

OAC-114-019 (研究報告)

船舶設計之 AI 助教
(成果報告)

海洋委員會補助研究

中華民國 一一四年 九月

「本研究報告僅供海洋委員會施政參考，並不代表該會政策，該會保留採用與否之權利。」

OAC-114-019 (研究報告)

船舶設計之 AI 助教
(成果報告)

學校：國立臺灣海洋大學

指導教授：翁維珠

學生：黃翊倫

研究期程：中華民國 114 年 3 月至 114 年 10 月

研究經費：新臺幣 101,000 萬元

海洋委員會補助研究

中華民國 一一四年 九月

「本研究報告僅供海洋委員會施政參考，並不代表該會政策，該會保留採用與否之權利。」

「本研究報告絕無侵害他人智慧財產權之情事，如有違背願自負民、刑事責任。」

目次

表次.....	2
圖次.....	3
摘要.....	4
第一章 前 言.....	5
第一節 問題背景與研究緣起.....	5
第二節 文獻回顧與現況分析.....	6
第三節 研究目的與研究重點.....	9
第四節 預期目標.....	11
第二章 研究方法及過程.....	12
第一節 系統架構與代理流程.....	12
第二節 文件提取與版面分析.....	17
第三節 關係提取與向量索引.....	22
第四節 策略檢索與決策生成.....	26
第五節 使用介面與系統部屬.....	28
第三章 結果與討論.....	29
第一節 文件分析訓練結果.....	29
第二節 系統檢索生成結果.....	35
第四章 結 論.....	37
參考資料.....	39
附錄.....	41

表次

表 2-1 系統硬體與軟體配置表.....	13
表 2-2 Yolov10-large 模型訓練參數表	18
表 3-1 不同模型在文件版面分析性能與速度比較表	30
表 3-2 不同模型在問答測試集表現比較表	35

圖次

圖 2-1 系統架構圖.....	12
圖 2-2 Graph Agent 任務流程圖.....	14
圖 2-3 Query Agent 任務流程圖.....	15
圖 2-4 典型文件版面類型示意圖.....	19
圖 2-5 典型文件閱讀順序示意圖.....	20
圖 2-6 特別文件版面類型示意圖.....	21
圖 2-7 文件層級化關係提取示意圖.....	22
圖 3-1 Layout-Model 訓練損失函數曲線圖.....	29
圖 3-2 Layout-Model 訓練評估指標曲線圖.....	29
圖 3-3 Layout-Model 分析結果圖（一）.....	31
圖 3-4 Layout-Model 分析結果圖（二）.....	32
圖 3-5 Layout-Model 分析結果圖（三）.....	33
圖 3-6 Layout-Model 分析結果圖（四）.....	34

摘要

船舶設計是一門高度跨領域的專業工程，各國與船級社因而制定龐大而複雜的規範，這些文件往往動輒數千頁，包含大量圖片、公式與交叉引用，對學生與工程師而言檢索困難，進而影響學習與設計效率。本研究旨在開發一套「船舶設計 AI 助教」系統，以整合法規知識、連結設計邏輯並提升檢索效率，使學習者能即時確認設計合規性，並透過自然語言查詢降低逐條閱讀或關鍵字比對的負擔。並且在不依賴雲端服務且僅具消費級算力的環境下，仍能實現結構化文件解析與智慧化法規檢索，未來可望進與自動審圖及監造流程整合，推動船舶產業的數位化與智慧化轉型。

關鍵詞：船舶工程、大型語言模型、檢索增強生成、文件分析

第一章 前言

第一節 問題背景與研究緣起

一、 問題背景

船舶設計是一門高度跨領域的專業工程，涵蓋結構力學、流體力學、材料科學、電機與輪機工程以及人因工程等多方面知識，而工程師在設計過程中，除了需要整合這些專業知識，還必須遵循嚴謹的安全規範與國際規則。自 1912 年鐵達尼號事件以來，國際社會開始重視船舶安全，並於 1914 年由國際海事組織 (International Maritime Organization, IMO) 制定《國際海上人命安全公約》 (International Convention for the Safety of Life at Sea, SOLAS)，奠定了現代船舶安全規範的基礎。隨後，隨著船舶科技的快速發展，全球各大船級社如中國驗船中心 (CR Classification Society)、美國船級社 (American Bureau of Shipping, ABS)、挪威船級社 (Det Norske Veritas, DNV) 等，以及各國政府，陸續制定龐大的技術規範，內容不僅涉及結構、流體、材料、電機與輪機，還進一步擴展至人因工程與環境保護，包括《國際防止船舶造成污染公約》 (International Convention for the Prevention of Pollution from Ships, MARPOL)、《新造船能源效率設計指數》 (EEDI) 等。

雖然這些規範為船舶安全提供明確依據，但也帶來前所未有的複雜性。首先，法規文件篇幅龐大，往往動輒數千頁，如中國驗船中心的《Rules for the Construction and Classification of Steel Ships 2026》，其語言嚴謹且專業，對學習者與跨領域人才而言是沉重的負擔。其次，不同條文之間經常交叉引用，查閱時必須同時參考其他章節、附錄，甚至不同國際公約或國家法規，才能獲得完整資訊。再者，規範中包含大量表格與圖說，如結構示意圖、載重線曲線及安全距離數據，僅憑文字檢索難以全面掌握。此外，船舶設計不同於車輛設計，後者多數僅需符合單一出口國規範，而船舶若航行國際航線，必須同時遵守船籍國、經過國及目的港國的規定，使得法規檢索與應用更顯複雜。^[16]

二、 研究緣起

在傳統的學習與實務場域中，學生與工程師多依賴紙本文獻、線上法規資料庫或搜尋引擎逐條比對，但這些方式難以解決條文冗長、專業艱澀以及缺乏跨章節與跨文件串聯的問題。初學者因此難以將法規轉化為具體的工程依據，而資深工程師雖然知道某些規範存在，卻常遺忘細節，不得不反覆查找，最終，學習者與實務工作者都容易陷入「知道有規定，卻無法靈活應用」的困境，進一步阻礙產學銜接與專業實務的推進。

基於此，本研究提出並開發「船舶設計 AI 助教」系統，其核心目標在於整合法規知識、連結設計邏輯並提升學習效率。系統將分散於各國的法規與技術條文進行結構化處理，降低檢索與理解的門檻；同時將力學、流體、電機等專業計算過程與法規要求對應起來，使學習者能在設計過程中即時確認其合規性；並透過自然語言互動，讓使用者能以問題驅動的方式檢索與應用法規，而不再受限於繁瑣的逐條閱讀或單純的關鍵字比對。

第二節 文獻回顧與現況分析

一、 大型語言模型的潛力與限制

在船舶設計領域中，各國船級社與政府對法規的發布與維護主要依賴 PDF 文件與線上資料庫，雖然部分資料庫提供關鍵字檢索功能，但在實務上仍難以滿足跨章節、跨文件及跨模態（文字、表格、圖示）的需求，然而隨著大型語言模型 (Large Language Model, LLM) 的出現，為知識與規範檢索帶來新的解決思路。Vaswani et al. (2017) 提出的 Transformer 架構，使模型能在詞向量空間與解碼器的協作下處理長篇文本，並以自然語言與使用者互動，大幅提升檢索與學習體驗，隨著技術進展，部分前沿模型已能一次處理相當於長篇小說的文本量，甚至具備理解多模態資訊（如圖片、影片等）的能力。^[10]

然而，LLM 的本質仍是基於機率的生成系統，存在「幻覺 (Hallucination)」風險，即生成看似合理卻實際錯誤的答案，此外，船舶法規文件龐雜冗長，其總

量遠超過現代最新模型單次可處理的上下文長度，因此單純將所有文件直接輸入模型並不可行。更值得注意的是，現行主流 AI 聊天平台（如 ChatGPT、Gemini、Claude、Grok 等）多是透過外部套件擷取 PDF 文件內容，而非真正理解文件結構，使得面對掃描型 PDF 或課本格式時，這種方法更顯侷限，因而限制其在專業領域的應用效果。

二、 檢索增強生成的提出與發展

為了克服這些限制，學術界提出「檢索增強生成」（Retrieval-Augmented Generation, RAG），其原理是將文件切分成文本片段（chunk）後轉換為向量，再透過相似度檢索篩選出相關內容，最後交由模型生成答案，以降低幻覺並提高可追溯性。Lewis et al. (2020) 的研究指出，僅依賴參數化記憶的預訓練模型在知識密集型任務上表現不足，因此提出結合 Sequence to sequence (seq2seq) 模型與檢索索引的方法，即透過嵌入向量進行檢索與生成，該方法在開放域問答任務中不僅超越純參數模型，也使生成內容更加具體且事實正確。^[7]

隨後，Edge et al. (2023) 提出圖形檢索增強生成（Graph-based Retrieval-Augmented Generation, GraphRAG），即透過從文件中萃取實體與關係來構建知識圖譜，並以社群總結與多段彙整的方式回應「全局性」問題，彌補傳統 RAG 在多跳推理與整體性任務上的不足；Han et al. (2024) 則對 RAG 與 GraphRAG 進行系統性比較，結果顯示 RAG 在精確檢索與局部問答任務中更為穩定，而 GraphRAG 在查詢導向摘要與跨章節問題上具有優勢，顯示兩者之間存在互補關係。在此基礎上，Sarmah et al. (2024) 提出 HybridRAG，將向量檢索與知識圖譜檢索結合，在財務長文與專業術語繁多的場景中優於單一路徑的方法，對於處理船級規範這類「非結構化條文＋結構化表格」並存的語料具有直接啟發。^{[3][5][8]}

三、 文件結構化與多模態檢索

然而，僅靠文字檢索仍難以應對船舶規範中大量表格與圖說的特性，更無法解決特定文件在排版與閱讀上的挑戰。Auer et al. (2023) 提出的 Docling 框架，

正是針對 PDF 文件轉換而設計，結合 DocLayNet（版面分析）與 TableFormer（表格結構辨識）等專用模型，能在資源有限的環境下高效運行，與傳統僅能抽取文字的工具不同，Docling 能將段落、標題、表格與圖示結構化，轉化為可被機器理解的元素，為後續檢索與語言模型生成奠定更精確的基礎。在文件結構化之外，研究者也意識到跨模態檢索的重要性。因此 Zhao et al. (2024) 綜整影像、表格、程式碼與音訊等多模態檢索增強方法，指出多模態技術能顯著提升事實一致性、推理能力與可解釋性，對於必須同時整合文字、圖示與表格資訊的工程法規尤其關鍵。^{[1][9][14]}

另一方面，檢索過程本身仍存在不完美性與潛在知識衝突。Wang et al. (2024) 指出，外部檢索結果往往包含雜訊，甚至與模型內部知識矛盾，為此，他們提出 Astute RAG，能在檢索過程中自適應調用內部知識，並透過來源感知與可靠度判斷來整合多源資訊，即使在最不利的情境下亦能維持可靠表現，對於合規要求極高的船舶法規應用而言，這類強化可信度與可追溯性的設計具有重要意義。^[12]

四、船舶法規檢索的關鍵挑戰

綜合技術演進與領域特性，船舶法規檢索面臨的關鍵挑戰已日益清晰：其一，如何設計版面與語義兼顧的切片策略（包含條號、標題階層、表格單元格與圖說標註），以兼顧召回率與可解釋性；其二，如何策略選擇在向量檢索、圖譜導引與多模態檢索間動態切換，以支援多跳推理與全局性提問；其三，如何在來源可驗證與衝突調解的機制下，抑制幻覺並提供可檢核的出處。

第三節 研究目的與研究重點

一、 研究目的

在當前船舶設計產業中，龐大且跨領域的法規檢索已成為學生與工程師的沉重負擔。由於文件結構複雜、交叉引用頻繁，專業人員往往需要耗費大量時間進行文件比對與查找，進而影響學習效率與設計開發的進度。基於此挑戰，本研究的目的在於開發一套能有效應對船舶法規檢索需求的船舶設計 AI 助教系統。

本研究所提出的系統旨在協助使用者快速定位並理解與設計任務相關的規範條文，使學生與工程師能將更多精力集中於設計與創新，而非繁瑣的文件查找工作。透過整合法規知識與設計邏輯，系統能將分散於各國的法規與技術條文進行結構化處理，降低檢索與理解的門檻；同時將力學、流體、電機等專業計算過程與法規要求相互對應，協助學習者在設計過程中即時確認合規性。此外，系統將結合自然語言互動，讓使用者能以問題驅動的方式檢索與應用法規，而不再受限於繁瑣的逐條閱讀或單純的關鍵字比對。

最終不僅能提升教育端的教學與學習效率，還能推動造船產業的數位化與智慧化轉型，並促進跨領域人才的流動與培育，縮短產學落差，加速技術知識的實際應用，期望透過 AI 助教系統，為船舶設計教育與產業實務提供一個兼具效率、準確性與可追溯性的智慧化解決方案。

二、 研究重點

雖然 LLM 與檢索增強生成技術 (RAG 與 GraphRAG) 為法規檢索帶來了新契機，但在實際應用中仍受限於多項挑戰。多數造船企業與學術單位因涉及軍事或資安要求，無法依賴外部雲端 API，同時硬體環境亦多受限於消費級設備，難以維持伺服器等級的算力資源。因此，本研究的重點不僅在於技術方法的創新，更在於如何將其落實於低資源的實務環境，確保系統能在有限算力下仍具備高效能與可擴展性。具體而言，本研究將聚焦於以下五個方向：

1. 文件解析與多模態結構化

本研究將開發能在本地環境中高效運行的文件處理模組，針對 PDF 課本與規範進行多模態結構化解析。除了文字抽取外，系統將結合版面分析與表格結構辨識，將段落、編號、圖表轉換為機器可理解的結構化元素。進一步將這些元素將映射為知識圖譜節點與關係，作為 GraphRAG 的輸入基礎，確保船舶法規的多模態內容能以統一的圖結構表示，提升檢索與生成的精確度。

2. 多粒度索引（Multi-granularity Indexing）與策略檢索

本研究將建構結合向量索引與知識圖譜索引的多層級檢索框架。在局部問題情境下，系統可透過向量與全文檢索快速定位相關文本片段；而在跨文件或多跳推理任務中，則依賴 GraphRAG 機制，利用知識圖譜中的節點關係進行擴展與推理。特別是表格單元格與圖說元素，將被轉化為圖譜中的屬性或關聯，實現跨文字、表格與圖示的整合檢索，支援更複雜的全局性問題。同時，系統亦將引入 Astute RAG 的來源感知與多源整合策略，使其能在檢索結果存在雜訊或與內部知識衝突時，透過可靠度判斷進行答案篩選，提升系統可信度。

3. 多模態檢索與自然語言生成

為回應船舶法規中「圖文並重」的特性，本研究將設計具備多模態理解能力的檢索與自然語言生成機制。系統將整合文字、表格、曲線圖與結構示意圖等不同模態資料，並在回答問題時同時利用數據與語言。例如，在涉及載重線曲線或安全距離計算的情境下，系統能直接檢索表格或圖示的結構化資訊，並在自然語言生成過程中加以解釋與引用，實現「數據+語言」結合的答案生成模式，提供更具專業性與可操作性的回應。

第四節 預期目標

一、 預期目標

本研究的目標在於建構一套智慧化的船舶設計 AI 助教系統，能夠自動化處理並理解使用者所提供的可攜式文件格式（Portable Document Format, PDF）。這些文件不僅包含一般課本教材，還涵蓋船舶法、各船級社規範以及國際公約（如 SOLAS 等），其內容形式多樣，涉及文字、圖片與圖表等多模態資訊。系統也將具備自主提取與解析能力，能將文件中的多模態元素進行結構化處理，萃取其中的實體與關係，並進一步嵌入至多模態向量空間，最終儲存於圖形資料庫中，以便後續的檢索與推理。

在應用端，使用者可透過網頁介面提出查詢，系統將依據問題的特性自動選擇合適的檢索策略，並綜合向量檢索、圖譜導引以及多模態檢索等方法，精準定位相關資訊。系統同時具備雜訊檢測與可信度判斷機制，能在外部檢索結果存在不確定性或與內部知識衝突時，進行檢索篩選，以確保生成回覆的可靠性與可追溯性。最終，系統將以自然語言生成方式回傳答案，幫助學生與工程師快速獲取法規依據，提升學習與設計效率，以回應船舶工程教育中所面臨的學科整合困難與教育與產業接軌間之落差問題，並推動船舶工程教育與產業的智慧化轉型。

第二章 研究方法及過程

第一節 系統架構與代理流程

1. 系統架構

與現行主流 AI 聊天平台多僅依賴外部套件擷取 PDF 文件內容，並直接以單一模型進行推理不同，本研究開發之系統採代理式人工智慧（AI Agent）與圖形資料庫相結合的多層結構，整體架構如圖 2-1 所示。

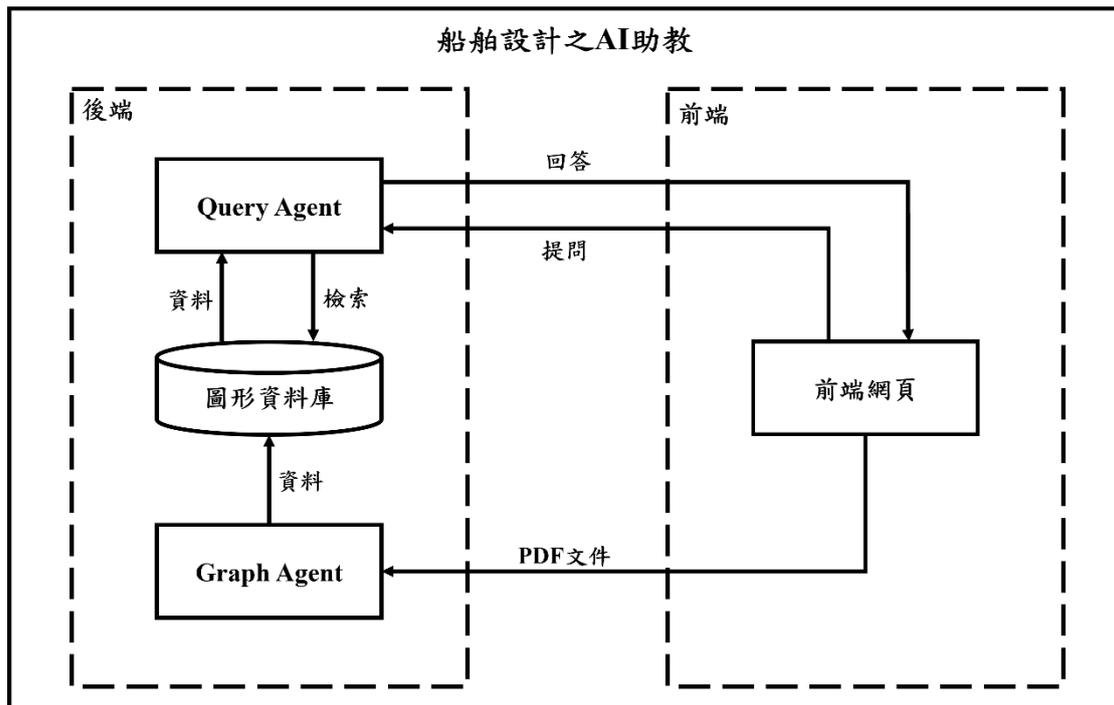


圖 2-1 系統架構圖

整體架構可分為四大核心組成：

1. 文件圖形化代理（Graph Agent）負責對使用者提供的 PDF 文件進行分析與內容提取，包括文字、表格與圖像等多模態元素的解析，並建立實體與關係的映射。這些元素經過轉換後，將被嵌入至多模態向量空間，並儲存於圖形資料庫。
2. 圖形資料庫承擔資料存取與中介的角色，保存嵌入向量、文字資料及元數據（如時間戳記、文件來源與版本資訊），不僅確保資料的完整與可追溯性，也作為兩個代理間的通訊橋樑。

3. 查詢代理 (Query Agent) 專注於檢索、生成與推理。當使用者提出查詢時，Query Agent 會先判斷是否需要檢索，並自動選擇最合適的檢索策略 (如向量檢索、知識圖譜導引或多模態檢索)。在檢索過程中，系統具備雜訊判斷與資料篩選能力，可排除不相關或矛盾資訊，並在整合檢索結果後進行推理，生成自然語言形式的回覆。
4. 前端網頁介面作為使用者與系統互動的橋樑，提供直觀的操作環境，讓使用者能以自然語言提出問題，並獲得結合檢索與生成的答案回應。

為了有效驗證系統在實際部署環境中的可行性，本研究採用消費級高階工作站作為實驗平台，以確保多模態文件解析與模型能在本地環境中高效運行。完整的系統硬體與軟體配置如表 2-1 所示。

表 2-1 系統硬體與軟體配置表

類別	規格
CPU	Intel Core™ i9-14900K
GPU	NVIDIA GeForce RTX4090
記憶體	DDR4 128GB
儲存空間	5TB SSD, 8TB HDD
作業系統	Ubuntu 22.04.05 LTS
Python	3.12.4
NVIDIA Driver	575.64.03
NVIDIA CUDA	12.9
NVIDIA cuDNN	v9

2. 代理流程

根據系統設計，兩個代理的任務流程如圖 2-2 與圖 2-3 所示：

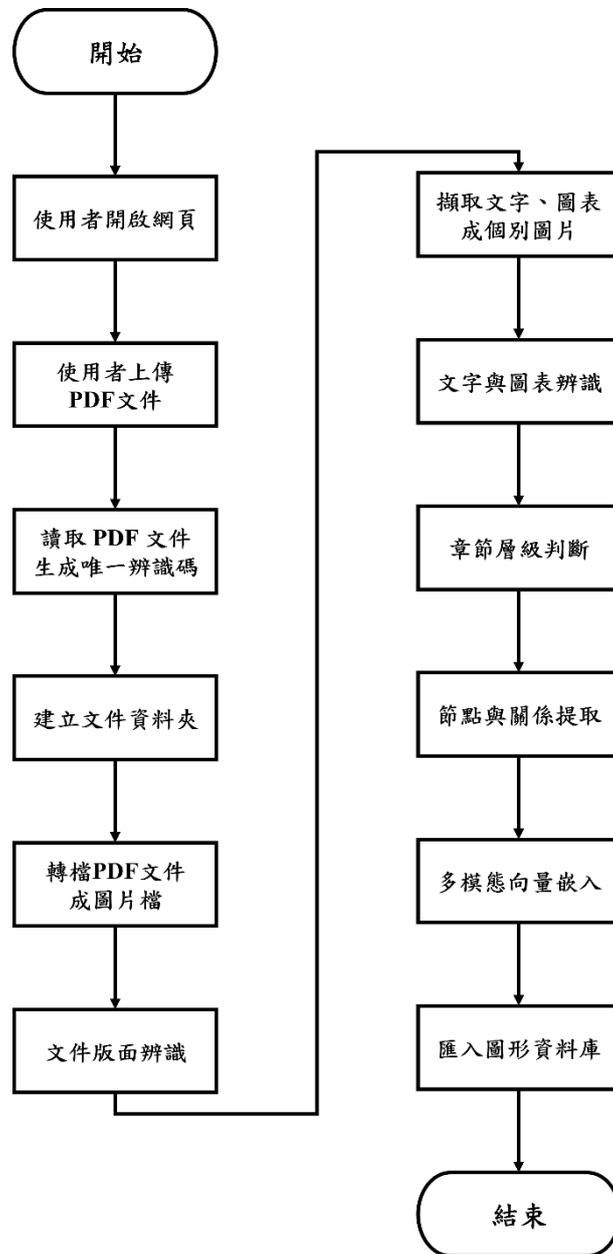


圖 2-2 Graph Agent 任務流程圖

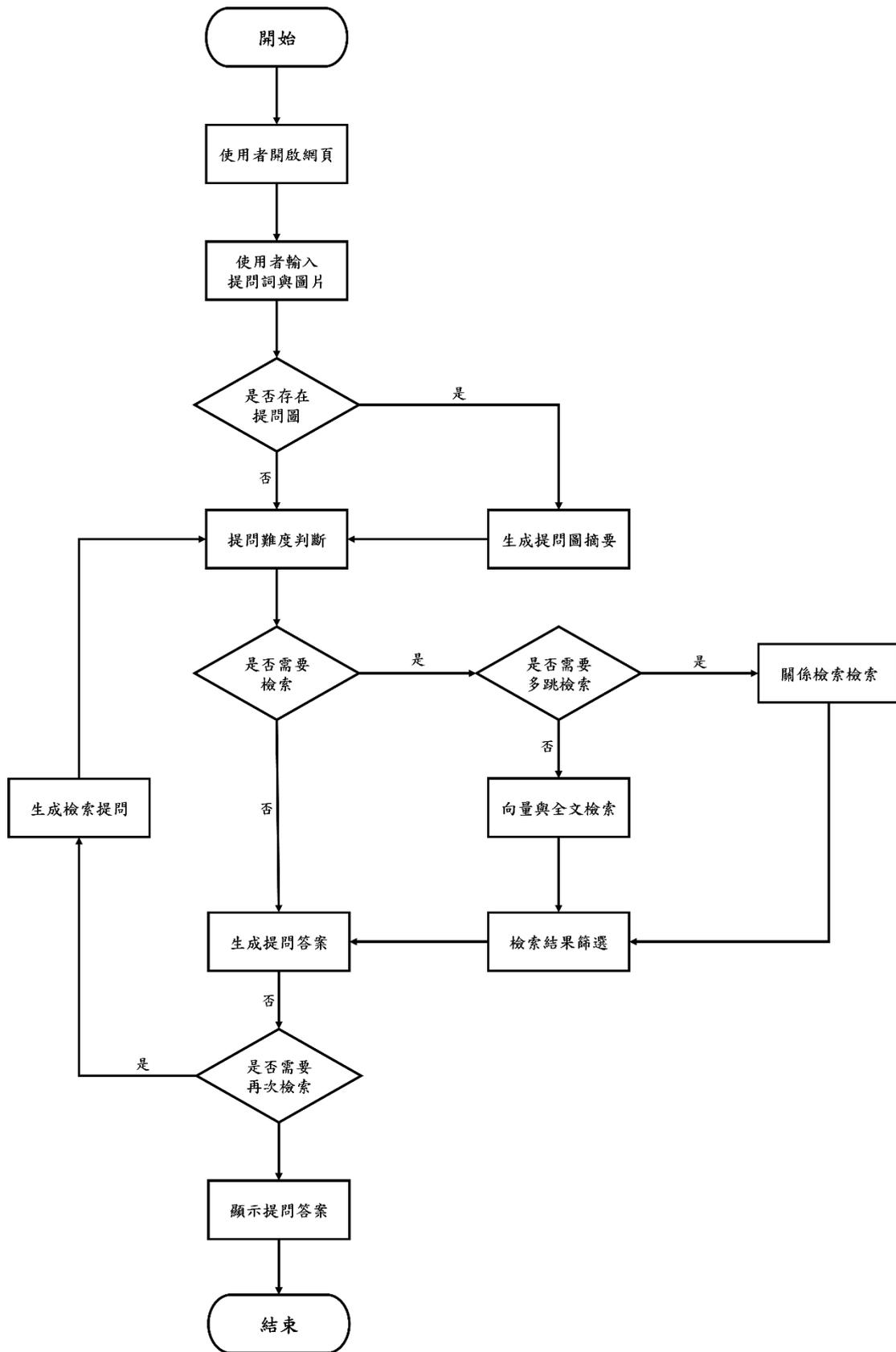


圖 2-3 Query Agent 任務流程圖

整個系統的流程從使用者開啟網頁開始。首先，使用者可以上傳 PDF 文件，系統會為每份文件生成唯一的識別碼，並建立對應的文件資料夾，以確保後續處理結果歸檔。接著，PDF 文件會被轉換成圖檔，以利後續的圖像與文字處理。

在轉檔完成後，系統會進行文字與圖表的辨識，透過 OCR 技術與版面分析，將文件中的文字、數學公式、圖片與表格逐一辨識出來，系統會將文件拆解為不同元素，將文字、圖表與圖像分別提取成獨立的區塊，並同時進行文件版面的辨識，以確定正確的閱讀順序。當元素被正確識別後，系統會進行章節層級判斷，釐清文件的結構，並自動提取節點與其之間的關係。這些章節、段落與條列項目會被映射成結構化層級，進而形成文件的知識樹。隨後，系統會將多模態資料進行向量嵌入，將它們轉換為可用於檢索的語義向量，並最終存入圖形資料庫。

在完成文件資料的處理與儲存後，系統的互動流程從使用者輸入問題開始。使用者可以輸入文字問題，並選擇性地上傳相關的提問圖片。系統會先檢查是否存在提問圖，如果有圖片輸入，會透過模型生成該圖片的語義摘要，並與文字問題整合為完整的查詢輸入。

接著，系統會根據輸入問題進行難度判斷。若問題屬於簡單型，系統可能直接回覆；若涉及較複雜的情境，則需判斷是否需要進行檢索。當需要檢索時，系統會依問題性質選擇適合的檢索策略：若僅需單一來源即可解答，則透過向量檢索與全文檢索完成；若屬於跨節點的多跳推理，則會透過關係檢索來跨越多個章節或條文，以獲得完整的答案來源。

檢索完成後，系統會將候選結果進行篩選，並生成檢索提問，再由模型生成最終答案。如果答案未能符合問題需求，系統會判斷是否需要再度檢索，若需要則重新進行檢索與生成的閉環流程，直到獲得合理回應為止。當最終答案生成完成後，系統會將結果顯示給使用者，並附上引用依據與可追溯的來源資訊。

在此基礎之上，將於後續說明核心技術的實現與具體應用，進一步展示系統如何支援高效、精準且可追溯的知識檢索與答案生成。

第二節 文件提取與版面分析

一、 文件提取

船舶相關法規與一般傳統法規存在顯著差異。船舶設計領域的專業知識與技術規範往往包含大量圖片（如結構示意圖、設計圖表等）、數學公式以及複雜的章節層級結構。同時，不同機構如各大船級社、各國政府及國際海事組織所制定的法規文件雖有相似之處，但在章節命名與結構上缺乏統一標準，造成跨文件整合上的困難。類似情況亦存在於工程相關教材之中，其內容同樣富含圖片、公式與表格，甚至包含結構示意與左右版面等排版形式，對文件處理帶來挑戰。

現有主流的大型語言模型，如 GPT-5、Claude Sonnet 4、Grok 4 等，在處理單一頁面的內容時，能給出完整且正確的回答；但面對動輒數千頁的船舶規範文件時，則受到上下文長度限制，難以涵蓋整體內容。若僅透過 PDF 文字抽取並進行向量嵌入，系統能回答文字類型的問題，卻無法針對圖片內容做出有效回應，尤其對於掃描型 PDF 文件，甚至無法產生任何回應；反之，若將 PDF 全部轉換為圖片並以多模態方式處理，雖能應付圖像提問，但在文字細節的理解與生成上則顯得不足，最終導致難以同時兼顧文字與圖像內容的精確處理的現狀。

為解決上述問題，本研究採用 Docling 框架作為基礎，並結合 YOLO 模型進行文件元素分析，將文字、公式、圖片與表格分別進行提取與標註。雖然 Docling 已內建性能強大的模型，但在消費級工作站的環境下，受限於套件與架構設計，推理速度與效能仍難以滿足需求。因此，本研究進一步利用 DocLayNet 資料集對 YOLOv10-large 模型進行訓練，以兼顧精度與推理效能。

DocLayNet 為一個人工標註的版面分析資料集，共含來自六種類型文件的 80,863 個頁面，並以邊界框形式提供 11 種類別標籤，進行逐頁版面切分。數據集劃分為訓練集 69,375 頁、驗證集 6,489 頁、測試集 4,998 頁，其標註類別如下：

1. Caption（圖表說明）
2. Footnote（註腳）

3. Formula (公式)
4. List-item (條列項目)
5. Page-footer (頁尾)
6. Page-header (頁首)
7. Picture (圖片)
8. Section-header (節標題)
9. Table (表格)
10. Text (正文)
11. Title (標題)

YOLOv10-large 模型不僅能滿足文件解析的訓練需求，其推理效能亦符合後續在本地端消費級環境中部署的需求。其訓練參數如表 2-2 所示。為了驗證模型的效能，本研究的評估方法採用 Precision、Recall 與 IoU (Intersection over Union) 作為主要指標，以全面衡量模型在檢測準確度、召回率及邊界框重疊度的表現。實驗比較基準包括：

- 原始 YOLOv10-large 模型：檢視訓練後的性能提升幅度。
- Qwen2.5-VL-7B 多模態模型：對比傳統目標檢測模型與多模態大型語言模型在文件解析任務上的差異與互補性。^[2]

表 2-2 Yolov10-large 模型訓練參數表

訓練參數	設定值
訓練集數量	69,375
驗證集數量	6,489
圖像尺寸	1024
批次大小 (Batch Size)	12
訓練週期 (Epoch)	30

二、 版面分析

然而，經過 DocLayNet 訓練的 YOLOv10-large 模型僅能提供文件元素的邊界框 (bounding box, bbox)，並未解決文件閱讀順序的問題。因此，本研究設計文件版面理解演算法，能依據元素類別與邊界框資訊，推導出合理的閱讀順序。典型文件版面類型示意圖如圖 2-4 所示。

圖 2-4 展示船舶設計相關法規與教材中常見的三種版面結構：

1. 左圖為一般規範文件之分布，其排版方式多為單欄結構，由頁首與章節標題開始，以下為大段正文，並輔以條列項目進行細節說明，最後以頁尾收尾。此類版面結構線性明確，通常可依自上而下的順序直接閱讀。
2. 中圖為課本例題分布，常見於教材中用於呈現範例或演算過程。此類版面通常為雙欄結構，左、右兩欄分別包含文字與圖片的交錯呈現，並於欄位下方附帶補充說明。閱讀時需依據欄位先左後右的順序來確保內容的完理解。
3. 右圖為課本左頁分布，此類版面通常出現在書本攤開的左頁位置。頁首與標題置於上方，正文與公式位於中央，兩側則插入圖片與說明。由於版心位置偏向書脊，閱讀時多先處理靠近書脊的右半區塊，再向左側圖片與附註延伸。

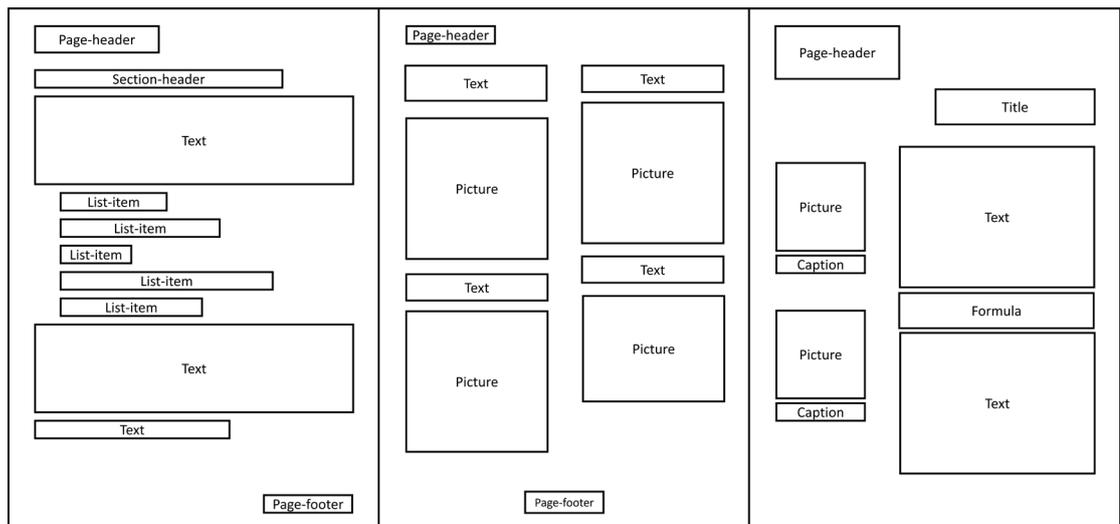


圖 2-4 典型文件版面類型示意圖

(左：單欄規範文件；中：課本例題雙欄排版；右：課本左頁圖文混排)

透過對三種典型版面分布的分析，可以看出不同文件類型在版面設計上的差異，也凸顯僅依據 bounding box 位置排序無法正確反映閱讀順序的問題，基於此，本研究設計的演算法可根據版面區分線，動態判斷閱讀區塊，如圖 2-5 所示：左圖是單欄規範文件，中圖是課本例題雙欄排版，右圖是課本左頁圖文混排。

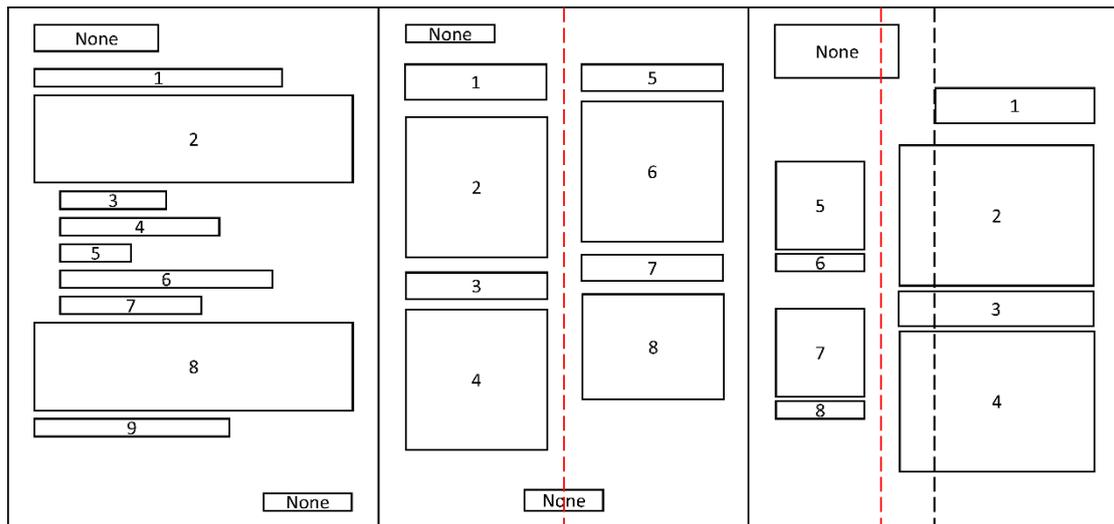


圖 2-5 典型文件閱讀順序示意圖

在文件閱讀順序的判斷上，演算法首先會排除頁首與頁尾元素。這些區塊通常僅包含頁碼或篇章標題，雖然在文件定位與結構標記上具有參考價值，但並非正文內容的一部分，若納入閱讀順序反而可能干擾系統的解析。因此，本研究將頁首與頁尾資訊保留於後續的元數據紀錄中，而不列入正文判讀流程。

在正文部分，演算法會根據元素的上下與左右關係進行判斷。若元素左右皆無其他區塊，則依由上至下的順序直接閱讀；若元素存在左右分布，則需進一步依據版面結構區分左右區塊。以圖 2-5 為例，左圖的所有元素皆呈現單欄結構，故可依直線式自上而下的方式判讀。而中圖與右圖則屬於多欄或跨頁結構，左右皆有元素分布，因此必須引入左右分組策略。

在中圖中，版面清楚劃分為左右兩欄，依據英文教材與規範的慣例，通常採先讀左欄、再讀右欄的順序。而右圖則顯示出右半頁的元素佔據較大區域，甚至

跨越頁面的一半以上，可判斷其為課本左頁的排版，因此在閱讀時需優先處理右半頁的內容，再回到左側補充資訊。最後，為處理更複雜的混合結構，本研究設計的演算法亦能動態分組，以應對不同的左右分布情境。此類特殊案例將於圖 2-6 中進一步展示，其中左圖顯示版面分布，右圖展示閱讀順序。

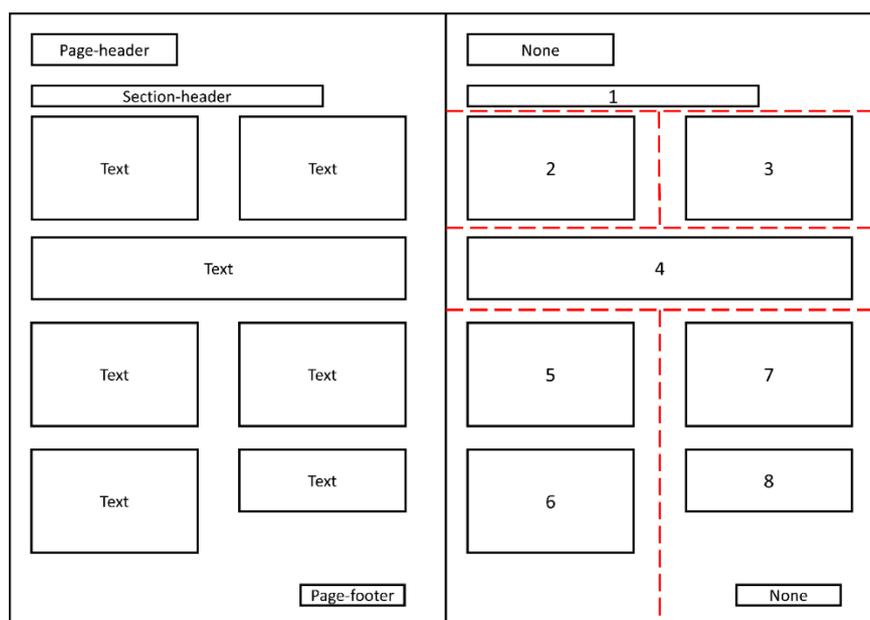


圖 2-6 特別文件版面類型示意圖

在圖 2-6 所示的此類版面中，雖無法以單一步驟直接區分左右欄位，但可透過元素分組策略進行判讀。具體而言，若元素左右無其他區塊，則可直接視為單一閱讀單位並優先處理；若元素存在左右分布，則依據由左至右的順序進行閱讀；而當元素再次呈現無左右相鄰的情況時，則重新作為下一個獨立的閱讀單位。此方法能有效應對不連續的左右半頁結構，使系統具備更靈活的版面理解能力，不僅超越原始功能的限制，亦能提供更具檢核性與一致性的閱讀邏輯。

最終，系統將辨識後的文件元素依照演算法進行逐頁排序，並對其中的圖片生成唯一辨識碼，以確保後續檢索與追蹤的正確性。隨後，系統透過 Qwen2.5-VL-7B 模型執行光學字元辨識 (OCR)，從中提取文字、數學公式與圖表數據。所有處理結果將統一儲存為 JSON 格式檔案，其中完整保留文件的閱讀順序、

內容提取結果以及圖片的唯一辨識碼，為後續檢索與生成提供結構化且可追溯的基礎。

第三節 關係提取與向量索引

一、 關係提取

在完成文件元素的抽取與閱讀順序判斷後，接下來需進行關係提取 (Relation Extraction)。現有主流方法大多將文件切分為固定長度的片段，再由開發者設定節點與關鍵字，交由大型語言模型自動尋找並建立關係。然而，這種方式存在三項限制：其一，模型推理過程屬於「黑箱」，缺乏可解釋性與可驗證性；其二，依賴人工關鍵字設定，難以適應不同來源與不同格式的大量文件；其三，對上下文長度與計算能力的需求極高，造成在消費級設備環境下難以高效部署。

為克服上述問題，本研究提出一種基於文件結構導向的關係提取方案。如圖 2-7 所示，系統在前一階段已完成文件元素的標註與閱讀順序推導，基於這些資料，將文件轉換為層級化樹狀結構 (Tree Structure)，並以此來表徵文件內容之間的邏輯與層級關係。

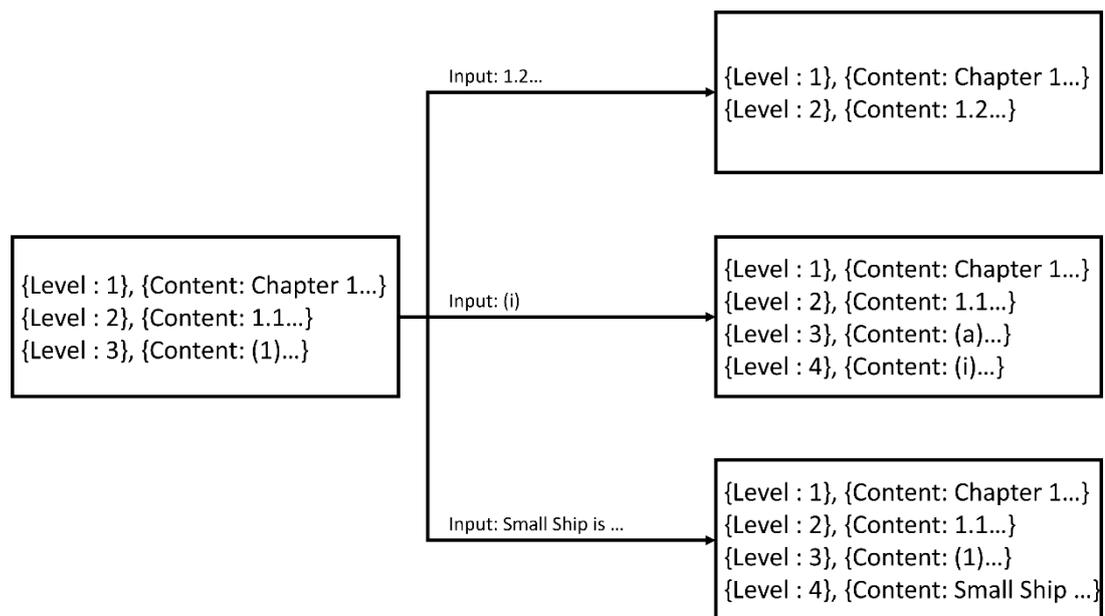


圖 2-7 文件層級化關係提取示意圖

在該結構中，每一層節點對應不同層級的文件單元：

- Level 1：對應章節標題（如 Chapter 1），作為最高層次的內容劃分；
- Level 2：對應子章節（如 1.1、1.2），承接章節內容的細分；
- Level 3：對應條列或子項目（如 (1)、(a)），進一步展開具體規範；
- Level 4：對應更細粒度的規範與具體條文（如小項 (i)，或明確定義如 Small Ship …）。

在文件結構關係的建立過程中，若遇到與現有層級相似或相同的結構，則會自動以更新與替代的方式進行整合，避免重複與冗餘；若判定為新內容（如圖片、文字段落、數學公式或新子層級），則會在節點樹中新增一個子層級，並將其掛載於對應的父節點之下。透過這種動態層級式映射，文件條文內容得以逐步自動嵌入節點樹，最終構建出一個能完整反映文件結構的樹狀關係網絡。

此種方法具備三大特點：首先，它能保證文件內部邏輯的延續性，父子節點間的關聯清楚反映了「章節—子章節—條列—細項」的逐層展開關係；其次，它在檢索與推理時能保留文件的完整脈絡，避免因傳統片段切分方法導致的上下文斷裂與信息遺失；最後，透過持續的節點更新與分層，樹狀結構能隨著文件處理的進行逐步擴展，使系統具備「時刻保留記憶」的特性，特別適合處理動輒數千頁的長篇技術文件或法規。

與傳統基於片段切分的 RAG 方法相比，本研究提出的文件結構導向關係提取方案展現出顯著優勢：

- 可解釋性：文件關係並非由模型隱式推導，而是根據章節號碼、條列符號、縮排與格式標記等顯性結構建立。研究人員與使用者可直接檢視節點樹，驗證其合理性，從而提升系統的透明度與可信度。
- 低算力需求：不再依賴大模型處理長上下文，而是透過文件本身的層級結構進行邏輯建模，極大降低了推理對 GPU 記憶體與計算力的要求，使其能在消費級硬體環境中穩定運行。

- 普適性：該方法不依賴關鍵字設計，能直接適用於多來源文件，包括船級社規範、國際海事公約、各國政府法規及工程教材。無論文件語言、排版或來源機構如何差異，只要具備層級化的章節或條列結構，系統皆能有效解析。
- 擴展性：節點樹能作為知識圖譜（Knowledge Graph）的基礎，進一步支援跨章節檢索、多跳推理與規範對照分析。未來亦能結合多模態信息，將圖片、表格與公式的解析結果納入圖譜節點中，實現跨模態知識整合。

換言之，圖 2-7 所示的層級化節點樹，不僅能準確映射文件的結構特徵，還能有效降低關係提取的計算成本，提升可解釋性，並且解決跨頁文件資訊斷裂的問題，最終也為後續的知識圖譜構建、跨文件鏈結以及 AI 助教系統的合規檢索提供可擴展的技術基礎。這種方法特別適合應用於船舶設計領域龐大而複雜的規範文件，可顯著提升文件檢索、跨章節比對與設計應用的效率，進而縮短學習曲線，促進產學銜接。

更進一步地，本研究提出的方法不僅限於船舶領域，還能推廣至其他需遵循嚴格規範的高專業領域，例如航空航天、能源工程、建築結構設計與醫療器械檢驗等。這些領域同樣存在結構龐雜、跨模態、跨文件的法規與技術文件，因此，文件結構導向的關係提取方案將具有跨領域的推廣價值與學術貢獻。

在完成文件結構的層級映射後，系統會將結果轉換為 Neo4j 圖形資料庫中的節點（Node）與關係（Relationship），使文件不再僅是線性文本，而是可查詢、可推理的圖結構。每個節點除了儲存其文字內容外，還會附帶元數據（Metadata），例如元素類型、所屬文件名稱、頁碼位置以及圖片的唯一辨識碼等。

透過這種設計，研究人員或使用者不僅可以檢索某一章節的文字內容，還能進一步查詢該章節對應的圖表或公式；反之，也能從某一圖片節點追溯至其所屬條文或章節，從而實現多方向、多層級的檢索與推理。這種方法使船舶法規與教材能以結構化的方式被理解和應用，並為後續的知識圖譜構建提供基礎支撐。

二、 向量嵌入

在完成文件結構的關係提取後，系統進一步將節點所包含的文字與圖片進行向量嵌入，以支援後續的檢索、比對與推理。本研究選用 Jina Embeddings v4 作為核心嵌入模型，該模型是一種專為多模態與多語言檢索設計的通用嵌入方法，能夠將文字、圖像、圖表等元素映射至統一的語義空間，並在語言與視覺之間建立橋樑。對於結構龐大且資訊豐富的船舶法規與工程教材而言，單純依賴文字嵌入往往不足以捕捉圖表和公式的意涵，而 Jina Embeddings v4 所提供的多模態嵌入特性，正好能補足缺陷，使系統能對文字、圖片和表格同時進行建模。^[4]

模型嵌入的維度預設為 2048 維，在表示能力上能保證充分的語義覆蓋；然而，對於資源有限的消費級硬體而言，2048 維的計算負擔仍然不容忽視。因此，本研究採用其可調式嵌入機制，在必要時將嵌入維度縮減至 128 維，並透過模型內建的降維優化方法確保效能損失最小化，不僅大幅減少運算時間與儲存空間需求，也使系統能在單機環境下處理數千頁以上的文件，為部署在學術與企業環境提供了可行性。

此外，本研究在嵌入策略上並未採用「一體適用」的模式，而是根據文件元素類型進行差異化設計。對於純文字節點，僅進行文字嵌入，保證語義準確並維持上下文的一致性；對於純圖片節點，則專注於圖像嵌入，保留其在形態、色彩與結構上的特徵；對於圖表節點，則採取雙重嵌入策略，將圖表內部的文字與圖形部分分別嵌入。原因在於，圖表承載的資訊具有雙重特性：一方面，文字內容通常包含數據標註、單位與解釋，對理解條文具有關鍵價值；另一方面，圖形部分則呈現出趨勢、比例與結構，相似度檢索往往需要依賴圖形特徵來判斷不同圖表之間的關聯。若僅嵌入文字，將可能遺漏圖形層面的相似性；若僅嵌入圖形，則無法捕捉語義層面的精準細節。因此，雙重嵌入策略能最大限度地保留圖表的多重信息，確保檢索結果的全面性與可靠性。

這種依節點類型調整的差異化嵌入策略，使得系統能夠在檢索過程中展現更高的靈活性。例如，當使用者查詢「載重線計算方法」時，系統可以返回包含文字解釋的條文節點，同時檢索出對應的曲線圖表，形成「文字+圖形」的完整答案；若使用者查詢「救生艇配置圖」的相似案例，系統則能透過圖像嵌入找到結構相似的配置圖，並追溯到對應的法規條文，以此更接近專業人員在實務中查閱資料的思維模式。

最終，系統會將生成的嵌入向量整合進節點的元數據 (Metadata) 中，並連同節點之間的關係一併匯入圖形資料庫。如此一來，每個節點不僅保有文字內容、頁碼、圖片唯一辨識碼等結構化資訊，還附加了對應的語義向量，使其能同時支援結構化查詢與語義檢索，兼顧可追溯性與靈活性。

本研究選用 Neo4j 作為核心圖形資料庫，其在圖結構存儲與檢索上具有成熟的生態系統，提供 Cypher 查詢語言，能直觀表達節點與關係的模式；另一方面，Neo4j 與 Python 的整合度高，可透過官方驅動程式與多種外掛套件，實現與文件處理模組、檢索模組的無縫銜接，並且近年逐步支援向量檢索 (Vector Search)，允許將嵌入向量直接儲存在節點屬性中，並基於向量相似度執行檢索，這使得結構化圖譜與語義檢索能夠融合運作。

第四節 策略檢索與決策生成

一、策略檢索

前兩節已經說明 Graph Agent 如何透過文件解析與關係提取建立完整的知識圖譜，接下來的流程將系統帶入檢索階段。當使用者提交查詢時，輸入內容可能同時包含文字與圖片，因此第一步是進行多模態訊息的偵測與處理。如果查詢中帶有截圖，系統會啟動 Qwen2.5-VL-32B 模型對圖片進行語義摘要，並將該摘要與使用者的文字問題合併成一個完整的輸入。之所以採用 Qwen2.5-VL-32B，是因為它具備優異的多模態處理能力，能同時理解圖文並生成高品質摘要，特別適合技術性問題中常見的表格或截圖。[2][12]

在獲得整合後的查詢輸入後，Qwen3-32B 負責進一步分析。這個模型因其語言理解與推理的強大能力，被用來判斷問題是否需要進入資料庫檢索；若問題與知識圖譜內容相關，系統就進入檢索流程，若屬於一般性知識，則直接利用模型的內建參數回答，以減少檢索的性能消耗。選擇 Qwen3-32B 的原因，在於它在問題類型判斷、檢索策略規劃與結構化查詢語句生成上的穩定性與精準性。

接著，系統會依問題屬性而定檢索策略的決策，若是單純的文字匹配會啟用全文檢索，語義相關的問題則使用向量檢索，以確保找到語義上相似的內容；至於多跳問題，系統會基於知識圖譜生成對應的 Cypher 語句，藉由關係檢索跨越多個節點完成推理。為了避免檢索結果過於稀疏或充斥雜訊，系統會動態調整 Top-K 參數：Top-K 過大會帶來大量不相關內容，過小則可能導致上下文不足，因此需要在召回率與精確度之間保持平衡。

二、 決策生成

完成檢索之後，系統將取得的文字與圖片內容交由 Qwen3-32B 生成答案，並再次透過同一模型檢查答案是否切實回應原始問題。如果判斷不符合，系統會重新調整檢索策略並再次檢索；若生成的答案合理，則直接回傳給使用者，通過「生成—驗證—再檢索」的閉環流程，能確保回覆既準確又具備可追溯性。

在整體流程的最終評估中，系統會引入 ROUGE-L 指數作為生成答案的語言品質衡量依據，並設計不同層次的題庫來檢驗系統效能。題庫規模包括一百道《船舶法》英文版相關問題、五百道《Rules for the Construction and Classification of Steel Ships 2026》問題，以及五百道《Engineering Mechanics: Statics, Fourteenth edition in SI Units》問題，題型涵蓋簡單定義、數學定義、多跳推理、圖片與圖表解讀，甚至跨文件整合。透過 GPT-5 直接處理 PDF 文件並以 Qwen3-32B 進行向量檢索的比較，可以全面測試系統在不同層級與不同任務上的表現，進一步驗證檢索與生成流程的有效性。^{[6][15][16]}

第五節 使用介面與系統部屬

一、 使用介面

在後端完成檔案解析、知識圖譜建構與檢索策略的流程設計後，系統需要一個直觀的入口，讓使用者能夠輕鬆地與之互動。前端介面採用 React 架構開發，結合 Tailwind CSS 等元件庫，實現簡潔而具延展性的設計風格。介面功能包含文字輸入框與圖片上傳區，使用者可同時輸入文字問題並附加截圖，系統會自動進行多模態處理，並在結果區域返回答案與相關文件片段。為了提升使用體驗，前端設計即時回應提示的可視化呈現，使使用者能理解系統在處理過程中的狀態。同時，介面亦提供檢索結果的追蹤功能，使用者可以展開檢視系統引用的文本來源，確保回覆具備可追溯性與透明度。

為了與後端保持高效通訊，系統使用 FastAPI 建立中介層，負責處理使用者請求與後端模型的交互。FastAPI 能支援高併發請求，方便開發與維護。前端的每一次查詢請求，會經由 FastAPI 轉送至檢索與生成模組，再將結果封裝為 JSON 格式返回，確保整個互動流程快速且穩定。

二、 系統部屬

在系統部署的規劃上，為了同時兼顧效能與可擴展性，我們採用容器化技術與雲端資源相結合的方式。後端服務以及模型推理模組會事先封裝為標準化的 Docker 映像檔以及上傳至 Hugging Face 網站，這樣能有效避免不同環境間的依賴衝突，確保開發、測試與正式運行環境的一致性。透過容器化的設計，系統不僅能快速在各種雲端平台或本地伺服器上部署，也能在需求增加時透過橫向擴展輕鬆啟動更多實例以分擔負載，進一步提升整體的可移植性與彈性。

第三章 結果與討論

第一節 文件分析訓練結果

一、 訓練結果

以 DocLayNet 資料集訓練 YOLOv10-large 模型，稱為 Layout-Model，訓練過程與評估指數如圖 3-1 與圖 3-2 所示：

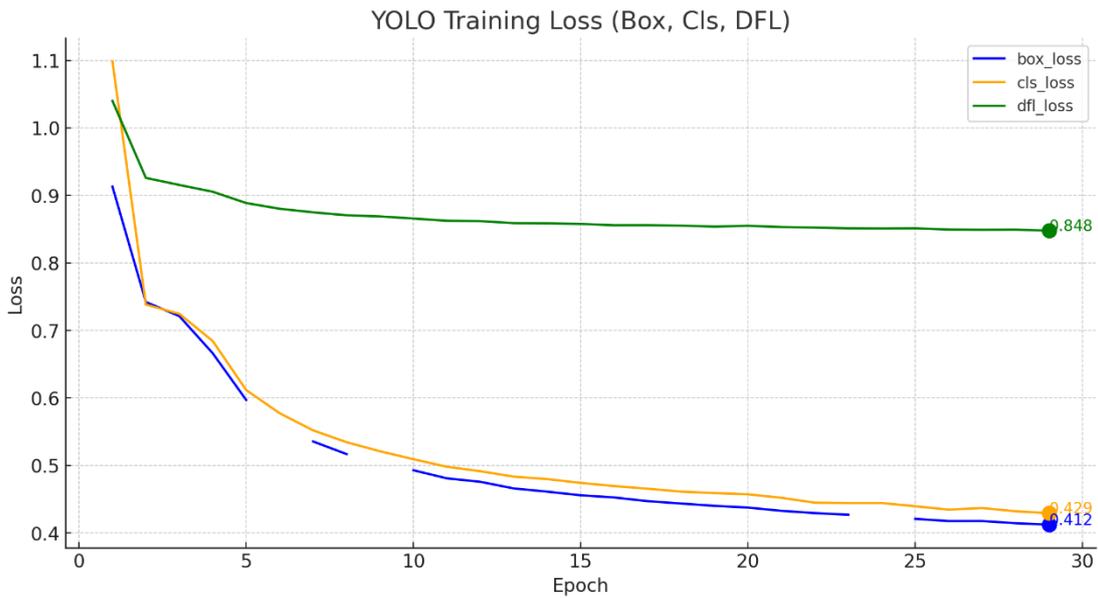


圖 3-1 Layout-Model 訓練損失函數曲線圖

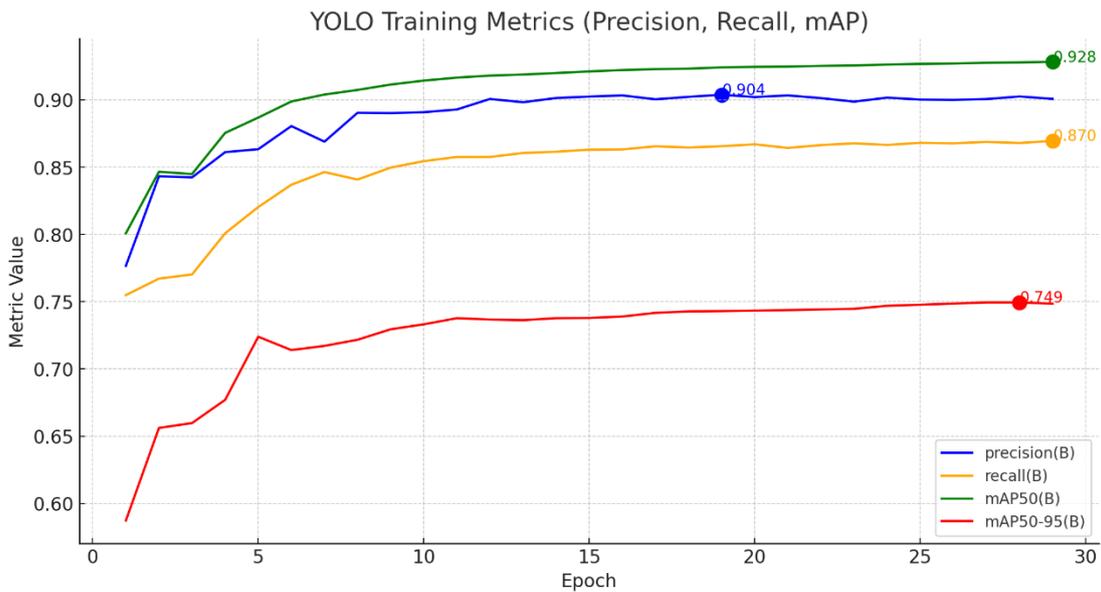


圖 3-2 Layout-Model 訓練評估指標曲線圖

根據訓練結果可見，Layout-Model 在約 Epoch 25 後逐漸收斂，損失函數 (Loss) 趨於平穩，Precision、Recall 與 mAP 等評估指標亦持續提升並保持穩定，顯示模型在文件版面分析任務上具備良好的學習效果。為進一步驗證本研究模型的效能，與未經訓練的 YOLOv10-large 以及 Qwen2.5-VL-7B 模型對 DocLayNet 驗證集進行比較，結果如表 3-1 所示。

表 3-1 不同模型在文件版面分析性能與速度比較表

	Precision	Recall	mAP50	mAP50-95	Speed
YOLOv10-large	0.428	0.234	0.453	0.167	10.3 ms
Qwen2.5-VL-7B	0.139	0.152	0.105	0.003	30 s
Layout-Model (Our)	0.902	0.871	0.931	0.747	10.2 ms

在性能表現上，未經訓練的 YOLOv10-large 模型僅能提供基礎的檢測能力，缺乏在文件場景上的專門適配，表現有限但合理。而 Qwen2.5-VL-7B 雖具備強大的多模態能力，但其在文件版面分析中的表現不佳，Precision 僅有 0.139，mAP50 與 mAP50-95 更分別下降至 0.105 與 0.003。造成這一現象的主要原因在於 Qwen2.5-VL-7B 傾向於生成過大的 bounding box，導致一定程度的高召回率 (Recall = 0.152)，但在精細定位與準確率上大幅落後。

相比之下，本研究提出的 Layout-Model 展現出顯著優勢，其 Precision 達到 0.902，Recall 為 0.871，mAP50 與 mAP50-95 分別為 0.931 與 0.747，在速度上亦保持與 YOLOv10-large 相近的推理效能 (約 10.2 ms)。表示本模型不僅能在精確度與召回率上大幅提升，同時兼顧了推理效率，適合部署於消費級硬體環境中。

然而，實驗結果亦顯示仍存在進一步的優化空間。雖然 Layout-Model 在 Precision 與 Recall 上已達高水準，但在 mAP50-95 指標上仍有改進潛力，顯示模型對於小尺度物件或複雜排版的處理尚不夠完美。

二、 實際結果

PART I CHAPTER I
1.1 General

Section header: 0.30
Chapter I
Classification of High-Speed Craft

1.1 General

Text: 0.90
1.1.1 High-Speed Craft (HSC) built and surveyed in accordance with the Rules for the Construction and Classification of High-Speed Craft (hereinafter referred to as the Rules) published by the CR Classification Society (hereinafter referred to as the Society) or alternatives found to represent an overall safety standard equivalent to that of the Rules (See 1.1.6 of this Part) will be assigned a class in the Register of Ships (hereinafter referred to as the Register) and will continue to be classed so long as they are found, upon examination at the prescribed surveys, to be maintained in a fit and efficient condition and in accordance with the requirements of the Rules.

Text: 0.88
The Rules is applicable to high-speed craft constructed of steel, aluminum, or FRP and having V/\sqrt{L} not less than 2.36 where L is as defined in 1.1.2 of Part III and V is as defined in 1.1.18 of Part III. Applicable craft type and length are as follows:

Table: 0.88

Application Limits	
Craft Type	Applicable Length
Mono-hull	< 130 m
Multi-hull	< 100 m
Surface Effects Ship (SES)	< 90 m
Hydrofoil	< 60 m

Text: 0.93
Craft which have lengths that are in excess of the above table may be classed using the requirements in the Rules, as applicable, providing that requirements in 2.1 and 2.2 of Part III of the Rules are verified by theoretical prediction, first principles analysis, and model testing.

Text: 0.84
The craft will at all times be in reasonable proximity to a place of refuge, having due regard to the following:

List-Item: 0.87
(a) Passenger craft which do not proceed in the course of their voyage more than 4 hours at 90% of maximum speed from a place of refuge; and

List-Item: 0.86
(b) Cargo craft of 500 gross tonnage and upwards which do not proceed in the course of their voyage more than 8 hours at 90% of maximum speed from a place of refuge when fully laden.

Text: 0.92
1.1.2 Classification will be conditional upon compliance with the Rules in respect of both hull and machinery (i.e., main and auxiliary engines, boilers, essential appliances, pumping arrangements and electrical equipment). The products including the equipment, components, systems and materials intended for classed craft or craft to be classed are to comply with the requirements in "Guidelines for Survey of Product for Marine Use".

Text: 0.83
1.1.3 The Rules are framed on the understanding that craft will be properly loaded and handled; they do not, unless stated in the class notation, provide for special distributions or concentrations of loading. The Society may also require additional strengthening to be fitted in any craft which, in their opinion, may be subjected to severe stresses due to particular features in her design, or when it is desired to make provisions for exceptionally loaded or ballasted conditions. In these cases particulars are to be submitted for consideration.

Text: 0.89
1.1.4 The stability of the craft is to be sufficient and in compliance with the requirements of governmental authority or International Conventions and Codes and Part II Chapter 30 and 30A of the Rules for Steel Ships, as applicable.

Text: 0.90
1.1.5 The Rules do not cover certain technical characteristics, such as trim, hull vibration, etc., but the Society is willing to advise on such matters although it cannot assume responsibility for them.

Page footer: 0.70
1.1.6 Alternatives are to be accepted, provided that they are considered by the Society to be equivalent to the Rules.

Page header: 0.73
CR CLASSIFICATION SOCIETY
RULES FOR THE CONSTRUCTION AND CLASSIFICATION OF HIGH-SPEED CRAFT 2026

- 1 -

圖 3-3 Layout-Model 分析結果圖（一）

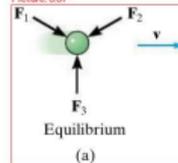
Text 0.06

Newton's Three Laws of Motion. Engineering mechanics is formulated on the basis of Newton's three laws of motion, the validity of which is based on experimental observation. These laws apply to the motion of a particle as measured from a *nonaccelerating* reference frame. They may be briefly stated as follows.

Text 0.03

First Law. A particle originally at rest, or moving in a straight line with constant velocity, tends to remain in this state provided the particle is *not* subjected to an unbalanced force, Fig. 1-1a.

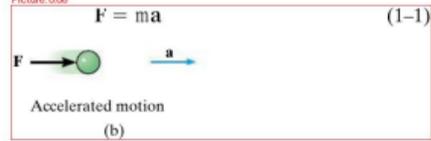
Picture 0.07



Text 0.06

Second Law. A particle acted upon by an *unbalanced force* \mathbf{F} experiences an acceleration \mathbf{a} that has the same direction as the force and a magnitude that is directly proportional to the force, Fig. 1-1b.* If \mathbf{F} is applied to a particle of mass m , this law may be expressed mathematically as

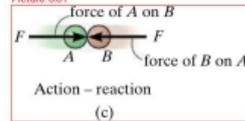
Picture 0.08



Text 0.03

Third Law. The mutual forces of action and reaction between two particles are equal, opposite, and collinear, Fig. 1-1c.

Picture 0.01



Caption 0.00

Fig. 1-1

Text 0.03

*Stated another way, the unbalanced force acting on the particle is proportional to the time rate of change of the particle's linear momentum.

圖 3-4 Layout-Model 分析結果圖 (二)

Picture: 090



Text: 059

第 15-1 條

Text: 091

1 自國外輸入現成船，除法規另有規定者外，應於輸入前檢附買賣意向書或契約書、船舶規範、船舶證明及權責機關同意等相關文件，向航政機關申請核定。

Text: 090

2 自國外輸入之現成船或使用目的變更者，其船齡不得超過允許輸入之年限。輸入現成船年限表，由主管機關公告之。

Text: 091

3 第一項之船舶證明指船舶國籍證書。但船舶不適用該國船舶法規，致無船舶國籍證書者，得以造船廠之建造證明文件替代。

Section-header: 030

第 16 條

Text: 092

船舶依第十五條規定登記後，航政機關除依船舶登記法之規定核發登記證書外，並核發船舶國籍證書；必要時，得先行核發臨時船舶國籍證書。

Text: 063

第 17 條

Text: 094

船舶所有人在所認定之船籍港以外港口取得船舶者，得檢附取得船舶或原船籍國之相關證明文件，向船舶所在地或船籍港航政機關申請核發臨時船舶國籍證書，並應自領得該證書之日起三個月內，依第十五條規定申請登記。

Text: 061

第 18 條

Text: 095

1 在船籍港以外港口停泊之船舶，遇船舶國籍證書遺失、破損，或證書上登載事項變更者，該船舶之船長或船舶所有人得自發覺或事實發生之日起三個月內，向船舶所在地或船籍港航政機關，申請核發臨時船舶國籍證書。

Text: 091

2 船舶在航行中發生前項情形時，該船舶之船長或船舶所有人應向到港港或船籍港航政機關為前項申請。

Text: 093

3 依前二項規定申請臨時船舶國籍證書者，船舶所有人應自領得該證書之日起三個月內，向船籍港航政機關申請換發或補發船舶國籍證書。

Text: 059

第 19 條

Text: 094

臨時船舶國籍證書之有效期間，在國外航行之船舶不得超過六個月；在國內航行之船舶不得超過三個月。但有正當理由者，得敘明理由，於證書有效期間屆滿前，向船舶所在地或船籍港航政機關重行申請換發；重行換發證書之有效期間，不得超過一個月，並以一次為限。

Text: 041

第 20 條

Text: 095

1 經登記之船舶，遇滅失、報廢、喪失中華民國國籍、失蹤滿六個月或沉沒不能打撈修復者，船舶所有人應自發覺或事實發生之日起四個月內，依船舶登記法規定，向船籍港航政機關辦理船舶所有權註銷登記；其原領證書，除已遺失者外，並應繳銷。

Text: 091

2 船舶改裝為遊艇或小船者，船舶所有人應自改裝完成之日起三個月內，依規定辦理登記或註冊，其原領證書除已遺失者外，應予繳銷。

圖 3-5 Layout-Model 分析結果圖（三）

<small>Text: 0.96</small>	surveyors may not be engaged in work related to the duty of a ship surveyor such as private or public shipping carrier business, shipping agency business or shipbuilding factories, etc.
<small>List-Item: 0.87</small>	Article 86 The practicing certificate of a ship surveyor is valid for 5 years. Licensed ship surveyors shall apply for renewal of the practicing certificate one month before the expiry of the license, enclosing the original practicing certificate and supporting documents of the service experience.
<small>List-Item: 0.52</small>	Article 87 Persons Under any one of the following circumstances may not take the practice as a marine surveyor. For those who have already been a practicing ship surveyor, their practicing certificate shall be withdrawn or annulled: 1. Those who have been convicted of offenses against internal and external security. 2. Those that are sentenced to one-year imprisonment or above due to business-related criminal acts and no probation is pronounced. 3. Those whose exam qualification is withdrawn or annulled pursuant to the provisions of Examination Act.
<small>Text: 0.96</small>	4. Those who are diagnosed to have psychiatric disorder or abnormal physical or mental status, and verified by a relevant specialist doctor designated by the competent authority as unfit for the practice.
<small>Text: 0.89</small>	5. Those whose pronouncement of custody or assistance has not been withdrawn.
<small>List-Item: 0.88</small>	Article 88 The competent authority shall undertake to prescribe and give effect to the regulations on the issuance, renewal/replacement, annulment, cancellation or withdrawal of the practicing certificate for a ship surveyor, imposition of fees and other matters for compliance.
<small>Section-header: 0.82</small>	Chapter 10 Penalties
<small>List-Item: 0.87</small>	Article 89 In case of violation of the provisions of Article 8, the shipping administration authority shall impose a fine between NT\$30,000 and NT\$300,000 on the owner or master of a small ship, or deck

圖 3-6 Layout-Model 分析結果圖（四）

由以上結果圖所示，可以看出 Layout-Model 在不同類型的文件上的提取能力都十分優異，但是在部分文件格式上還是有待加強，如圖 3-6 的段落標號並未被辨識到。

第二節 系統檢索生成結果

一、系統檢索比較

以本研究開發之系統，以及網頁版 GPT-5，以及傳統向量檢索的 Qwen3-32B 模型以一百道《船舶法》英文版相關問題、五百道《Rules for the Construction and Classification of Steel Ships 2026》問題，以及五百道《Engineering Mechanics: Statics, Fourteenth edition in SI Units》問題進行比較，比較表如表 3-2 所示：

表 3-2 不同模型在問答測試集問答表現比較表

	定義問題	數學問題	多跳問題	多模態問題	跨文本問題
《船舶法》英文版					
GPT-5	0.63	<u>0.79</u>	<u>0.68</u>	None	<u>0.62</u>
RAG	0.84	0.81	0.35	None	0.32
Our	<u>0.79</u>	0.65	0.69	None	0.71
《Rules for the Construction and Classification of Steel Ships 2026》					
GPT-5	0.65	<u>0.78</u>	0.71	<u>0.21</u>	<u>0.59</u>
RAG	0.81	0.81	0.29	0.14	0.21
Our	<u>0.75</u>	0.63	<u>0.52</u>	0.51	0.61
《Engineering Mechanics: Statics, Fourteenth edition in SI Units》					
GPT-5	0.54	0.83	0.67	<u>0.24</u>	<u>0.64</u>
RAG	0.74	<u>0.61</u>	0.21	0.13	0.24
Our	<u>0.71</u>	0.57	<u>0.57</u>	0.41	0.67

根據表 3-2 的比較結果可觀察到，傳統 RAG 系統在回答定義型問題與數學計算問題時表現最佳，其原因在於這類問題通常可透過單段文字檢索與直接匹配獲得答案。然而，當問題涉及多跳推理、多模態內容或跨文件檢索時，RAG 系統便顯現出瓶頸，分數顯著下降，顯示其在複雜任務中的局限性。

相比之下，GPT-5 與本研究所開發的系統在多跳推理與跨文件問題上展現出相對良好的表現，兩者表現相近。然而，在多模態問題上，本研究系統明顯優於 GPT-5，能更有效處理圖文並重的輸入。進一步分析 GPT-5 的回答發現，雖然其輸出往往涵蓋標準答案的核心內容，但同時生成了大量冗餘敘述，導致自動化評分的相似度分數偏低；而本研究系統則主要受限於 OCR 誤讀與文字提取錯誤，導致答案與標準參考不符，這部分仍有進一步優化的空間。

值得注意的是，若輸入資料為掃描型 PDF，GPT-5 與傳統 RAG 的表現均大幅下滑，原因在於二者均無法直接解析影像格式的文件，僅能依賴模型內建知識生成回覆，缺乏真實檢索能力。相對而言，本研究系統由於結合文件結構化處理與多模態檢索，能維持一定的效能，展現出其在實務應用中的優勢。此外，在實驗過程中也發現 GPT-5 在推理階段會意外生成 Python 程式碼，嘗試以程式方式檢索或驗證答案，這一行為雖然出乎預料，但顯示其具備潛在的工具化推理能力，值得後續深入探討。

第四章 結 論

一、 結 論

本研究開發一套可在消費級硬體環境下運行的船舶設計 AI 助教系統，能針對龐雜的法規文件進行結構化解析、知識圖譜建構與多模態檢索，並以自然語言提供可追溯的回答。系統由 Graph Agent 負責文件元素提取與閱讀順序建立，再轉換為知識圖譜；Query Agent 則依據問題類型選擇向量檢索、圖譜檢索或多模態檢索，並進行答案生成。

在文件解析上，系統結合 Docling 框架與 YOLOv10-large，利用 DocLayNet 資料集訓練以提升文件元素檢測能力，並設計演算法還原閱讀順序，最終轉換為 Neo4j 節點與關係。多模態處理方面，系統採用 Jina Embeddings v4，針對文字、圖片與圖表分別設計嵌入策略，確保圖文並重的法規內容能被完整捕捉。

比較實驗顯示，傳統 RAG 在定義與數學問題表現最佳，但在多跳、跨文件與多模態問題上明顯受限；GPT-5 與本系統則在多跳與跨文件問題上表現相近，但 GPT-5 因生成冗語導致自動評分偏低，而本系統則受限於 OCR 錯誤。值得注意的是，本系統在多模態問題的處理優於 GPT-5，展現結構化檢索與生成的優勢。

整體而言，本研究證明了在不依賴雲端 API 的條件下，仍能建構可應用於教育與產業的 AI 法規助教系統。在教育端，可幫助學生快速掌握規範要點；在產業端，則能提升工程師的檢索效率並降低錯誤風險。

二、 未來工作

本研究已成功驗證在消費級硬體環境下建構船舶設計 AI 助教系統的可行性，但仍有若干面向值得持續發展與優化。首先，在文件解析方面，未來可引入更高精度的 OCR 模型與多語言支援，以進一步提升文字辨識的正確率，特別針對掃描型 PDF 與低品質影像文件，減少因文字誤讀導致的知識圖譜錯誤。其次，在多模態處理上，後續研究可嘗試更細緻的融合策略，讓文字、圖片、數學公式

與圖表能夠在同一語義空間中更深度交互，並透過自適應權重的方式強化不同模態間的貢獻度，提升跨模態檢索的表現。

在檢索策略上，未來系統可結合強化學習或自適應檢索決策機制，根據問題複雜度與資料庫狀態動態調整 Top-K 值與檢索流程，以進一步降低雜訊並提升回答準確性。此外，可探索「檢索規劃 (Retrieval Planning)」的應用，讓系統能像專業工程師般逐步查閱不同法規，形成更符合人類推理習慣的回答過程。

在產業應用層面，未來工作可將系統與船舶設計軟體或自動審圖系統整合，實現從設計輸入到法規驗證的全自動化流程，並拓展至監造與維修階段的規範檢核，為產業鏈提供更完整的智慧化支援。同時，在教育端，可開發結合教學案例與模擬訓練的互動功能，讓學生在真實的設計情境中使用 AI 助教進行法規檢索與設計驗證，縮短學習曲線並強化專業能力。

最後，在跨領域推廣上，系統架構可延伸至航空航天、能源工程、建築結構設計與醫療器械等領域，這些領域同樣面臨複雜規範與跨模態文件的挑戰，因此本研究方法具有高度的可移植性與學術價值。未來若能結合開放資料集與跨領域合作，將有助於建立一個普適的結構化法規檢索框架，推動智慧化法規應用的全面發展。

參考資料

1. Auer, C., M. Lysak, A. Nassar, M. Dolfi, N. Livathinos, P. Vagenas, C. B. Ramis, M. Omenetti, F. Lindlbauer, K. Dinkla, L. Mishra, Y. Kim, S. Gupta, R. T. de Lima, V. Weber, L. Morin, I. Meijer, V. Kuropiatnyk and P. W. J. Staar (2024). "Docling Technical Report." arXiv preprint arXiv:2408.09869.
2. Bai, S., K. Chen, X. Liu, J. Wang, W. Ge, S. Song, K. Dang, P. Wang, S. Wang, J. Tang, H. Zhong, Y. Zhu, M. Yang, Z. Li, J. Wan, P. Wang, W. Ding, Z. Fu, Y. Xu, J. Ye, X. Zhang, T. Xie, Z. Cheng, H. Zhang, Z. Yang, H. Xu, J. Lin et al. (2025). "Qwen2.5-VL Technical Report." arXiv preprint arXiv:2502.13923.
3. Edge, D., H. Trinh, N. Cheng, J. Bradley, A. Chao, A. Mody, S. Truitt, D. Metropolitanaky, R. O. Ness, and J. Larson (2024). "From Local to Global: A Graph RAG Approach to Query-Focused Summarization." arXiv preprint arXiv:2404.16130.
4. Günther, M., S. Sturua, M. K. Akram, I. Mohr, A. Ungureanu, B. Wang, S. Eslami, S. Martens, M. Werk, N. Wang, and H. Xiao (2025). jina-embeddings-v4: Universal Embeddings for Multimodal Multilingual Retrieval. arXiv preprint arXiv:2506.18902.
5. Han, H., H. Shomer, Y. Wang, Y. Lei, K. Guo, Z. Hua, B. Long, H. Liu, and J. Tang (2025). "RAG vs. GraphRAG: A Systematic Evaluation and Key Insights." arXiv preprint arXiv:2502.11371.
6. Hibbeler, R. C., Engineering Mechanics: Statics, Fourteenth Edition in SI Units (England: Pearson, 2017), 1-684
7. Lewis, P., E. Perez, A. Piktus, F. Petroni, V. Karpukhin, N. Goyal, H. Küttler, M. Lewis, W. Yih, T. Rocktäschel, S. Riedel, and D. Kiela (2020). "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks." arXiv preprint arXiv:2005.11401.
8. Pfizmann, B., C. Auer, M. Dolfi, A. S. Nassar, and P. W. J. Staar (2022). "DocLayNet: A Large Human-Annotated Dataset for Document-Layout Analysis." arXiv preprint arXiv:2206.01062.
9. Sarmah, B., B. Hall, R. Rao, S. Patel, S. Pasquali, and D. Mehta (2024). HybridRAG: Integrating Knowledge Graphs and Vector Retrieval Augmented Generation for Efficient Information Extraction. arXiv preprint arXiv:2408.04948.

10. Vaswani, A., N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser and I. Polosukhin (2017). "Attention Is All You Need." arXiv preprint arXiv:1706.03762.
11. Wang, A., H. Chen, L. Liu, K. Chen, Z. Lin, J. Han and G. Ding (2024). "YOLOv10: Real-Time End-to-End Object Detection." arXiv preprint arXiv:2405.14458.
12. Wang, F., X. Wan, R. Sun, J. Chen, and S. Ö. Arik (2025). Astute RAG: Overcoming Imperfect Retrieval Augmentation and Knowledge Conflicts for Large Language Models. arXiv preprint arXiv:2410.07176.
13. Yang, A., A. Li, B. Yang, B. Zhang, B. Hui, B. Zheng, B. Yu, C. Gao, C. Huang, C. Lv, C. Zheng, D. Liu, F. Zhou, F. Huang, F. Hu, H. Ge, H. Wei, H. Lin, J. Tang, J. Yang, J. Tu, J. Zhang, J. Yang, J. Yang, J. Yang, J. Zhou, J. Zhou, J. Lin, K. Dang, K. Bao, K. Yang, L. Yu, L. Deng, M. Li, M. Xue, M. Li, P. Zhang, P. Wang, Q. Zhu, R. Men, R. Gao, S. Liu, S. Luo, T. Li, T. Tang, W. Yin, X. Ren, X. Wang, X. Zhang, X. Ren, Y. Fan, Y. Su, Y. Zhang, Y. Zhang, Y. Wan, Y. Liu, Z. Wang, Z. Cui, Z. Zhang, Z. Zhou and Z. Qiu (2025). "Qwen3 Technical Report." arXiv preprint arXiv:2505.09388.
14. Zhao, P., H. Zhang, Q. Yu, Z. Wang, Y. Geng, F. Fu, L. Yang, W. Zhang, J. Jiang, and B. Cui (2024). "Retrieval-Augmented Generation for AI-Generated Content: A Survey." arXiv preprint arXiv:2402.19473.
15. "The Law of Ships," Laws & Regulations Database of The Republic of China (Taiwan), 28 Nov. 2018, <https://law.moj.gov.tw/ENG/LawClass/LawAll.aspx?pcode=K0070004> (9 Apr. 2025).
16. "Rules for the Construction and Classification of Steel Ships," CR Classification Society, <https://www.crclass.org/cr-rules/> (9 Apr. 2025).

附錄

1. OCR 提示詞

Below is the image of one page of a document.

Just return the text representation of this document as if you were reading it naturally.

All tables should be presented in HTML format like `<tr><td> </td></tr>`.

All Formula should be presented in HTML format like `$ $`.

Do not hallucinate.

2. 關係提取提示詞

You are an expert in hierarchical document structure analysis.

Input:

- 1) A Level List: a list of dictionaries, each with : { "Level": `<int>`, "Content": `<string>` }
- 2) A paragraph of text (string).

Task:

Determine the correct structural LEVEL of the given paragraph in the document.

Rules:

1. Iterate through the Level List from bottom to top (start checking from the most recent item).

2. Sibling Rule (Same Level):

- If the paragraph text starts with a heading token or marker that follows the same pattern as a sibling (e.g., "1.2" after "1.1", "(b)" after "(a)", "Article 2" after "Article 1"), then assign the SAME level as that sibling.

3. Child Rule (New Sub-Level):

- If no sibling match is found, but the paragraph begins with a recognizable heading token (Chapter, Article, Section, Part, Clause, Paragraph, number, roman numeral, or parenthesized marker), then treat it as a child of the nearest valid parent above it.

- Assign `Level = parent.Level + 1`.

4. Plain Text Rule:

- If the paragraph is plain text (no heading token, no number, no parentheses at the start):

- First, check if the nearest previous plain-text item exists. If yes → SAME level as that item.

- If no previous plain-text sibling is found → nest it under the nearest valid parent (Level = parent.Level + 1).

5. Priority of Parent-Child Relation:

- Always prefer sibling detection first, before assigning as child.

- If neither sibling nor child rule applies, fallback to plain-text rule.

Examples:

Case 1:

```
{"Level" : 1, "Content" : "Chapter 1"}
```

```
{"Level" : 2, "Content" : "1.2"}
```

```
{"Level" : 3, "Content" : "1.2.3"}
```

Paragraph: "1.3 The Content is ..."

=> Level = 2 (sibling of "1.2")

Case 2:

```
{"Level" : 1, "Content" : "Article 3"}
```

```
{"Level" : 2, "Content" : "3.2"}
```

```
{"Level" : 3, "Content" : "(a)"}
```

Paragraph: "The Content is ..."

=> Level = 4 (child of "(a)")

Output:

Return only a JSON object in the exact format: {"level": <integer>}

Restrictions:

- Do not add any other keys or text.

- Do not output reasoning.

- Do not hallucinate.

3. 答案生成提示詞

You are an expert answer generator in the domain of naval architecture and shipbuilding engineering.

Input:

1) A question (string).

2) A list of results (list of strings).

Task:

Formulate a concise and accurate answer to the question using only the provided results.

Output:

Return only a JSON object in the exact format: {"answer": "<string>"}

Restrictions:

- Do not add any other keys or text.
- Do not output reasoning.
- Do not hallucinate.