

Título del Trabajo final de Grado

Trabajo Final de Grado

Grado en Multimedia

Apellidos: Ordóñez Brumós Nombre: Roger

Plan: 2009

Director: Raventós Mayoral, Arnau

**Índex**

[Resumen 3](#_Toc7627754)

[Palabras clave 4](#_Toc7627755)

[Enlaces 4](#_Toc7627756)

[Índex de tablas 5](#_Toc7627757)

[Índex de figuras 6](#_Toc7627758)

[Glosario 7](#_Toc7627759)

[Agradecimientos 8](#_Toc7627760)

[1. Introducción 9](#_Toc7627761)

[1.1 Motivación 11](#_Toc7627762)

[1.2 Formulación del problema 12](#_Toc7627763)

[1.3 Objetivos generales del TFG 13](#_Toc7627764)

[1.4 Objetivos específicos del TFG 13](#_Toc7627765)

[1.5 Alcance del proyecto 14](#_Toc7627766)

[2. Estado del arte 15](#_Toc7627767)

[2.1 Detección de caras 16](#_Toc7627768)

[2.1 Reconocimiento de caras 18](#_Toc7627769)

[2.3 Estudi de Mercat 18](#_Toc7627770)

[3. Gestió del projecte 20](#_Toc7627771)

[3.1 Procediment i Eines per al seguiment del projecte 20](#_Toc7627772)

[3.1.1 GANTT 20](#_Toc7627773)

[3.1.2 Trello 20](#_Toc7627774)

[3.1.3 GitHub repositori en xarxes, Git eines de control de versions 20](#_Toc7627775)

[3.2 Eines de validació 21](#_Toc7627776)

[3.3. DAFO 21](#_Toc7627777)

[3.4. Riscos i pla de contingències 22](#_Toc7627778)

[3.5. Anàlisi inicial de costos 22](#_Toc7627779)

[4. Metodología 23](#_Toc7627780)

[5. Desenvolupament del projecte 24](#_Toc7627781)

[6. Conclusions i treballs futurs 25](#_Toc7627782)

[7. Bibliografia 26](#_Toc7627783)

[8. Annexos 28](#_Toc7627784)

# Resumen

El resum (de 250 a 500 paraules) ha de reflectir el que és present en el text i no inclou referències bibliogràfiques. De manera resumida ha de descriure el problema, plantejar els objectius i l’abast de la investigació, així com la metodologia que s'ha emprat; resumint el resultat i sintetitzant les conclusions.

# Palabras clave

Entre 5 i 10 paraules que defineixin conceptes clau del treball. Han d’anar separades per comes.

# Enlaces

URL on està la maqueta, web, APP, videojoc, vídeo, etc.

# Índex de tablas

Taula 1: Nombre de la taula....................................................................................Pag. 11

Taula 2: Nombre de la taula....................................................................................Pag. 41

# Índex de figuras

Figura 1: Nombre de la figura.................................................................................Pag. 10

Figura 2: Nombre de la figura.................................................................................Pag. 56

# Glosario

# Agradecimientos

Quiero dedicar este apartado para agradecer:

Primero, a Arnau Raventós Mayoral, director de este Trabajo de final de grado, por su inestimable propuesta de Trabajo de final de grado, su confianza, disposición y guía en la realización del Trabajo de final de grado.

A mis padres, que me han apoyado durante la etapa de estudios.

A mi hermana, que me ha motivado a esforzarme en el trabajo de final de grado.

A vosotros,

Gracias.

# 1. Introducción

Estamos viviendo en un momento de la historia donde cada vez más estamos más conectados entre nosotros gracias a la tecnología. En poco más de quince años hemos pasado de una sociedad donde teníamos internet solo en casa a poder disponer de dispositivos móviles que nos conectan constantemente con internet i la información allí donde vamos.

Durante este proceso las persones cada vez han tenido un acceso más fácil y rápido a la información, con más contenidos y de mayor calidad, sobretodo en cuanto a material multimedia. La tecnología ha evolucionado a gran velocidad permitiendo dispositivos con mayor capacidad para almacenar archivos y a su vez aumentando la velocidad de las conexiones entre dispositivos.

Esta facilidad que nos ofrece este avance en tecnología ha implicado que aumente la demanda de contenidos multimedia, provocando que se genere una cantidad de información muy grande. En un día en internet se publica más contenido multimedia del que puede consumir una persona en toda su vida. Esto implica que hay mucha información entre la que escoger y encontrar aquello que queremos usar entre todo el conjunto de contenidos disponibles es una de las principales tareas que ofrecen diferentes tipos de buscadores inteligentes a partir de etiquetado de contenido.

El etiquetado en si nos permite clasificar diferentes contenidos para poder filtrarlos según nuestras necesidades de una manera rápida. Este etiquetado comúnmente se ha realizado de forma manual o a partir de mecanizaciones utilizando software diseñado para el etiquetado.

Este trabajo de final de grado propone implementar un sistema de indexación automática de contenidos audiovisuales, reconociendo los actores que participan en series o películas.

Para ello se analizara y se extraerá las caras de los actores a partir de videos de estos contenidos para generar librerías de caras. Estas librerías se utilizaran para detectar que actores son dentro del contexto de la película con la ayuda de la base de datos IMDB.

IMDB nos ofrece información referente a películas y con la cual podemos saber los actores que participan.

Este proceso está planteado a una escala reducida, conociendo inicialmente la película y actores que participan, creando un prototipo escalable que permita reconocer con precisión los actores con unos requisitos de procesado reducidos, tanto en cantidad de imágenes necesarias para procesar como en almacenamiento de esta información.

Este trabajo de final de grado está dividido en ocho apartados:

1. **Introducción:** Resumen de la problemática y contexto del proyecto.
2. **Estado del arte:** Referentes y tecnologías utilizadas actualmente que nos ayudan a resolver el problema.
3. **Gestión del proyecto:** Procesos y herramientas para la comunicación y gestión del proyecto.
4. **Metodología:** fases del proyecto que se realizaran para realizar el trabajo de final de grado.
5. **Desarrollo del proyecto:** Desarrollo de las fases del proyecto documentada.
6. **Conclusiones y trabajos futuros:** Información de cierre del trabajo con la visión personal de la utilidad y posibles aplicaciones futuras del trabajo de final de grado.
7. **Bibliografía**
8. **Anexos**

## 1.1 Motivación

El planteamiento de este proyecto surgió de una práctica durante unas sesiones de clase relacionadas con la biometría, tecnología de identificación basada en el reconocimiento de una característica física e intransferible de las personas, como por ejemplo, la huella digital.

En esta práctica generamos un sistema que comparaba diferentes imágenes y reconocía imágenes parecidas a partir de algoritmos sencillos y comparaciones de cuan diferentes eran las imágenes entre ellas.

Al considerar que esta área muy interesante y con mucha capacidad para crecer decidí ampliar mis conocimientos en sistemas de análisis de la información y Arnau Raventós Mayoral vio este interés y me propuso realizar el trabajo de final de grado bajo su tutela.

Actualmente hay compañías millonarias que analizan contenidos multimedia, pero sus sistemas de procesado y etiquetaje de estos contenidos estás aun mejorándose y no hay un techo en cuanto a un sistema perfeccionado que ofrezca una solución para etiquetar correctamente todo el contenido que se genera. Por eso mismo al ser una tecnología que se está investigando y mejorando da opción a entrar y aportar dentro del conjunto de trabajos que se están realizando actualmente en esta materia.

Por último, quería apartarme del resto de compañeros en la carrera y no quería generar más contenido multimedia sino una herramienta para poder analizar los contenidos que se generan y marcar un punto diferencial respecto a ellos.

## 

## 1.2 Formulación del problema

Actualmente en un día se genera y publica en internet más contenido multimedia del que es capaz de visualizar una persona en toda su vida. **[11]**

Una de las plataformas de video más conocidas, YouTube, recibe videos por parte de sus usuarios y genera en un día 4 millones de horas de contenido, equivalente a 452 años y medio al día.

Solo con el contenido que se genera en un día en YouTube ya se visualiza que es imposible para una persona revisar tal cantidad de contenidos y no es la única plataforma donde se generan.

Esta imposibilidad de revisar todo el contenido que se genera implica que se requiere de sistemas automáticos para poder comprobar estos contenidos y poder etiquetarlos para que los usuarios de las plataformas puedan encontrarlos y visualizarlos de forma rápida.

Estos sistemas de etiquetado están basados en aprendizaje supervisado **[12]** o semi-supervisado **[13]**, permitiendo generar un modelo que genera una función a partir de datos de entrenamiento. Esta función nos permite a partir de datos de entrada como videos recibir datos de salida, por ejemplo, etiquetas para clasificar estos videos.

Este trabajo de final de grado pretende generar un sistema para catalogar los actores que aparecen dentro de una película de forma automática utilizando algoritmos de inteligencia artificial ya existentes y fáciles de implementar gracias a APIs externas.

## 1.3 Objetivos generales del TFG

El objetivo general de este proyecto es la implementación de un sistema capaz de analizar y reconocer conjuntos de caras de actores extraídos de fragmentos de películas para etiquetarlos.

El proceso de etiquetado partirá del uso de algoritmos de Machine Learning, que permiten identificar de forma automática las caras de los actores en frames de películas utilizando datos de APIs externas como IMDB.

Para comprobar la validez del sistema se pretende analizar un conjunto reducido de películas para extraer los actores que participan y determinar quiénes son.

La validez del sistema se obtendrá de métricas extraídas en cuanto al porcentaje de acierto del sistema y del algoritmo en sí.

## 1.4 Objetivos específicos del TFG

Los objetivos específicos de este proyecto son:

1. Establecer una pequeña muestra de cinco películas para hacer pruebas con el algoritmo desarrollado.
2. Recopilar información sobre los actores que participan en las películas seleccionadas.
3. Implementar un sistema que genere automáticamente bases de datos de imágenes con las caras de los actores de las películas seleccionadas.
4. Utilizar APIs externas para analizar las películas seleccionadas y obtener clústers no supervisados de caras similares.
5. Comparar las bases de datos de actores con los clústers generados por el sistema no supervisado para reconocer a que caras corresponden los diferentes actores que participan.
6. Una vez seleccionado el actor con sus correspondientes caras, mostrar información de este actor y de la película.
7. Establecer y consolidar los conocimientos necesarios para resolver la problemática mediante la programación de un algoritmo capaz de solventar las necesidades previamente expuestas.

## 1.5 Alcance del proyecto

Este trabajo de final de grado pretende generar un algoritmo de análisis de películas para poder etiquetar información referente a los actores que participan en estas.

El algoritmo beneficiara a dos sectores diferentes de usuarios:

* Usuario que consume películas y quiere conocer más información referente a los actores que participan en ella. Actualmente para encontrar esta información se requiere de herramientas de terceros y búsquedas específicas que requieren conocimientos de datos como el nombre de los actores para encontrar esta información.
* Usuarios de plataformas de contenido que quieran catalogar sus series o películas de forma automática sin necesidad de emplear recursos humanos para ello.

La complejidad del proyecto presenta un conjunto de obstáculos en cuanto a refinar el prototipo debido a que las muestras serán limitadas y por lo tanto no se podrá g arantizar en ningún caso un acierto del algoritmo del 100%.

# 2. Estado del arte

Este proyecto pretende automatizar la detección de actores dentro de películas. Para ello se requiere en primera instancia detectar si aparece un actor en una escena y encontrar algún rasgo característico que sirva para identificar de quien se trata y diferenciarlo del resto de actores. Para ello se puede emplear la detección facial.

La detección facial es una rama de la biometría, ciencia que mide las propiedades físicas y de comportamiento de los seres vivos.

La biometría nos permite identificar o verificar un individuo del resto. Las características físicas de cada individuo lo caracterizan y en la mayoría de casos no se comparten con otros individuos. El ejemplo más concreto para diferenciar dos individuos sería una comparación de ADN. En el ámbito de la tecnología cuotidiana tenemos sistemas que analizan nuestras características físicas como por ejemplo los sistemas de desbloqueo de huella dactilar de los dispositivos móviles.

Existen diferentes sistemas biométricos que permiten identificar y verificar individuos. Todos ellos requieren de unos requisitos para funcionar, como en el caso del análisis de ADN una muestra de sangre, una imposición intrusiva que requiere el sistema para funcionar. En el caso del ADN además se requiere de un tiempo elevado para procesar la muestra y de un coste elevado de recursos.

Estas complejidades hacen que estos sistemas tan precisos como en el caso del análisis de ADN nos impidan realizar análisis biométricos a gran escala y en tiempos reducidos.

En el caso de este trabajo de final de grado utilizamos un sistema biométrico que trabaja a partir de fotografías de caras para analizar las características de estas, requisito no intrusivo en comparación y con un coste muy pequeño ya que al tener a disposición las películas podemos extraer esta información.

La detección facial no requiere por tanto de muestras físicas de los individuos i no requiere de la voluntad de estos para ser capturada, un claro ejemplo de esto es su aplicación en sistemas de seguridad.

Algunos de los usos comunes de la detección de caras en seguridad son:

* Identificación de criminales
* Registro de asistencia de persones a eventos o en ámbitos laborales.
* Verificación de identidad en aeropuertos, aduanas, entre otros.

Diversas tareas en el ámbito del procesamiento de imagen en caras son:

* Detección de caras
  + La capacidad del modelo para encontrar una cara dentro del contexto de una imagen.
* Verificación de caras
  + La capacidad de discernir si dos caras pertenecen a la misma persona.
* Reconocimiento de caras
  + La capacidad de decir a quien pertenece una cara, por lo tanto identificando a ese individuo.
* Clustering de caras
  + La capacidad de encontrar caras con características parecidas dentro de un gripo de caras.

En este trabajo de final de grado se utilizara la detección de caras conjuntamente con el reconocimiento de caras para localizar los actores dentro de las películas y decir a quien pertenecen esas caras.

A continuación se dará una visión global del estado del arte de estos dos procedimientos.

## 2.1 Detección de caras

La detección facial se ocupa de encontrar una cara dentro de una imagen digital, si es que hay alguna **[3]**. Es un requisito previo de cualquier sistema automático que requiera localizar caras dentro de imágenes. La detección facial está especialmente diseñada para funcionar con imágenes singulares, pero para este proyecto se utilizara para extraer caras de secuencias de video, consiguiendo una muestra de caras más grande y precisa de los diferentes actores que aparecen en las películas.

La detección facial es el proceso que se estudia en el ámbito de la computación de información y tiene un recorrido extenso en la historia. Se podría decir que, desde que existen imágenes digitales, se ha analizado la información que contienen.

Los métodos de detección clásicos están basados en estudios como ***“detección rápida de objetos mediante características sencillas”*** **Viola, P. Jones, M 2004 [5]** donde se analizan caras mediante el uso de filtros secuenciales que permiten encontrar las diferencias de contraste en intensidad de secciones de caras para poder generar unos vectores representativos que las caractericen. Es un sistema en cascada por etapas que trabaja la información por bloques de menor a mayor precisión.

Estos modelos han ido evolucionando y actualmente los más utilizados son:

* MTCNN **[4]**
  + Sistema que aprovecha las redes neuronales convolucionales para localizar patrones dentro de imágenes que pertenecen a elementos faciales y explota el uso de un sistema en cascada por etapas que trabaja por bloques la información de menor a mayor precisión.
* SSH **[6]**
  + Sistema més costós computacionalment que presenta millors resultats que els anteriors degut a que es basa en un mètode de detecció d’objectes .

## 

## 2.2 Reconocimiento de caras

El reconocimiento de una cara es el proceso por el cual podemos analizar las características propias de una cara para identificar o verificar la identidad de una persona. Esta técnica se conoce como reconocimiento facial. El análisis de las imágenes recogen las características faciales que describen una cara en un conjunto de datos que luego podemos comparar con conjuntos de datos conocidos para calcular la diferencia estos.

El uso de versiones clásicas como PCA **[9]** permite reducir la complejidad de la información que se presenta en una imagen reduciendo la redundancia y preservando los cambios notables en la imagen distribuyendo coeficientes de mayor a menor cantidad de información relevante. Estos coeficientes se pueden comparar con pares de imágenes i calcular si las dos imágenes son similares o no.

Actualmente en el sector del reconocimiento facial una de las diferentes opciones que hay dentro del estado del arte es FaceNet **[1]** de Google. La base de funcionamiento de FaceNet es la agrupación de grupos de caras den paquetes de 3 caras, dos parecidas y una diferente. Estos paquetes comparan las características de las caras que lo forman y generan dos vectores de distancia respecto a una de las tres caras generando un ancla. Posteriormente genera un par de distancias con estos vectores, situando las caras que son más parecidas a poca distancia y la cara que es más diferenciada a más distancia. Este método genera comparaciones donde tenemos directamente una salida de caras parecidas y no parecidas y podemos conseguir con mucha velocidad grupos de caras parecidas y diferentes dentro de clústers de imágenes.

Además de FaceNet, otras empresas como Facebook proponen soluciones como FaceMatch **[8]**, también centrada al reconocimiento de caras con otras aproximaciones, pero que aportan como el resto una salida de información en forma de valores numéricos o descriptores faciales.

Todos estos sistemas usan Redes Neuronales ya que permiten a partir de información conocida, por ejemplo en el caso del reconocimiento de caras clústers de caras, entrenar un sistema que aprende a realizar tareas de reconocimiento de patrones. En el caso del reconocimiento de caras se enseña a estos sistemas a reconocer patrones dentro de la estructura facial para generar en la salida los valores que determinan los descriptores faciales.

Este proceso de entrenamiento guiado permite definir modelos como los antes citados que generan en poco tiempo y con un porcentaje de error reducido la información necesaria para comparar las caras.

## 2.3 Estudio de Mercado

El propósito de este trabajo de final de grado no pretende crear una herramienta que genere una competencia contra otros productos o soluciones comerciales actuales.

Actualmente existen sistemas que muestran información de los artistas que aparecen en series o películas en plataformas de contenido a la carta como HBO o Netflix, los datos de las cuales son introducidos por operarios de estas empresas para poder ofrecer esta información.

Existen además sistemas comerciales como DIVE **[14]**, antes conocida como Touchvie que actualmente se ha integrado en las televisiones inteligentes de Samsung, capaces de reconocer una película o una seria y mostrar información de los actores o vestuario. Este sistema utiliza el audio de la película o el canal en el que aparece para reconocerla y mostrar información sobre esta. Una de las limitaciones que tiene este sistema es que requiere de datos previos introducidos. Esto se puede obtener de comentarios del CEO José Luis Flórez, cuando indicaba que el funcionamiento de sus sistema utilizaba el top 100 de películas de cada año, por lo tanto denotando que la información no se generaba dinámicamente y que por el contrario se introducía a mano a partir de bases de datos como IMDB o parecidas.

Amazon ofrece dentro de su contenido a la carta la opción X-ray **[10]** que muestra información de los contenidos que ofrece al usuario en diferentes plataformas y dispositivos para la mayoría de sus títulos. Como en los casos anteriores esta información la introducen manualmente operarios de Amazon.

En cuanto a sistemas no específicos de reconocimiento de actores tenemos por ejemplo Clarifai que permite detectar i verificar a que actor pertenece una imagen, pero a diferencia de lo requerido por el sistema no trabaja con video y solo opera con imágenes estáticas.

Estos sistemas en general presentan inconvenientes conocidos, o bien requieren de datos introducidos manualmente o por el contrario están limitados a los datos conocidos y requieren de un mantenimiento y ampliación de las bases de datos con las que operan para trabajar.

El sistema que se propone para este trabajo de final de grado pretende automatizar el reconocimiento y búsqueda de datos extrayendo información de bases de datos fiables como es el caso de IMDB. Esto permite reducir tanto el coste de mantenimiento del sistema como aumentar la velocidad respecto a los métodos donde se introduce la información manualmente y facilita que sea un sistema escalable y que pueda funcionar incluso con películas o series españolas, catalanas o poco conocidas.

# 3. Gestión del proyecto

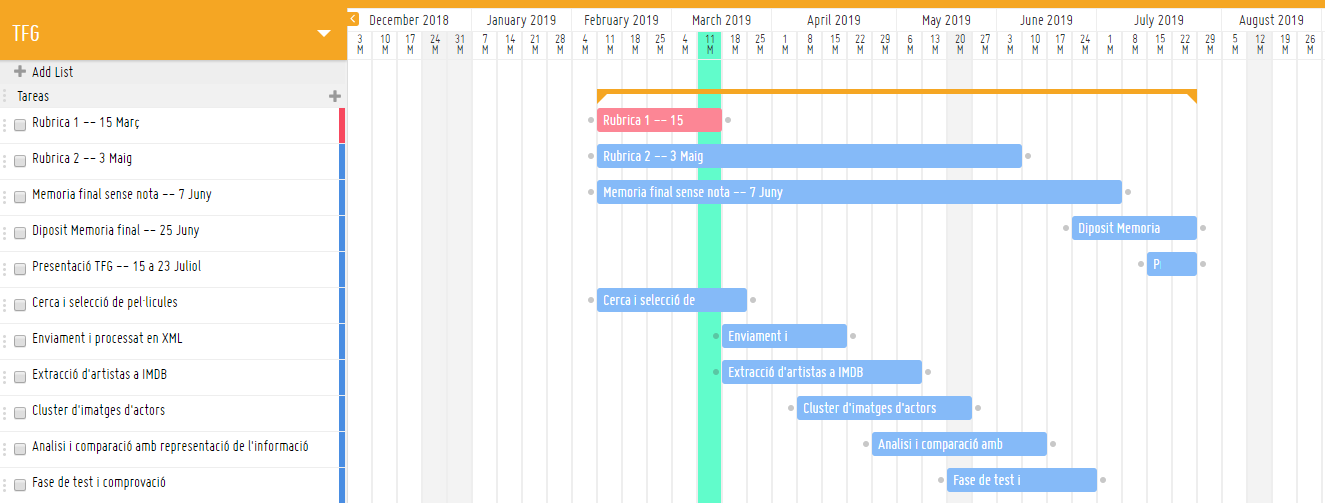
## 3.1 Procedimiento y Herramientas para el seguimiento del proyecto

El procedimiento estipulado para el seguimiento del proyecto consta de sesiones presenciales cada dos semanas y un seguimiento no presencial a partir de Mayo donde se evalúan las tareas con la supervisión del tutor y se revisa la evolución y etapas del proyecto.

A partir de Mayo el seguimiento pasa a ser telemático vía correo electrónico.

El proyecto esta desglosado en un grupo de metas que se describen de forma visual en el diagrama de Gantt del apartado 3.1.1 y que están descritas en profundidad en el apartado 4 Metodología.

### 3.1.1 GANTT



### 3.1.2 Trello

Trello es una plataforma web que ofrece un gestor de tableros basados en Kanban, sistema de información de tarjetas. Se utiliza para generar tareas o tarjetas y generar un seguimiento del desarrollo del proyecto simplificado. Esta metodología se usa ampliamente en el mercado para gestionar la productividad y el flujo de ejecución de las tareas de proyectos complejos.

Para este trabajo de final de carrera se han desglosado unas tareas que posteriormente han servido para situar tiempos de ejecución en un diagrama de Gantt que refleja el tiempo estimado de las tareas para poder realizar el trabajo de final de grado.

### 3.1.3 GitHub

GitHub es una plataforma web que ofrece un servicio de repositorios en la nube. Estos repositorios sirven como una copia tanto de seguridad, permitiendo obtener acceso a la información respaldada, como a su vez como una herramienta de evolución de contenido.

A partir de las herramientas de control de versión mediante Git se tiene no solo el contenido en la nube, sino todo el proceso de evolución del contenido y se puede saltar a cualquier punto del desarrollo del proyecto.

En si una versión de control ayuda a los desarrolladores de contenidos a llevar un registro de todos los cambios tanto en código como en contenido de un proyecto.

Git es un sistema de control de versión creada por Linus Torvalds en 2005.

Para este proyecto de final de grado es imprescindible el uso de un sistema de control de versiones ya que se requiere un control de la evolución de código y la capacidad de poder acceder a este en diferentes máquinas para su edición y testeo, tanto por parte de creación como de pruebas del sistema.

Todos los contenidos del proyecto que formen parte de producción de código estarán alojados en GitHub para mantener un control de versión del código y poderlo ofrecer posteriormente al finalizar como una herramienta de desarrollo y punto de arranque para futuros proyectos.

## 

## 3.2 Herramientas de validación

La validació del model funcional es farà a partir d’analitzar la fiabilitat i funcionament del model en sí amb un grup de pel·lícules reduït (5) en primeres fases de test i posteriorment amb un grup més ampli de pel·lícules per calcular el nivell de precisió del sistema.

Aquest test tindrà en compte el nombre de encerts i errors per tal de calcular el percentatge de fidelitat del model.

## 3.3. DAFO

|  |  |  |
| --- | --- | --- |
|  | **Positius** | **Negatius** |
| **Origen**  **Intern** | **Fortaleses**   1. Bon nivell de programació 2. Interès en aprendre i buscar informació 3. Bons resultats en estudis similars anteriors (treballs universitaris relacionats) | **Debilitats**   1. Disponibilitat horària reduïda per treball |
| **Origen**  **Extern** | **Oportunitats**   1. Gran quantitat d’informació i referents 2. Projectes demo Open Source 3. Àmbit d'ús del tutor del treball de final de grau. | **Amenaces**   1. costos derivats del estudi o infraestructura necessària per operar el sistema. |

## 

## 3.4. Riscos i pla de contingències

Els possibles riscos identificats d’aquest projecte, i les seves corresponents solucions son les següents, ordenades de menor a major importància:

|  |  |
| --- | --- |
| **Risc** | **Solució** |
| Impossibilitat de crear un model de reproducció de dades realtime. | Creació de model teòric o emulat que demostri el resultat general del sistema de detecció i classificació de pel·lícules. |
| Costos de manteniment web. | reducció de costos amb hosting propi amb maquinaria pròpia. |
| Necessitat de més temps per elaborar el model. | Traspassar la data d’entrega de TFG a la segona data posterior a agost. |

## 3.5. Anàlisi inicial de costos

El treball final de grau conforma un ús de 300 hores per tal de desenvolupar-se, per tant el primer gran cost inicial son aquestes hores que s’han de dedicar per la correcta finalització del treball final de grau.

Per altre part per tal de gestionar de manera adequada aquest temps el mètode de treball escollit aprofita al màxim el desenvolupament del treball de final de grau en tasques i fites assequibles i amb marge en el temps.

Com es tracta de un projecte prototip es farà ús de eines de programació de codi obert i sistemes d’emmagatzemament de recursos públics sense costos com ara GitHub i l’ús de hardware propi, per tant no hi ha necessitat d’una inversió ni en infraestructura ni en programari o emmagatzematge.

# 4. Metodología

La metodologia proposada per aquest projecte es Scrum, aquest procés s’aplica de manera que es pretén obtenir el millor resultat possible de un projecte. Com el treball requereix de unes fites complexes i es necessita obtenir resultats en un espai de temps reduït. La metodologia Scrum planteja la elaboració de fites i iteracions que normalment son de dues setmanes, temps que coincideix amb les sessions de seguiment amb el tutor del treball de final de grau, poden verificar la compleció de les tasques i planificar les següents en conseqüència. Cada iteració planteja proporcionar un compliment de la fita dintre dels quinze dies que formen la iteració de cada fita.

Les fases o fites generals que es desglossen en sub fases, posteriorment, seran:

* Cerca i selecció de pel·lícules per crear la mostra de treball per el model.
  + En aquesta fita s’ha de triar un conjunt petit de pel·lícules que serveixin com a mostra del sistema per fer les proves i extracció de dades inicials.
* Enviament i processat de les pel·lícules i extracció de dades en format XML.
  + En aquest punt s’utilitza una api funcional d'anàlisi d’escenes proporcionat per el tutor del treball de final de grau per tal de extreure la informació característica de les cares que apareixen a diferents moments de la pel·lícula.
* Extracció d’artistes i la seva informació a IMDB, base de dades de pel·lícules.
  + Com partim de conèixer inicialment el títol de la pel·lícula podem buscar els actors que hi participen i informació general com ara altres titols on han participat.
* Creació d’una base de dades d’imatges d’actors amb cercadors web.
  + Mitjançant l'anàlisi anterior es requereix imatges que pertanyin a actors de les pel·lícules per tal de generar descriptors emprant una base de dades d’imatges que es generarà amb l’us de cercadors web.
* Anàlisi i comparació del clúster amd els descriptors d’actors i representació de informació.
  + En aquesta fase es comparen els descriptors generats a partir del clúster de la pel·lícula amb els descriptors dels diferents actors que hi participen en ella.
  + Un cop identificat el actor es preveu mostrar informació variada d’aquest.
* Fase de test i comprovació.
  + Per últim s'avalua la capacitat del sistema i el percentatge de fiabilitat amb una mostra més gran de pel·lícules i s’analitza sí es un model correcte i funcional.

# 5. Desenvolupament del projecte

Aparts propis de aquest tema como pot ser usabilitat, maquetació, disseny, programació, proves, etc.

# 6. Conclusions i treballs futurs

# 7. Bibliografia

1. *FaceNet: A Unified Embedding for Face Recognition and Clustering. 17 Juny, 2015* [*https://arxiv.org/abs/1503.03832*](https://arxiv.org/abs/1503.03832) *Consultat. 8 Març, 2019*
2. *SSH: Single Stage Headless Face Detector. 18 Oct, 2017* [*https://arxiv.org/abs/1708.03979*](https://arxiv.org/abs/1708.03979) *Consultat. 8 Març, 2019*
3. *Face detection* [*http://faculty.ucmerced.edu/mhyang/papers/face-detection-chapter.pdf*](http://faculty.ucmerced.edu/mhyang/papers/face-detection-chapter.pdf) *Consultat. 12 Març, 2019*
4. *Joint Face Detection and Alignment using Multi-task Cascaded Convolutional Networks* [*https://arxiv.org/abs/1604.02878*](https://arxiv.org/abs/1604.02878) *Consultat. 12 març, 2019.*
5. *Rapid Object Detection Using a Boosted Cascade of Simple Features* [*http://www.merl.com/publications/docs/TR2004-043.pdf*](http://www.merl.com/publications/docs/TR2004-043.pdf) *Consultat. 12 Març, 2019.*
6. *SSH: Single Stage Headless Face Detector* [*http://openaccess.thecvf.com/content\_ICCV\_2017/papers/Najibi\_SSH\_Single\_Stage\_ICCV\_2017\_paper.pdf*](http://openaccess.thecvf.com/content_ICCV_2017/papers/Najibi_SSH_Single_Stage_ICCV_2017_paper.pdf) *Consultat. 12 Març, 2019.*
7. *The return of AdaBoost.MH: multi-class Hamming trees* [*https://arxiv.org/abs/1312.6086*](https://arxiv.org/abs/1312.6086) *Consultat. 12 Març, 2019.*
8. *Extret de la pagina* [*https://www.rankred.com/face-recognition-algorithms-techniques/*](https://www.rankred.com/face-recognition-algorithms-techniques/) *Consultat. 12 Març, 2019.*
9. *Eigenfaces for Recognition* [*http://www.face-rec.org/algorithms/PCA/jcn.pdf*](http://www.face-rec.org/algorithms/PCA/jcn.pdf) *Consultat. 13 Març, 2019*
10. *X-ray* [*https://www.amazon.com/adlp/xray*](https://www.amazon.com/adlp/xray) *Consultat. 14 Març, 2019*
11. *Información de contexto extraida de* <https://www.forbes.com/sites/bernardmarr/2018/05/21/how-much-data-do-we-create-every-day-the-mind-blowing-stats-everyone-should-read/#683d7b7760ba> Consultado. 20 Abril, 2019
12. *Aprendizaje supervisado* <https://es.wikipedia.org/wiki/Aprendizaje_supervisado> Consultado, 20 Abril, 2019
13. *Aprendizaje semi-supervisado* <https://es.wikipedia.org/wiki/Aprendizaje_semisupervisado> Consultado, 20 Abril, 2019
14. *Dive* <https://dive.tv/> Consultado, 1 Mayo, 2019

# 8. Annexos

**Exemples**

Guió Literari

Guió Tècnic

gantt anex <https://app.trellogantt.com/gantt>