

L'AUTOMAZIONE DEL CENSIMENTO TECNOLOGICO URBANO

Il centro storico di Betlemme

AUTOMATION OF URBAN TECHNOLOGICAL CENSUS

The historical centre of Bethlehem

Elisabetta Doria

ABSTRACT

Il contributo riporta gli esiti di una ricerca sull'automazione del censimento di elementi tecnologici in ambito urbano a supporto della progettazione di un sistema di gestione e monitoraggio del Patrimonio architettonico. Lo studio si interessa di alcuni particolari elementi (i serbatoi idrici) presenti sulle coperture del centro storico di Betlemme e utilizza acquisizioni fotogrammetriche close-range come base di addestramento di modelli di Deep Learning. Il ciclo di vita dei modelli di Deep Learning, dalla fase di addestramento fino alla restituzione degli output e l'archiviazione di immagini e metadati è effettuato sfruttando la scalabilità di un'infrastruttura cloud. Il monitoraggio con ispezioni periodiche permette di confrontare condizioni differenti e valutare situazioni di criticità rilevando variazioni quali sostituzione, aggiunta e rimozione degli elementi ricercati. Il progetto ha l'obiettivo di definire un protocollo per automatizzare l'identificazione di elementi ricorrenti e monitorarli nel tempo.

The research proposal reports the outcomes of a research track concerning the automation of the architectural census of technological elements in urban environments, aiming at the development of a monitoring and management system for the built heritage. The proposal is focused on a set of specific elements (water tanks) stacked on the coverings of the historical centre of Bethlehem and leverages close-range photogrammetric acquisitions to train Deep Learning models. The model lifecycle management, from training to prediction and deployment, as well as the storage of both image data and metadata, is performed through the scalability of a Cloud enterprise architecture. Periodical scheduled monitoring enables comparisons across different periods, allowing the detection of modifications, removals, and additions, therefore identifying the insurgence of potential criticalities. The goal of the project is the definition of a protocol to automate the identification of recurrent elements and monitor their evolution through time.

KEYWORDS

fotogrammetria, Betlemme, monitoraggio urbano, riconoscimento di oggetti, censimento architettonico

photogrammetry, Bethlehem, urban monitoring, object recognition, architectural census



Elisabetta Doria is an Architect/Engineer and PhD Candidate at the Department of Civil Engineering and Architecture, University of Pavia (Italy). Her research activity concerns operational protocols for the preservation of buildings in historic, monumental and UNESCO sites for the enhancement of architectural and landscape heritage through the management of digital databases. E-mail: elisabetta.doria@unipv.it

Il contributo illustra gli esiti di una ricerca per l'automazione del censimento di elementi tecnologici e architettonici in ambito urbano tramite tecniche di Object Detection finalizzate a mettere a punto un sistema per la gestione del Patrimonio architettonico. L'acquisizione di informazioni tecnologiche, costruttive e materiche sul Patrimonio architettonico costituisce un'importante fase nel processo di conservazione, poiché la 'documentazione integrale' del Patrimonio stesso consente di guidare con maggiore consapevolezza le attività di gestione e di valorizzazione (Gasparoli, 2012; Montella, 2009). In particolare, gli strumenti UAV (Unmanned Aerial Vehicle), insieme alle tecnologie di post-produzione per la generazione di gemelli digitali, consentono di sviluppare metodi di monitoraggio innovativi in cui la presenza degli operatori sul campo può essere notevolmente ridotta in favore di una maggiore sicurezza sul lavoro e ottimizzazione dei costi di gestione e manutenzione.

L'ambito di indagine è circoscritto alle coperture del centro storico della città di Betlemme (Fig. 1) e, in particolare, all'analisi delle cisterne d'acqua a uso civile presenti sui tetti. La mappatura di queste, a partire dalle acquisizioni fotogrammetriche close-range da UAV, diviene di particolare importanza in quanto esse costituiscono elementi di carico su edifici storici in territorio sismico (Fig. 2). L'obiettivo è quello di ottenere un prodotto digitale informativo facilmente fruibile dai tecnici della Municipalità di Betlemme per monitorare gli sviluppi dell'edificato della città e connettere successivamente questi dati a ulteriori informazioni acquisite sul campo relative allo stato di conservazione degli edifici. La ricerca proposta si inserisce nel contesto del progetto di cooperazione e ricerca '3D Bethlehem – Management and control of urban growth for the development of Heritage and Improvement of life in the city of Bethlehem' cofinanziato da AICS e coordinato dal Dipartimento di Ingegneria Civile e Architettura dell'Università di Pavia (Parrinello, 2019).¹

Il Progetto, avviato nel 2018 e concluso nel gennaio 2020, ha visto lo sviluppo di un rilievo digitale, della schedatura e della modellazione digitale degli edifici del centro storico per la realizzazione di uno strumento tridimensionale GIS dal quale monitorare gli interventi sull'edificato. Le numerose attività di rilevamento fotogrammetrico close-range hanno prodotto un dataset di oltre 9.000 immagini delle coperture, georeferenziate, acquisite con strumentazione UAV quale DJI Phantom 4 PRO, DJI Mavic, DJI Mavic Mini (Parrinello and Picchio, 2019; Picchio, 2019). Tale dataset è stato utilizzato come base per la sperimentazione qui proposta, autofinanziata e sviluppata a conclusione del progetto 3D Bethlehem in accordo con il responsabile scientifico Prof. Sandro Parrinello (Fig. 3).

La sperimentazione ha visto l'utilizzo dei dati fotogrammetrici close-range delle coperture come base di addestramento di modelli di Deep Learning. Questi modelli necessitano di dataset massivi per convergere, ovvero per raggiungere uno stato in cui il risultato si assesta entro un intervallo di errore paragonabile al valore finale. La scelta del dataset è stata effettuata a seguito della compilazione della schedatura per la strutturazione del sistema informativo GIS, da cui è emersa la complessità urbana e le criticità legate alle coperture (Doria, 2022). L'obiettivo è stato quello di attivare un pro-

cesso di automazione del censimento degli elementi tecnologici di copertura, volto a soddisfare la necessità di monitoraggio e di azioni di intervento programmatico sul Patrimonio. Gli elementi tecnologici scelti per questa prima sperimentazione sono le cisterne d'acqua a uso civile presenti in maniera estensiva e non regolamentata sulle coperture del centro storico. Le cisterne d'acqua sono un sistema di immagazzinamento particolarmente prezioso nei periodi di razionamento dell'acqua ma, allo stesso tempo, sono una importante criticità a livello urbano: esse rappresentano elementi di carico concentrato sul 69,3% delle coperture degli edifici censiti e sono posizionate senza le opportune verifiche strutturali compromettendo l'integrità degli edifici in un territorio sismico (Abueladas et alii, 2021; De Marco, 2022).

Il caso di studio | Le cisterne diffuse nel centro storico di Betlemme e nel territorio circostante si suddividono in differenti categorie per dimensione e materiale. Per la ricerca, le cisterne sono state catalogate e raggruppate in tre macrocategorie (Fig. 4): cisterne cilindriche in materiale plastico di fattura recente e con un buono stato di conservazione (54%), cisterne metalliche rettangolari e cisterne metalliche cilindriche di fattura meno recente e in peggiore stato di conservazione (46%)². Durante le fasi di approvvigionamento d'acqua, che avviene in maniera non costante e solo in specifici giorni, le cisterne in cattivo stato di conservazione poste sui tetti e che presentano problemi di controllo del flusso hanno perdite d'acqua che fluiscono sulle facciate degli edifici e sugli impianti presenti lungo il percorso dell'acqua. Il flusso di acqua causa patologie negli edifici quali crescita di vegetazione – sia intensiva che estensiva, in base alla posizione e allo stato d'uso dell'architettura – e aree umide con evidenti alterazioni del materiale della facciata.

Il territorio della città di Betlemme ha una morfologia collinare e lo stesso centro storico è caratterizzato sia da strade carrabili ripide che da numerose scalinate pedonali che hanno funzione di collegamento tra i quartieri posti su livelli diversi della collina. Questa morfologia fa sì che le cisterne posizionate sulle coperture siano visibili in tutta la città, tanto quelle poste su edifici di recente costruzione quanto quelle su edifici storici. Il posizionamento delle taniche in queste due classificazioni di edifici – recenti e storici – è variabile; negli edifici di recente costruzione le cisterne sono collocate in copertura secondo una maglia regolare; negli edifici storici, che hanno invece coperture voltate con estradosso visibile e una morfologia complessa tipica della tipologia abitativa locale³, sono disposte in maniera disordinata. In molti di questi edifici storici la copertura viene inoltre utilizzata come luogo di stoccaggio di detriti e scarti, rendendo complessa la visibilità delle cisterne e la loro manutenzione.

Stato dell'arte | Il processo analitico di comprensione del contesto urbano è oggi connesso alla costruzione di banche dati affidabili volte a sviluppare da un lato strategie di gestione del Patrimonio Storico, dall'altro la comunicazione e fruizione virtuale degli spazi digitalizzati. I centri storici conservano le informazioni relative ai numerosi eventi che hanno interessato la città nel corso degli anni definendo un insieme di dati eterogeneo e varia-

bile che può e deve essere sistemizzato per una corretta gestione (Bocconcino and Manzone, 2019). L'aumento delle informazioni sul processo decisionale pubblico ha fatto emergere la necessità di adottare nuove tecnologie e tecniche di gestione delle informazioni. Le Tecnologie dell'Informazione e della Comunicazione (ICT), con il loro rapido sviluppo, influenzano il quadro decisionale se applicate alla gestione pubblica con l'obiettivo di realizzare strumenti di monitoraggio per strategie gestionali all'avanguardia (Kouziokas and Perakis, 2017).

Lo sviluppo tecnologico dei metodi di acquisizione di dati ha permesso di semplificare e accelerare la raccolta di grandi quantità di informazioni e metadati, che possono essere collegati tra loro per costruire e arricchire dataset massivi. Se i processi inferenziali per la conoscenza del Patrimonio partono dalle informazioni acquisite, è anche vero che la raccolta di dati non genera conoscenza senza un processo capace di individuare ed esplicitare le correlazioni tra le variabili. I computer possono utilizzare la crescente quantità di informazioni non ordinate definendo un nuovo modo di produrre conoscenza basata sull'identificazione, a priori o a posteriori, di pattern ricorrenti e correlazioni all'interno dei dati. In questo contesto diventa cruciale lo sviluppo di ricerche legate all'Intelligenza Artificiale (IA), termine coniato dal Professor John McCarthy e definito come «[...] the science and engineering of making intelligent machines» (McCarthy, 2007, p. 2) che oggi comprende un insieme differenziato di tecniche e metodi con l'obiettivo di automatizzare le capacità cognitive (Corea, 2019).

Fin dai suoi esordi la ricerca sull'IA si è concentrata sulla creazione di modelli che simulano le dinamiche dell'intelligenza umana, basati sul confronto e la rielaborazione di stimoli esterni, che possono essere classificati in base alle principali direzioni di ricerca attive (Buratti, Conte and Rossi, 2021). In particolare l'apprendimento automatico è un termine che racchiude molteplici tecniche e approcci; questi permettono di migliorare le prestazioni di un determinato algoritmo nell'identificare pattern nei dati di input, in modo non dissimile da come il cervello umano apprende e perfeziona una nuova capacità.

Le tecniche di apprendimento basate su reti neurali, modelli di Deep Learning (DL) che presentano connessioni casuali tra elementi computazionali le cui relazioni vengono perfezionate tramite apprendimento, possono gestire casi d'uso di complessità crescente all'aumentare della complessità computazionale e al numero di interconnessioni delle reti stesse (Mishra, 2021). La capacità di questi approcci di scalare, ossia di lavorare quantità crescenti di dati, li rende particolarmente adatti a lavorare con dataset massivi, sia di dati strutturati (tabelle, numeri) sia non strutturati (linguaggio naturale, immagini). La branca delle ICT che ha come obiettivo l'estrazione di informazioni dalle immagini è la Computer Vision che interseca la IA in tutte quelle applicazioni in cui la capacità di apprendimento e di gestione di dati non strutturati permette di risolvere problemi complessi, come la segmentazione o l'identificazione di oggetti in immagini.

Il riconoscimento automatico degli elementi tecnologici, sviluppato nella presente ricerca, si basa sull'Object Detection, una branca della Com-



Fig. 1 | UAV photographic images of Bethlehem's Old Town showing the morphological complexity of the urban settlement (credit: '3D Bethlehem' Project, E. Doria, 2022).

puter Vision che permette di identificare oggetti nelle immagini. Il compito dell'algoritmo implementato è duplice: isolare l'oggetto di interesse in una scena anche complessa e poi identificare l'oggetto isolato con la classe corretta. A partire dai primi studi e utilizzi delle Convolutional Neural Networks – CNN (Lecun and Bengio, 1998; Cosovic and Jankovic, 2020), i modelli di DL sono diventati lo standard per le alte prestazioni raggiunte (Redmon et alii, 2016). In particolare, è stato utilizzato AutoML, una famiglia di modelli e approcci di feature engineering che sfruttano tecniche di ensemble e transfer learning per consentire una costruzione più rapida di modelli di Machine Learning (ML) e che spingono verso una maggior democratizzazione dell'Intelligenza Artificiale (He, Zhao and Chu, 2021).

La gestione del dataset fotografico, potenzialmente di dimensione crescente, e la complessità computazionale intrinseca del lavoro con immagini pongono un vincolo non trascurabile agli approcci disponibili. Per una gestione on-premises con un dataset crescente, le risorse di archivio e la potenza di calcolo disponibili devono essere ampliate continuamente acquistando nuovo hardware. La ricerca vuole evitare questa criticità utilizzando non un'infrastruttura fisica ma una di Cloud Computing come Integrated Machine Learning Platform, che permette l'astrazione dalla potenza di calcolo disponibile in loco, eliminando i limiti imposti dall'hardware fisico (Rivera, 2020).

Ricerche correlate | Le tecniche di Object Detection precedentemente illustrate sono state sviluppate per apprendere il riconoscimento di un'entità, sulla base di dati inseriti, e per restituire previsioni. A partire dai dati immagini, tre sono gli obiettivi principali che vengono perseguiti: la classificazione, il riconoscimento e la segmentazione (Xiao et alii, 2020). Gli approcci allo sviluppo di questi obiettivi sono applicati in letteratura in differenti ambiti e sono ormai riconosciuti come strumento di supporto per la gestione e la documentazione dei Beni Culturali e per la pianificazione urbana (Li et alii, 2020).

In ambito urbano è stata verificata l'utilità dell'utilizzo di questa tecnologia per il monitoraggio di situazioni di criticità quali infrastrutture e cantieri (Dandabathula et alii, 2019; Wang et alii,

2022), gestione dei rifiuti urbani (Majchrowska et alii, 2022), per il controllo dell'adempimento dei regolamenti edilizi (Cunha et alii, 2021) e il monitoraggio ambientale e agricolo (Carbone et alii, 2022). Sono riscontrabili, inoltre, casi di applicazione con risultati ad alta prestazione nel campo della Patrimonio Culturale costruito, sia per l'individuazione di Beni non ancora identificati tramite immagini satellitari e close-range (Due Trier, Reksten and Løseth, 2021) sia per interventi su Patrimonio già documentato (Monna et alii, 2021; Zou et alii, 2019; Hatir and Ince, 2021) come supporto alla scelta delle pratiche di intervento.

Il presente contributo pone l'attenzione sull'utilizzo delle tecniche di Object Detection per la pianificazione e gestione del territorio urbano utilizzando piattaforme cloud e con tecnologie open source, intersecando l'ambito di controllo dell'esistente in relazione ai temi di sicurezza strutturale e impiantistica. I modelli di machine learning utilizzati e i flussi per la loro gestione, basandosi su tecnologia open source quali librerie Python e Kubeflow, sono esportabili e condivisibili con collaboratori esterni alla piattaforma. I collaboratori esterni, sia tecnici che fruitori come Università, aziende e municipalità, possono quindi avere a disposizione i modelli esportati utilizzandoli su infrastrutture cloud differenti da quella di sviluppo, facilitando la collaborazione tra realtà differenti sia nella fase di implementazione sia nella fase di utilizzo finale, oltre che poter fruire di modelli pre-addestrati esposti via web tramite l'Interfaccia di Programmazione dell'Applicazione (API).

Metodologia e fasi della ricerca | Per strutturare un processo di censimento automatico ripetibile è stato necessario disporre di immagini georeferenziate. Per fare questo, oltre che preferire l'utilizzo di piani di volo da UAV (DJI Phantom 4 PRO), si è scelto di utilizzare come base per il riconoscimento delle immagini il fotopiano delle coperture correttamente scalato e georeferenziato.

La prima fase della ricerca, sviluppata a partire dai dati acquisiti nella campagna di rilievo del progetto 3D Bethlehem, ha riguardato l'analisi dei dati prodotti dall'acquisizione fotogrammetrica close-range delle coperture e la post-produzione dei dati per ottenere il fotopiano dell'area di interesse. Il modello Structure from Motion (SfM) ottenuto,

da cui è stata esportata l'ortoimmagine delle coperture (Picchio, 2020), è stato realizzato con piani di volo pianificati dall'operatore⁴ con acquisizioni zenitali e inclinate sui 4 lati di ogni area per un totale di 5 piani di volo per le aree di minore dimensione (Figg. 5-7). I modelli fotogrammetrici ottenuti sono stati scalati metricamente utilizzando come sistema di riferimento la nuvola di punti realizzata con Terrestrial Laser Scanner (TLS) – Faro CAM2 S150. Il fotopiano georeferenziato è stato utilizzato come base per il riconoscimento automatico delle cisterne; da questo è possibile riprodurre a intervalli temporali scadenzati la stessa immagine, con caratteristiche di ripresa ripetibili laddove diviene necessaria una nuova campagna di acquisizione.

La seconda fase della ricerca è relativa alla scelta dell'infrastruttura su cui svolgere i test. Per procedere con l'addestramento dei modelli è stata utilizzata un'infrastruttura cloud, per garantire la scalabilità del prodotto: l'infrastruttura cloud, costruita sulla piattaforma enterprise Google Cloud Platform, permette una gestione end-to-end del dato. Questo processo parte dalla ingestione delle immagini, input, passando alla costruzione di dataset, alla popolazione di database, all'addestramento e deployment dei modelli per giungere alla fruizione finale del dato, output (Fig. 8). L'infrastruttura che supporta questo progetto è completamente serverless, ovvero al presentarsi della necessità è possibile aggiungere delle macchine virtuali per supplire alla eventuale mancanza di storage o potenza di calcolo, adattando e costruendo l'architettura su misura in base al mutare delle necessità del caso di studio.

Nel caso della Object Detection, facente largo uso di reti neurali convolutive in cui una sliding window viene applicata a una porzione di immagine, la complessità cresce all'aumentare del numero e della risoluzione delle immagini di input. Il dato processato da questi layer convolutivi serve ad alimentare gli stadi successivi delle reti fino ad arrivare al classificatore finale, che attribuirà l'appartenenza o meno di una immagine ad una delle classi identificate.

Nella terza fase della ricerca è avvenuta la scelta e l'addestramento dei modelli. I modelli di Deep Learning sono costruiti a livelli, che sono da intendersi come strutture di neuroni, funzioni matematiche che rappresentano la reazione della rete agli input, e interconnessioni, che descrivono le relazioni tra neuroni dello stesso layer e tra i layer adiacenti (Fig. 9). L'addestramento delle reti è avvenuto tramite la tecnica di retro-propagazione dell'errore, in cui il risultato viene confrontato con le etichette del dataset, propagando una correzione che modifica la struttura della rete stessa e con l'obiettivo di minimizzare l'errore all'iterazione seguente. Le iterazioni successive di questo processo permettono al modello di convergere stabilizzandosi verso una struttura in grado di risolvere con successo il problema, che in questo caso è l'identificazione nell'immagine delle cisterne. Le fotografie da UAV vengono caricate sui contenitori virtuali, detti bucket, dell'infrastruttura e vengono etichettate manualmente costituendo il dataset taggato attraverso gli strumenti di Vertex AI, piattaforma di Machine Learning Lifecycle Management (Lakshmanan, 2022).

Per taggare il dataset le fotografie sono state etichettate manualmente secondo il seguente

procedimento: tramite Vertex AI sono state riconosciute tutte le cisterne che erano visibili nelle fotografie, evidenziandone i perimetri con dei riquadri di delimitazione; a ogni riquadro di delimitazione è stata associata l'etichetta che classifica il contenuto in base a cosa è, alla forma e al materiale. Gli elementi tecnologici sono stati etichettati secondo le seguenti categorie: cisterne metalliche cilindriche, cisterne metalliche con forma di parallelepipedo, cisterne in plastica cilindriche e parabole metalliche. L'etichettatura delle foto fornisce la soluzione nota ai modelli

(Fig. 10), ovvero ciò con cui avviene il confronto nella fase di test per ottenere i valori della performance. Le immagini da UAV, che permettono di osservare le cisterne in prospettive differenti, consentono l'etichettatura delle cisterne da differenti punti di vista.

Questo processo, sebbene possa apparire scontato, aumenta notevolmente la riconoscibilità delle cisterne e permette di mitigare l'effetto di occlusione della scena dovuto alla presenza di detriti sui tetti, fornendo ai modelli di Deep Learning differenti prospettive dei medesimi oggetti. Sono

stati identificati ed etichettati 6.257 elementi tecnologici su 41 edifici in fotografie da UAV. La scelta di etichettare anche le parabole oltre che le cisterne è stata dettata dalla volontà di agevolare l'identificazione. Se viste dall'alto esse sono simili alle cisterne metalliche cilindriche e la loro etichettatura permette di ottimizzare il modello ed effettuare verifiche su falsi positivi e negativi.

La quarta fase della ricerca riguarda la sperimentazione dei modelli addestrati sul dataset etichettato. Il test e il deployment, ovvero la distribuzione del modello di machine learning, sono

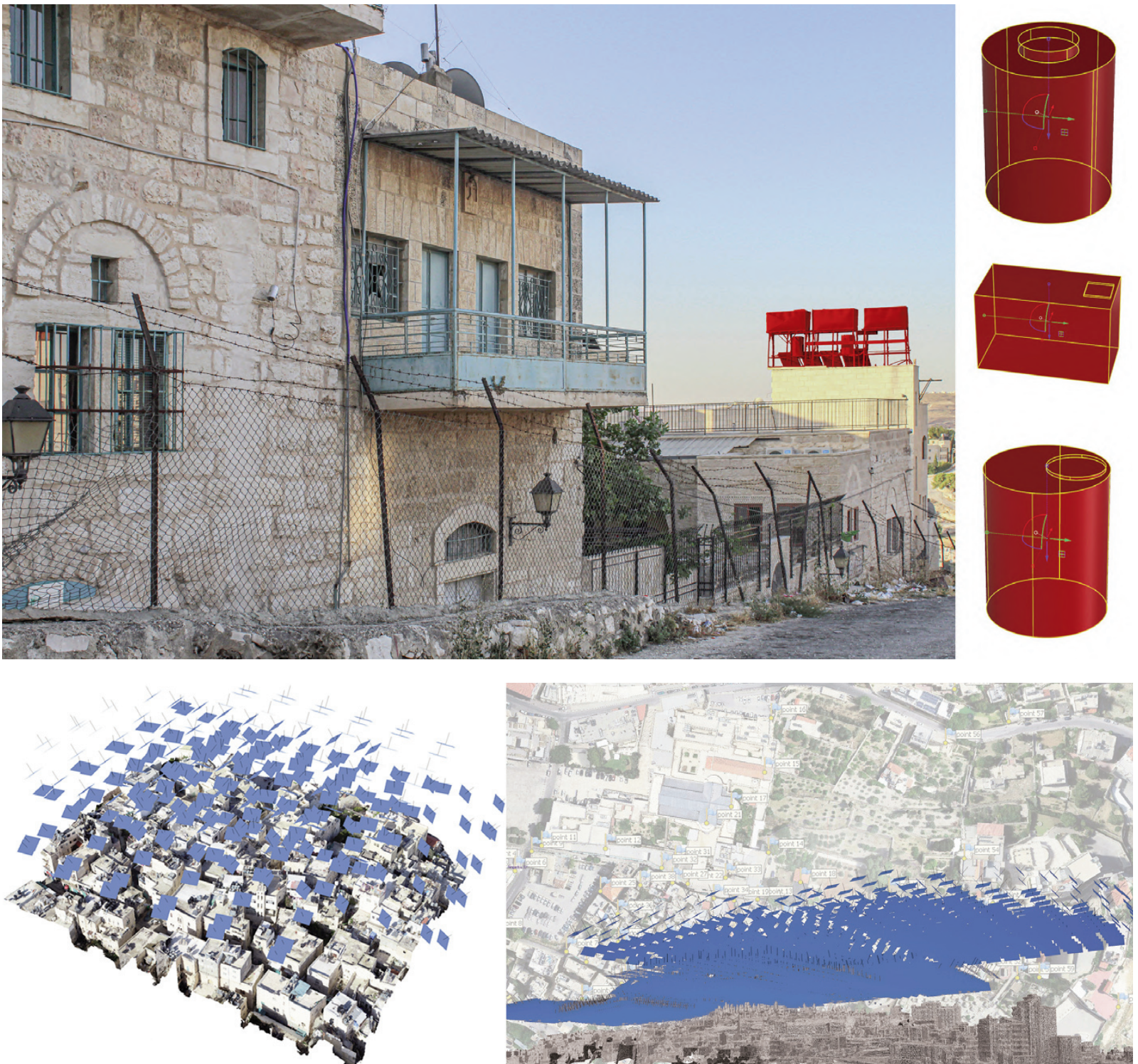


Fig. 2 | Example of buildings with flat roofs, with visible cisterns and accumulations of objects: on the left, elevated cisterns and photovoltaic panels (highlighted in red) placed on the edge of a historic building; on the right, the three-dimensional schematic of the cisterns analyzed (credit: E. Doria, 2022).

Fig. 3 | Example of a Structure from Motion photogrammetric capture made in the 3D Bethlehem project in which the location where individual photographs were taken is evident (in blue); some of the control points chosen to dimension and verify the photogrammetric model are visible in the background (credit: '3D Bethlehem' Project, E. Doria, 2022).



Fig. 4 | The cisterns on buildings in the historic center, highlighted in red; the photographs were taken from the ground and show how cisterns are strongly characterizing elements of the city's image (credit: E. Doria, 2022).

stati realizzati direttamente sul fotopiano ottenuto dal modello fotogrammetrico georeferenziato.

L'utilizzo del fotopiano per il deployment è uno dei punti di forza della ricerca perché permette di associare ai riquadri di delimitazione delle coordinate spaziali collocando le predizioni dell'algoritmo in un sistema con georeferenziazione spaziale. Il modello, addestrato e validato, è interrogabile con una chiamata RESTful alle API del progetto. Esso restituisce come output un file contenente le coordinate degli oggetti identificati con indicata la confidenza con cui ciascuno oggetto viene identificato come facente parte di una delle classi per cui il modello è stato addestrato. I risultati ottenuti sono visibili anche sotto forma di immagine, nella quale le cisterne e le parabole riconosciute sono evidenziate con dei rettangoli colorati in base all'etichetta che gli è stata associata dal modello (Fig. 11).

Risultati e criticità | La piattaforma utilizzata ha permesso la progettazione e la costruzione di un'architettura cloud scalabile e serverless, che può fungere da supporto per questo e per i futuri progetti che includeranno fasi di processamento algoritmico delle immagini. L'infrastruttura è accessibile e fruibile da remoto tramite l'utilizzo di un browser e pertanto risulta possibile utilizzarla direttamente in situ durante la campagna di acquisizione, accelerando le procedure di post processing del materiale acquisito. L'impiego di immagini fotografiche da UAV scattate da diverse angolazioni fornisce ai modelli molteplici punti di vista contribuendo a ottimizzare le performance in caso di occlusione dell'elemento da identificare rispetto al contesto, come in presenza di detriti o zone d'ombra.

La performance del modello viene valutata tramite due metriche, Recall e Precision, che mi-

surano la correttezza del risultato dal punto di vista sia della completezza degli oggetti identificati sia della corretta classificazione degli oggetti. Maggiore è la precisione minore è il numero di falsi positivi, maggiore è la recall minore è il numero di falsi negativi o di previsioni mancate. La soglia rappresenta il valore di confidenza oltre il quale il modello assegna un'etichetta, variabile tra 0 e 1 in quanto identifica una probabilità. Con una soglia impostata a 0,5 i risultati ottenuti sono stati 80,5% di precision e 28,0% di recall. Dato che all'aumentare del numero di oggetti etichettati migliora la convergenza del modello, per cercare di ottenere migliori risultati è stato fatto un ulteriore passaggio triplicando le dimensioni del dataset di addestramento. Il modello è stato quindi riaddestrato e, valutandolo sempre con soglia a 0,5, ha evidenziato prestazioni particolarmente migliorate dando come risultati 90,8% di precision e 61,0% di recall (Fig. 12). In fase di post-produzione dei risultati, un componente software si occupa di collegare le informazioni georeferenziate agli output del modello di DL abilitando la localizzazione degli elementi identificati.

Dai risultati ottenuti e dai processi metodologici sviluppati nella sperimentazione, il riconoscimento automatico degli elementi soddisfa i requisiti che erano stati valutati nella fase di progettazione della sperimentazione ma si possono evidenziare alcune criticità, suddivisibili in due categorie, di metodo e di dato. Le criticità di metodo riguardano principalmente la fase di acquisizione fotogrammetrica UAV, che deve essere accuratamente pianificata con piani di volo archiviabili, in modo da poterli riutilizzare nelle acquisizioni successive, e la costruzione e aggiornamento del dataset di addestramento. L'etichettatura del dataset è un processo manuale temporalmente di-

spendioso e che necessita di particolare cura da parte dell'operatore, in quanto rischia di introdurre sbilanciamenti nel dataset che si propagherebbero al modello in fase di training.

Le criticità del dato fotografico emerse sono legate a condizioni ambientali quali la luce e l'ambiente in cui sono poste le cisterne. Luce intensa e ombre sulle coperture creano zone di contrasto più difficilmente interpretabili e che necessitano di una pianificazione personalizzata in base al territorio e al periodo dell'anno. L'ambiente in cui si trovano le cisterne influisce sul riconoscimento per la presenza di detriti o depositi di rifiuti che rendono poco leggibili gli elementi tecnologici. La piattaforma cloud permette alti livelli di gestione della privacy, consentendo un controllo degli accessi alle diverse componenti dell'infrastruttura e di gestire separatamente l'esposizione dei dati fotografici delle analisi condotte e della gestione dei modelli.

Conclusioni e sviluppi | Il censimento automatizzato degli elementi tecnologici può rappresentare un supporto per la Municipalità al processo di gestione e manutenzione del centro storico, specialmente se utilizzato come strumento di monitoraggio dello stato di fatto e come piattaforma di confronto tra le condizioni temporali documentate. Utilizzando un'immagine georeferenziata come supporto dell'output, ogni identificazione degli elementi può essere ricollegata alle sue versioni temporali. È possibile pianificare le attività di monitoraggio delle coperture tramite strumentazione UAV a intervalli temporali a cadenza regolare (gestione ordinaria) e una tantum, nel caso di un intervento straordinario. Per ogni acquisizione è possibile utilizzare l'algoritmo già addestrato, applicandolo direttamente ai nuovi fotopiani ottenuti. Questo diviene possibile non solo per la città di

Betlemme, ma per tutti i territori limitrofi con elementi tecnologici paragonabili, con l'obiettivo di ottenere un archivio storico del rilevamento degli elementi tecnologici sul quale sviluppare analisi decisionali per la corretta gestione della città.

La metodologia presentata è applicabile in numerosi contesti, grazie alla combinazione della capacità di acquisizione dei dispositivi UAV e alla flessibilità delle piattaforme e dei modelli. Con sufficienti esempi iniziali per allenare il modello è possibile aumentare gli oggetti noti ai dataset etichettati, aggiungendo elementi al novero di quelli identificabili dai modelli quali piscine, parcheggi, coperture per effettuare confronti con dati catastali. Agendo poi sui dati e aggiornando il dataset senza modificare l'infrastruttura o la costruzione dei modelli, è possibile adattare facilmente il contesto di applicazione e la destinazione d'uso dei modelli stessi. Vista la natura disaccoppiata e modulare del progetto di ricerca, basato su molteplici componenti software interconnessi – stoccaggio dei dati, flussi di dati, addestramento e distribuzione dei modelli – è possibile agire separatamente su essi per aggiornamenti del dataset o cambio di tecnologie in maniera puntuale, facilitando la manutenzione e la risoluzione di problematiche che possono emergere.

Inoltre, per come è strutturato il processo descritto, è fondamentale sottolineare come sia possibile abilitare alla fruizione delle previsioni del modello diverse entità in modo 'puntuale', ovvero è possibile decidere di volta in volta chi ha accesso alle componenti del progetto e soprattutto a quali. La componente di innovazione è legata alla scelta dell'infrastruttura su cui si è proceduto a sviluppare la ricerca; utilizzando le infrastrutture cloud che sono oggi a disposizione, con un budget iniziale riferito allo strumento con cui si effettuano le acquisizioni close-range e alla workstation di sviluppo dei processi, si ottiene uno strumento poco soggetto a obsolescenza tecnologica, immune dai danneggiamenti hardware ed economico grazie all'utilizzo dello storage cloud, con particolare riferimento alla fase di prototipazione dei processi.

Il prossimo obiettivo della ricerca è rendere fruibili questi dati all'interno del GIS tridimensionale della città, già realizzato nel progetto 3D Bethlehem, per poter analizzare i singoli elementi tecnologici direttamente dal sistema informativo (Figg. 13, 14). In una logica di gestione e conservazione del Patrimonio costruito, i processi di automazione permettono di facilitare l'attivazione del monitoraggio delle situazioni di criticità e di mantenimento delle condizioni virtuose (Miceli, Morandotti and Parrinello, 2020), semplificando e incrementando la pianificazione di protocolli per la manutenzione preventiva del patrimonio costruito.

The proposal reports the outcome of research focused on the automation of the census of architectural elements in the urban environment via Object Detection techniques, aimed at the development of a system for Built Heritage management. The acquisition of technological, building and material information is a key phase in the conservation process, since the 'complete documentation' of the Heritage allows researchers to lead with additional consciousness management and

enhancement activities (Gasparoli, 2012; Montella, 2009). In this context UAVs (Unmanned Aerial Vehicles), along with post-production technologies that enable the construction of digital twins enable innovative monitoring methods in which the involvement of human technicians can be minimised towards increased workplace safety and the optimisation of maintenance and management costs.

The project scope is focused on the covering of the historical centre of Bethlehem (Fig. 1) with a particular emphasis on the water tanks placed by civilians on the roofs. Surveying such elements from close-range UAV acquisitions allows detection of loads on the historical buildings located in a seismic area (Fig. 2). The goal is the development of a usable informative digital product to be made available to technicians belonging to the Municipality of Bethlehem. Such a tool will allow the monitoring of the built environment of the city and the connection of derived information with pieces of data related to the state of conservation

subsequently surveyed in the field. The proposal fits within the scope of a broader research and cooperation project '3D Bethlehem – Management and control of urban growth for the development of Heritage and Improvement of life in the city of Bethlehem' co-financed by AICS and coordinated by the Department of Civil Engineering and Architecture – DICAr (Parrinello, 2019).¹

The Project, spanning from 2018 to January 2020 includes the digital survey, the filing and digital modelling of the historical city led to the development of a three-dimensional GIS aimed at monitoring interventions performed on the built heritage. Multiple photogrammetric surveying activities contributed to a dataset composed of more than 9,000 images of the coverings, acquired with UAV technologies (Fig. 3) such as DJI Phantom 4 PRO, DJI Mavic, DJI Mavic Mini (Parrinello and Picchio, 2019; Picchio, 2019). This dataset was the basis of the proposal, self-financed and developed after the conclusion of the '3D Bethlehem' project in agreement with the coordinator Prof. Sandro Parrinello.



Fig. 5 | On the left, an orthoimage obtained by processing a three-dimensional mesh model from a photogrammetric point cloud from UAV; on the right, a portion of the textured model generated from photogrammetric acquisitions of Bethlehem's historic center (credit: E. Doria, 2022).

Fig. 6 | Photographic images of the same building taken from different angles; specifically, it is a building in Manger Square, opposite the Basilica of the Nativity, which houses municipal offices, a bank, and some tertiary functions (credit: E. Doria, 2022).

Fig. 7 | Photographic images of the same building taken from different angles; in particular, the main building in which the municipality offices are located and on which some experimentation with BIM modelling integrated into the GIS model is being carried out is shown (credit: E. Doria, 2022).

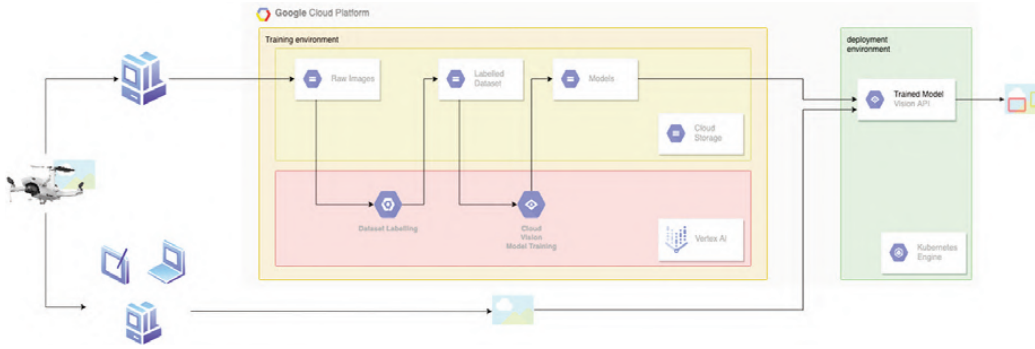


Fig. 8 | Architectural diagram of the project, showing the environment (on premises/cloud) and technologies involved (credit: L. Carcano and E. Doria, 2022).

Close range photogrammetric acquisitions of the coverings were used as the training dataset for Deep Learning models, requiring massive amounts of data to converge towards a solution, namely reaching a status in which the predicted results settle within an acceptable range of error from the known ground truth. The choice of dataset was performed following the filing for the GIS system, from which a clear picture of the inherent complexity of the urban environment, especially focusing on coverings, took shape (Doria, 2022). The goal is to be ultimately found in the automation of the process of the census of technological covering elements, supporting monitoring and programmatic intervention actions on the Heritage. Chosen technological elements for the first instance of the experiment are water tanks, placed in a widespread and unregulated fashion upon the coverings in the city. Water tanks are both an asset for the population during periods of water rationing and constructive criticalities at urban level; such tanks are concentrated load elements on 69.3% of the surveyed buildings, deployed unevenly and without adequate soundness analyses and potentially compromising the structural integrity of buildings in a seismic environment (Abueladas et alii, 2021; De Marco, 2022).

The case study | Water tanks across the centre of Bethlehem and in the surroundings can be categorised by means of shape, size and material. For the scope of these experiments the tanks were divided into three macro-categories (Fig. 4): cylindrical plastic tanks of recent construction and good state of conservation (54%), older cylindrical and squared metal tanks in a far worse state of conservation (46%)². During water sourcing, performed in a non-consistent fashion and only during given days, tanks in worse states of conservation, placed on roofs and manifesting flow regulation issues, leak on building facades and on the piping and plants found along the path of the water. Water leaks are among the causes of pathologies in the buildings such as the proliferation of vegetation – both intensive and extensive, depending on the placement and state of use of the built architectures – and moist areas with evident alterations of facade materials.

The city of Bethlehem is built in hilly territory, and the historical city centre is crossed by steep roads and multiple staircases connecting neighbourhoods built at different heights on the hills. This morphology allows the water tanks to be seen from all over the city, both for those located on re-

cent and historic buildings. The placement of tanks in historical and recent constructions varies: in recent buildings tanks follow a regular pattern; in historical constructions, which show vaulted covering with visible extrados and complex morphology typical of local housing³, the tanks are placed in an uneven distribution across the coverings. In many historical buildings, the covering is used as storage for debris and scrap, interfering with tank visibility and hindering maintenance.

State of the art | The analytical process aimed at understanding the urban context is now connected to the construction of trustworthy databases that allow both strategies to manage the Built Heritage and the communication and virtual fruition of digitised spaces. Within city centres information concerning the events that affected the city is preserved across the years defining a variable ensemble of data that can and should be digitised to allow for proper management (Bocconcino and Manzone, 2019). The increment in the amount of information involved in the public decision process required the implementation of novel technologies and techniques for information management. Information and Communication Technologies (ICT), quickly developing due to technological advancements, influence the decision-making framework when applied to public management, creating monitoring tools for technologically advanced managing (Kouziokas and Perakis, 2017).

In order to gather a large amount of data and metadata, to be subsequently linked when creating massive datasets, the technological development in acquisition methods allowed for the simplification and acceleration of processes. Even though inference processes for the knowledge of heritage start from acquired information, data collection itself is not a source of knowledge per se, requiring a variable correlation process. Computers can leverage the ever-growing amount of non-ordered information, defining novel ways of generating knowledge by analysing recurrent patterns and correlations either a-priori or ex-post. Developments in the field of Artificial Intelligence (AI), a term coined by Prof. John McCarthy and defined as «[...] the science and engineering of making intelligent machines» (McCarthy, 2007, p. 2), become crucial and nowadays include a wide set of technologies and techniques aimed at automating cognitive tasks (Corea, 2019).

Since their beginnings, the research topics related to AI focused on the creation of models imitating the dynamics of human intelligence, based

on comparison and the reworking of external stimuli, that can be reworked and classified based on active research tracks (Buratti, Conte and Rossi, 2021). Machine learning is an umbrella term including multiple techniques and approaches allowing improvement in the performance of algorithms in identifying patterns in input data, mimicking the approach of the human brain when learning and perfecting new skills.

Approaches based on neural networks and Deep Learning (DL) models start from random interconnections across computational elements, perfected through the learning process, enabling the management of use cases of growing complexity by increasing both computational complexity and interconnections among network elements (Mishra, 2021). Scalability allows DL models to work with growing amounts of data and, therefore is well suited for massive structured (tabular data and numbers) and unstructured (natural language, images) datasets. The branch of ICT concerned with the extraction of information from images is Computer Vision, intersecting AI in all the applications in which the capability to learn from and manage unstructured data allows the solution of complex problems such as segmentation and object detection.

The automated identification of technological elements presented in this proposal is based on object detection, the branch of computer vision that allows the identification of objects within images. The scope of the implemented algorithm is dual: on one side, isolate the object of interest in a complex scene, on the other correctly classify the nature of the object. Since the origins and development of Convolutional Neural Networks – CNN (Lecun and Bengio, 1998; Cosovic and Jankovic, 2020), such deep learning models became the state-of-the-art approach thanks to the high performance within this scope of tasks (Redmon et alii, 2016). In this specific case, the chosen approach is that of AutoML, a family of models which features engineering approaches that exploit ensemble and transfer learning techniques to enable quicker development of baseline Machine Learning (ML) models pushing towards the democratisation of AI (He, Zhao and Chu, 2021).

Managing the photographic dataset, its potentially growing size and the therefore increased computational complexity of the object detection task put a hard boundary on the actionable approaches. To enable on-premises management of a growing dataset, storage and computational resources would have needed progressive improvement via the acquisition of new hardware. This proposal aims at dismissing this boundary by deploying the experiments on a Cloud Computing infrastructure, used as an Integrated Machine Learning Platform, allowing abstraction from the computational power available in situ bypassing the limitations imposed by physical hardware and infrastructures (Rivera, 2020).

Related research | The aforementioned Object Detection techniques were developed to learn to identify an entity based on existing data and allowing predictions on unknown data. Starting from image data, three main goals can be identified: classification, detection and segmentation (Xiao et alii, 2020). Different approaches to the pursuit of these goals are well documented as a

support tool for the management and documentation of Cultural Heritage and urban planning (Li et alii, 2020).

In the urban environment, the use of such techniques is documented for the monitoring of criticalities of infrastructures and construction sites (Dandabathula et alii, 2019; Wang et alii, 2022), urban waste management (Majchrowska et alii, 2022), the control of compliance with building regulations (Cunha et alii, 2021) and environmental and agricultural monitoring (Carbone et alii, 2022). The application of such techniques, leading to high performance, is documented in the field of Built cultural heritage, both in the identification of unknown Heritage thanks to satellite and close-range imaging (Due Trier, Reksten and Løseth, 2021) and on already documented heritage (Monna et alii, 2021; Zou et alii, 2019; Hatir

and Ince, 2021) as a support to decision making when planning interventions.

The presented research proposal focuses on the use of Object Detection techniques for the planning and management of urban territory using cloud platforms and open-source technologies, intersecting the existing control area in relation to structural and plant safety issues. Machine learning models and pipelines are based on open-source technologies such as Python libraries and Kubeflow and can therefore be exported and shared with collaborators outside the platform. External collaborators, both technical and non-technical, such as Universities, companies and municipalities can leverage the exported models in cloud infrastructures different from the one presented in this work, allowing for high degrees of cooperation across multiple entities. Such coop-

eration can happen both in the development phase and in the final usage phase, and pre-trained models can be queried via the dedicated Application Programming Interface (API) used to expose them.

Methodology and research phases | Georeferenced images are the backbone on which to structure and define a repeatable census process, developing preferred UAV flight plans (DJI Phantom 4 PRO) and computing the predictions of object detection models on the scaled and georeferenced orthoimage.

The first phase of the research, based on data gathered during the surveying phase of the 3D Bethlehem project, included the analysis of close-range photogrammetric images of coverings and post-production of data aimed at the construction

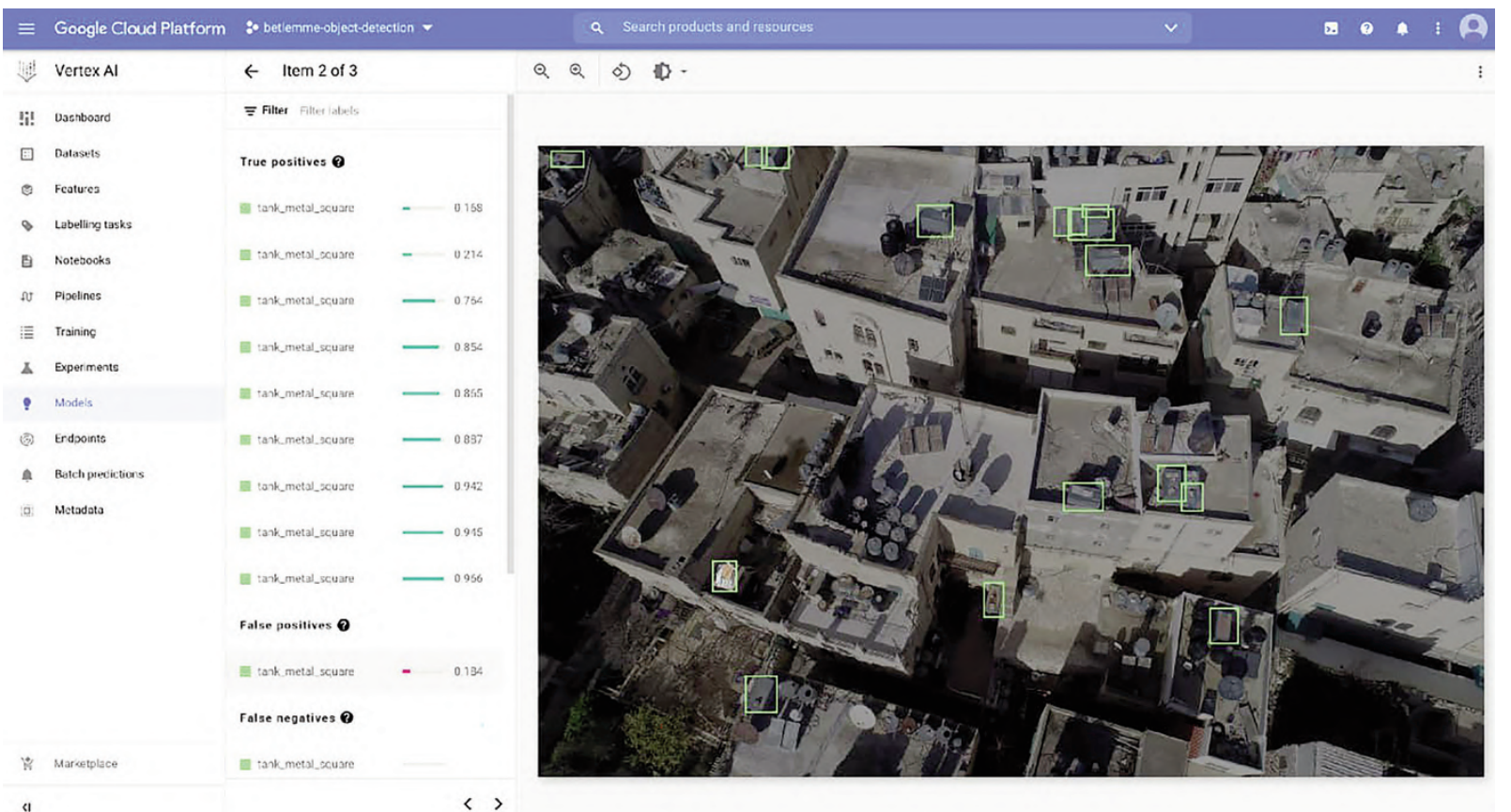
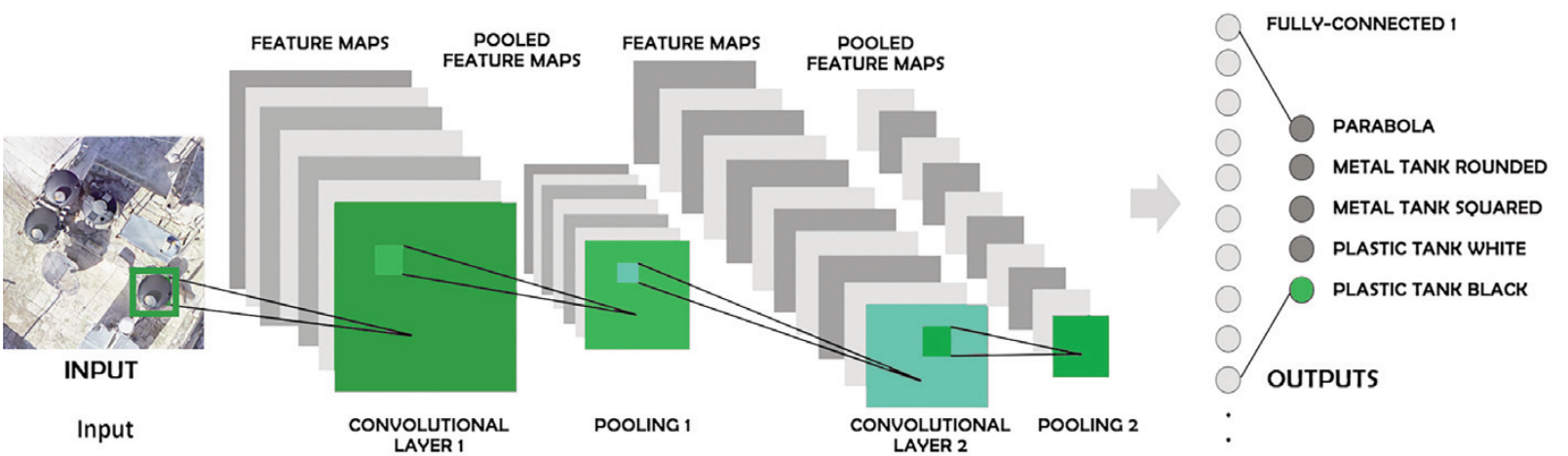


Fig. 9 | Scheme of the basic architectural components of a CNN: the signal feeds a convolutional layer, passes through a pooling layer that reduces the size and advances the signal to the underlying layers of the network (credit: E. Doria, 2022).

Fig. 10 | Model testing phase carried out on the photos in which the technological elements were labelled to get feedback on the values obtained (credit: E. Doria, 2022).

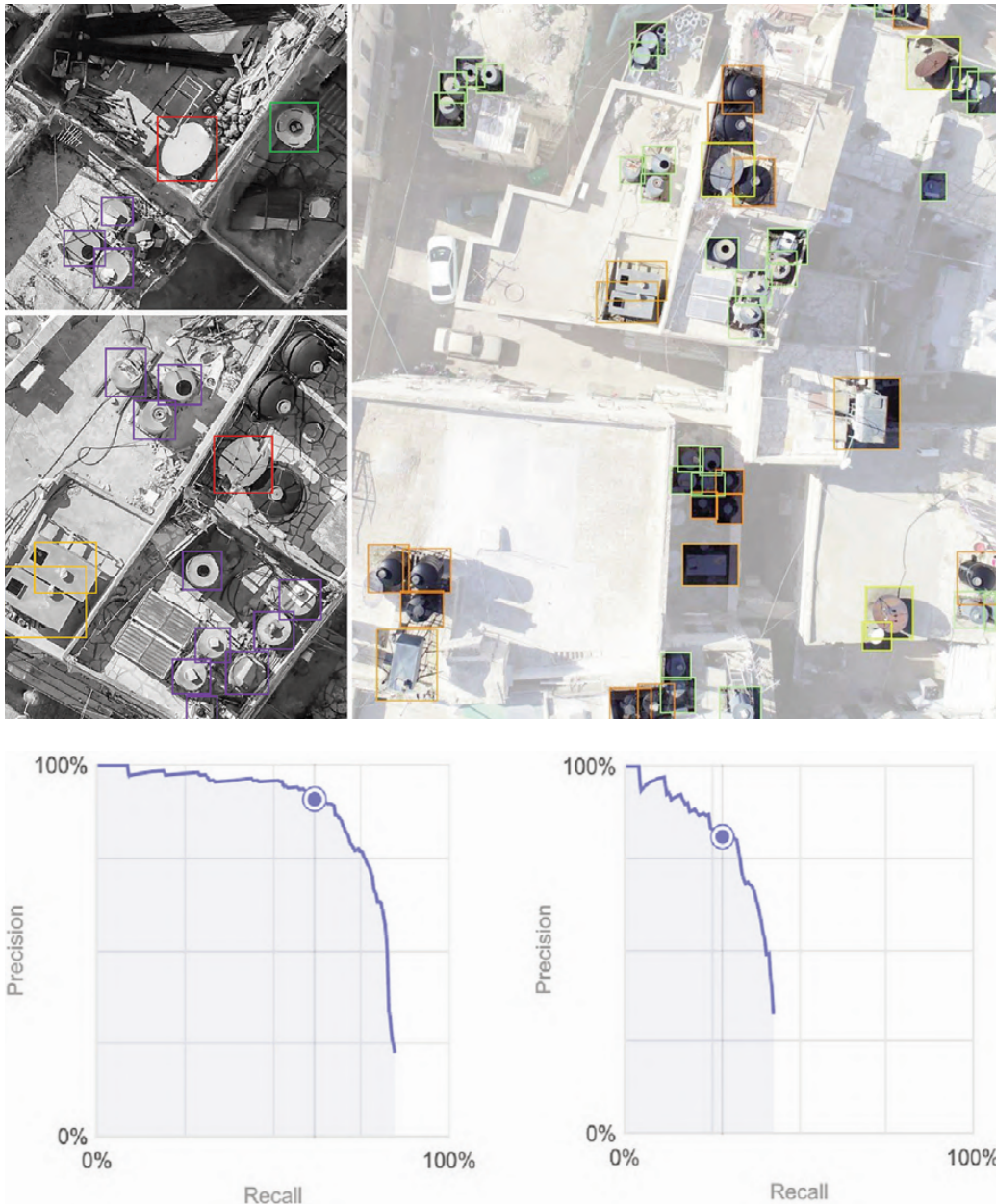


Fig. 11 | Some images returned by the model with bounding boxes of recognized elements: in the lower left, some black plastic tanks that were not recognized in the first phase of testing are visible; to improve this performance, it was increased in initial dataset by increasing the labelled elements and the number of photographs used for training (credit: E. Doria, 2022).

Fig. 12 | Graph of precision and recall after dataset increase compared with before; training images increased from 30 to 85 and the increase in reliability seen is evident (credit: E. Doria, 2022).

of ortho-images. The ortho-image of coverings was extracted from the Structure from Motion (SfM) model (Picchio, 2020): such a model was constructed from flight planned by the operator⁴ with acquisitions performed both zenital and oblique with respect to the four sides of each area, totalling five plans per each smaller sized area (Figg. 5-7). Photogrammetric models have been metrically scaled using the point cloud built with Terrestrial Laser Scanner (TLS) – Faro CAM2 S150 as a reference. The georeferenced ortho-image was used as the base for identifying water tanks. Referenced ortho images allow repeatability of the acquisition at multiple time intervals when the need for a new surveying campaign arises.

The second phase concerned the choice and development of the infrastructure upon which experiments are constructed. A Cloud Computing infrastructure was used for model training, in order

to take advantage of and ensure scalability. Such infrastructure, built on the enterprise-grade Google Cloud Platform allows end-to-end management of data lifecycle, from image ingestion, input management, dataset construction, database enrichment, training and deployment of machine learning models, up until model output fruition (Fig. 8). The infrastructure on which this project relies is completely serverless, which allows scaling via the addition of storage or virtual machines to transparently cope with the need for additional storage capability or computing power, adapting the architecture to the precise need of the case study at each moment in time.

Object Detection tasks rely heavily on convolutional neural networks, in which a sliding window is applied to portions of the image: computational complexity grows with the number and resolution of the images used for training. Data processed

by convolution layers is used to feed subsequent layers in the network architecture, towards the classifier that will assign a label to the images.

Phase three concerned the choice and training of the models. Deep Learning models are built with layers, to be considered as neuron structures, mathematical functions that represent the reaction of the network to incoming input signals and interconnections, representing the relationships between neurons of the same layer and across adjacent layers (Fig. 9). Network training was performed with error back-propagation, evaluating the predicted solution against the ground truth, propagating a correction in network weights to modify the network structure to minimise error at the next training epoch. Multiple training epochs allow the model to converge and settle towards a structure providing a level of performance suited to the needs of the use case, here the identification of tanks in the images. UAV photographs are loaded on storage buckets (containers for data) and are manually labelled building the tagged dataset using the tools provided by the Vertex AI suite, Google's platform for Machine Learning Lifecycle Management (Lakshmanan, 2022).

To tag the dataset, images have been manually labelled using the following process: in Vertex Ai human operators identified all visible tanks, delimiting them with bounding boxes: each bounding box was then labelled with object type, material, and shape. Labelled technological elements fell into one of four categories: square metal tanks, round metal tanks, round plastic tanks, and metal satellite dishes. This process constructed the so-called ground truth for the models (Fig. 10), which was used for comparison purposes and performance evaluation. UAV images allowed the observation of elements from multiple points of view.

Despite seeming obvious, this characteristic enabled improved recognisability of tanks, providing different perspectives and angles of the same objects to Deep Learning models. For this case study, a total of 6,257 technological elements were identified in 41 buildings, from UAV photographs. Labelling satellite dishes alongside tanks was aimed at identification; the similarity to round metal tanks from a zenital perspective meant the labelling process created further focus on the minimisation of false positives and false negatives.

Phase four concerned the testing of models trained on labelled dataset. Test and deployment, as per the distribution and fruition of the machine learning model, were performed on the ortho-image from the georeferenced photogrammetric model. Using referenced ortho-images allowed linking of bounding boxes with spatial coordinates precisely locating predictions in a referenced geospatial system. The model once trained and validated was deployed and made available via an 'http' call to project RESTful APIS. The output file contains the coordinates of bounding boxes per each identified object and the confidence with which class labels are assigned to each identified element. Results can be presented as images, in which identified elements, tanks and dishes, are highlighted with bounding boxes colour coded based on the labels (Fig. 11).

Results and criticalities | The platform used allowed the design and construction of scalable and serverless cloud architecture, supporting this and

future projects including algorithmic image processing steps. The infrastructure is accessible and actionable remotely via a web browser and can therefore be used directly in situ during the acquisition campaign, accelerating post-processing procedures of the acquired material. The use of photographic images from UAVs taken from dif-

ferent angles provided the models with multiple points of view, helping to optimise performance in the event of occlusion of the target elements, such as in the presence of debris or shaded areas.

Model performance is evaluated by means of two metrics, Recall and Precision, measuring the correctness of the result in terms of the complete-

ness of the identified objects and the correct classification of such objects respectively. The higher the precision, the lower the number of false positives, the higher the recall, the lower the number of false negatives or missed predictions. The threshold represents the confidence value above which the model assigns a label, varying between 0 and

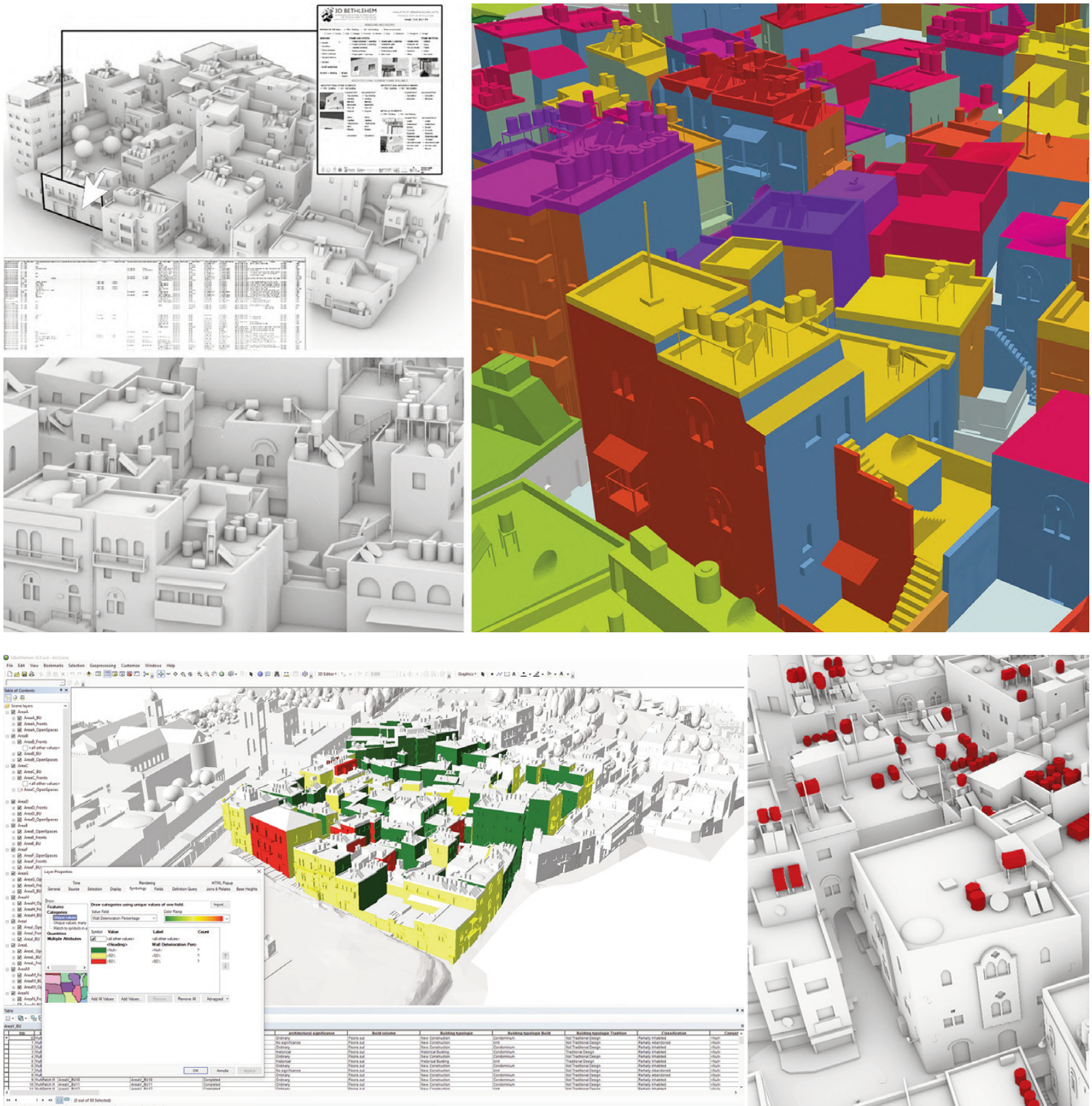


Fig. 13 | Three-dimensional model created during the 3D Bethlehem project: on the left, the model imported into the GIS environment with sample diagram of the urban census connections to the platform; on the right, an image of the application of thematic maps to the model, which are derived from the visualization of census data (credit: E. Doria, 2022).

Fig. 14 | Screenshot of the three-dimensional geo-referenced information system in which the integration of the presented research data is planned, thanks to which it will be possible to have punctual information on the positioning of the cisterns and their state of preservation; the integration in the information system will allow for analysis by combining the data of the cisterns with those of the buildings on which they are positioned (credit: E. Doria, 2022).

1 as it identifies a probability. With a threshold set at 0.5, the results obtained were 80.5% precision and 28.0% recall. As the number of labelled objects increased, the convergence of the model improved, and to obtain better results a further step was taken by increasing the size of the training dataset by a factor of three. The model was then retrained and when re-evaluated with a threshold of 0.5, showing notable performance improvements, yielding 90.8% precision and 61.0% recall (Fig. 12). When post-producing the results, a software component took care of linking georeferenced information to the output of the DL model, enabling location of the identified elements.

From the results obtained and the methodological processes developed in the experiment, the automatic feature recognition satisfies the requirements that were assessed in the design phase of the experiment. Even though results satisfy expectations, some criticalities can be highlighted in the method and data. The method criticalities mainly concern the UAV photogrammetric acquisition phase, which must be carefully planned with archivable flight plans, allowing repeatability for subsequent acquisitions, and the construction and updating of the training dataset. The labelling of the dataset is a time-consuming manual process that requires special care, as it might introduce biases that may propagate to the model in the training phase.

The critical aspects of the photographic data that emerged were related to environmental conditions such as light and the environment in which the tanks are placed. Intense light and shadows on the roofs create areas of contrast that are more difficult to interpret and require customised planning, taking into consideration the terrain and time of year. The environment in which the tanks are located affects recognition due to the presence of debris or scrap deposits that hinder the identification of technological elements. The cloud platform enables high granularity of access man-

agement, allowing access control to the different components of the infrastructure and separate management of the photographic data of the conducted analyses and model management.

Conclusions and future developments | The automated census of technological elements can support the municipality in the process of the management and maintenance of the historic centre, especially when employed as a monitoring tool of the as-is state and as a platform for the comparison of the such state through time. Using a georeferenced image as output support, each element identification can be linked back to existing states in time. Covering monitoring activities can be planned using UAV instrumentation at regular time intervals (ordinary) and one-offs, in the case of extraordinary interventions. For each acquisition, it is possible to use pre-trained algorithms, computing predictions on the newly obtained ortho-images. Such a process is enabled not only for the city of Bethlehem but for all neighbouring territories with comparable technological elements, aiming at the construction of a historical archive of technological elements detections on which to develop decision-making analyses for the proper management of the city.

The presented methodology is applicable in numerous contexts due to the combination of the acquisition capability of UAV devices and the flexibility of platforms and models. With sufficient examples to train the model, it is possible to append additional known objects to the labelled dataset, adding elements to the list of identifiable ones such as swimming pools, car parks, and roofs to make comparisons with cadastral data. By acting on the data and updating the dataset, without changing the infrastructure or construction of the models, it is possible to easily adapt the application context and the designated use of the models. Given the decoupled and modular nature of the research project, based on multiple intercon-

nected software components – data storage, data flows, model training and deployment – it is possible to act separately to update datasets or swap technologies as plug-ins, facilitating the maintenance and bug fixing of issues that may arise.

Furthermore, given the way the described process is structured, it is essential to emphasise how it is possible to enable different entities to use the model predictions in an ‘on-demand’ manner, i.e. it is possible to decide from time to time who has access to which components of the project and, above all, to which ones. The innovation component is linked to the choice of the infrastructure on which the research was carried out; by relying on the currently available cloud infrastructures, with an initial budget for the close-range acquisition tool and comparing to the process development workstation, the prototyping phase with the developed infrastructure is more convenient as the built cloud infrastructure is subject to little technological obsolescence and immune to hardware damage.

The next research objective is to enrich the three-dimensional GIS of the city, in scope of the 3D Bethlehem project, with the obtained predicted data, to be able to analyse individual technological elements directly from the information system (Figg. 13, 14). In the realm of management and conservation of the built heritage, automation processes facilitate the activation of the monitoring of critical situations and the maintenance of virtuous conditions (Miceli, Morandotti and Parrinello, 2020), simplifying and increasing the planning of protocols for the preventive maintenance of the built heritage.

Notes

1) ‘3D Bethlehem – Management and control of urban growth for the development of heritage and improvement of life in the city of Bethlehem’ is a cooperation project co-financed by the Italian Agency for Cooperation and Development (AICS). The project was coordinated by the Municipality of Pavia, with a partnership composed of the Municipality of Bethlehem, the University of Pavia (scientific coordination), Bethlehem University, the Province of Pavia, the Pavia Province Order of Engineers, SISTERR Pavia Territorial System for International Cooperation, ANCI Lombardia, VIS – International Voluntary Service for Development NGO and the Palestinian Engineers Association - Jerusalem Centre. The project is scientifically coordinated by Prof. Sandro Parrinello and the research laboratory DAda-LAB of DICAr – Department of Civil Engineering and Architecture, University of Pavia. As of 2022, a second phase of the project is active: ‘Bethlehem SMART City – Environmental Monitoring System and Technological Renewal in the City of Bethlehem,’ which will help promote sustainable and resilient urban and terri-

torial development in the Bethlehem area under the scientific leadership of Prof. Parrinello and co-financed by AICS.

2) Information on the qualitative (state of preservation) and quantitative (percentages of their distribution and proportions) aspects of civil tanks was collected during the urban census carried out in situ using UAV instrumentation in the Bethlehem 3D Old Town project ‘3D Bethlehem’ Project. Due to the sensitivity of the data with respect to the territories in which they were collected and based on agreements with project partners, it is not possible to share the result of the census in its entirety in open access, but in this contribution significant data used for the results presented are extracted.

3) Historical neighbourhoods present a morphology stemming from typical building – awash (Dabdoub and Hifz al-Turāth al-Thaqāfi, 2005).

4) The close-range photogrammetric shooting operations via UAV of the Bethlehem Old Town were divided into different phases given the extent of the urban centre and were carried out per neighbourhood with common landmarks, in order to merge the different models into a single database.

Specifically, the shots used for the proposed research were taken by certified pilot Prof. F. Picchio, University of Pavia.

References

- Abueladas, A.-R. A., Niemi, T. M., Al-Zoubi, A., Tibor, G., Kanari, M. and Ben-Avraham, Z. (2021), “Liquefaction susceptibility maps for the Aqaba-Elat region with projections of future hazards with sea-level rise”, in *Quarterly Journal of Engineering Geology and Hydrogeology*, vol. 54, issue 2, pp. 1-12. [Online] Available at: [lyellcollection.org/doi/10.1144/qjegh2020-039](https://doi.org/10.1144/qjegh2020-039) [Accessed 30 September 2022].
- Bocconcino, M. M. and Manzone, F. (2019), “Sistemi informativi e strumenti grafici per la manutenzione di manufatti complessi | Information systems and graphic tools for the maintenance of complex buildings”, in Garda, E., Mele, C. and Piantanida, P. (eds), *Colloqui.AT.e 2019 – Ingegno e costruzione nell’epoca della complessità – Atti del Convegno – Torino, 25-28 Settembre 2019*, Edizioni Politecnico di Torino, Torino, pp. 678-688. [Online] Available at: researchgate.net/publication/348334768_Sistemi_informa

tivi_e_strumenti_grafici_per_la_manutenzione_di_manufatti_complessi_-_Information_systems_and_graphic_tools_for_the_maintenance_of_complex_buildings [Accessed 30 September 2022].

Buratti, G., Conte, S. and Rossi, M. (2021), "Artificial Intelligence, Big Data and Cultural Heritage", in Giordano, A., Russo, M. and Spallone, R. (eds), *Representation challenges – Augmented Reality and Artificial Intelligence in Cultural Heritage and Innovative Design Domain*, FrancoAngeli, Milano, pp. 29-33. [Online] Available at: re.polic.polimi.it/retrieve/e0c31c11-a30c-4599-e053-1705fe0aef77/Buratti_Conte_Rossi.pdf [Accessed 30 September 2022].

Carbone, C., Fawakherji, M., Trianni, V. and Nardi, D. (2022), "Photorealistic simulations of crop fields for remote Sensing with UAV swarms", in Parrinello, S., Barba, S., Dell'Amico, A. and di Filippo, A. (eds), *D-SITE. Drones – Systems of Information on Cultural Heritage for a spatial and social investigation*, Pavia University Press, Pavia, pp. 576-581. [Online] Available at: paviauniversitypress.it/catalogo/d-site/6693 [Accessed 12 October 2022].

Corea, F. (2019), *An Introduction to Data – Studies in Big Data*, vol. 50, Springer, Cham. [Online] Available at: doi.org/10.1007/978-3-030-04468-8 [Accessed 30 September 2022].

Ćosović, M. and Janković, R. (2020), "CNN Classification of the Cultural Heritage Images", in *2020 19th International Symposium INFOTEH-JAHORINA (INFOTEH)*, IEEE, pp. 1-6. [Online] Available at: doi.org/10.1109/INFOTEH48170.2020.9066300 [Accessed 30 September 2022].

Cunha, H. S., Schlauser, B. S., Wildemberg, P. F., Fernandes, E. A. M., dos Santos, J. A., de Oliveira Lage, M., Lorenz, C., Barbosa, G. L., Quintanilha, J. A. and Chiaravallotti-Neto, F. (2021), "Water tank and swimming pool detection based on remote sensing and deep learning – Relationship with socioeconomic level and applications in dengue control", in *PLoS ONE*, vol. 16, issue 12, pp. 1-24. [Online] Available at: doi.org/10.1371/journal.pone.0258681 [Accessed 12 October 2022].

Dabdoub Nasser, C. and Hifz al-Turāth al-Thaqāfi, M. (2005), *Anatreh Quarter – An Urban and Architectural Study of a Bethlehem Quarter*, Centre for Cultural Heritage Preservation, Bethlehem.

Dandabathula, G. B., Saini, S., Parikh, D., Sharma, P., Khandelwal, S. and Rao, S. S. (2019), "Automatic detection of overhead water tanks from satellite images using faster-RCNN", in *International Journal of Advanced Research in Computer Science*, vol. 10, n. 5, pp. 34-37. [Online] Available at: dx.doi.org/10.26483/ijarcs.v10i5.6466 [Accessed 12 October 2022].

De Marco, R. (2022), "Dal modello informativo ad un repository sulla stabilità delle strutture urbane – Una riflessione sulla valutazione della sicurezza negli aggregati storici della città [From the information model to a repository on the stability of urban structures – A reflection on the evaluation of safety in the historical aggregates of the city]", in Parrinello, S. (ed.), *3D Bethlehem – Gestione e controllo della crescita urbana per lo sviluppo del patrimonio e miglioramento della vita nella città di Betlemme | Management and control of urban growth for the development of heritage and improvement of life in the city of Bethlehem*, vol. 3, Edifir Edizioni, Firenze, pp. 203-224. [Online] Available at: researchgate.net/publication/361305496_Dal_modello_informativo_ad_un_repository_sulla_stabilita_delle_strutture_urbane_Una_riflessione_sulla_valutazione_della_sicurezza_negli_aggregati_storici_della_citta_From_the_information_model_to_a_re [Accessed 30 September 2022].

Doria, E. (2022), "Strategie di strutturazione della conoscenza – Innovazioni e attività programmabili tramite il censimento tecnologico del centro storico di Betlemme | Knowledge structuring strategies – Innovations and programmable activities via the technological census of the historic center of Bethlehem", in Parrinello, S. (ed.), *3D Bethlehem – Gestione e controllo della crescita urbana per lo sviluppo del patrimonio e miglioramento della vita nella città di Betlemme | Management and control of urban growth*

for the development of heritage and improvement of life in the city of Bethlehem, vol. 3, Edifir Edizioni, Firenze, pp. 51-67. [Online] Available at: researchgate.net/publication/361217094_Strategie_di_strutturazione_della_conoscenza_innovazioni_e_attivita_programmabili_tramite_il_censimento_tecnologico_del_centro_storico_di_Betlemme_Knowledge_structuring_strategies_innovations_and_prog [Accessed 30 September 2022].

Due Trier, Ø., Reksten, J. H. and Løseth, K. (2021), "Automated mapping of cultural heritage in Norway from airborne lidar data using faster R-CNN", in *International Journal of Applied Earth Observation and Geoinformation*, vol. 95, article 102241, pp. 1-19. [Online] Available at: doi.org/10.1016/j.jag.2020.102241 [Accessed 12 October 2022].

Gasparoli, P. (2012), "La manutenzione preventiva e programmata del patrimonio storico tutelato come prima forma di valorizzazione | Planned preventive maintenance in the preservation of historic heritage sites as a first form of valorisation", in *Techné | Journal of Technology for Architecture and Environment*, vol. 03, pp. 148-157. [Online] Available at: doi.org/10.13128/Techné-10840 [Accessed 30 September 2022].

Hatir, M. E. and İnce, I. (2021), "Lithology mapping of stone heritage via state-of-the-art computer vision", in *Journal of Building Engineering*, vol. 34, article 101921, pp. 1-12. [Online] Available at: doi.org/10.1016/j.job.2020.101921 [Accessed 12 October 2022].

He, X., Zhao, K. and Chu, X. (2021), "AutoML – A survey of the state-of-the-art", in *Knowledge-Based Systems*, vol. 212, article 106622, pp. 1-27. [Online] Available at: doi.org/10.1016/j.knosys.2020.106622 [Accessed 30 September 2022].

Kouziokas, G. N. and Perakis, K. (2017), "Decision Support System Based on Artificial Intelligence, GIS and Remote Sensing for Sustainable Public and Judicial Management", in *European Journal of Sustainable Development*, vol. 6, issue 3, pp. 397-404. [Online] Available at: doi.org/10.14207/ejds.2017.v6n3p397 [Accessed 30 September 2022].

Lakshmanan, V. (2022), *Data Science on the Google Cloud Platform*, Editions O'Reilly, Paris.

Lecun, Y. and Bengio, Y. (1998), "Convolutional networks for Images, Speech, and Time-Series", in Arbib, M. A. (ed.), *The Handbook of Brain Theory and Neural Networks*, MIT Press, pp. 255-258.

Li, K., Wan, G., Cheng, G., Meng, L. and Han, J. (2020), "Object detection in optical remote sensing images – A survey and a new benchmark", in *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 159, pp. 296-307. [Online] Available at: doi.org/10.1016/j.isprsjprs.2019.11.023 [Accessed 12 October 2022].

Majchrowska, S., Mikołajczyk, A., Ferlin, M., Klawikowska, Z., Plantykowski, M. A., Kwasigroch, A. and Majek, K. (2022), "Deep learning-based waste detection in natural and urban environments", in *Waste Management*, vol. 138, pp. 274-284. [Online] Available at: doi.org/10.1016/j.wasman.2021.12.001 [Accessed 12 October 2022].

McCarthy, J. (2007), *What is Artificial Intelligence?*, Stanford University, Stanford, pp. 1-15. [Online] Available at: jmc.stanford.edu/articles/whatisai/whatisai.pdf [Accessed 30 September 2022].

Miceli, A., Morandotti, M. and Parrinello, S. (2020), "3D survey and semantic analysis for the documentation of built heritage – The case study of Palazzo Centrale of Pavia University", in *Vitruvio | International Journal of Architecture Technology and Sustainability*, vol. 5, n. 1, pp. 65-80. [Online] Available at: doi.org/10.4995/vitruvio-ijats.2020.13634 [Accessed 30 September 2022].

Mishra, M. (2021), "Machine learning techniques for structural health monitoring of heritage buildings – A state-of-the-art review and case studies", in *Journal of Cultural Heritage*, vol. 47, pp. 227-245. [Online] Available at: doi.org/10.1016/j.culher.2020.09.005 [Accessed 30 September 2022].

Monna, F., Rolland, T., Denaire, A., Navarro, N., Granjon, L., Barbé, R. and Chateau-Smith, C. (2021), "Deep

learning to detect built cultural heritage from satellite imagery – Spatial distribution and size of vernacular houses in Sumba, Indonesia", in *Journal of Cultural Heritage*, vol. 52, pp. 171-183. [Online] Available at: doi.org/10.1016/j.culher.2021.10.004 [Accessed 12 October 2022].

Montella, M. (2009), *Valore e valorizzazione del patrimonio culturale storico*, Mondadori Electa, Milano.

Parrinello, S. (ed.) (2019), *3D Bethlehem – Gestione e controllo della crescita urbana per lo sviluppo del patrimonio e miglioramento della vita nella città di Betlemme | Management and control of urban growth for the development of heritage and improvement of life in the city of Bethlehem*, vol. 3, Edifir Edizioni, Firenze.

Parrinello, S. and Picchio, F. (2019), "Integration and comparison of close-range SfM methodologies for the analysis and the development of the historical city center of Bethlehem", in *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science*, vol. XLII-2/W9, pp. 589-595. [Online] Available at: doi.org/10.5194/isprs-archives-XLII-2-W9-589-2019 [Accessed 30 September 2022].

Picchio, F. (2020), "Acquisition protocols for UAV photogrammetric data – Comparison in methodological SfM procedures from architectural till urban scale", in Barba, S., Dell'Amico, A., Limongello, M. and Parrinello, S. (eds), *D-SITE – Drones-Systems of Information on cultural hEritage – For a spatial and social investigation*, Pavia University Press, Pavia, pp. 70-79.

Picchio, F. (2019), "La fotogrammetria close range per la costituzione di modelli SfM affidabili sul sistema urbano", in Parrinello, S. (ed.), *3D Bethlehem – Gestione e controllo della crescita urbana per lo sviluppo del patrimonio e miglioramento della vita nella città di Betlemme | Management and control of urban growth for the development of heritage and improvement of life in the city of Bethlehem*, vol. 1, Edifir Edizioni, Firenze, pp. 119-138.

Redmon, J., Divvala, S., Girshick, R. and Farhadi, A. (2016), "You Only Look Once – Unified, Real-Time Object Detection", in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR) – Las Vegas, NV, USA, June 27-30, 2016*, IEEE Computer Society, pp. 779-788. [Online] Available at: cv-foundation.org/openaccess/content_cvpr_2016/papers/Redmon_You_Only_Look_CVPR_2016_paper.pdf [Accessed 30 September 2022].

Rivera, J. D. D. S. (2020), "Object detection with a model trained in Google Cloud AutoML", in *Practical TensorFlow.js*, Apress, Berkeley, pp. 164-184. [Online] Available at: doi.org/10.1007/978-1-4842-6273-3_7 [Accessed 30 September 2022].

Wang, X., Wang, H., Zhang, C., He, Q. and Huo, L. (2022), "A Sample Balance-Based Regression Module for Object Detection in Construction Sites", in *Applied Science*, vol. 12, n. 13, article 6752, pp. 1-15. [Online] Available at: doi.org/10.3390/app12136752 [Accessed 12 October 2022].

Xiao, Y., Tian, Z., Yu, J., Zhang, Y., Liu, S., Du, S. and Lan, S. (2020), "A review of object detection based on deep learning", in *Multimedia Tools Application*, vol. 79, pp. 23729-23791. [Online] Available at: doi.org/10.1007/s11042-020-08976-6 [Accessed 12 October 2022].

Zou, Z., Zhao, X., Zhao, P., Qi, F. and Wang, N. (2019), "CNN-based statistics and location estimation of missing components in routine inspection of historic buildings", in *Journal of Cultural Heritage*, vol. 38, pp. 221-230. [Online] Available at: doi.org/10.1016/j.culher.2019.02.002 [Accessed 12 October 2022].