Art I ndet

	Service Framework Contract for the provision of a Study for "Artificial Intelligence for Automatic Detection Recognition, Identification (DRI) and Tracking - (ARTINDET)" **Executive summary**	

Summary: Executive summary

Record of Modifications

Version	Description [or description of changes]	Author	Date of Creation	Approved by	Date of Approval
1.0	First version	everisADS	03/11/2020		
2.0	Second version	everisADS	20/11/2020		
3.0	Third version	everisADS	26/11/2020		

This study was commissioned by the European Defence Agency in response to the invitation to tender No. 19.RTI.OP.177. The study does not, however, express the Agency's official views. The views expressed and all recommendations made are those of the authors.

1. Introduction

This project analyses new techniques of imagery-based systems relying on Artificial Intelligence (AI) by including some techniques based on the deep learning paradigm. For that aim, images captured by high-resolution cameras and operating with Wide Field of View (WFOV) feed the different algorithms tested. In addition, applications for Unmanned Air Vehicle (UAV) mounting high-resolutions imagery cameras cannot stream all the gathered information to the ground station in real-time. Therefore, embedded systems are crucial to allow the feasibility of the implemented technology. The outputs of the study consist of:

- Selecting, describing and analysing the most promising AI-based automatic DRIT techniques.
- Clearly and quantitatively showing the improvement in term of performances in specific selected military use cases also emphasizing the added value for the final users for the situational awareness interpretation.
- Define a roadmap for results exploitation aiming at defining new project initiatives and at developing new EU defence capabilities.

2. Objectives

The general objective of this study is to analyse the most promising AI-based techniques for Automatic Detection Recognition Identification and Tracking (A-DRIT) at least in the following operative conditions:

- Spatially distributed small targets in strong background.
- High resolution imaging systems (passive or active) with wide Field of View (Wide region surveillance).
- Multidimensional imagery (multiband/ multispectral/hyperspectral, multiresolution).
- Multiple cooperating platforms.

The specific objectives of the study are:

- Identify which AI techniques can be used for implementing A-DRIT and provide the new related image processing system schemes.
- Select some specific military use cases where AI can provide tangible benefits.
- Make a performance analysis of the selected techniques in the specific identified use cases relevant to the above listed operative conditions, in comparison with the current most common DRIT techniques.
- Provide a clear roadmap about exploitation of results for new Cat B project and new EU defence capabilities.

3. Methodology

The different steps implemented in the project define a methodology based on AI for military purposes that has been summarized in Figure 1.

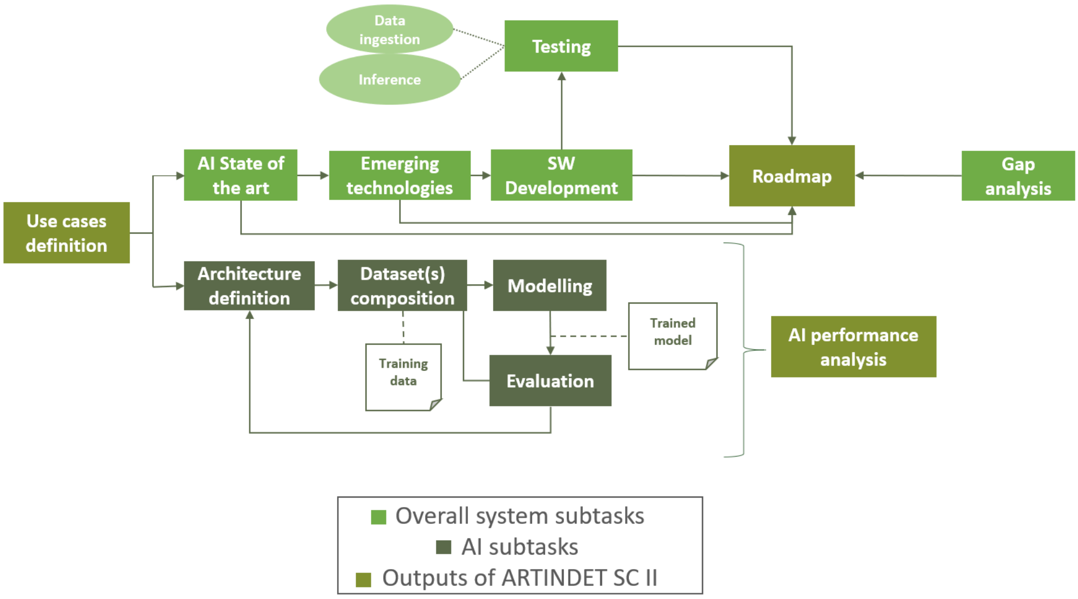

Figure 1 Methodology followed in ARTINDET

4. Project results

The project has gathered results in seven areas:

- **Use cases** (*D2 Definition and description of military use cases*)
- **Dataset composition** (*D4 Performance analysis in the selected use cases*)
- **Detection and recognition** (*D4 Performance analysis in the selected use cases*)
- **Tracking** (*D4 Performance analysis in the selected use cases*)
- **Re-Identification** (*D4 Performance analysis in the selected use cases*)
- **Gap definition** (*D5 Results exploitation and roadmap definition*)
- **Roadmap** (*D5 Results exploitation and roadmap definition*)

4.1 Use cases

Two use cases have been defined with the collaboration of different stakeholders that participated in the workshop gathered. They are specified in Table 1.

	MISSION	SCENARIO	PLATFORM	SENSOR	TARGET
UC1	Situation awareness	ROAD /URBAN	Small UAV	EO, IR	People
UC2	Camp protection	ROAD / RURAL	Small UAV	EO, IR	Car / Trucks / Tanks

Table 1 Use cases specification

4.2 Dataset composition

Different natures of datasets have been explored for all the stages included in the A-DRIT system. It shall be remarked the lack of available military datasets, especially in the infrared domain where the generic available datasets are less frequent. For that reason, custom datasets specific for the military domain have been built. The dataset for use case 1 relies on a simulation tool to generate synthetic data in the visible domain, while in the infrared domain images acquired by Everis ADS are merged together with images from a publicly available dataset. For the vehicle detection, which is the use case 2, the dataset for the visible domain relies on video sequences published in YouTube that have been adapted and ingested to be enrolled in the corresponding training procedures and the composition of the infrared dataset follows the same scheme as in use case 1.

4.3 Detection and recognition

Different approaches for transfer learning have been evaluated and the influence of the balance of the data included in the training step has been evaluated (see Table 2). The Faster RCNN is the detector that yields the best results for use case 1 and they confirm that this network usually takes more iterations to achieve sufficient accuracy. The results obtained for use case 2 and Single Stage Detector (SSD) are very significant among all the results gathered in this stage. In addition, from the IR datasets provided by Everis ADS different approaches have been tested, providing more reliable results for the solution of the use case 1.

Architecture	Use Case	Domain	F1-score (%)
Faster RCNN	UC1	Visible	45.21
SSD	UC2	Visible	71.33
SSD	UC1, UC2	Infrared	94.05

Table 2 Summary of best results obtained for detection and recognition in terms of F1-Score (%) per use case and domain.

4.4 Tracking

Two approaches have been tested for the tracking stage and a compilation of the best results can be observed in Table 3. The first is a classical tracking algorithm and it is based on the best model learned for the detection stage by the SSD architecture. It offers the most reliable results for use case 2, since the detector trained for this use case offered significant results. The second is based in a joint detection and tracking and offers the best results for the solution of the use case 1. However, the range obtained with the corresponding metrics is still far from the results gathered for use case 2.

Tracking method	Use case	AP (%)	Recall (%)
Kernel-based (KCF)	UC1	46	60
Kernel-based (MOSSE)	UC2	73	89
Joint	UC1	61.46	70

Table 3 Summary of best results obtained for tracking in terms of AP (%) and Recall (%). In the case of kernel-based tracking, results are given every 5 frames.

4.5 Re-identification

One baseline solution for human re-identification, and one Deep Learning-based solution relying on spatial temporal features have been tested. The baseline solution was aimed to test different backbones, and the second solution offered the most reliable results for both use case 1 (with a public dataset) and use case 2 (with a custom dataset built).

ReID method	Use Case	Top 1 (%)	Top 5 (%)	Top 10 (%)	mAP (%)
Baseline (PCB)	UC1	89.6	-	-	74.5
Baseline (PCB)	UC2	75.95	75.95	75.95	75.95
Spatio temporal	UC1	98.0	98.9	99.1	95.5
Spatio temporal	UC2	74.32	75.96	75.96	75.24

Table 4 Summary of best results obtained for re-identification in terms of mAP (%).

4.6 Full performance of the processing chain

The different tasks of the A-DRIT chain are independently studied in the previous sections. Based on these results, a proposal to test the full performance of the processing chain is presented in Table 5. For the UC1, it is important to consider the impact of the different sizes and aspect ratio on the results and it might be required a pre-processing step. By creating custom datasets for UC2, results are proven to be more robust and reliable in the independent steps.

Use Case	Detection	Tracking	Re-Identification
UC1	Faster RCNN	Joint	Spatio-Temporal
UC2	SSD	Kernel-based (MOSSE)	Baseline (PCB)

Table 5 Proposal for processing chain given the results independently studied.

In addition, the values of the metrics selected for the best results for the different tasks tested in the visible domain (the whole processing chain in the IR domain cannot be analysed due to the unavailability of data) are exposed in Table 6.

Use Case	Detection	Tracking		Re-Identification
	F1-Score (%)	mAP (%)	Recall (%)	mAP (%)
UC1	45.21	61.46	70	95.5
UC2	70.11	73	89	75.24

Table 6 Summary of the most relevant metrics for the different tasks in the use cases selected.

It can be observed that for UC1 the value of the metric F1-Score selected for the detection is the lowest of the results presented. The detection stage has been joint to the tracking in the solution presented in the Subsection **Error! Reference source not found.** improving the values obtained until a 61.46% of mAP and 70% of recall for the sequence with the best performance. However, this still is a value not significant to deploy the DRIT system in a real scenario. Even more, considering the difference in the values of the metrics for the sequences tested (See Section 4 for more details). Regarding the Re-Identification of UC1, this type of target is the most studied in the literature queried, which derives in the best results obtained for this task with a 95.5% of mAP. It can be concluded that a least a 70% should be achieved for the metrics selected for the different tasks to deploy the system for this use case.

The best results for the detection and tracking tasks are obtained by UC2. The values of all the metrics selected are above 70% which makes them attractive to perform a deployment in a real scenario. The accuracy of the whole chain should be at least around this value for the metrics selected (F1-Score, mAP and Recall), although a more intensive testing is required to adapt the solution proposed to the target scenario and the conditions that should be faced.

4.7 Gap definition

An analysis of the DRIT systems in the military domain has been performed, and the following areas have been covered:

- Business
- Data
- Application

- Technology

From that analysis different areas of interest such as analysis of optical features, optimization of supervised algorithms, Human-Machine Interfaces, deployment in operative field… have emerged. These areas are the seed to build the roadmap that is outlined in the next subsection.

4.8 Roadmap

The roadmap built has been divided in two terms:

- Short-term
- Mid-term

The first is located on a lower TRL (3-5), and its objective is to perform a proof of concept. The second is located on a higher TRL (4-6) and aims to cover all the functionalities and perform field tests.

The step challenges defined for them are enumerated below:

- Short-term
  - Dataset definition and classification in the military domain
  - Applying supervised AI for target detection and tracking in military scenarios
  - Human-machine interaction in the military domain
  - Real unmanned systems with AIDRIT capabilities
- Mid-term
  - Human-machine interaction in the military domain
  - Real unmanned systems with AIDRIT capabilities
  - AI Certification for ADRIT purposes

5. Conclusions

The study performed in ARTINDET has provided a more precise knowledge regarding the feasibility and the steps required to adopt A-DRIT systems in the military domain. The project has gone through all the stages required for this kind of analysis:

- **State of the art**
- **Use case definition**
- **Artificial Intelligence emerging technologies and techniques**
- **Performance analysis**

Exploitation and roadmap

The main outcomes of the project are:

Definition of two use cases

Two use cases (UC1 and UC2) are defined base on stakeholder consultation via an ad hoc workshop and questionnaires. UC1 is defined for situation awareness in a road/urban environment and the target is person detection while UC2 is defined for camp protection and as the target are cars, trucks and tanks, the environment is road/rural. Both use cases are stablished for visible and infrared domains (see section 4.1). The summary of the use cases can be noticed in Table 7.

	MISSION	SCENARIO	PLATFORM	SENSOR	TARGET
UC1	Situation awareness	ROAD /URBAN	Small UAV	EO, IR	People
UC2	Camp protection	ROAD / RURAL	Small UAV	EO, IR	Car / Trucks / Tanks

Table 7 Outputs of ARTINDET SC2: use cases.

Database construction

It is analysed in section 4.2, where it can be seen how the lack of images of the specific use cases may be an issue and different strategies must be followed to compose the right datasets. These strategies rely on merging the part of public datasets that fulfil with the requirements of the use cases, acquiring additional images with own means, generating synthetic images and extracting and annotating public images. The summary of the datasets built can be observed in Table 8.

Dataset name	Use Case	Classes	Dataset composition
v3.1	UC1	person	Stanford, Visdrone, YouTube
v3.2	UC1	person	Stanford, Visdrone.
v2.1	UC2	car, truck, tank	DOTA, VEDAI, Visdrone, YouTube, Simulated
v2.2	UC2	car, truck	DOTA, VEDAI, Visdrone

Table 8 Outputs of ARTINDET SC2: Datasets built.

Feasibility of the different stages required for the use cases

The goal of small targets in strong background is fulfilled by letting the algorithms learn from images of public dataset with these requirements and the evaluation is performed in the infrared domain as it is the architecture with best results, which achieves F1-score = 83.31% for the best model in Table 2. All images are full HD, so the use of High-resolution imaging systems is accomplished, and they are also multidimensional imagery, multiband in particular as the agreed use cases state.

- **Selection of most relevant techniques for A-DRIT purposes and performance analysis**

  For each task in the A-DRIT chain, different methods are applied and their behaviour on this particular environment are analysed for each use case. They are summarized in Table 9, Table 10 and Table 11.

  - **Detection**

Architecture	Use Case	Domain	F1-score (%)
Faster RCNN	UC1	Visible	45.21
SSD	UC2	Visible	71.33
SSD	UC1, UC2	Infrared	94.05

Table 9 Outputs of ARTINDET SC2: Detection results.

  - **Tracking**

Tracking method	Use case	AP (%)	Recall (%)
Kernel-based (KCF)	UC1	46	60
Kernel-based (MOSSE)	UC2	73	89
Joint	UC1	61.46	70

Table 10 Outputs of ARTINDET SC2: Tracking results.

  - **Re-identification**

ReID method	Use Case	Top 1 (%)	Top 5 (%)	Top 10 (%)	mAP (%)
Baseline (PCB)	UC1	89.6	-	-	74.5
Baseline (PCB)	UC2	75.95	75.95	75.95	75.95
Spatio temporal	UC1	98.0	98.9	99.1	95.5
Spatio temporal	UC2	74.32	75.96	75.96	75.24

Table 11 Outputs of ARTINDET SC2: Re-identification results.

- **Selection of technologies and UAV architecture proposal**

  Based on the results independently obtained for each step, a proposal of the best techniques to use in the full processing chain is presented in Table 12.

Use Case	Detection	Tracking	Re-Identification
UC1	Faster RCNN	Joint	Spatio-Temporal
UC2	SSD	Kernel-based (MOSSE)	Baseline (PCB)

Table 12 Outputs of ARTINDET SC2: Selection of DRIT techniques.

When implementing this proposal, it is important to bear in mind that UC1 might require an additional pre-processing step due to high variation in size and aspect ratio of targets in the datasets. For UC2, the custom datasets introduce a factor of reliability and robustness in the results of the independent steps.

- **Next steps to be carried out for A-DRIT systems in the military domain.** A roadmap in the scope of DRIT systems in the military domain has been defined based on the work performed in ARTINDET SC 1 and SC2. This proposal covers the areas that require a higher effort in a short and mid term perspective by analysing the business, data, application and technology gaps in the ambit defined by this project. The result includes a budget proposal, and the funding sources available to realize the corresponding projects. The summary of challenges and aims included in the roadmap are summarized in Table 13.

Challenge	Aim
Dataset definition and classification in the military domain	Methodology to select relevant visual features for military datasets
	Real military dataset compliance
	Simulated and synthetic military data generation
Applying supervised AI for target detection and tracking in military scenarios	Deep learning modelling for different use cases and automatic updating for new features in the operational scenario
	AI deployment in embedded systems for real-time inference in different operational scenarios
Human-machine interaction in the military domain	Human-machine interface development for AIDRIT applications and user-training
	Semi-automatic systems based on decision-making to manage operations
Real unmanned systems with AIDRIT capabilities	RPAS operative missions
	Integration of smart unmanned systems in the military scope with DRIT purposes
	Field tests in the military scope
AI Certification for ADRIT purposes	Taxonomy definition
	Content definition

Table 13 Outputs of ARTINDET SC2: New challenges and aims.

6. Acronyms

- **A-DRIT** - Automatic Detection, Recognition, Identification and Tracking
- **AP** - Average Precision
- **AI** - Artificial Intelligence
- **DRIT** - Detection, Recognition, Identification and Tracking
- **EO** - Electro-Optical
- **EU** - European Union
- **F1-Score** - Harmonic mean of precision and recall
- **HD** - High Definition
- **IR** – Infrared
- **mAP** - Mean Average Precision
- **SC** - Specific Contract
- **SSD** - Single Stage Detector
- **TRL** - Technology Readiness Level
- **UAV** - Unmanned Air Vehicle
- **UC** - Use Case
- **WFOV** - Wide Field of View