Optical Character Recognition (OCR)

字元辨識流程說明

大綱

- 影像前處理 (Pre-processing)
  - ✓ Adaptive/Otsu thresholding, image inverse
- 影像分割 (Segmentation)
  - ✓ Connected component labelling
- 特徵擷取 (Feature Extraction)
  - ✓ Image Resize (Shrink)
- 樣式分類 (Pattern Classification)
  - ✓ Multilayer Perceptron (MLP)

流程簡介

原始影像

Image inverse
Adaptive/Otsu thresholding

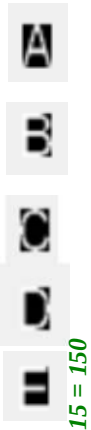
前處理

Connected component labelling

影像分割

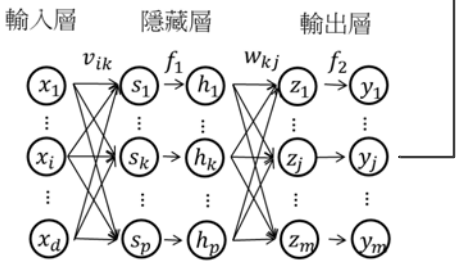
Resize to 10 x 15 pixel

特徵擷取



Input vector = 10 x 15 = 150

樣式分類

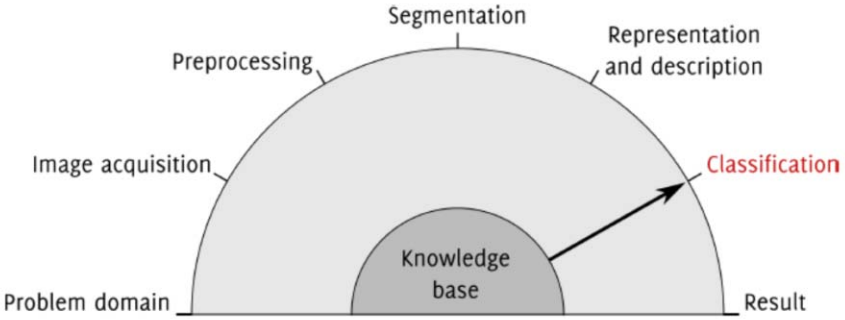


Multilayer perceptron (MLP)

Output vector = 16 bits Unicode

Unicode Text

A	0000 0000 0100 0001
S	0000 0000 0101 0011
C	0000 0000 0100 0011
I	0000 0000 0100 1001
I	0000 0000 0100 1001



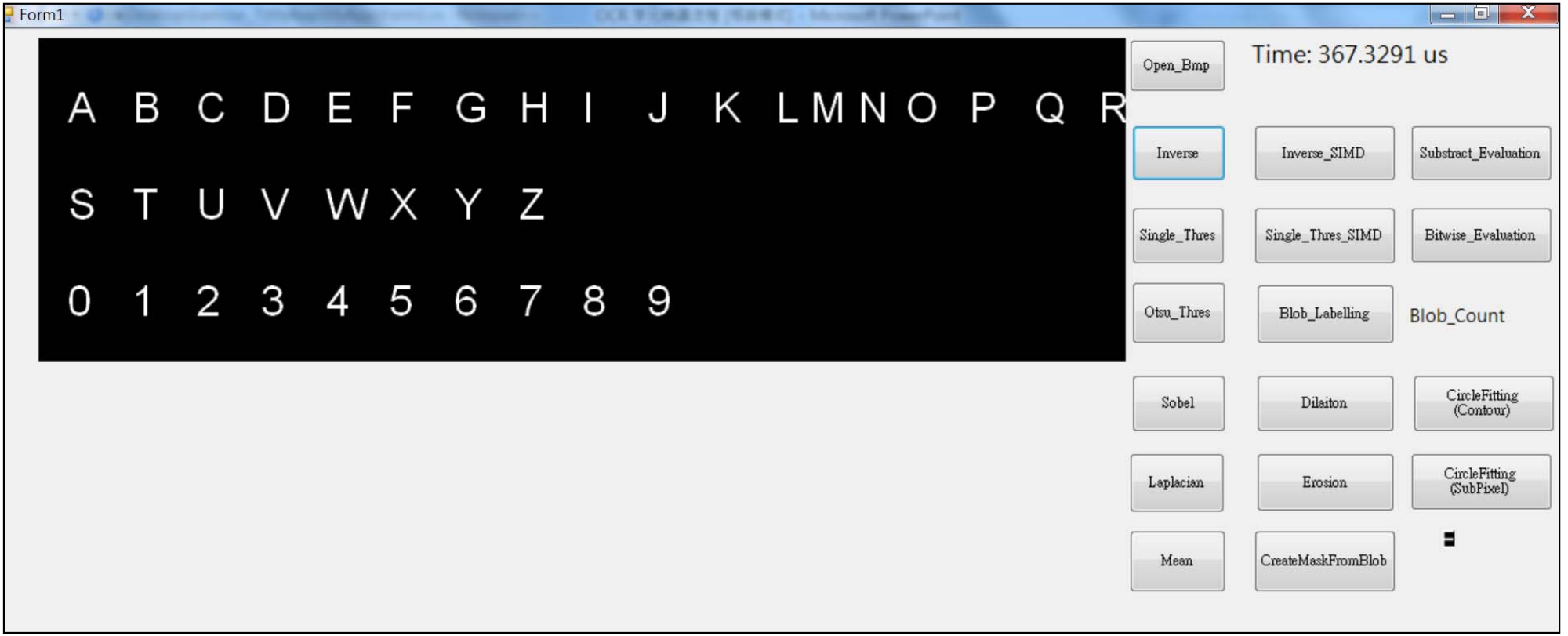
影像前處理

- 採用影像反向、 Otsu 二值化或適應性二值化。
- Otsu 二值化的方法與範例詳見 **Lab_5** 。
- 影像反向詳見 **Lab_5** 。
- 適應性二值化請參考作業一。

```
// 影像前處理：Otsu 二值化  Lab 5_Using PInvoke for NImgProcess
NImgProcessDLL.OtsuThresholding(m_Img, m_Otsu_Img, m_ImgPro);

// 影像前處理：影像反向 Lab 5_Using PInvoke for NImgProcess
NImgProcessDLL.Inverse(m_Otsu_Img, m_ImgPro);
```

影像前處理：程式驗證

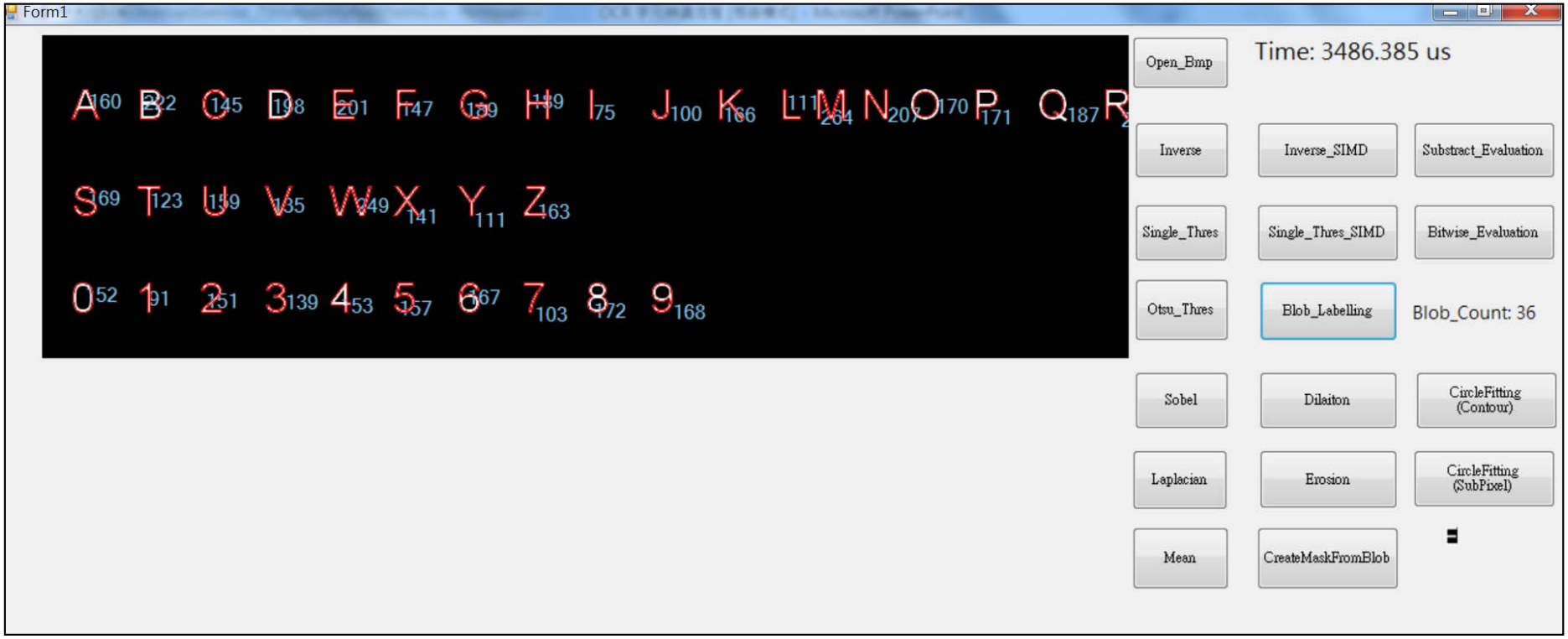

Form1
A B C D E F G H I J K L M N O P Q R
S T U V W X Y Z
0 1 2 3 4 5 6 7 8 9
Open_Bmp
Time: 367.3291 us
Inverse
Inverse_SIMD
Substract_Evaluation
Single_Thres
Single_Thres_SIMD
Bitwise_Evaluation
Otsu_Thres
Blob_Labelling
Blob_Count
Sobel
Dilaiton
CircleFitting (Contour)
Laplacian
Erosion
CircleFitting (SubPixel)
Mean
CreateMaskFromBlob

影像分割

- 採用線段編碼式的連結體標籤 (Connected component labelling)。
- 演算法方法與範例詳見 **Lab_12** 。
- 在期中測驗，有說明可利用幾何特徵來篩選待分類的候選物件。

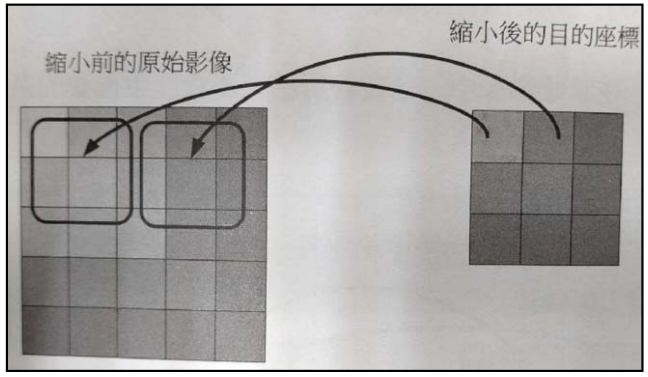
```
// 影像分割：Connected component labelling  Lab 12_Using PInvoke in C# for Object Detection
int blob_count = NObjectDLL.Blob_Labelling(m_Otsu_Img, m_Obj);
```

影像分割：程式驗證


Form1
Open_Bmp
Time: 3486.385 us
Inverse
Inverse_SIMD
Substract_Evaluation
Single_Thres
Single_Thres_SIMD
Bitwise_Evaluation
Otsu_Thres
Blob_Labelling
Blob_Count: 36
Sobel
Dilaiton
CircleFitting (Contour)
Laplacian
Erosion
CircleFitting (SubPixel)
Mean
CreateMaskFromBlob

特徵擷取

- 把影像縮小至 10 x 15 pixels ，並輸出為 150 元素的特徵向量。
- 特徵向量在輸入時先經過正規化處理。
- 影像縮小是從眾多像素中抽樣的一種處理，無法避免失真。
- 利用縮小來源區域中包含的像素平均值計算。




```
//影像分割：切割字元影像 Lab 18_Optical Character Recognition
if (NImgProcessDLL.Split_Image(m_Otsu_Img, start_x, start_y, rect_w, rect_h, m_Split_Img, m_ImgPro))
{
    //特徵擷取：Resize image to 10*15  Lab 18_Optical Character Recognition
    if (NImgProcessDLL.Small_Transform(m_Split_Img, m_OCR_Img, m_ImgPro))
    {
        //特徵擷取：Transfer to 1-D feature vector Lab 18_Optical Character Recognition
        NImgProcessDLL.FromImageToVector(m_OCR_Img, ref vector[0], 150, m_ImgPro);

        for (int j = 0; j < 150; j++)
            samples[j, i] = vector[j];
    }
}
```

樣式分類

- 延伸 **Lab_15** 的 ***MLP*** 分類器至字元辨識。
- 調整輸入特徵 = 150 。
- 調整輸出特徵 = 16 。

```
const int number_of_layers = 3;                    //3
const int number_of_input_nodes =  150;            //150
const int number_of_output_nodes = 16;             //16
const int maximum_layers = 250;                    //250
const int maximum_number_of_sets = 100;            //100
int number_of_input_sets;
int epochs = 600;                                  //600
const double error_threshold = 0.0002;             //0.0002F
```

樣式分類

- 在 **Lab_18** 新增儲存及讀取訓練網路參數。
- ***MLP*** 類別有以下四個介面函式：

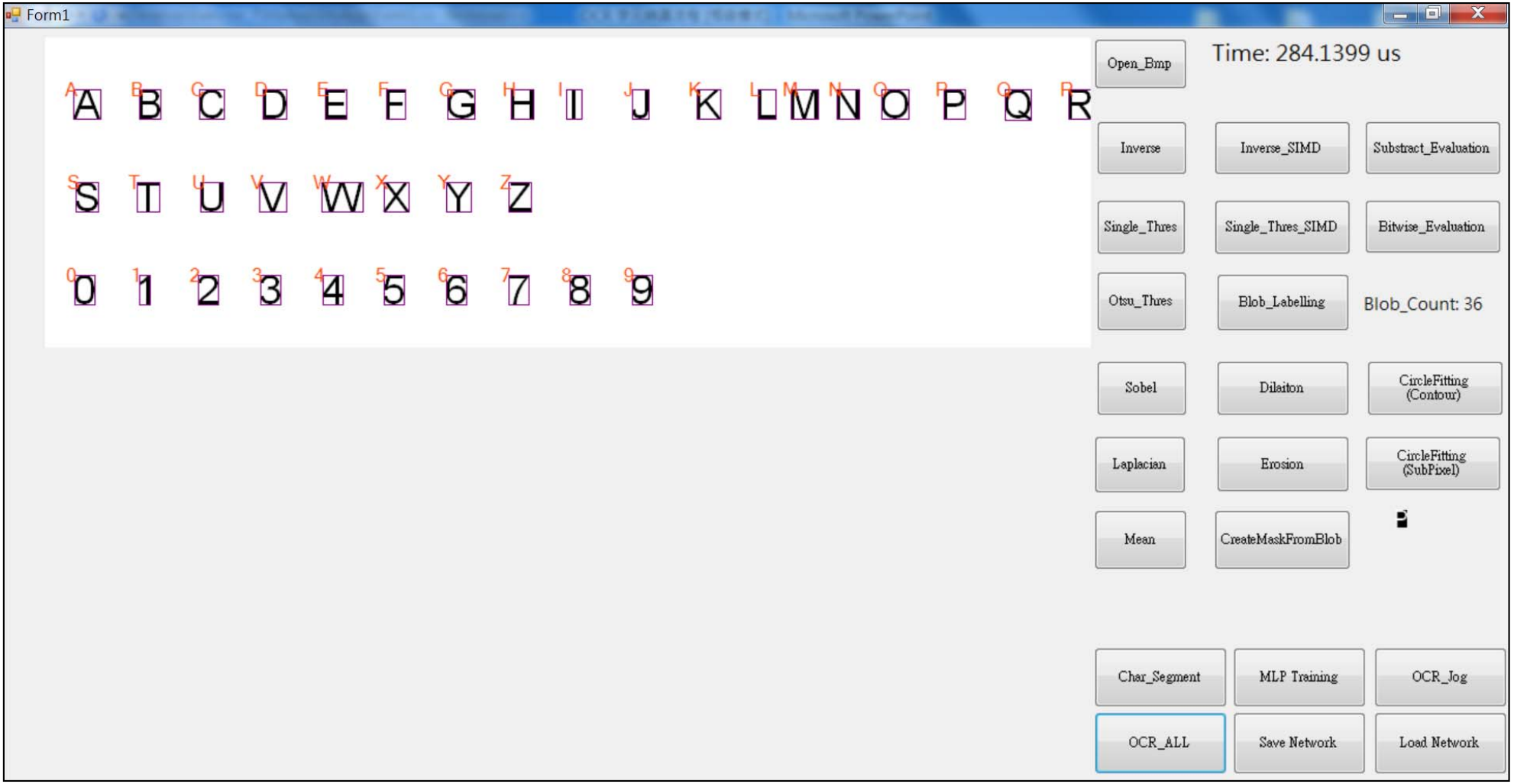
```
//樣式分類：MLP 訓練 Lab 18_Optical Character Recognition
//samples 為輸入學習字元組的影像特徵
//trainer_string 為每組學習影像特徵對應的字串(後續需轉為 16 bits 的 Unicode 當成輸出值)
//字串及物體個數需匹配才能開始學習
if (number_of_input_sets == blob_count)
    mlp.Training(samples, trainer_string, blob_count);

//樣式分類：MLP 推論或分類 Lab 18_Optical Character Recognition
//samples 為輸入字元的影像特徵
string result = mlp.Classify(sample);

//樣式分類：儲存訓練後的 MLP 參數 Lab 18_Optical Character Recognition
mlp.Save_Network(network_save_file_stream);

//樣式分類：讀取訓練後的 MLP 參數 Lab 18_Optical Character Recognition
mlp.Load_Network(network_load_file_stream);
```

樣式分類：程式驗證


Form1
A B C D E F G H I J K L M N O P Q R
S T U V W X Y Z
0 1 2 3 4 5 6 7 8 9
Open_Bmp
Time: 284.1399 us
Inverse
Inverse_SIMD
Substract_Evaluation
Single_Thres
Single_Thres_SIMD
Bitwise_Evaluation
Otsu_Thres
Blob_Labelling
Blob_Count: 36
Sobel
Dilaiton
CircleFitting (Contour)
Laplacian
Erosion
CircleFitting (SubPixel)
Mean
CreateMaskFromBlob
Char_Segment
MLP Training
OCR_Jog
OCR_ALL
Save Network
Load Network