|  |  |
| --- | --- |
| Internship Project Title | Automate identification and recognition of handwritten text from an image |
| Student Name | Aman Kumar |
| Name of the Company | TCS ion |
| Name of the Industry Mentor | Anamika chatterjee/Debashis Roy |
| Name of the Institute | GIET University |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Start Date | End Date | Total Effort (hrs.) | Project Environment | Tools used |
| 19th may | 2nd july | 90 | Anaconda(spyder) | Scipy,pandas,numpy  ,TensorFlow, |
| Project Synopsis: pre processing the images and apply Artificial Neural Network | | | | |
| Solution Approach:   1. Pre-processing of sample image 2. Feature extraction 3. Creating and training of network 4. Simulation 5. Result 6. conclusion | | | | |
| Assumptions: : The assumptions considered are as follows:  1. The handwritten text across must be in English.  2. The text across the input image must be clearly handwritten in order to achieve good results.  3. All machine dependencies must be installed properly. | | | | |
| Project Diagrams: Logic Flow of Convolution Neural Networks    Logic Flow of Pytesseract OCR Engine | | | | |
| **Algorithms**: The algorithm used to implement the project is the Pytesseract ocr engine which uses convolutional neural network algorithm which is used by the tesseract optical character recognition engine in python. There are four layered concepts we should understand in convolutional neural networks:  1. Convolution  2. Rectified Linear Unit  3. Pooling Layers  4. Full Connectedness (Fully Connected Layer)  We have 4 steps for convolution:  • Line up the feature and the image  • Multiply each image pixel by corresponding feature pixel  • Add the values and find the sum  • Divide the sum by the total number of pixels in the feature    The output signal strength is not dependent on where the features are located, but simply whether the features are present. Hence, an alphabet could be sitting in different positions and the convolutional neural network algorithm would still be able to recognize it.  **Rectified Linear Unit**: Transform function only activates a node if the input is above a certain quantity, while the input is below zero, the output is zero, but when the input rises above a certain threshold, it has a linear relationship with the dependent variable.    The main aim is to remove all the negative values from the convolution. All the positive values remain the same but all the negative values get changed to zero as shown below:      **Pooling Layer**: In this layer the shrink the image stack into a smaller size. Pooling is done after passing through the activation layer. We do this by implementing the following 4 steps: • Pick a window size (usually 2 or 3) • Pick a stride (usually 2) • Walk your window across your filtered images • From each window, take the maximum value | | | | |
| **Outcome**: The algorithm is able to detect and segment handwritten text from an image. The model successfully able to detect maximum words in a given line of sentence or words, which makes it about 90% accurate while implementation and testing. For example the input image having the handwritten text is given as following: | | | | |
| Exceptions considered:: The exceptions considered are as follows  1. The text across the input image must be of the same color not multicolor handwritten text.  2. The image doesn’t have too aggressive multicolor backgrounds across the text of the image.  3. The image doesn’t have any kind’s objects in the background across the text of the image. | | | | |
| Enhancement Scope: The enhancement scope of this project are follows:  1. The accuracy of the model can increased with predefined models and powerful machine learning GPU processors can be used to attain a good percentage of accuracy.  2. In future we can use this algorithm with more than one particular language. | | | | |
|  | | | | |