COMPS461F Data Science Project

Topic:

Deep Learning-Based System for Handwritten Character Recognition in Exam Papers

By Group 3

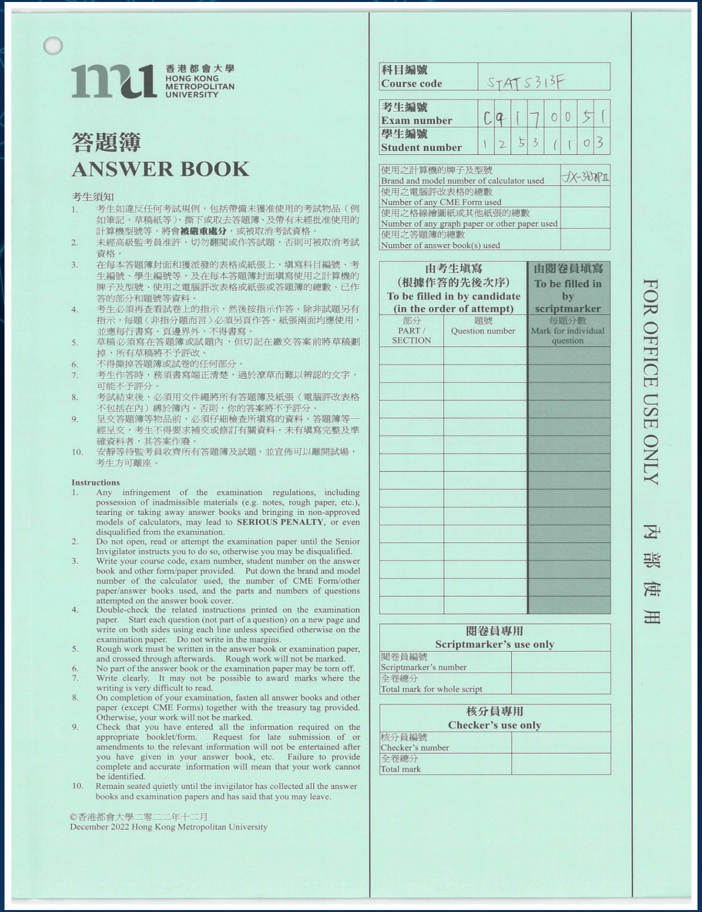
Chu Ying Ying, Tam Oi Laam, Tsao Sai Chak,Wong Hok Man,
Cheung Yau Cheuk, Lung Kwan Chak

Supervisor: Dr. Jimmy Kang

Date: 17 APR 2025

1. Introduction

1.1 Background


香港都會大學
HONG KONG METROPOLITAN UNIVERSITY

答題簿
ANSWER BOOK

考生須知
1. 考生如違反任何考試規例，包括帶備未獲准使用的考試物品（例如筆記、草稿紙等）、撕下或取去答題簿、及帶有未經批准使用的計算機型號等，將會**被嚴重處分**，或被取消考試資格。
2. 未經高級監考員准許，切勿翻閱或作答試題，否則可被取消考試資格。
3. 在每本答題簿封面和獲派發的表格或紙張上，填寫科目編號、考生編號、學生編號等，及在每本答題簿封面填寫使用之計算機的牌子及型號、使用之電腦評改表格或紙張或答題簿的總數、已作答的部分和題號等資料。
4. 考生必須再查看試卷上的指示，然後按指示作答。除非試題另有指示，每題（非指分題而言）必須另頁作答。紙張兩面均應使用，並應每行書寫。頁邊界外，不得書寫。
5. 草稿必須寫在答題簿或試題內，但切記在繳交答案前將草稿劃掉，所有草稿將不予評改。
6. 不得撕掉答題簿或試卷的任何部分。
7. 考生作答時，務須書寫端正清楚，過於潦草而難以辨認的文字，可能不予評分。
8. 考試結束後，必須用文件繩將所有答題簿及紙張（電腦評改表格不包括在內）縛於簿內。否則，你的答案將不予評分。
9. 呈交答題簿等物品前，必須仔細檢查所填寫的資料。答題簿等一經呈交，考生不得要求補交或修訂有關資料。未有填寫完整及準確資料者，其答案作廢。
10. 安靜等待監考員收齊所有答題簿及試題，並宣佈可以離開試場，考生方可離座。

Instructions
1. Any infringement of the examination regulations, including possession of inadmissible materials (e.g. notes, rough paper, etc.), tearing or taking away answer books and bringing in non-approved models of calculators, may lead to **SERIOUS PENALTY**, or even disqualified from the examination.
2. Do not open, read or attempt the examination paper until the Senior Invigilator instructs you to do so, otherwise you may be disqualified.
3. Write your course code, exam number, student number on the answer book and other form/paper provided. Put down the brand and model number of the calculator used, the number of CME Form/other paper/answer books used, and the parts and numbers of questions attempted on the answer book cover.
4. Double-check the related instructions printed on the examination paper. Start each question (not part of a question) on a new page and write on both sides using each line unless specified otherwise on the examination paper. Do not write in the margins.
5. Rough work must be written in the answer book or examination paper, and crossed through afterwards. Rough work will not be marked.
6. No part of the answer book or the examination paper may be torn off.
7. Write clearly. It may not be possible to award marks where the writing is very difficult to read.
8. On completion of your examination, fasten all answer books and other paper (except CME Forms) together with the treasury tag provided. Otherwise, your work will not be marked.
9. Check that you have entered all the information required on the appropriate booklet/form. Request for late submission of or amendments to the relevant information will not be entertained after you have given in your answer book, etc. Failure to provide complete and accurate information will mean that your work cannot be identified.
10. Remain seated quietly until the invigilator has collected all the answer books and examination papers and has said that you may leave.

科目編號 Course code	STATS313F
考生編號 Exam number	C 9 1 7 0 0 5 1
學生編號 Student number	1 2 5 3 1 1 0 3

使用之計算機的牌子及型號 Brand and model number of calculator used	FX-3650PII
使用之電腦評改表格的總數 Number of any CME Form used	
使用之格線繪圖紙或其他紙張的總數 Number of any graph paper or other paper used	
使用之答題簿的總數 Number of answer book(s) used	

由考生填寫 (根據作答的先後次序) To be filled in by candidate (in the order of attempt)		由閱卷員填寫 To be filled in by scriptmarker
部分 PART / SECTION	題號 Question number	每題分數 Mark for individual question

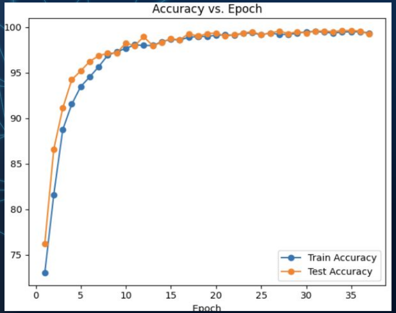
閱卷員專用 Scriptmarker's use only	
閱卷員編號 Scriptmarker's number	
全卷總分 Total mark for whole script	

核分員專用 Checker's use only	
核分員編號 Checker's number	
全卷總分 Total mark	

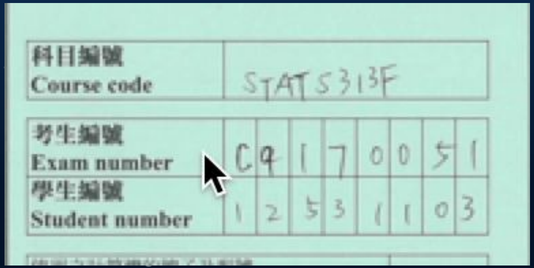
FOR OFFICE USE ONLY 內部使用

- High variability, noise, and low contrast in exam papers cause manual data entry errors
- Manual region selection on exam papers
- Improving recognition accuracy can enhance efficiency in educational data entry.

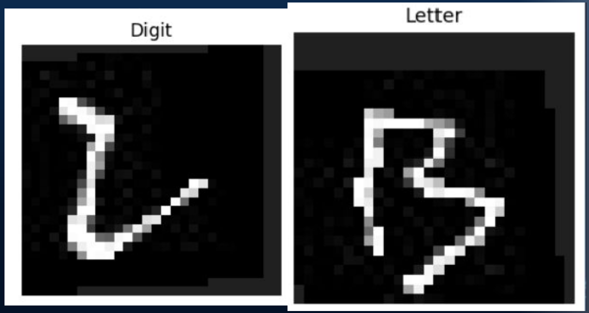
1.2 Objectives



All model accuracy >= 95%



Enable the system to crop specific regions(e.g., Course code, Exam number, Student number.)



Recognize English letters (A-Z) and digits (0-9)

Develop a GUI application for easy PDF processing.

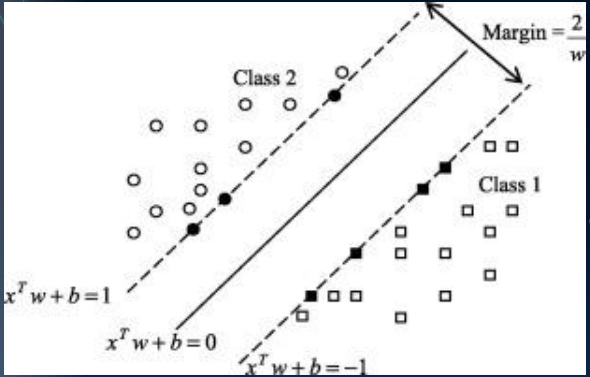
Export the result to Excel

2. Literature Review

2.1 Evolution of Handwritten Character Recognition (HCR)

1979s

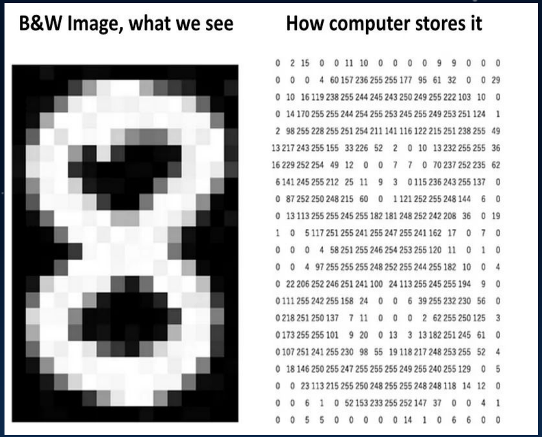
Traditional Methods - SVM



Support Vector Machine (SVM) classification visualization (Boukharouba & Bennia, 2015)

1986s

Traditional Methods -HOG

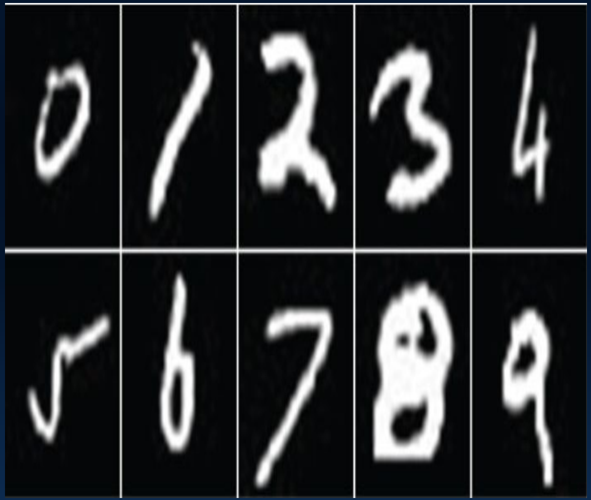


Ai, S. (2022, August 13). Histogram of Oriented Gradients (HOG) — Simplest Intuition. *Medium.* https://medium.com/@skillcate/histogram-of-oriented-gradients-hog-simplest-intuition-2392995f8010

2.1 Evolution of Handwritten Character Recognition (HCR)

1998s

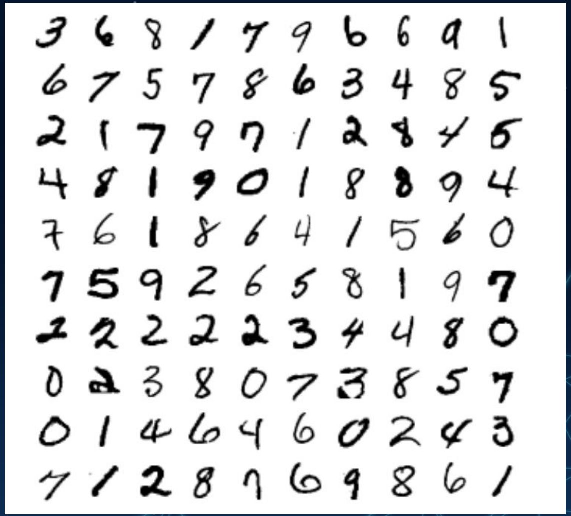
MNIST Dataset



Ahmed, S. S., Mehmood, Z., Awan, I. A., & Yousaf, R. M. (2023). A novel technique for handwritten digit recognition using deep learning. Journal of Sensors, 2023, 1–15. https://doi.org/10.1155/2023/2753941

2017s

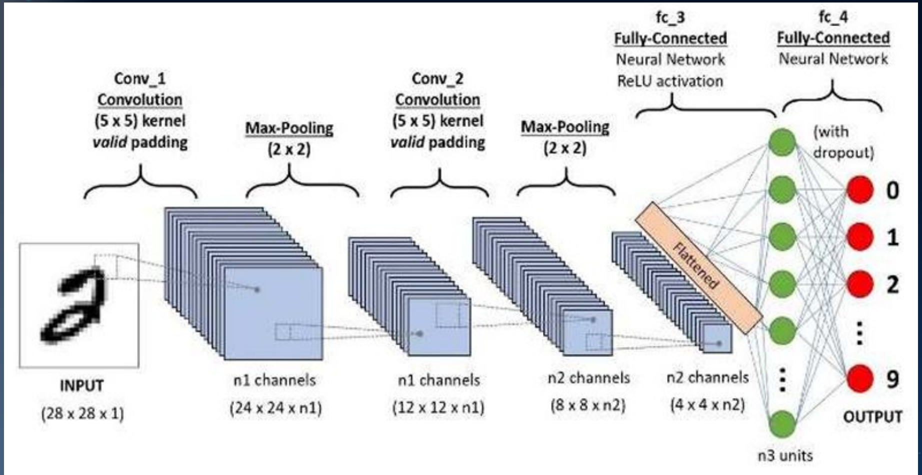
EMNIST Dataset



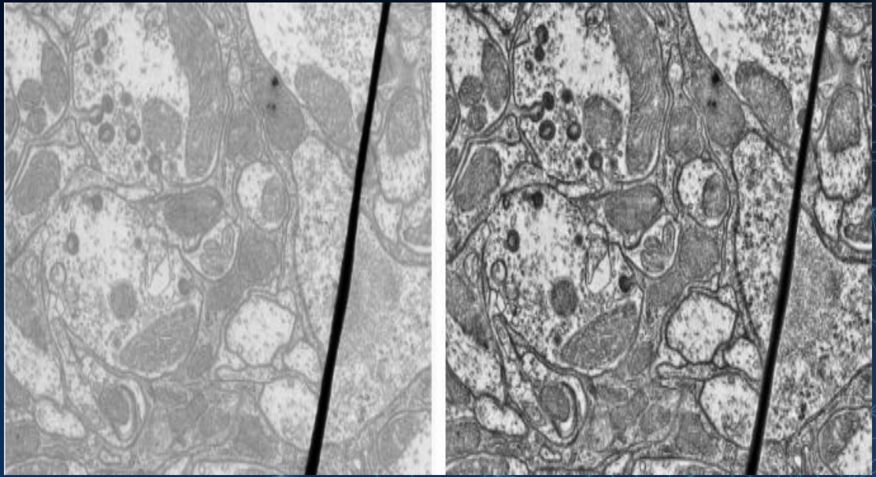
Cohen, G., Afshar, S., Tapson, J., & van Schaik, A. (2017). EMNIST: an extension of MNIST to handwritten letters.

2.2 Deep Learning and Preprocessing in HCR

- CNNs enabled automatic feature extraction, high accuracy (LeCun et al., 1998).
- Dropout and ReLU improved CNN robustness (Krizhevsky et al., 2012; Park & Kwak, 2017).
- Preprocessing (CLAHE, adaptive thresholding) enhances noisy exam paper images (Nockels et al., 2024).



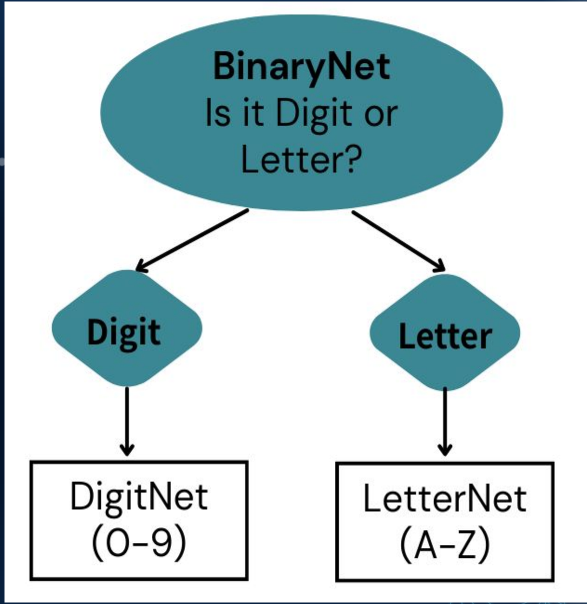
Khandokar, I., Hasan, M., Ernawan, F., Islam, S., & Kabir, M. N. (2021). Handwritten character recognition using convolutional neural network. Journal of Physics Conference Series, 1918(4), 042152. https://doi.org/10.1088/1742-6596/1918/4/042152



刘斯宁. (n.d.). Image enhancement - CLAHE. 知乎专栏. https://zhuanlan.zhihu.com/p/150381937

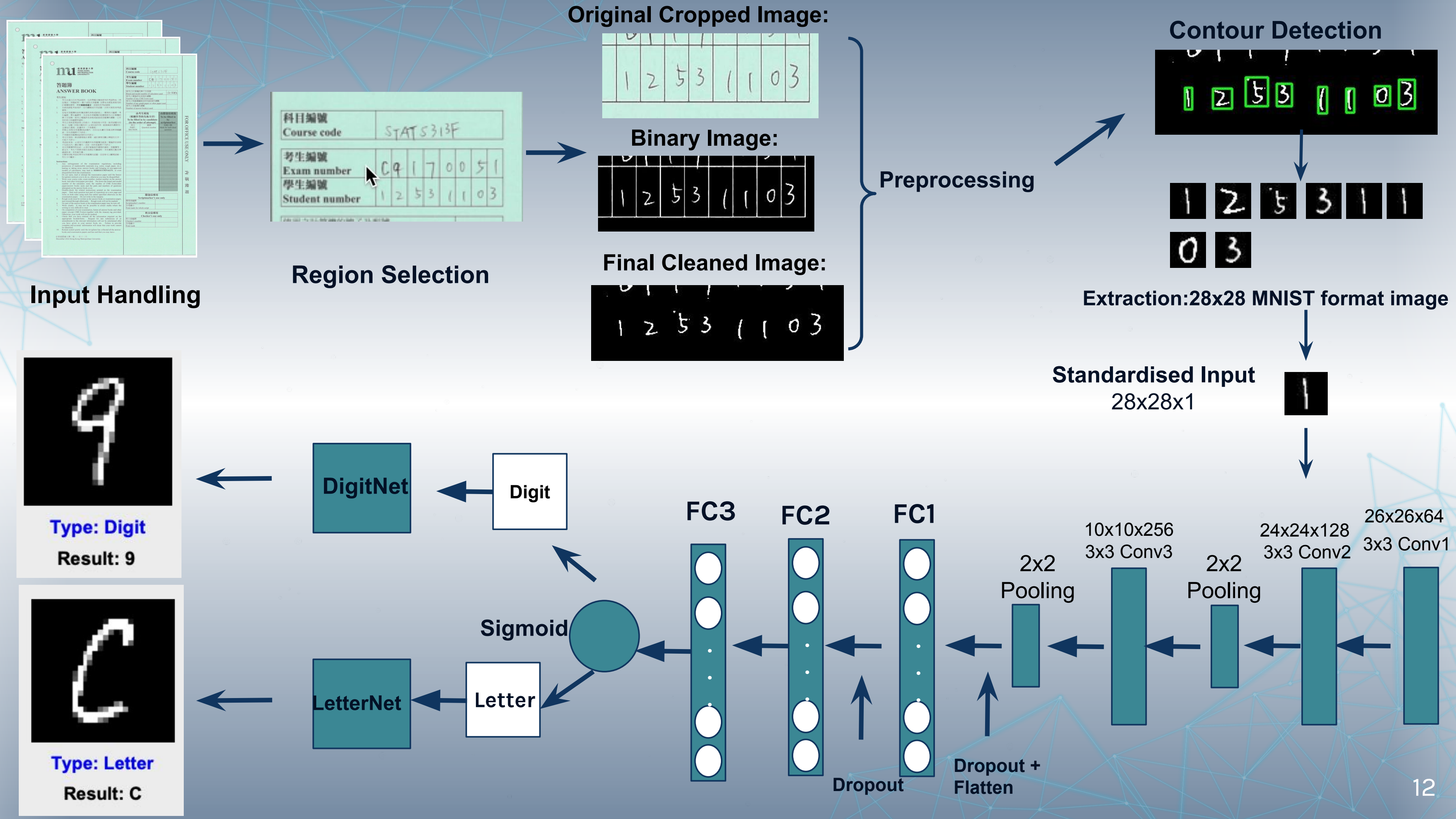
2.3 Hierarchical Classification and Challenges

- Hierarchical classification (BinaryNet → DigitNet/LetterNet) improves accuracy (Yu & Zhu, 2020).
- Achieved 99.58% accuracy in distinguishing digits vs. letters in this study.
- **Challenges**: handwriting variability, noisy inputs, need for user-friendly GUIs (Iwana & Uchida, 2021).

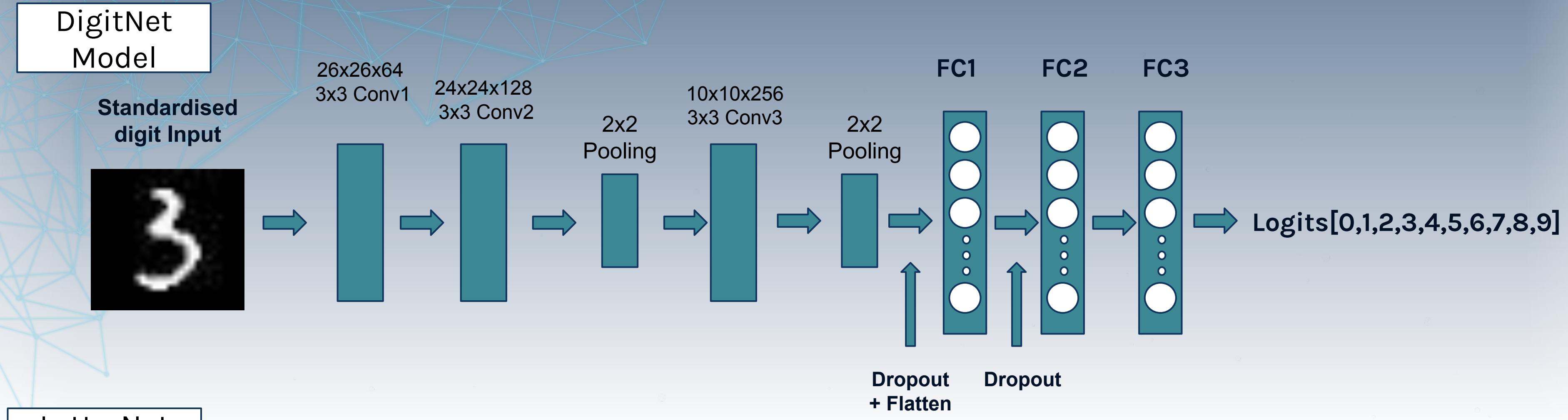


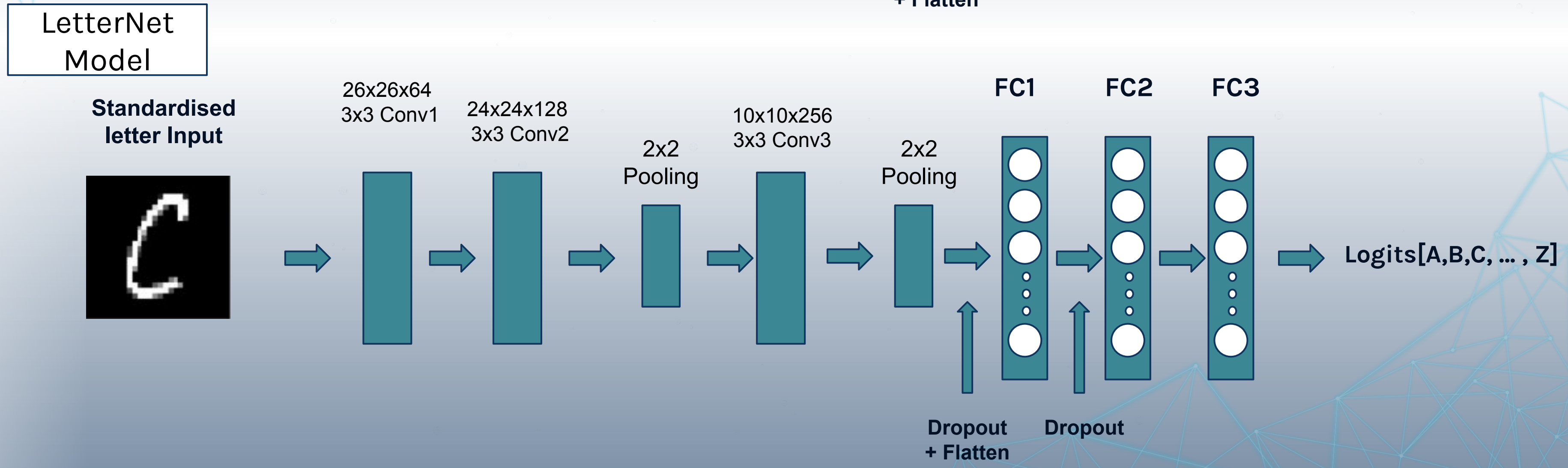
3. Method

3.1 Whole System Flow


Input Handling
ANSWER BOOK
科目編號
Course code
STAT S313F
考生編號
Exam number
學生編號
Student number
Region Selection
Original Cropped Image:
Binary Image:
Final Cleaned Image:
Preprocessing
Contour Detection
Extraction:28x28 MNIST format image
Standardised Input
28x28x1
26x26x64
3x3 Conv1
24x24x128
3x3 Conv2
2x2
Pooling
10x10x256
3x3 Conv3
2x2
Pooling
Dropout +
Flatten
FC1
Dropout
FC2
FC3
Sigmoid
Digit
DigitNet
Letter
LetterNet
Type: Digit
Result: 9
Type: Letter
Result: C

3.2 Detailed Model Architecture


DigitNet Model
Standardised digit Input
26x26x64 3x3 Conv1
24x24x128 3x3 Conv2
2x2 Pooling
10x10x256 3x3 Conv3
2x2 Pooling
FC1
FC2
FC3
Dropout + Flatten
Dropout
Logits[0,1,2,3,4,5,6,7,8,9]


LetterNet Model
Standardised letter Input
26x26x64 3x3 Conv1
24x24x128 3x3 Conv2
2x2 Pooling
10x10x256 3x3 Conv3
2x2 Pooling
FC1
FC2
FC3
Dropout + Flatten
Dropout
Logits[A,B,C, ... , Z]

3.3 Model Techniques

Loss Functioin:

Cross-entropy loss
Binary CrossEntropy Loss

Activation:

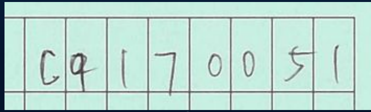
Sigmoid

Transformers

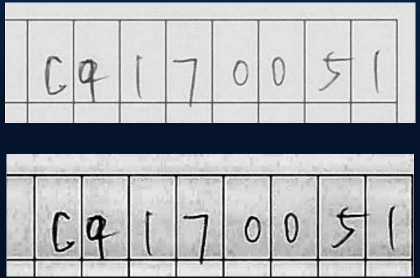
- **Preprocessing**
  - Grayscale
  - Resize: Scale images to 28x28 pixels
  - ToTensor: Convert images to PyTorch tensors
  - Normalize: mean=0.1307, std=0.3081
- **Augmentation**
  - Random affine transformation (e.g., ±5° rotation, 10% translation)
  - Color jitter (brightness/contrast tweaks, p=0.3)
  - Salt-and-pepper noise (p=0.2) to mimic real world distortions

3.4 Preprocessing

Original cropped image

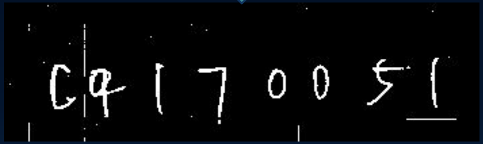


1



2

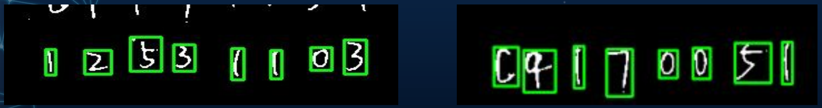
3



4

1. Converts PDFs to enhanced **grayscale** images using CLAHEA
2. Adaptive thresholding
3. Remove vertical and horizontal lines
4. Clean up noises

3.5 Contour Detection and Classification



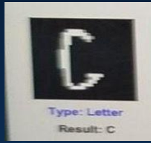
1.Finds character contours and merges nearby ones

2.Resizes regions to 28x28 (MNIST format)

BinaryNet
(Digit or Letter?)

DigitNet

LetterNet



3.Classifies using a two-step process: binary (digit/letter), then specific model.

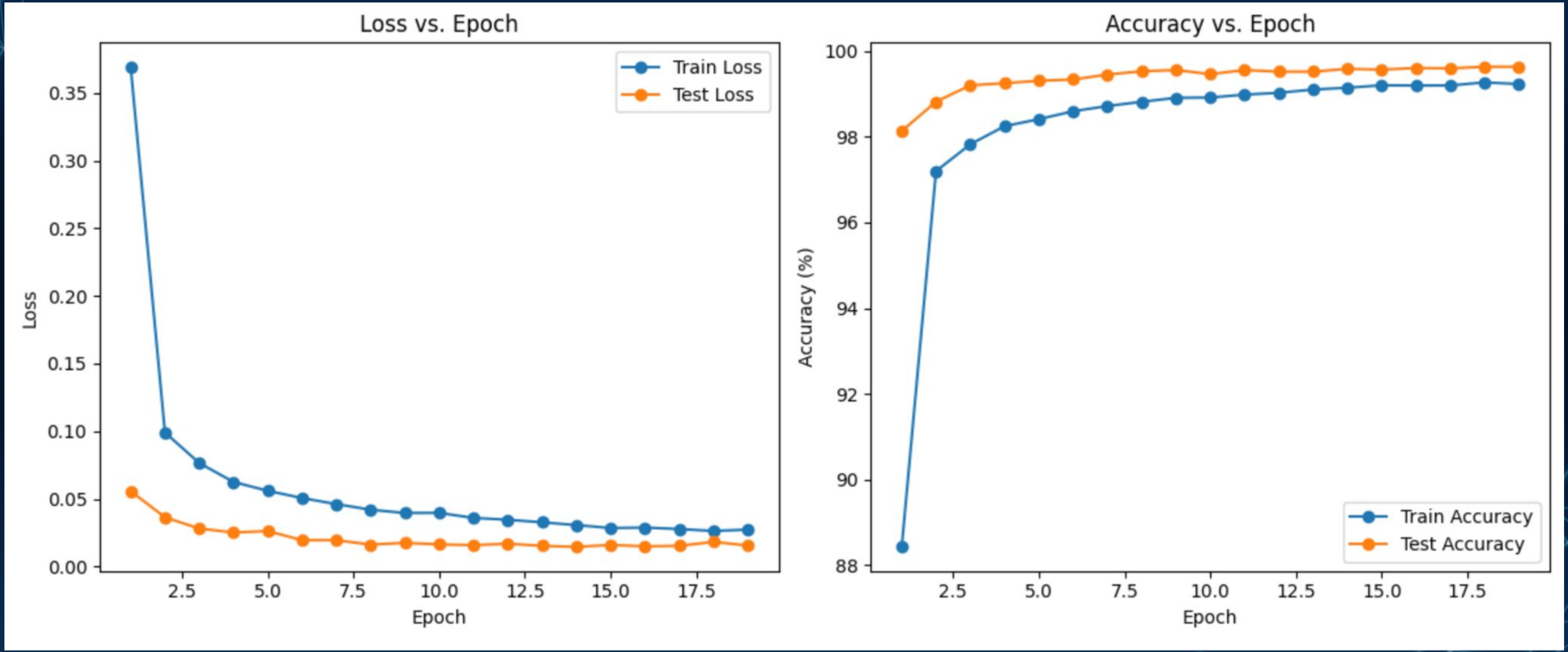
4. Demo

Scoring Criteria

Accuracy	Measures correct classification, targeting >95% for reliable exam paper processing.
Loss	Indicates training stability, aiming for low values (<0.03) to ensure convergence.

5. Results

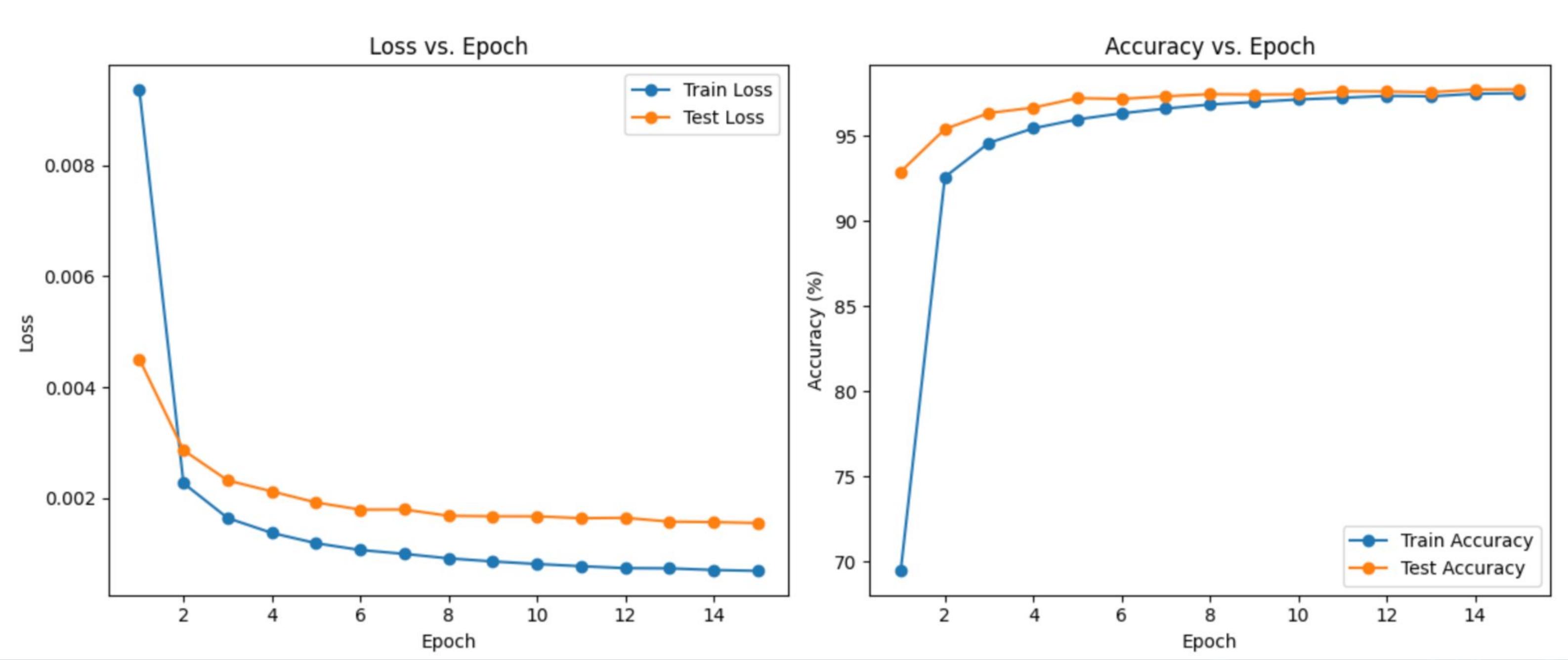
Model: DigitNet
Dataset: MNIST



Average loss: 0.0155, Accuracy: 0.99 (99%)

5. Results

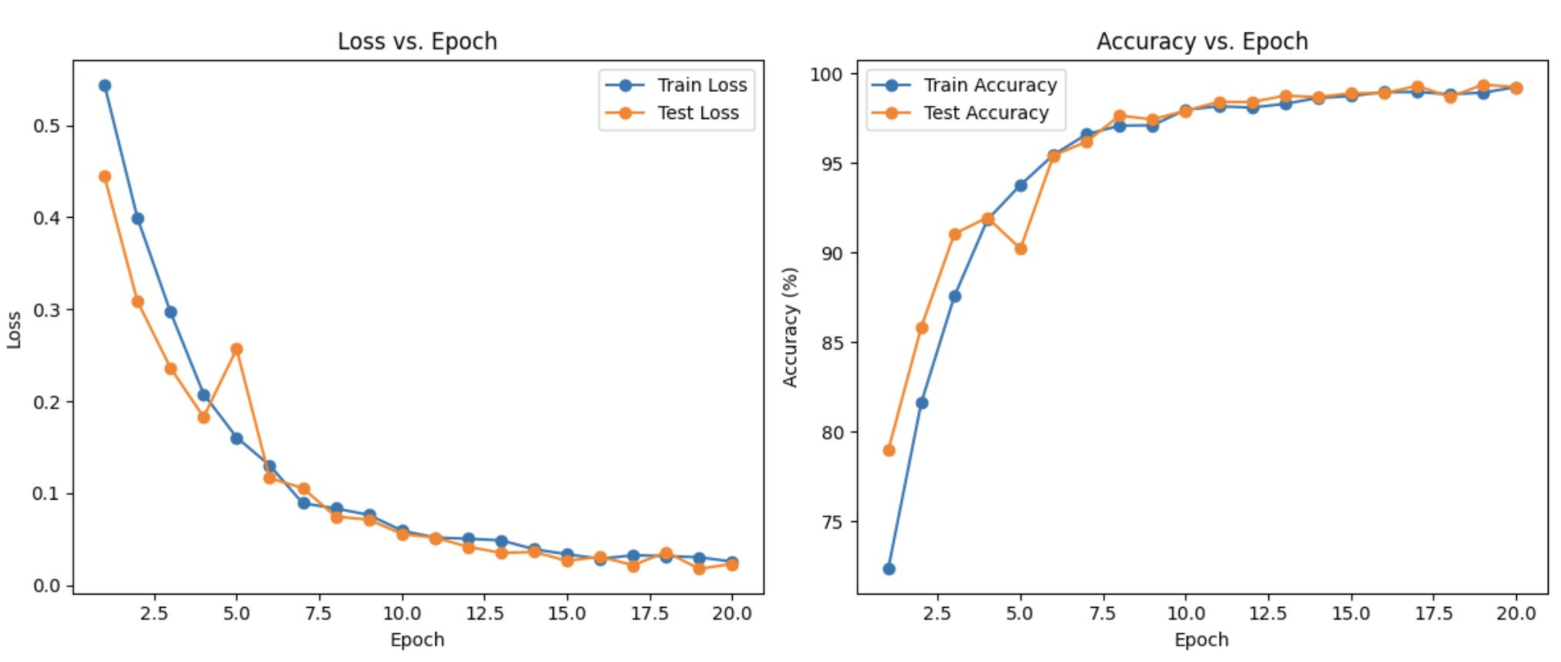
Model: LetterNet
Dataset: EMNIST



Average loss: 0.0015, Accuracy: 0.98 (98%)

5. Results

Model: BinaryNet
Dataset: EMNIST



Average loss: 0.0132, Accuracy: 0.99 (99%)

5. Results

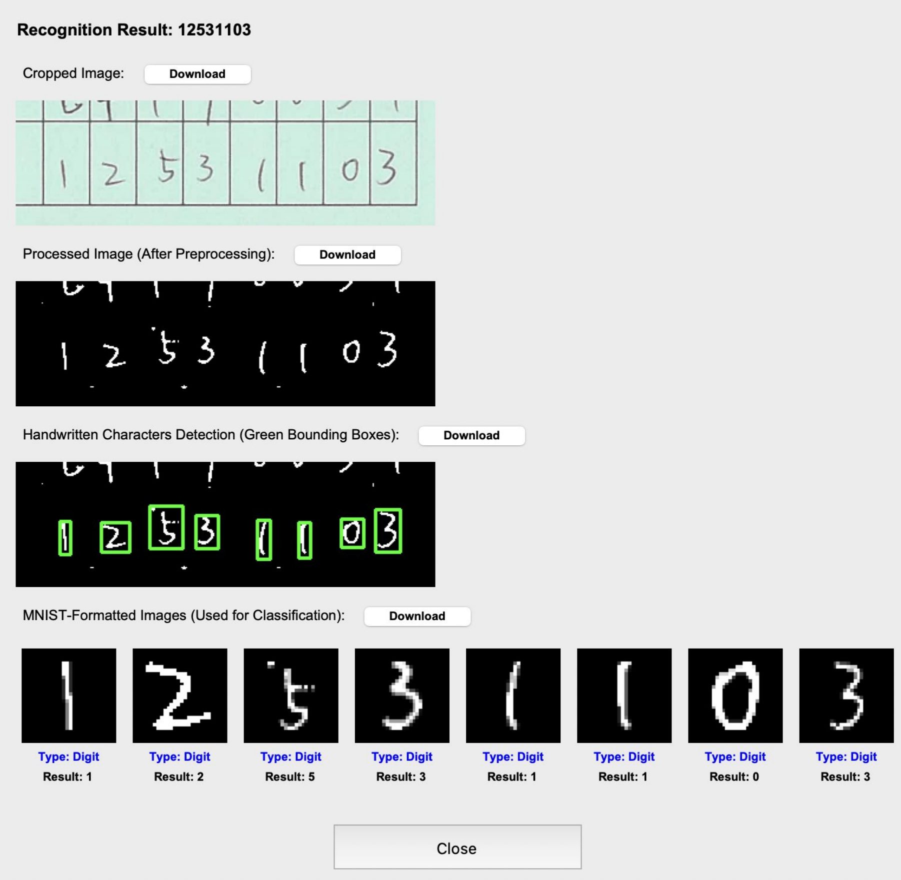
- All models (BinaryNet, DigitNet, LetterNet) exceed 95% accuracy with loss below 0.02.

- LetterNet has the lowest loss at 0.0015; BinaryNet and DigitNet lead in accuracy at 99%.

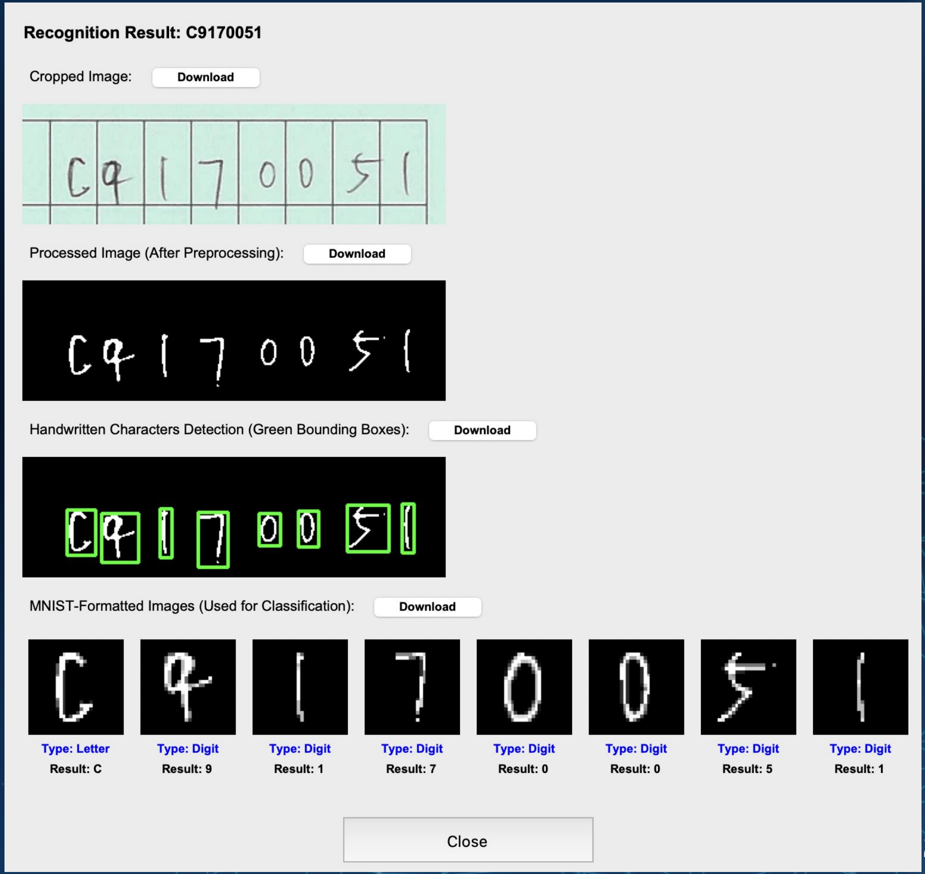
- Strong generalizability ensures reliability for exam paper processing.

5. Results

Student Number

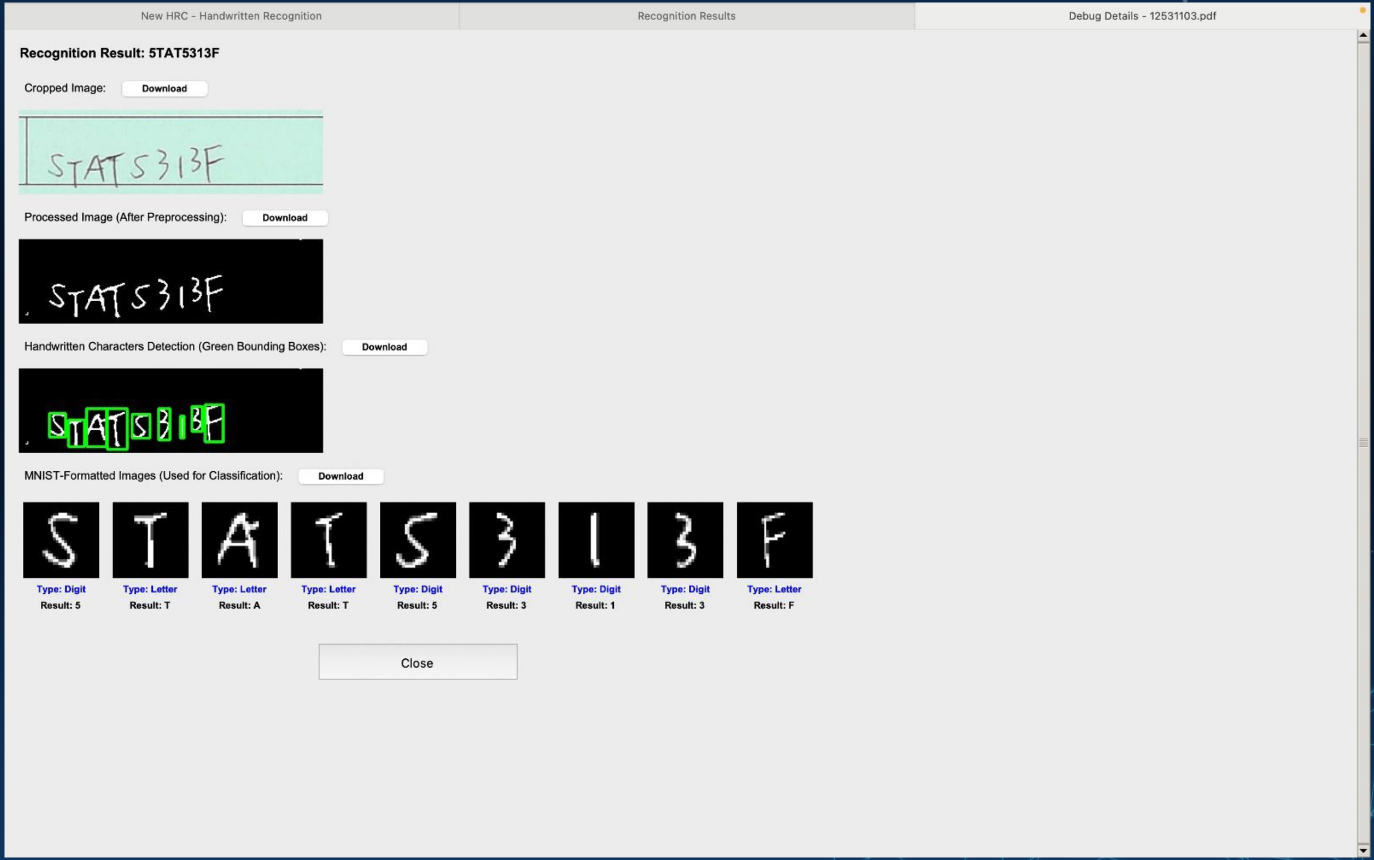


Exam Number



5. Results

Course Code



6. Discussions

Strengths	Challenges
- High accuracy (>95%) after model optimization. - User-friendly GUI for easy PDF uploads and region selection. - Robust preprocessing for noise handling.	- Initial deep learning model accuracy below 95% with softmax - Difficulty detecting course code "STAT S313F" due to noise and spacing. - Misclassification of letters as digits, e.g., "S" as "5."

6. Discussions

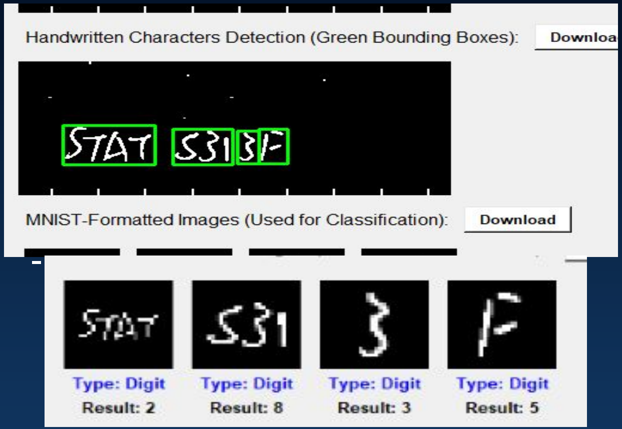
Improving Deep Learning Model Accuracy

- Switch to logits with cross-entropy loss and the accuracy boost above 95%

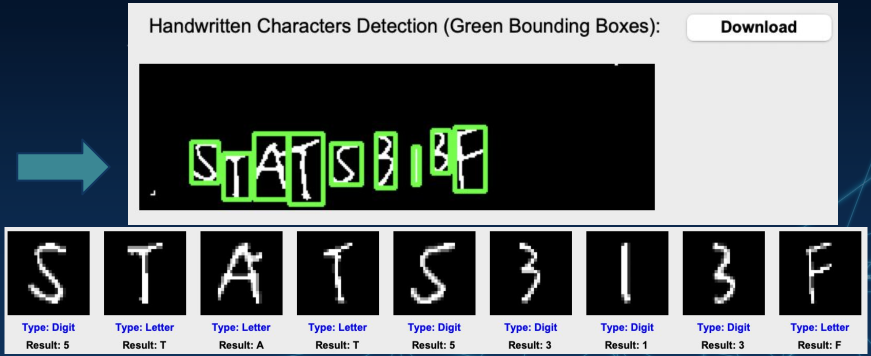
Enhancing Course Code Sequence Detection

- Refine contour detection to better handle spaces (e.g., in "STAT S313F") and apply advanced noise reduction techniques like Gaussian blur.

Before



After



6. Discussions

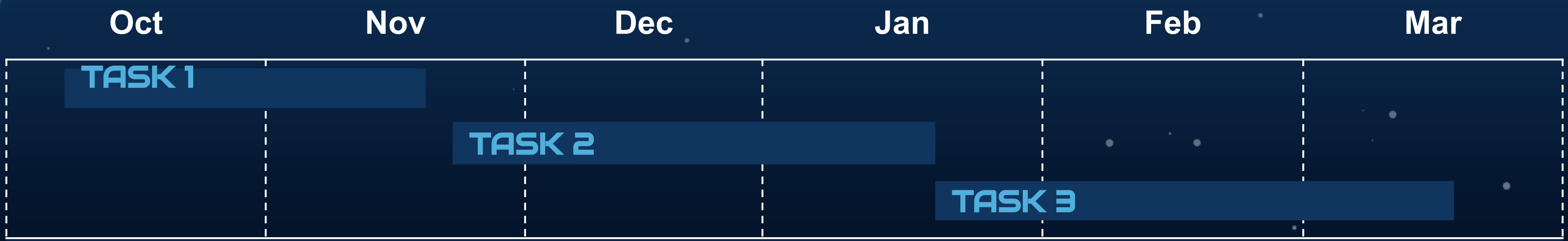
Reducing Misclassification Errors

- Retrain BinaryNet with more handwritten letter samples and add post-processing rules to enforce expected course code formats.

7. Conclusion

- Deep learning system for handwritten character recognition
- CNN models (BinaryNet, DigitNet, LetterNet) with >95% accuracy
- Robust preprocessing, user-friendly GUI, and Excel export
- Overcame challenges to meet all objectives
- Future goal: Automate cropping for faster workflows

8. Team Contribution



Task	Description	Members	Date	Status
Task 1	Literature review and dataset preparation	Tsao Sai Chak	Oct-Nov	Completed
Task 2	Model development	Wong Hok Man	Nov-Jan	Completed
Task 3	Image preprocessing and Contour Detection	Chu Ying Ying	Jan-Mar	Completed

Task	Description	Members	Date	Status
Task 4	GUI Development	Tam Oi Lam	Jan-Mar	Completed
Task 5	Classification and Result Handling	Cheung Yau Cheuk	Feb-April	Completed
Task 6	Presentation and Documentation	Lung Kwan Chak	April	Completed

References

Ahmed, S. S., Mehmood, Z., Awan, I. A., & Yousaf, R. M. (2023). A novel technique for handwritten digit recognition using deep learning. Journal of Sensors, 2023, 1–15. https://doi.org/10.1155/2023/2753941

Ai, S. (2022, August 13). Histogram of Oriented Gradients (HOG) — Simplest Intuition. Medium. https://medium.com/@skillcate/histogram-of-oriented-gradients-hog-simplest-intuition-2392995f8010

Boukharouba, A., & Bennia, A. (2015). Novel feature extraction technique for the recognition of handwritten digits. Applied Computing and Informatics, 11(1), 57-66. https://doi.org/10.1016/j.aci.2015.05.001

Cohen, G., Afshar, S., Tapson, J., & Van Schaik, A. (2017). EMNIST: Extending MNIST to handwritten letters. In *2017 International Joint Conference on Neural Networks (IJCNN)* (pp. 2921-2926). IEEE. https://doi.org/10.1109/IJCNN.2017.7966217

Grüning, T., Leifert, G., Strauß, T., Michael, J., & Labahn, R. (2019). A two-stage method for text line detection in historical documents. *International Journal on Document Analysis and Recognition (IJDAR)*, *22*(3), 285-302. https://doi.org/10.1007/s10032-019-00336-2

Iwana, B. K., & Uchida, S. (2021). An empirical survey of data augmentation for time series classification with neural networks. *Plos One*, *16*(7), Article e0254841. https://doi.org/10.1371/journal.pone.0254841

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. Neural Information Processing Systems, 25, 1097–1105. http://books.nips.cc/papers/files/nips25/NIPS2012_0534.pdf

Liu, S. (n.d.). Image enhancement - CLAHE. Zhihu. https://zhuanlan.zhihu.com/p/150381937

LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the* IEEE, *86*(11), 2278-2324. https://doi.org/10.1109/5.726791

Memon, J., Sami, M., Khan, R. A., & Uddin, M. (2020). Handwritten optical character recognition (OCR): A comprehensive systematic literature review (SLR). *IEEE Access*, *8*, 142642-142668. https://doi.org/10.1109/ACCESS.2020.3012542

Nockels, J., Gooding, P., & Terras, M. (2024). The implications of handwritten text recognition for accessing the past at scale. *Journal of Documentation*, *80*(7), 148-167. https://doi.org/10.1108/JD-09-2023-0183

Park, S., & Kwak, N. (2017). Analysis on the dropout effect in convolutional neural networks. In *Computer Vision—ACCV 2016: 13th Asian Conference on Computer Vision, Taipei, Taiwan, November 20-24, 2016, Revised Selected Papers, Part II 13* (pp. 189-204). Springer. https://doi.org/10.1007/978-3-319-54184-6_12

Yu, T., & Zhu, H. (2020). Hyper-parameter optimization: A review of algorithms and applications. *arXiv*. https://doi.org/10.48550/arXiv.2003.05689