International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Indian Sign Language Alphabet Recognition Using Hybrid ML

Aditya Singh[1], Aryan Kumar[2], Harshit Singh[3], Priyanshu Chaudhary[4]

[1,2,3,4] Dept. of Information Technology, RKGIT Ghaziabad, India

sengar.aditya2@gmail.com[1], *choudharyaryan3011@gmail.com*[2], *singhharshit3132@gmail.com*[3], chaudharypriyanshu481@gmail.com[4]

ABSTRACT

You know what? Let's talk about a real bummer our society - the struggle that folks with disabilities face when trying to open up to everyone else. It's a tough spot. But hey, people with disabilities have come up with an awesome way to bridge that gap by using sign (gesture) languages. Pretty cool, right?

Now, picture this: there's this rad project in the works that aims to create a model that can take those hand gestures and turn them into both text and sound fancy machine learning stuff. It's all making sure folks with disabilities feel heard and understood like everyone else. This project is putting in the work using Convolution Neural Networks based on the Inception V3 deep learning model - yeah, it's super high-tech.

But why is this even needed? Well, let's get real for a sec - communication is like air for us humans; we need it to breathe! But for some of us out there dealing with hearing loss or speech issues, well, they're stuck without this lifeline. Imagine being unable to hear or speak – it's tough stuff. And when you throw deafness into the mix from childhood? Learning languages gets really tricky.

Sign language is like magic for these communities! It helps individuals who can't hear or talk express themselves through gestures made up of different hand shapes and motions - kinda like having secret codes just for them!

Think about it — there are over 460 million people globally facing hearing loss challenges. Crazy numbers! And unfortunately, not many folks know sign language because well...it's not exactly global news material! The Deaf community often struggles with written communication too since writing spoken languages isn't always easy peasy.

Did you know less than one percent of people worldwide have hearing or speech impairments according to the UN Statistics Division? That's massive odds stacked against them! Not enough skilled sign language interpreters out there either - it's like finding good Wi-Fi in a remote village!

The Indian scene isn't much better - imagine longing for connection but living worlds apart from understanding due to zero resources around! And don't even get me started on skilled teachers shortages hitting schools for the deaf and dumb hard...

But hey, here comes researchers playing superheroes trying out new ways—fancy tech tools galore—to break down barriers between those who understand sign language and those who don't. Think sensors, cameras, AI tech – basically modern-day wizardry at work!

Let's dig deeper into sign language vibes — adding emotion through facial expressions…that sacred body movement plus hand shapes combo; these are vital cues used by our peeps needing Sign Language daily!

Alrighty – three types of sign languages are rocking their world:

Fingerspelling one letter at a time

Used everyday words

Non-manual features like facial expressions and full-on body moves

A whole world right under our noses!

Sign languages are as diverse as the people who use them. From intricate hand shapes to animated facial expressions, each sign language has its own story to tell. But when it comes to popularity, American Sign Language (ASL) steals the show.

ASL isn't just another language; it's a whole new world of. Your hands become the storytellers, and your expressions bring those stories to life in English can never match. Can you imagine painting pictures with your face and weaving tales with your fingers? ASL makes it all possible!

For deaf individuals across North America, AS is more than just a way communicate - it's a lifeline. But guess what? Even hearing folks are jumping on the ASL train because let's be real, it's pretty darn awesome!

So where did this gem of a language come from? Some say it has roots tracing back over two centuries, blending different sign languages with French Sign Language along the way. The next time you flash a peace sign or greet someone using ASL, remember that each gesture carries centuries of history within it. Now that's something truly mind-blowing!

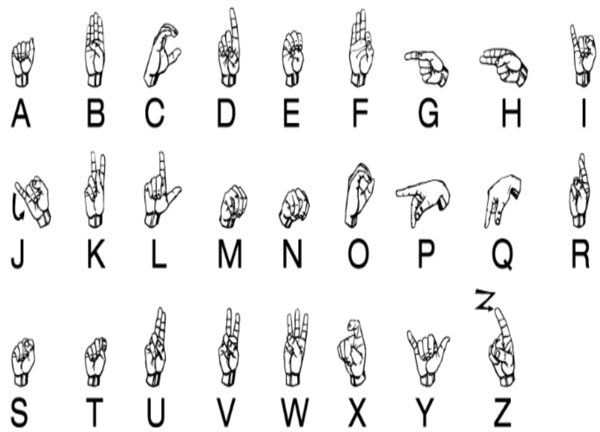

Figure 1

Sign language is just like any other language - it changes and grows over time. You'd be surprised to learn that many schools and universities in the US actually consider it a legit "foreign" language for their courses.

Indian Sign Language (ISL)

Now, let's talk about Indian Sign Language (ISL). It's basically THE go-to sign language across India. In some of the big cities, they even see it as the mother tongue because almost everyone uses it! Picture this: ISL is a mix of different sign languages that have been developing for ages. It's everywhere!

Here's the cool part - India's sign language? Super scientific. Yup, it has its very own grammar rules. Imagine that! So next time you think about sign languages, remember – there’s more to them than meets the eye!

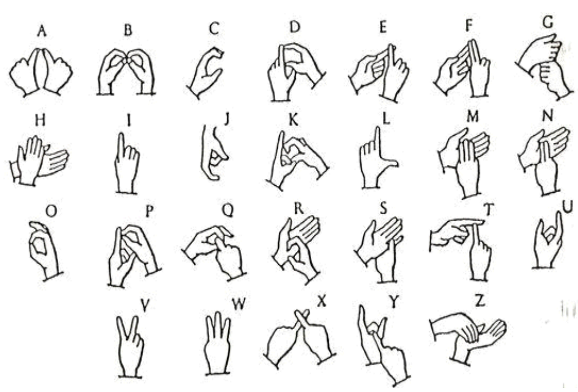

Figure 2

There's a split in ISL - manual and non-manual stuff. Check out Figure 2 for a visual on that.Now, let's talk about:

i. Manual - where you can it with one hand both.
ii. ii) Non-manual - facial expressions can come into play.

SYSTEM OVERVIEW

So, what's the big idea? To build a system smart enough to spot and classify sign language moves from these collected datasets (take a peek at Figure 3). The plan is to go with the Inception v3 model. It's like the big shot in image recognition, nailing an accuracy of almost 99% for American Sign Language. Cool beans, right?

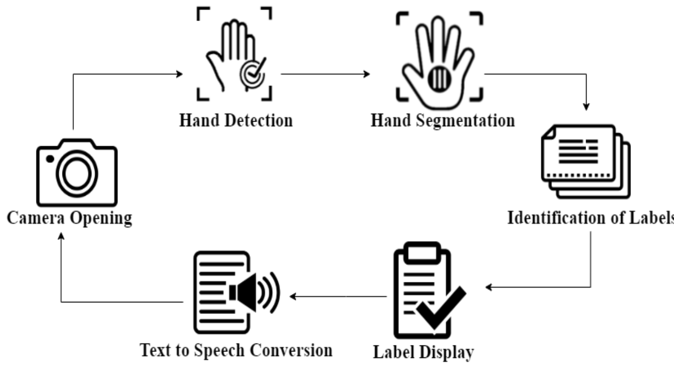

Figure 3

LITERARY SURVEY

Ever wondered how to bridge the gap between sign language and spoken/written languages in today's digital age? Communication is key, but for those with hearing and speaking impairments, it can be a real struggle. People steer clear of learning sign language, leaving many individuals in silence. But fear not, innovative minds have been at work trying to crack this code!

Various methods and algorithms have been tossed into the ring to tackle sign recognition in videos and images. We've seen quite a few proposals floating around out there. The lies in transforming those beautiful gestures of sign language into good old text and speech that we all understand.

Now, picture this - there are three stages involved in this whole shebang: input, processing, and output. Sounds like a complex equation doesn't it? Well buckle up because we got two different approaches here- one playing around with image processing & machine learning while the other dives deep into sensors & microcontrollers.

Let's break it down - the image-based method is light on your pockets and easy to lug around wherever you go. But hold up! Dealing with varying light settings can throw a spanner in the works sometimes. On the flip side, our sensor-powered pal comes dressed as a glove ready to rock n' roll directly translating hand movements into digital signals without all those pesky conversions.

While both sides have their strong suits (pun intended), they also pack some weaknesses along for the ride. Cost could be an issue for one team whereas the other might find itself tied down by wearing gloves 24/7 like it's going out of style.

But hey - silver lining alert! These solutions do manage to hit close to bullseye when it comes to accuracy rates soaring above 98%. So yes folks, communication may just be getting a little easier for our friends facing these challenges thanks to these groundbreaking methods on the table!

RELATED WORK

Deep Learning Model

Deep learning has been killing it in the computer vision world lately. The method is all about diving deep into rich feature extraction, killer modeling skills, and making training look easy. But what exactly is a deep learning model? Well, it's basically a brain-inspired algorithm that uses artificial neural networks to work its magic.

Think of deep learning as the secret behind self-driving cars being able to tell the difference between a stop sign and a lamppost. It's also the reason why your voice-controlled gadgets like phones and smart speakers actually understand what you're saying .Adays, deep learning is getting all the attention - and rightfully so! It's breaking barriers and achieving things we once thought impossible. Imagine teaching a computer to categorize stuff just by looking at images, text, or sound. That's the power of deep learning right there.

When it comes to understanding language, deep learning is poised to shake things up big time. Picture this: systems using recurrent neural networks (RNNs) mastering words or even whole texts by focusing on one piece at a time. And get this - deep learning models can hit accuracy levels that sometimes even outshine us humans!

But how do they pull off these mind-blowing feats? Well, it all boils down to loads of labeled data and some seriously intricate neural network setups. It's like giving these models an insane amount of knowledge until they become total pros at what they do.

Convolutional Neural Network (CNN)

Convolutional networks are used in the majority of cutting-edge computer vision solutions for a variety of issues [22]. Convolutional neural networks are among the most widely used deep neural network models that we have used (CNN). Convolutional neural networks were modelled after the neocognitron46, which had a similar architecture but no backpropagation, an end-to-end supervised learning process. To detect phonemes and short sentences, a simple 1D CNN called a time-delay neural net was employed [29][30]. Convolution is the mathematical linear operation between matrices that gives it its name. The pooling layer, fully-connected layer, non-linearity layer, and convolutional layer are the layers that make up CNN. The pooling and non-linearity layers are parameter-free, while the convolutional and fully-connected layers are. The assignment may involve localising specific objects within a scene, classifying pictures, or segmenting numerous classes [26]. CNNs are commonly used to tackle challenging image-driven pattern recognition tasks, and they facilitate the introduction of ANNs due to their precise yet straightforward architecture [14].

Transfer Learning

Since transfer learning requires substantially less data to improve performance and significantly reduces training time, its popularity has skyrocketed. By transferring knowledge from different but related source domains, transfer learning aims to enhance target learners' performance on target domains [27]. Transfer learning can be very helpful in knowledge engineering in a variety of situations. One example is the classification of web documents [33] [34].

Lately, transfer learning methodologies have been effectively implemented in numerous real-world settings. Transfer learning techniques have been proposed by Raina et al. [35] and Dai et al. [36] [37] for learning text data across domains. We utilised the Inception v3 model for our project; it incorporates transfer learning, which requires a significant amount of processing power to train the model on millions of photos—a feat that is challenging to accomplish from scratch [15].

Our strategy involved employing a model that had previously been trained on photos with a lot of processing capacity. We then trained the model according to our needs and goals by giving it specific image datasets, and the outcomes were outstanding.

D. Model on Inception

There are two types of inception models:

- Inception V1: Overfitting occurred when a model contained very deep layers of convolutions of the data happened. The idea V1 model uses many filters of different sizes on the same level to prevent this. As a result, our inception models have parallel layers rather than deep ones, which makes our model bigger rather than deeper.
- Inception V3

When it comes to object recognition, the Inception-v3 model performs better than GoogleNet (Inception-v1) [38]. Three parts make up the Inception-v3 model: a classifier, an improved Inception module, and a basic convolutional block. A 48-layer deep convolutional neural network is called Inception-v3. The ImageNet database contains a pre-trained version of the network that has been trained on more than a million images. Factorising convolutions in Inception v3 aims to minimise the number of connections and parameters while maintaining network efficiency [40].

- Features of Inception V3

The following are the features of Inception v3 model –

Compared to the Inception V1 and V2 models, it is faster, more efficient, and has a larger network. It also costs less to compute and uses auxiliary classifiers as it regularises.

Since the inception V3 model is essentially an enhanced version of the inception V1 model, that is why we selected it as our model for this project. The Inception V3 model optimises the network in a variety of ways for increased model adaptability.

SUGGESTIVE SYSTEM

The proposed system's design consists of the five phases listed below, as well as a flow diagram depicting the steps required is shown in the figure 4 below –

- Database Gathering
- Model Training
- Hand Segmentation and Pre-Processing
- Feature Extraction
- Classification

  Word to Speech Conversion

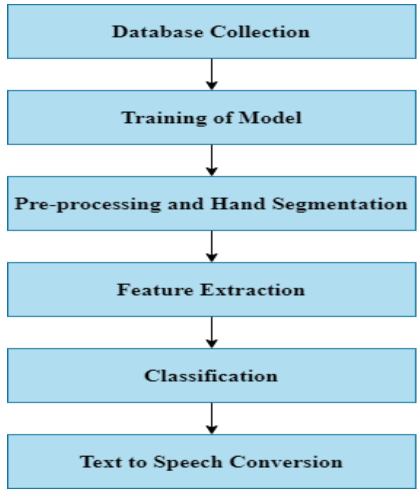

Figure 4: Flowchart for System

In the following sub-sections of this section, each of the steps is detailed.

Database Gathering

Every project needs a database collection since we train our models using databases, and getting databases from reputable sources is crucial to guaranteeing the accuracy of the models. We have gathered the dataset from Kaggle, the world's largest data science community, which offers an abundance of tools and services to help you accomplish your data science goals, for this research as we have worked on both American and Indian sign languages [8].Two different kinds of data sets have been used: one for American Sign Language and the other for Indian Sign Language.

To improve accuracy and efficiency, we have trained our model using around 1200 photos for each alphabet. Hence, 26x1200 is a sizable dataset that can help us train our model more effectively, enabling us to achieve higher accuracy. Additionally, because the signer's hands are in different positions and orientations for the same sign, the system is more flexible. There are a total of 26 alphabets, and each alphabet has 1200 data sets.

Model Training

We've used the Inception V3 model to achieve our objective. The deep learning model Inception V3, which is used for picture classification, makes use of convolutional neural networks. The Inception V3, a refined iteration of the foundational model Inception V1, was first made available as GoogLeNet in 2014. It was created by a Google team, as the name would imply [11].

- Step 1: The model training process is to load the dataset via the working folder's url.
- Step 2: Open a variable and read the dataset photos.
- Step 3: Establish a testing data size of 20% and a training data size of 80%.
- Step 4: To create training and testing samples, randomly divide the photos.
- Step 5: Apply image manipulation techniques such as normalisation, random rotation.
- Step 6: After the transformation, verify that the photos are legitimate.
- Step 7: Check different label classes by iterating through the training photos.
- Step 8: Assign class names to labels so that the result may be predicted.
- Step 9: Increase the epoch (training cycle) count to 25.
- Step 10: Model Training for each cycle to increase precision at each stage.
- Step 11: Test the model to determine the loss of training and accuracy.
- Step 12: Store the created model locally.

Hand Segmentation and Pre-processing

The preliminary processing procedure begins with this. This is how picture sensing works. Every picture frame is pre-processed to eliminate noise. In this project, the camera module is started using Open CV. A rectangular box that aids in hand gesture detection in the absence of background noise is visible when the cameras open to capture a picture of a hand gesture.

- Open CV

In 1999, Gary started OpenCV at Intel with the intention of accelerating computer vision investigation and business use globally while simultaneously increasing demand for ever-more-powerful processors for Intel[20]. A sizable open-source library for image processing, machine learning, and computer vision is called OpenCV. Python, when paired with other libraries like NumPy, can process the OpenCV array structure for analysis. To find visual patterns and their different features, we utilise vector space and apply mathematical operations to these features [16].

- Segmentation

Segmentation is the process of breaking an image up into smaller parts in order to extract more precise picture features. The representation and description of the image will be accurate if the segments are sufficiently autonomous, meaning that no two segments of the image should have the same information. Rough segmentation will produce inaccurate results.[19]

To guarantee that high-quality features are recovered from the input image, a number of actions are carried out on it prior to feature extraction. Hand segmentation is done using threshold-based segmentation [18].

Feature Extraction

Feature extraction is an important step in the categorization of pictures. It allows for the most accurate portrayal of the visual content [21]. A component in the dimensionality reduction process that separates and organises massive amounts of raw data is called feature extraction. With classes reduced to smaller, more manageable groups, processing would be easier.

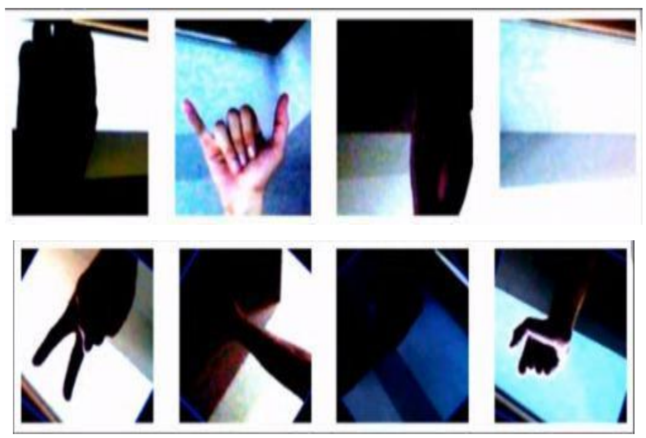

Figure 5

Figure 5 illustrates how different image pre-processing methods, such as scaling, normalisation, thresholding, and binarization, are applied to the sampled image prior to obtaining features. After that, features that can be utilised to categorise and identify photos are extracted using feature extraction techniques [25]. The most important aspect of these enormous data sets is their abundance of variables. Processing these variables requires a large computational resource. As a result, function extraction helps to extract the best feature from enormous data sets by choosing and combining variables into functions. These features reduce the amount of data while precisely and uniquely identifying the data collection process. They are also simple to utilise.

Classification

The most important problem of identification accuracy has led to the development of numerous photo classification models. With so many useful applications, image categorization is an important topic in computer vision [28]. To train our model, we employed a transfer training approach. This project uses the Inception V3 Model, an image classifier model that is pre-trained on a sizable amount of data and runs on a Convolutional Neural Network, or CNN. Therefore, by transfer learning, we imply that we have used our target dataset of sign languages to train the current Inception V3 model. As of right now, we have predicted the different sign language labels using our alphabet recognition model. Using the user image as input, the predict function maps it to the appropriate label using the learned model. In the end, as Figure 6 illustrates, the output returns the correct label.

Alphabet - R

Alphabet - O

Alphabet - C

Alphabet – W

Alphabet – V

Figure 6: Output Labels

Word to Speech Conversion

One of the oldest and most instinctive ways that people communicate is through speech. Over the course of time [23]. Text-to-speech (TTS) is the process of converting words into a voice audio format. The programme, tool, or software receives text input from the user and applies natural language processing techniques to infer the language's grammatical structure and make logical deductions from it. The following block receives this processed text and applies digital signal processing on it. After processing, a number of methods and changes are applied to convert this text into a speech format. The entire process involves the synthesis of speech.

We utilised the gTTS module in this project to turn the text into speech. A Python library and command-line tool called Google Text-to-Speech (gTTS) can be used to communicate with the Google Translate text-to-speech API [10]. The gTTS module will import the gTTS library, which is useful for voice translation [9].

The process of text-to-speech (TTS) synthesis consists of two phases. The first phase is text analysis, which converts the input text into a phonetic or other linguistic representation. The second phase involves creating speech waveforms by using the prosodic and phonetic information to produce the desired output. Usually, "high-level synthesis" and "low-level synthesis" are used to describe these two phases [24].

The gTTS module can also be useful for languages like Hindi, French, and German. This is very helpful when there is a barrier to communication and the user cannot convey his messages to others. For people who are blind or have other disabilities, text-to-speech is a great help since it can help with text-to-speech translation. There are numerous opportunities when using the gTTS module for additional languages.

i) Features in GTTS Customisable Sentence tokenizer tailored to speech: Fig. 7 illustrates the text pre-processors' ability to be customised to provide features like pronunciation using the GTTS library. The sentence tokenizer can read any length of text while maintaining appropriate intonation, abbreviations, decimals, and other features.

.

```
mytext = 'Identified Label is {}'.format(label)
language = 'en'
myobj = gTTS(text=mytext, lang=language, slow=False)
myobj.save("voice.mp3")
os.system("voice.mp3")
```

Figure 7: Text to Speech part

RESULTS

As seen in Figure 8, the Inception v3 model performed exceptionally well in categorising sign language movements. We were able to attain 98.99% accuracy for American Sign Language with a training loss of 1.46%. Another problem was that when the researchers used letters like {C, L, M, N, R, U, Y} in our project, they were not getting good accuracy. We have reached the previously mentioned accuracy with the inclusion of these seven letters, for a total of 26 letters, {C, L, M, N, R, U, Y}. Using American Sign Language, we tested our method on all 26 alphabets, including the letters C, L, M, N, R, U, and Y. We achieved 99.99% accuracy.

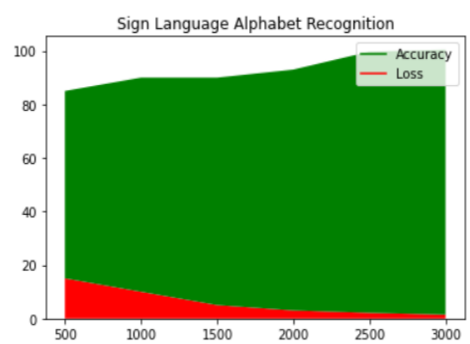

Figure 8

SUMMARY & FUTURE GOALS

Through this study, we have shown how effective it is to use transfer learning. Using a dataset of 3000 images per alphabet for sign language, we trained the pre-trained Inception V3 model—which is based on CNN and Deep Neural Network algorithms—on sign language data. We were able to improve our accuracy in sign language recognition because to this sizable dataset.

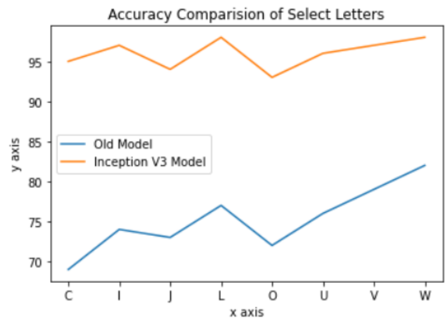

Figure 9: Select Alphabet Accuracy comparison

As seen in Fig. 9, we have obtained same accuracy for every alphabet, which solves the issue of earlier work having inferior accuracy on some selected single hand alphabets.

The project's future goals include developing a mobile application that can be loaded on portable devices like smartphones and smart watches and recognise words and phrases with the same level of accuracy, allowing users to use it freely in their daily lives.

REFERENCES

1. A real-time continuous gesture identification system for sign language was developed by [1]Liang R-H and Ouhyoung M in 1998. In: 1998 IEEE International Conference on Automatic Face and Gesture Recognition. Process. IEEE, Third, pp. 558–567
2. Liang R-H (1997) Taiwanese sign language continuous gesture recognition system. University of National Taiwan
3. [Dieleman S., Kindermans PJ, Schrauwen B., Pigou L. (2015) Convolutional Neural Networks for Recognition of Sign Language. In: Bronstein M., Rother C., Agapito L. (eds) Computer Vision - Workshops at ECCV 2014. Ily is alive.
4. Starner T, Weaver J, Pentland A (1998) Desk and wearable computer based video for real-time American sign language detection. 20(12):1371–1375 in IEEE Trans Pattern Anal Mach Intell
5. Metaxas D. and Vogler C. (1997) modifying hidden markov models for three-dimensional computer vision techniques in order to recognise ASL. Published in: IEEE SYSTEMS MAN AND CYBERNETICS, vol. 1, IEEE, pp. 156–161.
6. Jack MA, Ariki Y, and Huang XD (1990) Markov hidden-models for audio identification
7. Hendriks EA, Reinders MJT, and Lichtenauer JF (2008) Combining independent classification with statistical DTVW allows for the recognition of sign language. 30(11):2040–2046 IEEE Transactions on Pattern Analysis & Machine Intelligence
8. Databases at https://www.kaggle.com/docs/gTTS at https://pypi.org/project/gtts projects: https://readthedocs.org/downloads/pdf/latest/ http://iq.opengenus.org/model-architecture-inception-v3/
9. What is deep learning, according to https://machinelearningmastery.com/
10. "Understanding of a convolutional neural network," S. Albawi, T. A. Mohammed, and S. Al-Zawi, 2017 International Conference on Engineering and Technology (ICET), 2017, pp. 1-6, doi: 10.1109/ICEngTechnol.2017.8308186
11. Keiron O'Shea and Ryan Nash. (2015) Convolutional Neural Networks: An Introduction. ArXiv electronic prints.
12. This article describes how to use Inception v3 for image classification in transfer learning for image classification (86700411251b).
13. OpenCV Overview: https://www.geeksforgeeks.org
14. Itmconf_icacc2021_03004 (https://www.itmconferences.org/articles/itmconf/pdf/2021/05/itmconf). Pdf
15. A unique Extreme Learning Machine method for ISL alphabet recognition Anand Kumar (1) Kumar Ravinder2. Accepted on September 26, 2020 / Received on December 5, 2019 The 2020 Institute of Computer Applications and Management at Bharati Vidyapeeth The file is called IJAER18/IJAERv13n9_90.pdf.
16. Shervin Emami and Valentin Suciu (2012). OpenCV facial recognition technology. The journal for distributed, embedded, and mobile systems.
17. Ahmed Medjahed, Seyyid (2015) An Analysis of Feature Extraction Techniques for Image Classification in Comparison. International Journal of Signal Processing, Images, and Graphics

18 Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, and ZB Wojna (2016). reevaluating the computer vision-focused Inception architecture. CVPR.2016.308 10.1109/.

19. Ifeanyi Nwakanma, Ikenna Oluigbo, and Okpala Izunna (2014). 2. 154–163. Text–To–Speech Synthesis (TTS). S. Lemmetty (1999). Evaluation of Speech Synthesis Technology. Helsinki University of Technology, Master's Diss.
20. Gaurav Kumar and Pradeep Bhatia (2014). An in-depth analysis of image processing systems' feature extraction capabilities. doi:10.1109/ACCT.2014.74.
21. Convolutional neural networks (CNN) in vegetation remote sensing: a review by Teja Kattenborn, Jens Leitloff, Felix Schiefer, and Stefan Hinz, ISPRS Journal of Photogrammetry and Remote Sensing, Volume 173,2021, Pages 24-49, ISSN 0924-2716, https://doi.org/10.1016/j.isprsjprs.2020.12.010.
22. In Proceedings of the IEEE, vol. 109, no. 1, pp. 43-76, Jan. 2021, F. Zhuang et al., "A Comprehensive Survey on Transfer Learning," doi: 10.1109/JPROC.2020.3004555.
23. Mahbub Hussain, Jordan Bird, and Diego Faria (2018). A Study on Image Classification Using CNN Transfer Learning.
24. Waibel, A., Hanazawa, T., Hinton, G. E., Shikano, K. & Lang, K. Time-delay neural networks for phoneme recognition. IEEE Transactions on Acoustics and Speech Signal Processing 37 (1989), 328–339.
25. Lienard, J., Blanchet, P., Fogelman-Soulié, F., and Bottou, L. Trials utilising time delay networks and dynamic time warping to recognise solitary digits independently of speakers. In Proc. EuroSpeech 89 (1989), 537–540.
26. Bengio, Y., Cho, K., and Bahdanau, D. neural machine translation through collaborative alignment and translation learning. Published in Proceedings of the International Conference on Learning Representations (2015), arxiv.org/abs/1409.0473.
27. K. Xu et al. Neural image caption generation combining visual attention to "show," "attend," and "tell." Title: Proceedings of the International Conference on Learning Representations (2015); arXiv:1502.0344.
28. ks. IEEE Trans. Acoustics Speech Signal Process. 37, 328–339; G. P. C. Fung, J. X. Yu, H. Lu, and P. S. Yu, "Text classification without negative examples revisit," IEEE Transactions on Knowledge and Data Engineering, vol. 18, no. 1, pp. 6–20, 2006.
29. H. Al-Mubaid and S. A. Umair, IEEE Transactions on Knowledge and Data Engineering, vol. 18, no. 9, pp. 1156–1165, 2006, "A new text categorization technique using distributional clustering and learning logic."
30. Construction of informative priors using transfer learning, R. Raina, A. Y. Ng, and D. Koller, [35]Proceedings of the 23rd International Conference on Machine Learning, Pittsburgh, Pennsylvania, USA, June 2006, pp. 713–720
31. The paper "Co-clustering based classification for out-of-domain documents" was presented at the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining in August 2007 in San Jose, California, USA. [36] W. Dai, G. Xue, Q. Yang, and Y. Yu.
32. Transferring naive bayes classifiers for text classification," by W. Dai, G. Xue, Q. Yang, and Y. Yu, in Proceedings of the 22nd AAAI Conference on Artificial Intelligence, Vancouver, British Columbia, Canada, July 2007, pp.
33. In 2019 Period. Polytech. Transp. Eng., vol. 47, no. 3, pp. 242–250; C. Lin, L. Li, W. Luo, K. C. P. Wang, and J. Guo, "Transfer Learning Based Traffic Sign Recognition Using Inception-v3 Model."
34. In the 2019 International Conference on Artificial Intelligence and Advanced Manufacturing (AIAM), S. He presents "Research of a Sign Language Translation System Based on Deep Learning," with pp. 392–396, doi: 10.1109/AIAM48774.2019.00083.
35. "Deep neural mapping support vector machines," Li Y., Zhang T., Neural Networks, Vol. 93, pp. 185-194, 2006.