



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA



UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Escuela Técnica Superior de Ingeniería Geodésica,
Cartográfica y Topográfica

Reescalado de imágenes fotográficas por inteligencia
artificial aplicado a fotogrametría

Trabajo Fin de Grado

Grado en Ingeniería Geomática y Topografía

AUTOR/A: Chumillas García, Sergio

Tutor/a: Buchón Moragues, Fernando Francisco

Cotutor/a: Balaguer Puig, Matilde

CURSO ACADÉMICO: 2022/2023



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA



ESCUELA TÉCNICA SUPERIOR
DE INGENIERÍA GEODÉSICA
CARTOGRÁFICA Y TOPOGRÁFICA
UNIVERSITAT POLITÈCNICA DE VALÈNCIA

REESCALADO DE IMÁGENES FOTOGRAFICAS POR INTELIGENCIA ARTIFICIAL APLICADO A FOTOGRAMETRÍA

GRADO EN INGENIERÍA GEOMÁTICA Y TOPOGRAFÍA

CURSO 2022/2023

TRABAJO FIN DE GRADO

Autor: Sergio Chumillas García

Tutores: Fernando Buchón, Matilde Balaguer

AGRADECIMIENTOS

A mi familia, por su apoyo y amor incondicional y por haberme guiado en la vida pero dejándome elegir mi propio camino.

A mis profesores, por haberme enseñado tanto y por la calidad y pasión de su docencia.

A mis compañeros y amigos, por el increíble viaje que han sido estos cuatro años de carrera a su lado.

COMPROMISO

"El presente documento ha sido realizado completamente por el firmante; no ha sido entregado como otro trabajo académico previo y todo el material tomado de otras fuentes ha sido convenientemente entrecorillado y citado su origen en el texto, así como referenciado en la bibliografía"

Valencia - 21/06/2023

RESUMEN

La inteligencia artificial o IA es una tecnología en ciernes que se está desarrollando a velocidad vertiginosa y que sin duda está suponiendo el inicio de una nueva era. Como tal, es de alto interés estudiar el alcance de dicha tecnología en todos los campos posibles.

Este proyecto busca evaluar la viabilidad de utilizar el reescalado por inteligencia artificial al conjunto de fotografías empleadas para generar un modelo 3D por fotogrametría.

Se basa en la premisa de que, a mayor resolución de las imágenes, menor *ground sampling distance (GSD)* y mayor cantidad de píxeles, y por tanto mayor cantidad de puntos homólogos. Se pretende determinar si puede resultar un activo útil a la hora de mejorar los resultados.

Para ello se debe realizar una toma de fotografías de un objeto modelo, a las que posteriormente se les aplica un reescalado por IA para aumentar su resolución. Con dichas fotografías aumentadas se realiza un proceso de correlación fotogramétrica con el que se generan unos mapas de profundidad y en base a ellos se obtiene un modelo tridimensional.

Este modelo 3D obtenido por fotogrametría se compara con un modelo 3D de referencia tomado como verdadero obtenido con un escáner de luz estructurada. Esto con el fin de poder cuantificar los resultados. Ambos modelos deben ser escalados antes de poder ser comparados.

Este proceso se va a estudiar para varias resoluciones y para varias calidades de mapas de profundidad.

Palabras clave: Inteligencia Artificial, IA, Reescalado de imágenes, Fotogrametría.

RESUM

La intel·ligència artificial o IA és una tecnologia en embrió que s'està desenvolupant a velocitat vertiginosa i que sens dubte està suposant l'inici d'una nova era. Com a tal, és d'alt interès estudiar l'abast d'aquesta tecnologia en tots els camps possibles.

Aquest projecte busca avaluar la viabilitat d'utilitzar el reescalat per intel·ligència artificial al conjunt de fotografies emprades per a generar un model 3D per fotogrametria.

Es basa en la premissa que, a major resolució de les imatges, menor *ground sampling distance (GSD)* i major quantitat de píxels, i per tant major quantitat de punts homòlegs. Es pretén determinar si pot resultar un actiu útil a l'hora de millorar els resultats.

Per a això s'han de realitzar unes de fotografies d'un objecte model, a les quals posteriorment se'ls aplica un reescalat per IA per a augmentar la seua resolució. Amb aquestes fotografies augmentades es realitza un procés de correlació fotogramètrica amb el qual es generen uns mapes de profunditat i sobre la base d'ells s'obté un model tridimensional.

Aquest model 3D obtingut per fotogrametria es compara amb un model 3D de referència pres com a vertader obtingut amb un escàner de llum estructurada. Això amb la finalitat de poder quantificar els resultats. Tots dos models han de ser escalats abans de poder ser comparats.

Aquest procés s'estudiarà per a diverses resolucions i per a diverses qualitats de mapes de profunditat.

Paraules clau: Intel·ligència Artificial, IA, Reescalat d'imatges, Fotogrametria.

ABSTRACT

Artificial intelligence or AI is a budding technology that is developing at breakneck speed and is undoubtedly ushering in a new era. As such, it is of high interest to study the scope of such technology in all possible fields.

This project seeks to evaluate the feasibility of using artificial intelligence to upscale to the set of photographs used to generate a 3D model by photogrammetry.

It is based on the premise that the higher the resolution of the images, the lower the ground sampling distance (GSD) and the greater the number of pixels, and therefore greater the number of homologous points. The aim is to determine if it can be a useful asset to improve the results.

In order to do it, photographs of a model object must be taken, which are then upscaled by AI to increase their resolution. With these enlarged photographs, a photogrammetric correlation process is carried out to generate depth maps and, based on them, a three-dimensional model is obtained.

This 3D model obtained by photogrammetry is compared with a reference 3D model taken as true obtained with a structured light scanner. This in order to be able to quantify the results. Both models must be scaled before they can be compared.

This process will be studied for various resolutions and for various depth map qualities.

Keywords: Artificial Intelligence, AI, Image upscaling, Photogrammetry.

TERMINOLOGÍA UTILIZADA

A continuación se detallan una serie de términos que se utilizarán de forma repetitiva a lo largo del presente documento por estos facilitar la lectura y entendimiento de los temas a exponer:

- **Downscaling:** Disminuir la resolución de una imagen. Este proceso no involucra inteligencia artificial.
- **Upscaling:** Aumentar la resolución de una imagen. Este proceso siempre involucra inteligencia artificial (en este trabajo).
- **AI-Upscaler:** Referido a un modelo de inteligencia artificial que ha sido entrenado con el único objetivo de aumentar la resolución de imágenes.
- **Reescalar:** Referido a cambiar la resolución de una imagen, ya sea mediante *downscaling* o mediante *upsampling*.

Aclaraciones y consideraciones:

En el presente documento se utiliza como separador decimal “.”, y como separador de miles en adelante “,”.

Hardware empleado:

- CPU: Intel i7-6700K 4 GHz.
- RAM: 2x8 GB DDR5 2400 MHz.
- GPU: AMD RADEON RX 5700 XT 8GB VRAM.

ÍNDICE DE FIGURAS

<i>Ilustración 1: Proceso Structure from motion. Fuente: Basado en ilustración de Ozgur Yilmaz en Researchgate.</i>	1
<i>Ilustración 2: Geometría epipolar. Fuente: Colorado School of Mines, Engineering Division.</i>	2
<i>Ilustración 3: Ejemplo de las capacidades de un AI-Upscaler. Fuente: (https://www.hitpaw.com)</i>	3
<i>Ilustración 4: Ejemplo de uso de un escáner de luz estructurada. Fuente: CreaForm</i>	4
<i>Ilustración 5: Fotografía del objeto modelo. Fuente: Propia.</i>	5
<i>Ilustración 6: Conjunto de fotos tomadas. Fuente: Propia.</i>	7
<i>Ilustración 7: Entorno de trabajo de chaiNNer. Fuente: Propia.</i>	8
<i>Ilustración 8: Flujo de trabajo para obtención del set de fotos x0.25 en chaiNNer. Fuente: Propia.</i>	10
<i>Ilustración 9: Flujo de trabajo para obtención del set de fotos x0.50 en chaiNNer. Fuente: Propia.</i>	11
<i>Ilustración 10: Flujo de trabajo para obtención del set de fotos x0.75 en chaiNNer. Fuente: Propia.</i>	12
<i>Ilustración 11: Flujo de trabajo para obtención del set de fotos x1 en chaiNNer. Fuente: Propia.</i>	13
<i>Ilustración 12: Flujo de trabajo para obtención del set de fotos x2 en chaiNNer. Fuente: Propia.</i>	14
<i>Ilustración 13: Flujo de trabajo para obtención del set de fotos x3 en chaiNNer. Fuente: Propia.</i>	14
<i>Ilustración 14: Flujo de trabajo para obtención del set de fotos x4 en chaiNNer. Fuente: Propia.</i>	15
<i>Ilustración 15: Comparación de tres fotos (x0.25, x1 y x4 respectivamente) al hacer zoom al mismo elemento. Fuente: Propia.</i>	15
<i>Ilustración 16: Parámetros del comando Align Photos dentro de Metashape. Fuente: Propia.</i>	17
<i>Ilustración 17: Modelo 3D generado en Metashape. Fuente: Propia.</i>	19
<i>Ilustración 18: Parámetros del comando Build Mesh dentro de Metashape. Fuente: Propia.</i>	20
<i>Ilustración 19: Markers y scale bars en el modelo para escalarlo. Fuente: Propia.</i>	22
<i>Ilustración 20: Malla refinada. Fuente: Propia.</i>	23
<i>Ilustración 21: Especificaciones técnicas CreaForm Academia 50. Fuente: Emco UK.</i>	24
<i>Ilustración 22: Escáner CreaForm Academia 50. Fuente: CreaForm.</i>	25
<i>Ilustración 23: Modelo de referencia obtenido con el escáner de luz estructurada. Fuente: Propia.</i>	25
<i>Ilustración 24: Alineamiento inicial erróneo por la falta de correspondencia geométrica entre modelos. Fuente: Propia.</i>	27
<i>Ilustración 25: Tabla comparativa de los distintos modelos 3D obtenidos. Fuente: Propia.</i>	32
<i>Ilustración 26: Modelos obtenidos con diferentes resoluciones en los que se aprecia un aumento del detalle así como del ruido. Fuente: Propia.</i>	33
<i>Ilustración 27: Tabla comparativa de las comparaciones realizadas entre los modelos obtenidos por IA y el modelo de referencia. Fuente: Propia.</i>	34
<i>Ilustración 28: Evolución del porcentaje dentro de tolerancia a lo largo de las resoluciones para distintas calidades de mapas de profundidad. En verde se marcan los máximos Fuente: Propia.</i>	35
<i>Ilustración 29: Resultado de la comparación entre el modelo x1_high y el modelo x1_original_high. Unidades en mm. Fuente: Propia.</i>	36
<i>Ilustración 30: Agrupamiento de los puestos de trabajo por grupo profesional y nivel salarial. Fuente: XX convenio colectivo nacional de empresas de ingeniería; oficinas de estudios técnicos; inspección, supervisión y control técnico y de calidad. BOE.</i>	37
<i>Ilustración 31: Tabla salarial para el año 2022 para el convenio de ingenieros y oficina técnica. Fuente: XX convenio colectivo nacional de empresas de ingeniería; oficinas de estudios técnicos; inspección, supervisión y control técnico y de calidad. BOE.</i>	38

ÍNDICE DE TABLAS

<i>Tabla 1: Características y parámetros comunes de la cámara. Fuente: Propia.</i>	6
<i>Tabla 2: Conjuntos de fotos y sus resoluciones. Fuente: Propia.</i>	8
<i>Tabla 3: Información acerca de los conjuntos de fotos. Fuente: Propia.</i>	30
<i>Tabla 4: Tiempos de procesamiento para mapas de profundidad en calidad lowest. Fuente: Propia.</i>	30
<i>Tabla 5: Tiempos de procesamiento para mapas de profundidad en calidad low. Fuente: Propia.</i>	30
<i>Tabla 6: Tiempos de procesamiento para mapas de profundidad en calidad medium. Fuente: Propia.</i>	31
<i>Tabla 7: Tiempos de procesamiento para mapas de profundidad en calidad high. Fuente: Propia.</i>	31
<i>Tabla 8: Correspondencia entre los diferentes sets y sus factores de escala respecto al set base de x0.25. Fuente: Propia.</i>	36
<i>Tabla 9: Desglose de tareas y horas, así como su coste. Fuente: Propia.</i>	38
<i>Tabla 10: Disponibilidad y precios del software empleado. Fuente: Propia.</i>	39
<i>Tabla 11: Cotizaciones a la seguridad social 2023. Fuente: (https://www.seg-social.es)</i>	39

ÍNDICE

1. INTRODUCCIÓN	1
2. OBJETIVOS	5
3. DATOS.....	5
3.1 Objeto a reconstruir.....	5
3.2 Fotografías.....	5
3.3 Modelo de inteligencia artificial para reescalado	6
4. METODOLOGÍA	7
4.1 Descripción general.....	7
4.2 Obtención del set de fotos original del objeto.	7
4.3 Reescalado de fotografías.....	7
4.3.1 Set de fotos x0.25, x0.50, x0.75 y x1.....	9
4.3.2 Set de fotos x2, x3 y x4	13
4.4 Obtención de los modelos 3D por fotogrametría	16
4.4.1 Importar y alinear fotografías.....	16
4.4.2 Generar malla 3D	18
4.4.3 Escalar modelo	21
4.4.4 Refinar malla	22
4.5. Modelo 3D de referencia por luz estructurada.....	23
4.6 Comparación de modelos.....	26
4.6.1 Alineación de modelos	26
4.6.2 Comparación de modelos.....	28
5. RESULTADOS	30
5.1 Fotografías.....	30
5.2 Modelos 3D por fotogrametría	32
5.3 Comparación entre modelos	34
6. PRESUPUESTO.....	37
7. CONCLUSIONES.....	40
8. BIBLIOGRAFÍA	41

1. INTRODUCCIÓN

La fotogrametría es una técnica utilizada en muchos ámbitos, desde la geomática para generar modelos de elevaciones, hasta videojuegos e incluso películas para generar objetos tridimensionales digitales. Se trata de una técnica que recibe de entrada un conjunto de imágenes y a partir de estas es capaz de reconstruir un modelo 3D.

El hecho de poder reconstruir un modelo tridimensional a partir de imágenes radica en el uso de dos técnicas: *structure from motion (SFM)* y *multi-view stereo (MVS)*.

Estas técnicas trabajan de forma complementaria y cada una produce un resultado distinto. El flujo de trabajo de dichas técnicas es el siguiente:

Las técnicas o algoritmos SFM reciben como entrada un conjunto de imágenes, a continuación se detectan elementos 2D en cada imagen de entrada y se relacionan dichos elementos entre pares de imágenes. Las imágenes deben poseer elementos suficientemente diferenciables entre sí, de tal forma que haya elementos únicos en forma, tamaño y textura, sin que estos sean repetitivos, pues si hay un elemento de dichas características, por ejemplo una textura que sigue un patrón regular, la relación de un mismo elemento entre diferentes imágenes puede no ser consistente.

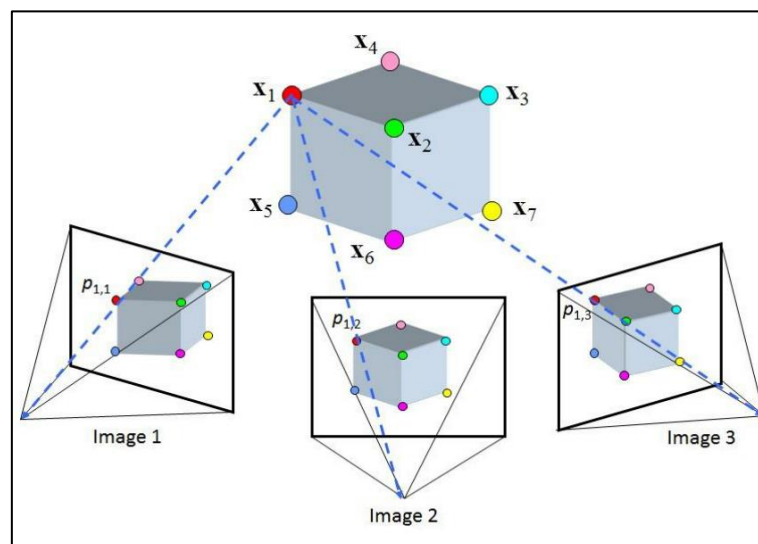


Ilustración 1: Proceso Structure from motion. Fuente: Basado en ilustración de Ozgur Yilmaz en Researchgate.

Una vez hay suficientes elementos detectados en distintas imágenes, se puede calcular una matriz de transformación que permite conocer la posición tridimensional relativa entre dos posiciones de cámara. Esto también posibilita la obtención de los parámetros de orientación interna y externa de la cámara para cada imagen.

Los elementos resultantes de dicho proceso son una nube poco densa de puntos y los ya mencionados parámetros de orientación interna y externa.

Posteriormente, las técnicas MVS utilizan los parámetros de orientación de cada imagen obtenidos previamente para refinar el resultado producido, dando lugar a una nube de puntos densa o, en su defecto, mapas de profundidad.

La mejora en eficacia a la hora de detectar elementos homólogos en distintas imágenes con las técnicas MVS respecto a las SFM radica en que, conocidos los parámetros de orientación interna y externa de la cámara para cada imagen, y sabiendo que el mismo punto de un mismo elemento en dos imágenes distintas se encuentra en una línea (principio de la geometría epipolar), se consigue reducir la búsqueda de elementos homólogos de un entorno plano (2D) a una línea (1D).

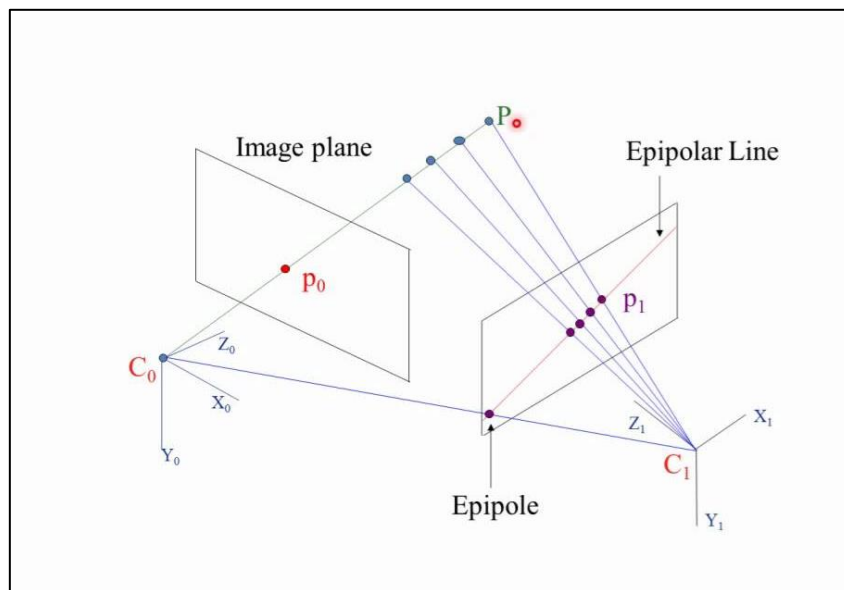


Ilustración 2: Geometría epipolar. Fuente: Colorado School of Mines, Engineering Division.

Al aumentar la resolución de las imágenes de entrada, cada píxel representa una porción más pequeña del objeto fotografiado a reconstruir. Esto aumenta el coste computacional pero a su vez permite detectar un mayor número de elementos homólogos entre imágenes, lo que resulta en un mejor resultado.

Reescalar una imagen consiste en aumentar o disminuir su resolución o número de píxeles. Es en esta empresa donde entran en juego los modelos de inteligencia artificial conocidos como *AI Up-Scalers*, que permiten aumentar mucho la resolución de una imagen de forma efectiva y sin apenas perder detalles.

Estas IAs se tratan de modelos de machine-learning de aprendizaje reforzado que han sido entrenadas con el único propósito de reescalar imágenes para aumentar su resolución.

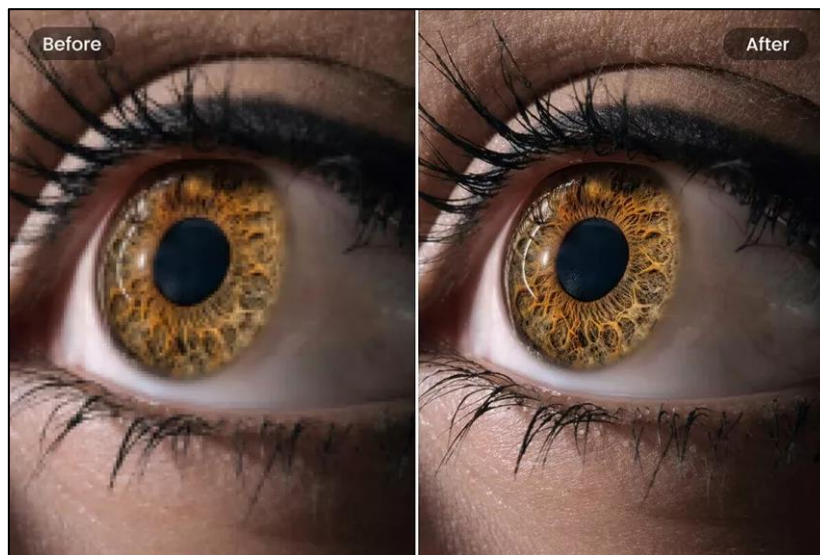


Ilustración 3: Ejemplo de las capacidades de un AI-Upscaler. Fuente: (<https://www.hitpaw.com>)

Esto supone un gran hito pues hasta hace un par de años siempre se había podido reescalar una imagen a una menor resolución mediante algoritmos de interpolación, pero a la hora de querer aumentar la resolución de una imagen, dichos métodos producían resultados no satisfactorios. Pues en cierto modo se está intentado obtener detalle de donde no lo hay.

Estos *AI Up-Scalers* se están utilizando de forma multidisciplinar en distintos ámbitos, desde el más lógico como puede ser el artístico para mejorar la resolución de fotografías o dibujos artísticos, hasta ámbitos policiales para permitir identificar matrículas en aquellas fotografías de menor calidad de los radares de velocidad, pasando incluso por el ámbito de los videojuegos y el mundo audiovisual al poder producir texturas de mayor resolución con facilidad.

Otra forma de obtener un modelo 3D de un objeto es mediante el uso de un *escáner de luz estructurada*. Se trata de un dispositivo, generalmente de mano, que emite patrones calibrados de luz a alta frecuencia y un receptor incorporado captura la deformación de dichos patrones sobre la superficie en que se proyecta.

Como los patrones están calibrados sobre una superficie plana, al ser proyectados a una superficie tridimensional se pueden cuantificar de forma

precisa las deformaciones sufridas. Dichas deformaciones permiten generar el modelo 3D a reconstruir. Son herramientas cuya precisión suele rondar las décimas de milímetro.



Ilustración 4. Ejemplo de uso de un escáner de luz estructurada. Fuente: CreaForm

En el presente trabajo se desarrolla la siguiente idea: *¿Se puede utilizar la IA para aumentar la resolución de las fotografías y producir mejores resultados en la fotogrametría? Y si es así, ¿cómo de alejados de la realidad están dichos resultados?*

2. OBJETIVOS

En base a lo expuesto previamente, lo que se busca conseguir con este trabajo es lo siguiente:

- 1- Determinar si es posible mejorar la precisión y detalle de la malla 3D al involucrar el uso de un *AI-Upscaler* en el set de fotografías usadas para el proceso fotogramétrico.
- 2- Determinar un umbral a partir del cual seguir aumentando la resolución de las fotografías no mejora virtualmente los resultados. Pues se espera que la mejora en los resultados se comporte de forma similar a una función logarítmica.

3. DATOS

3.1 Objeto a reconstruir

Se busca generar un modelo 3D de la siguiente figura de unos 20cm de alto:



Ilustración 5: Fotografía del objeto modelo. Fuente: Propia.

3.2 Fotografías

Se parte de un conjunto de 75 fotografías tomadas con un teléfono móvil. Dichas fotos poseen una resolución original de 2976 x 3968 píxeles.

De ahora en adelante se referirá a los distintos conjuntos de fotografías por su factor de escala ($x0.25$, $x0.50$, $x0.75$, $x1$, $x2$, $x3$, $x4$). Siendo el conjunto original $x1_original$.

Cada imagen posee metadatos incrustados en el mismo archivo referentes a la ISO, distancia focal, velocidad del obturador o tiempo de exposición, y apertura del diafragma.

La siguiente tabla muestra características y parámetros comunes de la cámara a la hora de realizar la toma de todas las fotografías:

Cámara	WAS-LX1A
Megapíxeles	12
Apertura del diafragma	f/ 2.2
Tiempo de exposición	1/33 s
ISO	120 – 400
Distancia focal	3.83 mm

Tabla 1: Características y parámetros comunes de la cámara. Fuente: Propia.

3.3 Modelo de inteligencia artificial para reescalado

Como modelo de *AI-Upscaler* se ha optado por el modelo *Remacri_x4*. Actualmente no existen documentos que comparen cuantitativa y estadísticamente la calidad de los resultados para cada modelo publicado en internet, y menos aplicado a fotogrametría.

Se ha escogido, por tanto, dicho modelo por la opinión popular de ser el que mejores resultados ofrece.

4. METODOLOGÍA

4.1 Descripción general

La metodología a desarrollar puede resumirse en los siguientes apartados:

- 1- Obtención del set de fotografías
- 2- Reescalado del set de fotos por inteligencia artificial
- 3- Generar modelos 3D por fotogrametría con las fotos reescaladas
- 4- Obtención de modelo referencia tomado como verdadero con escáner de luz estructurada.
- 5- Comparación de todos los modelos 3D obtenidos por fotogrametría con el modelo de referencia.

A continuación se ahonda en el flujo de trabajo seguido:

4.2 Obtención del set de fotos original del objeto.

Se han realizado 75 fotografías al objeto modelo con un teléfono móvil siguiendo un patrón circular en anillo en torno al objeto, desde todos los ángulos posibles y buscando el mayor solapamiento entre tomas.

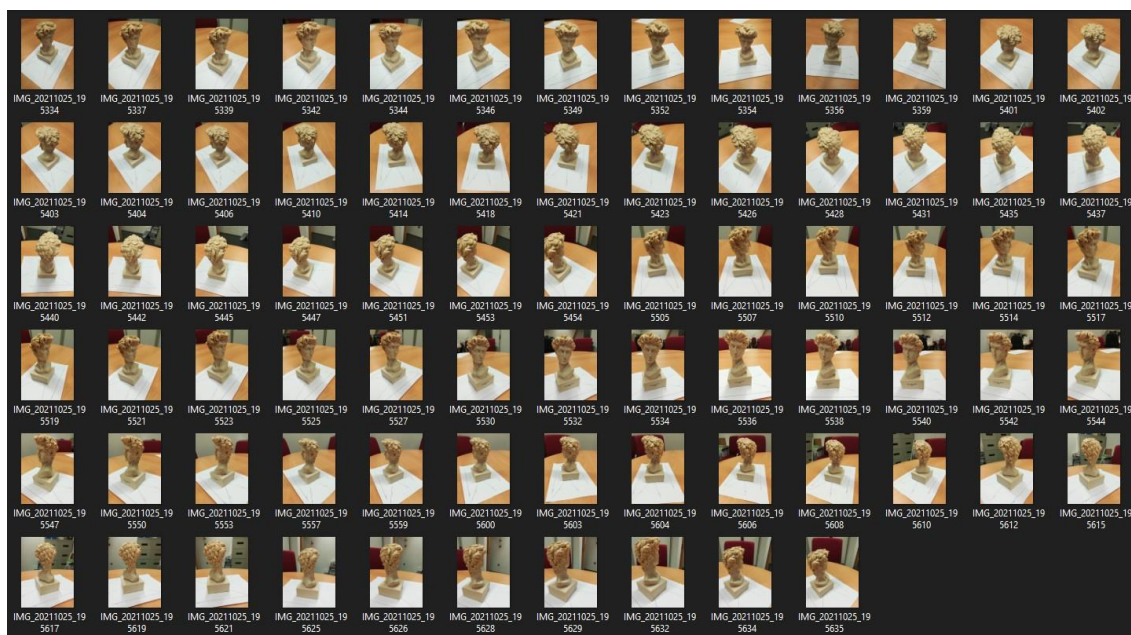


Ilustración 6: Conjunto de fotos tomadas. Fuente: Propia.

Con el fin de mejorar la calidad del set inicial, se han tomado las fotos con un ISO bajo, una iluminación fija e invariable y una distancia focal constante.

4.3 Reescalado de fotografías

Los modelos de IA se pueden operar con *Python* utilizando algunas librerías específicas. Este proceso puede probar ser tedioso y poseer una curva de aprendizaje elevada. Es por este motivo que la comunidad crea proyectos de

interfaces gráficas para poder operar los modelos de IA de una forma fácil e intuitiva.

El proceso de reescalado se ha llevado a cabo utilizando el proyecto de interfaz gráfica *ChaiNner*, de código libre. Se trata de una interfaz gráfica que permite manipular los modelos de IA mediante un flujo de trabajo basado en nodos.

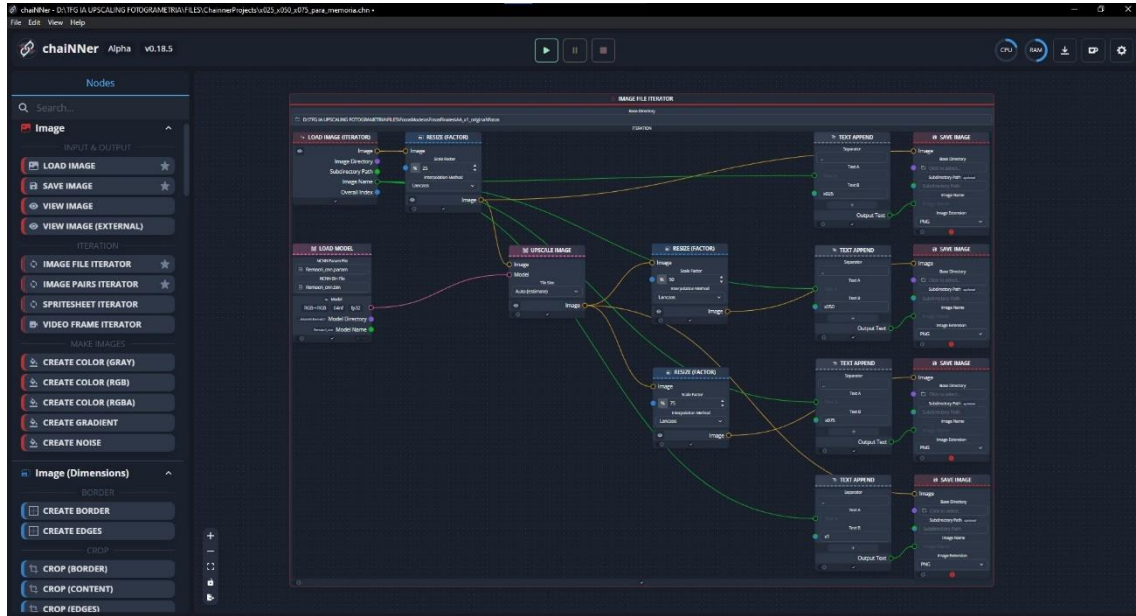


Ilustración 7: Entorno de trabajo de *chaiNner*. Fuente: Propia.

Un flujo de trabajo de programación a través de nodos consiste en una serie de bloques (que representan órdenes, operaciones o procesos) los cuales se pueden conectar entre sí a través de nodos. Estos bloques reciben los datos de entrada por los nodos de izquierda y producen los resultados por los nodos de la derecha de manera secuencial. Conectando y concatenando bloques entre sí se pueden lograr algoritmos tan complejos y extensos como se desee.

Los conjuntos o sets de fotografías de distintas resoluciones tendrán su propia terminología. El nombre con el cual se va a referir a ellos y las resoluciones que presentan se muestran en la siguiente tabla:

Nombre	Resolución	Involucra IA
x0.25	744 x 992	No
x0.50	1488 x 1984	Sí
x0.75	2232 x 2976	Sí
x1	2976 x 3968	Sí
x1_original	2976 x 3968	No
x2	5952 x 7936	Sí
x3	8928 x 11904	Sí
x4	11904 x 15872	Sí

Tabla 2: Conjuntos de fotos y sus resoluciones. Fuente: Propia.

En función de la calidad establecida para los mapas de profundidad, como se verá más adelante, se añadirá el sufijo correspondiente. E.g: *x3_high*, *x4_low*.

4.3.1 Set de fotos x0.25, x0.50, x0.75 y x1

Debido a la alta resolución del set de fotografías original (casi 3000x4000 píxeles), se ha decidido reducir la resolución de dichas imágenes a un 25% del tamaño original. Llamaremos a este conjunto de fotos como el conjunto x0.25. Dicho set de fotos será el que se utilizará como base para obtener el resto de resoluciones.

Escalar las imágenes originales por un factor de x0.25 se realiza con el fin de dificultar el proceso de *upscaling* a la IA, pues la calidad que estos modelos son capaces de generar es proporcional a la calidad de la imagen original. Por lo tanto al reducir la resolución de las imágenes originales (a aproximadamente 750x1000 píxeles) se está poniendo a prueba la capacidad del modelo de inteligencia artificial para trabajar en entornos poco favorables.

El set de fotos x0.25 se obtiene directamente a partir del set de imágenes originales x1_original, aplicando un *downscaling* por el método lanczos por el cual se reescala cada imagen por un factor de 0.25.

En consecuencia, el set base de fotos x0.25 es el único conjunto de imágenes que no involucra inteligencia artificial, además del conjunto original.

Se ha escogido el método de interpolación lanczos por ser este el que mejores resultados ofrece, a costa de suponer un ligero aumento en la carga computacional. Se trata de una técnica de interpolación compleja que se basa en la transformada de Fourier.

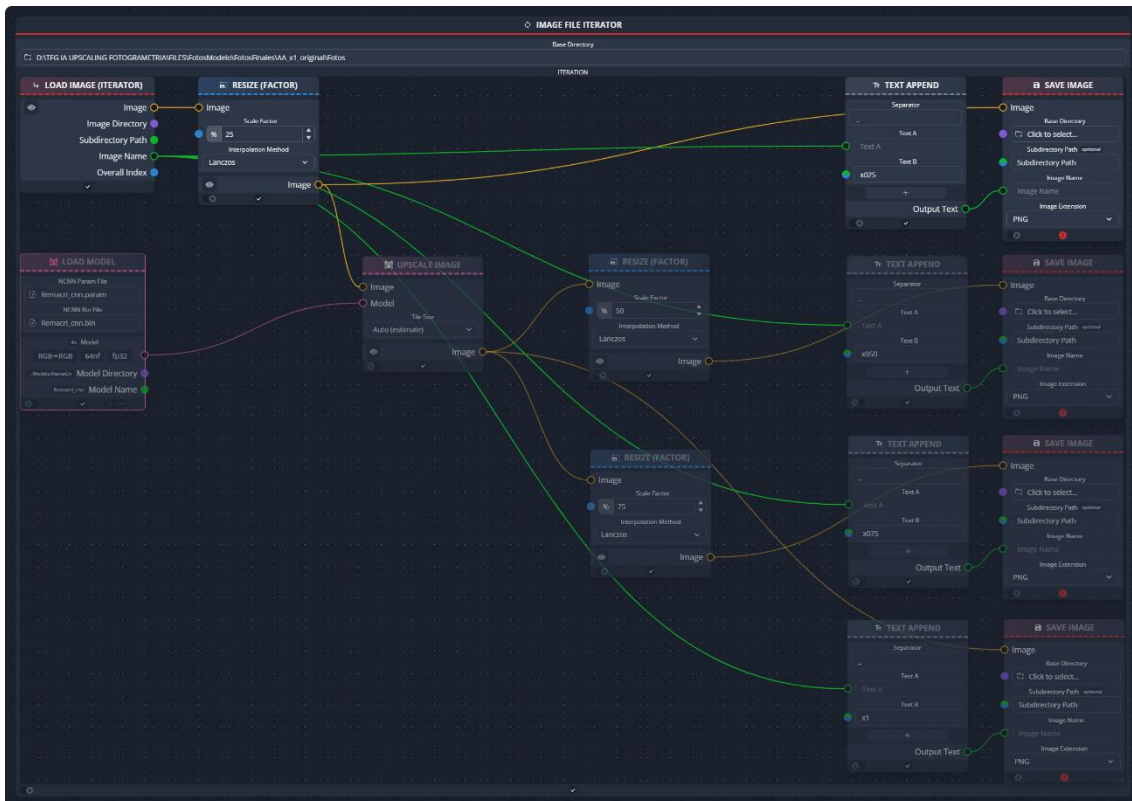


Ilustración 8: Flujo de trabajo para obtención del set de fotos x0.25 en chainNer. Fuente: Propia.

El proceso de reescalado mostrado en la imagen es el siguiente:

- Se introduce un directorio base en el bloque iterador y este va cargando las imágenes presentes en dicho directorio de manera secuencial.
- La imagen es enviada desde el bloque iterador a través de un nodo al bloque *resize*, donde se le aplica un factor de escala del 25% según el método de interpolación escogido. En este caso lanczos.
- Al mismo tiempo que el bloque iterador envía la imagen al bloque *resize*, envía también el nombre que recibe el archivo de la imagen al bloque *text append*, donde se concatena cualquiera que sea el nombre original de la foto, con el texto deseado. En este caso concatena el texto x025 utilizando “_” como separador. De la forma que el nombre del archivo resultaría “NombreDeImagen_x025”.
- Una vez los bloques anteriores han terminado de ejecutarse, el bloque *save image* recibe por un lado la imagen reescalada del bloque *resize*, y por otro el nuevo nombre para la imagen del bloque *text append*. Este bloque guardará cada imagen en el directorio especificado.
- Una vez especificado el directorio, solo queda ejecutar el bloque iterador principal, y este irá realizando todo este proceso para cada una de las imágenes en el directorio base.

La obtención del conjunto de fotos x0.50 y x0.75 difiere un poco respecto al proceso seguido para el resto de escalas debido a una limitación del modelo de IA utilizado: dicho modelo solo permite reescalar por un factor de 4 (4 por cada lado. Distinto a *metashape*, como se describe más adelante). Por lo que para obtener los sets de x0.50 y x0.75 a partir del set base de x0.25, se debe primero reescalar por un factor de 4 (se obtienen imágenes con la misma resolución que el set original) y posteriormente reescalar por un factor de 0.50 y 0.75 respectivamente.

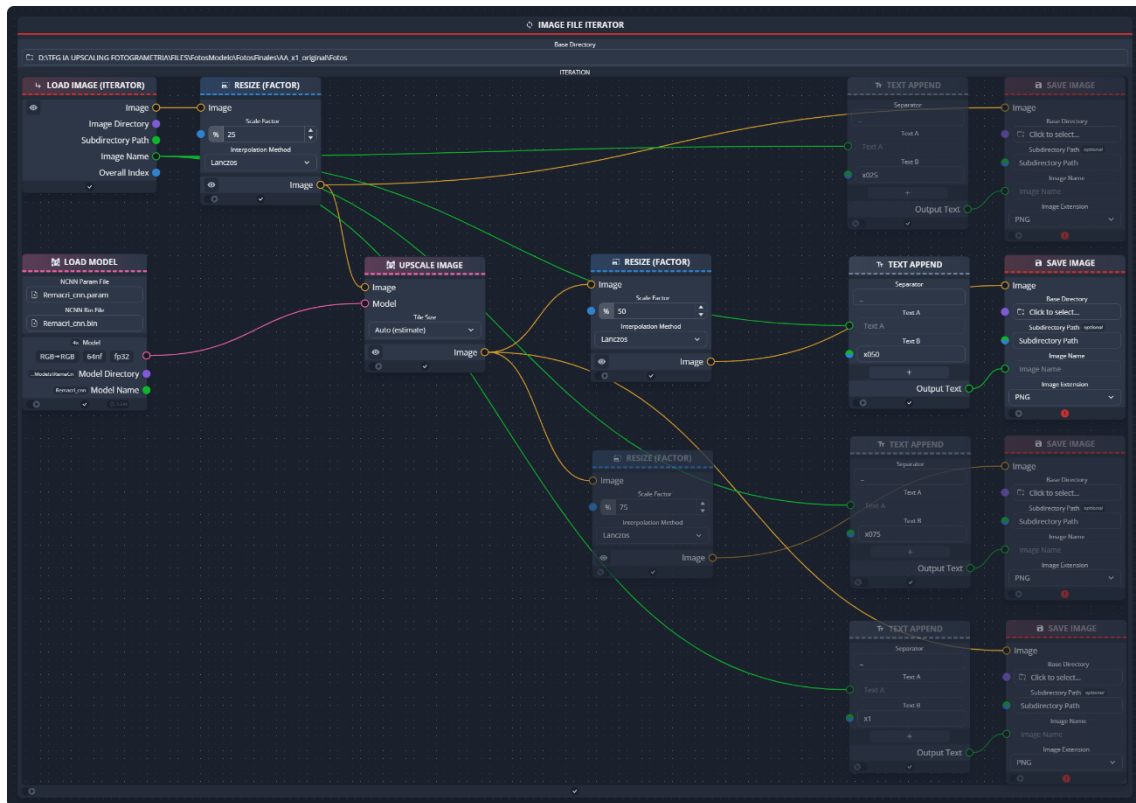


Ilustración 9: Flujo de trabajo para obtención del set de fotos x0.50 en chainNer. Fuente: Propia.

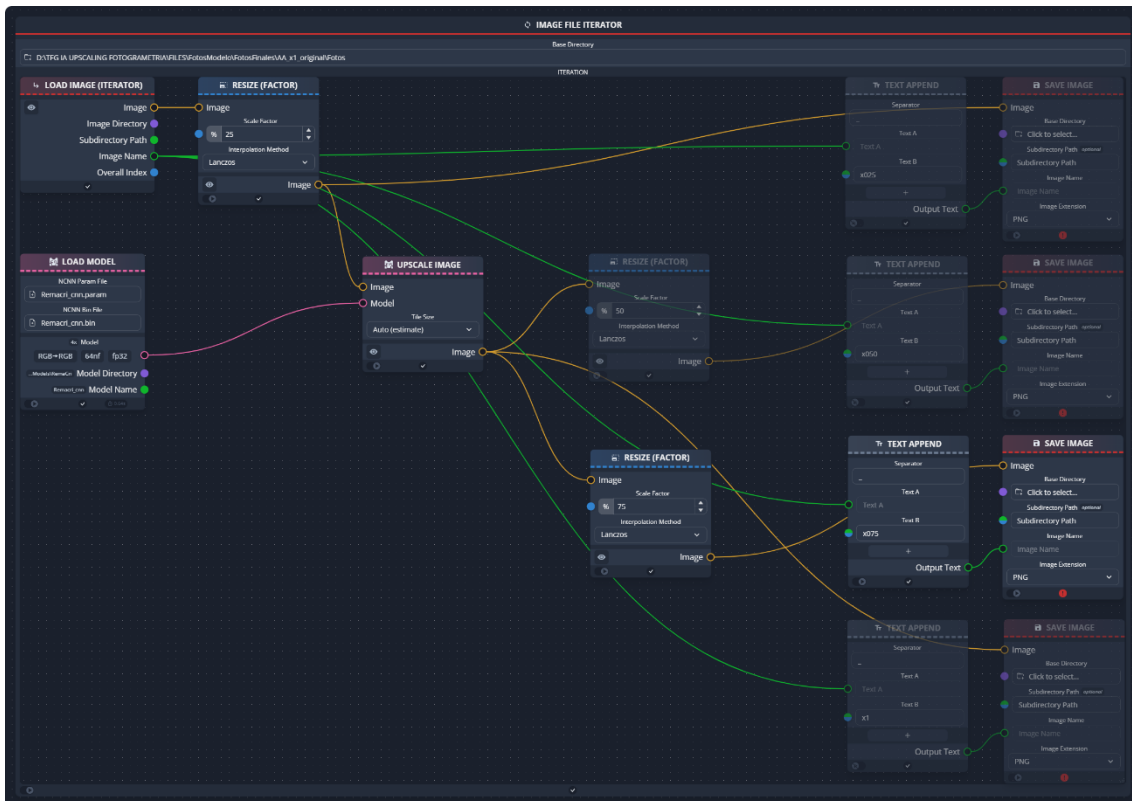


Ilustración 10: Flujo de trabajo para obtención del set de fotos x0.75 en chaiNNer. Fuente: Propia.

En este caso el proceso es igual que el seguido para las imágenes de x0.25, con la salvedad de que se añaden dos bloques intermedios entre el bloque *resize* y el bloque *save image*. La funcionalidad de estos bloques es la siguiente:

- La imagen reescalada al 25% llega al bloque *upscale image*. Este bloque también debe recibir como dato de entrada el modelo de IA utilizado para reescalar, en este caso dicho dato viene dado por el bloque *load model*, que permite cargar un modelo ubicado localmente en ordenador. El modelo utilizado para este caso es el modelo *Remacri*. Una vez este bloque posee todo lo necesario, reescala la imagen de entrada utilizando el modelo asignado. Este proceso es altamente costoso computacionalmente y requiere de varios minutos para cada fotografía según la resolución de la imagen de entrada.
- Una vez la imagen x0.25 ha sido reescalada por un factor de 4 con el modelo *Remacri*, y por tanto el resultado es una imagen con la misma resolución que las imágenes originales, se envía cada una a un bloque *resize* diferente con sus factores de escala pertinente.

El resto del proceso es similar al seguido para obtener las imágenes x0.25.

Por último, para la obtención del conjunto de fotos x1, de misma resolución que el conjunto original, se realiza un proceso similar al previo, con la diferencia de que no se aplica una reducción de la resolución tras el reescalado x4:

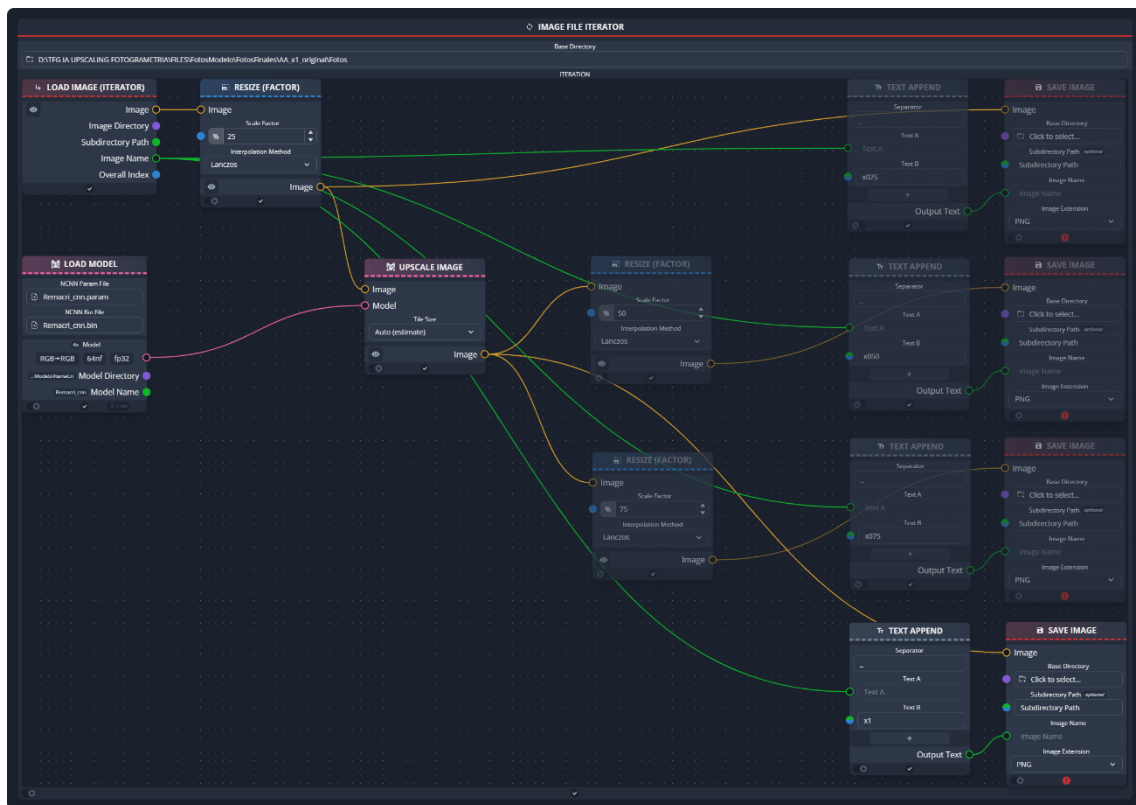


Ilustración 11: Flujo de trabajo para obtención del set de fotos x1 en chaiNner. Fuente: Propia.

4.3.2 Set de fotos x2, x3 y x4

La obtención de los conjuntos de fotos x2, x3 y x4 es un proceso directo, pues solamente se deben reescalar por un factor de 4 los conjuntos de fotografías x0.50, x0.75 y x1 respectivamente. Para ello se sigue una metodología parecida a las explicadas previamente, con la diferencia de que como no parten de la misma fotografía para reescalar, no se pueden realizar los tres procesados a la vez dentro de un mismo bloque iterador, como sí se había hecho para las resoluciones x0.25, x0.50, x0.75 y x1 con el fin de ser más eficiente en cuanto al tiempo de procesamiento.

Es indispensable cambiar en cada proceso el directorio base del iterador en consecuencia.

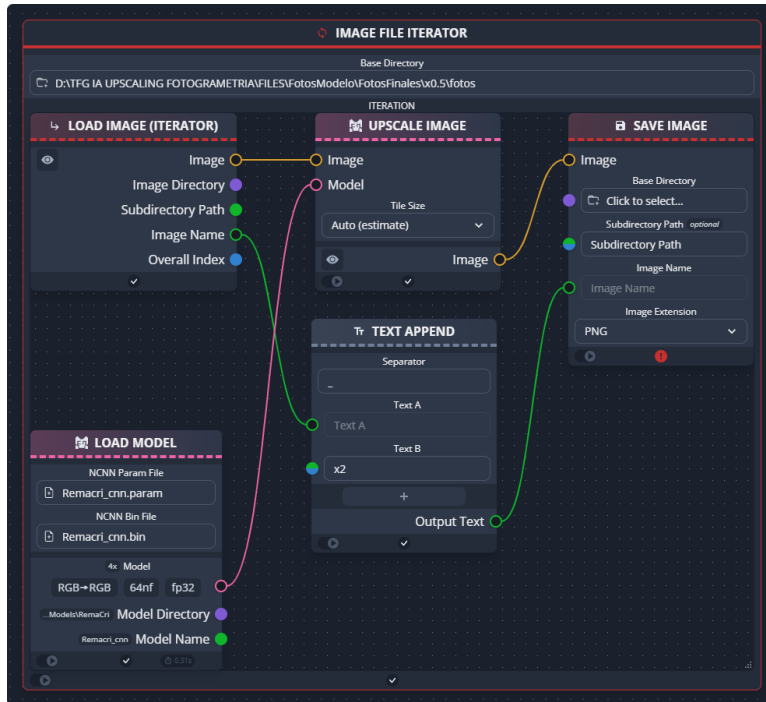


Ilustración 12: Flujo de trabajo para obtención del set de fotos x2 en chaiNNer. Fuente: Propia.

Se aprecia que, conforme aumenta la resolución de las fotografías base a reescalar, aumenta el tiempo de procesado. En este caso 48 segundos por foto.

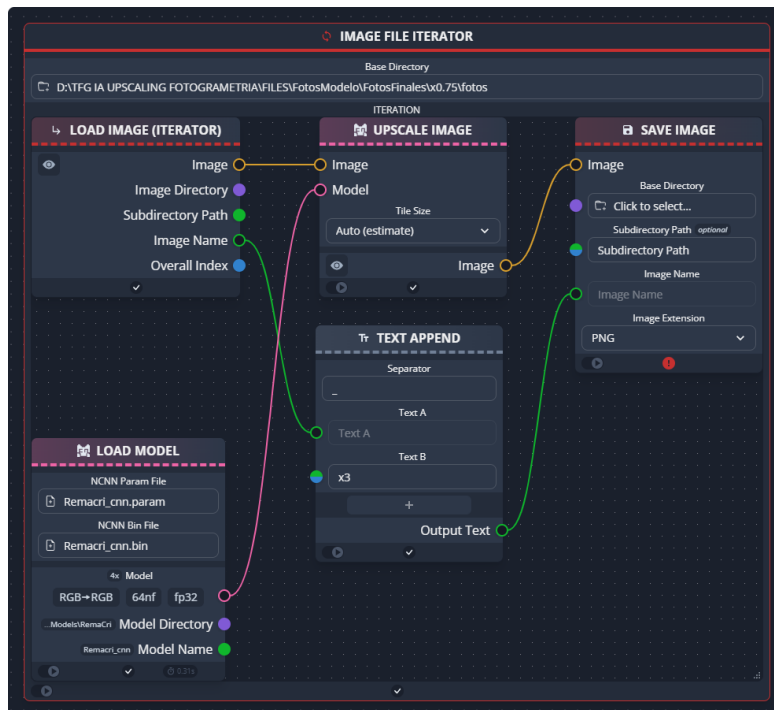


Ilustración 13: Flujo de trabajo para obtención del set de fotos x3 en chaiNNer. Fuente: Propia.

Tiempo de procesado promedio por foto de 2 minutos.

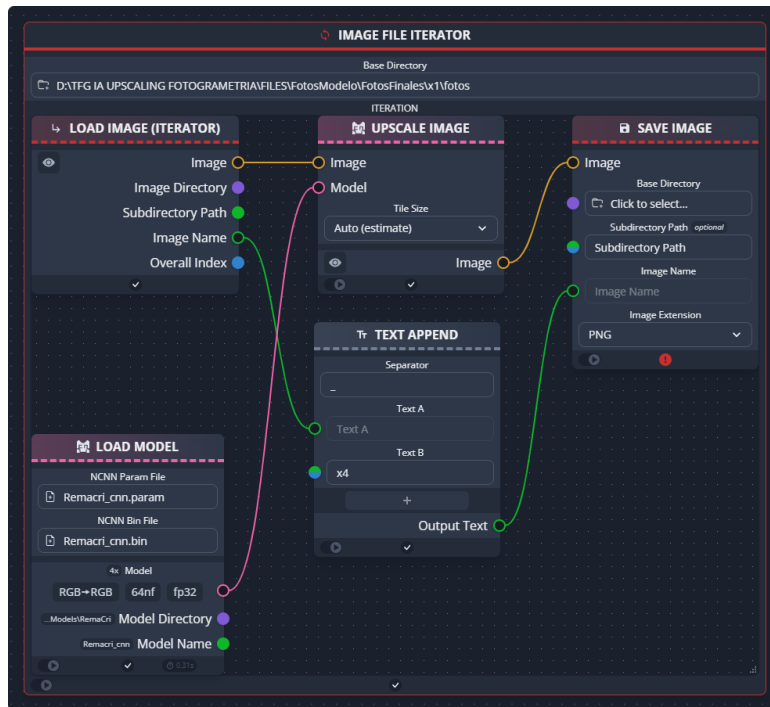


Ilustración 14: Flujo de trabajo para obtención del set de fotos x4 en chainNer. Fuente: Propia.

La obtención del conjunto de fotos x4 es el que más tiempo requiere, con un tiempo promedio de procesado de 3.2 minutos por fotografía. Teniendo en cuenta que el conjunto es de 75 fotos, solo obtener este set ha requerido de un tiempo de procesamiento de 4 horas.

En total, el proceso completo de *upscaling* y obtención de todas las resoluciones deseadas costó aproximadamente 16 horas de procesamiento.

Los resultados, si bien no se aprecian visualmente a simple vista, se hacen muy evidentes al hacer zoom a la misma zona en las distintas resoluciones:



Ilustración 15: Comparación de tres fotos (x0.25, x1 y x4 respectivamente) al hacer zoom al mismo elemento. Fuente: Propia.

4.4 Obtención de los modelos 3D por fotogrametría

Para la obtención de los diferentes modelos 3D con distintas resoluciones y calidades se ha utilizado el software *Metashape*.

Agisoft Metashape es un programa independiente que realiza un procesado fotogramétrico de imágenes digitales y genera datos espaciales 3D para ser utilizados en ámbitos como los sistemas geográficos de información y efectos visuales, entre otros.

El flujo de trabajo en metashape es el siguiente:

4.4.1 Importar y alinear fotografías

Importar las fotografías consiste en cargarlas dentro del software para que estas puedan ser procesadas. Se pueden cargar los archivos de imagen por separado, o se puede seleccionar la carpeta que contenga todas las imágenes. Se ha optado por lo segundo.

Una vez cargado el conjunto de imágenes, se orientan y alinean las imágenes con el comando *align photos*.

El programa se refiere a alinear fotografías al proceso de determinar la posición y orientación espacial de la cámara para cada imagen. Esta información viene definida por los parámetros de orientación interna y externa de la cámara. Los parámetros de orientación externa definen la posición y rotación de la cámara respecto a un sistema de coordenadas global común para todas las cámaras. Los parámetros de orientación interna definen la geometría de los componentes internos de la cámara tales como el punto central, la distancia desde la lente a dicho punto central, y las distorsiones de la lente.

Previo a comenzar el procesado el programa necesita una aproximación inicial de los parámetros de orientación interna y externa. Dicha aproximación se realiza de forma automática a partir de los datos EXIF (metadatos) presentes en cada archivo de fotografía.

Al ejecutar el comando, se realiza una búsqueda de *feature* o *key points* entre imágenes mediante algoritmos basados en *scale invariant feature transform (SIFT)*. Dichos algoritmos comienzan buscando *binary large objects (blobs)* en las imágenes, se trata de agrupaciones grandes de píxeles con un valor parecido.

Posteriormente se calcula el gradiente para cada píxel de dicho *blob* y se genera un histograma que agrupa todas las direcciones de cada píxel, se establece como orientación del *blob* a la dirección más repetida en el histograma. Esto soluciona el problema de que el mismo *blob* en cada imagen puede aparecer con distintas orientaciones, es decir, permite deshacer las rotaciones entre *blobs* homólogos.

Además, los histogramas son normalizados con el fin de lograr que la escala de las imágenes tampoco influya en la detección de *blobs*.

Comparando dichos histogramas normalizados es que se consigue correlacionar *blobs* entre imágenes. Los *key points* son el resultado de calcular el centroide de cada *blob*. Cuando la correlación entre *blobs* es satisfactoria, dichos *key points* pasan a considerarse *tie points*

Una vez obtenidos los *tie points* se realiza un *bundle block adjustment (BBA)*, como resultado de este proceso, se obtienen parámetros de orientación externa más fiables y precisos que la aproximación inicial determinada a partir de los metadatos de las fotos. Además de una nube de puntos poco densa o *sparse point cloud*.

Todo este proceso se conoce como *structure from motion (SFM)*.

Los parámetros escogidos para el alineamiento son los siguientes:

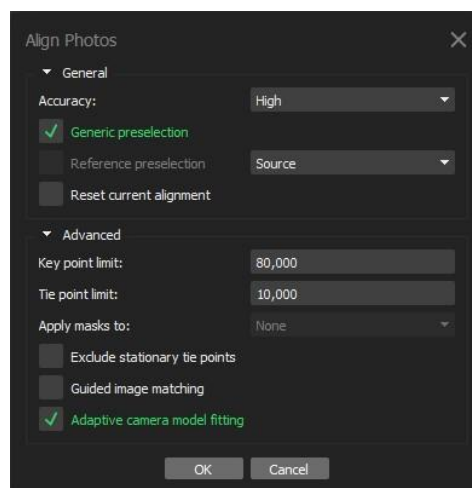


Ilustración 16: Parámetros del comando *Align Photos* dentro de *Metashape*. Fuente: Propia.

- *Accuracy: High*. Sirve para obtener estimaciones más precisas de la posición de la cámara. Al establecer la calidad en alto, el programa trabaja con las fotografías en tamaño original. Al aplicar peores calidades, metashape aplica un *downscaling* por un factor de 4 (2 por cada lado) al set de fotos utilizado por cada calidad inferior consecutiva.
- *Generic preselection: Checked*. Alinear las fotografías puede resultar un proceso costoso en cuanto al tiempo de cálculo, gran parte de este tiempo se debe al proceso de correlacionar elementos detectados entre fotos. Esta opción puede reducir el tiempo empleado debido a que se selecciona previamente un subconjunto de imágenes para correlacionar elementos.
- *Reference preselection: Unchecked*. Para reducir el tiempo del alineamiento si se cuenta con datos externos de las posiciones de las cámaras.
- *Key point limit: 80,000*. Indica el límite superior de *feature points* a tener en cuenta en cada imagen para la fase de procesamiento actual. Se ha

establecido en 80,000 por considerar que puede ofrecer mejores resultados en los casos en los que se procesen fotografías de resoluciones altas.

- *Tie point limit: 10,000.* Indica el límite superior de homólogos detectados por imagen.
- *Exclude stationary tie points: Unchecked.* Excluye *tie points* que permanecen invariables entre las distintas fotografías. Esta opción permite alinear las fotografías en el escenario en que la cámara permanece fija y el objeto está posado sobre una base rotatoria.
- *Guided image matching: Unchecked.* Permite establecer una cantidad de *tie points* proporcional a los megapíxeles de cada imagen. Se ha decidido desmarcar esta opción debido a que los conjuntos de fotos de diversas resoluciones varían enormemente en cuanto a número de píxeles. Por lo que en imágenes de poca resolución se determinarían muy pocos *tie points*, y en imágenes de alta resolución se produciría un exceso.
- *Adaptative camera model fitting: Checked.* Permite la selección automática de parámetros de la cámara para ser incluidos en el ajuste inicial basado en sus estimaciones de calidad. Esto ayuda a ajustar más parámetros durante al alineamiento inicial de la cámara.

4.4.2 Generar malla 3D

Tras obtener la nube de puntos poco densa (formada por *tie points*) se debe generar el modelo 3D del objeto mediante el comando *build mesh*.

Este comando realiza un proceso conocido como *multi-view stereo (MVS)*. Las técnicas MVS utilizan los parámetros de orientación de cada imagen obtenidos previamente para refinar el resultado producido, dando lugar a una nube de puntos densa o, en su defecto, mapas de profundidad.

La mejora en eficacia a la hora de detectar elementos homólogos en distintas imágenes con las técnicas MVS respecto a las SFM radica en que, conocidos los parámetros de orientación interna y externa de la cámara para cada imagen, y sabiendo que el mismo punto de un mismo elemento en dos imágenes distintas se encuentra en una línea (principio de la geometría epipolar), se transforman y re proyectan pares de imágenes y se consigue reducir la búsqueda de elementos homólogos entre estas de un entorno plano (2D) a una línea (1D).

La mejora en la cantidad de *tie points* detectados es lo que da lugar a la nube de puntos densa o *dense point cloud*.

Los mapas de profundidad son imágenes que representan la profundidad de cada píxel respecto al sistema local de coordenadas de su fotografía base, se obtienen calculando el desplazamiento de puntos homólogo entre pares de imágenes (disparidad o paralaje). A partir de dicha disparidad se puede calcular la profundidad en cada píxel.

En base a los mapas de profundidad (uno por cada fotografía), el programa reconstruye mediante algoritmos internos la geometría 3D del objeto y produce como resultado una malla o modelo tridimensional del mismo.

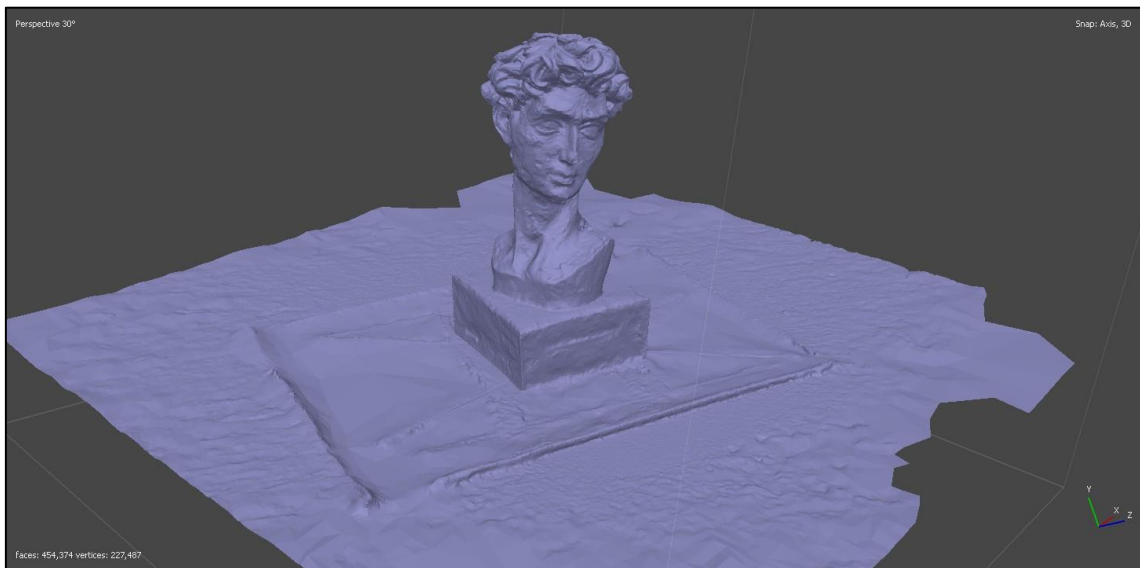


Ilustración 17: Modelo 3D generado en Metashape. Fuente: Propia.

Los parámetros establecidos para la generación de la malla son los siguientes:

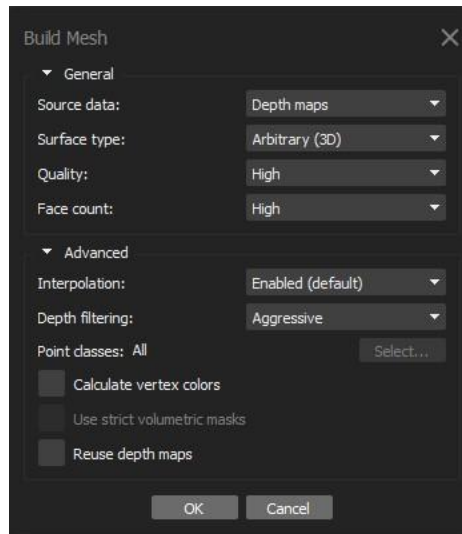


Ilustración 18: Parámetros del comando Build Mesh dentro de Metashape. Fuente: Propia.

- *Source data: Depth maps.* Datos a partir de los cuales se va a generar el modelo 3D. Se puede escoger utilizar la nube poco densa de puntos para hacer una primera aproximación del resultado.
- *Surface type: Arbitrary (3D).* Adecuado para modelar cualquier tipo de objeto. Pensado para objetos cerrados tales como estatuas, edificios, etc. Es posible seleccionar *Height field* para aquellos casos en que se busca reconstruir superficies planas como modelos del terreno.
- *Quality: Variable para cada resolución.* Especifica la calidad deseada en la reconstrucción de los mapas de profundidad. Si selecciona *ultra high*, el algoritmo trabaja con las imágenes en su resolución original, cada calidad consiguiente implica realizar un *downscaling* por un factor de 4 (2 por cada lado).
- *Face count: High.* Permite decidir de forma genérica la cantidad del número de caras, y por tanto vértices, del modelo 3D resultante.
- *Interpolation: Enabled.* Al estar activado, el programa interpolará superficies dentro de un círculo de un determinado radio en rededor de cada punto de la nube de puntos.
- *Depth filtering: Aggressive.* El filtrado agresivo es capaz de ofrecer mejores resultados si se aprecia un exceso de geometría (tales como mallas aisladas alrededor del objeto reconstruido). Se ha escogido esta opción por el modelo 3D de las fotografías de mayor calidad presentar una gran cantidad de excesos de geometría.
- *Reuse depth maps: Unchecked.* Permite reutilizar los mapas de profundidad si estos han sido generados previamente. Al activar esta opción se establece de forma predeterminada un filtrado de mapas de profundidad suave.

4.4.3 Escalar modelo

Con el objetivo de poder escalar el modelo 3D generado, se han realizado y medido unas marcas con forma de línea en la base sobre la que se apoya el objeto. Debido a que las marcas están escritas sobre una superficie plana, es necesario generar la textura del modelo (comando *build texture*) para poder diferenciarlas y, por ende, escalar el modelo en consecuencia.

El proceso seguido para escalar los modelos, de forma general, es el siguiente:

- 1- Se genera la textura del modelo para poder apreciar las marcas
- 2- Se crean puntos o *markers* en cada extremo de las marcas
- 3- Se seleccionan los *markers* en pares que coincidan con las marcas presenten, y se crea una *scale bar* por cada par.
- 4- Con las *scale bars* creadas, se define la distancia que miden en la realidad
- 5- Una vez todas las *scale bars* están definidas con las distancias que les corresponde, se actualiza la transformación geométrica del modelo y este queda escalado.

Se ha repetido el mismo proceso de generación de malla y escalado para todas las resoluciones obtenidas (x0.25, x0.50, x0.75, x1, x2, x3, x4) y para todas las calidades de mapas de profundidad que posibilitan el hardware utilizado. Estas son: *high, medium, low, lowest*. En total se ha hecho este proceso 28 veces, más 1 con el set de fotos x1_original en calidad *high*.

El escalado de algunos modelos obtenidos a partir de fotografías de mayor resolución (x3 y x4 con calidad de los Depth maps en *high*) ha debido realizarse siguiendo otra metodología, debido a que presentaban agujeros en la malla en la zona del soporte del objeto donde se encontraban las marcas para escalado, por lo que no es posible establecer un *marker* en uno de los dos extremos de las líneas y, por ende, no se pueden crear *scale bars* fiables.

Para escalar dichos modelos, se ha medido la distancia de esquina a esquina de la base del objeto en la vida real con un pie de rey. Posteriormente y dentro del software, se ha creado un *marker* en cada una de las dos esquinas medidas y se ha definido una *scale bar* a la que se le ha introducido la distancia medida con el pie de rey. Tras esto se actualiza la transformación geométrica y los modelos quedan escalados.

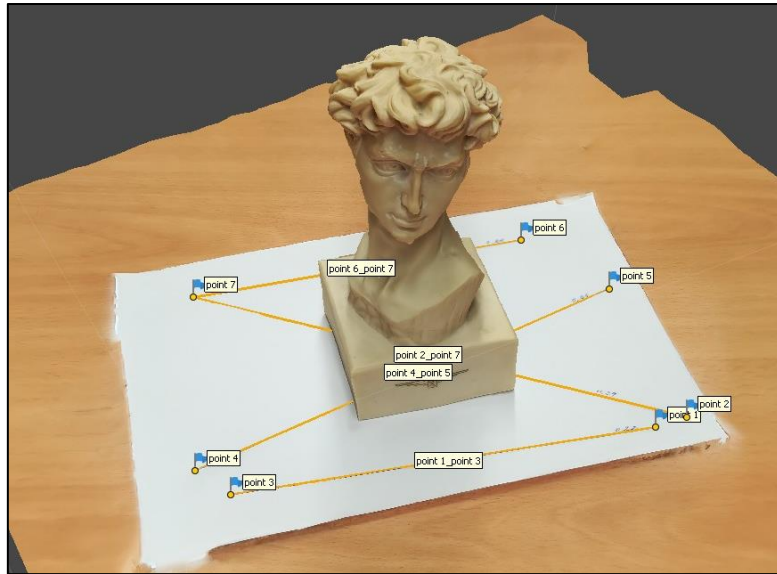


Ilustración 19: Markers y scale bars en el modelo para escalarlo. Fuente: Propia.

Si bien esta metodología permite menos precisión que el uso de marcas, la magnitud del error generado es despreciable debido al objetivo que se busca conseguir con este trabajo.

4.4.4 Refinar malla

Los modelos generados por fotogrametría presentan una base plana sobre la que se apoya el objeto. Dicho plano de apoyo no se encuentra presente en el modelo generado por el escáner de luz estructurada, lo que puede conducir a errores en el alineamiento de los modelos durante la fase de comparación de mallas.

Por este motivo, los modelos obtenidos por fotogrametría deben refinarse eliminando los vértices de la malla correspondientes a las partes del modelo que no son comunes con el modelo generado por el escáner.

Para eliminar dichos vértices simplemente se seleccionan manualmente en cada modelo y se borran.



Ilustración 20: Malla refinada. Fuente: Propia.

Una vez terminados los modelos, se exportan como archivos de formato .obj, pues es un formato de archivo de uso altamente extendido que facilita la interoperabilidad con otros programadas que vean involucrados modelos 3D digitales.

4.5. Modelo 3D de referencia por luz estructurada

Una vez obtenidos los modelos 3D que involucran IA, es necesario poseer un modelo de referencia tomado como verdadero respecto al cual poder realizar comparaciones.

Además de la fotogrametría, existen dos principales tecnologías que capacitan la captura tridimensional de elementos. Estas son: Lidar y luz estructurada.

Lidar es el acrónimo para *Ligth Detection and Ranging*, se trata de una técnica de *remote sensing*. Utiliza luz en forma de un láser que emite pulsos a una superficie y estos se ven reflejados de vuelta y captados por un sensor. La distancia que recorre el láser se puede calcular o bien por tiempo de vuelo del láser o bien por diferencia de fase. Estos pulsos de luz, en combinación con otros datos captados por el sistema del aparato que esté emitiendo y recibiendo el láser, genera información tridimensional precisa de la superficie.

El dispositivo que utiliza la tecnología lidar para la toma de datos desde la superficie (pues el lidar puede ser aerotransportado) se denomina *laser scanner*. Si bien este dispositivo resulta muy eficaz a la hora de captar información

tridimensional de objetos relativamente grandes como edificios, calles, estatuas, etc. No está pensado para objetos pequeños que se deban medir a una distancia próxima.

Por otro lado, la luz estructurada es la tecnología empleada por el dispositivo que recibe el nombre de escáner de luz estructurada. Se trata de un dispositivo que emite un patrón en pulsos de luz a alta frecuencia el cual ha sido perfectamente medido y calibrado sobre una superficie plana, por lo que cuando este es proyectado sobre la superficie a reconstruir, el sensor receptor del escáner capta un patrón deformado. Al comparar y cuantificar dichas deformaciones con respecto al patrón sin deformar, se puede deducir matemáticamente la forma de la superficie. El resultado de la medición es la obtención directa de un modelo 3D del objeto.

El escáner de luz estructurada está pensado para medir objetos relativamente pequeños, estos poseen una gran precisión volumétrica, y una velocidad de captura de datos elevada. Es por ello que resulta idóneo para la obtención del modelo de referencia.

El escáner de luz estructurada escogido para esta empresa se trata del CreaForm Academia 50, por su disponibilidad y precisión suficiente para el trabajo a desarrollar, pues posee una precisión volumétrica de 0.5mm.

ACADEMIA 3D SCANNER™	
Part size range (recommended)	0.3 – 3.0 m (1 – 10 ft)
Accuracy	Up to 0.250 mm (0.01 in)
Resolution	0.500 mm (0.020 in)
Measurement rate	550,000 measurements/s (1)
Volumetric accuracy (based on part size)	0.500 mm/m (0.006 in/ft)
Scanning area	380 x 380 mm (15.0 x 15.0 in)
Stand-off distance	400 mm (15.75 in)
Depth of field	250 mm (10.0 in)
Light source	White light (LED)
Positioning methods	Geometry and/or targets
Weight	850 g (1.9 lb)
Dimensions (LxWxH)	96 x 140 x 258 mm
Operating temperature range	5–40 °C (41–104 °F)
Operating humidity range (non-condensing)	10–90%
Certifications	EC Compliance (Electromagnetic Compatibility Directive, Low Voltage Directive), IP50, WEEE

Ilustración 21: Especificaciones técnicas CreaForm Academia 50. Fuente: Emco UK.



Ilustración 22: Escáner CreaForm Academia 50. Fuente: CreaForm.

Para operar el escáner y exportar el modelo 3D se ha utilizado el software de la misma casa (*CreaForm*) *VX Elements*.

El escáner se debe calibrar apuntando a un patrón propio. Posteriormente se deben poner unos objetivos específicos en derredor al objeto para que el escáner pueda tener una orientación espacial relativa mientras escanea.

Como resultado se obtiene un modelo 3D preciso del objeto desprovisto de escala.



Ilustración 23: Modelo de referencia obtenido con el escáner de luz estructurada. Fuente: Propia.

Para escalar el modelo generado, se han medido en dicho modelo las mismas dos esquinas que se han medido previamente para el escalado de los modelos *x3_high* y *x4_high* y se ha calculado el factor de escala necesario a aplicar. Como último paso, se debe exportar a formato .obj.

4.6 Comparación de modelos

Una vez obtenidos todos los datos necesarios, solo resta realizar la comparación. Esta tarea se ha llevado a cabo con el software de metrología *Geomagic Control X*.

La filosofía de trabajo en este programa consiste en que un modelo 3D se debe establecer como objeto referencia, al hacer esto, dicho objeto quedará inmutable y todos los alineamientos y comparaciones se harán respecto a él.

4.6.1 Alineación de modelos

El primer paso en este proceso consiste en cargar ambos modelos (previamente escalados) y mover el modelo de referencia obtenido con el escáner al espacio de *datos de referencia*. Los modelos obtenidos por fotogrametría permanecen en el espacio *datos medidos*.

Una vez cargados los modelos, se realiza una primera alineación inicial. El algoritmo analiza rápida y someramente elementos en el modelo medido y los hace coincidir con los datos de referencia. Se trata de un algoritmo propietario que no presenta más información de su funcionamiento interno. Esta alineación inicial no requiere de ninguna alineación previa, es decir, no importa cuánto difiera la posición espacial y rotación del modelo medido respecto a la referencia. Sino que esta alineación inicial procura dicha alineación previa.

En los casos en los que el modelo 3D medido difiere en gran medida con el de referencia (e.g. el modelo *x0.25_lowest*, o si no se hubiese refinado la malla eliminando la base sobre la que está apoyada el objeto) puede resultar en una alineación inicial poco satisfactoria que puede inducir a producir resultados erróneos en la futura alineación precisa.

En estos casos en los que el modelo medido no se prealinea correctamente, se debe realizar un alineamiento de forma manual, trasladando y rotando el modelo medido en consecuencia para que se su posición coincida de la mejor forma posible con el modelo de referencia.



Ilustración 24: Alineamiento inicial erróneo por la falta de correspondencia geométrica entre modelos.
Fuente: Propia.

Una vez el modelo medido está correctamente prealineado al de referencia, se realiza una *alineación de ajuste perfecto*. Se trata de un proceso computacional costoso pero que minimiza la desviación general entre ambos modelos y sí que requiere de una alineación previa.

Si bien el programa utiliza algoritmos propietarios y no especifica en qué fundamentos se basa o qué métodos sigue, por lo general estos alineamientos 3D precisos parten todos de una técnica denominada *Iterative Closest Point (ICP)*. El funcionamiento de dicha técnica a grosso modo es el siguiente:

- 1- Con los dos modelos prealineados, se realiza una primera asociación o transformación del modelo medido para alinearlo al de referencia comparando los vértices más cercanos utilizando el algoritmo *simple nearest neighbour*. Es decir, se mueve y rota el modelo medido (sin deformarlo) para que los vértices de dicho modelo coincidan, en la medida de lo posible, con sendos vértices más cercanos del modelo de referencia.

- 2- A partir de la primera asociación de datos, se intenta encontrar la transformación geométrica tal que minimice las distancias entre cada par de vértices relacionados (mínimos cuadrados). Para ello primero se calcula el centro de masas de ambas nubes de puntos (los vértices de los modelos 3D) y se hacen coincidir, moviendo el modelo medido en consecuencia. Y posteriormente se calcula la matriz de rotación óptima a aplicar gracias un complejo algoritmo llamado *singular value decomposition (SVD)*. Uno de los usos del SVD es que permite resolver sistemas lineales de ecuaciones matriciales $Ax = b$ para matrices A no

cuadradas, esto resulta de especial utilidad en modelos de regresión lineal.

3- Una vez aplicada la transformación, se iteran o repiten los pasos anteriores, partiendo cada vez del resultado del paso anterior, hasta que el error mínimo cuadrático en el ajuste de ambos modelos resulte invariable.

Es importante aclarar que, en el caso de tener que realizar una alineación previa manual, esto no influye en el resultado obtenido por la alineación de precisión, pues se trata de un resultado obtenido por mínimos cuadrados.

4.6.2 Comparación de modelos

Una vez el modelo medido se encuentra perfectamente alineado con respecto al modelo de referencia, se procede a realizar la comparación. Esto se lleva a cabo con el comando *comparación en 3D*.

Esta herramienta analiza la desviación entre el modelo medido y el modelo de referencia proyectando todos los vértices emparejados sobre los datos de referencia. La desviación se muestra como un mapa de colores que permite identificar tanto la magnitud como el sentido de las desviaciones.

Los parámetros utilizados son los siguientes:

- Relación de muestreo: 100%. Porcentaje de vértices del modelo medido a utilizar en la comparación.
- Método: Forma. Pues interesa las desviaciones entre las caras del modelo, no las diferencias en grosor.
- Dirección de proyección: Más corto. Proyecta los puntos o vértices del modelo medido sobre la superficie del modelo de referencia a través de la distancia más corta. La medida de esa distancia es el resultado de la desviación. La opción *a lo largo de la normal* permite medir las distancias a lo largo de los vectores normales de cada vértice del modelo medido.
- Rango máximo de barra de colores: $\pm 4\text{mm}$. Se ha decidido que este se trata de un rango óptimo que permite apreciar las desviaciones tanto en modelos que se desvían mucho como en los que tienen un alto porcentaje de correlación.
- Tolerancia específica: $\pm 0.5\text{mm}$. Por ser esta la precisión volumétrica que ofrece el escáner de luz estructurada. Todos los vértices que difieran menos que dicha cantidad se pueden considerar virtualmente idénticos a los vértices del modelo de referencia.
- Resto de parámetros: Por defecto. No se ha considerado necesario aplicar ningún cambio pues los parámetros preestablecidos resultan adecuados.

Como resultado de la comparación se obtienen diversas métricas, la que más interesa para el objetivo de este trabajo es el porcentaje de vértices que entra en tolerancia con respecto al modelo de referencia.

5. RESULTADOS

5.1 Fotografías

A continuación se muestra una tabla que recoge información acerca de los sets de fotografías obtenidos reescalando por inteligencia artificial:

Nombre conjunto	Resolución	Tamaño promedio por foto	Tamaño conjunto	Tiempo de reescalado
x0.25	744 x 992	724 KB	54.3 MB	1 min
x0.50	1488 x 1984	2.25 MB	169 MB	16 min
x0.75	2232 x 2976	5.17 MB	388 MB	16 min
x1	2976 x 3968	8.59 MB	644 MB	16 min
x2	5952 x 7936	25.87 MB	1.94 GB	1 h
x3	8928 x 11904	65.6 MB	4.92 GB	2.5 h
x4	11904 x 15872	118.27 MB	8.87 GB	4 h

Tabla 3: Información acerca de los conjuntos de fotos. Fuente: Propia.

En las siguientes tablas se recogen datos aproximados de los tiempos de procesamientos para cada resolución en cada calidad de mapa de profundidad:

Tiempos de procesamiento para mapas de profundidad en calidad *lowest*:

	Alineamiento	Mapas de profundidad	Malla	Textura	Total
x0.25	14 s	4 s	1 s	26 s	45 s
x0.50	34 s	7 s	1 s	51 s	1.5 min
x0.75	1.5 min	9 s	2 s	1 min	2.5 min
x1	2 min	14 s	3 s	2 min	4 min
x2	4.5 min	40 s	8 s	3 min	8.5 min
x3	6 min	1.5 min	16 s	19 min	26.5 min
x4	8.5 min	6.5 min	23 s	30 min	45 min

Tabla 4: Tiempos de procesamiento para mapas de profundidad en calidad *lowest*. Fuente: Propia.

Tiempos de procesamiento para mapas de profundidad en calidad *low*:

	Alineamiento	Mapas de profundidad	Malla	Textura	Total
x0.25	14 s	4 s	1 s	38 s	57 s
x0.50	34 s	8 s	3 s	3 min	4 min
x0.75	1.5 min	15 s	6 s	2.5 min	4 min
x1	2 min	23 s	9 s	6.5 min	9 min
x2	4.5 min	1 min	25 s	35 min	41 min
x3	6 min	2.5 min	45 s	17.5 min	21 min
x4	8.5 min	8 min	1 min	33 min	50 min

Tabla 5: Tiempos de procesamiento para mapas de profundidad en calidad *low*. Fuente: Propia.

Tiempos de procesamiento para mapas de profundidad en calidad *medium*:

	Alineamiento	Mapas de profundidad	Malla	Textura	Total
x0.25	14 s	6 s	3 s	28 s	51 s
x0.50	34 s	17 s	9 s	1 min	2 min
x0.75	1.5 min	30 s	16 s	1 min	3 min
x1	2 min	48 s	24 s	3 min	6 min
x2	4.5 min	3 min	1 min	4 min	12.5 min
x3	6 min	9 min	2 min	16 min	33 min
x4	8.5 min	19 min	2.5 min	45 min	75 min

Tabla 6: Tiempos de procesamiento para mapas de profundidad en calidad medium. Fuente: Propia.

Tiempos de procesamiento para mapas de profundidad en calidad *high*:

	Alineamiento	Mapas de profundidad	Malla	Textura	Total
x0.25	14 s	14 s	9 s	46 s	1.5 min
x0.50	34 s	41 s	24 s	2 min	3.5 min
x0.75	1.5 min	1.5 min	45 s	2 min	6 min
x1	2 min	2.5 min	1 min	2.5 min	8 min
x2	4.5 min	13 min	3 min	6 min	26.5 min
x3	6 min	29.5 min	4.5 min	31 min	71 min
x4	8.5 min	53 min	7 min	43 min	111.5 min

Tabla 7: Tiempos de procesamiento para mapas de profundidad en calidad high. Fuente: Propia.

Cabe destacar que los tiempos de procesamiento obtenidos están directamente influidos por el hardware utilizado. Dicho hardware se encuentra constatado al inicio del presente documento.

5.2 Modelos 3D por fotogrametría

Los resultados de los modelos 3D en las distintas resoluciones y calidades se han recogido en la siguiente tabla:

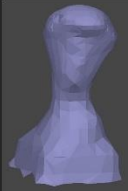























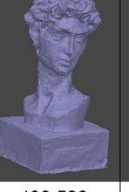
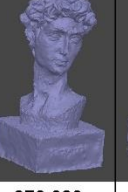

Conjunto de fotos	x0.25	x0.50	x0.75	x1	x2	x3	x4	
Depth map quality	Lowest							
		396	1,425	2,400	5,048	16,988	32,144	51,507
	Low							
		1,375	5,212	8,710	17,003	51,882	107,539	164,118
	Medium							
5,145		17,316	29,124	51,755	160,025	353,321	501,824	
High								
	17,426	53,918	98,901	155,902	499,538	870,620	979,181	

Ilustración 25: Tabla comparativa de los distintos modelos 3D obtenidos. Fuente: Propia.

El número debajo de cada figura indica el número de vértices del modelo 3D.

En los resultados de los modelos se aprecia a simple vista una mejora en cuanto al detalle que los sets de fotos a los que se les ha aplicado *upscaling* han logrado obtener.

Se aprecia que, como es obvio, la resolución de las imágenes de entrada está directamente relacionada con la cantidad de vértices que posee el modelo resultante, y la cantidad de vértice a su vez está relacionada en cierta forma con el detalle conseguido.

La cantidad de detalle como tal no es algo que pueda ser cuantificado pues se trata de un concepto abstracto, pero sin duda sí se aprecia una mejora, esto resulta especialmente evidente a lo largo de los modelos de la calidad *lowest*.

Si bien el detalle conseguido aumenta conforme aumenta la resolución de las fotografías, en los conjuntos de fotos de altas resoluciones y cuando la calidad de los mapas de profundidad es alta, se comienzan a generar artefactos o pequeñas mallas aisladas del modelo principal, así como deformaciones y ruido:

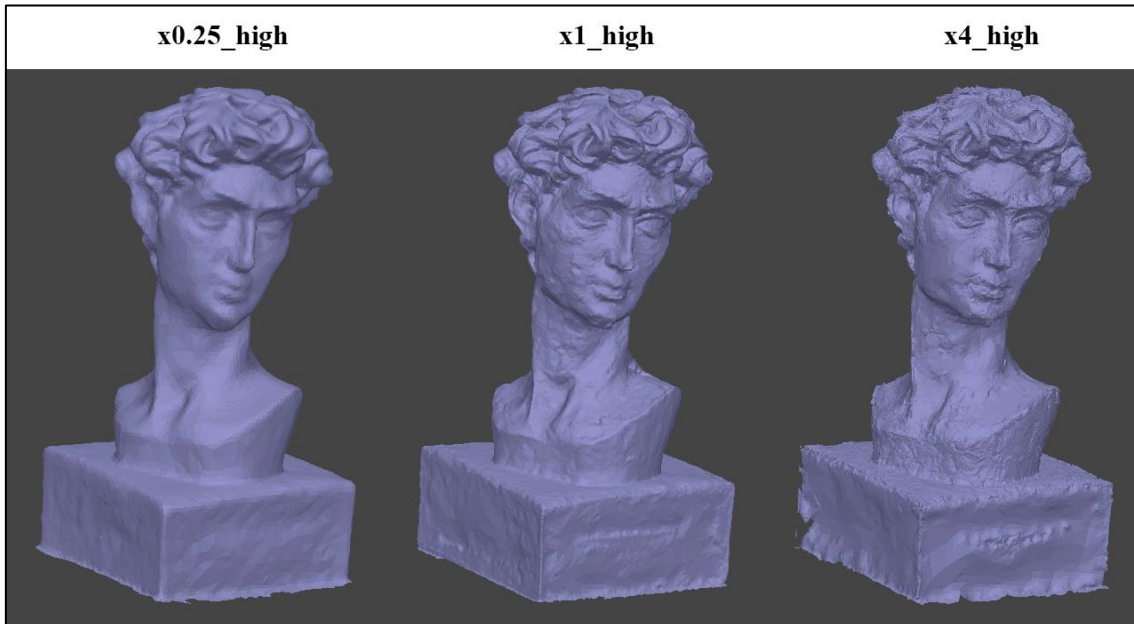


Ilustración 26: Modelos obtenidos con diferentes resoluciones en los que se aprecia un aumento del detalle así como del ruido. Fuente: Propia.

5.3 Comparación entre modelos

De forma similar, la siguiente tabla presenta las comparaciones de cada uno de los modelos mostrados en la tabla anterior con respecto al modelo de referencia obtenido con el escáner de luz estructurada:

Conjunto de fotos	x0.25	x0.50	x0.75	x1	x2	x3	x4	Diferencia (mm)
Depth map quality								
Lowest	 13.67%	 29.97%	 42.25%	 50.88%	 74.70%	 66.80%	 73.71%	
Low	 32.44%	 51.63%	 59.99%	 74.94%	 82.74%	 69.06%	 76.18%	
Medium	 51.36%	 71.55%	 74.15%	 81.51%	 83.34%	 66.57%	 66.86%	
High	 69.25%	 78.59%	 77.55%	 83.30%	 80.66%	 65.35%	 75.38%	

Ilustración 27: Tabla comparativa de las comparaciones realizadas entre los modelos obtenidos por IA y el modelo de referencia. Fuente: Propia.

El porcentaje debajo de cada figura representa el porcentaje de vértices del modelo obtenido por IA que se desvían menos de 0.5mm respecto del modelo de referencia y, por tanto, se pueden considerar virtualmente idénticos. Los porcentajes en negrita indican el porcentaje más alto, y por tanto mejor, para cada calidad de mapas de profundidad.

Si bien la correlación entre los modelos generados por IA y el modelo de referencia aumenta conforme aumenta la resolución, contrario a la hipótesis de partida, al aumentar mucho la resolución, lejos de converger la mejora y estabilizarse, comienza de nuevo a empeorar la correlación respecto al modelo de referencia. Esto se debe con total seguridad a los artefactos generados en resoluciones altas mencionados anteriormente:

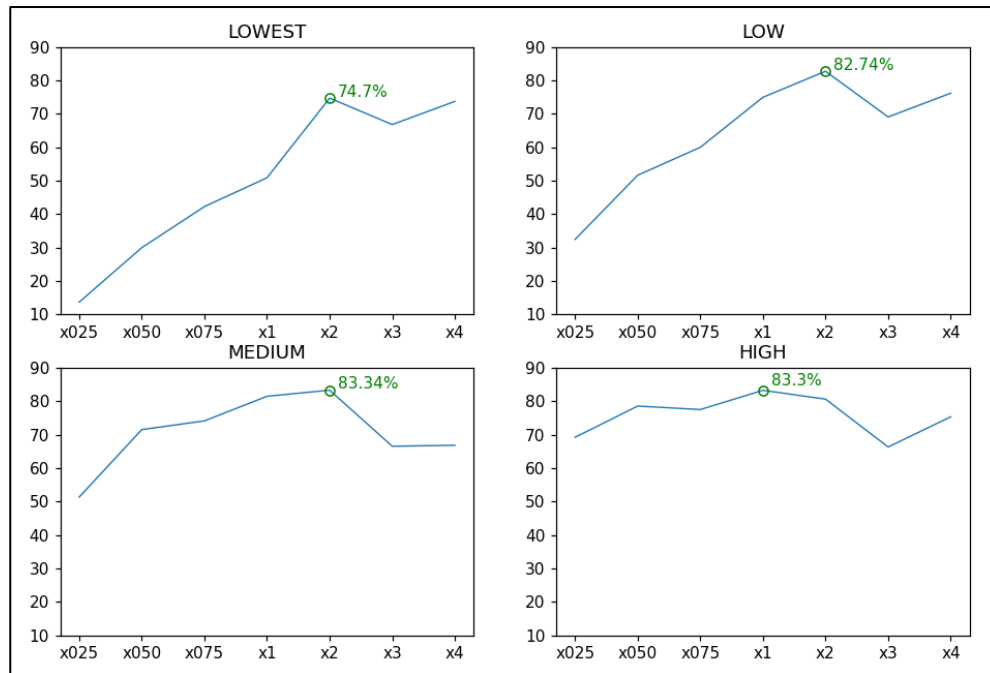


Ilustración 28: Evolución del porcentaje dentro de tolerancia a lo largo de las resoluciones para distintas calidades de mapas de profundidad. En verde se marcan los máximos Fuente: Propia.

Por lo que en lugar de determinar un umbral a partir del cual el modelo ya prácticamente no mejora porque converge, es más adecuado definirlo como determinar un umbral a partir del cual el modelo vuelve a empeorar. El *factor de reescalado máximo útil*.

Debido a la falta de estudio con objetos diferentes sobre esta metodología, no es posible afirmar que dicho factor máximo útil sea el mismo para todos los conjuntos de fotos. En este caso y para este objeto, el factor que mejores resultados procura, el factor de reescalado máximo útil, está entre 4 y 8. No confundir con el nombre que se les ha dado a los conjuntos de fotos o resoluciones. Se debe tener en cuenta que todos los conjuntos se han obtenido a partir de reescalar el set de x0.25, por lo tanto, el set de x0.25 es el set base.

La siguiente tabla expresa la correspondencia entre el nombre del set y su factor de escala respecto al set de x0.25:

Nombre del set de fotos	x0.25	x0.50	x0.75	x1	x2	x3	x4
Factor de escala respecto al set base x0.25	1	2	3	4	8	12	16

Tabla 8: Correspondencia entre los diferentes sets y sus factores de escala respecto al set base de x0.25.
Fuente: Propia.

Además de las comparaciones entre los modelos obtenidos por fotogrametría y el modelo de referencia, se ha realizado una comparación entre el modelo obtenido con el set de fotos x1_high y el modelo obtenido con el set de fotos x1_original_high, con el fin de cuantificar en cuan medida se acerca el resultado obtenido utilizando IA al resultado obtenido sin IA:

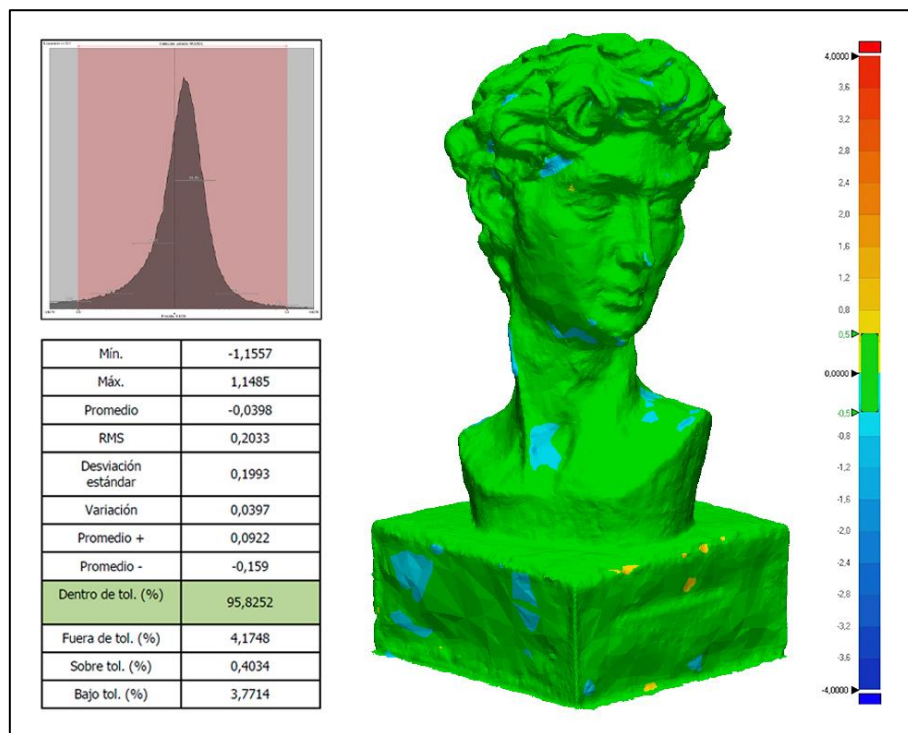


Ilustración 29: Resultado de la comparación entre el modelo x1_high y el modelo x1_original_high.
Unidades en mm. Fuente: Propia.

Como se puede apreciar, el modelo generado a partir de las fotografías x1 originales y el modelo generado a partir de las fotografías x1 reescaladas con IA, coinciden en prácticamente un 96%. Lo que revela que utilizar un *AI-Upscaler* puede producir resultados muy aproximados a la realidad.

6. PRESUPUESTO

Para el cálculo del presupuesto se ha tenido en cuenta el XX convenio colectivo nacional de empresas de ingeniería; oficinas de estudios técnicos; inspección, supervisión y control técnico y de calidad.

Además, se ha calculado bajo la premisa de que se trata de un trabajo realizado por cuenta ajena bajo una empresa contratante determinada.

<i>Esquema del agrupamiento de los puestos de trabajo por grupo profesional y nivel salarial</i>			
Grupo profesional	Nivel salarial	Personal Técnico	Personal Administrativo
		Puestos de trabajo (relación no exhaustiva)	Puestos de trabajo (relación no exhaustiva)
I	1	INGENIERO; ARQUITECTO; DOCTOR; LICENCIADO; TITULADO 2.º Y 3.º CICLO UNIVERSITARIO; GRADUADO UNIVERSITARIO CON MÁSTER UNIVERSITARIO OFICIAL HABILITANTE O MÁSTER UNIVERSITARIO OFICIAL (MÍN. 60 ECTS), CUANDO APORTE ESPECIALIZACIÓN Y COMPETENCIAS PROFESIONALES NECESARIAS PARA EL DESEMPEÑO DEL PUESTO DE TRABAJO. ANALISTA.	LICENCIADO; TITULADO 2.º Y 3.º CICLO UNIVERSITARIO; GRADUADO UNIVERSITARIO CON MÁSTER UNIVERSITARIO OFICIAL HABILITANTE O MÁSTER UNIVERSITARIO OFICIAL (MÍN. 60 ECTS), CUANDO APORTE ESPECIALIZACIÓN Y COMPETENCIAS PROFESIONALES NECESARIAS PARA EL DESEMPEÑO DEL PUESTO DE TRABAJO.
	2	GRADUADO UNIVERSITARIO; INGENIERO TÉCNICO; ARQUITECTO TÉCNICO; APAREJADOR; DIPLOMADO UNIVERSITARIO; TITULADO 1er. CICLO UNIVERSITARIO.	GRADUADO UNIVERSITARIO; DIPLOMADO UNIVERSITARIO; TITULADO 1er. CICLO UNIVERSITARIO
II	3	TÉCNICO DE CÁLCULO O DISEÑO; PROGRAMADOR INFORMÁTICO.	JEFE 1.ª ADMINISTRATIVO
	4	DELINEANTE-PROYECTISTA.	JEFE 2.ª ADMINISTRATIVO
III	5	DELINEANTE; TÉCNICO 1.ª; TÉCNICO MODELADOR BIM; TÉCNICO INFORMÁTICO.	OFICIAL 1.ª ADMINISTRATIVO; TRADUCTOR E INTÉRPRETE NO JURADO DE UNO O MÁS IDIOMAS EXTRANJEROS
	6	TÉCNICO 2.ª	OFICIAL 2.ª ADMINISTRATIVO

Ilustración 30: Agrupamiento de los puestos de trabajo por grupo profesional y nivel salarial. Fuente: XX convenio colectivo nacional de empresas de ingeniería; oficinas de estudios técnicos; inspección, supervisión y control técnico y de calidad. BOE.

Se ha considerado adecuado el nivel salarial 2 como referencia por recibir el título de ingeniero técnico al finalizar los estudios de grado.

En la siguiente tabla se muestran los salarios por convenio:

Año 2022				
Nivel salarial	Tabla salarial art. 33		Plus convenio según art. 38 convenio	Total anual
	Mes x 14	Anual		
1	1.765,51	24.717,14	2.396,68	27.113,82
2	1.331,06	18.634,84	2.396,68	21.031,52
3	1.283,52	17.969,28	2.396,68	20.365,96
4	1.176,74	16.474,36	2.396,68	18.871,04
5	1.051,43	14.720,02	2.396,68	17.116,70
6	905,87	12.682,18	2.396,68	15.078,86
7	875,48	12.256,72	2.396,68	14.653,40
8	864,28	12.099,92	2.396,68	14.496,60
9	857,97	12.011,58	2.396,68	14.408,26

Ilustración 31: Tabla salarial para el año 2022 para el convenio de ingenieros y oficina técnica. Fuente: XX convenio colectivo nacional de empresas de ingeniería; oficinas de estudios técnicos; inspección, supervisión y control técnico y de calidad. BOE.

En consecuencia, el salario bruto anual será de 21,031.52€.

En el artículo 22 del convenio se estipula la jornada ordinaria máxima de trabajo efectivo en 1,792 horas.

Por lo que el salario bruto por hora resulta en: $21,031.52 / 1,792 = 11.74€$

Teniendo en cuenta el cómputo de horas totales que se ha empleado para la realización de este trabajo:

Tarea	Tiempo (días)	Tiempo (horas)	Coste (€)
Toma de fotografías	0.5	4	46.96
Reescalado de fotografías	2	16	187.84
Obtención modelos 3D por fotogrametría	2	16	187.84
Obtención modelo 3D con escáner	0.5	4	46.96
Desarrollo de comparaciones	1	8	93.92
Total:	6	48	563.52

Tabla 9: Desglose de tareas y horas, así como su coste. Fuente: Propia.

El software empleado en el desarrollo del trabajo también se debe tener en cuenta:

Software	Tipo de software	Precio de compra	Alquilable
ChaiNNer	Software libre	Gratis	
Metashape Pro Edition	Software propietario	3212,89 €	No
Control X Essentials	Software propietario	6426,69 €	No

Tabla 10: Disponibilidad y precios del software empleado. Fuente: Propia.

Como se puede apreciar, tanto el software utilizado para llevar a cabo la obtención de modelo 3D como el utilizado para realizar las comparaciones poseen un precio de compra muy elevado y sin opción de alquiler temporal del programa. Por suerte, ambos poseen una versión de prueba completamente gratuita y que habilita todas las funcionalidades del programa. Dicha versión de prueba es válida durante 30 días para el caso de Metashape, y de 15 días para el caso de Control X. Más que suficiente en ambos casos para llevar a cabo todo el trabajo expuesto en este documento.

Por tanto, teniendo en cuenta la naturaleza puntual del presente trabajo, pues está más en línea con el ámbito de la investigación, se han utilizado las versiones de prueba de ambos programas, con lo que no ha supuesto un aumento en el coste de desarrollo del mismo.

El coste del desarrollo del trabajo para la empresa contratante es, hasta este paso: 563.52 €

Por último, falta aplicar las cotizaciones a la seguridad social.

Según el régimen general de la seguridad social para el 2023, los porcentajes a pagar a nombre de la empresa son:

Cotización	Porcentaje (%)	Relativo al coste (€)
Contingencias comunes	23.60	133
Para desempleo	5.5	31
Para FOGASA	0.2	1.13
Para formación profesional	0.6	3.38
Total:	29.9	168.51

Tabla 11: Cotizaciones a la seguridad social 2023. Fuente: (<https://www.seg-social.es>)

Como resultado, el coste total del presente trabajo es de: $563.52 + 168.51 = 732.03$ €

7. CONCLUSIONES

Las herramientas de Inteligencia artificial han probado ser muy útiles en diversos ámbitos. Con este estudio también han demostrado ser útiles en el mundo de la fotogrametría, pudiendo generar unos resultados que apenas divergen de los obtenidos con un flujo de trabajo clásico.

Este procedimiento puede resultar útil en situaciones en las que se busque mejorar la precisión de los resultados cuando se parte de un conjunto de imágenes de calidad mediocre. De todas formas, es importante tener en cuenta que se trata de un proceso que requiere una gran cantidad de tiempo y potencia computacional.

También puede utilizarse con el fin de almacenar un mayor número de fotografías (cuando no se busca una rigurosidad absoluta) si el espacio de almacenamiento es limitado, pudiendo reescalarlas más adelante para recuperar detalles.

8. BIBLIOGRAFÍA

- 3D Systems Inc. (26 de Junio de 2023). *Control X Knowledge Base*. Obtenido de 3D Systems: <https://support.3dsystems.com/s/article/Control-X-Knowledge-Base>
- Agisoft LLC. (2023). *Guía de usuario*. Obtenido de Agisoft Metashape: <https://www.agisoft.com>
- Brunton, S. &. (Febrero de 2019). *Singular Value Decomposition (SVD)*. Obtenido de ResearchGate: https://www.researchgate.net/publication/331230334_Singular_Value_Decomposition_SVD
- Duchon, C. E. (7 de Mayo de 1979). *Lanczos Filtering in One and Two Dimensions*. Obtenido de School of Meteorology, University of Oklahoma: https://www.researchgate.net/publication/252898828_Lanczos_Filtering_in_One_and_Two_Dimensions
- Foolhardy. (2021). *Remacri Upscaler Model*. Obtenido de Upscale wiki: https://upscale.wiki/wiki/Model_Database
- Joey Ballentine, M. S. (2023). *ChaiNNer*. Obtenido de GitHub: <https://github.com/chaiNNer-org/chaiNNer>
- Jonathan L. Carrivick, M. W. (2016). *Structure from Motion in the Geosciences*. Wiley Blackwell.
- Procházková, J. &. (2018). *Notes on Iterative Closest Point Algorithm*. Obtenido de Slovak University of Technology in Bratislava: https://www.researchgate.net/publication/324500004_Notes_on_Iterative_Closest_Point_Algorithm
- Shree Nayar, Columbia University. (2021). *SIFT Detector*. Obtenido de First Principles of Computer Vision, Youtube: <https://youtube.com/playlist?list=PL2zRqk16wsdqXEMpHrc4Qnb5rA1Cylrhx>
- Thomas Luhmann, S. R. (2020). *Close-Range Photogrammetry and 3D Imaging*. De Gruyter.