Deep Learning Based Annotation Tool For Optical Character Recognition

¹Ms. Bhargavi N.G, ²Mr.Darshan K.V, ³Mr.Aravind Ravindran

¹M.Tech.PG Student, ²Assistant Professor, ³Software Engineer
¹,²Department of DECS, ³Qualitas Technologies Pvt.Ltd
¹,²JNNCE, Shivamogga

Abstract: This paper discusses about the Annotation tool for the optical Character Recognition Technique using deep learning. In today's world, with the advancement in the technology, everything is become digitized. This has led to converting all the important documents in digital format. It is at this point where OCR (Optical Character Recognition) plays a very important role. For OCR annotation purpose, we are creating a user interface with the help of WPF (Windows Presentation Foundation) markup language. WPF helps to the user to look good UI design for a more professional look and feel to the user interface. Windows Presentation Foundation (WPF) is a software technology contains the developments of desktop applications, web applications and mobile applications

Keywords: optical character recognition; annotation; user interface; wpf;

I. INTRODUCTION

Optical Character Recognition (OCR) deals with converting printed text into text characters so that it can machine read, which helps for the blind peoples to read and for converting the telegraphic messages that are received as well as read characters and convert them into telegraphic code. In the 1950's OCR was commonly used in data entry applications. OCR also helps in transform of foreign text into another language. Handwritten and printed documents which are either scanned or photographed cannot be used well as they are in an image format which proves as a lot of barrier and involves labor to convert them into computer editable data. This conversion can prove costly for organizations that handle large amounts of data and also can generate huge amounts of human errors in the resulted documents. Deep learning which is also known as the deep structured learning or hierarchical learning is a broader family of machine learning methods based on learning data representations as opposed to task specific algorithms. A computer aided conversion such optical documents can prove when labor is scarce and efficiency needs to increased. People have been struggling in the past with data entering and digitization of data. Existing techniques to overcome these problems such as OMR sheets and form data have been outdated due to the changes in type of data. To overcome these problems a neural network based approach is required to convert data of varied types in to digital information.

II. LITERATURE SURVEY

For the purpose of creating a annotation tool for ocr we have created the user interface for the annotations purpose. Here we have used wpf language for creating this. Windows Presentation Foundation (WPF) is a user interface framework for building Windows client applications with immersive and intuitive user experiences [3]. WPF combines the application user interface, 2D graphics, 3D graphics, documents and multimedia into one single framework to help developers create rich and interactive applications. This framework was created to help developers meet increasing expectations of the experience and usability of software applications.[1]. The operator interface is the critical link between a test system and its operator. When a test fails, the operator must quickly process the results and decide whether to troubleshoot, rerun or halt the test based on information displayed by the software.[4] An effective and well-designed operator interface can increase productivity, reduce testing time and operator error as well as improve adoption of the software. Whether the interface displays a simple pass/fail status or offers sophisticated troubleshooting operations, implementing a good user interface experience can be a challenging task [2]. In today's world, with the advancement in the technology, everything is become digitized. This has led to converting all the important documents in digital format. It is at this point where OCR (Optical Character Recognition) plays a very important role.[5] Lots of researchers have been working tirelessly to improve the accuracy of the recognition rate. There have been a lot of improvement in searching various methods that gives a very high accuracy where the recognition is concerned.

III. :SYSTEM OVERVIEW

The operator interface is the critical link between a system and its operator. When a test fails, the operator must quickly process the results and decide whether to troubleshoot, rerun or halt the test based on information displayed by the software. An effective and well-designed operator interface can increase productivity, reduce testing time and operator error as well as improve adoption of the software. Whether the interface displays a simple pass/fail status or offers sophisticated troubleshooting operations, implementing a good user interface experience can be a challenging task.

Windows Presentation Foundation (WPF) is a user interface framework for building Windows client applications with immersive and intuitive user experiences. WPF combines the application user interface, 2D graphics, 3D graphics, documents and multimedia into one single framework to help developers create rich and interactive applications. This framework was created to help developers meet increasing expectations of the experience and usability of software applications. WPF facilitates the creation of high-quality user interfaces that stand out amongst competitors and help test operators accomplish their tasks faster with less opportunity for error. Windows Presentation Foundation provides developers with the tools needed to more rapidly iterate on the UI and reach a better quality user interface in a shorter amount of time

A. Creating a simple Wpf Application

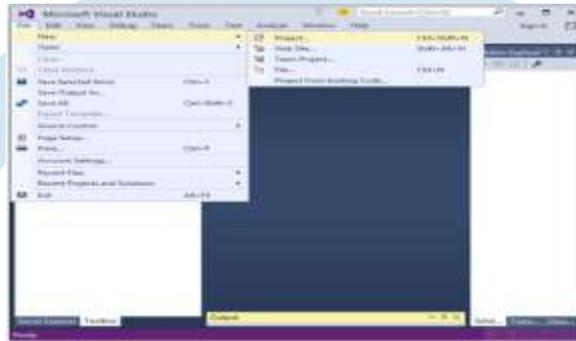- Open Visual Studio and then Click File > New > Project menu option.

Figure1: Creation of wpf application

- In the New Project dialog box, in the Templates pane, click on Visual C# and in middle pane, select WPF Application.

Figure2: Selecting wpf application

- Give the project a name. Type **HelloWorld** in the name field and click the OK button.

- By default, two files are created, one is the **XAML** file (mainwindow.xaml) and the other one is the **CS** file (mainwindow.cs)

- On mainwindow.xaml, you will see two sub-windows, one is the **design window** and the other one is the **source** (XAML) **window**.

- In WPF application, there are two ways to design an UI for your application. One is to simply drag and drop UI elements from the toolbox to the Design Window. The second way is to design your UI by writing XAML tags for UI elements. Visual Studio handles XAML tags when drag and drop feature is used for UI designing.

- In mainwindow.xaml file, the following XAML tags are written by default.

Figure3: Basic Xaml

- For example, If we want to create a button click event then either from the toolbox we have to drag the elements, or we have to create it in source window by writing xaml code.

- When the above code is compiled and executed, the output can be seen as below.

Figure4: Execution

IV. PROPOSED WORK

In this work,, various wpf controls and wpf panel is used for making an user interface , and source coding is done using csharp language .Here we annotate the characters of various images by their shapes by giving the proper annotation details ,and we also able to zoom the image using a slider control. We save the results of these annotation in a file by using an xml document and also we are able to create a log file for storing all the activities that we have done during output.

A. Annotation Tool

For creating a user control we have used the button control for various functions like opening the selected image, for storing the annotated image and for selecting the shapes and also for the tag name that we are giving as input. Here we are slider control for zooming of the image for correct annotation purpose. We have used grid panel here for making the panel responsive, and also to arrange the various controls in a proper way. We have used user controls here for the purpose of various controls like image source control, image viewer control, image view controller. The outlook of the project 'Annotation Tool for OCR' is as shown in figure.
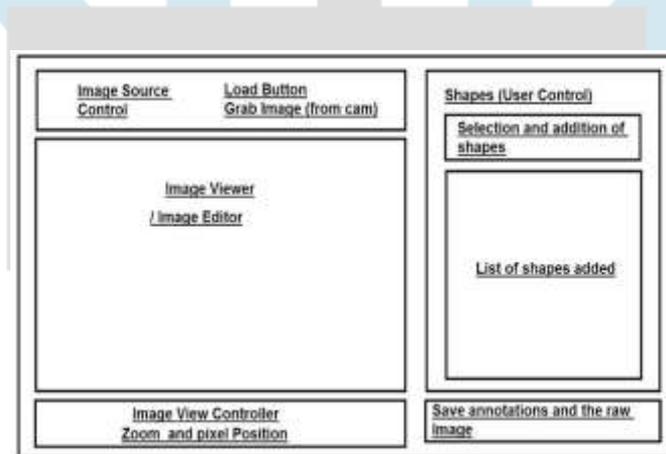

Figure5: Design of Annotation Tool

1) Image source controller

Here,for image source controller, we have arranged the elements in a grid panel since it makes the user interface very responsive .Inside the grid panel button control and also one expander control are used, button control for loading the image from the file ,and expander control for listing the images that we have used for annotation.

Figure6: Image Source Controller

2)   Image Viewer and Image View Controller(User control):

Image Viewer and Image viewer controller is made using the grid panel and and they are used as user controls. User Controls serve the purpose of reusing controls. A user control is basically a grouping of other existing control, intended as a reusable component (i.e. composite control). If you need to place the same group of controls on different windows you'd rather group them in a user control, adding things like data validation for instance, and then reuse this control whenever you need it.
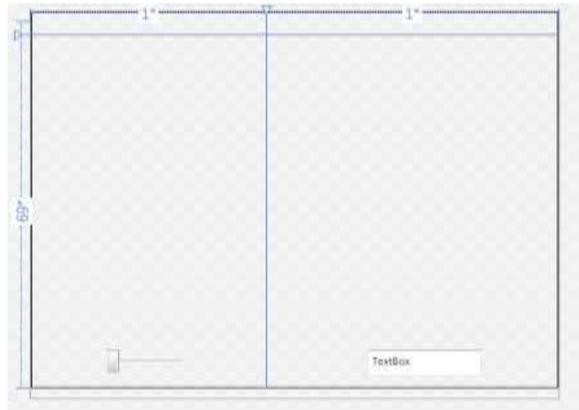

Figure7: Image Viewer

3)   List of Annotations(User Control)

Here , user control is created and consisting of a combobox, and buttons, and a listbox.Here also we have used grid panel foe arranging elements .a combo box is created for selection of shapes that we are adding, and add and delete buttons for adding and deleting the images that we are adding in the listbox . Below shows the design of the user control that we have created.
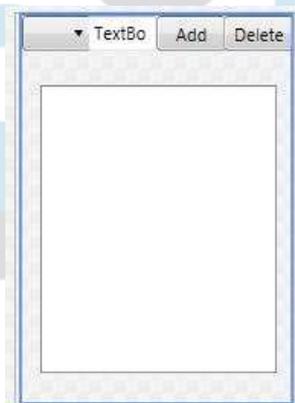

Figure8: List of Shapes (User Control)

V.   RESULTS AND DISCUSSIONS

Based on the controls we have used and the controls we have created a well designed xaml structure is shown in figure9.It also consists of the save button for saving the raw images that we have taken , and also the annotated image.

1)   Annotation Results

Based on the controls and user controls that we have created,an annotation tool for ocr system is shown in below figure.Upon clicking the load button , a open file dialog box is created for choosing the image that we want to annotate.More number of annotations can be done by choosing the different image samples, and the name and also the list of images that we have taken for annotations can be seen in the expander control.a zoom option is provided for zooming the images for proper annotations.
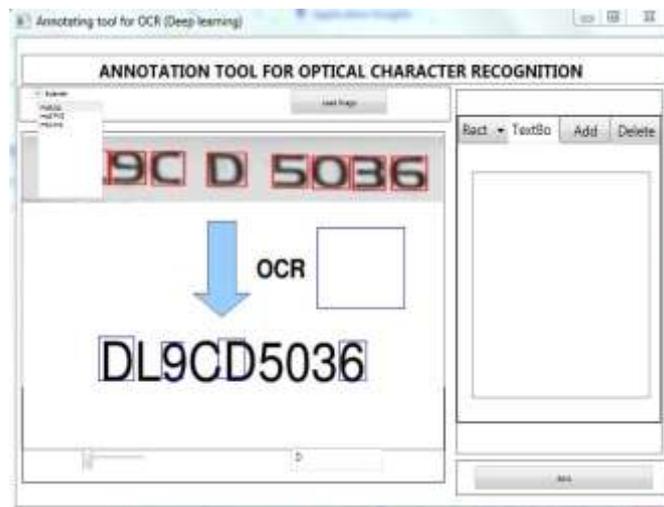
Figure 9: Annotation tool for OCR

The annotated character is displayed in the textbox and they are recorded in the xml file. An xml file consist of the information about the character that we are annotating like shape, character name, rectangle height, rectangle width etc. To annotate the characters firstly, mouse left button was clicked for taking the initial point of drawing rectangle, and then we have to drag the mouse until we need the correct final point of drawing rectangle. Mouse left button was raised and then it draws a rectangle on the characters for annotations. These mouse move events are recorded in a separate log file for future reference purpose. Hence we can annotate the number of characters.

2)   Log File

A log file is a file that records either events that occur in an operating system or in a software system\cite{test10}. In our project, log file is used for updating the mouse move events that are occurring inside the image control for annotation purpose.Below is the output of mouse move events in image control for annotations.
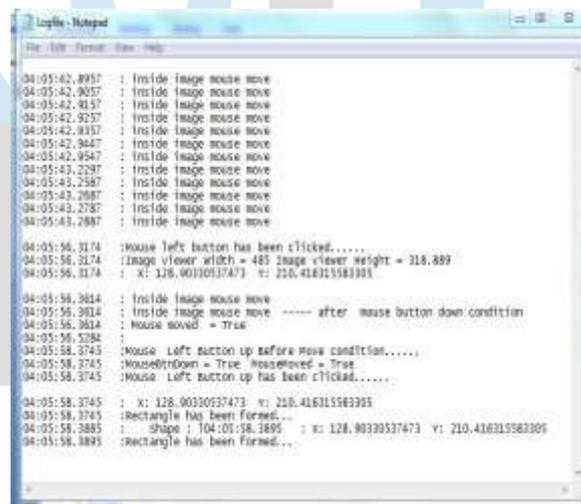

Figure10: Log file Output

3)   Xml Document

XML is a general purpose tag based language and very easy to transfer and store data across applications. Xml files are made up of tags that contains data. Generally a start tag and end tag to hold the data. For example, if you want to create an XML tag name "Header" , the start tag is like *< Header >* and the end tag is like *< Header >* . We can fill our information between these tags.Data that we have stored in xml document is shown below.

Figure11:XML Document

VI. CONCLUSION

Optical characters can be recognized in multiple ways, however using something that mimics the nature and characteristic of human brain is bound to succeed better than any other system. For character recognition purpose a user interface based system is created for the annotation of characters using Windows Presentation Foundation technology. Windows Presentation Foundation is Microsoft's premier technology for creating Windows graphical user interfaces. The user experience shapes the perception of quality of a software application because it is the only means to give commands and receive feedback from the system. An intuitive and productive user interface implies a thorough, professional development team whereas an outdated or difficult-to-use UI can cause frustration or even operator error. Moreover, the perceived standard of a good user experience has increased and continues to grow through the proliferation of sophisticated consumer products. As a result, test system developers will need to re-evaluate their UI design platforms and take advantage of frameworks such as Windows Presentation Foundation to create the next generation of test systems.

References

[1]   H. Chung, S. J. Lee, and J. G. Park, "Deep neural using trainable activation functions," in Neural Networks (IJCNN), 2016 International Joint Conference on, 2016, pp. 348–352.

[2]   Y. Tang, L. Peng, Q. Xu, Y. Wang, and A. Furuhata, "CNN Based Transfer Learning for Historical Chinese Character Recognition," in Proceedings of the 12th IAPR Workshop on Document Analysis Systems (DAS), 2016, pp. 25–29].

[3]   Jihad El-Sana, Klara Kedem, Ofer Biller (2014), "The Influence of language orthographic characteristics in digital word recognition", Proceedings of the 10th International Workshop on Frontiers of Handwriting and Recognition, pages 1009–1033.

[4]   O. A. W. Ahmed, "Application of artificial neural networks to optical character recognition", Master's Thesis, King Fahd University Of Petroleum & Minerals, Saudi Arabia, June 1994.

[5]   D. X. Ni, "Application of neural networks to character recognition", Proceedings of Students/Faculty Research Day, Pace University, pp.1-6, 2007.

[6]   Microsoft MSDN, "Introducing Windows Presentation Foundation," (http://msdn.microsoft.com/enus/library/aa663364.aspx [Accessed: June 8, 2012]).

[7]   Badawy, W. *"Automatic License Plate Recognition (ALPR): A State of the Art Review."* IEEE International Conference on Document Analysis and Recognition, 2012.