summarized as follows:

- Accuracy: above 90% under controlled conditions [4], [13], [25].
- Requirements: large volumes of data and connectivity with external servers [11], [17].
- Limitation: trained in international contexts, without considering particularities of the PSL [11], [17].

II. METHODOLOGY

Quality Function Deployment (QFD) was used as a comparative analysis tool, allowing the attributes that the proposed prototype must address to be evaluated against existing solutions. To this end, the main user requirements were reviewed: clear and accessible visualization, real-time gesture recognition, portability and ergonomics, simultaneous audio and video feedback, and adaptation to the PSL. Each requirement was assigned a weighting based on its relevance, and the performance of the proposed prototype was compared with that of several international patents [19], [20], [21].

The results show that the proposed system achieved the maximum score across all criteria, achieving 97% accuracy, operating in real time at 125 fps, working with a standard, low-cost, non-intrusive camera, offering immediate visual and auditory feedback, and being trained on a proprietary PSL dataset, thereby ensuring cultural relevance. The dataset was generated under controlled but variable conditions, using a 1080p webcam to capture the 28 letters of the PSL, including “Ñ”, with at least 50 images per class. Data acquisition sessions were conducted on different days, introducing variations in lighting, distance, and hand orientation to better simulate real-world environments. Gesture references were obtained through direct visual observation in a Peruvian Sign Language (PSL) school and were subsequently replicated during data collection, ensuring that the captured samples reflected the authentic characteristics of the local context. In contrast, the reviewed patents, while reporting accuracies above 90%, rely on gloves or other specialized devices, provide partial feedback, and focus on foreign sign languages, thereby limiting their ergonomics, accessibility, and applicability to the Peruvian context.

The proposed modular system for PSL translation consists of various phases that integrate computer vision and deep learning techniques. The main performance indicators targeted were: achieving an accuracy between 95% and 99 %, ensuring real-time operation with a minimum speed of 120 fps, maintaining a response time of less than 2 s, obtaining a WER of no more than 15% in dynamic sequences, and ensuring system non-intrusiveness through the use of a low-cost, standard camera adapted to the PSL context.

The experimentation phases followed a validation protocol widely used in computer vision [13], [14], consisting of dividing the dataset into 80% for training and 20% for testing, ensuring the evaluation of the model with data not seen during learning. In contrast, previous studies based on LSTM architectures have evaluated performance using sequence-based metrics such as BLEU, ROUGE, and CIDEr [15]. In real-time experiments, the system response rate in frames per second (fps) was also measured, and a five-frame moving average was applied to smooth predictions, thus stabilizing the output and reducing instantaneous

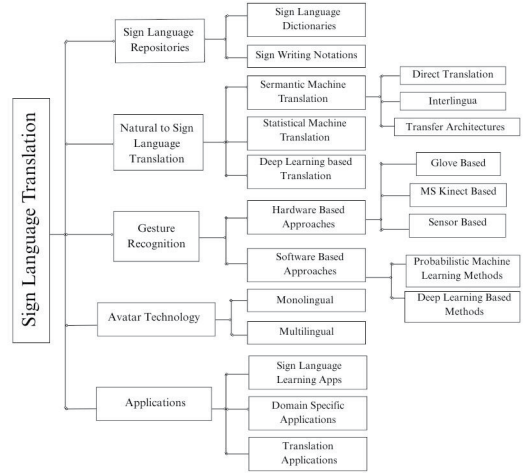


Fig. 1. Assistive technology for people with hearing impairments.

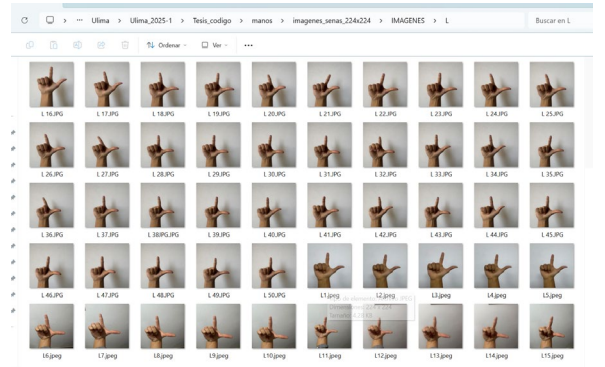


Fig. 2. Dataset of letter L.

classification fluctuations, improving output stability and reducing instantaneous classification fluctuations.

The experimental process was progressively designed in three phases, using 5, 15 and 50 images per class. This strategy enabled the evaluation of the impact of dataset size on the model's generalization capability. The deep learning literature indicates that small datasets tend to induce overfitting, which limits the system's ability to recognize new variations in the samples [13], [14]. In the first phase, using only five images per class, the model exhibited low accuracy and high classification instability. In the second phase, with 15 images per class, a substantial improvement in prediction stability was observed, reflecting increased diversity in the training samples. Finally, in the third phase, using 50 images per class, the system achieved its best performance, attaining a competitive level of accuracy consistent with that reported in international CNN-based sign recognition studies, where accuracies exceed 90% on larger datasets [4].

A. Data Collection and Use of Formulas

A proprietary dataset was constructed from video captures acquired with a high-resolution webcam (1080p at 30 fps), focusing exclusively on the user’s hands against a neutral background. For each of the 28 letters of the PSL alphabet, including “Ñ”, a minimum of 50 images per class were captured, as shown in Fig. 2. Sessions were conducted on different days, with intentional variations in lighting, distance, and orientation to ensure diversity and improve model generalization, as recommended in [13].

For letters involving movement, such as “J”, “Ñ”, and “Z”, video sequences were recorded and segmented into individual frames at a rate of 5-10 fps, following a strategy similar to that used in DeepASLR [14] and in the hybrid architectures described in [17], as illustrated in Fig. 3.

B. Preprocessing

The preprocessing pipeline consisted of several stages, including conversion to grayscale to reduce computational complexity, resizing to 64x64 pixels to standardize the input of the models, and normalization of pixel values in a range between 0 and 1. Additionally, data augmentation techniques were applied, such as random rotation, translation, horizontal mirroring, and brightness modification, with the aim of increasing system robustness to real-world environmental variations. This strategy follows the approach suggested in [23] for LSTM-based models trained with augmented data. The corresponding Python code is as follows:

```
def preprocesar_frame(frame):
    roi=frame[ROI_TOP_LEFT[1]:ROI_BOTTOM_
             RIGHT[1],ROI_TOP_LEFT[0]:ROI_BOTTOM_
             RIGHT[0]]
    resized=cv2.resize(roi,(IMG_WIDTH,IMG_HEIGHT))
    normalized = resized / 255.0
    return np.expand_dims(normalized, axis=0), roi
```

C. Model Architecture

The basic architecture for static sign recognition consisted of a CNN with three convolutional layers, intermediate pooling layers, ReLU activation functions, and a final fully connected layer with SoftMax for multiclass classification as illustrated in Fig. 4. To improve accuracy for gestures with a high degree of similarity (e.g., “M” and “N”), a SVM classifier was integrated as a post-feature extraction stage, functioning as a refined decision layer, as proposed in [14].

In parallel, CNN-LSTM architecture was implemented for dynamic letters. This model receives a sequence of images (consecutive frames of the same gesture) as input and learns to capture temporal progression through recurrent layers. This approach is supported by works such as [17].

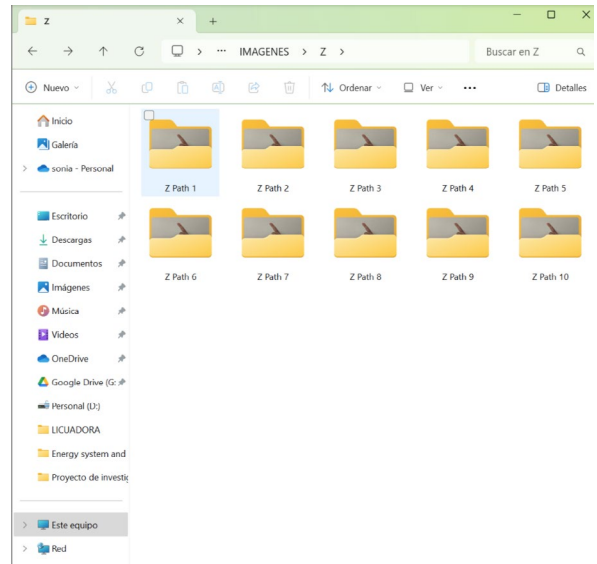


Fig. 3. Dataset of letter Z.

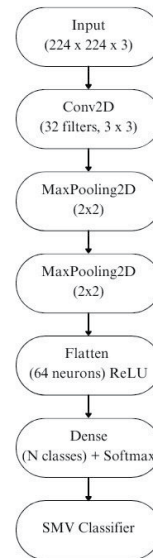


Fig. 4. Block diagram illustrating the operation of the model architecture.

D. Real-Time Integration and Execution

The system was implemented in Python using libraries such as OpenCV for video capture, TensorFlow/Keras for model training and deployment, and Pytsx3 for text-to-speech synthesis. During execution, the camera continuously detects the region of interest (ROI) corresponding to the user’s hands and processes each frame in real time. The model predicts the corresponding class and accumulates the recognized letters, as illustrated in Fig. 5. A temporal segmentation

algorithm based on pauses between gestures enables the differentiation of individual letters, words, or complete phrases. The resulting text is displayed on the screen using a simple graphical interface and can be automatically converted into speech, facilitating immediate communication between the user and their environment, as shown in Fig. 6.

E. Testing, Validation and Metrics

From a statistical perspective, the validity of the performance metrics is directly influenced by the size of the test sample. With the current configuration of 50 images per class (28 letters, for a total of 1400 images) and an 80/20 training-testing split, the test set comprised approximately 280 samples. This number allows estimating the overall accuracy with a confidence interval close to $\pm 6-7\%$ at 95% confidence, which is acceptable for an experimental prototype. However, the literature suggests that reducing the margin of error to $\pm 5\%$ requires at least 384 test samples [13], [14] Likewise, for per-class metrics such as precision or recall, it is recommended to have a sufficient number of validation samples per category to ensure the statistical reliability of per-class metrics such as precision and recall. With 50 images per class, the system obtains, on average, about ten test samples per letter, which limits the statistical stability of these metrics. Therefore, as future work, it is proposed to expand the dataset to 100-150 images per class or to apply stratified *k*-fold cross-validation in order to increase the reliability of results.

The prototype was also subjected to alpha and beta testing with volunteer users to evaluate performance under both controlled conditions and real-world environments. Accuracy, sensitivity, and the confusion matrix were used for static models, while for dynamic letters, the WER was used together with sequence-level precision and recall, following the guidelines proposed in DeepASLR [14]. Response time was measured by directly timing the interval between image acquisition and prediction output, to determine an average processing speed of 125 fps, which confirms the system's real-time viability, as shown in Fig. 7.

III. RESULTS

The prototype developed for the automatic translation of the PSL finger alphabet achieved an overall accuracy of 97% using a dataset of 50 images per class, together with a WER of 15% and an average response time of 1.8-2.0 s with a processing speed of 125 frames per second, thereby confirming its viability for real-time operation. Tests were conducted under controlled laboratory conditions, with stable lighting (300 lumens/250 lux), a uniform background, and minimal visual interference, thereby ensuring that the observed variations were attributable to the model's performance rather than to external factors.

The class-wise analysis showed that letters with distinctive features (L, T, and Y) achieved accuracies above 95%,

The image shows a MATLAB script for training a neural network. It includes steps for loading data, normalizing, splitting into training and testing sets, defining a neural network architecture with layers like 'conv_1', 'conv_2', 'conv_3', 'conv_4', 'conv_5', 'conv_6', 'conv_7', 'conv_8', 'conv_9', 'conv_10', 'conv_11', 'conv_12', 'conv_13', 'conv_14', 'conv_15', 'conv_16', 'conv_17', 'conv_18', 'conv_19', 'conv_20', 'conv_21', 'conv_22', 'conv_23', 'conv_24', 'conv_25', 'conv_26', 'conv_27', 'conv_28', 'conv_29', 'conv_30', 'conv_31', 'conv_32', 'conv_33', 'conv_34', 'conv_35', 'conv_36', 'conv_37', 'conv_38', 'conv_39', 'conv_40', 'conv_41', 'conv_42', 'conv_43', 'conv_44', 'conv_45', 'conv_46', 'conv_47', 'conv_48', 'conv_49', 'conv_50', 'conv_51', 'conv_52', 'conv_53', 'conv_54', 'conv_55', 'conv_56', 'conv_57', 'conv_58', 'conv_59', 'conv_60', 'conv_61', 'conv_62', 'conv_63', 'conv_64', 'conv_65', 'conv_66', 'conv_67', 'conv_68', 'conv_69', 'conv_70', 'conv_71', 'conv_72', 'conv_73', 'conv_74', 'conv_75', 'conv_76', 'conv_77', 'conv_78', 'conv_79', 'conv_80', 'conv_81', 'conv_82', 'conv_83', 'conv_84', 'conv_85', 'conv_86', 'conv_87', 'conv_88', 'conv_89', 'conv_90', 'conv_91', 'conv_92', 'conv_93', 'conv_94', 'conv_95', 'conv_96', 'conv_97', 'conv_98', 'conv_99', 'conv_100', 'conv_101', 'conv_102', 'conv_103', 'conv_104', 'conv_105', 'conv_106', 'conv_107', 'conv_108', 'conv_109', 'conv_110', 'conv_111', 'conv_112', 'conv_113', 'conv_114', 'conv_115', 'conv_116', 'conv_117', 'conv_118', 'conv_119', 'conv_120', 'conv_121', 'conv_122', 'conv_123', 'conv_124', 'conv_125', 'conv_126', 'conv_127', 'conv_128', 'conv_129', 'conv_130', 'conv_131', 'conv_132', 'conv_133', 'conv_134', 'conv_135', 'conv_136', 'conv_137', 'conv_138', 'conv_139', 'conv_140', 'conv_141', 'conv_142', 'conv_143', 'conv_144', 'conv_145', 'conv_146', 'conv_147', 'conv_148', 'conv_149', 'conv_150', 'conv_151', 'conv_152', 'conv_153', 'conv_154', 'conv_155', 'conv_156', 'conv_157', 'conv_158', 'conv_159', 'conv_160', 'conv_161', 'conv_162', 'conv_163', 'conv_164', 'conv_165', 'conv_166', 'conv_167', 'conv_168', 'conv_169', 'conv_170', 'conv_171', 'conv_172', 'conv_173', 'conv_174', 'conv_175', 'conv_176', 'conv_177', 'conv_178', 'conv_179', 'conv_180', 'conv_181', 'conv_182', 'conv_183', 'conv_184', 'conv_185', 'conv_186', 'conv_187', 'conv_188', 'conv_189', 'conv_190', 'conv_191', 'conv_192', 'conv_193', 'conv_194', 'conv_195', 'conv_196', 'conv_197', 'conv_198', 'conv_199', 'conv_200', 'conv_201', 'conv_202', 'conv_203', 'conv_204', 'conv_205', 'conv_206', 'conv_207', 'conv_208', 'conv_209', 'conv_210', 'conv_211', 'conv_212', 'conv_213', 'conv_214', 'conv_215', 'conv_216', 'conv_217', 'conv_218', 'conv_219', 'conv_220', 'conv_221', 'conv_222', 'conv_223', 'conv_224', 'conv_225', 'conv_226', 'conv_227', 'conv_228', 'conv_229', 'conv_230', 'conv_231', 'conv_232', 'conv_233', 'conv_234', 'conv_235', 'conv_236', 'conv_237', 'conv_238', 'conv_239', 'conv_240', 'conv_241', 'conv_242', 'conv_243', 'conv_244', 'conv_245', 'conv_246', 'conv_247', 'conv_248', 'conv_249', 'conv_250', 'conv_251', 'conv_252', 'conv_253', 'conv_254', 'conv_255', 'conv_256', 'conv_257', 'conv_258', 'conv_259', 'conv_260', 'conv_261', 'conv_262', 'conv_263', 'conv_264', 'conv_265', 'conv_266', 'conv_267', 'conv_268', 'conv_269', 'conv_270', 'conv_271', 'conv_272', 'conv_273', 'conv_274', 'conv_275', 'conv_276', 'conv_277', 'conv_278', 'conv_279', 'conv_280', 'conv_281', 'conv_282', 'conv_283', 'conv_284', 'conv_285', 'conv_286', 'conv_287', 'conv_288', 'conv_289', 'conv_290', 'conv_291', 'conv_292', 'conv_293', 'conv_294', 'conv_295', 'conv_296', 'conv_297', 'conv_298', 'conv_299', 'conv_300', 'conv_301', 'conv_302', 'conv_303', 'conv_304', 'conv_305', 'conv_306', 'conv_307', 'conv_308', 'conv_309', 'conv_310', 'conv_311', 'conv_312', 'conv_313', 'conv_314', 'conv_315', 'conv_316', 'conv_317', 'conv_318', 'conv_319', 'conv_320', 'conv_321', 'conv_322', 'conv_323', 'conv_324', 'conv_325', 'conv_326', 'conv_327', 'conv_328', 'conv_329', 'conv_330', 'conv_331', 'conv_332', 'conv_333', 'conv_334', 'conv_335', 'conv_336', 'conv_337', 'conv_338', 'conv_339', 'conv_340', 'conv_341', 'conv_342', 'conv_343', 'conv_344', 'conv_345', 'conv_346', 'conv_347', 'conv_348', 'conv_349', 'conv_350', 'conv_351', 'conv_352', 'conv_353', 'conv_354', 'conv_355', 'conv_356', 'conv_357', 'conv_358', 'conv_359', 'conv_360', 'conv_361', 'conv_362', 'conv_363', 'conv_364', 'conv_365', 'conv_366', 'conv_367', 'conv_368', 'conv_369', 'conv_370', 'conv_371', 'conv_372', 'conv_373', 'conv_374', 'conv_375', 'conv_376', 'conv_377', 'conv_378', 'conv_379', 'conv_380', 'conv_381', 'conv_382', 'conv_383', 'conv_384', 'conv_385', 'conv_386', 'conv_387', 'conv_388', 'conv_389', 'conv_390', 'conv_391', 'conv_392', 'conv_393', 'conv_394', 'conv_395', 'conv_396', 'conv_397', 'conv_398', 'conv_399', 'conv_400', 'conv_401', 'conv_402', 'conv_403', 'conv_404', 'conv_405', 'conv_406', 'conv_407', 'conv_408', 'conv_409', 'conv_410', 'conv_411', 'conv_412', 'conv_413', 'conv_414', 'conv_415', 'conv_416', 'conv_417', 'conv_418', 'conv_419', 'conv_420', 'conv_421', 'conv_422', 'conv_423', 'conv_424', 'conv_425', 'conv_426', 'conv_427', 'conv_428', 'conv_429', 'conv_430', 'conv_431', 'conv_432', 'conv_433', 'conv_434', 'conv_435', 'conv_436', 'conv_437', 'conv_438', 'conv_439', 'conv_440', 'conv_441', 'conv_442', 'conv_443', 'conv_444', 'conv_445', 'conv_446', 'conv_447', 'conv_448', 'conv_449', 'conv_450', 'conv_451', 'conv_452', 'conv_453', 'conv_454', 'conv_455', 'conv_456', 'conv_457', 'conv_458', 'conv_459', 'conv_460', 'conv_461', 'conv_462', 'conv_463', 'conv_464', 'conv_465', 'conv_466', 'conv_467', 'conv_468', 'conv_469', 'conv_470', 'conv_471', 'conv_472', 'conv_473', 'conv_474', 'conv_475', 'conv_476', 'conv_477', 'conv_478', 'conv_479', 'conv_480', 'conv_481', 'conv_482', 'conv_483', 'conv_484', 'conv_485', 'conv_486', 'conv_487', 'conv_488', 'conv_489', 'conv_490', 'conv_491', 'conv_492', 'conv_493', 'conv_494', 'conv_495', 'conv_496', 'conv_497', 'conv_498', 'conv_499', 'conv_500', 'conv_501', 'conv_502', 'conv_503', 'conv_504', 'conv_505', 'conv_506', 'conv_507', 'conv_508', 'conv_509', 'conv_510', 'conv_511', 'conv_512', 'conv_513', 'conv_514', 'conv_515', 'conv_516', 'conv_517', 'conv_518', 'conv_519', 'conv_520', 'conv_521', 'conv_522', 'conv_523', 'conv_524', 'conv_525', 'conv_526', 'conv_527', 'conv_528', 'conv_529', 'conv_530', 'conv_531', 'conv_532', 'conv_533', 'conv_534', 'conv_535', 'conv_536', 'conv_537', 'conv_538', 'conv_539', 'conv_540', 'conv_541', 'conv_542', 'conv_543', 'conv_544', 'conv_545', 'conv_546', 'conv_547', 'conv_548', 'conv_549', 'conv_550', 'conv_551', 'conv_552', 'conv_553', 'conv_554', 'conv_555', 'conv_556', 'conv_557', 'conv_558', 'conv_559', 'conv_560', 'conv_561', 'conv_562', 'conv_563', 'conv_564', 'conv_565', 'conv_566', 'conv_567', 'conv_568', 'conv_569', 'conv_570', 'conv_571', 'conv_572', 'conv_573', 'conv_574', 'conv_575', 'conv_576', 'conv_577', 'conv_578', 'conv_579', 'conv_580', 'conv_581', 'conv_582', 'conv_583', 'conv_584', 'conv_585', 'conv_586', 'conv_587', 'conv_588', 'conv_589', 'conv_590', 'conv_591', 'conv_592', 'conv_593', 'conv_594', 'conv_595', 'conv_596', 'conv_597', 'conv_598', 'conv_599', 'conv_600', 'conv_601', 'conv_602', 'conv_603', 'conv_604', 'conv_605', 'conv_606', 'conv_607', 'conv_608', 'conv_609', 'conv_610', 'conv_611', 'conv_612', 'conv_613', 'conv_614', 'conv_615', 'conv_616', 'conv_617', 'conv_618', 'conv_619', 'conv_620', 'conv_621', 'conv_622', 'conv_623', 'conv_624', 'conv_625', 'conv_626', 'conv_627', 'conv_628', 'conv_629', 'conv_630', 'conv_631', 'conv_632', 'conv_633', 'conv_634', 'conv_635', 'conv_636', 'conv_637', 'conv_638', 'conv_639', 'conv_640', 'conv_641', 'conv_642', 'conv_643', 'conv_644', 'conv_645', 'conv_646', 'conv_647', 'conv_648', 'conv_649', 'conv_650', 'conv_651', 'conv_652', 'conv_653', 'conv_654', 'conv_655', 'conv_656', 'conv_657', 'conv_658', 'conv_659', 'conv_660', 'conv_661', 'conv_662', 'conv_663', 'conv_664', 'conv_665', 'conv_666', 'conv_667', 'conv_668', 'conv_669', 'conv_670', 'conv_671', 'conv_672', 'conv_673', 'conv_674', 'conv_675', 'conv_676', 'conv_677', 'conv_678', 'conv_679', 'conv_680', 'conv_681', 'conv_682', 'conv_683', 'conv_684', 'conv_685', 'conv_686', 'conv_687', 'conv_688', 'conv_689', 'conv_690', 'conv_691', 'conv_692', 'conv_693', 'conv_694', 'conv_695', 'conv_696', 'conv_697', 'conv_698', 'conv_699', 'conv_700', 'conv_701', 'conv_702', 'conv_703', 'conv_704', 'conv_705', 'conv_706', 'conv_707', 'conv_708', 'conv_709', 'conv_710', 'conv_711', 'conv_712', 'conv_713', 'conv_714', 'conv_715', 'conv_716', 'conv_717', 'conv_718', 'conv_719', 'conv_720', 'conv_721', 'conv_722', 'conv_723', 'conv_724', 'conv_725', 'conv_726', 'conv_727', 'conv_728', 'conv_729', 'conv_730', 'conv_731', 'conv_732', 'conv_733', 'conv_734', 'conv_735', 'conv_736', 'conv_737', 'conv_738', 'conv_739', 'conv_740', 'conv_741', 'conv_742', 'conv_743', 'conv_744', 'conv_745', 'conv_746', 'conv_747', 'conv_748', 'conv_749', 'conv_750', 'conv_751', 'conv_752', 'conv_753', 'conv_754', 'conv_755', 'conv_756', 'conv_757', 'conv_758', 'conv_759', 'conv_760', 'conv_761', 'conv_762', 'conv_763', 'conv_764', 'conv_765', 'conv_766', 'conv_767', 'conv_768', 'conv_769', 'conv_770', 'conv_771', 'conv_772', 'conv_773', 'conv_774', 'conv_775', 'conv_776', 'conv_777', 'conv_778', 'conv_779', 'conv_780', 'conv_781', 'conv_782', 'conv_783', 'conv_784', 'conv_785', 'conv_786', 'conv_787', 'conv_788', 'conv_789', 'conv_790', 'conv_791', 'conv_792', 'conv_793', 'conv_794', 'conv_795', 'conv_796', 'conv_797', 'conv_798', 'conv_799', 'conv_800', 'conv_801', 'conv_802', 'conv_803', 'conv_804', 'conv_805', 'conv_806', 'conv_807', 'conv_808', 'conv_809', 'conv_810', 'conv_811', 'conv_812', 'conv_813', 'conv_814', 'conv_815', 'conv_816', 'conv_817', 'conv_818', 'conv_819', 'conv_820', 'conv_821', 'conv_822', 'conv_823', 'conv_824', 'conv_825', 'conv_826', 'conv_827', 'conv_828', 'conv_829', 'conv_830', 'conv_831', 'conv_832', 'conv_833', 'conv_834', 'conv_835', 'conv_836', 'conv_837', 'conv_838', 'conv_839', 'conv_840', 'conv_841', 'conv_842', 'conv_843', 'conv_844', 'conv_845', 'conv_846', 'conv_847', 'conv_848', 'conv_849', 'conv_850', 'conv_851', 'conv_852', 'conv_853', 'conv_854', 'conv_855', 'conv_856', 'conv_857', 'conv_858', 'conv_859', 'conv_860', 'conv_861', 'conv_862', 'conv_863', 'conv_864', 'conv_865', 'conv_866', 'conv_867', 'conv_868', 'conv_869', 'conv_870', 'conv_871', 'conv_872', 'conv_873', 'conv_874', 'conv_875', 'conv_876', 'conv_877', 'conv_878', 'conv_879', 'conv_880', 'conv_881', 'conv_882', 'conv_883', 'conv_884', 'conv_885', 'conv_886', 'conv_887', 'conv_888', 'conv_889', 'conv_890', 'conv_891', 'conv_892', 'conv_893', 'conv_894', 'conv_895', 'conv_896', 'conv_897', 'conv_898', 'conv_899', 'conv_900', 'conv_901', 'conv_902', 'conv_903', 'conv_904', 'conv_905', 'conv_906', 'conv_907', 'conv_908', 'conv_909', 'conv_910', 'conv_911', 'conv_912', 'conv_913', 'conv_914', 'conv_915', 'conv_916', 'conv_917', 'conv_918', 'conv_919', 'conv_920', 'conv_921', 'conv_922', 'conv_923', 'conv_924', 'conv_925', 'conv_926', 'conv_927', 'conv_928', 'conv_929', 'conv_930', 'conv_931', 'conv_932', 'conv_933', 'conv_934', 'conv_935', 'conv_936', 'conv_937', 'conv_938', 'conv_939', 'conv_940', 'conv_941', 'conv_942', 'conv_943', 'conv_944', 'conv_945', 'conv_946', 'conv_947', 'conv_948', 'conv_949', 'conv_950', 'conv_951', 'conv_952', 'conv_953', 'conv_954', 'conv_955', 'conv_956', 'conv_957', 'conv_958', 'conv_959', 'conv_960', 'conv_961', 'conv_962', 'conv_963', 'conv_964', 'conv_965', 'conv_966', 'conv_967', 'conv_968', 'conv_969', 'conv_970', 'conv_971', 'conv_972', 'conv_973', 'conv_974', 'conv_975', 'conv_976', 'conv_977', 'conv_978', 'conv_979', 'conv_980', 'conv_981', 'conv_982', 'conv_983', 'conv_984', 'conv_985', 'conv_986', 'conv_987', 'conv_988', 'conv_989', 'conv_990', 'conv_991', 'conv_992', 'conv_993', 'conv_994', 'conv_995', 'conv_996', 'conv_997', 'conv_998', 'conv_999', 'conv_1000', 'conv_1001', 'conv_1002', 'conv_1003', 'conv_1004', 'conv_1005', 'conv_1006', 'conv_1007', 'conv_1008', 'conv_1009', 'conv_1010', 'conv_1011', 'conv_1012', 'conv_1013', 'conv_1014', 'conv_1015', 'conv_1016', 'conv_1017', 'conv_1018', 'conv_1019', 'conv_1020', 'conv_1021', 'conv_1022', 'conv_1023', 'conv_1024', 'conv_1025', 'conv_1026', 'conv_1027', 'conv_1028', 'conv_1029', 'conv_1030', 'conv_1031', 'conv_1032', 'conv_1033', 'conv_1034', 'conv_1035', 'conv_1036', 'conv_1037', 'conv_1038', 'conv_1039', 'conv_1040', 'conv_1041', 'conv_1042', 'conv_1043', 'conv_1044', 'conv_1045', 'conv_1046', 'conv_1047', 'conv_1048', 'conv_1049', 'conv_1050', 'conv_1051', 'conv_1052', 'conv_1053', 'conv_1054', 'conv_1055', 'conv_1056', 'conv_1057', 'conv_1058', 'conv_1059', 'conv_1060', 'conv_1061', 'conv_1062', 'conv_1063', 'conv_1064', 'conv_1065', 'conv_1066', 'conv_1067', 'conv_1068', 'conv_1069', 'conv_1070', 'conv_1071', 'conv_1072', 'conv_1073', 'conv_1074', 'conv_1075', 'conv_1076', 'conv_1077', 'conv_1078', 'conv_1079', 'conv_1080', 'conv_1081', 'conv_1082', 'conv_1083', 'conv_1084', 'conv_1085', 'conv_1086', 'conv_1087', 'conv_1088', 'conv_1089', 'conv_1090', 'conv_1091', 'conv_1092', 'conv_1093', 'conv_1094', 'conv_1095', 'conv_1096', 'conv_1097', 'conv_1098', 'conv_1099', 'conv_1100', 'conv_1101', 'conv_1102', 'conv_1103', 'conv_1104', 'conv_1105', 'conv_1106', 'conv_1107', 'conv_1108', 'conv_1109', 'conv_1110', 'conv_1111', 'conv_1112', 'conv_1113', 'conv_1114', 'conv_1115', 'conv_1116', 'conv_1117', 'conv_1118', 'conv_1119', 'conv_1120', 'conv_1121', 'conv_1122', 'conv_1123', 'conv_1124', 'conv_1125', 'conv_1126', 'conv_1127', 'conv_1128', 'conv_1129', 'conv_1130', 'conv_1131', 'conv_1132', 'conv_1133', 'conv_1134', 'conv_1135', 'conv_1136', 'conv_1137', 'conv_1138', 'conv_1139', 'conv_1140', 'conv_1141', 'conv_1142', 'conv_1143', 'conv_1144', 'conv_1145', 'conv_1146', 'conv_1147', 'conv_1148', 'conv_1149', 'conv_1150', 'conv_1151', 'conv_1152', 'conv_1153', 'conv_1154', 'conv_1155', 'conv_1156', 'conv_1157', 'conv_1158', 'conv_1159', 'conv_1160', 'conv_1161', 'conv_1162', 'conv_1163', 'conv_1164', 'conv_1165', 'conv_1166', 'conv_1167', 'conv_1168', 'conv_1169', 'conv_1170', 'conv_1171', 'conv_1172', 'conv_1173', 'conv_1174', 'conv_1175', 'conv_1176', 'conv_1177', 'conv_1178', 'conv_1179', 'conv_1180', 'conv_1181', 'conv_1182', 'conv_1183', 'conv_1184', 'conv_1185', 'conv_1186', 'conv_1187', 'conv_1188', 'conv_1189', 'conv_1190', 'conv_1191', 'conv_1192', 'conv_1193', 'conv_1194', 'conv_1195', 'conv_1196', 'conv_1197', 'conv_1198', 'conv_1199', 'conv_1200', 'conv_1201', 'conv_1202', 'conv_1203', 'conv_1204', 'conv_1205', 'conv_1206', 'conv_1207', 'conv_1208', 'conv_1209', 'conv_1210', 'conv_1211', 'conv_1212', 'conv_1213', 'conv_1214', 'conv_1215', 'conv_1216', 'conv_1217', 'conv_1218', 'conv_1219', 'conv_1220', 'conv_1221', 'conv_1222', 'conv_1223', 'conv_1224', 'conv_1225', 'conv_1226', 'conv_1227', 'conv_1228', 'conv_1229', 'conv_1230', 'conv_1231', 'conv_1232', 'conv_1233', 'conv_1234', 'conv_1235', 'conv_1236', 'conv_1237', 'conv_1238', 'conv_1239', 'conv_1240', 'conv_1241', 'conv_1242', 'conv_1243', 'conv_1244', 'conv_1245', 'conv_1246', 'conv_1247', 'conv_1248', 'conv_1249', 'conv_1250', 'conv_1251', 'conv_1252', 'conv_1253', 'conv_1254', 'conv_1255', 'conv_1256', 'conv_1257', 'conv_1258', 'conv_1259', 'conv_1260', 'conv_1261', 'conv_1262', 'conv_1263', 'conv_1264', 'conv_1265', 'conv_1266', 'conv_1267', 'conv_1268', 'conv_1269', 'conv_1270', 'conv_1271', 'conv_1272', 'conv_1273', 'conv_1274', 'conv_1275', 'conv_1276', 'conv_1277', 'conv_1278', 'conv_1279', 'conv_1280', 'conv_1281', 'conv_1282', 'conv_1283', 'conv_1284', 'conv_1285', 'conv_1286', 'conv_1287', 'conv_1288', 'conv_1289', 'conv_1290', 'conv_1291', 'conv_1292', 'conv_1293', 'conv_1294', 'conv_1295', 'conv_1296', 'conv_1297', 'conv_1298', 'conv_1299', 'conv_1300', 'conv_1301', 'conv_1302', 'conv_1303', 'conv_1304', 'conv_1305', 'conv_1306', 'conv_1307', 'conv_1308', 'conv_1309', 'conv_1310', 'conv_1311', 'conv_1312', 'conv_1313', 'conv_1314', 'conv_1315', 'conv_1316', 'conv_1317', 'conv_1318', 'conv_1319', 'conv_1320', 'conv_1321', 'conv_1322', 'conv_1323', 'conv_1324', 'conv_1325', 'conv_1326', 'conv_1327', 'conv_1328', 'conv_1329', 'conv_1330', 'conv_1331', 'conv_1332', 'conv_1333', 'conv_1334', 'conv_1335', 'conv_1336', 'conv_1337', 'conv_1338', 'conv_1339', 'conv_1340', 'conv_1341', 'conv_1342', 'conv_1343', 'conv_1344', 'conv_1345', 'conv_1346', 'conv_1347', 'conv_1348', 'conv_1349', 'conv_1350', 'conv_1351', 'conv_1352', 'conv_1353', 'conv_1354', 'conv_1355', 'conv_1356', 'conv_1357', 'conv_1358', 'conv_1359', 'conv_1360', 'conv_1361', 'conv_1362', 'conv_1363', 'conv_1364', 'conv_1365', 'conv_1366', 'conv_1367', 'conv_1368', 'conv_1369', 'conv_1370', 'conv_1371', 'conv_1372', 'conv_1373', 'conv_1374', 'conv_1375', 'conv_1376', 'conv_1377', 'conv_1378', 'conv_1379', 'conv_1380', 'conv_1381', 'conv_1382', 'conv_1383', 'conv_1384', 'conv_1385', 'conv_1386', 'conv_1387', 'conv_1388', 'conv_1389', 'conv_1390', 'conv_1391', 'conv_1392', 'conv_1393', 'conv_1394', 'conv_1395', 'conv_1396', 'conv_1397', 'conv_1398', 'conv_1399', 'conv_1400', 'conv_1401', 'conv_1402', 'conv_1403', 'conv_1404', 'conv_1405', 'conv_1406', 'conv_1407', 'conv_1408', 'conv_1409', 'conv_1410', 'conv_1411', 'conv_1412', 'conv_1413', 'conv_1414', 'conv_1415', 'conv_1416', 'conv_1417', 'conv_1418', 'conv_1419', 'conv_1420', 'conv_1421', 'conv_1422', 'conv_1423', 'conv_1424', 'conv_1425', 'conv_1426', 'conv_1427', 'conv_1428', 'conv_1429', 'conv_1430', 'conv_1431', 'conv_1432', 'conv_1433', 'conv_1434', 'conv_1435', 'conv_1436', 'conv_1437', 'conv_1438', 'conv_1439', 'conv_1440', 'conv_1441', 'conv_1442', 'conv_1443', 'conv_1444', 'conv_1445', 'conv_1446', 'conv_1447', 'conv_1448', 'conv_1449', 'conv_1450', 'conv_1451', 'conv_1452', 'conv_1453', 'conv_1454', 'conv_1455', 'conv_1456', 'conv_1457', 'conv_1458', 'conv_1459', 'conv_1460', 'conv_1461', 'conv_1462', 'conv_1463', 'conv_1464', 'conv_1465', 'conv_1466', 'conv_1467', 'conv_1468', 'conv_1469', 'conv_1470', 'conv_1471', 'conv_1472', 'conv_1473', 'conv_1474', 'conv_1475', 'conv_1476', 'conv_1477', 'conv_1478', 'conv_1479', 'conv_1480', 'conv_1481', 'conv_1482', 'conv_1483', 'conv_1484', 'conv_1485', 'conv_1486', 'conv_1487', 'conv_1488', 'conv_1489', 'conv_1490', 'conv_1491', 'conv_1492', 'conv_1493', 'conv_1494', 'conv_1495', 'conv_1496', 'conv_1497', 'conv_1498', 'conv_1499', 'conv_1500', 'conv_1501', 'conv_1502', 'conv_1503', 'conv_1504', 'conv_1505', 'conv_1506', 'conv_1507', 'conv_1508', 'conv_1509', 'conv_1510', 'conv_1511', 'conv_1512', 'conv_1513', 'conv_1514', 'conv_1515', 'conv_1516', 'conv_1517', 'conv_1518', 'conv_1519', 'conv_1520', 'conv_1521', 'conv_1522', 'conv_1523', 'conv_1524', 'conv_1525', 'conv_1526', 'conv_1527', 'conv_1528', 'conv_1529', 'conv_1530', 'conv_1531', 'conv_1532', 'conv_1533', 'conv_1534', 'conv_1535', 'conv_1536', 'conv_1537', 'conv_1538', 'conv_1539', 'conv_1540', 'conv_1541', 'conv_1542', 'conv_1543', 'conv_1544', 'conv_1545', 'conv_1546', 'conv_1547', 'conv_1548', 'conv_1549', 'conv_1550', 'conv_1551', 'conv_1552', 'conv_1553', 'conv_1554', 'conv_1555', 'conv_1556', 'conv_1557', 'conv_1558', 'conv_1559', '

However, the statistical reliability at the class level was limited by the number of available samples (approximately ten test images per letter). Therefore, future work includes expanding the dataset to 100–150 images per class, incorporating additional metrics such as recall and F1-score, validating the system in real educational scenarios with end users, and evaluating its performance under more variable lighting conditions to ensure robustness in uncontrolled contexts. The use of five images per class was intended to represent a minimal scenario for observing overfitting and classification instability under limited data conditions, whereas the configuration with 15 images per class provided a mid-level setting that increased sample variability and improved prediction stability. Finally, using 50 images per class provided a more representative dataset, enabling the system to achieve competitive accuracy levels consistent with those reported in the literature for CNN-based sign recognition [13]–[15]. The corresponding test results are summarized in Table I, whereas the performance curve is shown in Fig. 8.

A. Tests With Five Images per Class

In an initial exploratory phase, the model was trained using five images per class of the finger alphabet. The results showed high instability in predictions, especially for visually similar letters. Overall accuracy was low, and the system exhibited limited generalization to slight variations in posture, lighting, and background.

B. Tests With 15 Images per Class

In the second iteration, the dataset was increased to 15 images per class. This expansion resulted in substantial improvements in model learning, with a significant increase in overall accuracy and reduced sensitivity to external conditions. However, some misclassifications persisted among classes with similar morphologies.

C. Tests With 50 Images per Class

In the third phase, the model was trained using 50 images per class. Under this configuration, the system achieved its best overall performance. High recognition accuracy, a decreased error rate, and a more stable response were observed under controlled lighting conditions (300 lumens/250 lux). Furthermore, the average processing speed remained at 125 fps.

Performance varied slightly depending on the letter. The most difficult classes to identify were those with similar handforms, such as “M” and “N,” or “U” and “V,” which exhibited higher error rates. In contrast, letters with distinctive features such as “L,” “T,” or “Y” achieved accuracies above 95%.

Under the configuration with 50 images per class, the model achieved accuracies above 90% for 21 of the 26 evaluated letters and maintained a minimum accuracy of 85% for

TABLE I
PERFORMANCE COMPARISON OF THE THREE MAIN TESTS

	Images by class	Accuracy (%)	Loss	Response time (FPS)	Observations
1	5	63.4%	1.2	5.4	Low learning with high errors
2	15	79.6%	0.8	3.2	Better learning and control
3	50	97%	0.2	1.8	High model accuracy and stability

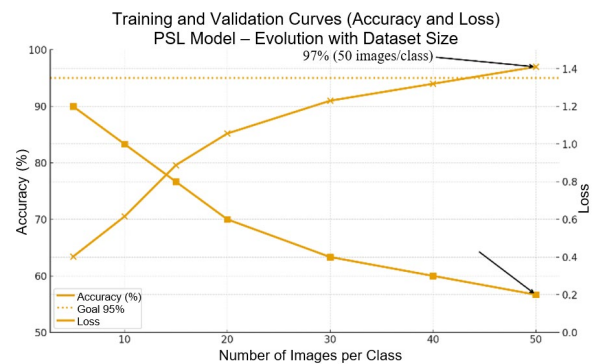


Fig. 8. Training and validation curve of the PSL model.

the remaining classes, demonstrating strong generalization capability under optimal environmental conditions.

In addition to overall accuracy, other relevant metrics were calculated to further assess the system’s overall performance:

- WER: 15% with 50 images/class
- Average accuracy of 97%
- Average processing speed: 125 fps

These metrics confirm that the system is functional in real time and suitable for educational environments provided that controlled conditions—such as a uniform background, stable lighting, and minimal visual interference—are maintained.

IV. DISCUSSION

The results obtained from the experimental tests enable a comparison of the system’s performance with respect to the proposed objectives and existing technological solutions. The initial objective of this prototype was to achieve an accuracy between 95% and 99% in recognizing letters of the PSL finger alphabet while maintaining a real-time

response time of less than 2 s, thereby allowing the system to be considered viable for inclusive educational settings.

Despite the promising results, this project presents several limitations. In particular, the dataset size was limited to 50 images per class, which restricts the statistical reliability of per-class performance metrics. All experiments were conducted under controlled conditions of lighting and background conditions, which limits the system's robustness in real-world environments. In addition, the prototype focused primarily on static letters of the PSL alphabet and a limited set of dynamic gestures, without addressing the recognition of complete words or sentences. Finally, validation with end users in real classroom settings has not yet been conducted, which is necessary to assess usability and scalability.

A. Comparison With Accessibility Mobile Applications

National platforms such as [2] and [3] provide basic accessibility features through manual selection or voice-to-text conversion. However, both are limited to static, predefined vocabularies without visual gesture recognition. In contrast, the proposed prototype achieved an average accuracy of 97% in real-time operation at 125 fps, providing automatic visual-to-speech translation of live gestures and overcoming the static nature and limited scalability of these applications.

B. Comparison With Systems Based on Physical Sensors

Wearable-device-based systems, such as those reported in [5] and [6], achieve accuracy ranges under standard conditions and up to 96% when specialized sensors are used. Although technically effective, these solutions are intrusive, costly, and require frequent calibration, which makes them impractical for routine classroom use. In contrast, the proposed model relies solely on a standard camera and achieves a real-time correct prediction rate of 91.3% using 50 images per class, thus eliminating the discomfort and cost associated with gloves and sensors.

C. Comparison With International CNN Recognition Models

International proposals such as [4], [13], and [15] reported accuracies in the range of 93–98% using large ASL or ISL datasets under controlled conditions. Similarly, advanced hybrid models such as those presented in [11], [14], [16], and [17] demonstrate robust performance; however, they typically depend on large-scale datasets, cloud-based infrastructure, and application contexts different from Peru. By contrast, the proposed prototype achieved an accuracy of 97% and a WER of 15% using a locally constructed Peruvian dataset, thereby ensuring cultural relevance while operating on low-cost hardware.

D. Comparison With Patented Solutions

Patented devices such as those described in [19]–[21] focus on glove-based or camera-assisted translation of foreign sign languages and report accuracies above 90%.

However, these solutions rely on specialized hardware and do not address the specific requirements of the Peruvian context. This prototype differs by integrating CNN and SVM models trained on a PSL dataset, providing a scalable, camera-only solution tailored to local conditions.

V. CONCLUSIONS

In summary, the prototype meets the established objectives by exceeding 95% accuracy, achieving an average accuracy of 97%, and operating in real time at a processing speed of 125 frames per second. This performance is competitive with international models and offers a practical advantage over intrusive solutions such as sensor-enhanced gloves. Furthermore, the objective of developing a non-invasive and low-cost solution was achieved, with an approximate 30% savings compared to sensor-based systems. Its contextualization within the Peruvian environment was guaranteed by using the Bilingual Dictionary of PSL as a reference.

Its accessible and educational nature reinforces its potential as an inclusive tool. However, it is acknowledged that the results are based on tests under controlled conditions and with a dataset limited to 50 images per class. Therefore, future work will focus on validating the system in real-life classroom settings, using a larger sample size and incorporating additional evaluation metrics such as recall and F1 score.

REFERENCES

- [1] Ombudsman's Office, "Situation of inclusive education in Peru," Ombudsman's Office, Lima, Peru, 2022.
- [2] Pontificia Universidad Católica del Perú, "Diccionario LSP al Español [Virtual Bilingual Dictionary of Peruvian Sign Language]," 2022. [Online]. Available: <https://diccionariolsp.pucp.edu.pe/search-by-text>
- [3] *Yapaykuy*. (2023). App Store. Lima, Peru. [Mobile app]. Available: <https://apps.apple.com/pe/app/yapaykuy/id6478118708>
- [4] K. Abbas, R. Kumar, and P. Jain, "Sign language recognition using CNN," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 5, pp. 1941–1953, May 2021.
- [5] J. DelPreto, J. Hughes, M. D'Aria, M. de Fazio, and D. Rus, "A wearable smart glove and its application of pose and gesture detection to sign language classification," *IEEE Robot. Autom. Lett.*, vol. 7, no. 4, pp. 10589–10596, 2022, doi: 10.1109/LRA.2022.3191232.
- [6] A. Ji, Y. Wang, X. Miao, T. Fan, B. Ru, L. Liu, R. Nie, and S. Qiu, "Dataglove for sign language recognition of people with hearing and speech impairment via wearable inertial sensors," *Sensors*, vol. 23, no. 15, Art. no. 6693, 2023, doi: 10.3390/s23156693

- [7] A. Radzi, S. Salim, M. H. Abd Wahab, M. M. Abdul Jamil, and T. C. Phing, "Development of a wearable sensor glove for real-time sign language translation." *Ann. Emerg. Technol. Comput.*, vol. 7, no. 5, pp. 25–38. doi: 10.33166/AETiC.2023.05.003
- [8] World Health Organization, *World Report on Hearing*. Geneva, Switzerland: WHO, 2021. [Online]. Available: <https://www.who.int/publications/i/item/world-report-on-hearing>
- [9] H. Knoors, and M. Marschark, "Language planning for the 21st century: Revisiting bilingual language policy for deaf children," *J. Deaf Stud. Deaf Educ.*, vol. 17, no. 3, pp. 291–305, 2012, doi: 10.1093/deafed/ens018.
- [10] C. Tannenbaum-Baruchi, B. Ahad Ha'am, and R. Harari, "Accessibility and support in higher education: A case study of deaf and hard of hearing students' perceptions," *High. Educ.*, 2025, doi: 10.1007/s10734-025-01586-x.
- [11] H. Wang, et al., "Enhanced open biomass burning detection: The BrantNet approach using UAV aerial imagery and deep learning for environmental protection and health preservation," *Ecol. Indic.*, vol. 154, Art. no. 110788, 2023, doi: 10.1016/j.ecolind.2023.110788.
- [12] L. Findlater et al., "Design preferences and barriers for deaf and hard of hearing users of mobile technology," in *Proc. ASSETS 2019*, Pittsburgh, PA, USA, Oct. 28–30 2019.
- [13] A. Gangal, A. Kuppahally, and M. Ravindran, *Sign Language Recognition with Convolutional Neural Networks*, Stanford CS231n Project Report, 2024. [Online]. Available: <https://cs231n.stanford.edu/2024/papers/sign-language-recognition-with-convolutional-neural-networks.pdf>
- [14] A. Kasapbaşı, A. E. A. Elbushra, O. Al-Hardanee, and A. Yilmaz, "DeepASLR: A CNN-based human-computer interface for American Sign Language recognition for hearing-impaired individuals," *Comput. Methods Programs Biomed. Update*, vol. 2, p. 100048, Jan. 2022, doi: 10.1016/j.cmpbup.2021.100048.
- [15] S. Subburaj, S. Murugavalli, and B. Muthusenthil, "Sign language video to text conversion via optimised LSTM with improved motion estimation," *J. Exp. Theor. Artif. Intell.*, vol. 37, no. 1, pp. 163–182, 2024, doi: 10.1080/0952813X.2024.2380991.
- [16] Y. Wang, H. Jiang, Y. Sun, and L. Xu, "A static sign language recognition method enhanced with self-attention mechanisms," *Sensors*, vol. 24, no. 21, p. 6921, 2024, doi: 10.3390/s24216921.
- [17] A. Elhagry and R. G. Elrayes, "Egyptian Sign Language recognition using CNN and LSTM," *arXiv*, Jul. 28, 2021. [Online]. Available: <https://arxiv.org/abs/2107.13647>
- [18] S. M. S. Swetha, M. S. Muneshwara, V. Venkatesh, S. M. S. Kudremane, R. S. Nair, and V. H. Vismitha, "Developing a translation model to bridge ASL and Indian regional languages using neural networks," in *Proc. 2025 Int. Conf. Comput. Sustain. Intell. Future (COMP-SIF)*, Bangalore, India, Mar. 21–22, 2025, pp. 1–6.
- [19] Indian Patent Office, *A sign language translator glove*, Patent application IN201841026260 A, Jul. 20, 2018. [Online]. Available: https://patentscope.wipo.int/search/es/detail.jsf?docId=IN224134191&_cid=P20-MFKLFW-88048-1
- [20] Indian Patent Office, *Sign language translator for deaf and speech impaired*, Patent application IN202321041213 A, 2023. [Online]. Available: https://patentscope.wipo.int/search/es/detail.jsf?docId=IN414860224&_cid=P20-MFKLWX-89227-1
- [21] Korean Intellectual Property Office, *Sign language translator and method thereof, particularly for translating the video of a finger language into a voice or text*, Patent KR1020100026701 A, 2010. [Online]. Available: https://patentscope.wipo.int/search/es/detail.jsf?docId=KR5842487&_cid=P20-MFKLHM-90082-1
- [22] J. A. Czechowicz et al., "Hearing impairment and poverty: The epidemiology of ear disease in Peruvian schoolchildren," *Otolaryngol. Head Neck Surg.*, vol. 142, no. 2, pp. 272–277, 2010. doi: 10.1016/j.otohns.2009.10.040
- [23] B. Alsharif, A. S. Altaher, A. Altaher, M. Liyas, and E. Alalwany, "Deep learning technology to recognize American Sign Language alphabet," *Sensors*, vol. 23, no. 19, p. 7970, 2023. doi: 10.3390/s23197970
- [24] D. Kothadiya, C. Bhatt, K. Sapariya, K. Patel, A.-B. Gil-González, and J. M. Corchado, "Deepsign: Sign language detection and recognition using deep learning," *Electronics*, vol. 11, no. 11, p. 1780, 2022. doi: 10.3390/electronics11111780
- [25] K. Assaleh and M. Al-Rousan, "Recognition of Arabic sign language alphabet using polynomial classifiers," *EURASIP Journal on Advances in Signal Processing*, vol. 2005, Art. no. 507614, 2005, doi: 10.1155/ASP.2005.2136