Using GeoAI and Mixed-Data to Classify Built Heritage

by

Siyu Li

A thesis

presented to the University of Waterloo

in fulfillment of the

thesis requirement for the degree of

Master of Science

in

Geography

Waterloo, Ontario, Canada, 2024

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Built heritage, which comprises structures that are valued for their historical or architectural characteristics, is important to community's identity, sense of place, and, in many cases, economy. Many communities in Ontario have used Heritage Act provisions to protect locally significant buildings through local heritage registers or local bylaws that officially designate them as heritage structures. Recent changes introduced by Ontario's Bill 23, the 2022 More Homes Built Faster Act, have significantly altered the heritage designation process by limiting how long a property can remain on a local heritage register. This change highlights the need for a more efficient method for municipalities to identify and classify potential heritage properties. Geospatial Artificial Intelligence (GeoAI), an innovative approach integrating AI with Geographic Information Science (GIS), has potential to automate heritage identification and classification tasks, and assist heritage planners.

This thesis explores the potential of GeoAI to streamline heritage property identification and designation through three main models. These models leverage Convolutional Neural Networks (CNN) and Multi-Layer Perceptrons (MLP) to classify architectural styles, identify potential heritage properties, and predict heritage designations. The models were trained on non-spatial and geospatial datasets, including archival photographs and region-specific data from Ontario, enhancing their ability to detect architectural details and heritage features unique to the area.

The results demonstrate the effectiveness of these models, with the Architectural Style Classification Model achieving a 89% accuracy, despite challenges with similar styles. The Heritage Identification Model significantly improved efficiency with a 96.62% accuracy rate, while the Heritage Property Designation Prediction Model, combining CNN and MLP approaches, achieved 96% accuracy. The findings highlight the potential of AI and GeoAI to aid heritage practices with new technological methods. This research contributes to the broader knowledge base by providing refined tools for decision-making in heritage conservation and also suggests directions for future research to further optimize the integration of GeoAI in heritage tasks.

Keywords: Geospatial Artificial Intelligence (GeoAI), built heritage, Convolutional Neural Networks (CNN), Geographic Information Science (GIS)

Acknowledgements

As I approach the conclusion of my thesis, it marks the nearing end of my two-year journey at the University of Waterloo, where I pursued a Master of Science in Geography. Looking back, I stand on the brink of 2024, addressing myself as a recent 2022 graduate, who was then torn between entering the industry and accepting an offer to explore areas of interest as a graduate student. I have no regrets about my decision.

I owe my deepest gratitude to my supervisor, Dr. Rob Feick, whose guidance was indispensable throughout the process of writing my thesis. Dr. Feick's excellent mentorship illuminated my path whenever confusion clouded my direction, allowing me to proceed correctly and complete my thesis successfully. His insights significantly enhanced my research and academic writing skills. I extend sincere thanks to my committee members, Dr. Peter Deadman and Dr. Majid Hojati, for their valuable suggestions and for dedicating time to review my thesis repeatedly. Their contributions have been pivotal in refining my research.

My research in the second year of my graduate studies was greatly supported by the Earth Observation to Decision Information Services (DOTS) program. I am thankful for the funding they provided, which alleviated concerns about tuition and living expenses and allowed me to focus on my research. The training sessions, symposiums, and workshops they organized not only facilitated valuable academic exchanges with peers but also introduced new techniques and ideas applicable to my thesis.

On a more personal note, I would like to express my appreciation to Xuan Wang for his tremendous mental support and encouragement throughout my thesis writing process. A heartfelt thank you to my parents for their financial support, patience, and love during my studies. I am also grateful to all my friends and relatives who lightened my spirits and offered relief during challenging times in my thesis journey. Lastly, I cannot forget to thank my two pet rabbits, Charcoal and Mochi, whose company brought me much joy while I worked on my thesis.

In conclusion, I hope this thesis does not represent the end of my exploration in the fields of my interest but rather the beginning of a new chapter. My life in research and discovery is far from over, and I look forward to the opportunities and challenges that lie ahead.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

CNN	Convolutional Neural Network
DCNN	Deep Convolutional Neural Network
DL	Deep Learning
FT	Fine-Tuning / Fine-Tuned
GeoAI	Geospatial Artificial Intelligence
GIS	Geographic Information Science
HCD	Heritage Conservation District
MLP	Multi-layer Perceptron
ML	Machine Learning
OHA	Ontario Heritage Act
ROC	Receiver Operating Characteristic
RUC	Relative Uncertainty Coefficient
UNESCO	United Nations Educational, Scientific and Cultural Organization

Chapter 1
Introduction

1.1 Background and Problem Statement

Built heritage refers to entities which range from individual buildings to entire complexes of structures, monuments, installations, and even historical remnants which hold significance due to their intricate ties with the architectural, cultural, social, political, economic, and military narratives of history (Wicklow County Council, 2010). As defined by the World Heritage Convention (UNESCO, 1972), cultural heritage comprises monuments, building groups, and sites that possess outstanding universal value from historical, artistic, or scientific perspectives. Serving as a cornerstone of multiculturalism, built heritage stands as a witness in the evolution of human civilization and is a key symbol of urban culture, societal progress, and economic development. They are not only renowned for their architectural excellence but also recognized for their roles as cultural signifiers (Costa & Carneiro, 2021), drivers of economic growth (Albu, 2021), valuable educational tools (Lapadula & Quiroga, 2012), and proponents of environmental sustainability through their conservation and adaptive reuse (Sigmund, 2016). By encapsulating diverse design philosophies, historical narratives, and socio-cultural evolutions, built heritage provides deep insights into our collective heritage.

Growing awareness about the importance of protecting heritage has prompted the use of 'Designation,' which is a formal action in legal and planning communities that publicly recognizes the significance of heritage properties, typically established through the adoption of a municipal by-law aimed at preservation (Heritage Brockville, n.d.). It is also commonly known as "Heritage Designation" or "Heritage Protection." (Global Heritage Fund, 2010). Globally, countries including the UK, the US, China, and Canada have adopted a variety of legislative actions aimed at safeguarding their unique cultural heritage. For example, the Ontario Heritage Act (OHA) in Canada, initially published on March 5, 1975, empowers municipal governments, citizens, and the Minister of Culture and Multiculturalism to designate properties of cultural heritage value or interest. Recognizing the significance of heritage sites to the collective memory of all Ontarians, the OHA mandates the province to maintain formal records of all properties and heritage conservation districts designated under the act, as well as any other properties the Minister deems to have cultural heritage value or interest (Ontario Heritage Trust, n.d.-b).

Despite the attention paid to heritage conservation, the identification of heritage architectures is still a challenging task. This process demands a detailed evaluation of each building's historical values, architectural values, and environmental contexts (City of Markham, 1991). Normally, this evaluation has depended on the professional knowledge of domain experts (Al-Sakkaf et al., 2020). However, as many cities increase the pace and density of urban development unrecognized historic structures may be demolished or inappropriately altered (Wang et al., 2019).

Considering these difficulties, leveraging advanced technological solutions emerges as a possible strategy to aid building classification and possible heritage recognition. Artificial Intelligence (AI), a technology that enables computers to perform tasks typically requiring human intelligence (IBM, 2021), is promising since it can learn from vast amounts of data, identify patterns, make predictions, and automate data classification processes.

Geospatial Artificial Intelligence (GeoAI) is a term applied to the integration of AI with Geographic Information Science (GIS). This burgeoning field represents a synthesis of spatial science with the analytical prowess of AI, data mining capabilities, and the processing power of high-performance computing (Kamel Boulos et al., 2019). By leveraging this powerful combination, GeoAI excels at mining information from the extensive datasets of spatial big data. Its introduction marks a shift towards employing sophisticated algorithms and ML techniques to decode spatial complexities, decreasing reliance on manual labour and enhancing operational efficiency (Li & Hsu, 2022). This shift towards automated and intelligent analysis of spatial information represents a transformative way to manage and interpret our global spatial environment. It aligns with the continuous efforts to protect and preserve architectural heritage in our rapidly changing urban landscapes.

Convolutional Neural Networks (CNNs) are a specialized branch of deep learning (DL) that are highly effective at extracting detailed features from image data and are particularly adept at identifying semantic (meaning and context) and structural information (arrangement and relationships) within image sources (Rawat & Wang, 2017). This ability has received broad recognition and success in the domain of computer vision to classify remote sensing imagery, facilitate analysis and feature extraction (Song et al., 2019). The capability of CNNs to identify complex patterns and analyze visual data, particularly in differentiating historically significant features from the more commonplace, has been demonstrated to be beneficial (Z. Li et al., 2024).

As a branch of DL, CNNs are proficient in extracting detailed information from image data and identifying subtle semantic and structural information (Li & Hsu, 2022).A growing number of researchers have turned to CNNs to analyze the complexities of street-level imagery to classify diverse architectural styles and extract features (Kang et al., 2018; B. Wang et al., 2023). This precise identification by CNNs plays a vital role in elevating our appreciation and understanding of architectural heritage, offering a more nuanced view of our built environment.

While existing studies have highlighted the potential of AI, particularly the use of CNN, in classifying buildings, there are still limitations. A shortcoming of current models is their dependence on datasets that may not fully capture the architectural variety specific to a study area, along with an absence of mechanisms for directly evaluating a building's potential as built heritage. Moreover, the current models often give precedence to the external appearance of buildings, overlooking the historical and cultural importance that characterizes heritage buildings. To address these issues, this research focuses on Ontario, introduces approaches by developing : an Architectural Style Model and a Heritage Identification Model. The Architectural Style Model is designed to precisely identify the architectural style of buildings from images, and the Heritage Identification Model ascertains their heritage status, thereby indicating if they qualify as potential built heritage. This initial step is vital as it establishes a foundation for appreciating the visual and historical elements that contribute to a building's heritage value.

The study then progresses to formulate a Designation Status Prediction Model that utilizes the combined strengths of CNN and Multilayer Perceptron (MLP). This approach is underpinned by a detailed geospatial dataset, encompassing street view imagery, archival photographs, and geospatial information tailored to Stratford, which trains the model to recognize the unique heritage architectural features found in southern Ontario. The CNN component is responsible for analyzing image data to identify visual features that signify heritage value, such as specific architectural details and styles characteristic of the region. Simultaneously, the MLP component integrates geospatial data, concentrating on the geographical distribution of both designated and non-designated properties. This amalgamation of numerical and categorical data is intended to boost the model's predictive accuracy, offering a refined perspective on a building's potential heritage value based on its location and the context of surrounding heritage sites. This method of incorporating additional spatial datasets is aimed at improving classification accuracy and enhancing the ability to predict which buildings in urban settings might be recognized as heritage structures. By supplying data-driven insights to government bodies and heritage committees, this model aids in a more efficient and thorough identification of buildings that hold potential heritage value.

To conclude, this thesis addresses the prevailing challenges in heritage planning and promotes a deeper, more nuanced preservation of our architectural heritage. By reducing the risk of losing valuable heritage assets, it advocates for a more inclusive and effective heritage management approach. Through this research, the integration of GeoAI, CNN, and MLP not only overcomes technical hurdles but also deepens the heritage conservation discourse. It introduces a broader understanding of the architectural and cultural significance of buildings, enhancing our appreciation and protection of these treasures.

1.2 Motivation

The inspiration for this thesis stems from significant legislative changes introduced by recent amendments to the OHA through Bill 23, the More Homes Built Faster Act, 2022 (Government of Ontario, 2022). These changes have dramatically altered the landscape for the management and recognition of heritage properties. Previously, properties could be listed indefinitely as non-designated (not legally protected as heritage buildings) on the heritage register, allowing municipalities ample time to consider their designation status. However, the new legislation now limits this listing to a two-year period, after which a decision regarding their designation must be made (City of Toronto, 2023). This change imposes a strict timeline for the re-evaluation of currently listed properties and mandates a proactive approach to identifying and designating heritage properties.

The implications of these amendments are wide-ranging. First of all, there is an increased pressure on municipalities to speed up the process of identifying potential heritage properties due to the approaching deadline of January 1, 2025. This urgency is compounded by a stipulation that once a property is removed from the register, it cannot be re-listed for five years, potentially leaving significant heritage assets unprotected. Moreover, the decision-making process within municipalities regarding which properties deserve official designation is now constrained by critical time pressures, necessitating more efficient and accurate methodologies for heritage identification.

This study addresses the need for innovative solutions to tackle the challenges brought about by these legislative changes and the preexisting need to improve building classification methods. The adoption of GeoAI and mixed-data approaches can offer a new and potentially labour-saving method to classify built heritage. Leveraging advanced algorithms and diverse data sets, the aim is to significantly test methods that may enhance the efficiency and accuracy of identifying buildings with heritage value. Furthermore, this research highlights how municipalities can use computer vision methods as a scalable and dynamic tool for heritage property identification. This is particularly relevant in Ontario, where municipalities maintain extensive heritage registers that include many more properties than those officially designated as heritage sites. Cities such as Ottawa, Toronto, London, Hamilton, and Guelph have long lists that far exceed the number of properties formally recognized as heritage sites (Polowin & Polowin, n.d.).

In summary, this research stands at the intersection of heritage conservation and technological innovation, offering a timely and forward-looking approach to the classification of heritage properties. Its significance is highlighted by its potential to influence policy, guide municipal decision-making, guide municipal decision-making, and contribute to the sustainable preservation of cultural heritage in Ontario and beyond.

1.3 Research Questions and Objectives

This thesis is motivated by significant legislative changes in the management of heritage properties in Ontario, which necessitate the adoption of advanced methodologies for their identification and designation. The study explores the use of GeoAI and mixed-data approaches, aiming to enhance both the efficiency and effectiveness of the heritage designation process.

Central Research Question:

- How can GeoAI technology accelerate and optimize the current heritage designation process in Ontario?

To provide detailed insights into this overarching question, the research is structured around two specific sub-questions:

1. In what ways can GeoAI models assist in the initial screening stage of the heritage designation process? Additionally, how accurately and efficiently can these models identify architectural styles and investigate potential heritage buildings?
2. How does the integration of archival photographs and geospatial information enhance the efficiency of the evaluation of heritage designation process?

Based on the research questions, this research seeks to achieve the following research objectives:

a. To develop an Architectural Style Model that accurately identifies the architectural style of buildings from images, employing well-known and commonly used CNN algorithms for comparative analysis and optimization.
b. To create a Heritage Identification Model that ascertains the heritage status of buildings using different advanced DL techniques
c. To create a Designation Status Prediction Model that integrates visual features from CNN-processed image data with geospatial information through an MLP, enhancing the prediction of heritage designation by incorporating both architectural and location-specific data.
d. To evaluate the effectiveness and accuracy of the proposed models in classifying architectural styles and predicting heritage status through rigorous testing and validation processes.

1.4 Thesis Structure

This thesis is organized into five chapters, each addressing a distinct aspect of the integration of GeoAI in heritage conservation. The next chapter presents a review of existing literature related to heritage designation and GeoAI. It covers global and regional legal frameworks and protective measures, specific protective plans for heritage properties, the heritage designation process in Ontario, including Ontario evaluation criteria and identified gaps. This chapter also delves into Ontario's architectural styles and examines the styles that are more likely to be classified as heritage. Advanced GeoAI techniques in architectural and heritage studies are also introduced, with a focus on deep CNNs for image classification and applications of hybrid CNN-MLP models and mixed data inputs.

Chapter 3 describes the data and the methodologies employed in the study. It includes an overview of the study areas, the use of TensorFlow and Keras frameworks, and the development of three main models for this research: DCNN models for identifying architectural styles and built heritage in Ontario, and hybrid MLP-CNN model for predicting heritage property designation. The case study on Stratford, including the datasets used and the specific models developed, as discussed as well.

The Result and Discussion chapter presents and analyzes the models' performance followed by a discussion of the implications of the findings for model use in heritage conservation practice. The final chapter summarizes the key findings and contributions of the thesis by revisiting the research questions given in the first chapter. It reflects on how the study advances our understanding and management of heritage properties through the use of advanced technologies like GeoAI. The chapter also outlines theoretical and practical implications, ethical consideration and social impacts, as well as limitations and potential areas for future research.

Chapter 2
Literature Review

2.1 Chapter Overview

This chapter provides a comprehensive literature review to lay the foundation for understanding the thesis. It starts by analyzing global and regional legal frameworks for protecting built heritage, focusing specifically on Ontario's conservation framework, legal regulations, designation process, and evaluation criteria. The review identifies gaps in Ontario's heritage designation process and analyzes architectural styles prevalent in Ontario, including a statistical analysis of Stratford's styles. This exploration provides context for why this research topic and specific research questions were chosen and sets the groundwork for the solutions detailed in Chapter 3. The latter part of the chapter introduces the innovative application of GeoAI technology in heritage planning, exploring DL algorithms and architectures essential for image classification and their role in identifying heritage buildings. It also demonstrates how AI and mixed data can enhance the efficiency and accuracy of heritage conservation, while also providing an overview of the knowledge necessary for understanding the methodology explained in Chapter 3, ensuring that readers from various backgrounds can follow the thesis with ease.

2.2 Heritage Planning

2.2.1 Global and Regional Legal Frameworks and Protective Measures for Built Heritage

Over the past century, the recognition of the value of built heritage has evolved, which led the public to value historic structures more as they should be preserved. The Venice Charter (1964) initially provides a detailed definition of Architectural Heritage, stating that the concept of historical monuments encompasses not only individual architectural works but also urban or rural settings where evidence of a particular civilization, significant development, or historical event can be found. This applies to great works of art and more ordinary works that have indicated better cultural significance over time (International Charter for the Conservation and Restoration of Monuments and Sites (The Venice Charter), 1964). Subsequently, with an intensified public awareness of architectural heritage, The Convention Concerning the Protection of the World Cultural and Natural Heritage, established in 1972 by UNESCO, defined what constitutes architectural heritage with a general but more straightforward description. It included buildings, monuments, and ensembles with outstanding universal value from historical, artistic, or scientific perspectives (Convention Concerning the Protection of the World Cultural and Natural Heritage, 1972). The in-depth exploration of heritage preservation and reuse methods has led to a more profound research of heritage values, as conducted by David Throsby (2001), who identified various heritage values such as aesthetic, historical, spiritual, social, and symbolic (Throsby, 2001).

People recognized the increasing threats to architectural heritage and acknowledged a human responsibility to transmit its authenticity and richness to future generations, thereby creating numerous laws and regulations. The 1972 Convention Concerning the Protection of the World Cultural and Natural Heritage (UNESCO Convention) emphasizes that it is the primary responsibility of each signatory country to identify, protect, conserve, present, and transmit to future generations the cultural heritage outlined within its territories; significant resources must be committed, and when necessary, assistance from other countries must be sought. (Convention Concerning the Protection of the World Cultural and Natural Heritage, 1972). Following this, the European Charter of the Architectural Heritage (1975) attached the importance of historical buildings and towns, introducing policies for holistic urban conservation. The Amsterdam Declaration of the same year further stipulated that the success of any holistic conservation policy should depend on incorporating social factors, meaning that a conservation policy also implies integrating architectural heritage into social life (European Charter of the Architectural Heritage, 1975; Amsterdam Declaration, 1975).

The fundamental tenets of the Athens Charter were incorporated into the Venice Charter of 1964, which updated and revised the guidelines for preserving historic buildings and environments. It strongly emphasizes the comprehensive and integral conservation of historical monuments, which covers entire urban or rural settings in addition to individual buildings. In order to preserve monuments' value as historical artifacts and works of art, the charter promotes the use of all scientific and technological methods for their conservation and restoration. Permanent maintenance is advocated, and any changes to the structure or decoration during use are proscribed to preserve their cultural and historical characteristics (International Charter for the Conservation and Restoration of Monuments and Sites (The Venice Charter), 1964).

Moreover, the Florence Charter, which ICOMOS registered in 1982, was first created to protect historic gardens; however, its strategies and guiding principles were later expanded to involve the conservation of architectural heritage. The 2011 Paris Charter highlights the significance of heritage in promoting social cohesion, well-being, creativity, economic attraction, and intercommunal understanding. It is directed towards heritage preservation, development, and tourism stakeholders. The Paris Declaration stresses the economic, social, and cultural values of heritage, which are pivotal to community development. It promotes balanced urban development strategies and planning to guarantee suitable activity zones to protect historic neighbourhoods and support their rehabilitation and revitalization (Paris Charter, 2011).

Conventions and the implementation of plans for architectural heritage protection are widely accepted worldwide. China has made notable efforts to well-structured heritage protection practices. The Principles for the Conservation of Heritage Sites in China (2015) specify a detailed protection procedure, including heritage surveys, assessment, designation of protection levels, drafting of protection plans, implementation, and regular inspection. In a similar vein, Japan's architectural heritage has been continuously registered, maintained, and repaired, with over 4,800 buildings designated as important cultural properties, including approximately 280 buildings of special cultural significance designated as National Treasures. Furthermore, the preservation and restoration of architectural heritage in Japan have developed into a highly organized and professional practice that guarantees the preservation of the conservation work's structural integrity and cultural authenticity (The Japanese Association for Conservation of Architectural Monuments, n.d.).

In conclusion, this section has traced the evolution of architectural heritage conservation, highlighting the expanded understanding and appreciation of heritage that now includes broader urban and rural environments. It has emphasized the significant global efforts to protect architectural heritage through detailed legal frameworks and international conventions. These collective efforts emphaze that heritage plays in sustainable development and show a strong worldwide commitment to protecting our architectural legacy for the coming generations.

2.2.2 Heritage Conservation Framework in Ontario: Legal Regulations, Designation Process, and Evaluation Criteria

2.2.2.1 Legal Regulations in Ontario

In Ontario, heritage protection adheres precisely to international agreements by utilizing a well-established legislative structure. The combination of OHA (Government of Ontario, 1990) and the Planning Act (Government of Ontario, 1990) creates a comprehensive system for conserving heritage within the province.

Heritage Conservation Districts (HCDs) are mostly protected under the OHA. It stipulates that any alterations within these districts require a heritage permit from the relevant city planning department. This process promotes not only public participation but also awareness of cultural conservation while simultaneously protecting heritage treasures. Sections 27 and 41 of the Act elaborate on more key provisions, such as the requirement for heritage permits and the responsibility of local governments in supervising heritage protection (Government of Ontario, 1990).

The Planning Act provides a solid framework for land use planning that expands the OHA. It highlights the value of architectural heritage resources and describes decision-makers' responsibilities in land use, conflict resolution, and public participation (Government of Ontario, 1990). This ensures the smooth integration of historical structures and cultural legacies into modern urban development, striking a balance between preservation and the demands of urban growth.

To support the implementation of these legislative measures, the Ministry of Heritage, Sport, Tourism and Culture Industries in Ontario has established eight guiding principles for the conservation of built heritage properties. These principles are derived from a century of international charters and stress maintaining buildings' integrity and cultural heritage value. They advocate for minimal intervention, reversible alterations, and a clear differentiation between old and new elements, ensuring that heritage properties are preserved responsibly and sustainably for future generations (Government of Ontario, n.d.-a).

2.2.2.2 Ontario Designation Process

Heritage designation is a method by which municipalities acknowledge properties that possess cultural heritage value or interest, and it also serves as an approach for property owners to demonstrate their pride in the heritage significance of their properties (Government of Ontario, n.d.-

b). The OHA shows that heritage designation not only contributes to society but also serves individual interests. From a government or societal perspective, the Act enables municipalities to legally preserve properties and districts of historical significance. Sections 29 and 41 of the OHA empower municipalities to designate individual properties and heritage conservation districts, respectively (Government of Ontario, 1990). This legal framework ensures the protection of unique community characters and cultural heritage, which can foster local pride and contribute to cultural tourism, such as walking tours and Doors Open Ontario festivals. Additionally, municipalities can use designation to prevent the demolition of heritage properties by neglect, using heritage property standards bylaws (Ontario Heritage Trust, n.d.-a). Overall, the designation is a cost-effective method compared to expropriation or purchase, allowing for the preservation of various real properties, including buildings, landscapes, and natural features.

Property owners' benefits of designation under the OHA are substantial, although there is no compensation (Ontario Heritage Trust, n.d.-a). The designation provides property owners access to various heritage incentives that offset maintenance costs and increase property value. These incentives include grants, planning incentives, and property tax rebates, which make investing in and maintaining heritage properties more appealing. Combined with local commemorative, interpretive, and educational programs, designation not only enhances the value of the properties but also promotes a deeper connection and pride in one's community heritage (Ontario Heritage Trust, n.d.-a). This dual benefit framework reinforces the importance of heritage preservation both for enhancing community identity and for the personal benefit of property owners.

To better manage Ontario heritage buildings, Section 27 of the OHA mandates that each municipality maintain a Municipal Heritage Register that is open to the public (Government of Ontario, 1990). Generally, municipalities must list all properties designated under Parts IV or V of the OHA in the municipal register (Government of Ontario, 1990). For a property that has not been designated, the municipal council must recognize the property as having cultural heritage value or interest and provide a sufficient description to identify the property, such as its street address (Government of Ontario, 2016).

The process of identifying potential heritage properties for inclusion in a municipal register under the OHA typically involves a combination of municipal initiatives and inputs from property owners and the community. This collaborative approach ensures that property values or cultural heritage interests are recognized proactively by government entities and through community participation. Initially, municipalities play a proactive role in discovering potential heritage properties. Criteria can be set out in Ontario Regulation 9/06 for Determining Cultural Heritage Value or Interest to guide their decisions. While comprehensive research and evaluation of each property are not required at this stage, municipalities are advised to provide a brief rationale that explains why they believe a property qualifies as having heritage value. Although consulting with property owners or the public is not mandatory before adding non-designated properties to the register, municipalities often notify property owners and may engage the community, enhancing transparency and public involvement (Government of Ontario, 1990).

In addition to government initiatives, property owners and community members contribute to identifying potential heritage properties. For instance, in the Town of New Tecumseth, community members, town staff, or members of the Heritage Advisory Committee are encouraged to recommend buildings for heritage registration and possibly for designation (The Town of New Tecumseth, 2019). Moreover, municipalities like the City of Kenora actively involve the public in these decisions through forums that discuss which properties should be included in the register (Government of Ontario, 2016). This process underscores significant community engagement in heritage matters, allowing various perspectives to influence heritage conservation efforts. To ensure a request for designation is processed correctly, a completed and signed application form must be submitted along with an aerial photograph or location map and photographs (both historical and/or current) of the building’s exterior, landscape features, and any interior features relevant to the designation application (City of Ottowa, n.d.). As a result, this combination of municipal proactive discovery and community or owner submissions creates a comprehensive and inclusive approach to identifying and conserving heritage properties, ensuring that both the historical significance and community value of these assets are recognized and preserved.

Under Section 29 of the OHA, the designation process involves a structured sequence to ensure properties of cultural heritage value or interest are appropriately recognized and protected (Government of Ontario, 1990). A detailed flowchart to show the entire process for designation is shown in Figure 2-1.

The initial step in the designation process requires the municipality to issue a notice of intention to designate, which must be served on the property owner and published in a newspaper with general circulation within the municipality. This notice includes a description of the property, a statement of its cultural heritage value or interest, and an explanation of the heritage attributes that justify the designation. Furthermore, it informs that objections to the designation can be submitted to the municipal clerk within 30 days of the notice's publication (Government of Ontario, 1990).

If objections are raised, the municipal council is obliged to review and make a decision on the designation within 90 days following the objection period. If the council decides to proceed with the designation despite objections, it may pass a by-law formally designating the property. This by-law must include a detailed statement of the property's cultural heritage value and a description of its heritage attributes. Once passed, the designation by-law is served on the property owner and published, triggering another 30-day period during which appeals can be made to the Tribunal (Government of Ontario, 1990).

Should there be no objections or appeals, or if all appeals are resolved in favour of designation, the by-law comes into force, and the property is formally recognized as a designated heritage property. The municipality must then register the by-law against the property in the appropriate land registry office to ensure future compliance and preservation under the new status (Government of Ontario, 1990).

This thorough process is a balanced approach to heritage conservation, accommodating public and private interests while providing mechanisms for objection and appeal to ensure fairness in the designation process.

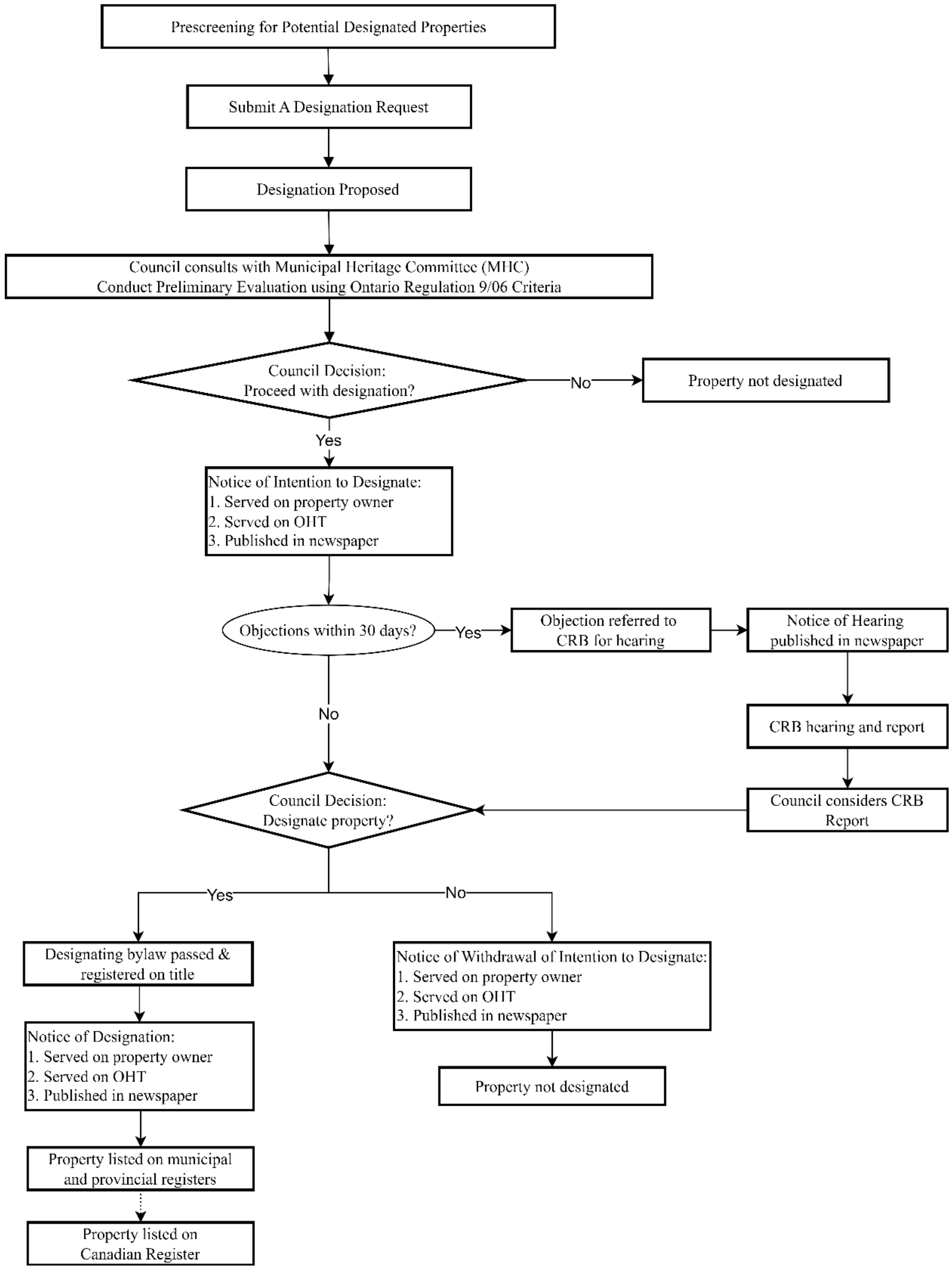

Figure 2-1 Flowchart of Ontario Designation Process (adapted from Ontario Heritage Act)

2.2.2.3 Evaluation Criteria

The Act facilitates the individual designation of properties, empowering City Councils to recognize and protect sites of cultural heritage value. This ensures their preservation for both current and future generations. The criteria for individual designation are detailed in Ontario Regulation 9/06, which includes considerations for design or physical value, historical or associative value, and contextual value. A property may be designated based on its rare, unique, or early style, type, or construction method or for showcasing notable craftsmanship, artistic merit, or technical achievement.

By examining the pivotal document issued by the Heritage Section of the Planning and Urban Design Department of Markham in Evaluating Heritage Resources in the Town of Markham (1991), a standardized system for assessing and categorizing buildings of potential architectural and/or historical significance is detailed expressed. Comparable methodologies employed by the cities of Toronto, Vaughan, and Mississauga were also reviewed to ensure a comprehensive understanding of evaluation practices across different jurisdictions.

To minimize personal bias and ensure a thorough and objective assessment, it is suggested that the evaluation and scoring need to be conducted by more than one individual (City of Markham, 1991). The evaluation criteria employed aim to encapsulate the attributes that underscore the heritage significance of each building and are divided into three primary categories:

- Historical Value is assessed based on the building's age, its association with notable figures or events, and its capacity to exemplify cultural, social, political, military, industrial, or agricultural history trends. Its educational value in illustrating cultural history and its potential for tourism promotion are also considered (City of Markham, 1991).
- Architectural Value evaluates stylistic purity or rarity, design and craftsmanship quality, the significance of the architect or builder, structural condition, and the preservation state or integrity of the building (City of Markham, 1991).
- Environmental Context examines the building's contribution to the identity of the community or landscape, including its design compatibility with the streetscape or surroundings, community context, landmark status, and site characteristics (City of Markham, 1991).

Detailed criteria are shown in Table:

Table 2-1 Criteria for Evaluating Heritage Resources in Markham

Category	Criteria	Description
Historical Value	**Date of Construction**	Older structures that reflect significant historical periods should be acknowledged to determine the potential historical significance based on the building's age, even if not directly linked to a specific event or person.
	Association with Historic Trends/Patterns/Themes	Evaluates whether the building reflects specific social, economic, political, or cultural patterns characteristic of Markham's history or that of a potential heritage conservation district.
	Association with Events	Considers buildings linked to significant local, regional, or national events, emphasizing the long-term consequences of such events on the community.
	Association with a Person or Group	It focuses on buildings linked to individuals, groups, institutions, or corporations significantly contributing to local, provincial, or national history.
	Archaeological Resources (Bonus)	Assesses the presence of archaeological sites that might provide valuable historical information about the property or community.
	Historic Grouping (Bonus)	Considers buildings that are part of a historically associated group of buildings, which illustrate important trends or patterns in the community.
Architectural Value	**Design**	Evaluate the building's design quality, considering its artistic merit, uniqueness, and how well it fits the local architectural context, including any alterations that might affect its original qualities.
	Building Style	Assesses the building's architectural style, comparing it to other buildings of the same style in the area to understand its stylistic significance.

	Architectural Integrity	It looks at how well the building's original stylistic elements have been preserved through alterations or additions.
	Physical Condition	Considers the overall structural condition of the building, evaluating the extent of necessary repairs.
	Designer/Builder	Recognizes buildings designed or constructed by notable architects, engineers, or builders significant to local, regional, or national history.
	Interior Elements	(Bonus) Assesses the historical significance, attractiveness, or uniqueness of the building's original interior elements.
Environmental Context	**Design Compatibility with Streetscape/Environs**	Evaluate how well the building fits with its surrounding environment, including built and natural landscapes.
	Community Context	Considers the building's role and symbolic value within the community, reflecting its historical and ongoing contributions to community life.
	Landmark Status	Assesses the building's role as a cultural or historical landmark within the community, noting its visibility and recognizability.
	Site	Evaluate whether the building remains on its original site and how well its layout has been preserved or altered.

Note. Adapted from "Evaluating Heritage Resources in the Town of Markham," by City of Markham, 1991.

2.2.2.4 Designated vs. Non-Designated Properties

After the evaluation, the Municipal Register of Cultural Heritage Properties in Ontario was designed to create two significant lists, designated and non-designated properties. By definition, designated properties are those that have been formally recognized under the OHA. This list has proved and can provide detailed information about each property's historical, architectural, or cultural significance and ensures the properties are legally protected against alterations that could compromise their integrity (City of Guelph, n.d.).

On the other hand, although non-designated properties may not have formal protection, they are still recognized for their cultural heritage value or interest. Listing these properties is a preliminary safeguard requiring property owners to give the city 60 days' notice before demolishing or altering these structures. A non-designated properties list is crucial for the government with sufficient time to re-evaluate the property's heritage value and re-consider whether it warrants full designation to ensure long-term preservation (City of Guelph, n.d.).

As a result, this approach in the Municipal Register allows for a dynamic and responsive heritage conservation strategy, ensuring that both designated and potentially significant non-designated properties are considered and protected in the face of urban development and change.

2.2.3 Gaps in Heritage Designation Process in Ontario

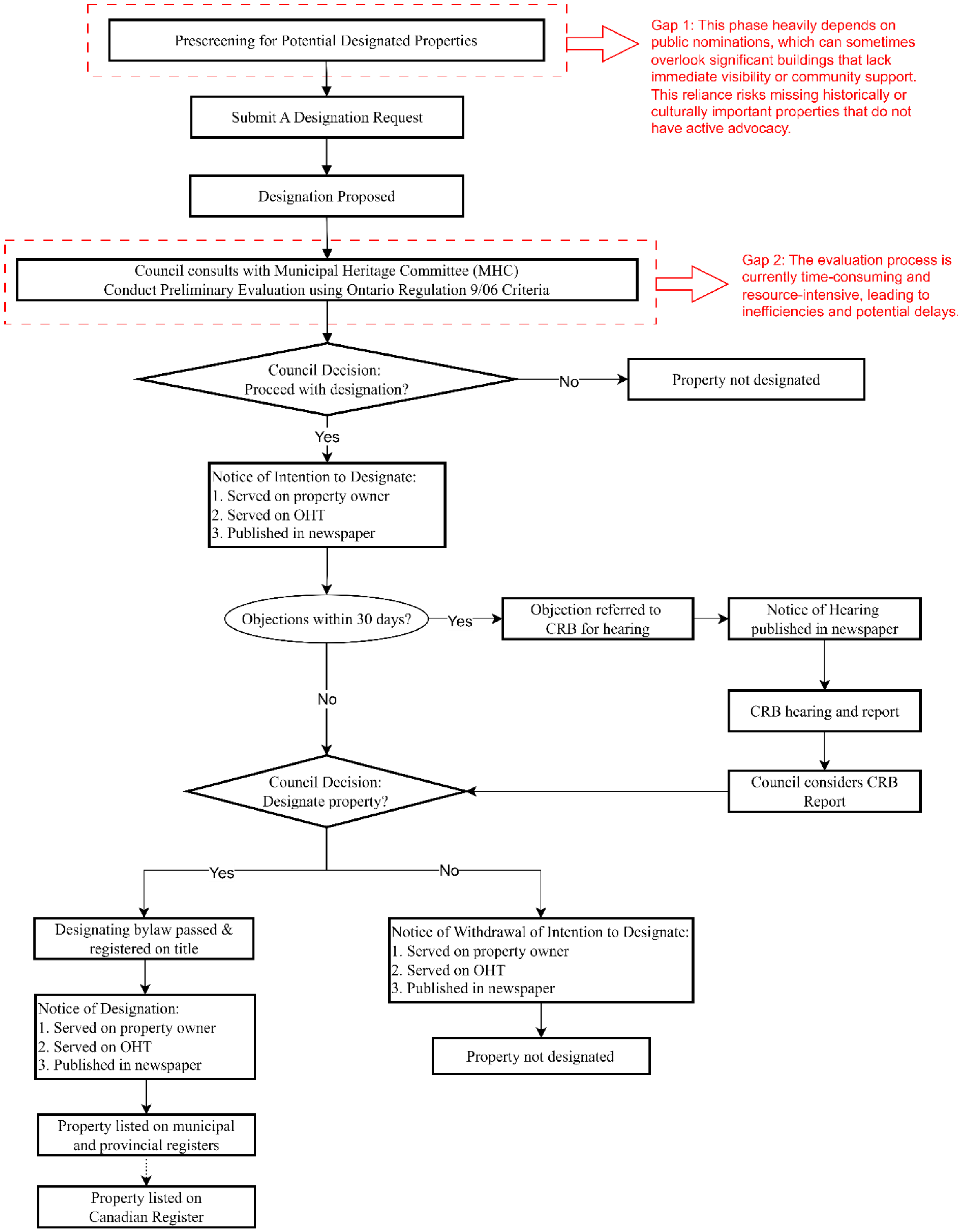

Figure 2-2 Gaps in Ontario Heritage Designation Process

Ontario’s heritage designation process has great contributions to Ontario’s heritage protection, but it still has several gaps. Figure 2-2 shows the gaps in the Ontario Heritage Designation Process, explicitly highlighting the steps where these gaps occur. A significant gap in the heritage designation process is prescreening potential heritage buildings. Current methods depend on discoveries made by local governments or municipal heritage committees during urban planning, historical research, or suggestions from the public, historians, or local history and heritage organizations (The Town of New Tecumseth, 2019). This method relies heavily on unsolicited nominations and may overlook essential buildings that lack community support or visibility, potentially missing buildings of significant historical or cultural importance.

Further gaps become apparent during the “Conduct Preliminary Evaluation using Ontario Regulation 9/06 Criteria” phase. The identification and assessment of a building’s heritage value depend heavily on significant human involvement and labour, leading to inefficiencies and potential inconsistencies in the process. This manual evaluation process is time-consuming and labour-intensive, which strains resources, especially under the new legislative pressures of Bill 23 that mandate shorter time frames for heritage designations without any corresponding increase in resources’ (City of Toronto, 2023, p. 23).

In response to these challenges, this research utilizing street view imagery and geospatial artificial intelligence (GeoAI) technology offers an innovative solution to aid in heritage designation. Through automated image recognition and geospatial data analysis, GeoAI can effectively screen and identify buildings with heritage value and judge whether the property is more likely to be designated or non-designated. This technology supports existing heritage assessment processes and provides data-driven decision support for governments and heritage committees, reducing the potential loss of heritage assets and enhancing the effectiveness and comprehensiveness of heritage management.

2.3 Ontario Architectural Styles

Understanding popular architectural styles in Ontario was crucial to creating the Ontario Architectural Style and Heritage Dataset. The Ontario Architectural Style Guide prepared by the University of Waterloo’s HPI Nominations Team is a valuable reference. This document offers a comprehensive overview of popular architectural styles in Ontario from the late 18th century to the early 20th century, detailing more than 20 styles. It provides in-depth descriptions of key features such as form, exterior, roof, and decorative elements, which are essential for accurately identifying and classifying Ontario's historic buildings. The detailed information is shown in Table 2-2.

Table 2-2 Chronological Overview of Architectural Styles in Ontario

Year	Name of Style	Main Characteristics
1780s - 1860s	Georgian	Rectangular and symmetrical, two to three storeys, hip or end gable roof, small-paned sash windows
1780s - 1860s	Log Houses and Structures	Gable-roofed, one-and-a-half-storey, lean-to roofs, plain horizontal log facade
1810 - 1850	Neoclassical	Box-like, symmetrical, refined detailing around doors and quoins, decorative pilasters
1820s - 1870s	Regency	Symmetrical and low, large windows often have wide verandahs
1820s - 1860s	Greek Revival	Symmetrical, temple shape, columns, rectangular form
1820s - 1900	Mennonite Georgian	Full-width front porches, attached "doddy houses" built into a hillside, plaster under porches
1840 - 1900	Romanesque Revival	Use of round-headed windows and arches, massive structure, often in stone
1840s - 1870s	Gothic Revival	Pointy, picturesque, steeply pitched roof, arched windows, decorative bargeboards
1840 - 1885	Italianate	Ornate, controlled design with heavy cornice brackets, paired windows
1860s - 1880s	Second Empire	Elaborate, mansard roof with dormer windows, similar to Italianate but more complex.
1880s - 1910s	Queen Anne	Irregular, busy, and ornate with complexity in detail often features a turret.
1880s - 1930s	Beaux Arts	Eclectic and classic, often with a temple-like façade and columns, found in public buildings.
1900s - 1940s	Tudor Revival	Medieval look, half-timbering on façades, steeply pitched roofs
1910 - 1940	Art Deco	Vertical, geometric with design motifs, use of stucco and smooth concrete
1930 - 1950	Art Moderne	Streamlined, wrap-around windows, horizontal lines, smooth and clean façade
1930s - 1960s	Modern/International	Horizontal lines, flat façades, and landscape orientation often feature glass and flat roofs.
1830 - 1900	Octagonal and Round	Unique shapes with distinctive roofing and windows used varied materials
1940 - 1960	Victory Housing	Simple, modest post-war housing, often rectangular with side-hall plans and steeply pitched gable roofs

1890s-1940s	Colonial/Georgian Revival	Simple, symmetrical suburban architecture, often with small dormers and elaborate downspouts
1890s-1940s	Late Gothic Revival	Evokes castles, forts, and churches, used frequently in large institutions, flat roofs with parapets
1900 - 1920	Edwardian	Simple, classical, balanced, precursor to simplified styles of the 20th century
1900s - 1930s	Prairie/Craftsman/Bungalow	Arts and crafts, horizontal emphasis, extensive verandas, overhanging roofs

Note: This table presents a concise overview of the leading architectural styles in Ontario, organized chronologically based on their first documented appearance. The architectural details outlined in this table are derived from a comprehensive guide by the Heritage Resource Centre, which provides an in-depth look at the evolution of architectural styles within the region. For further details on these styles and their historical significance, refer to the Heritage Resource Centre's Architectural Styles Guide (HPI Nomination Team & University of Waterloo, 2009).

2.3.1 Statistical Analysis of Stratford's Architectural Styles of Built Heritage

Table 2-3 Initial Counts of Architectural Styles in Stratford

Architectural Style	Count
Italianate	25
Gothic Revival	21
Queen Anne	18
Ontario Cottage	8
Vernacular	8
Georgian	5
Edwardian	4
Gothic Revival/Ontario Cottage	4
Italianate/Gothic Revival	3
Second Empire	3
Craftsman	2
Regency Cottage	2
Tudor Revival	2
Arena	1
Bandshell	1
Collegiate Gothic	1
Colonial Revival	1
Fieldstone Construction with Square Beaded Masonry	1

Flemish Renaissance	1
Georgian/Gothic Revival	1
Industrial	1
Italianate with Queen Anne elements	1
Italianate/Queen Anne	1
Italianate/Second Empire	1
Jacobethean Revival	1
Late Victorian	1
Neo-Classical	1
Prairie Style	1
Regency/Ontario Cottage	1
Romanesque Revival	1
Romanesque Revival/Queen Anne	1
Second Empire/Italianate	1
Stone Arch Bridge	1
Grand Total	125

Table 2-4 Adjusted Counts of Architectural Styles in Stratford

Rank	Architectural Style	Count
1	Italianate	32
2	Gothic Revival	30
3	Queen Anne	21
4	Regency (Ontario) Cottage	16
5	Vernacular	8
6	Georgian	6
7	Second Empire	5
8	Edwardian	4
9	Craftsman	2
9	Tudor Revival	2
9	Romanesque Revival	2
12	Arena	1
12	Bandshell	1
12	Colonial Revival	1
12	Fieldstone Construction with Square Beaded Masonry	1
12	Flemish Renaissance	1
12	Industrial	1
12	Jacobethean Revival	1
12	Late Victorian	1
12	Neo-Classical	1
12	Prairie Style	1

12	Stone Arch Bridge	1
	Grand Total	139

To effectively meet the research objectives concerning examining Ontario's built heritage, it is crucial to deeply investigate the prevalent architectural styles in the region. Stratford was selected as a case study to establish a foundational quantitative overview of architectural diversity due to the extensive data available. Information on both designated and non-designated properties was collected from Stratford's official website and summarized in Tables 3 and 4. The collection techniques are mainly manual extraction. A straightforward quantitative method to count the occurrences of each architectural style was applied, and its initial statistical results are displayed in Table 3.

The study also tackled the challenge posed by composite style descriptions frequently found in the data, such as "Italian style with Gothic Revival characteristics" or "Italian style/Second Empire." A method was developed to separate and individually count for these composite styles. For example, a property listed as "Gothic Revival/Ontario Cottage" was counted once for both Gothic Revival and Ontario Cottage. The results are presented in Table 4, which offers adjusted counts for each architectural style, thereby providing a clearer and more precise view of the diversity and distribution of architectural styles within the dataset. The data analysis reveals that the key styles of historical architecture in Stratford include Italianate, Gothic Revival, Queen Anne, Regency (Ontario) Cottage, Vernacular, Georgian, Second Empire, Edwardian, Craftsman, Tudor Revival, and Romanesque Revival. These styles are marked by their classical influence and traditional design elements, emphasizing the city's historical and architectural significance.

2.4 Advanced GeoAI Techniques for Multimodal Image Classification in Heritage Studies

2.4.1 AI and GeoAI

In recent years, AI has become a hot topic due to its wide applications spanning multiple disciplines and industries. However, traditional AI often falls short when dealing with spatial relationships and geographic contexts. To address these challenges, GeoAI emerged. Wenwen Li and Chia-Yu Hsu characterized GeoAI as research that seeks to enhance AI technologies for spatial analysis by utilizing geospatial methods and big data. As depicted in their research, GeoAI melds AI with geographic knowledge and spatial thinking to offer new capabilities for understanding and analysing the natural world and human society (Li & Hsu, 2022). Figure 2-3 shows GeoAI as an interdisciplinary field that merges artificial intelligence techniques with geography to analyze and interpret geospatial big data.

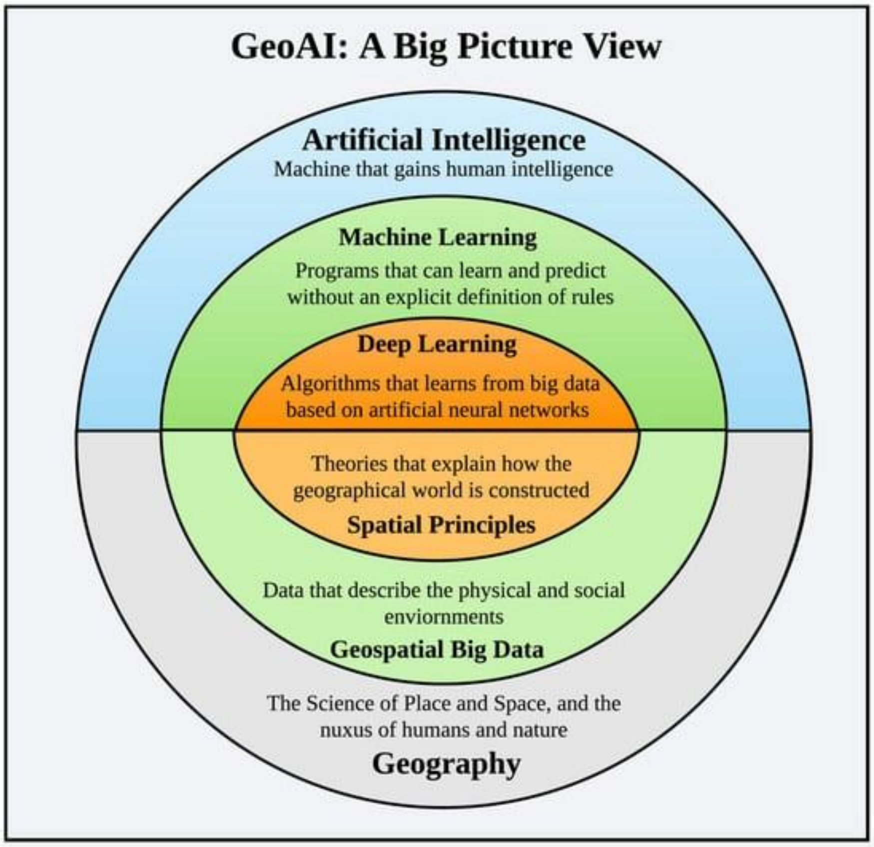

Figure 2-3 A Big Picture View of GeoAI (Li & Hsu, 2022)

Li and Hsu (2022) further highlight how GeoAI diverges from and connects with traditional AI. Unlike traditional AI, which typically handles structured and non-spatial data such as text, images, and numerical data for applications in fields like image recognition, natural language processing, and predictive analytics, GeoAI focuses on geospatial big data. This includes remote sensing imagery, geo-scientific data, street view images, and topographic maps. These data types often exhibit complex spatial dependencies and require the integration of GIS and spatial data analysis techniques(Li & Hsu, 2022).

GeoAI demonstrates significant advantages over traditional AI and standalone GIS. GeoAI methods are capable of processing large-scale spatial and temporal data, surpassing the limitations of traditional approaches and enhancing the robustness and applicability of GeoAI models. Furthermore, GeoAI facilitates automated feature extraction from raw data, and in some instances can reduce the need for manual definition of independent variables. This not only improves model accuracy but can also uncover previously unknown insights. The GeoAI models are also responsive to subtle variations and robust against data noise, marking a clear advancement over conventional analysis techniques (Li & Hsu, 2022). These capabilities enable GeoAI to significantly improve object detection, geographic information retrieval, and predictive analytics, delivering precise, actionable insights (Janowicz et al., 2020). However, GeoAI also encounters notable challenges, particularly its dependency on extensive, high-quality data which can be limited and inconsistently distributed. The intricate nature of urban social structures often necessitates spatially explicit models, potentially introducing biases in data quality and representation. The black-box nature of deep learning models used in GeoAI hinders transparency and complicates the interpretation of results, raising concerns about the validity and reliability of outcomes in urban geography applications (Liu & Biljecki, 2022). Ethical considerations are also crucial, demanding vigilance from both individual researchers and the broader community to ensure responsible GeoAI research (Janowicz et al., 2020).

In conclusion, while GeoAI presents transformative potential for geospatial analysis, it also demands ongoing refinement to address its current limitations and ethical implications. Future research should focus on enhancing data quality, improving model transparency, and ensuring equitable technology deployment. By addressing these challenges, GeoAI can more effectively harness its capabilities to offer sustainable solutions in geo-related studies and beyond, cultivating a deeper and more comprehensive understanding of geospatial phenomena.

2.4.2 Deep CNNs for Image Classification

AI refers to the capacity of machines to carry out cognitive tasks typically associated with human intelligence, including perception, reasoning, learning, environmental interaction, problem-solving, decision-making, and exhibiting creativity (Rai et al., 2019). The genesis of AI can be traced back to 1950 when Alan Turing published “Computing Machinery and Intelligence,” proposing the Turing Test to evaluate a machine’s ability to exhibit intelligent behavior indistinguishable from that of a human (Turing, 1980). This foundational work laid the groundwork for AI’s theoretical development.

Since the mid-20th century, AI has fascinated numerous scholars and technologists due to its renowned concepts and ideal development prospects. Through continuous exploration, AI has evolved significantly in past decades, transforming into a multifaceted field that drives advancements across diverse sectors. Presently, AI encompasses several critical subfields: Machine Learning (ML) enables systems to learn from data and make decisions autonomously through supervised, unsupervised, and reinforcement learning techniques. Neural Networks (NN), including CNNs and Recurrent Neural Networks (RNNs), mimic the human brain's structure to process complex information. DL, a subset of ML and NNs, is designed to utilize layered neural networks for high-dimensional data analysis. The sub-disciplines of Robotics, Computer Vision, and Natural Language Processing extend AI's applications to physical tasks, visual data interpretation, and human language understanding (Athanasopoulou et al., 2022). These AI technologies pave the way for new applications and innovations and significantly broaden the scope of what machines can achieve in assisting humans.

Image classification, defined as categorizing images into one of several predefined classes, has been one of the earliest and most popular research directions in computer vision. Traditionally, image classification relied on a two-stage approach to address classification problems. Feature descriptors were used to extract handcrafted features from images, which were then input into trainable classifiers (Rawat & Wang, 2017). However, the accuracy of this method heavily depended on the design of the feature extraction stage (Rawat & Wang, 2017), leading to poor generalization and portability (Chen et al., 2021).

The emergence of DL has significantly advanced this field in recent years. DL models utilizing multi-layered nonlinear information processing for feature extraction, transformation, and pattern analysis, have proven capable of overcoming these challenges (Rawat & Wang, 2017). Researchers have developed computer neural networks inspired by biological vision systems, leading to the advent of CNNs. CNNs have become the leading architecture for most image recognition, classification, and detection tasks (Chen et al., 2021; Rawat & Wang, 2017). Early CNN models, such as LeNet-5 developed by Lecun et al. in 1998 were limited by theoretical foundations and computational power, resulting in suboptimal recognition of complex images. The breakthrough by Hinton et al. (2006) then introduced an effective learning algorithm for deep convolutional neural networks (DCNNs), marking the beginning of a new era in DL.

Building on their foundational strengths, DCNNs have significantly advanced the field of image classification, consistently outperforming traditional methods across various domains. In their survey on remote sensing image classification using DCNNs, Song et al. (2019) outline several key advantages of DCNNs over shallow structure models. First, DCNNs apply convolution operations directly to image pixels to extract abstract features, offering powerful generalization capabilities across various scenarios. Second, DCNNs can represent image information distributedly, quickly processing large volumes of data and effectively handling complex nonlinear problems like image rotation and translation. Third, the architecture of DCNNs, characterized by sparse connections, weight sharing, and spatial subsampling, results in a simpler but more adaptable network structure. To better understand DCNN-based image classification, this section will introduce the structure and training methods of DCNNs, followed by an overview of several popular DCNN models in computer vision (Song et al., 2019).

Leveraging their advantages, DCNNs have demonstrated remarkable success across various domains for image classification, often surpassing traditional methods. DCNNs have revolutionized diagnostic imaging in the medical field by providing tools to detect anomalies such as tumours with high precision. For instance, according to research by Esteva et al. (2017), DCNNs have been used to identify skin cancer with dermatologist-level accuracy using a dataset of 2,032 diseases, significantly improving early diagnosis and treatment outcomes. In environmental and remote sensing applications, CNNs analyze satellite and aerial imagery to classify land cover and monitor environmental changes effectively. Compared to classical methods that rely on handcrafted features, DCNNs provide superior accuracy and efficiency in addressing remote sensing challenges, particularly those involving complex physical models or poorly understood phenomena that resist generalization (Zhu et al., 2017).

In social media content moderation, DCNNs have made significant strides in detecting and classifying inappropriate video content. The study by Yousaf and Nawaz (2022) proposed a novel DL-based architecture utilizing EfficientNet-B7 and BiLSTM networks to detect inappropriate content in YouTube videos. This architecture effectively classifies malicious uploads, such as animated videos with inappropriate content targeted at children.

In industrial applications, DCNNs play a crucial role in ensuring product quality and safety across various sectors. For instance, in the food packaging industry, CNNs are employed to perform real-time detection of contamination in heat-sealed food trays using hyperspectral imaging. This system captures detailed information about the trays and utilizes CNNs to identify and discard faulty products, achieving a detection accuracy of over 94% (Medus et al., 2021). Additionally, manufacturing automated optical inspection (AOI) systems leverage multi-stage CNN models for defect detection. These models identify and classify defects with high precision and pinpoint defect locations, thereby reducing error rates and labour requirements.

To conclude, the applications in diverse fields show the versatility and transformative impact of CNNs in advancing image classification technology. Their ability to autonomously extract features and adapt to various data scales makes them indispensable in academic research and practical applications.

2.4.2.1 Basic Structure of DCNNs

DCNNs are specialized multilayer perceptrons designed to identify two-dimensional shapes by mapping input images to desired outputs. Each neuron in a DCNN is connected to neurons in a local area of the previous layer, reducing the number of weights and enhancing efficiency. Like traditional NNs, DCNNs have a hierarchical structure comprising convolutional, pooling, fully connected layers, and an output layer (Song et al., 2019). In a comprehensive view, Song et al. (2019) detail that CNN architectures primarily include three fundamental layer types: convolutional, pooling, and fully connected layers. Understanding these components is crucial for comprehending how CNNs effectively process and analyze visual information.

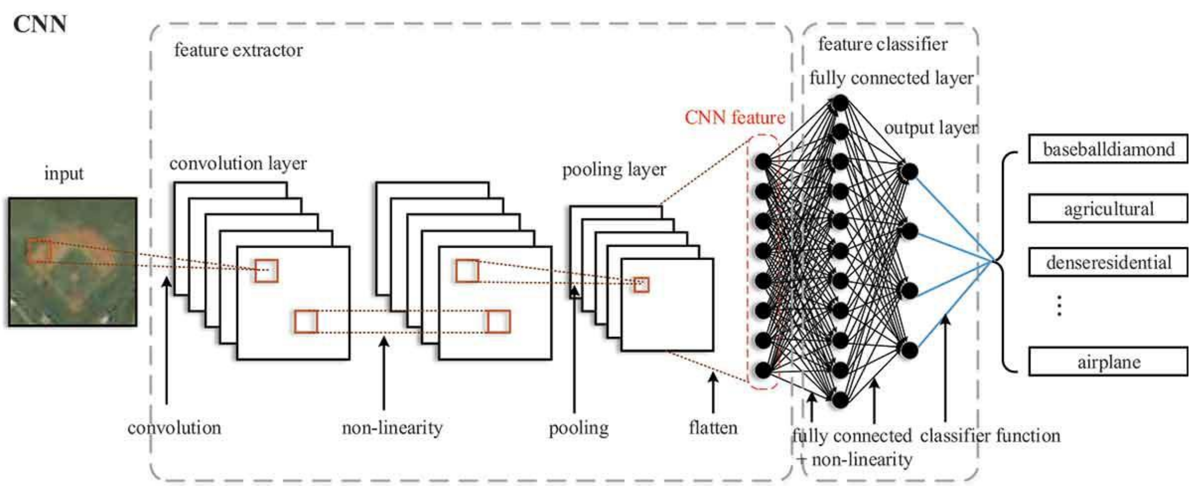

Figure 2-4 Basic Structure of CNN (adapted from Song et al., 2019)

A. Convolutional Layer

In CNN architecture, the convolutional layer is the most crucial component. This layer works as the fundamental building block of a CNN, utilizing convolutional kernels, which are learnable weight matrices, to process image data. Each kernel operates on a localized area of the image, capturing patterns such as edges and textures. The output, which is known as a convolved feature map, then goes through a nonlinear activation function to introduce non-linearity, enabling the network to learn complex features (Song et al., 2019). This process can be repeated across multiple layers to extract more sophisticated input features.

B. Pooling Layer

Following the convolutional layers, pooling layers reduce the spatial dimensions of the feature maps, thus decreasing the computational load and enhancing the network's focus on significant features. Common types of pooling include max pooling, average pooling, and random pooling, each contributing differently to the down-sampling process. Pooling layers help make feature detection invariant to scale and translation, providing robustness to variations within the input data (Song et al., 2019).

C. Fully Connected Layer and Output Layer

Deep in the architecture, fully connected layers integrate all features from the preceding layers to perform classification. These layers flatten the output of the last pooling or convolutional layers into a single vector, which is then used to map the learned features to the output space, such as the categories in classification tasks. The output layer typically uses activation functions like Softmax for multi-class classification tasks, translating the outputs into probability scores indicating class membership (Song et al., 2019).

D. Activation and Loss Functions

Activation functions such as ReLU or Sigmoid introduce nonlinearities into the model, which are essential for learning complex patterns. The choice of activation function affects the efficiency and effectiveness of training. Similarly, the network uses loss functions like cross-entropy to measure the error in predictions, which is crucial for training through backpropagation. Regularization techniques like L1 and L2 are implemented within the loss functions to prevent overfitting, ensuring that the model generalizes well to unseen data (Song et al., 2019).

In summary, this structure enables DCNNs to perform feature extraction and classification in an efficient way, making them powerful tools for tasks that require understanding and interpreting vast amounts of visual data.

2.4.2.2 Common Architectures in CNN

DCNNs have experienced significant evolution over the past decade, leading to the development of several influential architectures that have set new benchmarks in the field of image classification. These architectures have introduced various innovations, addressed limitations of their predecessors, and enhanced the performance of CNNs across multiple applications. A short introduction for each model will be included in this section.

One of the earliest and most influential DCNN models is AlexNet, which was created in 2012. AlexNet gained prominence after winning the ImageNet Large Scale Visual Recognition Challenge that year with a top-5 error rate of 15.3%, which shows its outstanding effectiveness of DL for image classification tasks. As an innovated model, AlexNet introduced several key innovations; it successfully used ReLU as the activation function, employed Dropout for regularization to prevent overfitting, and utilized max pooling layers to retain the highest responses in the feature maps. Additionally, AlexNet introduced the Local Response Normalization (LRN) layer, which increased the model's capacity for generalization by establishing a system of competition among local neurons. This mechanism amplified the most sensitive neurons while suppressing the less responsive ones. (Krizhevsky et al., 2012).

Following AlexNet, the Network-In-Network (NIN) model was introduced by Lin et al. (2014). NIN enhanced network structure by incorporating multilayer perceptrons within each convolutional layer to boost non-linearity. For innovation, this model replaced fully connected layers with global average pooling, mitigating overfitting caused by an excess number of parameters.

Another widely used model is VGG-Net, developed by Simonyan and Zisserman (2015). VGG-Net is notable for its simplicity and depth, utilizing a 19-layer network with smaller convolutional kernels (3×3). Its straightforward and uniform architecture significantly improves classification accuracy without adding computational complexity.

In the same year, Szegedy et al. introduced GoogLeNet, featuring a novel deep-learning structure that efficiently utilized computational resources to extract more features with the same amount of computation, thereby improving training results. The inception module was a significant innovation that efficiently optimized computational resources by increasing network depth and width. This 22-layer network achieved high performance in the ImageNet competition, demonstrating its advanced architecture (Szegedy et al., 2015).

ResNet, developed by He et al. (2016) at Microsoft Research, employs residual connections to address the vanishing gradient problem in deep networks. This innovation enables the construction of networks with over 1,000 layers, greatly enhancing their training and performance capabilities.

Following ResNet, DenseNet was proposed by Huang et al. (2017). DenseNet features dense cross-layer connections that ensure each layer is connected to all previous layers. This architecture allows for efficient feature reuse and faster model convergence, addressing the vanishing gradient issue and improving overall network performance.

EfficientNet, developed by the Google Brain team, is a highly efficient CNN architecture designed for image classification and recognition tasks. It achieves efficient model design by uniformly scaling the network's depth, width, and resolution, maintaining accuracy while reducing computation and parameter counts. This makes it an excellent solution for high-performance tasks under limited computational resources (Tan & Le, 2020).

In summary, these advancements highlight the continuous evolution of CNN architectures, each building upon the strengths of previous models to achieve superior performance in image classification. As a result, CNNs have become indispensable tools in computer vision, driving innovations in various applications and setting new standards in the field.

2.4.3 Deep Learning in Heritage Image Classification

The classification of architectural heritage, a significant research topic within cultural preservation, has been revolutionized by the advent of DL techniques, especially CNNs. This section delves into the diverse methodologies, applications, and existing gaps in the use of CNNs to classify architectural styles and elements, offering a nuanced understanding of the field's advancements and challenges.

A. CNN Models

The use of CNNs such as ResNet, GoogLeNet, and AlexNet in the classification of architectural heritage has been critical, but it also shows the need for constant improvement and innovation. Studies by Abed, Al-Asfoor, and Hussain, along with Cosovic and Jankovic (2020), have showcased the effectiveness of CNNs in handling a variety of architectural styles, emphasizing the importance of specialized datasets for training models to accurately recognize and categorize diverse architectural features. Despite these advancements, there is a noticeable gap in the creation of CNN architectures specifically designed for the unique challenges of heritage building classification. This gap is partially bridged by the application of transfer learning, as demonstrated in the works of Belhi et al. (2021) and Jankovic Babic (2023), where well-established models like ResNet and GoogLeNet are repurposed for heritage image classification. While these approaches utilize the strengths of existing networks, they may not fully capture the specific nuances of architectural heritage images.

The integration of CNNs with other data sources and methodologies, as exemplified by Kang et al. (2018) and Trier et al. (2021) who combined CNNs with street view images and airborne lidar data, respectively, represents a move towards more comprehensive and accurate classification methods. This trend suggests a growing emphasis on creating holistic and context-aware systems in heritage classification. However, the field still grapples with challenges in model specificity and adaptability, as the reliance on models designed for broad categorization might overlook finer architectural details. This underscores the necessity for more targeted and specialized CNN models. The limited exploration of advanced techniques like channel-spatial attention (Wang et al., 2023) and federated learning (Mehta et al., 2023) in heritage classification further highlights this need. These innovative approaches, which offer improved feature extraction and decentralized model training, open new pathways for future research.

In conclusion, while CNNs have significantly advanced the field of architectural heritage classification, there remains substantial scope for the development of more specialized, integrative, and context-sensitive models. Such progress would not only enhance the efficiency of heritage building classification but also align with the overarching objectives of cultural preservation and heritage conservation.

B. Datasets and Classification Focus

The research focus within this domain exhibits a clear dichotomy. On one hand, studies like those of Kang et al. (2018) and Wang et al. (2023) have concentrated on the classification of broader architectural styles. On the other hand, research efforts by Belhi, Bouras, and Foufou (2018a, 2018b) have been more narrowly focused on specific architectural elements. This distinction is crucial as it aligns with the overarching need for both a broad categorization of architectural styles and a detailed analysis of individual architectural features, which are essential for comprehensive heritage conservation. However, a significant gap remains in the comprehensiveness of datasets, particularly in terms of representing a wider array of architectural styles and elements from diverse geographical regions. This gap highlights the need for more inclusive and varied datasets that can provide a more holistic understanding of global architectural heritage.

C. Outcomes and Research Gaps

The outcomes of various studies employing CNNs for architectural heritage classification have been largely encouraging, highlighting the effectiveness of CNNs in accurately identifying heritage buildings. However, the disparity in classification results across different research projects points to the inherent complexity of this task. This variation underscores the necessity for ongoing refinement and enhancement of both the models and methodologies used in this domain. A critical area needing attention is the detailed analysis of specific architectural elements and spatial characteristics, which are essential for the precise classification of heritage buildings.

Furthermore, there is a notable gap in research utilizing street view imagery combined with DL for the identification of potential heritage buildings. This gap, as indicated by the works of Prasomphan (2022) and Vu et al. (2018), presents a significant opportunity for future research. On another front, the development of a CNN model incorporating channel-spatial attention by Wang et al. (2023) and the application of DL in the remote sensing of historical architecture by Yazdi et al. (2022) have demonstrated the versatility and global relevance of these methods in diverse contexts.

In summary, the integration of CNNs in architectural heritage classification has been a significant step forward, yet it simultaneously unveils new research frontiers. The development of comprehensive datasets, tailored CNN architectures for heritage buildings, and the fusion of CNNs with advanced technologies are pivotal areas for future exploration. These advancements are essential for refining the accuracy and efficiency of classification methods in heritage conservation.

2.4.4 Integrated CNN and MLP Models for Enhancing Image Classification with Mixed Data

A Multi-Layer Perceptron (MLP) is a type of artificial neural network designed to mimic the human brain's structure, consisting of an input layer, multiple hidden layers, and an output layer (Singhal & Sharma, 2023). Each neuron in one layer is fully connected to every neuron in the next layer, with connections weighted and adjusted during training to minimize errors using backpropagation. MLPs use nonlinear activation functions like sigmoid or ReLU to capture complex patterns, making them effective for tasks such as disease diagnosis from voice signals by identifying correlations between voice features and health conditions (Del Frate et al., 2007; Singhal & Sharma, 2023).

The integration of MLPs with DCNNs, which leverage mixed data for image classification, has increasingly been applied across various specialized fields. Zhang et al. (2018) raised an idea that MLP and CNN classifiers can offer distinct and complementary feature representations. As a result, combining these classifiers in an ensemble can potentially improve overall classification performance.

This hybrid approach has proven effective in enhancing the accuracy and efficiency of image classification systems, demonstrating significant advancements across different domains. A core innovation in the application of these models is the strategic fusion of CNN and MLP architectures to handle a diverse array of data inputs. This methodology utilizes the unique properties of each data type, such as numerical, categorical, and image data, to enhance the model's ability to identify complex and nuanced patterns within the dataset. For instance, in the medical domain, the integration of patient clinical data with diagnostic images has facilitated the development of highly accurate diagnostic tools. Studies like those conducted by Ahsan et al. (2020) have applied this approach to to differentiate between COVID-19 and non-COVID-19 patients by combining numerical and categorical data, such as age, gender, and temperature, with chest X-ray images. This integration has led to enhanced diagnostic capabilities, achieving an overall average accuracy of 94.6% ± 3.42%.

In environmental monitoring and remote sensing, the addition of geographical and temporal data to traditional image inputs has significantly improved the precision of classification models. For example, Sharifzadeh et al. (2019) developed a novel hybrid CNN-MLP classifier for identifying ship types in synthetic aperture radar (SAR) images. By integrating SAR image data with maritime traffic patterns, their model could more accurately identify and classify different ship types compared with only applying MLP and CNN respectively, demonstrating the model's benefits in complex maritime environments. Zhang et al. (2018) further explored this approach by incorporating very fine resolution remotely sensed images with land use data to enhance land cover classification tasks. Their method involved sophisticated preprocessing techniques to normalize and align the mixed data types, facilitating more effective integration and subsequent analysis.

Findings from the above studies consistently demonstrate that CNN-MLP models can produce better categorization results. These models not only increase the accuracy of classification tasks but also bolster the robustness of the systems, enabling them to effectively manage complex and varied datasets. This is particularly evident in challenging environments where traditional single-data-type models may fail to capture the nuances necessary for accurate classification. Overall, the integration of MLP and CNN models using mixed data inputs represents a significant evolution in image classification technologies. By effectively combining the capabilities of both architectures and utilizing the rich information inherent in mixed data types, these models offer more precise, efficient, and adaptable solutions for a wide range of complex classification tasks.

2.4.5 Summary and Future Research Focus

This review of global and Ontario heritage planning frameworks highlights the importance of preserving architectural heritage through robust legal regulations, comprehensive designation processes, and precise evaluation criteria. Despite significant advancements, there are notable gaps in the heritage designation process, indicating the need for further research and improvement.

The review also summarizes Ontario's architectural styles, particularly through the statistical analysis of Stratford's built heritage, which clearly reveals the diversity and significance of local architectural elements. DCNNs have also been proven highly effective in image classification, especially in heritage studies, offering superior accuracy and efficiency compared to traditional methods. The basic structure and common architectures of DCNNs, including ResNet, VGGNet, and MobileNet, are examined in order to demonstrate the importance of DCNNs in the advancement of this topic. The integration of DCNNs with MLPs further demonstrates enhanced classification capabilities by leveraging mixed data inputs, applicable to heritage architecture classification.

Looking ahead, several research directions are crucial for advancing heritage image classification. Firstly, testing or developing CNN architectures tailored specifically for heritage identification can be beneficial for improving the efficiency of the Ontario Heritage Designation

Process. Comprehensive datasets that encompass a wider variety of architectural styles and elements from the Ontario region are necessary to facilitate more accurate and representative classification models within the study area of Ontario. Additionally, integrating more advanced techniques, such as transfer learning and combining CNNs with MLPs, can further improve classification performance for capturing architectures with historical values. This approach leverages data beyond traditional street-view images and archived photographs by adding detailed geospatial data of specific cities, making the model more accurate and adaptable to the unique characteristics of each city in Ontario.

In conclusion, the integration of advanced DCNN models with appropriate datasets in heritage image classification presents significant opportunities for improving the accuracy and efficiency of heritage conservation practices. This thesis aims to address these gaps by developing specialized CNN architectures, specialized datasets, and exploring innovative techniques to tackle the complex challenges of heritage image classification. The next chapter will delve into the methodology adopted in detail.

Chapter 3
Data and Method

3.1 Chapter Overview

Chapter 3 provides an in-depth examination of the methodologies and data used to develop models that substantially enhance the heritage designation process. It introduces three specialized AI or GeoAI models that are crucial for identifying architectural styles, assessing potential heritage properties, and predicting heritage designation status. The chapter explains how these models are designed and integrated into existing frameworks, highlighting their role in automating and refining decision-making processes. The approaches discussed include the direct application of DL techniques, along with innovative uses of transfer learning and mixed data strategies. It demonstrates how each model addresses specific challenges within the heritage designation process, ultimately showing how advanced GeoAI methods can be effectively utilized in heritage preservation.

3.2 Model Overview

The integration of DL models, which includes the Architectural Style Identification Model, Potential Heritage Identification Model, and Heritage Property Designation Prediction Model, may contribute to enhancing the heritage designation framework by providing data-driven automated insights. Figure 3-1 illustrates the benefits these models bring to the Ontario heritage designation process. At the pre-screening and evaluating stage for potential designated properties, these models work as automated tools for municipal heritage committees, property owners, and local governments.

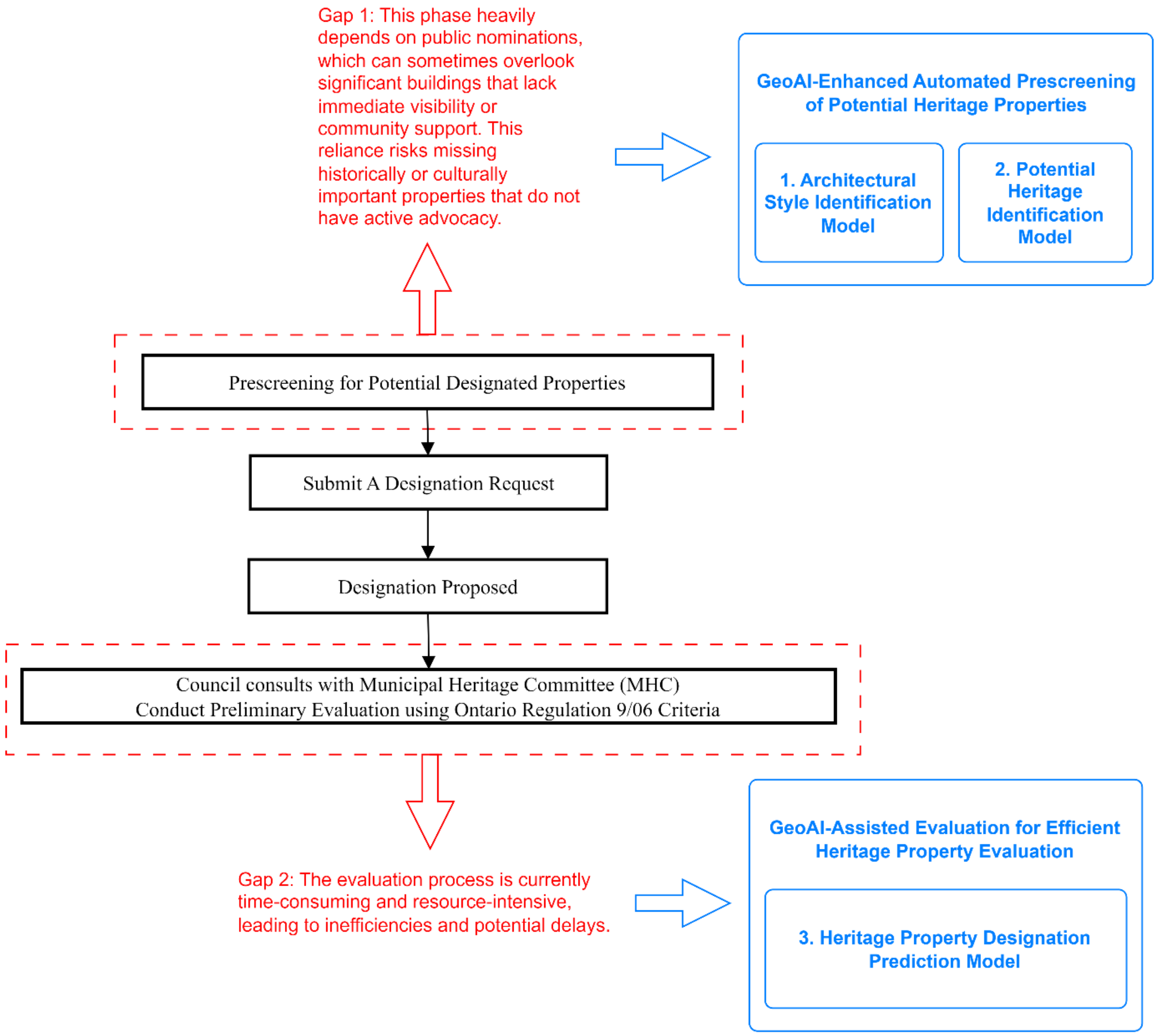

Figure 3-1 Link Between Ontario Heritage Designation Process and Solutions

As proposed, the Architectural Style Model automatically categorizes buildings into Ontario's prevalent architectural styles. This functionality is significant as certain styles, as discussed in Chapter 2, Section 2.3 on Ontario Architectural Styles, are more likely to be designated as heritage sites in Stratford. By recognizing these styles, the model provides essential data to heritage planners, aiding in informed decision-making during the preliminary screening phase and potentially guiding the focus of subsequent models.

The Potential Heritage Identification Model continues to assist in the pre-screening process by analyzing images of buildings to evaluate their potential heritage status, employing pattern recognition and feature extraction to assist in the early identification of buildings with heritage value. This proactive approach minimizes the risk of overlooking significant buildings that lack visible support or community advocacy, thereby enhancing the process's efficiency and accuracy.

Lastly, the Heritage Property Designation Prediction Model which applied GeoAI knowledge, utilizes archival photos and associated geospatial information to forecast whether a building is more likely to be designated as a heritage property on the local register. This model aims to address the inefficiencies caused by its high time consumption and level of human resources. By providing predictive insights on the likelihood of building listings, it also avails a basic decision-making process that can enable staff to make informed decisions within the legislative timelines.

The following Table 3-1 introduces three models, each designed to address requirements in the heritage designation process and optimized to leverage different datasets at different geographic scales. The Architectural Style Model is deployed at the provincial level, using the Ontario Architectural Style Dataset to automatically classify buildings into 21 architectural styles. This dataset includes a wide-ranging collection of architectural images, classified according to the Ontario Architectural Style Guide by the HPI Nomination Team at the University of Waterloo in Section 2.3.

At the same provincial level, the Heritage Identification Model employs the Ontario Built Heritage Dataset to identify buildings potentially holding heritage value by returning the label of 'Heritage' or 'Non-Heritage' for each architecture. This dataset captures images of legally protected architecture as well as officially designated and non-designated properties throughout Ontario. It categorizes images into 'Heritage' and 'Non-Heritage' groups, with the latter mainly consisting of more modern buildings and houses selected from the Ontario real estate listings and Bing search engine. This model is crucial for bridging the initial gap in heritage recognition, significantly boosting the efficiency and accuracy of heritage status evaluations and reducing the likelihood of overlooking significant properties.

For more localized analysis, the Heritage Property Designation Prediction Model focuses on a specific city in Ontario, which is Stratford. It uses the Stratford Heritage Geospatial Dataset to predict and indicate whether buildings are more likely to be 'Designated' or 'Non-Designated' as heritage properties. This dataset includes images and tabular geospatial data for each property listed on the Stratford Heritage List, covering attributes such as address, building type, architectural style, year of construction, etc. Collectively, these models significantly broaden the scope and improve the effectiveness and accurateness of heritage designation in Ontario, ensuring that decisions are not only faster but also well-supported by thorough analytical evidence. This overview highlights the stages at which these models are beneficial and outlines the essential components of each model. Detailed discussions on data sources, dataset composition, and model development are explored in the following sections.

Table 3-1 Summary of Proposed Model, Study Area, Dataset Used, and Expected Outcome

Model	Study Area	Dataset Used	Expected Outcome
Architectural Style Model	Ontario (Province-level)	Ontario Architectural Style Dataset	Returns one of the 21 specific architectural styles for the input image of an architecture.
Heritage Identification Model	Ontario (Province-level)	Ontario Built Heritage Dataset	Returns the label of 'Heritage' or 'Non-Heritage' for the input image of an architecture.
Heritage Property Designation Prediction Model	Stratford (City-level)	Stratford Heritage Geospatial Dataset	Predicts and returns the label of 'Designated' or 'Non-Designated' for the potential heritage properties with geospatial information details.

Lastly, the Heritage Property Designation Prediction Model is specifically designed for the evaluation process, particularly addressing the expedited needs brought about by the implementation of "The More Homes, Built Faster Act" (Bill 23). This model is critical for accelerating the process to designate existing properties which were not previously listed as designated, and is thus only applicable at this final stage of the process.

The three models developed for the Ontario heritage process are designed to be used in a specific sequence to address different stages of the heritage designation workflow. The Architectural Style Model is the first step, aimed at enhancing the accuracy of the subsequent Potential Heritage Identification Model. It is important to note that while the Architectural Style Model identifies the architectural style of a building, it does not determine whether a building is a heritage property. This is the role of the Potential Heritage Identification Model, which is designed for the initial auto pre-screening of potential heritage properties. This model provides a comprehensive but basic retrieval of data, assisting planners in identifying possible heritage properties but not in making final designation decisions.

Together, these models form a cohesive suite of tools that support planners through the structured, sequential stages of the heritage designation process, ensuring that each stage is informed by the insights and data provided by the preceding model.

3.2.1 TensorFlow and Keras

TensorFlow and Keras are chosen for all DL models in this research due to their flexibility, scalability, and extensive community support. As a brief introduction, TensorFlow, developed by Google, is an open-source DL software library used for defining, training, and deploying ML models (Goldsborough, 2016). Initially an independent project, Keras is now a high-level neural network API that has been incorporated into TensorFlow (Gulli et al., 2019). This integration enhances its accessibility and simplifies the experimentation process with DL models.

Building on the enhanced accessibility and simplified model experimentation provided by TensorFlow and Keras, one of the benefits to use Keras in this research is the valuable feature of Keras Applications. These applications provide pre-trained DL models along with their weights. They can be employed directly for tasks like prediction and feature extraction, or serve as a foundation for further training, a process referred to as fine-tuning (Keras Team, n.d.). Specifically, architectures such as ResNet, VGG-16, and MobileNet are highly effective for image recognition tasks, which are essential for accurately classifying built heritage.

Numerous studies across various domains have demonstrated the effectiveness of using Keras Applications and pre-trained architectures in achieving research objectives through DL. In meteorological image analysis, Joshi and colleagues (2021) leveraged these tools to enhance the classification of cloud street images, utilizing advanced preprocessing and hyperparameter tuning techniques. Similarly, in the medical field, Brady Kieffer and associates (2017) employed pre-trained networks for the classification of histopathology images, illustrating the value of these models in handling complex datasets with limited sample sizes. Additionally, Deng Xing’s team in 2020 applied Keras pre-trained architectures for COVID-19 detection using chest X-ray and CT images, showcasing their application in urgent healthcare diagnostics. These studies collectively highlight the adaptability and effectiveness of Keras pre-trained models in achieving specific research goals across different scientific and practical contexts. Thus, the choice of TensorFlow, specifically Keras, is well proved.

3.3 Deep Convolutional Neural Network Models for Architectural Style Identification

3.3.1 Study Area

Serving as both an industrial core and a vital cultural and media center for English-speaking Canada, Ontario presents a unique and varied geographical landscape. According to the National Trust for Canada (n.d.), the province boasts over 31,500 properties listed on municipal Heritage Registers, featuring a wide range of architectural styles and building types. These sites serve as cultural landmarks and are crucial to the local tourism sector, significantly boosting the local economies of cities rich in heritage.

However, these valuable assets are increasingly threatened by urban expansion (Caldwell et al., 2022). Traditionally, the OHA has played a crucial role in protecting these cultural treasures. Despite these protective measures, the rapid urbanization of southern Ontario, particularly in major metropolitan areas such as Toronto and Ottawa, continues to pose significant challenges. For example, Toronto's population increased by 2.3% from 2016 to 2021, exacerbating pressures on both infrastructure and historic sites (City of Toronto, 2022). Additionally, the recent implementation of the More Homes Built Faster Act (Bill 23) complicates the heritage designation process (City of Toronto, 2023), posing particular risks to sites significant to Indigenous and underrepresented communities.

Given such circumstances, Ontario is the region of interest due to the high density of national historic sites. This province has become a perfect testing ground for innovative strategies against the backdrop of rapid developmenohat and the massive heritage preservation effort. The large stock of cultural heritage resources, combined with a complex urban landscape, is ideal for the application of new technologies based on GeoAI. The use of these technologies allows more efficacious heritage conservation, enabling accurate identification and preservation of historic sites within the course of ongoing developmental changes.

3.3.2 Ontario Architectural Style Dataset

The Ontario Architectural Style Dataset was designed and created with the specific purpose of supporting the identification of architectural styles of potential built heritage in Ontario. The dataset includes a wide-ranging collection of architectural images, classified into 21 distinct styles according to the Ontario Architectural Style Guide by the HPI Nomination Team at the University of Waterloo in Section 2.3. Figure 3-2 offers a representative example of each architectural style that is included in the dataset.

Figure 3-2 Representative Examples of 21 Architectural Styles in Ontario

The general workflow for creating the dataset is shown in Figure 3-3. The dataset's foundation is built on Marian Dumitru Danci's Architectural Style Dataset, a comprehensive array of 10,113 images across 25 styles (Marian Dumitru Danci, 2020). This initial selection was enriched with images sourced from Google and integrated with insights from the Architectural Style Dataset created by Xu et al. in 2014. In addressing the representation needs of styles that were not included in Danci and Xu's collection, such as Italianate and Second Empire, efforts were made to gather additional images. Images not included in the open datasets were sourced from Bing through a Python script designed to query and scrape Microsoft Bing image search using specific keywords. The keywords consisted of the name of each architectural style (e.g., "Queen Anne") combined with "Ontario Exterior," ensuring a thorough collection of these styles. All downloaded images were checked and cleaned manually to ensure the quality of the dataset. After the final selection, images where the scenery dominated over the buildings were cropped to emphasize the architectural features.

Effective training of DL models is also highly dependent on the quantity of the data available. In this dataset, architectural images were initially found in different numbers. Some styles, such as Neoclassical and Octagonal, had as few as 80 and 60 images, respectively, while others, like Queen Anne, had more than 400 images. To address this issue, data augmentation techniques were applied. This strategy served not only to enrich the dataset but also to standardize the representation across styles. Utilizing the 'ImageDataGenerator' from TensorFlow, a series of detailed transformations were applied to each image, which included horizontal flipping to create a mirror image, rotation within a 15-degree range to introduce angular diversity, zooming to capture different size perspectives, and brightness adjustments to simulate varying lighting conditions. These enhancements were vital, fortifying the dataset and training our models to recognize architectural styles in diverse visual conditions. The effectiveness of these augmentations is evident from Table 3-2, which documents the augmented count for each style nearing our goal of 2000 images. This equilibrium was imperative. By balancing the image counts, it is guaranteed that no architectural style was overshadowed in the dataset, thus mitigating model bias. Such a uniform distribution is essential for a fair learning process, allowing the models to develop an even-handed recognition ability across the spectrum of architectural diversity.

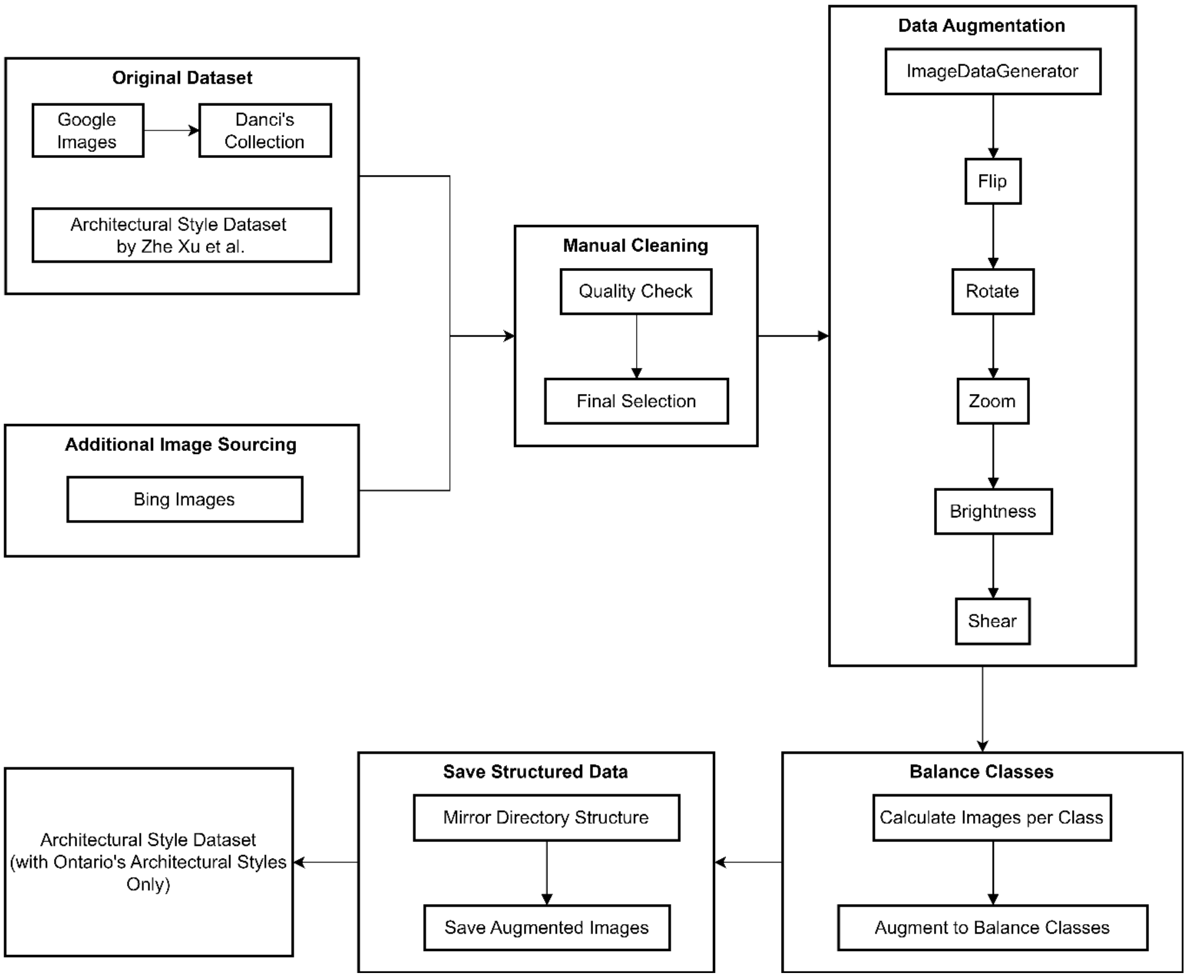

Figure 3-3 Methodological Flowchart for Ontario Architectural Style Dataset Creation

Table 3-2 Image Count Comparison Before and After Data Augmentation for Different Architectural Styles

Architectural Style	Original Count	Augmented Count
Art (Streamline) Moderne	284	1982
Art Deco	473	1981
Beaux Arts	470	1988
Colonial (Georgian Revival)	303	1963
Edwardian	409	1934
Georgian (combined Mennonite)	381	1995
Gothic Revival	283	1954
Greek (Classical) Revival	513	1922

Italianate	372	1970
Log Houses	366	1954
Modern (International)	417	1937
Neoclassical	80	1981
Octagonal	60	1973
Prairie Craftsman Bungalow	371	1986
Queen Anne	420	1972
Regency	249	1988
Romanesque Revival	248	1972
Second Empire	373	1995
Tudor	293	1915
Victorian	194	1992
Victory Housing	123	1986

3.3.3 Architectural Style Model Design

A variety of pre-trained architectures from the Keras library, including EfficientNetB0, DenseNet121, InceptionV3, ResNet50, VGG16, NASNetMobile, MobileNetV3Large, ResNet101, and VGG19, are utilized for their superior ability to extract features effectively, particularly in tasks related to image recognition. Then, the top four architectures demonstrating the best initial performance undergo further optimization through fine-tuning, enhancing their capability to classify architectural styles accurately.

The figure below provides a comprehensive flowchart of the DCNN model for classifying architectural styles, including all significant steps from data preparation to fine-tuning. This process, which outlines the basic structure of the architectural style model, is applicable across different algorithms, though specific parameters may vary. Each of the seven critical stages of the process, from data preparation to model fine-tuning, is detailed in the following sections to ensure a thorough understanding of the entire model development cycle.

The process of developing and refining the model begins with data preparation. The images are organized and divided into three distinct subsets: training, validation, and test datasets. Following a strategic allocation of 70% for training and an equitable division of the remaining images (15% for validation and 15% for testing), each category within the dataset is treated with this stratified approach, guaranteeing uniformity across all architectural styles. This ensures consistency and compatibility with the CNNs used in the training process, as the input data is standardized to a uniform shape of (224, 224, 3), a typical format for RGB images that represents the dimensions of the image in image processing tasks, as highlighted by Krizhevsky et al. (2012).

The base model, pre-trained on the ImageNet dataset by Keras, is then loaded. This model is initially tailored for the task by freezing the top layer designed for ImageNet classification. This adjustment is necessary because the original output layers are configured to accept only a predefined input size, and any alteration to the input shape from earlier layers would render the learned weights incompatible due to matrix size mismatches (Simonyan & Zisserman, 2015).

Feature extraction and classification involve each selected model serving as a feature extractor. During the feature extraction phase, convolutional and pooling layers are utilized to identify patterns and structures within the images. After the base model processes the images, a GlobalAveragePooling2D layer is applied to average the spatial dimensions of the feature maps across each channel, condensing them into a single value per channel. This method, recommended by Lin et al. in 2014 for its effectiveness in reducing overfitting and computational complexity, is similarly employed in my model to achieve these goals.

The classification phase incorporates fully connected layers and a softmax layer to map the extracted features to specific classes. A dense layer with 1024 neurons and ReLU activation is added, representing the "Fully Connected" part of the network. This layer converts the feature vector output from the feature extractor into a hidden layer, a method detailed by LeCun et al. (2015). Finally, an output dense layer with a number of neurons equal to the number of classes and softmax activation is used. This layer converts the output into a probability distribution, facilitating multi-class classification, as supported by the research of Szegedy et al. (2015).

Model compilation and optimization involve defining key parameters that govern how the model learns from data during training. This process includes selecting an optimizer, specifying a loss function, and choosing evaluation metrics. The Adam optimizer, discussed by Kingma and Ba (2017), is chosen for its adaptive learning rate capabilities and efficiency in handling sparse gradients. Sparse categorical cross-entropy is selected as the loss function due to its suitability for multi-class classification tasks where target labels are integers. This approach is particularly practical for handling such tasks as it simplifies the calculation of the loss for integer labels, thus improving computational efficiency and performance, as detailed by Li et al. (2021) in their discussion on improved categorical cross-entropy loss functions for DCNNs with noisy labels. Accuracy is used as the evaluation metric, providing a straightforward measure of the model's performance in making correct classifications.

The training process incorporates callbacks such as ModelCheckpoint and EarlyStopping to monitor progress effectively. ModelCheckpoint saves the best model observed during training, measured by validation accuracy, ensuring retention of the best-performing version. EarlyStopping monitors validation accuracy and halts training if no improvement is seen over a specified number of epochs, preventing overfitting. These techniques are well-documented for their effectiveness in model training, as noted by Prechelt (1998).

Fine-tuning further enhances the performance of the pre-trained models by selectively unfreezing the top layers of the model, specifically the top 20 layers. This approach is based on the hypothesis that the upper layers capture more abstract representations, which differ between ImageNet and architectural style classification tasks and thus benefit from fine-tuning, as Yosinski et al. (2014) suggested. Fine-tuning is performed over 35 epochs with an initial learning rate of $1e^{-5}$, allowing for subtle adjustments to the model weights. This structured approach ensures that the model not only adapts to the specific nuances of architectural styles but also exhibits robustness and high generalization capability when deployed in practical scenarios.

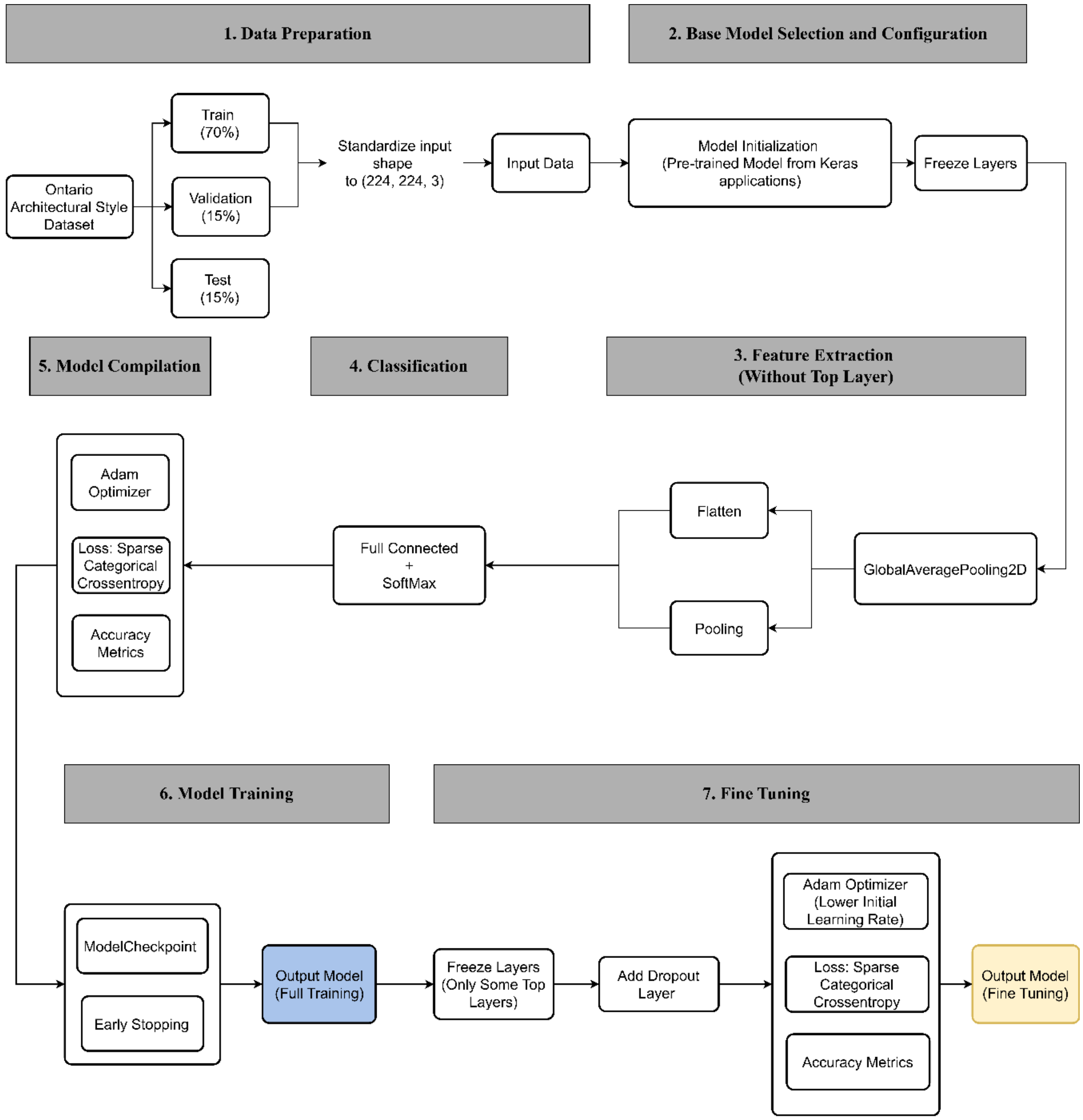

Figure 3-4 Flowchart of the Architectural Style Model Design

3.4 Deep Convolutional Neural Network Models for Heritage Identification Model

3.4.1 Ontario Built Heritage Dataset

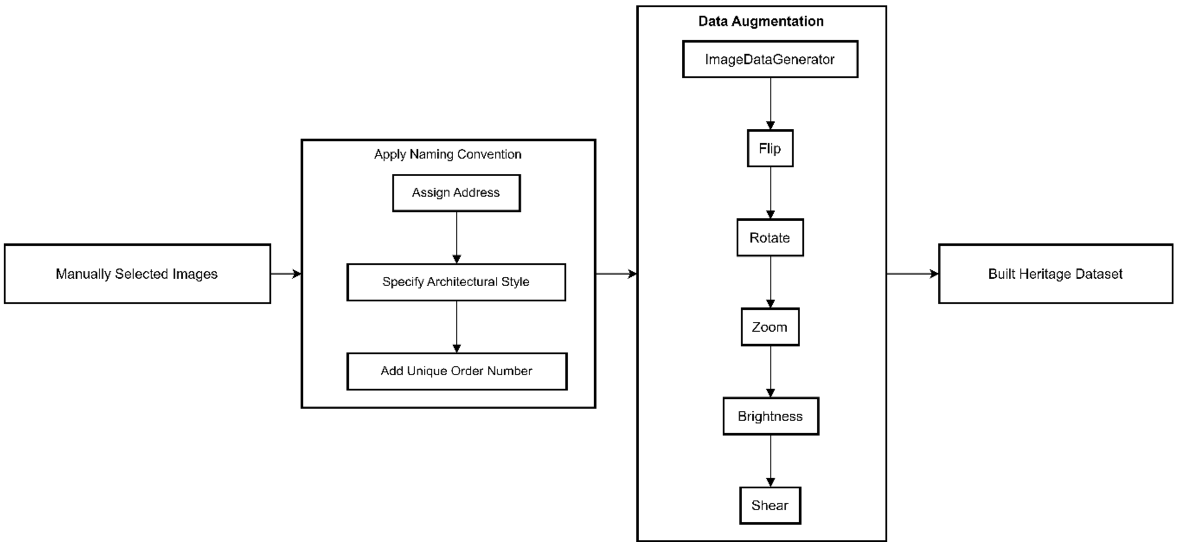

Figure 3-5 Methodological Flowchart for the Creation of the Ontario Heritage Dataset

Parallel to the Ontario Architectural Style Dataset, the Ontario Heritage Dataset has been assembled with a similar methodology. This dataset includes images of legally protected architectural sites, as well as both officially designated and undesignated properties across Ontario. The images of legally protected architectural sites are primarily sourced from the Ontario Heritage Trust, a non-profit agency of the Ontario Ministry of Tourism and Culture, and use publicly available data from their official website. It categorizes images into 'Heritage' and 'Non-Heritage' groups, and the latter mainly consists of more modern buildings and houses selected from the Ontario real estate listings. Each image in the Heritage Dataset is systematically named in a format that captures critical information concisely. This structured approach not only streamlines data management but also underpins the labelling process, which is vital for the successful implementation of multi-input model architectures.

Mirroring the procedure taken with the architectural style dataset, the heritage dataset has similarly undergone a data augmentation process, which is thoroughly described in Section 3.3.2. This step increased the image counts in our study, with the heritage category images rising from 267 to 777, and the non-heritage category images increasing from 311 to 817. These enhancements have boosted the visual diversity of the dataset, playing a crucial role in strengthening the model's ability to generalize.

3.4.2 Built Heritage Model Design

Two different versions of the model were developed to address the challenge of classifying images into heritage and non-heritage categories using the Ontario Heritage Dataset (Figure 3-7). Both versions explore DL strategies, ranging from direct training to transfer learning, tailored to detect potential architectural built heritage effectively.

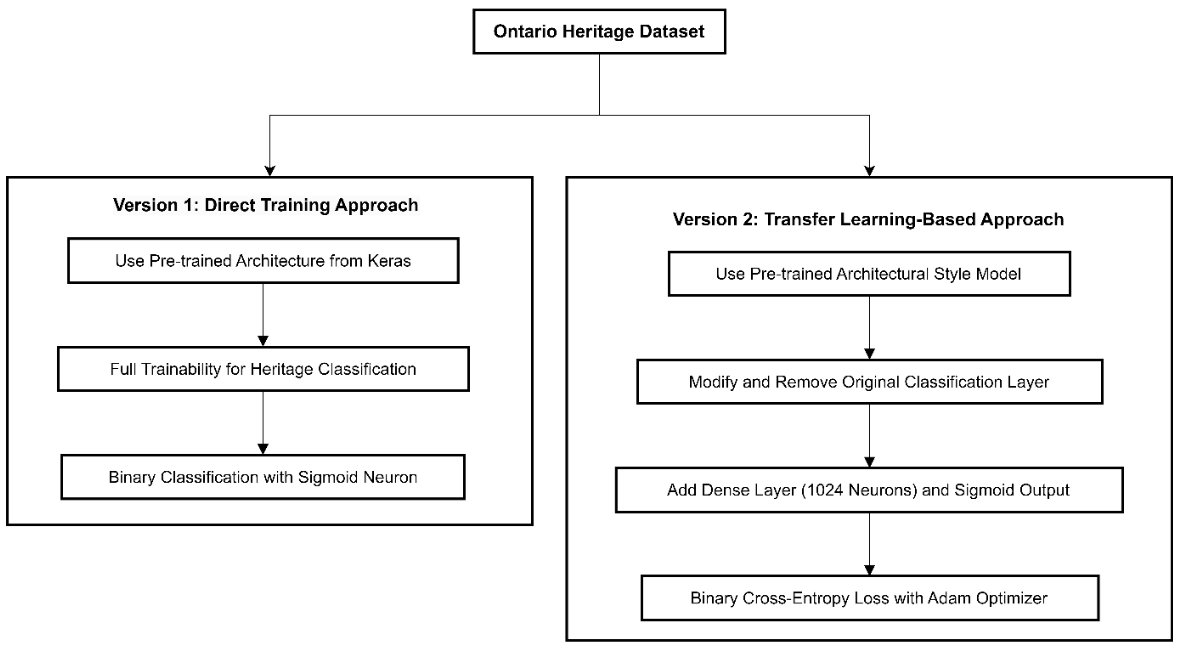

Figure 3-6 Dual Model Development Approaches for Heritage Classification

Version 1 of the study implemented a direct training approach, employing pre-trained architectures from the Keras library. This approach does not rely on further pre-trained models or external datasets but focuses on learning directly from the provided Ontario Heritage Dataset. The full trainability of the base model enables it to be fine-tuned to the specific requirements of heritage classification, ensuring that the neural network accurately identifies and learns the unique architectural elements of Ontario heritage buildings. Version 1 is almost identical to Section 3.3.3 in building the model. The only difference is that the configuration of this model is unique, as it ends with a single neuron with a sigmoid activation function for binary classification. It is compiled using the Adam optimizer and a binary cross entropy loss function, matching the binary classification nature of the task.

Version 2 adopts a transfer learning-based approach to refine the classification of images into heritage and non-heritage categories, primarily leveraging the capabilities of the best pre-trained architectural style recognition model as a feature extractor. Transfer learning is a ML method where a model developed for one task is reused as the starting point for a model on a second task, effectively leveraging pre-trained networks to improve detection capabilities in new, related domains (Pan & Yang, 2010). This approach was mainly focused on the subtle detection of architectural heritage. For this task, the version 2 model first incorporates an architecture pre-trained on ImageNet and then adapts it to focus on extracting and repurposing features specifically for heritage recognition. The strategy involves modifying the pre-trained model to remove its original classification layers, retaining its ability to extract various architectural features. These features are then channelled into a newly added dense layer with 1024 neurons, culminating in a sigmoid output layer tailored for binary classification. The model is also compiled using the Adam optimizer and trained using the binary cross entropy loss function.

3.5 Hybrid MLP-CNN Model for Heritage Property Designation Prediction

3.5.1 Study Location of Stratford

Figure 3-7 Map of Stratford (Google, 2024).

Stratford is a small city situated in the heart of southwestern Ontario. Geographically situated along the Avon River, Stratford's development and community life have been significantly influenced by the river. The city's origins trace back to a settlement established in the winter of 1831–32 by William Seargeant (or Sargent), who built the Shakespeare Hotel near the Avon. Initially named Little Thames, the river and settlement were renamed by 1835, likely due to the influence of William Dunlop of the Canada (development) Company, to honour the birthplace of William Shakespeare, Stratford-upon-Avon in England, highlighting the city's profound cultural ties (Britannica, 2024). Culturally, Stratford has been enriched by the diverse influx of English, German, Scottish, and Irish immigrants in the 1830s and 1840s. This multicultural foundation has transformed Stratford into a microcosm of Canadian diversity and history, reflecting the broader narrative of Canada's development. The city's status as the seat of Perth County since 1853, located in the heart of dairy-farming country, adds another layer to its rich historical and cultural identity, making Stratford a unique blend of natural beauty, historical depth, and cultural diversity (Perth County, 2017).

By the 20th century, the city had become a nexus for furniture manufacturing and railway locomotive repairs, with the Grand Trunk Railway locomotive repair shops becoming a pivotal source of employment, at one point engaging around 40% of the residents (Simpson, 2020). This industrial heritage laid the groundwork for Stratford's recognition as a center of historical significance. In 1976, Stratford City Hall was honoured as a National Historic Site of Canada. Following this, the Stratford Armoury and the city's VIA Rail Station were respectively acknowledged in 1992 and 1993 as Federal Heritage buildings by the Registrar of the Government of Canada Heritage Buildings (Government of Canada, n.d.). These designations underscore Stratford's rich history, marking it as a representative of Canada's culture and architecture.

In addition to its rich and diverse built heritage, the Stratford City Centre Core has been recognized as an HCD under the OHA. This designation involves properties marked on the city's official maps and subjects them to HCD District Standards. These publicly available guidelines aim to ensure that any development within the district aligns with Stratford's unique heritage character, thus preserving the city's distinctive architectural and historical identity for future generations (The Corporation of the City of Stratford, 2019a). The map illustrates that the HCD is bounded by St. Patrick Street, Downie Street, and Lake Victoria Street. This downtown core HCD contains 190 commercial buildings, which highlight the architectural diversity and historical richness of Stratford.

Further reinforcing Stratford's preservation endeavours, Heritage Stratford, instituted under By-law 133-2004, serves as an advisory body to the City Council on matters of cultural heritage value designation and conservation. Composed of dedicated citizen volunteers, this committee collaborates with property owners and the community, holding monthly public meetings to foster a transparent and inclusive heritage conservation process. This initiative not only promotes community engagement but also ensures the collective effort to preserve Stratford's historical essence (The Corporation of the City of Stratford, 2019b).

The decision to select Stratford as the study area for this research is grounded in several compelling reasons. Firstly, Stratford has a rich diversity of built heritage, indicating a wide array of architectural styles and historical periods. This variety provides a comprehensive backdrop for examining heritage conservation practices and methodologies. Secondly, the availability of extensive, open-access heritage geospatial data on the government's official website significantly bolsters the research framework. This data includes archived photographs of Designated and Non-Designated

Properties, detailed information such as addresses, individuals associated with each building, construction years, whether the property is located on its original site, and the identities of the builders. Such detailed information is invaluable for constructing and testing GeoAI and mixed-data methodologies, offering a robust foundation for analyzing and classifying heritage properties.

3.5.2 Stratford Heritage Geospatial Dataset

The Stratford Heritage Geospatial Dataset is a well-structured collection that forms the foundation of the hybrid MLP-CNN model, which assesses Stratford buildings based on multiple factors. The dataset comprises geospatial and attribute data for each heritage property in Stratford, including both designated and non-designated properties listed on the City of Stratford Heritage Pages. The dataset contains a total of 125 properties within the study area of Stratford, with 89 designated properties and 36 non-designated properties.

3.5.2.1 Stratford Heritage Tabular Data

The tabular dataset was created following the evaluation criteria detailed in Section 2.2.2.3 of Chapter 2, aligning with the principles established by the OHA and the guidelines set forth in the foundational document “Evaluating Heritage Resources in the Town of Markham.” The criteria without any data were excluded from this dataset. Attributes such as ‘Address,’ ‘Building Type,’ ‘Architectural Style,’ ‘Year of Construction,’ ‘Persons Associated with The Building,’ ‘Designer/Builder,’ ‘Site,’ and ‘Designated Status’ were extracted from the Stratford government’s designated and non-designated properties list through meticulous summarization.

The attributes ‘Within Heritage Conservation District’ and ‘Distance to Closest Built Heritage (meter)’ were calculated in ArcGIS Pro. To achieve this, geocoding was applied to convert each building’s address into coordinates, and then these coordinates were added as a layer. Using the map of the City of Stratford HCD boundary from the Stratford government’s website, a feature dataset which has the same boundary was created (Figure 3-8). By using the “select by location” function, the buildings within the HCD boundary were selected and coded as “yes,” while others were coded as “no.” The ‘Distance to Closest Built Heritage (meter)’ attribute was calculated using buffer analysis, which measures the distance and additional proximity information between the input features and the nearest feature in another layer or feature class.

Figure 3-8 Stratford's HCD Boundary

These processes capture the diverse features necessary for the GeoAI model to predict the designation status of buildings accurately. Table 3-3 displays the dataset by listing all the attributes along with their descriptions, data types, and examples. Figure 3-9 in the next section exemplifies the detailed archival information of a building on Stratford's designated heritage list, offering a snapshot of each building's details and historical background. The table was saved in the .csv format for better loading and preprocessing.

Table 3-3 Heritage Property Attributes Table

Attribute Name	Description	Data Type	Example
Address	Physical location of the building.	Text	10 Downie Street

Building Type	Classification of the building based on its usage or function.	Categorical	Residential
Architectural Style	Design style of the building reflects historical and cultural context.	Categorical	Late Victorian
Year of Construction	The year the building was initially constructed.	Numerical	1894
Persons Associated with The Building	Indicates if there are notable persons linked to the history of the building.	Boolean	Yes
Designer Builder	Indicates if the original designer or builder is known.	Boolean	No
Site	Describe if the building is located on its original site or if it has been moved.	Categorical	Original Site
Designated Status	Heritage status of the building (e.g., designated as a heritage building).	Categorical	Designated
Within Heritage Conservation District	Specifies if the building is within a heritage conservation district.	Boolean	Yes
Distance to Closest Built Heritage (meters)	The measured distance is in meters to the closest designated heritage building.	Numerical	34.2551716

3.5.2.2 Stratford Heritage Image Data

Image data is another key component of the Stratford Heritage Geospatial Dataset that collects images of designated and non-designated properties listed on the official website for the City of Stratford. As shown in Figure 3-9, the official documentation for each heritage building on the list contains archived photographs provided by the official authorities, which range from one to several. To ensure the quality and consistency of the data, the photographs were rigorously screened and selected to ensure that the clearest images, those showing the most details of the house, and those where the house made up the largest percentage of the photograph were extracted and downloaded. To facilitate correspondence with other data in the Stratford Heritage Table, each photograph was named following the address of its corresponding building. This not only ensures data consistency and traceability, but also provides high-quality image data for subsequent data analysis and model training. Figure 3-10 provides a general overview of an example image in the dataset.

Address: 51 Avon Street, Stratford

Legal Description: Lot 2 Plan 130

Designation By-law: # 209-87

By-law Designation Date: November 9, 1987

Date of Construction: 1870s

Architectural Style: Second Empire/Italianate

Architectural Description: Two storey, buff brick, three bay, mansard roof with a flat section in the centre; three round- arched hooded dormers on slope of roof with 2/2 windows, decorative brackets under the eaves; quoins at the corners and at corners of centre projection; three rectangular 2/2 windows on second storey; one three sided bay window on each side of centre entrance, bay windows have a flat roof, brackets under roof, one tall narrow arched window at each side of bay, larger arched window with multiple panes in centre of bay window; centre entrance with curved transom and side lights and small rope shaped columns on either side of door, door has decorative glass in upper portion with a wooden panel below

Property has a National Historic Plaque. This was the home of Miss Annie Macpherson who brought many orphaned children over from London England, between 1883 and 1919, to live in the house until they found employment on farms and in homes where they worked as domestics.

Figure 3-9 Information page for the designated property at 51 Avon Street (City of Stratford, n.d.).

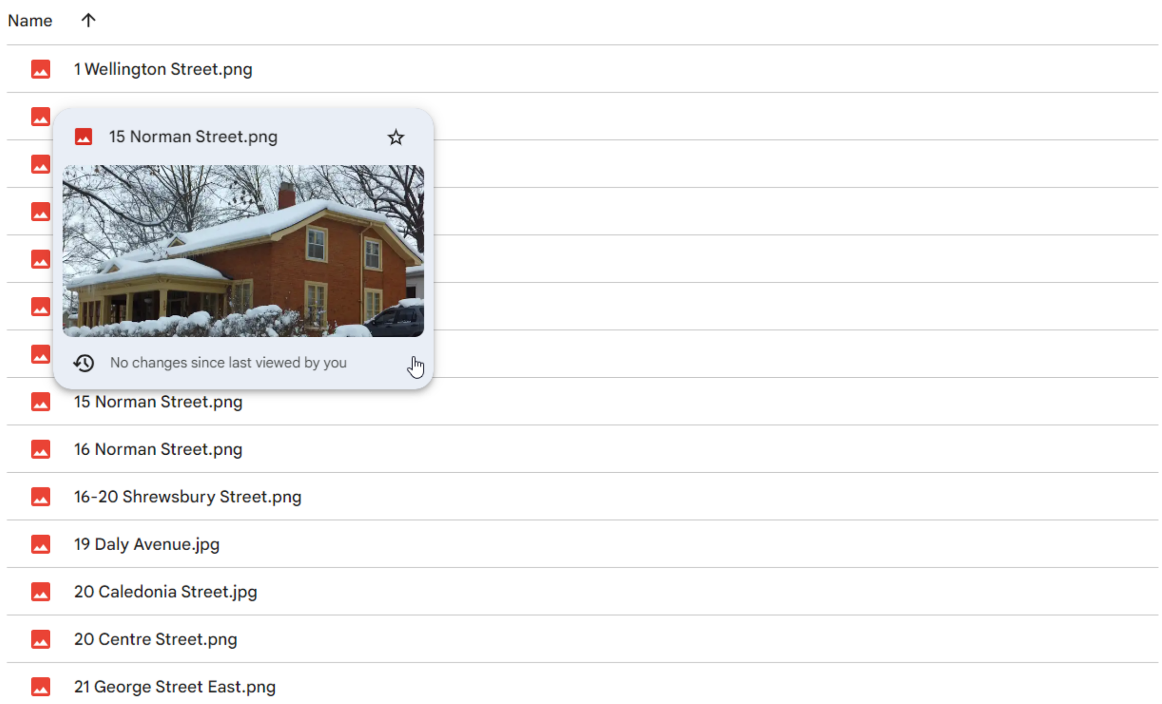

Figure 3-10 Sample of Stratford Heritage Image Data

3.5.3 Hybrid CNN–MLP Network

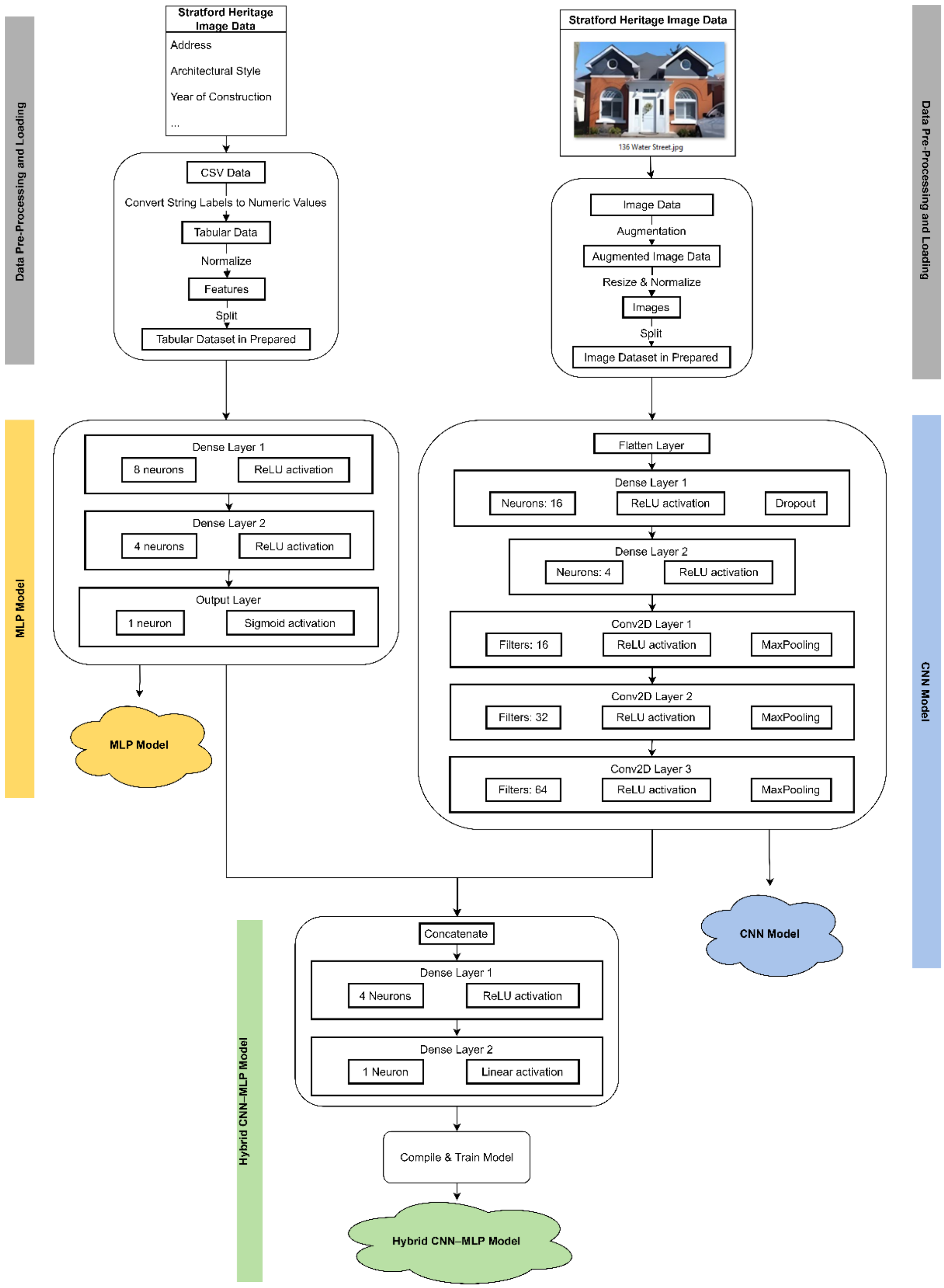

Figure 3-11 Flowchart of Proposed MLP-CNN Model

The hybrid MLP-CNN model designed in this study aims to combine the advantages of MLP and CNN to process different types of data input to improve the model's understanding and processing capabilities of data. The model construction is mainly inspired by the research by Ahsan et al. (2020) and the tutorial by Rosebrock (2019). The entire process of data preparation, model architecture design, model training and evaluation is visualized as a flowchart (Figure 3-11).

As the first step, the data preparation phase involves extracting information from the Stratford Heritage Tabular Dataset. The dataset is derived from a CSV file containing various information about the building. The target variable is "Designation Status," which indicates whether the building is designated as heritage or not. In order to facilitate the ML model, a method called one-hot coding was used to convert the categorical features in the dataset into numerical features. Subsequently, the target variable was binarized using Label Binarizer to make it suitable for the classification task.

In terms of data partitioning, an 80-20 ratio was used to divide the dataset into a training set and a test set. This ensures that the model sees enough data during training while retaining a portion of the data for evaluating model performance. In order to make the individual features have the same scale, we use the Min-Max normalization method to scale the feature values between 0 and 1. This normalization step is essential to ensure the stability of the model training process and to improve the model performance. Loading and preprocessing of image data is a key aspect of this study. Based on the path information in each data record, the corresponding image files were loaded. These images are uniformly resized to 64×64 pixel size and normalized so their pixel values are between 0 and 1. This preprocessing helps to accelerate the training process of the model and improves the convergence of the model.

In terms of model architecture design, three models, including a CNN model, an MLP model, and a hybrid model combining CNN and MP, are created. The multilayer perceptron contains two fully connected layers with eight and four neurons, using the ReLU activation function designed to capture complex patterns in tabular data. The CNN, on the other hand, consists of three convolutional layers with filter sizes of 16, 32, and 64 in that order, and each layer is followed by ReLU activation, batch normalization, and maximum pooling layers for extracting spatial features in the image data. The output of the convolutional layer is converted into a one-dimensional feature vector through a spreading layer, followed by two fully connected layers containing 16 and 4 neurons, respectively, using the ReLU activation function. In addition, a Dropout layer is added to prevent overfitting. In order to combine the advantages of MLP and CNN, the hybrid model splices the outputs of both and further processes them through two additional fully-connected layers, ultimately predicting the specified state through an output layer with a linear activation function.

The model training phase begins with model compilation using the Adam optimizer, which is suitable for large datasets and non-smooth objectives, allowing the model to perform parameter updates efficiently. Since this task is a binary classification problem, binary cross entropy is chosen for the loss function. In order to enhance the generalization ability of the model, data enhancement of the image data is performed by applying stochastic transformations such as rotation, scaling, translation, shearing and horizontal flipping. These data enhancement techniques help to simulate various real-world scenarios and improve the robustness of the model. In addition, a learning rate scheduler is used to reduce the learning rate when the verification loss is no longer decreasing, resulting in finer tuning during model convergence. Model training was performed over 100 cycles with a batch size of 8. During the training process, both image data and tabular features were input, and the performance of the validation set was monitored to prevent overfitting.

3.6 Performance Metrics and Interpretation

In the evaluation stage, it is essential to employ a comprehensive set of metrics that accurately reflect the effectiveness and efficiency of the developed predictive models. These metrics are integral to validating the models' capabilities and ensuring their applicability in real-world scenarios. This section defines each metric, explains how it is calculated, and interprets what the results signify about the model's performance.

- **Confusion Matrix:** It is a table most frequently used to describe the performance of a classification model on a test dataset. It contains four key components: true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN). Figure 3-12 is an excellent visual presentation of the matrix. In this research, they are defined as below:
  - True Positive (TP) = Correctly classified instances of each architectural style.
  - False Positive (FP) = Instances incorrectly classified as the target architectural style.
  - True Negative (TN) = Instances correctly classified as not the target architectural style.
  - False Negative (FN) = Instances of the target architectural style that were not recognized.

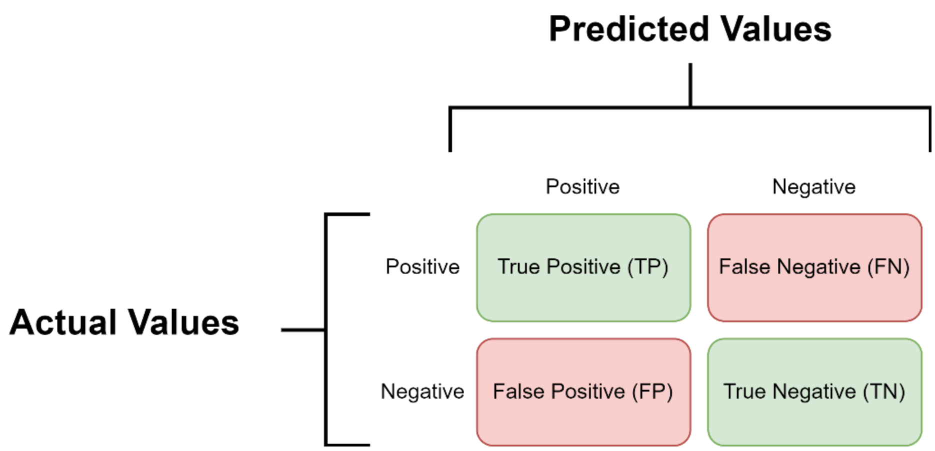

Figure 3-12 Confusion Matrix for Model Performance Evaluation

- **Test Accuracy:** This metric provides the overall rate of correct predictions made by the model. Higher accuracy indicates better overall performance but should be considered alongside other metrics to understand potential biases. The formula is shown below:

$$\text{Test Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$

- **Test Loss:** Test loss quantifies the average error per data point and aims to represent the disparity between the predicted values and the actual values. Lower test loss values indicate a model that predicts more closely to the actual labels. This metric is beneficial during the training phase to monitor optimization progress.

- **Precision:** Precision measures the accuracy of positive predictions. This metric is crucial in scenarios where the consequence of a false positive is significant. High precision indicates a low rate of false positives, suggesting that it is likely to be correct when the model predicts a positive result. The formula for calculating the precision is:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

- **Recall (Sensitivity):** It measures the model's ability to identify all actual positives from the data. It is especially important in situations where failing to detect positives could have severe consequences. The formula for recall is:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

- **F1-Score:** This metric is particularly useful when the data is unbalanced, as it maintains a balance between the precision and the recall. A high F1-Score indicates that the model successfully balances both precision and recall; conversely, a low F1-Score suggests that the model is performing poorly in either precision or recall, indicating a potential bias or inefficiency in handling one class over another. It is calculated as:

$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

- **Receiver Operating Characteristic (ROC) Curve and Area Under the ROC (AUC):** An ROC curve represents the average sensitivity value over all possible specificity values or vice versa. It generally interprets the probability that a randomly selected pair of patients can be correctly classified based on the given test results.

- AUC evaluates the model across all classification thresholds by measuring the area underneath the ROC curve. This curve plots the Recall against the false positive rate at various threshold settings. A higher AUC value (close to 1) suggests a better model, which correctly classifies positives and negatives across all thresholds. A model with an AUC close to 0.5 does not perform better than random guessing.

By thoroughly understanding and analyzing these metrics, model performances can be evaluated, and informed decisions about model adjustments and potential improvements can be made.

Chapter 4
Results and Discussion

4.1 Chapter Overview

This chapter comprehensively analyzes the results of the three differently-purposed DCNN models developed in Chapter 3 for architectural style classification, potential heritage identification, and heritage attribute designation prediction. Each section of the chapter is dedicated to the discussion of a specific model, evaluating its performance against a range of performance metrics outlined in Section 3.6. These results are visualized by detailed performance metrics in a variety of forms, including tables, charts, and confusion matrices, in order to provide a clear, quantitative understanding of the effectiveness of each model.

Firstly, this chapter provides a comprehensive evaluation of the frequently used DCNN models, analyzing the base performance of each model in preparation for subsequent fine-tuning. Subsequently, four architectures selected and fine-tuned are analyzed and compared. For the next step, two versions of the heritage designation models were evaluated to determine if they performed as expected. Finally, the hybrid MLP-CNN model for heritage designation prediction is analyzed, demonstrating its predictive ability in a real heritage conservation setting. Preliminary conclusions will be provided at the end of each section.

After presenting the quantitative results of each model, the chapter delves into the qualitative implications of these findings. This includes a critical evaluation of model performance, practical challenges encountered during model training and deployment, and the implications of these results for GeoAI and the field of heritage conservation. The discussion also covers the strategic decisions made during the model selection and fine-tuning process (as described in Chapter 3) and reflects on the impact of these decisions on the results observed in this chapter. Finally, the reflection section analyzed the challenges and areas for improvement in research design and dataset creation, which provided valuable insights for future studies to enhance model accuracy, robustness, and applicability in architectural style and heritage identification.

4.2 Architectural Style Classification Results

4.2.1 Model Selection Analysis

This section presents a general analysis of the performance of various DL architectures used in this study. The architectures were evaluated based on their final training loss, final training accuracy, final validation loss, final validation accuracy, and average epoch time. The results are summarized in Table 4-1. The validation accuracy over the epochs for each architecture is presented in Figure 4-1. This figure visually demonstrates the progression and stabilization of validation accuracy across the training epochs.

Table 4-1 Performance Metrics of Different Architectures

Architecture Name	Final Training Loss	Final Training Accuracy	Final Validation Loss	Final Validation Accuracy	Average Epoch Time (s)
EfficientNetB0	0.1255	0.9602	0.4063	0.8854	~492
MobileNetV3Large	0.1129	0.9638	0.6846	0.8207	~270
ResNet101	0.1867	0.9375	0.766	0.7878	~1293
ResNet50	0.2095	0.9269	0.7991	0.7754	~766
VGG16	0.2504	0.915	1.1339	0.7231	~1244
VGG19	0.2708	0.9097	1.2732	0.7056	~1689
DenseNet121	1.6356	0.4871	1.8405	0.428	~913
NASNetMobile	2.1741	0.3393	2.7888	0.2107	~350
InceptionV3	2.7954	0.1252	2.8296	0.1278	~371

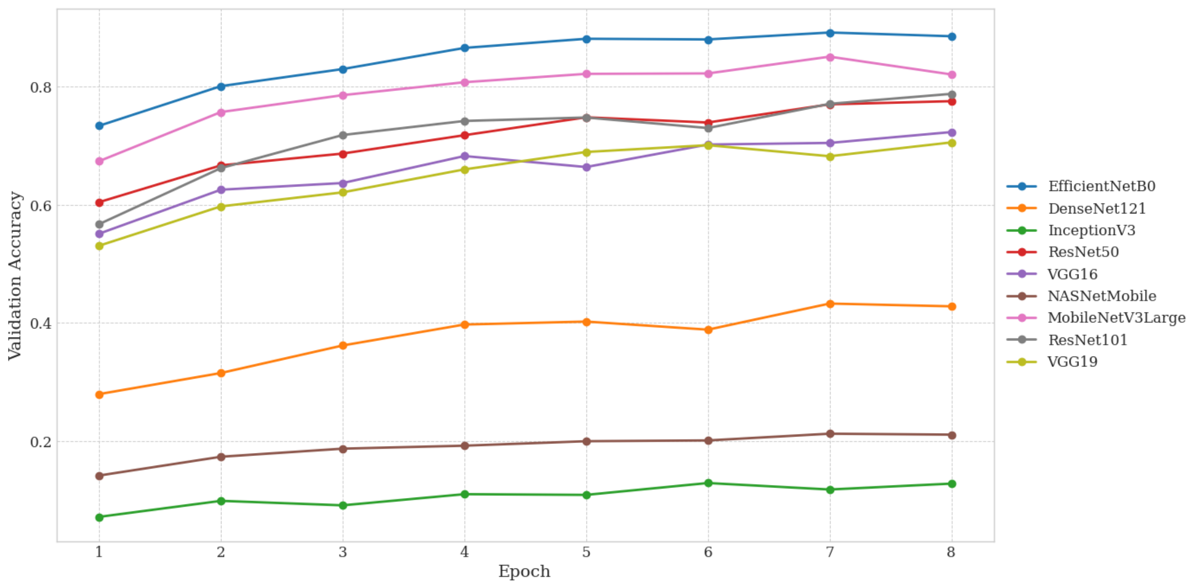

Figure 4-1 Validation Accuracy Across Epochs for Various Architectures

This step aims to identify the top four architectures demonstrating the best initial performance for further optimization through fine-tuning, thereby enhancing their capability to accurately classify architectural styles. After evaluating the DL architectures using the Ontario Architectural Style Dataset, the performance metrics were recorded over eight epochs. After each epoch, the architectures were tested with validation data, and the validation accuracy was calculated. Typically, the accuracy of the validation dataset is slightly lower than that of the training data, which provides insights into the architectures' generalization capabilities.

Among all the architectures, EfficientNetB0 demonstrated the best overall performance, achieving the highest final validation accuracy of 88.54% and a relatively low final validation loss of 0.4063. This architecture also maintained a high training accuracy of 96.02% and a low training loss of 0.1255, indicating excellent generalization and architecture efficiency. The average epoch time for EfficientNetB0 was moderate at approximately 492 seconds, making it a balanced choice in terms of both performance and training time. MobileNetV3Large also performed well, achieving a final validation accuracy of 82.07% and a very low training loss of 0.1129. With the shortest average epoch time of approximately 270 seconds, this architecture was highly efficient in terms of training. ResNet101 and ResNet50 also achieved good training accuracies of 93.75% and 92.69%, respectively, but showed a slight drop in final validation accuracy to 78.78% and 77.54%, respectively.

Some architectures did not perform well or were inefficient for this task. VGG16 and VGG19 had higher final validation losses of 1.1339 and 1.2732, respectively, and lower final validation accuracies of 72.31% and 70.56%, respectively, compared to other architectures. Their longer average epoch times of approximately 1244 seconds and 1689 seconds suggest they are less efficient in training, making them less desirable for this task. DenseNet121, NASNetMobile, and InceptionV3 exhibited significantly higher final validation losses and lower validation accuracies. DenseNet121, in particular, had a validation accuracy of only 42.8%, while NASNetMobile and InceptionV3 had even lower accuracies of 21.07% and 12.78%, respectively. These architectures were less effective for this particular dataset and task.

From the analysis, EfficientNetB0 stands out as the best-performing architecture, considering its validation accuracy and reasonable training time. MobileNetV3Large is a close second, especially for scenarios where training time is a significant concern. ResNet101 and ResNet50 also offer good performance but require more training time. Therefore, the top four architectures selected for further optimization through fine-tuning are EfficientNetB0, MobileNetV3Large, ResNet101, and ResNet50. These architectures provide a balance of high accuracy, efficiency, and potential for further performance improvements. Fine-tuning these architectures will be highly possible to enhance their capability to accurately classify architectural styles and identify potential heritage properties, making them suitable for deployment in real-world applications.

4.2.2 Fine-Tuned Model Performance Analysis

When creating the architectural style model, we adopted four architectures (EfficientNetB0, MobileNetV3Large, ResNet50, and ResNet101), pretrained and fine-tuned them. Table 4-2 shows the test accuracy and image size of each model after pre-training and fine-tuning. First, the test accuracy of EfficientNetB0 in the pre-training stage was 0.8958, but after fine-tuning, the test accuracy dropped to 0.8181. This indicates that the fine-tuning process did not effectively improve model performance but may have led to overfitting or data set mismatch problems. In contrast, the test accuracy of MobileNetV3Large in the pre-training stage was 0.8327, which improved to 0.8542 after fine-tuning. Although the improvement was not significant, it shows that fine-tuning has a certain optimization effect on model performance.

The test accuracy of ResNet50 in the pre-training stage was 0.8042, which improved significantly to 0.8764 after fine-tuning, indicating that the model benefited a lot from the fine-tuning process and could better handle specific data sets. The pre-training test accuracy of ResNet101 was 0.8060, which increased to 0.8872 after fine-tuning, which was the largest improvement among all models, showing the advantages of its deep structure in feature extraction and learning capabilities. Overall, ResNet101 performed best after fine-tuning, followed by ResNet50. Both models showed significant performance improvements during the fine-tuning process. The performance of EfficientNetB0 decreased after fine-tuning, and the fine-tuning strategy needs to be further optimized to improve its effect. Although MobileNetV3Large has not improved much, its performance after fine-tuning is better than that in the pre-training stage, indicating that it has certain adaptability to fine-tuning.

In this evaluation, ResNet101 and ResNet50 have emerged as the top performers for the task of architectural style classification. ResNet101, with its robust all-around performance, is deemed the optimal choice, closely followed by ResNet50, which excels particularly in precision and F1-Score. MobileNetV3Large offers a well-rounded alternative, striking a balance between efficiency and accuracy, suitable for scenarios where both are prioritized. EfficientNetB0, however, struggles with complex classifications and has been ruled out for this specific project.

Table 4-2 Pre-trained and Fine-tuned Test Accuracy

Model	Pre-trained Test Accuracy	Fine-tuned Test Accuracy	Image Size
EfficientNetB0	0.8958	0.8181	224×224
MobileNetV3Large	0.8327	0.8542	224×224
ResNet50	0.8042	0.8764	224×224
ResNet101	0.8060	0.8872	224×224

Given that my primary goal is to develop a high-accuracy classification model using a large dataset, accuracy is the paramount metric. ResNet101 stands out for its exceptional accuracy, making it the preferred model despite the strengths of the others.

Table 4-3 shows the performance of Fine-tuned (FT) Models on the test dataset. In terms of test accuracy, the FT-ResNet101 model performed the best, reaching 0.8872, indicating that it has the best overall performance in identifying architectural styles. FT-ResNet50 followed closely with a test accuracy of 0.8764, while FT-MobileNetV3Large and FT-EfficientNetB0 had test accuracies of 0.8548 and 0.8181. In terms of test loss, FT-ResNet101 once again took the lead with the lowest value of 0.4413, further proving its stability and reliability on the test set. In contrast, FT-EfficientNetB0 has a test loss of 0.5565, slightly lower than FT-MobileNetV3Large's 0.5595, while ResNet50 has a test loss of 0.6318.

In terms of precision, both FT-ResNet101 and FT-ResNet50 perform well, 0.8870 and 0.8764, respectively, showing their superiority in correctly classifying positive samples. FT-MobileNetV3Large has a precision of 0.8545, while FT-EfficientNetB0 has a precision of 0.8232. In terms of recall, FT-ResNet101 leads again with a high value of 0.8872, indicating that it works best when detecting architectural style categories that exist. FT-ResNet50 has a recall of 0.8764, while FT-MobileNetV3Large and FT-EfficientNetB0 have a recall of 0.8541 and 0.8180, respectively. As an important indicator that combines precision and recall, F1-Score still performs well for FT-ResNet101, reaching 0.8869, followed by FT-ResNet50 at 0.8759 and FT-MobileNetV3Large at 0.8538, and FT-EfficientNetB0 at 0.8166.

Finally, the AUC indicator shows the performance of the classifier in distinguishing different categories. FT-ResNet101 ranks first with an AUC value of 0.9408, demonstrating its excellent performance in classification tasks. The AUCs of FT-ResNet50 and FT-MobileNetV3Large are 0.9351 and 0.9234, respectively, while the AUC of FT-EfficientNetB0 is 0.9045.

In this evaluation, FT-ResNet101 and FT-ResNet50 have emerged as the top performers for the task of architectural style classification. FT-ResNet101, with its robust all-around performance, is deemed the optimal choice, closely followed by FT-ResNet50, which excels particularly in precision and F1-Score. FT-MobileNetV3Large offers a well-rounded alternative, striking a balance between efficiency and accuracy, suitable for scenarios where both are prioritized. FT-EfficientNetB0, however, struggles with complex classifications and has been ruled out for this specific project.

Table 4-3 Fine-tuned (FT) Model Performance on the Test Set

Model	**FT-EfficientNetB0**	**FT-MobileNetLarge**	**FT-ResNet50**	**FT-ResNet101**
Test Accuracy	0.8181	0.8548	0.8764	0.8872
Test Loss	0.5565	0.5595	0.6318	0.4413
Precision	0.8232	0.8545	0.8764	0.8870
Recall	0.8180	0.8541	0.8764	0.8872
F1-Score	0.8166	0.8538	0.8759	0.8869
AUC	0.9045	0.9234	0.9351	0.9408

Figure 4-2 further shows the training time of each model across different training rounds, while Figure 4-3 illustrates the changes in accuracy for each model over these rounds. The computational experiments were conducted on a machine powered by a 12th Generation Intel Core i7-12700K processor with a base clock speed of 3.60 GHz and equipped with 32 GB of RAM, providing ample processing power for complex model training. All operations were executed using the CPU of this 64-bit, x64-based system.

There are significant differences in training times among the models. FT-ResNet50 and FT-ResNet101 have much longer training times per round compared to FT-EfficientNetB0 and FT-MobileNetV3Large, with each round exceeding 30 minutes. Specifically, FT-ResNet101 initially takes up to 120 minutes per round, though this drops significantly after the eighth round. Subsequent training times for FT-ResNet101 fluctuate widely, likely due to dynamic adjustments in training resources or other process changes. In contrast, FT-MobileNetV3Large has the shortest training time, stabilizing at about 10-15 minutes per round, demonstrating high training efficiency.

From the perspective of accuracy, the accuracy of each model increases with the increase in training rounds, but the improvement speed and final accuracy of different models are different. The accuracy of FT-EfficientNetB0 and FT-MobileNetV3Large steadily increases with the increase of training rounds, while the accuracy of FT-ResNet50 and FT-ResNet101 quickly approaches the perfect level in the early stage of training and remains stable. Overall, FT-ResNet50 and FT-ResNet101 have advantages in terms of accuracy, but in terms of training efficiency, FT-MobileNetV3Large shows superior performance.

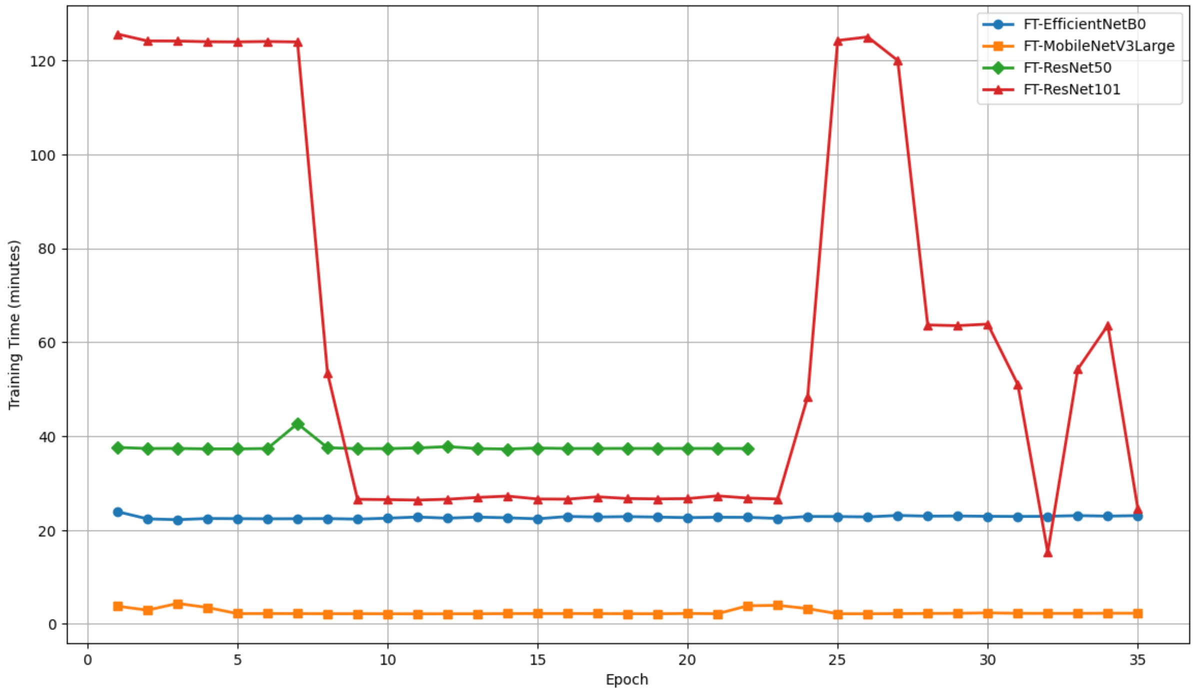

Figure 4-2 Comparison of Training Times per Epoch for Different Models

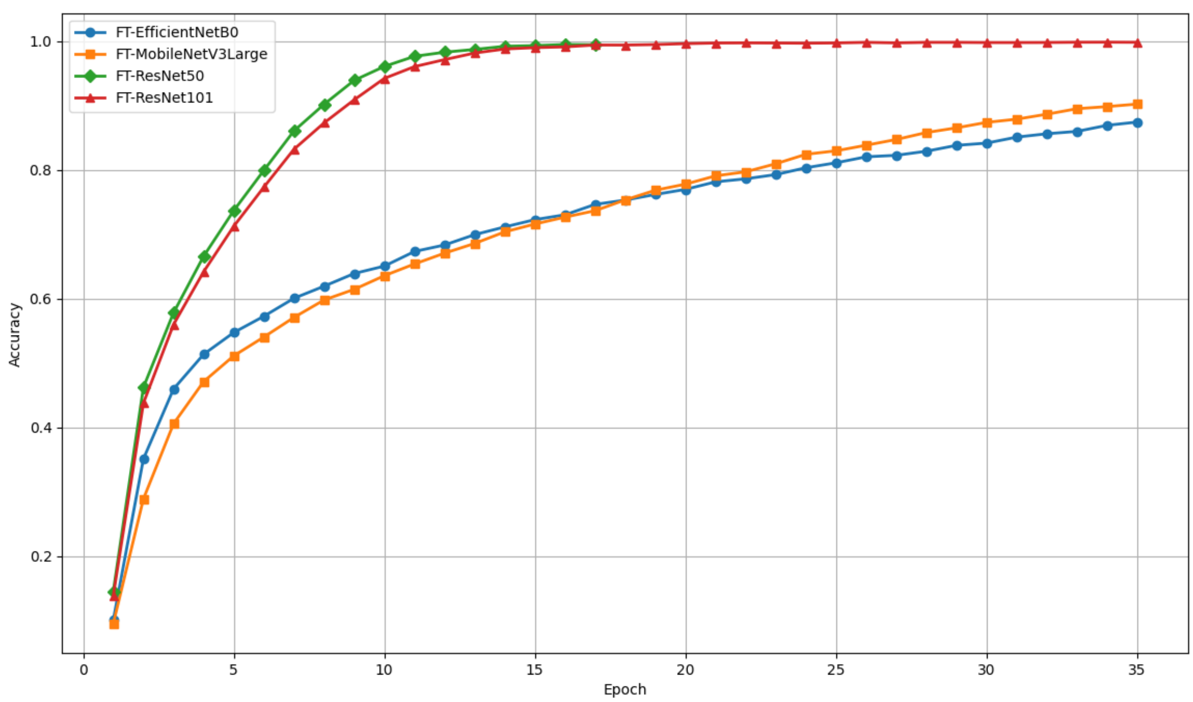

Figure 4-3 Comparison of Accuracy per Epoch for Different Models

Table 4-4 provides a comprehensive view of the accuracy of each model's predictions for different architectural styles. Overall, the Octagonal style had a classification accuracy of 100% across all models, indicating that the style was well captured and classified by the model due to the unique features of the octagonal staircase shape. The Neoclassical style also showed high classification accuracy, with all models achieving near-perfect scores. In contrast, the architectural styles with the lowest average accuracy were Edwardian and Victorian. The Edwardian style had an average accuracy of 71.81%, while the Victorian style had an average accuracy of 81.33%. The features of these styles are more complex or similar to other styles, resulting in a high misclassification rate for the models when identifying these styles. The Romanesque Revival style was also difficult for all models, with an average accuracy lower than the highest-accurate ResNet50 at only 83.45%. This indicates that the models face challenges in generalizing the features of this particular style.

Comparing the performance of different models in classifying architectural styles, ResNet50 performed well in most architectural styles, but performed poorly in styles with more complex features (such as Queen Anne and Victorian), with an accuracy of less than 80%. ResNet101 provides high accuracy in most styles, but there is room for improvement in some specific styles, such as Romanesque Revival. This indicates that additional training or data augmentation may be needed. EfficientNetB0 performs well, performing well in styles such as Victory House and Octagon, but shows weaknesses in more historically complex styles such as Edwardian. MobileNetV3Large, while slightly less consistent than ResNet101, performs well in Neoclassical and Octagonal styles, with accuracy close to or reaching 100%. However, performance is less good in Edwardian and Victorian styles.

Table 4-4 Accuracy of 21 Architecture Styles Across Four Fine-Tuned Models

Architecture Style	FT- **ResNet50**	FT- **ResNet101**	FT- **EfficientNetB0**	FT- **MobileNetV3Large**
Art (Streamline) Moderne	96.31%	94.63%	94.97%	88.93%
Art Deco	85.91%	91.28%	88.59%	87.25%
Beaux Arts	88.29%	86.96%	81.27%	81.94%
Colonial (Georgian Revival)	83.05%	86.78%	81.69%	81.02%
Edwardian	76.29%	75.60%	66.32%	68.04%
Georgian (combined Mennonite)	85.67%	87.67%	85.67%	81.67%
Gothic Revival	86.05%	84.01%	85.37%	85.37%

Greek (Classical) Revival	81.66%	86.85%	82.70%	79.58%
Italianate	84.80%	83.11%	85.47%	79.05%
Log Houses	93.52%	93.86%	95.90%	93.17%
Modern (International)	93.47%	97.94%	92.10%	94.50%
Neoclassical	99.66%	99.33%	100.00%	99.66%
Octagonal	100.00%	100.00%	100.00%	100.00%
Prairie Craftsman Bungalow	85.91%	89.60%	88.26%	83.56%
Queen Anne	73.99%	78.72%	75.34%	77.70%
Regency	86.62%	87.29%	90.64%	89.30%
Romanesque Revival	83.45%	83.11%	77.03%	77.03%
Second Empire	83.67%	86.67%	78.33%	79.00%
Tudor	95.83%	93.06%	94.79%	91.67%
Victorian	78.26%	79.26%	78.26%	76.59%
Victory Housing	97.99%	97.32%	98.99%	98.66%

4.2.3 Comparison of Architectural Style Identification Model

Table 4-5 Comparative Analysis of Architectural Style Classification Models

Models	**Key Strengths**	**Key Weaknesses**
FT-**ResNet101**	Highest overall accuracy	Longest training time
FT-**ResNet50**	High precision and F1-Score	Longer training time than some models
FT-**MobileNetV3Large**	Fastest training time	Lower accuracy compared to ResNet models
FT-**EfficientNetB0**	Good for simpler styles	Underperforms with complex styles

Table 4-5 shows the key strengths and weaknesses of each model based on the results. After conducting a comparative analysis of various CNN models used for architectural style classification, it becomes apparent that the selection of the most appropriate model heavily depends on the specific application needs and the resources available. The FT-ResNet101 model emerges as a top contender due to its highest overall accuracy, making it exceptionally suitable for this scenario. Given its robust performance, FT-ResNet101 is particularly valuable in detailed architectural analysis and preservation efforts, where even minor errors can lead to significant misjudgments in heritage conservation.

However, in terms of the efficiency versus accuracy trade-offs, FT-MobileNetV3Large presents a compelling case. It offers a very balanced approach, with commendable accuracy and exceptionally fast training times. This architecture is particularly advantageous for applications requiring rapid decision-making, such as real-time recognition of architectural styles in mobile apps or during preliminary surveys where the speed of analysis may be more critical than extremely high precision.

In conclusion, the decision may favour more resource-efficient architectures in cases where resources are severely limited, whether due to time, money, or computational issues. In these kinds of situations, FT-MobileNetV3Large provides a good compromise. Additionally, FT-EfficientNetB0 could be considered for simpler tasks. While it underperforms with complex styles, FT-EfficientNetB0 is cost-effective for classifying less intricate designs and might be ideal for educational purposes where the utmost precision is not the primary concern.

Given that accuracy is the primary concern in my current research, FT-ResNet101 is the preferred architecture due to its superior performance in classifying complex architectural styles accurately. Consequently, for the development of the next version of the potential heritage identification model, particularly in the application of transfer learning techniques, I have opted to utilize FT-ResNet101 with all considerations.

4.3 Heritage Identification Models Evaluation

4.3.1 Version 1 – without Transfer Learning

The performance evaluation of the Heritage Identification Model (Version 1) is presented in this section. Based on the obtained metrics shown in Table 4-6, the model achieved an overall accuracy of 96.62%, indicating a very high level of correctness in its predictions. The associated loss was recorded at 0.0915, which also demonstrates the model's efficiency during the training process.

The Version 1 model's performance was further examined using precision, recall, and F1-score metrics for both heritage and non-heritage categories. For the heritage class, the precision was 0.97, indicating that 97% of the instances predicted as heritage were correctly identified. The recall was 0.96, showing that the model correctly identified 96% of actual heritage instances. The F1-score, which balances precision and recall, was also 0.97 for the heritage category. Similarly, the non-heritage category exhibited a precision of 0.96 and a recall of 0.97, with an F1-score of 0.97, showcasing the model's balanced and consistent performance across both classes.

Additionally, the training and validation accuracy over epochs, as shown in Figure 4-4, shows that the model converges well with high accuracy for both training and validation sets. The training accuracy quickly reaches near 100%, while the validation accuracy stabilizes around 97%. The corresponding loss curves have a rapid decrease in both training and validation losses, with the training loss approaching zero and the validation loss stabilizing at a low value. These trends suggest that the model is well-fitted to the training data and generalizes effectively to unseen data.

Table 4-6 Classification Metrics of Heritage Identification Model (Version 1)

	precision	recall	f1-score	support	accuracy	loss
heritage	0.97	0.96	0.97	117		
non-heritage	0.96	0.97	0.97	120		
overall					0.9662	0.0915

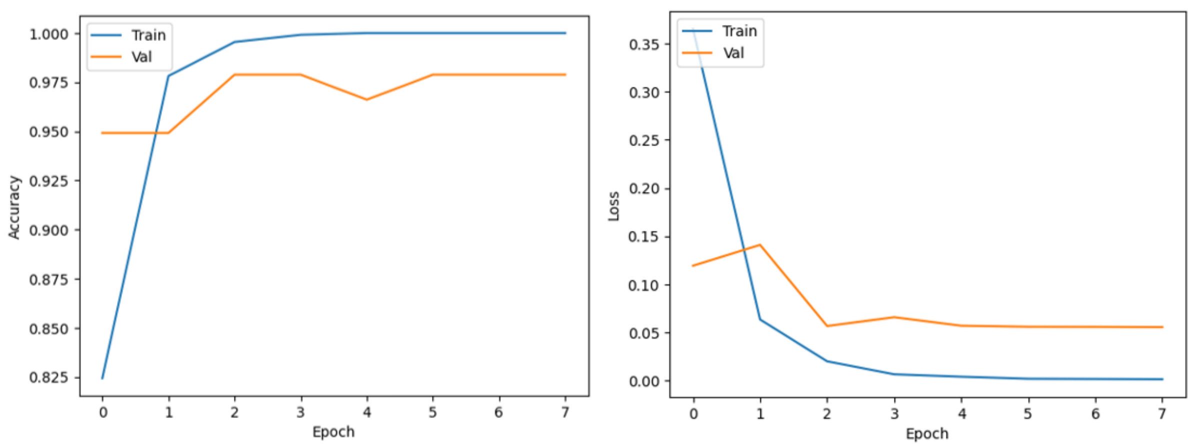

Figure 4-4 Accuracy and Loss Graph of Heritage Identification Model (Version 1)

The confusion matrix (Table 4-7) also provides a clear view of the model performance in terms of correctly and incorrectly classified instances. The true positive count for heritage predictions was 112, while the false negative count was 5. For non-heritage predictions, the true negative count was 117, with a false positive count of 3. These results collectively indicate that the model has a high level of reliability with minimal critical errors.

Table 4-7 Confusion Matrix on Test Dataset

	Predicted Heritage	Predicted Non-Heritage
True Heritage	112 (TP)	5 (FN)
True Non-Heritage	3 (FP)	117 (TN)

To further illustrate the model's performance, Figure 4-5 shows a selection of images with their actual and predicted classifications. This visual representation complements the quantitative analysis, demonstrating the model's ability to accurately distinguish between heritage and non-heritage instances in real-world scenarios. The majority of the predictions align correctly with the actual classifications, reinforcing the model's effectiveness.

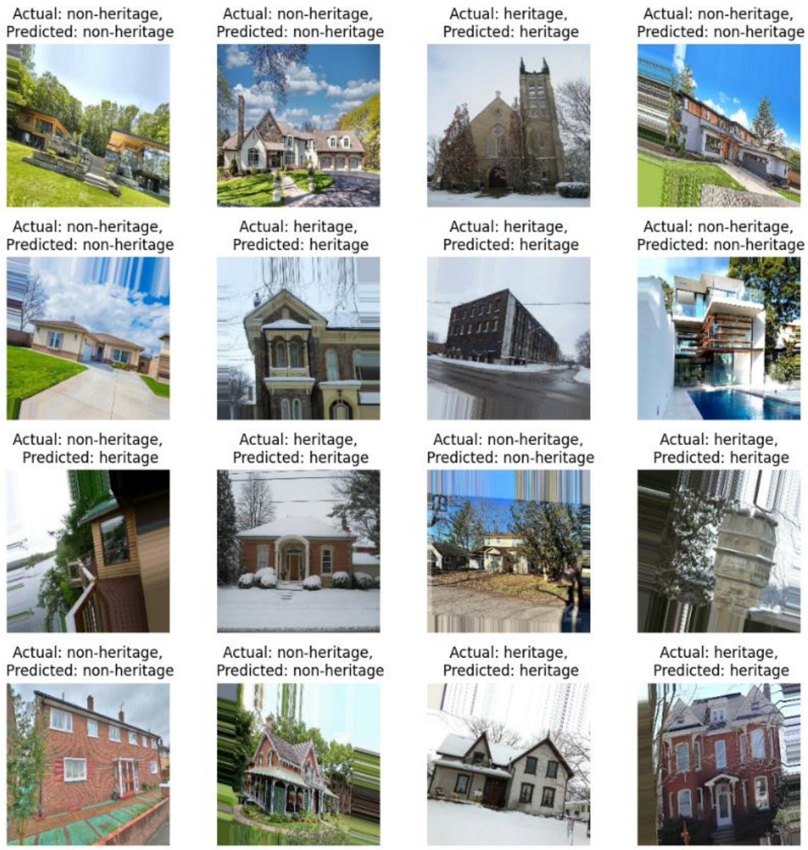

Figure 4-5 Visualization of Classification Result of Heritage Identification Model (Version 1)

4.3.2 Version 2 – with Transfer Learning

The results of Version 2 did not meet expectations. The performance of Version 2 in identifying "Heritage" is only 0.49. This indicates that while the model is able to efficiently identify most samples in the heritage category, its accuracy is only 49%. This means that about half of the results are misclassifications, with non-heritage samples being incorrectly categorized as heritage. For the category of "Non-Heritage", the model's performance is even more disappointing. The precision is only 0.44, the recall is as low as 0.03, and the F1 score is just 0.06. This means the model hardly recognizes non-heritage samples correctly, with most non-heritage samples being misclassified as heritage. Figure 4-6 shows the classification results of Version 2 by printing the actual and predicted labels of some images in the test dataset. The visualization results show that the model's classifications are more concentrated on heritage data but scattered on non-heritage data, which is significantly inconsistent with the actual situation. This further confirms the model's inadequacy in non-heritage classification.

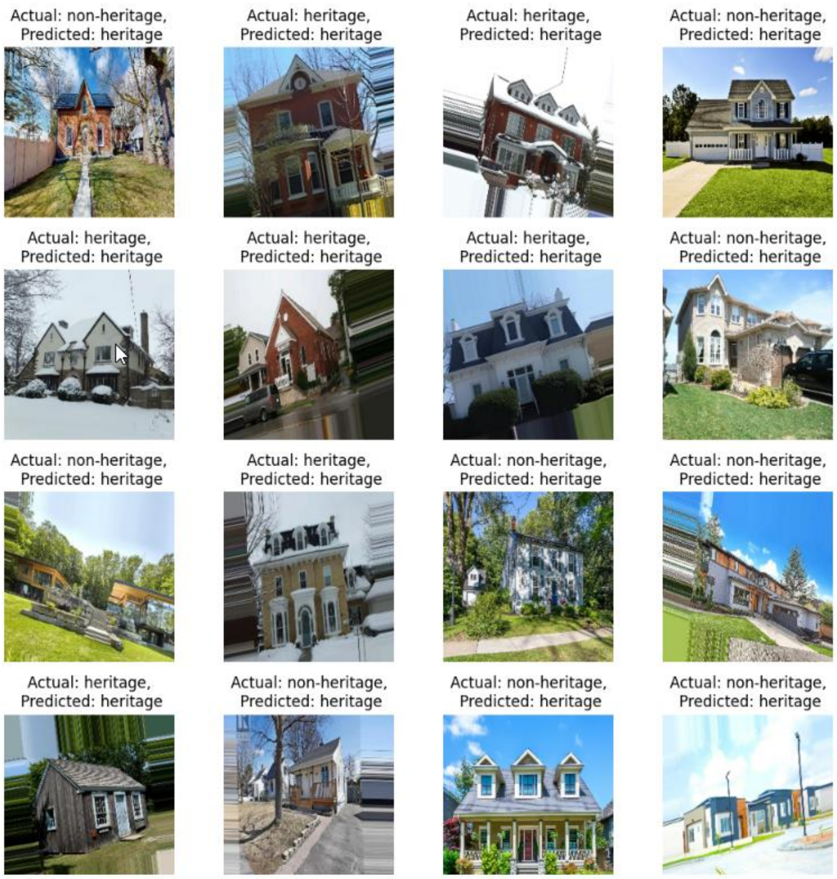

Figure 4-6 Visualization of Classification Result of Heritage Identification Model (Version 2)

Table 4-8 Classification Metrics of Heritage Identification Model (Version 2)

	precision	**recall**	**f1-score**	**support**	**accuracy**	**loss**
heritage	0.49	0.96	0.65	117		
non-heritage	0.44	0.03	0.06	120		
overall					0.4895	2.2640

4.3.3 Discussion on Heritage Identification

In this study, Version 1 employs a direct training approach, utilizing a pre-trained architecture in the Keras library to learn directly from the Ontario Heritage dataset, without relying on further pre-trained models or external datasets. This approach performs very well on the task of distinguishing heritage from non-heritage, showing high accuracy and low error rates. This suggests that a training strategy customized for a specific dataset can be very effective, especially when dealing with tasks with clear classification goals.

The exact opposite is true for version 2, which employs a migration learning approach based on an Architectural Style Classification Model in the last section. The aim of this strategy is to use the existing knowledge base in order to enhance the performance of the model on the heritage classification task. This model, despite its ability to recognize heritage categories, however, the application of transfer learning did not achieve the expected results with a low accuracy rate. This may be since the features of the pre-trained model are not fully adapted to the heritage data during the tuning process, or because the migration learning strategy is not fine-grained enough at the tuning level to efficiently transform the learned features to match the specific needs of heritage recognition. In addition, the low recall and low F1 scores also indicate the inadequacy of the model in recognizing non-legacy samples, which may be related to the architecture of the model or the choice of the loss function.

These findings suggest that in future model design and training, it is important to not only choose appropriate training methods, but also pay more attention to model construction and data adaptation. For direct training methods, although it performs well on specific datasets, it may lack generalization ability. For the migration learning-based approach, although it can theoretically improve the generalization ability of the model, in practice, it is necessary to ensure that the migrated features are highly relevant to the target task and the adaptation strategy can effectively cope with new classification challenges.

To summarize, these two approaches have their own strengths and limitations, and future research should explore how to combine the advantages of these two strategies or develop new approaches to better adapt to specific heritage recognition tasks. Through continuous optimization of the model construction details and in-depth study of the training strategies, we can further improve the performance and applicability of the model and provide more reliable technical support for the conservation and research of heritage buildings.

4.4 Heritage Property Designation Prediction Model Evaluation

4.4.1 Comparative Analysis of CNN, MLP, and Hybrid Model

The evaluation of three models—CNN only, MLP only, and a hybrid CNN-MLP—for heritage property designation prediction indicates some performance differences (Table 4-9). The hybrid model exhibits the highest test accuracy at 96% and the lowest test loss at 0.23, outperforming the CNN-only model with a test accuracy of 92% and a test loss of 0.25. The MLP-only model has the lowest performance, with a test accuracy of 76% and a test loss of 0.60, indicating its limitations in handling this classification task.

The precision, recall, and F1-score metrics further demonstrate the hybrid model's superiority. The hybrid model achieves perfect precision (100%) in the "Designated" category and high precision (95%) in the "Non-Designated" category. In addition, the hybrid model's recall rates are impressive, with 86% for "Designated" and 100% for "Non-Designated." Overall, the hybrid model is robust enough to identify true positives across both categories.

The CNN-only model also performs well but is slightly behind the hybrid model. It correctly classified most samples, but its precision and recall are marginally lower. Specifically, it classifies 17 out of 18 non-designated samples and 6 out of 7 designated samples correctly, indicating strong generalization but some challenges with certain edge cases.

The MLP-only model, however, struggles significantly, misclassifying a notable number of samples. It correctly identifies 15 non-designated samples but incorrectly labels three as designated, and while it correctly identifies five designated samples, it misclassifies two as non-designated. This suggests that the MLP-only approach may lack the capacity to handle the complexity of the image data effectively.

Table 4-9 Confusion Matrix, Evaluation Metrics, Accuracy, and Loss for CNN, MLP, and Hybrid CNN-MLP Models

Method Applied	Confusion Matrix			Evaluation Parameter			Accuracy	Loss
		True Class						
CNN only		Positive	Negative	Precision	Recall	F1-Score	0.92	0.25

	Positive	15	2	0.86	0.86	0.86		
	Negative	3	5	0.94	0.94	0.94		
MLP only		True Class		Precision	Recall	F1-Score	0.76	0.60
		Positive	Negative					
	Positive	14	2	0.56	0.71	0.82		
	Negative	4	5	0.88	0.78	0.82		
Hybrid CNN-MLP Model		True Class		Precision	Recall	F1-Score	0.96	0.23
		Positive	Negative					
	Positive	18	1	1.00	0.86	0.92		
	Negative	0	6	0.95	1.00	0.97		

Figure 4-7 illustrates these three models' training and validation accuracy trends. The CNN model achieved 100% training accuracy by the 24th epoch, maintaining between 99% and 100% thereafter, showing rapid convergence and efficient learning. The MLP model reached 100% training accuracy at the 50th epoch, suggesting a more gradual learning process. The Hybrid model displayed fluctuating training accuracy, but it remained above 90% for most epochs, indicating strong learning capability despite instability.

Regarding validation accuracy, the CNN model improved after the 24th epoch, reaching 92% by the 35th epoch and remaining stable, indicating good generalization. The MLP model's validation accuracy gradually increased from 76% to 80% between the 30th and 40th epochs but stabilized at 76% after the 70th epoch, showing slower improvement in generalization. The Hybrid model's validation accuracy was initially volatile but showed a consistent upward trend, reaching 96% by the 40th epoch and maintaining this level, indicating excellent generalization after initial instability.

Comparing training to validation accuracy reveals overfitting in the CNN model, with 100% training accuracy versus 92% validation accuracy, suggesting good training data learning but slightly lower performance on unseen data. The MLP model also showed overfitting, with 100% training accuracy and 76% validation accuracy, indicating reasonable generalization despite overfitting. The Hybrid model demonstrated minimal overfitting, with both training and validation accuracies above 90% and validation accuracy reaching 96%, showing robust generalization.

The CNN model exhibited fast convergence, with stable validation accuracy by the 35th epoch, indicating high efficiency and reliability in training. The MLP model converged more slowly, with stable validation accuracy only after the 70th epoch, reflecting lower stability and slower learning. The Hybrid model, despite initial fluctuations, showed high stability and fast convergence after the 40th epoch, with consistently high validation accuracy at 96%, indicating efficient learning and robust performance.

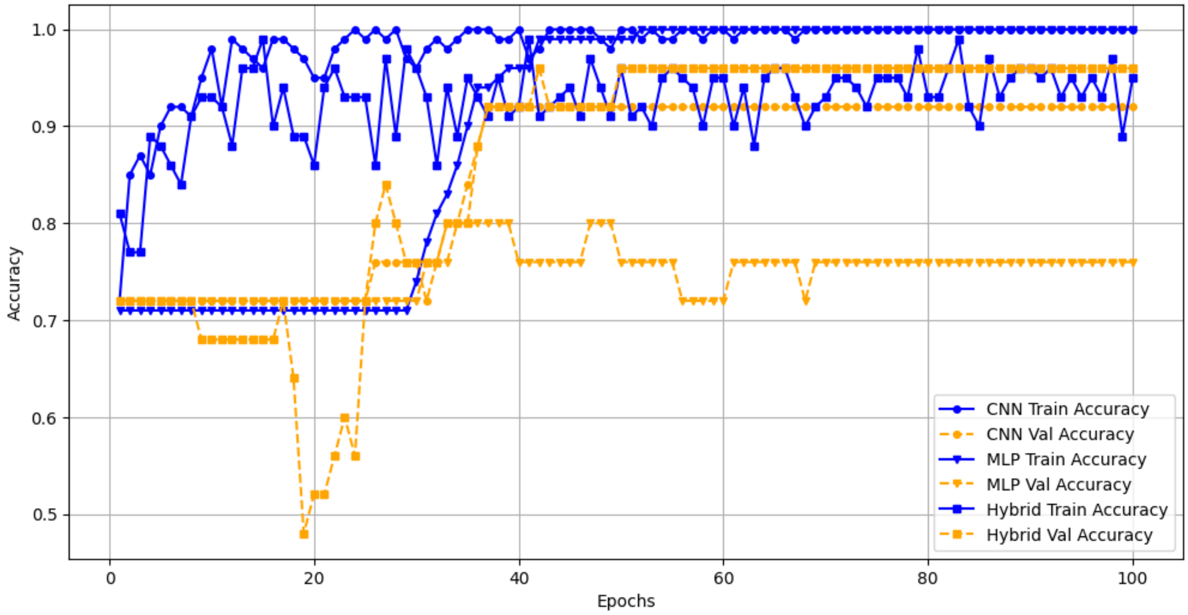

Figure 4-7 Loss and Accuracy Graph of CNN, MLP and Hybrid CNN-MLP Model

4.4.2 Discussion on the Heritage Property Designation Prediction Task

This study compares the performance of three models (CNN, MLP, and hybrid) in the task of predicting whether a building is more likely to be designated as a heritage property. Each model has its strengths and weaknesses under different evaluation criteria, as shown in Table 4-10. Overall, the CNN model has the advantage of fast convergence, being the fastest of the three models to reach 100% training accuracy and maintaining high accuracy thereafter. Much of this fast learning is attributed to the powerful feature extraction capabilities of CNNs, which are particularly suitable for image-based data. However, the slight overfitting observed suggests that while the model performs well on training data, its ability to generalize to unseen data could be improved. This overfitting can be mitigated by techniques such as dropout, data augmentation, and regularization, thus improving the model's performance on the validation set.

In comparison, the MLP model shows a more gradual learning process, reaching 100% training accuracy at the 50th epoch. The slower convergence may be related to the simpler architecture of MLP as it lacks the complex feature extraction layer of CNN.The validation accuracy of MLP eventually stabilized at 76%, suggesting that although it is capable of learning, it may have difficulty in dealing with complex patterns in the data. This implies that MLP could better capture data complexity through feature engineering or introducing more layers.

The hybrid model combines the strengths of CNN and MLP to show the best overall performance. It achieves high training accuracy and maintains stable validation accuracy despite initial fluctuations. The model leverages the feature extraction capabilities of CNN and the decision-making capabilities of MLP for robust learning and generalization. The hybrid approach effectively mitigates the limitations of each model, providing a balanced solution that can handle both feature extraction and efficient classification.

The superior performance of the hybrid model is attributed not only to its combination of the strengths of CNN and MLP but also to the effective utilization of mixed data, including image and tabular data. The hybrid model's ability to process both image and tabular data (both categorical and numerical) makes it perform more comprehensively in feature extraction and decision-making capabilities. By utilizing the powerful image feature extraction capability of CNN, the model can extract important visual features from building images, while by processing the tabular data through MLP, the model is able to effectively incorporate and utilize non-image information, such as the year and style of the building, whether it is associated with a famous person or event, and whether it is located in a HCD. This advantage allows the hybrid model to significantly outperform models with a single data type regarding comprehensive performance, demonstrating higher accuracy and stability when dealing with complex and diverse datasets.

The results of the study also validate Ahsan et al. (2020). Despite the different application areas, mixed data and mixed models can create highly accurate models on small and balanced datasets. This approach performs well when dealing with complex and diverse datasets, showing higher accuracy and stability.

To summarize, applying hybrid models in assisted heritage designation is an innovative attempt, but there are some limitations. The currently used hybrid model has not yet experimented with different architectures and integration methods to optimize performance. Future research should explore various regularization techniques and advanced training strategies (e.g., learning rate scheduling and integration learning), apply the hybrid model to architectural heritage datasets from other cities, etc. This would help to enhance the model's robustness and generalizability, providing more reliable and accurate heritage designation predictions across diverse urban contexts.

Table 4-10 Evaluation of Different Techniques and Recommendations

Evaluation Criteria	Recommended Model	Reason
Training Speed	CNN	Achieved 100% training accuracy within 24 epochs, fastest convergence speed.
Validation Accuracy	Hybrid	Reached 96% validation accuracy after 40 epochs and maintained stability.
Stability	Hybrid	Despite initial fluctuations, high stability in validation accuracy was maintained after 40 epochs.
Overfitting	Hybrid	Minimal overfitting with both training and validation accuracies above 90%.
Overall Performance	Hybrid	Highest precision, recall, and F1 score, best overall performance.
Feature Extraction Ability	CNN	The strongest feature extraction capability is due to CNN architecture.
Suitability for Simple Tasks	MLP	Simple architecture, suitable for lower complexity tasks, easier to understand and implement.
Suitability for Complex Tasks	Hybrid	It combines the strengths of CNN and MLP and is suitable for tasks requiring high accuracy and stability.
Resource-Constrained Environments	MLP	Requires fewer computational resources due to its simple structure, suitable for resource-constrained environments.
Future Improvement Potential	Hybrid	Potential for further optimization in architecture and training strategies for better performance in various tasks.

Chapter 5
Conclusion

5.1 Summary of Key Findings and Revisiting Research Questions

This section summarizes the key findings from evaluating the models developed in this study and revisits the central research questions to assess how effectively these models addressed the initial challenges and objectives. The primary question this research aimed to answer is: How can GeoAI technology accelerate and optimize the current heritage designation process in Ontario? To explore this overarching question, the study focused on two specific sub-questions:

1. How do GeoAI models function in the initial screening stage of the heritage designation process? Additionally, how accurately and efficiently can these models identify architectural styles and assess potential heritage buildings?
2. How does the integration of archival photographs and geospatial information improve the efficiency of the heritage designation evaluation process?

To address these sub-questions, the study developed and evaluated three models: the Architectural Style Classification Model, the Heritage Identification Model, and the Heritage Property Designation Prediction Model. These models utilized the Ontario Architectural Style Dataset, the Ontario Built Heritage Dataset, and the Stratford Heritage Geospatial Dataset, respectively.

The Architectural Style Classification Model demonstrated high accuracy in identifying architectural styles from images, with the best performance achieved by ResNet101, which reached nearly 89% accuracy after fine-tuning. This model proved particularly effective in processing and classifying complex image data, showcasing AI’s capability to recognize detailed architectural features. However, the model faced challenges when dealing with architectural styles that have subtle or similar features, such as Edwardian and Victorian styles, sometimes resulting in misclassifications. Overall, this model significantly answered the first sub-question, highlighting AI’s potential in accurately identifying architectural styles.

The Heritage Identification Model, particularly Version 1, improved the accuracy and efficiency of identifying heritage buildings, achieving an accuracy rate of 96.62%. Contrary to expectations, Version 2, which employed transfer learning from the Architectural Style Classification

Model, did not perform as well, with an accuracy rate falling below 50%. This discrepancy may be due to the pre-trained model features not fully adapting to the heritage data during tuning or the transfer learning strategy not being fine-grained enough to translate learned features to meet specific heritage identification needs effectively. This finding suggests that while AI significantly enhances heritage identification, further optimization of model parameters, algorithms, and dataset expansion is necessary to achieve optimal performance.

The Heritage Property Designation Prediction Model, utilizing geospatial information and archival photographs, successfully predicted the designation status of buildings with an impressive 96% accuracy rate. This model stands out from previous models that primarily leveraged AI and image classification by integrating sophisticated geospatial analysis. Significantly, this GeoAI model effectively addressed the challenges posed by the second sub-question, demonstrating the critical role that geospatial data plays in refining the heritage designation process. The successful application of geospatial data, analyzed through GeoAI, highlights the unique and substantial benefits of GeoAI, confirming its effectiveness and superiority in addressing specific research queries in heritage classification.

Overall, the findings from these models provide substantial evidence supporting the potential of GeoAI in heritage research. By demonstrating high accuracy and efficiency, the Architectural Style Classification Model and the Heritage Identification Model successfully addressed the first sub-question. The Heritage Property Designation Prediction Model, through integrating multiple data sources, illustrated how the heritage designation process can be improved, effectively responding to the second sub-question. These findings offer important insights into the capabilities and limitations of GeoAI in heritage conservation.

5.2 Methodological and Applied Contributions

This thesis has made several significant contributions to architectural heritage preservation and the advancement of GeoAI technology. By moving beyond the predominant research focus on architectural style analysis and specific element extraction, this research introduces innovative models that determine whether buildings may qualify as heritage buildings directly which could be of practical value for government-led conservation efforts. This shift from style identification to comprehensive heritage assessment points to a range of new opportunities to advance heritage research and practice with recent AI and GeoAI advancements.

Moreover, while most past studies employed AI techniques that depend solely on image analysis for heritage classification, this thesis demonstrates how GeoAI approaches can improve identification and classification accuracy. This deeper contextual understanding, unattainable through traditional image analysis alone, leverages GeoAI to analyze complex spatial datasets in conjunction with archival photographs and historical data.

Further advancing these technological innovations, the thesis showcases the adaptability and optimization of GeoAI models for heritage tasks. The successful deployment of these models illustrates their versatility and establishes a blueprint for their application in other areas such as urban planning, archaeological exploration, and placemaking. The integration of diverse data sources into GeoAI models not only enhances their performance and accuracy but also exemplifies a practical method for managing complex datasets related to heritage management. Building on these applications, it is clear that the GeoAI techniques have become particularly valuable in evolving some policy contexts, such as Ontario's "The More Homes, Built Faster Act" (Bill 23). It exemplifies how GeoAI can adapt to and positively impact heritage management amidst shifting regulatory landscapes, contributing to more informed and efficient policy implementations.

In conclusion, this thesis not only contributes to the field of heritage preservation by integrating advanced technological methods but also offers a comprehensive framework for future research and practical applications. It may help the effective and equitable protection of cultural heritage, advocating a significant shift in how heritage preservation is approached and executed.

5.3 Limitations and Future Work

This study provides technically feasible ideas for identifying and protecting built heritage using GeoAI models; however, it also faces several limitations. Firstly, each model was designed to assist planners with specific Ontario Heritage Designation Process tasks. For example, the precondition for using the Heritage Property Designation Prediction Model is that the building in question must already be listed as a non-designated property on the Municipal Heritage Register. A comprehensive consideration for future work would involve expanding the applicability of these models to accommodate a broader range of scenarios and adjusting them to cater to diverse situations effectively. This expansion could enhance the models' utility, making them more adaptable and useful across different contexts within the heritage designation process. Additionally, to further enhance accessibility and ease of use, a web application could be developed. This would allow planners without a technical background to utilize the models more easily, thereby streamlining the process and making it more user-friendly.

Dataset creation is also one of the primary challenges, particularly incomplete and noisy data, which directly affects the reliability of the model results. Our architectural style recognition model relies heavily on a dataset collected automatically through a Python script in Bing Images. Keyword searches sometimes lead to image category confusion, such as the same results appearing for Queen Anne-style and Italianate-style buildings. Although we attempted to correct these errors through manual inspection, the sheer volume of data made it challenging to avoid omissions entirely.

Additionally, data collection was slow for the potential cultural heritage identification model, and the dataset size was small due to time constraints and the lack of publicly available datasets from many cities. Although data enhancement techniques partially addressed this issue, they were insufficient to fully overcome the dataset limitations. In future research, it is suggested that more accurate data collection tools be developed, data sources expanded, and collaborations with research organizations, universities, and communities be established.

Another limitation is the lack of diverse images for each building. We initially included only single photos of building fronts, aiming to quickly capture the most architectural details. However, this approach did not fully capture the diversity and complexity of the buildings. Different viewing angles can provide unique and complementary perspectives on a building's style and structure, thus enhancing the model's ability to recognize architectural details and stylistic variations. Therefore, utilizing Street View imagery as a data source is a superior option as it can provide more comprehensive and three-dimensional visual information by showing buildings from multiple directions. We propose collecting and including photographs from multiple angles for each building in future dataset construction. This approach will significantly enrich the information density of the dataset and help improve model performance.

This study also has limitations and room for improvement at the model-building level. Firstly, we used pre-trained models in Keras to identify architectural styles and potential heritage sites. These models, based on the ImageNet dataset, can be quickly implemented. However, the ImageNet dataset is not fully suited for architectural image analysis, which leads to limitations in this domain and affects the effectiveness of parameter tuning. Moreover, we selected nine prominent models from existing literature and research for the architectural style recognition task. However, optimizing parameters and algorithms remains a key challenge, requiring significant computational resources and rigorous experimentation.

Moving forward, exploring hybrid models that combine the strengths of different models is a promising direction. By integrating the accuracy of ResNet101 with the efficiency of MobileNetV3Large, we can develop models that excel in both accuracy and speed. Furthermore, we can explore more advanced architectures, such as the Transformer model, which captures long-range dependencies within the data, thereby enhancing performance in image recognition tasks (Dai et al., 2021). This approach has already been successfully applied to medical and remote-sensing images, producing remarkable results (Bazi et al., 2021; Dai et al., 2021).

5.4 Ethical Considerations and Societal Impact

If AI models become integral to heritage identification and preservation, it is essential to address the ethical complexities associated with automated decision-making. This includes potential biases within the models, the accuracy and fairness of their decisions, and transparency. First, imbalances in the training data or biases in the historical data may lead to biases in the model outputs and under- or overestimate certain architectural styles. In addition, models may make errors in recognizing and classifying heritage buildings, and these errors may impact conservation efforts. Therefore, it is necessary to establish a strict review and calibration mechanism to ensure the accuracy and fairness of modeling decisions.

Another issue of concern is the conflict between automated decision-making and human work. Although automation and AI technologies can greatly improve efficiency, they always operate based on human standards. For example, in the OHA, there are rigorous review steps and dozens of detailed evaluation criteria for the heritage designation process that were developed by human experts. As such, automation and machines should be used as aids, not as complete replacements for human work. Currently, AI models still have many shortcomings and cannot completely replace humans in making perfect decisions. Therefore, when using these technologies, it is important to ensure that they are designed to support and augment the work of human experts and not replace their professional judgment.

In cultural heritage preservation, enhancing the transparency of the model decision-making process and promoting public participation are crucial. To achieve this, it is important to disclose decision standards, the sources of training data, and the methods used for model evaluation, thereby explaining to the public how the model works and the basis for its decisions.

Lastly, by widely soliciting public opinions and feedback, we can better understand how the public values built heritage may enhance the public's sense of responsibility and involvement in built heritage preservation. This will not only enhance the public's trust in conservation work but also increase their participation in heritage protection, thereby achieving more comprehensive and sustainable cultural heritage protection.

5.5 Concluding Remarks

The application of artificial intelligence in the preservation of architectural cultural heritage holds immense potential and opportunities. By leveraging AI technologies, it is possible to protect and manage architectural heritage more efficiently, gain deeper insights into historical and cultural developments, and promote the transmission of urban culture. This thesis has demonstrated how GeoAI models can enhance the accuracy and effectiveness of heritage identification and preservation, bridging the gap between traditional practices and modern technological advancements. As we continue to improve these technologies and address ethical considerations, the integration of AI will undoubtedly play a pivotal role in safeguarding our cultural heritage for future generations.

References

Abed, M. H., Al-Asfoor, M., & Hussain, Z. M. (n.d.). *Architectural heritage images classification using deep learning with CNN*.

Ahsan, M. M., E. Alam, T., Trafalis, T., & Huebner, P. (2020). Deep MLP-CNN Model Using Mixed-Data to Distinguish between COVID-19 and Non-COVID-19 Patients. *Symmetry*, *12*(9), Article 9. https://doi.org/10.3390/sym12091526

Albu, S. (2021). The Economic Value and Valuation of Architectural Heritage. *Journal of Building Construction and Planning Research*, *9*(1), Article 1. https://doi.org/10.4236/jbcpr.2021.91001

Al-Sakkaf, A., Zayed, T., & Bagchi, A. (2020). *A Review of Definition and Classification of Heritage Buildings and Framework for their Evaluation*.

Amsterdam Declaration. (1975).

Athanasopoulou, K., Daneva, G., Adamopoulos, P., & Scorilas, A. (2022). Artificial Intelligence: The Milestone in Modern Biomedical Research. *BioMedInformatics*, *2*, 727–744. https://doi.org/10.3390/biomedinformatics2040049

Bazi, Y., Bashmal, L., Rahhal, M. M. A., Dayil, R. A., & Ajlan, N. A. (2021). Vision Transformers for Remote Sensing Image Classification. *Remote Sensing*, *13*(3), Article 3. https://doi.org/10.3390/rs13030516

Belhi, A., Ahmed, H. O., Alfaqheri, T., Bouras, A., Sadka, A. H., & Foufou, S. (2021). Study and Evaluation of Pre-trained CNN Networks for Cultural Heritage Image Classification. In A. Belhi, A. Bouras, A. K. Al-Ali, & A. H. Sadka (Eds.), *Data Analytics for Cultural Heritage* (pp. 47–69). Springer International Publishing. https://doi.org/10.1007/978-3-030-66777-1_3

Britannica. (2024, January 14). *James Crerar Reaney | Canadian Poet, Playwright, Critic | Britannica*. https://www.britannica.com/place/Stratford-Ontario

Caldwell, W., Epp, S., Wan, X., Singer, R., Drake, E., & Sousa, E. C. (2022). Farmland Preservation and Urban Expansion: Case Study of Southern Ontario, Canada. *Frontiers in Sustainable Food Systems*, *6*. https://doi.org/10.3389/fsufs.2022.777816

Chen, L., Li, S., Bai, Q., Yang, J., Jiang, S., & Miao, Y. (2021). Review of Image Classification Algorithms Based on Convolutional Neural Networks. *Remote Sensing*, *13*(22), Article 22. https://doi.org/10.3390/rs13224712

City of Guelph. (n.d.). *Municipal Register Review Process*. Retrieved 30 April 2024, from https://guelph.ca/wp-content/uploads/MunicipalRegisterReviewProcess.pdf

City of Markham. (1991, October). *Evaluating Heritage Resources in the Town of Markham*. https://www.markham.ca/wps/wcm/connect/markham/f352bd16-30c0-4dc5-a547-ee6ad38db23e/HERITAGE+BUILDING+EVALUATION+Markham.pdf?MOD=AJPERES&CONVERT_TO=url&CACHEID=ROOTWORKSPACE.Z18_2QD4H901OGV160QC8BLCRJ1001-f352bd16-30c0-4dc5-a547-ee6ad38db23e-muuQFq2

City of Ottowa. (n.d.). *Application Form—Heritage designation under Part IV of the Ontario Heritage Act*.

City of Toronto. (2022, February 9). *2021 Census: Population and Dwelling Counts*. https://www.toronto.ca/wp-content/uploads/2022/02/92e3-City-Planning-2021-Census-Backgrounder-Population-Dwellings-Backgrounder.pdf

City of Toronto. (2023). *Implementing Bill 23 – Update on Amendments to the Ontario Heritage Act*. https://www.toronto.ca/legdocs/mmis/2023/ph/bgrd/backgroundfile-240818.pdf

Convention Concerning the Protection of the World Cultural and Natural Heritage. (1972). UNESCO.

Cosovic, M., & Jankovic, R. (2020). CNN Classification of the Cultural Heritage Images. *2020 19th International Symposium INFOTEH-JAHORINA (INFOTEH)*, 1–6. https://doi.org/10.1109/INFOTEH48170.2020.9066300

Costa, M., & Carneiro, M. J. (2021). The influence of interpretation on learning about architectural heritage and on the perception of cultural significance. *Journal of Tourism and Cultural Change*, *19*(2), 230–249. https://doi.org/10.1080/14766825.2020.1737705

Dai, Y., Gao, Y., & Liu, F. (2021). TransMed: Transformers Advance Multi-Modal Medical Image Classification. *Diagnostics*, *11*(8), Article 8. https://doi.org/10.3390/diagnostics11081384

Del Frate, F., Pacifici, F., Schiavon, G., & Solimini, C. (2007). Use of Neural Networks for Automatic Classification From High-Resolution Images. *IEEE Transactions on Geoscience and Remote Sensing*, *45*(4), 800–809. IEEE Transactions on Geoscience and Remote Sensing. https://doi.org/10.1109/TGRS.2007.892009

Deng, X., Shao, H., Shi, L., Wang, X., & Xie, T. (2020). A Classification–Detection Approach of COVID-19 Based on Chest X-ray and CT by Using Keras Pre-Trained Deep Learning Models. *Computer Modeling in Engineering & Sciences*, *125*(2), 579–596. https://doi.org/10.32604/cmes.2020.011920

European Charter of the Architectural Heritage. (1975).

Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, *542*(7639), 115–118. https://doi.org/10.1038/nature21056

Global Heritage Fund. (2010, October 19). *Safeguarding Endangered Cultural Heritage Sites in the Developing World*.

Goldsborough, P. (2016). *A Tour of TensorFlow* (arXiv:1610.01178). arXiv. https://doi.org/10.48550/arXiv.1610.01178

Google. (2024). *Map of Stratford, Ontario* [Map]. Retrieved from https://www.google.ca/maps/place/Stratford,+ON/@43.3669008,-81.004242,14z/data=!3m1!4b1!4m6!3m5!1s0x882eadf9b80320f1:0xc4f78a388b586b05!8m2!3d43.3705082!4d-80.9821162!16s%2Fg%2F11kj0gwbp4?hl=en&entry=ttu

Government of Canada. (n.d.). *Stratford City Hall National Historic Site of Canada*. Retrieved 13 March 2024, from https://www.pc.gc.ca/apps/dfhd/page_nhs_eng.aspx?id=522

Government of Ontario. (n.d.-a). *Eight guiding principles in the conservation of built heritage properties | ontario.ca*. Retrieved 30 April 2024, from http://www.ontario.ca/page/eight-guiding-principles-conservation-built-heritage-properties

Government of Ontario. (n.d.-b). *Heritage properties and insurance | ontario.ca*. Retrieved 29 April 2024, from http://www.ontario.ca/page/heritage-properties-and-insurance

Government of Ontario. (2016, Fall). *The Municipal Register of Heritage Properties*. https://www.tiny.ca/sites/default/files/2021-12/Provincial%20Information%20Sheet.pdf

Gulli, A., Kapoor, A., & Pal, S. (2019). *Deep Learning with TensorFlow 2 and Keras: Regression, ConvNets, GANs, RNNs, NLP, and more with TensorFlow 2 and the Keras API, 2nd Edition*. Packt Publishing Ltd.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770–778. https://doi.org/10.1109/CVPR.2016.90

Heritage Brockville. (n.d.). *WHAT IS HERITAGE DESIGNATION? A Primer for Property Owners*. Retrieved 11 March 2024, from https://www.heritagebrockville.ca/index.cfm?ID=109&Download=&1["57B=&*=&f=7

Hinton, G. E., Osindero, S., & Teh, Y.-W. (2006). A Fast Learning Algorithm for Deep Belief Nets. *Neural Computation*, *18*(7), 1527–1554. https://doi.org/10.1162/neco.2006.18.7.1527

HPI Nomination Team & University of Waterloo. (2009, January). *Ontario Architectural Style Guide*. https://www.therealtydeal.com/wp-content/uploads/2018/06/Heritage-Resource-Centre-Achitectural-Styles-Guide.pdf

Huang, G., Liu, Z., van der Maaten, L., & Weinberger, K. Q. (2018). *Densely Connected Convolutional Networks* (arXiv:1608.06993). arXiv. https://doi.org/10.48550/arXiv.1608.06993

IBM. (2021, March 3). *IBM Documentation*. https://ibm.com/docs/en/control-desk/7.6.1.2?topic=structure-views

International Charter for the Conservation and Restoration of Monuments and Sites (The Venice Charter). (1964).

ICOMOS. (1982). Florence Charter.

ICOMOS. (2011). Paris Charter.

Jankovic Babic, R. (2023). A comparison of methods for image classification of cultural heritage using transfer learning for feature extraction. *Neural Computing and Applications*, 1–11. https://doi.org/10.1007/s00521-023-08764-x

Janowicz, K., Gao, S., McKenzie, G., Hu, Y., & Bhaduri, B. (2020). GeoAI: Spatially explicit artificial intelligence techniques for geographic knowledge discovery and beyond. *International Journal of Geographical Information Science*, *34*(4), 625-636. https://doi.org/10.1080/13658816.2019.1684500

Joshi, S., Owens, J. A., Shah, S., & Munasinghe, T. (2021). Analysis of Preprocessing Techniques, Keras Tuner, and Transfer Learning on Cloud Street image data. *2021 IEEE International Conference on Big Data (Big Data)*, 4165–4168. https://doi.org/10.1109/BigData52589.2021.9671878

Kamel Boulos, M. N., Peng, G., & VoPham, T. (2019). An overview of GeoAI applications in health and healthcare. *International Journal of Health Geographics*, *18*(1), 7. https://doi.org/10.1186/s12942-019-0171-2

Kang, J., Körner, M., Wang, Y., Taubenböck, H., & Zhu, X. X. (2018). Building instance classification using street view images. *ISPRS Journal of Photogrammetry and Remote Sensing*, *145*, 44–59. https://doi.org/10.1016/j.isprsjprs.2018.02.006

Keras Team. (n.d.). *Keras documentation: Transfer learning & fine-tuning*. Retrieved 1 May 2024, from https://keras.io/guides/transfer_learning/

Kieffer, B., Babaie, M., Kalra, S., & Tizhoosh, H. R. (2017). Convolutional neural networks for histopathology image classification: Training vs. Using pre-trained networks. *2017 Seventh International Conference on Image Processing Theory, Tools and Applications (IPTA)*, 1–6. https://doi.org/10.1109/IPTA.2017.8310149

Kingma, D. P., & Ba, J. (2017). *Adam: A Method for Stochastic Optimization* (arXiv:1412.6980). arXiv. http://arxiv.org/abs/1412.6980

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. *Advances in Neural Information Processing Systems*, *25*. https://proceedings.neurips.cc/paper_files/paper/2012/hash/c399862d3b9d6b76c8436e924a68c45b-Abstract.html

Lapadula, M. I., & Quiroga, C. (2012). Heritage as a pedagogical resource and platform for exploration in architectural design education. *The Journal of Architecture*, *17*(4), 591–607. https://doi.org/10.1080/13602365.2012.709028

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444. https://doi.org/10.1038/nature14539

Lecun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, *86*(11), 2278–2324. Proceedings of the IEEE. https://doi.org/10.1109/5.726791

Li, W., & Hsu, C.-Y. (2022). GeoAI for Large-Scale Image Analysis and Machine Vision: Recent Progress of Artificial Intelligence in Geography. *ISPRS International Journal of Geo-Information*, *11*(7), Article 7. https://doi.org/10.3390/ijgi11070385

Li, Z., He, W., Li, J., Lu, F., & Zhang, H. (2024). *Learning without Exact Guidance: Updating Large-scale High-resolution Land Cover Maps from Low-resolution Historical Labels* (arXiv:2403.02746; Version 2). arXiv. https://doi.org/10.48550/arXiv.2403.02746

Lin, M., Chen, Q., & Yan, S. (2014). *Network In Network* (arXiv:1312.4400). arXiv. https://doi.org/10.48550/arXiv.1312.4400

Liu, P., & Biljecki, F. (2022). A review of spatially-explicit GeoAI applications in Urban Geography. *International Journal of Applied Earth Observation and Geoinformation*, *112*, 102936. https://doi.org/10.1016/j.jag.2022.102936

Marian Dumitru Danci. (2020). *Architectural styles*. https://www.kaggle.com/datasets/dumitrux/architectural-styles-dataset

Medus, L. D., Saban, M., Francés-Víllora, J. V., Bataller-Mompeán, M., & Rosado-Muñoz, A. (2021). Hyperspectral image classification using CNN: Application to industrial food packaging. *Food Control*, *125*, 107962. https://doi.org/10.1016/j.foodcont.2021.107962

Mehta, S., Kukreja, V., & Yadav, R. (2023). Multi-Classification of Heritage Buildings using Federated Learning CNN: A Comparative Analysis of Client-Side and Global Model Performance. *2023 Second International Conference on Augmented Intelligence and Sustainable Systems (ICAISS)*, 179–183. https://doi.org/10.1109/ICAISS58487.2023.10250461

National Trust for Canada. (n.d.). The Ontario Heritage Act, Future Designations, and Over 31,500 ‘Listed’ Heritage Properties in Ontario. *National Trust for Canada*. Retrieved 26 May 2024, from https://nationaltrustcanada.ca/nt-endangered-places/the-ontario-heritage-act-future-designations-and-over-31500-listed-heritage-properties-in-ontario

Ontario Heritage Act, R.S.O. 1990, c. O.18

Ontario Heritage Trust. (n.d.-a). *Benefits of heritage designation under the Ontario Heritage Act*. Ontario Heritage Trust. Retrieved 29 April 2024, from https://www.heritagetrust.on.ca/pages/tools/tools-for-conservation/benefits-of-heritage-designation-under-the-ontario-heritage-act

Ontario Heritage Trust. (n.d.-b). *Introduction*. Ontario Heritage Trust. Retrieved 13 March 2024, from https://www.heritagetrust.on.ca/pages/tools/ontario-heritage-act-register/introduction

Pan, S. J., & Yang, Q. (2010). A Survey on Transfer Learning. *IEEE Transactions on Knowledge and Data Engineering*, *22*(10), 1345–1359. IEEE Transactions on Knowledge and Data Engineering. https://doi.org/10.1109/TKDE.2009.191

Perth County. (2017, February 25). *About Perth County*. Perth County Central News. https://web.archive.org/web/20170225155418/http://www.perthcounty.ca/about_perth_county

Polowin, J., & Polowin, M. S. (n.d.). *Changes coming to heritage registers across Ontario*. Gowling WLG. Retrieved 13 March 2024, from https://gowlingwlg.com/en/insights-resources/articles/2023/changes-coming-to-heritage-registers-in-ontario/

Prasomphan, S. (2022). Ensemble Classification Technique for Cultural Heritage Image. In X. Jiang (Ed.), *Machine Learning and Intelligent Communications* (pp. 17–27). Springer International Publishing. https://doi.org/10.1007/978-3-031-04409-0_3

Prechelt, L. (1998). Early Stopping—But When? In G. B. Orr & K.-R. Müller (Eds.), *Neural Networks: Tricks of the Trade* (pp. 55–69). Springer. https://doi.org/10.1007/3-540-49430-8_3

ICOMOS China. (2015). *Zhongguo wen wu gu ji bao hu zhun ze = Principles for the conservation of heritage sites in China* (Rev. ed.). Wen wu chu ban she. http://hdl.handle.net/10020/gci_pubs/china_principles_2015

Rai, A., Constantinides, P., & Sarker, S. (2019). Next generation digital platforms: Toward human-AI hybrids. *MIS Quarterly*, *43*(1), Article 1.

Rawat, W., & Wang, Z. (2017). Deep Convolutional Neural Networks for Image Classification: A Comprehensive Review. *Neural Computation*, *29*(9), 2352–2449. Neural Computation. https://doi.org/10.1162/neco_a_00990

Rosebrock, A. (2019, February 4). Keras: Multiple Inputs and Mixed Data. *PyImageSearch*. https://pyimagesearch.com/2019/02/04/keras-multiple-inputs-and-mixed-data/

Sharifzadeh, F., Akbarizadeh, G., & Seifi Kavian, Y. (2019). Ship Classification in SAR Images Using a New Hybrid CNN–MLP Classifier. *Journal of the Indian Society of Remote Sensing*, *47*(4), 551–562. https://doi.org/10.1007/s12524-018-0891-y

Sigmund, Z. (2016). Sustainability in architectural heritage: Review of policies and practices. *Organization, Technology & Management in Construction : An International Journal*, *8*(1), 1411–1421. https://doi.org/10.1515/otmcj-2016-0007

Simonyan, K., & Zisserman, A. (2015). *Very Deep Convolutional Networks for Large-Scale Image Recognition* (arXiv:1409.1556). arXiv. https://doi.org/10.48550/arXiv.1409.1556

Simpson, M. (2020, October 31). *Grand Trunk Railway and Canadian National Railway Shops—Stratford, ON* - *Simpson, Morgan—Local Landscape Report*.

https://www.guidetags.com/mindmaps/explore//4460-simpson-morgan-local-landscape-report-grand-trunk-railway-and-canadian-national-railway-shops-stratford-on

Singhal, A., & Sharma, D. K. (2023). Chapter 3—Voice signal-based disease diagnosis using IoT and learning algorithms for healthcare. In C. Chakraborty, S. K. Pani, M. Abdul Ahad, & Q. Xin (Eds.), *Implementation of Smart Healthcare Systems using AI, IoT, and Blockchain* (pp. 59–81). Academic Press. https://doi.org/10.1016/B978-0-323-91916-6.00005-9

Song, J., Gao, S., Zhu, Y., & Ma, C. (2019). A survey of remote sensing image classification based on CNNs. *Big Earth Data*, *3*(3), 232–254. https://doi.org/10.1080/20964471.2019.1657720

Szegedy, C., Wei Liu, Yangqing Jia, Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2015). Going deeper with convolutions. *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 1–9. https://doi.org/10.1109/CVPR.2015.7298594

Tan, M., & Le, Q. V. (2020). *EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks* (arXiv:1905.11946). arXiv. http://arxiv.org/abs/1905.11946

The Corporation of the City of Stratford. (2019a). *Heritage Conservation District*. https://www.stratford.ca/en/live-here/heritage-conservation-district.aspx

The Corporation of the City of Stratford. (2019b). *Heritage Stratford*. https://www.stratford.ca/en/live-here/heritage-stratford.aspx

The Japanese Association for Conservation of Architectural Monuments. (n.d.). Retrieved 27 April 2024, from https://www.bunkenkyo.or.jp/en/conserva.html

The Town of New Tecumseth. (2019). *Heritage Designation*. https://www.newtecumseth.ca/en/parks-recreation-and-culture/heritage-designation.aspx

Throsby, D. (2001). *Economics and Culture*. Cambridge University Press.

Trier, Ø. D., Reksten, J. H., & Løseth, K. (2021). Automated mapping of cultural heritage in Norway from airborne lidar data using faster R-CNN. *International Journal of Applied Earth Observation and Geoinformation*, *95*, 102241. https://doi.org/10.1016/j.jag.2020.102241

Turing, A. M. (1980). Computing Machinery and Intelligence. *Creative Computing*, *6*(1), 44–53.

UNESCO. (1972). Convention Concerning the Protection of the World Cultural and Natural Heritage.

Vu, M.-T., Beurton-Aimar, M., & Le, V.-L. (2018). Heritage Image Classification by Convolution Neural Networks. *2018 1st International Conference on Multimedia Analysis and Pattern Recognition (MAPR)*, 1–6. https://doi.org/10.1109/MAPR.2018.8337517

Wang, B., Zhang, S., Zhang, J., & Cai, Z. (2023). Architectural style classification based on CNN and channel–spatial attention. *Signal, Image and Video Processing*, *17*(1), 99–107. https://doi.org/10.1007/s11760-022-02208-0

Wang, Y., Matsumoto, K., & Masanori, S. (2019). *Examination of Unrecognized Historic Area from the Perspective of Overall Conservation for Historic Urban Area Taking 11 Old Concession Settlement Cities in China as an Example*.

Wicklow County Council. (2010, October 4). *Wicklow County Development Plan 2010-2016 | Wicklow.ie*. https://www.wicklow.ie/Living/Services/Planning/Development-Plans-Strategies/National-Regional-County-Plans/Wicklow-County-Development-Plan/Wicklow-County-Development-Plan-2010-2016

Xu, Z., Tao, D., Zhang, Y., Wu, J., & Tsoi, A. C. (2014). Architectural Style Classification Using Multinomial Latent Logistic Regression. In D. Fleet, T. Pajdla, B. Schiele, & T. Tuytelaars (Eds.), *Computer Vision – ECCV 2014* (Vol. 8689, pp. 600–615). Springer International Publishing. https://doi.org/10.1007/978-3-319-10590-1_39

Yazdi, H., Sad Berenji, S., Ludwig, F., & Moazen, S. (2022). Deep Learning in Historical Architecture Remote Sensing: Automated Historical Courtyard House Recognition in Yazd, Iran. *Heritage*, *5*(4), 3066–3080. https://doi.org/10.3390/heritage5040159

Yosinski, J., Clune, J., Bengio, Y., & Lipson, H. (2014). How transferable are features in deep neural networks? *Advances in Neural Information Processing Systems*, *27*. https://proceedings.neurips.cc/paper_files/paper/2014/hash/375c71349b295fbe2dcdca9206f20a06-Abstract.html

Yousaf, K., & Nawaz, T. (2022). A Deep Learning-Based Approach for Inappropriate Content Detection and Classification of YouTube Videos. *IEEE Access*, *10*, 16283–16298. IEEE Access. https://doi.org/10.1109/ACCESS.2022.3147519

Zhang, C., Pan, X., Li, H., Gardiner, A., Sargent, I., Hare, J., & Atkinson, P. M. (2018). A hybrid MLP-CNN classifier for very fine resolution remotely sensed image classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, *140*, 133–144. https://doi.org/10.1016/j.isprsjprs.2017.07.014

Zhu, X. X., Tuia, D., Mou, L., Xia, G.-S., Zhang, L., Xu, F., & Fraundorfer, F. (2017). Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources. *IEEE Geoscience and Remote Sensing Magazine*, *5*(4), 8–36. IEEE Geoscience and Remote Sensing Magazine. https://doi.org/10.1109/MGRS.2017.2762307