



Implementación de técnicas de visión artificial para la identificación de frutos de duraznero

Autor:

Ing. Sergio Hinojosa

Director:

Ing. Maxim Dorogov (FIUBA)

*Esta planificación fue realizada en el curso de Gestión de proyectos
entre el 18 de junio de 2024 y el 13 de agosto de 2024.*

Índice

1. Descripción técnica-conceptual del proyecto a realizar	5
1.1 Contexto de la implementación	5
1.2 Objetivo del proyecto	5
2. Identificación y análisis de los interesados	6
3. Propósito del proyecto	6
4. Alcance del proyecto	6
5. Supuestos del proyecto.	7
6. Requerimientos	7
7. Historias de usuarios (<i>Product backlog</i>).	8
8. Entregables principales del proyecto	9
9. Desglose del trabajo en tareas	10
10. Diagrama de Activity On Node.	11
11. Diagrama de Gantt	12
12. Presupuesto detallado del proyecto	14
13. Gestión de riesgos	14
14. Gestión de la calidad	16
15. Procesos de cierre	18

Registros de cambios

Revisión	Detalles de los cambios realizados	Fecha
0	Creación del documento	18 de junio de 2024
1	Se completa hasta el punto 5 inclusive	2 de Julio de 2024
2	Se completa hasta el punto 9 inclusive	9 de Julio de 2024
3	Se completa hasta el punto 12 inclusive	22 de Julio de 2024
4	Se completa el plan	31 de Julio de 2024
4.1	Se completa el plan	4 de Agosto de 2024

Acta de constitución del proyecto

Buenos Aires, 18 de junio de 2024

Por medio de la presente se acuerda con Ing. Sergio Hinojosa que su Trabajo Final de la Carrera de Especialización en Inteligencia Artificial se titulará “Implementación de técnicas de visión artificial para la identificación de frutos de duraznero” y consistirá en la implementación de un algoritmo que permita identificar y cuantificar frutos de duraznos a partir de imágenes de árboles tomadas a campo. El trabajo tendrá un presupuesto preliminar estimado de 600 horas y un costo estimado de US\$18267, con fecha de inicio el 18 de junio de 2024 y fecha de presentación pública en el mes de diciembre de 2024.

Se adjunta a esta acta la planificación inicial.

Dr. Ing. Ariel Lutenberg
Director posgrado FIUBA

Dr. Gerardo Sánchez
INTA

Ing. Maxim Dorogov
Director del Trabajo Final

1. Descripción técnica-conceptual del proyecto a realizar

1.1. Contexto de la implementación

Contar de forma manual la cantidad de frutos que posee un árbol en una plantación agrícola resulta una tarea laboriosa, lenta y propensa a errores. No obstante, conocer estos datos tiene diversas aplicaciones importantes que, debido a su dificultad, no están siendo abordadas. Para un productor, por ejemplo, saber la cantidad de frutos que posee una muestra de su lote durante el raleo* le permite calcular la intensidad a aplicar. Además, al momento de la cosecha, le permite tener una estimación de su producción.

1.2. Objetivo del proyecto

Este proyecto busca proporcionar al productor:

- Precisión y eficiencia: conteo de manera precisa y rápida.
- Ahorro de tiempo y costos: reducción del tiempo y de costos asociados con la mano de obra requerida para esta tarea.
- Optimización del rendimiento agrícola: permitir toma de decisiones informadas sobre la gestión del cultivo, la cosecha y la planificación de la mano de obra.
- Monitoreo temprano de la producción.

Por otro lado, los algoritmos a desarrollar resultan una herramienta muy útil, ya que permiten la obtención de un gran volumen de datos que pueden vincularse con diferentes características de interés, tales como:

- Porcentaje de cuajado: relacionado con la producción.
- Potencial de raleo: capacidad de genotipo a soportar mayor o menor raleo.
- Rendimiento.

A esto se lo denomina fenotipo y esta herramienta fenómica será fundamental para alimentar modelos de IA que combinan datos genómicos con variables ambientales.

El siguiente diagrama de bloques ilustra el pipeline del sistema, los diferentes subsistemas involucrados y el flujo de datos:

*Descarte de una parte de los frutos para que los restantes tomen más nutrientes de la planta y puedan alcanzar tamaño comercial

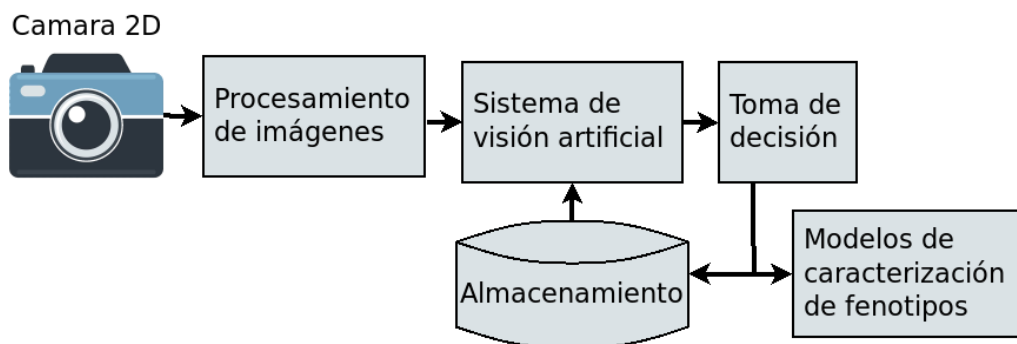


Figura 1. Diagrama en bloques del sistema.

2. Identificación y análisis de los interesados

Rol	Nombre y Apellido	Organización	Puesto
Cliente	Dr. Gerardo Sánchez	INTA	Director Científico Biotango Technologies SAS
Responsable	Ing. Sergio Hinojosa	FIUBA	Alumno
Orientador	Ing. Maxim Dorogov	FIUBA	Director del Trabajo Final

Cliente: el Dr. Gerardo Sanchez tiene más de 20 años de experiencia en biotecnología frutícola, colaborará en los requerimientos y en la validación de los modelos desarrollados.

Orientador: el Ing. Maxim Dorogov posee gran experiencia y conocimientos en el área de visión artificial. Es orientador desde los aspectos técnicos y participa en definir los objetivos del proyecto.

3. Propósito del proyecto

El propósito de este proyecto es desarrollar un sistema informático capaz de detectar y contar los frutos de un duraznero en el árbol a partir de imágenes tomadas por una cámara 2D.

4. Alcance del proyecto

El proyecto incluye:

- Procesamiento de la imagen tomada por un celular.
- Detección del árbol a analizar.
- Conteo de los frutos.

El proyecto no incluye:

- Procesamiento de imágenes en blanco y negro.
- Procesamiento de video.
- Implementación industrial. No se realizará un deploy en ningún sistema embebido ni en la nube, el sistema se mantendrá en un entorno de desarrollo.

5. Supuestos del proyecto

Para el desarrollo del presente proyecto se supone que se contará con:

- Horas de trabajo del desarrollador y del director del proyecto.
- Un ordenador con placa gráfica dedicada.
- Banco de imágenes de campo reales para el entrenamiento del sistema.
- No será necesario una inversión en hardware.
- No se espera una precisión del 100 %. Se definirá un margen aceptable.

6. Requerimientos

1. Requerimientos funcionales:

- 1.1. El sistema debe ser capaz de identificar y cuantificar frutos tanto en estado inmaduro o maduro a partir de fotos de árboles tomadas en la plantación.
- 1.2. El sistema debe ser capaz de analizar fotos en formato .jpeg.
- 1.3. El sistema deberá indicar el grado de confianza de la detección y el error esperado en las mediciones.
- 1.4. El sistema debe mostrar los resultados en una interfaz gráfica y debe exportar los resultados a formato .csv
- 1.5. Los dataset y código en repositorios deben mantenerse NO abiertos, de modo de mantener la confidencialidad del trabajo.
- 1.6. El usuario debe poder analizar una foto individual o un grupo de fotos, por lo menos grupos de 20 imágenes.
- 1.7. El usuario debe ser capaz de cambiar el nivel de confianza para modificar la detección.

2. Requerimientos de documentación:

- 2.1. Se debe proveer el códigos base.
- 2.2. Se deben describir los ensayos y condiciones para la reproducción de los mismos.
- 2.3. Se debe generar el informe de avance y la memoria final del trabajo.
- 2.4. Se debe proveer un manual de instrucciones para la instalación y uso.

3. Requerimientos de testing:

- 3.1. El modelo debe ser entrenado a partir de modelos pre-entrenados utilizando la técnica de *transfer learning*.

- 3.2. El entrenamiento del modelo debe realizarse a partir de un dataset creado con imágenes propias y de terceros respetando los derechos de uso establecidos por los mismos.
- 3.3. El entrenamiento y las pruebas en los datasets de validación y testing deben ser llevadas adelante en notebooks escritas en lenguaje de programación Python 3.8 o superior.
4. Requerimientos de la interfaz:
 - 4.1. La interfaz debe ser intuitiva, con capacidad de cargar las fotos de manera sencilla.

7. Historias de usuarios (*Product backlog*)

El criterio establecido para la asignación de los puntajes de las historias de usuario se basa en los criterios:

1. Esfuerzo requerido en tiempo.
2. Complejidad de la tarea.
3. Riesgo asociado.

Cada uno de los criterios individuales puede tener un valor entre 0 (bajo) y 5 (alto). El puntaje final será la suma de los puntajes individuales y se le asignará el valor más próximo de los valores de la serie de Fibonacci: 1, 2, 3, 5, 8, 13, 20.

- Como usuario, quiero un sistema de visión artificial que detecte y cuente automáticamente la cantidad de duraznos en mis árboles, para poder monitorear y gestionar mejor mi producción.
 - Complejidad: 4
 - Dificultad: 3
 - Incertidumbre: 3
 - *Story points*: 13
- Como usuario, quiero poder capturar imágenes de mis árboles utilizando una cámara y poder subirla al sistema para el conteo de los frutos.
 - Complejidad: 1
 - Dificultad: 1
 - Incertidumbre: 2
 - *Story points*: 5
- Como usuario, quiero poder configurar parámetros del sistema (como la sensibilidad de detección), para ajustar la precisión del conteo según las condiciones específicas de mi plantación.
 - Complejidad: 4

- Dificultad: 2
- Incertidumbre: 4
- *Story points*: 8

- Como usuario, quiero recibir un reporte con la cantidad de duraznos detectados en cada árbol, para poder tomar decisiones informadas sobre mi cultivo.
 - Complejidad: 2
 - Dificultad: 2
 - Incertidumbre: 2
 - *Story points*: 8

- Como desarrollador, quiero establecer un entorno de desarrollo escalable, bien documentado y organizado para facilitar la mejora continua del sistema.
 - Complejidad: 3
 - Dificultad: 5
 - Incertidumbre: 3
 - *Story points*: 13

- Como desarrollador, quiero crear un sistema de notificaciones de error y logeo que alerte sobre los problemas en la detección y sea más sencillo remediar bugs o corregir parámetros de configuración.
 - Complejidad: 3
 - Dificultad: 3
 - Incertidumbre: 2
 - *Story points*: 8

- Como desarrollador, quiero diseñar una base de datos para almacenar las imágenes y los resultados de detección, para poder acceder y analizar los datos históricos.
 - Complejidad: 2
 - Dificultad: 2
 - Incertidumbre: 2
 - *Story points*: 8

8. Entregables principales del proyecto

Los entregables del proyecto son:

- Documento de arquitectura de software.
- Notebooks desarrolladas para el entrenamiento del modelo.

- Datasets utilizados.
- Manual de usuario.
- Informe de avance.
- Informe final.

9. Desglose del trabajo en tareas

1. Gestión del proyecto (50 h)
 - 1.1. Estudio de necesidades. (6 h)
 - 1.2. Análisis de requerimientos. (6 h)
 - 1.3. Análisis de factibilidad. (6 h)
 - 1.4. Entregables GdP hasta sección 5. (4 h)
 - 1.5. Entregables GdP hasta sección 9. (4 h)
 - 1.6. Entregables GdP hasta sección 12. (6 h)
 - 1.7. Entregables GdP hasta sección 15. (6 h)
 - 1.8. Elaboración presentación GdP. (6 h)
 - 1.9. Revisión y correcciones entregables GdP. (6 h)
2. Investigación preliminar (110 h)
 - 2.1. Investigación de modelos de detección y clasificación de objetos. (40 h)
 - 2.2. Desarrollo de tests de performance para la aplicación. (40 h)
 - 2.3. Análisis de los resultados y selección del modelo final. (30 h)
3. Recopilación de imágenes para los datasets (90 h)
 - 3.1. Búsqueda de datasets de terceros. (20 h)
 - 3.2. Captura de imágenes. (30 h)
 - 3.3. Preparación del dataset. (25 h)
 - 3.4. Refinamiento y selección final. (15 h)
4. Entrenamiento de la red neuronal (130 h)
 - 4.1. Desarrollo del modelo. (50 h)
 - 4.2. Entrenamiento del modelo. (40 h)
 - 4.3. Ajuste de hiperparámetros. (30 h)
 - 4.4. Refinamiento del dataset. (10 h)
5. Desarrollo del sistema global (120 h)
 - 5.1. Desarrollo de una interfaz de comandos de alto nivel. (40 h)
 - 5.2. Desarrollo del sistema de registro en base de datos. (20 h)
 - 5.3. Evaluación de desempeño del sistema. (20 h)
 - 5.4. Desarrollo de demos. (40 h)

6. Generación de entregables y proceso de cierre (100 h)

- 6.1. Inicio elaboración memoria técnica - Taller de Trabajo Final A. (40 h)
- 6.2. Revisión y correcciones de la memoria. (5 h)
- 6.3. Fin de elaboración memoria técnica - Taller de Trabajo Final B. (40 h)
- 6.4. Revisión y correcciones de la memoria. (5 h)
- 6.5. Elaboración presentación final. (8 h)
- 6.6. Revisión y correcciones de la presentación final. (2 h)

Cantidad total de horas: 600 horas.

10. Diagrama de Activity On Node

En la figura 2, se muestra el diagrama Activity on Node del proyecto. Los tiempos de las tareas están expresados en horas.

Si bien se observa alguna paralelización de tareas, estas se realizarán de forma secuencial o intercalada debido a que solo habrá una persona trabajando en ellas.

En el diagrama se indica con un mayor grosor de flecha el camino crítico.

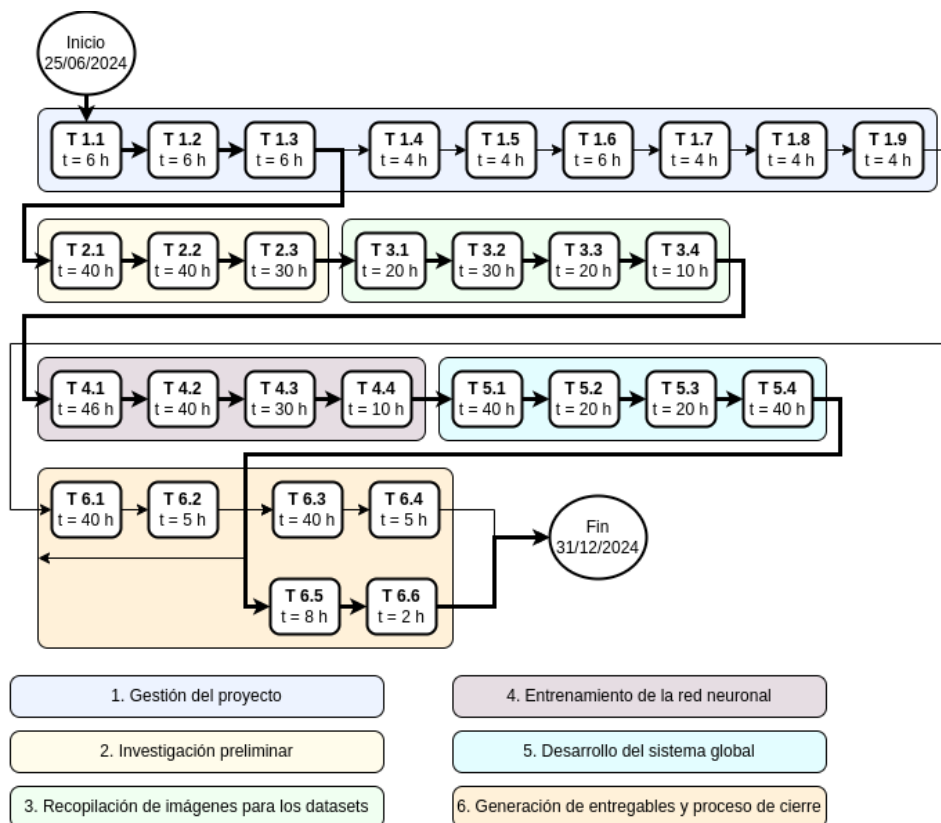


Figura 2. Diagrama de *Activity on Node*.

Duración del camino crítico: 484 h.

11. Diagrama de Gantt

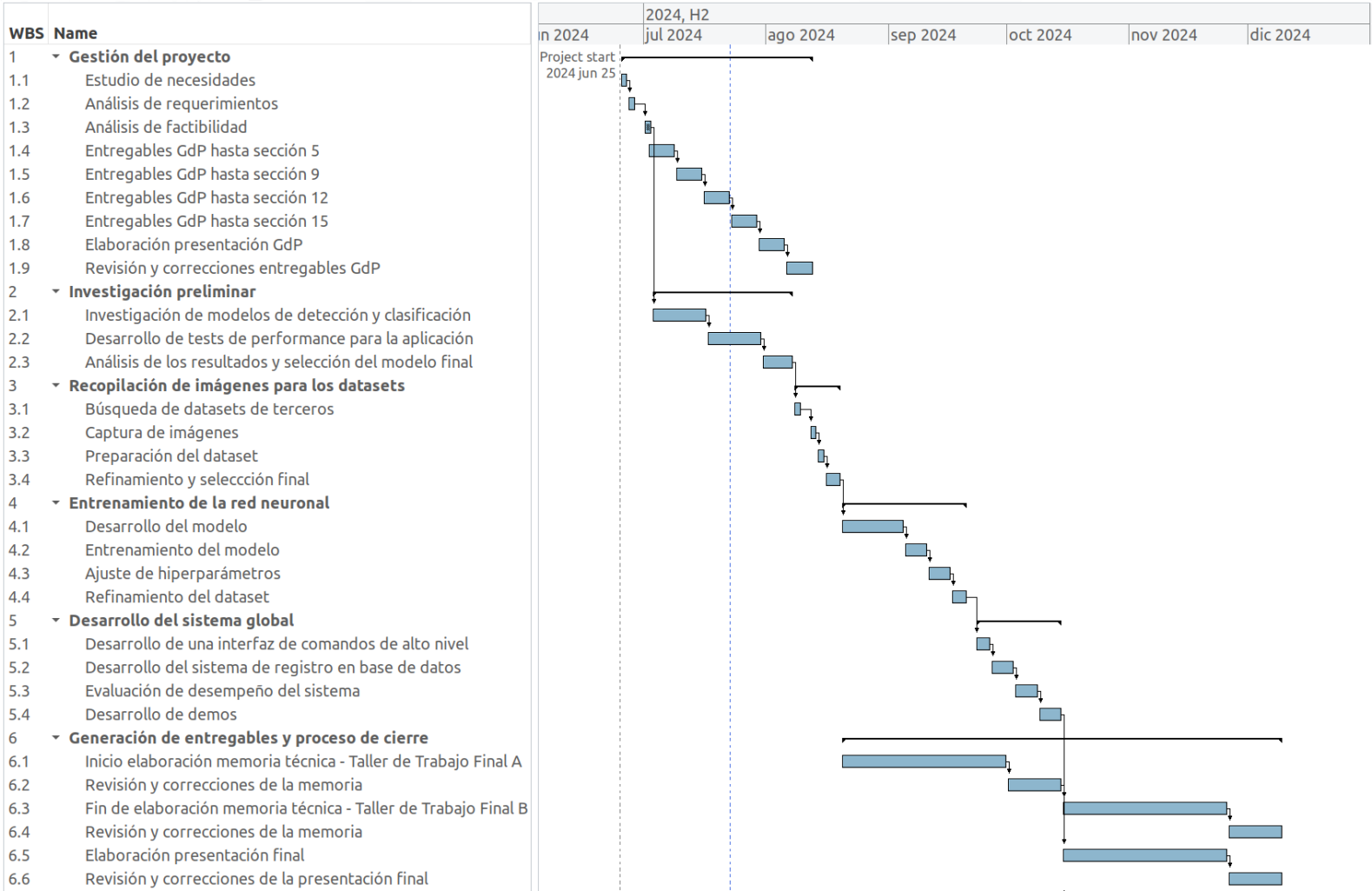


Figura 3. Diagrama de Gantt.

12. Presupuesto detallado del proyecto

A continuación se presenta el presupuesto detallado del proyecto expresado en dólares estadounidenses:

COSTOS DIRECTOS			
Descripción	Cantidad	Valor unitario	Valor total
Horas de ingeniería	580	30	17400
Cómputo - Google cloud	6	45	270
SUBTOTAL			17670
COSTOS INDIRECTOS			
Descripción	Cantidad	Valor unitario	Valor total
Viáticos	8	12	96
30 % del costo directo	-	-	501
SUBTOTAL			597
TOTAL			18267

Al día de la fecha, 28 de Julio de 2024, el costo total es de \$17353650 (pesos argentinos).

13. Gestión de riesgos

En la identificación de los riesgos se utiliza una escala de 1 a 10 para cuantificar la probabilidad de ocurrencia y severidad. Se considera 1 como la mínima probabilidad y 10 la máxima.

Riesgo 1: registro de imágenes insuficientes para el entrenamiento del modelo.

- Severidad (S): 9. El rendimiento del modelo entrenado depende directamente de la calidad del dataset recopilado.
- Probabilidad de ocurrencia (O): 6. La recopilación de imágenes en campo puede ser dificultosa y de baja calidad.

Riesgo 2: pobre resolución de las imágenes.

- Severidad (S): 8. No solo es necesario una cantidad significativa de imágenes si no que también es importante la calidad de las mismas. Los modelos no se desempeñarían correctamente ocasionando un mal funcionamiento global del sistema.
- Probabilidad de ocurrencia (S): 3. La cámara sería seleccionada teniendo en cuenta los requerimientos y un margen de seguridad adecuado que cubra cualquier eventual disminución de resolución causada por el encargado de tomar las imágenes y el preprocesamiento de las mismas.

Riesgo 3: mala precisión de los modelos entrenados.

- Severidad (S): 8. Los malos resultados del modelo en la etapa de prueba pueden requerir mayores tiempos de entrenamiento e impactar en la viabilidad técnica del proyecto.

- Probabilidad de ocurrencia (O): 4. Los tiempos de entrenamiento son muy extensos y muchas veces son una variable de ajuste necesaria.

Riesgo 4: no contar con suficiente capacidad de cómputo.

- Severidad (S): 8. Trabajar con procesamiento de imágenes implica gran consumo de recursos de hardware. No contar con la capacidad de cómputo adecuada implicaría que las tareas se hagan más largas de lo estimado o que se requiera contratar servicios de cómputo más costoso.
- Probabilidad de ocurrencia (O): 3. El responsable posee un equipo adecuado y se contrata un servicio de cómputo en la nube que en principio es suficiente para el desarrollo.

Riesgo 5: falta de tiempo y/o recursos humanos para el desarrollo.

- Severidad (S): 7. Se atrasarían las tareas, no se cumpliría con la planificación y quizás no se lograría completar el proyecto para la fecha de presentación establecida.
- Probabilidad de ocurrencia (O): 4. El proceso de etiquetado del dataset suele ser muy extenso y tedioso por lo que la estimación de tiempo para esa etapa podría ser errónea. La elección del modelo y el ajuste de hiperparámetros es también una tarea que podría variar a lo estimado por la complejidad de su desarrollo.

b) Tabla de gestión de riesgos: (El RPN se calcula como $RPN = S \times O$)

Riesgo	S	O	RPN	S*	O*	RPN*
Registro de imágenes insuficientes para el entrenamiento del modelo	9	6	54	5	6	30
Pobre resolución de las imágenes	8	3	24	-	-	-
Mala precisión de los modelos entrenados	8	4	32	6	3	18
No contar con suficiente capacidad de cómputo	8	3	24	-	-	-
Falta de tiempo y/o recursos humanos para el desarrollo	7	5	35	7	4	28

Criterio adoptado: se tomarán medidas de mitigación en los riesgos cuyos números de RPN sean mayores a 30.

Plan de mitigación de los riesgos que originalmente excedían el RPN máximo establecido:

Riesgo 1 - Datos insuficientes. Plan de mitigación: lograr el mayor avance posible con datos suplementarios. Existen fuentes de datos públicos que se pueden adaptar a esta problemática, el durazno es una fruta muy común.

- Severidad (S): 5. Es posible favorecer el rendimiento de los modelos con técnicas como *transfer learning* o *one-shot learning* con la variabilidad de datos proporcionado con datasets públicos, aunque no sean específicamente de lo especificado en la problemática a solucionar. De todos modos no es el escenario más favorable.
- Probabilidad de ocurrencia (O): 6. La probabilidad de ocurrencia no se modifica.

Nueva asignación de S y O: 30.

Riesgo 3. Bajo desempeño de modelos. Plan de mitigación: Profundizar en las arquitecturas y fundamentos matemáticos de cada modelo para poder hacer ajustes y correcciones. Usar otros trabajos como referencia para incorporar mejoras.

- Severidad (S): 6. Realizar modificaciones sobre arquitecturas conocidas complejiza la solución e implica un esfuerzo adicional (puede ser más difícil convertir formatos y usar herramientas estándar, por ejemplo), pero compromete en menor grado los objetivos principales.
- Probabilidad de ocurrencia (O): 3. Una buena comprensión de los algoritmos permite diagnosticar problemas y encontrar estrategias para mitigarlos.

Nueva asignación de S y O: 18.

Riesgo 5: Falta de tiempo y/o recursos humanos para el desarrollo. Plan de mitigación: se incluyó un tiempo de contingencia en la planificación de las tareas que conforman el camino crítico, de manera de mitigar posibles demoras causadas por imprevistos como enfermedades, viajes o mudanzas por trabajo, etc.

- Severidad (S): 7 la severidad se mantiene constante.
- Probabilidad de ocurrencia (O): 4 el plan de mitigación reduce la probabilidad de que las demoras eventuales afecten el desarrollo del proyecto.

Nueva asignación de S y O: 28.

14. Gestión de la calidad

- Req #1.1: el sistema debe ser capaz de identificar y cuantificar frutos tanto en estado inmaduro o maduro a partir de fotos de árboles tomadas en la plantación.
 - Verificación: se ensayará con un set de imágenes con las casuísticas que se pueden presentar. Por ejemplo, árboles con todos frutos inmaduros, con todos maduros, etc. y se medirá la performance del sistema.
 - Validación: el cliente procesará un conjunto nuevo de imágenes con la solución desarrollada. Se inspeccionará la salida obtenida para todos los casos posibles.
- Req #1.3: el sistema deberá indicar el grado de confianza de la detección y el error esperado en las mediciones.
 - Verificación: se ejecutará el sistema con el dataset de test y se confirmará que el resultado sea el esperado.
 - Validación: el cliente procesará un conjunto nuevo de imágenes y se medirá que el resultado esté en el grado de confianza y error esperado que se hayan configurado.
- Req #1.5: los dataset y código en repositorios deben mantenerse NO abiertos, de modo de mantener la confidencialidad del trabajo.
 - Verificación: se controlarán las configuraciones de seguridad del repositorio y de las unidades de almacenamiento compartido donde se haya guardado la información del trabajo.

- Validación: el cliente comprobará que tiene acceso al repositorio y podrá validar con un tercero ajeno al proyecto que el ingreso se encuentra bloqueado para este.
- Req #1.6: el usuario debe poder analizar una foto individual o un grupo de fotos, por lo menos grupos de 20 imágenes.
 - Verificación: se verificará el funcionamiento del sistema utilizando el dataset de test en grupos de diferentes tamaños.
 - Validación: el cliente procesará un conjunto nuevo de imágenes en grupos de diferentes tamaños.
- Req #1.7: el usuario debe ser capaz de cambiar el nivel de confianza para modificar la detección.
 - Verificación: se verificará el funcionamiento del sistema utilizando el dataset de test con diferentes valores de nivel de confianza.
 - Validación: el cliente procesará un conjunto nuevo de imágenes con diferentes valores de nivel de confianza.
- Req #2.2. se deben describir los ensayos y condiciones para la reproducción de los mismos.
 - Verificación: los ensayos con el dataset de test se diseñarán y documentarán contemplando diferentes escenarios, se verificará que se puedan reproducir.
 - Validación: el cliente procesará el dataset de test siguiendo la documentación de los ensayos para verificar su reproducción.
- Req #3.1: el modelo debe ser entrenado a partir de modelos pre-entrenados utilizando la técnica de *transfer learning*.
 - Verificación: análisis de las notebooks de validación de los algoritmos, el código fuente implementado y los datasets utilizados durante el entrenamiento de los modelos.
 - Validación: verificación de métricas superiores a los umbrales mínimos definidos.
- Req #4.1: la interfaz debe ser intuitiva, con capacidad de cargar las fotos de manera sencilla.
 - Verificación: se implementará la interfaz de manera que la interacción con el usuario sea mínima pero conteniendo los diferentes comandos y tareas definidas, se ensayará cada uno de ellos en los diferentes escenarios.
 - Validación: se realizará una demostración global del sistema al cliente donde él pueda verificar la funcionalidad de la interfaz.

15. Procesos de cierre

Finalizado el proyecto, el proceso de cierre contemplará las siguientes actividades:

- Reunión de evaluación final:
 - Será virtual, contará con la participación de todos los interesados del proyecto. En la agenda se incluirán las siguientes actividades:
 - Se realizará el contraste entre el plan y la ejecución.
 - Revisión general del informe, ensayos y resultados.
 - Demostración del repositorio con la documentación y código. Se repasará la instalación/ejecución del sistema.
 - Repaso de las recomendaciones para la captura de imágenes para el uso con el sistema.
 - Análisis de las lecciones aprendidas (discutir qué problemas surgieron durante el proyecto y qué soluciones se aplicaron).
 - Se proporcionará un espacio para que los interesados puedan realizar consultas y evacuar dudas sobre los temas discutidos o los entregables.
- Defensa pública del proyecto: se realizará en modalidad virtual y participarán los interesados, jurados y personal docente. Al final de la presentación se agradecerá formalmente a todos aquellos que colaboraron durante el proyecto.