



**Universidad**  
Zaragoza

# Trabajo Fin de Grado

## Percepción de materiales: estudio del espacio latente

Autor

Ignacio Galve Ceamanos

Director

Diego Gutiérrez Pérez

Escuela de Ingeniería y Arquitectura

2020

# AGRADECIMIENTOS

Me gustaría agradecer la ayuda recibida de Diego Gutiérrez, y algunos miembros del grupo de investigación *Graphics and Imaging Lab* (Johanna Delanoy, Manuel Lagunas, Ana Serrano y Belén Masia), que me han guiado a lo largo del proyecto de la Beca de Colaboración, la base del TFG. En especial a Johanna Delanoy, con quien he podido comentar y discutir muchas de las ideas que iban surgiendo durante el desarrollo del trabajo.

# RESUMEN

Identificar y distinguir materiales es una tarea que los humanos llevamos a cabo de forma rutinaria, sin apenas esfuerzo. Pese a ser una tarea que nos resulta sencilla, el proceso que lleva a cabo nuestro sistema visual no es trivial, ya que la escena que observamos depende de varios factores como las fuentes de luz, la geometría del objeto o el punto de vista, y nuestro sistema visual tiene que separar dichos factores del propio material. Este problema ha sido objeto de estudio desde hace muchos años y, aunque existen numerosos trabajos sobre el tema, todavía no se ha establecido una teoría única. Dentro del estudio de la percepción de materiales, el área de la similitud entre materiales es una parte muy importante con numerosas aplicaciones, como el mapeo de gamas de colores entre dispositivos o la asignación de materiales a escenas complejas. El objetivo de este trabajo es tratar de comprender mejor el modelo interno de similitud entre materiales que tenemos en nuestro sistema visual. Al ser un campo abierto, cualquier posible avance puede ser de gran interés y servir de base para aplicaciones de reconocimiento o edición de materiales.

Para abordar el problema se han llevado a cabo diferentes análisis y pruebas sobre un espacio multidimensional construido para representar la similitud entre materiales. Este ha sido generado por una red neuronal entrenada de manera que las distancias en el espacio concuerden lo mejor posible con juicios humanos de similitud entre materiales.

En primer lugar, se han realizado diferentes clasificaciones de materiales de acuerdo con algunos atributos físicos, como su nivel de brillo o metalicidad. La mayoría de las pruebas han logrado generar hiperplanos que dividen exitosamente el espacio, separando con una tasa de acierto alta las muestras que poseen un atributo de las que no. Estos resultados apuntan a que el espacio tiene algún tipo de organización que agrupa los materiales por características, principalmente las relacionadas con el nivel de brillo.

En base a las observaciones anteriores, para comprender mejor cómo se distribuyen los materiales en el espacio, se ha buscado una forma de poder visualizarlo en dos dimensiones. El método empleado ha sido el de PCA (*Principal Component Analysis*), que permite reducir los datos a cualquier número de dimensiones obteniendo las principales direcciones de cambio en las muestras. Para todos los conjuntos de materiales analizados se ha obtenido que la componente más importante explica en torno al 90% de la variabilidad de los datos. Además, hay una visible correlación entre el nivel de brillo del material y la componente principal. Todo esto apunta a que el espacio de similitud entre materiales es más simple de lo esperado, y a que las personas damos mucha importancia a la capacidad de reflejar luz de un material para distinguirlo.

Como objetivo final, se ha desarrollado un sistema que permite reproducir el espacio latente estudiado, de manera que se pueden generar imágenes realistas de materiales a partir de puntos del espacio. Además, gracias a los resultados obtenidos de los análisis previamente comentados, se ha podido diseñar un sistema intuitivo para la edición de materiales.

## Contenido

1. Introducción .....	5
Objetivos del proyecto .....	7
2. Trabajos relacionados .....	8
2.1. Percepción humana .....	8
2.2. Reconstrucción y descomposición del espacio latente .....	8
3. Conjuntos de datos y arquitectura de la red .....	9
4. Entendiendo el espacio latente .....	11
4.1. Clasificación semántica de materiales .....	12
4.1.1. Materiales con forma e iluminación constante .....	12
4.1.2. Materiales con diversas formas e iluminaciones .....	13
4.2. Clasificación semántica con PCA .....	16
5. Análisis interno de la red .....	25
6. Reconstrucción del espacio latente .....	28
6.1. Edición material .....	31
6.1.1. Sistema de edición intuitivo .....	32
7. Descomposición semántica del espacio latente .....	35
8. Trabajo futuro .....	37
9. Conclusiones .....	38
Bibliografía.....	40
Anexo 1: Clasificación semántica.....	42
Metal .....	42
Brillo .....	43
Mate .....	44
Definición de los reflejos .....	45
Áspero .....	46
Luminosidad.....	48
Anexo 2: Arquitectura del sistema generador de imágenes.....	49
Codificador.....	49
Discriminador.....	50
Generador.....	50
Anexo 3: Estudio del espacio latente reducido con PCA .....	52
Ajuste de las componentes para los diferentes entorno y formas .....	52

Tercera componente principal .....	53
Anexo 4: Desarrollo del proyecto .....	54

## 1. Introducción

Los humanos son capaces de diferenciar materiales y extraer de ellos sus atributos de alto nivel sin apenas esfuerzo. Se trata de un proceso interno del cerebro bastante complejo que no ha sido descifrado completamente todavía. A la hora de identificar un material entran en juego varios factores como las fuentes de luz, la geometría del objeto o el punto de vista, y es el sistema visual el que tiene que separar todos esos factores del propio material. No es interesante únicamente cómo se interpreta la información recibida por los estímulos visuales, sino también cómo se agrupan los materiales de acuerdo con lo percibido. Por diferentes capacidades físicas y niveles de conocimiento de los individuos, no todas las personas perciben de la misma forma, habiendo discrepancias entre ellas. El objetivo de este trabajo es tratar de entender un poco mejor la percepción humana de los materiales. Se engloban diferentes aspectos como la forma en la que agrupamos los materiales, qué características principales detectamos para decidir el tipo de material que estamos viendo, y la posible influencia en la percepción según las condiciones de iluminación y forma en las que se presente el material.

Hoy en día existen herramientas para la clasificación y el reconocimiento de materiales, sin embargo, su entrenamiento requiere de bastante tiempo y recursos. Además, los resultados que se obtienen con redes neuronales, que se basan únicamente en el espacio de píxeles de la imagen que analizan, no produce unos resultados que coincidan con las valoraciones que podría dar una persona. Para llevar a cabo el estudio, se va a analizar el espacio latente generado por una red entrenada con juicios humanos sobre la similitud entre materiales. Se va a considerar el espacio latente generado como una representación genérica de los materiales que las personas podemos tener en el cerebro.

Otro de los objetivos de este proyecto es, una vez estudiado el espacio latente, comprobar si este espacio construido para representar materiales es lo suficientemente bueno para ser reconstruido, y así generar las imágenes de los materiales correspondientes. Si se logra reconstruir, se podrían aplicar los resultados del análisis a la edición de puntos del espacio para modificar materiales, sin llegar a tener que editar directamente las imágenes.

En el contexto del proyecto, el espacio latente se entiende como una forma de representación comprimida de los datos de entrada de la red entrenada (Figura 1). Es una simplificación multidimensional del conjunto de datos a estudiar (imágenes de materiales), que los distribuye en el espacio según su similitud, entendiendo esta similitud como la relación entre los atributos aprendidos por la red, que han sido condicionados por valoraciones humanas. Un código latente es un vector  $n$ -dimensional que representa de forma comprimida un dato de entrada (una imagen de un material), normalmente de forma que pueda ser recuperado y convertido nuevamente en un dato del mismo tipo que la entrada. Cada código latente corresponde con un punto del espacio latente y representa una muestra.

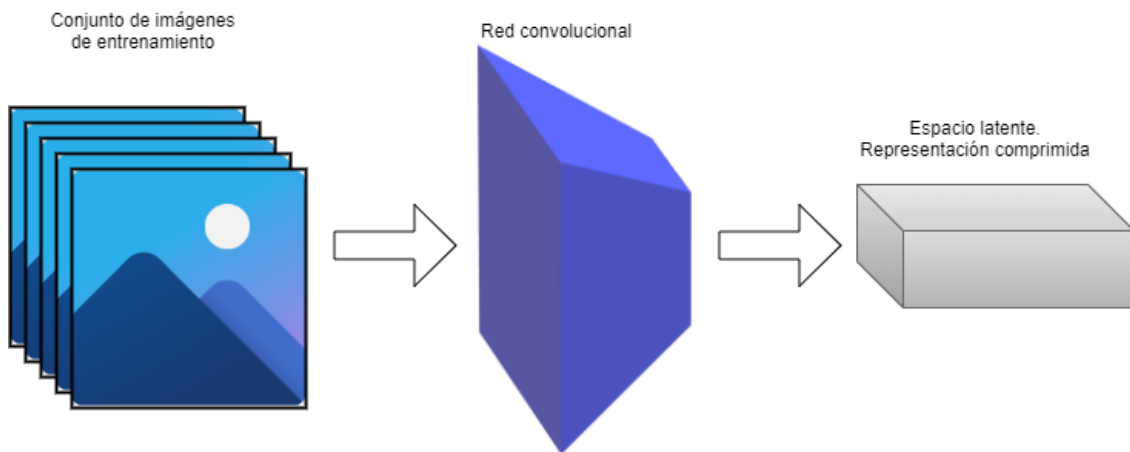


Figura 1. Espacio latente. Tomando como entrada imágenes, la red convolucional reduce los datos de entrada de alta dimensionalidad a una representación más simple con menos dimensiones.

Para poder alcanzar los objetivos del proyecto, el trabajo se ha dividido en dos partes principales. Estas partes corresponden al enfoque con el que se abordó el problema: primero un análisis en profundidad del espacio latente y los atributos aprendidos por la red, y segundo la reconstrucción del espacio y su edición.

El análisis y estudio del espacio latente ha tenido varias fases que abarcaban diferentes aspectos de su contenido y obtención. En primer lugar, se ha estudiado si este espacio generado a partir de la percepción humana de similitud entre materiales es semánticamente separable según ciertos atributos de alto nivel (metalicidad, plasticidad, brillantez...). En segundo lugar, se ha llevado a cabo una reducción de las dimensiones del espacio mediante el método de PCA (*Principal Component Analysis*) para comprobar la importancia de las diferentes componentes y tratar de lograr una visualización bidimensional representativa, que facilitase el entendimiento de la distribución de los materiales. Por último, se ha analizado el interior de la red empleada para ver que atributos o características aprendía en las diferentes capas. Para este último análisis se han estudiado las activaciones de cada capa de la red y se ha empleado la técnica de propagación hacia atrás del modo propuesto en el trabajo de Chu, Yang and Tadinada (2017).

La segunda parte se centra en el entrenamiento de un sistema GAN (*Generative Adversarial Network*) que se empleará para generar las imágenes correspondientes a un código latente que representa un material. El entrenamiento se ha dividido en dos, encontrar la mejor configuración de sistema para reconstruir el espacio, y ajustar lo mejor posible la función de coste y los hiperparámetros, tanto del generador como del discriminador. Esta fase es esencial para después poder obtener unos resultados de calidad y coherentes. Una vez conseguido un generador de materiales lo suficientemente bueno, se ha explorado la alteración de los puntos del espacio latente para ser capaces de encontrar una relación entre las modificaciones realizadas y los resultados en las imágenes generadas. A partir de ahí se podría desarrollar un sistema intuitivo de edición que permita modificar las características principales de un material.

Debido a las limitaciones de tiempo y recursos, la investigación se ha llevado a cabo en materiales renderizados. No se emplean fotografías del mundo real por la dificultad que supone obtener una gran variedad de materiales y controlar las condiciones de iluminación a la hora de capturar las imágenes. El trabajo se ha realizado con la colaboración del grupo de investigación [Graphics and Imaging Lab](#). Gracias a esta colaboración ha sido posible el uso de los datos ya recopilados por el grupo en trabajos previos sobre materiales o percepción humana.

### Objetivos del proyecto

En este subapartado se van a especificar de forma concisa los principales objetivos a tratar de alcanzar durante la realización del trabajo. Dada la naturaleza de investigación del proyecto, no se trata de objetivos específicos para superar barreras o el estado del arte.

- Estudiar la distribución de los materiales en el espacio latente mediante su clasificación semántica en distintas propiedades.
- Estudiar la complejidad del espacio latente mediante la reducción de dimensiones para una clara visualización.
- Estudiar el interior de la red que ha generado el espacio latente, para buscar patrones detectados por las neuronas que ayuden a entender cómo se ha formado.
- Comprobar la viabilidad de reconstruir el espacio latente de similitud material.
- Buscar un posible método de edición de los códigos latentes para obtener nuevos materiales.



## 2. Trabajos relacionados

### 2.1. Percepción humana

Existen numerosos trabajos que estudian el funcionamiento del cerebro humano a la hora de realizar diferentes tareas relacionadas con la percepción, una gran parte de ellos desde un punto de vista neurológico. Pero también es notable la variedad de estudios que buscan relacionar la percepción humana con técnicas informáticas que puedan replicarla o que estén guiadas por modelos de percepción humana. En Lagunas *et al.* (2019) se propone un sistema para medir la similitud entre dos materiales, en base al análisis de juicios humanos de similitud. Se propone en Serrano *et al.* (2016) un modelo de edición de BRDF que emplea valoraciones humanas de diferentes atributos de los materiales para encontrar correlaciones entre ellos, y así desarrollar un método más simple para alterar una BRDF. Un trabajo reciente [Zhang *et al.*, 2018] ha mostrado que existe cierta correlación entre la percepción humana y los atributos aprendidos por una red a la hora de comparar elementos. En un trabajo más teórico y menos centrado en aspectos informáticos [Fleming, 2014], se describe el proceso de percepción como un modelo generativo que capta las diferencias entre elementos para distinguirlos. Descarta la idea de que sea un proceso que calcule los parámetros físicos de los materiales, y propone cómo este modelo estadístico de apariencia podría ser computado. En Celebi and Aslandogan (2004) se propone un sistema de obtención de imágenes médicas que incorpora un modelo de percepción humana para devolver mejores resultados.

En este trabajo se va a intentar ir un paso más allá analizando una red entrenada a partir de valoraciones humanas sobre similitud entre materiales. El objetivo no es simular el proceso humano, sino intentar comprender mejor la distinción interna entre materiales de las personas.

### 2.2. Reconstrucción y descomposición del espacio latente

En los últimos años ha aumentado la cantidad de trabajos sobre el estudio semántico o descomposición del espacio latente, para editar de forma controlable los códigos latentes. Shen *et al.* (2019) introduce un sistema que permite la edición de ciertos atributos faciales de alto nivel (edad, género...) a partir del estudio de la distribución en el espacio latente de las muestras que presentan dichos atributos. Otro enfoque interesante en este ámbito es el de Donahue *et al.* (2018), propone un método para dividir directamente el espacio latente en dos vectores, uno correspondiente a la identidad de la persona en la imagen, y otro que engloba todos los atributos externos. Esto permite separar los fragmentos de identidad, de modo que se pueden sustituir unas identidades por otras manteniendo las características (iluminación, posición, sonrisa...) de la imagen destino. Otro trabajo reciente [Jiang *et al.*, 2019] estudia la descomposición de modelos tridimensionales de caras en identidad del sujeto y las expresiones faciales.

Estos estudios muestran una clara predominancia por la investigación de técnicas de descomposición dirigidas a las caras y la identidad de una persona. En este proyecto, como meta final, se va a tratar de generar imágenes de materiales pudiendo alterar o controlar algunas de las características de más alto nivel de los materiales.

### 3. Conjuntos de datos y arquitectura de la red

Para poder llevar a cabo una investigación representativa de la percepción material es necesario tener una colección de materiales lo más grande posible. Conseguir suficientes objetos que puedan representar correctamente el espacio de materiales es muy complicado. Además, es necesario encontrar elementos que tengan la misma forma o una similar, para evitar introducir en el estudio los sesgos que puede conllevar el uso de geometrías muy diferentes. Por tanto, se ha decidido emplear imágenes de materiales renderizados.

Para conseguir los diferentes materiales se han empleado las BRDF (*Bidirectional Reflectance Distribution Function*) de un conjunto de datos generado en Serrano *et al.* (2016). Una BRDF es una función que define cómo es reflejada la luz sobre la superficie de materiales no transparentes. Se trata de un conjunto obtenido a partir de la manipulación de las muestras de la base de datos MERL [Matusik *et al.*, 2003] (base de datos con 100 funciones de reflectancia de distintos materiales, guardadas en forma de BRDF densamente medidas). El conjunto de datos consta de las 100 muestras que están en MERL, junto con otras 306 resultado de interpolaciones de las 100 primeras. Este proceso de obtención de nuevos materiales se realiza sobre una reducción del espacio de las BRDF mediante PCA. El espacio es muestreado aleatoriamente y se genera un nuevo material interpolando las BRDF de los tres puntos más cercanos. Cómo se explica en el trabajo oficial de MERL y en el de Serrano [Matusik *et al.*, 2003; Serrano *et al.*, 2016], es posible realizar las interpolaciones porque el espacio de BRDF matemáticamente válidas es convexo, y por tanto cualquier combinación convexa de dos puntos dará una BRDF que seguirá siendo matemática y físicamente válida. Uno de los principales motivos para elegir este conjunto es que se tiene para cada material el valor de numerosos atributos (Tabla 1), permitiendo más adelante clasificar semánticamente los materiales para ocho de las propiedades: *plástico, metálico, mate, áspero, brillante, luminoso, dureza y nitidez de los reflejos*.

<b>Plástico</b>	El material tiene una apariencia similar a un plástico.
<b>Metálico</b>	El material tiene una apariencia similar a un metal.
<b>De goma</b>	El material parece hecho de goma.
<b>De tejido</b>	El material parece un tejido.
<b>Suave</b>	El material parece suave al tacto.
<b>Duro</b>	El material parece duro al tacto.
<b>Mate</b>	El material no tiene brillo.
<b>Áspero</b>	La superficie del material parece carente de suavidad, rígida.
<b>Brillante</b>	El material es lustroso.
<b>Luminoso</b>	El material refleja mucha luz.
<b>Dureza de los reflejos</b>	Los reflejos en la superficie del material se ven claramente.
<b>Nitidez de los reflejos</b>	Los reflejos del material están muy definidos.
<b>Reflejos coloreados</b>	Los reflejos del material están teñidos con su propio color.

Tabla 1. Atributos de los materiales medidos en Serrano *et al.* (2016). Representan características físicas de los materiales clasificadas según la percepción humana.

Con el objetivo de obtener un conjunto de datos más grande y al mismo tiempo tener más variación en las muestras, los 406 materiales se han renderizado con dos formas: esfera y blob; y bajo cuatro entornos de iluminación diferentes (Figura 2): *Saint Peters*, *Doge*, *Ennis* y *Grace*. Se trata de formas geométricas sencillas sin ninguna arista o borde definido como los de un cubo. Un mapa de entorno o iluminación es una textura preparada de una vista panorámica de una escena del mundo real, esto permite tener reflejos con una apariencia más realistas sobre los materiales sin tener que realmente generar la escena. Los mapas de entorno elegidos (Figura 2, primera fila) presentan distintos grados de intensidad lumínica y número de fuentes de luz. Con esto se consigue tener un total de 3248 imágenes diferentes para representar el espacio material mejor, y de forma que se pueda comprobar el efecto que tiene en un material las condiciones con las que se presenta.

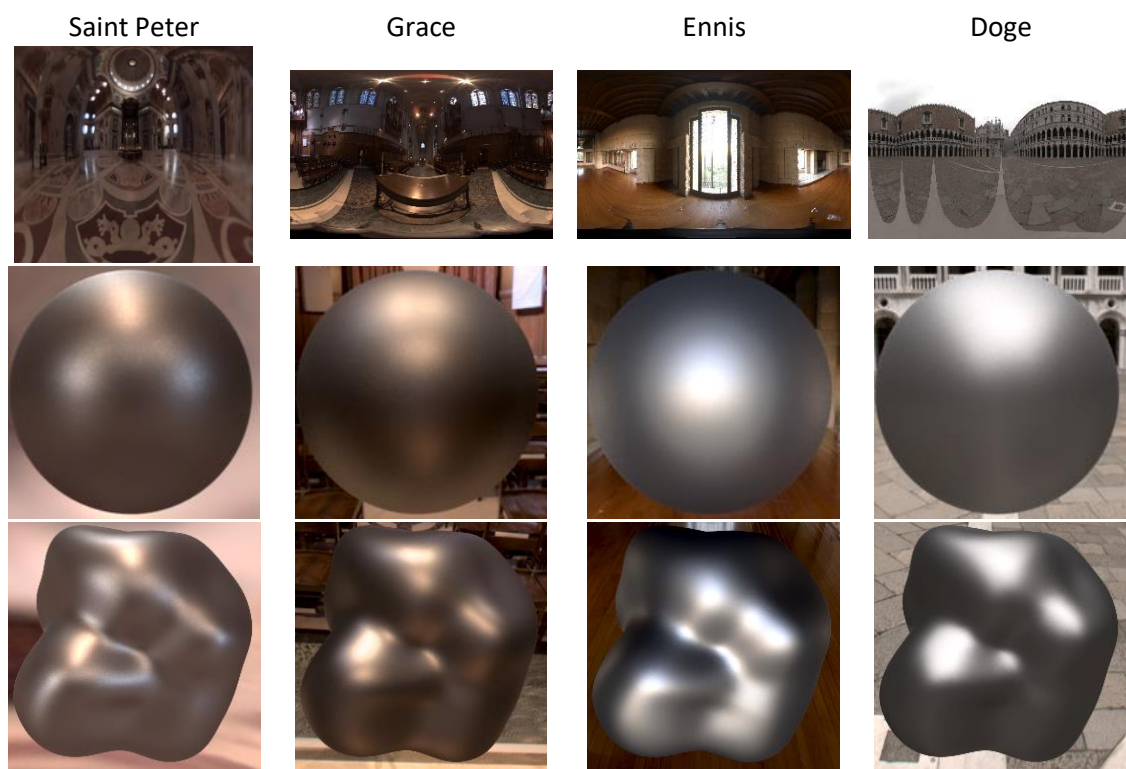


Figura 2. Ejemplo de material (*silver-metallic-paint*) con las diferentes posibilidades de forma y entorno. La primera fila corresponde con los mapas de iluminación. Las otras dos son ejemplos para dichos mapas con la forma de esfera y blob respectivamente.

La red empleada para obtener el código latente correspondiente a cada imagen es un modelo de red convolucional que se denomina ResNet-34 (Figura 3). Se trata de una red neuronal residual [He *et al.*, 2016] cuya principal diferencia frente a una red convolucional tradicional es la existencia de conexiones de salto entre capas, haciendo que la salida de un bloque sea una función que tiene en cuenta también los valores de su propia entrada. La implementación exacta de este modelo de red se ha tomado del trabajo de Lagunas *et al.* (2019), cuya principal modificación respecto al modelo original (sin considerar la calibración de parámetros) es la sustitución de la última capa por una completamente conexa, que devuelve el vector de atributos correspondiente a un código del espacio latente. En el trabajo de referencia esta red es entrenada con juicios

humanos que puntúan la similitud entre materiales, y para este proyecto se va a emplear esa misma red, con el modelo de pesos obtenido en el trabajo de Lagunas *et al.*, para generar los vectores de nuestro conjunto de datos.

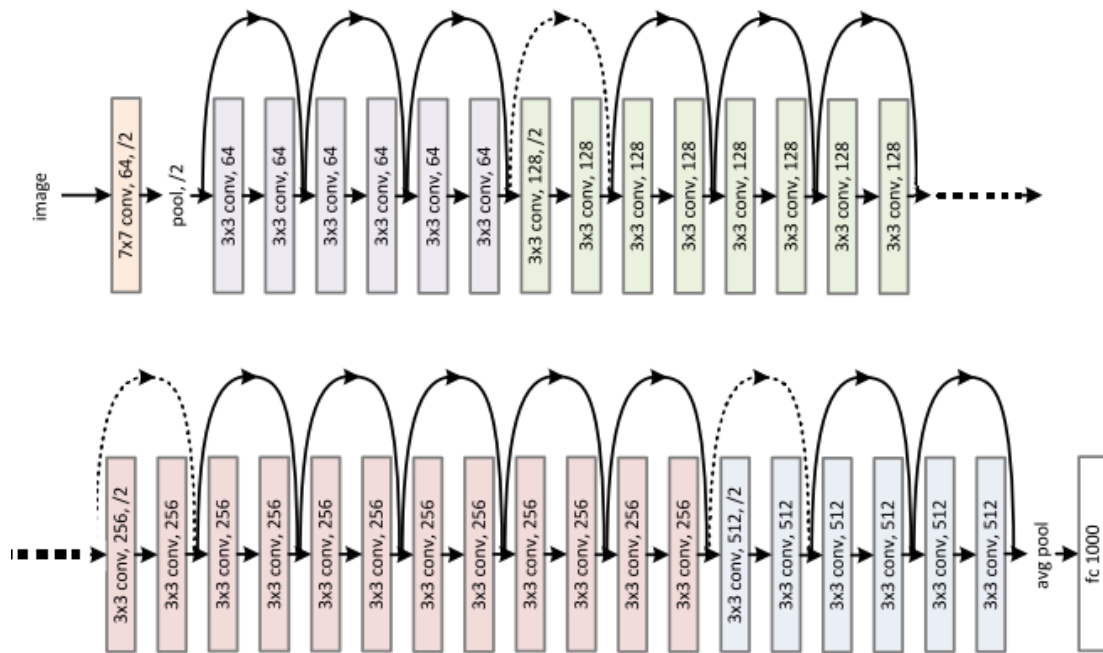


Figura 3. Representación de las capas y conexiones de la red convolucional residual usada, ResNet-34 [He *et al.*, 2016].

En las dos próximas secciones del trabajo se van a analizar los principales aspectos de la distribución interna y la construcción del espacio latente del trabajo antes mencionado [Lagunas *et al.*, 2019]. Se ha obtenido entrenando una red convolucional con una función de coste que minimiza el error a la hora de distribuir los materiales respecto a datos recopilados de juicios humanos de similitud entre materiales. Para llevar a cabo este estudio es necesario analizar tanto la agrupación semántica de los materiales, como estudiar cual es la actividad en el interior de la red al pasar las imágenes por ella.

#### 4. Entendiendo el espacio latente

El objetivo de esta sección es entender lo mejor posible la distribución de los materiales en el espacio latente generado por la red neuronal de Lagunas *et al.* (2019). Se va a tratar de averiguar si tiene una distribución que pueda permitir una clasificación semántica de los materiales, en atributos como metalicidad o brillantez, según su posición en el espacio. Al tratarse de una representación multidimensional, los primeros análisis básicos que se van a realizar no permiten una visualización intuitiva para personas, ofreciendo solamente resultados numéricos. En los análisis finales se emplearán técnicas de reducción de dimensionalidad para poder generar representaciones visibles que faciliten la interpretación de los resultados.

Para la clasificación de las imágenes se va a emplear un método similar al usado en Shen *et al.* (2019), consiste en el entrenamiento de un clasificador SVM (*Support Vector Machines*) para clasificar los códigos latentes de las muestras de materiales según sus propiedades. Con esto se consigue definir un hiperplano que debería separar los

materiales que poseen una cierta propiedad de los que no. Este hiperplano permite calcular la distancia a cualquier punto del espacio latente, y comprobar que este valor es coherente con el resto de los materiales del mismo tipo. Para poder llevar a cabo esta separación de los materiales es necesario tenerlos etiquetados para cada uno de los atributos que se desee analizar. Esto ha sido posible gracias a los datos recopilados en Serrano *et al.* (2016), que contiene las etiquetas de los atributos que se quiere estudiar para nuestro conjunto de datos. Esta recopilación de puntuación de atributos está basada en juicios humanos, y por lo tanto no se pierde el enfoque del proyecto de entender y analizar la percepción humana. Debido a que los atributos no son binarios, sino que tienen un rango normalizado de cero a uno, se ha decidido considerar que un material pertenece a una clase si el valor de su etiqueta es mayor que 0.5.

#### 4.1. Clasificación semántica de materiales

Para poder ver si los resultados obtenidos de una prueba a otra son coherentes, se han elegido 12 materiales de referencia (Figura 4) para representar de la forma más uniforme posible el espacio. Debido a que hay muchos atributos, esta selección de materiales está dirigida a cubrir principalmente las propiedades más claras visualmente, como lo metálico o plástico que es, o su nivel de brillo o mate. A su vez esos atributos tienen bastante correlación con otros más específicos, como la dureza y la nitidez de los reflejos.

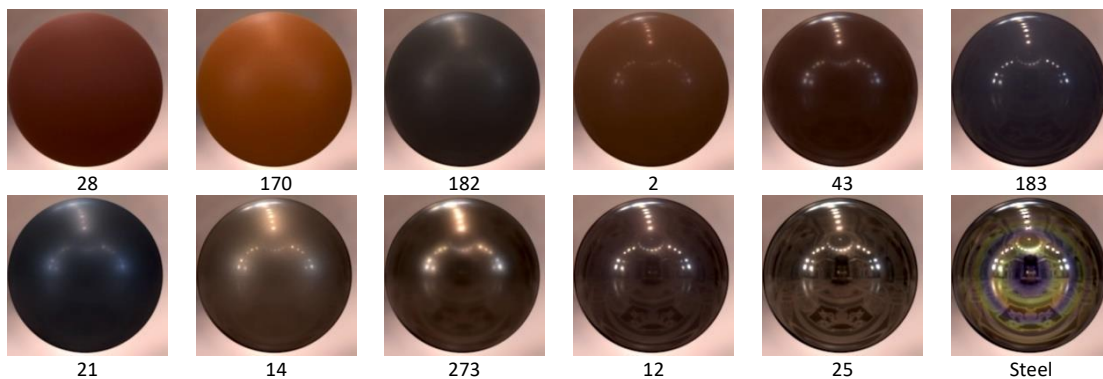


Figura 4. Materiales elegidos para representar el espacio. La primera fila representa de plástico mate a más brillante. La segunda fila representa de menos a más metálico. Debajo de cada imagen está el nombre de cada material según se nombran en Serrano *et al.* (2016), los que tienen por nombre un número son materiales interpolados. En las tablas posteriores el número/nombre del material se empleará para referirse al mismo, en vez de emplear imágenes.

##### 4.1.1. Materiales con forma e iluminación constante

Antes de comenzar con todo el conjunto de datos disponible, se van a analizar los materiales en unas condiciones sencillas, que van a ser la forma de esfera y en el entorno de *Saint Peters*. Se hace este primer análisis reducido para asegurar el correcto funcionamiento en un caso base, dado que no tendría sentido tratar de clasificar todo el conjunto de datos, si no se obtienen buenos resultados para una pequeña parte.

Las propiedades que se van a evaluar en esta primera parte son tres: metalicidad, brillo y mate. Se trata de tres atributos de los materiales que para las personas son más fáciles de clasificar o distinguir, y dado que se está empleando una red entrenada con datos de percepción humana cabría esperar que pudiera generar un espacio en el que estas tres cualidades sean separables.



Los resultados obtenidos para estas tres clasificaciones son bastante buenos, en todos los casos se encuentra un hiperplano capaz de separar correctamente las muestras. La variación de las distancias también concuerda con lo que se percibe de los materiales, y para reflejar mejor esa variación se ha empleado un código de colores como gradiente de valores más negativos (rojos) a más positivos (verdes). En algunos casos puede parecer que las muestras centrales no están en el lado correcto del hiperplano. Sin embargo, al tratarse de un espacio guiado por la percepción humana, la opinión de un individuo a otro puede diferir haciendo que materiales ambiguos no se sitúen bien. En esos casos las muestras aparentemente mal clasificadas estarán cerca del límite de separación, y el resto de los materiales estarán desplazados en el sentido en el que se haya equivocado el clasificador.

	28	170	182	2	43	183	21	14	273	12	25	Steel
GT	0,07	0,00	0,38	0,38	0,42	0,32	0,45	0,63	0,83	0,46	0,70	1,00
P1	-59,29	-30,37	-11,17	-15,33	-6,13	-2,27	-2,04	0,67	0,92	5,08	9,48	13,06
P2	-49,61	-21,41	-3,71	-10,49	-2,29	1,40	2,98	5,45	4,95	9,26	14,21	17,86

Tabla 2. Distancia al hiperplano de materiales metálicos.

	28	170	182	2	43	183	21	14	273	12	25	Steel
GT	0,07	0,25	0,29	0,38	0,75	0,75	0,31	0,67	0,68	0,64	0,89	0,90
P1	-34,04	-11,13	1,89	-4,23	1,12	3,92	5,90	8,13	7,00	10,75	15,25	18,44
P2	-58,81	-30,47	-11,92	-15,24	-6,17	-2,48	-2,53	0,10	0,37	4,41	8,42	11,85

Tabla 3. Distancia al hiperplano de materiales brillantes.

	28	170	182	2	43	183	21	14	273	12	25	Steel
GT	0,65	1,00	0,71	0,50	0,33	0,15	0,5	0,25	0,35	0,25	0,14	0,00
P1	58,97	29,12	9,62	14,17	5,15	1,28	0,68	-2,14	-2,09	-6,33	-11,00	-14,82
P2	52,37	23,28	5,87	11,11	2,77	-0,91	-1,78	-4,54	-4,10	-8,63	-13,33	-16,93

Tabla 4. Distancia al hiperplano de materiales mate.

Tabla 2-4. Distancias a hiperplanos en el entorno de Saint Peters. GT representa el valor de la etiqueta del atributo. Los colores indican la distancia al hiperplano, a más verde más lejos en el sentido positivo y lo contrario a más rojo. Los materiales están ordenados y el gradiente de color muestra en los tres casos un cambio gradual y coherente.

De cada atributo se han realizado múltiples pruebas para comprobar la consistencia de las clasificaciones, y se muestran para cada atributo dos de esos resultados en las tablas anteriores (Tabla 2-4). Si bien no es siempre el mismo valor de un caso a otro, se mantiene una clasificación muy similar y coherente entre pruebas. El hecho de que los resultados varíen apunta a que el espacio no es perfecto y de que, al tratarse de un hiperplano en un espacio de 128 dimensiones, ligeras variaciones en cualquiera de ellas conllevan cambios visibles en los resultados numéricos. Todo esto sugiere que la red entrenada genera un espacio latente en el que los materiales están agrupados por atributos, o al menos están lo suficientemente separados unos de otros como para poder clasificarlos según algunas de sus propiedades de alto nivel más sencillas.

#### 4.1.2. Materiales con diversas formas e iluminaciones

Una vez visto que el espacio es separable para al menos un caso pequeño y no muy complejo, hay que comprobar que los resultados siguen siendo similares y no se ven demasiado alterados por el cambio en la geometría de los materiales, o por las

condiciones de iluminación bajo las que se han renderizado las muestras. Para realizar estas clasificaciones se han llevado a cabo dos tipos de pruebas principales, por un lado, se ha entrenado y obtenido el límite de un atributo para una forma y un entorno, y después se ha comprobado cómo se clasifican las imágenes del resto de entornos para el hiperplano encontrado. Por otro lado, se ha entrenado el SVM con todos los códigos latentes para ver si aún es capaz de encontrar un límite bueno para separar los materiales.


Para este análisis y búsqueda de un hiperplano, se va a forzar que el resultado final encontrado tenga como mínimo unos valores de acierto del 75% y el 65% para validación y test respectivamente. Estos valores se han fijado después de numerosas pruebas en las que se ha visto que para los atributos más complejos y ambiguos como la luminosidad resultaba muy complicado encontrar límites con mejores valores. Para la mayoría de los atributos esto no es necesario, pues los valores medios obtenidos son altos en todos los casos (Tabla 5).

		Metal	Brillo	Mate	Áspero	Frz. Ref.	Nit. Ref.	Plástico	Lumin.
<i>St. Peter</i>	val.	1,00	0,97	0,98	0,88	1,00	1,00	0,92	0,88
	test	0,80	0,79	0,82	0,74	0,82	0,78	0,57	0,58
<i>Grace</i>	val.	0,89	1,00	0,97	0,77	0,99	1,00	0,89	0,81
	test	0,76	0,74	0,78	0,72	0,75	0,72	0,58	0,56
<i>Ennis</i>	val.	0,99	1,00	0,96	0,89	1,00	0,97	0,82	0,73
	test	0,75	0,76	0,78	0,75	0,73	0,72	0,59	0,58
<i>Doge</i>	val.	1,00	1,00	1,00	0,90	1,00	1,00	0,82	0,82
	test	0,80	0,81	0,80	0,77	0,79	0,79	0,59	0,61

*Tabla 5. Tasa de acierto media de validación (val.) y test para la clasificación en diferentes atributos de las muestras de materiales, separados en los diferentes mapas de iluminación. La tasa de acierto en test es mucho más baja en los atributos de luminosidad y plástico, se trata de los atributos probablemente más complejos que ofrecerán unos resultados menos claros. El resto de los atributos muestran valores altos para todos los entornos de iluminación.*

Como se puede observar en las tablas del Anexo 1, para el atributo de metalicidad de los materiales, independientemente del entorno con el que se ha entrenado el clasificador, todas las muestras se clasifican de forma similar. Los valores numéricos de un caso a otro cambian, pero se mantienen las relaciones de distancia entre las muestras. El único caso que muestra un comportamiento menos estable, y con mayor variación en la relación de distancias, es el de los materiales en el entorno *Ennis*. Esto es debido a que la iluminación de ese escenario es muy fuerte, y el color y reflejo de las muestras se ve alterado en gran medida haciendo difícil el reconocimiento de las imágenes. Para otros atributos como el brillo, la definición de los reflejos o el nivel de mate del material los resultados obtenidos son muy buenos, de forma similar a la clasificación de metales.

Sin embargo, para atributos más complejos como la plasticidad o la luminosidad, los resultados obtenidos son menos claros. La clasificación de plásticos tiene problemas con los colores, y se debe probablemente a que a la hora de evaluar un material muchas veces nos dejamos llevar por la experiencia y esta nos dice que los materiales de colores más claros suelen ser plásticos. Algo similar ocurre con la luminosidad, que se trata de un atributo bastante ambiguo que puede verse muy afectado por el color y la persona a valorarlo. Como se muestra en la Figura 5, el valor para estos dos atributos en materiales fenólicos especulares (un tipo de plástico) varía mucho entre diferentes colores. Estos materiales tienen unas propiedades físicas y una BRDF casi idénticas, y por tanto el valor de los atributos debería ser similar de unos a otros.



luminosidad	0,36	0,71	0,75	0,81	0,58	0,82	0,91
plástico	0,29	0,31	0,35	0,82	0,54	0,83	0,45

Figura 5. Valoración del atributo de brillo y plasticidad de materiales casi idénticos físicamente, exceptuando el color. Se puede observar que hay gran diferencia de unos a otros, cuándo cabría esperar valores similares.

El uso del blob en vez de esferas no introduce demasiado cambio en los resultados. Aunque en algunos casos el uso de esta forma hace que los resultados varíen menos entre entornos, a nivel general no hay gran influencia en la obtención de hiperplanos. Los atributos que ya se clasificaban bien siguen prácticamente igual, y solo en alguno de los atributos más complejos, como si el material tiene aspecto de plástico o no, los resultados son ligeramente mejores. Para el atributo que se considera más complicado, la luminosidad, los resultados siguen sin ser claros o consistentes.

Para las pruebas entrenando el clasificador con todas las muestras, los resultados son en media más estables. Por ejemplo, el atributo áspero (Tabla 6) muestra un comportamiento diferente según el entorno de entrenamiento y el de test, pero empleando todas las muestras para entrenamiento el resultado se vuelve mejor. Aunque siguen existiendo algunos casos peor clasificados para el entorno *Ennis*, al fin y al cabo, esas muestras parecen representar el mayor grado de dificultad en la clasificación. Para la luminosidad (Tabla 7) se siguen clasificando de forma inconsistente los materiales debido, probablemente, a la falta de consistencia en su etiquetación, pero al menos presentan un mejor comportamiento entre conjuntos de muestras.

		28	170	182	2	43	183	21	14	273	12	25	steel
todos	stpeter	2,702	1,242	1,612	0,897	1,235	0,455	1,218	0,649	0,713	0,472	0,357	0,168
todos	stpeter	2,757	1,475	1,834	1,080	1,317	0,703	1,270	0,831	0,794	0,689	0,546	0,335
todos	doge	3,775	1,600	1,436	1,560	1,286	-0,315	1,112	0,313	0,362	-0,462	-0,416	-0,808
todos	doge	3,172	1,560	1,389	1,433	1,484	0,097	1,155	0,564	0,611	-0,179	-0,185	-0,312
todos	ennis	1,498	1,244	1,441	0,672	0,914	0,311	0,858	0,608	0,236	0,581	0,240	-0,137
todos	ennis	1,520	1,203	1,751	0,724	0,785	0,081	1,054	0,547	0,082	0,168	0,053	-0,103
todos	grace	2,981	1,397	1,338	0,893	1,065	0,735	0,990	0,394	0,363	0,601	0,439	0,024
todos	grace	3,130	1,389	1,341	0,919	1,016	0,730	1,005	0,409	0,372	0,623	0,460	-0,017

Tabla 6. Distancia al hiperplano de materiales ásperos. De cada entorno usado como conjunto de test se han probado dos hiperplanos entrenados con todas las muestras. Aunque la variación de distancias sea coherente con la percepción de los materiales, en algunos casos no se llega a poner ningún material en el lado negativo (de los usados de referencia).



		28	170	182	2	43	183	21	14	273	12	25	steel
todos	stpeter	0,477	1,064	0,348	0,935	0,549	0,574	0,304	1,059	0,550	0,666	0,661	0,731
todos	stpeter	0,631	1,472	0,914	1,248	0,825	0,803	0,894	1,466	0,906	1,089	1,141	1,430
todos	doge	0,460	0,803	0,021	0,445	-0,252	0,298	-0,372	0,811	0,447	0,244	0,135	0,255
todos	doge	0,566	0,971	0,205	0,370	-0,019	0,229	-0,010	1,253	0,604	0,403	0,320	0,601
todos	ennis	0,408	0,648	0,182	0,772	0,643	0,343	-0,405	0,427	0,553	0,307	0,779	0,650
todos	ennis	0,468	0,835	0,405	0,894	0,593	0,580	-0,077	1,040	1,293	0,794	1,464	1,260
todos	grace	0,455	1,052	0,382	0,747	0,479	0,670	0,228	0,706	0,439	0,748	0,929	0,887
todos	grace	0,509	1,195	0,477	0,790	0,480	0,583	0,215	0,673	0,365	0,604	0,787	0,844

Tabla 7. Distancia al hiperplano de materiales luminosos. De cada entorno usado como conjunto de test se han probado dos hiperplanos entrenados con todas las muestras. No hay un comportamiento claro entre materiales y entornos.

De estos análisis la principal conclusión que se puede extraer es que el espacio que se está estudiando parece tener un buen grado de organización, que permite clasificar materiales según sus propiedades con gran éxito general. Es importante destacar también el hecho de que, al tratarse de un espacio de materiales únicamente, el efecto que tiene el cambio de condiciones de iluminación para la mayoría de los casos no es muy significativo, y lo mismo para el cambio de forma. Si por el contrario esos cambios fuesen muy notables, significaría probablemente que el espacio no ha sido del todo bien generado, ya que se influye mucho por las condiciones en las que se presenta un material. Para todos aquellos atributos que no se llegan a clasificar bien por su dificultad o dependencia con las condiciones con las que se ha renderizado el material, hay un punto en común, las distancias de los materiales al hiperplano son siempre muy bajas en comparación con las distancias que se generan para atributos como la metalicidad. Esto puede significar que a la hora de valorar la distancia entre materiales esos atributos no tengan mucho peso en la decisión, haciendo que la forma del espacio no separe los materiales de acuerdo con esas propiedades. Otra de las causas podría ser que las etiquetas dadas para esos atributos no sean precisas, porque para las personas sean propiedades subjetivas y difíciles de valorar.

#### 4.2. Clasificación semántica con PCA

Los resultados anteriores apuntan a que hay cierta organización en el espacio latente que permite clasificar los materiales de acuerdo con algunas propiedades físicas, con mayor o menor exactitud dependiendo de qué atributo se esté clasificando. Sin embargo, esto solo ofrece una intuición de cómo podrían estar distribuidas las muestras y no da lugar a una clara o sencilla visualización mental. Para tratar de entender mejor como se está distribuyendo el espacio, la mejor opción es ser capaces de representarlo en dos o tres dimensiones, de modo que se puedan generar las imágenes correspondientes para ver su forma. Existen diversas maneras de visualizar espacios de más de tres dimensiones, y de entre ellas una de las mejores opciones es el uso del análisis de las componentes principales (PCA) que permite reducir un espacio de  $n$  dimensiones al número de dimensiones deseado. El problema de la reducción mediante PCA es que las nuevas dimensiones, generadas del análisis de los datos, representan el grado de variabilidad de las muestras en diferentes direcciones, y si con las dos o tres componentes más importantes no se consigue explicar un porcentaje aceptable de la variación entre muestras, los resultados generados no serán representativos del espacio.

Por tanto, el primer paso ha sido calcular, para cada uno de los conjuntos de datos (materiales en un entorno y con una forma) y para el total de las muestras que se tienen, cuanta variabilidad se explica con las primeras componentes. Los resultados obtenidos de este análisis (Tabla 8) han sido sorprendentes y completamente inesperados. Para cualquiera de los conjuntos con tan solo tres de las principales componentes se explica alrededor del 98% de la variabilidad. Y lo que es aún más sorprendente es que con tan solo la primera componente (una dimensión) se explica en torno al 90% de la variabilidad. Esto parece indicar que el espacio generado por la red entrenada con juicios humanos de similitud entre materiales es mucho más simple de lo que se creía.

Entorno	Forma	Primera	Segunda	Tercera	2 Comp.	3 Comp.
St. Peter	Esfera	0,9477	0,0297	0,0110	0,9774	0,9885
Doge	Esfera	0,9345	0,0302	0,0224	0,9646	0,9871
Grace	Esfera	0,9298	0,0301	0,0225	0,9599	0,9824
Ennis	Esfera	0,9017	0,0525	0,0241	0,9541	0,9782
Media		0,9284	0,0356	0,0200	0,9640	0,9840
Todos	Esfera	0,9350	0,0315	0,0191	0,9665	0,9856
St. Peter	Blob	0,9578	0,0273	0,0054	0,9851	0,9905
Doge	Blob	0,9234	0,0442	0,0163	0,9676	0,9839
Grace	Blob	0,9369	0,0300	0,0142	0,9669	0,9811
Ennis	Blob	0,9058	0,0513	0,0215	0,9571	0,9787
Media		0,9310	0,0382	0,0144	0,9692	0,9836
Todos	Blob	0,9394	0,0363	0,0094	0,9757	0,9851

Tabla 8. Análisis de las componentes principales (PCA). Para cada conjunto de imágenes (una forma y un mapa de iluminación) se han calculado los porcentajes de variabilidad que explican las tres primeras componentes y su suma (dos últimas columnas). Los valores obtenidos de la primera componente son muy altos en todos los casos.

Una vez visto que en todos los casos el análisis hecho mediante PCA da unos valores numéricos muy buenos, se han generado diferentes imágenes con las muestras en el espacio reducido para poder visualizar más claramente la agrupación de los materiales. Con esto se puede conseguir ver de qué modo varían las muestras a lo largo de las componentes más importantes y establecer cuál o cuáles son los atributos principales en esas dimensiones. Como se puede observar en la Figura 6, a lo largo de la primera dimensión los materiales pasan de ser muy especulares o metálicos a cada vez más difusos y con menos reflejos especulares, hasta llegar a los materiales casi completamente difusos. En la segunda dimensión como el porcentaje de variación es menor, no se ve tan claramente que orden siguen los materiales, pero está claro que el color juega un papel importante, dejando los materiales de colores más oscuros abajo, y los más claros o coloridos arriba.

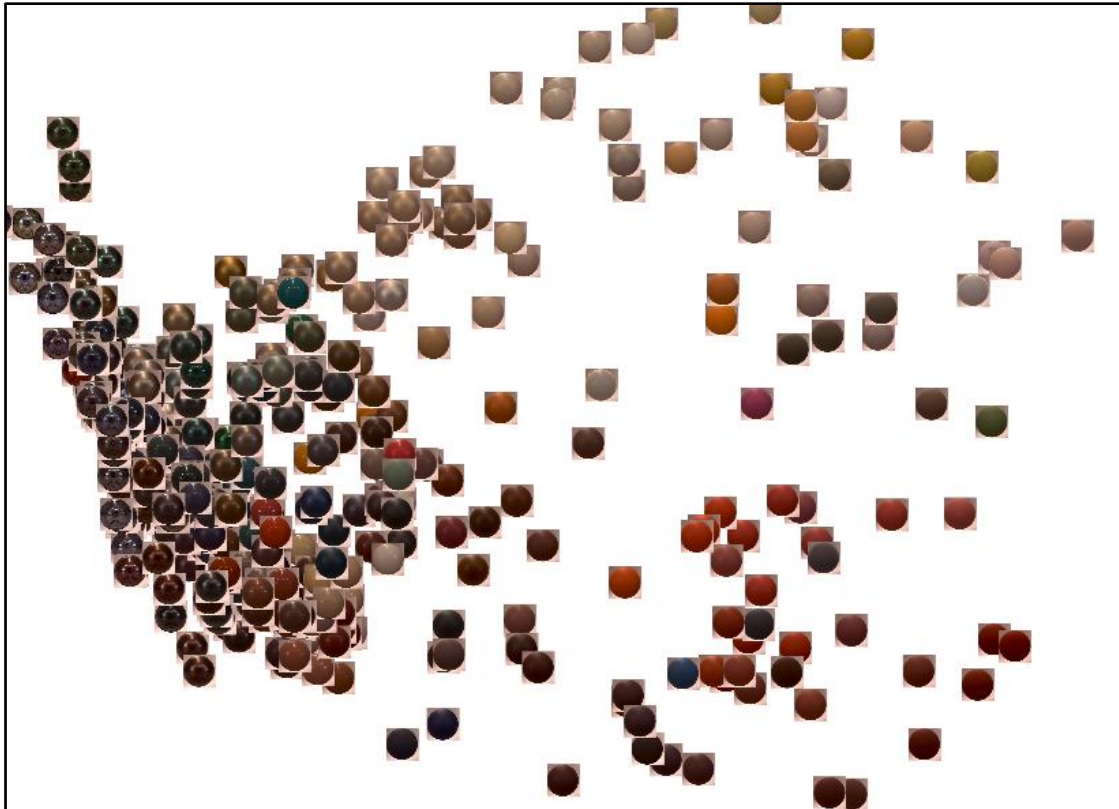


Figura 6. Gráfico PCA de las dos primeras componentes ajustadas con las esferas en el entorno de Saint Peter. La primera componente (eje horizontal) tiene una clara correlación con el nivel de brillo de los materiales, mientras que la segunda distribuye los materiales de acuerdo con los colores.

El uso de PCA con otros entornos y formas tiene resultados con una estructura muy similar (Figura 7), distribuyendo los materiales por su brillo en el eje horizontal y en el eje vertical distribuyendo las muestras de acuerdo con su color o tono. Es el caso del entorno de *Ennis* (Anexo 3) el que genera el resultado más diferente, da más importancia al eje vertical debido a que la fuente de luz es tan intensa que el color de todos los materiales se esclarece de forma muy apreciable, afectando así a su apariencia y por tanto percepción.

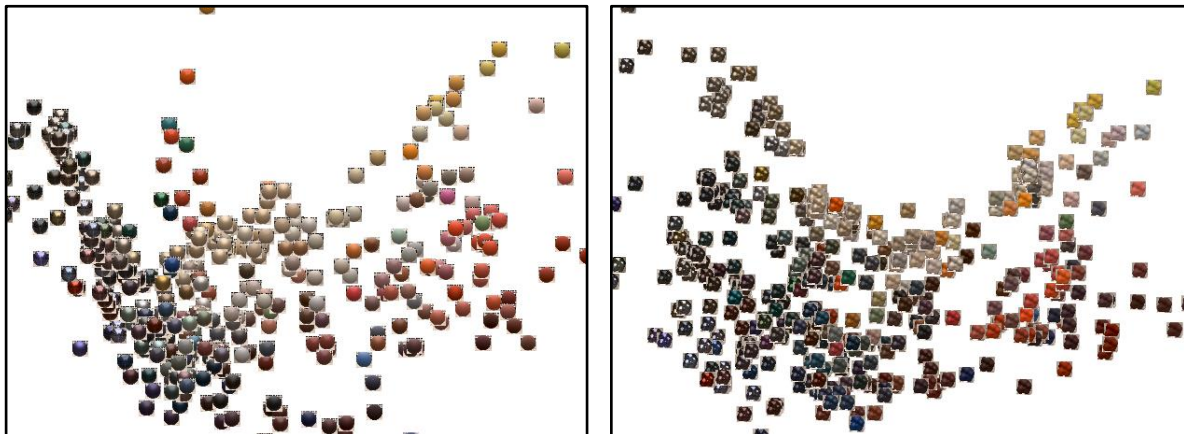


Figura 7. Gráfico de PCA de las dos primeras componentes para los materiales en el entorno de Doge. A la izquierda los materiales tienen forma de esfera y a la derecha de blob. Ambos casos mantienen las mismas direcciones principales. El eje horizontal distribuye materiales por su especularidad y el eje vertical por el color.

Visto que el resultado de ajustar las componentes principales con los diferentes conjuntos de imágenes sigue un patrón similar en todos los casos, se puede realizar el mismo análisis semántico del espacio realizado en el subapartado anterior. Para buscar los hiperplanos de separación entre clases, se ha empleado directamente el espacio reducido de dos dimensiones. Dado que anteriormente se ha mostrado que de un intento a otro el límite puede variar, se ha entrenado varias veces cada clasificador para mostrar que el hecho de que varíe no significa que se esté haciendo mal, puesto que los resultados para un mismo atributo, aunque diferentes, separan los mismos grupos de imágenes. Como se puede ver en la Figura 8, se muestran dos límites para separar materiales metálicos (azul) y brillantes (negro) que claramente son dos propiedades muy correlacionadas y corresponden a cortes en el eje horizontal. El resultado obtenido es bastante bueno, ya que se puede ver que hay materiales más allá de los metálicos que también son pertenecientes a la categoría de brillantes. Los dos resultados para metalicidad son muy similares, y los de brillo tienen más variación debido probablemente a que las muestras situadas cerca del límite tienen valores del atributo muy similares, haciendo más difícil la clasificación. Además, al tratarse de una reducción del espacio con una pérdida de hasta el 3% de información, es normal que aparezca en general algo más de error.

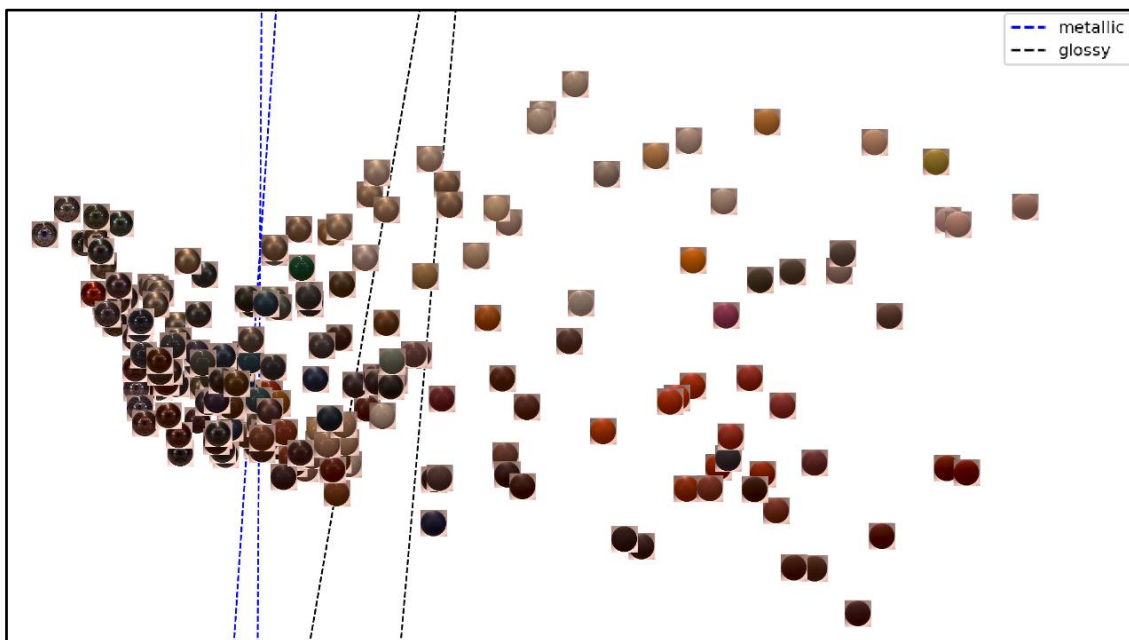


Figura 8. Límites materiales metálicos (azul) y brillantes (negro) sobre las esferas en el entorno de Saint Peter. Se corta el eje horizontal que distribuye los materiales de acuerdo con su especularidad. El límite de materiales brillantes está desplazado a la derecha de forma que incluye como brillantes aquellos materiales no metálicos que son especulares.

Para el atributo de luminosidad el resultado obtenido difiere bastante de lo esperado, ya que los valores de tasa de acierto medidos para este caso eran bastante bajos. Se aprecia en la Figura 9 que lo que este límite trata de separar son las muestras por su color. Esta discrepancia entre el resultado numérico, que parecía apuntar a una clasificación errónea o sin coherencia, y el resultado visual puede significar que, aunque el atributo no esté claramente puntuado en los diferentes materiales, el espacio esté

más organizado de lo que pueda parecer y que las personas dan importancia al color de un objeto para medir la distancia entre materiales.

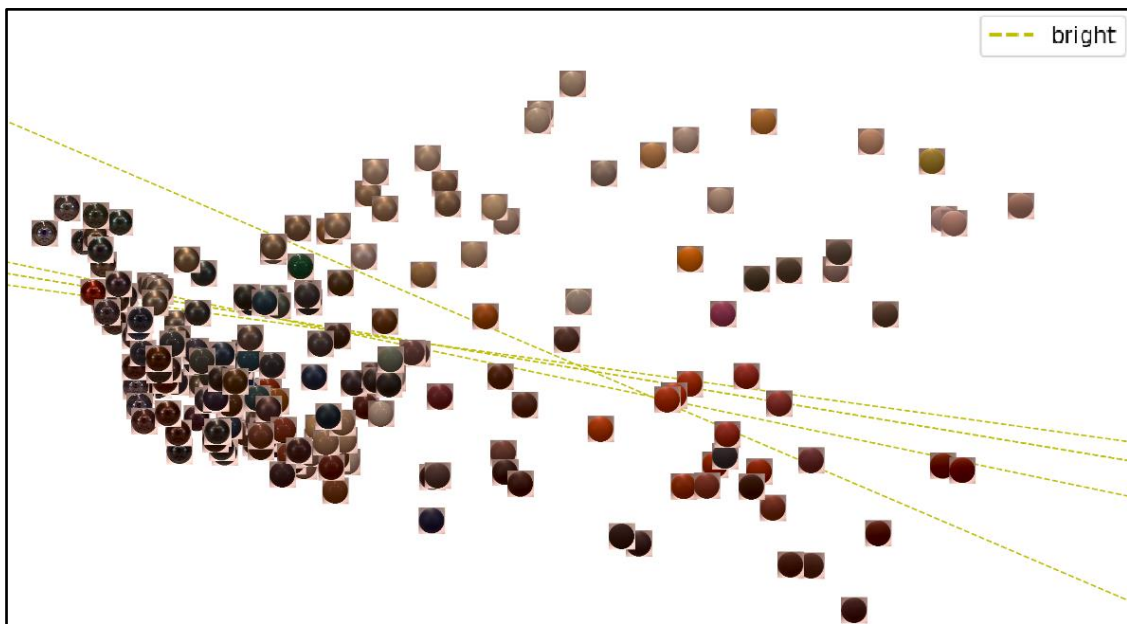


Figura 9. Límites para materiales luminosos en el entorno de Saint Peter. Se puede apreciar una clara dependencia con el color del material, dejando a un lado los colores claros.

Los planos de separación obtenida con otros entornos y formas son similares y consistentes respecto a los mostrados en las imágenes anteriores. Aunque, en ciertos entornos se obtienen mejores resultados para algunos atributos gracias a unas condiciones de iluminación favorables que pueden estar facilitando su reconocimiento. Esto se puede observar en la Figura 10 en la que se han obtenido varios hiperplanos para separar las muestras por su nivel de mate. En el caso de *Doge* los resultados varían más en comparación con *Saint Peter*. Probablemente este debido a que la luz es más tenue en el mapa de iluminación de *Doge*, dando la falsa impresión de que los materiales reflejan menos luz. Aun así, en ambos casos la separación encontrada es bastante buena, ya que tiene una disposición muy similar a la de los planos de separación de materiales brillantes. Se trata de dos propiedades opuestas que deberían tener una clasificación similar, pero con sentidos contrarios.

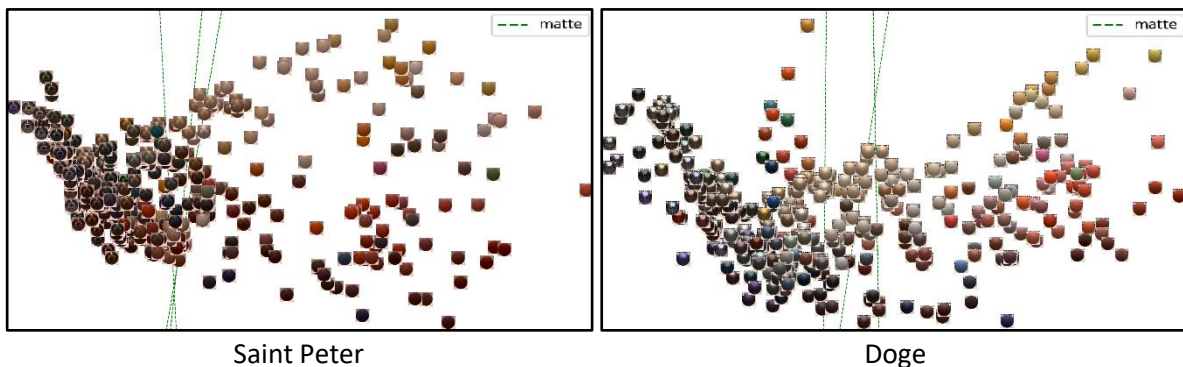


Figura 10. Límite de separación para materiales mate en dos entornos diferentes. Ambos ejemplos tienen un resultado muy similar, separando los mismos grupos de materiales. Sin embargo, se aprecia que en el caso del entorno de Doge los planos varían más. Esto puede deberse a la influencia que ejercen las condiciones de iluminación.



Para el atributo de plástico, aparentemente difícil de clasificar de acuerdo con lo visto en el subapartado anterior, los resultados varían en gran medida de un entorno a otro (Figura 11). En el caso del entorno *Doge* los planos generados son todos muy similares a los de materiales metálicos, ya que a priori se puede simplificar que un material si no es un metal es un plástico. Sin embargo, en el caso de *Saint Peter* los límites encontrados son casi perpendiculares a los de metalicidad, lo que resulta contradictorio. Este resultado difiere mucho de lo esperado, y puede estar algo sesgado por el color ya que asociamos más fácilmente materiales claros con plásticos. Otro motivo puede ser el 2% de información perdida por emplear dos dimensiones, que en el caso de clasificación de plásticos en *Saint Peter* puede ser importante.

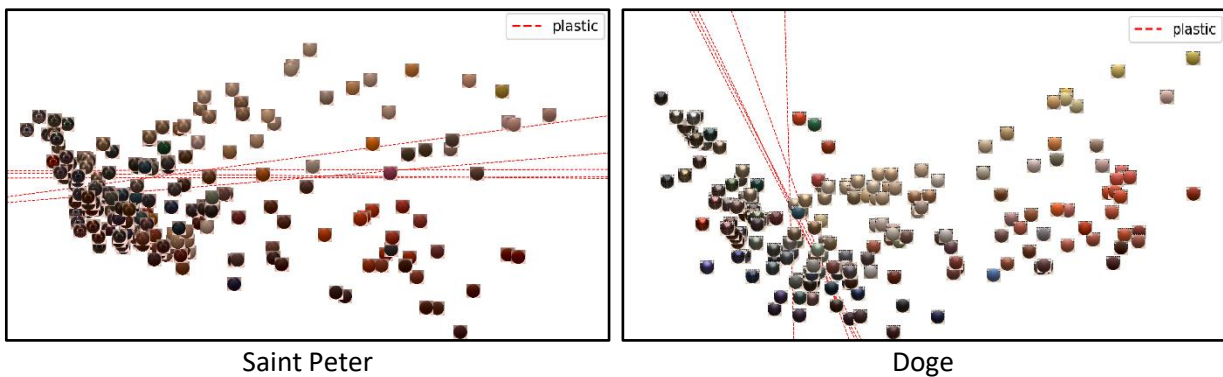


Figura 11. Límite de separación de plásticos en dos entornos distintos para esferas. Al tratarse de un atributo complejo el efecto de cambiar de entorno de iluminación tiene mayor efecto del esperado. Parece que el sesgo de color afecta a la clasificación de plásticos en *Saint Peter* por ser las muestras más oscuras en media que las de *Doge*.

Después de haber ajustado las componentes principales para las muestras de cada forma y entorno de iluminación por separado, se han ajustado empleando todas las muestras disponibles para comprobar que efectivamente todos los materiales se distribuyen de forma similar y uniforme. Si fuese un caso ideal y perfecto, los códigos latentes de un material variando su iluminación y forma deberían ser los mismos y, por tanto, al hacer uso de PCA aparecer en el mismo punto. Sin embargo, se trata de un espacio generado por una red entrenada con juicios humanos. Estos pueden llegar a variar mucho de una persona a otra, y la red entrenada no ha llegado a tener una tasa de acuerdo del 100% con las valoraciones humanas de distancia entre materiales. Además, al hacer uso de solo dos de las componentes, aunque se explique el 98% de la variabilidad de los datos, no es el total y por tanto se está perdiendo información. Por esto es esperable que un mismo material bajo diferentes condiciones varíe. En la Figura 12 se muestra cómo queda el espacio final conforme se van añadiendo los distintos conjuntos de imágenes.

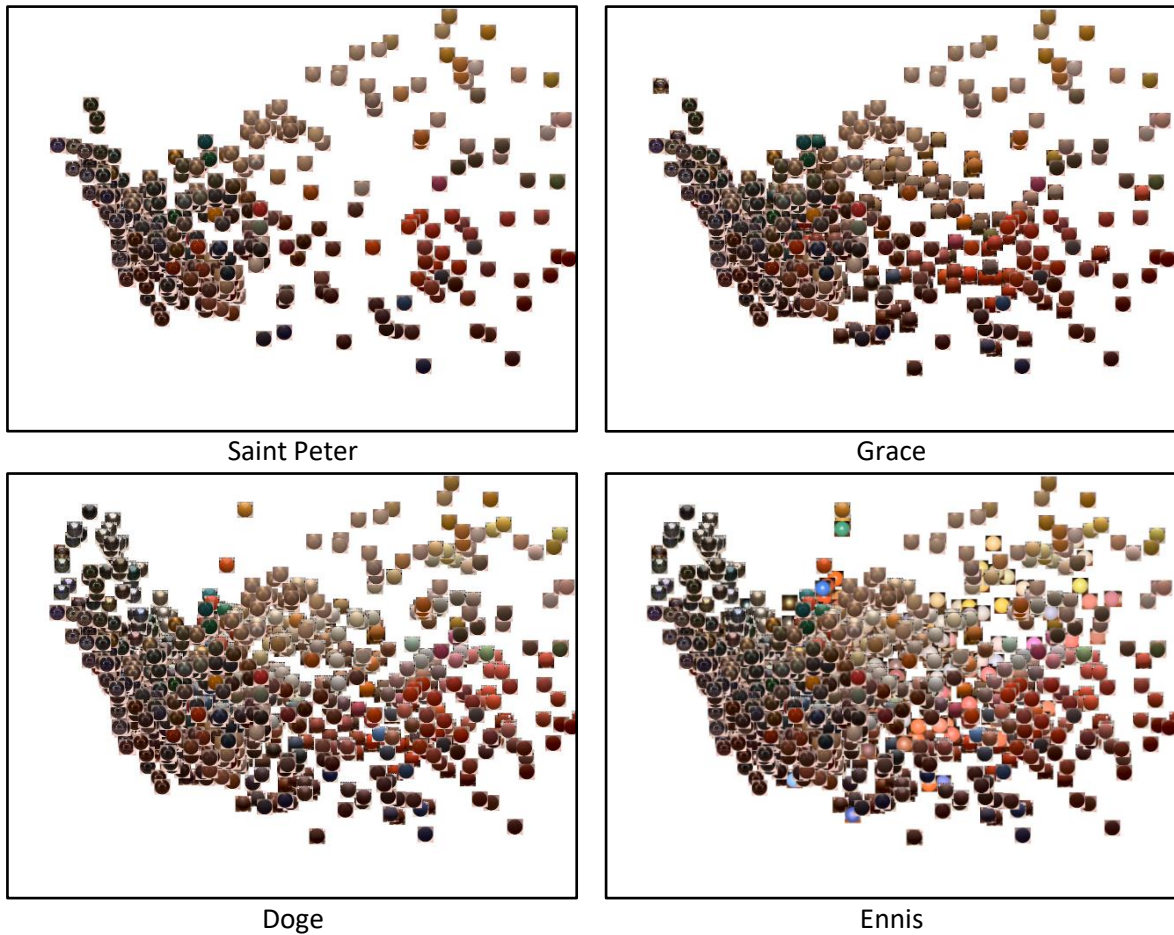
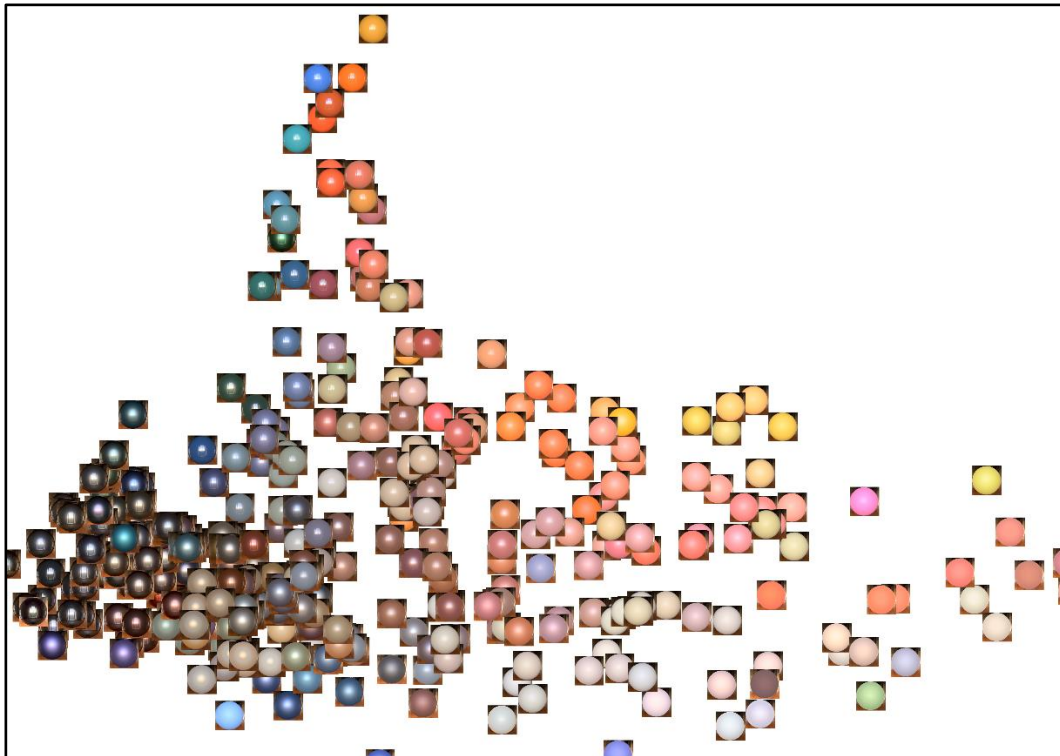


Figura 12. Componentes ajustadas con todas las muestras. De izquierda a derecha y de arriba a abajo en cada figura se han añadido las muestras de un entorno más. El entorno añadido corresponde con el nombre bajo la imagen. Se aprecia como el espacio se va rellenando por colores y nivel de brillo de acuerdo con la distribución general del espacio.

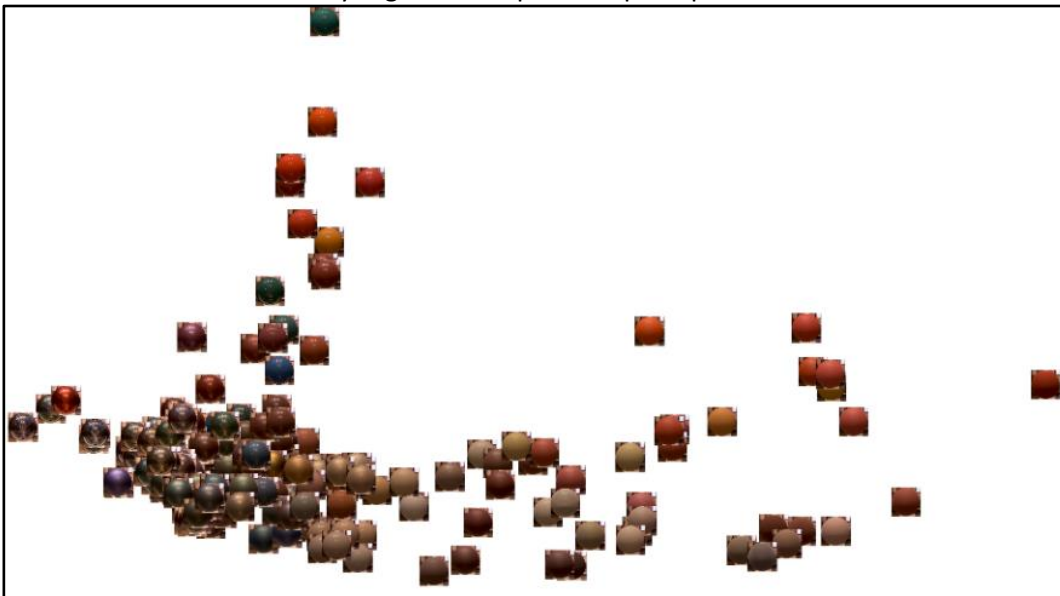
A medida que se añaden las muestras con diferentes iluminaciones se puede ver que la distribución general de materiales es la misma, y que con cada nuevo grupo de imágenes se completa un poco más el espacio, llenando los huecos por gamas de colores y reflejos. Este es el comportamiento esperado, pues se había visto ya que los dos ejes iban determinados por el nivel de brillo material y el color. Por eso, debido a que de un entorno a otro cambia el tipo de iluminación, los materiales aparecen más o menos brillantes y con un tono más o menos claro rellenando coherentemente el espacio. Al igual que en todos los casos anteriores, las imágenes de *Ennis* con los colores más vivos aparecen algo descolocadas.

Como se puede ver en la Tabla 8, en algunos casos la tercera componente principal llega a suponer más de un dos por ciento de la variabilidad de los datos, teniendo un valor cercano al de la segunda componente. Debido a esto resulta interesante estudiar si existe también un patrón que explique cómo varían las muestras en esa dimensión. Para realizar esta prueba se han generado los gráficos de forma similar a los anteriores, pero mostrando la primera y tercera componente en vez de la segunda. De este modo se puede ver la forma de la tercera componente, y si guarda alguna relación con la primera, que explica más del 90%. En la Figura 13 se muestra un ejemplo, y en este se puede apreciar que el factor de variabilidad del tercer eje es aparentemente lo plástico que es

un material, haciendo diferenciación entre los materiales especulares de colores más vivos y oscuros. En cierto modo el aspecto del gráfico de *Grace* para las componentes uno y tres recuerda al aspecto de las dos primeras componentes de *Ennis*. Esto puede significar que en *Ennis* la segunda componente corresponde con la tercera del resto de entornos, que tienen menos luz y los materiales parecen más oscuros.



Primera y segunda componente principal en *Ennis*



Primera y tercera componente principal en *Grace*

*Figura 13. Comparación entre componentes de PCA. La primera imagen muestra la disposición del espacio para las dos primeras componentes en el entorno de Ennis. La segunda imagen muestra el espacio para la primera y la tercera componente de Grace. Se aprecia una clara correlación entre las imágenes. Esto sugiere que tal vez en Ennis al haber más luz, y los materiales parecer más claros, la segunda componente coincide con la tercera componente del resto de mapas de iluminación.*



Los análisis empleando PCA han permitido entender mejor cómo es el espacio de similitud entre materiales, y que posibles bases toma la percepción humana para medir o separar materiales por su similitud. El primer punto que destacar es la sencillez aparente del espacio, debido a que con una sola dimensión se puede explicar más del 90% de la variabilidad de los datos. Además, de forma más o menos clara, las muestras se colocan a lo largo de la primera dimensión de acuerdo con su nivel de brillo.

El segundo punto que destacar es la disposición de los límites semánticos, que coincide en gran medida con lo esperado, la mayoría son planos verticales que cortan el eje horizontal a diferentes alturas. Este hecho concuerda con los análisis realizados en el trabajo de Serrano *et al.* (2016), que muestra una fuerte correlación entre la mayoría de atributos empleados en este trabajo. La principal sorpresa de los planos de separación del espacio se ha visto en el límite que clasifica las muestras de acuerdo con su luminosidad. Es paralelo al eje horizontal, y parece separar aquellos materiales de colores claros, como el amarillo o gris, de los materiales oscuros.

En tercer lugar, resulta interesante el modo en el que las muestras de los diferentes entornos completan poco a poco el espacio (Figura 12) cuando se han ajustado las componentes de PCA con todas las imágenes. Esto es debido probablemente a que los cambios de iluminación en las personas sí que tiene un efecto en la percepción, que se ve reflejado en el espacio al agrupar las imágenes por colores.

Finalmente, destaca la distribución de los materiales ajustando las componentes de PCA con las imágenes de *Ennis*. La forma que tiene el espacio en ese caso es muy similar a la que tienen el resto de las muestras en los otros mapas de iluminación para la primera y tercera componente. Puede estar debido a que el efecto de la fuente de luz en *Ennis* es tan alto que, al alterar significativamente la percepción de los materiales, la componente que corresponde con la tercera para otros entornos cobra más importancia y se convierte en la segunda.

## 5. Análisis interno de la red

Hasta ahora solo se ha analizado el espacio latente de la red de Lagunas *et al.*, con esto se ha podido entender o aproximar de qué manera percibimos y distinguimos unos materiales de otros. Los diferentes análisis han dado unos resultados mejores de lo esperado, en algunos casos mostrando que tal vez el espacio de similitud entre materiales sea más simple de lo que se cree. Sin embargo, no hay que olvidar que se trata de un espacio generado por una red neuronal que ha aprendido a detectar patrones y elementos de las imágenes de entrada para reconocer los materiales. En esta sección se va a estudiar el interior de la red para buscar características detectadas en las distintas capas, que puedan explicar o aclarar cómo se ha construido el espacio latente.

Al tratarse de un tipo de red convolucional con muchas capas intermedias y con un número de canales que aumenta conforme más adentro de la red se está, el total de neuronas a revisar es demasiado grande. Para poder tratar este problema de una forma eficiente que permita observar las activaciones más importantes, el modelo a seguir va a ser el de *top-9* activaciones [Yosinski *et al.*, (2015); Chu, Yang and Tadinada (2017)]. Consiste en seleccionar de cada canal solamente las nueve neuronas con los valores de activación más altos. Con esto se reduce en gran medida la cantidad de datos generada y al mismo tiempo permite revisar solo las neuronas más relevantes. El método exacto para visualizar que ha provocado las activaciones es la retropropagación guiada en redes convolucionales [Springenberg *et al.*, 2015]. Esta técnica es una combinación de la deconvolución (pasada hacia atrás en la red) y la retropropagación clásica (cálculo del gradiente de la función de coste a partir de la entrada). Dado que en las redes convolucionales una neurona solamente busca características en un corte de la imagen, los resultados generados corresponden con la parte de la imagen que ve la neurona. Para cada activación se generan dos imágenes (Figura 14), el recorte correspondiente resultante de la retropropagación guiada (izquierda de la Figura 14) y el de la imagen original. De esta manera resulta más fácil visualizar que provoca las distintas activaciones.

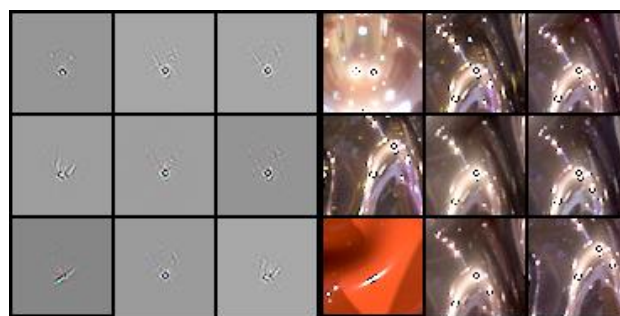


Figura 14. Ejemplo de *top-9* activaciones por reflejos especulares. Cada cuadrado de la mitad izquierda corresponde con el resultado de la retropropagación guiada, del corte de imagen de la derecha.

En las redes convolucionales conforme más profunda es la capa en la que se encuentra una neurona mayor es la sección de imagen que analiza. Las neuronas de las primeras capas, que solo observan un pequeño trozo de la imagen de entrada, suelen corresponder con características concretas y más fácilmente identificables. Por el contrario, las neuronas de las últimas capas pueden llegar a estar viendo la imagen

entera y buscar propiedades más abstractas de la imagen, que puedan ser menos claras a primera vista para una persona. Con esto en mente se han analizado las imágenes en busca de patrones dentro de las distintas capas o diferencias entre los primeros y los últimos niveles. Se han encontrado diferentes especializaciones detectadas por las neuronas de la red, las principales son: formas, colores, reflejos y patrones en los reflejos.

En cada uno de estos tipos de especialización las neuronas se activan por características concretas de las imágenes. En la Figura 15 se muestran en cada fila dos ejemplos de las clases principales:

- **Forma:** Resalta parte del contorno del objeto central, el material, normalmente con forma redondeada.
- **Patrón:** Resalta elementos del entorno de iluminación y no características del material. La mayoría de los casos son patrones tipo cuadrícula.
- **Color:** Resalta colores específicos de los materiales.
- **Reflejo:** Resalta reflejos especulares definidos de las fuentes de luz.

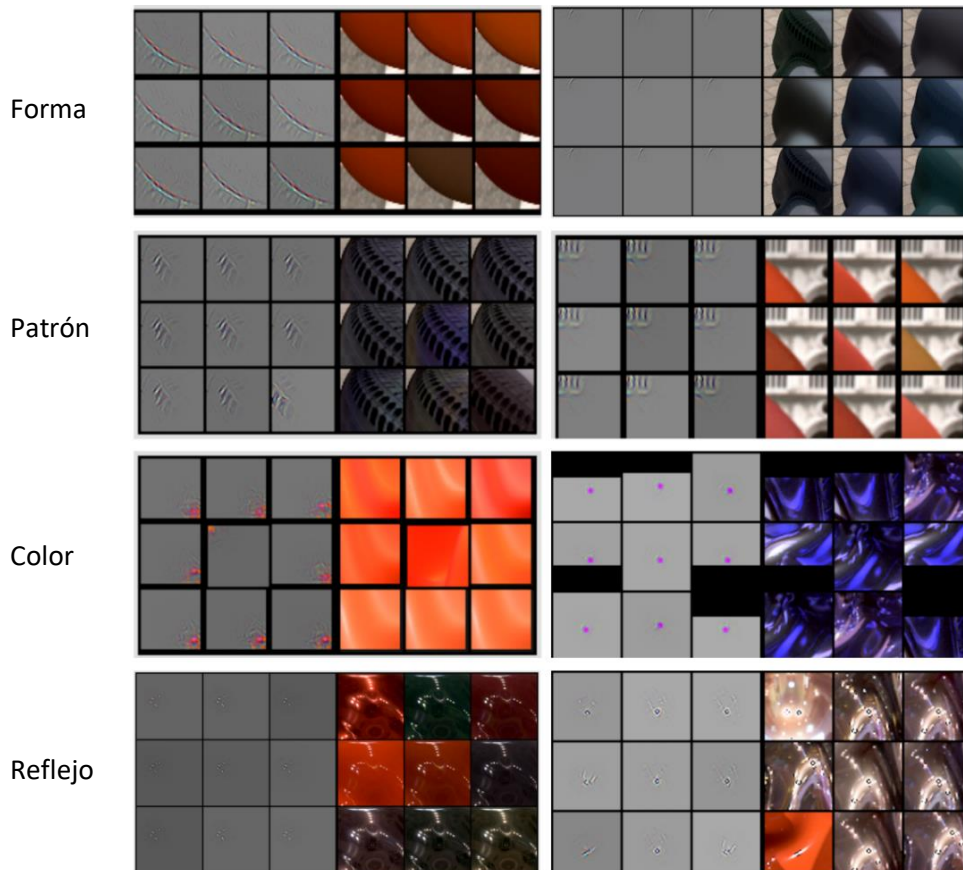


Figura 15. Tipos de especialización de activaciones. Cada fila resalta un tipo de activación diferente, y corresponde con los tipos de activaciones más informativos y repetidos en el interior de la red.

La única capa que presenta un comportamiento común en las *top-9 activaciones* de los distintos canales es la última convolución. Hay muchos canales, de los 512 de la capa, en los que las mejores nueve activaciones corresponden con un tipo claro de material. Se agrupan según propiedades como el color o el aspecto más o menos metálico entre

otros. En la Figura 16 se pueden ver dos ejemplos, el primero está detectando materiales con aspecto mate. El segundo ejemplo, es una selección de materiales rojizos difusos. Esto apunta a que hasta la última capa no se empiezan a agrupar las muestras por sus propiedades, y se genera la distribución que tendrán los diferentes materiales en el espacio latente.

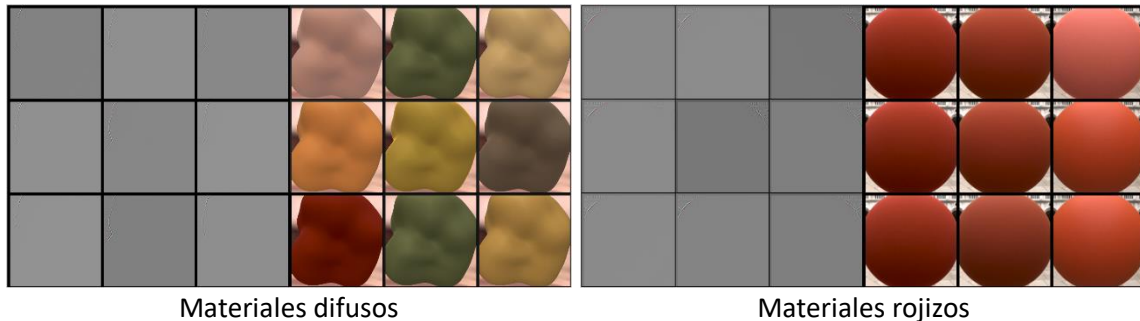


Figura 16. Ejemplos de clusterización de materiales de acuerdo con alguna propiedad física en los canales de la última capa de la red. En la derecha se muestra un ejemplo en la que se agrupan materiales de distintos tonos de rojo. En la izquierda se muestra un ejemplo de agrupación de materiales con la superficie difusa o mate.

Se ha visto que en algunas capas hay mayor presencia de alguna de las especializaciones mencionadas, sobre todo hacia las últimas capas aumenta el número de activaciones centradas en brillos especulares nítidos. Aunque al ser en las capas finales, en las que una neurona ve una parte muy grande de la imagen, puede que la red esté detectando un tipo de característica de más bajo nivel que no seamos capaces de percibir visualmente. En algunos canales se han encontrado activaciones que no se tiene muy claro a que corresponden, o que entre las mejores nueve difiere lo detectado, dejando ver que el interior de la red tiene partes complejas que podemos no llegar a entender completamente.

Los resultados obtenidos no han sido tan esclarecedores como se esperaba, no se ha podido encontrar un comportamiento claro en el interior de la red. Pero, sí que se ha podido ver que para generar el espacio latente en algunas partes de la red se están activando neuronas por la presencia de ciertos colores, o de elementos relacionados con el entorno en el que está el material. Esto concuerda con lo visto en el apartado anterior y con lo esperado, pues se trata de un espacio que representa la similitud entre materiales basada en la percepción humana, que depende de cada individuo y hasta cierto punto de las condiciones de iluminación y forma.

## 6. Reconstrucción del espacio latente

En los apartados anteriores se ha analizado la percepción humana de similitud entre materiales mediante el estudio del espacio latente y el interior de la red que lo ha generado. Los mejores resultados se han obtenido del análisis del espacio mediante técnicas de reducción de dimensiones, que han mostrado una mayor simplicidad de lo esperado en cómo se compone. Con estos resultados en mente, se ha reconstruido el espacio latente implementando un sistema GAN (*Generative Adversarial Network*) con un codificador de entrada incorporado, para poder generar imágenes de materiales a partir de un código latente de nuestro espacio.

La arquitectura final de sistema (Figura 17) consta de tres elementos principales:

- **Codificador:** Red convolucional residual ya entrenada que a partir de una imagen genera un código latente de 128 dimensiones.
- **Generador:** Red convolucional que genera una imagen, del mismo tipo que las de entrada, a partir de un código latente de 128 dimensiones.
- **Discriminador del generador:** Red neuronal para clasificar entre imágenes reales y generadas. Apoya el aprendizaje del generador.

A excepción del codificador los otros componentes del sistema han tenido que ser entrenados, y ha sido necesaria una etapa de ajuste de los hiperparámetros para tratar de mejorar todo lo posible el rendimiento. El codificador no ha necesitado de entrenamiento ya que se trata de la red de Lagunas *et al.* (2019), es decir, la red empleada en los apartados anteriores para el estudio del espacio latente.

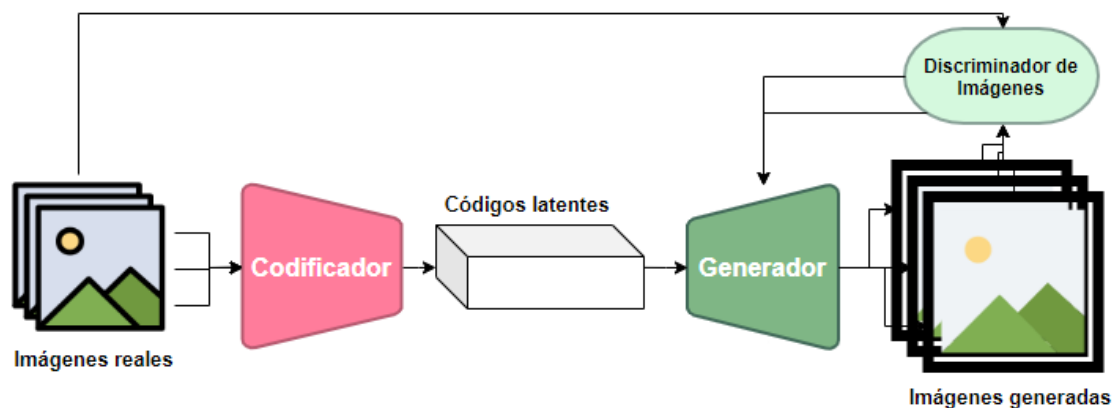


Figura 17. Arquitectura del sistema empleado para entrenar un generador de imágenes a partir del espacio latente de la red neuronal estudiada.

Para el diseño de los nuevos elementos del sistema, al tratarse de una prueba para ver si es posible reconstruir con facilidad el espacio, se han empleado arquitecturas de red no muy complejas que permitan entrenar más deprisa las diferentes configuraciones probadas. Se trata de redes convolucionales simples para obtener un DCGAN (*Deep Convolutional Generative Adversarial Network* (Radford, Metz and Chintala, 2016)), que es un tipo de arquitectura recomendada para generación de imágenes. Al tratarse de un sistema poco tradicional por el hecho de emplear un codificador, lo habitual sería tener

algún sistema que se encargase de ajustarlo y que aportara una parte del peso en la función de error del generador, para incrementar la calidad de las imágenes generadas. Sin embargo, al tratarse de un caso especial en el que se emplea un codificador ya entrenado y ajustado [Lagunas *et al.*, 2019] no es necesario.

Las principales modificaciones respecto a un sistema DCGAN se encuentran en el generador, ya que su función de coste se ha modificado añadiendo nuevas partes. La mitad de los códigos latentes que se le pasan al generador para entrenarlo corresponde directamente con los códigos de las imágenes que se introducen en el codificador, y por tanto se sabe exactamente el valor de los píxeles de las imágenes a generar. Gracias a esto se ha podido emplear una función de distancia entre píxeles para medir el error de la imagen generada. Esto permite aprender más deprisa a reproducir las imágenes de entrenamiento. Sin embargo, no se busca únicamente reproducir el conjunto de datos, sino ser capaces de generar imágenes realistas de materiales no vistos durante el aprendizaje. Para poder mejorar esto, se han empleado interpolaciones lineales entre los códigos latentes de las imágenes codificadas, para que el generador trate de aprender a generar estos materiales intermedios. Finalmente, tras varias pruebas, el principal obstáculo para obtener interpolaciones de calidad entre distintos materiales era la aparición de mucho ruido en las imágenes nuevas, y que suponía un claro problema para el realismo de los materiales. Para paliar este problema todo lo posible, la medida adoptada ha sido añadir a la función de coste del generador la pérdida de variación total (*Total Variation Loss*). Esta función calcula la suma en valor absoluto de la diferencia entre píxeles vecinos de una imagen para poder medir el nivel de ruido. Los detalles de los valores de los hiperparámetros usados, las funciones de coste y los algoritmos de optimización están detallados en el Anexo 2 del documento.

Con el objetivo de aumentar la apariencia de las muestras para el entrenamiento, se ha renderizado de nuevo el total de 406 materiales en el entorno de *Doge* y *Saint Peters*, generando para cada material 10 imágenes desde diferentes puntos de vista (Figura 18). Con esto se consigue que para un mismo material haya varias imágenes con diferentes grados de iluminación. Además, se ha sustituido el fondo original de las imágenes por un fondo blanco, facilitando la tarea del generador y haciéndola más coherente con el objetivo, que es poder generar materiales, no ser capaces de recrear el fondo de una imagen. Por el mismo motivo de centrar la tarea solo en generar materiales, la geometría elegida es la esfera.

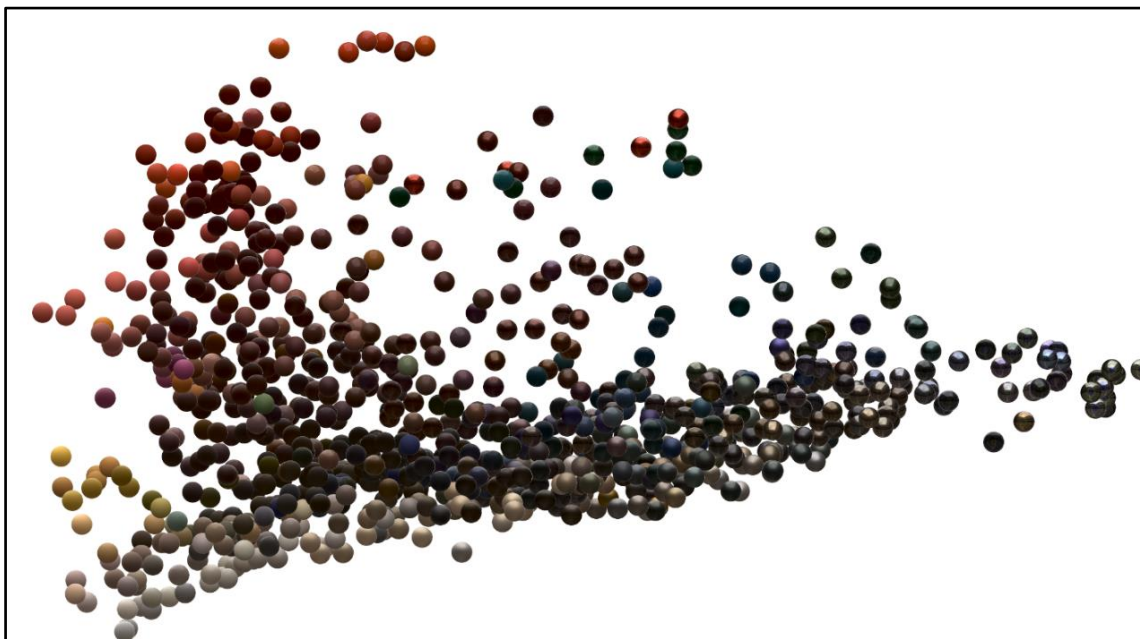


Figura 18. Cuatro imágenes del mismo material (*Tungsten Carbide*) desde diferentes ángulos en el entorno de *Doge*.

Para asegurar que el nuevo conjunto de datos comparte las propiedades encontradas durante el estudio del espacio latente, se han realizado pruebas similares a las de los apartados previos. Es importante que el conjunto tenga una clara similitud con los



resultados de PCA obtenidos hasta ahora. Por tanto, se ha calculado la variabilidad que explican las principales componentes del nuevo conjunto, y con las tres primeras se explica el 96.79% de la variabilidad, el 87% correspondiente a la primera componente. Si bien no es tan elevado el valor de la primera componente como en las pruebas iniciales, sigue siendo un valor muy alto con unos resultados similares y acordes a lo ya observado (Figura 19). La disminución del valor que representa la primera componente es debido al aumento de las muestras y de la variabilidad con la que se presenta el entorno. La mayor diferencia apreciable en la visualización del espacio con las dos primeras componentes es que el sentido de los ejes se ha invertido. Pero, la explicación de cada dimensión sigue siendo la misma en cuanto a la apariencia de los materiales. El cambio de sentido no es realmente relevante, ya que con PCA se encuentran las direcciones de mayor cambio en un conjunto de datos.



*Figura 19. Gráfico PCA con las 2 primeras componentes en el entorno Doge. El resultado comparte la estructura general de las pruebas previas, cambia el sentido de las dimensiones respecto a los primeros análisis.*

Finalmente, tras diversas pruebas y ajustes de los hiperparámetros de las diferentes partes del sistema, los resultados logrados son muy buenos (Figura 20), con cierta reducción de la nitidez en algunos de los materiales más especulares, ya que suponen la mayor dificultad para el generador. Es posible reconstruir fielmente el espacio latente incluso forzando a aprender todo el espacio y no solo las imágenes empleadas para el entrenamiento.

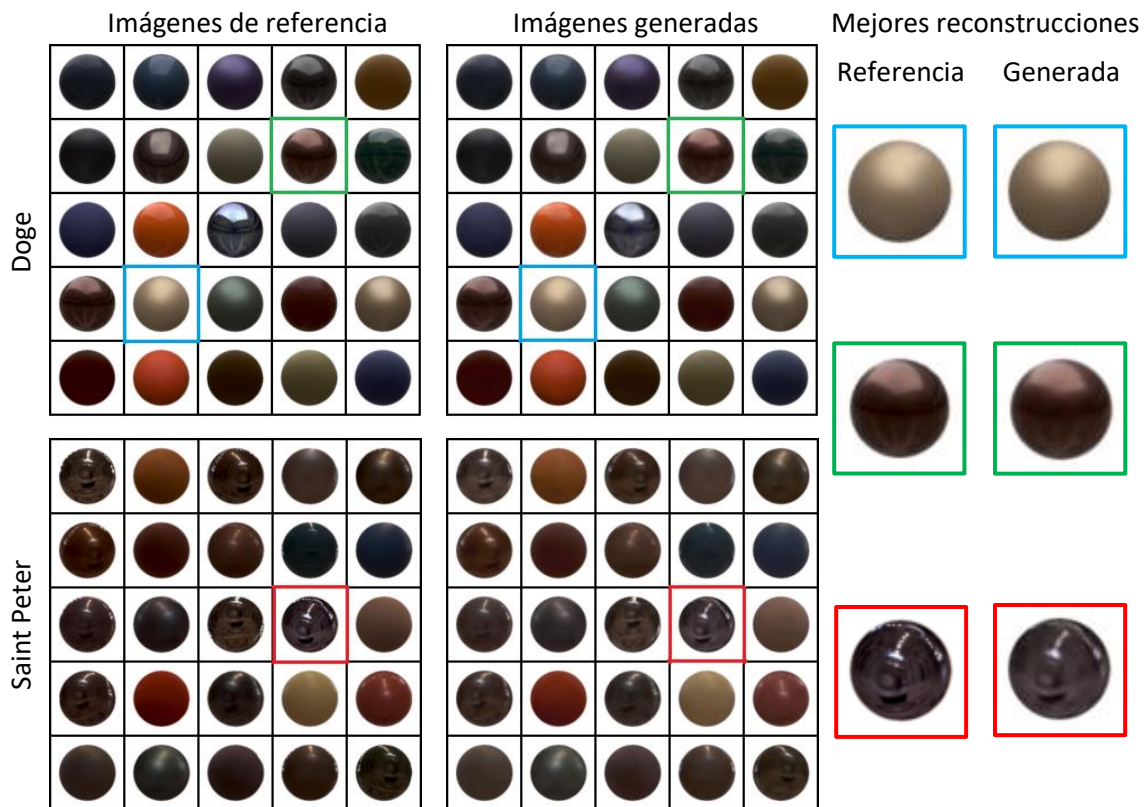


Figura 20. Resultados del entrenamiento del sistema. En la columna central aparecen las imágenes obtenidas con el generador, y en la primera columna las imágenes reales correspondientes. Se han remarcado de color tres ejemplos para visualizar más claramente las reconstrucciones.

### 6.1. Edición material

Como se puede ver en la Figura 20, los resultados generados después del entrenamiento son bastante precisos, solo se aprecia la disminución de nitidez en alguno de los materiales con los reflejos más definidos. Esto se debe principalmente al uso de la función de coste de variación total, que a cambio de reducir el ruido en las imágenes disminuye ligeramente la calidad. Sin embargo, como se ha comentado previamente, no solo se busca poder reproducir las imágenes del entrenamiento, sino generar imágenes de materiales nuevos. Para poder hacer esto de la mejor forma posible es necesaria esta función de coste, o de otro modo en las interpolaciones de los materiales más lejanos aparece mucho ruido (Figura 21).



Figura 21. Ambas imágenes representan el resultado de interpolar los dos mismos materiales. Solo se diferencian en el uso de la función de coste de variación total en el entrenamiento del generador de imágenes.



Pese a la mejora en la capacidad del generador para reconstruir el espacio, hay ciertos tipos de interpolaciones que no son posibles sin algo de ruido visible. Se trata principalmente de interpolaciones entre materiales especulares metálicos y difusos con tonos de color diferentes (Figura 22.a), esto se debe a que la red tiene que suponer como sería el reflejo del entorno sobre la esfera a la vez que cambia el color de esta. También está causado porque en el conjunto de datos no se representa completamente el espacio latente, y por tanto la red no tiene información suficiente para aprender que aspecto deberían tener los materiales en esos casos. Para las interpolaciones entre materiales que comparten la mayoría de sus propiedades, los resultados son muy buenos, con cambios paulatinos y suaves entre imágenes (Figura 22.b-f).

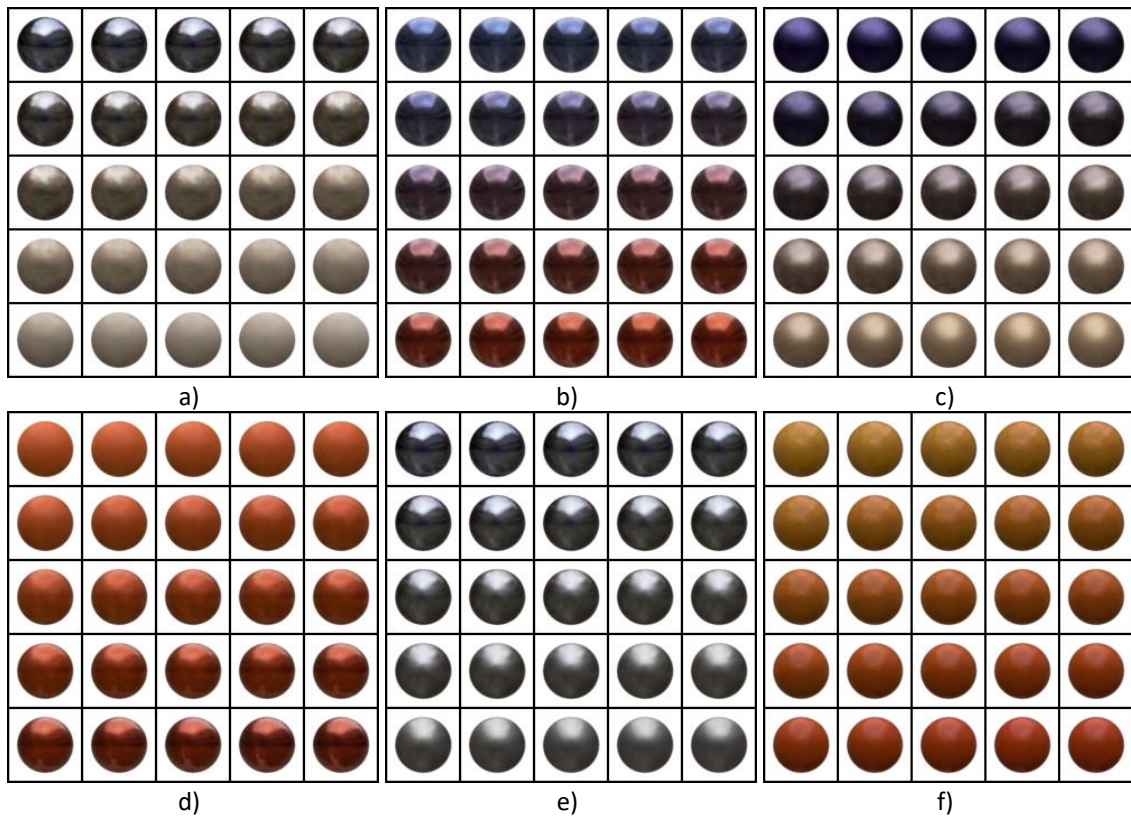


Figura 22. Resultados del generador al usar distintos tipos de materiales para interpolar sus códigos latentes. **a:** materiales muy diferentes. **b:** materiales especulares. **c:** materiales difusos. **d:** materiales rojos. **e:** materiales metálicos. **f:** materiales plásticos.

### 6.1.1. Sistema de edición intuitivo

Tras las observaciones hechas a lo largo del proyecto y visto que la reconstrucción del espacio latente es posible, se ha intentado encontrar un método intuitivo de edición material sin manipular directamente los píxeles de la imagen, alterando las propiedades de un material a través de modificaciones sobre su código latente. En concreto resulta de gran interés tratar de llevar a cabo esta tarea en el espacio reducido con PCA, ya que tenemos una idea más clara de cómo debería variar un material modificándolo en las dos primeras dimensiones.

Para llevar a cabo este experimento, el procedimiento ha consistido en ajustar las componentes de PCA para los códigos latentes del nuevo conjunto de imágenes, coger la imagen que se desea editar y transformar su código latente al espacio reducido. Al

tratarse de modificaciones en un espacio diferente del original, si no se trabaja con todas las dimensiones de PCA, es decir, la transformación del código latente de 128 dimensiones se hace a un vector de menos dimensiones, habrá información que se perderá. Esta pérdida de información se reflejará en que al volver al espacio base el punto diferirá de su posición original, y la imagen generada no corresponderá exactamente con la de referencia. Tras varias pruebas (Figura 23) se ha visto que cuantas menos dimensiones se emplean, menos características del material son representadas, y por tanto más difiere del original. Por eso, un punto medio permite mantener la apariencia original del material, a la vez que se facilita la edición de los códigos latentes.

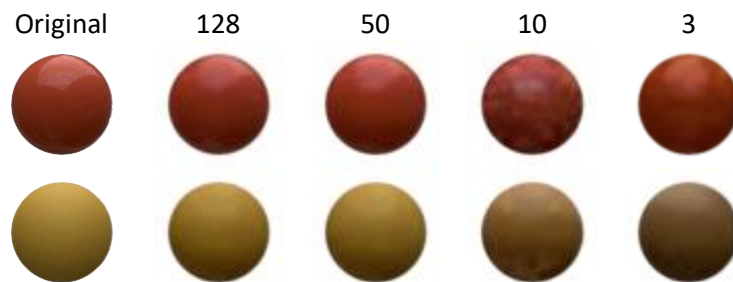


Figura 23. Reconstrucción de los materiales con diferente número de dimensiones. En cada columna se ha reconstruido el material con un número distinto de componentes extraídas de PCA.

Finalmente, tras observar los resultados de reducir a diferente número de dimensiones se ha decidido emplear 50, que representan el 99.68% de la variabilidad de los datos. Supone una pérdida de información mínima que facilita la edición de los materiales. Una vez transformado el vector se pueden editar las dos primeras componentes y observar la variación de la imagen resultante (Figura 24). Una modificación de la primera componente supone una alteración del brillo o los reflejos del material principalmente, mientras que una modificación de la segunda se ve reflejada en un cambio de color. Al tratarse de una representación con pérdida de información, las modificaciones sobre una componente no son perfectas, y en algunos casos conllevan la aparición de ruido. La mayor dificultad para el generador son los movimientos hacia las zonas con menos densidad de materiales, ya que no sabe con tanta exactitud cómo es esa parte del espacio.

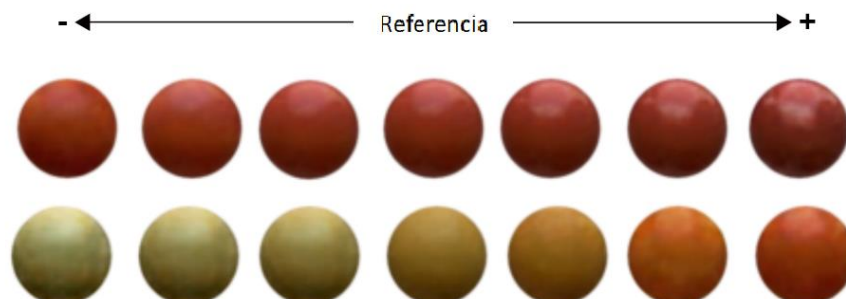


Figura 24. Edición material en el espacio PCA. La primera fila corresponde con cambios en la primera componente, y la segunda fila con cambios en la segunda componente. En ambos casos los cambios en la apariencia de las imágenes de referencia concuerdan con lo esperado. La primera componente afecta a los reflejos, y la segunda al color.

Los resultados de esta sección apuntan a que el espacio basado en percepción humana de similitud entre materiales se puede reconstruir de forma precisa, de acuerdo con los objetivos que se están tratando de alcanzar. Se ha visto que es posible editar los materiales alterando los códigos latentes (vectores de 128 dimensiones), sin necesidad de interactuar directamente con el espacio de píxeles de las imágenes de entrada, que supondría trabajar en un espacio de  $64 \times 64 \times 3$  dimensiones (número de píxeles de las imágenes). Además, las modificaciones sobre los códigos latentes reducidos mediante PCA producen unos resultados coherentes y consistentes con lo esperado dada la distribución mostrada en la Figura 19.

## 7. Descomposición semántica del espacio latente

Adicionalmente, se ha realizado una prueba para obtener un sistema GAN que permita la descomposición semántica del espacio latente en dos partes, el material y la forma del objeto a generar. Este sistema se basa en el estudio de Donahue *et al.* (2018) que desarrolla un sistema, SDGAN (*Semantically Decomposed GAN*), que permite separar un espacio latente de imágenes de caras de personas en dos partes: la identidad de la persona y cualquier otro rasgo (por ejemplo, el tono de piel o la iluminación). Para este estudio se ha adaptado la implementación ofrecida por los autores del trabajo. Tras numerosas pruebas y ajustes, los resultados obtenidos han sido de bastante calidad con un buen nivel de realismo para las imágenes generadas (Figura 25). Pudiendo finalmente generar diferentes materiales con la misma forma, o generar el mismo material con diferentes formas.



*Figura 25. Resultados del entrenamiento del SDGAN. Cada fila comparte el mismo código latente para la identidad del material. Cada columna comparte el mismo código latente para la forma. Se puede apreciar como ambos aspectos se conservan correctamente en las dos direcciones.*

Una vez entrenado el generador se ha analizado el espacio latente aprendido para la parte de identidad, y aunque no se esperaba que tuviera una disposición tan buena como la del espacio generado con la red de Lagunas *et al.*, podía haberse esperado cierta organización entre imágenes similares. Sin embargo, el espacio tiene una distribución aleatoria sin apenas correlación entre materiales, lo que no permite ningún tipo de manipulación intuitiva. La única forma de controlar hasta cierto punto la edición de los materiales es mediante la interpolación de códigos latentes, para poder comprobar si se ha aprendido correctamente todo el espacio. Para llevar a cabo esta prueba, dado que el espacio latente está dividido en dos partes completamente diferentes (identidad y forma), se han realizado interpolaciones manteniendo fija una de las dos partes, interpolando entre dos identidades o entre dos geometrías. Un ejemplo de estas pruebas se puede ver en la Figura 26, que muestra unos resultados intermedios muy suaves. Se mantiene perfectamente la forma al variar la identidad, y se mantiene la identidad al cambiar la forma.



*Figura 26. Interpolaciones de identidad y observación. En la primera fila se han interpolado dos tipos de material, manteniendo la forma. En la segunda fila se han interpolado dos observaciones manteniendo la identidad.*

Con estos resultados se ha visto que es posible entrenar un sistema capaz de dividir semánticamente el espacio latente en dos, de manera que se puede manipular independientemente la identidad y la forma de un material. Sin embargo, por la naturaleza aleatoria del muestreo del espacio latente para el entrenamiento, no es posible un control intuitivo de los códigos latentes para generar nuevas imágenes.

## 8. Trabajo futuro

Con vistas a futuros trabajos, la parte que resulta más interesante de continuar explorando es la referente a la generación y edición de imágenes de materiales. Se ha visto que es posible reconstruir y editar con éxito el espacio latente basado en percepción humana, y se ha comprobado también que es posible el aprendizaje de un espacio aleatorio que separa identidad y forma del material. Por tanto, resultaría de gran interés tratar de combinar ambos modelos en uno solo, de forma que la parte de identidad aprendida en el SDGAN correspondiera con el espacio latente conocido y estudiado, permitiendo posteriormente editar libremente la identidad del material. Esta propuesta es plausible ya que tras diversos análisis todo parece apuntar que el espacio latente no es tan complejo, y tiene una distribución coherente que puede facilitar el proceso de aprendizaje. Adicionalmente, para tener un sistema más completo se podría tratar de entrenar una pequeña red que pudiera codificar imágenes de modelos 3D, obligando al generador a aprender esa distribución de formas de objetos. Con esto se alcanzaría una herramienta con la que controlar totalmente la edición de materiales sin necesidad de edición directa de los píxeles de la imagen. De lograr entrenar con éxito este sistema similar a un SDGAN, se tendría un espacio controlable de forma intuitiva que permitiría generar modelos de objetos con diferentes materiales fácilmente editables. Se muestra en la Figura 27 una vista de alto nivel de la arquitectura que podría tener este sistema.

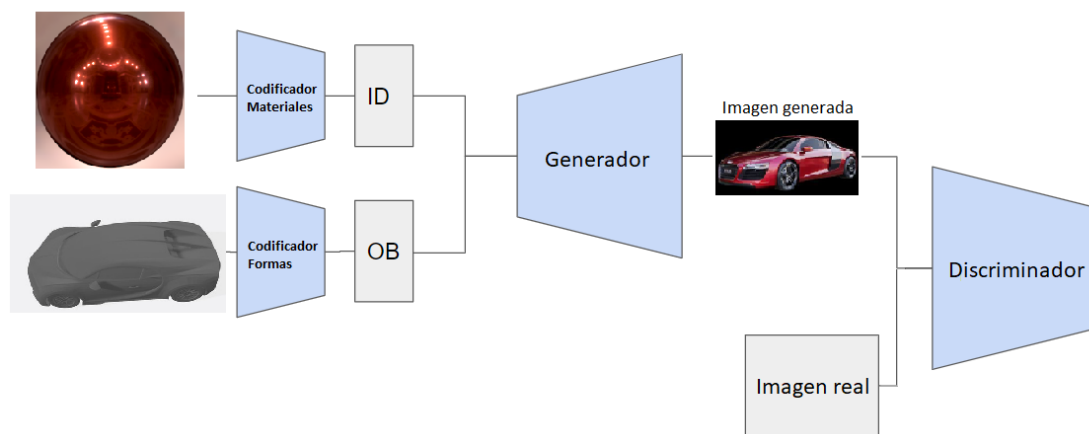


Figura 27. Modelo ideal de sistema para generar materiales y geometrías de forma manipulable. Se tiene un codificador de materiales para controlar la identidad (ID) y un codificador de formas para controlar el aspecto de la imagen (OB). Estos dos elementos permitirían tener un espacio controlable para manipular la identidad y forma de los objetos a generar.

## 9. Conclusiones

Tras realizar diferentes análisis sobre el espacio latente y la red neuronal, los resultados apuntan a que la percepción humana de similitud entre materiales, aunque compleja, lleva a una agrupación de materiales relativamente sencilla. El resultado más sorprendente ha sido que con una sola componente del PCA se explica en torno al 90% de la variabilidad de los datos. Esto puede explicar que la agrupación de materiales sea lo suficientemente clara como para separar unos materiales de otros de acuerdo con ciertas propiedades físicas. Además, la organización de los materiales depende casi enteramente de su aspecto especular, lo que probablemente ha facilitado la reconstrucción del espacio, siendo capaces incluso de manipularlo.

Con los análisis de la primera parte, se ha podido ver que el espacio estudiado es separable mediante hiperplanos para propiedades de los materiales como el nivel de brillo o mate. Sin embargo, para algunas propiedades como la luminosidad de un material que es más compleja, y de una persona a otra las valoraciones cambian bastante, los resultados no son claros, apuntando a que el espacio no tiene separados los materiales para este atributo.

Con un segundo análisis, empleando PCA se ha observado que con la primera componente más importante se puede explicar alrededor del 90% de la variabilidad de las muestras, y con las dos primeras en torno al 95%. Un aspecto que resulta de mucho interés es que la primera dimensión tiene una visible correlación con el nivel de brillo y especularidad del material. Esto parece sugerir que para distinguir materiales la decisión principal gira en torno al reflejo de la luz. También se ha visto que la segunda componente se relaciona con el color o luminosidad del material, reafirmando la idea de que las personas emplean el color para distinguir materiales.

Los análisis con los resultados menos claros han sido los referentes al estudio de las activaciones internas de la red empleada para generar el espacio. Si bien, se ha visto que parte de las activaciones dependen del entorno en el que está el material y de su color.

Tras haber observado que el espacio latente tenía unas propiedades aparentemente buenas, se ha decidido crear un sistema que permita generar materiales a partir de puntos del espacio. Este experimento ha sido exitoso, las imágenes generadas por la red entrenada son muy precisas. El generador no solo ha aprendido a reproducir el conjunto de datos de entrenamiento, sino que ha aprendido la distribución del espacio de tal forma que es posible realizar interpolaciones realistas entre casi cualquier tipo de material. Además, se ha encontrado un posible método intuitivo de edición de materiales a partir de las observaciones sobre el espacio reducido con PCA.

A nivel personal este trabajo me ha permitido entender mejor en que consiste el desarrollo de un proyecto como parte de un grupo de investigación. Supone un modo de trabajar muy diferente al que acostumbramos en las prácticas dirigidas por un guion. El trabajo de investigación me ha permitido también descubrir nuevas áreas de estudio y trabajos con aplicaciones muy interesantes. Todo esto me ha ayudado a tener más claro que hacer al año que viene, seguir estudiando en la rama de computación haciendo

un máster en la Universidad de Zaragoza, y continuar investigando en el grupo de *Graphics and Imaging Lab*.



## Bibliografía

- Celebi, M. E. and Aslandogan, Y. A. (2004) 'Content-based image retrieval incorporating models of human perception', in *International Conference on Information Technology: Coding Computing, ITCC*. doi: 10.1109/ITCC.2004.1286639.
- Chu, B., Yang, D. and Tadinada, R. (2017) 'Visualizing Residual Networks', *arXiv preprint arXiv:1701.02362*. Available at: <http://arxiv.org/abs/1701.02362>.
- Donahue, C. *et al.* (2018) 'Semantically decomposing the latent spaces of generative adversarial networks', *6th International Conference on Learning Representations, ICLR 2018 - Conference Track Proceedings*, pp. 1–19.
- Fleming, R. W. (2014) 'Visual perception of materials and their properties', *Vision Research*. Elsevier Ltd, 94(June), pp. 62–75. doi: 10.1016/j.visres.2013.11.004.
- He, K. *et al.* (2016) 'Deep residual learning for image recognition', in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 770–778. doi: 10.1109/CVPR.2016.90.
- Jiang, Z. H. *et al.* (2019) 'Disentangled representation learning for 3D face shape', *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2019-June, pp. 11949–11958. doi: 10.1109/CVPR.2019.01223.
- Kingma, D. P. and Ba, J. L. (2015) 'Adam: A method for stochastic optimization', *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, pp. 1–15.
- Lagunas, M. *et al.* (2019) 'A similarity measure for material appearance', *ACM Transactions on Graphics*, 38(4), pp. 1–12. doi: 10.1145/3306346.3323036.
- Matusik, W. *et al.* (2003) 'A data-driven reflectance model', *ACM SIGGRAPH 2003 Papers, SIGGRAPH '03*, 22, pp. 759–769. doi: 10.1145/1201775.882343.
- Radford, A., Metz, L. and Chintala, S. (2016) 'Unsupervised representation learning with deep convolutional generative adversarial networks', *4th International Conference on Learning Representations, ICLR 2016 - Conference Track Proceedings*, pp. 1–16.
- Serrano, A. *et al.* (2016) 'An intuitive control space for material appearance', *ACM Transactions on Graphics*, 35(6). doi: 10.1145/2980179.2980242.
- Shen, Y. *et al.* (2019) 'Interpreting the Latent Space of GANs for Semantic Face Editing', *arXiv preprint arXiv:1907.10786*. Available at: <http://arxiv.org/abs/1907.10786>.
- Springenberg, J. T. *et al.* (2015) 'Striving for simplicity: The all convolutional net', *3rd International Conference on Learning Representations, ICLR 2015 - Workshop Track Proceedings*, pp. 1–14.
- Yosinski, J. *et al.* (2015) 'Understanding Neural Networks Through Deep Visualization', *arXiv preprint arXiv:1506.06579*. Available at: <http://arxiv.org/abs/1506.06579>.
- Zhang, R. *et al.* (2018) 'The Unreasonable Effectiveness of Deep Features as a Perceptual Metric', in *Proceedings of the IEEE Computer Society Conference on*

*Computer Vision and Pattern Recognition*, pp. 586–595. doi:  
10.1109/CVPR.2018.00068.

## Anexo 1: Clasificación semántica

En este anexo se van a exponer las tablas con los resultados de los cálculos de diferentes hiperplanos de separación de materiales, que se han comentado en el subapartado *Materiales con diversas formas e iluminaciones*. En cada tabla aparecen los resultados de entrenar el clasificador para un entorno en concreto y emplear las muestras de los otros cuatro entornos para comprobar los resultados. Con cada entorno se obtienen dos hiperplanos que corresponden con las dos pruebas hechas a cada mapa de iluminación.

### Metal

train	test	28	170	182	2	43	183	21	14	273	12	25	steel
todos	stpeter	-48,061	-24,84	-9,3962	-13,285	-5,9455	-2,823	-2,405	-0,455	-0,049	2,9492	6,51	9,0705
todos	stpeter	-18,495	-8,1702	-0,7423	-3,7949	-1,1198	0,4606	1,2753	2,355	2,0327	3,1554	5,2896	6,726
todos	doge	-10,511	-5,1627	-1,9268	-4,8214	-2,7535	-1,184	-0,403	1,2802	0,9447	0,3215	1,6548	4,2258
todos	doge	-9,4974	-4,1774	-1,2964	-3,8287	-2,4112	-0,649	-0,171	1,6493	1,4385	0,845	2,2931	4,6287
todos	ennis	-24,902	-13,772	-6,9288	-10,967	-9,4276	-4,752	-3,438	-2,261	-0,248	-1,294	4,1775	7,0116
todos	ennis	-22,77	-13,866	-9,8951	-12,384	-11,004	-7,717	-6,546	-5,803	-4,59	-4,291	-2,116	0,8475
todos	grace	-10,467	-2,718	0,05955	-1,072	-1,0178	-0,165	0,3526	0,1722	0,7043	0,978	1,1281	4,2712
todos	grace	-19,055	-8,1196	-3,1277	-3,6282	-2,9584	-2,084	-2,163	-1,687	-0,891	-0,61	-1,105	3,4863
doge	stpeter	-39,931	-15,059	-1,068	-6,646	-0,017	3,178	4,638	7,179	6,334	10,524	15,129	18,503
doge	stpeter	-59,980	-31,755	-12,267	-15,950	-6,852	-3,133	-2,924	-0,165	0,202	4,007	8,269	11,861
doge	doge	-48,897	-28,317	-12,351	-24,038	-10,805	-4,655	-5,682	2,014	3,003	4,227	7,418	18,331
doge	doge	-47,842	-27,168	-12,038	-23,156	-10,304	-3,969	-5,612	2,724	3,715	5,354	8,534	19,457
doge	ennis	-25,841	-11,034	-3,032	-7,239	-5,379	0,618	0,937	2,603	5,332	3,568	12,310	14,079
doge	ennis	-25,841	-11,034	-3,032	-7,239	-5,379	0,618	0,937	2,603	5,332	3,568	12,310	14,079
doge	grace	-33,396	-8,762	-0,318	-3,257	-0,658	2,735	1,120	3,867	4,333	6,230	5,843	17,415
doge	grace	-29,637	-6,409	0,842	-1,876	0,775	4,335	2,090	5,397	5,517	7,622	7,418	18,830
ennis	stpeter	-56,976	-27,465	-8,595	-13,440	-4,647	-0,835	-0,098	2,733	2,613	6,945	11,740	15,554
ennis	stpeter	-52,952	-25,578	-8,705	-12,613	-4,159	-0,549	-0,256	2,303	2,252	6,486	10,702	14,073
ennis	doge	-45,530	-26,507	-12,096	-22,648	-9,820	-3,887	-5,889	2,038	3,153	5,423	8,234	18,113
ennis	doge	-45,381	-27,123	-12,509	-23,080	-10,124	-4,332	-6,230	1,318	2,402	4,706	7,375	16,979
ennis	ennis	-29,362	-14,220	-7,756	-10,365	-7,902	-2,342	-2,726	-1,060	1,668	1,169	8,081	10,331
ennis	ennis	-34,228	-18,793	-10,972	-15,354	-12,921	-6,959	-5,683	-4,302	-1,767	-2,134	4,110	7,795
ennis	grace	-38,025	-14,820	-4,770	-6,406	-4,015	-1,845	-2,639	-0,786	0,349	1,482	0,724	11,140
ennis	grace	-37,754	-13,614	-3,815	-5,794	-3,236	-0,612	-1,788	0,533	1,430	2,852	2,233	13,347
grace	stpeter	-50,444	-25,179	-10,178	-12,718	-4,798	-1,417	-1,807	0,809	0,961	4,901	8,568	11,500
grace	stpeter	-38,660	-15,628	-3,761	-7,735	-1,308	1,835	2,217	4,337	3,966	7,953	11,771	14,427
grace	doge	-30,485	-12,829	-4,816	-11,229	-4,121	1,495	-1,126	7,256	8,182	9,823	12,711	21,282
grace	doge	-42,878	-25,934	-12,166	-22,111	-9,808	-4,606	-6,381	0,412	1,245	3,421	5,663	14,681
grace	ennis	-16,037	-4,948	1,518	-1,250	-0,061	4,511	3,941	5,688	8,078	5,763	13,867	14,220
grace	ennis	-28,546	-14,108	-8,419	-10,491	-8,132	-2,843	-3,266	-1,744	0,870	0,595	6,721	9,020
grace	grace	-23,648	-3,575	1,906	-0,511	1,496	5,025	2,762	5,715	5,659	7,729	7,683	17,828
grace	grace	-26,851	-5,562	0,861	-1,491	0,665	4,136	1,961	4,892	5,018	7,046	6,889	17,453
stpeter	stpeter	-15,739	-8,942	-4,342	-5,674	-3,225	-2,031	-1,917	-1,407	-0,831	0,149	1,292	2,208
stpeter	stpeter	-8,616	-2,304	0,521	-2,085	-0,513	0,444	1,014	1,801	1,778	2,817	4,272	4,911
stpeter	doge	-6,348	-4,593	-2,587	-4,099	-1,827	-1,912	-2,176	-0,969	-0,595	-0,299	0,069	1,623
stpeter	doge	-39,583	-21,218	-9,112	-18,115	-7,623	-1,605	-3,576	4,143	5,126	6,861	9,820	19,724
stpeter	ennis	-4,651	-2,676	-0,301	-2,743	-2,337	-1,204	0,350	-0,130	0,173	-0,356	1,194	2,100
stpeter	ennis	-6,936	-5,861	-3,538	-6,085	-5,293	-4,501	-2,461	-3,368	-3,162	-3,159	-2,905	-1,173
stpeter	grace	-6,865	-3,048	-0,756	-1,781	-1,590	-1,213	-0,331	-0,824	-0,200	-0,411	-0,488	1,283
stpeter	grace	-9,161	-5,812	-3,382	-3,242	-2,651	-2,674	-2,720	-2,451	-1,827	-1,769	-2,038	-0,766

Tabla 9. Distancia al hiperplano de materiales metálicos.

Siguiendo el gradiente de color, para visualizar los valores de forma más cómoda, se puede ver un claro comportamiento estable independientemente del entorno a clasificar o el entorno base empleado para generar el hiperplano. Como ya se ha comentado, el hecho de que los valores numéricos varíen de un caso a otro no representa un problema, siempre y cuando en los diferentes casos las proporciones de distancias sean similares.

### Brillo

train	test	28	170	182	2	43	183	21	14	273	12	25	steel
todos	stpeter	-6,443	-0,467	0,990	0,213	1,871	2,430	2,421	2,805	2,547	4,006	4,650	5,231
todos	stpeter	-36,026	-13,068	-1,915	-6,323	-0,117	2,894	3,462	5,506	4,909	9,053	12,800	15,360
todos	doge	-8,264	-4,058	-2,392	-3,756	-0,264	1,172	-1,134	1,914	2,785	4,989	5,172	7,531
todos	doge	-32,187	-19,903	-9,927	-17,111	-6,689	-2,998	-5,333	0,502	1,489	4,079	5,488	12,404
todos	ennis	-2,245	1,535	0,994	3,470	4,209	4,896	2,714	3,230	4,123	4,309	6,107	4,867
todos	ennis	-2,941	0,996	0,718	3,000	3,752	4,559	2,517	3,018	3,931	4,095	5,963	4,821
todos	grace	-5,045	-0,632	0,866	0,649	2,186	3,294	1,317	3,386	2,816	3,835	3,580	6,488
todos	grace	-11,509	3,063	5,078	2,974	5,047	8,405	5,227	8,635	7,714	10,101	10,145	17,966
doge	stpeter	-40,207	-14,725	-1,330	-6,873	-0,032	3,280	4,294	6,635	5,835	10,308	14,664	17,764
doge	stpeter	-34,044	-11,128	1,891	-4,232	1,119	3,922	5,904	8,131	7,000	10,751	15,252	18,439
doge	doge	-36,432	-17,285	-7,472	-15,311	-5,316	0,328	-2,517	6,358	7,495	9,720	12,408	21,689
doge	doge	-50,448	-30,541	-13,677	-25,918	-12,117	-5,937	-6,791	0,307	1,092	2,415	5,343	15,824
doge	ennis	-15,417	-4,599	2,787	-0,674	0,267	5,021	4,740	6,273	8,782	6,240	15,065	15,194
doge	ennis	-15,417	-4,599	2,787	-0,674	0,267	5,021	4,740	6,273	8,782	6,240	15,065	15,194
doge	grace	-38,458	-13,792	-3,752	-5,774	-3,356	-0,688	-1,613	0,485	1,449	2,790	2,227	13,302
doge	grace	-25,308	-3,320	2,448	-0,491	1,668	5,285	3,504	6,377	6,279	8,234	8,311	18,983
ennis	stpeter	-47,540	-22,646	-6,682	-10,787	-3,216	-0,182	0,564	2,710	2,483	6,084	9,945	13,114
ennis	stpeter	-47,540	-22,646	-6,682	-10,787	-3,216	-0,182	0,564	2,710	2,483	6,084	9,945	13,114
ennis	doge	-41,520	-23,873	-10,625	-20,359	-8,736	-3,041	-4,415	2,274	3,062	4,515	6,873	15,940
ennis	doge	-41,520	-23,873	-10,625	-20,359	-8,736	-3,041	-4,415	2,274	3,062	4,515	6,873	15,940
ennis	ennis	-31,698	-16,276	-8,612	-12,456	-10,158	-4,097	-3,406	-2,042	0,648	-0,080	7,014	9,946
ennis	ennis	-25,901	-12,556	-5,994	-9,026	-7,016	-1,796	-1,428	-0,307	2,070	1,328	7,955	10,087
ennis	grace	-37,217	-13,549	-3,839	-5,720	-3,155	-0,652	-1,638	0,482	1,320	2,647	1,983	12,582
ennis	grace	-31,978	-10,827	-2,482	-4,336	-1,904	0,604	-0,515	1,608	2,173	3,484	2,943	12,561
grace	stpeter	-25,305	-5,336	3,846	-1,847	2,775	5,486	6,717	8,484	7,405	11,362	15,273	17,643
grace	stpeter	-36,816	-13,087	-1,120	-6,217	0,181	3,389	4,110	6,231	5,526	9,917	14,058	16,831
grace	doge	-35,360	-18,334	-8,290	-15,950	-5,598	-0,666	-3,418	4,525	5,678	8,177	10,355	18,981
grace	doge	-34,547	-16,798	-7,423	-14,821	-5,023	0,178	-2,786	5,556	6,696	9,178	11,485	20,219
grace	ennis	-10,749	-0,398	3,738	3,453	4,624	8,170	6,116	7,704	9,980	8,207	15,782	14,883
grace	ennis	-8,250	1,196	5,111	4,940	5,916	9,166	7,024	8,576	10,752	8,814	16,378	15,149
grace	grace	-32,456	-9,802	-1,717	-3,584	-1,011	1,777	-0,035	2,816	3,179	4,815	4,227	14,781
grace	grace	-19,071	-0,555	3,497	1,174	3,545	7,124	4,063	7,552	7,005	9,424	9,339	18,818
stpeter	stpeter	-40,400	-17,531	-5,158	-8,552	-1,747	1,214	1,380	3,325	3,080	7,027	10,398	12,966
stpeter	stpeter	-55,478	-26,227	-7,941	-12,701	-4,102	-0,346	0,458	3,169	2,964	7,310	11,972	15,621
stpeter	doge	-49,349	-31,400	-14,478	-26,543	-12,596	-6,894	-7,653	-1,417	-0,679	0,763	3,322	13,114
stpeter	doge	-49,152	-31,093	-14,352	-26,304	-12,381	-6,651	-7,557	-1,157	-0,383	1,157	3,722	13,544
stpeter	ennis	-32,398	-17,062	-9,227	-13,140	-10,793	-4,738	-4,066	-2,591	0,168	-0,565	6,396	9,454
stpeter	ennis	-32,869	-18,959	-13,407	-16,068	-13,537	-8,483	-7,793	-6,683	-4,529	-3,805	0,076	3,727
stpeter	grace	-25,804	-7,718	-1,395	-2,627	0,069	2,505	0,101	3,003	3,012	4,918	4,283	12,963
stpeter	grace	-25,555	-6,947	-0,766	-2,185	0,531	3,189	0,667	3,702	3,618	5,645	5,062	14,087

Tabla 10. Distancia al hiperplano de materiales brillantes.

Al igual que el caso anterior los resultados continúan siendo consistentes entre pruebas y entornos de iluminación. Lo más importante de esta prueba es que hay una clara

correlación entre lo brillante que es un material y lo metálico que es. Se puede ver que la clasificación de las muestras es muy similar en ambos casos.

### Mate

train	test	28	170	182	2	43	183	21	14	273	12	25	steel
todos	stpeter	26,15	8,2844	-1,052	3,5526	-1,033	-3,259	-4,456	-5,78	-5,284	-8,3552	-11,528	-13,42
todos	stpeter	26,15	8,2844	-1,052	3,5526	-1,033	-3,259	-4,456	-5,78	-5,284	-8,3552	-11,528	-13,42
todos	doge	25,26	12,014	5,5711	10,744	3,1353	-0,874	2,075	-4,895	-5,838	-8,2532	-9,9495	-16,628
todos	doge	25,26	12,014	5,5711	10,744	3,1353	-0,874	2,075	-4,895	-5,838	-8,2532	-9,9495	-16,628
todos	ennis	8,7846	0,7011	-1,801	-1,208	-1,936	-4,616	-3,721	-4,377	-5,684	-4,9034	-9,1992	-8,8121
todos	ennis	25,63	13,757	9,4603	10,826	8,6174	4,2982	4,5663	3,4883	1,3192	0,89575	-3,3849	-5,6855
todos	grace	19,535	2,4706	-2,453	-0,118	-1,739	-4,473	-3,206	-5,078	-5,125	-6,5114	-6,2596	-14,273
todos	grace	11,768	2,3729	-0,005	0,1213	-1,335	-2,695	-0,589	-2,914	-2,662	-3,9157	-3,5298	-8,126
doge	stpeter	52,366	23,279	5,868	11,113	2,769	-0,914	-1,783	-4,545	-4,104	-8,627	-13,330	-16,932
doge	stpeter	40,207	14,725	1,330	6,873	0,032	-3,280	-4,294	-6,635	-5,835	-10,308	-14,664	-17,764
doge	doge	49,508	31,859	14,558	26,815	12,714	7,155	7,668	1,564	0,871	-0,469	-3,028	-12,852
doge	doge	48,807	27,152	11,701	23,240	10,209	3,735	5,056	-3,216	-4,202	-5,532	-8,807	-19,951
doge	ennis	27,764	12,421	4,848	8,304	6,255	0,316	0,435	-1,332	-4,201	-2,906	-11,256	-13,138
doge	ennis	35,200	20,442	12,717	17,140	14,738	8,869	7,358	6,065	3,673	3,757	-1,654	-5,709
doge	grace	37,593	15,555	5,448	6,729	4,299	2,273	3,167	1,305	0,242	-0,926	-0,128	-9,982
doge	grace	35,121	10,343	1,636	3,854	1,122	-1,965	-0,191	-3,056	-3,545	-5,422	-4,918	-16,454
ennis	stpeter	54,567	25,547	7,759	12,522	3,892	0,058	-0,473	-3,088	-2,891	-7,330	-11,897	-15,389
ennis	stpeter	47,509	20,950	6,141	10,337	2,644	-0,798	-1,056	-3,472	-3,210	-7,574	-11,602	-14,550
ennis	doge	33,849	20,659	10,200	17,802	6,980	3,203	5,467	-0,590	-1,594	-4,159	-5,629	-12,818
ennis	doge	35,402	21,926	10,782	18,833	7,581	3,685	5,823	-0,184	-1,179	-3,697	-5,216	-12,598
ennis	ennis	22,821	8,928	4,609	5,478	3,302	-1,266	0,121	-1,435	-3,870	-3,492	-9,873	-10,970
ennis	ennis	29,896	14,048	6,902	10,026	7,700	1,734	1,834	0,246	-2,574	-1,701	-9,291	-11,624
ennis	grace	30,693	11,428	3,679	4,523	2,040	0,076	1,851	-0,579	-1,110	-2,623	-1,761	-10,519
ennis	grace	37,308	14,248	4,424	5,958	3,382	1,016	2,255	0,007	-0,882	-2,306	-1,526	-12,058
grace	stpeter	36,737	13,131	0,912	6,024	-0,324	-3,512	-4,361	-6,499	-5,748	-10,042	-14,261	-17,144
grace	stpeter	43,711	18,997	5,075	9,083	1,881	-1,356	-1,795	-4,038	-3,713	-7,824	-11,844	-14,861
grace	doge	38,500	21,118	9,517	18,147	7,241	1,866	4,381	-3,283	-4,367	-6,881	-9,203	-17,942
grace	doge	38,500	21,118	9,517	18,147	7,241	1,866	4,381	-3,283	-4,367	-6,881	-9,203	-17,942
grace	ennis	25,014	10,614	4,192	6,308	4,062	-1,369	-0,293	-1,986	-4,787	-3,713	-11,355	-12,756
grace	ennis	25,013	10,613	4,191	6,307	4,062	-1,369	-0,294	-1,986	-4,787	-3,713	-11,355	-12,756
grace	grace	29,582	8,713	1,376	3,036	0,395	-2,401	-0,250	-3,169	-3,365	-5,228	-4,698	-14,649
grace	grace	28,801	9,355	2,302	3,490	0,916	-1,385	0,667	-2,084	-2,322	-4,005	-3,289	-12,359
stpeter	stpeter	25,580	9,106	-1,242	3,874	-0,186	-2,226	-3,915	-5,054	-4,660	-6,968	-10,274	-12,268
stpeter	stpeter	25,580	9,106	-1,242	3,874	-0,186	-2,226	-3,915	-5,054	-4,660	-6,968	-10,274	-12,268
stpeter	doge	22,964	9,256	2,708	8,371	3,125	-0,684	-0,234	-5,486	-6,025	-6,237	-8,516	-14,897
stpeter	doge	46,403	25,555	10,658	21,841	9,954	3,631	4,328	-3,305	-4,117	-4,952	-8,330	-18,852
stpeter	ennis	12,394	4,014	-1,861	1,722	1,248	-2,454	-3,304	-3,949	-5,473	-3,585	-9,445	-9,988
stpeter	ennis	22,904	9,992	5,862	6,587	4,256	0,015	1,341	-0,040	-2,393	-2,392	-7,932	-9,214
stpeter	grace	35,576	10,488	1,585	4,253	1,724	-1,328	-0,180	-2,547	-3,174	-4,785	-4,305	-15,826
stpeter	grace	36,295	16,405	6,002	7,422	5,481	4,004	3,710	2,809	1,461	1,046	1,793	-6,873

Tabla 11. Distancia al hiperplano de materiales mate.

Para el atributo de nivel de mate el resultado es muy similar al obtenido para el atributo de brillo, pero cambiando los signos de las distancias. Esto es coherente con el hecho de que son dos atributos casi contrarios, por lo que el hiperplano debería estar en una posición parecida, pero cambiando el sentido del vector normal que lo define.



## Definición de los reflejos

train	test	28	170	182	2	43	183	21	14	273	12	25	steel
todos	stpeter	-4,7748	-1,8644	-1,1315	-1,2131	-0,7433	-0,088	-0,71	0,262	0,3739	0,9778	1,5186	1,9199
todos	stpeter	-5,0945	-2,088	-1,2372	-1,3236	-0,8468	-0,321	-0,813	0,0985	0,2744	0,7226	1,2543	1,4365
todos	doge	-2,4483	-1,8842	-1,1393	-1,7405	-1,3909	-0,798	-1,015	0,0622	-0,218	-0,226	0,0566	1,1485
todos	doge	-6,0661	-5,2839	-2,9663	-4,3823	-2,7833	-2,609	-2,496	-1,478	-1,813	-2,251	-1,601	0,2334
todos	ennis	-2,6429	-1,6069	-2,0409	-2,1043	-2,1547	-1,901	-2,382	-1,206	-1,165	-0,975	-0,472	-0,102
todos	ennis	-2,6182	-1,5255	-1,7229	-1,8204	-2,0254	-1,71	-1,993	-0,857	-0,732	-0,657	0,0537	0,2608
todos	grace	-2,4286	-1,3677	-0,8172	-0,2692	-0,7267	-0,869	-1,033	-0,519	-0,313	-0,593	-0,591	-0,196
todos	grace	-1,6069	-0,8543	-0,5203	0,01366	-0,3876	-0,356	-0,761	-0,095	0,1629	0,0086	0,0822	0,5691
doge	stpeter	-40,207	-14,725	-1,330	-6,873	-0,032	3,280	4,294	6,635	5,835	10,308	14,664	17,764
doge	stpeter	-40,207	-14,725	-1,330	-6,873	-0,032	3,280	4,294	6,635	5,835	10,308	14,664	17,764
doge	doge	-36,432	-17,285	-7,472	-15,311	-5,316	0,328	-2,517	6,358	7,495	9,720	12,408	21,689
doge	doge	-36,432	-17,285	-7,472	-15,311	-5,316	0,328	-2,517	6,358	7,495	9,720	12,408	21,689
doge	ennis	-35,073	-19,883	-11,605	-16,362	-14,004	-7,872	-6,337	-4,973	-2,446	-2,857	3,368	7,308
doge	ennis	-32,526	-16,977	-9,862	-13,237	-10,759	-4,834	-4,516	-2,941	-0,257	-0,610	5,969	8,941
doge	grace	-28,166	-5,748	0,740	-1,533	1,227	4,650	2,084	5,466	5,372	7,677	7,325	18,212
doge	grace	-28,166	-5,748	0,740	-1,533	1,227	4,650	2,084	5,466	5,372	7,677	7,325	18,212
ennis	stpeter	-56,598	-27,440	-8,780	-13,383	-4,450	-0,669	-0,101	2,633	2,541	6,912	11,546	15,252
ennis	stpeter	-26,731	-8,560	-0,723	-4,125	0,858	3,249	3,459	5,056	4,489	8,078	10,973	12,869
ennis	doge	-25,049	-12,490	-6,174	-11,159	-3,197	0,500	-2,588	4,180	5,238	7,979	9,435	15,813
ennis	doge	-27,360	-13,967	-6,804	-12,385	-3,842	0,137	-2,902	4,099	5,178	7,902	9,518	16,355
ennis	ennis	-22,497	-8,955	-5,101	-5,686	-3,492	0,846	-0,591	0,911	3,231	3,083	8,936	10,056
ennis	ennis	-22,416	-9,330	-6,017	-6,276	-4,022	0,071	-1,396	0,012	2,220	2,377	7,546	8,780
ennis	grace	-37,108	-13,023	-3,342	-5,280	-2,593	0,193	-1,247	1,255	2,023	3,668	3,036	14,277
ennis	grace	-20,856	-4,517	0,259	-0,983	1,482	4,043	1,412	4,464	4,206	6,242	5,771	13,943
grace	stpeter	-57,254	-27,947	-8,371	-13,458	-4,657	-0,819	0,130	2,931	2,793	7,041	11,961	15,895
grace	stpeter	-36,965	-16,100	-5,435	-8,191	-1,909	0,907	0,670	2,477	2,342	6,064	9,072	11,221
grace	doge	-48,323	-30,477	-14,089	-25,782	-11,871	-6,489	-7,438	-1,002	-0,151	1,447	3,873	13,763
grace	doge	-38,426	-21,534	-9,971	-18,594	-7,310	-2,252	-4,732	2,752	3,819	6,299	8,396	17,169
grace	ennis	-23,100	-9,009	-3,099	-4,862	-2,670	2,472	1,125	2,865	5,640	4,531	12,236	13,299
grace	ennis	-33,352	-17,825	-10,875	-14,208	-11,667	-5,779	-5,273	-3,777	-1,088	-1,298	4,825	8,086
grace	grace	-37,947	-16,063	-5,727	-7,194	-5,094	-3,249	-3,502	-2,101	-0,771	0,000	-0,740	8,938
grace	grace	-30,267	-6,854	0,367	-2,018	0,670	4,191	1,772	5,053	5,167	7,361	6,994	18,369
stpeter	stpeter	-53,302	-23,832	-5,673	-11,403	-2,982	0,929	2,040	4,795	4,356	8,928	13,967	17,821
stpeter	stpeter	-53,830	-25,606	-6,865	-12,416	-4,193	-0,646	0,813	3,266	3,109	7,019	11,845	15,591
stpeter	doge	-35,548	-20,365	-9,685	-17,623	-6,785	-2,180	-4,682	2,331	3,366	5,911	7,859	15,766
stpeter	doge	-45,559	-31,647	-14,521	-26,385	-13,481	-8,891	-8,096	-4,382	-3,976	-3,376	-1,176	7,138
stpeter	ennis	-31,733	-17,393	-8,270	-13,743	-11,704	-5,590	-3,818	-2,607	-0,019	-1,241	5,937	9,288
stpeter	ennis	-22,904	-9,992	-5,862	-6,587	-4,256	-0,015	-1,341	0,040	2,393	2,392	7,932	9,214
stpeter	grace	-36,005	-10,548	-1,477	-3,995	-1,178	2,243	0,432	3,307	3,878	5,877	5,437	17,397
stpeter	grace	-35,738	-10,612	-1,594	-4,263	-1,678	1,372	0,206	2,605	3,220	4,827	4,322	15,879

Tabla 12. Distancia al hiperplano de materiales con los reflejos definidos.

Al igual que propiedades como metalicidad y brillo, este atributo está muy bien separado, es muy probable que todos estos atributos estén bien separados porque están muy correlacionados, y si uno de ellos está claramente distribuido el resto debe estarlo también.

## Áspero

train	test	28	170	182	2	43	183	21	14	273	12	25	steel
todos	stpeter	2,702	1,242	1,612	0,897	1,235	0,455	1,218	0,649	0,713	0,472	0,357	0,168
todos	stpeter	2,757	1,475	1,834	1,080	1,317	0,703	1,270	0,831	0,794	0,689	0,546	0,335
todos	doge	3,775	1,600	1,436	1,560	1,286	-0,315	1,112	0,313	0,362	-0,462	-0,416	-0,808
todos	doge	3,172	1,560	1,389	1,433	1,484	0,097	1,155	0,564	0,611	-0,179	-0,185	-0,312
todos	ennis	1,498	1,244	1,441	0,672	0,914	0,311	0,858	0,608	0,236	0,581	0,240	-0,137
todos	ennis	1,520	1,203	1,751	0,724	0,785	0,081	1,054	0,547	0,082	0,168	0,053	-0,103
todos	grace	2,981	1,397	1,338	0,893	1,065	0,735	0,990	0,394	0,363	0,601	0,439	0,024
todos	grace	3,130	1,389	1,341	0,919	1,016	0,730	1,005	0,409	0,372	0,623	0,460	-0,017
doge	stpeter	4,586	1,609	1,556	0,102	0,149	-0,570	0,185	-1,127	-0,821	-1,546	-2,330	-2,896
doge	stpeter	6,799	3,365	2,387	1,416	1,059	0,051	0,997	-0,287	-0,141	-0,680	-1,501	-2,007
doge	doge	7,548	3,606	1,883	3,258	2,916	0,756	1,407	-0,089	0,194	0,106	-0,713	-2,097
doge	doge	7,059	2,989	1,848	2,832	2,374	0,118	1,441	-0,779	-0,597	-1,219	-2,003	-3,427
doge	ennis	25,103	13,322	10,525	10,889	8,702	4,941	5,893	4,468	2,640	1,681	-1,501	-3,784
doge	ennis	5,091	3,627	0,851	2,109	2,094	0,779	0,782	-0,025	-0,933	0,166	-2,129	-2,466
doge	grace	18,849	9,819	4,675	4,411	3,597	3,302	3,105	2,568	1,866	1,912	2,223	-1,576
doge	grace	6,004	2,159	1,175	0,853	1,019	0,690	0,736	0,397	0,066	0,040	-0,205	-1,691
ennis	stpeter	0,929	-5,028	-5,442	-3,008	-3,344	-4,134	-4,624	-5,115	-4,149	-5,570	-6,634	-7,251
ennis	stpeter	0,177	-1,044	-0,424	-0,597	-0,623	-0,739	-0,643	-0,582	-0,350	-0,555	-0,441	-0,444
ennis	doge	1,022	0,275	0,690	0,060	-0,689	-1,400	0,273	-0,277	-0,593	-1,764	-1,161	-1,027
ennis	doge	1,850	0,282	1,117	0,498	-0,625	-1,705	0,358	-1,353	-1,699	-3,339	-3,149	-3,390
ennis	ennis	-0,395	-0,898	0,002	-1,598	-1,839	-2,141	-1,093	-0,911	-1,053	-1,373	-1,479	-0,942
ennis	ennis	-0,427	-0,950	0,051	-1,614	-1,775	-2,079	-1,016	-0,880	-0,996	-1,309	-1,389	-0,853
ennis	grace	15,384	3,866	0,270	0,871	-0,893	-3,020	-0,860	-3,348	-3,153	-4,668	-4,539	-10,391
ennis	grace	15,384	3,866	0,270	0,871	-0,893	-3,020	-0,860	-3,348	-3,153	-4,668	-4,539	-10,391
grace	stpeter	4,141	2,296	1,580	0,881	0,522	0,359	0,619	0,416	0,440	0,410	0,524	0,643
grace	stpeter	3,756	1,671	1,606	0,898	0,366	0,198	0,301	0,158	-0,091	-0,330	-0,481	-0,523
grace	doge	3,015	2,443	1,779	2,328	1,489	1,044	1,452	1,478	1,518	0,889	1,150	0,914
grace	doge	3,973	2,243	1,587	2,305	1,288	0,459	0,965	0,883	0,860	0,194	0,518	0,186
grace	ennis	3,452	1,796	2,937	0,821	0,290	0,114	1,369	1,238	0,548	0,109	0,620	0,511
grace	ennis	17,643	9,608	9,340	8,086	6,396	3,660	5,365	4,400	3,122	1,790	0,716	-0,904
grace	grace	6,077	4,932	3,418	2,032	1,680	2,359	2,642	2,321	2,085	2,336	2,887	2,980
grace	grace	3,869	2,049	1,405	0,464	-0,005	0,375	1,151	0,362	0,377	0,436	0,866	0,437
stpeter	stpeter	4,200	2,738	2,094	1,477	1,609	0,935	1,120	0,941	0,811	0,524	0,277	0,000
stpeter	stpeter	5,289	2,293	1,704	0,810	0,646	-0,223	0,402	-0,410	-0,211	-0,955	-1,388	-1,589
stpeter	doge	2,384	1,832	1,504	1,648	1,957	0,796	1,488	0,813	1,172	0,602	0,213	0,030
stpeter	doge	3,148	1,583	0,979	1,539	1,523	0,358	1,102	-0,343	0,030	-0,188	-0,817	-1,533
stpeter	ennis	0,968	0,996	0,799	0,521	0,505	0,261	0,567	0,322	-0,067	0,444	-0,077	-0,295
stpeter	ennis	1,803	0,656	2,065	-0,116	-0,422	-0,486	1,271	0,365	-0,089	-0,323	-0,588	-0,338
stpeter	grace	3,561	2,398	1,490	0,836	0,774	0,443	1,229	0,374	0,359	0,431	0,440	-0,253
stpeter	grace	3,044	1,927	1,456	0,864	1,112	0,444	1,018	0,511	0,455	0,348	0,206	-0,127

Tabla 13. Distancia al hiperplano de materiales ásperos en esferas.

		28	170	182	2	43	183	21	14	273	12	25	steel
doge	stpeter	3,232	1,682	1,165	0,778	0,502	0,724	1,121	0,763	0,623	0,135	0,252	-0,037
doge	stpeter	3,194	1,756	1,299	0,918	0,490	0,810	1,205	1,032	0,957	0,271	0,536	0,297
doge	doge	1,894	1,582	0,612	0,287	0,274	0,556	0,668	0,534	0,532	0,319	0,379	0,168
doge	doge	1,694	1,552	0,829	0,365	0,322	0,724	0,891	0,766	0,821	0,457	0,632	0,477
doge	ennis	1,995	0,837	0,934	0,092	0,414	0,288	0,637	0,657	0,647	0,209	0,151	-0,014
doge	ennis	2,090	1,035	1,025	0,139	0,321	0,133	0,817	0,851	0,903	0,250	0,387	0,262
doge	grace	2,271	1,684	1,352	1,087	0,519	-0,111	0,948	0,904	0,228	-0,272	-0,093	-0,177
doge	grace	2,100	1,789	1,728	1,313	0,421	0,137	1,166	1,212	0,540	-0,045	0,172	0,055
ennis	stpeter	7,809	5,100	3,706	3,822	0,895	1,637	2,616	2,585	2,233	-0,104	0,930	1,024
ennis	stpeter	10,780	7,381	5,083	5,772	2,505	2,707	3,922	4,059	3,671	1,011	2,628	2,582
ennis	doge	3,014	2,487	2,241	1,070	0,671	1,814	1,904	1,347	1,710	1,037	1,323	1,042
ennis	doge	4,626	4,227	3,840	2,377	2,103	3,149	3,634	3,025	3,546	2,280	2,682	2,503
ennis	ennis	5,143	3,114	3,311	1,245	0,975	-0,736	3,231	2,661	2,530	-0,112	0,540	0,582
ennis	ennis	7,078	4,793	5,825	3,529	3,894	1,672	5,505	5,191	4,918	1,859	2,772	2,061
ennis	grace	5,705	4,326	5,249	4,422	1,350	2,502	3,902	3,447	2,335	1,805	1,771	0,820
ennis	grace	7,879	6,047	6,647	5,851	3,000	3,656	5,442	5,747	4,666	3,090	3,189	2,040
grace	stpeter	0,031	-0,418	-0,284	-0,634	-0,652	-0,359	-0,396	-0,831	-0,803	-0,920	-1,658	-2,215
grace	stpeter	0,478	-0,414	-0,743	-1,288	-1,365	-1,022	-1,167	-2,003	-2,155	-1,592	-2,750	-3,714
grace	doge	0,318	-0,500	-0,842	-1,194	-1,025	-1,535	-0,877	-0,910	-1,265	-1,740	-2,151	-2,149
grace	doge	0,642	-0,352	-0,977	-1,605	-1,695	-2,037	-1,576	-2,172	-2,498	-2,232	-2,840	-3,300
grace	ennis	-0,227	-0,929	-1,163	-1,260	-1,354	-1,327	-1,205	-1,277	-1,188	-1,125	-1,941	-1,723
grace	ennis	0,257	-1,147	-1,572	-1,621	-1,977	-1,738	-1,829	-2,140	-2,167	-1,900	-3,066	-2,815
grace	grace	0,206	-0,246	-0,615	-0,491	-0,408	-0,881	-0,453	-1,415	-1,380	-0,747	-1,103	-1,060
grace	grace	0,054	-0,520	-1,052	-0,904	-0,731	-1,368	-1,077	-2,193	-2,800	-1,775	-2,336	-2,517
stpeter	stpeter	0,936	0,088	0,427	-1,362	-0,652	-0,877	0,367	-1,210	-2,050	-2,337	-3,007	-2,875
stpeter	stpeter	2,878	1,785	1,364	0,685	0,761	-0,221	1,083	-0,061	-0,871	-1,064	-1,167	-1,553
stpeter	doge	0,331	-0,230	-0,853	-0,977	-1,004	-1,083	-1,395	-2,101	-2,328	-1,454	-2,291	-2,540
stpeter	doge	1,458	0,998	0,137	0,053	0,054	-0,297	-0,276	-0,829	-0,826	-0,611	-0,982	-1,178
stpeter	ennis	0,656	-0,617	-1,078	-1,381	-1,726	-2,079	-1,042	-2,088	-2,117	-2,351	-2,817	-2,179
stpeter	ennis	1,857	0,397	0,357	0,155	0,205	-0,392	0,106	-0,562	-0,666	-0,670	-1,000	-0,885
stpeter	grace	0,515	-0,063	0,391	0,086	-0,278	-0,721	0,168	-1,408	-2,604	-1,126	-1,561	-2,307
stpeter	grace	2,049	0,936	1,227	1,113	0,855	0,082	0,863	0,010	-0,848	-0,302	-0,338	-1,167

Tabla 14. Distancia al hiperplano de materiales ásperos en blobs.

Para este atributo se muestran los resultados de los materiales con las dos formas, y en ambos casos se puede apreciar que las distancias guardan más correlación con los entornos de iluminación de lo esperado. Además, los resultados obtenidos con las esferas varían claramente respecto de los de blob. Probablemente se deba al hecho de que detectar visualmente si un material es áspero es muy difícil, y por tanto no esté claramente valorado como para poder ser independiente de la forma al hacer la clasificación de los materiales.



## Luminosidad

train	test	28	170	182	2	43	183	21	14	273	12	25	steel
todos	stpeter	0,477	1,064	0,348	0,935	0,549	0,574	0,304	1,059	0,550	0,666	0,661	0,731
todos	stpeter	0,631	1,472	0,914	1,248	0,825	0,803	0,894	1,466	0,906	1,089	1,141	1,430
todos	doge	0,460	0,803	0,021	0,445	-0,252	0,298	-0,372	0,811	0,447	0,244	0,135	0,255
todos	doge	0,566	0,971	0,205	0,370	-0,019	0,229	-0,010	1,253	0,604	0,403	0,320	0,601
todos	ennis	0,408	0,648	0,182	0,772	0,643	0,343	-0,405	0,427	0,553	0,307	0,779	0,650
todos	ennis	0,468	0,835	0,405	0,894	0,593	0,580	-0,077	1,040	1,293	0,794	1,464	1,260
todos	grace	0,455	1,052	0,382	0,747	0,479	0,670	0,228	0,706	0,439	0,748	0,929	0,887
todos	grace	0,509	1,195	0,477	0,790	0,480	0,583	0,215	0,673	0,365	0,604	0,787	0,844
doge	stpeter	0,750	3,054	2,078	2,704	1,593	1,458	1,865	2,720	2,164	2,230	3,051	3,705
doge	stpeter	0,447	2,076	0,852	2,093	1,159	0,930	0,685	1,273	0,836	0,714	0,876	1,303
doge	doge	-0,835	2,132	0,875	1,279	0,262	1,602	0,523	3,002	2,936	2,804	3,610	5,155
doge	doge	-0,251	1,645	0,533	0,941	0,034	1,131	0,109	1,459	1,013	1,386	1,559	1,569
doge	ennis	1,611	1,517	1,782	2,040	1,544	1,548	0,744	1,781	1,849	1,384	2,267	1,695
doge	ennis	2,242	1,508	1,856	2,079	1,571	1,601	0,729	1,826	1,921	1,438	2,389	1,743
doge	grace	-0,085	0,799	0,098	0,821	0,558	0,660	0,212	0,541	0,239	0,440	0,672	0,847
doge	grace	0,289	0,592	0,196	0,582	0,140	0,334	0,212	0,456	0,305	0,548	0,941	0,985
ennis	stpeter	12,102	11,785	10,117	6,927	5,368	5,297	7,051	7,203	5,943	6,545	7,518	7,827
ennis	stpeter	8,555	7,718	7,002	4,642	3,244	3,123	4,468	4,405	3,523	3,606	4,014	4,119
ennis	doge	5,918	10,516	6,189	8,229	4,510	5,691	5,073	8,774	8,544	7,310	8,698	10,144
ennis	doge	10,277	14,357	7,799	11,233	6,517	6,889	5,906	8,680	8,600	7,785	8,718	9,190
ennis	ennis	7,198	5,738	7,109	6,138	5,347	5,063	4,953	5,338	5,059	3,934	5,572	4,152
ennis	ennis	10,609	9,345	10,599	9,867	9,044	8,547	8,051	8,621	8,762	6,813	10,008	7,669
ennis	grace	5,760	9,261	6,180	4,784	4,720	5,986	5,044	6,011	5,173	6,113	6,674	8,176
ennis	grace	2,547	4,398	3,314	2,128	2,177	3,040	2,762	3,136	2,784	3,128	3,576	4,271
grace	stpeter	0,546	-0,050	-0,335	0,391	-0,256	0,044	-0,429	0,008	-0,085	-0,293	-0,214	-0,412
grace	stpeter	-0,496	-0,614	-0,668	-0,156	-0,385	-0,148	-0,477	-0,060	-0,109	-0,411	-0,313	-0,386
grace	doge	-0,959	-0,809	-1,031	-0,946	-1,515	-0,846	-1,359	-0,607	-1,011	-0,837	-0,831	-0,861
grace	doge	-0,615	-0,210	-0,389	-0,552	-0,806	-0,430	-0,697	0,035	-0,218	-0,296	-0,309	-0,226
grace	ennis	0,732	1,285	0,284	1,183	0,872	0,704	-0,148	0,937	1,133	0,660	1,055	0,687
grace	ennis	-0,951	-0,396	-1,346	-0,718	-1,068	-1,030	-1,269	-0,358	-0,315	-0,585	-0,965	-0,675
grace	grace	-0,422	-0,528	-0,670	-0,063	-0,618	-0,463	-0,586	-0,380	-0,365	-0,495	-0,325	-0,704
grace	grace	1,262	0,516	0,108	0,670	0,247	0,309	0,070	0,245	0,167	0,178	0,310	0,299
stpeter	stpeter	3,398	3,204	1,417	2,001	1,320	1,346	1,132	1,666	1,360	1,588	1,516	1,283
stpeter	stpeter	8,835	7,744	4,215	4,048	2,899	2,696	3,068	3,164	2,518	2,718	2,629	2,289
stpeter	doge	5,283	5,861	2,483	4,698	2,449	3,414	1,471	3,459	3,483	4,549	4,575	4,527
stpeter	doge	6,157	7,198	2,873	5,651	3,174	3,957	1,643	4,242	4,302	5,415	5,726	5,720
stpeter	ennis	4,993	3,768	2,242	4,346	4,015	3,494	1,697	2,593	3,122	2,450	3,555	2,496
stpeter	ennis	4,702	3,714	1,968	4,344	4,207	3,621	1,636	2,529	3,029	2,482	3,240	2,244
stpeter	grace	1,510	0,954	0,082	0,822	0,337	0,116	-0,208	0,599	0,471	0,474	0,887	0,839
stpeter	grace	2,612	3,339	1,786	1,834	1,794	2,100	1,326	2,713	2,225	2,458	2,887	3,784

Tabla 15. Distancia al hiperplano de materiales luminosos.

Tal y como cabía esperar, la clasificación de este atributo (*brightness*) no es en absoluto constante de un caso a otro. *Saint Peters* es el entorno que consigue hacer que las distancias se establezcan más. Se puede apreciar una relación entre el color del material y los valores obtenidos, los que son más oscuros son menos luminosos. Aunque la mayoría de las distancias son muy pequeñas en comparación con las distancias en atributos más sencillos. Esto puede significar que a la hora de valorar la distancia entre materiales este atributo no tenga mucho peso en la decisión, haciendo que la forma del espacio no separe los materiales de una forma clara para esta propiedad.

## Anexo 2: Arquitectura del sistema generador de imágenes

En este anexo se van a describir detalladamente las características de las partes principales del sistema entrenado para generar imágenes, empleado en *Reconstrucción del espacio latente*. En especial, se van a detallar las funciones de coste empleadas para regular el aprendizaje del generador y del discriminador. Para este trabajo se ha entrenado tanto el discriminador como el generador para trabajar con imágenes en color con tres canales de dimensión 64x64.

La arquitectura de red empleada para el generador y el discriminador sigue el mismo modelo, se ha diseñado un DCGAN [Radford, Metz and Chintala, 2016] que se considera una de las mejores opciones para generación básica de imágenes. Se trata de dos redes convolucionales que constan de varias convoluciones 2D (tres convoluciones para el generador y cinco para el discriminador), con normalizaciones 2D por lotes y activaciones ReLU entre cada convolución (Figura 28). Se siguen las recomendaciones del trabajo original en cuanto al algoritmo de optimización a emplear. Se usa el optimizador Adam [Kingma and Ba, 2015] manteniendo la configuración por defecto de los hiperparámetros, exceptuando el ritmo de aprendizaje que se ha disminuido para el discriminador (de  $2e-4$  a  $1e-5$ ). En caso contrario se volvía muy bueno antes de que el generador pudiera terminar de aprender, afectando negativamente al rendimiento del sistema.

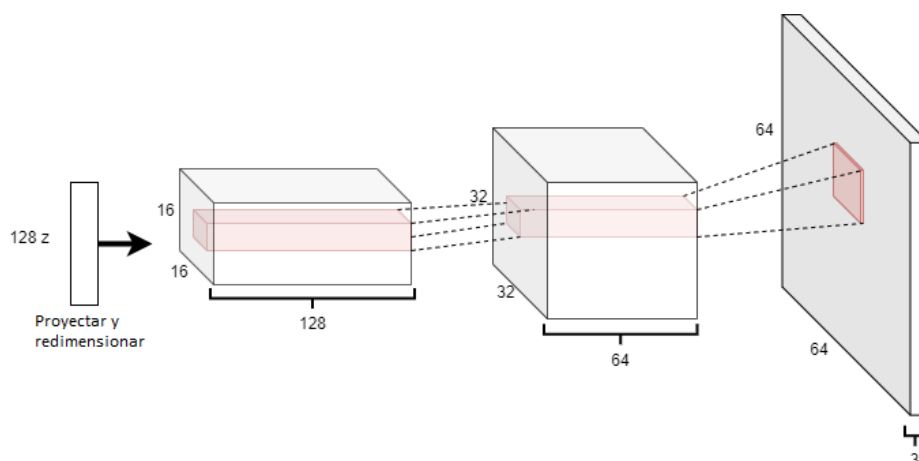


Figura 28. Arquitectura de la red convolucional del generador, con un tamaño de kernel (parte en naranja) de  $3 \times 3$ .

Las ecuaciones expuestas a continuación se han extraído de la documentación de *Pytorch*, que ha sido el framework empleado para implementar el sistema de redes convolucionales y las funciones de coste.

### Codificador

El codificador (*encoder*) es la parte del sistema encargada de convertir las imágenes de entrada en códigos latentes de 128 dimensiones. Se trata de la red descrita en *Conjuntos de datos y arquitectura de la red*, y que corresponde con la red entrenada en *Lagunas et al. (2019)*. Esta parte no tiene ningún tipo de función de coste o algoritmo de optimización, pues se trata del modelo de pesos exacto obtenido en el trabajo del que se ha tomado de referencia.

## Discriminador

El discriminador es la parte de la red que se encarga de evaluar las imágenes resultantes del generador, actúa a modo de clasificador entre imágenes reales y generadas. La salida de la red es la probabilidad de que la imagen de entrada sea real, pasando la probabilidad por una función de activación sigmoideal.

El objetivo de entrenar el discriminador es maximizar la probabilidad de que la clasificación de las imágenes sea correcta. Para ello la función de coste empleada es la función de Entropía Cruzada Binaria (BCE, *Binary Cross Entropy Loss*), Ecuación 1.

$$\ell(x, y) = \text{mean}(L) = \text{mean}(\{l_1, \dots, l_N\}_T) \quad (1)$$

$$l_n = - [y_n \cdot \log(x_n) + (1 - y_n) \cdot \log(1 - x_n)] \quad (2)$$

Para facilitar la tarea del discriminador y que pueda aprender mejor a clasificar las imágenes es recomendable no pasar por la red a la vez imágenes falsas y generadas. Es preferible hacer dos pasadas, una para las reales de entrenamiento y otra para las generadas. El discriminador va a tratar de maximizar la Ecuación 3.

$$\log(D(x)) + \log(1 - D(G(z))) \quad (3)$$

Dónde  $D(x)$  corresponde con la probabilidad de que una imagen ( $x$ ) sea real, y  $G(z)$  corresponde con una imagen falsa generada a partir de un código latente ( $z$ ).

## Generador

Para el generador (*decoder*) de imágenes la función de aprendizaje es algo más compleja que la propuesta para un DCGAN convencional, por la necesidad de que aprenda la distribución exacta del espacio latente del codificador empleado. En primer lugar, está la parte principal de un generador de este tipo de sistemas que aprende a partir de la salida del discriminador, empleando la misma función de coste (Ecuación 1), BCE, que el discriminador para aprender reproducir mejor las imágenes. Para esta parte se calculan dos costes diferentes: el de los resultados generados para los códigos latentes de las imágenes pasadas por el codificador, y el de las imágenes generadas para las interpolaciones de códigos latentes (parte muy importante para aprender el espacio uniformemente). En segundo lugar, como se sabe cuál es el resultado exacto que debería producir el generador para los códigos latentes obtenidos del codificador, se emplea un coste de distancia entre píxeles. La función empleada es la distancia L1 que mide la media de la distancia absoluta entre imágenes (Ecuación 4). En último lugar, se emplea una función que mide la variación total de una imagen por zonas para medir el nivel de ruido de la imagen, se trata de la función coste de Variación Total (Ecuación 5), que como se expone en el apartado *Edición material* tiene un gran peso en la calidad de la generación de imágenes nuevas.

$$\ell_1(x, y) = \text{mean}(L) = \text{mean}(\{l_1, \dots, l_N\}_T), \quad l_n = |x_n - y_n| \quad (4)$$

$$TV(image) = \sum_{i,j \in image} |y_{i+1,j} - y_{i,j}| + |y_{i,j+1} - y_{i,j}| \quad (5)$$

Con estas funciones de coste, finalmente el generador tiene que tratar de minimizar la siguiente función (Ecuación 6):

$$\ell_1(G(z), x) + \log(D(G(z))) + \log(D(G(z'))) + TV(G(z)) + TV(G(z')) \quad (6)$$

Donde  $G(z)$  corresponde con las imágenes generadas a partir de los códigos latentes obtenidos del codificador y  $G(z')$  corresponde con las imágenes generadas a partir de la interpolación de los códigos latentes originales,  $z$ .

### Anexo 3: Estudio del espacio latente reducido con PCA

En este anexo se van a mostrar más imágenes de las distintas pruebas realizadas sobre el espacio reducido a dos dimensiones con el análisis de las componentes principales.

#### Ajuste de las componentes para los diferentes entorno y formas

A continuación, se muestran los resultados de ajustar las componentes principales para las imágenes de los diferentes entornos de iluminación y las diferentes geometrías de los materiales.

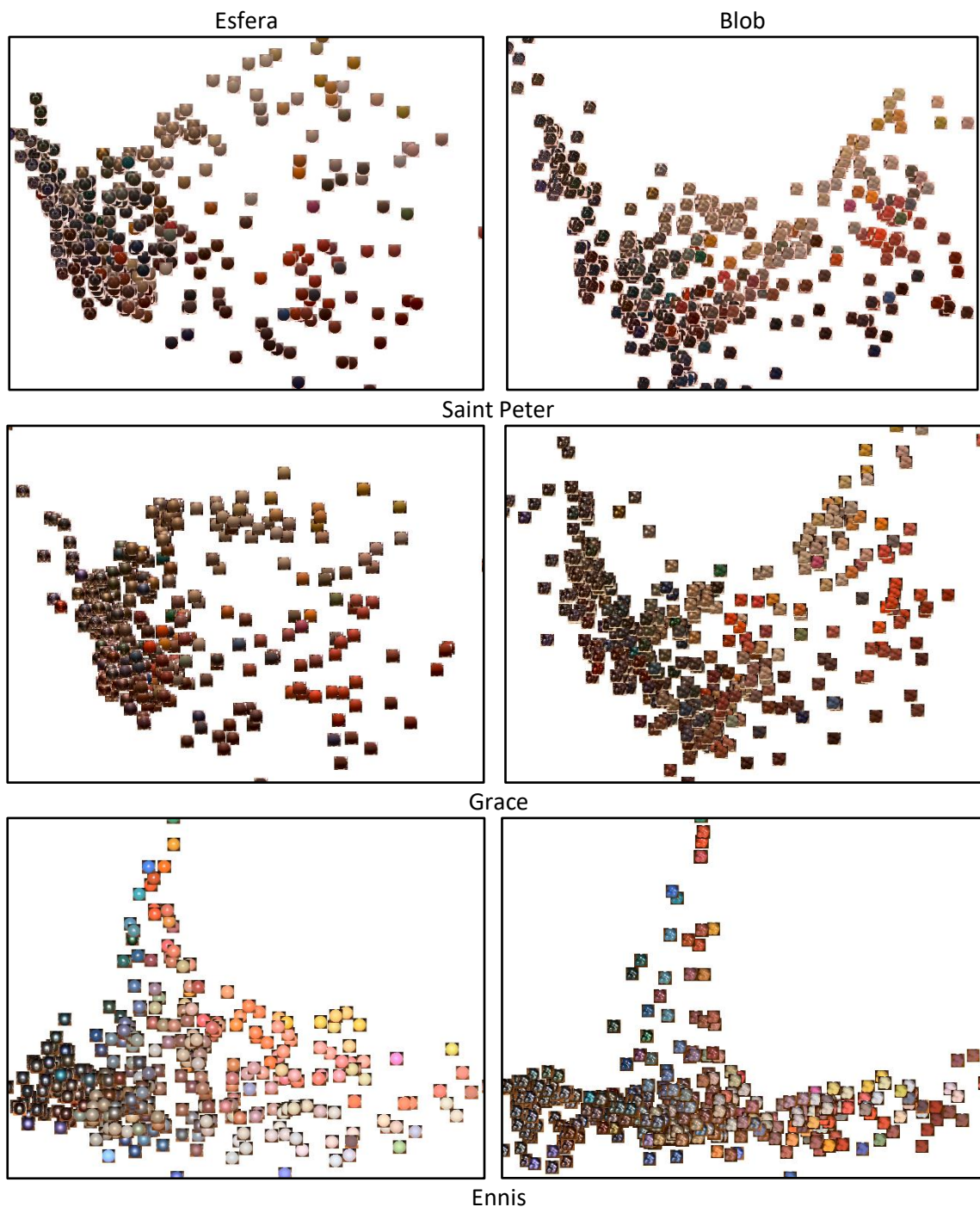


Figura 29. Representación del espacio reducido con PCA. En cada fila se han empleado las imágenes de un entorno de iluminación. En la primera columna se han empleado las muestras con forma de esfera y en la segunda con forma de blob. En todos los casos los eje principal (horizontal) es muy similar.

### Tercera componente principal

A parte de las muestras del entorno de *Grace*, las muestras del entorno de *Doge* también tienen más de un 2% en la variabilidad explicada por la tercera componente. De forma similar al caso de *Grace* en *Doge* (Figura 30) también se aprecia una correlación entre la tercera componente y los materiales especulares que no son metálicos y con colores vivos.

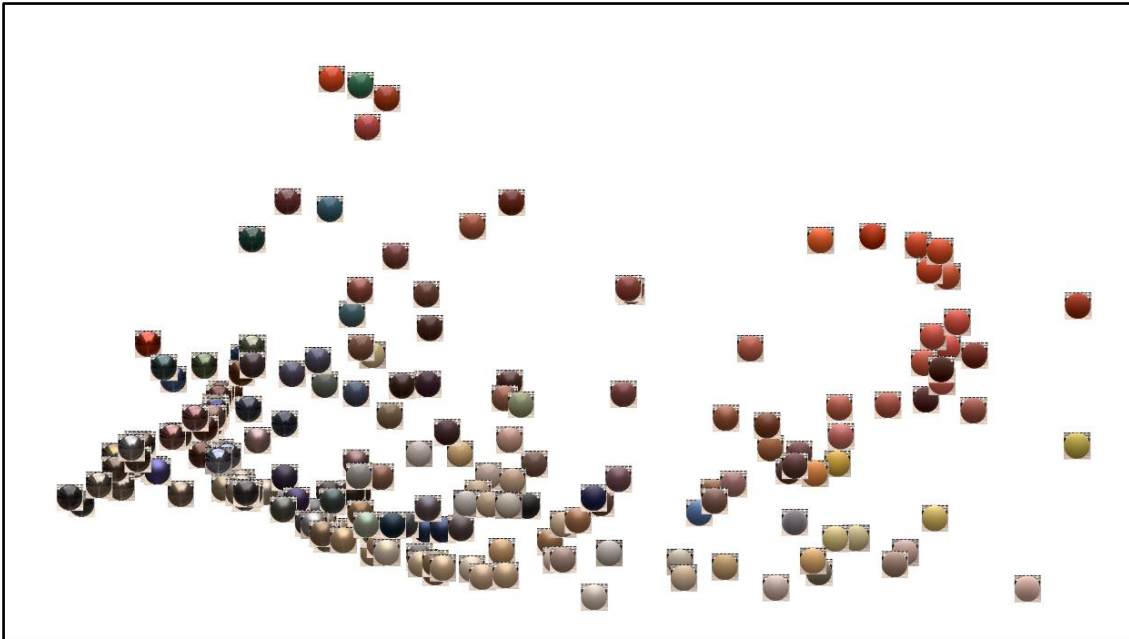


Figura 30. Vista de la primera y tercera componente principal de las imágenes del entorno Doge. Se aprecia como los materiales de colores más vivos se separan del resto, y sobre todo los materiales especulares no metálicos de los metálicos.

## Anexo 4: Desarrollo del proyecto

En este anexo se expone el tiempo dedicado a cada una de las tareas principales del proyecto (Tabla 16), así como en qué momento se ha realizado cada una de ellas (Figura 31). Al tratarse de un proyecto realizado dentro de la investigación de la Beca de Colaboración se ha desarrollado a lo largo de todo el curso, comenzando en octubre.

Tarea	Horas
Análisis semántico del espacio latente	55
Análisis mediante reducción con PCA	65
Análisis interno de la red	40
Entrenamiento del sistema generador	70
Pruebas de edición de imágenes	30
Entrenamiento del SDGAN	30
Memoria	85
<b>Total</b>	<b>375</b>

Tabla 16. Tareas del proyecto y número de horas invertido en cada una.

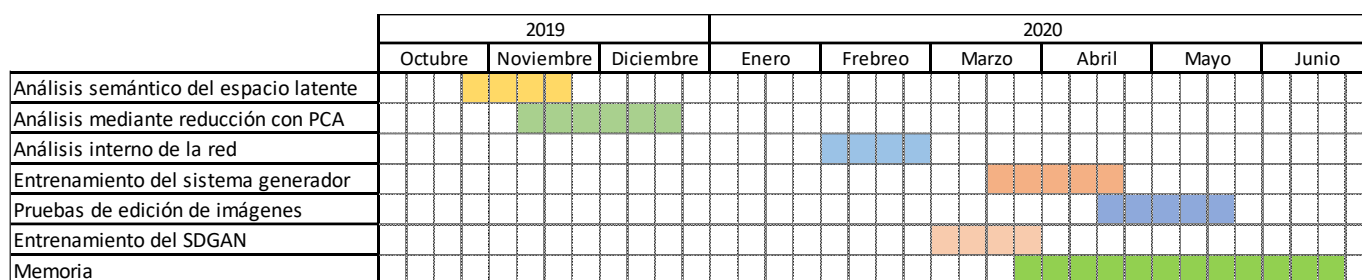


Figura 31. Diagrama de Gantt con las fases del desarrollo del trabajo.