

Faculty of Engineering and Environment

**KF7028 Research and Project Management**

**Assignment: *Research Proposal***

**Handwritten Text Recognition using Deep Learning.**

**Module**: KF7028 – Research and Project Management

**Module tutor**: Shelagh Keogh

**1St Supervisor name:** Bo wei

**2nd Supervisor name:** Yilun Shang

**Assignment title:** MSc Research Proposal ‘Handwritten Text Recognition using Deep Learning’.

**Student name and your university identifier:** Ujwal Nagesh Sri Krishna Kumar Mathamsetti - 20027087

**Programme of study/course:** MSc Advance Computer Science with Advance Practice

**Confirmation that you have completed the draft of your ethical approval process:**

Word Count: 2936

Table of Contents

[1 Aim/Research question 4](#_Toc68323322)

[1. 1 Aim 4](#_Toc68323323)

[1.2 Research Questions 4](#_Toc68323324)

[2 Background and Motivation 5](#_Toc68323325)

[2.1 Background 5](#_Toc68323326)

[2.2 Motivation 6](#_Toc68323327)

[2.3 Literature Review 6](#_Toc68323328)

[2.3.1 Summary of Literatures 8](#_Toc68323329)

[3 Scope, Objectives and Risk 10](#_Toc68323330)

[3.1 Scope 10](#_Toc68323331)

[3.2 Objectives 11](#_Toc68323332)

[3.3 Risk Factors 11](#_Toc68323333)

[4 Ethics, Legal, Social, Security and Professional Issues 13](#_Toc68323334)

[4.4 Ethics Issues 13](#_Toc68323335)

[4.2 Social Issues 13](#_Toc68323336)

[4.3 Security Issues 13](#_Toc68323337)

[4.4 Professional Issues 13](#_Toc68323338)

[5 Schedule of Activities 13](#_Toc68323339)

[5.1 Work Breakdown Structure 13](#_Toc68323340)

[5.2 List of Tasks to be Done 14](#_Toc68323341)

[5.3 Gantt Chart 15](#_Toc68323342)

[6 References 16](#_Toc68323343)

[7 Appendix A 18](#_Toc68323344)

**List of Figures**

Fig-1: Methodology Proposed by Joshi & Risodkar (2018) 6

Fig-2: Framework and Confusion Matrix by Yin et al. (2019) 7

Fig-3: Breakdown of Work 14

**Handwritten Text Recognition using Deep Learning.**

# 1 Aim/Research question

## 1. 1 Aim

Handwritten text recognition is a technology through which electronic devices and gadgets like smartphone, tablet pc etc. can identify any type of data. This dissertation aims to propose a model using machine learning algorithms Neural Network (Deep Learning) so that the algorithm can identify the pattern of the text, special symbols and can detect them. Another important thing is that while recognizing the text, some characters may be missed out due to the wrong detection which should be minimized. So, the second aim of this dissertation is to minimize the Missing Character Rate while detecting the entire set of texts. In this context, the emphasis will be done on the patterns of the texts like alphabets or numbers that will be written by hand.

## 1.2 Research Questions

The aim of the dissertation has been discussed in the last section. The detection of the handwritten text will be aimed to be detected using Neural Network (Cai, Chang, Tang, Xue, & Wei, 2018). To fulfil the aims of this project and dissertation, the following research questions have been taken into consideration:

1. What is the best way to detect the handwritten text?
2. Can both the alphabet and numbers be detected precisely or there will be any difference inaccuracy?
3. How much alphanumeric text sequence can be recognized at a time?

These research questions will be addressed while at the time of analysis of the handwritten text recognition. In the first question, the main issue is the choice of the method or algorithm to detect the handwritten text. There are several methodologies are available to execute the scenario like by applying machine learning or using deep learning. If the recognition will be done using deep learning, there are also different algorithms are available like convolutional neural network or recurrent neural network or artificial neural network. So, using the analysis and proposed methodology, the appropriate method will be found to detect the handwritten text with higher accuracy. The second research question is addressing to the type of handwritten text to be detected. As the texts may be a mixture of alphanumeric characters, the accuracy of detection of the alphabet and numeric characters may be different using the same algorithm. Here, it will be found whether the accuracy of the detection of the alphabet and numeric characters will be the same or there will be a difference. However, there may be a limitation of the recognition of the number of alphanumeric characters sequence at a time. So, the third research question is focused on the detection of the alphanumeric characters sequence whether it will be limited to certain characters or any length of alphanumeric characters sequence can be detected precisely.

# 2 Background and Motivation

## 2.1 Background

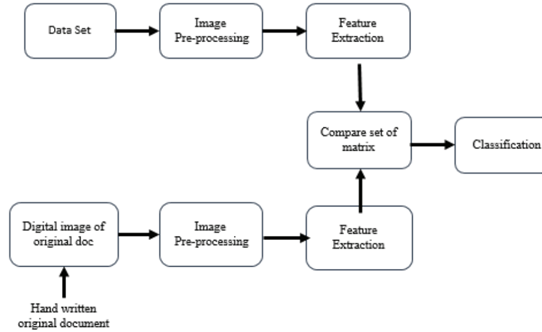
Text recognition is a soft technology that can solve the complex problem in the detection of handwritten characters in a suitable way. It facilitates the user to visualize the detected characters which they have entered. It facilitates more the busiest person who has no time to open the keypad and to type the text on mobile. However, this technology and application are no more bounded in this area only (Huang, Zhou, Li, & Li, 2019). It is applicable to different domain and field to facilitate the fastness of the detection. For this reason, this technology becomes famous in the last few years. This technology is the subdomain of Optical Character Recognition when it is applied offline. In the offline mode, the recognition of handwritten texts is detected by Optical Character Recognition or ORC technology. However, handwritten text recognition can be done both in online or offline mode (Sahu, et al., 2017). In the online mode of detection, the remote database is used where the pre-trained algorithm is invoked. when a user will write something or scan an image with a certain group of text, the online system can recognize the texts and show the detected text on the screen. in the offline system, the predicting algorithm is invoked in the application which will be responsible for the detection of handwritten texts. Many researchers have contributed their idea and framework to improve the procedure of detection (Razavian, Azizpour, & Carlsson, 2014).

## Motivation

Handwritten Text Recognition is a well-known technology that is used in a variety of fields today. However, all types of fonts and text styling are now recognisable. This is primarily because the algorithm is trained on specific font types and can only predict text based on that data. As a result, this could be the limited functionality of the technology that needs to be generalised. The font type and structure of the text differ from person to person. So, if technology and methodology can be proposed to recognise handwritten text in various fonts and structures, this will be a research topic. This is the driving force behind this research (Huang, et al., 2019)

## Literature Review

Joshi & Risodkar (2018) have proposed the method to detect handwritten text using deep learning. In this research, they have used deep learning and machine learning algorithms to recognize handwritten texts. In this context, they have applied the text images that contain a different combination of texts and have extracted the features from it using the following methodology:

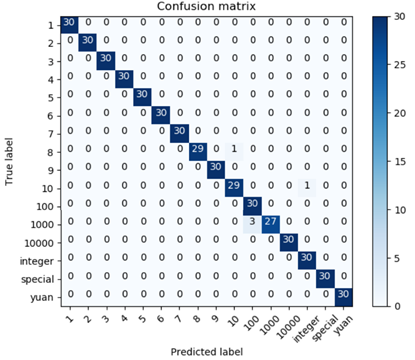
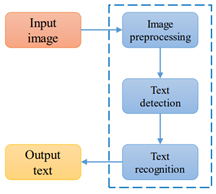


**Fig-1: Methodology Proposed by Joshi & Risodkar (2018)**

They have used Neural Network Classifier (NNC) and K-Nearest Neighbour Classifier (KNN) and finally compared the accuracy for both two cases. In the meantime, they have combined these two algorithms and checked the accuracy of detection of handwritten text image. In this experiment, they have got 80% accuracy in the detection of handwritten texts using the Neural Network Classifier. So, finally, they have proposed the model using Neural Network Classifier.

Borad, Dethaliya & Mehta (2020) have proposed the method for the detection of Handwritten Characters using deep learning algorithms. In this research, they have collected the image database for Handwritten Character Recognition and applied Convolutional Neural Network to that. They have especially emphasized the detection of handwritten Gujarati texts. At the time of execution, they have used five measures for image preprocessing namely Segmentation, Equalization, Skeletonization, Dilation, and Merging. After preprocessing the Handwritten Character Recognition database, they have applied Convolutional Neural Network and used MATLAB as a simulation tool. In this context, they have got average of 92% accuracy for handwritten text recognition.

Yin et al. (2019) have proposed the framework for the detection of handwritten text using Optical Character Recognition technology. They have applied the deep learning algorithm for the better detection of handwritten text. They have collected the database for OCR to train the deep learning algorithm that they have used (Convolutional Neural Network). Using Convolutional Neural Network, they have proposed the following framework and have got the confusion matrix for detecting handwritten text:



**Fig-2: Framework and Confusion Matrix by Yin et al. (2019)**

Using this framework, they have first extracted the features from the text image and tested those using a Convolutional Neural Network. In this research, they have got 96% accuracy with an error rate of 1.26%.

Sahu et al. (2017) have proposed the framework for the recognition of handwritten text using the deep learning algorithm. In this context, they have two deep learning algorithms namely the combination of Recurrent Neural Network and Long Short Term Memory (RNN-LSTM) and the Bidirectional Long Short Term Memory (BLSTM). In this research, they have used a very less number of images to train the algorithms which shows a new direction and improved version of the previous projects. They have applied RNN-LSTM and BLSTM on the collected text image data and trained those. In the next step, they have tested both two algorithms. Finally, they have got the highest accuracy using Bidirectional Long Short Term Memory by 92%.

### 2.3.1 Summary of Literatures

The review of the literature has been done in the last subsection. In this section, the comparison of the literature will be done in a tabular format concerning the brief of the research. The comparative study of the previous researches is shown below:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Authors** | **Year** | **Title** | **Brief** | **Algorithms** | **DOI** |
| Joshi & Risodkar (2018) | 2018 | Deep Learning-Based Gujarati Handwritten Character Recognition | * They have proposed the method to detect handwritten text using deep learning. * they have applied the text images that contain a different combination of texts and have extracted the features from it. * They have combined NNC & KNN and checked the accuracy of detection of handwritten text image. * They have got 80% accuracy in the detection of handwritten texts using the Neural Network Classifier. | Neural Network Classifier (NNC) and K-Nearest Neighbour Classifier (KNN) | 10.1109/ICACCT.2018.8529410 |
| Borad, Dethaliya & Mehta (2020) | 2020 | Augmentation based Convolutional Neural Network for recognition of Handwritten Gujarati Characters | * They have proposed the method for the detection of Handwritten Characters using deep learning algorithms. * They have collected the image database for Handwritten Character Recognition. * They have applied Convolution Neural Network to the images of the collected database. * They have used five measures for image preprocessing namely Segmentation, Equalization, Skeletonization, Dilation, and Merging. * They have got average of 92% accuracy for handwritten text recognition. | Convolutional Neural Network | 10.1109/INOCON50539.2020.9298192 |
| Yin et al. (2019) | 2019 | Deep Learning-Aided OCR Techniques for Chinese Uppercase Characters in the Application of Internet of Things | * They have proposed the framework for the detection of handwritten text using Optical Character Recognition technology. * They have collected the database for OCR to train the deep learning algorithm that is Convolutional Neural Network. * They have extracted the features from the text image to facilitate the analysis and detection. * Finally, they have got 96% accuracy with an error rate of 1.26%. | Convolutional Neural Network | 10.1109/ACCESS.2019.2909401 |
| Sahu et al. (2017) | 2017 | Personalized Hand Writing Recognition Using Continued LSTM Training | * They have proposed the framework for the recognition of handwritten text using the deep learning algorithm. * They have used a very less number of images to train the algorithms. * They have applied RNN-LSTM and BLSTM on the collected text image data and trained & tested those. * Finally, they have got the highest accuracy using Bidirectional Long Short Term Memory by 92%. | Recurrent Neural Network and Long Short Term Memory (RNN-LSTM) and the Bidirectional Long Short Term Memory (BLSTM) | 10.1109/ICDAR.2017.44 |

# 

# 3 Scope, Objectives and Risk

## 3.1 Scope

The scope of this dissertation is to detect handwritten text using the deep learning algorithm. In the review of the previous researches, it has been found that different researchers have applied different deep learning algorithms to detect handwritten text. Some of the found algorithms are Convolutional Neural Network, Bidirectional Long Short Term Memory, Recurrent Neural Network and Long Short Term Memory (Yin, et al., 2019). The researchers have applied feature extraction procedures also to extract the important features from ten text images like segmentation, edge detection, pattern identification etc. all those have been done to improve the accuracy of detection. So, the main scope that is found and can be done is to improve the accuracy of recognition of handwritten text by extracting the important features and by identifying the pattern so that the text can be recognized irrespective of the pattern applied (Joshi & Risodkar, 2018). The summarization of available scores are as follows:

1. To scan text image which will be subjected for detection.
2. To extract features from the text image.
3. To identify the patterns of the text in different shapes.
4. To identify the good previous models that have achieved higher accuracy and precision.
5. To apply the deep learning algorithms to train it with the extracted features.
6. To record the performance of the applied algorithms and to compare the accuracy.
7. To propose the framework using the best model concerning accuracy.

## 

## 3.2 Objectives

The aim and the research questions have been discussed earlier for this dissertation and execution. The primary aim is to detect the handwritten text and the secondary aim is to minimize the error in the detection (Razavian, Azizpour, & Carlsson, 2014). Now, to fulfil the aim and to address the research questions, the following objectives have been taken concerning the scopes that have been mentioned in the last subsection. The objectives are discussed below:

1. To review the previous research and to identify the good models that have been proposed by the former researchers.
2. To conduct a comparative study regarding previous works to identify the god framework concerning accuracy and precision of the detection.
3. To select the deep learning algorithms for the detection of handwritten text.
4. To collect the data that will contain the handwritten text of different format and shape.
5. To extract features from the database to identify the pattern of the text.
6. To train the deep learning algorithms using the training set of the image database.
7. To test the algorithms using the test set of the image database.
8. To compare the accuracy and find the best algorithm.
9. To validate the best algorithm by detecting the images in real-time.
10. To record the classification matrices of the model regarding accuracy, precision, recall and f1-score.
11. To compare the previous best model with the present and proposed model for the detection of handwritten text.
12. To design and propose the framework for the detection of handwritten text.

## 

## 3.3 Risk Factors

The present research will be conducted based on the collected or secondary database of text images. Thus, certain risk factors will be involved in the research and those will be discussed as follows:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Risk Name | Risk Aspect | Process of Managing the Risk | Risk Score  (Out of 10) | Risk Strength |
| Data Security | The image database will be collected from any open source repository and will be utilized for the detection procedure. If certain data will be hampered due to mismanagement, the detection procedure may go the wrong way. | The risk can be minimized or reduced by applying the data protection strategy so that the data loss will not have occurred. | **7** | **High** |
| Data Source | The data will be collected from any open-source repository. However, not all database is good for this purpose because of the text image quality | This risk can be minimized by the selection of the data source so that the detection can be done with effectiveness | **6** | **High** |
| Data Understanding and Utilization | The database needs to be understood first as not all database will be available in the same way or orientation of images. If the project will be processed without understanding the data, it may produce the wrong result. | This risk can be resolved by thoroughly studying the data and the types of images that have been stored there. | **5** | **Medium** |
| Understanding methodology | The research will be focused on the method through which the entire execution will be done. So, if the methodology will be wrong, the procedure of detection will be wrong. | The risk factor can be managed by understanding the data and by focusing on the previous researches | **7** | **High** |
| Financial Risk | There is no such risk found | N/A | **3** | **Low** |
| Social Risk | There is no such risk found | N/A | **3** | **Low** |
| Environmental Risk | There is no such risk found | N/A | **3** | **Low** |

# 4 Ethics, Legal, Social, Security and Professional Issues

The issues considering the ethics, legal cases, security etc. will be discussed in this section below.

## 4.4 Ethics Issues

The previous researches will play a key role in identifying and understanding the methods and algorithms to be used. So, before proceeding to the research and execution, the model and the deep learning algorithms will be studied and focused on making the detection procedure with effectivity (Zhang, 2020).

## 4.2 Social Issues

The physical involvement will not be there for the collection of data. The data will be collected from the open-source repository and so, the survey for taking the pictures of the handwritten text will not be applied to the project. Thus, there will be no possibility of social issues objection in doing so (Xiao, 2019).

## 4.3 Security Issues

The data and the images will be collected only concerning the handwritten text and no other images will be involved in this research. Additionally, the data will be collected from a well-known repository like Kaggle and so, there will be no possibility for external security issues like the attack of malware which can be seen while visiting suspicious websites (Shi, Xu, Pan, Yan, & Zhang, 2018).

## 4.4 Professional Issues

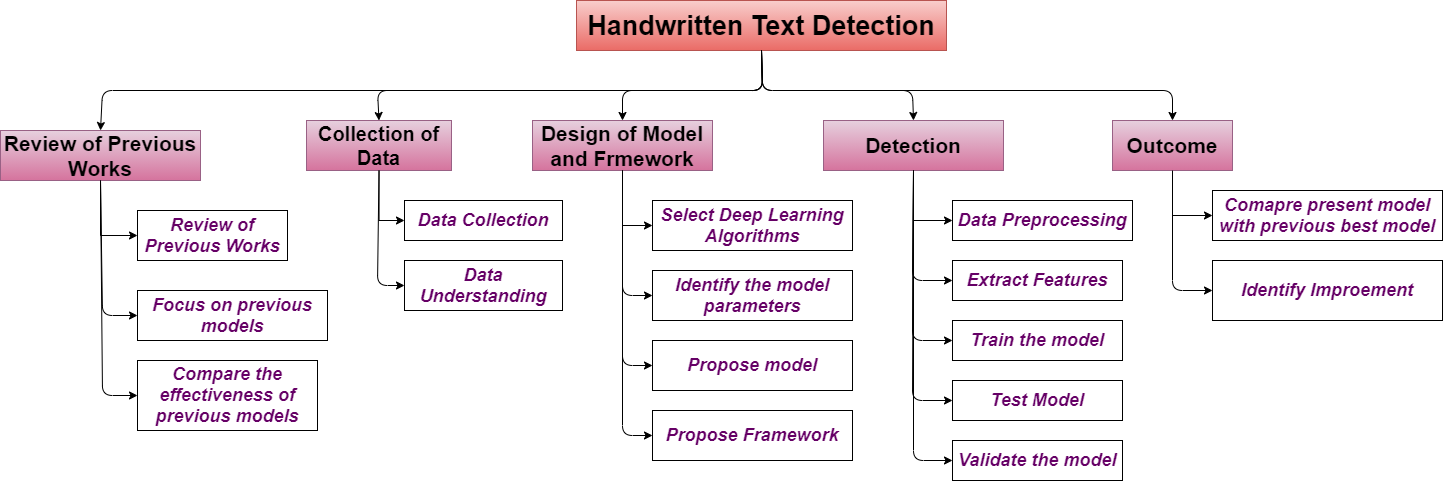
The dissertation will be conducted by reviewing the previous researches and by emphasizing the previous models. In this context, the idea of the models will be applied in the dissertation and at all of those places, the citation will be provided for those research papers from where the idea has been taken (Shi, et al., 2019).

# 

# 5 Schedule of Activities

## 5.1 Work Breakdown Structure

The breakdown structure of work for this dissertation is shown below:



**Fig-3: Breakdown of Work**

## 

## 5.2 List of Tasks to be Done.

The dissertation will be started on <place your date> and will be completed tentatively <place your date>. The list of tasks that will be performed in this dissertation to detect the object is as follows:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Task Id | Name of Tasks | Duration | Start date | End Date |
| 1 | **Review of Previous Works** | **15** | **17/01/2022** | **31/01/2022** |
| 1.1 | Review of Previous Works | 10 | **17/01/2022** | 26/01/2022 |
| 1.2 | Focus on previous models | 3 | 27/01/2022 | 29/01/2022 |
| 1.3 | Compare the effectiveness of previous models | 2 | 30/01/2022 | 31/01/2022 |
| 2 | **Collection of Data** | **5** | **01/02/2022** | **05/02/2022** |
| 2.1 | Data Collection | 3 | 01/02/2022 | 03/02/2022 |
| 2.2 | Data Understanding | 2 | 04/02/2022 | 05/02/2022 |
| 3 | **Design of Model and Framework** | **15** | **06/02/2022** | **21/02/2022** |
| 3.1 | Select Deep Learning Algorithms | 4 | 06/02/2022 | 09/02/2022 |
| 3.2 | Identify the model parameters | 3 | 10/02/2022 | 12/02/2022 |
| 3.3 | Propose model | 3 | 13/02/2022 | 16/02/2022 |
| 3.4 | Propose framework | 5 | 17/02/2022 | 21/02/2022 |
| 4 | **Detection** | **35** | **22/02/2022** | **28/03/2022** |
| 4.1 | Data Pre-processing | 8 | 22/02/2022 | 01/03/2022 |
| 4.2 | Extract Features | 8 | 02/03/2022 | 09/03/2022 |
| 4.3 | Train the model | 8 | 10/03/2022 | 17/03/2022 |
| 4.4 | Test Model | 7 | 18/03/2022 | 24/03/2022 |
| 4.5 | Validate the model | 4 | 25/03/2022 | 28/03/2022 |
| 5 | **Outcome** | **20** | **29/03/2022** | **17/04/2022** |
| 5.1 | Compare the present model with the previous best model | 10 | 29/03/2022 | 07/04/2022 |
| 5.2 | Identify Improvement | 10 | 08/04/2022 | 17/04/2022 |

## Gantt Chart

The Gantt Chart for this thesis is shown below:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Day**  **1-10** | **Day 11-25** | **Day**  **26-30** | **Day**  **31-45** | **Day**  **46-70** | **Day**  **71-90** |
| **Introduction** |  |  |  |  |  |  |
| **Review of the previous works and Compare models** |  |  |  |  |  |  |
| **Data Collection for handwritten text image** |  |  |  |  |  |  |
| **Select algorithms of deep learning** |  |  |  |  |  |  |
| **Analysis and Detection of handwritten text** |  |  |  |  |  |  |
| **Outcomes and compare the result with previous models** |  |  |  |  |  |  |

**Fig-4: Gantt Chart**

# 6 References

Aly, S., & Almotairi, S. (2020). Deep Convolutional Self-Organizing Map Network for Robust Handwritten Digit Recognition. *IEEE Access*, 8.

Borad, P., Dethaliya, P., & Mehta, A. (2020). Augmentation based Convolutional Neural Network for recognition of Handwritten Gujarati Characters. *IEEE International Conference for Innovation in Technology (INOCON)*, 122-126.

Cai, J., Chang, O., Tang, X.-L., Xue, C., & Wei, C. (2018). Facial Expression Recognition Method Based on Sparse Batch Normalization CNN. *37th Chinese Control Conference (CCC)*, 117-123.

Ejaz, M. S., & Islam, M. R. (2019). Masked Face Recognition Using Convolutional Neural Network. *International Conference on Sustainable Technologies for Industry 4.0 (STI)*, 1-6.

Fernández, D., Porras, F., Gilman, R., Mondonedo, M., & Zimic, S. P. (2020). A Convolutional Neural Network for gaze preference detection: A potential tool for diagnostics of autism spectrum disorder in children. *arXiv preprint arXiv:2007.14432*, 220-227.

Ho, M.-J., Lin, Y.-C., Hsu, H.-C., & Sun, T.-Y. (2019). An Efficient Recognition Method for Watermelon using Faster R-CNN with Post-Processing. *Innovation Communication and Engineering (ICICE) 2019 8th International Conference*, 86-89.

Huang, C.-W., & Narayanan, S. S. (2017). Deep convolutional recurrent neural network with attention mechanism for robust speech emotion recognition. *IEEE International Conference on Multimedia and Expo (ICME)*, 583-588.

Huang, J., Zhou, W., Li, H., & Li, W. (2019). Attention-Based 3D-CNNs for Large-Vocabulary Sign Language Recognition. *IEEE Transactions on Circuits and Systems for Video Technology, 29*(9), 2822 - 2832.

Joshi, D. S., & Risodkar, Y. R. (2018). Deep Learning Based Gujarati Handwritten Character Recognition. *International Conference On Advances in Communication and Computing Technology (ICACCT)*, 563-566.

Lee, S.-M., Yoon, S. M., & Cho, H. (2017). Human activity recognition from accelerometer data using Convolutional Neural Network. *IEEE International Conference on Big Data and Smart Computing (BigComp)*, 668-672.

Razavian, A. S., Azizpour, H., & Carlsson, J. S. (2014). CNN features Off-the-shelf: An astounding baseline for recognition. *IEEE Conf. Comput. Vis. Pattern Recognit. Workshops*, 806–813.

Sahu, P. P., Singh, V., Kiran, I., Veera, V., Abhinav, T., Vijay, A., & Venkatesan, S. M. (2017). Personalized Hand Writing Recognition Using Continued LSTM Training. *14th IAPR International Conference on Document Analysis and Recognition (ICDAR)*, 218-223.

Shi, Z., Hao, H., Zhao, M., Feng, Y., He, L., & Suzuki, Y. W. (2019). A deep CNN based transfer learning method for false positive reduction. *Multimedia Tools Appl*, 1017–1033.

Shi, Z., Xu, M., Pan, Q., Yan, B., & Zhang, H. (2018). LSTM-based Flight Trajectory Prediction. *Neural Networks (IJCNN) 2018 International Joint Conference*, 1-8.

Xiao, X. (2019). An Image-Inspired and CNN-Based Android Malware Detection Approach. *34th IEEE/ACM International Conference on Automated Software Engineering (ASE)*, 1259-1261.

Yin, Y., Zhang, W., Hong, S., Yang, J., Xiong, J., & Gui, G. (2019). Deep Learning-Aided OCR Techniques for Chinese Uppercase Characters in the Application of Internet of Things. *IEEE Access* , 47043 - 47049.

Zhang, J. (2020). DeepMal: A CNN-LSTM Model for Malware Detection Based on Dynamic Semantic Behaviours. *International Conference on Computer Information and Big Data Applications (CIBDA)*, 1278-1281.

# 7 Appendix A – Ethics Form

Graphical user interface, text, application

Description automatically generated

Graphical user interface, text, application

Description automatically generatedText

Description automatically generatedGraphical user interface, application

Description automatically generatedGraphical user interface, text, application

Description automatically generatedGraphical user interface, text, application

Description automatically generatedGraphical user interface, application

Description automatically generatedGraphical user interface, text, application

Description automatically generatedA picture containing graphical user interface

Description automatically generated

Graphical user interface, text

Description automatically generated