



PID2021-125051OB-I00 HVD (2022-2025)

*Harvesting Visual Data: enabling computer vision in unfavourable data
scenarios*

D3.2v1

Results Report

Video Processing and Understanding Lab

Escuela Politécnica Superior

Universidad Autónoma de Madrid

Supported by



AUTHORS LIST

<i>José M. Martínez</i>	josem.martinez@uam.es
<i>Juan C. SanMiguel</i>	Juancarlos.sanmiguel@uam.es
<i>Álvaro García Martín</i>	alvaro.garcia@uam.es
<i>Marcos Escudero Viñolo</i>	Marcos.escudero@uam.es
<i>Pablo Carballeira López</i>	pablo.carballeira@uam.es

HISTORY

Version	Date	Editor	Description
0.1	20/08/2023	José M. Martínez	First Draft version for contributions
0.2	07/09/2023	Juan C. SanMiguel	Contributions
0.3	11/09/2023	Álvaro García Martín	Contributions
0.4	12/09/2023	Marcos Escudero-Viñolo	Contributions
0.5	15/09/2023	Pablo Carballeira López	Contributions
0.6	18/09/2023	José M. Martínez	Working Draft edition
0.7	27/09/2023	José M. Martínez	Final Working Draft
1.0	30/09/2023	José M. Martínez	Editorial checking

EXECUTIVE SUMMARY

This Document compiles the results of the project during the first 9 months. It is structured following the sections of the “INFORME CIENTIFICO-TECNICO INTERMEDIO – Proyectos de Generación de Conocimiento 2021”.

CONTENTS:

1. A. DATOS DEL PROYECTO.....	3
2. B. PERSONAL ACTIVO EN EL PROYECTO	3
3. C. PROGRESO Y RESULTADOS DEL PROYECTO	5
4. D. DIFUSIÓN DE LOS RESULTADOS DEL PROYECTO	17

1. A. Datos del proyecto

Relacione los datos actuales del proyecto. En caso de que haya alguna modificación, indíquelo en la casilla A2

A1. Datos del proyecto	
REFERENCIA: PID2021-125051OB-I00	
TITULO: RECOLECCION DE DATOS VISUALES: PERMITIENDO LA VISION POR COMPUTADOR EN ESCENARIOS CON DATOS DESFAVORABLES (HVD – Harvesting Visual Data: enabling computer vision in unfavourable data scenarios)	
Modalidad	Proyectos de investigación orientada
Área/Subarea	Tecnologías de la Información y de las Comunicaciones / Tecnolgías de las Comunicaciones
Prioridad temática*	Clima, energía y movilidad
IP1	José María Martínez Sánchez
Código Orcid:	0000-0002-2236-1769
IP2 (si procede)	Álvaro García Martín
Código Orcid:	0000-0002-1705-3972
Entidad Beneficiaria	Universidad Autónoma de Madrid
Centro	Escuela Politécnica Superior – Depto. Tecnologíass Eleltróncia y d elas Comunicaciones – Video Processing and Understadnandin gLab (VPULab)
Fecha de inicio	01/09/2022
Fecha final	31/08/2025
Duración	36 meses
Total concedido (costes directos)	97.000,00 €

*Para modalidad Investigación Orientada

A2. Descripción de modificaciones en los datos iniciales del proyecto (*Cambio de IP, entidad, centro, modificación del periodo de ejecución...*)

Ningún cambio por el momento.

2. B. Personal activo en el proyecto

Tiene que relacionar la situación de todo el personal de las entidades participantes que haya prestado servicio en el proyecto en el periodo que se justifica, o que no haya sido declarado anteriormente, y cuyos costes (dietas, desplazamientos, etc.) se imputen al mismo.

B.1. Equipo de investigación

<i>Incluido en la solicitud original</i>					
	Nombre	NIF/NIE	Función en el proyecto	Fecha de baja	Observaciones
1	Jesús Bescós Cano	33.507.77 6-L	Resto de equipo		
2	Pablo Carballeira López	53.178.56 1-P	Resto de equipo		
3	Marcos Escudero Viñolo	02.267.76 9-S	Resto de equipo		
4	Miguel Ángel García García	39.686.96 9-D	Resto de equipo		Tiempo parcial
5	Álvaro García Martín	46.920.26 4-B	IP2		
6	José María Martínez Sánchez	51.380.80 9-M	IP1		
7	Juan Carlos San Miguel Avedillo	70.070.73 9-C	Reso de equipo		

<i>No incluido en la solicitud original</i>						
	Nombre	NIF/NIE	Función en el proyecto	Fecha de alta	Fecha de baja	Observaciones
1						
2						

B.2. Personal incluido inicialmente en el Equipo de Trabajo, o incorporado posteriormente, que no cumpla alguno de los requisitos para formar parte del equipo de investigación

	Nombre	NIF/NIE	Función en el proyecto	Inicio	Fin	Observaciones
1	Roberto Alcover Couso		Personal investigador en formación			Incluido inicialmente en equipo de trabajo. Beca FPI-UAM desde 01/03/2023
2	Javier Montalvo Rodrigo		Personal investigador en formación			Incluido inicialmente en equipo de trabajo (tiempo parcial).

						Ayudante desde 01/09/2023
3	Juan Ignacio Bravo Pérez-Villar		Personal investigador en formación	01/09/2022		Doctorado Industrial IND2020/TIC-17515
4	Paula Moral de Eusebio		Personal investigador en formación	01/09/2022		Beca FPI-UAM
5	Kirill Sirotkin		Personal investigador en formación	01/09/2022		Beca FPI-UAM
6	Rosa Graciela Guerrero Idrovo		Técnico de Apoyo para la gestión, mantenimiento y actualización de las infraestructuras de computación. TS1 (T1.1)	11/04/2023	11/08/2023.	Contrato con cargo al proyecto
7	Cecilia Diana Albelda		Personal investigador en formación	01/09/2023		Doctorado a tiempo parcial

Nota: Cree tantas filas como necesite.

La solicitud de “Altas” y “Bajas” de personal investigador en el **equipo de investigación** debe tramitarse de acuerdo con **las instrucciones de ejecución y justificación** expuestas en la página web de la convocatoria. La incorporación de personal que participe en el proyecto en el **apartado B.2** no necesita autorización por parte de la AEI, pero su actividad debe incluirse y justificarse en este informe.

3. C. Progreso y resultados del proyecto

C1. Desarrollo de los objetivos específicos planteados	
<i>Describe los objetivos específicos y el grado de cumplimiento de los mismos (porcentaje estimado respecto al objetivo específico planteado y, en su caso, indique lo que queda por realizar en cada uno de ellos).</i>	
Objetivo 1: O1. To maintain and upgrade the system infrastructure: storage (large datasets), communication and computing infrastructures. (T1.1): this objective considers the infrastructure required for achieving the objectives and goals of the project, and the development of the associated software for deep learning algorithms.	<i>Progreso y consecución del objetivo 1:</i> Se ha adquirido y puesto en marcha nueva infraestructura para computación de alto rendimiento (T1.1) que permite entrenar modelos complejos para las tareas de análisis y simulación del proyecto. Se actualizado la documentación existente con estas adquisiciones en el deliverable D1.1v1: “System infrastructure”. <i>Grado de cumplimiento:</i> 65%

	<p><i>Pendiente:</i> actualmente se encuentra en fase de desarrollo un sistema de gestión de colas tipo SLURM para que los investigadores puedan emplear todos los recursos de la infraestructura de manera unificada. Adicionalmente, también se pretende desarrollar una metodología de desarrollo software, que permita unificar los trabajos realizados en el proyecto.</p>
<p>Objetivo 2: O3. To collect/generate datasets and an evaluation methodology. (T1.2): this objective consists of the support throughout the project for collecting and providing test and train data, which can be used for the developed approaches in the Research and Innovation objectives and for the case studies.</p>	<p><i>Progreso y consecución del objetivo 2:</i> se han recopilado diversos <i>datasets</i> y <i>challenges</i> relevantes para el proyecto que se están usando en las diversas tareas del WP2. Adicionalmente se han generado nuevos <i>datasets</i> que están pendientes de publicación. El D1.2 se ha retrasado hasta enero 2024 para incorporar los nuevos <i>datasets</i>.</p> <p><i>Grado de cumplimiento: 30%</i></p> <p><i>Pendiente:</i> la publicación de D1.2v1 y de los <i>datasets</i> generados. Continuar con la recopilación y generación de <i>datasets</i> en base a las necesidades de WP2.</p>
<p>Objetivo 3: O3. To collect/implement state-of-the-art image and video analysis tools for deep learning algorithms. (T2.1, T2.2, T2.3): this objective identifies suitable supervised and unsupervised state-of-the-art image and video analysis tools (e.g., semantic segmentation, scene recognition, tracking, re-identification), implements and evaluates their performance as a baseline, for a comparison with the achievements in WP2.</p>	<p><i>Progreso y consecución del objetivo 3:</i> se han identificado, implementado y evaluado algoritmos supervisados y no supervisados para segmentación semántica en video-juegos [J1], segmentación semántica entornos urbanos [J3][C2][O3], detección y clasificación de elementos urbanos [O4], detección y estimación de pose [J4][O1], seguimiento de objetos [C1], reidentificación de vehículos [J2][C4], estimación de profundidad en entornos urbanos [C3].</p> <p><i>Grado de cumplimiento: 40%</i></p> <p><i>Pendiente:</i> actualmente se continúa trabajando en identificar soluciones del estado del arte tanto supervisadas como no supervisadas. Además de continuar con las líneas ya iniciadas, se está trabajando en algoritmos para reidentificación de elementos urbanos, segmentación semántica en imagen médica y segmentación semántica de vegetación en entornos urbanos.</p>
<p>Objetivo 4: O4. To manage the project and disseminate the project results. (T3.1, T3.2, T3.3): this objective covers all the tasks</p>	<p><i>Progreso y consecución del objetivo 4:</i> este objetivo progresa adecuadamente, realizando las labores regulares de gestión del proyecto, así como de diseminación. Se han incorporado tres nuevos EPOs (Deimos, Sener,</p>

<p>associated to the management and dissemination of results (publications, public reports, organization of workshops, industrial days, and newsletters).</p>	<p>Servicio de Radiodiagnóstico del Hospital Universitario La Princesa/Facultad de Medicina UAM).</p> <p><i>Grado de cumplimiento: 33%</i></p> <p><i>Pendiente: continuar con las actividades regulares de seguimiento y diseminación.</i></p>
<p>Objetivo 5: O5. To analyze relevant use-cases and develop technology transfer initiatives. (T3.4): this objective aims to clearly define the selected use cases of the project based on the results achieved by WP2. This objective also comprehends the development of technology transfer initiatives with our Observing Partners or external companies.</p>	<p><i>Progreso y consecución del objetivo 5</i> La tarea asociada a este objetivo no se inicia hasta la segunda anualidad.</p>
<p>Objetivo 6: R1. To quantify the feasibility of adapting or reusing available models for new data-scarce scenarios. (T2.1): in this objective, we will explore the learning capabilities of unsupervised learning regimes, searching for the hypothetical learning limit of a given architecture when trained with unlabeled images of a given domain. We plan to explore different self-supervised alternatives in order to reduce or mitigate the gap between supervised vs unsupervised learning capabilities.</p>	<p><i>Progreso y consecución del objetivo 6:</i> se han desarrollado esquemas de aprendizaje no supervisado o levemente supervisado para segmentación semántica en entornos urbanos [J3], adaptación de dominio no supervisada para detección y estimación de pose [J4][O1], aprendizaje auto-supervisado para estimación de profundidad en entornos urbanos [C3], adaptación de dominio no supervisada para reidentificación de vehículos [C4].</p> <p><i>Grado de cumplimiento: 40%</i></p> <p><i>Pendiente:</i> actualmente se continúa trabajando en soluciones no supervisadas o auto supervisadas. Además de continuar con las líneas ya iniciadas, se está trabajando completar desarrollos basados en curriculum y self-paced learning para segmentación semántica en entornos urbanos.</p>
<p>Objetivo 7: R2. To develop continual lifelong learning alternatives for scenarios with few available data or incremental data requirements. (T2.1 and T2.2): this objective focuses on strategies to learn representations from stationary batches of training data, accounting for situations in which information becomes incrementally available over</p>	<p><i>Progreso y consecución del objetivo 7:</i> se ha implementado una solución basada en aprendizaje por refuerzo para segmentación semántica en video-juegos [J1], se ha implementado una solución basada en aprendizaje adversario auto supervisado para la estimación de profundidad en entornos urbanos [C3] [O1], robusto a cambios en la distancia focal de la cámara.</p>

<p>time. We plan to explore three different continual lifelong learning alternatives: incremental domain learning, incremental class learning and incremental task learning.</p>	<p><i>Grado de cumplimiento: 30%</i></p> <p><i>Pendiente:</i> actualmente se continúa trabajando en soluciones incrementales. Además de continuar con las líneas ya iniciadas, se está trabajando en el aprendizaje continuado para la detección y clasificación de elementos urbanos.</p>
<p>Objetivo 8: R3. To generate realistic simulations of relevant use-cases for the training of computer vision algorithms. (T1.2 and T2.2): we plan to use the Unity3D framework to generate complete virtual environments in order to automatically generate ground-truth data for the detection, segmentation, and tracking of multiple moving objects. We will also consider different configurations by the control of the environmental conditions, camera types (moving and static) and virtual objects.</p>	<p><i>Progreso y consecución del objetivo 8:</i> se han desarrollado nuevas metodologías de evaluación [C1], y se han desarrollado varios sistemas basados en Unity3D y Unreal para simulación de captación de datos visuales en los siguientes casos: vehículos autónomos en entornos urbanos [J3][C3][O1][O2], y en el ámbito de videojuegos [J1]. Adicionalmente se ha comenzado a explorar algoritmos generativos basados en difusión para obtener datos sintéticos de manera controlada [O3].</p> <p><i>Grado de cumplimiento: 60%</i></p> <p><i>Pendiente:</i> completar desarrollos iniciados para la generación de un dataset para estimación de pose en satélites, generación de un dataset sintético para segmentación semántica utilizando LIDAR y esquemas de explicabilidad que permitan controlar el dato sintético obtenido por algoritmos generativos.</p>
<p>Objetivo 9: R4. To quantify the advantages of adapting models trained using synthetic data to real data-scarce scenarios. (T2.2): we plan to explore the use of synthetic training data to increase the performance of computer vision algorithms, e.g., detection, segmentation, re-identification, tracking. We plan to explore two different learning alternatives: curriculum learning and domain adaptation.</p>	<p><i>Progreso y consecución del objetivo 9:</i> se han desarrollado esquemas de entrenamiento sintético-real para segmentación semántica en video-juegos [J1] y entornos urbanos [J3]; y reidentificación de vehículos [J2]. También se han explorado esquemas de adaptación de dominio sintético-real para estimación de poses [J4], reconocimiento de anomalías [C2] y re-identificación de vehículos [C4]. Adicionalmente, se han comenzado desarrollos basados en curriculum y self-paced learning para segmentación semántica.</p> <p><i>Grado de cumplimiento: 60%</i></p> <p><i>Pendiente:</i> Completar desarrollos basados en curriculum y self-paced learning para segmentación semántica en entornos urbanos. Explorar el entrenamiento de modelos supervisados utilizando datos obtenidos por modelos generativos.</p>
<p>Objetivo 10: R5. To design and evaluate profiling methods to</p>	<p><i>Progreso y consecución del objetivo 10:</i> se han estudiado y extraído las fronteras de decisión de modelos</p>

<p>explain model predictions and detect and correct potential social biases of trained models. (T2.3): this objective aims to define methodologies for the visualization of DL models, focusing on aiding human interpretation of their operation in order to: detect model biases, help the foretelling of their behavior in real-world applications, and design effective training mechanisms in accordance with their knowledge encoding processes.</p>	<p>entrenados usando técnicas de aprendizaje auto-supervisado [O9].</p> <p><i>Grado de cumplimiento: 20%</i></p> <p><i>Pendiente:</i> Completar los esquemas de análisis de espacios de decisión y suministrar explicaciones cualitativas y cuantitativas de estos. Detectar y corregir o minimizar sesgos potenciales, Mejorar el aprendizaje auto-supervisado de acuerdo con la distribución de datos.</p>
---	--

Nota: Cree tantas filas como necesite

<p>C2. Actividades realizadas y resultados alcanzados <i>Describe las actividades científico-técnicas realizadas para alcanzar los objetivos planteados en el proyecto. Indique para cada actividad los resultados alcanzados y los miembros del equipo que han participado. Extensión máxima 2 páginas</i></p>	
<p>Actividad 1 (T1.1: Infrastructure update and maintenance (M1-M28): Arrangement and configuration of the available equipment, and the acquisition of complementary equipment, for establishing the necessary infrastructure to meet the objectives of the project (storage, communication, and computing infrastructures).</p>	<p>Miembros del equipo participantes*: JCSM, AGM, TS1, all Adquisición y configuración de nuevo equipamiento de investigadores y para procesado HPC basado en GPU de imagen y vídeo.</p>
<p>Actividad 2: T1.2: Design and generation of datasets (M5-M32) Support to other tasks for generating test data and defining evaluation methodologies. It includes the selection of appropriate datasets (sequences and associated ground-truth) and their generation, if required, as well as the research, selection, and proposal of evaluation metrics and benchmarks.</p>	<p>Miembros del equipo participantes*: JMS, all Se han recopilado diversos <i>datasets</i> y <i>challenges</i> relevantes para el proyecto que se están usando en las diversas tareas del WP2. Adicionalmente se han generado nuevos <i>datasets</i> que están pendientes de publicación. El D1.2 se ha retrasado hasta enero 2024 para incorporar los nuevos <i>datasets</i>.</p>
<p>Actividad 3: T2.1: Real data in absence of annotations (M1-M32) To quantify the impact of new learning-based methods in absence of large human-annotated real datasets. In this direction, we propose: (1)</p>	<p>Miembros del equipo participantes*: AGM, TS1, all Desarrollo y evaluación de algoritmos supervisados y no supervisados [J3][J4][C4]. Desarrollo de mecanismos de aprendizaje auto supervisados e incrementales [J1][C3][O1].</p>

<p>searching for the learning capabilities of different unsupervised and self-supervised learning alternatives in order to reduce the gap between supervised vs unsupervised learning capabilities; (2) studying the use of multiple complementary self-supervised pretext tasks of different nature, and assess their contribution to the knowledge acquired by the trained model; (3) evaluating the impact of data in self-supervised learning regimes, analyzing the effect of the size, diversity and noise of the data in the training of self-supervised learning models, and quantifying their impact in the knowledge encoded by the trained model; (4) exploring situations in which information becomes incrementally available over time with three different continual lifelong learning alternatives: incremental domain learning, incremental class learning and incremental task learning.</p>	
<p>Actividad 4: T2.2: Creation and use of synthetic data (M1-M32) To explore the creation and use of synthetic datasets to complement the training process. Jointly with the existing real datasets, the synthetic data will provide diversity in the training process helping to obtain more robust visual models. In this direction, we propose: (1) the automatic generation of complexity-variable scenarios, associated synthetic data and ground-truth for detection, tracking and semantic segmentation tasks; (2) exploring supervised and weakly supervised domain adaptation approaches for video object tracking and semantic segmentation tasks by analyzing adversarial strategies; (3) to develop new unsupervised domain adaptation approaches for video object</p>	<p>Miembros del equipo participantes*: JCSM, MEV, AGM, JBC, PCL, all</p> <p>Simulación de vehículos autónomos en entornos urbanos [J3][C3][O1][O2], y en el ámbito de videojuegos [J1]. Explicabilidad de algoritmos generativos basados en difusión para obtener datos sintéticos [O3].</p>

<p>tracking and semantic segmentation tasks based on distance metrics, mutual information and self-training strategies.</p>	
<p>Actividad 5: T2.3: Model profiling (M5-M32) To explore methodologies for the visualization of DL models, focusing on aiding human interpretation. Utilize model profiling tools to scrutinize the knowledge encoded in self-supervised trained modes in order to: (1) detect and mitigate possible biases; (2) improve the self-supervised learning process considering the characteristics of the target visual task and target data distribution; (3) provide qualitative and quantitative explanations of the model predictions by measuring the attribution of the model predictions on the target data.</p>	<p>Miembros del equipo participantes*: MEV</p> <p>Extracción del margen decisión de modelos entrenados mediante la adición de ruido a las muestras y búsqueda de cambio en la clasificación. Extracción de distribución estadística del margen para muestras correcta e incorrectamente clasificadas y varios modelos auto-supervisados, encontrando diferencias entre los modelos analizados [O9].</p>
<p>Actividad 6: T3.1 Management (M1-M36) This task has the following activities: monthly follow up of Project progress and achievements, workplan milestones and outcomes deadlines control, workplan updates, corrective actions, and administrative issues.</p>	<p>Miembros del equipo participantes*: AGM, JMS; all Se realiza un seguimiento semanal entre los IPs y reuniones de seguimiento mensuales para mantener al equipo al tanto. Se han realizado pequeños ajustes de asignación de tareas. En Junio 2023, fecha prevista de su publicación, D3.1 y D3.2v1 se retrasaron a Septiembre. D3.1 pasa a ser un documento con versiones (para cada nuevo <i>dataset</i>). D</p>
<p>Actividad 7: T3.2 Communication (M1-M36) This task coordinates the communication of the project advances to Observing Partners and general public. A web site will include a description of the project, status, advances, and announcements, as well as the biannual HVD Newsletters.</p>	<p>Miembros del equipo participantes*: AGM, JMS, all Durante esta primera anualidad se ha actualizado regularmente la página web, se han publicado dos <i>Newsletters</i> y se ha informado a los EPOs de los avances del proyecto. Se han incorporado 3 nuevos EPOs: Sener en Abril 2023, Deimos en Marzo 2023 y el Servicio de Radiodiagnóstico del Hospital Universitario La Princesa (Facultad de Medicina UAM) en Julio 2023.</p>
<p>Actividad 8: T3.3 Dissemination (M1-M36) This task coordinates the compilation and internal publication of intermediate results, in order to have a clear plan for dissemination based on a commitment for bi-monthly reports' updates that will constitute seeds for</p>	<p>Miembros del equipo participantes*: JMS, PCL, MEV; all Se ha publicado el D3.2v1 (este documento). Durante este primer año se han realizado 4 publicaciones en revista, 4 publicaciones en congresos internacionales y 5 contribuciones software en github. Se ha elaborado y publicado el <i>Data Management Plan</i> (DMP) (D3.1v1), que se irá actualizando (frente a la idea</p>

<p>publications, as well as the analysis of target journals and conferences for the different works in progress. It will also coordinate the dissemination of results via Research Reports, GitHub software repositories and generated datasets, to be made available at the project's web page, via links to open repositories. This task also handles the Data Management plan</p>	<p>de un único documento) con cada nuevo <i>dataset</i> publicado. El DMPv1 contiene las características de los sistemas de almacenamiento, seguridad y acceso que se utilizan para los <i>datasets</i> generados en el proyecto. Además, incluye una plantilla para detallar las características específicas de formatos, privacidad y consideraciones éticas específicas a cada <i>dataset</i> generado en el proyecto.</p>
<p>Actividad 9: T3.4 Use cases and technology transfer (M12-M36) Using WP2 outcomes, use cases will be defined with the help and recommendations of the Observing Partners (as described in section 4.3). Information meetings will be held with the Observing Partners and up to two Workshops, open to other companies and interested parties, will be held in the form of Industry Days. This task also considers, during the last year, the implementation of the software required for each use case (Minimum Viable Product prototype) as a tool for boosting the technology transfer initiatives with our Observing Partners and other stakeholders.</p>	<p>Miembros del equipo participantes*: AGM, JMS, TS4, all</p> <p>Esta tarea no ha comenzado durante la primera anualidad.</p>

Notas: Cree tantas filas como necesite. En caso de incluir figuras, cítelas en el texto e insértelas en la última página
*Resalte en negrita las actividades realizadas por las personas que son IP.

<p>C3. Problemas y cambios en el plan de trabajo <i>Describe las dificultades y/o problemas que hayan podido surgir durante el desarrollo del proyecto. Indique cualquier cambio que se haya producido respecto a los objetivos o el plan de trabajo inicialmente planteado, así como las soluciones propuestas para resolverlos. Extensión máxima 1 página</i> <i>*Se recuerda que la aceptación de la propuesta de resolución implica el compromiso del cumplimiento de todos los objetivos planteados en la solicitud*.</i></p> <p>Se han reorganizado tareas en base al recorte sobre el presupuesto solicitado, sobre todo a nivel técnicos de apoyo, modificando someramente participación y responsables. Se han sufrido algunos retrasos en la adquisición de inventariable debido a la imposibilidad de comprar antes de fin de año por los plazos de arranque de proyecto y cierre de compras en la universidad. Durante este primer año se han retrasado 3 deliverables: D3.1 y D3.2v1 de junio a septiembre 2023; D1.2v1 de septiembre 2023 a enero 2024 por el retraso en la publicación de datasets geerados y las implicaciones derivadas del “Data Management Plan” que pasa a ser un deliverable con versiones para cada nuevo dataset.</p>

C4. Colaboraciones con otros grupos de investigación directamente relacionadas con el proyecto

Relacione las colaboraciones con otros grupos de investigación y el valor añadido que aportan al proyecto. Describa, si procede, el acceso a equipamientos y/o infraestructuras de otros grupos o instituciones.

Se han establecido contactos con el grupo de investigación del Centro propio de Investigación para la Integración en Paleobiología (CIPB-UAM) para identificar sinergias entre el proyecto y las necesidades del CIPB, especialmente en el procesamiento de imágenes capturadas desde un dron en el yacimiento de Las Hoyas (Cuenca), que presenta condiciones similares a las definidas en el proyecto (i.e., baja disponibilidad de datos de entrenamiento).

Para la segunda anualidad estancia de Marcos Escudero Viñolo en el ISR (Institute for Systems and Robotics) Lisboa, instituto de RD&I afiliado al Instituto Superior Técnico (IST) de la Universidad de Lisboa. (T2.3: Model profiling)

Preparación de una propuesta MSCA-JDN con las Universidades de Burdeos (UBx), Dublin City (DCU), y el instituto de RD&I suizo IDIAP.

C5. Colaboraciones con empresas o sectores socioeconómicos directamente relacionados con el proyecto. Relacione las colaboraciones con empresas o sectores socioeconómicos y el valor añadido que aportan al proyecto.

Se han incorporado tres nuevos EPOs: Deimos (marzo 2023), SENER (abril 2023), y el Servicio de Radiodiagnóstico del Hospital Universitario de La Princesa/Facultad de Medicina UAM (julio 2023).

Fruto de los desarrollos alcanzados en el proyecto y del gran avance de la IA generativa, se ha comenzado una colaboración con la empresa SENER Aeroespacial S.A., que ha fructificado en la propuesta de proyecto “Generación Avanzada de Imágenes Sintéticas Verosímiles para Inteligencia Artificial” para el programa COINCIDENTE 2023 del Ministerio de Defensa. Actualmente se están buscando otros programas de financiación que sean de interés ara ambas partes.

C6. Actividades de formación y movilidad de personal directamente relacionadas con el proyecto

Indique las actividades de formación y movilidad de personal relacionadas con el desarrollo del proyecto. Describa, además, si procede, las actividades realizadas en colaboración con otros grupos o con actividades de formación en medianas o grandes instalaciones.

	Nombre	Tipo de personal (becario/a, técnico/a, contratado/a con cargo al proyecto, posdoctoral, otros)	Descripción de las actividades de formación o motivo de la movilidad
	Acciones de formación: tesis doctorales		
1	Roberto Alcover Couso (directores de tesis: Juan	Becario FPI-UAM (miembros del equipo	Detection and Segmentation in Images using Synthetic Data

	Carlos San Miguel, Marcos Escudero)	investigador)	(prevista para finales 2025) <i>Tesis Doctoral (PhD Thesis), Programa de Doctorado en Ingeniería Informática y de Telecomunicación, Univ. Autónoma de Madrid.</i>
2	Javier Montalvo Rodrigo (directores de tesis: Álvaro García-Martín, Pablo Carballeira)	Ayudante LOU (miembros del equipo investigador)	Deep Learning with Adversarial Networks and Synthetic Data (prevista para finales 2025) <i>Tesis Doctoral (PhD Thesis), Programa de Doctorado en Ingeniería Informática y de Telecomunicación, Univ. Autónoma de Madrid.</i>
	Juan Ignacio Bravo Pérez-Villar (directores de tesis: Jesús Bescós, Álvaro García Martín)	Becario Doctorado Industrial IND2020/TIC-17515 (miembro del equipo investigador)	Técnicas de Visión Artificial Verificables y Explicables Para Navegación Basada en Visión Alrededor de Objetos Desconocidos (prevista para mediados 2024) <i>Tesis Doctoral (PhD Thesis), Programa de Doctorado en Ingeniería Informática y de Telecomunicación, Univ. Autónoma de Madrid.</i>
	Paula Moral de Eusebio (directores de tesis: Álvaro García-Martín, José M. Martínez)	Becaria FPI-UAM (miembros del equipo investigador)	Detection and association of objects in multiview scenarios (prevista para finales 2024) <i>Tesis Doctoral (PhD Thesis), Programa de Doctorado en Ingeniería Informática y de Telecomunicación, Univ. Autónoma de Madrid.</i>
	Kirill Siroktin (directores de tesis: Marcos Escudero, Pablo Carballeira).	Becario FPI-UAM (miembros del equipo investigador)	Self-supervised deep learning from unlabeled medical images (prevista para finales 2024) <i>Tesis Doctoral (PhD Thesis), Programa de Doctorado en Ingeniería Informática y de Telecomunicación, Univ. Autónoma de Madrid.</i>
	Cecilia Diana Alberla (directores de tesis: Jesús Bescós, Álvaro García Martín)	Doctoranda (miembros del equipo investigador)	Análisis de imagen médica mediante técnicas de aprendizaje profundo (Proyecto de tesis Enero 2024) <i>Tesis Doctoral (PhD Thesis), Programa de Doctorado en Ingeniería Informática y de</i>

					<i>Telecomunicación, Univ. Autónoma de Madrid.</i>
Acciones de formación: Trabajos Fin de Máster					
	Cecilia Diana Alberla (tutor: Juan Ignacio Braco Pérez-Villas; ponente: Álvaro García-Martín)	Estudiante de Máster (miembros de equipo investigador)			Self-supervised monocular Depth estimation and visual odometry on unseen synthetic cameras <i>Trabajo Fin de Máster (Master Thesis), Máster Universitario en Deep Learning for Audio and Visual Signal Processing, Univ. Autónoma de Madrid, Jun. 2023.</i>
	Andrés Gómez-Caraballo Yélamos (tutor: Javier Montalvo Rodrigo; ponente: Pablo Carballeira López)	Estudiante de Máster (miembros de equipo investigador)			Deep Learning modelling od intensity signals for synthetic Lidar data generation <i>Trabajo Fin de Máster (Master Thesis), Máster Universitario en Deep Learning for Audio and Visual Signal Processing, Univ. Autónoma de Madrid, Jun. 2023.</i>
	Pablo Marcos Manchón (tutor: Juan Carlos San Miguel Avedillo)	Estudiante de Máster (miembro de equipo investigador)			Synthetic data generation using latent diffusion models for semantic segmentation of urban scenes <i>Trabajo Fin de Máster (Master Thesis), Máster Universitario en Deep Learning for Audio and Visual Signal Processing, Univ. Autónoma de Madrid, Jun. 2023.</i>
	Juan José López Castilla (tutor: Álvaro García-Martín)	Estudiante de Máster (miembro de equipo investigador)			Design and evaluation of a dataset for detection and geopositioning of urban elements. <i>Trabajo Fin de Máster (Master Thesis), Máster Universitario en Deep Learning for Audio and Visual Signal Processing, Univ. Autónoma de Madrid, Sep. 2023.</i>
Acciones de formación: Trabajos Fin de Grado					
	Alejandro Guzmán Huerta (tutor: Miguel Ángel García)	Estudiante de Grado (miembro de equipo investigador)			<i>Clasificación auto-supervisada de imágenes mediante aprendizaje contrastivo (Auto-supervised image classification based on contrastive learning)</i> <i>Trabajo Fin de Grado (Graduate Thesis), Grado en Ingeniería Informática, Univ. Autónoma de</i>

			<i>Madrid, Feb. 2023.</i>
	Franciso Lerma Martínez (tutor: Miguel Ángel García García)	Estudiante de Grado (miembro de equipo investigador)	Clasificación auto-supervisada de imágenes mediante aprendizaje no- contrastivo (Auto-supervised image classification based on non- contrastive learning) Trabajo Fin de Grado (Graduate Thesis), Grado en Ingeniería Informática, Univ. Autónoma de Madrid, May 2023.
	Víctor Sánchez de la Roda Nuñez (tutor: Miguel Ángel García),	Estudiante de Grado (miembro de equipo investigador)	Clasificación auto-supervisada de imágenes mediante contraste de momento (Auto-supervised image classification based on momentum contrast) Trabajo Fin de Grado (Graduate Thesis), Grado en Ingeniería Informática (Programa de Doble Grado Ing. Informática y Matemáticas), Univ. Autónoma de Madrid, Jul. 2023.
	Anisha Anil Sujanani (tutor: Kiril Sirotkin; ponente: Marcos Escudero-Viñolo)	Estudiante de Grado (miembros de equipo investigador)	Interpretabilidad de redes neuronales de clasificación de imágenes mediante el cálculo de sus márgenes (Aiding interpretability of deep learning model by calculation of the decision margins) Trabajo Fin de Grado (Graduate Thesis), Grado en Ingeniería de Tecnologías y Servicios de Telecomunicación, Univ. Autónoma de Madrid, Sept. 2023.
	Actividades de formación: Seminarios		
	Actividades de movilidad		

Nota: Cree tantas filas como necesite

C7. Actividades de internacionalización y otras colaboraciones relacionadas con el proyecto
Indique si ha colaborado con otros grupos internacionales. Consigne si ha concurrido, y con qué resultado, a alguna convocatoria de ayudas (proyectos, formación, infraestructuras, otros) de programas europeos y/o otros programas internacionales, en temáticas relacionadas con la de

este proyecto. Indique el programa, socios, países y temática y, en su caso, financiación recibida.

Se ha obtenido un proyecto Erasmus+ Erasmus Mundus Joint Master “Image Processing and Computer Vision AI (IPCV AI)” que financiará 4 cohortes de un programa conjunto de máster (de dos años cada cohorte) junto a la Univesite de Bordeaux (UBx) de Burdeos (Francia) y la Pazmany Peter Katolikus Egytem (PPKE) de Budapest (Hungría)

(Concedida, a realizar) 2/2024–7/2024 Beca José Castillejo. “Decision regions of automatic deep learning models based on visual signal analysis for the formal quantification of their operability certainty and the interpretability of their predictions”. Lead researcher: M. Escudero Viñolo. Host researcher: C. Barata. Computer Vision and Robot Vision Lab. Institute for Systems and Robotics. Lisbon. 15.992 €. Ministry of Universities of the Spanish Government. Internship.

(A entregar) 11/2023. Entrega de la propuesta de Doctorado Conjunto Europeo MSCA-EJDN (UBx, DCU, IDIAP, UAM).

4. D. Difusión de los resultados del proyecto

Nota: Relacione **únicamente** los resultados derivados de este proyecto.

D1. Publicaciones científico-técnicas directamente derivadas de los resultados del proyecto.

Autores/as, título, referencia de la publicación*...

- [J1] Javier Montalvo, **Álvaro García-Martín**, and Jesus Bescós, “Exploiting Semantic Segmentation to Boost Reinforcement Learning in Video Game Environments”. Multimedia Tools and Applications, 82, 10961–10979, March 2023. Electronic ISSN 1573-7721, Print ISSN 1380-7501. DOI [10.1007/s11042-022-13695-1](https://doi.org/10.1007/s11042-022-13695-1) [*OpenAccess*]
This work is part of the preliminary tasks related to the Harvesting Visual Data (HVD) project (PID2021-125051OB-I00) funded by the Ministerio de Ciencia e Innovación of the Spanish Government.
- [J2] Paula Moral, **Álvaro García-Martín**, **José M. Martínez**, and Jesus Bescós, “Enhancing Vehicle Re-Identification Via Synthetic Training Datasets and Re-ranking Based on Video-Clips Information”. Multimedia Tools and Applications, Online 21 March 2023. Electronic ISSN 1573-7721, Print ISSN 1380-7501. DOI [10.1007/s11042-023-14511-0](https://doi.org/10.1007/s11042-023-14511-0) [*OpenAccess*]
This work is part of the preliminary tasks related to the Harvesting Visual Data (HVD) project (PID2021-125051OB-I00) funded by the Ministerio de Ciencia e Innovación of the Spanish Government.
- [J3] Roberto Alcover-Couso, Juan C. SanMiguel, Marcos Escudero-Viñolo, and **Alvaro Garcia-Martín**, “On exploring weakly supervised domain adaptation strategies for semantic segmentation using synthetic data”. Multimedia Tools and Applications, Online 11 March 2023. Electronic ISSN 1573-7721, Print ISSN 1380-7501. DOI [10.1007/s11042-023-14662-0](https://doi.org/10.1007/s11042-023-14662-0) [*OpenAccess*]
This work is part of the preliminary tasks related to the Harvesting Visual Data (HVD) project (PID2021-125051OB-I00) funded by the Ministerio de Ciencia e Innovación of the Spanish Government.
- [J4] Juan Ignacio Bravo Pérez-Villar, **Álvaro García-Martín**, Jesús Bescós, Marcos Escudero-Viñolo, "Spacecraft Pose Estimation: Robust 2D and 3D-Structural Losses and Unsupervised Domain Adaptation by Inter-Model Consensus", IEEE Transactions on Aerospace and Electronic Systems, Online 21 August 2023. Electronic ISSN: 1557-9603 , Print ISSN: 0018-9251. DOI [10.1109/TAES.2023.3306731](https://doi.org/10.1109/TAES.2023.3306731)

*Resalte en negrita las realizadas por las personas que son IP.

D2. Patentes directamente derivadas de los resultados del proyecto. Indicar si las patentes están licenciadas y/o en explotación.

D3. Asistencia a congresos, conferencias o workshops relacionados con el proyecto

Nombre del congreso, tipo de comunicación (invitada, oral, póster), autores/as....

[C1] Juan C. SanMiguel, Jorge Muñoz, and Fabio Poiesi, "Detection-aware multi-object tracking evaluation," 2022 18th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), Madrid (Spain), Nov. 2022, pp. 1-8. (ISBNprint 978-1-6654-6383-6, ISBNelectronic 978-1-6654-6382-9) (DOI 10.1109/AVSS56176.2022.9959412)

This work is part of the preliminary tasks related to the Harvesting Visual Data (HVD) project (PID2021-125051OB-I00) funded by the Ministerio de Ciencia e Innovación of the Spanish Government.

[C2] Roberto Alcover, Juan C. SanMiguel, Marcos Escudero, "Biased Class disagreement: detection of out of distribution instances by using differently biased semantic segmentation models", IEEE International Conference on Computer Vision Workshops 2023 (ICCV), Workshop on Uncertainty Quantification for Computer Vision, Paris (France), Oct. 2023, pp. 4580-4588.

Xuanling Yu et al., "The Robust Semantic Segmentation UNCV2023 Challenge Results", IEEE International Conference on Computer Vision Workshops 2023 (ICCV), Workshop on Uncertainty Quantification for Computer Vision, Paris (France), Oct. 2023, pp. 4618-4628.

[C3] Cecilia Diana Albelda, Juan Ignacio Bravo Pérez-Villar, Javier Montalvo, Álvaro García Martín, Jesús Bescós Cano, "Self-Supervised Monocular Depth Estimation on Unseen Synthetic Cameras", 26th Iberoamerican Congress on Pattern Recognition (CIAPR), Coimbra (Portugal), Nov. 2023.

[C4] Paula Moral, Alvaro Garcia-Martin, Jose M. Martinez, "Vehicle Re-Identification based on unsupervised domain adaptation by incremental generation of Pseudo-labels", 26th Iberoamerican Congress on Pattern Recognition (CIAPR), Coimbra (Portugal), Nov. 2023.

D4. Tesis doctorales finalizadas relacionadas con el proyecto

Nombre del doctor/a, director/a de tesis, título, calificación, organismo...

D5. Otras publicaciones derivadas de colaboraciones mantenidas durante la ejecución del proyecto y que pudieran ser relevantes para el mismo, así como artículos de divulgación, libros, conferencias, ...

Trabajos Fin de Máster

- [O1] Self-supervised monocular Depth estimation and visual odometry on unseen synthetic cameras, Cecilia Diana Albelda (advisor: Juan Ignacio Braco Pérez-Villas; lecturer: Álvaro García-Martín), Trabajo Fin de Máster, Máster Universitario en Deep Learning for Audio and Visual Signal Processing, Univ. Autónoma de Madrid, Jun. 2023.
- [O2] Deep Learning modelling of intensity signals for synthetic Lidar data generation, Andrés Gómez-Caraballo Yélamos (advisor: Javier Montalvo Rodrigo; lecturer: Pablo Carballería López), Trabajo Fin de Máster, Máster Universitario en Deep Learning for Audio and Visual Signal Processing, Univ. Autónoma de Madrid, Jun. 2023.
- [O3] Synthetic data generation using latent diffusion models for semantic segmentation of urban scenes, Pablo Marcos Manchón (advisor: Juan Carlos San Miguel Avedillo), Trabajo Fin de Máster, Máster Universitario en Deep Learning for Audio and Visual Signal Processing, Univ. Autónoma de Madrid, Jun. 2023.
- [O4] Design and evaluation of a dataset for detection and geopositioning of urban elements, Cecilia Diana Albelda (advisor: Álvaro García-Martín), Trabajo Fin de Máster, Máster Universitario en Deep Learning for Audio and Visual Signal Processing, Univ. Autónoma de Madrid, Sep. 2023.

Trabajos Fin de Grado

- [O5] Clasificación auto-supervisada de imágenes mediante aprendizaje contrastivo, Alejandro Guzmán Huerta (advisor: Miguel Ángel García), Trabajo Fin de Grado (Graduate Thesis), Grado en Ingeniería Informática, Univ. Autónoma de Madrid, Feb. 2023.
- [O6] Clasificación auto-supervisada de imágenes mediante aprendizaje no-contrastivo, Francisco Lerma Martínez (advisor: Miguel Ángel García García), Trabajo Fin de Grado, Grado en Ingeniería Informática, Univ. Autónoma de Madrid, May 2023.
- [O7] Clasificación auto-supervisada de imágenes mediante contraste de momento, Víctor Sánchez de la Roda Nuñez (advisor: Miguel Ángel García), Trabajo Fin de Grado, Grado en Ingeniería Informática (Programa de Doble Grado Ing. Informática y Matemáticas), Univ. Autónoma de Madrid, Jul. 2023.
- [O8] Contribuciones a la simulación de sistemas multi-cámara basado en UNITY, Miguel Bellido González del Alba (advisor: Juan C. SanMiguel), Trabajo Fin de Grado, Grado en

Tecnologías y Servicios de Telecom., Univ. Autónoma de Madrid, Jul. 2023.

[O9] Interpretabilidad de Redes Neuronales de clasificación de imágenes mediante el cálculo de sus márgenes, Anisha Anil Sujanani Sujanani (advisor: K. Sirtokin, lecturer: M. Escudero-Viñolo), Trabajo Fin de Grado, Grado en Tecnologías y Servicios de Telecom., Univ. Autónoma de Madrid, Sept. 2023.

D5.1 Software

- [Software for experimental evaluation Framework for MOT](#) This repository is an experimental framework for Multi Object Tracking analysis. Multiple detectors, trackers, and evaluation metrics can be found, and, also, documentation allowing to implement further models in the framework.
Software related to: Juan C. SanMiguel, Jorge Muñoz, and Fabio Poiesi, "Detection-aware multi-object tracking evaluation", 2022 18th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), Madrid, Spain, November 2022, pp. 1-8. (ISBNprint 978-1-6654-6383-6, ISBNelectronic 978-1-6654-6382-9) (DOI [10.1109/AVSS56176.2022.9959412](https://doi.org/10.1109/AVSS56176.2022.9959412))
- [Exploiting Semantic Segmentation To Boost Reinforcement Learning In Video Game Environments](#) Pytorch implementation of our paper:
Javier Montalvo, Álvaro García-Martín, and Jesus Bescós, "Exploiting Semantic Segmentation to Boost Reinforcement Learning in Video Game Environments". Multimedia Tools and Applications, 82, 10961–10979, March 2023. Electronic ISSN 1573-7721, Print ISSN 1380-7501. DOI [10.1007/s11042-022-13695-1](https://doi.org/10.1007/s11042-022-13695-1)
- [VPULab-CityFlow-ReID-Evaluation-Framework](#) This Evaluation Framework allows researchers to keep on evaluating the performance of their approaches and the ones from the State-of-the-Art over the CitiFlow-ReID dataset. As this dataset has not public ground truth annotations, AI City Challenge provided an on-line evaluation server during the challenge, which is no more available.
Software related to: Paula Moral, Álvaro García-Martín, José M. Martínez, and Jesus Bescós, "Enhancing Vehicle Re-Identification Via Synthetic Training Datasets and Re-ranking Based on Video-Clips Information". Multimedia Tools and Applications, Online 21 March 2023. Electronic ISSN 1573-7721, Print ISSN 1380-7501. DOI [10.1007/s11042-023-14511-0](https://doi.org/10.1007/s11042-023-14511-0)
- [On exploring weakly supervised domain adaptation strategies for semantic segmentation using synthetic data](#) Pytorch implementation of our paper:
Roberto Alcover-Couso, Juan C. SanMiguel, Marcos Escudero-Viñolo, and Alvaro Garcia-Martin, "On exploring weakly supervised domain adaptation strategies for semantic segmentation using synthetic data". Multimedia Tools and Applications, Online 11 March 2023. Electronic ISSN 1573-7721, Print ISSN 1380-7501. DOI [10.1007/s11042-023-14662-0](https://doi.org/10.1007/s11042-023-14662-0)
- [Spacecraft Pose Estimation: Robust 2D and 3D-Structural Losses and Unsupervised Domain Adaptation by Inter-Model Consensus](#). Pytorch implementation of our paper:
Juan Ignacio Bravo Pérez-Villar, Álvaro García-Martín, Jesús Bescós, Marcos Escudero-Viñolo, "Spacecraft Pose Estimation: Robust 2D and 3D-Structural Losses and Unsupervised Domain Adaptation by Inter-Model Consensus", IEEE Transactions on Aerospace and Electronic Systems, Online 21 August 2023. Electronic ISSN: 1557-9603 , Print ISSN: 0018-9251. DOI [10.1109/TAES.2023.3306731](https://doi.org/10.1109/TAES.2023.3306731)

D5.2 Datasets

D5.3 Deliverables

- D1.1 “System Infrastructure” version 1 (September 2023)
- D1.2 “Evaluation frameworks: datasets, metrics, and benchmarks” version 1 (September 2023) [delayed till January 2024]
- D3.1v1 “Data management plan” (September 2023) [originally planned June 2023]
- D3.2 “Results Report” version 1 (September 2023) [originally planned June 2023]