



UNIVERSITÀ DEGLI STUDI DI SALERNO

Dipartimento di Informatica

Corso di Laurea Triennale in Informatica

TESI DI LAUREA

Tecniche di Prevenzione di Attacchi Terroristici: Un Confronto Empirico tra le Tecniche Esistenti

RELATORE

Prof. Fabio Palomba

Università degli studi di Salerno

CANDIDATO

Roberto Esposito

Matricola: 0512106041

Anno Accademico 2020-2021

"Le persone lavorano meglio quando sanno qual è l'obiettivo e perché. È importante che la gente non veda l'ora di venire al lavoro la mattina e si diverta a lavorare"

-Elon Musk

Sommario

Oggigiorno le forze dell'ordine ricoprono un ruolo importante nella sicurezza delle varie città, le quali sono sempre sotto la costante minaccia dei terroristi. Le autorità, per rendere le strade, che percorriamo quotidianamente, più sicure, individuano le possibili vulnerabilità dei vari edifici, mediante simulazioni di attacchi terroristici. Lo scopo di questa tesi è quello di analizzare e confrontare le soluzioni esistenti evidenziando i vantaggi e gli svantaggi di ciascuna di esse. Si tende a sottolineare, in modo particolare, la problematica relativa alla creazione dei dataset contenenti scenari terroristici, i quali vengono generati in maniera artificiale attraverso tecniche di gamification a causa della scarsità di tali informazioni. Ci si focalizza anche sulla difficoltà nel mantenere la posizione degli edifici fissa all'interno dello spazio di rappresentazione che si può riscontrare nel caso in cui ci si approccia al problema mediante l'utilizzo della Generative Adversarial Network e delle sue varianti. In futuro si desidera sviluppare uno strumento che permetta alle autorità locali di simulare attentati terroristici. A partire da una semplice immagine, relativa ad esempio ad una piazza, lo strumento realizzato aggiunge autonomamente vari elementi come armi, veicoli, ostaggi e forze dell'ordine in modo da ottenere uno scenario di carattere terroristico, che successivamente le autorità locali possono analizzare. L'esigenza dello sviluppo di uno strumento di questo tipo nasce dal fatto che attualmente non esistono delle tecniche di prevenzione che abbiano un'elevata efficacia e facilità d'uso contro eventuali attacchi terroristici.

Indice	ii
Elenco delle figure	iv
Elenco delle tabelle	vi
1 Introduzione	1
1.1 Contesto applicativo	2
1.2 Obiettivi della tesi	4
1.3 Metodologia e risultati	5
1.4 Struttura della tesi	6
2 Analisi e confronto della letteratura su riproduzione di attacchi terroristici	7
2.1 Analisi di tecniche già esistenti	7
2.1.1 PRA - Probabilistic Risk Analysis	8
2.1.2 Agent-Based Simulation	14
2.1.3 GIS e Random Forest - Geographic Information System	18
2.1.4 ALTER - Adversial Learning for counTerrorism	24
2.1.5 Temporal Meta-Graph	31
2.2 Confronto fra le tecniche analizzate	35
3 Sviluppo di nuove tecniche per la prevenzione di attacchi terroristici	39
3.1 Analisi di tecniche sviluppate	39
3.1.1 Approccio mediante la rete generativa avversaria StyleGAN2	40

3.1.2	Approccio mediante la rete generativa avversaria BigGAN	43
3.1.3	Approccio mediante algoritmi genetici	44
3.2	Confronto fra le varie tecniche	46
3.2.1	Performance	47
3.2.2	Capacità di cross-context	48
3.3	Analisi dei risultati	48
4	Osservazioni e limitazioni	50
5	Conclusioni	52
	Ringraziamenti	55

Elenco delle figure

1.1	Sulla sinistra sono presenti immagini relative a spazi urbani di GTA V, al centro sono presenti rappresentazioni delle strutture degli edifici di Malaga e a destra sono presenti immagini prese dal mondo reale relative alla centro città di Malaga	4
2.1	Albero di probabilità	11
2.2	Albero di decisione	12
2.3	Rete Bayesiana associata all'attacco di Sarin	13
2.4	Variabili nel modello di simulazione	15
2.5	Formalizzazione delle strategie relative al movimento dei guardiani	17
2.6	La figura mostra come utilizzare il modello della Random Forest per la simulazione. Vengono introdotte differenti caratteristiche utilizzate dal classificatore durante la predizione.	20
2.7	Nella figura sulla sinistra vengono mostrati in rosso tutti i luoghi dove si sono verificati numerosi attacchi, mentre nella figura a destra sono mostrate le zone ad alto rischio di attacchi.	22
2.8	Nella figura viene rappresentata la frequenza con la quale si sono verificati attacchi terroristici in ogni Stato dal 1970 al 2016.	23
2.9	Esempi di scenari terroristici nel videogioco Grand Theft Auto V.	25
2.10	Sulla sinistra viene mostrato uno screenshot catturato dal videogioco GTA V; sulla destra è presente uno scenario di vita quotidiana nella città di New York.	25
2.11	Volti di persone generati grazie alla Stylegan.	27
2.12	Differenze fra l'architettura di una normale GAN e la Stylegan.	28

2.13 Esempi di immagini scattate dall'alto.	30
2.14 Esempi di immagini contenenti solo la struttura degli edifici.	30
2.15 Rappresentazione grafica del risultato ottenuto dall'operazione di processing dei dati.	33
2.16 Esempio grafico della trasformazione dei meta-grafi temporali.	34
2.17 Evoluzione temporale della centralità della dimensione relativa agli obiettivi per i casi dell'Afghanistan e per quelli dell'Iraq.	35
3.1 Immagini catturate da uno scenario simulato nel videogioco Grand Theft Auto.	42
3.2 Immagini generate da un Big Generative Adversial Network.	44
3.3 Struttura base di un algoritmo genetico	46
3.4 Confronto fra la Stylegan e la Stylegan2.	48

Elenco delle tabelle

CAPITOLO 1

Introduzione

"Col passare degli anni, la maggioranza degli americani è stata in grado di tornare a vivere una vita normale, come prima dell'undici settembre. Ma io no. Ogni mattina ho ricevuto i briefing sulle minacce alla nostra nazione. E ho giurato che avrei fatto tutto quanto in mio potere per mantenerci al sicuro."

-George W. Bush

L'11 settembre 2001 è una data molto importante in quanto si è verificato uno degli avvenimenti storici più drammatici di questo secolo: l'attacco terroristico alle Torri Gemelle di New York. L'evento si è verificato durante la mattina quando due aerei di linea dirottati da un comando terrorista di Al Qaeda si schiantarono sui grattacieli del World Trade Center, causandone così il crollo.

Questo attentato ha avuto un grande significato simbolico visto che abbattere le torri newyorkesi, simbolo del commercio e dell'intera economia occidentale, significava infliggere un colpo non solo all'America, ma all'intero mondo dell'Occidente.

Vi furono all'incirca 3,000 vittime innocenti, 342 pompieri deceduti e 6,000 feriti, senza tenere conto dei gravi danni psicologici che hanno subito alcuni dei sopravvissuti.

Episodi di questo genere hanno visto protagonisti non solo l'America, ma anche l'Italia. Basti ricordare l'attacco a Nassiriya del 12 novembre 2003 quando, nel sud dell'Iraq, persero la vita circa 30 militari italiani, i quali si trovavano all'interno della base militare creata per l'operazione Antica Babilonia. Lo scopo della missione era quello di addestrare la polizia locale, mantenere l'ordine pubblico e gestire gli aiuti.

Malgrado gli avvenimenti appena descritti si siano verificati circa due decenni fa, negli ultimi anni si continua a sentire parlare nei notiziari di attacchi terroristici, nonostante siano state sviluppate da parte delle forze dell'ordine una serie di misure di sicurezza per salvaguardare la vita di intere popolazioni.

Sono stati numerosi gli episodi di carattere terroristico, basti pensare che tra il 2018 e il 2019 sono stati arrestati, solamente in Europa, circa 1000 individui come sospetti terroristi.

Durante questi episodi, i luoghi che vengono presi di mira dai terroristi non sono solo quelli pubblici, come ad esempio le strutture governative oppure le piazze e le strade particolarmente trafficate, ma si possono verificare anche in prossimità di edifici affollati, come ad esempio i centri commerciali.

Se si prendono in considerazione grandi città come ad esempio Milano, Londra o New York è evidente come sia difficile essere in grado di prevenire un attacco terroristico poiché bisogna considerare tutte le possibili forme sotto le quali esso si può presentare (utilizzo di armi da fuoco, armi chimiche, esplosione di uno o più edifici, utilizzo di veicoli per investire le persone) e tutti i possibili edifici nei quali si può verificare.

Tale problema diventa ancora più difficile da gestire in città meno popolate dove, seppure la probabilità che si verifichi un attentato risulti essere minore rispetto ad una città altamente abitata, mancano infrastrutture e strumenti tecnologici da utilizzare a proprio vantaggio in situazioni di pericolo per poter sviluppare eventuali piani di contingenza.

1.1 Contesto applicativo

A seguito degli attentati di Parigi del 13 novembre 2015 al Bataclan, vi sono stati numerosi sforzi da parte delle agenzie governative per riuscire a stimare quali potrebbero essere i possibili danni causati dagli attacchi, alle persone e agli edifici.

A tal proposito le autorità locali hanno deciso di aumentare le misure di sicurezza nei luoghi che potrebbero entrare nel mirino dei terroristi, effettuando una serie di simulazioni per poter individuare, qualora ci fossero, i punti deboli, che possono essere sfruttati per dare vita ad un attacco.

Per raggiungere questo scopo le forze dell'ordine hanno deciso di investire nello sviluppo e nell'uso di strumenti che permettano loro di simulare scenari di carattere terroristico e, di conseguenza, creare una serie di piani di emergenza da poter attuare nel caso in cui lo scenario da loro simulato dovesse verificarsi.

Tali piani, sviluppati dalle agenzie governative, hanno lo scopo di riconoscere subito una

situazione nella quale si sta per iniziare un attacco terroristico e interromperla immediatamente. Qualora ciò non fosse possibile l’obiettivo è quello di ridurre al minimo il numero di persone che potrebbero essere coinvolte in questi eventi, salvando la loro vita.

A fronte di ciò è evidente che strumenti, i quali permettono la simulazione di attacchi terroristici, possono essere particolarmente utili. Ma sorge un altro problema: la difficoltà da parte delle forze dell’ordine di riuscire a sviluppare un sistema in grado di individuare tutti i possibili attentati poiché vi sono numerose variabili da tenere in considerazione, sia in città molto abitate sia in città meno abitate, come riportato in precedenza.

Un ulteriore aspetto da non sottovalutare sono i costi effettivi per riuscire ad ottenere le strutture e gli strumenti tecnologici adeguati in modo da contrastare tali attività.

Con gli anni e con lo sviluppo dell’intelligenza artificiale, in particolare, sono stati proposti, dopo una serie di studi, alcuni strumenti da offrire alle autorità locali per riuscire a scovare ed anche a predire potenziali attacchi terroristici, riducendo il più possibile i danni nei confronti degli individui e delle strutture coinvolte.

Tali tool sono risultati utili alle agenzie governative nella simulazione e predizione di attività sospette, ma ogni tecnica adottata presenta una serie di vantaggi e svantaggi che ne influenzano il risultato finale.

Fra i vari progressi che si sono raggiunti nel campo dell’intelligenza artificiale sicuramente va evidenziato quello ottenuto con la visione artificiale la cui applicazione nell’ambito della prevenzione di attacchi terroristici verrà discussa in modo approfondito successivamente.

Nell’immagine mostrata in figura è possibile osservare gli scenari adottati per il training del modello di deep learning.

L’addestramento della rete neurale richiede una grande quantità di immagini affinché il modello risulti essere addestrato correttamente. Dato che non è stato possibile reperire datasets contenti scenari di carattere terroristico allora si è ricorso mediante tecniche di gamification alla costruzione di un dataset artificiale.

Per creare questa banca dati si sono raccolte una serie di foto provenienti dal gameplay del famoso videogioco di avventura Grand Theft Auto V. Questo approccio si è mostrato adatto alla risoluzione di questo tipo di problema poiché il mondo virtuale in GTA V si avvicina molto al mondo reale soprattutto per l’architettura degli edifici e la loro distribuzione nello spazio: alcuni luoghi di questo videogioco riproducono in maniera fedele molte strade o piazze delle città, come accade ad esempio per la città di Malaga.

Inoltre, questo approccio è risultato particolarmente conveniente perché è possibile simulare,

mediante i giocatori, sparatorie e attacchi nei confronti dei cittadini e delle forze dell'ordine. Nonostante un dataset reale con immagini di attacchi terroristici fosse stato più efficace nel training della rete neurale, questa tecnica di gamification si è rivelata molto utile, in quanto le informazioni in questo campo scarseggiano, inoltre godono di una certa riservatezza poiché considerate informazioni sensibili.

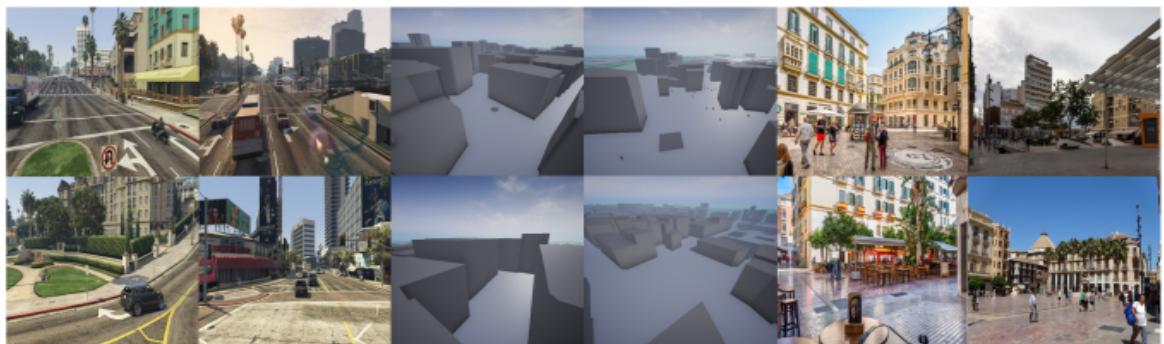


Figura 1.1: Sulla sinistra sono presenti immagini relative a spazi urbani di GTA V, al centro sono presenti rappresentazioni delle strutture degli edifici di Malaga e a destra sono presenti immagini prese dal mondo reale relative alla centro città di Malaga

1.2 Obiettivi della tesi

Dopo aver individuato il problema da risolvere relativo agli attacchi terroristici e dopo averne analizzato il contesto applicativo, sottolineando in maniera molto generale la complessità del problema a causa della grande quantità di variabili che si devono tenere in considerazione durante la simulazione di scenari terroristici, indichiamo anche l'obiettivo che si vuole raggiungere con questo elaborato.

Lo scopo della tesi è quello di riuscire ad analizzare e confrontare nel dettaglio tecniche di prevenzione già esistenti, sottolineando le caratteristiche principali in modo da mostrare gli aspetti positivi e negativi per ciascuna di esse.

Successivamente, si vuole elaborare una o più tecniche, basandosi sia sulla letteratura sia su particolari approcci che non sono stati ancora testati o utilizzati, con la possibilità di effettuare anche dei cross fra più studi, scegliendo le caratteristiche che si ritengono essere le più indispensabili.

Si effettuerà un'analisi delle performance per ciascuna di queste tecniche in quanto, come verrà mostrato in seguito, alcuni di questi approcci ricorrono all'uso di reti neurali particolar-

mente complesse.

Il problema principale delle tecniche di deep learning consiste nell'elevata potenza computazionale richiesta per poter allenare il modello adottato, rendendo così il problema da risolvere del tutto intrattabile in alcuni casi.

In conclusione, ci si propone di sviluppare una serie di tecniche per poi scegliere la migliore in termini di prestazioni che possa aiutare le forze dell'ordine nella simulazione di attacchi terroristici. L'ottimalità di un modello viene definita sulla base di vari aspetti, uno fra i più rilevanti è il tempo, poiché si desidera minimizzare la complessità computazionale, in particolar modo il training, nel caso in cui si tratti di una rete neurale.

Altro aspetto da tenere in considerazione di notevole importanza è l'efficacia del tool, dato che lo strumento proposto deve essere in grado di riuscire a generare immagini di scenari terroristici che risultino essere quanto più possibili vicini alla realtà.

Ad esempio, se si utilizzano tecniche di visione artificiale, nella creazione di una sparatoria, non dovrebbero essere generate forze dell'ordine che si sparano fra di loro oppure terroristi che sono collocati sull'acqua anziché sulla terra ferma.

Obiettivi della tesi

- Analisi delle tecniche già esistenti di simulazione di attacchi terroristici;
- Confronto fra le tecniche analizzate, sottolineando le caratteristiche principali in modo da mostrare vantaggi e svantaggi di ciascun approccio;
- Sviluppo di nuove tecniche di prevenzione di attacchi terroristici.

1.3 Metodologia e risultati

In generale il lavoro svolto segue una metodologia il cui scopo è quello di analizzare in maniera oggettiva e approfondita tutte le tecniche che sono state prese in considerazione, partendo dalla Probabilistic Risk Analysis fino ad arrivare ai Meta-Grafi Temporali. Ciascun approccio si distingue dagli altri poiché presenta aspetti unici, come ad esempio decision tree, ambiente di simulazione LEADSTO, Random Forest, Generative Adversarial Network e Meta-Grafi.

Nel terzo capitolo vengono proposte nuove tecniche e si segue sempre la stessa metodologia adottata nell’analisi della letteratura sulla riproduzione di attacchi terroristici. In questo caso vengono proposti nuovi approcci tenendo in considerazione i vantaggi e gli svantaggi delle tecniche già analizzate.

Nell’ultimo paragrafo del capitolo relativo allo sviluppo di nuove tecniche per la prevenzione di attacchi terroristici vengono analizzati i risultati ottenuti con l’utilizzo delle nuove tecniche. Fra di essi si è mostrato particolarmente innovativo quello relativo alla generazione di scenari terroristici con le Generative Adversarial Network, che fa riferimento al framework ALTER analizzato nel secondo capitolo.

1.4 Struttura della tesi

All’interno di questa sezione verrà esposta la struttura della tesi la quale è organizzata in cinque capitoli ognuno dei quali a sua volta è costituito dalle rispettive sotto sezioni.

Il primo capitolo è di introduzione e parte dal contesto applicativo, si illustra in maniera generale il campo di applicazione della tesi, indicandone il problema e successivamente si mostrano alcuni esempi. Successivamente vengono mostrati gli obiettivi dell’elaborato all’interno di questo paragrafo si specifica in maniera più dettagliata lo scopo finale della tesi. Il secondo capitolo descrive in maniera approfondita cinque tecniche volte alla simulazione e prevenzione di attacchi terroristici ognuna delle quali affronta il problema da un punto di vista differente, partendo da un approccio di tipo probabilistico fino ad uno più elaborato come quello della visione artificiale.

Sempre all’interno di questo capitolo si effettua anche un confronto fra le varie tecniche analizzate in modo tale che nel terzo capitolo, sulla base delle informazioni ottenute dal confronto, vengono proposte nuove tecniche di simulazione di attacchi terroristici.

Negli ultimi due capitoli vengono poste le osservazioni e soprattutto le limitazioni riscontrate in questo particolare ambito di applicazione e come in futuro con lo sviluppo della tecnologia e il progresso dell’intelligenza artificiale questo problema potrà essere risolto in modo più efficiente.

CAPITOLO 2

Analisi e confronto della letteratura su riproduzione di attacchi terroristici

Dopo un'attenta analisi si è riscontrato che in letteratura non sono state sviluppate tecniche particolarmente efficaci volte alla prevenzione e alla simulazione di attacchi terroristici. Pertanto le tecniche esposte in letteratura non sono state adottate come strumenti principali dalle agenzie governative.

All'interno di questo capitolo verranno analizzate alcune delle tecniche più caratteristiche, già presenti, sulla riproduzione di attacchi terroristici in modo da evidenziarne vantaggi e svantaggi, per poi, nel capitolo successivo, sviluppare nuovi modelli in grado di risolvere gli aspetti negativi riscontrati nelle ricerche precedenti.

2.1 Analisi di tecniche già esistenti

Vi sono varie tecniche che verranno prese in considerazione, ognuna delle quali utilizza approcci differenti:

- Approccio di tipo probabilistico, sviluppato da Ezell, Charles et al. [1];
- Approccio mediante un agente basato sul modello, utilizzando il linguaggio LEADSTO; tale tecnica è stata sviluppata da Bosse, Tibor et al. [2];
- Approccio attraverso tecniche di ensemble learning, quali la Random Forest, condotto da Jiang, Dong et al. [3];

- Approccio con la computer vision, combinandola con le Generative Adversarial Network, in particolare con la StyleGAN; tale studio è stato condotto da Palomba, Fabio et al. [4];
- Approccio mediante lo sviluppo di Temporal Meta-Graph; questa tecnica è stata proposta da Campedelli, Gian Maria et al. [5];

Di seguito viene effettuata un'analisi fra le varie tecniche in dettaglio per poi in seguito effettuare un confronto fra di esse.

2.1.1 PRA - Probabilistic Risk Analysis

Fra le varie tecniche proposte la prima che viene analizzata è quella relativa all'analisi del rischio probabilistico, spesso denominata come PRA (Probabilistic Risk Analysis).

Per più di 30 anni tale approccio è stato lo strumento principale utilizzato da alcuni governi, come ad esempio quello americano, e nel settore dell'industria, per la valutazione dei rischi e per la scelta di decisioni legate al risk management (gestione del rischio).

In realtà, applicazioni della PRA legate ai rischi di terrorismo sono nuove.

A tal proposito sono stati proposti più tipi di approcci di carattere probabilistico i quali utilizzano vari strumenti:

- Albero degli eventi (event tree);
- Albero dei guasti (fault tree);
- Albero di decisione (decision tree);

L'ostacolo principale riscontrato nell'analisi del rischio del terrorismo è dato dal fatto che i terroristi, a differenza della natura o dei sistemi ingegneristici, sono avversari intelligenti e quindi sono in grado di adattarsi alle nuove misure difensive che vengono adottate di volta in volta, per questo motivo Ezell ha introdotto concetti di teoria dei giochi e sistemi dinamici, seguendo il metodo mini-max.

Inoltre si vuole evidenziare come non esiste un singolo modello o approccio in grado di effettuare correttamente un'analisi di un intero scenario terroristico e individuare le decisioni da prendere dalle forze dell'ordine.

Per questo motivo, seppure vi siano degli strumenti promettenti, sarà necessario integrare anche sviluppi addizionali prima di poter utilizzare tali tool nelle applicazioni del mondo reale.

Come riportato in precedenza, un aspetto essenziale da tenere in considerazione nel momento

in cui si utilizza la Risk Analysis nell’ambito terroristico è quello relativo alla modellazione degli avversari (i terroristi).

La difficoltà riscontrata risiede nel convertire le informazioni ottenute dall’Intelligence Community (IC) in dati significanti che fungano da input nella Risk Analysis.

Per questo motivo i terroristi vengono individuati come avversari intelligenti perché si assume che ad ogni decisione nella pianificazione di un attacco, l’avversario effettuerà la scelta che massimizza i suoi obiettivi. Il comitato della NRC, il cui compito è quello di difendere gli Stati Uniti d’America da attacchi bioterroristici, ha proposto di modellare l’interazione fra i difensori e gli attaccanti come un albero di decisione, all’interno del quale le scelte dell’attaccante tentano sempre di massimizzare la propria funzione obiettivo, mentre le scelte dei difensori vengono descritte come eventi incerti.

Vi sono anche altre tre possibilità:

1. Un albero di decisione all’interno del quale le scelte dei difensori vengono modellate come decisioni che massimizzano la propria funzione di utilità e quelle degli attaccanti vengono descritte come eventi incerti che sono influenzati dalle decisioni dei difensori;
2. Un albero di decisione in cui sia le scelte degli attaccanti che le scelte dei difensori tendono a massimizzare la propria funzione utilità;
3. Un albero degli eventi che modella le scelte degli attaccanti e le risposte dei difensori come eventi incerti;

Tra i vari modelli che sono stati individuati Ezell sceglie quello proposto dal comitato della NRC. A tal proposito bisogna indicare, dal punto di vista probabilistico, come viene modellato il rischio: secondo Willis, McGill ed altri ricercatori del terrorism risk lo si può descrivere come il prodotto fra la minaccia, la vulnerabilità e le conseguenze.

Gli elementi che costituiscono il rischio sono i seguenti:

- Una minaccia, che viene definita come la probabilità di un attacco;
- La vulnerabilità, definita come la probabilità che abbia successo un attacco il quale viene effettuato;
- Le conseguenze, definite come le perdite (morti, feriti, impatti economici diretti e indiretti) che si hanno nel caso in cui l’attacco abbia successo.

Se si vuole esprimere ciò in termini matematici si ha:

$$Risk = P(A) \times P(S|A) \times C \quad (2.1.1)$$

All'interno di questa equazione la probabilità di un attacco terroristico, $P(A)$, è molto difficile da stimare poiché per conoscere tale valore bisogna avere un'approfondita conoscenza dei dati, delle motivazioni, degli intenti e delle capacità dei terroristi.

Soltamente quando gli analisti dell'Intelligence effettuano una stima del valore di $P(A)$ calcolano la credenza di ciò che un terrorista può fare sulla base delle informazioni disponibili fornite dall'Intelligence Community, dalle loro personali esperienze e giudizi.

Inoltre si vuole evidenziare che queste probabilità non sono statiche, ma, come detto in precedenza, l'avversario è intelligente ed osserva sempre le azioni intraprese dall'altro giocatore e per questo motivo le sue scelte sono sempre in continua evoluzione. Tale aspetto rende ancora più difficile la modellazione di questo problema.

Un ulteriore problema riscontrato nella PRA è legato al fatto che le probabilità associate ad eventi complessi sono spesso difficili da calcolare e per questo motivo si preferisce scomporre questi eventi in più componenti per poi determinare complessivamente la probabilità di ciascuna componente, assemblandole in seguito.

Esistono vari strumenti che vengono adottati nella decomposizione, come ad esempio gli alberi logici, influence diagram, modelli di sistemi dinamici e le reti Bayesiane.

Inoltre è possibile suddividere gli alberi logici in due categorie:

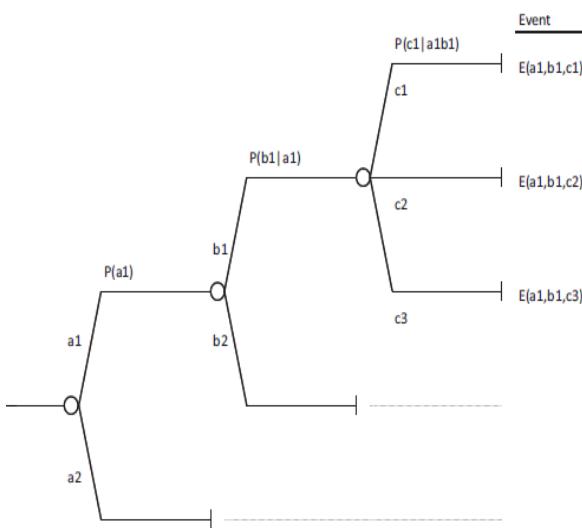
- Alberi di probabilità, eventi e decisione
- Alberi di guasti, attacco e successo

Nonostante Ezell et al. [1] abbia proposto questa serie di approcci, verranno analizzati solo alcuni di essi: la prima categoria degli alberi logici e la modellazione attraverso le reti Bayesiane.

Gli alberi di probabilità modellano una sequenza di eventi incerti con lo scopo di calcolare le probabilità degli eventi all'interno dello spazio dei risultati.

Un albero di probabilità è una successione di nodi circolare, rappresentanti variabili il cui stato è incerto, con vari rami.

I rami uscenti da ogni nodo rappresentano i diversi possibili valori delle variabili incerte associate con il nodo. Come mostrato in figura.

**Figura 2.1:** Albero di probabilità

Gli alberi degli eventi, invece, modellano le sequenze degli eventi che conducono alle conseguenze. Tale struttura non è altro che un'estensione degli alberi di probabilità aggiungendo semplicemente l'evento iniziale, i vari eventi durante l'attacco e le conseguenze che sono aggiunte per ogni cammino di probabilità.

Nella PRA questa struttura opera con lo scopo di identificare la verosimiglianza di ogni cammino di probabilità dato (dall'evento iniziale fino ad arrivare alle foglie dei vari rami).

Gli alberi degli eventi hanno avuto applicazioni in vari campi, come ad esempio negli studi di sicurezza sui reattori nucleari oppure per la valutazione del rischio di alcuni progetti della NASA come le missioni su Marte o la costruzione della Stazione Spaziale.

In realtà, si preferisce creare alberi degli eventi che siano di piccole dimensioni e compatti così da poterli descrivere in modo semplice per questo motivo sono inadeguati per la rappresentazione di eventi incerti. La PRA sviluppata dalla NASA ha 5,000 alberi degli eventi, 6,000 eventi e 2,000,000 rami; a fronte della PRA sviluppata contro gli attacchi bioterroristici con 16 eventi e 74 rami. È chiaro che la prima applicazione presenta una struttura altamente complessa e di difficile comprensione.

Infine vi sono gli alberi di decisione i quali non sono altro che un diagramma di decisione, letto da sinistra verso destra, ed il nodo più a sinistra indica la radice dell'albero e solitamente rappresenta il nodo di decisione (disegnato come un quadrato).

I rami che partono dal nodo di decisione e si diramano verso destra rappresentano l'insieme di decisioni alternative che sono disponibili. I cerchi nell'albero sono i nodi chance e rappresentano l'incertezza del risultato. Le foglie di ogni cammino attraverso l'albero vengono

chiamate endpoint. Ogni endpoint rappresenta il risultato finale del cammino dalla radice dell'albero fino a quel determinato endpoint. Di seguito viene riportato un albero di decisione, molto semplice, relativo alla Risk Analysis di un attacco bioterroristico.

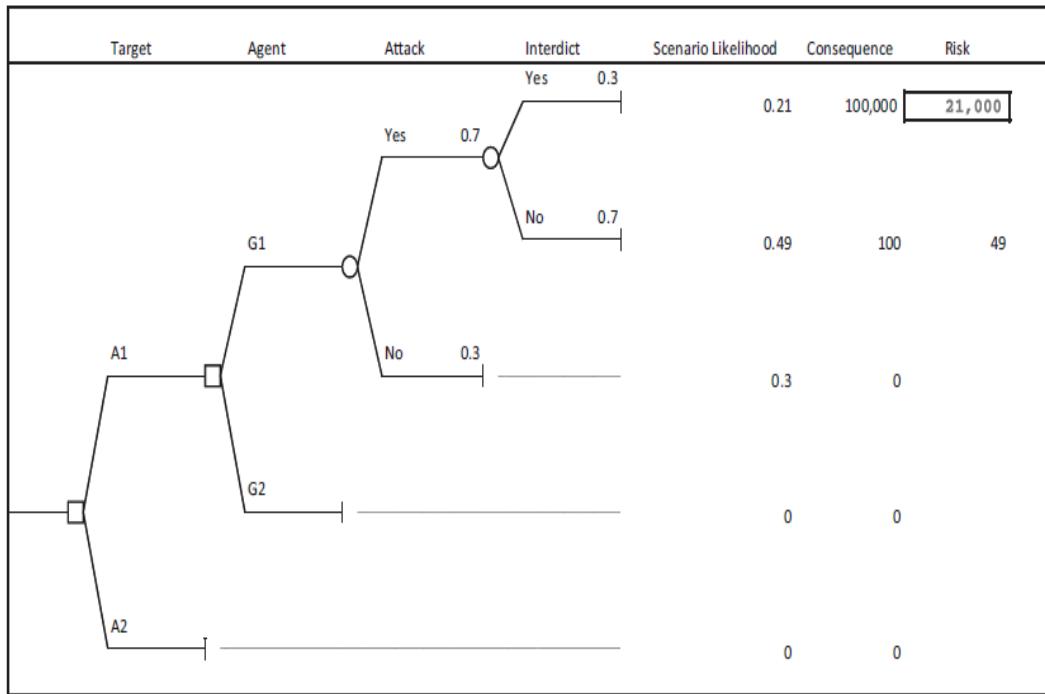


Figura 2.2: Albero di decisione

Tale strumento di supporto permette di individuare una soluzione alternativa con il miglior valore atteso. In questo specifico contesto, gli alberi di decisione possono strutturare le azioni degli attaccanti come decisioni e quelle dei difensori come nodi chance.

Gli alberi logici presentano molte limitazioni in questa applicazione. La prima di esse è data dal fatto che il valore assegnato alle probabilità si basa nella maggior parte dei casi su un giudizio soggettivo perché spesso i dati in possesso risultano essere limitati.

Un'altra limitazione, più volte sottolineata, sin dall'inizio, è che le scelte dei terroristi non sono eventi casuali e il loro comportamento non è statico poiché cambia in base alle loro preferenze.

Inoltre è stato detto che lo scopo dei terroristi è quello di massimizzare la propria funzione obiettivo, ma l'avversario può avere più di uno obiettivo. Lo scopo degli attaccanti potrebbe essere quello di massimizzare il numero di morti e/o feriti, impattare semplicemente sull'economia di un paese oppure sulla psicologia degli individui.

Ad esempio, nel 2001 vi è stato un attacco di antrace negli Stati Uniti, dove vennero uccise 4 persone e vi furono 18 infetti. Si è notato che vi è stata un'enorme diffusione di panico, sono stati chiusi edifici governativi e ci sono state molte misure preventive, più di quanto

fosse necessario. Per questo motivo per i terroristi potrebbe essere stato più importante danneggiare l'economia e la psicologia delle persone piuttosto che massimizzare il numero di uccisioni.

Un altro approccio proposto da Ezell et al.[1], sempre di carattere probabilistico, è quello relativo alle reti Bayesiane.

Una rete Bayesiana è un grafo diretto aciclico i cui nodi rappresentano le variabili casuali e gli archi diretti indicano le relazioni di dipendenza fra i vari nodi del grafo.

I genitori di un nodo non sono altro che variabili che puntano al nodo figlio. Gli antenati di un nodo corrispondono a tutti quei nodi che puntano in modo diretto o indiretto, attraverso i nodi figlio, ad altri nodi del grafo. Le reti Bayesiane sono state utilizzate nello sviluppo di modelli di antiterroristici e per la predizione della distribuzione per l'esposizione letale a particolari agenti chimici come la Sarin. Di seguito viene riportata la rete Bayesiana costruita a seguito dell'attentato alla metropolitana di Tokyo durante il quale fu rilasciato il gas nervino sarin.

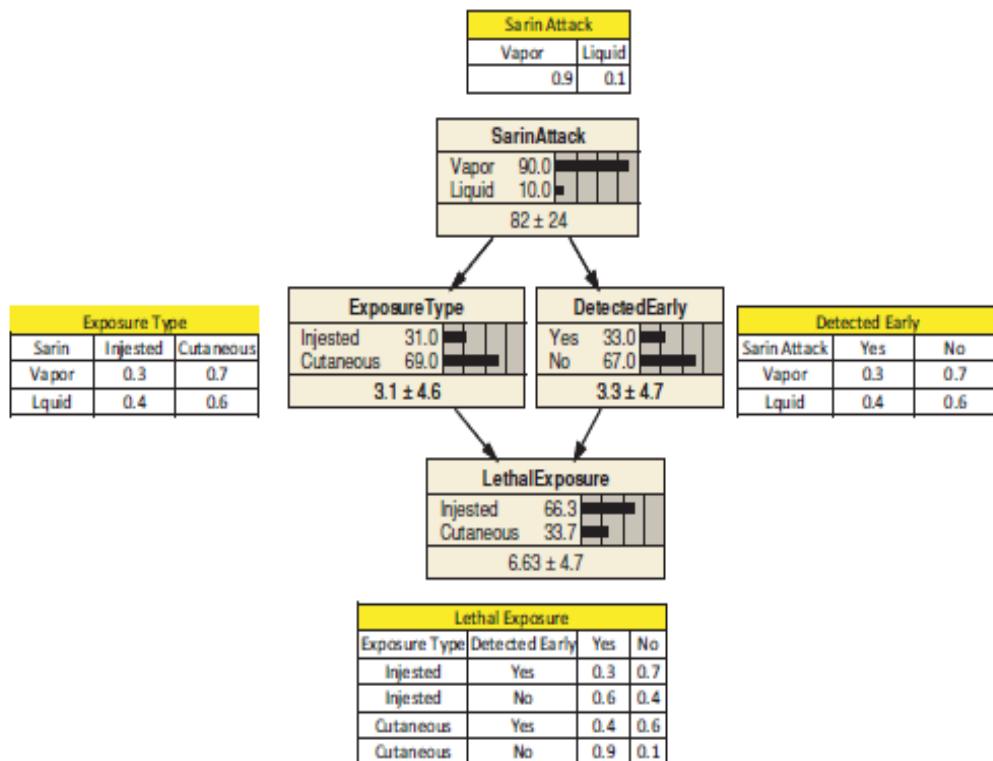


Figura 2.3: Rete Bayesiana associata all'attacco di Sarin

Nell'esempio mostrato in figura, il nodo Sarin Attack corrisponde alla radice della rete ed è il

genitore dei nodi Exposure Type e Detected Early i quali sono entrambi i genitori del nodo Lethal Exposure.

Uno dei vantaggi principali nell'utilizzo di una rete Bayesiana deriva dal fatto che eventuali modifiche ad essa sono veramente facili da apportare e aggiornare.

Dal punto di vista implementativo, soprattutto per quanto riguarda lo spazio di memoria occupato, si può migliorare la rete Bayesiana ricorrendo all'uso del cosiddetto Noisy Or (Or rumoroso), il quale permette di ridurre notevolmente le dimensioni delle tabelle CPT associate ai nodi del grafo. Si può implementare anche facilmente ricorrendo all'uso di alcuni software che permettono la creazione di reti Bayesiane ed interagire e scambiare informazioni con altri modelli.

Sia l'approccio mediante gli alberi logici sia quello mediante la rete Bayesiana risultano essere corretti, ma una soluzione più efficiente potrebbe essere quella di combinare più tecniche insieme, cercando di scomporre il problema in sottoproblemi.

2.1.2 Agent-Based Simulation

La seconda tecnica che viene analizzata è quella proposta da Bosse et al.[2].

In questo studio non si parte dal terrorismo, ma da un concetto più ampio: la criminalità. Dal momento in cui con "terroismo" ci si riferisce ad "azioni criminali violente, premeditate aventi lo scopo di suscitare terrore nella popolazione" è possibile studiare e contestualizzare tale tecnica anche all'interno di questo elaborato. Lo scopo principale di Bosse è quello di sviluppare un modello agent-based di simulazione della dislocazione del crimine che permetta di individuare i cosiddetti "hot spots" (punti caldi) dei crimini, ossia i luoghi dove vi è un elevato tasso di criminalità.

Questo modello può essere usato non solo per la simulazione delle dinamiche spazio-temporali del crimine, ma anche per l'analisi ed il controllo di queste.

Nel modello di simulazione sulla dislocazione del crimine ci sono tre tipi di agenti:

- Criminali;
- Guardiani;
- Passanti.

La scelta di questi tre agenti fa riferimento alla Teoria delle Attività Routinarie, secondo la quale i crimini si verificano quando un criminale motivato incontra un passante adatto,

mentre non c'è alcun guardiano presente.

La maggior parte dei crimini spesso si verifica nei luoghi che vengono definiti con il nome di "hot spot", come ad esempio una stazione ferroviaria oppure un centro commerciale. Questi luoghi di solito hanno numerose caratteristiche in comune, fra queste vi è la presenza di molti passanti e la mancanza di sistemi di videosorveglianza. Inoltre, con il trascorrere del tempo le situazioni spesso cambiano poiché le attività criminali si spostano in altri luoghi. Tale fenomeno probabilmente è causato dal miglioramento dei sistemi di videosorveglianza in quel luogo oppure da un incremento del numero degli ufficiali di polizia.

Un altro aspetto da tenere in considerazione è la "reputazione": un luogo dopo aver ricevuto un certo numero di assalti inizia a sviluppare una cattiva reputazione quindi i passanti iniziano a spostarsi da quel luogo.

Per descrivere i pattern nel dislocamento dei crimini è necessario conoscere il numero degli agenti per ogni tipologia (criminali, guardiani e passanti). In seguito bisogna essere in grado di individuare la densità di criminali, guardiani e passanti ed infine è necessario avere informazioni relative alla reputazione ("attractiveness") di un luogo.

I concetti appena esposti vengono formalizzati attraverso la seguente tabella che mostra tutte le variabili presenti all'interno del modello di simulazione.

Name	Explanation
c	Total number of criminals
g	Total number of guardians
p	Total number of passers by
$c(L, t)$	Density of criminals at location L at time t.
$g(L, t)$	Density of guardians at location L at time t.
$p(L, t)$	Density of passers-by at location L at time t.
$\beta(L, a, t)$	Attractiveness of location L at time t for type a agents: c (criminals), p (passers-by), or g (guardians)
$ba(L, a, t)$	Basic attractiveness of location L at time t for type a agents: c (criminals), p (passers-by), or g (guardians)
$assault_rate(L, t)$	Number of assaults taking place at location L per time unit.

Figura 2.4: Variabili nel modello di simulazione

Il calcolo del numero di agenti nei vari luoghi viene effettuato determinando il movimento di essi in funzione della reputazione del luogo. Ad esempio se si vuole calcolare la densità di criminali in un luogo L ad un istante di tempo $t + \Delta t$ allora la formula è la seguente:

$$c(L, t + \Delta t) = c(L, t) + \eta * (\beta(L, c, t) * c - c(L, t)) \Delta t \quad (2.1.2)$$

In altri termini, la densità $c(L, t + \Delta t)$ dei criminali nel luogo L all'istante $t + \Delta t$ è pari alla densità dei criminali al medesimo luogo all'istante t più una costante η la quale esprime la velocità con la quale i criminali si spostano per ogni unità di tempo.

Nel caso in cui, invece, si volesse calcolare la densità relativa ai passanti, la formula risulta essere pressoché identica, ottenendo:

$$p(L, t + \Delta t) = p(L, t) + \eta * (\beta(L, p, t) * p - p(L, t))\Delta t \quad (2.1.3)$$

Tale formula non può essere applicata per determinare anche la densità dei guardiani poiché quest'ultima è dinamica.

Per individuare l'"attractiveness" di un luogo è necessario ricorrere all'utilizzo di due combinazioni lineari di densità. Di seguito vengono mostrate:

$$\beta(L, c, t) = \beta_{c1} * (1 - g(L, t)/g) + \beta_{c2} * p(L, t)/p + \beta_{c3} * ba(L, c, t) \quad (2.1.4)$$

$$\beta(L, p, t) = \beta_{p1} * (1 - c(L, t)/c) + \beta_{p2} * g(L, t)/g + \beta_{p3} * ba(L, p, t) \quad (2.1.5)$$

In altri termini, si può notare come i criminali si allontanino dai guardiani, ma sono attratti dai passanti. Allo stesso modo i passanti sono allontanati dai criminali, mentre attratti dai guardiani.

Per sviluppare un modello di prevenzione corretto bisogna studiare una serie di strategie per le azioni intraprese dai guardiani, le quali si distinguono in strategie reattive e di anticipo.

In totale, sono state sviluppate otto differenti strategie

- La prima strategia è la "Baseline strategy". In questo caso i guardiani non effettuano alcuno spostamento e la loro densità nei differenti luoghi è la medesima in tutti gli istanti di tempo;
- La seconda strategia è di tipo reattivo e prende il nome di "Reactive 1". In questo caso la quantità di guardiani che si muove verso un nuovo luogo è proporzionale alla densità dei criminali presenti in questo nuovo luogo;
- La terza strategia, denominata "Reactive 2", afferma che il numero di guardiani che si spostano verso un nuovo luogo è proporzionale alla percentuale di attacchi che vi sono stati recentemente;
- La quarta strategia prende il nome di "Reactive 3" ed è simile alla terza con l'unica differenza che, invece di calcolare la percentuale di attacchi nell'ultimo periodo, si fa riferimento a tutti gli attacchi che vi sono stati;

- La quinta strategia è l'ultima di tipo reattivo ("Reactive 4") e anziché far riferimento alla densità di guardiani è proporzionale alla densità dei passanti, perciò il numero di guardiani che si dirige verso un nuovo luogo cresce all'aumentare della densità dei passanti in quel punto;
- La sesta strategia viene definita "di anticipo" ("Anticipate 1"), la quantità di guardiani che si sposta verso una nuova posizione è proporzionale alla densità di criminali che in futuro ci si aspetta di trovare in quel luogo;
- La settima strategia, denominata ("Anticipate 2"), differisce dalla precedente in quanto non si fa più riferimento alla densità dei criminali, bensì a quella dei passanti che ci si aspetta di trovare presso quella posizione in futuro;
- L'ottava strategia prende il nome di "Anticipate 3". In questo caso il numero di guardiani è direttamente proporzionale al numero di assalti che si potrebbero verificare in una specifica posizione nel futuro. Tale valore è calcolato tramite una media delle densità attese dei criminali e dei passanti, in altri termini non è altro che una combinazione lineare di esse, pertanto questo numero si evince dall'unione della sesta e della settima strategia.

Di seguito viene riportata la tabella che formalizza quanto appena detto, sviluppando così le equazioni per ciascuna strategia descritta.

Strategy	Formalisation
baseline	$g(L, t + \Delta t) = g(L, t)$
reactive 1	$g(L, t + \Delta t) = g(L, t) + \eta \cdot ((c(L, t)/c) \cdot g - g(L, t)) \Delta t$
reactive 2	$g(L, t + \Delta t) = g(L, t) + \eta \cdot (aar(L, t) \cdot g - g(L, t)) \Delta t$
reactive 3	$g(L, t + \Delta t) = g(L, t) + \eta \cdot (taar(L, t) \cdot g - g(L, t)) \Delta t$
reactive 4	$g(L, t + \Delta t) = g(L, t) + \eta \cdot ((p(L, t)/p) \cdot g - g(L, t)) \Delta t$
anticipate 1	$g(L, t + \Delta t) = g(L, t) + \eta \cdot (c(L, t) + \eta_2 \cdot (\beta(L, c, t) \cdot c - c(L, t)) \cdot \Delta t) / c \cdot g - g(L, t) \Delta t$
anticipate 2	$g(L, t + \Delta t) = g(L, t) + \eta \cdot (p(L, t) + \eta_2 \cdot (\beta(L, p, t) \cdot p - p(L, t)) \cdot \Delta t) / p \cdot g - g(L, t) \Delta t$
anticipate 3	$g(L, t + \Delta t) = g(L, t) + \eta \cdot ((c(L, t) + \eta_2 \cdot (\beta(L, c, t) \cdot c - c(L, t)) \cdot \Delta t) / c + (p(L, t) + \eta_2 \cdot (\beta(L, p, t) \cdot p - p(L, t)) \cdot \Delta t) / p) / 2 \cdot g - g(L, t) \Delta t$

Figura 2.5: Formalizzazione delle strategie relative al movimento dei guardiani

LEADSTO - a Language and Environment for Analysis of Dynamics by SimulaTiOn

Fino a questo istante il modello di simulazione è stato presentato da un punto di vista puramente matematico.

L'implementazione viene effettuata attraverso il software LEADSTO, il quale è un linguaggio

e un ambiente di sviluppo per l'analisi della simulazione dinamica.

Nella simulazione di processi dinamici è possibile distinguere due approcci:

- approccio logic-oriented,
- approccio matematico.

La prima tecnica viene utilizzata spesso quando si vogliono esprimere relazioni di tipo qualitativo, ma risulta essere meno adatta quando vi sono relazioni di tipo quantitativo.

Nel momento in cui il modello sviluppato presenta relazioni quantitative si ricorre ad un approccio matematico.

Lo scopo del linguaggio LEADSTO è quello di combinare le specifiche sia delle relazioni quantitative che qualitative, pertanto risulta essere una soluzione particolarmente efficace su modelli di simulazione dinamici.

Ad esempio è possibile modellare equazioni differenziali e combinarle con approcci di modellazione qualitativa, ricordando che, trattandosi di un modello dinamico, vi è un'evoluzione degli stati nel corso del tempo.

2.1.3 GIS e Random Forest - Geographic Information System

La prossima tecnica che si analizza è quella proposta da Hao et al.[3] la quale nasce con l'intento di prevenire attacchi terroristici solamente nella penisola dell'Indocina.

L'approccio stabilito da Hao ricorre all'utilizzo del sistema informativo geografico, spesso abbreviato come GIS (Geographic Information System) e della Random Forest.

Il primo strumento permette di effettuare una serie di operazioni (acquisizione, registrazione, analisi, etc.) su informazioni derivanti da dati geografici spesso definiti come "geo-riferiti". Il software utilizzato in questo studio è ArcGIS e permette di creare e utilizzare mappe sulla base delle informazioni che vengono fornite.

Il secondo strumento non è altro che un metodo di machine learning che permette di predire il rischio potenziale di attacchi terroristici nella penisola dell'Indocina. Il metodo della Random Forest ha mostrato delle buone performance con valori AUC di 0.839. L'AUC lo si può interpretare come la probabilità che il modello di machine learning classifichi un esempio positivo casuale più alto di un esempio negativo casuale. L'AUC ha un valore compreso tra 0 e 1. Di conseguenza se le previsioni fornite da un modello sono errate al 100% allora l'AUC avrà valore di 0.0; contrariamente se le previsioni sono corrette al 100% il modello avrà un AUC di 1.0.

In altri termini l'AUC corrisponde all'area sotto la curva ROC, la quale mostra graficamente le prestazioni di un modello di classificazione e permette di tracciare:

- Tasso di veri positivi;
- Tasso di falsi positivi.

La combinazione del GIS con la Random Forest porta con sé un grande potenziale nella simulazione degli attacchi terroristici.

Lo scopo dello studio era quello di capire l'evoluzione spazio-temporale degli attacchi terroristici nella penisola dell'Indocina e predire le zone che si trovano a rischio. Come affermato già dall'inizio individuare una soluzione al problema del terrorismo è molto complesso perché si tratta di numerosi eventi dinamici che non è possibile risolvere solo grazie ad approcci puramente matematici.

Lo studio condotto da Hao et al.[3] lo si può schematizzare attraverso una serie di steps:

1. Estrarre scenari terroristici che si sono verificati nella penisola dell'Indocina dal database del terrorismo globale, spesso definito come GTD (Global Terrorism Dataset), e usare il software ArcGIS per localizzare gli attacchi terroristici sulla mappa;
2. Utilizzare la funzione "Kernel Density" su ArcGIS e OriginLab (software per la rappresentazione grafica dei dati) e analizzare l'evoluzione degli attacchi a seconda delle variabili dello spazio e del tempo;
3. Preparare i dati geografici e individuare i corrispondenti raster data dell'attacco terroristico. Con l'espressione "raster data" si fa riferimento ad una matrice di pixel organizzata in righe e in colonne dove ogni cella contiene un valore che rappresenta un'informazione, come ad esempio la temperatura. Solitamente si utilizza questa struttura dati nel caso in cui le informazioni da elaborare constano di fotografie aeree, immagini prese dai satelliti oppure mappe;
4. Costruire l'algoritmo della Random Forest per predire attacchi terroristici su una scala geografica spaziale nella penisola dell'Indocina.

Di seguito viene mostrata l'architettura del sistema che viene utilizzato per la predizione di attacchi terroristici.

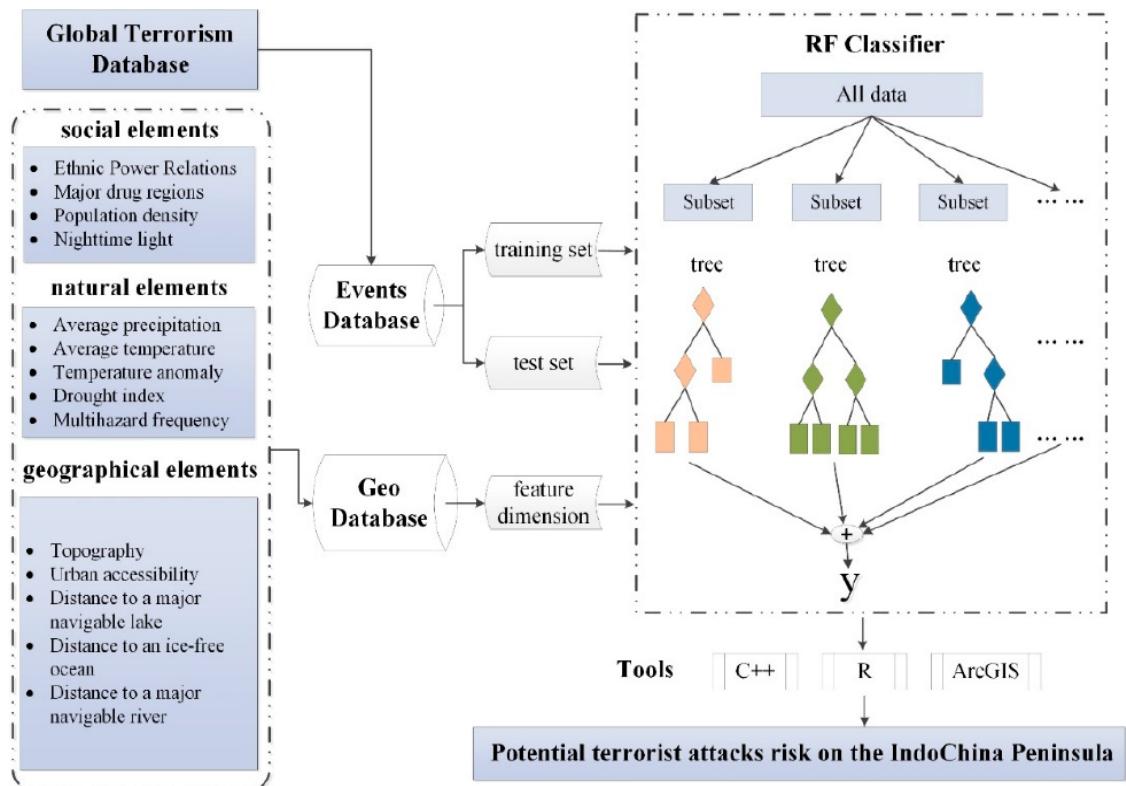


Figura 2.6: La figura mostra come utilizzare il modello della Random Forest per la simulazione. Vengono introdotte differenti caratteristiche utilizzate dal classificatore durante la predizione.

Come si può evincere dall'immagine mostrata sopra vi sono tre categorie principali di elementi:

- Elementi sociali;
- Elementi naturali;
- Elementi geografici.

Gli attacchi terroristici sono un fenomeno sociale molto complesso poiché sono guidati da numerosi fattori, partendo dagli elementi sociali, naturali e geografici. Inoltre vi possono essere anche influenze politiche e religiose.

Creazione del dataset

I dati utilizzati nella ricerca sono stati estratti dal Global Terrorist Dataset, il quale è un database open-source disponibile in rete e contiene informazioni su attentati terroristici che

si sono verificati nel mondo fra il 1970 e il 2016. Si può accedere a questo dataset al seguente indirizzo: Global Terrorist Dataset.

Il database in questione si basa su una copia della banca dati creata dalla Pinkerton Global Intelligence Service (PGIS). Ogni elemento del GTD contiene informazioni relative alla data dell'attentato e altre caratteristiche, come le armi utilizzate, gli ostaggi presi dai terroristi, il luogo e i responsabili, qualora tali informazioni fossero disponibili.

Inoltre per avere coerenza e consistenza fra i dati, le informazioni relative al dataset sono state convertite in un raster data con la stessa risoluzione spaziale dei dati geografici. Laddove vi sono stati attacchi terroristici il pixel corrispondente nel raster viene considerato un'area ad alto rischio e gli viene assegnato come valore 1, altrimenti assume valore 0.

Stima kernel di densità

La stima kernel di densità è un metodo che viene utilizzato per il riconoscimento dei pattern e per la classificazione attraverso una stima, come suggerisce il termine, di densità negli spazi metrici e permette di convertire un insieme di punti in un raster. In altri termini, per ogni x all'interno dello spazio metrico l'algoritmo permette di calcolare la probabilità di appartenere ad una classe X , considerando la densità di X in un intorno h del punto x .

La funzione proposta da Hao et al.[3] è la seguente:

$$\hat{f}_h(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - X_i}{h}\right) \quad (2.1.6)$$

dove h rappresenta l'ampiezza dell'intorno; $f_h(X)$ è la stima del kernel nel punto X con ampiezza dell'intorno h ; $x - X_i$ è la distanza fra il punto x e il punto X_i ; in conclusione, K è la funzione Kernel.

L'ampiezza dell'intorno è stata scelta con valore pari a 50km. Grazie al software ArcGIS è possibile calcolare tale funzione con lo strumento "Densità Kernel", inoltre, sempre grazie al software, è possibile distribuire geograficamente i punti degli attacchi terroristici.

Algoritmo della Random Forest

La Random Forest è stata implementata grazie all'utilizzo dell'ambiente di sviluppo di R. Questo approccio non è altro che una tecnica di ensemble learning sviluppata sulla base

di un grande insieme di alberi di decisione. Ogni albero è addestrato selezionando solo un numero casuale di variabili e di campioni dal dataset di training. Per utilizzare questa tecnica è necessario scegliere il valore di tre parametri:

- il numero di campioni di bootstrap;
- il numero di variabili campionate ad ogni divisione;
- la dimensione minima dei nodi terminali, valore al di sotto del quale le foglie non vengono più suddivise.

Inoltre per evitare che il modello sviluppato mostri un fenomeno di overfitting sul dataset si può applicare il metodo di 10-fold cross-validation che consiste nel suddividere la banca dati in campioni da training e campioni da validazione.

Dopo aver eseguito tutti gli step elencati precedentemente si possono sviluppare grazie al software ArcGIS le distribuzioni spaziali con le relative zone ad alto rischio di attentati terroristici. Di seguito viene riportato il risultato ottenuto.

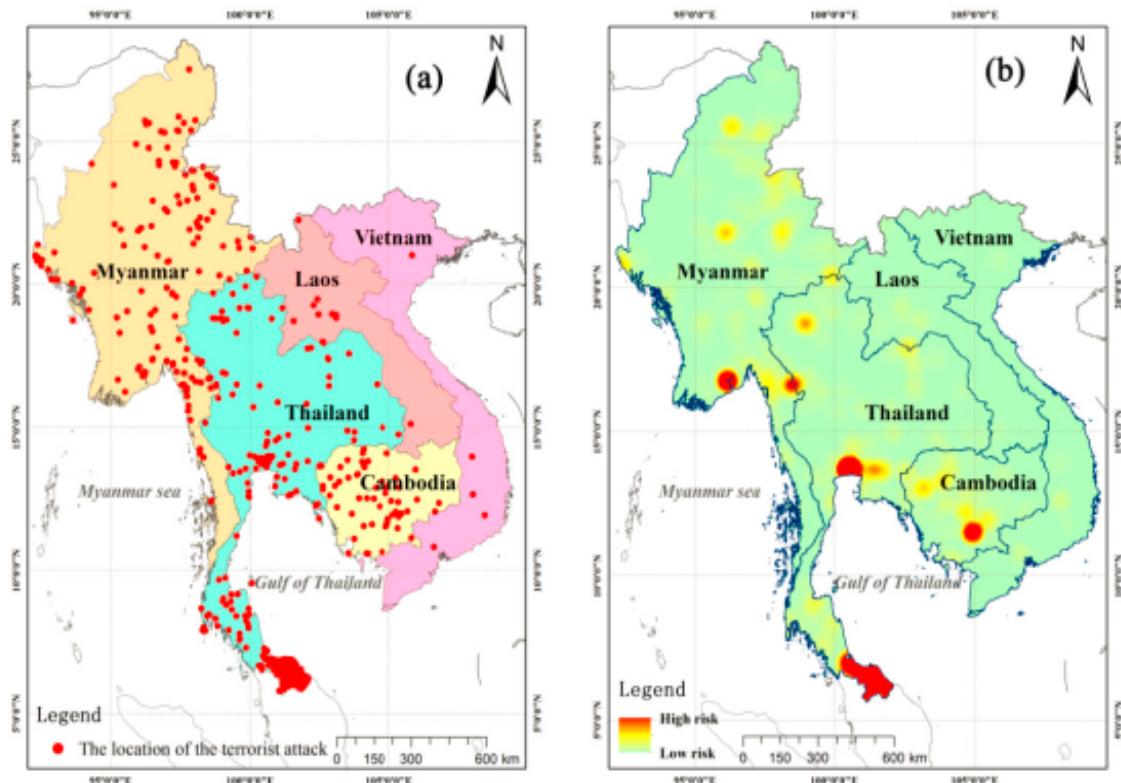


Figura 2.7: Nella figura sulla sinistra vengono mostrati in rosso tutti i luoghi dove si sono verificati numerosi attacchi, mentre nella figura a destra sono mostrate le zone ad alto rischio di attacchi.

Sono stati individuati in totale cinque hot spot distribuiti negli Stati di Cambogia, Myanmar e Tailandia. Si vuole evidenziare come la maggior parte di essi si trovino sulle frontiere.

Basandosi sui cinque Stati presenti nella penisola dell'Indocina e dalla rappresentazione grafica delle frequenze di attacchi terroristici in ciascun paese, è chiaro che lo Stato più pericoloso fra tutti, negli ultimi anni, sia la Tailandia, lo si può evincere molto facilmente dal grafico disegnato grazie al software OriginLab.

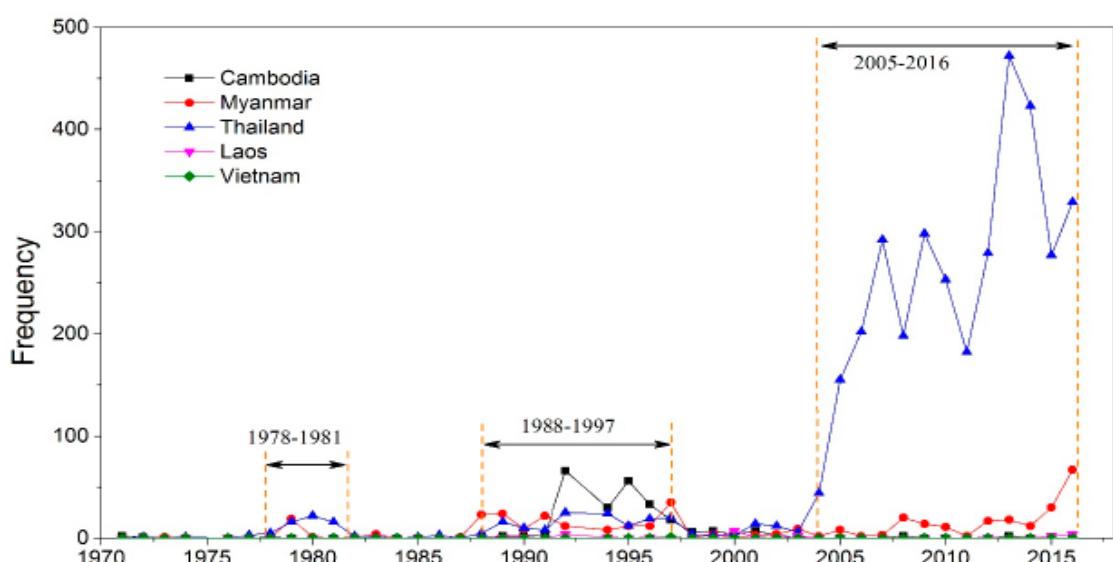


Figura 2.8: Nella figura viene rappresentata la frequenza con la quale si sono verificati attacchi terroristici in ogni Stato dal 1970 al 2016.

In conclusione si può affermare che i risultati ottenuti con questo studio si sono rivelati particolarmente efficienti e ciò è dovuto alla combinazione di tecniche di machine learning, come la Random Forest, con sistemi di geo-informazione per la simulazione della distribuzione del rischio di attacchi terroristici.

Si vuole evidenziare come questo approccio non è context-dependent e quindi non è legato solamente all'ambito del terrorismo, ma può essere utilizzato senza alcun problema laddove la variabile geo-spatiale assuma un ruolo di principale importanza.

2.1.4 ALTER - Adversial Learning for counTerrorism

L'approccio che viene analizzato in questa sezione prevede la simulazione di scenari terroristici attraverso tecniche di computer vision e deep learning. Tale studio è stato condotto da Palomba, Fabio et al.[4] con lo scopo di permettere alle forze dell'ordine di prevedere le conseguenze che si potrebbero verificare con questi attacchi, individuando anticipatamente quali zone rafforzare e come rispondere in caso di attentato.

La modellazione del problema avviene mediante tecniche di computer vision nelle quali viene addestrata una Generative Adversarial Network su immagini di attacchi terroristici.

La GAN (rete generativa avversaria) è costituita da due reti neurali, le quali vengono addestrate in maniera competitiva seguendo le caratteristiche del gioco minimax.

La prima rete neurale è una rete generativa il cui compito è quello di creare un'imitazione basandosi sulle foto iniziali. Questa rete crea, quindi, un'immagine completamente nuova che non è un duplicato di uno dei dati di partenza.

La seconda rete è il discriminatore e riceve in input sia i dati di base che le informazioni generate dalla rete generativa. Lo scopo del discriminatore è quello di verificare se i dati ricevuti siano autentici oppure falsi. Un'immagine viene classificata come falsa non solo quando si discosta in modo particolare dai dati di base, ma anche quando è troppo perfetta, evitando così che si accettino immagini che non hanno un effetto naturale.

Se si utilizza questo approccio è necessario considerare numerosi problemi poiché il trasferimento degli scenari non si limita ad "inserire" semplicemente delle persone da un'immagine ad un'altra. Bisogna garantire anche una consistenza dell'immagine, come ad esempio una persona che corre sull'acqua non è un evento che si può verificare nella realtà. La rete non dovrebbe dare origine ad immagini di questo tipo.

Lo sviluppo delle GAN ha avuto una crescita esponenziale negli ultimi anni e, fra le varie architetture esistenti, è stata selezionata la StyleGAN.

Questa rete neurale è stata sviluppata da NVIDIA e Karras, Tero et al.[6] ne descrive accuratamente ogni aspetto.

Per l'addestramento della StyleGAN è necessario un dataset di training e per la generazione di questi dati occorre un ambiente di simulazione. La maggior parte delle informazioni riguardanti gli attacchi terroristici sono sensibili e non sono quindi disponibili in grandi quantità, per ovviare a questo problema è stato creato un dataset sintetico.

Con il videogioco Grand Theft Auto V sono stati generati scenari in cui si simulava un attacco terroristico attraverso la registrazione del gameplay.

Nella figura successiva vengono mostrati alcune scenari catturati dal videogioco, nei quali vengono intraprese azioni criminali.



Figura 2.9: Esempi di scenari terroristici nel videogioco Grand Theft Auto V.

Il motivo per il quale è stato scelto GTA V è dato dal fatto che è un gioco open world con una grande accuratezza nei minimi dettagli e la rappresentazione del mondo virtuale è molto vicina a quella del mondo reale.

Ad esempio sono presenti sia componenti statiche come edifici, alberi e strade, sia componenti dinamiche come persone e macchine in movimento. Bisogna ammettere però che il numero di passanti e il numero di macchine in circolazione nelle strade è molto basso, mentre nel mondo reale è molto più elevato. Non è possibile trovare all'interno del videogioco strade affollate con migliaia di persone, cosa che nel mondo reale, nelle città densamente abitate, si verifica. Ciò è uno degli aspetti più importanti da considerare poiché uno dei fattori considerati dai terroristi è il numero di persone che potrebbero essere coinvolte nell'attentato. Di seguito viene mostrato un confronto mediante un'immagine.



Figura 2.10: Sulla sinistra viene mostrato uno screenshot catturato dal videogioco GTA V; sulla destra è presente uno scenario di vita quotidiana nella città di New York.

Dopo aver effettuato il training della rete, il passo successivo è quello di eseguirne il reverse e mappare le caratteristiche dell’immagine nelle loro variabili all’interno dello spazio latente. Per questo motivo è stata proposta la creazione di un encoder che mappa a partire dalle immagini create la loro rappresentazione all’interno dello spazio.

In altri termini la StyleGAN possiede tre componenti principali:

- l’input delle rete;
- i parametri;
- l’output della rete.

Solitamente per ottimizzare la rete neurale si fissano l’input e l’output e si scelgono i valori dei parametri mediante la backpropagation dall’input all’output. Alcuni studiosi, però, hanno proposto di mantenere fissi l’output e i parametri e di mappare le immagini ottenute in output all’interno dello spazio latente.

Per fare ciò non basta altro che generare un vettore casuale in un certo intervallo, ad esempio $N(0, 1)$, e inviarlo in input al generatore. In seguito si calcola la loss function e si cerca di minimizzarne il valore facendo un confronto fra l’immagine generata e l’immagine target. Grazie a questa comparazione si può sfruttare la backpropagation all’interno della rete per modificare il valore delle variabili latenti. Si segue questo passaggio fin quando non si minimizza la funzione di perdita.

Architettura della StyleGAN

Negli ultimi anni sono stati raggiunti importanti traguardi nell’ambito del neural style transfer poiché sono state sviluppate nuove tecniche e nuove architetture. Fra queste sicuramente va evidenziata l’adaptive instance normalization (AdaIN), un metodo di normalizzazione, che allinea la media e la varianza delle funzionalità del contenuto con quelle delle funzionalità di stile.

L’azienda Statunitense NVIDIA ha dato origine ad una nuova architettura chiamata StyleGAN, in cui si combinano le caratteristiche della AdaIN con quelle delle GANs.

In particolare, si è mostrato che le tecniche utilizzate nello style transfer possono essere applicate non solo nel dominio in cui sono nate, cioè nella creazione di dipinti, ma anche in altri domini.

Da ciò Karras et al.[6] sono riusciti a creare volti finti indistinguibili dai volti reali semplice-

mente prendendo alcune caratteristiche, come occhi, naso e bocca da un insieme di immagini reali.

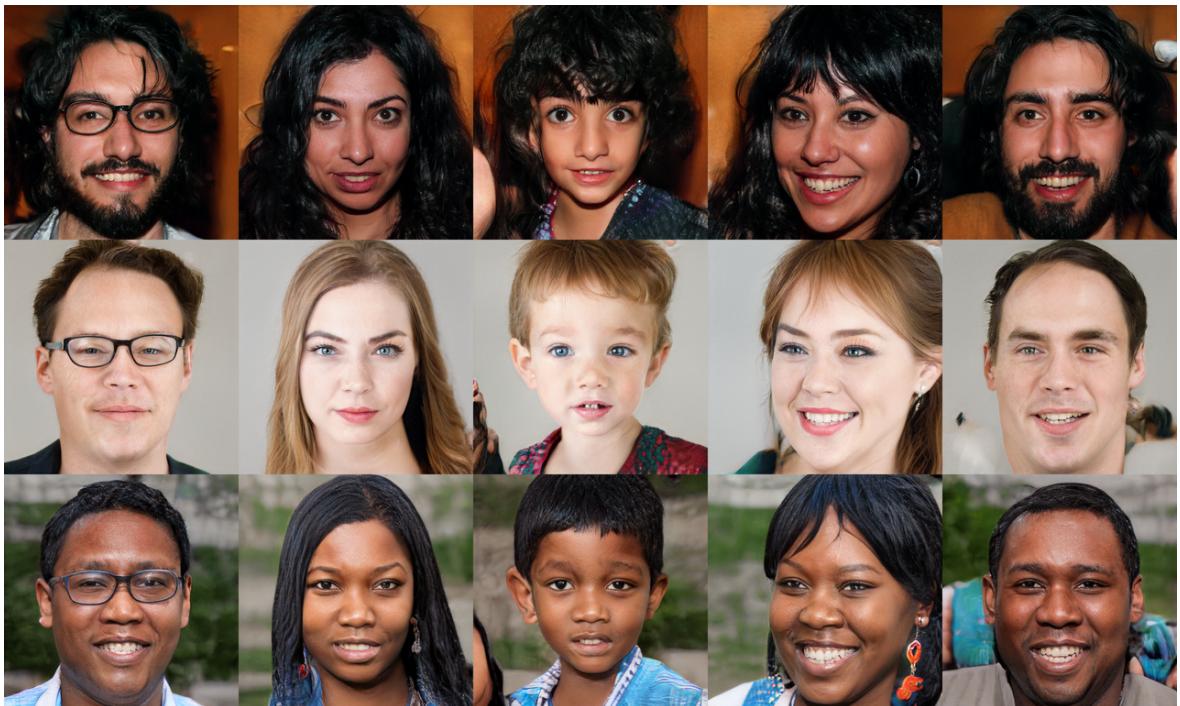


Figura 2.11: Volti di persone generati grazie alla Stylegan.

Il progresso dell'AdaIN è il motivo principale alla base dello sviluppo della rete neurale che si sta analizzando. Questa crescita progressiva ha permesso di migliorare la qualità della rete neurale. L'operazione di normalizzazione AdaIN ha permesso di controllare le funzionalità della rete neurale attraverso il codice nascosto nei differenti livelli della StyleGAN.

Ciò è stato possibile in quanto l'architettura della rete presenta tre spazi latenti anziché uno solo, tale caratteristica è unica.

I tre spazi latenti sono i seguenti: oltre al normale spazio latente Z è presente anche un nuovo spazio latente libero che viene definito spazio intermedio W ; inoltre è presente anche lo spazio esteso $W+$, il quale si ottiene dalla media di tutti i differenti livelli di W .

In una normale GAN il vettore nascosto Z è dato in pasto solo all'input layer, mentre nella StyleGAN il vettore iniziale Z è prima mappato in un vettore intermedio W grazie a una rete neurale feed-forward di otto livelli. Questo vettore W è in seguito fornito come input ad ogni livello convoluzionale nella StyleGAN.

Per comprendere meglio questo concetto da un punto di vista grafico di seguito viene mostrata una figura rappresentante le differenze fra l'architettura di una normale GAN e la StyleGAN.

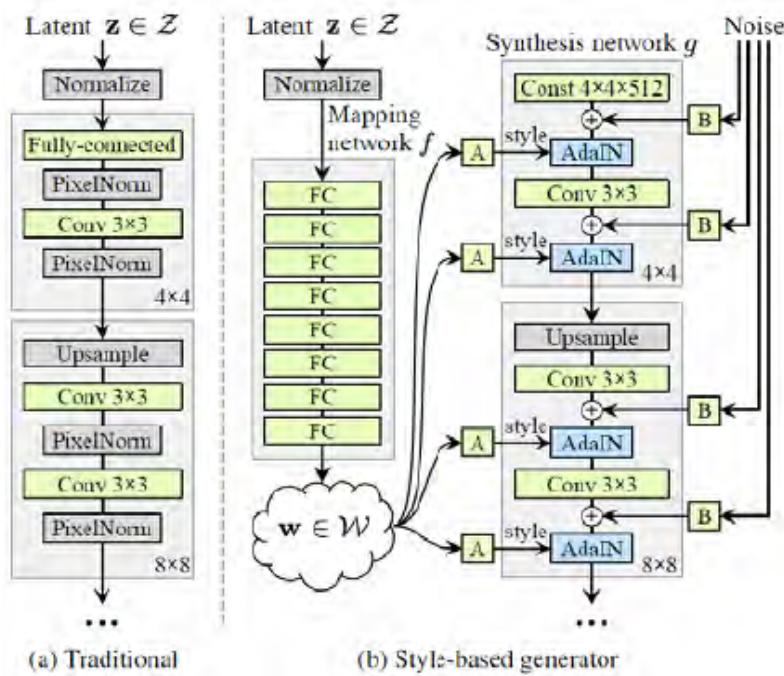


Figura 2.12: Differenze fra l’architettura di una normale GAN e la Stylegan.

L’operazione di normalizzazione è una delle caratteristiche principali di questa architettura e a tal proposito Karras et al.[6] procedono definendo l’AdaIN come:

$$\text{AdaIN}(x_i, y) = y_{s,i} \frac{x_i - \mu(x_i)}{\sigma(x_i)} + y_{b,i} \quad (2.1.7)$$

Dove $y = (y_s, y_b)$ controlla i parametri della normalizzazione, la x_i rappresenta la mappa delle funzionalità che è normalizzata separatamente e la y rappresenta le variabili scalari.

Dopo aver definito l’architettura della rete neurale e l’operazione di normalizzazione, si può passare alla definizione del processo di inversione, il quale è un problema di minimizzazione rappresentato come:

$$z^* = \min_z - E_x \log[G(z)] \quad (2.1.8)$$

Dove z appartiene a Z e corrisponde allo spazio latente, x appartiene allo spazio R^{mxm} , ovvero l’immagine target e $G(z)$ è il grafo computazionale della GAN.

Funzione obiettivo

Un altro parametro da individuare per procedere con la simulazione di scenari terroristici è la funzione obiettivo. Lo scopo è quello di minimizzare tale valore e rappresenta la qualità

dell’immagine che viene generata. L’idea adottata nello studio condotto da Palomba, Fabio et al.[4] è quella di utilizzare una combinazione lineare di più loss function.

Si possono distinguere due categorie di funzioni:

- Differenza assoluta;
- Differenza percettiva.

La prima si basa sul confronto dei pixel delle due immagini che si analizzano, la seconda, invece, si basa sulla differenza percettiva delle immagini. Per la differenza assoluta si considera la Pixel-wise loss, la quale si calcola tramite il logaritmo del coseno iperbolico della predizione dell’errore. Si preferisce questa funzione alla funzione dell’errore quadratico medio perché meno sensibile ad eventuali outlier.

Per la differenza percettiva si considerano le seguenti funzioni: VGG loss, LPSIS loss e MS-SIM loss. Senza entrare nel dettaglio di ciascuna di esse, si riporta solamente la combinazione delle quattro funzioni obiettivo:

$$\begin{aligned}
 w^* = \min_w & \lambda_{Pixel-wise} * \frac{1}{N} * \log(\cosh(G(w) - I) \\
 & + \lambda_{VGG} * L_{VGG}(G(w), I) \\
 & + \lambda_{LPSIS} * L_{LPSIS}(G(w), I) \\
 & + \lambda_{MS-SIM} * L_{MS-SIM}(G(w), I)
 \end{aligned} \tag{2.1.9}$$

Affinché la Stylegan utilizzata nella ricerca possa convergere, la funzione obiettivo del generatore deve assumere un valore prossimo a -0.5, mentre la funzione obiettivo del discriminatore deve valere circa 0.5. Durante il training del modello si è notato che la rete neurale ha avuto difficoltà nel convergere ai valori sopracitati.

In generale, l’addestramento della rete con immagini ad una risoluzione più bassa (128px) ha mostrato una progressione migliore rispetto alle immagini con una risoluzione più elevata. Entrambi i casi sembravano avvicinarsi ai valori indicati, ma ad un certo punto la curva subiva un’inversione e se ne allontanava nuovamente, generando così scarsi risultati.

Dopo aver riconfigurato molte volte i parametri della rete ci si è resi conto che le prestazioni continuavano a non essere delle migliori e grazie ad un’attenta analisi si è notato che le immagini riprese dall’alto degradavano le performance della rete e per questo motivo sono state rimosse dal dataset.

Nella figura sotto vengono mostrate alcune delle immagini che hanno provocato particolari

problemi nell'efficacia del modello di simulazione.

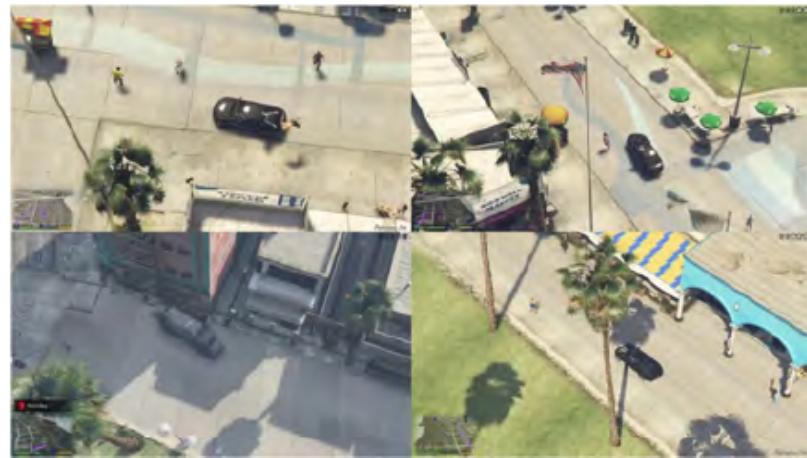


Figura 2.13: Esempi di immagini scattate dall'alto.

A seguito delle modifiche appena descritte i risultati non hanno subito un grande miglioramento e per questo motivo è stata modificata la banca dati utilizzando le immagini prese dal dataset Unreal, il quale trascura una notevole quantità di dettagli. Sicuramente ci si aspetta che le performance miglioreranno, ma non si sarà in grado di codificare immagini prese dal mondo reale, ad esempio nella seguente immagine si può notare un campione estratto dallo studio. Qui è evidente come le immagini siano molto più elementari e completamente differenti da quelle del dataset di partenza costruito artificialmente mediante il videogioco GTA V.

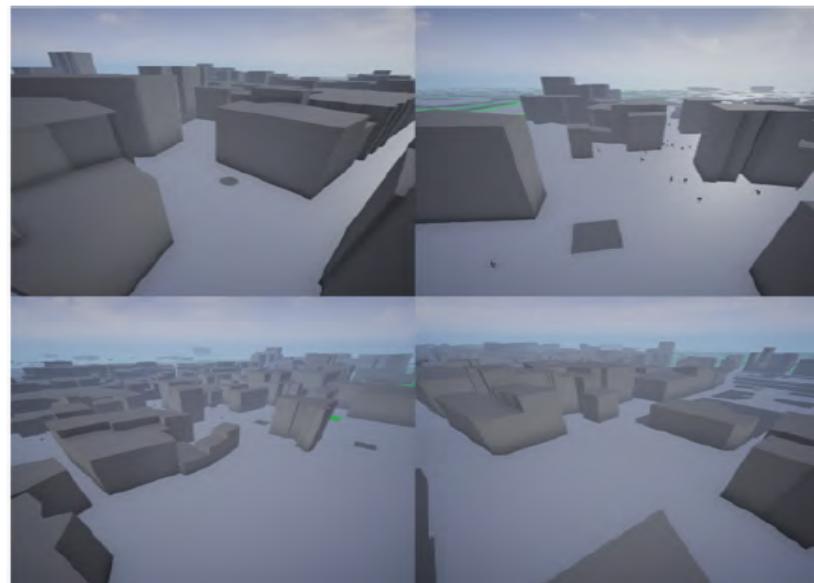


Figura 2.14: Esempi di immagini contenenti solo la struttura degli edifici.

In conclusione si è osservato che la combinazione lineare delle loss function non ha portato ad un vantaggio, bensì ad uno svantaggio.

Solo una funzione converge ed è la Pixel-wise, appartenente alla categoria differenza assoluta. Ciò si è verificato poiché la rete non è in grado di codificare correttamente tutte le caratteristiche che costituiscono l'immagine.

L'idea alla base di questa tecnica risulta essere particolarmente rivoluzionaria, è davvero molto utile ai fini della sicurezza dei cittadini del mondo. In futuro con l'avvento di nuove tecnologie tale tecnica potrà essere più facilmente applicata.

2.1.5 Temporal Meta-Graph

L'ultima tecnica che viene analizzata è quella proposta da Campedelli et al.[5] con la quale si propone l'uso dei meta-grafi temporali e del deep learning per prevenire i futuri target dei terroristi.

Mentre il terrorismo resta caratterizzato da elevati livelli di incertezza e imprevedibilità, la ricerca nell'ambito dell'intelligenza artificiale può aiutare a fornire soluzioni basate sui dati, che sono volte a contrastare questo fenomeno. In questa tecnica si sfruttano così i seguenti elementi: il potere dei dati, i quali oggigiorno sono sempre più preziosi, i potenti modelli computazionali e le teorie relative al comportamento dei terroristi.

Alla luce di ciò, lo studio punta all'unione fra l'intelligenza artificiale e la prevenzione di attacchi terroristici proponendo un nuovo framework basato sui meta-grafi, sulle serie storiche e sugli algoritmi di previsione utilizzati nel campo del machine e deep learning.

I dati sui quali è stato condotto lo studio sono stati prelevati dal Global Terrorism Database, già utilizzato nello studio condotto da Kao et al.[3], ma in questo caso ci si è focalizzati sugli attacchi avvenuti in Afghanistan e Iraq tra il 2001 e il 2018.

L'evento terroristico presenta tre dimensioni:

- armi utilizzate;
- tattiche schierate;
- target individuati.

Una volta che i meta-grafi sono creati, si derivano le serie storiche mappando la centrality di ogni nodo nella dimensione corrispondente. In seguito le serie vengono utilizzate per in-

dividuare pattern ricorrenti per prevenire i prossimi target che potrebbero essere presi di mira.

Sono state proposte differenti teorie per descrivere e spiegare le azioni dei terroristi. Queste ultime è possibile dividerle in tre macrocategorie:

- psicologiche;
- organizzative;
- strategiche.

Le teorie psicologiche ambiscono a spiegare le singole cause che hanno portato i criminali a prendere parte nelle attività terroristiche. Le teorie organizzative si focalizzano sulla struttura interna e sul simbolismo formale di ciascun gruppo cercando di comprenderne il comportamento. Infine, le teorie strategiche indirizzano il processo decisionale dei terroristi e sulla base di ciò originano lo studio dei conflitti.

La letteratura ha mostrato come gli attacchi terroristici non avvengono in maniera casuale, ma vi sono dei cluster temporali dai quali si possono inferire informazioni sui prossimi attacchi, analizzando i pattern ricorrenti. Per catturare le connessioni fra gli eventi e le loro caratteristiche non è sufficiente suddividere il problema in serie storiche. Per questo motivo si è introdotto un nuovo framework che sfrutta i vantaggi delle serie storiche derivanti dai grafi.

Per prima cosa, per ogni unità di tempo vengono generati i grafi pesati, i quali rappresentano le connessioni esistenti nelle tre dimensioni di dati considerate (tattiche, armi e obiettivi). Una volta che il seguente step è stato completato, si calcola, per ogni dimensione e per ogni unità di tempo, il grado normalizzato della centrality di ogni caratteristica.

L'attività di pre-processing dei dati avviene a partire dal dataset D_{Axz} che contiene $|A|$ attacchi terroristici e $|z|$ variabili associate ad ogni attacco, corrisponde esattamente al formato originale del Global Terrorism Database.

Arrivati a questo punto, si filtrano i dati prendendo in considerazione solamente gli attacchi che si sono verificati in Afghanistan e Iraq fra il 2001 e il 2018. Si ottengono così due dataset separati: D_{txz}^{AFG} e D_{txz}^{IRA} .

I due nuovi dataset sono costituiti da $|t|$ osservazioni e $|z|$ caratteristiche. Il valore che possono assumere tali caratteristiche non è altro che il numero di volte che quella caratteristica era presente negli attacchi eseguiti in quell'unità di tempo, come ad esempio il singolo

giorno.

Come detto fin’ora, le dimensioni utilizzate in questo studio sono tre quindi è possibile individuare tre insiemi: l’insieme delle caratteristiche tattiche (X), l’insieme delle caratteristiche delle armi (W) e l’insieme delle caratteristiche degli obiettivi (Y). Se si indica con C il paese di riferimento, anziché utilizzare AFG oppure IRA, il dataset lo si può scrivere come $D^C = \{D_X, D_W, D_Y\}$.

Per ogni $\{D_X, D_W, D_Y\}$ si crea una finestra temporale U tale che $u = 2t$. In altre parole, per ogni dimensione, si collassano i dati in unità di tempo di due giorni e ciascun dato è pari alla somma del numero di ogni caratteristica nei due giorni.

Il motivo dietro alla creazione di unità di tempo basate su due giorni è duplice. In primis, fare affidamento su serie storiche di un singolo giorno aumenta il rischio di avere serie troppo scarse, con grafi molto piccoli che non produrrebbero alcuna informazione.

D’altro canto, nella realtà le risorse richiedono tempo per elaborarle e avere un sistema di previsione che opera di giorno in giorno produrrebbe informazioni che sarebbe difficile trasformare in decisioni concrete e piene di significato. Per questo motivo un’architettura con unità di tempo due giorni risulta essere un buon compromesso.

Nella seguente figura viene schematizzato graficamente la struttura che si ottiene a seguito dell’operazione di processing dei dati.

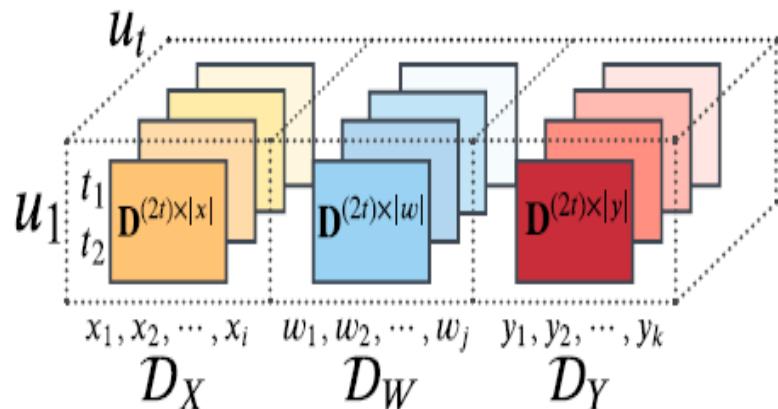


Figura 2.15: Rappresentazione grafica del risultato ottenuto dall’operazione di processing dei dati.

In quest’altro grafico si mostra un esempio, sempre grafico, dei meta-grafi temporali che vengono creati. I vari u_i rappresentano le differenti unità di tempo e poi in giallo sono presentati i grafi relativi alla prima dimensione, le tattiche; in blu sono rappresentati i grafi relativi alla seconda dimensione, le armi; infine, in rosso sono rappresentati i grafi dell’ultima

dimensione, gli obiettivi.

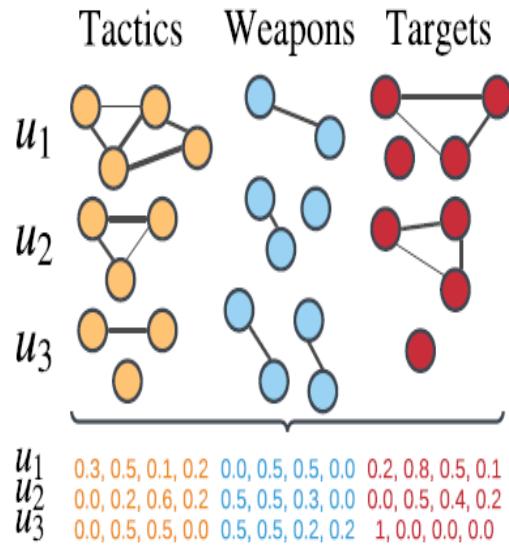


Figura 2.16: Esempio grafico della trasformazione dei meta-grafi temporali.

Infine, per rendere l’idea di come evolve il modello per ogni unità di tempo, si riportano di seguito le immagini relative all’evoluzione della centralità della terza dimensione.

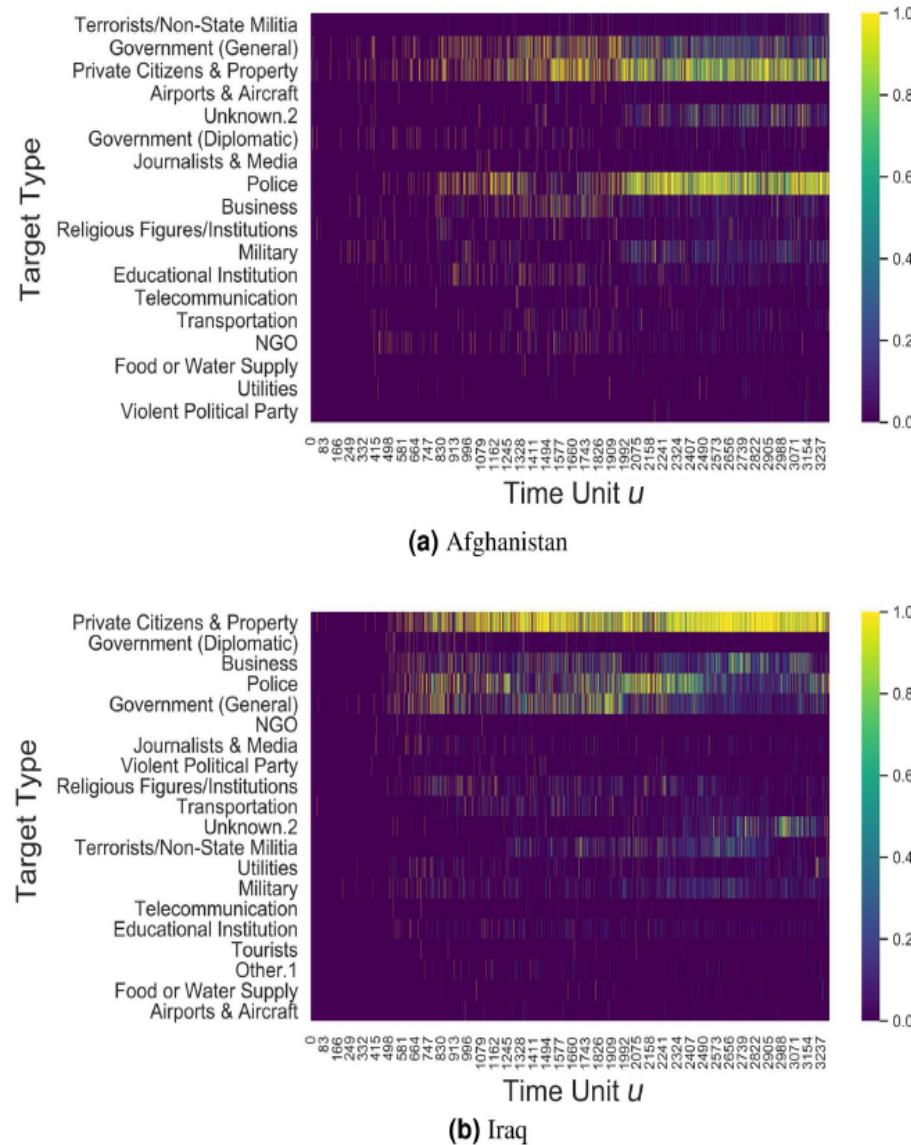


Figura 2.17: Evoluzione temporale della centralità della dimensione relativa agli obiettivi per i casi dell’Afghanistan e per quelli dell’Iraq.

2.2 Confronto fra le tecniche analizzate

All’interno di questa sezione vengono confrontate le tecniche esposte precedentemente per riuscire ad individuare eventuali vantaggi e svantaggi di ciascuna di esse e sulla base di ciò nel prossimo capitolo vengono proposte nuove possibili soluzioni.

Gli approcci considerati sono:

- Probabilistic Risk Analysis;
- Agent-Based Simulation;

- GIS e Random Forest;
- Computer Vision e Deep Learning;
- Temporal Meta-Graph.

Ognuno di essi presenta caratteristiche differenti, ad esempio l'approccio proposto da Kao et al.[3] e quello proposto da Campedelli et al.[5] sono legati strettamente ad alcune aree geografiche. La tecnica che utilizza la Random Forest è stata modellata su una specifica zona, ovvero la penisola dell'Indocina, così come nello studio che sfrutta i meta-grafi temporali i luoghi analizzati sono l'Afghanistan e l'Iraq.

Il problema relativo alla prevenzione di attacchi terroristici dovrebbe utilizzare una tecnica che non sia context-based; in altri termini, l'approccio adottato non deve essere limitato ad un solo contesto, ma dovrebbe essere versatile.

Inoltre fra le tecniche analizzate risalta in modo particolare quella relativa alla StyleGAN in quanto, nel caso in cui si riescano a superare le difficoltà riscontrate, risulterebbe essere particolarmente efficace.

Basti pensare che nel momento in cui si sviluppa tale framework, le forze dell'ordine e le agenzie governative non devono fare altro che scattare una foto con un qualsiasi dispositivo, ad esempio uno smartphone, e fornirla in input al software. Una volta che l'immagine viene elaborata si ha in output una foto che mostra eventuali pericoli e punti deboli poiché aggiunge elementi allo scenario di partenza, come criminali, agenti di polizia e passanti.

A questo punto è evidente come uno strumento di questo tipo risulti essere particolarmente utile e soprattutto generale, visto che può essere utilizzato in qualsiasi Stato e in qualsiasi città, sia in una grande metropoli che in una piccola cittadina.

Si può sostenere che, grazie al progresso nell'ambito dell'intelligenza artificiale e in modo particolare con le Generative Adversarial Network, la fattibilità di un software che permetta un'analisi di questo tipo aumenti sempre di più.

Considerando la prima tecnica analizzata bisogna evidenziare che è stata utilizzata da alcuni governi per alcuni decenni, ma è completamente diversa da quella proposta da Palomba, Fabio et al.[4] in quanto lo studio sviluppato da Ezell et al.[1] si focalizza solamente su un'analisi del rischio di tipo probabilistico.

Una volta sviluppati gli alberi di decisione si calcolano le probabilità che si verifichi uno scenario piuttosto che un altro. Nonostante questo approccio possa apparire banale, la dif-

ficoltà maggiore sta nel codificare le informazioni ottenute dall'Intelligence Community in probabilità. Nel caso in cui si calcola erroneamente questo valore i risultati che ne seguono non sono in alcun modo attendibili.

La tecnica che fa riferimento al sistema informativo geografico combinata con algoritmi di machine learning come la Random Forest si è rivelata particolarmente utile per individuare i cosiddetti "punti caldi", luoghi nei quali vi è un elevato rischio di attacco terroristico. La caratteristica peculiare di questo studio è stato il sistema Geo-Informativo, il quale si è adattato perfettamente nella simulazione di attentati.

Nonostante non siano state individuate nello specifico le vulnerabilità degli hot spot, le agenzie governative sono riuscite comunque a comprendere che la zona più pericolosa è Bangkok e quindi è necessario migliorare la sicurezza in questo luogo.

Un'altra tecnica innovativa è quella proposta da Bosse et al.[7]. Qui l'autore si focalizza sull'obiettivo dei guardiani, i quali scelgono una strategia fra alcune a loro disposizione per mantenere il tasso di criminalità il più basso possibile.

I risultati migliori di questa tecnica si ottengono con la strategia secondo cui i guardiani si spostano verso nuovi luoghi in base al numero di passanti che ci si aspetta di trovare nel periodo futuro. Tuttavia bisogna ammettere che anche le altre strategie elaborate da Bosse hanno prodotto buoni risultati.

Sebbene quest'analisi sia particolarmente positiva, bisogna ricordare che i risultati non vanno generalizzati poiché sono stati raggiunti in un ambiente di simulazione nel quale sono stati semplificati molti vincoli. Pertanto il livello di complessità di questi scenari è minore rispetto alla complessità degli scenari del mondo reale.

Ad esempio, da un punto di vista pratico non è facile calcolare l'esatto valore dell'attractiveness di un luogo oppure il numero di assalti che si possono verificare. Tuttavia i risultati di queste simulazioni sono stati utili per le forze dell'ordine, in quanto essendoci molteplici strategie si può scegliere quella più congeniale alla situazione in cui ci si trova.

Nella seguente tabella vengono mostrate tutte le tecniche e per ciascuna di esse vengono mostrate in modo schematico vantaggi e svantaggi.

Tabella 2.1: Tabella che indica pro e contro di ciascuna tecnica analizzata.

Tecnica analizzata	Pro	Contro
PRA - Probabilistic Risk Analysis	Semplicità d'uso	Difficoltà nel codificare le informazioni in probabilità
Agent-Based Simulation	Differenti strategie da adottare	Molti dettagli della realtà vengono trascurati
GIS e Random Forest	Grande efficacia nell'individuare gli hot spot	Tecnica context-based limitata solo alla penisola dell'Indocina
ALTER – Adversial Learning for counTER-torism	Forte efficacia e facilità d'uso per le forze dell'ordine	Elevata complessità computazionale richiesta
Temporal Meta-Graphs	Grande adattamento al modello dinamico che evolve nel tempo	Tecnica context-based limitata solo all'Afghanistan e all'Iraq

CAPITOLO 3

Sviluppo di nuove tecniche per la prevenzione di attacchi terroristici

All'interno di questo capitolo si discute dello sviluppo di nuove tecniche volte alla prevenzione di attacchi terroristici.

Sulla base di quanto detto nel capitolo precedente ogni tecnica analizzata presenta i suoi vantaggi e svantaggi e a fronte di ciò vengono proposte nuove tecniche. Fra i vari approcci, risulta particolarmente innovativo quello che utilizza le Generative Adversial Network.

In particolare, nello studio condotto da Palomba, Fabio et al.[4] si utilizza la StyleGAN, da qui la NVIDIA ha sviluppato una nuova versione di questa rete neurale, denominata StyleGAN2 la quale risolve alcune difficoltà riscontrate nel modello precedente. Tali soluzioni verranno analizzate di seguito.

3.1 Analisi di tecniche sviluppate

Fra i vari approcci che sono stati sviluppati sicuramente uno di quelli più innovativi è quello che fa riferimento alla visione artificiale.

In modo particolare si propone l'utilizzo di differenti reti neurali, come ad esempio la StyleGAN2 e la BigGAN.

Un'altra tecnica che verrà trattata successivamente è quella che fa riferimento agli algoritmi genetici. Questi algoritmi sono stati sviluppati basandosi sulle teorie evoluzionistiche di Darwin, infatti, essi si basano sul principio darwiniano secondo il quale gli elementi più

adatti all’ambiente hanno maggiore possibilità di sopravvivere e di trasmettere le loro caratteristiche ai successori: vi è una popolazione di individui, i quali evolvono di generazione in generazione attraverso meccanismi di riproduzione e di mutazione.

In questo modo si avrà una ricerca euristica che privilegia le zone dello spazio di ricerca dove è maggiore la possibilità di trovare le migliori soluzioni, non trascurando altre zone a più bassa probabilità di successo in cui saranno impiegate comunque risorse, ma in numero minore.

3.1.1 Approccio mediante la rete generativa avversaria StyleGAN2

La rete neurale generativa avversaria presa in analisi è stata sviluppata da Karras et al.[8] e lo scopo principale è quello di permettere il training della rete con un numero limitato di dati.

Solitamente il training delle GAN, nel caso in cui si usino dataset di piccole dimensioni, conduce a fenomeni di overfitting e l’addestramento non converge. L’idea proposta dalla NVIDIA è quella di introdurre un meccanismo adattivo del discriminatore, migliorandone le performance quando vi è mancanza dei dati.

Questo approccio non influenza in alcun modo la scelta delle loss function o l’architettura della rete; per questo motivo è stato mostrato che su numerosi dataset è possibile ottenere buoni risultati usando solamente poche migliaia di immagini per il training del modello.

Il numero di campioni richiesti dalla rete diminuisce notevolmente visto che in precedenza erano richieste dalle 10^5 alle 10^6 immagini per addestrare una Generative Adversarial Network di alta qualità.

Idealmente lo scopo di Karras è quello di evitare la modifica manuale dei parametri relativi al discriminatore e rendere il controllo dinamico in base al grado di overfitting della rete sul dataset.

Una soluzione applicabile per far sì che non si verifichi il fenomeno di overfitting è quella di utilizzare un insieme di validazione separato e osservarne il comportamento in relazione all’insieme di training. In altri termini, da un punto di vista matematico, si indicano gli output del discriminatore come D_{train} , $D_{validation}$ e $D_{generated}$ rispettivamente per il training set, validation set e immagini generate e si indica la media su N mini batch consecutive definite come $E[\cdot]$.

In pratica scegliendo il valore di N pari a 4 si ottengono $4 \times 64 = 256$ immagini. Le osservazioni condotte all’interno dello studio hanno dato origine a due formule plausibili di overfitting:

$$r_v = \frac{E[D_{train}] - E[D_{validation}]}{E[D_{train}] - E[D_{generated}]} \quad (3.1.1)$$

$$r_t = E[\text{sign}(D_{train})] \quad (3.1.2)$$

Per entrambe le euristiche appena mostrate, $r = 0$ sta ad indicare che non c'è alcun fenomeno di overfitting e $r = 1$ indica che la rete si è completamente sovra-adattata sui dati.

Lo scopo è quello di scegliere al meglio l'euristica che si adatti bene al modello del discriminatore. La prima euristica, r_v , esprime l'output per il convalidation set relativo al training set e alle immagini generate. Dal momento in cui si presuppone l'esistenza di un convalidation set separato viene incluso principalmente come metodo di confronto. La seconda euristica, r_t , stima la porzione del training set che ha in output dal discriminatore valori positivi.

Dopo aver descritto l'architettura della nuova rete neurale, è necessario affrontare altre due problematiche relative alla scelta del dataset e alla scelta della loss function.

Nonostante la StyleGAN2 permetta di effettuare il training della rete neurale con un numero notevolmente ridotto di immagini, a causa della riservatezza delle foto relative a scenari terroristici, non è possibile creare un dataset reale, ma bisogna ricorrere ad uno artificiale.

La scelta sulla generazione dei dati ricade nuovamente sul videogioco Grand Theft Auto, il quale si è mostrato particolarmente adatto alla generazione di scenari terroristici ad eccezione di alcune problematiche, come ad esempio l'utilizzo di immagini riprese dall'alto che non hanno fatto altro che peggiorare le prestazioni della rete neurale.

Un'altra problematica riscontrata, utilizzando il dataset artificiale, è data dalla risoluzione delle immagini: essendo la risoluzione particolarmente alta implica un elevato costo computazionale. A tal proposito una soluzione più efficiente potrebbe essere quella di far riferimento ad una vecchia versione del videogioco, in quanto nelle immagini vengono trascurati molti dettagli e quindi vi è una risoluzione più bassa.

Si ricordi che nella costruzione del modello di simulazione non sono richiesti molti dettagli di un edificio. Ad esempio è più che sufficiente solo la struttura dell'edificio con porte ed eventuali finestre.

Di seguito vengono riportate alcune immagini riprese dal videogioco Grand Theft Auto la cui risoluzione è minore e la complessità computazionale si riduce notevolmente.



Figura 3.1: Immagini catturate da uno scenario simulato nel videogioco Grand Theft Auto.

Un’ultima caratteristica da analizzare è la scelta della loss function: nello studio condotto da Palomba, Fabio et al.[4] inizialmente era stata scelta come funzione obiettivo una combinazione lineare di più loss function, ma si era stato notato che la rete aveva difficoltà a convergere ai risultati sperati. Successivamente utilizzando una sola loss function, la Pixel-wise, si è notato che le prestazioni sono migliorate notevolmente.

Tale funzione viene utilizzata per individuare le differenze tra le immagini a livello di pixel. La loss function considerata misura le differenze tra i valori dei pixel di output in un’immagine. Sebbene la funzione sia utile per comprendere l’interpolazione a livello di pixel, il processo presenta degli svantaggi. Ad esempio, alcuni data scientist sostengono che la loss function per pixel non affronta in maniera accurata la qualità dell’immagine il quale è invece

un aspetto di particolare importanza.

3.1.2 Approccio mediante la rete generativa avversaria BigGAN

Il secondo approccio che è stato proposto presenta le medesime caratteristiche della tecnica appena descritta con l'unica differenza che l'architettura utilizzata è diversa: anziché ricorrere alla StyleGAN si utilizza la BigGAN.

Il training del modello è dinamico e particolarmente sensibile ai valori dei parametri, ad esempio i parametri di ottimizzazione, ma lo scopo dello studio condotto da Brock et al.[9] è quello di rendere il training stabile. Con le Big Generative Adversarial Network si colma il gap fra le immagini generate dalla rete neurale e le immagini prese dal dataset costruito artificialmente.

Si è notato che le BigGAN presentano dei vantaggi notevoli rispetto alle normali GAN. Per questo motivo Brock et al.[9] ha effettuato il training del modello su ImageNet (database con un numero elevato di immagini realizzata per il riconoscimento di oggetti nell'ambito della visione artificiale) ad una risoluzione di 128x128 e le prestazioni sono incrementate notevolmente ottenendo come Inception Score (IS) un valore compreso fra 52.52 e 166.5 e come Fréchet Inception Distance (FID) un valore compreso fra 18.65 e 7.4.

In generale l'obiettivo delle GAN è quello di trovare l'equilibrio di Nash all'interno di un problema di min-max con due giocatori:

$$\min_G \max_D E_{x \sim q_{data(x)}} [\log D(x)] + E_{z \sim p(z)} [\log(1 - D(G(z)))] \quad (3.1.3)$$

dove z appartiene ad R^{d_z} ed è una variabile ottenuta dalla distribuzione $p(z)$, G e D rappresentano le due reti neurali convoluzionali e corrispondono rispettivamente al generatore e al discriminatore.

La BigGAN è progettata per la generazione di immagini class-conditional. In altri termini la rete presa in considerazione si focalizza sullo scaling up delle GAN infatti presenta le seguenti caratteristiche: più parametri, un'ampiezza maggiore delle batch e alcune modifiche dal punto di vista dell'architettura.

Sicuramente si può affermare che, per quanto riguarda le prestazioni, tale tecnica risulta essere più efficace rispetto ad una Generative Adversarial Network basilare, ma allo stesso tempo tutto ciò comporta un elevato costo computazionale di un peso non trascurabile.

Di seguito viene mostrato un risultato ottenuto mediante l'interpolazione tra i campioni di una BigGAN.



Figura 3.2: Immagini generate da un Big Generative Adversarial Network.

3.1.3 Approccio mediante algoritmi genetici

L'ultima tecnica che è stata individuata è quella relativa allo sviluppo di algoritmi genetici, i quali vengono utilizzati per risolvere problemi di ottimizzazione, che nel nostro caso sono la prevenzione di attacchi terroristici. L'aggettivo "genetico" è ispirato al principio della selezione naturale teorizzato da Charles Darwin e deriva dal fatto che gli algoritmi genetici utilizzano dei meccanismi concettualmente simili a quelli dei processi della genetica.

Prima di indicare il funzionamento di questa tipologia di algoritmi è necessario indicare gli elementi che li compongono.

Il cromosoma è una delle soluzioni al problema preso in considerazione cioè la prevenzione di attacchi terroristici e non sarebbe altro che una strategia che le forze dell'ordine potrebbero adottare. A questo punto se si individua un insieme di soluzioni relative a questo problema si parla popolazione.

Il gene è la parte di un cromosoma e rappresenta una caratteristica della soluzione che si sta considerando. Solitamente consiste di una o più parti del vettore le quali codificano il cromosoma. Ad esempio nel problema che si sta considerando un gene potrebbe essere il numero di agenti che intervengono durante l'attacco oppure qual è l'edificio sul quale è necessario migliorare la sicurezza, utilizzando camere di videosorveglianza.

Si parla di valore fitness quando si indica il grado di valutazione associato ad una soluzione, in altri termini corrisponde a quanto è considerata "buona" tale soluzione. La valutazione avviene in base ad una funzione appositamente sviluppata che prende il nome di funzione di fitness.

Un altro aspetto da definire è il crossover e corrisponde alla generazione di una nuova soluzione combinando soluzioni già esistenti. In questo caso tale operazione consiste nel

combinare due strategie difensive scegliendo i geni di ciascun cromosoma che risultano essere più efficaci nel contrasto dell'attentato terroristico.

Infine si ricorre alla mutazione per alterare in maniera casuale una soluzione, evitando così di focalizzare la ricerca in una sola zona dello spazio di ricerca.

Una codifica che potrebbe essere utilizzata è quella vettoriale reale in cui i cromosomi sono costituiti da un vettore di n campi di numeri reali. Si ha così il vantaggio di introdurre maggiore espressività e versatilità nella codifica, portando però una maggiore complessità.

Nel caso in cui si decida di implementare questa tecnica la difficoltà principale risiede nel riuscire ad elaborare la funzione di fitness, la quale dovrebbe essere in grado di stabilire se una soluzione risulti essere efficiente o meno, ma tale calcolo è particolarmente complesso da effettuare.

Nel momento in cui bisogna procedere all'operazione di crossover è necessario indicare la logica di selezione con la quale vengono prelevati alcuni dei campioni fra la popolazione. Esistono varie tecniche di "selezione", ad esempio la selezione a roulette, la selezione per categoria oppure la selezione a torneo. Quest'ultima è quella che potrebbe risultare adatta in questo contesto poiché le soluzioni vengono raggruppate e si procede a valutarle secondo un algoritmo, il quale consiste nello scegliere in maniera casuale gli individui appartenenti alla popolazione, in seguito si sceglie l'individuo migliore e si imposta la sua probabilità di scelta a p . Il secondo individuo che viene scelto è il secondo migliore e la sua probabilità di scelta è pari a $p * (1 - p)$ e si procede in questo modo fino ad esaurire le soluzioni scelte.

Nel caso in cui si voglia elaborare una tecnica più complessa, ma sicuramente allo stesso tempo anche più efficiente, è possibile utilizzare un algoritmo genetico multiobiettivo.

Da un punto di vista pratico l'algoritmo segue gli stessi step di un algoritmo genetico a singolo obiettivo, ma si individuano e si conservano un certo numero di soluzioni ottimali, le quali verranno poste su un fronte di Pareto.

La differenza principale risiede nel fatto che ora esistono due o più funzioni fitness da valutare, ad esempio massimizzare la sicurezza di un edificio, minimizzare il numero di individui coinvolti nell'attentato oppure minimizzare i danni alle strutture prese di mira dai terroristi.

Di seguito viene mostrato un flowchart relativo alla composizione di un algoritmo genetico.

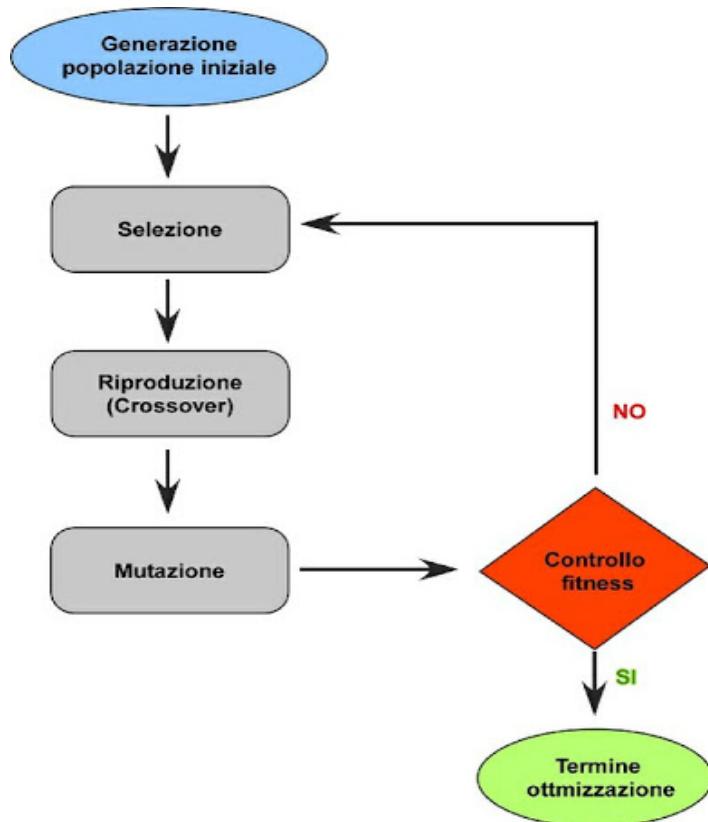


Figura 3.3: Struttura base di un algoritmo genetico

3.2 Confronto fra le varie tecniche

In questa sezione si effettua un confronto fra le tecniche descritte in precedenza. Principalmente è possibile suddividere gli approcci in due categorie: la prima fa riferimento alle Generative Adversarial Network e alla visione artificiale, la seconda, invece, fa riferimento ad un particolare tipo di algoritmi i quali prendono il nome di algoritmi genetici.

Nel primo caso le tecniche presentano le medesime caratteristiche con l'unica differenza che l'architettura della GAN è differente, prima si propone la Stylegan2 e poi la BigGAN.

Le due reti neurali seppure entrambe presentano la stessa struttura, la Stylegan2 è nata con l'intento della generazione di volti artificiali e la generazione di dipinti. Per questo motivo non si adatta particolarmente alla generazione di scenari terroristici visto che la struttura di un edificio presenta feature ben diverse da quelle di un grattacielo.

La BigGAN invece potrebbe risultare più efficace in questo ambito anche se è necessario evidenziare che il costo computazionale di questa rete neurale è notevolmente elevato.

Se si vuole utilizzare una metrica per confrontare le prestazioni fra le due Generative Adversarial Network si può considerare il FID (Fréchet Inception Distance), il quale viene utilizzato

per indicare la qualità delle immagini create dal generatore.

Tale valore si ottiene da:

$$FID = |\mu - \mu_w|^2 + \text{tr}(\sum + \sum_w - 2(\sum \sum_w)^{1/2}) \quad (3.2.1)$$

Nel caso in cui si effettui il confronto fra le due reti neurali utilizzando questa metrica si ha che la Stylegan2 ha un valore compreso fra 5.59 e 2.42 a fronte del valore di 7.4 ottenuto dalla BigGAN e quindi la rete sviluppata dalla NVIDIA dovrebbe risultare più potente come architettura.

L'approccio che fa riferimento agli algoritmi genetici è ben diverso da quelli appena confrontati. Gli algoritmi genetici permettono di modificare più parametri e più tecniche in modo da personalizzare ed adattare l'algoritmo al contesto nel quale ci si trova. Ciò deriva dal fatto che esistono vari modi per codificare i cromosomi, è possibile sviluppare più funzioni di fitness, ci sono più tecniche di crossover o di selezione.

3.2.1 Performance

Nel momento in cui si effettua un confronto fra le tecniche proposte basandosi sulle performance, come il costo computazionale allora in tal caso la scelta più efficiente sono gli algoritmi genetici poiché le Generative Adversarial Network hanno un elevato costo computazionale.

Basti pensare che il training della rete su immagini con una risoluzione 1024x1024 richiede una o più GPU con almeno 16GB di memoria. Ad esempio i tempi di training su immagini con una risoluzione 1024x1024 con una GPU richiedono dai 14 giorni ai 69 giorni.

A tal proposito è evidente come un risultato di questo genere è particolarmente costoso a fronte del costo computazionale richiesto dall'esecuzione di uno o più algoritmi genetici.

Nella figura mostrata di seguito viene rappresentato un grafico che effettua un confronto in termini di performance fra la Stylegan e la Stylegan2 evidenziandone le differenze.

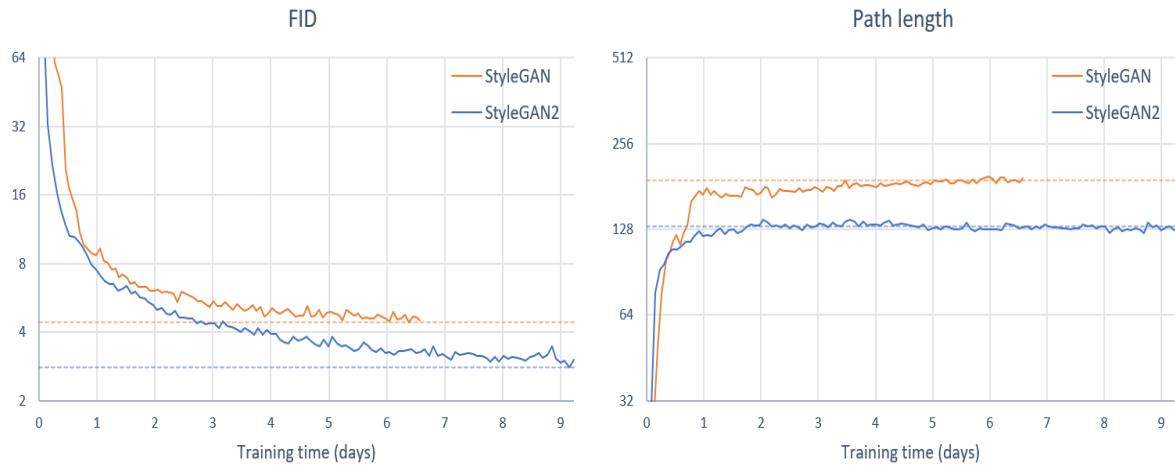


Figura 3.4: Confronto fra la Stylegan e la Stylegan2.

3.2.2 Capacità di cross-context

Un altro aspetto da tenere in considerazione per le tecniche sviluppate appena descritte è la capacità di cross-context. In altri termini la possibilità di adottare questi approcci catturati anche da altri contesti ed effettuare un’intersezione fra di loro.

Fra le tecniche proposte sicuramente è possibile aggiungerne delle nuove attraverso delle combinazioni. Ad esempio se si prende in considerazione la terza tecnica sviluppata è possibile combinare altre strategie con essa come quella proposta da Ezell et al.[1] che utilizza un approccio probabilistico volto all’analisi del rischio di attacchi terroristici.

L’approccio mediante la visione artificiale può essere applicato in più contesti, sempre legati alla sicurezza dei luoghi pubblici, ma presenta alcuni limiti poiché i parametri che costituiscono la Generative Adversarial Network sono particolarmente legati alla prevenzione di attività criminali.

3.3 Analisi dei risultati

All’interno di questa sezione vengono analizzati i risultati ottenuti dopo aver sviluppato le differenti tecniche il cui scopo è quello di prevenire attentati terroristici attraverso la simulazione.

L’approccio che si è rivelato più adatto è stato quello che utilizza l’architettura della Stylegan2 anche se gli strumenti disponibili oggigiorno non permettono di eseguire in maniera efficiente tali tecniche.

Sicuramente si può affermare che è stato introdotto, già con Palomba, Fabio et al.[4], un

nuovo modo per la simulazione di scenari terroristici utilizzando la computer vision e le GAN le quali imparano in maniera automatica le feature degli scenari.

Nel mentre il training della Stylegan2 non si è mostrato particolarmente adatto con la struttura scelta, l'approccio della generazione di scenari terroristici sulla base di un ambiente artificiale, elaborato all'interno di un videogioco, ha portato alla luce buoni risultati.

CAPITOLO 4

Osservazioni e limitazioni

All'interno di questo capitolo vengono descritte alcune osservazioni e limitazioni riscontrate nello studio di nuove tecniche volte alla prevenzione di attacchi terroristici mediante la simulazione.

Lo scopo di questo lavoro è stato quello di individuare un approccio che permettesse alle forze dell'ordine di simulare scenari di attacchi terroristici nella vita reale in modo tale da poter sviluppare piani di contingenza.

Ogni approccio ha presentato delle limitazioni, partendo dalla PRA, ovvero la Probabilistic Risk Analysis, la quale ha creato problemi nel momento in cui bisognava codificare le informazioni ottenute dall'Intelligence Community in probabilità.

Se si prende in considerazione la tecnica che ricorre alla simulazione di un agente basato su modello seppure vi siano numerose strategie che è possibile adottare, il modello di simulazione viene semplificato notevolmente rispetto alla realtà poiché viene considerato solo ad alto livello uno scenario terroristico. Nel mondo reale, invece, ci sono molti più dettagli da considerare che però vengono trascurati.

Nello studio che fa riferimento alla Random Forest, l'algoritmo di machine learning è stato combinato con un geo-information system per simulare la distribuzione del rischio di attacchi terroristici su una scala di pixel. Prima della simulazione, la penisola dell'Indocina era stata analizzata utilizzando un metodo kernel density. Si è notato che nell'Indocina sono presenti cinque hot spot e il luogo più soggetto ad attività criminali è Bangkok e le città limitrofe.

L'approccio che utilizza le Generative Adversarial Network risulta essere particolarmente

innovativo e con lo sviluppo della tecnologia utilizzata sicuramente sarà più efficace e avrà un'elevata facilità d'uso. La limitazione principale è data dal fatto che la complessità computazionale richiesta è molto complessa e con gli strumenti disponibili oggigiorno non è possibile sviluppare un framework che garantisca delle ottime prestazioni in tempo polinomiale. Il lavoro svolto da Campedelli et al.[5] evidenzia il potenziale delle soluzioni di machine learning e deep learning nelle strategie di previsione di attacchi terroristici. Questi risultati promettenti presentano alcune limitazioni come ad esempio la generalizzazione dei risultati in altri contesti geografici perché lo scopo degli studiosi è quello di assumere che gli eventi occorrono nello stesso paese e siano parte dei processi di decisione ad alto livello strategico. Nel caso in cui si tenti questo approccio in Stati dove vi è un numero maggiore di attori, di gruppi o organizzazioni allora le problematiche aumentano e le limitazioni si iniziano a notare.

CAPITOLO 5

Conclusioni

L'obiettivo della tesi era quello di ideare e sviluppare una o più tecniche volte alla simulazione di attacchi terroristici in modo tale da elaborare eventuali piani di contingenza. Si è partiti dall'analisi di cinque tecniche differenti fino ad elaborare tre tecniche delle quali due di esse presentano le medesime caratteristiche ad eccezione dell'architettura della rete neurale adottata.

Alla fine del percorso, possiamo affermare di essere riusciti nel conseguimento del nostro obiettivo. Siamo partiti dall'analisi delle diverse metodologie e successivamente, dopo averne analizzate e descritte le caratteristiche positive e negative, abbiamo studiato ed esplorato nuove tecniche.

In ultima battuta abbiamo ideato e descritto una possibile implementazione di un framework il cui scopo è quello di individuare i punti deboli, qualora vi fossero, degli edifici e le strategie per minimizzare il numero di danni nei confronti dei passanti. Uno strumento di questo tipo potrebbe risultare più che utile alle agenzie governative e alle forze dell'ordine per rendere le strade che percorriamo quotidianamente più sicure.

Bibliografia

- [1] B. C. Ezell, S. P. Bennett, D. Von Winterfeldt, J. Sokolowski, and A. J. Collins, "Probabilistic risk analysis and terrorism risk," *Risk Analysis: An International Journal*, vol. 30, no. 4, pp. 575–589, 2010. (Citato alle pagine 7, 10, 13, 36 e 48)
- [2] T. Bosse and C. Gerritsen, "Comparing crime prevention strategies by agent-based simulation," in *2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology*, vol. 2, pp. 491–496, IEEE, 2009. (Citato alle pagine 7 e 14)
- [3] M. Hao, D. Jiang, F. Ding, J. Fu, and S. Chen, "Simulating spatio-temporal patterns of terrorism incidents on the indochina peninsula with gis and the random forest method," *ISPRS International Journal of Geo-Information*, vol. 8, no. 3, p. 133, 2019. (Citato alle pagine 7, 18, 19, 21, 31 e 36)
- [4] G. Cascavilla, J. Slabber, F. Palomba, D. Di Nucci, D. A. Tamburri, and W.-J. van den Heuvel, "Counterterrorism for cyber-physical spaces: A computer vision approach," in *Proceedings of the International Conference on Advanced Visual Interfaces*, pp. 1–5, 2020. (Citato alle pagine 8, 24, 29, 36, 39, 42 e 48)
- [5] G. M. Campedelli, M. Bartulovic, and K. M. Carley, "Learning future terrorist targets through temporal meta-graphs," *Scientific reports*, vol. 11, no. 1, pp. 1–15, 2021. (Citato alle pagine 8, 31, 36 e 51)
- [6] T. Karras, S. Laine, and T. Aila, "A style-based generator architecture for generative adversarial networks," 2019. (Citato alle pagine 24, 26 e 28)

- [7] T. Bosse, C. M. Jonker, L. van der Meij, and J. Treur, "Leadsto: A language and environment for analysis of dynamics by simulation," in *Innovations in Applied Artificial Intelligence* (M. Ali and F. Esposito, eds.), (Berlin, Heidelberg), pp. 363–366, Springer Berlin Heidelberg, 2005. (Citato a pagina 37)
- [8] T. Karras, M. Aittala, J. Hellsten, S. Laine, J. Lehtinen, and T. Aila, "Training generative adversarial networks with limited data," 2020. (Citato a pagina 40)
- [9] A. Brock, J. Donahue, and K. Simonyan, "Large scale gan training for high fidelity natural image synthesis," 2019. (Citato a pagina 43)

Ringraziamenti

Ringrazio il Professor Fabio Palomba per avermi dato l'opportunità di lavorare con lui ad un progetto, per me, particolarmente interessante e significativo, guidandomi con pazienza e professionalità alla stesura di questa tesi.

Desidero ringraziare tutti i miei amici e colleghi, che sono sempre stati al mio fianco, pronti a festeggiare per ogni esame passato. In modo particolare desidero ringraziare Andrea Terlizzi prima come compagno di vita e poi come collega.

Il ringraziamento più grande va a tutta la mia famiglia: a mia sorella Alessandra per essere stata, da sempre, al mio fianco ed è stata per me esempio di forza; a mia madre, Tiziana, una costante, senza la quale non ce l'avrei mai fatta e infine all'uomo più importante della mia vita, mio padre Salvatore, che mi ha sempre trasmesso saggezza e grazie a lui sono la persona che sono ora.