

國立臺灣大學工學院土木工程學研究所



碩士論文

Department of Civil Engineering

College of Engineering

National Taiwan University

Master Thesis

運用 AI 進行不同料源粗粒料辨識

Artificial Intelligence for Recognition and  
Classification of Different Sources of Coarse  
Aggregates

項灝先

Luo-Xian Xiang

指導教授：詹穎雯 博士

Advisor: Yin-Wen Chan, Ph.D.

中華民國 112 年 8 月

August, 2023

國立臺灣大學碩士學位論文  
口試委員會審定書  
MASTER'S THESIS ACCEPTANCE CERTIFICATE  
NATIONAL TAIWAN UNIVERSITY



運用 AI 進行不同料源粗粒料辨識

Artificial Intelligence for Recognition and  
Classification of Different Sources of Coarse  
Aggregates

本論文係項灝先(R10521240)在國立臺灣大學土木工程學系結構工程組完成之碩士學位論文，於民國 112 年 7 月 28 日承下列考試委員審查通過及口試及格，特此證明。

The undersigned, appointed by the Department / Institute of Civil Engineering Structural Engineering on 28 July, 2023 have examined a Master's thesis entitled above presented by Yu-Chiuan Tsai (R10521243) candidate and hereby certify that it is worthy of acceptance.

口試委員 Oral examination committee:

詹穎雯  
(指導教授 Advisor)

廖文正

楊仲家

系主任

葛宇甯

Handwritten signature of 詹穎雯 in black ink.

Handwritten signature of 廖文正 in black ink.

Handwritten signature of 楊仲家 in blue ink.

Handwritten signature of 葛宇甯 in blue ink.



## 誌謝

兩年的碩士生活轉眼即逝，在這兩年期間不只有學習研究與課業，對於做人處事以及生活態度有很大的改變與進步，也認識了許多人並獲得很多人的幫助。

首先，感謝我的指導教授 詹穎雯老師，在研究上提供我許多方向，教導我做研究的想法以及如何整理文獻的成果，給我不同觀點思考研究，使我的研究更豐富。

感謝廖文正老師，教導我研究中對於研究目的的深入以及數據解釋與意義，建構更完整的研究框架，讓我的研究更精進；感謝楊仲家老師在口試時的寶貴意見，指點出我撰寫論文時疏漏以及不足之處，使得我的研究更加完整。

感謝銘智學長，提供我非常多的研究資料，也會與我討論研究的細項以及發展，給了我許多研究上的幫助。感謝威佑學長在這兩年來對我的關心與幫助。另外，感謝營建院季軒學長與瑞紋學姊幫我處理實驗中所遇到的狀況，不管是材料運輸或是報帳的事情，使我能順利完成實驗。

感謝 812 的學長姐承禹、易霖、丞駿、璧銘、宗佑、淳憶，在碩一時提供考古題或是研究上的幫助，還有一起打橋牌跟電動的時光，讓我的碩士生活變得更佳豐富。

感謝 812 的同學文國、威翰、彥輔、灝先、宇筌、奕親、澧恩、映捷、喬艾，在碩一修課過程中，大家一起討論互相幫助，碩二期間，每天在研究室一起吃飯、做研究、打牌，都是珍貴的回憶，願大家未來一帆風順。

感謝 812 的學弟竣帆、峻輔、濂淞、尚恩、瀚仁、庭安、奕霖，謝謝你們在我實驗需要時給予幫助，希望各位在往後的研究一切順利。

最後感謝我的家人在這兩年來支持與鼓勵，也感謝我的女友瑞婷，在我壓力大的時候給予我支持，總是陪在我身邊，讓我能夠克服困難，順利畢業。



## 摘要

粗粒料是水泥預拌混凝土中的重要組成部分，其級配曲線和粗粒料之扁平、細長和扁長率對混凝土的工作性能和力學性能具有重要影響。因此，粒徑識別和分析成為了確保混凝土品質和性能的關鍵步驟。

傳統上，兩者的評估主要依賴於機械篩析法以及比例測徑器，需要耗費大量時間且操作繁瑣。然而，隨著影像處理和機器學習技術的發展，影像的粒徑分析方法逐漸受到關注。通過分析粒料影像來獲取粒徑分佈信息，具有快速、準確的優勢。

本研究針對北部預拌混凝土廠常用的花蓮和閩侯料源之粗粒料進行尺寸辨識，旨在有效即時監測三分石和六分石的資訊。為了獲取粒徑分佈的準確信息，本研究採用了影像辨識軟體 Image Pro Plus 來獲取粒料的影像資料，使用 Python 程式語言編碼機器學習模型進行數據處理以及分類，對粒料的料源和尺寸進行了詳盡的分析。利用分類結果，以及粗粒料篩析法，以獲取篩號停留重量，並建立了與影像數據的關係，最終得到不同料源之間的占比以及生成級配曲線，其成果較傳統文獻上的誤差更小，甚至在 D90 和 D50 有接近零誤差的表現。

粗粒料的扁平、細長和扁長率在傳統的影像辨識方法通常需要多個攝影機來準確捕捉。本研究通過大量平面資料的採樣以及實際重量的估算，建立了影像平均厚度與真實量測粒料厚度之間的關係，推估粒料在平面影像上的厚度，並最終獲得粒料的扁平、細長和扁長率結果，可以在短時間內得到與實際試驗結果相近，特別是扁平率以及細長率誤差較小。

綜上所述，本研究結合了影像辨識和機器學習技術，通過分析粗粒料的影像資料來評估其粒徑分佈和形狀特徵。該方法不僅具有高準確性和效率，而且能夠提供即時的資訊，支援混凝土生產和品質管理的過程，對於提高生產效率、節省成本並確保混凝土品質具有重要意義。

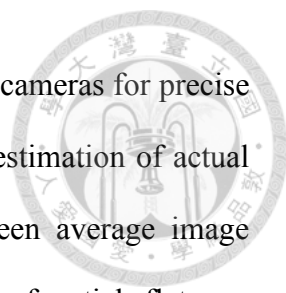
**關鍵字：**影像辨識、機器學習、級配曲線、粗粒料、扁平率



## Abstract

Coarse aggregates are essential components of concrete. Their gradation curves and flat and elongated particles, significantly affect the workability and mechanical properties of concrete. Therefore, accurate particle size distribution and analysis are crucial for ensuring concrete quality and performance.

Traditionally, particle size and shape evaluation relied on labor-intensive methods such as sieving method and the use of proportioning calipers. However, with advancements in image processing and machine learning, image-based particle analysis methods have gained attention. These methods utilize digital image processing techniques to analyze particle images and provide instant and accurate information on particle size distribution. This study focuses on the commonly used coarse aggregates from Hualien and Minhou in ready concrete plants. The aim is to effectively monitor the information of three-quarter-inch and one-half-inch aggregates, as these particle sizes are representative of three-quarter-inch and one-half-inch aggregates. To obtain accurate particle size distribution information, the image processing software Image Pro Plus is used to acquire particle image data. Through machine learning classification methods, comprehensive analysis of the particle sources and sizes is performed. The classification results reveal the proportions among different sources and facilitate the generation of gradation curves. Additionally, traditional sieve analysis is employed to determine the weights of particles in each sieve layer, and the relationship between sieve weights and image data is established. This allows a thorough understanding of the proportions among different sources and the generation of accurate gradation curves, with minimal errors as compared to traditional methods, including approach zero errors in D90 and D50.



Assessing flat and elongated particles traditionally requires multiple cameras for precise capture. In this study, a large amount of planar data sampling and estimation of actual particle weights are conducted to establish the relationship between average image thickness and measured particle thickness. This enables the estimation of particle flatness, elongation, and angularity from planar images, yielding results that closely match actual test results, particularly with small errors in flatness and elongation ratios.

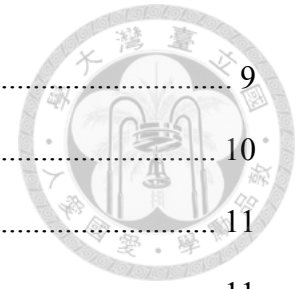
In conclusion, this study combines image recognition and machine learning techniques to evaluate the particle size distribution and shape characteristics of coarse aggregates using image data analysis. The proposed method demonstrates high accuracy and efficiency, providing real-time information to support concrete production and quality management. It is of significant importance in improving production efficiency, reducing costs, and ensuring concrete quality.

**Keywords:** image recognition, machine learning, gradation curve, coarse aggregates, flat and elongated particle

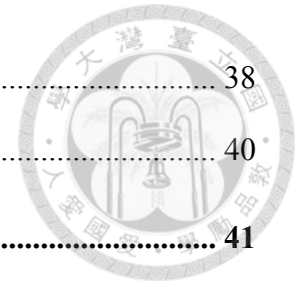
# 目錄



誌謝 .....	i
摘要 .....	ii
Abstract .....	iii
目錄 .....	v
圖目錄 .....	ix
表目錄 .....	xi
第一章、緒論 .....	1
1.1. 研究動機 .....	1
1.2. 研究範圍與內容 .....	2
1.3. 研究流程 .....	2
第二章、文獻回顧 .....	5
2.1. 材料介紹 .....	5
2.1.1 粒料 .....	5
2.1.2 石英斑岩 .....	5
2.1.3 石灰石 .....	5
2.2. 粗粒料篩析法 .....	6
2.2.1 級配曲線 .....	6
2.3. 粗粒料中扁平、細長或扁長顆粒含量試驗法 .....	6
2.3.1 扁平、細長或扁長顆粒定義 .....	7
2.3.2 扁平、細長或扁長顆粒對混凝土之影響 .....	7
2.4. 影像辨識 .....	7
2.4.1 粒料尺寸分類 .....	8

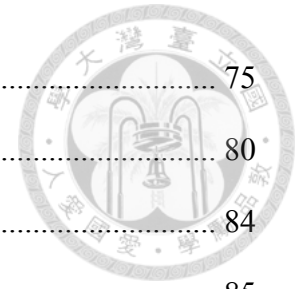


2.4.2 影像平面與立體之關係.....	9
2.4.3 影像粒料形狀.....	10
2.5. 機器學習.....	11
2.5.1 前言.....	11
2.5.2 機器學習類型.....	13
2.5.3 機器學習流程.....	14
2.5.4 機器學習評估.....	15
2.6. 特徵工程.....	17
2.6.1 特徵選擇.....	17
2.7. 演算法.....	19
2.7.1 LogisticRegression.....	19
2.7.2 DecisionTree.....	19
2.7.3 RandomForest.....	20
2.7.4 Xgboost.....	21
2.7.5 Stacking.....	22
<b>第三章、粗粒料試驗與攝影.....</b>	<b>23</b>
3.1. 實驗材料.....	23
3.2. 粗粒料試驗.....	32
3.2.1 器材.....	32
3.2.2 試驗流程.....	32
3.2.3 粗粒料篩析法.....	34
3.2.4 粗粒料中扁平、細長或扁長顆粒含量試驗法.....	36
3.3. 影像拍攝.....	37
3.3.1 器材.....	37



3.3.2 攝影流程 .....	38
3.3.3 取樣 .....	40
<b>第四章、分析計畫與方法 .....</b>	<b>41</b>
4.1. 分析計畫背景 .....	41
4.2. 影像前處理 .....	41
4.2.1 處理流程 .....	42
4.2.2 模型資料 .....	43
4.2.3 資料特徵 .....	44
4.3. 料源辨識模型 .....	46
4.3.1 特徵工程 .....	47
4.3.2 分類模型 .....	50
4.3.3 成效評估 .....	53
4.4. 尺寸分類模型 .....	58
4.4.1 特徵工程 .....	59
4.4.2 分類模型 .....	62
4.4.3 成效評估 .....	66
4.5. 質量轉換係數 .....	69
4.5.1 質量轉換係數定義 .....	69
4.5.2 質量轉換係數精度 .....	69
4.6. 影像平均厚度 .....	71
4.6.1 影像平均厚度定義 .....	71
4.6.2 影像平均厚度結果 .....	72
<b>第五章、分析結果與討論 .....</b>	<b>75</b>
5.1. 粗粒料篩析法結果 .....	75

5.1.1 粗粒料篩析法結果.....	75
5.1.2 粗粒料篩析法分析.....	80
5.2. 扁平、細長或扁長顆粒.....	84
5.2.1 扁平、細長或扁長顆粒結果.....	85
5.2.2 扁平、細長或扁長顆粒分析.....	86
<b>第六章、 結論與建議.....</b>	<b>89</b>
6.1. 結論.....	89
6.2. 建議.....	91
<b>參考文獻.....</b>	<b>93</b>





## 圖目錄

圖 1.3-1 訓練資料建立模型與參數流程圖.....	3
圖 1.3-2 測試資料實際操作流程圖.....	4
圖 2.4-1 影像處理分析粒料參數[4].....	8
圖 2.4-2 影像粒徑近似橢圓示意圖 [6].....	9
圖 2.4-3 WipShape 形狀量測系統[7].....	10
圖 2.5-1 偏差和方差權衡圖 [11].....	13
圖 3.3-1 拍攝環境.....	38
圖 3.3-2 粒料實際拍攝影像.....	39
圖 4.2-1 粒料影像圖.....	42
圖 4.2-2 辨識結果.....	43
圖 4.2-3 粗粒料訓練組別以及測試組別.....	44
圖 4.3-1 料源辨識模型之各項特徵 F-value.....	47
圖 4.3-2 花蓮地區.....	49
圖 4.3-3 閩侯地區.....	49
圖 4.3-4 料源辨識模型之學習曲線 演算法：Stacking.....	52
圖 4.3-5 混淆矩陣 組別 4.....	54
圖 4.3-6 混淆矩陣 組別 5.....	55
圖 4.3-7 混淆矩陣 組別 6.....	56
圖 4.4-1 尺寸分類模型之各項特徵 F-value 花蓮料源.....	59
圖 4.4-2 尺寸分類模型之各項特徵 F-value 閩侯料源.....	60
圖 4.4-3 尺寸分類模型之學習曲線 花蓮料源 演算法：RandomForest.....	64
圖 4.4-4 尺寸分類模型之學習曲線 閩侯料源 演算法：RandomForest.....	65
圖 4.4-5 最終分類結果分區分析示意圖.....	67

圖 4.4-6 最終分類結果之混淆矩陣 組別 4.....	67
圖 4.4-7 最終分類結果之混淆矩陣 組別 5.....	68
圖 4.4-8 最終分類結果之混淆矩陣 組別 6.....	68
圖 4.5-1 花蓮地區之質量轉換係數.....	70
圖 4.5-2 閩侯地區之質量轉換係數.....	70
圖 5.1-1 級配曲線結果 花蓮 4.....	75
圖 5.1-2 級配曲線結果 花蓮 5.....	76
圖 5.1-3 級配曲線結果 花蓮 6.....	77
圖 5.1-4 級配曲線結果 閩侯 4.....	78
圖 5.1-5 級配曲線結果 閩侯 5.....	79
圖 5.1-6 級配曲線結果 閩侯 6.....	80



# 表目錄



表 2.7-1 混淆矩陣.....	16
表 3.1-1 試驗結果一覽表 花蓮料源.....	23
表 3.1-2 粗粒料比重(密度)及吸水性(率)試驗報告 花蓮料源(三分石).....	24
表 3.1-3 粗粒料比重(密度)及吸水性(率)試驗報告 花蓮料源(六分石).....	25
表 3.1-4 粗粒料中扁平、細長、扁長粒料試驗報告-1 花蓮料源(三分石).....	25
表 3.1-5 粗粒料中扁平、細長、扁長粒料試驗報告-2 花蓮料源(三分石).....	25
表 3.1-6 粗粒料中扁平、細長、扁長粒料試驗報告-1 花蓮料源(六分石).....	26
表 3.1-7 粗粒料中扁平、細長、扁長粒料試驗報告-2 花蓮料源(六分石).....	26
表 3.1-8 粗粒料篩分析試驗報告 花蓮料源(三分石).....	26
表 3.1-9 粗粒料篩分析試驗報告 花蓮料源(六分石).....	27
表 3.1-10 試驗結果一覽表 閩侯料源.....	27
表 3.1-11 粗粒料比重(密度)及吸水性(率)試驗報告 閩侯料源(三分石).....	28
表 3.1-12 粗粒料比重(密度)及吸水性(率)試驗報告 閩侯料源(六分石).....	29
表 3.1-13 粗粒料中扁平、細長、扁長粒料試驗報告-1 閩侯料源(三分石).....	29
表 3.1-14 粗粒料中扁平、細長、扁長粒料試驗報告-2 閩侯料源(三分石).....	29
表 3.1-15 粗粒料中扁平、細長、扁長粒料試驗報告-1 閩侯料源(六分石).....	30
表 3.1-16 粗粒料中扁平、細長、扁長粒料試驗報告-2 閩侯料源(六分石).....	30
表 3.1-17 粗粒料篩分析試驗報告 閩侯料源(三分石).....	30
表 3.1-18 粗粒料篩分析試驗報告 閩侯料源(六分石).....	31
表 3.2-1 粗粒料試樣最少量.....	33
表 3.2-2 粗粒料篩析法之結果 花蓮 1.....	34
表 3.2-3 粗粒料篩析法之結果 花蓮 2.....	35
表 3.2-4 粗粒料篩析法之結果 花蓮 3.....	35

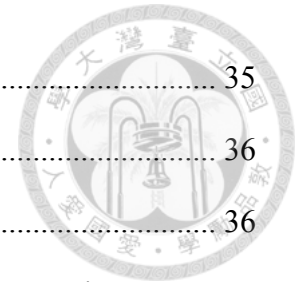


表 3.2-5 粗粒料篩析法之結果 閩侯 1.....	35
表 3.2-6 粗粒料篩析法之結果 閩侯 2.....	36
表 3.2-7 粗粒料篩析法之結果 閩侯 3.....	36
表 3.2-8 扁平、細長或扁長顆粒含量試驗法之結果 左：花蓮、右：閩侯 組別 1 .....	36
表 3.2-9 扁平、細長或扁長顆粒含量試驗法之結果 左：花蓮、右：閩侯 組別 2 .....	37
表 3.2-10 扁平、細長或扁長顆粒含量試驗法之結果 左：花蓮、右：閩侯 組別 3 .....	37
表 4.3-1 F-value 最高三名特徵之統計數據 花蓮料源 .....	48
表 4.3-2 F-value 最高三名特徵之統計數據 閩侯料源 .....	49
表 4.3-3 每個模型之 5 次迭代以及總平均的 Accuracy .....	51
表 4.3-4 料源辨識模型評估報告 組別 4.....	53
表 4.3-5 料源辨識模型評估報告 組別 5.....	54
表 4.3-6 料源辨識模型評估報告 組別 6.....	55
表 4.3-7 Density(red)、Density(min)之統計數據 花蓮料源 .....	57
表 4.3-8 Density(red)、Density(min)之統計數據 閩侯料源 .....	58
表 4.4-1 Diameter (min)之統計數據 花蓮料源 .....	61
表 4.4-2 Diameter (min)之統計數據 閩侯料源 .....	62
表 4.4-3 每個模型之 5 次迭代以及總平均的 Accuracy 花蓮料源 .....	64
表 4.4-4 每個模型之 5 次迭代以及總平均的 Accuracy 閩侯料源 .....	65
表 4.5-1 訓練資料建立之質量轉換係數.....	70
表 4.6-1 $\bar{t}_{a,max}$ 、 $\bar{t}_i$ 、 $t_{ratio}$ 結果-花蓮 1.....	72
表 4.6-2 $\bar{t}_{a,max}$ 、 $\bar{t}_i$ 、 $t_{ratio}$ 結果-花蓮 2.....	72
表 4.6-3 $\bar{t}_{a,max}$ 、 $\bar{t}_i$ 、 $t_{ratio}$ 結果-花蓮 3.....	73



表 4.6-4 $\bar{t}_{a,max}$ 、 $\bar{t}_i$ 、 $t_{ratio}$ 結果-閩侯 1.....	73
表 4.6-5 $\bar{t}_{a,max}$ 、 $\bar{t}_i$ 、 $t_{ratio}$ 結果-閩侯 2.....	73
表 4.6-6 $\bar{t}_{a,max}$ 、 $\bar{t}_i$ 、 $t_{ratio}$ 結果-閩侯 3.....	73
表 4.6-7 $t_{ratio}$ 三組平均之結果.....	74
表 5.1-1 粗粒料篩析法以及影像辨識結果 花蓮 4.....	76
表 5.1-2 粗粒料篩析法以及影像辨識結果 花蓮 5.....	76
表 5.1-3 粗粒料篩析法以及影像辨識結果 花蓮 6.....	77
表 5.1-4 粗粒料篩析法以及影像辨識結果 閩侯 4.....	78
表 5.1-5 粗粒料篩析法以及影像辨識結果 閩侯 5.....	79
表 5.1-6 粗粒料篩析法以及影像辨識結果 閩侯 6.....	80
表 5.1-7 分類最終成果總表 單位：顆.....	82
表 5.1-8 D10、D50、D90 結果比較 花蓮 4.....	83
表 5.1-9 D10、D50、D90 結果比較 花蓮 5.....	83
表 5.1-10 D10、D50、D90 結果比較 花蓮 6.....	83
表 5.1-11 D10、D50、D90 結果比較 閩侯 4.....	84
表 5.1-12 D10、D50、D90 結果比較 閩侯 5.....	84
表 5.1-13 D10、D50、D90 結果比較 閩侯 6.....	84
表 5.2-1 扁平、細長或扁長顆粒結果 左：真實、右：影像 花蓮 4.....	85
表 5.2-2 扁平、細長或扁長顆粒結果 左：真實、右：影像 花蓮 5.....	85
表 5.2-3 扁平、細長或扁長顆粒結果 左：真實、右：影像 花蓮 6.....	85
表 5.2-4 扁平、細長或扁長顆粒結果 左：真實、右：影像 閩侯 4.....	86
表 5.2-5 扁平、細長或扁長顆粒結果 左：真實、右：影像 閩侯 5.....	86
表 5.2-6 扁平、細長或扁長顆粒結果 左：真實、右：影像 閩侯 6.....	86



# 第一章、緒論


## 1.1. 研究動機

混凝土在建築和基礎工程中扮演著重要的角色。為了確保混凝土的工作性和硬固性質，級配曲線是一個關鍵的因素。級配曲線指的是粒料中不同顆粒大小的比例和分佈。如果級配不合規範或不均勻，可能導致混凝土的工作性下降和強度耐久性的不足。當級配不理想時，為了維持混凝土的工作性，施工人員可能需要增加拌和水量，過多的水會對混凝土的品質產生負面影響，並可能導致工作性的劣化。此外，即使增加水量，也無法完全調整回級配到理想範圍內。在這種情況下，還需要調整砂子的比例(砂率)，以及增加粉體的用量，這將增加成本和碳排放。

在混凝土生產中，了解粒料的料源對於級配的影響是非常重要的，因為不同砂石廠的料源特性存在差異。因此，可以預期級配的良好與否，並採取相應的防範措施。在實際應用中，有時需要進行料源的轉換，即由舊料轉換為新料。然而，在進行轉換時常常會出現問題。雖然我們知道需要進行轉換，但是在砂石庫存量很大且不確定何時會用到的情況下，轉換可能帶來困擾。

預拌混凝土廠之粗粒料由三分石(4.75~9.5mm)以及六分石(9.5~19mm)組成，若能在生產過程中即時監測和量測三分石和六分石的級配，就可以及時進行微調比例，以確保混合料的級配符合規範要求。例如，微調三分石和六分石的比例，使級配符合規範需求，可避免為了調整水膠比而增加成本和碳排放，若無法及時發現級配偏差，可能導致混凝土不符合工程需求，需要退料或可能導致工程缺失。

此外，粗粒料的顆粒扁平、細長和扁長率也會影響混凝土用水量。當級配不合理時，需要調整砂率和粉體用量。如果無法及時發現粒料問題，將對品質和成本產生負面影響。因此，希望能夠即時監控所使用的粒料，一旦發現問題，立即進行製程管控調整，方能節省成本並減少碳排放。



基於以上的考慮，本研究的動機是開發一種能夠即時監測和量測三分石和六分石級配以及粒料的顆粒扁平、細長和扁長率的方法。透過影像辨識和機器學習，期望能夠實現快速、準確的測量和調整。這將為預拌廠生產提供寶貴的信息，幫助調整配比，確保混凝土的工作性和硬固性質符合要求，同時也能節省成本和減少碳排放，也可以避免混凝土不符合工程需求和減少工程缺失的風險。

## 1.2. 研究範圍與內容

本研究的範圍涵蓋了預拌混凝土廠常用的花蓮和來自中國福建省閩侯料源之粗粒料，旨在有效監測三分石和六分石的資訊。研究重點在於探討 1/2"、3/8"和#4 停留之粒料，因為停留之粒料大小涵蓋了三分石和六分石。由於尚未建立台灣粗粒料的資料庫，本研究採用影像辨識軟體 Image Pro Plus 來獲取粒料的影像資料。透過機器學習的分類方法，對粒料的料源和尺寸進行了分類，同時也進行了粗粒料篩析法以獲取各層重量，並建立了與影像數據的關係。這樣做可以了解不同料源之間的占比以及生成級配曲線。

傳統的影像辨識方法通常需要多個攝影機來準確捕捉顆粒的扁平、細長和扁長率。本研究採用大量平面資料的採樣以及實際重量的估算，建立了影像平均厚度與真實量測粒料厚度之間的關係。憑藉這種方法，能夠推估粒料在平面影像上的厚度，並最終獲得粒料的扁平、細長和扁長率等特徵。

總結而言，本研究的範圍包括了預拌混凝土廠常用的花蓮和閩侯料源之粗粒料。利用影像辨識和機器學習的方法對粒料進行分析，以了解不同料源之間的占比和生成的級配線。同時，提出了一種基於平面影像資料的方法，推估粒料的扁平、細長和扁長率等。這些研究成果將對混凝土生產和品質管理提供有價值的信息，有助於調整配比、確保混凝土的工作性和硬固性質符合要求。

## 1.3. 研究流程

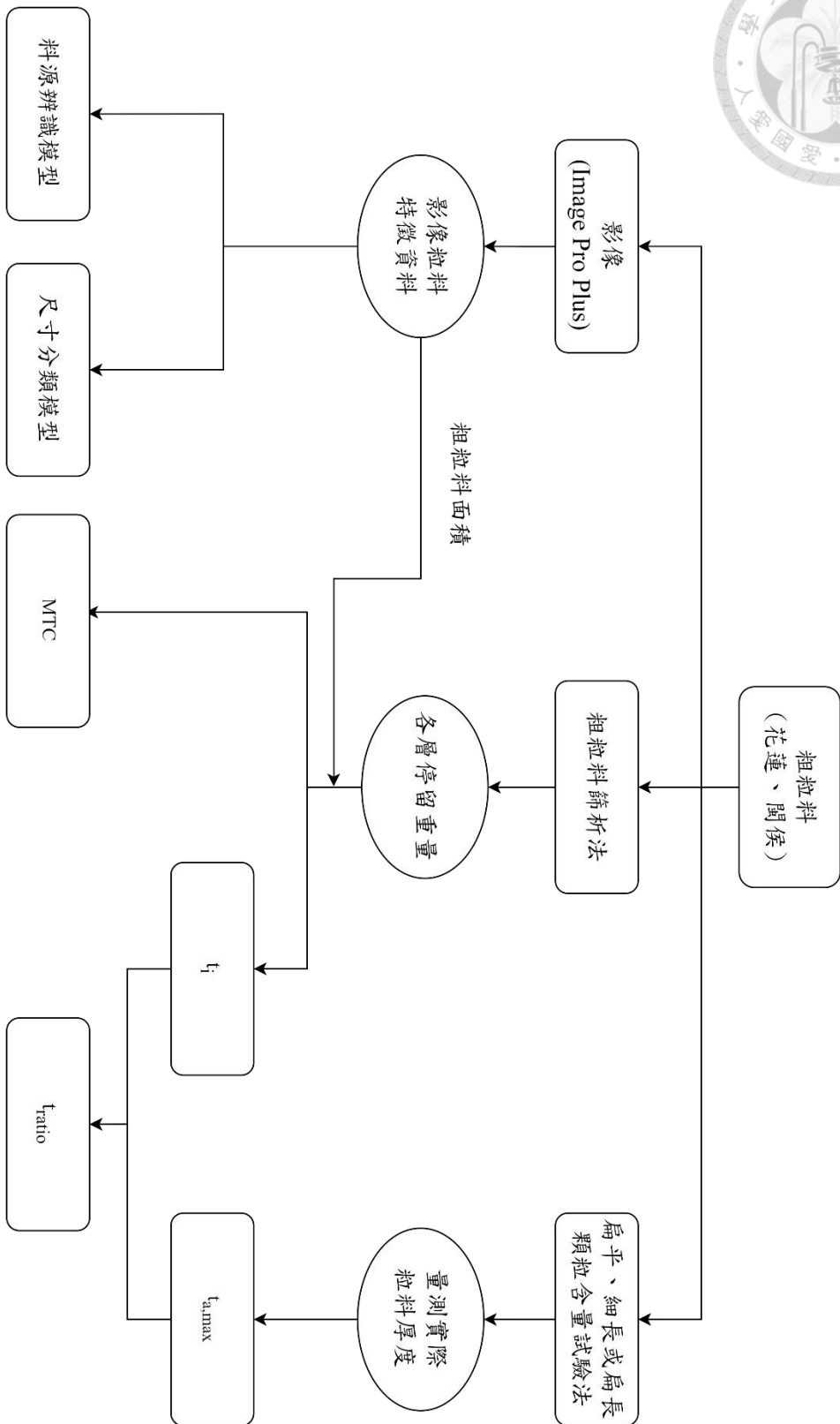


圖 1.3-1 訓練資料建立模型與參數流程圖

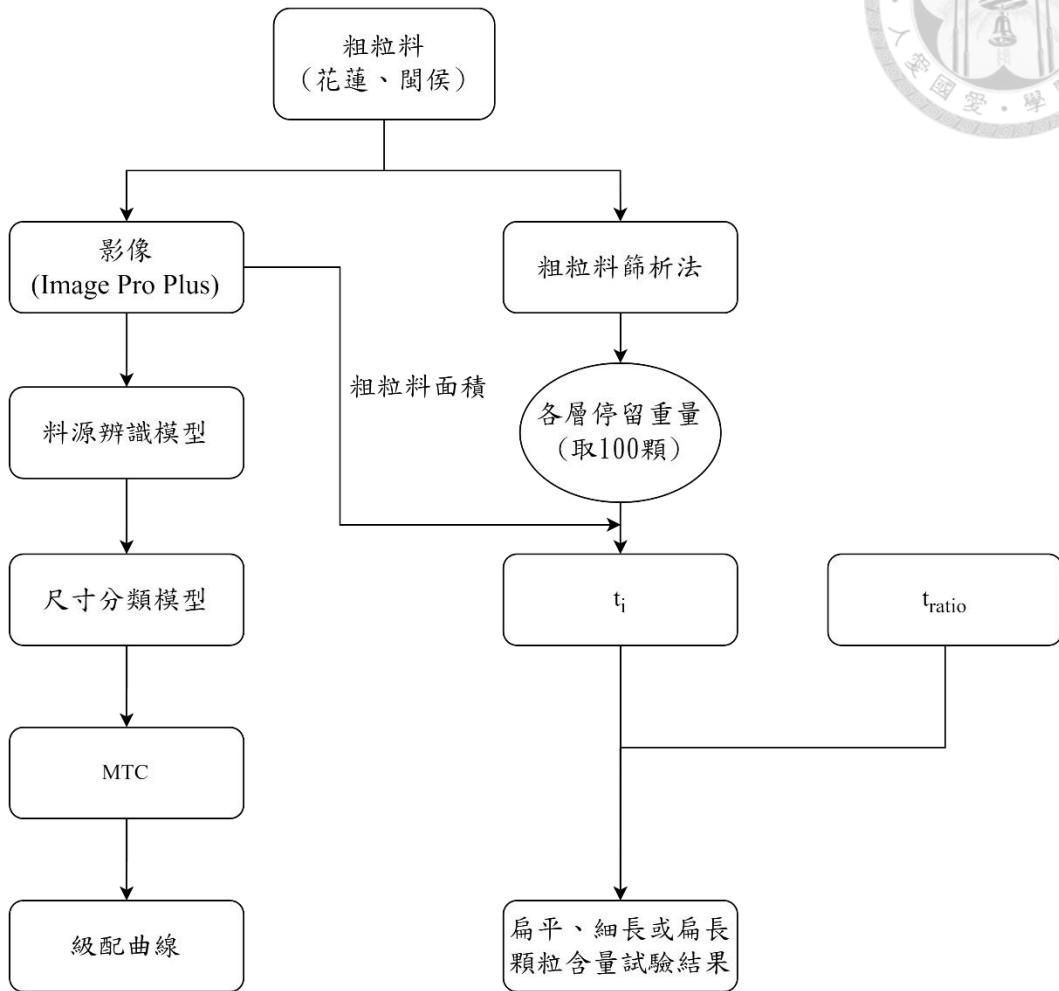


圖 1.3-2 測試資料實際操作流程圖

## 第二章、文獻回顧



### 2.1. 材料介紹

#### 2.1.1 粒料

粒料是一種用於建築、道路、水泥等建設工程的材料，如砂、礫石、碎石、破碎的水硬性混凝土、或是爐石等。粒料在建築工程中通常用作混凝土、瀝青混凝土、砂漿等材料的主要成分，約占混凝土體積 60%~70%。粒料的選擇與篩分尺寸的準確控制是十分重要的，粒料除了會影響混凝土塑性，也會影響混凝土耐久性、強度與熱學性質等。粒料依顆粒大小可分為細粒料和粗粒料兩類，粗粒料與細粒料以 4 號篩(#4, 4.75mm)來做區分，通過#4 者稱為細粒料；留在#4 以上者稱為粗粒料。

#### 2.1.2 石英斑岩

閩侯料源之粒料材質主要為石英斑岩，石英斑岩是一種由石英和長石等礦物組成的火成岩，它具有特殊的斑狀結構和顯著的石英含量。在石英斑岩中，石英是最主要的礦物成分之一，它以顆粒狀或斑晶狀的形式存在，呈現出豐富的顏色和紋理變化。石英斑岩中的長石則通常以鉀長石和鈉長石為主，它們與石英共同形成了岩石的基本骨架。

石英斑岩的顏色和外觀取決於其中的礦物成分和結晶結構。一些石英斑岩可能呈灰色、粉紅色、黃色或綠色，並且在岩石的表面可以觀察到斑麼狀的紋理。這種斑狀結構是由於岩漿在地下冷卻時，礦物晶體逐漸生長形成的。

#### 2.1.3 石灰石

花蓮料源之粒料材質主要為石灰石，石灰石是一種常見的沉積岩，主要由碳酸鈣( $\text{CaCO}_3$ )組成。它形成於海洋或湖泊中，是由生物殘骸、貝殼、珊瑚、藻類等有機物在長時間的壓力和化學作用下沉積而成。

石灰石的外觀通常呈白色或淺灰色，但也有其他顏色變體，如黃色、粉紅色、灰褐色等，這些顏色的變化主要由於其中的雜質成分。石灰石的質地可以從細膩到粗糙不等，取決於沉積過程中的壓力和結晶度



## 2.2. 粗粒料篩析法

粒料篩分析是一種重要的測試方法，用於評估和分析石材、砂子、礫石等材料的顆粒大小分佈。這些材料的顆粒大小分佈直接影響混凝土的力學性質，例如密度、孔隙率、強度和耐久性。

### 2.2.1 級配曲線

級配曲線(Gradation curve)，又稱粒度分佈曲線(Particle size distribution curve)，是描述一種材料中不同粒徑顆粒的分佈情況的曲線。它是工程和材料科學中常用的一種圖示方式，用於評估和分析材料的顆粒組成。

級配曲線通常將顆粒的粒徑大小(通常以直徑或孔徑表示)作為橫軸，將累積顆粒百分比(按質量或體積計算)作為縱軸。曲線上的每個點表示該粒徑以下的顆粒累積百分比。例如，曲線在直徑為 1 毫米的位置處於 50%的過篩率，則代表該材料中 50%的顆粒直徑小於 1 毫米，其中這種指標又稱為 D50。

級配曲線對於工程設計和材料性質的評估非常重要。例如，在道路建設中，根據材料的級配曲線可以確定最適合的骨料組合，以獲得良好的工程性能。在混凝土和瀝青混凝土的配製中，級配曲線可用於確定粒料的最佳比例，以實現所需的強度和耐久性。級配曲線是評估和描述材料顆粒組成的重要工具，它提供了有關顆粒大小分佈的定量和定性信息，有助於工程設計和材料性能的分析 and 優化。

## 2.3. 粗粒料中扁平、細長或扁長顆粒含量試驗法



粗粒料中扁平、細長或扁長顆粒含量試驗法，用於評估粗粒料樣品中具有不規則形狀的顆粒的含量比例。這些不規則形狀的顆粒可能是扁平、細長或扁長的，在粗粒料中可能對材料特性和性能產生影響。

### 2.3.1 扁平、細長或扁長顆粒定義

根據中華民國國家標準 CNS15171[1]定義扁平、細長或扁長顆粒，分述如下：

1. 扁長顆粒:顆粒長厚比值大於規定者。
2. 扁平或細長顆粒:顆粒寬度與厚度比值或長度與厚度比值大於規定者。
3. 長度:粒料之最大尺寸。
4. 寬度:垂直於長度之平面中之最大尺寸。
5. 厚度:垂直於長度及寬度之最大尺寸。

### 2.3.2 扁平、細長或扁長顆粒對混凝土之影響

ASTM D4791[2]以及 CNS15171[1]提到粗粒料中的扁平 and 細長顆粒比其他粒狀顆粒更容易斷裂。當粗粒料斷裂時，篩分析結果可能會發生變化，對混凝土產生不利影響。此外，對於某些構造物，粗粒料中的扁平 and 細長顆粒可能會干擾密實性並增加澆置之困難度。

## 2.4. 影像辨識

影像辨識是一種人工智慧技術，可以通過影像的方式獲取拍攝到的物體數據。在文獻中，已有研究使用影像辨識技術對粗粒料進行分析，以獲得級配曲線以及顆粒的扁平、細長和扁長率等。然而，文獻中的影像辨識方法往往需要進行許多手動影像前處理，並使用多個攝影器材，這些過程耗時且複雜。

本研究參考了文獻中的影像辨識方法，但對其中的數值處理進行了改進且將粒料分開拍攝，以獲得更好的結果且可省去影像前處理的步驟，通過適當的數值處理方法，能夠提高影像辨識的準確性和效率。



## 2.4.1 粒料尺寸分類

在影像辨識粒料中繪製級配曲線，其關鍵的一個因素就是粒料的尺寸分類，要確實分類粒料的尺寸到相對應之篩號，以便進一步分析和評估粒料的重量。傳統文獻在影像的粒料辨識領域中，主要利用粒料的單一特徵，亦是單因子的評估方法分類粒料的尺寸，例如 Size(width)、短軸(Minor Axis)，來量化其尺寸，下述為不同文獻的方法：

1. Maiti Abhik et al.[3]透過比較主軸、次軸、邊界框長度和等效直徑等決定粒料尺寸，20mm 以下粒料使用次軸，20mm 以上使用等效直徑有較好的分類成效，D50 最小誤差 5.9%。
2. Mert Aydin et al.[4]使用 Fmax 作為粒料尺寸分類，其結果表明，使用影像處理技術獲得的結果與篩分析實驗數有 95.22%的相似性。

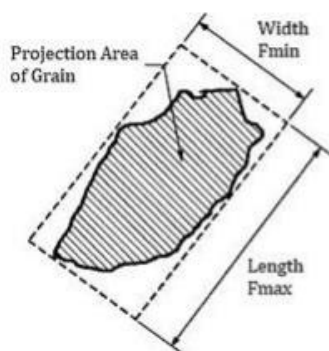


圖 2.4-1 影像處理分析粒料參數[4]

3. Xiaohan Yang et al.[5]利用數位影像處理技術進行煤碎片的粒徑篩分析曲線，其中煤碎片的尺寸使用等校面積橢圓之短軸表示，因為通常使用橢球體來表示碎片的不規則形狀。
4. 江田正敏[6]粒徑的判斷方法是通過最小二乘法將粒料進行橢圓逼近，並將該橢圓的短軸作為粒徑進行評估，該研究顯示尺寸分類準確率約 80%，如圖 2.4-2 所示。

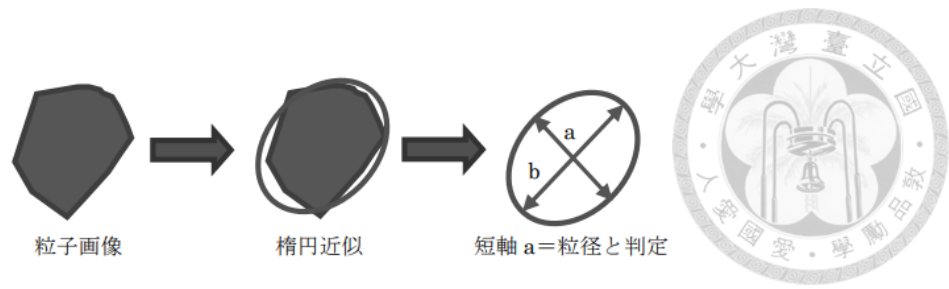


圖 2.4-2 影像粒徑近似橢圓示意圖 [6]

透過上述文獻可以觀察到使用單因子評估粒料的尺寸分類具有侷限性，相較於單因子評估方法，本研究使用多因子的評估粒料尺寸，能夠更全面地反映不同形狀參數對粒料尺寸分類的影響，提供更精確的結果。

### 2.4.2 影像平面與立體之關係

在處理立體物體時，傳統的平面影像辨識方法存在一定的局限性。平面影像無法捕捉到對象的立體信息和空間結構，從而導致在一些場景中的誤判和辨識錯誤，在實際應用層面，影像辨識之物體本身具有立體結構和空間資訊。為了更全面地理解影像數據並實現更精確的辨識，建立平面資料與立體之間的關係至關重要。部分文獻中認為相同料源的同一篩號停留之粒料，在大量採樣的情況下會有相似的趨勢，因粒料特性以及破碎行為相似，下述為不同文獻的方法：

1. 質量轉換係數：江田正敏[6]將各層停留粒料重量與影像辨識面積相除，獲得質量轉換係數(Mass transfer coefficient, MTC)。
2. 重量顆粒比：Maiti Abhik et al.[3]通過計算每個篩網中顆粒的數量，並將總重量除以顆粒數量，獲得 W/P(重量/顆粒)比值。
3. 體積取代法：Mert Aydin et al.[4]透過影像辨識獲得粒料的面積以及其長寬，將粒料寬度直接取代平面上無法取得的厚度資訊，計算各層各顆粒料的體積，並使用體積計算累積過篩比，繪製成級配曲線。

重量顆粒比假設為同層顆粒的重量相等，但實際上同層之停留粒料單顆之間變化性可能很高；體積取代法的假設粒料的形狀是均勻的，且可以用長度和寬度來代表其立體形狀。實際上粒料的形狀可能會有很大變化，並且不僅僅由長度和寬

度來描述。這種假設可能無法準確地反映出粒料的實際體積。質量轉換係數方法考慮了重量和面積之間的關係。通過計算粒料在影像中的面積並與其對應的重量進行比較，可以獲得質量轉換係數。這種方法更全面地考慮了重量和形狀之間的關係，並且可以提供更準確的粒料重量估算。

本研究選擇使用質量轉換係數作為粒料分析的方法。透過考慮重量和面積之間的關係，可以更準確地評估粒料的重量分佈。

### 2.4.3 影像粒料形狀

在混凝土結構中，粒料的形狀評估至關重要。現有的粒料形狀測試方法需要大量的人工操作，耗時且費力，並且容易出現由人工判斷而引起的誤差。因此，需要一種替代方法，能夠以客觀、經濟且及時的方式測試粒料的形狀。[7]–[9]使用一種基於影像辨識的粒料形狀測量替代方案。該系統利用雙重、同步、雙速逐行掃描的攝影器材從兩個方向捕獲粒料顆粒的影像。使用專用的影像採集卡同時對這兩個影像進行數字化，並進行實時閾值處理，生成二值影像，然後傳輸至主機計算機進行進一步分析，如圖 2.4-3 所示。

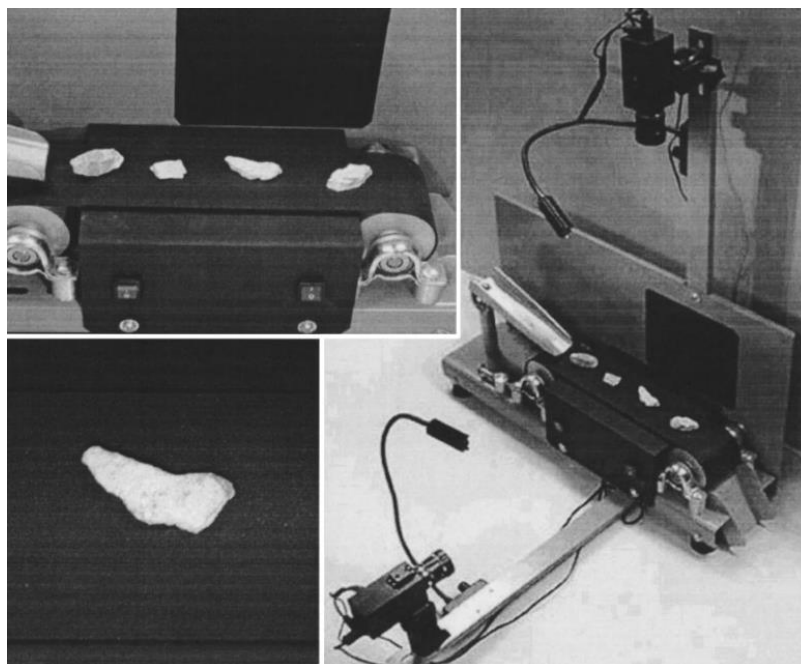



圖 2.4-3 WipShape 形狀量測系統[7]



這種基於視覺的方法相對於傳統方法具有多個優勢。它提供了客觀的粒料形狀測量結果，消除了人工測量中的主觀性和變異性，並能更全面地評估形狀參數。此外，使用雙攝影器材可以更準確地從多個視角表示粒料顆粒的形狀。然而，由於涉及專用設備和軟體，實施這種影像辨識的系統可能需要更高的資本投入，故本研究提出僅只有單個攝影器材評估粒料在影像當中的厚度，同時達到快速且客觀的成果，也不須對硬體設備以及軟體要求嚴苛。

## 2.5. 機器學習

機器學習是一門研究如何從數據中自動學習並改進性能的領域，被廣泛應用於人工智慧和數據科學領域。該領域的目標是開發出能夠從數據中學習的算法和模型，並利用這些模型來進行預測、分類、集群和優化等任務。


[10]提到機器學習(Machine Learning)被視為人工智慧(Artificial Intelligence)的一個子集，它展現了與人類智能相關的經驗性「學習」能力，同時通過使用演算法能夠學習和改進其分析能力。這些演算法使用大量的數據輸入和輸出來識別模式，從而有效地「學習」以訓練機器進行自主建議或決策。在足夠的重複和修改算法的過程中，機器能夠接受一個輸入並預測一個輸出。然後，將輸出與一組已知結果進行比較，以評估算法的準確性，並進一步迭代調整以完善預測未來結果的能力。

機器學習在人工智慧領域中起著重要作用，因為它使機器能夠根據過去的經驗和數據做出智能決策和預測。它被廣泛應用於各個領域，如圖像識別、語音識別、自然語言處理、推薦系統、醫學診斷、金融預測等。機器學習的不斷發展和創新推動了人工智慧領域的進步，使機器能夠更好地理解 and 處理複雜的現實世界問題。

### 2.5.1 前言

機器學習領域有許多專有名詞以及術語，描述如下：

1. 數據(Data)：數據是指收集到的樣本，可以是結構化或非結構化的信息。數據通常包含特徵和標籤，用於訓練和驗證機器學習模型。

- 
2. 特徵(Feature)：特徵是指用來描述和表示數據的屬性或觀測值，可以是數值型、類別型或其他形式的數據，作為模型的輸入。
  3. 標籤(Label)：標籤是指與每個數據樣本相關聯的目標變量或輸出變量。標籤用於訓練模型，使模型能夠預測新的數據樣本的標籤或類別。
  4. 模型(Model)：模型是根據給定的數據和特徵透過演算法學習到的規則或模式。可以是數學函數、統計模型、機器學習演算法等形式，用於對新的數據進行預測或分類。
  5. 擬合(Fitting)：擬合是指將模型與訓練數據訓練的過程。當模型能夠很好地捕捉到訓練數據中的模式和關係時，模型在訓練數據上達到了良好的擬合。
  6. 偏差(Bias)：指模型的預測值與真實值之間的平均差異。表示模型對問題的錯誤假設或簡化程度。高偏差代表模型對問題的解釋能力較弱，可能無法捕捉到真實模式，甚至在訓練數據上也表現不佳。
  7. 方差(Variance)：指模型對於不同訓練數據集的預測結果之間的變異程度。表示模型對訓練數據的敏感性和過擬合程度。高方差意味著模型對訓練數據過於敏感，導致對於新的數據表現不穩定，無法很好地泛化。
  8. 欠擬合(Underfitting)：欠擬合發生在模型無法捕捉到訓練數據中的模式和關係時。模型在訓練數據上的性能表現不佳，無法適應數據的變異性和複雜性，通常是由於模型過於簡單或不足以擬合訓練數據的結構，也就是高偏差低方差的結果，如圖 2.5-1[11]所示。
  9. 過擬合(Overfitting)：過擬合發生在模型過於複雜或過度擬合訓練數據時。模型在訓練數據上表現非常好，但對於新的、沒有見過的數據則預測能力較差。通常是由於模型過於複雜，過多地學習了訓練數據中的噪聲或細節，也就是低偏差高方差的結果。

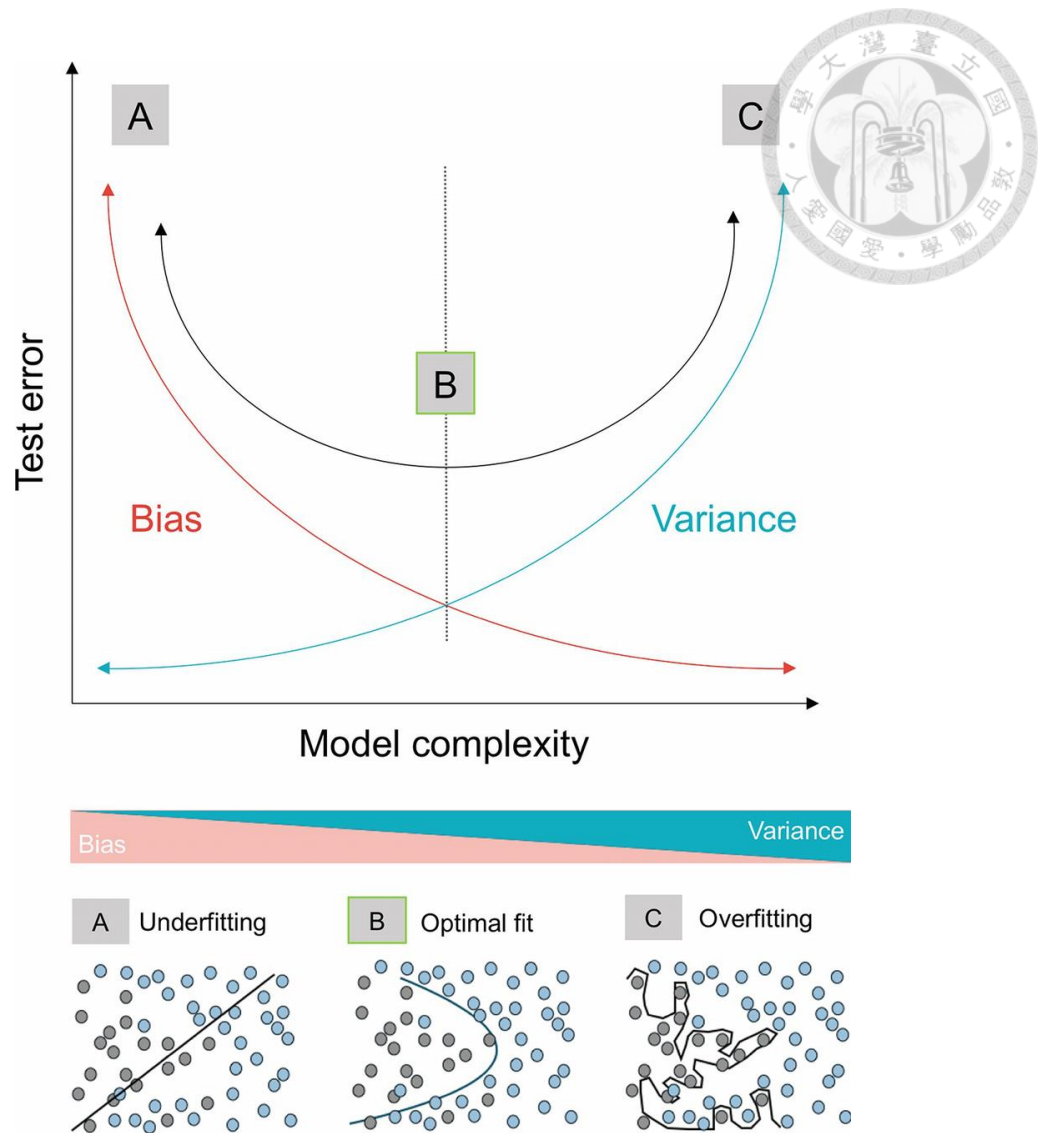


圖 2.5-1 偏差和方差權衡圖 [11]

圖 2.5-1 顯示了三種擬合狀態的結果，其中具有高偏差和低方差的預測模型(A)，以更簡單的模型(例如超平面)始終近似於潛在的數據生成過程，形成欠擬合的結果；(B)U 形的決策邊界代表了在這種情況下的最佳解答，偏差和方差都很低，從而產生了最低的測試誤差；(C)過度靈活的模型會導致過擬合。模型學習到了與潛在信號無關的數據特異性和隨機非預測結構。

## 2.5.2 機器學習類型

機器學習可以分為以下幾種主要類型：



1. 監督式學習(Supervised learning)：使用帶有標籤的數據進行訓練，通過演算法將輸入數據和對應的輸出標籤(Label)相匹配來學習。機器可以根據這個模型對新的輸入進行準確的預測。
2. 非監督式學習(Unsupervised learning)：使用未標籤的數據進行訓練，機器通過發現數據中的結構和模式來進行分類和規律性發現。這種方法不需要人工提供標籤，讓機器自行學習和探索。無監督學習通常用於聚類、降維和規律性發現等。
3. 半監督式學習(Semi-supervised learning)：結合了監督和非監督學習的特點，使用帶有標籤和未標籤的數據進行訓練。機器首先利用帶有標籤的數據建立模型，然後利用未標籤的數據進一步優化模型性能。常用於當有標籤數據相對較少時，利用未標籤數據來提升模型的性能和泛化能力。
4. 強化式學習(Reinforcement learning)：機器透過與環境的交互學習，根據正負反饋來調整自身的行為。機器根據環境給予的獎勵或懲罰，逐步優化策略，以達到最大化預期利益的目標。強化式學習常應用於需要做出一系列決策的問題，如遊戲策略、機器人控制和自動化交易等。

本研究建立之機器學習模型皆為監督式分類模型。

### 2.5.3 機器學習流程

機器學習的基本流程分為以下幾個步驟：

1. 數據收集與準備：收集與問題相關的數據，確保數據的品質和完整性。其中涉及數據探索、數據清理、數據轉換等過程。
2. 特徵選擇和特徵轉換：根據問題的要求和數據的特性，選擇最具代表性的特徵並進行相應的轉換，以提高模型的表現和泛化能力。



3. 模型選擇和建立：根據問題的類型和目標，選擇適合的演算法，例如 DecisionTree、RandomForest、Support Vector Machine 等，建立模型並設置相應的參數。
4. 模型訓練：使用訓練集對選定的模型進行訓練，使其能夠學習數據中的模式和關係。訓練的過程通常涉及對模型進行迭代的參數調整，以最小化預測誤差。
5. 模型評估與優化：使用驗證集評估訓練的模型性能，通過調整模型的超參數或使用交叉驗證等方法進行模型的優化，以提高模型的準確性和泛化能力。
6. 模型測試與部署：使用測試集對最終的模型進行性能評估，確定模型的實際效果。在模型通過測試後，可以將其部署到實際應用中進行預測或決策支持。
7. 監控與維護：持續監控模型在實際應用中的性能和效果，根據需要進行模型的更新和維護，以確保模型的穩定性和準確性。

#### 2.5.4 機器學習評估

機器學習效能評估是評估機器學習模型在處理特定任務或問題時的表現和性能的過程，分別有準確率、精確率、召回率、F1 分數、混淆矩陣和學習曲線等，分述如下：

1. 準確度(Accuracy)：衡量模型在分類問題中正確預測的比例。定義為預測結果中正確預測的樣本數除以總樣本數，如 2.7-1 所示。

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{FN} + \text{TN}} \quad \text{式 2.7-1}$$

其中真陽性、真陰性、假陽性和假陰性分述如下：

- (1) 真陽性(True Positive, TP)：模型正確地將正類樣本(Positive)分類為正類，實際觀察值為正類，且模型也正確預測為正類。



(2) 真陰性(True Negative, TN)：模型將負類樣本(Negative)分類為負類，實際觀察值為負類，且模型也正確預測為負類。

(3) 假陽性(False Positive, FP)：模型將負類樣本錯誤地分類為正類，實際觀察值為負類，但模型錯誤地預測為正類。

(4) 假陰性(False Negative, FN)：模型將正類樣本錯誤地分類為負類，實際觀察值為正類，但模型錯誤地預測為負類。

2. 精確率(Precision)：評估正類樣本中被正確預測的比例。定義為真陽性數量除以真陽性和假陽性的總和，如式 2.7-2 所示。

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad \text{式 2.5-1}$$

3. 召回率(Recall)：評估真實正類樣本中被正確預測的比例。定義為真陽性數量除以真陽性和假陰性的總和，如 2.7-3 所示。

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad \text{式 2.7-3}$$

4. F1 分數(F1 score)：結合精確度和召回率的指標，可以綜合評估模型的性能，為精確度和召回率的調和平均值，如式 2.7-4 所示。

$$\text{F1 score} = \frac{2 \times (\text{precision} \times \text{recall})}{(\text{precision} + \text{recall})} \quad \text{式 2.5-4}$$

5. 混淆矩陣(Confusion Matrix)：用於衡量模型在分類問題中的預測結果，將真實標籤和預測標籤組合成四個類別：真陽性、真陰性、假陽性和假陰性，如表 2.7-1 所示。

6. 學習曲線(Learning Curve)：用於評估機器學習模型性能和訓練進展的工具，描述了模型在不同訓練樣本數量下的訓練和驗證表現之間的關係。

表 2.5-1 混淆矩陣

	預測陽性	預測陰性
實際陽性	真陽性(TP)	假陰性(FN)

實際陰性	假陽性(FP)	真陰性(TN)
------	---------	---------



## 2.6. 特徵工程

特徵工程為機器學習中一項重要環節，是指對原始數據進行轉換、選擇和創建特徵的過程，旨在提取出對標籤具有預測能力且最有價值的特徵資料。特徵工程的重要性可以分成四個面向，描述如下：

1. 改善模型性能：通過適當的特徵工程，可以提取出與標籤密切相關的特徵，從而提高模型的預測能力和準確性。
2. 降低過擬合風險：過多的特徵可能導致模型過於複雜和維度過高，容易出現過擬合現象。特徵工程可以幫助選擇和創建最具有預測能力的特徵，從而減少過擬合風險。
3. 增強模型解釋性：通過對特徵的轉換，可以使特徵具有更好的解釋性，使模型的預測結果更容易被理解和解釋。
4. 節省計算資源：通過特徵選擇和降維等方法，可以減少特徵空間的維度，從而節省計算資源和模型訓練時間。

### 2.6.1 特徵選擇

根據特徵選擇的形式，可以將特徵選擇方法分為三種類型：

1. 過濾法(Filter)：過濾法根據特徵的發散性或相關性對每個特徵進行評分，並根據設定的閾值或選擇的特徵數量來選擇特徵。
2. 包裝法(Wrapper)：包裝法根據預測效果評分作為目標函數，每次選擇一些特徵或排除一些特徵，並通過反覆迭代選擇最佳特徵組合。
3. 集成法(Embedded)：集成法先使用機器學習算法和模型進行訓練，獲得各個特徵的權重係數，然後根據相關係數從大到小選擇特徵。與過濾法相似，但是通過訓練過程來確定特徵的重要性。

綜上所述，特徵工程中的特徵選擇對於構建高效的機器學習模型至關重要。透過選擇具有意義且與標籤相關性高的特徵，能夠提高模型的預測能力和解釋性。


本研究使用 sklearn 的 SelectKBest 函式，由 Trevor Hastie[12]提出，其為過濾法的特徵選擇方式，用於從給定的特徵集中選擇最佳的 K 個特徵。它基於統計學原理的單變量統計檢驗，通過對每個特徵進行評分，然後根據評分選擇 K 個最佳特徵，選擇得分最高的 K 個特徵作為最終選擇的特徵集。Score 使用預設的 ANOVA F-value 為方差分析(Analysis of Variance, ANOVA)中的一個統計量，用於評估不同組之間的變異程度是否有顯著差異。ANOVA F-value 衡量的是組間變異和組內變異之間的比值，通過計算 F-value 來判斷組間的變異是否顯著大於組內的變異。ANOVA F-value 的步驟如下：

1. 計算組間變異(Between-group variation)：將每個組的平均值與整體平均值的差異進行平方，然後將平方差加總起來，得到組間變異。
2. 計算組內變異(Within-group variation)：將每個組內的觀測值與其所在組的平均值的差異進行平方，然後將平方差加總起來，得到組內變異。
3. 計算 ANOVA F-value：將組間變異除以組內變異，得到 F-value。F-value 越大表示組間變異相對於組內變異越大，即不同組之間的平均值差異越顯著。

當 F-value 越大時，表示該特徵的變異解釋標籤的能力越強，與其他特徵相比，它更能區分不同的目標類別的變異。因此，在使用 SelectKBest 選擇特徵時，較大的 F-value 通常被視為更重要的特徵，可能被選擇進入最終的特徵集合。

Manikandan Ayyanar et al.[13]建立五個機器學習模型預測可能罹患心臟病之高風險人群，並使用 SelectKBest 的特徵選擇方式降低資料的特徵數，其中 LGBM 模型在特徵選擇後其 accuracy, recall 和 F1-score 皆有所提升。

Alexander Powell et al.[14]使用多種特徵選擇方式以及多種機器學習演算法，交互比較在不同數據集下其準確性以及處理時間，發現在減少總處理時間的同時，



準確率也保持在一定水準之上，其中 Logistic Regression 和 SelectKBest 的組合，將特徵數量從 80 個減少至 10 個，雖然準確率從 91% 下降至 88%，但處理時間大幅減少 65.04 秒；Decision Tree 和 SelectKBest 的組合，計算時間從 0.78 秒降至 0.11 秒，成為組合當中最短的計算時間，則準確率從 99% 降至 97%。

Md. Sabab Zulfiker et al.[15]訓練六種不同的機器學習分類器，利用各種社會人口學和心理社會信息來檢測受試者是否患有抑鬱症。此外，還使用了三種不同的特徵選擇方法，包括 SelectKBest、最小冗餘和最大相關(Minimum Redundancy and Maximum Relevance)以及 Boruta 特徵選擇算法。結果顯示 AdaBoost 和 SelectKBest 表現優於其他組合，準確率為 92.56%。

## 2.7. 演算法

機器學習演算法是一類能夠自動學習和提取數據中隱含模式的程序。演算法能夠通過觀察大量的數據樣本，發現特徵和關聯性，並將這些學習應用於新的未見數據上。機器學習演算法的目標是通過優化模型參數和函數來最大程度地提高預測和分類的準確性。

### 2.7.1 LogisticRegression

LogisticRegression(邏輯回歸)是由 David Cox[16]提出，為一種統計模型，用於預測二元結果變量的機率。它基於對線性回歸模型的改進，通過將線性組合轉化為概率值，並利用邏輯函數(如 sigmoid 函數)將結果映射到[0, 1]區間。

邏輯回歸具有可解釋性強、計算效率高、對特徵工程友好等特點。它可以處理大量特徵和大規模數據集，並且對於二元分類問題表現良好。

### 2.7.2 DecisionTree

Decision Tree (決策樹)是由多位學者在不同時期提出的，最早由 J. Ross Quinlan[17]於 1986 年提出，並且在之後的研究中得到了廣泛的應用和改進。

決策樹方法是一種基於樹狀結構的監督學習模型，用於建立基於多個協變量的分類系統或用於開發標籤的預測算法。該方法將種群分類為類似分支的片段，這些片段構建了一個具有根節點、內部節點和葉節點的倒置樹。該演算法為非參數型，可以有效地處理大型、複雜的數據集，而無需強加複雜的參數結構。

決策樹的建構過程為以下步驟：

- (1) 選擇一個特徵作為根節點。
- (2) 根據選定的特徵將數據集分成不同的子集。
- (3) 對每個子集，重複上述步驟，選擇最佳的特徵進行節點分裂，直到滿足停止條件。

決策樹具有易於理解和解釋的特點，能夠處理數值和類別特徵，並能夠處理缺失數據。它也可以處理大型數據集，並且對於探索性數據分析和特徵選擇非常有用。

### 2.7.3 RandomForest

隨機森林是由 L.Breiman[18]在 2001 年提出的一種集成學習方法，其特色是由多個隨機決策樹組成，通過投票或平均來進行預測。每棵樹都是基於不同的隨機樣本和特徵子集來建立的，具有抗過擬合能力。

隨機森林的基本流程如下：

1. 定義大小為  $n$  的隨機樣本：從資料集中隨機選取  $n$  個資料，並放回原始資料集。每次選取的樣本可能有重複。
2. 訓練決策樹：從上述選取的  $n$  個資料中，建立一個決策樹。在建立決策樹的過程中，對於每個節點，進行以下步驟：
  - (1) 隨機選取  $d$  個特徵：從所有特徵中隨機選取  $d$  個特徵。
  - (2) 使用特徵分割該節點：使用選取的  $d$  個特徵來尋找最佳的分割方式，通常使用信息增益或基尼係數等指標來評估分割的好壞。
3. 重複  $k$  次步驟 2：重複執行步驟 2，建立  $k$  棵決策樹，每棵決策樹都是使用不同的隨機樣本和特徵進行訓練。

4. 集成預測結果：將所有決策樹的預測結果匯總起來，以多數決的方式來決定最終的分類結果。



## 2.7.4 Xgboost

XGBoost 是由陳天奇[19]於 2016 年提出的，是一種梯度提升樹模型，它通過集成多棵樹來進行預測，並使用梯度下降法進行模型優化。它具有高效性和準確性，並且能夠處理多類別分類和回歸問題。

以下是 XGBoost 的主要特點和優勢：

1. 梯度提升算法：XGBoost 是一種梯度提升算法，它通過連續地迭代增加模型的預測能力。它在每一輪迭代中，根據之前迭代的結果計算出殘差，然後擬合新的弱學習器來預測殘差。通過不斷迭代，XGBoost 能夠有效地提升模型的準確性。
2. 正則化：XGBoost 提供了多種正則化技術，如 L1 正則化和 L2 正則化，以控制模型的複雜度和避免過擬合。這些正則化技術有助於提高模型的泛化能力，使其能夠更好地應對新的未知數據。
3. 特徵處理：XGBoost 具有良好的特徵處理能力。它能夠自動處理缺失值、處理類別型特徵、進行特徵轉換等。這使得用戶能夠更輕鬆地應對真實世界中的數據，並節省了特徵工程的時間和精力。
4. 高效性能：XGBoost 通過多種優化技術實現了高效的計算性能。它使用並行計算和近似算法來加快模型的訓練速度，並且在大型數據集上也能夠有效運行。
5. 可解釋性：XGBoost 提供了對模型的可解釋性。它能夠計算特徵的重要性，顯示每個特徵對於模型的貢獻程度，這有助於用戶理解模型的預測過程和結果。



## 2.7.5 Stacking

Stacking(堆疊法)是由 David H. Wolpert[20]於 1992 年提出，是一種集成學習方法，它通過結合多個基模型的預測結果來進行最終預測，再使用另一個模型（元模型）來結合基模型，以提高預測性能。

Stacking 分為兩個階段：

1. 訓練階段：訓練數據分成若干份，將其中一部分用於訓練不同的基本模型。這些基本模型可以是不同類型的機器學習算法，如決策樹、支持向量機、隨機森林等。每個基本模型都使用部分訓練數據進行訓練，並生成預測結果。
2. 預測階段：將測試數據提供給每個基本模型進行預測，預測結果將作為輸入，再與原始特徵一起提供給元模型。元模型的目標是根據基本模型的預測結果和原始特徵生成最終的預測結果。通常，元模型使用的是簡單的演算法如線性回歸、邏輯回歸等。

堆疊模型能夠利用不同模型的優點，並通過適當的組合和權重調整來改善預測結果。它可以有效地捕捉到數據的非線性關係和複雜模式，並具有較好的泛化能力。

### 第三章、粗粒料試驗與攝影



本研究實驗分成兩個部分，第一部分為粗粒料試驗，分別有粗粒料篩析法以及粗粒料中扁平、細長或扁長顆粒含量試驗法。第二部分為粗粒料影像拍攝，結合兩個部分的結果建立機器學習模型的訓練資料以及測試資料。實驗組別總共有 12 組，6 組為花蓮地區之粗粒料；6 組為閩侯地區之粗粒料。

#### 3.1. 實驗材料

##### 1. 粗粒料

採用財團法人台灣營建研究院提供之花蓮石以及閩侯石。

表 3.1-1 試驗結果一覽表 花蓮料源

試驗項目		試驗值				試驗方法
		A料 六分石	B料 三分石	A:B 30%:70%	*規範值	
篩 分 析 (通過 百分 比)	1 1/2"	100	100	100	100	CNS 486(2001)
	1"	100	100	100	95-100	
	3/4"	79	100	92	—	
	1/2"	18	75	52	25-60	
	3/8"	10	49	33	—	
	#4	4	9	7	0-10	
	#8	3	5	4	0-5	
通過#200 號篩材 料, %		1.0	1.0	1.0	≤1.0	CNS 491(2001)
S.S.D. 比重		2.70	2.66	2.68	—	CNS 488(2008)
吸水率, %		0.8	1.1	1.0	—	
輕質粒料含量, %		0.0	0.0	—	≤0.5	CNS 10990(1998)
粘土塊及易碎顆 粒, %		0.1	0.2	—	≤3.0	CNS 1171(1995)
健康試驗損失 量, % (硫酸鈉溶液五循環)		1.4	0.9	—	≤12	CNS 1167(1995)
洛杉磯 磨損值	100 轉	—	—	6.5	—	CNS 490(2009)
	500 轉	—	—	29.3	≤50	



(B級), %						
磨損均勻硬度比	100/500轉	—	—	0.22	—	
粒料單位重 kg/m	搗實方	—	1644	1676	—	CNS 1163(2008)
	鬆方	—	1482	1500	—	
扁平粒料(寬:厚=5:1),%		0.03	0.3	—	≤10	CNS 15171(2008)
細長粒料(長:寬=5:1),%		0.0	0.0	—		
扁長粒料(長:厚=5:1),%		1.4	4.9	—		

表 3.1-2 粗粒料比重(密度)及吸水性(率)試驗報告 花蓮料源(三分石)

試驗項目		試驗值	*規範值	備註
烘乾狀態	密度, 23°C, kg/m <sup>3</sup>	2630	—	樣品狀態: <input type="checkbox"/> 經烘乾浸泡程序 <input checked="" type="checkbox"/> 未經烘乾程序、直接加
	相對密度, 23°C/23°C	2.64		
面乾內飽和狀態	密度, 23°C, kg/m <sup>3</sup>	2660	—	
	相對密度, 23°C/23°C	2.66		
視密度, 23°C, kg/m <sup>3</sup>		2710	—	
相對視密度 23°C/23°C		2.72	—	
吸水率, (%)		1.1	—	

表 3.1-3 粗粒料比重(密度)及吸水性(率)試驗報告 花蓮料源(六分石)

試驗項目		試驗值	*規範值	備註
烘乾狀態	密度, 23°C, kg/m <sup>3</sup>	2670	—	樣品狀態: <input type="checkbox"/> 經烘乾浸泡程序 <input checked="" type="checkbox"/> 未經烘乾程序、直接加
	相對密度, 23°C/23°C	2.67		
面乾內飽和狀態	密度, 23°C, kg/m <sup>3</sup>	2690	—	
	相對密度, 23°C/23°C	2.70		
視密度, 23°C, kg/m <sup>3</sup>		2720	—	
相對視密度 23°C/23°C		2.73	—	
吸水率, (%)		0.8	—	

表 3.1-4 粗粒料中扁平、細長、扁長粒料試驗報告-1 花蓮料源(三分石)

篩號 尺寸	個別停留		扁平粒料			細長粒料			扁長粒料		
	百分率 (%)	顆粒數	顆粒數	顆粒百分率 (%)	重量百分率 (%)	顆粒數	顆粒百分率 (%)	重量百分率 (%)	顆粒數	顆粒百分率 (%)	重量百分率 (%)
3/4"	0	0	0	0	0	0	0	0	0	0	0
1/2"	27	136	0	0	0	0	0	0	2	1.47	1.03
3/8"	29	157	2	1.27	0.43	0	0	0	9	5.73	3.45
#4	44	125	1	0.80	0.48	0	0	0	10	8.00	8.15

表 3.1-5 粗粒料中扁平、細長、扁長粒料試驗報告-2 花蓮料源(三分石)

試驗結果	扁平粒料		細長粒料		扁長粒料	
	累積權值	*規範值	累積權值	*規範值	累積權值	*規範值
質量百分率	0.3	≤10	0.0	≤10	4.9	≤10
顆粒百分率	—	—	—	—	—	—

表 3.1-6 粗粒料中扁平、細長、扁長粒料試驗報告-1 花蓮料源(六分石)

篩號 尺寸	個別停留		扁平粒料			細長粒料			扁長粒料		
	百分率 (%)	顆粒數	顆粒數	顆粒百分率 (%)	重量百分率 (%)	顆粒數	顆粒百分率 (%)	重量百分率 (%)	顆粒數	顆粒百分率 (%)	重量百分率 (%)
1"	0	0	0	0	0	0	0	0	0	0	0
3/4"	22	149	0	0	0	0	0	0	0	0	0
1/2"	63	160	0	0	0	0	0	0	3	1.88	1.37
3/8"	9	124	0	0	0	0	0	0	4	3.23	2.22
#4	6	111	1	0.90	0.54	0	0	0	4	3.60	5.11

表 3.1-7 粗粒料中扁平、細長、扁長粒料試驗報告-2 花蓮料源(六分石)

試驗結果	扁平粒料		細長粒料		扁長粒料	
	累積權值	*規範值	累積權值	*規範值	累積權值	*規範值
質量百分率	0.03	≤10	0.0	≤10	1.4	≤10
顆粒百分率	—	—	—	—	—	—

表 3.1-8 粗粒料篩分析試驗報告 花蓮料源(三分石)

篩網 稱號	個別留篩 百分率(%)	累 計		*規範值
		留篩重量百分率(%)	過篩重量百分率(%)	
25.0mm(1")	0.0	0	100	—
19.0mm(3/4")	0.1	0	100	—
12.5mm(1/2")	25.2	25	75	—
9.5mm(3/8")	25.7	51	49	—
4.75mm(#4)	39.8	91	9	—
2.36mm(#8)	3.8	95	5	—
F . M . =		—		—
小於試驗篩 75μm(No.200)材料含量百分率(%)			1.0	—

表 3.1-9 粗粒料篩分析試驗報告 花蓮料源(六分石)

篩網 稱號	個別留篩 百分率(%)	累 計		*規範值
		留篩重量百分率(%)	過篩重量百分率(%)	
25.0mm(1")	0.0	0	100	—
19.0mm(3/4")	21.2	21	79	—
12.5mm(1/2")	61.1	82	18	—
9.5mm(3/8")	7.9	90	10	—
4.75mm(#4)	6.0	96	4	—
2.36mm(#8)	0.7	97	3	—
F . M . =		—		—
小於試驗篩 75 $\mu$ m(No.200)材料含量百分率(%)			1.0	—

表 3.1-10 試驗結果一覽表 閩侯料源

試 驗 項 目		試 驗 值				試 驗 方 法
		A 料 六分石	B 料 三分石	A : B 30%:70%	*規範值	
篩 分 析 (通 過 百 分 比)	1 1/2"	100	100	100	100	CNS 486(2001)
	1"	100	100	100	95-100	
	3/4"	81	100	94	—	
	1/2"	28	73	60	25-60	
	3/8"	11	57	43	—	
	#4	3	9	7	0-10	
	#8	2	2	2	0-5	
通過#200 號篩材 料, %		1.0	1.0.80	0.9	$\leq 1.0$	CNS 491(2001)
S . S . D . 比重		2.65	2.64	2.64	—	CNS 488(2008)
吸 水 率, %		0.8	1.1	1.0	—	
輕質粒料含量, %		0.0	0.0	—	$\leq 0.5$	CNS 10990(1998)
粘土塊及易碎顆 粒, %		0.1	0.1	—	$\leq 3.0$	CNS 1171(1995)
健康試驗損失 量, %		0.9	1.2	—	$\leq 12$	CNS 1167(1995)

(硫酸鈉溶液五循環)						
洛杉磯 磨損值 (B 級) , %	100 轉	—	—	2.4	—	CNS 490(2009)
	500 轉	—	—	13.2	≤50	
磨損均 勻硬度 比	100/500 轉	—	—	0.18	—	
粒料單 位重 kg/m	搗實方	—	1588	1634	—	CNS 1163(2008)
	鬆方	—	1445	1468	—	
扁平粒料(寬:厚 =5:1),%		0.0	0.0	—	≤10	CNS 15171(2008)
細長粒料(長:寬 =5:1),%		0.0	0.0	—		
扁長粒料(長:厚 =5:1),%		0.0	1.6	—		

表 3.1-11 粗粒料比重(密度)及吸水性(率)試驗報告 閩侯料源(三分石)

試 驗 項 目		試驗值	*規範值	備 註
烘乾狀態	密度, 23°C, kg/m <sup>3</sup>	2610	—	樣品狀態: <input type="checkbox"/> 經烘乾浸泡 程序 <input checked="" type="checkbox"/> 未經烘乾程 序、直接加
	相對密度, 23°C/23°C	2.61		
面乾內 飽和狀態	密度, 23°C, kg/m <sup>3</sup>	2630	—	
	相對密度, 23°C/23°C	2.64		
視密度, 23°C, kg/m <sup>3</sup>		2680	—	
相對視密度 23°C/23°C		2.69	—	
吸水率, (%)		1.1	—	

表 3.1-12 粗粒料比重(密度)及吸水性(率)試驗報告 閩侯料源(六分石)

試驗項目		試驗值	*規範值	備註
烘乾狀態	密度, 23°C, kg/m <sup>3</sup>	2630	—	樣品狀態: <input type="checkbox"/> 經烘乾浸泡程序 <input checked="" type="checkbox"/> 未經烘乾程序、直接加
	相對密度, 23°C/23°C	2.63		
面乾內飽和狀態	密度, 23°C, kg/m <sup>3</sup>	2650	—	
	相對密度, 23°C/23°C	2.65		
視密度, 23°C, kg/m <sup>3</sup>		2680	—	
相對視密度 23°C/23°C		2.69	—	
吸水率, (%)		0.7	—	

表 3.1-13 粗粒料中扁平、細長、扁長粒料試驗報告-1 閩侯料源(三分石)

篩號尺寸	個別停留		扁平粒料			細長粒料			扁長粒料		
	百分率 (%)	顆粒數	顆粒數	顆粒百分率 (%)	重量百分率 (%)	顆粒數	顆粒百分率 (%)	重量百分率 (%)	顆粒數	顆粒百分率 (%)	重量百分率 (%)
3/4"	0	0	0	0	0	0	0	0	0	0	0
1/2"	30	132	0	0	0	0	0	0	0	0	0
3/8"	17	143	0	0	0	0	0	0	1	0.70	0.51
#4	53	120	0	0	0	0	0	0	3	2.50	2.93

表 3.1-14 粗粒料中扁平、細長、扁長粒料試驗報告-2 閩侯料源(三分石)

試驗結果	扁平粒料		細長粒料		扁長粒料	
	累積權值	*規範值	累積權值	*規範值	累積權值	*規範值
質量百分率	0.0	≤10	0.0	≤10	1.6	≤10
顆粒百分率	—	—	—	—	—	—

表 3.1-15 粗粒料中扁平、細長、扁長粒料試驗報告-1 閩侯料源(六分石)

篩號 尺寸	個別停留		扁平粒料			細長粒料			扁長粒料		
	百分率 (%)	顆粒數	顆粒數	顆粒百分率 (%)	重量百分率 (%)	顆粒數	顆粒百分率 (%)	重量百分率 (%)	顆粒數	顆粒百分率 (%)	重量百分率 (%)
1"	0	0	0	0	0	0	0	0	0	0	0
3/4"	20	90	0	0	0	0	0	0	0	0	0
1/2"	54	139	0	0	0	0	0	0	0	0	0
3/8"	18	152	0	0	0	0	0	0	0	0	0
#4	8	128	0	0	0	0	0	0	0	0	0

表 3.1-16 粗粒料中扁平、細長、扁長粒料試驗報告-2 閩侯料源(六分石)

試驗結果	扁平粒料		細長粒料		扁長粒料	
	累積權值	*規範值	累積權值	*規範值	累積權值	*規範值
質量百分率	0.0	≤10	0.0	≤10	0.0	≤10
顆粒百分率	—	—	—	—	—	—

表 3.1-17 粗粒料篩分析試驗報告 閩侯料源(三分石)

篩網 稱號	個別留篩 百分率(%)	累 計		*規範值
		留篩重量百分率(%)	過篩重量百分率(%)	
19.0mm(3/4")	0.0	0	100	—
12.5mm(1/2")	26.7	27	73	—
9.5mm(3/8")	16.5	43	57	—
4.75mm(#4)	47.6	91	9	—
2.36mm(#8)	7.0	98	2	—
F . M . =		—		—
小於試驗篩 75μm(No.200)材料含量百分率(%)			0.8	—

表 3.1-18 粗粒料篩分析試驗報告 閩侯料源(六分石)

篩網 稱號	個別留篩 百分率(%)	累 計		*規範值
		留篩重量百分率(%)	過篩重量百分率(%)	
25.0mm(1")	0.0	0	100	—
19.0mm(3/4")	19.2	19	81	—
12.5mm(1/2")	52.5	72	28	—
9.5mm(3/8")	16.9	89	11	—
4.75mm(#4)	8.5	97	3	—
2.36mm(#8)	0.7	98	2	—
F . M . =		—		—
小於試驗篩 75 $\mu$ m(No.200)材料含量百分率(%)			1.0	—



## 3.2. 粗粒料試驗

### 3.2.1 器材

本研究採用的儀器有試驗篩、機械式搖篩機、比例測徑器、游標尺、烘箱等設備儀器，描述如下：

1. 試驗篩：本研究使用 1”、3/4”、1/2”、3/8”、#4 的篩網，且規格符合 CNS 386[21]之規定。
2. 機械式搖篩機：使用時可使篩產生搖動，可使粒料震動、翻滾或轉動。
3. 比例測徑器：由高擎企業有限公司所製作，包含一個底板，兩個固定支柱和一個懸臂，懸臂安裝在兩個支柱之間，以保持臂與支柱之間的開口保持恆定比例。軸的位置可以調整，以提供所需的開口尺寸比例，可以設置 1:2、1:3、1:4 和 1:5 的比例。
4. 游標尺：由主尺和貼附在主尺上可部分滑動的游標兩部分構成。主尺以公分為單位。根據分格的不同，游標卡尺可分為十分度游標卡尺、二十分度游標卡尺、五十分度格游標卡尺等。
5. 烘箱：是用由建成科學儀器股份有限公司所製作的烘箱，最高溫可達約 900 °C，是由內部鐵圈通電加熱來提供熱能。

### 3.2.2 試驗流程

1. 本研究分別對花蓮、閩侯地區各取六組，一共十二組粗粒料試樣，且根據 CNS486[22]中規定最大粒徑為 19mm，每組試驗取最少量 5 公斤以上，如表 3.2-1 所示。

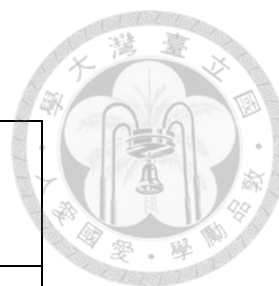
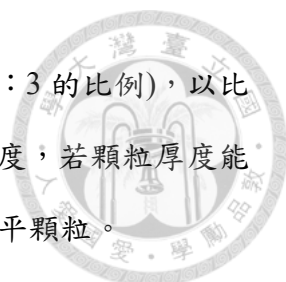


表 3.2-1 粗粒料試樣最少量

標稱最大粒徑 方孔篩孔寬，mm(吋)	試樣最少量，kg
9.5( $\frac{3}{8}$ )	1
12.5( $\frac{1}{2}$ )	2
19.0( $\frac{3}{4}$ )	5
25.0(1)	10
37.5( $1\frac{1}{2}$ )	15
50(2)	20
63( $2\frac{1}{2}$ )	35
75(3)	60
90( $3\frac{1}{2}$ )	100
100(4)	150
125(5)	300

2. 試樣置於  $110\pm 5^{\circ}\text{C}$  之烘箱烘乾 24 小時至恆重。
3. 篩網按照孔寬大小從高至低安裝至機械式搖篩機，每次搖篩前會先空搖(不包含試樣)五分鐘，檢查各層是否有無其他殘留物。
4. 每組試樣搖篩 20 分鐘，搖篩結束後各層取出秤重，並繪製篩分析級配曲線。
5. 根據 CNS 15174[1]各篩號之粗粒料停留顆粒數大於 100 顆者，隨機取樣 100 顆進行扁平、扁長、細長顆粒試驗。



6. 扁平顆粒試驗：設定固定軸至設定之比例(本研究使用 1：3 的比例)，以比例測徑器之較大開口端夾住顆粒寬度，使其等於顆粒寬度，若顆粒厚度能放置於比例測徑器之較小開口端內，則該顆粒被歸類為扁平顆粒。
7. 細長顆粒試驗：如步驟 6 設定固定軸為 1：3，比例測徑器之較大開口端夾住顆粒長度，使其等於顆粒長度，若顆粒寬度能放置於比例測徑器之較小開口端內，則該顆粒被歸類為細長顆粒。
8. 扁長顆粒試驗：如步驟 6 設定固定軸為 1：3，以比例測徑器之較大開口端夾住顆粒長度，使其等於顆粒長度，若顆粒厚度能放置於比例測徑器之較小開口端內，則該顆粒被歸類為扁長顆粒。該顆粒若已經歸類為扁平顆粒，就不計入扁長顆粒。
9. 本研究扁平、細長、扁長率計算以顆粒百分比[]計算，如式 3.2-1 所示。

$$Y_i(\%) = 100 \times \frac{U_i}{U_n} \quad \text{式 3.2-1}$$

其中：Y<sub>i</sub> 為停留個別篩號中扁平、細長或扁長顆粒之顆粒數量百分比(%)；  
 U<sub>i</sub> 為停留個別篩號中扁平、細長或扁長顆粒之顆粒數量(顆)；U<sub>n</sub> 為停留個別篩號中之全部顆粒數量(顆)。

### 3.2.3 粗粒料篩析法

表 3.2-2~表 3.2-7 為訓練資料之粗粒料篩析法之結果。

表 3.2-2 粗粒料篩析法之結果 花蓮 1

篩網 稱號	個別留篩 百分率(%)	累 計	
		留篩重量百分率(%)	過篩重量百分率(%)
25.0mm(1")	0.00	0	100
19.0mm(3/4")	3.78	3.78	96.22
12.5mm(1/2")	59.89	63.67	36.33
9.5mm(3/8")	21.66	85.33	14.67
4.75mm(#4)	12.95	98.28	1.72
底盤	1.72	100	0



表 3.2-3 粗粒料篩析法之結果 花蓮 2

篩網 稱號	個別留篩 百分率(%)	累 計	
		留篩重量百分率(%)	過篩重量百分率(%)
25.0mm(1")	0.00	0	100
19.0mm(3/4")	1.60	1.60	98.40
12.5mm(1/2")	46.36	47.96	52.04
9.5mm(3/8")	20.38	68.34	31.66
4.75mm(#4)	23.64	91.98	8.02
底盤	8.02	100	0

表 3.2-4 粗粒料篩析法之結果 花蓮 3

篩網 稱號	個別留篩 百分率(%)	累 計	
		留篩重量百分率(%)	過篩重量百分率(%)
25.0mm(1")	0.0	0	100
19.0mm(3/4")	1.61	1.61	98.39
12.5mm(1/2")	40.86	42.47	57.53
9.5mm(3/8")	21.03	63.50	36.50
4.75mm(#4)	26.75	90.25	9.75
底盤	9.75	100	0

表 3.2-5 粗粒料篩析法之結果 閩侯 1

篩網 稱號	個別留篩 百分率(%)	累 計	
		留篩重量百分率(%)	過篩重量百分率(%)
25.0mm(1")	0.0	0	100
19.0mm(3/4")	13.98	13.98	86.02
12.5mm(1/2")	49.20	63.18	36.82
9.5mm(3/8")	16.11	79.30	20.70
4.75mm(#4)	15.21	94.50	5.50
底盤	5.50	100	0



表 3.2-6 粗粒料篩析法之結果 閩侯 2

篩網 稱號	個別留篩 百分率(%)	累 計	
		留篩重量百分率(%)	過篩重量百分率(%)
25.0mm(1")	0.0	0	100
19.0mm(3/4")	8.35	8.35	91.65
12.5mm(1/2")	33.56	41.91	58.09
9.5mm(3/8")	17.02	58.92	41.08
4.75mm(#4)	31.26	90.19	9.81
底盤	9.81	100	0

表 3.2-7 粗粒料篩析法之結果 閩侯 3

篩網 稱號	個別留篩 百分率(%)	累 計	
		留篩重量百分率(%)	過篩重量百分率(%)
25.0mm(1")	0.0	0	100
19.0mm(3/4")	5.47	5.47	94.53
12.5mm(1/2")	41.79	47.26	52.74
9.5mm(3/8")	19.71	66.97	33.03
4.75mm(#4)	27.46	94.43	5.57
底盤	5.57	100	0

### 3.2.4 粗粒料中扁平、細長或扁長顆粒含量試驗法

表 3.2-8~表 3.2-10 為訓練資料之扁平、細長或扁長顆粒含量試驗法之結果。

表 3.2-8 扁平、細長或扁長顆粒含量試驗法之結果 左：花蓮、右：閩侯 組別 1

花蓮 1	1/2"	3/8"	#4	閩侯 1	1/2"	3/8"	#4
扁平率	3%	6%	5%	扁平率	0%	2%	5%
細長率	0%	0%	0%	細長率	0%	0%	0%
扁長率	14%	23%	23%	扁長率	3%	9%	23%

表 3.2-9 扁平、細長或扁長顆粒含量試驗法之結果 左：花蓮、右：閩侯 組別 2

花蓮 2	1/2"	3/8"	#4	閩侯 2	1/2"	3/8"	#4
扁平率	1%	6%	6%	扁平率	1%	4%	1%
細長率	0%	0%	0%	細長率	0%	0%	0%
扁長率	18%	26%	23%	扁長率	4%	11%	15%

表 3.2-10 扁平、細長或扁長顆粒含量試驗法之結果 左：花蓮、右：閩侯 組別 3

花蓮 3	1/2"	3/8"	#4	閩侯 3	1/2"	3/8"	#4
扁平率	2%	4%	5%	扁平率	0%	5%	3%
細長率	0%	0%	0%	細長率	0%	0%	0%
扁長率	16%	16%	23%	扁長率	3%	8%	%

### 3.3. 影像拍攝

本研究分別對兩個料源的粒料進行拍攝且利用影像辨識軟體 Image-Pro Plus，得到每一個粒料的特徵資料建立訓練資料以及測試資料。

#### 3.3.1 器材

1. 鐵盤:限制粒料拍攝範圍，且會在底板放置白色底板，增加對比度及可辨識度。
2. 拍攝儀器:iPhone 12 Pro Max(1200 萬像素相機系統)。
3. 燈條:可移動式燈具，方便調整光照方向以及控制陰影範圍(長燈條:40W、短燈條:20W)。



### 3.3.2 攝影流程

訓練資料根據步驟 1-5 取樣；測試資料根據步驟 1-3 取樣。

1. 將長燈條安置在固定高度，以及將拍攝儀器放置於固定的高度，長燈條高度需在拍攝儀器之下，如圖 3.3.1 所示。光源的擺設必須固定高度，且盡量光源照射方向與拍攝粒料平面的法向量平行，若放置側邊易產生過多陰影。

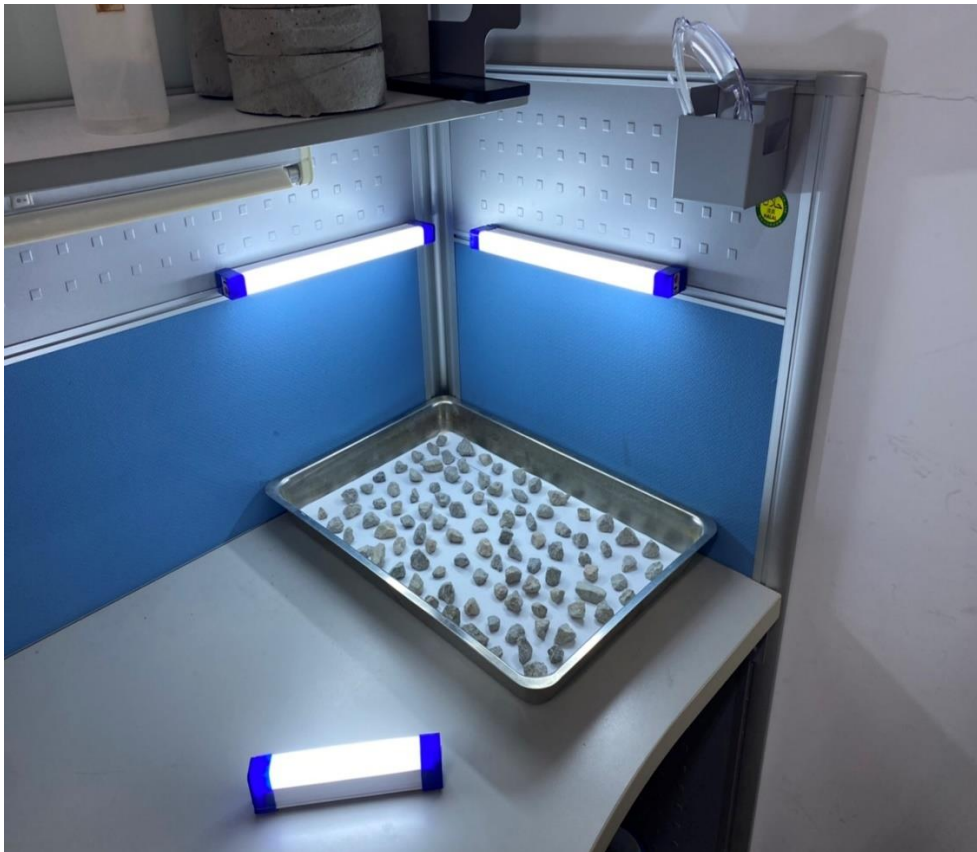


圖 3.3-1 拍攝環境

2. 為了最大程度的保留粒料特徵資料，粒料排列時須將每顆都分開，避免重疊以及遮擋比例尺的情況發生。
3. 手持短燈條調整高度使得陰影面積最小化，拍攝影像，如圖 3.3.2 所示。



圖 3.3-2 粒料實際拍攝影像

4. 因每顆粒料的陰影面積無法同時最小化，會保持粒料相同角度但微調短燈條高度拍攝，此步驟重複三次。
5. 翻動粒料且重新排列再根據上述步驟重複拍攝七次。此步驟目的是每一次粒料傾倒方向不盡相同，可以重複拍攝得到相同粒料但不同角度的特徵。



### 3.3.3 取樣

1. 訓練資料:分別對花蓮、閩侯地區各取 3 組試體，根據 3.3 粗粒料中扁平、細長或扁長顆粒含量試驗法，分別將 1/2”、3/8”、#4 三層中的停留粒料各層隨機取樣 100 顆並進行拍攝。每一張影像會有 100 顆粒料，些微調整光源角度則粒料保持相同角度拍攝 3 次，再翻動粒料改變粒料的角度進行拍攝，重複進行 7 次，相同的粒料總共會有 21 張影像。整個訓練資料集總共有 37800 筆粒料資料，如式 3.3.1 所示:

$$2 * 3 * 3 * 100 * 3 * 7 = 37800 \quad \text{式 3.3-1}$$

2. 測試資料: 分別對花蓮、閩侯地區各取 3 組試體，根據 3.2.2 粗粒料篩析法將 1/2”、3/8”、#4 的停留粒料取樣進行拍攝；另外，根據 3.3.2 粗粒料中扁平、細長或扁長顆粒含量試驗法扁平率試驗分別對三層中隨機 100 顆進行取樣並進行拍攝。

## 第四章、分析計畫與方法



### 4.1. 分析計畫背景

本研究的主要目的在於透過影像辨識的資料以機器學習的分類模型進行料源分類以及尺寸辨識，且使用質量轉換係數以及影像平均厚度推估粗粒料級配曲線與扁平、細長或扁長顆粒含量。

預拌廠的粗粒料儲存槽進行粒料補充時，通常會有上一批的殘留粒料，甚至料源會源自於不同地區，不同料源的粒料其特性皆不同，可透過機器學習分類模型可以對粒料進行料源分類，即時了解粒料各地區的佔比，修改拌合配比。

使用粒料影像平面資料繪製級配曲線的有兩個關鍵因素，第一個為粒料分布在不同尺寸篩號的數量，傳統的影像辨識在粒料分類上多半使用單因子當作分類標準，也有文獻提到使用多因子分類，本研究使用機器學習分類模型對粒料的尺寸進行分類，考慮的因子數量多達數十種以上，且利用集成式學習的演算法，能有效且更精確的分類粒料；第二個為粒料在各層篩號的重量，需要將影像平面資料轉換成重量，本研究參考江田政敏[6]所提出的質量轉換係數，其定義為真實重量與影像面積的比值，透過質量轉換係數推算各層重量從而得到級配曲線。

扁平率是評估粒料品質的一項指標，然而粒料的厚度是平面影像無法直接取得的資料，本研究延伸質量轉換係數的概念，將重量除以單位重得到粒料總體積，再除以影像辨識的粒料總面積，可以得到粒料影像平均厚度，以及電子游標尺量測全部粒料的最大厚度取其平均值，透過兩者之間的比值，再使用影像辨識的粒料長寬，推估影像粒料的扁平、細長、扁長率。

### 4.2. 影像前處理

本研究使用影像辨識軟體 Image-Pro Plus 分析影像，得到每一個粒料的平面特徵資料。



## 4.2.1 處理流程

1. 剪裁影像的邊緣處，使得影像只有粒料以及比例尺，如圖 4.2-1 所示。



圖 4.2-1 粒料影像圖

2. 將影像導入至 Image-Pro Plus。
3. 設定比例尺:Measure→Calibration→Spatial→選定比例尺名稱以及設定 Unit 為 centimeters→Image→使綠色標線與影像上比例尺重和→Scaling 設定與比例尺相同的數值(本研究比例尺為 10cm)→Apply。
4. 使用比例尺:Measure→Calibration→Select Spatial→選定設定的比例尺名稱。
5. 載入設定檔:Measure→Count/Size→File→Load Setting。Image-Pro Plus 提供 56 個特徵量測值可使用。
6. 辨識粒料:Measure→Count/Size→Automatic Dark Objects→Count，如圖 4.2-2 所示。

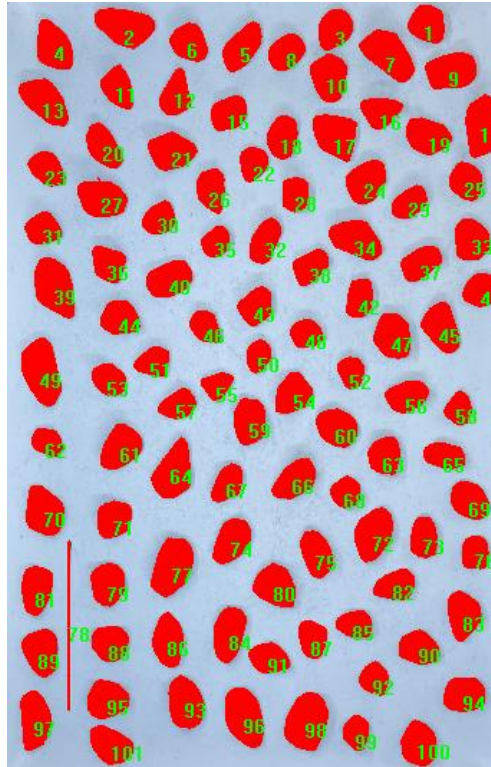


圖 4.2-2 辨識結果

7. 影像處理:同上一步的 Count/Size 中可使用 Edit 的 Toggle Objects On/Off 將比例尺、辨識不良或相連的粒料以及背景雜訊隱藏；亦可使用 Edit 的 Draw/Merge Objects 將未辨識完全的粒料手動補上，通常是貼在圖片邊緣處的粒料或本身色差過大的粒料會有辨識不良的情況發生。
8. 結果輸出:Measure→Count/Size→File→Export Data，可將資料輸出至 Excel。

#### 4.2.2 模型資料

本研究透過上述方式建立訓練資料以及測試資料。每筆粒料資料具有 45 個特徵。此外，有兩個標籤(Label)用於分類問題。第一個標籤是料源辨識，其中將花蓮地區定義為 0，閩侯地區定義為 1。第二個標籤是尺寸辨識，其中將#4 定義為 0，3/8"定義為 1，1/2"定義為 2。

為了訓練並評估模型，使用了包含 37800 筆的訓練資料集，將原始資料集按照 8:2 的比例分割為訓練集和驗證集。訓練集用於模型的訓練，驗證集用於模型的選



擇，其目的是訓練過程中評估模型在未見過的數據上的性能和泛化能力，可即時調整模型的超參數。

測試集資料用於評估最終模型的性能，其目的是衡量模型的泛化能力，即模型在新的未知數據上的預測準確性，其中測試資料的組別 4、組別 5 與訓練資料源自於同一袋粒料，目的是在面對相似資料的情況下測試模型的學習能力；組別 6 的粒料與其他五組源自於不同袋，目的是測試模型的泛化能力，能否辨識以及分類沒學習過的資料。

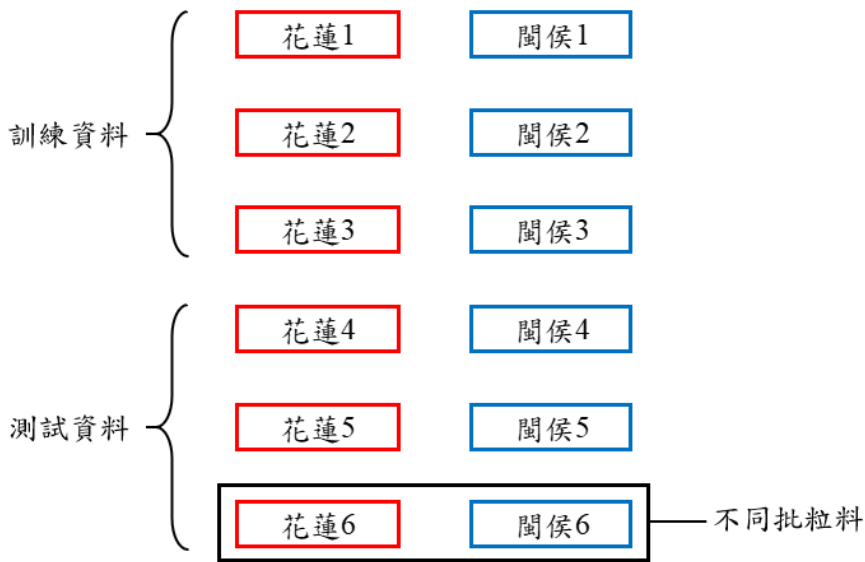


圖 4.2-3 粗粒料訓練組別以及測試組別

### 4.2.3 資料特徵

1. Area:物體面積。不包含物體上孔洞的面積。
2. Aspect:物體等效橢圓的主軸和次軸之間的比例。等效橢圓是一個具有相同面積和質量的橢圓形，可以通過將其擬合到物體的輪廓來生成。因此，Aspect 可以用來描述物體的形狀特徵。如果 Aspect 接近 1，則物體可以被視為圓形。
3. Area/Box:物體面積和其邊界框面積之比。
4. Box X/Y:物體邊界框寬度和高度之比。



5. Density (mean):物體平均光密度(或光強度)。
6. Hole Area:物體內部孔洞的面積。
7. Hole Ratio:物體除去孔洞面積與物體總面積之比。
8. Axis (major):具有與物體相同一次和二次矩的橢圓的主軸長度。
9. Axis (minor):具有與物體相同一次和二次矩的橢圓的次軸長度。
10. Diameter (max):物體輪廓上連接兩點的最長線段且通過重心的長度。
11. Diameter (min):物體輪廓上連接兩點的最短線段且通過重心的長度。
12. Diameter (mean):以角度 2 度間隔測量並通過物體重心的直徑的平均長度。
13. Radius (max):物體重心與輪廓之間的最大距離。
14. Radius (min):物體重心與輪廓之間的最小距離。
15. Holes:物體內的孔洞的數量。
16. Perimeter:物體的輪廓長度。
17. Radius Ratio:最大半徑與最小半徑的比值。
18. Roundness: $\text{Perimeter}^2 / (4 * \pi * \text{Area})$
19. Cluster:估計輪廓內包含的個別物體的數量。
20. Density (red):物體的平均紅色值。
21. Density (green):物體的平均綠色值。
22. Density (blue):物體的平均藍色值。
23. Per Area (Obj./Total): 物體面積與圖像總面積的比率。
24. Size (length):物體沿著主軸的 Feret 直徑(卡尺長度)。
25. Size (width):物體沿著次軸的 Feret 直徑(卡尺長度)。
26. Perimeter2:輪廓的鏈碼長度，如果存在孔洞，也包括孔洞的輪廓。比 Perimeter 更快，但精度較低。
27. IOD:積分光密度(或積分光強度)。也等於面積乘以平均光密度(或光強度)。
28. Perimeter (convex):物體凸輪廓的周長。



29. Perimeter (ellipse): 等效橢圓的周長。
30. Perimeter (ratio): 凸輪廓周長與周長之比。
31. Area (polygon): 包含在定義物體輪廓的多邊形中的面積。與周長相同的多邊形。
32. Count (adjusted): 根據大小加權的物體計數。僅在將 Clean Border 選項設置為「所有邊界」時有效。
33. Box Width: 物體外框的寬度。
34. Box Height: 物體外框的高度。
35. Feret (min): 最短卡尺長度。
36. Feret (max): 最長卡尺長度。
37. Feret (mean): 平均卡尺長度。
38. Density (min): 物體內部光強度(或光密度)的最小值。
39. Density (max): 物體內部光強度(或光密度)的最大值。
40. Density (std.dev.): 物體內部光強度(或光密度)的標準差。
41. Margination: 物體亮度在中心和邊緣之間的相對分佈，較亮的中心區域具有較大的值，均值的物體其值為 0.33。
42. Heterogeneity: 超過一定範圍（默認為 10%）與平均光強度有所偏差的像素比例。
43. Clumpiness: 膨脹後，相對於平均值偏離的像素比例，反映了紋理變化。
44. Density (sum): 物體內部光強度(或光密度)的總和。
45. Perimeter3: 修正的鏈碼周長，不計算孔洞，比 Perimeter2 更具客觀性。

### 4.3. 料源辨識模型

本研究旨在利用影像辨識所獲得的粒料特徵資料建立機器學習的監督式分類模型，開發料源辨識的模型。該模型的首要目標是以辨識不同種類的料源，即時獲



得各種料源的比例，以便進行配比的修正。不同種之料源的形狀特徵皆不相同，透過此料源辨識模型將料源分類可以更加精準地評估粒料的級配曲線以及扁平率等粒料特性。

### 4.3.1 特徵工程

特徵貢獻度為評估各個特徵對於標籤的重要性的指標之一，以協助選擇具有最大影響力的特徵。在本研究中，使用 sklearn 當中的函數 SelectKBest，其中的 score\_ 函數用於評估每個特徵的 ANOVA F-value。該分數衡量了特徵與標籤之間的關聯程度且使用了 Matplotlib 庫來繪製一個橫向的長條圖，顯示了各個特徵的 F-value 如圖 4.3-1 所示。當 F-value 越高時，表示特徵在分類模型中的解釋能力越強，與標籤之間的差異越大，這意味著該特徵在分類結果上具有更大的貢獻度。

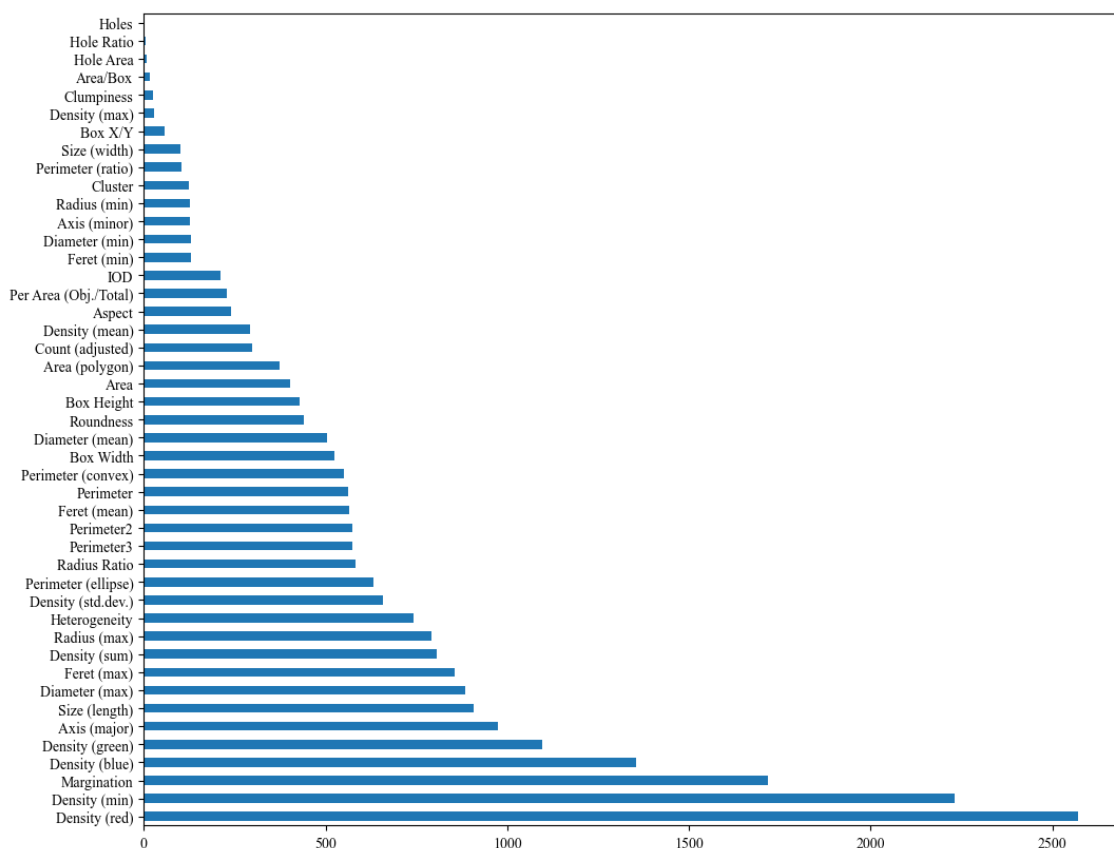


圖 4.3-1 料源辨識模型之各項特徵 F-value

訓練時若特徵過多會導致模型的複雜性也會增加，可能需要更多的計算資源和時間進行訓練和推斷，並使模型受到訓練數據中的噪聲和不具代表性的模式影響，而忽略了真正與標籤相關的特徵。使用 SelectKBest 當中的 k 值設定為 22，代表使用前 22 名的特徵資料(F-value 大於 500)作為此模型的訓練資料。圖 4.3-1 中顯示 F-value 前三名的分別為 Density(red)、Density(min)、Margination，表 4.3-1 與表 4.3-2 為各地區前三名特徵之詳細統計數據，資料顯示三原色當中的紅色以及粒料內顏色密度最小值為料源辨識主要關鍵，花蓮地區的紅色平均低於閩侯地區，整體色系會更接近灰色；閩侯地區的粒料內顏色密度最小值大多小於花蓮地區，可能因為閩侯地區粒料為石英斑岩，成分較複雜以及受高溫高壓生成的火山岩，比起花蓮地區顏色較為平滑的石灰石，閩侯地區的粒料上會有較多的淺色斑點，如圖 4.3-2 與圖 4.3-3 所示。Margination 兩個地區的數值幾乎相近，代表能夠有效幫助模型分類料源的特徵資料其效果有一定程度的侷限性。

表 4.3-1 F-value 最高三名特徵之統計數據 花蓮料源

花蓮	Density(red)	Density(min)	Margination
Mean	98.342	70.651	0.359
Std	10.708	18.566	0.013
Min	59.193	11.000	0.309
25%	91.264	58.000	0.350
50%	97.691	73.000	0.357
75%	104.605	84.333	0.366
max	144.025	160.000	0.430

表 4.3-2 F-value 最高三名特徵之統計數據 閩侯料源

閩侯	Density(red)	Density(min)	Margination
Mean	104.002	62.009	0.353
Std	11.194	17.337	0.013
Min	65.026	5.000	0.320
25%	96.584	50.333	0.345
50%	104.560	62.000	0.351
75%	111.133	74.000	0.360
max	141.044	117.000	0.424

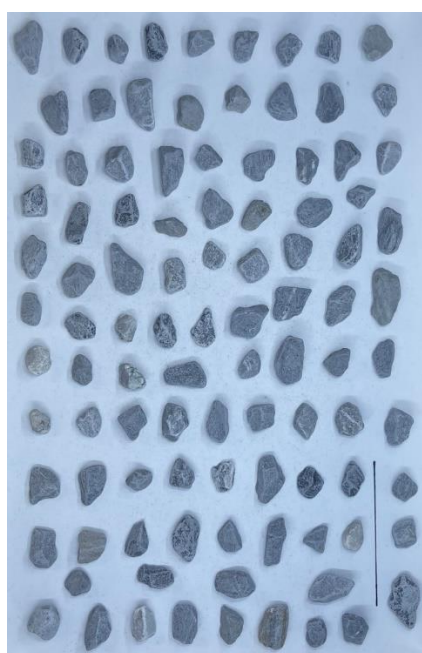


圖 4.3-2 花蓮地區

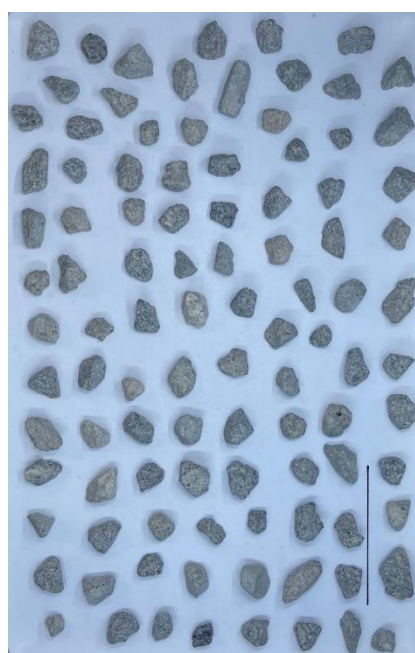


圖 4.3-3 閩侯地區



### 4.3.2 分類模型

本研究使用之演算法可以分為 RandomForest、LogisticRegression、DecisionTree、Xgboost 以及 Stacking，分述如下：

1. LogisticRegression 是由 David Cox[16]提出，其原理基於統計學中的廣義線性模型，適合用於分類問題，其中標籤為二元(如花蓮/閩侯)，並且假設特徵和標籤之間的關係可以用 Sigmoid 函數來表示。這個 Sigmoid 函數將連續的變數轉換為在 0 和 1 之間的概率值，用於表示某一個樣本屬於某一個類別的概率。特色為模型簡單易懂、計算效率高、適用於大型數據集等。
2. DecisionTree 最早為 J. Ross Quinlan[17]提出，基於樹狀結構的監督式機器學習算法，適用於解決分類和回歸問題。決策樹通過在每個節點上基於特徵值進行分割，從而將數據集分成多個子集，直到滿足停止條件。在每個節點的分割過程中，決策樹根據特徵的不同取值選擇最佳的分割標準，使得子集的純度或預測能力最大化。特色為易於理解和解釋、可處理類別型特徵、能夠捕捉非線性關係等。
3. RandomForest 是由 Leo Breiman[18]提出，其為自助重抽總合法(bagging)的集成式學習法，結合了決策樹和隨機特徵選擇的概念，由複數個決策樹組成，每個決策樹獨立地對數據進行學習和預測，根據多數投票原則(majority voting)決定最終的預測結果。特色為能夠處理高維數據和大型數據集，同時具有良好的泛化能力，能夠有效地應對過擬合問題。
4. Xgboost 是由陳天奇[19]提出，其全名為 eXtreme Gradient Boosting，為推升法(Boosting)的集成式學習法。XGBoost 的原理基於梯度提升樹(Gradient Boosting Tree)的思想，通過迭代地訓練多個弱學習器，並將它們組合成一個強大的集成式模型。特色為優化目標函數和正則化項，同時考慮模型的預測性能和複雜度，以達到更好的泛化能力和抗過擬合能力。



5. Stacking 是由 David H. Wolpert[20]提出，為堆疊式(Stacking)的集成式學習法。建立 Stacking 模型可以分為兩個主要步驟：

(1) 基礎模型(base model)的訓練和預測：使用訓練資料集將多個不同的基礎模型進行訓練，每個基礎模型獨立地從訓練數據中學習並生成預測結果。

(2) 元模型(meta model)的訓練和預測：將基礎模型的預測結果作為新的特徵資料，再加上原始的特徵資料，作為元模型的訓練數據。元模型使用這些訓練數據進行訓練，並進行最終的預測結果。

特色為使元模型能夠從基礎模型的預測中學習到更高層次的特徵表示(新特徵直接與標籤高度相關)，且可充分利用不同基礎模型之間的互補性，將各個模型的優點結合起來，提供元模型更強大和穩健的預測性能。本研究 Stacking 模型使用的基礎模型為：RandomForest、DecisionTree、Xgboost，元模型為：LogisticRegression。

使用上述五種演算法分別建立五個模型訓練，並且使用 K-Fold Cross-Validation 驗證模型的穩定性以及性能，k 值設定為 5，將數據集分成 5 個相等的部分，每次選取其中一部分作為驗證集，其他部分作為訓練集，進行 5 次交叉驗證，測試模型在不同的數據子集上的性能表現。結果呈現是使用 scikit-learn 的 cross\_val\_score 函數，其中 score 為 Accuracy，如表 4.3-3 所示。

表 4.3-3 每個模型之 5 次迭代以及總平均的 Accuracy

LogisticRegression	0.797	0.912	0.935	0.906	0.919	0.894
DecisionTree	0.724	0.807	0.847	0.862	0.832	0.814
RandomForest	0.734	0.828	0.894	0.877	0.810	0.839
Xgboost	0.775	0.867	0.926	0.906	0.867	0.862
Stacking	0.798	0.921	0.947	0.923	0.920	0.902

Stacking 模型在 K-Fold Cross-Validation 的 score 為最高，並使用 Stacking 模型繪製學習曲線觀察模型在訓練當中的有無過擬合或欠擬合的情況，透過不同大小下的訓練集和驗證集的評估指標來可視化模型的學習情況，使用 Accuracy(準確度)作為評分 score，衡量模型在預測中正確分類樣本的比例，如圖 4.3-4 所示。

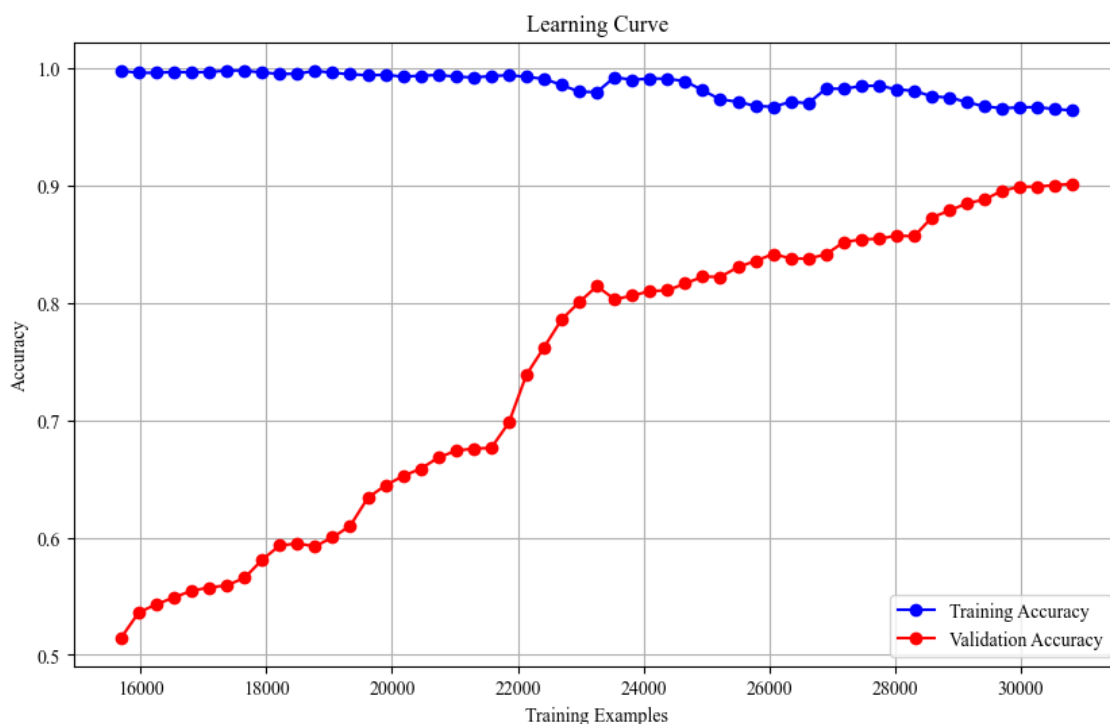


圖 4.3-4 料源辨識模型之學習曲線 演算法：Stacking

根據下述原因，選擇 Stacking 作為最終模型：

1. 複雜度：4.3.1 的特徵工程可以看出在料源辨識上大部分的特徵與標籤之間的關係較為薄弱，Stacking 可使訓練其基礎模型的特性，轉換成新的特徵資料至元模型以供訓練，且這些由基礎模型訓練的新特徵是直接與標籤高度相關的資料。
2. 穩定性：使用 K-Fold Cross-Validation 進行評估，在每次迭代以及整體平均表現上，Stacking 模型的準確度皆表現最佳。代表 Stacking 模型對於不同的交叉驗證切分具有穩定的預測能力，並且在整體上表現出色。

3. 泛化能力：觀察學習曲線的表現，可以看到 Stacking 模型的趨勢隨著樣本數的增加而逐漸收斂。這顯示模型已達到相對穩定的狀態，且訓練集和驗證集之間的差距表明模型沒有類似欠擬合或過擬合的情況。這證明了 Stacking 模型具有良好的泛化能力，可以對新的未見過的數據進行準確預測。

綜上所述，基於 Stacking 模型在特徵轉換、穩定性和泛化能力等方面的優勢，本研究選擇將其作為料源辨識之最終模型。

### 4.3.3 成效評估

測試資料作為最終的模型成效評估，其目的是測試模型在面對未訓練過的資料能否有良好的泛化能力，其中每一組別資料為花蓮以及閩侯粒料，組別 4、組別 5 與訓練組的取樣為同袋粒料，組別 6 為不同袋之粒料。使用 scikit-learn 中的 Classification Report，用於計算並生成分類模型的評估報告，以及混淆矩陣評估測試資料其成效。

表 4.3-4 料源辨識模型評估報告 組別 4

	Precision	Recall	F1 score	Support
花蓮 4	0.84	0.92	0.88	3452
閩侯 4	0.92	0.84	0.88	3838
Accuracy			0.88	7290
Macor avg	0.88	0.88	0.88	7290
Weighted avg	0.88	0.88	0.88	7290

其中 Support：表示每個類別在測試集中的樣本數量；Macro avg：Macro average 是計算每個類別的評估指標的平均值。對於每個類別，先計算該類別的評估指標，然後將所有類別的評估指標平均。對每個類別給予相等的權重，不考慮樣本數量的不平衡情況；Weighted avg：Weighted average 是計算每個類別的評估指標的加權

平均值，權重是根據每個類別在測試集中的樣本數量。樣本數量較多的類別在計算加權平均時具有更高的權重，而樣本數量較少的類別具有較低的權重。考慮了不同類別的樣本數量差異，並給予較大樣本數量的類別更高的影響力。

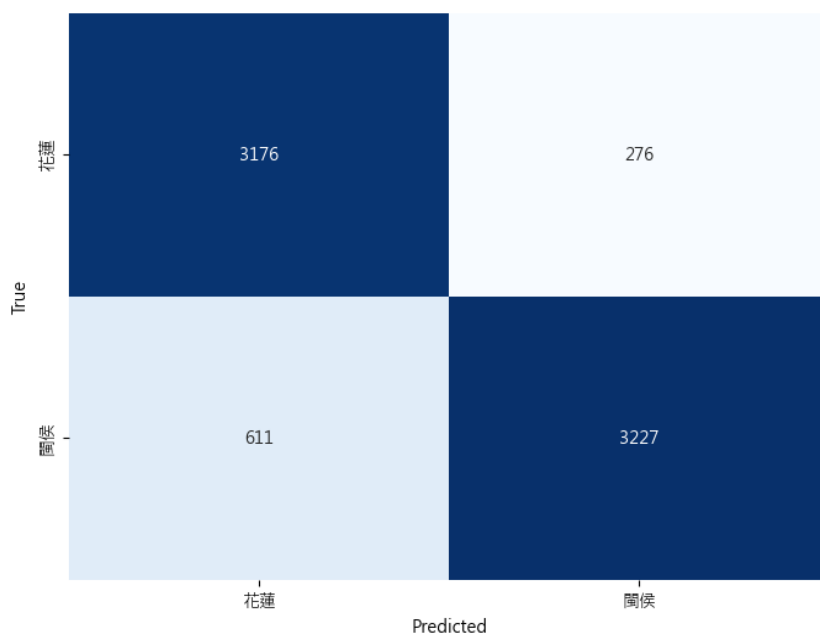


圖 4.3-5 混淆矩陣 組別 4

表 4.3-5 料源辨識模型評估報告 組別 5

	Precision	Recall	F1 score	Support
花蓮 5	0.79	0.83	0.81	2615
閩侯 5	0.88	0.85	0.86	3794
Accuracy			0.84	6409
Macor avg	0.84	0.84	0.84	6409
Weighted avg	0.84	0.84	0.84	6409

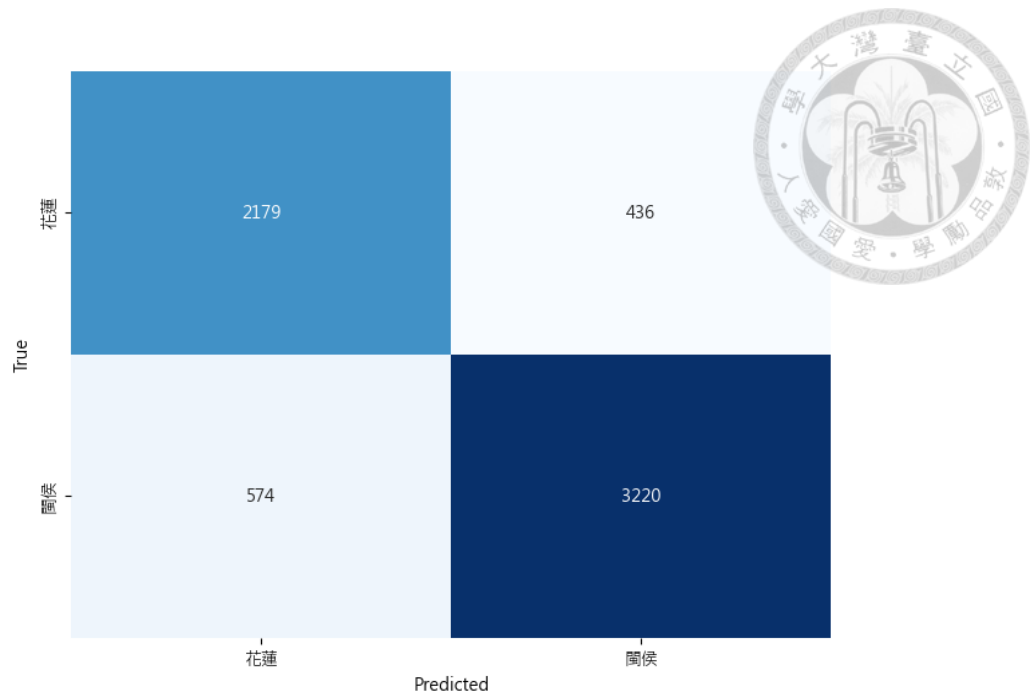


圖 4.3-6 混淆矩陣 組別 5

表 4.3-6 料源辨識模型評估報告 組別 6

	Precision	Recall	F1 score	Support
花蓮 6	0.78	0.59	0.67	2736
閩侯 6	0.48	0.70	0.57	1500
Accuracy			0.63	4236
Macor avg	0.63	0.65	0.62	4236
Weighted avg	0.68	0.63	0.64	4236

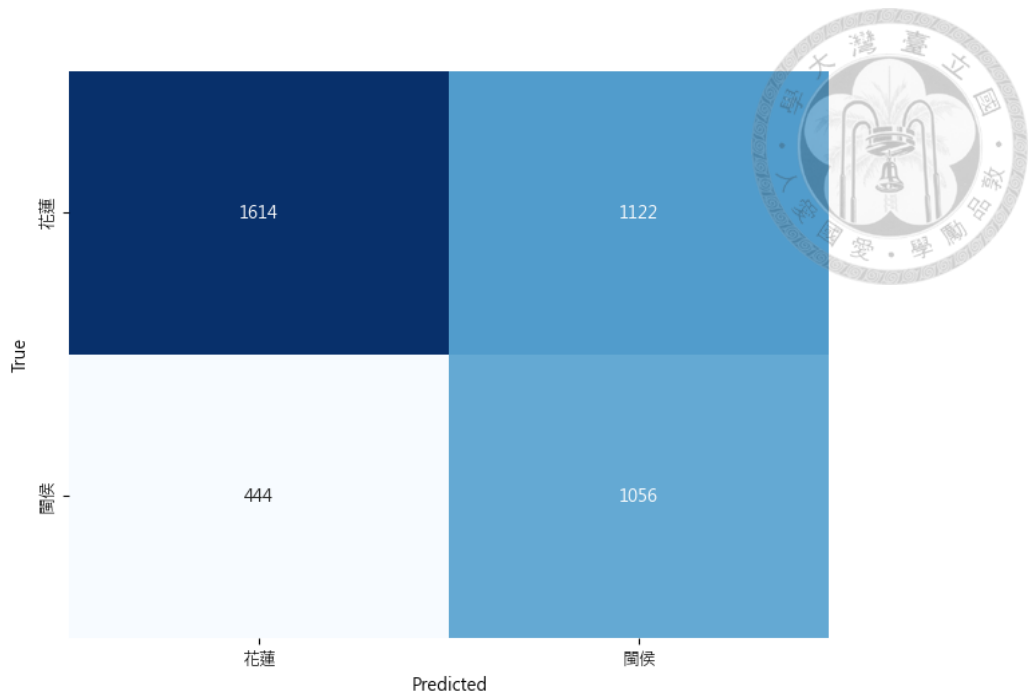


圖 4.3-7 混淆矩陣 組別 6

組別 4 以及組別 5 的 Precision、Recall、F1 score 和 Accuracy 皆有良好的表現，代表料源辨識模型在面對與訓練集相似的特徵資料能有優良的成效以及擬合能力。組別 6 使用與其他組別不同袋之粒料，花蓮 Recall 只有 0.59 以及閩侯 Precision 只有 0.48 的表現，整體的 F1 score 也只有 0.67 以及 0.57 的成效，代表組別 6 在料源辨識模型中表現不佳，造成組別 6 的表現較差可能原因如下：

1. 粒料特徵資料：組別 6 使用的是與其他組別不同袋的粒料，這些粒料可能具有特定的特徵資料，與其他組別的粒料有所不同。可能導致模型難以準確地對組別 6 的粒料進行分類，因為模型在訓練過程中主要學習到的是與其他組別相似的特徵。根據 4.3.1 對標籤貢獻度最高的三個特徵分別為 Density(red)、Density(min)、Margination，其中 Margination 的變異性相較前兩者相較小，故選擇前兩者之特徵資料，比較訓練組和組別 6 之差異性。表 4.3-7 與表 4.3-8 為花蓮料源與閩侯料源之前兩名特徵資料統計數據，數據結果表明，在 Density(red) 方面，花蓮和訓練組之間的差異不明顯，兩組數據的分佈相似且集中在相近的範圍內，並且沒有顯著的偏差；閩侯地區

組別 6 比訓練組整體上說數值更大，以訓練組的方面觀察兩料源之間閩侯地區原本數值整體較大，故 Density(red)影響較小。Density(min)在訓練組的閩侯料源之粒料因粒料性質會有較多的白色斑點，故數值小於花蓮地區，在組別 6 的數據可以觀察到花蓮地區的平均值略低於訓練組甚至更接近閩侯地區的訓練組，則閩侯地區高於訓練組也高於花蓮之訓練組。花蓮 6 相較於其他組的粒料，有更多白色以及其他色澤之粒料；閩侯 6 整體色澤較其他組別更深。綜上所述，通過對 Density(red)和 Density(min)等特徵的分析，組別 6 的特徵資料為可能造成模型分類結果不佳的成因之一。

表 4.3-7 Density(red)、Density(min)之統計數據 花蓮料源

花蓮	Density(red)	Density(red)	Density(min)	Density(min)
	訓練組	組別 6	訓練組	組別 6
Mean	98.342	98.627	70.651	65.762
Std	10.708	17.209	18.566	20.122
Min	59.193	61.920	11.000	8.333
25%	91.264	86.436	58.000	52.333
50%	97.691	95.266	73.000	64.333
75%	104.605	107.274	84.333	77.000
max	144.025	166.354	160.000	153.667

表 4.3-8 Density(red)、Density(min)之統計數據 閩侯料源

閩侯	Density(red) 訓練組	Density(red) 組別 6	Density(min) 訓練組	Density(min) 組別 6
Mean	104.002	112.083	62.009	81.234
Std	11.194	7.283	17.337	11.455
Min	65.026	89.700	5.000	44.333
25%	96.584	106.969	50.333	73.333
50%	104.560	111.227	62.000	80.667
75%	111.133	116.471	74.000	88.333
max	141.044	140.452	117.000	137.333

2. 資料分布不平衡：模型在訓練時學習到平衡資料的分佈和模式，當在測試時使用不平衡資料時，模型可能會遇到類別間的偏差，導致模型的表現下降。本研究訓練料源辨識模型使用的訓練資料比值為 1，故測試資料樣本數越平衡，則越能有效預測。組別 4 比值為 0.9，組別 5 為 0.7，比值越接近訓練資料則準確度越高，組別 6 的 Support 較少，且花蓮以及閩侯樣本之間的比值為 1.824，導致模型在預測過程中無法獲得足夠的資料量和模式來準確地分類組別 6。其中閩侯地區的粒料數量只有 1500，是因此批篩分析有 2kg 以上停留在 3/4”，其他組的 3/4”通常只有 300g 以內，且 3/4”不在本研究討論範圍，在取樣皆取 5kg 以上(5.5kg 以內)的前提下，使得閩侯 6 的粒料數量比例與其他組差異較大。

綜上所述，料源辨識模型在面對相似但未見過的資料能有一定的預測能力，則使用完全新的資料其預測能力略顯不足，若能夠增加不同袋粒料為新的訓練資料，模型能夠有穩定的效能以及擬和能力。

#### 4.4. 尺寸分類模型



本研究根據不同料源之間有著不同之特徵資料以及性質，故分別為花蓮地區以及閩侯地區訓練尺寸分類的監督式機器學習的分類模型，其目的是藉由多個因子(特徵資料)達到有效且精確的評估平面之粒料的篩號分類。

#### 4.4.1 特徵工程

與 4.1.1 使用相同的特徵選擇函數 SelectKBest 處理特徵選擇，以下分別為花蓮地區以及閩侯地區的特徵選擇：

##### 1. 花蓮地區：

依 F-value 大於 5000 為基準，k 值選定為 23，F-value 前三名的特徵分別為 Diameter(min)、Feret(min)、Axis(minor)，如圖 4.4-1 所示。其中 Diameter(min)

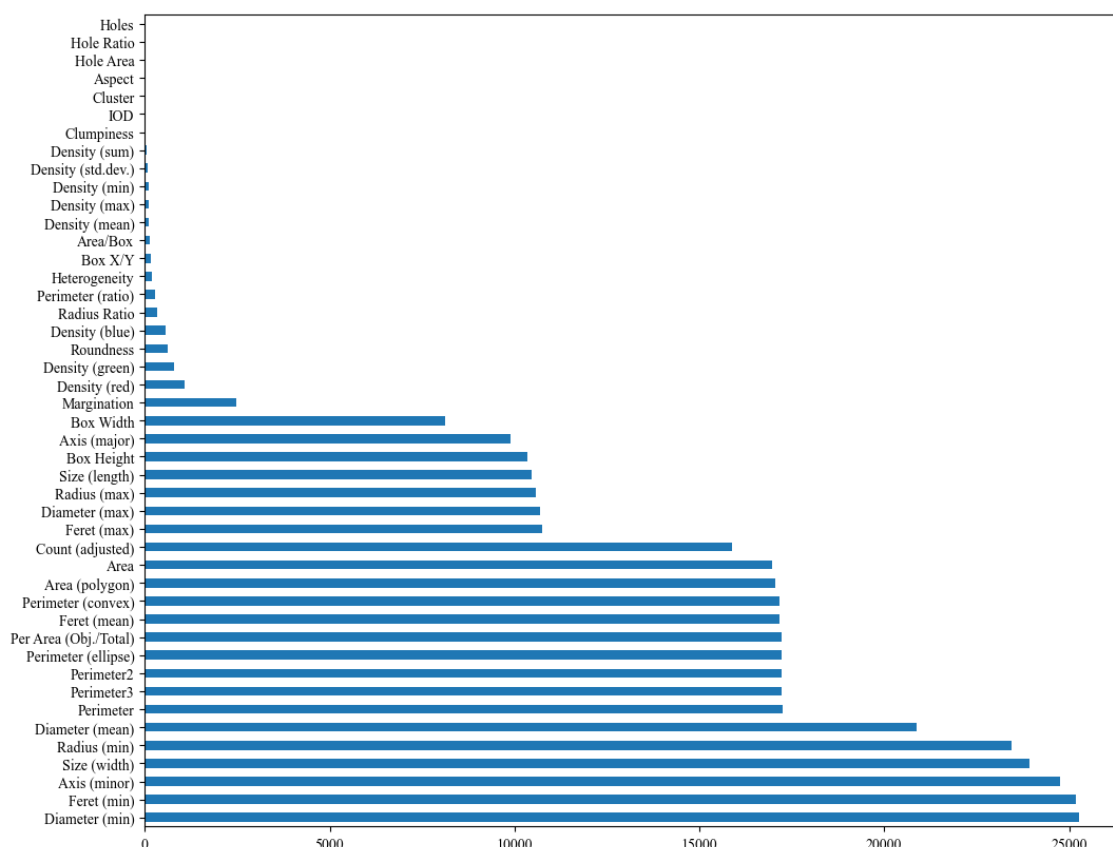


圖 4.4-1 尺寸分類模型之各項特徵 F-value 花蓮料源



## 2. 閩侯地區：

依 F-value 大於 5000 為基準，k 值選定為 24，F-value 前三名的特徵分別為 Diameter(min)、Axis(minor)、Feret(min)，如圖 4.4-2 所示。

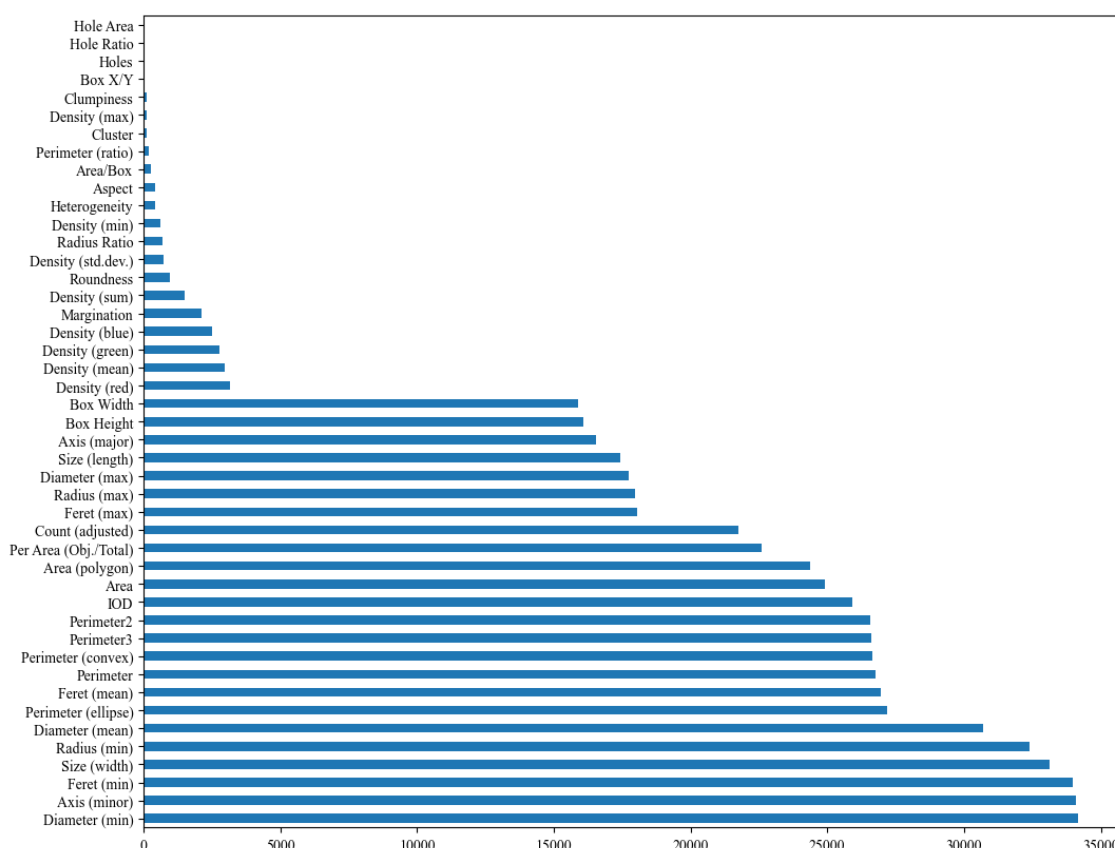


圖 4.4-2 尺寸分類模型之各項特徵 F-value 閩侯料源

上述之特徵工程皆顯示單顆粒料在影像平面中較小尺寸的維度能夠有效的分類尺寸，其中 Axis(minor)與 Size(width)為[5],[6]和[4]定義之粒料尺寸分類的因子。Diameter (min)為兩個地區模型的最高 F-value 特徵，其定義是物體輪廓上任意兩點的最短距離且通過重心的長度。篩號#4、3/8”、1/2”換算成公分單位分別為 0.475、0.952、1.27，根據表 4.4-1 與 4.4-2 可看出三組數據平均值之間的差異如同實際篩號大小的差異，則最小值與最大值的差距可以看出即使是位於同個篩號，可能差距會達到 1cm 以上甚至 2cm 以上，其造成差距可能原因如下：

1. 粒料角度：粒料傾倒時會有特定方向的呈現，即使反覆拍攝不同角度，重複拍攝到特定角度的機率較大，導致無法精確描述粒料之尺寸，少數情況會拍攝到粒料以豎立的方式立直，拍攝豎立的粒料會導致平面上特徵資料顯示粒料的尺寸較小，這種情況多發生在過長且寬厚相當之粒料。
2. 對角過篩：過篩時會產生對角過篩之現象，容易發現這種情況的粒料較為尖銳或者其一維度尺寸大於另外兩個維度過多，故在拍攝較小的篩號之粒料會出現部分尺寸過大之粒料。

透過 F-value 最高之特徵數據 Diameter (min)可以看出即使拍攝較小的篩號(如 #4)之粒料，也會有與較大的篩號(1/2")相當的粒料尺寸，甚至部分有超過的情況。故使用單因子的評估粒料之篩號會導致分類上的誤差，本研究之尺寸分類模型使用 20 種以上之評估因子評估粒料之尺寸分類。

表 4.4-1 Diameter (min)之統計數據 花蓮料源

Diameter (min)	#4	3/8"	1/2"
Mean	0.965	1.452	1.832
Std	0.177	0.220	0.249
Min	0.471	0.783	0.043
25%	0.845	1.298	1.658
50%	0.970	1.434	1.826
75%	1.088	1.595	1.991
max	1.568	2.331	2.979

表 4.4-2 Diameter (min)之統計數據 閩侯料源


Diameter (min)	#4	3/8"	1/2"
Mean	0.903	1.353	1.876
Std	0.171	0.173	0.277
Min	0.478	0.782	1.078
25%	0.792	1.227	1.671
50%	0.895	1.353	1.864
75%	1.030	1.479	2.069
max	1.452	2.043	3.245

#### 4.4.2 分類模型

同 4.3.2 使用 RandomForest、LogisticRegression、DecisionTree、Xgboost 以及 Stacking(基礎模型：RandomForest、DecisionTree、Xgboost，元模型：LogisticRegression)來分別為花蓮以及閩侯地區建立尺寸分類模型，K-Fold Cross-Validation 以及學習曲線來評估最終的模型選擇。表 4.4-3 與表 4.4-4 中兩料源之 K-Fold Cross-Validation 最高 Accuracy 皆為 RandomForest 模型，使用 RandomForest 模型分別繪製學習曲線，如圖 4.4-3 與圖 4.4-4 所示。

根據下述原因，選擇 RandomForest 作為兩種料源之最終模型：

1. 演算法特性：RandomForest 使用一種稱為特徵子集隨機化 (Feature Subsampling) 的技術。在每棵決策樹的訓練過程中，從原始特徵集合中隨機選擇一個子集，這個子集只包含部分特徵。每棵決策樹的判斷過程都只考慮部分特徵，從而降低了特徵之間的相關性，避免過度依賴單一特徵的情況。由於每棵決策樹都是在隨機選擇的特徵子集上進行訓練，每棵樹的預



測結果都有所不同。最終的預測結果是基於所有決策樹的集成，這樣可以減少單一決策樹的偏差，提高整個模型的準確性。

2. 穩定性：使用 K-Fold Cross-Validation 進行評估，在每次迭代以及整體平均表現上，RandomForest 模型的準確度皆表現最佳。其中 Stacking 以及 XGBoost 也有較高的準確度，Stacking 在訓練基礎模型以及元模型時，需將資料分割給每一個模型，在尺寸分類的訓練資料只有 18900 筆資料，故分割給每個模型的訓練資料過少，從而導致基礎模型以及元模型學習不佳；XGBoost 模型，其特點是將每個特徵都考慮進模型訓練中，可能導致對某些特徵過度依賴，對雜訊敏感，影響模型的穩定性。
3. 泛化能力：圖 4.4-3 與圖 4.4-4 為兩個模型之學習曲線，可以觀察到訓練集和驗證集的準確度隨著訓練樣本數的增加而提高，顯示隨著更多的訓練樣本被用於模型訓練，模型的性能有所提升。當訓練樣本數達到一定程度後，驗證集的準確度趨於平緩，沒有顯著的提升。這可能意味著模型已經達到了一個相對穩定的狀態。訓練樣本數到達 5000 以及 10000 時，驗證集皆有顯著的提升，可能原因是訓練資料集的排序，本研究合併訓練集是按照組別順序合併(如花蓮 1、花蓮 2、花蓮 3)，故在不同組之間的交際處會有顯著提升的現象。

綜上所述，RandomForest 模型的演算法特性、穩定性和泛化能力使其成為適合用於兩種料源的最終模型選擇。

表 4.4-3 每個模型之 5 次迭代以及總平均的 Accuracy 花蓮料源

LogisticRegression	0.816	0.825	0.807	0.763	0.725	0.787
DecisionTree	0.770	0.810	0.836	0.801	0.805	0.804
RandomForest	0.791	0.851	0.902	0.883	0.872	0.860
Xgboost	0.798	0.843	0.895	0.873	0.871	0.856
Stacking	0.810	0.864	0.893	0.875	0.807	0.850

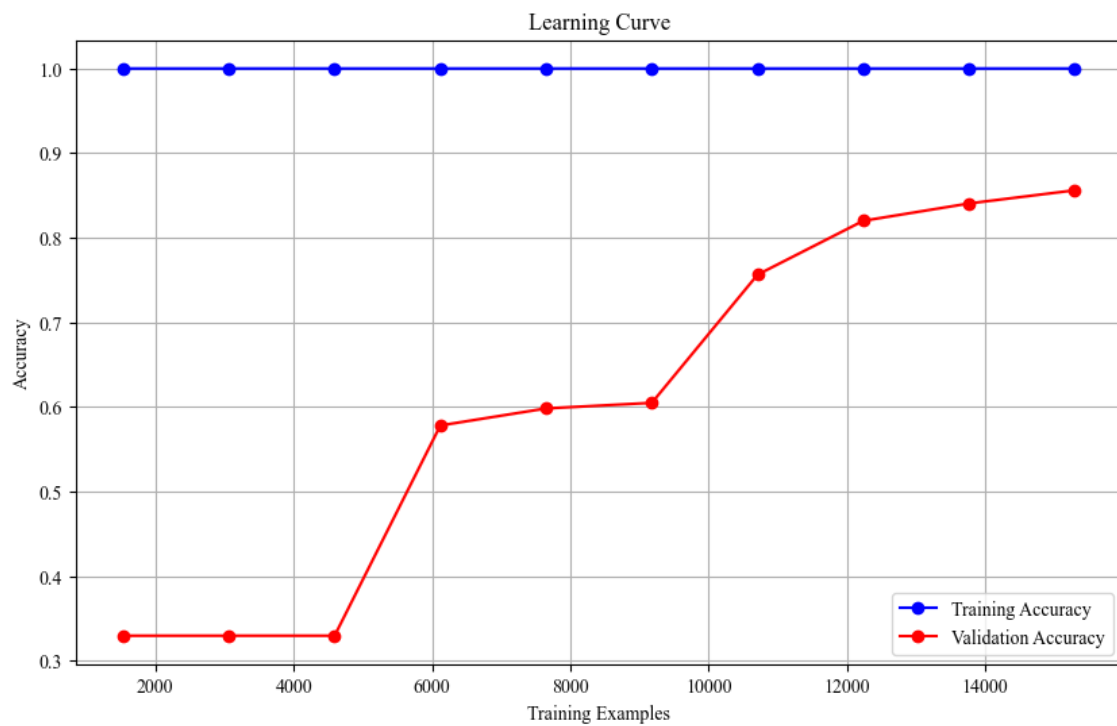


圖 4.4-3 尺寸分類模型之學習曲線 花蓮料源 演算法：RandomForest

表 4.4-4 每個模型之 5 次迭代以及總平均的 Accuracy 閩侯料源

LogisticRegression	0.826	0.848	0.867	0.874	0.886	0.860
DecisionTree	0.775	0.838	0.841	0.869	0.8841	0.841
RandomForest	0.841	0.881	0.877	0.910	0.886	0.879
Xgboost	0.813	0.881	0.885	0.903	0.887	0.874
Stacking	0.830	0.888	0.876	0.904	0.887	0.877

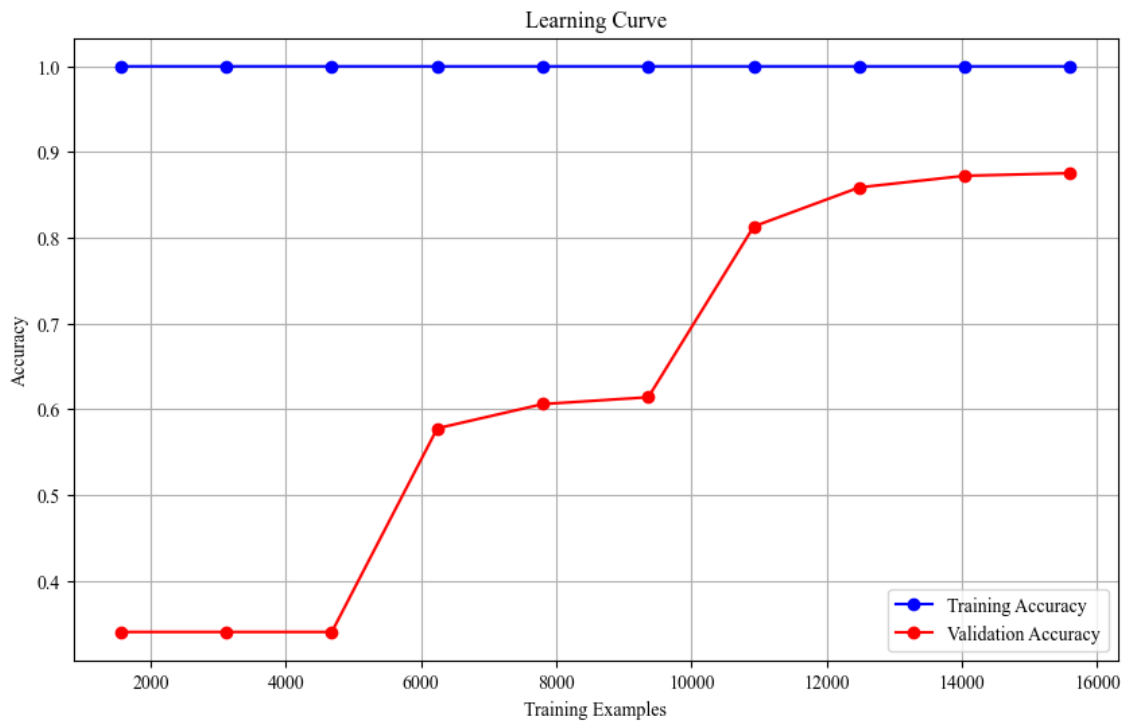


圖 4.4-4 尺寸分類模型之學習曲線 閩侯料源 演算法：RandomForest



### 4.4.3 成效評估

4.3.3 使用測試組分類完料源後，根據其預測的料源結果分別輸入花蓮尺寸分類模型以及閩侯尺寸分類模型。因粒料先經由料源辨識分類完料源後，再透過各料源之尺寸辨識模型分類尺寸，Precision、Recall、F1 score 和 Accuracy 原始定義無法精確描述最終分類結果，故建立 6×6 之混淆矩陣分析最終分類成果，如圖 4.4-6~4.4-8 所示。根據料源是否正確辨識分成兩類討論尺寸分類之成效，如圖 4.4-5 所示，其觀察結果如下：

1. 正確辨識料源(左上、右下)：組別 4~6 的結果顯示，兩料源之尺寸分類在 #4 篩粒料成效良好，F1 score 皆有 0.8 以上的表現。花蓮地區 3/8" 篩有過保守的趨勢，尤其在組別 4 和組別 6，其 Precision 只有 0.5 以及 0.52，則閩侯地區之 3/8" 篩皆有把部分 #4 預測為 3/8" 的表現，亦是組別 4 表現最差，其 Precision 只有 0.44。花蓮地區之 1/2" 篩有 Recall 較低的傾向，組別 4 只有 0.46 之表現，閩侯地區則是有 Precision 較低之趨勢，組別 4 以及組別 6 只有 0.66 以及 0.59。
2. 未正確辨識料源(左下、右上)：即便料源辨識錯誤，仍著重討論尺寸分類的正確性。觀察結果顯示，在 #4 篩的表現上與正確分類結果相似，整體 F1 score 表現皆有 0.8 以上之水準。兩個料源在 3/8" 篩分類中存在 Precision 較低的情況，而在 1/2" 篩分類中則呈現 Recall 較低的趨勢。

綜上所述，可以得出結論，無論料源是否正確辨識，尺寸分類的表現仍然具有一定的可靠性。然而在特定篩孔尺寸下，某些組別可能存在較保守的趨勢或分類錯誤的情況，這可能與特定篩孔尺寸的粒料特性有關。因此，在進行尺寸分類時，應特別關注此情況，並考慮進一步優化模型以提高分類的準確性。

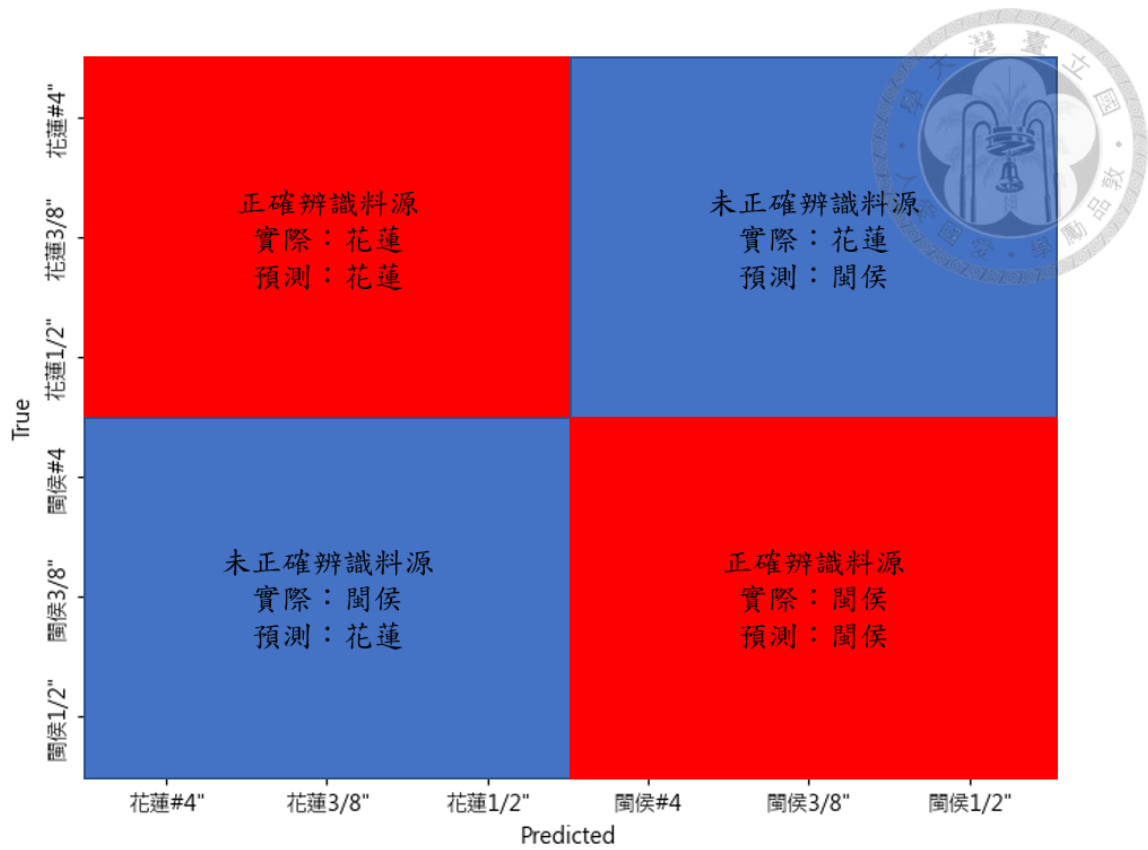


圖 4.4-5 最終分類結果分區分析示意圖

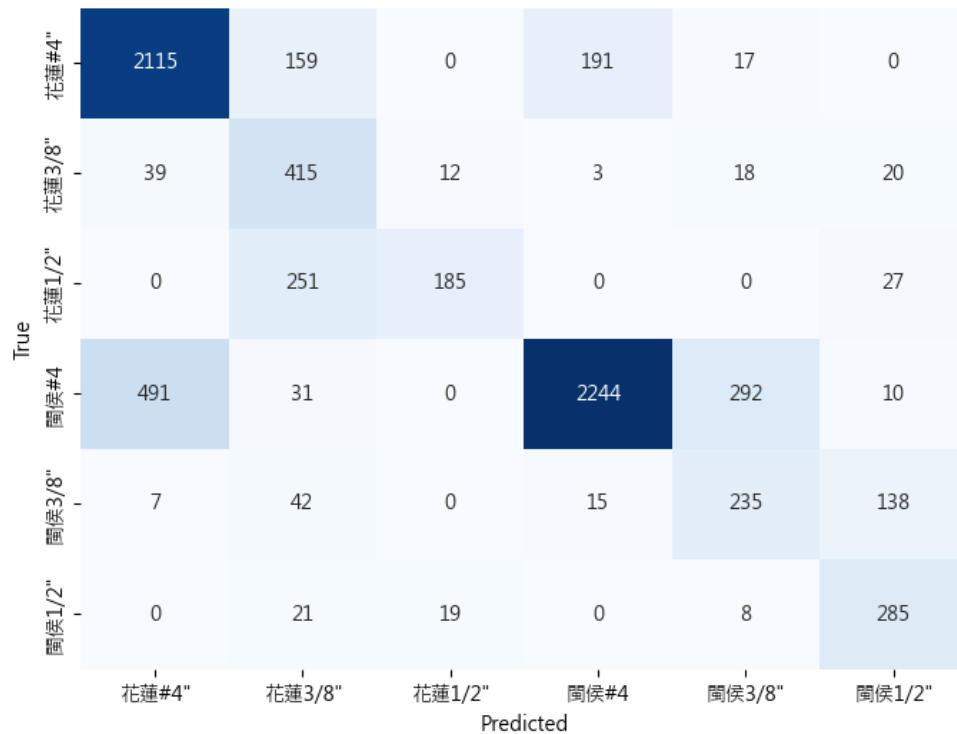


圖 4.4-6 最終分類結果之混淆矩陣 組別 4

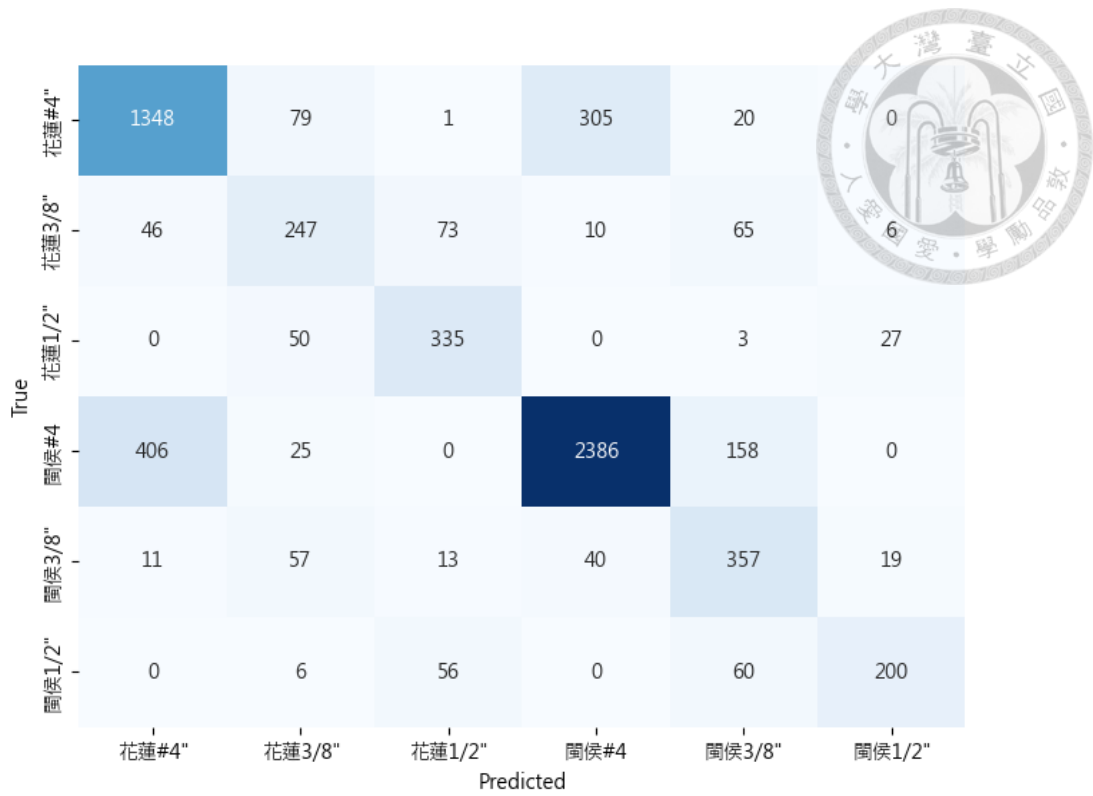


圖 4.4-7 最終分類結果之混淆矩陣 組別 5



圖 4.4-8 最終分類結果之混淆矩陣 組別 6



## 4.5. 質量轉換係數

料源辨識模型以及尺寸分類模型將粒料詳細分類成 6 類，要將粒料的影像數據轉換成實際使用的重量，需建立影像平面與實際立體之間的關係。本研究使用質量轉換係數，將各料源各層篩號之粒料資料轉換成各料源各層之重量，再將其結果繪製為級配曲線。

### 4.5.1 質量轉換係數定義

本研究參考[6]其中的質量轉換係數(Mass transfer coefficient, MTC)，透過大量攝影的影像粒料面積以及量測實際拍攝粒料的實際重量，透過兩者之間比值建立質量轉換係數，如式 4.5-1 所示。[6]提到相同料源之粒料因屬於同材質且使用同一粒料破碎機，故其形狀特徵在大量採樣下會有相似的特性或趨勢。本研究使用訓練資料分別對兩個料源的三層篩號分別為 1/2”、3/8”、#4 建立其質量轉換係數，分類完成的測試資料使用訓練資料建立之質量轉換係數將各層粒料面積轉換成各層重量，最後可繪製篩分析曲線。

$$MTC = \frac{W_{total}}{A_{total}} \quad \text{式 4.5-1}$$

其中：MTC 為質量轉換係數(Mass transfer coefficient)(g/cm<sup>2</sup>)；W<sub>total</sub>為該層粒料總重(Aggregates total weight)(g)；A<sub>total</sub>為該層粒料面積總和(Image aggregates total area)(cm<sup>2</sup>)。

### 4.5.2 質量轉換係數精度

質量轉換係數需大量採樣才能透過影像面積有效評估其質量，[6]提到分析精度隨著影像數量的增加而提高，大致在 50 張以上收斂。本研究建立質量轉換係數是對訓練資料之粒料進行重複採樣(Resampling)處理資料量不足，使用以下兩種方式進行重複採樣：



1. 對相同粒料且相同角度，改變手持短燈條之高度，使得影像之間的明亮程度以及影長有些微差異，重複拍攝 3 次。
2. 翻動粒料改變其角度，可使粒料呈現不同的角度，得到不同之特徵資料，重複拍攝 7 次。

透過上述兩種方式重複採樣可有效增加資料量，每組粒料會有 21 張影像，且訓練資料中有 3 組粒料，故每層質量轉換係數會有 63 張影像，再將其平均作為測試資料所使用之質量轉換係數，如表 4.5-1 所示。圖 4.5-1 與圖 4.5-2 為兩地區每張影像之質量轉換係數對各篩號之箱型圖，閩侯地區相較於花蓮地區之質量轉換係數較為收斂，其可能原因是閩侯地區粒料材質為石英斑岩，硬度遠高於花蓮地區之石灰石，故其粒料的形狀變化程度較小。

表 4.5-1 訓練資料建立之質量轉換係數

MTC(g/cm <sup>2</sup> )	1/2"	3/8"	#4
花蓮	1.599	1.064	0.770
閩侯	1.703	1.123	0.752

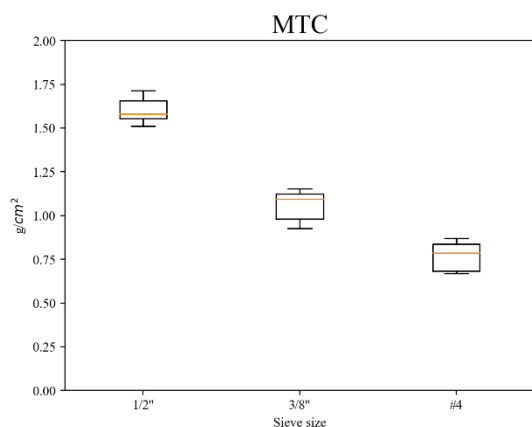


圖 4.5-1 花蓮地區之質量轉換係數箱型圖

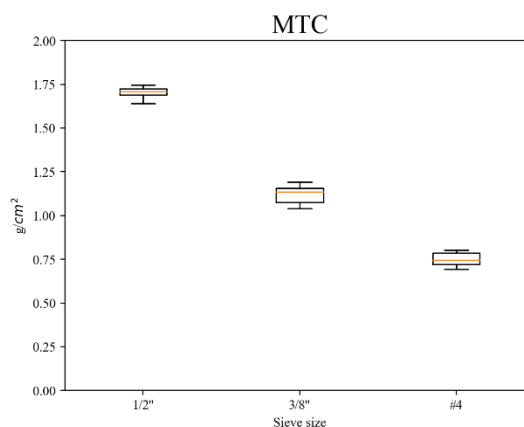


圖 4.5-2 閩侯地區之質量轉換係數箱型圖



## 4.6. 影像平均厚度

在傳統影像辨識文獻[7] - [9]辨識其粒料的扁平、扁長、細長顆粒，大多皆使用複數個相機拍攝粒料之三維形狀特徵，會導致拍攝環境更加嚴苛以及拍攝時間較為費時。本研究延伸質量轉換係數概念，在相同料源大量採樣之下，粒料之間會有相似的形狀特徵，提出影像平均厚度以及平均厚度比值估算粒料在平面影像當中之粒料厚度，再利用影像辨識之粒料長寬，從而推估粒料之扁平、扁長、細長率。

### 4.6.1 影像平均厚度定義

影像平均厚度是假設影像當中所有粒料厚度為一個定值，透過訓練資料當中各層的真实粒料重量除以其單位重，即可得到粒料總體積，再除以全部粒料面積總和，得到各層影像平均厚度，如式 4.6-1 所示。平均厚度比值為建立實際厚度以及影像厚度之間關係，其中實際粒料平均最大厚度是根據 2.3.1 中厚度定義以及使用電子游標尺量測粒料最大厚度將其平均，如式 4.6-2 所示。

$$t_i = \frac{W_{\text{total}}}{\gamma \times A_{\text{total}}} \quad \text{式 4.6-1}$$

其中： $\bar{t}_i$ 為影像平均厚度(Image average thickness)(cm)； $W_{\text{total}}$ 為粒料該層總重(Aggregates total weight)(g)； $\gamma$ 為乾單位重(Dry unit weight)(kgf/cm<sup>3</sup>)； $A_{\text{total}}$ 為粒料面積總和(Image aggregates total area)(cm<sup>2</sup>)。

$$t_{\text{ratio}} = \frac{\bar{t}_{a,\text{max}}}{\bar{t}_i} \quad \text{式 4.6-2}$$

其中： $t_{\text{ratio}}$ 為平均厚度比值(Thickness ratio)； $\bar{t}_{a,\text{max}}$ 為實際粒料平均最大厚度(Actual average maximum thickness)(cm)； $\bar{t}_i$ 為影像平均厚度(Image average thickness)(cm)。

根據 2.3.1 定義之粒料長寬，影像辨識之長邊使用 Feret(max)為卡尺直徑最大值，寬邊因影像辨識之特徵資料無相同定義之資料，故使用與實際寬度量測值平均

值最為接近的 Feret(mean)作為寬度之定義。影像辨識結果之顆粒扁平、細長和扁長定義為式 4.6-3~4.6-5。



$$\text{扁平} = \frac{\text{Feret}(\text{mean})}{t_i \times t_{\text{ratio}}}$$

$$\text{細長} = \frac{\text{Feret}(\text{max})}{\text{Feret}(\text{mean})} \quad \text{式 4.6-4}$$

$$\text{扁長} = \frac{\text{Feret}(\text{max})}{t_i \times t_{\text{ratio}}} \quad \text{式 4.6-5}$$

#### 4.6.2 影像平均厚度結果

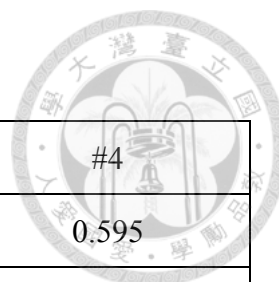
以下為訓練資料的平均影像厚度以及平均厚度比值之結果

表 4.6-1  $\bar{t}_{a,\text{max}}$ 、 $\bar{t}_i$ 、 $t_{\text{ratio}}$ 結果-花蓮 1

花蓮 1	1/2"	3/8"	#4
$\bar{t}_{a,\text{max}}$	1.110	0.833	0.612
$\bar{t}_i$	0.587	0.406	0.318
$t_{\text{ratio}}$	1.892	2.052	1.926

表 4.6-2  $\bar{t}_{a,\text{max}}$ 、 $\bar{t}_i$ 、 $t_{\text{ratio}}$ 結果-花蓮 2

花蓮 2	1/2"	3/8"	#4
$\bar{t}_{a,\text{max}}$	1.113	0.780	0.559
$\bar{t}_i$	0.582	0.411	0.297
$t_{\text{ratio}}$	1.912	1.896	1.885

表 4.6-3  $\bar{t}_{a,max}$ 、 $\bar{t}_i$ 、 $t_{ratio}$ 結果-花蓮 3

花蓮 3	1/2"	3/8"	#4
$\bar{t}_{a,max}$	1.176	0.871	0.595
$\bar{t}_i$	0.583	0.361	0.253
$t_{ratio}$	2.018	2.410	2.349

表 4.6-4  $\bar{t}_{a,max}$ 、 $\bar{t}_i$ 、 $t_{ratio}$ 結果-閩侯 1

閩侯 1	1/2"	3/8"	#4
$\bar{t}_{a,max}$	1.345	0.927	0.583
$\bar{t}_i$	0.649	0.442	0.288
$t_{ratio}$	2.073	2.096	2.026

表 4.6-5  $\bar{t}_{a,max}$ 、 $\bar{t}_i$ 、 $t_{ratio}$ 結果-閩侯 2

閩侯 2	1/2"	3/8"	#4
$\bar{t}_{a,max}$	1.266	0.877	0.597
$\bar{t}_i$	0.635	0.431	0.302
$t_{ratio}$	1.995	2.034	1.974

表 4.6-6  $\bar{t}_{a,max}$ 、 $\bar{t}_i$ 、 $t_{ratio}$ 結果-閩侯 3

閩侯 3	1/2"	3/8"	#4
$\bar{t}_{a,max}$	1.399	0.873	0.570
$\bar{t}_i$	0.652	0.402	0.271
$t_{ratio}$	2.145	2.173	2.103

表 4.6-7  $t_{ratio}$ 三組平均之結果

$t_{ratio}$	1/2"	3/8"	#4
花蓮	1.940	2.119	2.053
閩侯	2.071	2.101	2.034



## 第五章、分析結果與討論

### 5.1. 粗粒料篩析法結果

粗粒料篩析法結果透過影像辨識得到粒料特徵資料，再使用本研究的兩個模型，料源辨識模型以及尺寸分類模型分類粒料，以及 4.5.2 之質量轉換係數將各層粒料面積轉換成各層重量。本研究模擬真實篩分析之實驗，根據 CNS486 規定之最大粒徑 19mm，取 5 公斤以上之粒料，分別對花蓮料源以及閩侯料源各取 3 組做為測試組，對測試組進行粗粒料篩析法與影像拍攝和本研究建立之模型，最後將真實級配曲線與影像級配曲線相互比較以及分析其誤差可能原因，本研究探討對象為 1/2"、3/8"、#4 三層篩號之停留粒料，故 3/4" 之停留重量皆設定為 0。

#### 5.1.1 粗粒料篩析法結果

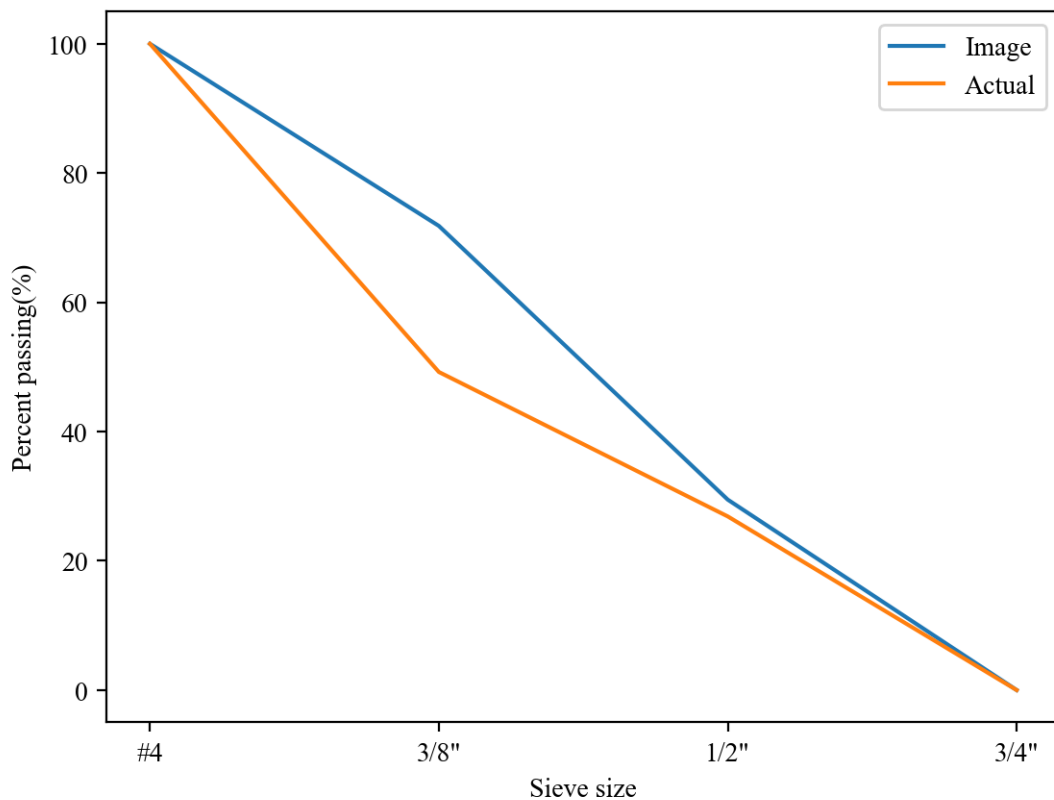


圖 5.1-1 級配曲線結果 花蓮 4

表 5.1-1 粗粒料篩析法以及影像辨識結果 花蓮 4

篩網 稱號	個別留篩 百分率(%) 真實	個別留篩 百分率(%) 影像	累 計	
			過篩重量百分率(%) 真實	過篩重量百分率(%) 影像
19.0mm(3/4")	0	0	100	100
12.5mm(1/2")	50.78	28.18	49.22	71.82
9.5mm(3/8")	22.35	42.36	26.87	29.46
4.75mm(#4)	26.87	29.46	0	0

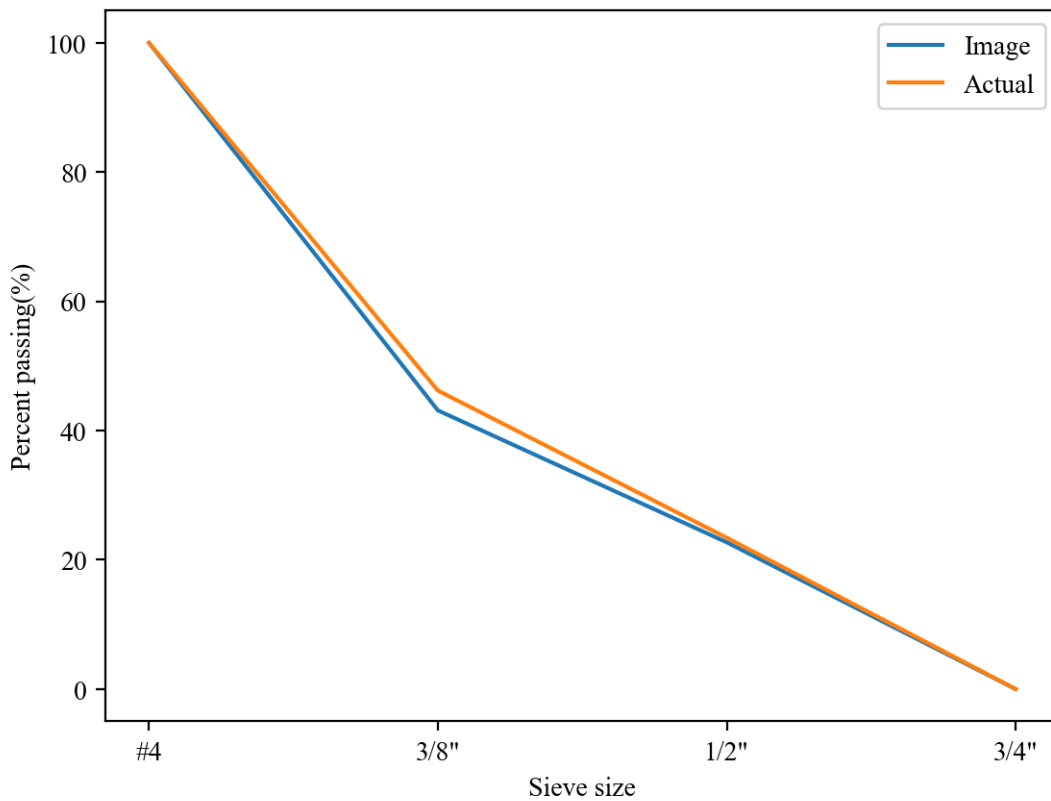


圖 5.1-2 級配曲線結果 花蓮 5

表 5.1-2 粗粒料篩析法以及影像辨識結果 花蓮 5

篩網 稱號	個別留篩 百分率(%) 真實	個別留篩 百分率(%) 影像	累 計	
			過篩重量百分率(%) 真實	過篩重量百分率(%) 影像
19.0mm(3/4")	0	0	100	100
12.5mm(1/2")	53.81	56.88	46.19	43.12
9.5mm(3/8")	22.78	20.42	23.41	22.71
4.75mm(#4)	23.41	22.71	0	0

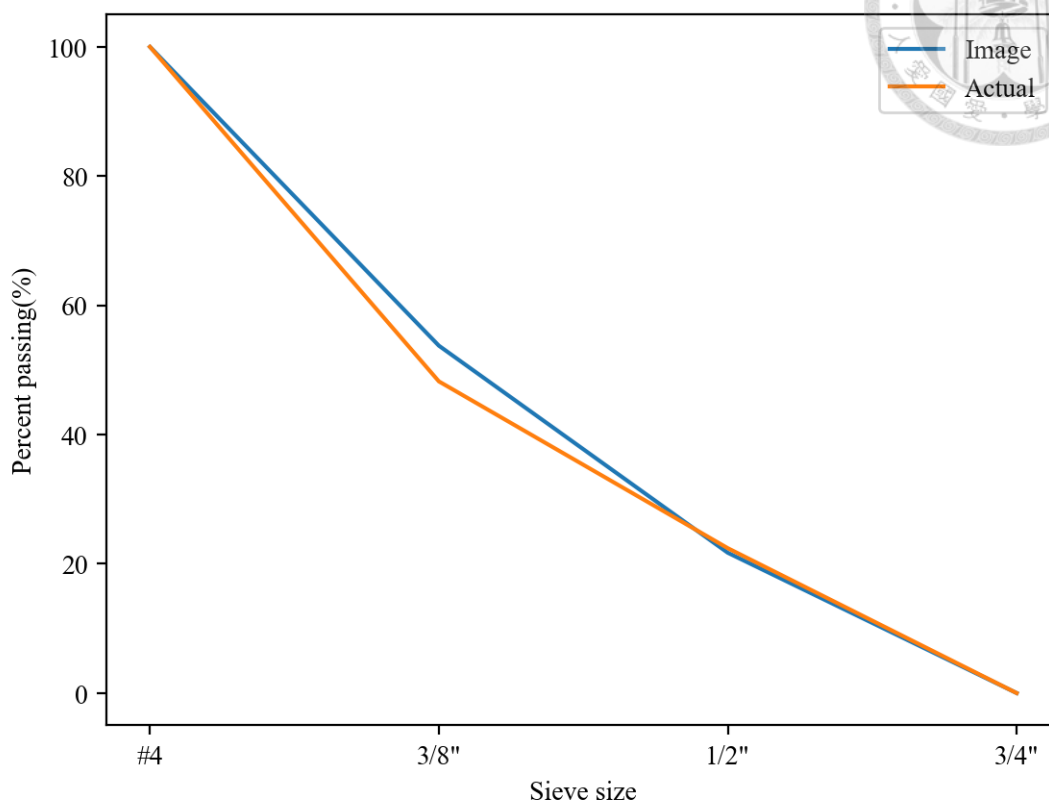
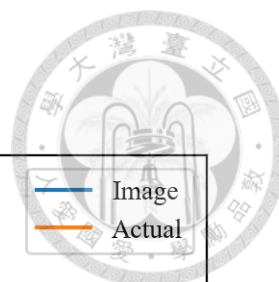


圖 5.1-3 級配曲線結果 花蓮 6

表 5.1-3 粗粒料篩析法以及影像辨識結果 花蓮 6

篩網 稱號	個別留篩 百分率(%) 真實	個別留篩 百分率(%) 影像	累 計	
			過篩重量百分率(%) 真實	過篩重量百分率(%) 影像
19.0mm(3/4")	0	0	100	100
12.5mm(1/2")	51.77	46.24	48.23	53.76
9.5mm(3/8")	25.81	32.04	22.42	21.71
4.75mm(#4)	22.42	21.71	0	0

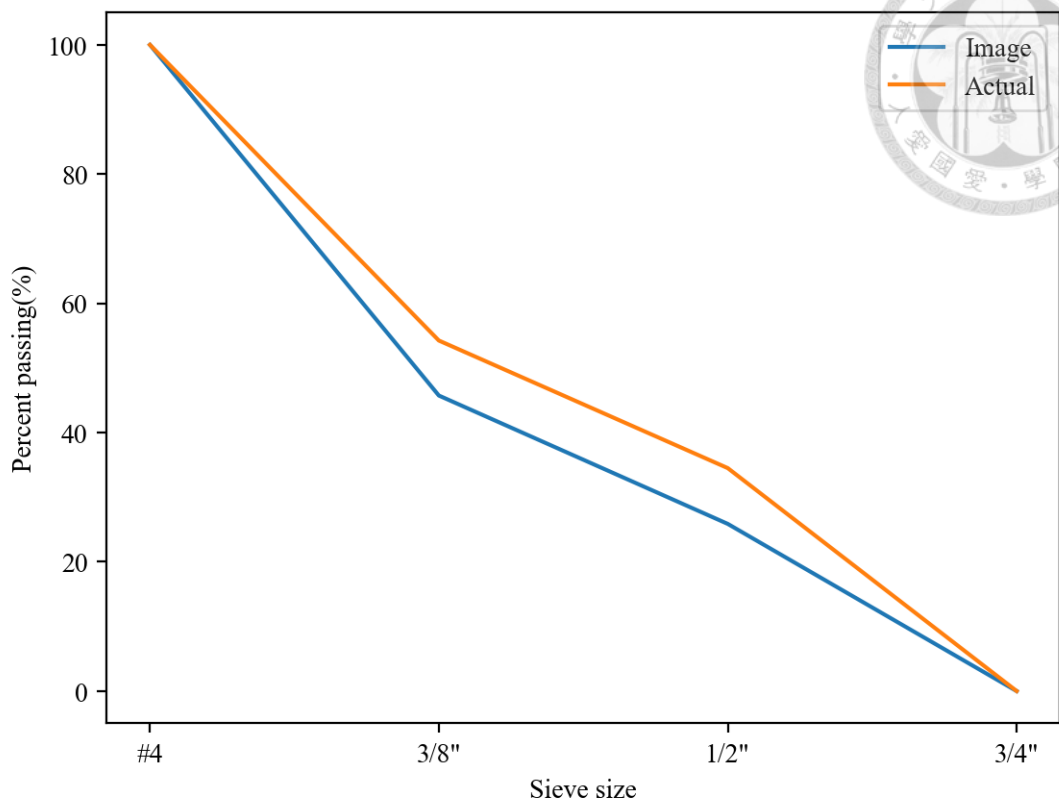


圖 5.1-4 級配曲線結果 閩侯 4

表 5.1-4 粗粒料篩析法以及影像辨識結果 閩侯 4

篩網 稱號	個別留篩 百分率(%) 真實	個別留篩 百分率(%) 影像	累 計	
			過篩重量百分率(%) 真實	過篩重量百分率(%) 影像
19.0mm(3/4")	0	0	100	100
12.5mm(1/2")	45.76	54.27	54.24	45.73
9.5mm(3/8")	19.74	19.86	34.50	25.87
4.75mm(#4)	34.5	25.87	0	0

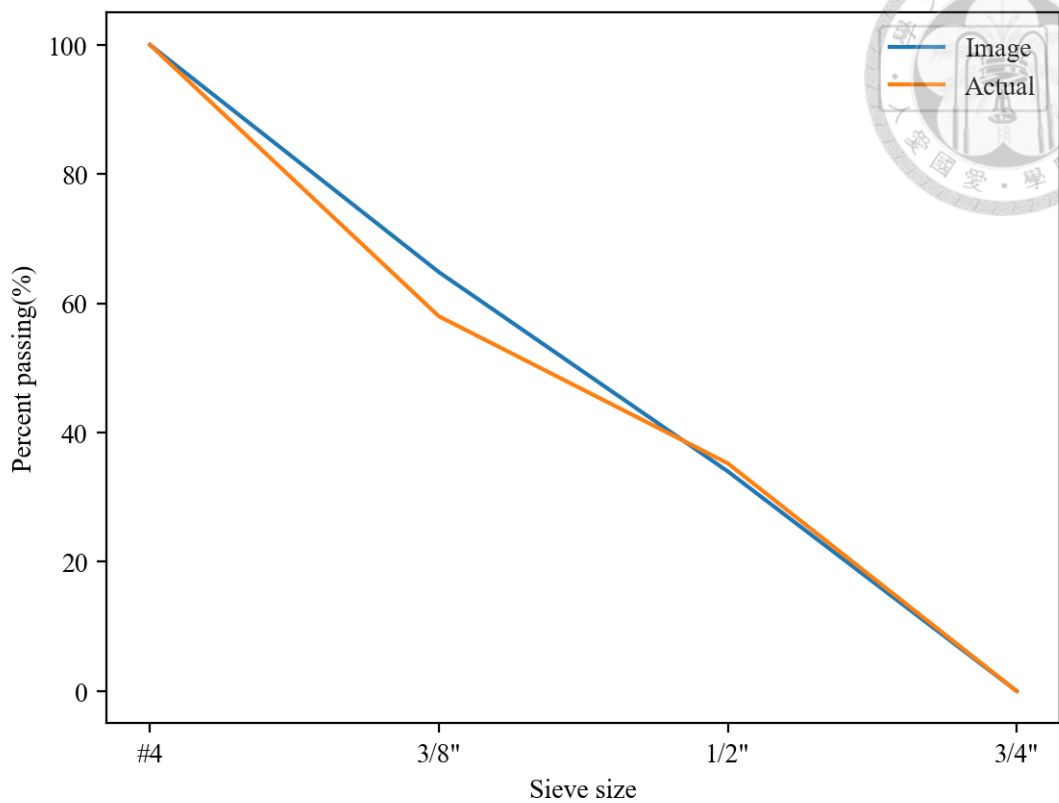


圖 5.1-5 級配曲線結果 閩侯 5

表 5.1-5 粗粒料篩析法以及影像辨識結果 閩侯 5

篩網 稱號	個別留篩 百分率(%) 真實	個別留篩 百分率(%) 影像	累 計	
			過篩重量百分率(%) 真實	過篩重量百分率(%) 影像
19.0mm(3/4")	0	0	100	100
12.5mm(1/2")	42.01	35.17	57.99	64.83
9.5mm(3/8")	22.75	30.83	35.24	34.00
4.75mm(#4)	35.24	34.00	0	0

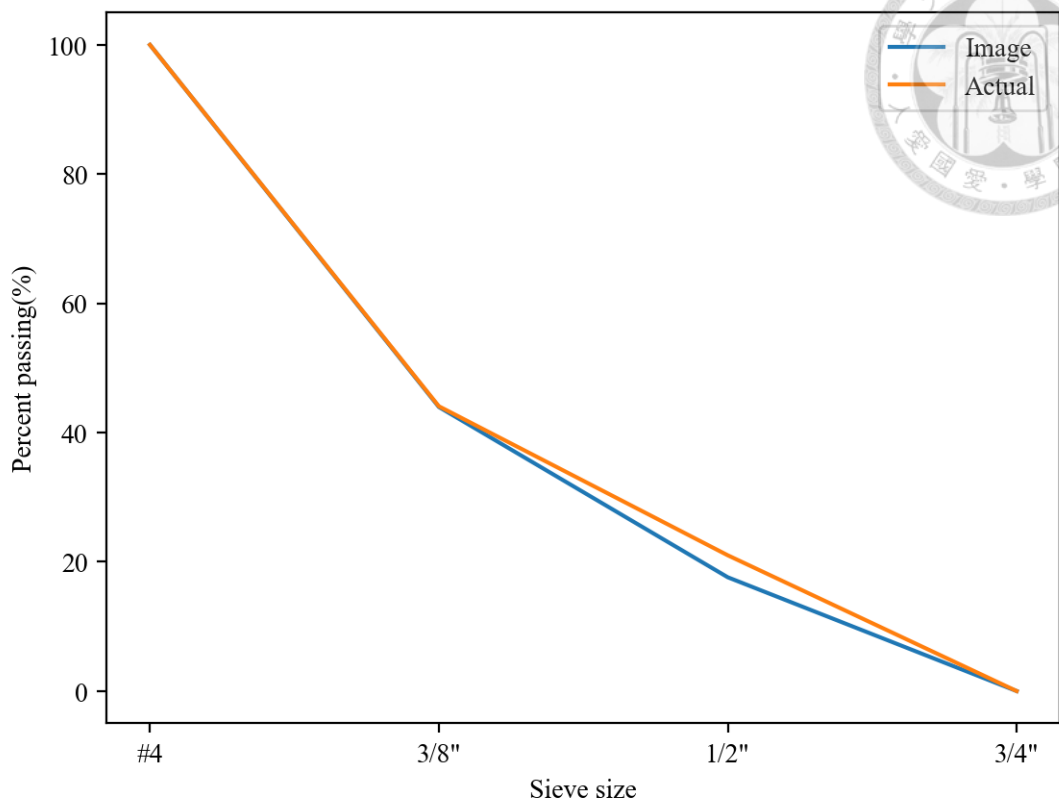


圖 5.1-6 級配曲線結果 閩侯 6

表 5.1-6 粗粒料篩析法以及影像辨識結果 閩侯 6

篩網 稱號	個別留篩 百分率(%) 真實	個別留篩 百分率(%) 影像	累 計	
			過篩重量百分率(%) 真實	過篩重量百分率(%) 影像
19.0mm(3/4")	0	0	100	100
12.5mm(1/2")	55.93	56.4	44.07	43.96
9.5mm(3/8")	23.06	26.35	21.01	17.61
4.75mm(#4)	21.01	17.61	0	0

### 5.1.2 粗粒料篩析法分析

以下使用絕對誤差百分比分析比較結果且統稱為誤差，其定義為 5.1-1 所示，以及使用 D10、D50、D90 比較其結果，分述如下：

1. D10：級配曲線上的 10%累積通過粒徑大小，10%的顆粒尺寸小於或等於 D10 的粒徑大小。



2. D50：級配曲線上的 50%累積通過粒徑大小，也稱為中值粒徑，50%的顆粒尺寸小於或等於 D50 的粒徑大小。
3. D90：級配曲線上的 90%累積通過粒徑大小，90%的顆粒尺寸小於或等於 D90 的粒徑大小。

$$\text{error}_{\text{abs}} = \left| \frac{\text{Pa} - \text{Pi}}{\text{Pa}} \right| \times 100 \quad \text{式 5.1-1}$$

其中： $\text{error}_{\text{abs}}$  為絕對誤差百分比(Absolute percentage error)(%)；Pa 為實際值；Pi 為影像值。

表 5.1-7~表 5.1-12 為兩地區各三組之試驗結果，D10、D50、D90 指標觀察級配曲線的分布，多數組別影像辨識結果與實際試驗值相近，甚至有些完全相符，花蓮料源中花蓮 5 誤差最小，最大誤差只有 2.433%，花蓮 4 誤差最大，其最大誤差有達 13.594；閩侯料源中閩侯 6 整體誤差最小，最大誤差只有 6.268%，閩侯 4 整體誤差最大，最大誤差為 9.817。其中可能造成誤差原因有尺規誤差、料源辨識、尺寸分類、質量轉換系數，其分述如下：

1. 尺規誤差：拍攝粒料時，底板的白只會有一條黑色比例尺線(10cm)，處理影像時比例尺為手動調整，可能導致粒料影像資料存在系統誤差。
2. 料源辨識：4.3.3 為料源辨識之成果，可以觀察到組別 4 之 Accuracy 高達 0.88；組別 5 之 Accuracy 有 0.84；組別 6 之 Accuracy 只有 0.63，但最後級配曲線的成果組別 4 反而是整體表現最差的，代表料源辨識造成之誤差較小。料源辨識使用的粒料特徵資料大多為顏色相關，與粒料形狀相關性較低。
3. 尺寸分類：4.4.3 為尺寸分類之結果，在組別 4 的混淆矩陣可以看到花蓮 1/2”篩有 463 顆，其中有 251 顆被誤判至 3/8”，閩侯 3/8”篩有 437 顆，其中有 138 顆誤判至 1/2”，導致 1/2”和 3/8”篩之間的比例差距過大，造成此結果可能原因為模型在 1/2”和 3/8”分類上有過擬合的情況，可能是因模型

的深度以及訓練資料不足造成，RandomForest 是建立多個 DecisionTree，其中每個 DecisionTree 的樹深度會影響模型性能，本研究模型平均深度為 30，樹的深度代表了模型可以進行分支和判斷的層數，深度較深的決策樹且訓練資料量有限時，模型容易過於複雜地擬合訓練資料中的噪音或細微變異，而無法泛化到新的資料，上述兩個原因可能導致模型在訓練資料上的表現非常好，但在遇到沒見過的數據可能會產生不佳的預測結果。組別 5 的花蓮地區以及組別 6 的閩侯地區結果顯示，最終分類結果與實際值趨勢相似，級配曲線的成果也會十分相近。組別 5 的閩侯地區以及組別 6 的花蓮地區，分類結果與實際值有部分差異，但其最大誤差只有 5.662%，整體的趨勢與實際值接近。表 5.1-7 可觀察到若 3/8”和 1/2”同時與實際值相差過大(±200 顆)，會造成較大誤差，因大顆粒誤差較大時更容易影響整體篩分析結果之重量比例。

表 5.1-7 分類最終成果總表 單位：顆

	實際 組別 4	預測 組別 4	實際 組別 5	預測 組別 5	實際 組別 6	預測 組別 6
花蓮#4	2482	2652	1753	1811	1729	1189
花蓮 3/8”	507	919	447	464	546	558
花蓮 1/2”	463	216	415	478	461	311
閩侯#4	3068	2453	2975	2741	918	1212
閩侯 3/8”	437	570	497	663	341	577
閩侯 1/2”	333	480	332	252	241	389

4. 質量轉換係數：分類為同一停留篩號粒料使用相同之質量轉換係數，然而 #4 篩停留粒料尺寸為 0.475~0.951cm；3/8”篩停留粒料尺寸為 0.952~1.26cm；

1/2”篩停留粒料尺寸為 1.27~1.904cm。1/2”以及#4 停留粒料範圍較大，導致使用定值評估整層篩號粒料會產生一定的誤差。

本研究中的影像辨識結果在評估級配曲線的分布上具有一定的可行性，即便有上述誤差來源以及尺寸分類模型有較大的誤差，試驗結果與實際試驗值趨勢幾乎一致。總體而言，本研究之模型在級配曲線中具有高水準的準確性，可透過進一步的研究和改進，提高模型的性能和準確性，可以實現更準確和可靠的結果。

表 5.1-8 D10、D50、D90 結果比較 花蓮 4

花蓮 4	真實(cm)	影像(cm)	error <sub>abs</sub> (%)
D10	0.653	0.637	2.450
D50	1.280	1.106	13.594
D90	1.780	1.678	5.730

表 5.1-9 D10、D50、D90 結果比較 花蓮 5

花蓮 5	真實(cm)	影像(cm)	error <sub>abs</sub> (%)
D10	0.679	0.685	0.884
D50	1.315	1.347	2.433
D90	1.787	1.793	0.336

表 5.1-10 D10、D50、D90 結果比較 花蓮 6

花蓮 6	真實(cm)	影像(cm)	error <sub>abs</sub> (%)
D10	0.688	0.695	1.017

D50	1.292	1.233	4.567
D90	1.782	1.768	0.786

表 5.1-11 D10、D50、D90 結果比較 閾侯 4

閾侯 4	真實(cm)	影像(cm)	error <sub>abs</sub> (%)
D10	0.613	0.659	7.504
D50	1.202	1.320	9.817
D90	1.767	1.789	1.245

表 5.1-12 D10、D50、D90 結果比較 閾侯 5

閾侯 5	真實(cm)	影像(cm)	error <sub>abs</sub> (%)
D10	0.610	0.615	0.820
D50	1.158	1.117	3.541
D90	1.754	1.724	1.710

表 5.1-13 D10、D50、D90 結果比較 閾侯 6

閾侯 6	真實(cm)	影像(cm)	error <sub>abs</sub> (%)
D10	0.702	0.746	6.268
D50	1.337	1.338	0.075
D90	1.791	1.792	0.056

## 5.2. 扁平、細長或扁長顆粒

本研究使用影像平均厚度以及平均厚度比值，模擬平面影像粒料之厚度，以及影像辨識得到粒料之長寬，計算且得到影像粒料之扁平、細長、扁長率，與



CNS15171 之扁平、細長或扁長顆粒試驗結果相互比較，本研究使用顆粒百分比作為結果呈現，且設定之比例為 1：3。

### 5.2.1 扁平、細長或扁長顆粒結果

表 5.2-1 扁平、細長或扁長顆粒結果 左：真實、右：影像 花蓮 4

真實	1/2"	3/8"	#4	影像	1/2"	3/8"	#4
扁平率	3%	5%	2%	扁平率	2%	3%	3%
細長率	0%	0%	1%	細長率	0%	0%	0%
扁長率	7%	24%	23%	扁長率	12%	30%	25%

表 5.2-2 扁平、細長或扁長顆粒結果 左：真實、右：影像 花蓮 5

真實	1/2"	3/8"	#4	影像	1/2"	3/8"	#4
扁平率	1%	4%	7%	扁平率	3%	5%	4%
細長率	0%	1%	0%	細長率	0%	0%	0%
扁長率	11%	28%	27%	扁長率	11%	22%	22%

表 5.2-3 扁平、細長或扁長顆粒結果 左：真實、右：影像 花蓮 6

真實	1/2"	3/8"	#4	影像	1/2"	3/8"	#4
扁平率	0%	7%	6%	扁平率	3%	2%	7%
細長率	0%	1%	0%	細長率	0%	0%	0%
扁長率	15%	15%	23%	扁長率	14%	20%	31%

表 5.2-4 扁平、細長或扁長顆粒結果 左：真實、右：影像 閩侯 4

真實	1/2"	3/8"	#4	影像	1/2"	3/8"	#4
扁平率	0%	3%	5%	扁平率	1%	1%	4%
細長率	0%	0%	0%	細長率	0%	0%	0%
扁長率	4%	6%	12%	扁長率	1%	6%	13%

表 5.2-5 扁平、細長或扁長顆粒結果 左：真實、右：影像 閩侯 5

真實	1/2"	3/8"	#4	影像	1/2"	3/8"	#4
扁平率	0%	2%	3%	扁平率	0%	0%	4%
細長率	0%	0%	0%	細長率	0%	0%	0%
扁長率	2%	13%	12%	扁長率	0%	12%	12%

表 5.2-6 扁平、細長或扁長顆粒結果 左：真實、右：影像 閩侯 6

真實	1/2"	3/8"	#4	影像	1/2"	3/8"	#4
扁平率	0%	1%	1%	扁平率	0%	0%	2%
細長率	0%	0%	0%	細長率	0%	0%	0%
扁長率	5%	5%	11%	扁長率	9%	2%	14%

## 5.2.2 扁平、細長或扁長顆粒分析

根據 5.2.1 的分析結果，為花蓮以及閩侯料源對於不同篩號尺寸的扁平率、細長率和扁長率的試驗結果。扁平、細長或扁長顆粒結果分析使用顆粒差異數且統稱為差值，單位為顆(每筆數據結果皆為每 100 顆之百分比)，以下是對其結果的觀察：


1. 扁平率：兩地區之扁平率整體誤差較小，則花蓮地區相較閩侯地區誤差較大，花蓮地區最大差值為 3，則閩侯地區為 2，代表影像方法對於扁平率的評估存在一定的偏差。



2. 細長率：兩料源在細長率之結果幾乎一致，最大差值只有 1。
3. 扁長率：影像方法對於扁長率的測量結果有較大之差異，花蓮地區之誤差大於閩侯地區。其中花蓮地區之最大差值達 8，閩侯地區為 4，代表影像方法對於扁長率的評估也存在一定的偏差。

綜合以上觀察，可以看出影像辨識在細長率有良好的表現，扁平率以及扁長率有部分誤差，其造成誤差可能原因如下：

1. 假設侷限性：影像平均厚度的前提是假設停留在同一篩號粒料皆為同一厚度，真實粒料厚度具有變異性，此假設使得影像厚度失去其變異性，可能為誤差來源之一。多數粒料為凸體，具有一定之厚度，少數粒料是因厚度極薄(與長寬之間比例超過 1：5)，則被歸類為扁平或扁長顆粒，這種極薄粒料其長寬可能與正常粒料相當，使用影像平均厚度評估粒料時，無法有效區分極薄粒料，可能為造成扁長率誤差較大原因之一。
2. 粒料形狀：[7]-[9]提到現今規範在粒料的長寬高描述不夠精準，因粒料在一些面上可能呈現凸多邊形、凹多邊形，甚至大多數粒料為非正幾何形狀。不規則形的粒料在進行試驗時，會有量測者主觀判斷寬度、厚度所造成的誤差，可能造成影像辨識與實際試驗值比較存在系統誤差。
3. 粒料特性：閩侯料源之粒料為石英斑岩，莫氏硬度為 7；花蓮料源之粒料主要為石灰石，其莫氏硬度為 3。粒料的軟硬程度可能會影響粒料破碎之變異性，粒料從原礦開採、破碎和搬運過程都會經歷許多的碰撞以及搖晃，故較軟的花蓮粒料比閩侯料源之粒料在形狀上可能具有更高的變異性，4.6.2 為兩地區訓練資料之平均厚度比值，數據顯示閩侯地區之  $t_{ratio}$  多為  $2 \pm 0.1$ ，則花蓮地區  $t_{ratio}$  最小值 1.885、最大值 2.349，以及圖 4.5-1 和圖 4.5-2 為兩地區質量轉換系數箱型圖，花蓮地區數據相較閩侯地區較為發散，上述兩點可觀察出閩侯地區相較花蓮地區變異性更小，故粒料特性可能造成花蓮地區整體誤差略高於閩侯地區。

- 
4. 平面影像：使用一台攝影儀器拍攝粒料具有侷限性，只能拍攝到粒料特定角度，可能導致在影像中得到的粒料長寬與真實粒料存在一定的差異。
  5. 評估因子：本研究使用的寬度定義為 Feret(mean)，卡尺直徑之平均值，因與實際量測寬度值比較，Axis minor、Feret(mean)和 Size(width)相比，Feret(mean)整體上結果最為接近實際量測值，故選其作為寬度。但 Feret(mean)與實際定義寬度不一，可能為其誤差來源之一。

使用影像平均厚度評估粒料形狀可以快速獲取結果且具有一定水準。然而，在扁平率和扁長率可能存在一些誤差，這些誤差可能由粒料的形狀特性、特定地區的粒料特性以及影像平面限制等因素所引起。後續研究工作可以進一步提高影像辨識的可靠性，改進評估因子的準確性，並深入探討不同形狀指標的有效性。

## 第六章、結論與建議



### 6.1. 結論

1. 本研究利用影像辨識以及機器學習模型，對粒料進行詳盡的分類，料源辨識模型的成果可以看出模型具有一定的準確性，有高達 0.88 之準確度，模型能夠快速而準確地識別不同料源的粒料，並提供其在整體粒料組成中的比例，對預拌混凝土的生產以及品質管理上有幫助，利用分類結果，可以及時了解預拌混凝土的工作性能和強度等特性的變化。例如，可以了解級配曲線、扁平率或岩心強度等與混凝土品質相關的參數。對於生產調整和品質控制提供了重要的信息。根據不同的料源之粒料特徵，兩料源的尺寸分類模型把粒料分類至各號篩，以及訓練資料建立之質量轉換係數，得到級配曲線其結果與真實試驗趨勢非常接近，D10、D50、D90 也與真實試驗值誤差較小。代表本研究之影像辨識以及模型可以在極短的時間內獲得與實際試驗相當的結果，有效節省人力以及時間且即時了解級配資訊可以隨時調整配比或者調整三分石以及六分石比例已符合規範需求等，達到製成管控的效果。本研究的結果顯示影像辨識和機器學習技術在預拌混凝土生產和品質管理領域具有巨大的潛力和應用價值。
2. 傳統的粒料顆粒扁平、細長和扁長率試驗通常需要花費 2 小時以上，耗費大量時間。本研究利用影像平均厚度和  $t_{ratio}$ ，以及影像辨識其粒料長寬並進一步計算粒料的扁平、細長和扁長率。結果顯示與實際試驗結果相近，特別是扁平率以及細長率誤差較小，且只需不到 30 分鐘的時間。使用影像辨識和計算方法大大節省了評估粒料形狀所需的時間成本，同時提供了快速而準確的結果。這對於預拌混凝土的生產和品質管理等應用領域具有重要意義，能夠提高生產效率並節省成本。這種基於影像辨識的方法不僅速度快、效率高，而且減少了人工操作的需求，降低了主觀誤差的可能性。因此，這種高效、快速的評估

方法在工程師和研究人員中具有實際的應用價值。提供了一個快速而可靠的工具，以評估粒料特性並進行相應的處理和決策。





## 6.2. 建議

1. 本研究對象為三分石以及六分石之粒料(4.75~19mm)，可以考慮將 3/4 英寸篩和#5 篩的停留粒料納入模型，可以增加級配曲線的完整性，提供更全面和準確的級配信息。擴展了研究的應用範圍，並有助於更好地了解和控制預拌混凝土的品質。
2. 實際應用於預拌廠時，粗粒料分為三分石以及六分石分開存放以及拍攝，模型使用時可能會遇到數據不平衡的情況，數據不平衡可能對模型的性能和準確性產生影響。由於樣本數量不平衡，模型可能會傾向於對樣本數量較多的類別進行較好的預測，而對樣本數量較少的類別預測較差。在應用本研究的模型時，需要特別關注數據不平衡的情況，並選擇合適的方法來處理這個問題。
3. 本研究在拍攝時為要求減少影像前處理的流程，粗粒料會逐一分開使得粗粒料沒有重疊情況發生。實際應用於水泥預拌廠時，會有很多粒料重疊的情況發生，可以設計演算法處理重疊粒料的特徵資料，特別是粒料形狀變化過大和重疊陰影過多的資料，使得數據處理流程更加自動化。
4. 本研究在顆粒扁平、細長和扁長率的比例選用 1：3，為使數據變化較大，故選擇較小的比例。現有的施工規範多為使用比例 1：5，往後的研究方向可以以該比例為重點目標。
5. 本研究使用特徵選擇的方式處理特徵工程，在料源辨識模型可以看到特徵與標籤之間的關係較為薄弱，可以使用其他特徵工程如特徵縮放、特徵提取等，可以改善模型的性能並增強特徵與標籤之間的關聯性。可能有助於提取更具代表性的特徵，使模型更好地學習和捕捉數據的模式和關係。
6. 本研究之機器學習模型建立時，並未考慮超參數調整，由於未進行超參數調整，模型可能存在一定程度的不穩定性和誤差。為了進一步提升模型的性能，建議在未來的研究中進行超參數調整，可以通過使用自動調參算法(如網格搜索或


隨機搜索)來實現，以提升機器學習模型的穩定性和準確性。將有助於改進模型的泛化能力，提高預測性能，並使其在實際應用中更加可靠和有效。





## 參考文獻

- [1] 「粗粒料中扁平、細長或扁長顆粒含量試驗法」. 2019 年.
- [2] 「ASTM D4791-19 - Standard Test Method for Flat Particles, Elongated Particles, or Flat and Elongated Particles in Coarse Aggregate」.  
<https://webstore.ansi.org/standards/astm/astmd479119> (2023, July 19).
- [3] A. Maiti, D. Chakravarty, K. Biswas and A. Halder, 「Development of a mass model in estimating weight-wise particle size distribution using digital image processing」, *International Journal of Mining Science and Technology*, vol. 27, no. 3, pp.435–443, May. 2017, doi: 10.1016/j.ijmst.2017.03.015.
- [4] M. Aydin and T. F. Kurnaz, 「An alternative method for the particle size distribution: Image processing」, *TUJE*, vol. 7, no. 2, Apr. 2023, doi: 10.31127/tuje.1053462.
- [5] X. Yang, T. Ren and L. Tan, 「Size distribution measurement of coal fragments using digital imaging processing」, *Measurement*, vol. 160, pp. 107867, Aug. 2020, doi: 10.1016/j.measurement.2020.107867.
- [6] 正敏江田, 博之新井, 三郎片山及等吉田, 「デジタル画像処理による連続粒度解析システムの開発」, *ダム工学*, vol. 24, 期 2, 頁 84–93, 2014, doi: 10.11315/jsde.24.84.
- [7] N. H. Maerz, 「Technical and Computational Aspects of the Measurement of Aggregate Shape by Digital Image Analysis」, *Journal of Computing in Civil Engineering*, vol. 18, no. 1, pp. 10–18, Jan. 2004, doi: 10.1061/(ASCE)0887-3801(2004)18:1(10).
- [8] M. Norbert and M. Lusher, 「Duplication for publication or sale is strictly prohibited without prior written permission of the Transportation Research Board Title: Measurement of Flat and Elongation of Coarse Aggregate Using Digital Image Processing」, Jul. 2023.
- [9] N. H. Maerz and W. Zhou, 「Calibration of optical digital fragmentation measuring systems」, *Fragblast*, vol. 4, no. 2, pp. 126–138, Jun. 2000, doi: 10.1076/frag.4.2.126.7450.
- [10] S. Badillo et al., 「An Introduction to Machine Learning」, *Clinical Pharmacology & Therapeutics*, vol. 107, no. 4, pp. 871–885, 2020, doi: 10.1002/cpt.1796.
- [11] J. M. Kernbach and V. E. Staartjes, 「Foundations of Machine Learning-Based Clinical Prediction Modeling: Part II—Generalization and Overfitting」,

- 
- Machine Learning in Clinical Neuroscience*, Dec. 2021, pp. 15–21. doi: 10.1007/978-3-030-85292-4\_3.
- [12] T. Hastie, J. Friedman and R. Tibshirani, *The Elements of Statistical Learning*. Springer Series in Statistics. New York, NY: Springer, 2001. doi: 10.1007/978-0-387-21606-5.
- [13] M. Ayyanar, S. Jeganathan, S. Parthasarathy, V. Jayaraman and A. R. Lakshminarayanan, 「Predicting the Cardiac Diseases using SelectKBest Method Equipped Light Gradient Boosting Machine」, *2022 6th International Conference on Trends in Electronics and Informatics (ICOEI)*, Apr. 2022, pp. 117–122. doi: 10.1109/ICOEI53556.2022.9777224.
- [14] A. Powell, D. Bates, C. V. Wyk and A. D. de Abreu, 「A cross-comparison of feature selection algorithms on multiple cyber security data-sets」.
- [15] Md. S. Zulfiker, N. Kabir, A. A. Biswas, T. Nazneen and M. S. Uddin, 「An in-depth analysis of machine learning approaches to predict depression」, *Current Research in Behavioral Sciences*, vol. 2, pp. 100044, Nov. 2021, doi: 10.1016/j.crbeha.2021.100044.
- [16] D. R. Cox, 「The Regression Analysis of Binary Sequences」, *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 20, no. 2, pp. 215–232, Jul. 1958, doi: 10.1111/j.2517-6161.1958.tb00292.x.
- [17] J. R. Quinlan, 「Induction of decision trees」, *Mach Learn*, vol. 1, no. 1, pp. 81–106, Mar. 1986, doi: 10.1007/BF00116251.
- [18] L. Breiman, 「Random Forests」, *Machine Learning*, vol. 45, no. 1, pp. 5–32, Oct. 2001, doi: 10.1023/A:1010933404324.
- [19] T. Chen and C. Guestrin, 「XGBoost: A Scalable Tree Boosting System」, *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco California USA: ACM, Aug. 2016, pp. 785–794. doi: 10.1145/2939672.2939785.
- [20] D. H. Wolpert, 「Stacked generalization」, *Neural Networks*, vol. 5, no. 2, pp. 241–259, Jan. 1992, doi: 10.1016/S0893-6080(05)80023-1.
- [21] 「試驗篩」. 1984 年.
- [22] 「粗細粒料篩析法」. 2001 年.