

## Technical Architecture:

Character : A  
Word :  
Sentence : HI I

Output text

*Recognized  
hand  
gesture*



User interface Capture  
the sign within ROI

Background is popped and original sign  
is extracted



Original Gray-scale Gaussian blur filter  
to get outline

End-user

Obtain Output text in the  
form of speech

**Table-1 : Components & Technologies:**

S.No	Component	Description
1.	User Interface	How user interacts with application <b>usage and clicking the corner</b>
2.	Application Logic-1	<ul style="list-style-type: none"> <li>• Camera <b>detects the sign</b></li> <li>• Captures <b>the sign with</b></li> </ul>
3.	Application Logic-2	<ul style="list-style-type: none"> <li>• <b>Background is popped</b></li> </ul>

**Table-2: Application Characteristics:**

S.No	Characteristics	Description	Technology
1.	Open-Source Frameworks	<ul style="list-style-type: none"> <li>• <b>Palm detector</b> operates on full images and outputs an oriented bounding box.</li> <li>• <b>Hand landmark</b> takes the cropped image defined by the palm detector and returns 3D hand key points.</li> <li>• <b>Gesture recognizer</b> then classifies the previously computed key point configuration into a discrete set of gestures</li> </ul>	<b>Media Pipe Framework is used.</b> Within this framework, the pipeline is built as a directed graph of modular components.

4.	Application Logic-3	<ul style="list-style-type: none"> <li>• Extract the <b>edges of the</b></li> </ul>
5.	Application Logic-4	<ul style="list-style-type: none"> <li>• Convert the <b>output text</b></li> </ul>
6.	Database	<ul style="list-style-type: none"> <li>• <b>Binary Large Object(BLOB)</b> images in the dataset</li> <li>• <code>/etc/mysql/my.cnf</code> is the default <b>MYSQL</b> that is used.</li> </ul>
7.	File Storage	<ul style="list-style-type: none"> <li>• <b>Create a BLOB column</b> <b>JPEG, PNG, PSD</b> or <b>into the table/column</b></li> </ul>
8.	Machine Learning Model	<b>Allows the user to feed a collection of data and have the computer recommendations and decisions</b>

2.	Scalable Architecture	<ul style="list-style-type: none"> <li>• It's a Three –Tier Architecture comprises the following technology, Convolutional neural network can be scaled in three dimensions: <i>depth, width, resolution</i>.</li> <li>• <b>Depth</b> of the network corresponds to the number of layers in a network.</li> <li>• <b>Width</b> is associated with the number of neurons in a layer. •</li> </ul> <p><b>Resolution</b> is the image resolution that is being passed to CNN. <b>Increasing the depth</b>, by stacking more convolutional layers, allows the network to learn <b>more complex features</b>.</p>	<b>Convolution Neural Networks is used.</b>
3.	Availability	Hand gestures are <b>the natural way of interactions when one person is communicating with one another</b> and therefore hand movements can be treated as a non verbal form of communication. Hand gesture recognition is a process of understanding and classifying meaningful movements by the human hands	<b>CNN, Media Pipe, Gaussian blur filter, Machine learning models along with Speech assistant</b> is used.