

Sistema de Integração Segura e Análise Automática de CCTV com Câmaras em Redes Não Controladas

Gustavo Caiano

**Dissertação para obtenção do Grau de Mestre em
Engenharia Informática, Área de Especialização em
Engenharia de Software**

**Orientador: Paulo Baltarejo Sousa
Supervisor: Francisco Loureiro**

Declaração de Integridade

Declaro ter conduzido este trabalho académico com integridade.

Não plagiei ou apliquei qualquer forma de uso indevido de informações ou falsificação de resultados ao longo do processo que levou à sua elaboração.

Portanto, o trabalho apresentado neste documento é original e de minha autoria, não tendo sido utilizado anteriormente para nenhum outro fim.

Declaro ainda que tenho pleno conhecimento do Código de Conduta Ética do P.PORTO.

ISEP, Porto, 21 de dezembro de 2025

Dedicatória

The dedicatory is optional. Below is an example of a humorous dedication.

"To my wife Marganit and my children Ella Rose and Daniel Adam without whom this book would have been completed two years earlier." in "An Introduction To Algebraic Topology" by Joseph J. Rotman.

Resumo

O presente trabalho de dissertação propõe o desenvolvimento do Flexible Universal Stream Engine (FUSE), uma plataforma de software inovadora concebida para centralizar e organizar sistemas de videovigilância heterogêneos, geograficamente dispersos e situados em redes não controladas. Esta investigação surge em resposta a desafios críticos identificados na segurança pública e privada, nomeadamente a fragmentação tecnológica, as vulnerabilidades de segurança inerentes às comunicações Peer-to-Peer (P2P) e a ineficiência operacional enfrentada por entidades como os Órgão de Polícia Criminals (OPCs) na análise forense de vídeo.

A arquitetura proposta integra uma camada de abstração de hardware que normaliza a comunicação com câmaras de diversos fabricantes através do protocolo Real Time Streaming Protocol (RTSP), garantindo a interoperabilidade. A segurança das comunicações é assegurada através de túneis Virtual Private Network (VPN) que protegem a integridade e confidencialidade dos dados em redes hostis, complementada por um modelo de controlo de acessos Action/Attribute Based Access Control (ABAC). Adicionalmente, o sistema incorpora um módulo de visão computacional progressivo, estruturado em três fases: deteção de atividade para filtragem temporal, classificação de objetos (e.g., humanos, veículos) e extração de atributos específicos, visando a automatização da análise forense e a redução significativa da carga de trabalho manual.

A metodologia adotada segue o paradigma de "Design and Creation", com a validação da solução a ser realizada através de um protótipo funcional (Minimum Viable Product (MVP)) em cenários simulados e reais. Os resultados esperados centram-se na demonstração da eficácia da integração segura de dispositivos, na robustez da proteção de dados e na otimização dos processos de investigação criminal, sempre em estrita observância dos princípios éticos e regulamentares, como o EU AI Act.

Palavras-chave: Videovigilância, Integração CCTV, Automatização, Deteção de Objetos, Software Seguro, IoT Communication

Abstract

This dissertation presents the development of FUSE, a software platform designed to unify and secure heterogeneous video surveillance systems dispersed across uncontrolled networks. Addressing the challenges of technological fragmentation and security vulnerabilities in current Closed Circuit Television (CCTV) infrastructures, FUSE provides a centralized solution for entities such as Law Enforcement Agencies (LEAs)

The proposed architecture features a hardware abstraction layer for normalizing camera communications via RTSP and secures data transmission through VPN tunneling. It implements ABAC as an access control method, and integrates a three-stage computer vision pipeline for automated activity detection, object classification, and attribute extraction. This approach aims to enhance interoperability, data security, and forensic analysis efficiency, adhering to ethical standards like the EU AI Act.

Agradecimientos

The optional Acknowledgment goes here. . . Below is an example of a humorous acknowledgment.

"I'd also like to thank the Van Allen belts for protecting us from the harmful solar wind, and the earth for being just the right distance from the sun for being conducive to life, and for the ability for water atoms to clump so efficiently, for pretty much the same reason. Finally, I'd like to thank every single one of my forebears for surviving long enough in this hostile world to procreate. Without any one of you, this book would not have been possible."in "The Woman Who Died a Lot"by Jasper Fforde.

Conteúdo

Lista de Figuras	xv
Lista de Tabelas	xvii
Lista de Algoritmos	xvii
Lista de Código	xvii
Lista de Símbolos	xix
1 Introdução	1
1.1 Introdução e Enquadramento	1
1.2 Pergunta de Investigação	2
1.3 Objetivos	3
2 Estudo do Estado da Arte	5
2.1 Processo de Investigação	5
2.2 Estado da Arte em Visão Computacional	6
2.2.1 Detecção de Movimento e Filtragem Temporal	7
2.2.2 Detecção e Classificação de Objetos (Object Detection)	7
2.2.3 Modelos de Visão-Linguagem (Vision-Language Models)	7
2.3 Identificação da Lacuna Literária	7
3 Metodologia	9
4 Planeamento de Trabalho	11
4.1 Definição do Âmbito e Entregáveis	11
4.2 Plano de Trabalho	11
4.3 Project Charter e Gestão de Riscos	12
5 Ética	15
Bibliografia	17
Apêndice A Questionário e Entrevistas	19

Lista de Figuras

3.1	Diagrama de "Research Proces" de B. J. Oates [12], aplicado ao contexto do projeto.	9
4.1	WBS do projeto.	11
4.2	Gantt Chart do projeto.	12

Lista de Tabelas

2.1	Questões de Pesquisa Orientadas à Revisão Literária	5
4.1	Stakeholders identificados	12
4.2	Identificação e gestão de riscos associados ao projeto	13

Lista de Símbolos

a	distance	m
P	power	W (J s^{-1})
ω	angular frequency	rad

Capítulo 1

Introdução

1.1 Introdução e Enquadramento

No panorama atual dos sistemas de videovigilância verifica-se uma grande dependência de hardware, como Network Video Recorders (NVRs) para agregar várias câmaras e disponibilizar streaming da sua imagem. A maioria destes dispositivos também possibilita o playback e exportação de vídeos gravados, sendo que alguns controlam ainda os acessos de utilizadores. A detecção de movimentos ou objetos é uma feature de apenas alguns modelos apesar de que, ultimamente, se têm verificado um grande crescimento e aposta tecnológica nesta área [1]. Assim, os NVRs, embora raramente reúnam individualmente todas estas funcionalidades num dispositivo só [2], são a peça central do processo de videovigilância dentro de uma rede controlada.

Em determinados cenários [3], existe a impossibilidade destas câmaras serem inseridas todas numa determinada Local Area Network (LAN) privada. Por este motivo e por estarem numa outra rede não controlada, é necessário aceder via Internet, que origina alguns problemas, tais como:

- Incompatibilidade entre câmaras de diferentes fornecedores, obrigando muitas vezes à utilização de mais do que um software, já que nem todas são totalmente compatíveis com a mesma aplicação.
- Falta de segurança e de privacidade dos dados, uma vez que muitas destas câmaras recorrem a conexões Peer-to-Peer (P2P) quando acedidas fora da LAN. Este mecanismo depende, geralmente, de servidores do fornecedor, sobre os quais não existe controlo nem garantia de proteção dos dados [4]. Além disso, a comunicação entre a aplicação e a câmara é frequentemente pouco segura e não encriptada, expondo a stream de vídeo e as credenciais a potenciais ataques.
- Falta de controlo de acessos e gestão de utilizadores, pois a maior parte das aplicações de acesso remoto a Closed Circuit Television (CCTV) são desenvolvidas para cenários onde estas atividades não são a prioridade [5], priorizando o acesso intuitivo e rápido.

Para além destas barreiras técnicas e de segurança, emerge um desafio operacional extremamente relevante, particularmente no contexto da segurança pública. Entidades como os Órgão de Polícia Criminals (OPCs) dependem de videovigilância para a obtenção de provas. O processo atual, contudo, assenta frequentemente na revisão humana através da visualização integral contínua das gravações efetuadas, distribuindo muitas vezes intervalos de tempo por vários agentes para que seja facilitada. Esta informação foi obtida através de um questionário/entrevista, realizado em conjunto com agentes da investigação criminal da Guarda Nacional Republicana (GNR) e da Polícia de Segurança Pública (PSP), no qual o

objetivo principal seria a validação da existência deste problema. O mesmo pode ser consultado no ANEXO A. Ainda assim, este processo torna-se demorado e intensivo, recorrendo a uma enorme quantidade de recursos humanos e tempo, que são valiosos e, infelizmente, escassos [6].

Uma vez que estes problemas trazem insegurança e ineficiência, resolvê-los é uma preocupação do encarregado de segurança e CCTV de uma entidade pública ou privada, tal como dos agentes encarregues pela obtenção de prova sob videovigilância dentro das polícias de investigação criminal. É neste contexto de múltiplos desafios que surge a proposta deste trabalho: o desenvolvimento do Flexible Universal Stream Engine (FUSE). O FUSE é idealizado como uma plataforma de software que centraliza e organiza sistemas de videovigilância heterogêneos e geograficamente dispersos, com um foco integrado na segurança dos dados, na automatização da análise de vídeo, e na manutenção de registos de auditoria e gestão de utilizadores.

Importa ressaltar que o presente trabalho decorre em contexto empresarial, sendo acolhido pela empresa StabilityBubble, Lda. O FUSE não se configura apenas como um exercício académico teórico, mas sim como a proposta de uma plataforma comercial destinada a responder a lacunas de mercado identificadas pela empresa. Assim, esta dissertação visa conciliar o rigor da investigação científica com os requisitos práticos e operacionais de um produto de software real, tirando partido da infraestrutura e know-how da entidade de acolhimento.

1.2 Pergunta de Investigação

A proposta desta plataforma levanta a seguinte pergunta principais de investigação:

RQ1: De que forma pode ser desenhada/projetada uma solução de software que permita aceder seguramente a câmaras CCTV, localizadas em redes externas não controladas, caracterizadas por uma elevada diversidade de arquiteturas, padrões tecnológicos e origem de fabrico?

Decompondo esta pergunta para uma melhor e mais estruturada compreensão, a investigação será guiada pelas seguintes sub-perguntas operacionais:

RQ1.1: Como pode ser desenhada uma camada de abstração de software que normalize as funcionalidades (visualização, controlo, gravação) de câmaras de diferentes interfaces e especificações técnicas distintas, garantindo a extensibilidade futura do sistema?

RQ1.2: Que mecanismos de rede e protocolos de comunicação são mais eficazes para garantir a confidencialidade e integridade da comunicação com câmaras localizadas em redes não fidedignas, sem introduzir vulnerabilidades na rede de destino e agregação das várias streams?

RQ1.3: Em que medida a integração de modelos de visão computacional para a automatização da deteção de eventos pode validar a utilização da plataforma proposta para otimização de processos de investigação criminal?

A resposta à questão central de investigação (RQ1) será materializada através da implementação e validação experimental da plataforma FUSE. Todo o processo encontra-se detalhado no capítulo 5 referente à Metodologia. No entanto, para fundamentar as decisões técnicas necessárias na implementação do Minimum Viable Product (MVP) analisar-se-ão as sub-questões operacionais, no capítulo 4 do Estudo do Estado da Arte.

1.3 Objetivos

Foram definidos os seguintes objetivos para o desenvolvimento e validação do FUSE:

- Desenvolver uma arquitetura de software extensível que, através de uma camada de abstração, normalize a comunicação com câmaras de diferentes fornecedores. O sistema deverá suportar um protocolo base de streaming como o Real Time Streaming Protocol (RTSP) e centralizar as funcionalidades de visualização em tempo real, gravação, e playback numa interface unificada.
- Implementar um modelo de comunicação seguro, Secured By Design, que utilize túneis Virtual Private Network (VPN) para isolar e criptografar a comunicação entre as câmaras externas e o servidor de implementação da aplicação.
- Desenvolver um sistema de controlo de acessos baseado em ações (modelo Action/Attribute Based Access Control (ABAC)) para gerir as permissões de utilizadores de forma granular e garantir a auditoria das ações.
- Integrar um módulo de análise de vídeo para a deteção automatizada de eventos complexos, implementando um pipeline de processamento progressivo em três fases:
 - Fase 1 (Deteção de Atividade): Identificação de movimento e atividade relevante nos streams de vídeo para filtrar segmentos de interesse.
 - Fase 2 (Classificação de Objetos): Análise dos segmentos filtrados para detectar e classificar objetos de categorias pré-definidas (e.g., humanos, veículos, animais).
 - Fase 3 (Extração de Atributos): Análise aprofundada dos objetos classificados para extrair características específicas e customizáveis, como matrículas e cores de veículos, ou atributos de vestuário e acessórios de pessoas, que servirão de base para a pesquisa de eventos complexos.
- Validar a viabilidade da arquitetura através de um protótipo funcional (MVP) que demonstre a integração bem-sucedida de, no mínimo, duas/três (DECIDIR AINDA) câmaras tecnologicamente diferentes, a segurança da comunicação e a eficácia da deteção de eventos num cenário simulado, pelo menos até à Fase 2 anteriormente mencionada.

Quanto aos principais contributos deste trabalho, espera-se que do ponto de vista técnico-científico surja uma arquitetura referência para integração segura e inteligente de sistemas CCTV distribuídos, principalmente quando há a presença de interoperabilidade entre redes não controladas e controladas. Por outro lado, do ponto de vista social e operacional, tenciona-se validar a ferramenta para que esta possa vir a aumentar significativamente a eficiência e eficácia da investigação criminal dos OPC, otimizando a alocação de recursos e reduzindo o tempo de análise manual do vídeo.

Capítulo 2

Estudo do Estado da Arte

Este capítulo apresenta a revisão de literatura fundamentada no âmbito do projeto, analisando o estado atual das tecnologias críticas para o desenvolvimento do FUSE. O objetivo principal desta revisão é responder às sub-questões operacionais previamente identificadas, nomeadamente a RQ1.1 e a RQ1.2, estabelecendo uma base teórica sólida para as decisões de arquitetura e segurança.

Adicionalmente, e dada a natureza aplicada da RQ1.3, é conduzido um estudo técnico aprofundado sobre os paradigmas atuais de visão computacional. Este estudo visa não apenas identificar o estado da arte, mas também comparar padrões, pipelines de processamento e modelos de Artificial Intelligence (AI) passíveis de integração na plataforma.

2.1 Processo de Investigação

O Processo de investigação foi guiado pelo mapeamento das duas primeiras sub-questões da presente dissertação em questões de pesquisa com foco na revisão literária, conforme descrito na Tabela 2.1.

Tabela 2.1: Questões de Pesquisa Orientadas à Revisão Literária

RQ ID	LRRQ ID	Description	Tópicos e Keywords de pesquisa
RQ1.1	LRRQ1	Quais são as arquiteturas de referência e padrões de design de software descritos na literatura para a abstração e interoperabilidade de dispositivos Internet of Things (IoT) heterogêneos?	IoT Interoperability, Hardware Abstraction Layer, Middleware patterns, Open Network Video Interface Forum (ONVIF) standardization, Heterogeneous device integration.
RQ1.2	LRRQ2	Qual o estado da arte em protocolos de comunicação segura e VPNs para acesso remoto e streaming?	Secure Tunneling Protocols, VPN performance analysis, Streaming encryption, Network Address Translation (NAT) Traversal, Zero Trust Network Access (ZTNA).

Cada questão de pesquisa será enquadrada de acordo com o modelo Population, Intervention, Comparison, Outcomes, Context, Study (PICOCs) [7], garantindo o rigor na seleção das fontes utilizadas. A informação será depois selecionada seguindo o fluxo Preferred Reporting

Items for Systematic reviews and Meta-Analysis (PRISMA) [8]. O processo inclui a definição de keywords de pesquisa, a aplicação estrita de critérios de inclusão e exclusão, seguida de uma filtragem em etapas e, finalmente, uma discussão crítica sobre a aplicabilidade dos estudos selecionados para a arquitetura do FUSE.

Como fator inclusivo de seleção, selecionou-se, por exemplo, a data de publicação posterior a 2018. A escolha deste intervalo visa garantir que as arquiteturas, frameworks e algoritmos analisados representam o atual estado da arte, evitando a adoção de paradigmas que, embora válidos no passado, não refletem as necessidades de desempenho e escalabilidade dos sistemas modernos. Foram privilegiados artigos revistos por pares, normas técnicas e literatura técnica amplamente reconhecida. Excluíram-se estudos puramente teóricos sem aplicação prática ao domínio da videovigilância ou da integração de dispositivos, tal como trabalhos relativos a soluções proprietárias fechadas sem documentação técnica pública suficiente.

2.2 Estado da Arte em Visão Computacional

Relativamente à terceira sub-questão, a RQ1.3, referente à automatização da análise de vídeo, optou-se por uma abordagem de Revisão Narrativa e Exploratória, como descrita por Maria J. Grant e Andrew Booth em [9] como “State-of-the-art review”. Esta modalidade metodológica caracteriza-se pela sua flexibilidade na análise crítica da literatura atual, permitindo identificar conceitos-chave, padrões arquiteturais e soluções técnicas emergentes sem a rigidez protocolar de uma revisão sistemática e formal.

Esta opção justifica-se pela vertiginosa evolução dos modelos de AI e pela necessidade de analisar não apenas literatura académica clássica, mas também documentação técnica de modelos recentes e benchmarks da indústria. A análise encontra-se segmentada em três domínios fundamentais que correspondem às fases da pipeline de processamento proposta para o FUSE:

1. Detecção de Movimento (Fase 1)
2. Detecção de Objetos (Fase 2)
3. Visual-Language Models - Vision-Language Models (VLMs) (Fase 3)

Apesar da natureza exploratória, esta pesquisa manterá o rigor científico na seleção de fontes, priorizando publicações de menos de 1 ano, dado o exponencial e recente crescimento tecnológico da área, e repositórios open-source com forte validação comunitária. A pesquisa foi orientada pelas seguintes keywords, agrupadas por domínio:

- **Fase 1 (Pré-processamento):** Background Subtraction, Motion Detection Algorithms, Frame Differencing efficiency, Video Activity Detection.
- **Fase 2 (Classificação):** Real-time Object Detection, You Only Look Once (YOLO) architecture, One-stage detectors, Convolutional Neural Network (CNN) inference optimization, Edge AI.
- **Fase 3 (Extração de Atributos):** Vision-Language Models (VLMs), Multimodal AI, Zero-Shot Learning, Open-vocabulary detection, Visual Question Answering.

2.2.1 Detecção de Movimento e Filtragem Temporal

A primeira etapa da revisão incidirá sobre métodos de pré-processamento e filtragem de vídeo, essenciais para a eficiência global do sistema. O estudo focará na comparação entre algoritmos clássicos de processamento de imagem (baseados em diferenças de pixels e estatística de cena) e abordagens mais modernas de “lightweight AI”. O objetivo principal será identificar técnicas capazes de filtrar eficazmente segmentos de vídeo sem atividade relevante, minimizando o uso de recursos computacionais (Central Processing Unit (CPU)/Graphics Processing Unit (GPU)) e garantindo robustez face a mudanças de iluminação ou ruído visual, sem descartar falsos negativos críticos. A literatura evidencia uma divisão clara entre abordagens clássicas de processamento de imagem, que privilegiam a eficiência computacional e simplicidade de implementação, e abordagens baseadas em modelos leves de AI, que oferecem maior robustez semântica à custa de maior complexidade. Esta tensão entre eficiência e capacidade de generalização constitui um dos principais trade-offs analisados nesta fase da pipeline.

2.2.2 Detecção e Classificação de Objetos (Object Detection)

Neste domínio, a literatura será analisada com foco em arquiteturas de Redes Neurais Convolucionais (CNNs) otimizadas para inferência em tempo real. Será dado destaque preponderante à análise da família de arquiteturas YOLO [10], atualmente considerada o padrão de indústria para o equilíbrio entre velocidade e precisão [11]. O estudo comparativo visará determinar qual a versão ou variação desta arquitetura melhor se adequa aos requisitos forenses do projeto, avaliando métricas como a capacidade de detecção de objetos pequenos ou distantes e o desempenho em hardware com recursos limitados. Os estudos analisados preliminarmente revelam um debate recorrente entre arquiteturas altamente precisas, mas computacionalmente exigentes, e modelos otimizados para inferência em tempo real.

2.2.3 Modelos de Visão-Linguagem (Vision-Language Models)

Por fim, explora-se a fronteira mais recente da Inteligência Artificial: a integração entre visão computacional e processamento de linguagem natural. A revisão abordará o estado da arte dos VLMs, analisando arquiteturas multimodais recentes, como por exemplo a família Qwen-VL, entre outras emergentes. O foco da investigação será compreender como estes modelos permitem a extração de atributos complexos e a realização de pesquisas em linguagem natural, avaliando a sua viabilidade de integração num pipeline local em termos de latência e exigência de memória.

2.3 Identificação da Lacuna Literária

Da análise preliminar da literatura resulta uma lacuna clara: embora existam estudos sólidos sobre integração de dispositivos IoT, protocolos de comunicação segura e modelos avançados de visão computacional, estes domínios são maioritariamente abordados de forma isolada. Verifica-se a ausência de uma arquitetura integrada que combine, de forma sistemática, a interoperabilidade segura de câmaras CCTV localizadas em redes não controladas com pipelines de análise automática de vídeo orientadas a contextos forenses. É precisamente nesta interseção, entre segurança de comunicação, abstração de hardware heterogêneo e inteligência analítica aplicada, que o presente trabalho se posiciona.

Capítulo 3

Metodologia

A metodologia adotada para a realização deste trabalho acadêmico baseia-se no modelo de “Research Process” proposto por B. J. Oates no livro “Researching Information Systems and Computing” [12], ilustrado na Figura 3.1.

Figura 3.1: Diagrama de “Research Proces” de B. J. Oates [12], aplicado ao contexto do projeto.

O ponto de partida da investigação enquadra-se em “Experiences and motivation”, dado que a génese deste projeto deriva diretamente da observação da ineficiência e limitações técnicas enfrentadas pelos OPCs e gestores de segurança na recolha de imagens de CCTV. Através de feedback adquirido através de questionários e entrevistas com Núcleos (GNR) e Esquadras (PSP) de Investigação Criminal (Anexo A), estes problemas são confirmados e registrados. Complementarmente, é realizada uma revisão de literatura preliminar para compreender o estado da arte em protocolos de comunicação e visão computacional. Esta análise permitiu identificar a ausência de soluções integradas que garantam simultaneamente segurança em redes hostis e inteligência analítica avançada. A junção desta necessidade prática (experiência) com a identificação desta lacuna teórica (literatura) resultou na formalização das perguntas de investigação apresentadas anteriormente.

A estratégia central adotada para este trabalho é a de “Design and Creation”. Esta abordagem é a mais adequada para projetos de Engenharia de Software cujo objetivo primário é o desenvolvimento de um protótipo que visa resolver muitos dos problemas práticos observados e identificados em determinado cenário, como forma de validação de futura implementação ou desenvolvimento de algo mais avançado. Dentro desta estratégia inserem-se quatro principais passos, nomeadamente Design, Desenvolvimento, Testes e Conclusões.

- **Design:** O foco será a proposta de arquitetura a seguir no passo seguinte de desenvolvimento.
- **Desenvolvimento:** Terá como principal objetivo a implementação do protótipo (MVP) e da pipeline de análise de vídeo.
- **Testes:** Incluirão uma avaliação da consistência e qualidade do código, tal como implementação em cenário real, que possibilitará a obtenção de resultados.
- **Conclusões:** Interpretação dos resultados obtidos para validação da solução.

Relativamente aos resultados e à sua avaliação, será usado o método de “Interviews” com utilizadores finais, sendo estes OPCs ou encarregados de segurança de uma outra entidade, de modo a recolher feedback e validação do funcionamento do protótipo e respetiva utilidade

do mesmo. Em adição, será também usado o método de “Observation”, onde se poderá obter e quantificar resultados e taxas de correspondência na análise automática de eventos e objetos.

Quanto à análise de dados, optou-se por uma abordagem mista que integra os dois métodos disponíveis: “Quantitative” e “Qualitative”. A análise quantitativa incidirá sobre as métricas técnicas recolhidas durante os testes do protótipo, tais como as taxas de precisão e alerta na deteção e classificação de objetos (Fase 2 e 3 da pipeline), a latência do streaming via túnel VPN, os tempos de processamento da análise de vídeo, entre outros. Já a análise qualitativa será aplicada à interpretação do feedback dos utilizadores e das observações de cenário real, avaliando o impacto operacional da ferramenta, a eficácia percebida na redução do tempo de investigação e a adequação da interface aos processos de trabalho dos OPCs.

Capítulo 4

Planeamento de Trabalho

Este capítulo descreve o planeamento deste projeto, que foi estruturado para garantir a exequibilidade dos objetivos propostos dentro do prazo académico estipulado. A organização das atividades divide-se entre a fase de preparação (unidade curricular de Preparação para Dissertação (PREPD)) e a fase de execução e escrita da dissertação (unidade curricular de Dissertação do Mestrado em Engenharia Informática (DIMEI)).

4.1 Definição do Âmbito e Entregáveis

O âmbito deste projeto foi decomposto através de um Work Breakdown Structure (WBS). Pelo mesmo, ilustrado na Figura 4.1, estão organizados os entregáveis previstos por cinco fases principais, sendo estas:

1. **Planeamento e Análise:** Focada na definição do problema e estado da arte. Inclui entregáveis como o Project Charter, o próprio WBS, um Gantt Chart, um questionário e respetivas respostas dos OPCs, este Extended Abstract e a revisão da literatura incluída no mesmo.
2. **Design:** Dedicada à modelação da solução. Os principais entregáveis previstos são a documentação da Arquitetura de Software (Vistas 4+1) e o Modelo de Domínio, assegurando que a estrutura do FUSE é robusta antes da implementação.
3. **Desenvolvimento:** Fase central do projeto, onde será implementado o código da aplicação. Divide-se em três módulos críticos: Comunicações Seguras, Camada de Abstração e Módulo de Análise de Vídeo.
4. **Testes:** Validação da solução através de testes unitários e funcionais, em conjunto com a validação do protótipo em ambiente controlado ou cenário real.
5. **Conclusões:** Considerações finais do projeto, incluindo a redação final da dissertação, a interpretação dos resultados e a apresentação final.

Figura 4.1: WBS do projeto.

4.2 Plano de Trabalho

O objetivo desta secção é ilustrar e descrever a calendarização das entregas dos respetivos entregáveis mencionados anteriormente, garantindo um acompanhamento rigoroso do

progresso do projeto. O planeamento temporal encontra-se representado no diagrama de Gantt, Figura 4.2.

Figura 4.2: Gantt Chart do projeto.

A execução do cronograma inicia-se com um primeiro marco de controlo a 15 de dezembro de 2025. O encerramento da fase de planeamento (PREPD) está estipulado para 7 de janeiro de 2026. O segundo grande bloco temporal, referente à unidade curricular de DIMEI, tem como data-alvo de entrega final o dia 22 de junho de 2026.

4.3 Project Charter e Gestão de Riscos

O Project Charter constituiu o artefacto inicial deste projeto, servindo como primeiro documento para a formalização da proposta de dissertação. A Tabela 4.1 apresenta o mapeamento dos Stakeholders, classificados quanto ao seu poder de influência e nível de interesse no sucesso do projeto, utilizando como base a “Mendelow’s Matrix” [13].

Tabela 4.1: Stakeholders identificados

Nome	Poder	Interesse	Justificação
StabilityBubble	Alto	Alto	Entidade promotora e futura fornecedora/exploradora comercial da solução FUSE.
OPCs	Alto	Alto	Utilizadores finais críticos; o seu feedback valida a utilidade operacional e requisitos forenses.
Paulo Baltarejo Sousa	Alto	Médio	Orientação académica e validação científica da metodologia e resultados.
Francisco Loureiro	Alto	Alto	Supervisão empresarial e alinhamento do produto com a estratégia de mercado.
Gestores de Segurança	Médio	Médio	Potenciais clientes empresariais que beneficiam da centralização de sistemas CCTV.

Relativamente à gestão de riscos, a Tabela 4.2 consolida os riscos identificados, atribuindo-lhes uma probabilidade e um impacto seguindo uma escala de Likert [14] de 1 a 5.

Pela análise à Tabela 4.2, destaca-se os riscos 1 e 4 como os mais elevados. A estratégia de mitigação assenta fundamentalmente na modularidade e no desenvolvimento incremental.

Tabela 4.2: Identificação e gestão de riscos associados ao projeto

ID	Descrição	P	I	P*I	Resposta
1	Indisponibilidade de Hardware (GPU)	3	4	12	Mitigar: Garantir integração modular e desacoplada.
2	Incompatibilidade de Protocolos	2	2	4	Aceitar: Focar o MVP na compatibilidade estrita com RTSP.
3	Falsos Positivos na Detecção	3	3	9	Mitigar: Implementar sistema de ajuste de sensibilidade configurável.
4	Atraso no Cronograma	3	5	15	Mitigar: Adotar abordagem de desenvolvimento incremental.

Capítulo 5

Ética

A realização deste trabalho rege-se pelos princípios de integridade, responsabilidade e rigor científico descritos no Código de Boas Práticas e de Conduta do P.PORTO [15].

Do ponto de vista profissional, o trabalho enquadra-se nos princípios estruturantes da engenharia de software, alinhados com referências internacionais como o Software Engineering Code of Ethics and Professional Practice, desenvolvido em conjunto pela Association for Computing Machinery (ACM) e pela IEEE Computer Society (IEEE-CS) [16]. A conformidade com estes princípios garante a qualidade e segurança dos sistemas desenvolvidos, a responsabilidade perante clientes e utilizadores, e a honestidade na comunicação de resultados, limitações e riscos associados às soluções tecnológicas.

O FUSE envolve uma pipeline de análise automática aplicada a contextos de videovigilância, potencialmente incidindo sobre a obtenção de dados pessoais identificáveis. Estes dados sensíveis apenas serão obtidos num contexto de validação do MVP em cenário real, que estará sempre salvaguardado por requerimento de despacho judicial para fornecimento de meios técnicos e respetiva autorização de captação de imagem. Ainda assim, a pipeline é desenhada em estrita observância do EU AI Act [17], garantindo que o sistema não incorre em “Unacceptable risks”. O sistema não se destina, nem permite, práticas proibidas tais como:

- Social scoring ou manipulação comportamental;
- Recolha indiscriminada (untargeted scraping) de imagens de CCTV para criação de bases de dados de reconhecimento facial;
- Identificação biométrica remota em tempo real em espaços públicos para fins policiais (real-time remote biometric identification).

No âmbito do último ponto, apesar do alvo ser legislar o reconhecimento biométrico em tempo-real, justifica-se a conformidade pela natureza da análise automática, que será realizada sob gravações dos eventos, e não sob o próprio streaming.

A análise focada na “Fase 3” da pipeline de deteção automática restringe-se à classificação de características objetivas (e.g., cor de vestuário, tipo de veículo) para auxiliar a pesquisa forense, mantendo sempre o princípio da validação humana, onde a decisão final cabe ao agente humano e não ao algoritmo.

Bibliografia

- [1] Honghai Liu, Shengyong Chen e Naoyuki Kubota. «Intelligent Video Systems and Analytics: A Survey». Em: *IEEE Transactions on Industrial Informatics* 9.3 (2013), pp. 1222–1233. doi: 10.1109/TII.2013.2255616.
- [2] Vlado Damjanovski. *CCTV: From light to pixels*. 3rd. Oxford: Butterworth-Heinemann, 2014. isbn: 978-0124046078.
- [3] W. Daniel Kissling et al. «Development of a cost-efficient automated wildlife camera network in a European Natura 2000 site». Em: *Basic and Applied Ecology* 79 (2024), pp. 141–152. issn: 1439-1791. doi: <https://doi.org/10.1016/j.baae.2024.06.006>. url: <https://www.sciencedirect.com/science/article/pii/S1439179124000458>.
- [4] Nozomi Networks. *New Reolink P2P Vulnerabilities Show IoT Security Camera Risks*. Acedido em: 2024. Jul. de 2021. url: <https://www.nozominetworks.com/blog/new-reolink-p2p-vulnerabilities-show-iot-security-camera-risks>.
- [5] Kaushik Ragothaman et al. «Access Control for IoT: A Survey of Existing Research, Dynamic Policies and Future Directions». Em: *Sensors (MDPI)* 23.4 (2023), p. 1805. doi: 10.3390/s23041805. url: <https://www.mdpi.com/1424-8220/23/4/1805>.
- [6] Diário de Notícias. *Falta de polícias e de viaturas e instalações e equipamentos em mau estado. Relatório aponta as falhas da PSP e GNR*. Baseado em relatório da IGA. Jun. de 2024. url: <https://www.dn.pt/sociedade/falta-de-policias-e-de-viaturas-e-instalacoes-e-equipamentos-em-mau-estado-relatorio-aponta-as-falhas-da-psp-e-gnr>.
- [7] Alberto Sampaio. «Improving Systematic Mapping Reviews». Em: *SIGSOFT Softw. Eng. Notes* 40.6 (nov. de 2015), 1–8. issn: 0163-5948. doi: 10.1145/2830719.2830732. url: <https://doi.org/10.1145/2830719.2830732>.
- [8] David Moher et al. «Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement». Em: *International Journal of Surgery* 8.5 (2010). PII: S1743-9191(10)00040-3, pp. 336–341. url: <https://www.sciencedirect.com/science/article/pii/S1743919110000403>.
- [9] Maria J. Grant e Andrew Booth. «A typology of reviews: an analysis of 14 review types and associated methodologies». Em: *Health Information and Libraries Journal* 26.2 (2009), pp. 91–108. doi: 10.1111/j.1471-1842.2009.00848.x. url: <https://onlinelibrary.wiley.com/doi/full/10.1111/j.1471-1842.2009.00848.x>.
- [10] Glenn Jocher, Ayush Chaurasia e Jing Qiu. *Ultralytics YOLO*. 2023. url: <https://github.com/ultralytics/ultralytics>.
- [11] Puneet Goswami et al. «Real-time evaluation of object detection models across open world scenarios». Em: *Applied Soft Computing* 163 (2024), p. 111921. issn: 1568-4946. doi: <https://doi.org/10.1016/j.asoc.2024.111921>. url: <https://www.sciencedirect.com/science/article/pii/S1568494624006951>.
- [12] Briony J. Oates. *Researching Information Systems and Computing*. SAGE Publications, 2006. isbn: 9781412902236. url: <https://us.sagepub.com/en-us/nam/researching-information-systems-and-computing/book226687>.

- [13] Aubrey L. Mendelow. «Environmental Scanning: The Impact of the Stakeholder Concept». Em: *ICIS 1981 Proceedings* (1981), p. 20. url: <https://aisel.aisnet.org/icis1981/20/>.
- [14] Rensis Likert. «A technique for the measurement of attitudes». Em: *Archives of Psychology* 22.140 (1932), pp. 1–55. url: <https://psycnet.apa.org/record/1933-01885-001>.
- [15] Politécnico do Porto. *Despacho P.PORTO/P-040/2020 - Código de Boas Práticas e de Conduta do P.PORTO*. 2020. url: <https://www.ipp.pt/comunidade/missao-equidade-diversidade-inclusao/DespachoP.PORTOP0402020CodigodeBoasPraticasedeConduita.pdf>.
- [16] Association for Computing Machinery. *ACM Code of Ethics and Professional Conduct*. 2018. url: <https://www.acm.org/code-of-ethics/software-engineering-code>.
- [17] European Union. *Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)*. 2024. url: <https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai>.

Apêndice A

Questionário e Entrevistas

O conteúdo deste anexo refere-se ao questionário e entrevistas realizados em conjunto com agentes da investigação criminal da GNR e da PSP para validação do problema identificado.