



# 目錄

目錄.....	I
中文摘要.....	II
ABSTRACT.....	III
一.前言.....	1
二.試驗設備與方法.....	4
三.結果與討論.....	16
四.結論.....	30
五.致謝.....	31
六.參考文獻.....	31

# 以近紅外線及影像技術鑑別水稻品種(II)

蕭介宗<sup>1</sup>，劉昌群<sup>2</sup>，彭敬益<sup>3</sup>，洪梅珠<sup>4</sup>

1. 國立台灣大學生物產業機電工程所教授
2. 國立台灣大學生物產業機電工程所博士生
3. 國立台灣大學生物產業機電工程所研究助理
4. 台中區農業改良場副研究員

## 中文摘要

本研究探討以近紅外線及影像處理技術來建立不同年份之台中區農業改良場栽種之86年一期、87年一期、88年一期的5品種類水稻的鑑別模式，以鑑別該場栽種之89年一期此5品種類水稻品種。

使用近紅外線分光光度計得到波長由1100nm到2500nm之原始光譜經一次差分處理後再進行主成份分析法，得到在主成份軸上的投影值，再引入倒傳遞類神經網路進行訓練，以建立鑑別模式。而影像外觀特徵值取面積、周長、形狀係數、面積與周長比值、最大寬度值、最長軸值、最長軸值與最大寬度值的比值、每粒稻穀平均紅色光度值(R)、平均綠色光度值(G)、平均藍色光度值(B)、水稻最長軸上50等分之寬度值共60個參數或其中43參數或其中6參數，經正規化，引入倒傳遞類神經網路進行模組訓練，以建立鑑別模式。

建立之近紅外線技術及影像技術之鑑別模式，對86、87、88、89年台中區農業改良場第一期稻作5品種水稻品種的平均鑑別率，以近紅外線技術為94.7%，標準偏差2.25%；以影像技術建立預測模式對89年台中區農業改良場第一期稻作5品種水稻品種的平均鑑別率，以60參數組為94%，標準偏差為9.97%；以43參數組為90.2%，標準偏差為12.51%；以6參數組為78.6%，標準偏差為29.84%。

關鍵字：近紅外線、影像、類神經網路、模式

# Classifying Paddy Rice by Near-infrared and Image Processing Technique

Jai-Tsung Shaw<sup>1</sup>, Chang-Chun Liu<sup>2</sup>, Keen-Yik Poong<sup>3</sup>, Mei-Chu Hong<sup>4</sup>

1. Professor, Dept. of Bio-Industrial Mechatronics Engineering, National Taiwan University.
2. Graduate Student, Dept. of Bio-Industrial Mechatronics Engineering, National Taiwan University.
3. Research Assistant, Dept. of Bio-Industrial Mechatronics Engineering, National Taiwan University.
4. Associate Agronomist, Taichung District Agricultural Improvement Station.

## ABSTRACT

By near-infrared (NIR) and image processing technique, the objective of this research is to establish a prediction model from five varieties of paddy rice grown in Taichung District Agricultural Improvement Station and harvested in the summer of 1997, 1998, 1999 to classify the paddy rice also grown the same improvement station and harvested in 2000 summer.

The original spectrum between 1100nm and 2500nm, from near-infrared analyzer treated by first difference method and analyzed by Principal Component Analysis (PCA). The projection values on the optimum principal component axis were input into backpropagation neural network for training to find the prediction model. The characteristics of image such as area, perimeter, shape factor, the ratio of area to perimeter, the maximum width, the maximum long axis, average color value of red (R), average color value of green (G), average color value of blue (B), and the value of widths along the long axis were divided into 50 divisions of each paddy rice kernel. These data were normalized and then input into backpropagation neural network for training to find the image prediction model.

The average classification rate is 94.7% with standard deviation 2.26% by the near-infrared model, and the average classification rate by the image processing model is 94% with standard deviation 9.97% for 60 parameters, 90.2% with standard deviation 12.54% for 43 parameters, and 78.5% with standard deviation 29.84% for 6 parameters.

Keywords : near-infrared 、 image 、 artificial neural network 、 model

## 一. 前言

傳統上鑑別水稻品種的方法，大都以肉眼觀察種子有無特殊的外觀，需要專業人員依經驗或是對照圖表做辨別，此種鑑別手段稱為外觀型態學(morphological method)。以水稻品種為例，觀察的性狀外型包括穀粒的長瘦、圓胖、芒的有無及長短、稃尖的顏色、護穎長度、內外穎顏色、絨毛的著生情形及 1000 粒重等(黃, 1987)，除非具有特殊外型性狀，否則同一類型的品種是不易以肉眼來辨別區分。如此可知，憑經驗或對照圖鑑別品種，不僅存在人為的誤差而且速度緩慢，而且此類專家之培育除了不容易也耗財費時。因此，有利用掃描式電子顯微鏡(Scanning Electronic Microscope; SEM)掃描種子表面的構造(葉, 1996; 蔡, 1988; 呂與林, 1980)，然而表面微細構造的差異卻難以量化。有以數位攝影機結合影像處理技術(謝等人, 1989; Zayas 等人, 1986; 陳, 1989; 陳等人, 1989)的研究結果認為應用影像處理分析方法可準確測量穀粒的外觀特徵藉以判別穀粒之品種。若以化學實驗取得品種之化學成分來分析，或是依去氣核糖核酸(DNA)序列差異鑑別作物，雖然準確但難免流於費時費力及費財了，不適合水稻之收購及計價系統。

使用非破壞性的檢驗和應用電腦分析的軟體，藉由數位影像攝取、近紅外線分光光度計和電腦，建立校正模式，可以快速地分析非破壞性的檢測資料，替我們省去了耗時及可能造成環境污染的化學分析，並且節省人員培訓之支出費用。

近紅外線分光光度計利用光譜反射值與其內部化學成分有高度的相關(夏賀等人, 1993; Yamashita, 1993; 劉與蕭, 1995; 李與蕭, 1996; 彭等人, 1997)，建立化學成分與光譜反射值之校正線，應用於偵測稻米內部成分如蛋白質含量、直鏈澱粉含量、脂肪酸度及含水率。張等人(2000)、李(1998)、張(2000)，以近紅外線與影像技術鑑別 86 年一期、87 年一期台中區農業改良場所提供的稻穀，計有台農秈 20 號、台農 67 號、台

農 70 號、台中秈 10 號、台粳 8 號、台粳 9 號共 6 種水稻品種，不論以近紅外線之技術或是使用影像技術之鑑別率分別為 89.3% 及 95.8% 標準偏差分別為 8.1% 及 5.3%。

主成分分析法(Principal Component Analysis PCA)是一種座標轉換的方式。資料數據經 PCA 法分析後，會建成一個以數主要成分軸為座標的系統來描述資料。在主要成分軸系統中，資料數據只需要有同主要成分軸數一樣多的變數就能加以描述，因而在表現上就減少了變數的個數，而這些變數都是完全獨立且為鑑別品種的主要變數。

類神經網路(artificial neural network)是以簡單的運算處理單元連結成網路，達到模擬生物體之大量基本神經元構成的類似腦神經網路及其學習的功能，由於類神經網路具有簡單的組織及學習能力，可用於解決原始定義與分類不十分清楚的領域。倒傳遞類神經網路為目前類神經網路中最具代表性而且應用也最廣的模式，為多層次前授式網路，採用監督式的訓練，而學習是以最小平方差為指標，運用梯度陡降法(the gradient steepest descent method) 搜尋最小誤差函數，導出網路連結加權值與閾值的修正值。記錄訓練完成的網路連結加權值與閾值，並輸入未知品種的參數向量以進行鑑別品種。Chen 等人(1995)利用類神經網路和近紅外線光譜分類 1987-1989 年硬紅小麥，校正及驗證準確度分別為 97.9% 及 96.8%。Delwiche 等人(1995)利用類神經網路和近紅外線光譜分類硬紅小麥準確度介於 95-98% 優於多重線性回歸、主成分分析及部分最小平方法。

影像處理是外觀型態學之應用亦是非破壞性的方法，陳等人(1989)以影像技術收集穀物外型特徵變數共 73 項，判別同屬粳稻之台農 67、台農 70、新竹 64 共 3 種水稻品種，資料經常態性測驗及逐步判別分析選取具常態性且辨別力較大的變數 11 個，而後再經判別分析，結果得 3 種水稻品種間的平均鑑別率為 77.3%。影像技術配合類神經網路分析，近年來相當多應用於穀物品種鑑別與分級、稻穀品質的檢定等，在品質鑑定難以量化的絕時上，影像技術結合類神經網路是相當好的方法(謝與林，1994；萬等人，1998；李，

1998)。

由以上可知，由於不同的水稻品種可能有不同的成分而存在不同的光譜反射值，故將探討以近紅外線光譜分析技術來鑑別品種，此方法猶如依內部化學成分之不同來鑑別品種，而影像視覺猶如外觀特性之量化，若能結合水稻之內(近紅外線光譜分析)、外(影像視覺)特性，來鑑別水稻品種，不但能消除傳統外觀型態學辨別品種之人為誤差，更能直接保障農民的投資及消費者的權益。

本研究之目的為：

1. 利用影像處理技術所得 86、87、88 年一期稻作之稻穀的外型影像特徵值，經正規化後，以倒傳遞內神經網路之方法建立一鑑別模式，利用此一模式預測 89 年一期稻作之品種。
2. 把從近紅外線分光光度計得到的原始反射式光譜經一次差分後利用主成分分析法 (Principle Component Analysis ; PCA) 分析) 求得 86、87、88、89 年一期稻作水稻品種之近紅外線光譜在最佳主成分軸上投影值，帶入倒傳遞類神經網路學習後建立鑑別模式，以鑑別不同之品種。

## 二. 試驗設備與方法

### (一) 試驗材料：

本研究所採用的樣品為台中區農業改良場 1997 年一期、1998 年一期、1999 年一期及 2000 年一期之台農秈 20 號、台農 67 號、台中秈 10 號、台稈 8 號、台稈 9 號共 5 種水稻品種。

### (二) 試驗設備：

本研究中所用到的主要設備，茲簡述如下，詳細資料參閱張(2000)：

1. 近紅外線分光光度計本體：Bran-Lubbe InfraLyzer 500，主要用於測定稻穀之反射式光譜。
2. InfraLyzer Data Analysis Software(IDAS)軟體，只利用此軟體來做反射光譜之收集。
3. Unscrambler<sup>®</sup>6.0 軟體，利用此軟體內的 Principal Component Analysis(PCA)運算模組，進行光譜主成分分析。
4. 風選機：用來去除稻穀樣本內之雜草、空稻殼等雜物。
5. 恆溫恆濕箱(Temperature and Humidity Chamber)溫度控制範圍為 0°C~100°C，相對濕度控制範圍為 20%-98%，主要用來控制稻穀之穀溫及含水率。稻穀樣品置於 25°C、RH70% 之條件下 4 天，調節成 13% 之含水率。
6. Shizudka Seiki CTR-800E 型：測量水稻含水率。
7. Wattec CCD (Charge Couple Device) 攝影機：可輸出 RGB 訊號或 NTSC (National Television Standards Committee, United States of America)複合彩色視訊，感光元件的解析度為 NTSC-512 X 512 像素點。
8. PXU200 影像擷取卡：解析度為 NTSC-640\*480 像素點。
9. 打光暗室：將樣本至於暗室內取其影像。
10. 影像分析軟體：以 Borland C++ 語言撰寫之影像分析軟體，分析水稻樣本得最長軸值、最大寬度值、最長軸值與最大寬度值的比值、面積、周長、面積與周長的比值、形狀係數(shape factor,  $4\pi \times \text{面積} / \text{周長}^2$ )、每粒稻穀平均紅色光度值(R)、平均綠色光度值(G)、平均藍色光度值(B)及長軸上等距之 50 個寬度值共 60 個外型影像特徵參數。



11. 倒傳遞類神經網路軟體：用 MathWorks 公司出版之數學運算軟體 Matlab 5.1 版為平台，運算自行編寫之倒傳遞類神經網路數學運算程式，以取得所需之網路連結加權值與閾值，並利用之預測未知品種。

### (三)試驗方法

#### 1. 近紅外線

- (1) 各品種樣本先以風選機風選數次，以去除空稻穀、稻穗、雜物等。將水稻樣本置於 25°C、RH70%的恆溫恆濕箱內約 4 天(100 小時)調質成 13%的含水率，以 Shizudka Seiki CTR-800E 檢測之。進行近紅外光照射約 100 組樣本，照射後之樣本不重複使用。將原始光譜資料經過界外點的利用馬氏距離大於 9/樣本數或 99.7% 信賴區外加以剔除(Vellwman and Welsch, 1981)，隨機挑選每品種每期 30 個樣本作為模式建立之訓練組，其餘的作為驗證組。
- (2) 將 1997 年、1998 年、1999 年及 2000 年 5 種共 600 個(表 1)水稻樣本的訓練組合併起來，結合成為倒傳遞類神經網路之訓練組。並將訓練組資料經一次差分後再經過主成分分析法(PCA)求得主成分模式，建議之最佳主成分軸數(7 軸)，再將各品種樣本在最佳主成分軸上的投影值代入倒傳遞類神經網路中，進行品種鑑別模式的訓練，直到全部誤差平方總和(SSE)=5 時停止訓練。並將四年之樣本共 1351 個合併成為一單獨之驗證組，如表 2。
- (3) 使用含引用 Matlab 軟體支援之自行撰寫之倒傳遞類神經網路程式，所使用之網路模式是一個 4 層的網路結構，其中第一層為輸入層，中間兩層為隱藏層—隱藏層第一層為 10 個節點，第二層為 30 個節點，第四層為輸出層，初始之網路連結加權值( $W_0$ )及閾值( $B_0$ )由程式以隨機產生，訓練組完成時記錄下網路連結加權值( $W$ )及閾值( $B$ )。
- (4) 將這些品種的驗證組共 1351 個樣本(表 1)依序組合經過一次差分，再使用訓練組做好之主成分模式，求得在最佳主成分軸上的投影值，代入由訓練組完成訓練時所得之網路連結加權值( $W$ )與閾值( $B$ )所建的倒傳遞類神經網路中，來計算各品種的鑑別率(classification rate)及平均鑑別率(average classification rate)。
- (5) 倒傳遞類神經網路輸出層目標輸出值見表 3 相應的目標輸出值為 1 時，代表辨別為正確的品種，其他品種則以 0 表示。

表 1 1997、1998、1999 及 2000 年一期之近紅外線漢式訓練組樣本數

Table 1 Trained samples harvested in 1997、1998、1999 and 2000 summer for near-infrared model

	1997	1998	1999	2000	總共 Total	編號 No.
台農私 20 號 Tainung Sen 20	30	30	30	30	120	1-120
台中私 10 號 Taichung Sen 10	30	30	30	30	120	121-240
台農 67 號 Tainung 67	30	30	30	30	120	241-360
台稔 8 號 Taikeng 8	30	30	30	30	120	361-480
台稔 9 號 Taikeng 9	30	30	30	30	120	481-600
總和 Total	150	150	150	150	600	

表 2 1997、1998、1999 及 2000 年一期之近紅外線漢式驗證組樣本數

Table 2 Validation samples harvested in 1997、1998、1999 and 2000 summer for  
near-infrared model

	1997	1998	1999	2000	總共 Total	編號 No.
台農私 20 號 Tainung Sen 20	68	66	67	70	271	1-271
台中私 10 號 Taichung Sen 10	68	68	64	68	268	272-539
台農 67 號 Tainung 67	66	65	67	69	267	540-806
台稈 8 號 Taikeng 8	68	68	70	68	274	807-1080
台稈 9 號 Taikeng 9	69	67	67	68	271	1081-1351
總和 Total	339	334	335	343	1351	

表 3 近紅外線模式之訓練組編號以及相應的目標輸出值

Table 3 code of trained set and corresponding output of near-infrared model

訓練組編號 code of trained set	品種 Varieties	相應的目標輸出值 Corresponding output				
1-120	台農私 20 號 Tainung Sen 20	-1	0	0	0	0
121-240	台中私 10 號 Taichung Sen 10	0	1	0	0	0
241-360	台農 67 號 Tainung 67	0	0	1	0	0
361-480	台稷 8 號 Taikeng 8	0	0	0	1	0
481-600	台稷 9 號 Taikeng 9	0	0	0	0	1

(6) 學習速率( $\eta$ )為 0.01。

(7) 停止訓練時之誤差平方總和(sum of square error : SSE) =5。

## 2. 稻穀外觀影像特徵值：

(1) 由各品種的近紅外線照射樣本中挑選出 1000 顆完整穀粒，置於打光暗室並以影像設備來擷取水稻樣本長軸 50 個寬度值、最長軸值、最大寬度值、最長軸值與最大寬度值的比值、面積、周長、面積與周長的比值、形狀係數(shape factor,  $4\pi \times \text{面積}/\text{周長}^2$ )、每粒稻穀平均紅色光度值(R)、平均綠色光度值(G)、平均藍色光度值(B)共 60 個外型影像特徵參數。以圖 1 為例，L 是最長軸、Y 是最大寬度、AC 的長度為第 4 個寬度值，BAOCB 為周長，BAOCB 所圍成的區域就是面積。

(2) 將 1997 年、1998 年、1999 年及 2000 年一期台農私 20 號、台農 67 號、台中私 10 號、台稈 8 號、台稈 9 號等 5 個品種每品種 1000 粒樣本數，以每 10 粒樣本平均成一個樣本，即每品種有 100 個樣本數。將此樣本資料利用馬氏距離大於 9/樣本數或 99.7% 信賴區外之樣本加以剔除(Vellwman and Welsch, 1981)，隨機挑選每品種每期 30 個樣本作為模式建立之訓練組，其餘的樣本作為驗證組。

(3) 將 1997 年、1998 年及 1999 年一期台農私 20 號、台農 67 號、台中私 10 號、台稈 8 號、台稈 9 號等 5 個品種之稻穀訓練組共 425 樣本(表 4)之 50 個寬度值、最長軸值、最大寬度值、最長軸值與最大寬度值的比值、面積、周長、面積與周長的比值、形狀係數(shape factor,  $4\pi \times \text{面積}/\text{周長}^2$ )、每粒稻穀平均紅色光度值(R)、平均綠色光度值(G)、平均藍色光度值(B)共 60 個外型影像特徵參數，取最大最小對映法(葉, 1993)，正規化後輸出值域介於 0~1 之間，代入倒傳遞類神經網路以進行品種鑑別的訓練，訓練完成記錄下連結加權值(W)與閾值(B)。

(4) 將 1997 年、1998 年及 1999 年一期台農私 20 號、台農 67 號、台中私 10 號、台稈 8 號、台稈 9 號等 5 個品種稻穀除訓練組外之樣本組成驗證組共 806 樣本(表 5)依序組合經正規化，並載入由訓練組倒傳遞類神經網路訓練求得之網路連結加權值(W)與閾值(B)，代入倒傳遞類神經網路來計算各品種的鑑別率。

(5) 倒傳遞類神經網路輸出層目標輸出值見表 6，相應的目標輸出值為 1 時，代表

辨別為正確的品種，其他品種則以 0 表示。

(6) 初始之網路連結加權值( $W_0$ )及閾值( $B_0$ )由 MATLAB 以隨機產生。

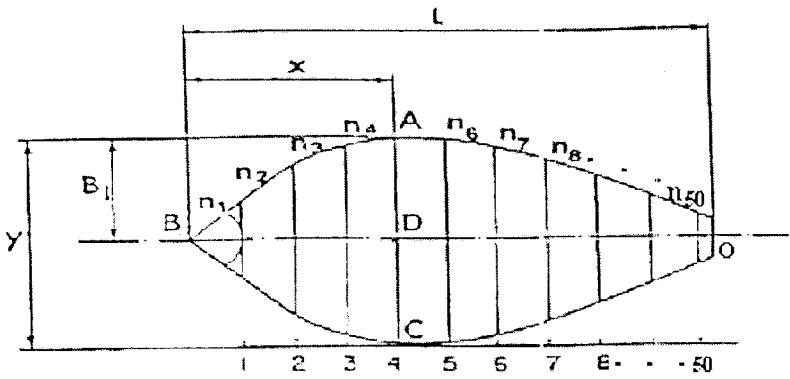


圖 1 水稻外型特徵值示意圖

Fig.1 Shape of paddy rice

表 4 1997、1998 及 1999 年一期之影像模式訓練組樣本數

Table 4 Trained samples harvested in 1997、1998 and 1999 summer for image model

	1997	1998	1999	總共 Total	編號 No.
台農私 20 號 Tainung Sen 20	25	30	30	85	1-85
台中私 10 號 Taichung Sen 10	25	30	30	85	86-170
台農 67 號 Tainung 67	25	30	30	85	171-255
台稔 8 號 Taikeng 8	25	30	30	85	256-340
台稔 9 號 Taikeng 9	25	30	30	85	341-425
總和 Total	125	150	150	425	

表 5 1997、1998 及 1999 年一期之影像模式驗證組樣本數

Table 5 Trained samples harvested in 1997、1998 and 1999 summer for image model

	1997	1998	1999	總共 Total	編號 No.
台農私 20 號 Tainung Sen 20	25	68	67	160	1-160
台中私 10 號 Taichung Sen 10	26	68	70	164	161-324
台農 67 號 Tainung 67	25	67	67	159	325-483
台稔 8 號 Taikeng 8	25	68	70	163	484-646
台稔 9 號 Taikeng 9	25	67	68	160	647-806
總和 Total	126	338	342	806	



表 6 影像模式其訓練組編號及相應的目標輸出值

Table 6 Code of trained set and corresponding output of image model

訓練組編號 code of trained set	品種 Varieties	相應的目標輸出值 Corresponding output				
1-85	台農私 20 號 Tainung Sen 20	1	0	0	0	0
86-170	台中私 10 號 Taichung Sen 10	0	1	0	0	0
171-255	台農 67 號 Tainung 67	0	0	1	0	0
256-340	台稔 8 號 Taikeng 8	0	0	0	1	0
341-425	台稔 9 號 Taikeng 9	0	0	0	0	1

- (7) 學習速率( $\eta$ )為 0.01。
- (8) 停止訓練時之誤差平方總和(sum of square error : SSE) =5。
- (9) 將 2000 年一期台農私 20 號、台農 67 號、台中私 10 號、台稔 8 號、台稔 9 號等 5 個品種之稻穀每品種 1000 粒樣本以每 10 粒平均為一個樣，即有 100 個樣本，再挑出值超出界外點之樣本(Vellman and Welsch, 1981)，共 493 個樣本作為預測組(prediction set)(表 7)。
- (10) 將 2000 年一期台農私 20 號、台農 67 號、台中私 10 號、台稔 8 號、台稔 9 號等 5 個品種稻穀共 493 個樣本(表 7)依序組合經正規化，並載入由訓練組倒傳遞類神經網路訓練求得之網路連結加權值(W)與閾值(B)，代入倒傳遞類神經網路來計算各品種的鑑別率。倒傳遞類神經網路輸出層目標輸出值見表 6，相應的目標輸出值為 1 時，代表辨別為正確的品種，其他品種則以 0 表示(表 7)。

表 7 影像模式其預測組數量及編號

Table 7 The sample size and code of prediction set of image model

品種 Varieties	數量 size	訓練組編號 code of trained set
台農私 20 號 Tainung Sen 20	100	1-100
台中私 10 號 Taichung Sen 10	98	101-198
台農 67 號 Tainung 67	99	199-297
台稈 8 號 Taikeng 8	98	298-395
台稈 9 號 Taikeng 9	98	396-493

### 三. 結果與討論

#### (一) 近紅外線跨年模式之鑑別結果：

結合 1997 年一期、1998 年一期、1999 年一期及 2000 年一期 5 種水稻的驗證組之光譜，經一次差分再由 PCA 分析求得在 7 個最佳主成分軸上的投影值，載入由訓練組完成時之網路連結加權值(W)及閾值(B)，代入倒傳遞類神經網路中，結果見表 8，5 種水稻平均鑑別率為 94.7%，標準偏差為 2.25%。此結果與張等人(2000)之平均鑑別率為 95.9%，標準偏差為 5.3，而成對 t 值檢定為 0.402 遠小於本研究所定顯著差異標準成對 t 值檢定  $t_{0.05, 2.5}=2.57$ ，顯示沒有顯著的差異，但最佳主成分軸由 3 軸增加到 7 軸。本研究與張等人(2000)的主要差別在於張沒有將數據資料進行一次差分，而本研究利用一次差分將稻穀本身大小的差異消弭並增加資料的可信度。由表 9，雖然由實驗 t 值知兩者間似無差別。用 3 主成分軸不經一次差分處理之結果如表 10，1997-1999 年的模式一和 1997-2000 年的模式二與張等人(2000)的模式間的實驗 t 值反而比較大，但仍沒有顯著差異。

表 11 與表 12 為近紅外線 7 個參數模式中的模式一及模式二中，5 種品種水稻各自誤判為其它品種的情況。從表 11 與表 12 中看出，私稻品種通常是被誤判為同為私稻的其它品種，而粳稻也有同樣的結果，其中比較特殊的是台粳 9 號有少數會被誤判為台中秈 10 號。台粳 8 號被誤判為台農 67 號的情況是 5 個品種中最多的也是唯一的，模式一中有 13 個，模式二有 23 個。

陳(1994)及 Chen 等人(1995)以 1987-1989 近紅外線樣本利用類神經網路所建立之模式鑑別 1990 年硬紅春麥及硬紅冬麥，鑑別率為 96.8%，本研究之 7 個參數類的模式二 5 種水稻品種的平均鑑別率 94.7%，標準偏差 2.25%，與 Delwiche 等人(1995)鑑別率介於 95-98% 之結果相當。

#### (二) 影像之跨年模式鑑別結果：

表 13，按照張等人(2000)作法挑選參數，並結合 1997、1998 和 1999 年的訓練組所建立的模型，驗證 1997、1998 和 1999 年 5 種水稻的鑑別率，平均鑑別率是 60 參數組為 90.3%，標準偏差為 9.22%；43 參數組為 93.1%，標準偏差為 9.80；6 參數組為 90.3%，標準偏差為 10.89%。而結合 1997、1998 和 1999 年所建立之模型鑑別 2000 年 5 種水稻

表 8 近紅外線模式之鑑別率

Table 8 Classification rates of NIR model

	模式一* Model 1* (%)	模式二** Model 2.** (%)
台農私 20 號 Tainung Sen 20	— 94.0	97.4
台中私 10 號 Taichung Sen 10	98.0	94.4
台農 67 號 Tainung 67	96.0	96.2
台有 8 號 Taikeng 8	93.7	91.6
台有 9 號 Taikeng 9	93.1	93.7
平均鑑別率 Average classification rate	95.0	94.7
標準偏差 Standard deviation	2.02	2.25

Notes(註)：\* included the paddy rice of 1997, 1998 and 1999 with 7 PCA projection values.

\* 包括 1997、1998 和 1999 年水稻品種之 7 主成分軸之投影值。

\*\* included the paddy rice of 1997, 1998, 1999 and 2000 with 7 PCA projection values.

\*\* 包括 1997、1998、1999 和 2000 水稻品種之 7 主成分軸之投影值。

表 9 與張等人(2000)之近紅外線鑑別模式之鑑別率比較

Table 9 Classification rate of NIR model to compare with Chang et al.(2000) model.

	張等人(2000)模式 Chang et al.(2000) model ** (%)	模式一 Model 1 (%)	模式二 Model 2 (%)
台農私 20 號 Tainung Sen 20	94.8	94.0	97.4
台中私 10 號 Taichung Sen 10	86.0	98.0	94.4
台農 67 號 Tainung 67	100	96.0	96.2
台有 8 號 Taikeng 8	100	93.7	91.6
台有 9 號 Taikeng 9	98.5	93.1	93.7
平均鑑別率 Average	95.9	95.0	94.7
標準偏差 STD	5.91	2.02	2.25
t 值 t-value		0.268	0.402
理論 t 值 Theoretical t-value	$t_{0.05; 2, 5} = 2.57^*$		

註(Notes)：\*表示處理間沒有顯著差異(significant)。

\*\*with 3 principal component axes (3 主成分軸)。

表 10 與張等人(2000)之近紅外線模式鑑別率\*之比較

Table 10 Classification rate of NIR model\* to compare with Chang et al.(2000) model.

	張等人(2000)模式 Chang et al.(2000) model (%)	模式一 Model 1 (%)	模式二 Model 2 (%)
台農私 20 號 Tainung Sen 20	94.8	96.5	98.9
台中私 10 號 Taichung Sen 10	86.0	91.0	94.4
台農 67 號 Tainung 67	100	97.5	90.3
台有 8 號 Taikeng 8	100	79.6	83.6
台有 9 號 Taikeng 9	98.5	80.3	87.1
平均鑑別率 Average	95.9	89.0	90.9
標準偏差 STD	5.91	8.61	6.0
t 值 t-value		1.32	1.05
理論 t 值 Theoretical t-value	$t_{0.05/2, 5} = 2.57$		

註(Notes)：\*均用 3 主成分軸之投影值。

\*with 3 principal component axes

表 11 NIR 鑑別模式一中誤判粒數

Table 11 Number of kernel by Classifying error in NIR model 1

7 個參數(7 parameters)					
	誤判成其他品種之顆數(Error by number of kernel)				
	台農私 20 號 Tainung Sen 20	台中私 10 號 Taichung Sen 10	台農 67 號 Tainung 67	台稈 8 號 Taikeng 8	台稈 9 號 Taikeng 9
台農私 20 號 Tainung Sen 20		12	0	0	0
台中私 10 號 Taichung Sen 10	4		0	0	0
台農 67 號 Tainung 67	0	0		3	5
台稈 8 號 Taikeng 8	0	0	13		0
台稈 9 號 Taikeng 9	0	3	6	5	



表 12 NIR 鑑別模式二中誤判粒數

Table 12 Number of kernel by Classifying error in NIR model 2

7 個參數(7 parameters)					
	誤判成其他品種之顆數(Error by number of kernel)				
	台農私 20 號 Tainung Sen 20	台中私 10 號 Taichung Sen 10	台農 67 號 Tainung 67	台稔 8 號 Taikeng 8	台稔 9 號 Taikeng 9
台農私 20 號 Tainung Sen 20		7	0	0	0
台中私 10 號 Taichung Sen 10	15		0	0	0
台農 67 號 Tainung 67	0	0		2	7
台稔 8 號 Taikeng 8	0	0	23		0
台稔 9 號 Taikeng 9	0	6	2	9	

表 13 不同影像參數對水稻品種鑑別率的影響(%)

Table 13 Different image parameters effecting classification rate of paddy rice(%)

	樣式驗證 1997、1998 和 1999 年品種 Validating 1997, 1998 and 1999 year's paddy rice (%)			樣式鑑別 2000 年品種 Classifying 2000 year's paddy rice (%)		
	60	43	6	60	43	6
輸入參數 Input parameters	60	43	6	60	43	6
台農私 20 號 Tainung Sen 20	98.1	98.8	97.5	100	100	100
台中私 10 號 Taichung Sen 10	98.8	100	96.3	100	100	100
台農 67 號 Tainung 67	88.0	91.2	89.3	93.0	76.0	29.0
台稔 8 號 Taikeng 8	76.1	76.7	71.8	77.0	77.0	91.0
台稔 9 號 Taikeng 9	90.6	98.7	96.9	100	98.0	73.0
平均鑑別率 average classification rate	90.3	93.1	90.3	94.0	90.2	78.6
標準偏差 standard deviation	9.22	9.80	10.89	9.97	12.54	29.84
t 值 (t-value)		1.980	0		1.144	1.112
理論 t 值 Theoretical t-value		$t_{0.05, 2.5} = 2.57$				

品種，平均鑑別率是 60 參數組為 94%，標準偏差為 9.97%；43 參數組為 90.2%，標準偏差為 12.54；6 參數組為 78.6%，標準偏差為 29.84%。可以看出在建立模式預測鑑別表 2000 的 5 個品種上 60 參數組和 43 參數組皆能達到 90% 以上的正確率，而模式鑑別 1997、1998 和 1999 年的 5 個品種上 60、43 和 6 參數組皆能達到 90% 以上的正確率。在模式預測 2000 年品種鑑別上，雖然張等人(2000)建議只利用 43 個參數來鑑別 5 種水稻品種，可是在 43 參數組和 6 參數組的標準偏差都分別高達 12.54% 和 29.84%。鑑別率也有參數越少就越低之趨勢，台農 67 號的鑑別率分別從 93.0% 降到 76.0% 及 29.0% 看出，減少參數也降低鑑別率。表 14、表 15 及表 16 皆為本研究與張等人(2000)在使用 60 個參數、43 個參數及 6 個參數上結果之比較，成對 t 值檢定上都沒有超過本研究所定的顯著差異標準  $t_{0.05/2,5}=2.57$ 。

由表 17、表 18 和表 19 中看出，稈稻彼此誤判的機率比較大，尤其是台農 67 號和台稈 8 號間互相誤判的情況較為嚴重，如 60 參數組裡台農 67 號有 7 個樣本給誤判為台稈 8 號，而台稈 8 號有 23 個樣本給誤判為台農 67 號；43 參數組裡，台農 67 號有 24 個樣本給誤判為台稈 8 號，而台稈 8 號有 23 個樣本給誤判為台農 67 號，此外台稈 9 號有 2 個樣本給誤判為台農 67 號；6 參數組裡，台農 67 號竟有 71 個樣本給誤判為台稈 8 號，而台稈 8 號有 8 個樣本給誤判為台農 67 號、1 個給誤判給台稈 9 號，台稈 9 號有 21 個樣本給誤判為台農 67 號、6 個給誤判給台稈 8 號。參數越少的情況下，稈稻誤判的情況越嚴重，雖然在 6 參數組的模式下台稈 8 號的鑑別率提高了，可是有 71 個樣本的台農 67 號和 21 個樣本的台稈 9 號給誤判給它了，顯然 6 個參數不足於讓模式正確的鑑別品種的差異；而台稈 8 號的鑑別率偏低，以它容易被誤判為台農 67 號的原因有很大關係。

減少參數對水稻品種鑑別率的影響，可參考表 13，在減少參數的各個模式中，模式鑑別 97、98 及 99 年 5 個品種的中不論 60、43 及 6 個參數組的模式，其平均鑑別率均高過 90%，而成對 t 值檢定也沒有顯著的差異，所以依張等人(2000)的意見挑選 43 個參數組，除了鑑別較高有 93.1% 外，引用的參數也比較少。若是以三年的模式鑑別 2000 年 5 個品種，顯然愈多的參數，對鑑別正確率的幫助愈大，引用 60 個參數的話可以減少某幾類的品種因為減少參數而鑑別率愈來愈低，如台農 67 號，雖然在平均鑑別率及成對 t 值檢定上，43 參數組與 60 參數組沒有太大差距，可是台農 67 號的鑑別率卻從 93% 掉到 76%，6 參數組則掉到 29%，為讓各品種能夠在足夠參數下表現彼此的差異，以便利鑑別，60 參數的全部利用是比較可接收的。

表 14 60 參數之影像模式與張等人(2000)模式之鑑別率比較

Table 14 Classification rate of image model of 60 parameters to compare with Chang et al.(2000) model.

60 參數( 60 parameters)			
	張等人(2000)模式 Chang et al.(2000) Model (%)	驗證 1997、1998 和 1999 年品種 Validating 1997,1998 and 1999 year's paddy rice (%)	鑑別 2000 年品種 Classifying 2000 year's paddy rice (%)
台農私 20 號 Tainung Sen 20	98.9	98.1	100
台中私 10 號 Taichung Sen 10	96.8	98.8	100
台農 67 號 Tainung 67	94.6	88.0	93.0
台有 8 號 Taikeng 8	79.6	76.1	77.0
台有 9 號 Taikeng 9	94.6	90.6	100
平均鑑別率 Average	92.9	90.3	94.0
標準偏差 STD	7.65	9.22	9.97
t 值 t-value		1.770	0.742
理論 t 值 Theoretical t-value		$t_{0.05,2,5}=2.57$	

表 15 43 參數之影像模式與張等人(2000)模式之鑑別率比較

Table 15 Classification rate of image model of 43 parameters to compare with Chang et al.(2000) model.

43 參數( 43 parameters)			
	張等人(2000)模式 Chang et.al.(2000) model (%)	驗證 1997、1998 和 1999 年品種 Validating 1997, 1998 and 1999 year's paddy rice (%)	鑑別 2000 年品種 Classifying 2000 year's paddy rice (%)
台農私 20 號 Tainung Sen 20	100	98.8	100
台中私 10 號 Taichung Sen 10	96.8	100	100
台農 67 號 Tainung 67	86.9	91.2	76.0
台有 8 號 Taikeng 8	79.6	76.7	77.0
台有 9 號 Taikeng 9	90.2	98.7	98.0
平均鑑別率 Average	90.7	93.1	90.2
標準偏差 STD	8.09	9.80	12.54
t 值 t-value		1.182	0.160
理論 t 值 Theoretical t-value	$t_{0.05/2,5}=2.57$		

表 16 6 參數之影像模式與張等人(2000)模式之鑑別率比較

Table 16 Classification rate of image model of 6 parameters to compare with Chang et al.(2000) image model.

6 參數( 6 parameters)			
	張等人(2000)模式 Chang et al.(2000) model (%)	模式驗證 1997、1998 和 1999 年品種 Validating 1997,1998 and 1999 year's paddy rice (%)	模式鑑別 2000 年品種 Classifying 2000 year's paddy rice (%)
台農私 20 號 Tainung Sen 20	95.8	97.5	100
台中私 10 號 Taichung Sen 10	93.7	96.3	100
台農 67 號 Tainung 67	75.5	89.3	29.0
台有 8 號 Taikeng 8	53.2	71.8	91.0
台有 9 號 Taikeng 9	91.4	96.9	73.0
平均鑑別率 Average	81.9	90.3	78.6
標準偏差 STD	17.94	10.89	29.84
t 值 t-value		2.53	0.235
理論 t 值 Theoretical t-value	$t_{0.05/2,5}=2.57$		

表 17 影像模式鑑別 2000 年品種中 60 參數組的誤判粒數

Table 17 Kernels of Classifying error in image model of classifying 2000 paddy rice with 60 parameters

60 個參數(60 parameters)					
	誤判成其他品種之顆數(Error by number of kernel)				
	台農秈 20 號 Tainung Sen 20	台中秈 10 號 Taichung Sen 10	台農 67 號 Tainung 57	台稈 8 號 Taikeng 8	台稈 9 號 Taikeng 9
台農秈 20 號 Tainung Sen 20		0	0	0	0
台中秈 10 號 Taichung Sen 10	0		0	0	0
台農 67 號 Tainung 67	0	0		7	0
台稈 8 號 Taikeng 8	0	0	23		0
台稈 9 號 Taikeng 9	0	0	0	0	

表 18 影像模式鑑別 2000 年品種中 43 參數組的誤判粒數

Table 18 Kernels of Classifying error in image model of classifying 2000 paddy rice with 43 parameters

43 個參數(43 parameters)					
	誤判成其他品種之類數(Error by number of kernel)				
	台農私 20 號 Tainung Sen 20	台中私 10 號 Taichung Sen 10	台農 67 號 Tainung 67	台稔 8 號 Taikeng 8	台稔 9 號 Taikeng 9
台農私 20 號 Tainung Sen 20		0	0	0	0
台中私 10 號 Taichung Sen 10	0		0	0	0
台農 67 號 Tainung 67	0	0		24	0
台稔 8 號 Taikeng 8	0	0	23		0
台稔 9 號 Taikeng 9	0	0	2	0	



表 19 影像模式鑑別 2000 年品種中 6 參數組的誤判粒數

Table 19 Kernels of Classifying error in image model of classifying 2000 paddy rice with 6 parameters

6 個參數(6 parameters)					
	誤判成其他品種之顆數(Error by number of kernel)				
	台農秈 20 號 Tainung Sen 20	台中秈 10 號 Taichung Sen 10	台農 67 號 Tainung 67	台稈 8 號 Taikeng 8	台稈 9 號 Taikeng 9
台農秈 20 號 Tainung Sen 20		0	0	0	0
台中秈 10 號 Taichung Sen 10	0		0	0	0
台農 67 號 Tainung 67	0	0		71	0
台稈 8 號 Taikeng 8	0	0	8		1
台稈 9 號 Taikeng 9	0	0	6	21	

## 四. 結論

1997 年一期、1998 年一期、1999 年一期及 2000 年一期之 5 種台中區水稻品種(台農私 20 號、台中私 10 號、台農 67 號、台有 8 號、台有 9 號)，利用近紅外線光譜及主成份分析，近紅外線光譜資料作一次差分，影像參數經過設大最小對映的正規化後，均用 0.01 之學習速率及誤差平方總和(SSE)設為 5，代入倒傳遞類神經網路訓練以建立鑑別模式，用來鑑別 2000 年水稻品種，結果如下：

1. 使用近紅外線原始光譜作一次差分再經 PCA 分析，求得 7 個最佳主成分軸的投影值，代入倒傳遞類神經網路訓練以建立鑑別模式，1997 年、1998 年、1999 年及 2000 年 5 種水稻樣本之平均鑑別率為 94.7%，標準偏差為 2.25%。雖與張等人(2000)模式比較平均鑑別率為 95.9%及標準偏差為 5.91%，並沒有顯著提昇鑑別率，但降低了標準偏差。
2. 60 參數影像鑑別模式雖然比張等人(2000)建議 43 參數多，明顯地有利於鑑別出各品種的差異，平均鑑別率為 94%，標準偏差為 9.97%；43 參數組平均鑑別率為 90.2%，標準偏差為 12.54%。
3. 近紅外線鑑別模式的平均鑑別率為 94.7%及標準偏差為 2.25%，比較影像技術的平均鑑別率為 94%及標準偏差為 9.97%，來說並沒有顯著的差異，但近紅外線鑑別模式的標準偏差優於影像技術模式。

## 五. 致謝

本文承蒙國科會計劃編號 NSC89-2313 B-002-241 之經費支援下方得以完成，謹致謝

意，在研究期間承蒙台大周瑞仁教授與台大生物產業機電工程所前研究生張鴻文先生提供寶貴意見及台中農業改良場米質實驗室提供試驗材料，一併致謝。

## 六. 參考文獻

1. 張鴻文、蕭介宗、李盛銘、洪梅珠。2000。以近紅外線及影像技術鑑別水稻品種。農機學刊 9(4):1-16。
2. 李汪盛、蕭介宗。1996。資料處理對於稻穀食味主要成分之近紅外線校正線之影響。農業機械學刊 5(4):19-34。
3. 彭智平、李汪盛、蕭介宗。1997。以近紅外線分光光度計偵測影響稻米食味主要成分。中國農業工程學報 43(2):56-72
4. 李盛銘。1998。以近紅外線與影像技術鑑別水稻品種。台北：台灣大學農機系研究所碩士論文。
5. 夏賀元康、川村周三、伊藤和彥。1993。近赤外分光法對穀物成分測定之精度(第二報)。農業機械學會誌 55(1):109-115。
6. 陳正義、謝英雄、曾美倉。1989。判別水稻穀粒外型之品種差異-利用影像處理技術。中國統計學報 27 (7): 21-36。
7. 陳育仁。1994。近紅外線光譜分析技術用於穀物品質之偵測與分級 I, II。穀物非破壞性檢測。第二冊。初版, 35-52。台北：行政院農委會降低稻米製銷成本技術服務團, 稻米加工自動化技術服務團。
8. 張鴻文。2000。以近紅外線及影像技術鑑別不同年份之水稻品種。台北：台灣大學農業機械工程研究所碩士論文。
9. 葉怡成著。1993。類神經網路模式應用與實作, 儒林圖書有限公司。
10. 黃正華。1987。台灣稻作品種圖誌。農委會、農林廳、亞太糧肥技術中心編印。
11. 萬一怒、盛中德、楊智超、何榮祥。1998。胴裂米以類神經網路檢測之研究。農業工程學報 44 (1): 67-80。
12. 曾美倉、謝英雄、陳正義。1990。利用穀粒外形影像特徵之樣本平均數來判別水稻栽培種。台大農學院研究報告 30(4): 58-65。
13. 劉民卿、蕭介宗。1995。以近紅外線光譜儀感測稻米之含水率及蛋白質含量。農業機械學刊 4(3):1-14。
14. 謝志誠、林勁助。1994。以類神經網路技術作青椒形狀及大小之分級。農業機械學

15. 蔡養正。1988。作物微細構造在研究上之應用。科學農業 36(5-6)：144-164。
16. 呂理燾、林鳳池。1980。一些屬於菊科沙草科和禾本科雜草種子之掃描式電子顯微鏡之研究。中華民國雜草學會會刊 1(1)：23-51。
17. 謝英雄、曾美倉、陳正義。1989。水稻穀粒表面特徵之計量方法。科學農業 37(3-4)：115-117。
18. 陳正義。1989。以電腦視覺探討水稻栽培種穀粒外型之差異。台北：台灣大學農藝學研究所碩士論文。
19. Chen, Y. R., S. R. Delwiche, and W. R. Hruschka. 1995. Classification of hard red wheat by feed forward backpropagation neural network. *Cereal Chem.* 72(3)：317-319.
20. Delwiche, S. R., Y. R. Chen, and W. R. Hruschka. 1995. Differentiation of hard red wheat by near-infrared analysis of bulk samples. *Cereal Chem.* 72(3)：243-247.
21. Travis, A. J. and S. R. Draper. 1985. A computer based system for the recognition of seed shape. *Seed Sci. & Technol.* 14：715-724.
22. Velleman, P. F. and R. E. Welsch. 1981. Efficient computing of regression diagnostics. *The Amer. Statistician* . 35:234-242.
23. Yamashita, R. 1993. New technology in grain postharvesting. Farm Machinery Industrial Research Corp., Tokyo 101, Japan.
24. Zayas, I., F. S. Lai and Y. Pomeranz. 1986. Discrimination between wheat classes and varieties by image analysis. *Cereal Chem.* 63(1)：52-66.