

東海大學資訊工程研究所

碩士論文

指導教授：蔡清欉 博士

機器學習的應用：視覺與影像廣告的互動分析

Machine Learning Application: Correlation between  
Visual and Video Advertisement

研究生：張書瑤

中華民國一百零八年六月

東海大學碩士學位論文考試審定書

東海大學資訊工程學系 研究所

研究生 張書瑤 所提之論文

機器學習的應用:視覺與影像廣告的互動分析

經本委員會審查，符合碩士學位論文標準。

學位考試委員會

召集人

蔡宏福 簽章

委員

朱正忠

林如生

指導教授

郭清枝 簽章

中華民國 108 年 6 月 17 日

## 摘要

隨著現代網路的普遍化，一般人使用網路的時間也大大的增加，社群媒體的成熟、資訊傳遞方法的改變，也使得廣告產業近幾年在網路上有著蓬勃的發展。

網路廣告則通常根據點擊率(Click through rate，縮寫 CTR)，又稱點閱率，也就是根據廣告被使用者點擊的次數與廣告顯示的次數，來衡量廣告的成效；然而實際上，當人們受到網路上的廣告吸引時，滑鼠的點擊卻並非必要的動作，「觀看者視線」是否有在廣告上才是真正的判斷依據，而現今的廣告成效計算方法並未考量到「觀看者視線」這個因素。

因此，本論文使用「眼動儀」進行眼睛追蹤，透過我們所開發的系統，直接蒐集觀看者觀看廣告的資料並整理成適合我們分析的資料集，從而使我們能以更專業且客觀的角度，來了解眼睛視線與廣告之間的互動關係，同時，我們還使用人工智慧機器學習方法，進一步的對於我們所蒐集的資料進行分析，讓機器去判斷、預測廣告的成效，使我們能知道觀看者是否對該廣告有了注意或是從中產生了興趣。

**關鍵字：**眼睛追蹤、影像廣告分析、人工智慧、機器學習、隨機森林演算法

## **Abstract**

Thanks to the popularity of modern networks, the number of Internet users has been growing high. More people watch videos on the Internet than that on televisions. Advertising industry on the Internet experience a vigorous growth in recent years. At the same time, people come to perceive different opinions and thoughts on the efficiency of video advertisements.

Online advertising is usually based on “CTR”, an abbreviation of “Click through Rate”, to measure the efficiency of the advertisements or the total length of footage that advertisements showed to the viewers and the number of times each advertisement is actually clicked by the viewers. However, these evaluation methods are generally considered only reference-worthy. There is a more promising way to confirm the advertisements are actually watched by people. It is eye tracking technology.

Our research discusses the application of eye tracking device, through the visual analysis system we developed, to collect the information from viewers' sight line, which directly allows us to recognize in which objects or details that video viewers are interested and to further acknowledge the efficiency of video advertisements. Furthermore, we also employ the technology of Artificial Intelligence - Machine Learning in the system, so the machine can analyze the final result and to further predict more information from data we have collected in the advertisements.

Keywords: Eye Tracker, Eye Tracking, Video Advertisement Analysis, Artificial Intelligence, Machine Learning, Random Forest Algorithm.

## 目錄

摘要 .....	I
Abstract .....	II
目錄 .....	III
圖目錄 .....	IV
表目錄 .....	V
第一章 前言 .....	1
第二章 文獻探討 .....	3
2.1 影像廣告的效益 .....	3
2.2 眼動追蹤技術 .....	4
2.3 影像廣告與眼動追蹤的互動關係 .....	7
2.4 Area of Interest (AOI) .....	9
2.5 人工智慧機器學習 .....	11
第三章 研究方法與設計 .....	17
3.1 視覺分析系統 .....	17
3.1.1 影像控制與 AOI 繪製 .....	18
3.1.2 視點紀錄與資料處理 .....	23
3.2 眼動儀的連結與資料蒐集的方法 .....	25
3.2.1 眼動儀的連結 .....	26
3.2.2 資料蒐集的方法 .....	27
3.3 資料前處理 .....	32
3.4 機器學習的訓練和演算法模型 .....	35
第四章 研究結果與發現 .....	37
第五章 結論 .....	40
參考文獻 .....	42
附錄一：廣告觀後問卷 .....	45

## 圖目錄

圖 2-1：Youtube 廣告收益數據分析介面 .....	4
圖 2-2：穿戴式的眼動儀.....	5
圖 2-3：非穿戴式的眼動儀.....	6
圖 2-4：Gazepoint GP3 Eye Tracker .....	7
圖 2-5：AIDA 消費者行動模式模型.....	9
圖 2-6：Area of Interest.....	10
圖 2-7：簡單決策樹 .....	14
圖 2-8：隨機森林架構圖.....	16
圖 3-1：視覺分析系統類別圖 .....	18
圖 3-2：載入影像 .....	19
圖 3-3：跳至要繪製 AOI 的影像時間點.....	19
圖 3-4：選擇 AOI 的號碼.....	20
圖 3-5：按下「Draw AOI」打開新視窗 .....	20
圖 3-6：繪製 AOI .....	21
圖 3-7：跨幀數 AOI 繪製.....	21
圖 3-8：儲存/載入 AOI .....	22
圖 3-9：loadFrame .....	23
圖 3-10：AOI 所圈選的內容.....	24
圖 3-11：Gazepoint Control 介面.....	26
圖 3-12：Gazepoint Control 校準結果.....	27
圖 3-13：問卷資料集.....	29
圖 3-14：眼動追蹤資料集示意圖 .....	30
圖 3-15：資料類別與資料數量 .....	30
圖 3-16：決策樹數量與準確度的折線圖(7 個特徵).....	35
圖 3-17：決策樹數量與準確度的折線圖(12 個特徵).....	36
圖 4-1：隨機森林演算法特徵重要度.....	38
圖 4-2：隨機森林演算法特徵重要度(注意層).....	39

## 表目錄

表 2-1：Gazepoint GP3 Eye Tracker 詳細規格 .....	6
表 2-2：混淆矩陣 .....	13
表 2-3：指標計算公式.....	13
表 3-1：廣告詳細資料.....	28
表 3-2：資料集特徵數值觀察(品牌和文字說明).....	30
表 3-3：資料集特徵數值觀察(商品和代言人) .....	31
表 3-4：資料集中的資料筆數 .....	32
表 3-5：ANOVA 計算結果 .....	34
表 4-1：7 個特徵與 12 個特徵的測試集準確度 .....	37
表 4-2：訓練集驗證結果(7 個特徵).....	37
表 4-3：測試集驗證結果(7 個特徵).....	38

## 第一章 前言

在現代日常生活中，人們使用網路的時間隨著社群媒體的成熟和科技的進步，人們上網觀看視頻的時間，甚至比在家觀看電視的時間都還要來得多(Chenyu Li et al., 2016)，再加上網路媒體的多元性，使得廣告產業近幾年，在網路上有著蓬勃的發展，同時，人們對於影像廣告成效的好壞也有了不同的見解(Ritu Lohtia et al., 2017; Sohil Jain, and Dr. Deepak Garg, 2014)。

傳統廣告成效的計算，通常是根據時段與收視率來評比，收視率越高，其成效和收益就會越高；網路廣告則通常是根據點擊率(Click through rate，縮寫 CTR)，又稱點閱率(B. Edelman et al., 2007)，根據廣告出現在網頁上的次數和廣告被「滑鼠點擊」，進入商品頁面的次數，來衡量廣告的成效；然而，廣告本身的好壞其實是無法單純透過點閱率去衡量的，因為即使廣告成功的讓觀看者對廣告產生了興趣，使觀看者看完了廣告，但觀看者若是對廣告中的「商品」沒有興趣，觀看者便不會去做「滑鼠點擊」的動作，「滑鼠點擊」沒有發生，點閱率自然就不會提升，一個能成功引起觀看者興趣的廣告，也是有可能會有低點閱率的狀況發生。為了避免這樣的狀況發生，了解客戶對於廣告中「商品」以外的刺激物的興趣程度也是相當重要的(B. Wooley, 2015; R. Alexander et al., 2000)，了解影音廣告對客戶的吸引力與廣告成效，「觀看者視線」(viewers' eye sight)因素將是關鍵的突破口。

本論文以眼動儀分析觀看者的眼球動作，抓取「觀看者視線」的資料，透過我們自己所開發的系統，把純粹的視線資料進一步整理成我們需要的資料形式，不只如此，我們還使用人工智慧機器學習方法，進一步的對於我們所蒐集的資料進行分析，幫助我們去預測影像廣告最後的成效和結果，換句話說，也就是要讓電腦能從



這些數據資料之中找到一個規律，使其能透過接收到的資料去預測出「受測者觀看廣告時，是否對廣告本身產生了興趣」、「受測者是否有看到廣告本身要觀看者注意到的內容」，從而能讓我們能以客觀的角度來了解一個眼睛視線與影像廣告之間的互動關係。

研究問題：人工智慧是否能透過人們的眼動軌跡資料，預測出他們是否對特定影像廣告產生了興趣。

## 第二章 文獻探討

在這個章節，我們將依下列排序來討論與本論文相關的文獻背景：

1. 計算影像廣告效益時會遇到的問題
2. 眼動追蹤技術和眼動儀的介紹
3. 影像廣告和眼動追蹤技術的影響關係
4. Area of Interest 的介紹
5. 機器學習和演算法模型的介紹

### 2.1 影像廣告的效益

目前網路上各大的搜尋引擎、社群媒體或是網路影片分享平台，例如：Youtube、Facebook 等等，它們的主要收入都是來自於廣告(K. C. Yang et al., 2014)，廣告大多是以平面廣告和影像廣告兩種形式呈現，但其中仍以影像形式呈現的廣告更容易吸引觀眾的目光。(Y. C. Hsieh, and K. H. Chen, 2011)

對於刊登廣告的人來說，追蹤並了解一個廣告在刊登期間所帶來的效益是很重要的，因為這樣才能夠知道自己所花的錢到底值不值得，或是需不需要再額外投資這個廣告(B. Edelman et al., 2007)，除此之外，他們還可以透過其效益結果，進一步的去審核這一則廣告的製作是否有瑕疵、廣告所要呈現的內容是否有被觀看者接收

到，諸如以上這些原因，了解一個廣告的效益對廣告製作者與刊登者來說都是十分重要的。

而目前能提供給廣告刊登者的資訊其實並不多，尤其是影像形式的廣告更是如此，除了最基本的廣告點擊率外，就是透過觀看者的資訊來進行分析，例如：廣告觸及用戶平均年齡、廣告觀看時間、廣告觀看總次數、廣告平均觀看時間長度、用戶常使用的搜尋關鍵字等等，這些數據其實是無法完全看出一個廣告效益的好壞的，因為即使我們能透過這些數據去推薦適合的廣告給用戶，但這些數據仍然無法讓我們知道觀看者當下或是觀後是否會對該廣告產生了興趣，雖然我們也可以直接在觀看後進行問卷調查或是直接詢問觀看者的觀後感想來解決這個問題，但這樣的方法所需要的時間和成本是龐大的。

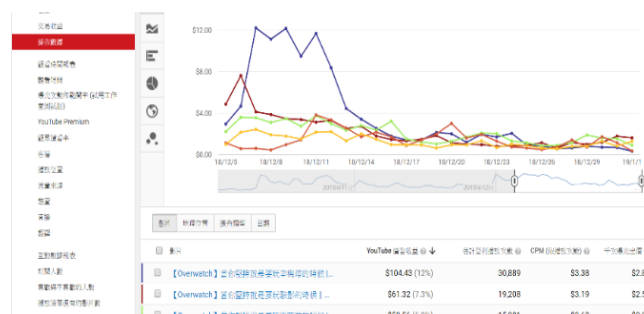


圖 2-1：Youtube 廣告收益數據分析介面

## 2.2 眼動追蹤技術

眼動追蹤技術，是一個讓我們可以了解人的眼睛視線位置以及眼睛視線是如何移動的一門技術，這樣的一門技術對於使用者體驗領域帶來了相當大的貢獻，也時常應用於改善網頁設計或是商品包裝等用途(Aga Bojko, 2013)，而「眼動儀」則是實現眼動追蹤技術的儀器。

眼動儀是一個可以記錄受測者眼睛視線移動過程的儀器，眼動儀會將紅外線照射到受測者的眼睛，然後記錄兩件事：來自視網膜的紅外光反射，這有助於找到瞳孔的中心，以及來自角膜的紅外光反射，眼動儀便是透過這兩者資料的相對位置來找出眼睛的視線(Aga Bojko, 2013)。

眼動儀蒐集的視點採樣率通常以赫茲(Hz)為單位，也就是一秒能蒐集幾筆視點資料，若只是要知道受測者在看哪裡，眼動儀的頻率其實是不用太高的，但高頻率的眼動儀比較能補足資料錯誤和缺失的問題，因為視點資料多多少少會因為光線的不足、眨眼、眼鏡反光等問題，產生錯誤和缺失，因此眼動儀的採樣率仍一般建議在 50—120Hz 之間(Aga Bojko, 2013)。

眼動儀通常分為兩大類，穿戴式和非穿戴式，穿戴式通常會是一個頭戴式的裝備，好處是受測者可以不受環境限制自由移動，通常應用於需要受測者在現實環境中活動的實驗；非穿戴式則是一個固定的偵測儀器，通常置於固定位置，例如：電腦螢幕前方，通常應用於可控性較高的模擬環境，所以基本上，我們會根據研究層面、目標和環境的不同，使用不同種類的眼動儀。



圖 2-2：穿戴式的眼動儀(<https://pupil-labs.com/>)

本論文所使用的眼動儀為 Gazepoint 所開發的 GP3 Eye Tracker，以下為詳細的設備

資料：



圖 2-3：非穿戴式的眼動儀(<https://www.gazept.com>)

表 2-1：Gazepoint GP3 Eye Tracker 詳細規格

Sample rate	60 Hz
Accuracy	0.5 – 1 degree
Operating Distance	50 cm – 80 cm
Tracking Range (Head Box)	25cm (horizontal) x 11cm (vertical)
Tracking Recovery Time	< 50 ms
Dimensions (W/H/D)	320 x 45 x 40 mm
Weight	145g
Supported OS System	Windows 7, 8.1 or 10



圖 2-4：Gazepoint GP3 Eye Tracker (<https://www.gazept.com>)

### 2.3 影像廣告與眼動追蹤的互動關係

根據許許多多文獻和書籍的研究，我們可以發現眼睛視線是可以表達人們一定程度的想法和動機的(Aga Bojko, 2013 ; Mao Wang et al., 2012; Wedel, M., and Pieters, R., 2007; Xuebai Zhang, and Shyan-Ming Yuan, 2018)，例如：如果受測者的目標是想了解商品的具體使用方法，他的視線便多會放在商品包裝的文字說明而非圖片上(Aga Bojko, 2013)；這也是為什麼在許多使用者體驗的測試和研究中，都會以眼動儀所蒐集的資料作為重要的參考依據，透過眼動追蹤，我們除了能即時性的觀察受測者當下的注意力和興趣趨向之外，同時，我們還能以一個更客觀的角度來分析資料所得出的結果。

但眼動儀在資料蒐集方面仍然是有所限制的，因為即使眼睛視點的軌跡可以讓我們了解個人的注意力和興趣外，它仍無法很完全的探查出個人的主觀深層想法或是眼睛視線與記憶的關係(Aga Bojko, 2013; Bergstrom, J. R. et al., 2012; Marco Porta et al., 2012; White SJ, 2012)。

AIDA 消費者行動模式模型(Attention、Interest、Desire、Action)，描述從消費者首次意識到產品或品牌(包含 Logo、文字名稱)，到消費者試用產品或作出購買決定

時所發生的步驟或階段，廣泛用於商業與廣告領域(Edward K Strong, 1925)；隨著網路科技發展以及消費模式的轉變，日本電通公司在 2009 年提出了新的消費者模型 AISAS(Attention、Interest、Search、Action、Share)(Satoshi Kono, 2009)，以下是具體詳細的內容：

- Attention(注意力)：消費者經由廣告的閱聽，是否注意到了商品或是品牌？通常與商品漂亮的包裝、代言人豔麗的外表或是聳動的標題有關，是讓消費者第一時間注意到廣告的主要誘因。
- Interest(興趣)：消費者注意到商品或是品牌後，是否產生興趣，從而想要更進一步的去了解其內容？像是精彩的廣告劇情或者是吸引人的畫面呈現，也是讓消費者願意繼續觀看廣告最主要的影響因素。
- Search(搜尋)：消費者產生了興趣之後，開始透過網路搜尋進一步的相關資訊。
- Action(行動)：是否實際付諸於行動？例如：進入了店家、參與了拍賣、購買了商品等等。
- Share(分享)：與他人分享內容並加以評論。

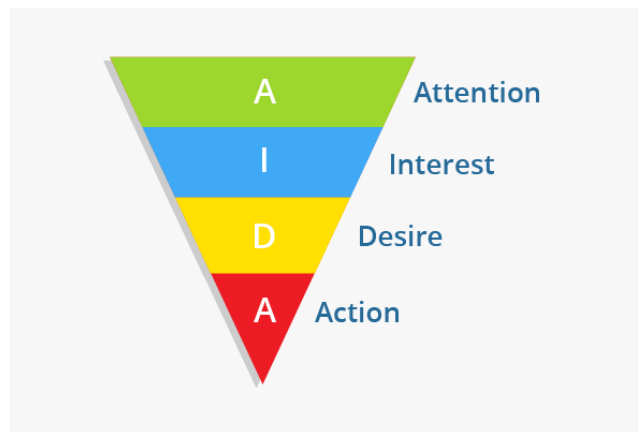


圖 2-5：AIDA 消費者行動模式模型(<https://en.ryte.com/wiki/AIDA>)

以 AIDA 模型為例，經過 Bojko (2013)的團隊的研究結果(Aga Bojko, 2013)，眼動追蹤技術只適合用於評估前兩個階段：也就是 Attention(注意)和 Interest(興趣)層級；例如當一個人的視線長時間停留在某一個物件上時，確實能表示這個人的注意力是在這個物件上的，或是說他對這個物件產生了興趣，但這並不能解釋為這個人是因為喜歡或是討厭這個物件才去注視它的，更別說要去推測一個人是否會因此而去做出更進一步的行動。眼動追蹤技術確實能解釋一些我們無法透過一般方法所得知的現象，例如：使用者為何找不到商品的使用說明？我們可以透過眼動儀得知使用者當下的眼睛視線來分析這個問題，但若是更進一步去探求說，使用者是否讀懂了使用說明或是接下來會做什麼行動，便不是眼動追蹤技術能辦到的。

## 2.4 Area of Interest (AOI)

在眼動儀紀錄眼動資料的過程中，它會把眼睛所有的移動過程全部記錄下來，但對於我們的研究來說，有許多資料並非是需要的，因此，我們將會透過一個我們事先規劃出來的區域來篩選我們的眼動資料，這就是所謂的「觀看興趣區」(Area of Interest, AOI) (Alex Poole et al., 2004)。



AOI 通常會以兩個座標點所形成的矩形方框呈現，用於標記我們要受測者觀看的物件、人物或是任何特定的刺激物上，過去多使用人工進行標記，目前也逐漸出現由電腦自動進行標記的技術出現，例如：You Only Look Once (Joseph Redmon et al., 2016)，但基本上，不管是人工或是電腦標記，其實都沒有一定正確的標記方法，因為即使是同樣的刺激物，也會根據不同的研究內容和目的而會有不同的 AOI (R. J. K. Jacob, and K. S. Karn, 2003)，例如：如果今天的研究目的為顏色，則顏色將會成會區別 AOI 之間的指標，紅色會歸為一類，藍色會歸為一類，但若今天的研究目的改為物件本身的差別，那不同的物件，則會各自歸屬於一類；因此，AOI 的區塊大小和位置都將是研究者該考慮的因素，AOI 的大小通常會根據眼動儀的距離、準確度高低去做調整，至少要 2.54cm x 2.54cm 以上(Aga Bojko, 2013)，同時考量到眼睛餘光和準確度誤差的問題，每個 AOI 周邊預留了一些空間(padding)，增加一個準確度容錯的機制。



圖 2-6：Area of Interest

通常研究者使用 AOI 所蒐集的視點資料會整理為兩種重要的形式：視點的個數 (Fixation Number)和視點的注視時間(Fixation Time)。

- 視點個數(Fixation Number)：計算視點在影像中出現的視點個數。
- 視點的注視時間(Fixation Time)：以時間作為單位，了解受測者視點滯留的時間長，人類視點的移動時常會在刺激物之間來回跳動(Aga Bojko, 2013)。

根據許多關於廣告內容影響的文獻研究(B. Wooley, 2015; Xuebai Zhang, and Shyan-Ming Yuan, 2018; Zhang, J. et al., 2009)，廣告 AOI 的選擇通常著重於三個要素：商品、商品代言人和品牌（包含 Logo、文字名稱），當然，除了這些要素外，廣告中的標題和文字說明也具有相當大的影響（Rik Pieters, and Michel Wedel, 2004）。

## 2.5 人工智慧機器學習

機器學習並非一個全新的技術，早在 1970 年代便有了第一個機器學習相關的演算法(Panos Louridas, and Christof Ebert, 2016)，機器學習近年來會變得這麼受歡迎，主要是由於電腦計算機的進步以及大數據概念的產生，才使得機器學習開始能解決更加複雜的問題並且實際應用在我們的生活之中。

機器學習基本的概念，就是藉由分析大量特定工作的資料，從其中的資料特徵找出一定的規律，從而讓機器可以去預測過去從未分析過的資料結果。依對資料類別的標記與否，機器學習又可以分為兩大類：監督式學習(Supervised Learning)和非監督式學習(Unsupervised Learning)，本論文便是採用監督式學習。

監督式學習和非監督式學習最大的差別就在於訓練資料是否有被標記過。所謂的資料標記，就是當輸入一筆資料到機器訓練好的演算法模型後，應該要產生的輸出結果，資料的標記對於監督式學習來說，是一件相當重要的事情，如果標記方法錯誤或是有瑕疵，將很有可能導致機器訓練到錯誤的資料，從而預測出有失偏頗的輸出結果。

另外，對於機器是否能正確地預測資料很重要的影響因素，便是資料量的大小和品質，資料特徵數量的不足或是資料量過小，都有可能使機器無法正確預測從未訓練過的資料，然而，預測從未訓練過的資料才是機器學習最主要的目標。

在機器學習當中要辨別一個機器的成效，我們將利用幾個指標去判定，指標有以下幾種：準確率(Accuracy)、精確率(Precision)、召回率(Recall)和 F1-Measure；這些指標將配合混淆矩陣(Confusion Matrix)和表 2-3 的公式計算而出。

表 2-2：混淆矩陣

	True condition	
Predicted condition	True Positives, TP	False Positives, FP
	False Negatives, FN	True Negatives, TN

表 2-3：指標計算公式

$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN}$
$\text{Precision} = \frac{TP}{TP + FP}$
$\text{Recall} = \frac{TP}{TP + FN}$
$\text{F1 Measure} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}$

在眾多演算法模型中，隨機森林演算法(Random Forest Algorithm)為目前最受歡迎的演算法模型之一，也是本研究訓練結果最好的演算法模型，因此，我們在這邊特別針對該演算法做介紹。

決策樹(Decision Tree)，也稱作為分類樹，是個常見用於解決分類與回歸問題的樹狀演算法模型，在每一個節點，又稱內部節點(Internal Node)，都會有一個應變數類別，也就是機器學習中資料的特徵，以該特徵作為資料的分割依據，分割後的分支又稱為葉節點(Leaf Node)，根據特徵的分割條件，資料將會被分割為兩個或多個葉節點；分割後所生成的葉節點又會成為下一層分割的內部節點，再次透過不同的

特徵進行另一層的分支產出新的葉節點，重複這樣的遞迴直到所有特徵都被作為其中一個節點的應變數類別，以此來完成資料的分類和預測(J. R. Quinlan, 1986)。

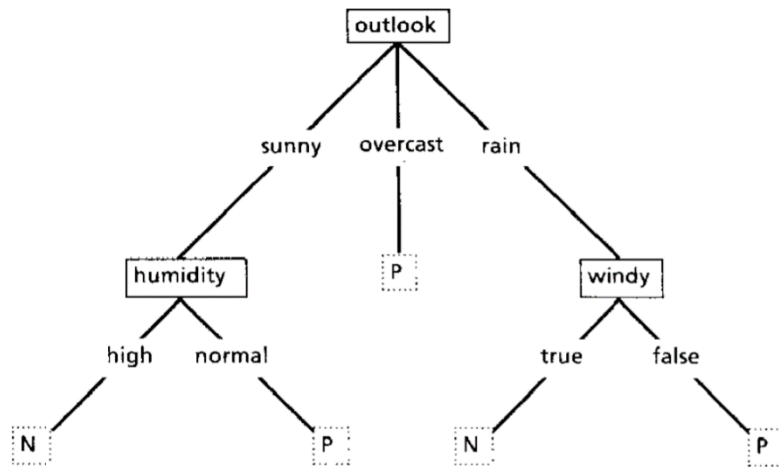


圖 2-7：簡單決策樹(J. R. Quinlan, 1986)

由於上層的特徵，將大大影響後續節點的分類，所以為了使決策樹能最佳化的做出分割，我們將透過計算節點的資訊增益(Information Gain)來選取各內部節點的資料特徵，使分類能力較好的特徵能處於決策樹的上層；資料增益越大，則代表特徵能使資料分類的越明確，而我們所獲得資訊量也就越大，公式如(2.1)：

$$IG(D_p, f) = I(D_p) - \sum_{j=1}^m \frac{N_j}{N_p} I(D_j) \quad (2.1)$$

IG：資訊增益，I：資訊量

$D_p$ ：整體的資料集， $D_j$ ：第 j 個子類別的資料集

$f$ ：應變數類別(資料特徵)，m：應變數類別總數量(資料特徵總數量)

$N_p$ ：資料總樣本數量， $N_j$ ：在第 j 個節點的樣本數量

由於我們希望獲得的資訊量要最大，因此經由分割後的資訊量要越小越好，決策樹常用的分割方法為熵(Entropy)，其公式如(2.2)：

$$I_H(t) = - \sum_{i=1}^c p(i|t) \log_2 p(i|t) \quad (2.2)$$

$I_H$ ：熵計算的資訊量

$p(i|t)$ ：位於  $t$  節點且屬於類別  $c$  的樣本機率

$c$ ：分類類別總數量， $i$ ：第幾個類別的資料總數

然而決策樹模型也是有缺點的，例如較高層次的應變數影響程度過大，使得下層的應變數類別容易受到其影響而產生有失偏頗的分割結果，訓練誤差值有時也會相對較大，除此之外，當有資料遺失的狀況時，也難以產出準確的預測結果；而這些問題將由隨機森林演算法來解決。

隨機森林(Random Forest)是一個在機器學習中十分常用的演算法之一，主要用於處理分類與回歸問題，是由大量的決策樹分類器所構成的集成學習模型(Tin Kam Ho, 1995)。

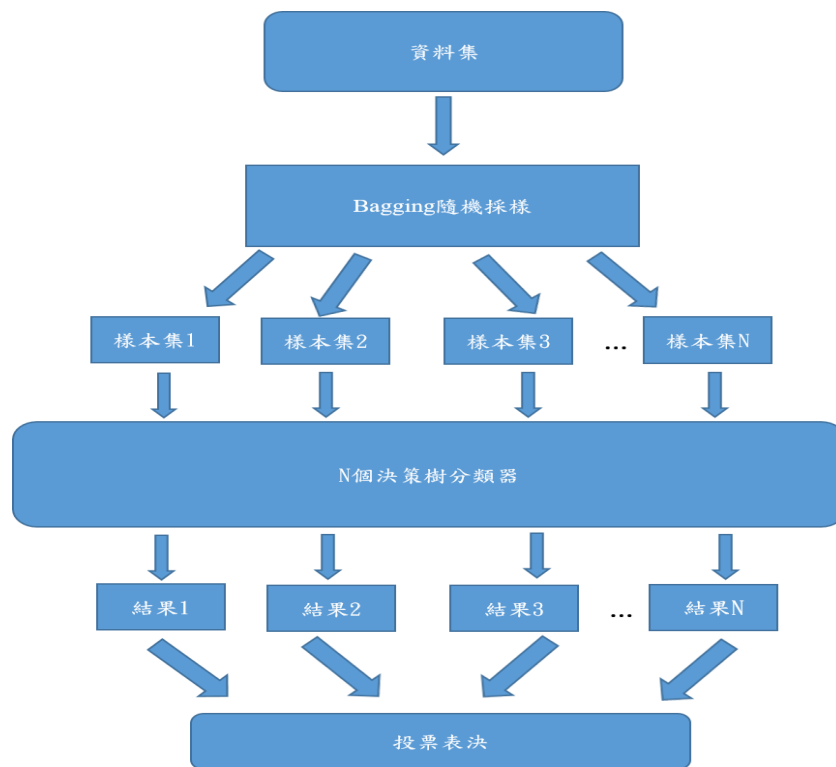


圖 2-8：隨機森林架構圖

隨機森林首先是利用一個稱為 Bagging (L. Breiman, 2001)的演算法進行樣本的隨機抽樣，從資料集中隨機抽出數個樣本，以抽出放回的方式重複進行取樣，形成一個個訓練樣本集，並把未抽到的樣本作為後續評估誤差的測試樣本集。獲得訓練樣本集後，我們設樣本的特徵維度為  $M$ ，指定一個常數  $m$ ，其中  $m$  應遠小於  $M$ ，並隨機地從  $M$  個特徵中選取  $m$  個特徵作為建立一棵決策樹的應變數類別；我們以同樣的方式，重複建立多種不同的決策樹，並透過平均各決策樹的預測結果作為整體隨機森林的預測結果，隨機森林的優點除了不易發生過度擬合(Overfitting)外，即便資料集有很大的數據缺失，也能維持很高的準確度，在處理非線性資料時的表現也十分優異，除此之外，還能藉由各棵決策樹的資訊增益計算結果，進一步對各個特徵的重要性(Importance)進行排名。

## 第三章 研究方法與設計

在本章節，我們將會介紹我們研究問題的具體過程，會依下列的排序進行說明：

1. 詳細介紹我們所開發的視覺分析系統
2. 如何利用眼動儀蒐集資料和標記資料的方法與過程
3. 資料的前處理
4. 如何透過機器學習演算法進行訓練和預測

### 3.1 視覺分析系統

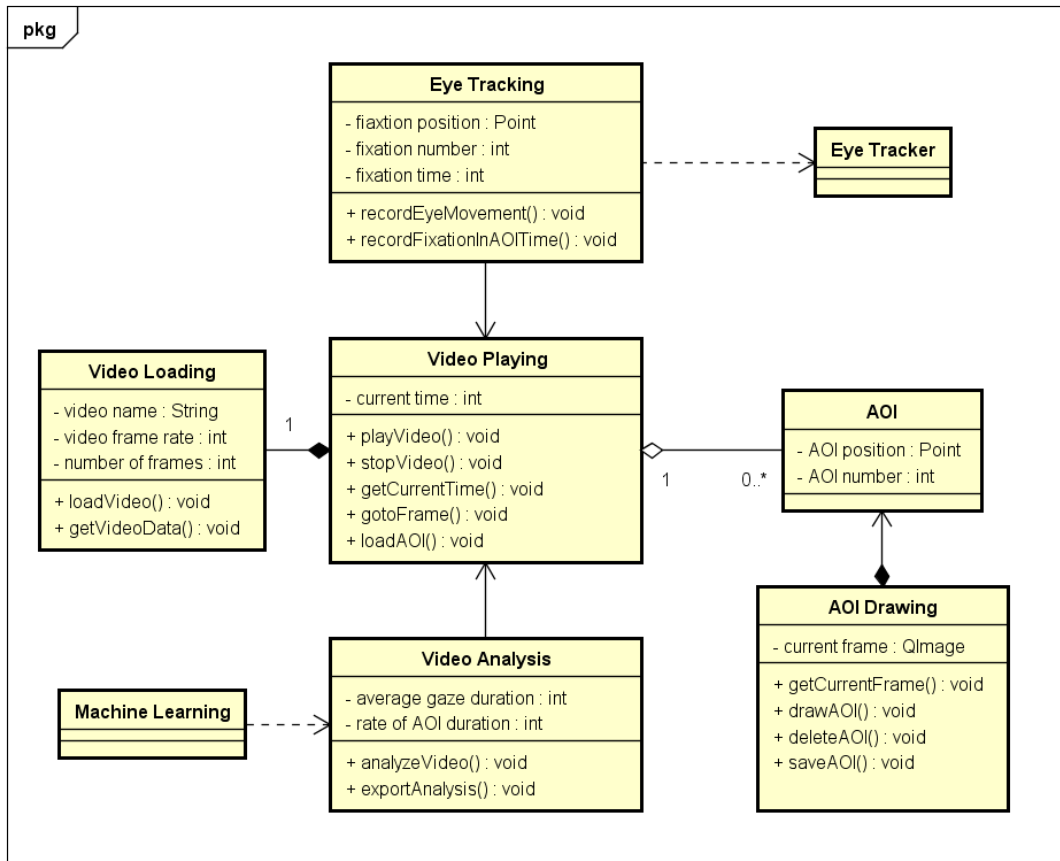
本系統我們使用 Qt 和 OpenCV 開發。Qt 是一個跨平台、開放原始碼的 C++ 應用程式開發軟體，它是 Digia 公司的產品，廣泛用於開發 GUI (Graphical User Interface)；OpenCV 則是一個跨平台的 C++ 電腦視覺庫，由 Intel 公司發起並參與開發，廣泛用於圖像處理、電腦視覺以及圖形辨識等領域。

本系統主要分成兩個部分：

- 影像控制與 AOI 繪製：選擇要觀看或要繪製的廣告，並進一步的控制廣告的播放/停止；抓取影像目前處於時間點(以幀為單位)的圖像，然後在該圖像上設定 AOI 的 x, y 座標位置。
- 視點紀錄與資料處理：播放廣告並同時記錄以每幀為單位的眼睛視點座標和座標處於各 AOI 的時間長度；彙整系統的紀錄資料並進行初步的資料處理。



本系統架構類別圖如下：



powered by Astah

圖 3-1：視覺分析系統類別圖

我們將透過這個系統，並配合眼動儀的使用進行資料蒐集，完成蒐集後，進行初步的資料處理，最後把結果匯出給機器學習訓練。

### 3.1.1 影像控制與 AOI 繪製

在使用眼動儀之前，我們要先選擇我們要受測者觀看的廣告並繪製屬於該廣告的 AOI。本系統提供影像的載入、影片的播放/暫停、影像時間軸拖拉(以秒為單位)

以及跳至特定影像幀數位置等功能，方便使用者進行操作；AOI 的繪製將以影像的幀數為單位，具體的操作步驟如下：

1. 按下「Load Video」載入影像，支援.avi、.mpg、.mp4 影像格式。

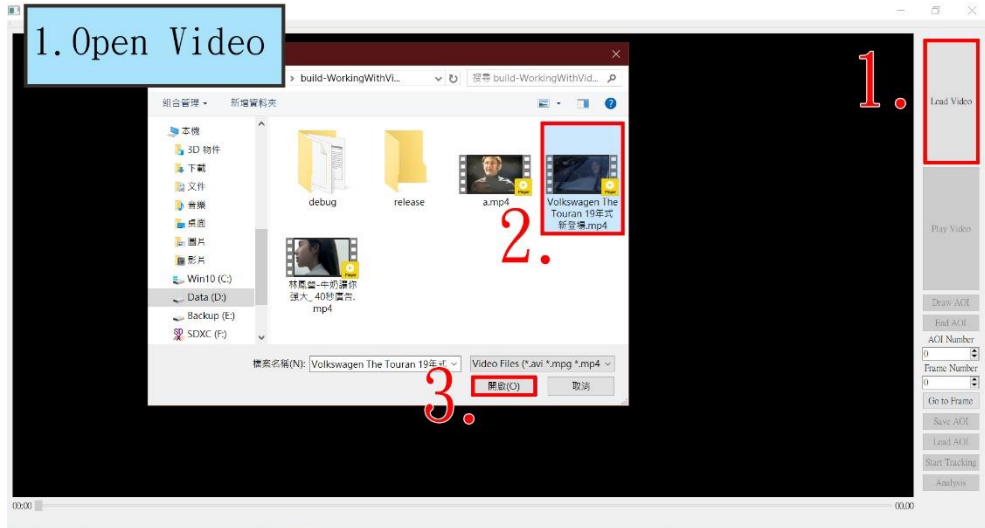


圖 3-2：載入影像

2. 跳至要繪製 AOI 的影像時間點：可以透過拖拉時間軸選擇對應秒數，或是在「Frame Number」處，直接輸入指定幀數，再按下「Go to Frame」，跳至指定影像幀數的位置。



圖 3-3：跳至要繪製 AOI 的影像時間點

3. 選擇 AOI 的號碼：處於不同時間點，但擁有相同號碼的 AOI，都將被識別為同一個 AOI，其蒐集的數據都會儲存於相同的位置。



圖 3-4：選擇 AOI 的號碼

4. 按下「Draw AOI」打開新視窗。

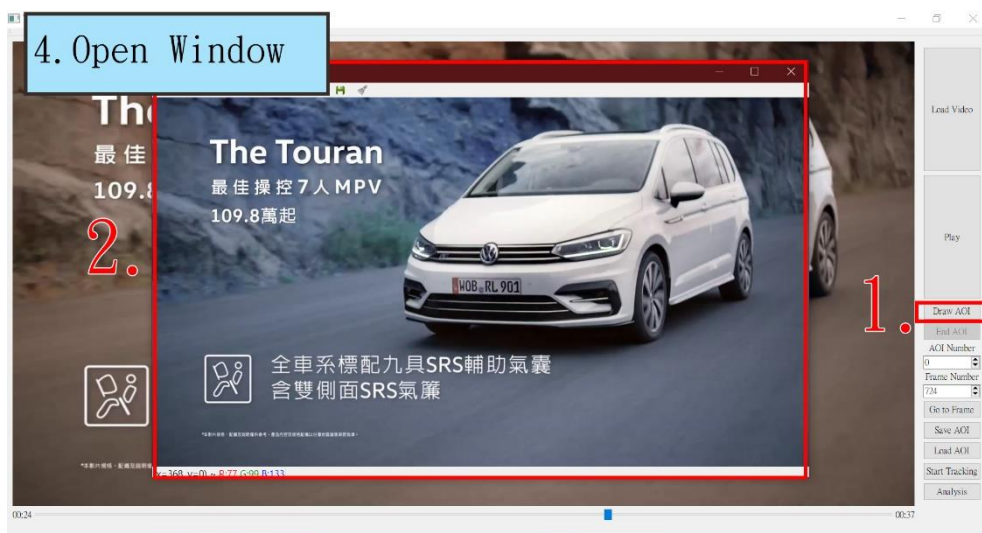


圖 3-5：按下「Draw AOI」打開新視窗

5. 繪製 AOI，按住左鍵決定 AOI 方框的左上角座標，放開左鍵決定 AOI 方框的右下角座標。

- 繪製錯誤，可以直接重複步驟 5. 進行重新繪製

- 連點左鍵兩下在指定的 AOI，可以直接刪除該 AOI 退出視窗，完成在該幀數處，繪製對應號碼的 AOI。



圖 3-6：繪製 AOI

6. 如果想以同樣的座標進行跨幀數的 AOI 繪製，則要先跳至指定幀數位置以選擇「AOI 的結束位置」，之後按下「End AOI」，這樣的動作可使「上一次繪製 AOI 的幀數位置」到「AOI 的結束位置」之間，都一併繪製出同樣座標的 AOI。

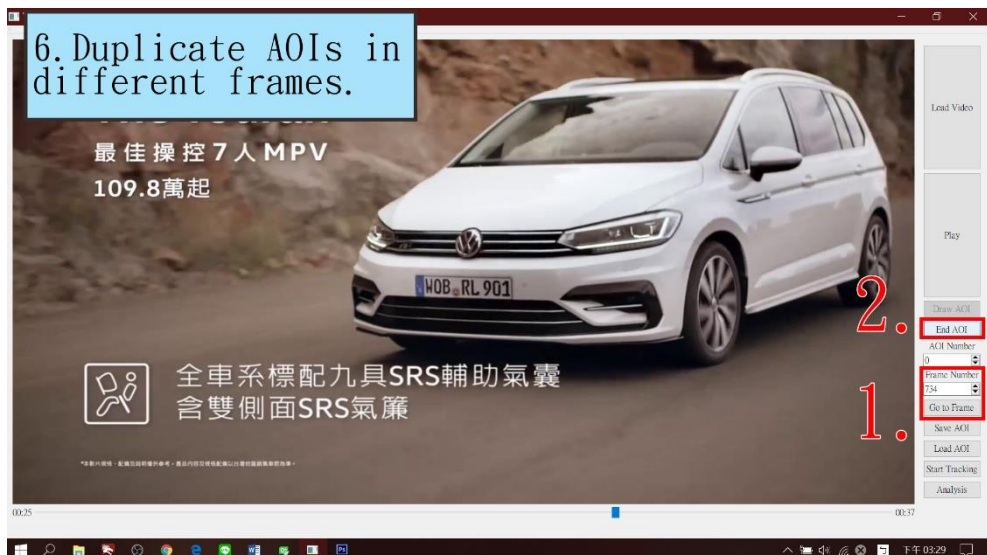


圖 3-7：跨幀數 AOI 繪製

7. 完成這個影像廣告的 AOI 繪製後，可以使用「Save AOI」功能進行 AOI 的儲存，每一幀的 AOI 座標和號碼都會被整理好並輸出成一個名為「loadFrame」的 Excel 檔，方便以後可以直接使用「Load AOI」功能載入 AOI 資料，免去需要再次繪製同一個影像廣告 AOI 的麻煩。

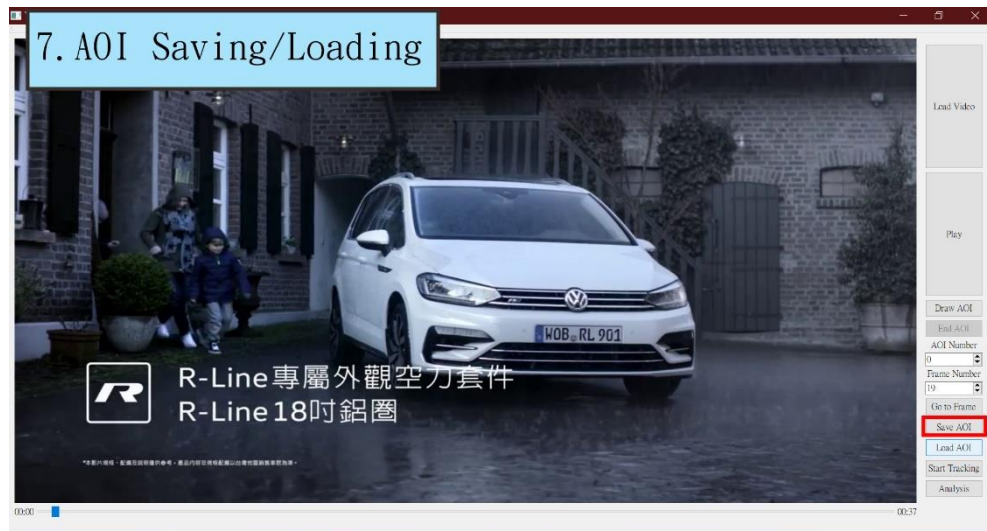


圖 3-8：儲存/載入 AOI

「loadFrame」的檔案格式如圖 3-9，「Frame Number」代表 AOI 的時間軸(以幀為單位)；「AOI Number」代表著 AOI 的類別，以此圖為例，「0」代表商標、「1」代表商品、「2」代表文字說明、「3」代表代言人；其餘的四個欄位代表 AOI 的座標位置，分別為左上方的 x 座標、左上方的 y 座標、右下方的 x 座標、右下方的 y 座標。

1	Frame Number	AOI Number	AOI_Left_x	AOI_Left_y	AOI_Right_x	AOI_Right_y
2	0	0	764	376	815	430
3	1	0	764	376	815	430
4	1	2	418	174	1006	582
5	1	3	97	180	265	510
6	2	0	764	376	815	430
7	2	2	418	174	1006	582
8	2	3	97	180	265	510
9	3	2	418	174	1006	582
10	3	3	97	180	265	510
11	4	0	764	376	815	430
12	4	1	68	480	719	621
13	4	2	418	174	1006	582
14	4	3	97	180	265	510
15	5	0	755	376	818	430
16	5	1	68	480	719	621
17	5	2	418	174	1006	582
18	5	3	97	180	265	510
19	6	0	755	376	818	430
20	6	1	68	480	719	621

圖 3-9：loadFrame

### 3.1.2 視點紀錄與資料處理

在這個小節，我們要來了解我們要用 AOI 圈選的刺激物，根據許多的文獻研究 (B. Wooley, 2015; Xuebai Zhang, and Shyan-Ming Yuan, 2018; Zhang, J., Wedel, M., and Pieters, R., 2009)，廣告 AOI 的選擇通常著重於三種刺激物：商品、商品代言人和品

牌，當然，除了這些刺激物外，廣告中的標題和文字說明也具有相當大的影響 (Rik Pieters, and Michel Wedel, 2004)。

- 商品(紅色)：影像廣告中最主要被銷售的內容。
- 商品代言人(黃色)：影像廣告中推銷商品的人。
- 品牌(綠色)：任何與品牌本身有關的圖像或是文字。
- 文字說明(紫色)：影像廣告中的廣告詞、產品說明或是標題等非品牌相關的文字。



圖 3-10：AOI 所圈選的內容

通常研究者使用 AOI 所蒐集的視點資料，會被整理成兩種重要的形式：視點個數(Fixation Number)和視點時間長(Fixation Time)(Aga Bojko, 2013; Marco Porta, Alice Ravarelli, and Francesco Spaghi, 2012)。

而為了更方便我們使用機器學習訓練資料，我們將資料更進一步的整理成三種特徵：Transformed Fixation Number(TFN)、Transformed Fixation Time(TFT)、Average

Gaze Duration(AGD)(Xuebai Zhang, and Shyan-Ming Yuan, 2018)，使資料不會因為影像廣告時間長度的不同、廣告內容曝光時間長度的不同，而有太大的差異。

- Transformed Fixation Number(TFN)：視點處於特定 AOI 內的個數(Fixation Number)與該 AOI 出現的時間長之間的比例；數值越高，就代表視點在該 AOI 內的個數越多。
- Transformed Fixation Time(TFT)：視點處於 AOI 內的總時間長(Fixation Time)與該 AOI 出現的時間長之間的比例；數值越高，就代表視點停留於該 AOI 的持續時間越久。
- Average Gaze Duration(AGD)：TFN 與 TFT 之間的比例；是一個十分普遍用在眼動追蹤研究中的特徵，時常用於表示視點的強度高低(Marco Porta, Alice Ravarelli, and Francesco Spaghi, 2012)。

所以，在之後的資料蒐集過程中，我們將會從廣告的四種刺激物(商品、代言人、品牌和文字說明)，分別蒐集上述所整理的三種特徵資料，換句話說，我們的資料集將總共會有 12 個特徵。

## 3.2 眼動儀的連結與資料蒐集的方法

經由視覺分析系統完成影像廣告的 AOI 繪製之後，我們將開始利用眼動儀抓取視點的位置，透過由 Gazepoint 所開發的眼動儀程式—Gazepoint Control，我們眼睛的視點將可以取代滑鼠的操作，使鼠標的位置等同於視點的位置，從而使電腦可以



間接獲得視點的座標位置，此章節將會詳細介紹我們連結眼動儀的具體步驟；受測者的選擇和資料的蒐集方法，也是決定我們資料品質好壞的重要因素，本章節也將詳細介紹其中的過程和構想。

### 3.2.1 眼動儀的連結

首先，我們得先安裝由 Gazepoint 所開發的眼動儀驅動程式 Gazepoint Control，完成安裝後開啟程式，具體操作如下：

1. 透過 USB 2.0 接頭使眼動儀連結上電腦。
2. 眼動儀開始抓取受測者的臉部，並抓取眼睛和瞳孔的位置，影像準確度與光線亮度和反光有關，配戴金屬框的眼鏡會有些許影響。
3. 視點校準：點擊按鈕「Calibrate」進入視點校準視窗，畫面將會出現 5 或 9 個校準點，受測者需要持續注視校準點並跟隨校準點移動視點。

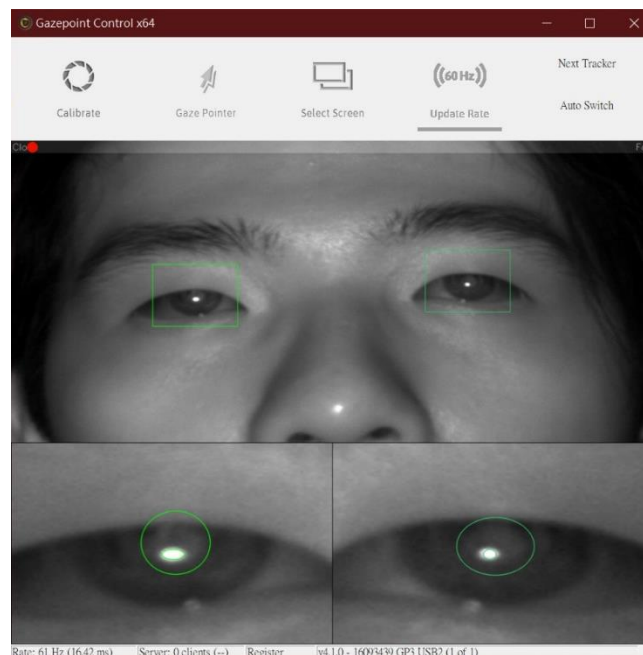


圖 3-11：Gazepoint Control 介面

4. 完成視點校準後，可稍加確認視點的精確度和靈敏度，如需要重新校準可按下快捷鍵 C，返回到步驟 3。
5. 取代滑鼠操作：點擊按鈕「Gaze Pointer」，眼動儀將取代滑鼠進行操作，重複



圖 3-12：Gazepoint Control 校準結果

按下按鈕或是用手遮住眼動儀即可返回滑鼠操作。

完成眼動儀的連結之後，我們便可以透過眼睛的視點來操作我們電腦的鼠標，我們看到哪裡，鼠標就會移動到哪裡，所以，當鼠標進入到其中一個我們所繪製好的 AOI 中時，就代表著我們的視點已經進入到了該 AOI 中，視覺分析系統便可以開始根據鼠標的座標和停留時間等等數據進行視點的蒐集和分析。

### 3.2.2 資料蒐集的方法

在 2.3 中，我們有提到眼動追蹤技術的限制以及它無法解釋的問題，包含受測者的動作、主觀想法和記憶等等，因此，為了盡可能以客觀的角度去分析我們的資料並方便機器學習的訓練，我們把資料中任何可控的變數，盡量控制在一個較小的範圍內，例如：受測者的年齡、廣告的種類、標記的方式等等；我們這邊分為兩個部分：廣告的選擇以及受測者的篩選。

首先，我們將會選擇四種廣告不同商品、不同品牌的影像廣告進行資料蒐集，以此來避免因為個人喜好而影響分析結果，廣告的時間長度皆控制在 30 到 40 秒間，同時廣告內容的語言也都控制在受測者可閱讀的語言，方便受測者觀看廣告，最後，我們也屏除了廣告中的聲音，讓資料的蒐集只專注在畫面所呈現的內容。

表 3-1：廣告詳細資料

廣告名稱	品牌	商品	片長
Volkswagen The Touran 19 年式 新登場 ( <a href="https://www.youtube.com/watch?v=WUZn9Xl4bUM">https://www.youtube.com/watch?v=WUZn9Xl4bUM</a> )	Volkswagen	汽車	36s
旺仔小饅頭：寶貝篇 ( <a href="https://www.youtube.com/watch?v=dBvJAXsGIbc">https://www.youtube.com/watch?v=dBvJAXsGIbc</a> )	旺仔	餅乾	30s
娛樂不插電 歡樂不受限 Flicks 行動電影院 ( <a href="https://www.youtube.com/watch?v=ZguNxvcIFrU">https://www.youtube.com/watch?v=ZguNxvcIFrU</a> )	Dashbon	投影機	30s
HTC RE - Miss Ko Eye 自拍 愛分享 ( <a href="https://www.youtube.com/watch?v=f2Wr-da_1eE">https://www.youtube.com/watch?v=f2Wr-da_1eE</a> )	HTC	自拍棒	30s

接下來關於受測者的篩選，我們將從 59 名不同年齡、不同性別的人蒐集資料，年齡控制在 18 歲~40 歲，男 31 人、女 28 人，每位受測者將會在不被告知廣告內容的情況下，以不同的順序去觀看這四種影像廣告，看完廣告後，我們將會透過受測者的問卷填寫結果，去標記我們資料的類別。

為了要預測廣告最後的成效高低，也就是要知道「受測者觀看廣告時，是否對廣告本身產生了興趣」、「受測者是否有看到廣告本身要觀看者注意到的內容」，我們將以在本論文 2.4 中所提到的 AIDA 消費者模型為基礎，並以 Attention(注意)層和

Interest(興趣)層作為我們要分析預測的結果，我們不以更深層的內容做為我們資料標記和製作問卷的參考，是因為眼動儀所蒐集的資料僅限用於分析在注意層和興趣層中的內容(Aga Bojko, 2013)，同理，機器學習也無法透過這些資料預測在注意層和興趣層以外的結果。

本論文將參考 Wei 和 Lu (Pei-Shan Wei, and Hsi-Peng Lu, 2013)所設計的問題進行資料蒐集和標記，但本論文的問卷(如附錄一)內容將不會涉及到 AIDA 或是 AISAS 模型中較深層的部分，而是專注於在注意層和興趣層，且為了方便之後的標記與分類，我們簡化了問卷的填寫和標記方法，並對於問題的內容進行了改良，讓問題能更明確的被填寫者回答。

我們的問卷將採用是非題，「同意」計 1 分，「不同意」計 0 分；為了方便後續的標記分類，當每一層的總分過半(大於等於 2 分)，便表示受測者滿足該層的條件，同時滿足了注意層和興趣層，則標記為「2」，沒有滿足興趣層的資料則都標記為「1」，分成兩種標記類別，但特別要注意的是 AIDA 模型是有順序性質的，只有在滿足注意層的情況下，才能有機會滿足興趣層，因此當注意層未被滿足時，興趣層亦無法獲得滿足。

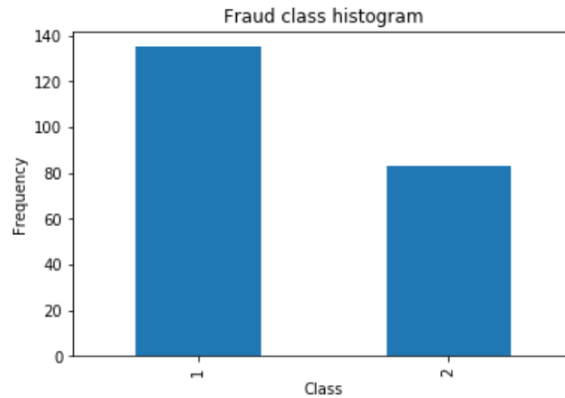
Label	Q1	Q2	Q3	Q4	Q5	Q6	Sex
1	是	是	否	否	否	是	女
2	是	是	是	是	否	是	男
1	否	否	是	是	否	否	女
1	是	否	是	否	否	是	男
1	否	是	否	是	否	否	男
1	否	是	否	否	否	否	女
1	是	否	否	否	否	是	男

圖 3-13：問卷資料集

本論文抽樣 59 位受測者，性別男性為 31 人，女性為 28 人，每一受測者都觀看四種廣告，資料集中總共有 236 筆資料；結合在章節 3.1.2 中所提到的特徵內容，資料集將如圖 3-14 所示，包含有 12 個特徵和 1 個標籤(label)，資料集中標籤為類別「1」資料 135 筆，而標籤為類別「2」資料有 83 筆。

Brand_T FT	Brand_T FN	Brand_A GD	Text_TF T	Text_TF N	Text_AG D	Product_ TFT	Product_ TFN	Product_ AGD	Endorser _TFT	Endorser _TFN	Endorser _AGD	Label
0.05	1.4	0.03	0.23	4.79	0.05	0.13	2.93	0.04	0.86	19.75	0.04	1
0.23	6.15	0.04	0.06	1.14	0.06	0.84	18.43	0.05	0.78	12.99	0.06	1
0.24	5.66	0.04	0	0.03	0.03	0.04	0.7	0.06	0.5	12.77	0.04	1
0.1	2.64	0.04	0.08	1.83	0.04	0.23	4.41	0.05	0.72	17.11	0.04	1
0.87	2.82	0.31	0.32	7.72	0.04	0.91	22.12	0.04	0.57	13.87	0.04	2

圖 3-14：眼動追蹤資料集示意圖



```

1    135
2     83
Name: Label, dtype: int64

```

圖 3-15：資料類別與資料數量

從表 3-2 和表 3-3 可以得知，特徵 TFN(視點個數)的平均值和最大值比起其他特徵都來的高，代表著眼睛視點停留於 AOI 的個數遠遠大於 AOI 出現的時間；但這樣的狀況，可能會帶來特徵重要性的偏見，從而影響機器學習的訓練。

表 3-2：資料集特徵數值觀察(品牌和文字說明)

	Brand_TFT	Brand_TFN	Brand_AGD	Text_TFT	Text_TFN	Text_AGD
<b>Average</b>	0.27	5.10	0.07	0.22	4.57	0.07

<b>Max</b>	1.70	26.67	0.85	0.69	16.34	1.36
<b>Min</b>	0.00	0.00	0.00	0.00	0.00	0.00

表 3-3：資料集特徵數值觀察(商品和代言人)

	Product_TFT	Product_TFN	Product_AGD	Endorser_TFT	Endorser_TFN	Endorser_AGD
<b>Average</b>	0.50	10.96	0.06	0.70	15.63	0.05
<b>Max</b>	1.00	26.27	1.16	1.17	24.40	0.98
<b>Min</b>	0.00	0.00	0.00	0.30	0.69	0.04

而從 TFT(注視時間)的平均值可以得知，「代言人」在廣告中有約 70%的時間是被觀看者注視的，而「品牌」和「文字說明」則相對容易被觀看者忽略，被注視到的時間比例分別為 27%和 22%。

不同的 AGD(視點強度)並沒有明顯的差異，可見在大部分的時間，人們觀看廣告時的視點強度並不會因為刺激物的不同而有明顯的差異；但在注視「文字說明」和「商品」時偶爾會出現明顯比較高的視點強度，這代表著「文字說明」和「商品」的出現可能會有一段時間提升觀看者視點的強度，從而引起注意力或興趣。幾乎所有的特徵之最小值都出現了「空值」的狀況，資料前處理的部分應針對這個問題去做處理。

### 3.3 資料前處理

在把收集好的資料進行機器學習之前，我們必須要先進行資料的前處理(Data Preprocessing)，以確保資料集的品質，使機器學習能有更完備的訓練和更精確的預測結果。

我們也必須用肉眼進行資料整理的動作，刪除一些可能有標記錯誤或是異常的資料，例如：受測者問卷中填寫說「有」看到目標，眼動儀的資料中卻顯示「沒有」。我們也不能光是依靠問卷的回答來刪除資料，對於目標的熟悉度、眼睛餘光的考量、眼動儀誤差等等因素，其實都有可能影響到問卷和眼動資料的結果，因此，為了盡量保留原始資料，只刪除眼動資料為 0 (完全沒有看到目標)、問卷中卻填寫「有」看到目標的異常資料；經過檢查後總共刪除了 18 筆異常資料，仍保留了 218 筆資料。

表 3-4：資料集中的資料筆數

性別	人數	資料數	刪減數	資料現有數
男	31 人	124	10	114
女	28 人	112	8	104

接下來我們進行的是特徵選取(Feature Selection)，也就是選取適合機器訓練的特徵，並屏除不適合訓練的特徵和雜訊，若給予訓練的特徵與模型無關或不重要是會影響訓練結果的。

本論文使用的是變異數分析(Analysis of Variance, ANOVA)進行特徵選取，該檢定方法透過計算樣本的「組間均方和」與「組內均方和」的相對大小來檢定不同組樣本之間的母體平均數差異，從而估計出 F 值(f-value)和 P 值(p-value)，通常 F 值越大，P 值就越小，樣本之間的差異也就越顯著；透過 ANOVA 可以使我們估計出特徵與要預測目標之間的顯著性，從而使我們能找出最適合用於訓練的特徵；我們保留 P 值為 0.05 (Tukur Dahiru, 2008)以下、顯著性較大的特徵，從原本 12 個特徵中，選取 7 個特徵給予模型做訓練，保留的特徵分別為 Brand\_TFT、Brand\_TFN、Brand\_AGD、Text\_TFT、Text\_TFN、Product\_TFT 和 Product\_TFN。



表 3-5：ANOVA 計算結果

	<b>F value</b>	<b>P value</b>
<b>Brand_TFT</b>	78.859	<0.05
<b>Brand_TFN</b>	46.848	<0.05
<b>Brand_AGD</b>	11.394	<0.05
<b>Text_TFT</b>	8.2643	<0.05
<b>Text_TFN</b>	10.062	<0.05
<b>Text_AGD</b>	0.634	0.426
<b>Product_TFT</b>	95.707	<0.05
<b>Product_TFN</b>	68.467	<0.05
<b>Product_AGD</b>	0.606	0.437
<b>Endorser_TFT</b>	0.613	0.434
<b>Endorser_TFN</b>	0.144	0.704
<b>Endorser_AGD</b>	0.014	0.902

但為了觀察各個特徵的互動關係和影響程度，在之後的機器學習訓練過程中，我們仍然會分別以未經特徵選取過的 12 個特徵和經特徵選取的 7 個特徵，各別做訓練，以此來互相做比較。

最後，我們對資料進一步做特徵歸一化(Normalization)進行特徵縮放的動作，使資料特徵按比例縮放，把特徵數據映射在 0~1 之間，除了能讓資料更快收斂外，也減少模型偏頗於部分特徵的可能性。

### 3.4 機器學習的訓練和演算法模型

本研究分別使用了於章節 2-5 所提到的隨機森林演算法(Random Forest Algorithm)、支援向量機(Support Vector Machine)以及羅吉斯回歸(Logistic Regression)三種不同的演算法來訓練我們的模型，資料集則以 8：2 的比例分作訓練集和測試集，訓練集提供給模型進行訓練，測試集則會經由交叉驗證(Cross-validation)來驗證我們訓練好的模型是否準確。

我們先以特徵選取的 7 種特徵和隨機森林演算法來建構模型。我們以 2、4、8、16、32、64、128、256 倍數方式，迭代觀察訓練後的準確度來決定我們模型中決策樹的數量；我們發現在使用約 64 棵決策樹時，準確度開始收斂，並在約 256 棵決策樹時擁有最高的準確度。同時，我們繼續使用 300、400 棵決策樹作訓練，根據圖 3-16 的結果可以發現，準確度在 256~300 棵決策樹時擁有最穩定且優異的表現，因此我們以 256 棵決策樹作訓練。

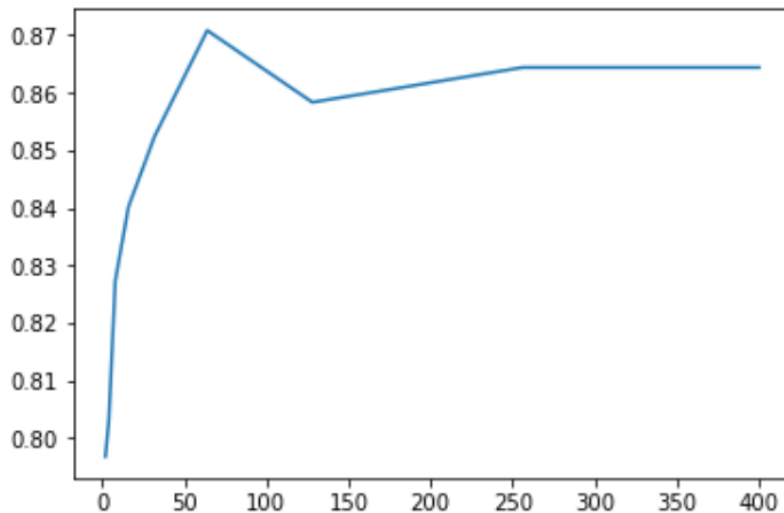


圖 3-16：決策樹數量與準確度的折線圖(7 個特徵)

接下來，我們以 12 種特徵來建構隨機森林演算法模型。以 2、4、8、16、32、64、128、256 倍數方式，迭代觀察訓練後的準確度來決定我們決策樹的數量；我們發現了在使用 256~400 棵決策樹時擁有最穩定且優異的表現，因此最後同樣以 256 棵決策樹作訓練。

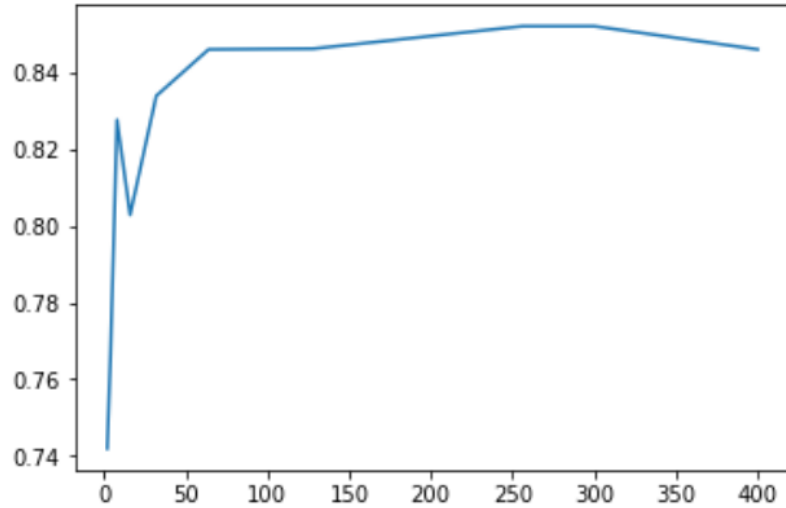


圖 3-17：決策樹數量與準確度的折線圖(12 個特徵)

支援向量機 SVM 和羅吉斯回歸我們同樣以特徵選取的 7 種特徵和原本的 12 種特徵分別做訓練。比較值得一提的是在支援向量機 SVM 的參數調整過程中，核函數 (kernel) 的選擇比較有明顯的差異，除了線性(linear)核函數的訓練效果比較好外，其他核函數的訓練結果就相對差一些。

## 第四章 研究結果與發現

經由交叉驗證(Cross-validation)來驗證準確度後，我們發現當隨機森林演算法和支援向量機使用特徵選取的 7 個特徵(Brand\_TFT、Brand\_TFN、Brand\_AGD、Text\_TFT、Text\_TFN、Product\_TFT、Product\_TFN) 做訓練時，其訓練結果比使用原本的 12 個特徵的訓練結果有比較高的準確度；但對羅吉斯演算法來說，使用未經特徵選取的 12 個特徵來訓練，卻有比較好的準確度。

我們推測隨機森林演算法和支援向量機之所以在使用 12 個特徵時會有比較差的準確度，是因為這兩種演算法本身就很擅長訓練擁有多維度特徵的資料，對於特徵的敏感度也相對較大，同理，它們受到不良特徵的影響也相對較大，再加上樣本數量少的關係，反而導致訓練結果不一定比羅吉斯演算法好，當然，此推論還需進一步做釐清。

表 4-1：7 個特徵與 12 個特徵的測試集準確度

演算法	7 個特徵	12 個特徵
Random Forest	80.21%	72.72%
Support Vector Machine	76.54%	74.54%
Logistic Regression	72.96%	74.39%

表 4-2：訓練集驗證結果(7 個特徵)

演算法	Accuracy	Precision	Recall	F1 Score
Random Forest	83.42%	83.68%	82.08%	82.44%
Support Vector Machine	80.98%	80.55%	79.52%	79.81%
Logistic Regression	82.17%	83.45%	79.74 %	80.55%

表 4-3：測試集驗證結果(7 個特徵)

演算法	Accuracy	Precision	Recall	F1 Score
<b>Random Forest</b>	80.21%	80.37%	76.72%	77.41%
<b>Support Vector Machine</b>	76.54%	77.68%	69.64%	70.75%
<b>Logistic Regression</b>	72.96%	72.44%	63.8%	64.63%

透過變異數分析和隨機森林演算法(如圖 4-1)我們可以發現，在各個刺激物中，品牌(Brand)和商品(Product)影響觀看用戶對於影像廣告的興趣程度是較為顯著的，而視點強度(AGD)相關特徵的重要度則普遍偏低。本研究所開發的視覺分析系統對於「視點停留」的構想和設定還需要進一步做改善。

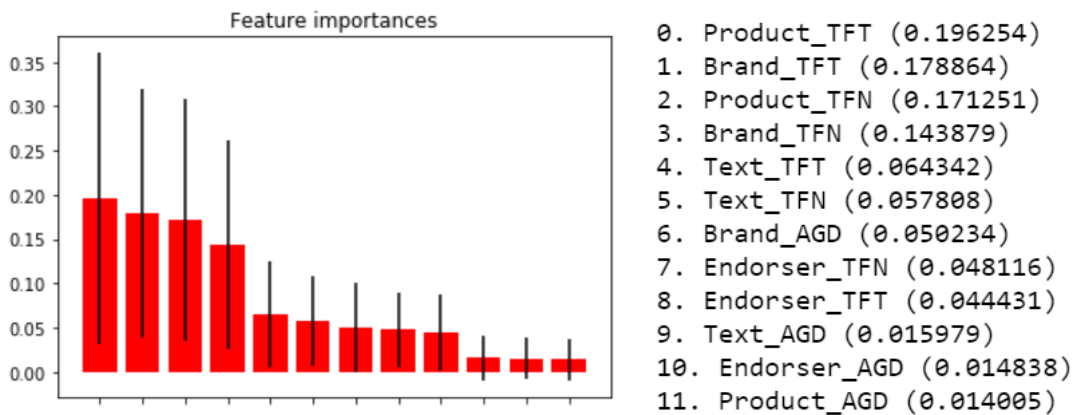


圖 4-1：隨機森林演算法特徵重要度

根據以上的發現，我們推測影像廣告中的代言人(Endorser)或許比較不會影響人們對於影像廣告的興趣，代言人的影響可能更著重於在吸引人們對於廣告的注意力，例如：人們在路上看到了廣告，會因為代言人的影響，而選擇駐足觀看廣告。

為了驗證代言人與注意層的關係，我們改變標記資料的方法，不滿足注意層的資料標記為「0」、滿足注意層或是同時滿足注意層與興趣層的資料則標記為「1」，以同樣的研究方法進行資料前處理，並使用目前準確度最高的隨機森林演算法進行訓練；經過訓練後，其測試集準確度為 90.84%，特徵重要程度排行如圖 4-2。

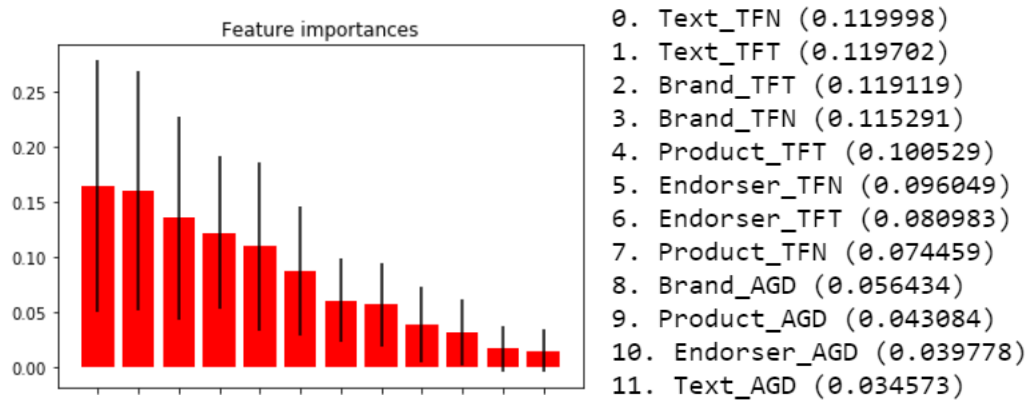


圖 4-2：隨機森林演算法特徵重要度(注意層)

從圖 4-2 可以得知，最能影響受測者與注意層之間關係的特徵為文字說明(Text)，代言人(Endorser)的影響仍然十分有限，所以我們仍然無法斷定代言人究竟影響了什麼；我們推測可能是因為受測者皆是由我們主動邀請而來，且受測者大多全程觀看了我們所提供的影像廣告，所以導致我們的研究方法可能無法驗證注意層問題。

## 第五章 結論

經由我們的研究，人工智慧確實可以透過眼動軌跡的資料，預測出他們是否對特定影像廣告產生了興趣，準確度達 80.21%：我們使用機器學習，以「隨機森林演算法」、「支援向量機」和「羅吉斯回歸」等演算法訓練模型，透過分析人們的眼睛視點在影像廣告中的品牌(Brand)、商品(Product)、代言人(Endorser)和文字說明(Text)等刺激物上的「視點個數」和「視點注視時間」，從而可以判斷人們對於特定影像廣告是否產生了興趣。本論文透過「變異數分析」和「隨機森林演算法」模型，藉由影像廣告中的品牌(Brand)和商品(Product)兩個刺激物，我們比較能了解一個人是否對於影像廣告產生了興趣，而藉由文字說明(Text)，比較了解一個人是否對於影像廣告有了注意力；至於代言人(Endorser)對於受測者和廣告的影響則還需要進一步去做驗證。雖然本論文已經有初步研究成果，但是實驗上或系統上，受限於眼動儀價格昂貴、眼動儀使用便利性、廣告型態多元和客戶數量，仍待近一步地探討。但我們相信隨著眼動儀以及影像辨識技術的進步，在不遠的未來仍然可以為廣告產業帶來極大的幫助。

本論文的研究結果，讓我們多了一個全新且客觀的方法去了解一個影像廣告與使用者的互動關係外，在未來或許還可以進一步的應用在非網路上的現實生活之中，例如：廣告看板，在目前市面上，並沒有一個很完善的方法去計算一個平面廣告看板或是電子廣告看板所產生的效益，我們只能花費人力去計算人潮的流量，再透過專家去大概推測看板的收益，這樣的方法，仍然無法得知一個看板實際上獲得了多少關注或是其無法受到關注的原因：是因為看板位置的關係？又或者是廣告內容本身的關係？本研究所開發的視覺分析系統若配合上人臉辨識技術，或許就能解決這方面的問題，讓我們能在節省人力的狀況下，就能知道一個影像廣告看板的曝光度外，還能更進一

步的去了解一個廣告看板實際上讓多少人產生了興趣、產生了多少效益，甚至有辦法直接針對客戶的視線和表情，即時去播放客戶可能會感興趣的廣告。



## 參考文獻

1. Aga Bojko (2013). *Eye Tracking the User Experience*. Rosenfeld Media, Brooklyn, New York.
2. Alex Poole, Linden J Ball, and Peter Phillips (2004). In search of salience: A response time and eye-movement analysis of book mark recognition. *People and Computers XVIII - Design for Life: Proceedings of HCI*, pp. 363-378.
3. B. Edelman, M. Ostrovsky, and M. Schwarz (2007). Internet Advertising and the Generalized Second Price Auction: Selling Billions of Dollars Worth Of Keywords. *Benjamin Edelman & Michael Ostrovsky & Michael Schwarz*, pp. 242-259.
4. B. Wooley (2015). The influence of dynamic content on visual attention during television commercials, Murdoch Univ., Dubai, United Arab Emirates.
5. Bergstrom, J. R., and Schall, A. (Eds.) (2012). *Eye tracking in user experience design*. Elsevier, pp.81-85.
6. Chenyu Li, Jun Liu, and Shuxin Ouyang (2016). Characterizing and Predicting the Popularity of Online Videos. *IEEE Access, (Volume: 4)*, pp. 1630-1641.
7. Edward K Strong (1925). *The Psychology of Selling and Advertising*. McGraw-Hill Book Co., New York.
8. Gazepoint GP3 Eye Tracker [Online]. Available: <https://www.gazept.com/product/gazepoint-gp3-eye-tracker/>
9. J. R. Quinlan (1986). Induction of decision trees. *Machine Learning*, Springer US, pp. 81-106.
10. Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi (2016). You Only Look Once: Unified, Real-Time Object Detection. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 27-30.
11. K. C. Yang, C. Yang, C. H. Huang, P. H. Shih, and S. Y. Yang (2014). Consumer attitudes toward online video advertising: An empirical study on YouTube as platform, *IEEE International Conference on Industrial Engineering and Engineering Management*, pp. 9-12.
12. L. Breiman (2001), Random Forests, *Machine Learning*, pp. 5-32.

13. Mao Wang, Yoichiro Maeda, and Yasutake Takahashi (2012). Human intention recognition via eye tracking based on fuzzy inference. *The 6th International Conference on Soft Computing and Intelligent Systems, and The 13th International Symposium on Advanced Intelligence Systems*, pp. 20-24.
14. Marco Porta, Alice Ravarelli, and Francesco Spaghi (2012). Online newspapers and ad banners: an eye tracking study on the effects of congruity. *Online Information Review, Vol. 37 Issue: 3*, pp.405-423.
15. Panos Louridas, and Christof Ebert (2016). Machine Learning. *IEEE Software*, pp. 110-115.
16. Pei-Shan Wei, and Hsi-Peng Lu (2013). An examination of the celebrity endorsements and online customer reviews influence female consumers' shopping behavior. *Computers in Human Behavior, 29*, pp.193-201.
17. R. Alexander et al. (2000). Systems and methods for displaying and recording control interface with television programs, video, advertising information and program scheduling information. *U.S. Patent 6177931*.
18. R. J. K. Jacob, and K. S. Karn (2003). Eye tracking in human-computer interaction and usability research: Ready to deliver the promises. *Mind, vol. 2, no. 3*, pp. 4.
19. Rik Pieters, and Michel Wedel (2004). Attention Capture and Transfer in Advertising: Brand, Pictorial, and Text-Size Effects. *Journal of Marketing*, pp. 36-50.
20. Ritu Lohtia, Naveen Donthu, and Idil Yaveroglu (2017). Evaluating the efficiency of Internet banner advertisements. *Journal of Business Research, Volume 60, Issue 4*, pp. 365-370.
21. Satoshi Kono (2009). From the marketers' perspective: The interactive media situation in Japan. *Television Goes Digital*, pp.57-59.
22. Sohil Jain, and Dr. Deepak Garg (2014). Evaluating Quality Score of New Ads. *International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, pp. 24-27.
23. Tin Kam Ho (1995). Random decision forests. *ICDAR '95 Proceedings of the Third International Conference on Document Analysis and Recognition*, pp. 278.
24. Tukur Dahiru (2008). P-value, a true test of statistical significance? a cautionary note. *Annals of Ibadan Postgraduate Medicine 6(1)*, pp. 21-26.

25. Wedel, M., and Pieters, R. (2007). A Review of Eye-Tracking Research in Marketing, *Review of Marketing Research*, pp. 123-147.
26. White SJ (2012). Eye Movement Control during Reading: Effects of Word Frequency and Orthographic Familiarity. *Journal of Experimental Psychology Human Perception & Performance*, pp.205-223.
27. Xuebai Zhang, and Shyan-Ming Yuan (2018). An Eye Tracking Analysis for Video Advertising: Relationship between Advertisement Elements and Effectiveness, *IEEE Access*, pp. 10699-10707.
28. Y. C. Hsieh, and K. H. Chen (2011). How different information types affect viewer's attention on Internet advertising. *Computers in Human Behavior*, vol. 27, no. 2, pp. 935-945.
29. Zhang, J., Wedel, M., and Pieters, R. (2009). Sales Effects of Attention to Feature Advertisements: A Bayesian Mediation Analysis. *Journal of Marketing Research*, pp. 669-681.

### 附錄一：廣告觀後問卷

注意層	
Q1. 我覺得廣告所呈現的畫面很吸引人	
Q2. 我有注意到廣告中的品牌	
Q3. 我有注意到廣告文字在寫些什麼	
興趣層	
Q4. 我對廣告中的產品有興趣	
Q5. 我想更進一步認識這個廣告品牌	
Q6. 我願意觀看同樣類似的廣告	