



Identificación de estados fenológicos de la flor de durazneros mediante visión por computadora

Autor:

Ing. Héctor Luis Sánchez Márquez

Director:

Nombre del Director (pertenencia)

Codirector:

Maximiliano Aballay (INTA)

*Esta planificación fue realizada en el curso de Gestión de proyectos
entre el 22 de agosto de 2023 y el 10 de octubre de 2023.*

Índice

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

Registros de cambios

Revisión	Detalles de los cambios realizados	Fecha
0	Creación del documento	22 de agosto de 2023
1	Se completa hasta el punto 5 inclusive	5 de septiembre de 2023
2	Aplicación de las correcciones de la revisión 1. Se completa hasta el punto 9 inclusive	12 de septiembre 2023

Acta de constitución del proyecto

Buenos Aires, 22 de agosto de 2023

Por medio de la presente se acuerda con el Ing. Héctor Luis Sánchez Márquez que su Trabajo Final de la Carrera de Especialización en Inteligencia Artificial se titulará “Identificación de estados fenológicos de la flor de durazneros mediante visión por computadora”, consistirá esencialmente en desarrollar un algoritmo que identifique flores y su estado a partir de fotos de varetas, y tendrá un presupuesto preliminar estimado de 600 horas de trabajo y **\$XXX**, con fecha de inicio 22 de agosto de 2023 y fecha de presentación pública TBD.

Se adjunta a esta acta la planificación inicial.

Dr. Ing. Ariel Lutenberg
Director posgrado FIUBA

Dr. Gerardo Sánchez
Instituto Nacional de Tecnología Agropecuaria (INTA)

Nombre del Director
Director del Trabajo Final

1. Descripción técnica-conceptual del proyecto a realizar

El Instituto Nacional de Tecnología Agropecuaria (INTA) es un organismo estatal descentralizado, con independencia operativa y financiera, que se encuentra adscrito a la Secretaría de Agricultura, Ganadería y Pesca del Ministerio de Economía de la Nación. Este ente nació en 1956 con el objetivo de impulsar la innovación y la transferencia de conocimientos en los sectores agroalimentario, agroindustrial y agropecuario a través de la investigación. Sus aportes permiten potenciar el país y generar nuevas oportunidades para acceder a mercados regionales e internacionales con productos y servicios de alto valor agregado.

1.1 Introducción general

La fenómica hace referencia a la obtención de un gran caudal de datos de las características de las plantas, lo que se denomina el fenotipo de la planta. Esta disciplina está en auge en la actualidad debido a sus aplicaciones potenciales. Por un lado, habilita el mejoramiento a gran escala debido a que es necesario vincular una gran cantidad de datos genéticos con datos fenotípicos para identificar la función de los genes. Por otro lado, si se incluyen otros conjuntos de datos como son los climáticos, permite realizar predicciones precisas sobre el comportamiento de las variedades, el cual es necesario para implementar lo que se conoce como agricultura de precisión. Sin embargo, la fruticultura no ha dado el salto hacia la fenómica.

En la Estación Experimental Agropecuaria (EEA) de San Pedro se ha logrado secuenciar el ADN de más de 250 variedades de duraznero (Aballay et al., 2021, Scientific Reports) disponiendo de una base de datos genómica de 75 gigabases (Gb) de ADN. Esta base permite identificar genes que controlan características del duraznero mediante algoritmos de inteligencia artificial (IA). Además, se dispone de datos climáticos diarios que se toman de forma automática que incluyen: las temperaturas medias, precipitaciones, horas de frío, radiación, etc. Esta información se combina con los datos genómicos y posteriormente, con modelos de IA se predice el comportamiento de las variedades en escenarios climáticos futuros.

Por otro lado, las heladas primaverales son actualmente el mayor problema de los frutales a nivel mundial. Este fenómeno ocurre cuando las flores abiertas (estado "F") se someten a temperaturas cercanas a los -2.5°C . Por este motivo, es necesario conocer el número de flores que se encuentran en estado vulnerable ante un pronóstico de heladas, así como también la densidad de flores. Es del interés del INTA determinar el estado fenológico a campo y mejorar para la tolerancia a heladas. Para lograrlo, se empleará IA y visión por computadora.

La visión por computadora es un subdominio de la IA que permite a las máquinas imitar el sistema visual del ser humano. De esta forma, es posible extraer información a partir de imágenes.

En la actualidad existen algoritmos capaces de detectar y clasificar de forma efectiva plantas a través de imágenes. Estos algoritmos se han utilizado para una gran variedad de aplicaciones, como es el caso de reconocimiento de enfermedades en plantas. Sin embargo, aún no se ha desarrollado un sistema para la extracción de estados fenológicos de la flor de durazneros.

La presente propuesta permitirá automatizar la toma de características de la flor de durazneros, a partir de fotos. De esta forma, se logrará aumentar el caudal de datos y mejorar los modelos de IA existentes.

1.2 Marco de la propuesta

La ejecución de este proyecto va alineada con el interés del INTA de determinar el estado de la flor de los durazneros a partir de imágenes.

Las especificaciones de las fotos de varetas que se disponen se detallan en la tabla 1. Por otro lado, las características a determinar se indican en la tabla 2. Se deberá evaluar diferentes modelos de IA a fin de definir el óptimo para esta tarea. Se deberá considerar la existencia de bases de datos de imágenes y/o modelos preentrenados disponibles.

Características de las fotos de duraznos			
Formato	Número	Tamaño	Observaciones
JPG	250	1 MB	Fotos de varetas de duraznero con flores en diferentes estados fenológicos. La mayoría está en estado “F” pero también están en estado “E” y “G”. Cada foto tiene una regla. Cada foto tiene entre 10 a 12 varetas.

Tabla 1

Características de interés a ser determinadas	
N° de Flores totales	Presencia de Flor.
Tipo de flor	En las fotos existen dos tipos de flores: campanulácea y rosásea.
Estado fenológico	Discriminar entre estado “F” o flor abierta respecto a los demás estados (“E” flor cerrada o “G” flor sin pétalos).
Densidad de flores	Determinar la cantidad de flores por cm de vareta.

Tabla 2

La solución que se propone hará uso de un modelo de detección de objetos óptimo para el problema planteado. Además, para la clasificación de flores, se utilizará un modelo preentrenado que se ajustará a las fotos proporcionadas. Como resultado final, se cumplirán todos los requisitos establecidos en la tabla 2. La figura 1 muestra esta solución a alto nivel.

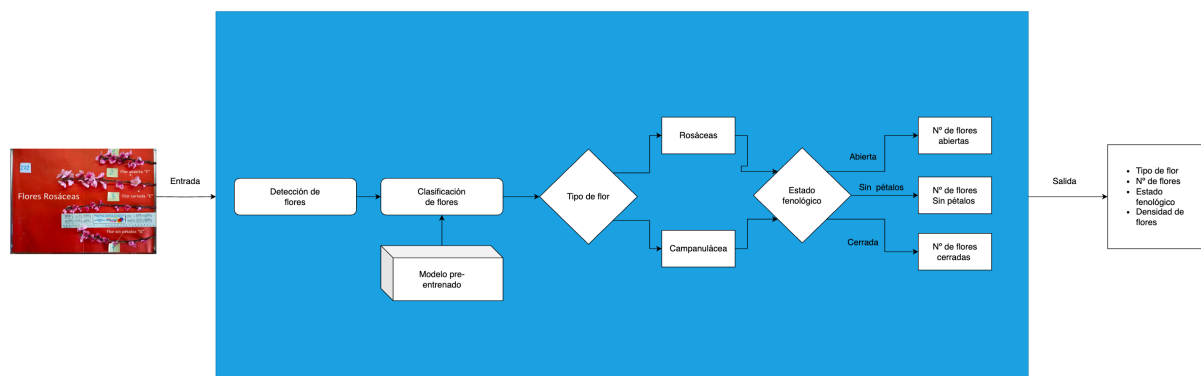


Figura 1. Diagrama en bloques del sistema.

2. Identificación y análisis de los interesados

Rol	Nombre y Apellido	Organización	Puesto
Auspiciante	Dr. Gerardo Sánchez	Instituto Nacional de Tecnología Agropecuaria (INTA)	-
Cliente	Dr. Gerardo Sánchez	Instituto Nacional de Tecnología Agropecuaria (INTA)	-
Impulsor	Dr. Gerardo Sánchez	Instituto Nacional de Tecnología Agropecuaria (INTA)	-
Responsable	Ing. Héctor Luis Sánchez Márquez	FIUBA	Alumno
Colaboradores	Maximiliano Aballay	Instituto Nacional de Tecnología Agropecuaria (INTA)	-
Orientador	Nombre del Director	pertenencia	Director Trabajo final
Equipo	Maximiliano Aballay	Instituto Nacional de Tecnología Agropecuaria (INTA)	-
Usuario final	INTA	INTA	-

3. Propósito del proyecto

El propósito de este proyecto es desarrollar un algoritmo que identifique flores de los durazneros y sus estados a partir de fotos de varetas, para aumentar el caudal de datos y mejorar los modelos de IA existentes.

4. Alcance del proyecto

El proyecto incluye:

- El preprocesamiento de las fotos para entrenar el modelo.
- La selección del modelo a entrenar.
- La elaboración del notebook de pruebas en Python.
- La implementación local del modelo.

El proyecto no incluye:

- La recolección de datos/fotos.
- La integración con otros modelos que utilice el cliente.

5. Supuestos del proyecto

Para el desarrollo del presente proyecto se supone que:

- Se contará con suficientes fotos de veretas para entrenar y evaluar el modelo.
- Se contará con soporte para temas no referentes al área de inteligencia artificial que se necesiten para la elaboración de este proyecto.
- Se tendrán imágenes con la calidad adecuada para resolver el problema planteado.
- Se contará con los recursos de hardware necesarios para el entrenamiento del modelo.
- El código se desarrollará en una notebook de Google Colab.

6. Requerimientos

Los requerimientos deben numerarse y de ser posible estar agruparlos por afinidad, por ejemplo:

1. Requerimientos funcionales
 - 1.1. El sistema debe...
 - 1.2. Tal componente debe...
 - 1.3. El usuario debe poder...
2. Requerimientos de documentación
 - 2.1. Requerimiento 1
 - 2.2. Requerimiento 2 (prioridad menor)
3. Requerimiento de testing...
4. Requerimientos de la interfaz...
5. Requerimientos interoperabilidad...
6. etc...

Leyendo los requerimientos se debe poder interpretar cómo será el proyecto y su funcionalidad.

Indicar claramente cuál es la prioridad entre los distintos requerimientos y si hay requerimientos opcionales.

No olvidarse de que los requerimientos incluyen a las regulaciones y normas vigentes!!!

Y al escribirlos seguir las siguientes reglas:

- Ser breve y conciso (nadie lee cosas largas).
- Ser específico: no dejar lugar a confusiones.
- Expresar los requerimientos en términos que sean cuantificables y medibles.

7. Historias de usuarios (*Product backlog*)

Descripción: En esta sección se deben incluir las historias de usuarios y su ponderación (*history points*). Recordar que las historias de usuarios son descripciones cortas y simples de una característica contada desde la perspectiva de la persona que desea la nueva capacidad, generalmente un usuario o cliente del sistema. La ponderación es un número entero que representa el tamaño de la historia comparada con otras historias de similar tipo.

El formato propuesto es: como [rol] quiero [tal cosa] para [tal otra cosa].”

Se debe indicar explícitamente el criterio para calcular los *story points* de cada historia

8. Entregables principales del proyecto

Los entregables del proyecto son (ejemplo):

- Manual de uso
- Diagrama de circuitos esquemáticos
- Código fuente del firmware
- Diagrama de instalación
- Informe final
- etc...

9. Desglose del trabajo en tareas

El WBS debe tener relación directa o indirecta con los requerimientos. Son todas las actividades que se harán en el proyecto para dar cumplimiento a los requerimientos. Se recomienda mostrar el WBS mediante una lista indexada:

1. Grupo de tareas 1
 - 1.1. Tarea 1 (tantas h)
 - 1.2. Tarea 2 (tantas hs)
 - 1.3. Tarea 3 (tantas h)
2. Grupo de tareas 2
 - 2.1. Tarea 1 (tantas h)
 - 2.2. Tarea 2 (tantas h)
 - 2.3. Tarea 3 (tantas h)

3. Grupo de tareas 3

- 3.1. Tarea 1 (tantas h)
- 3.2. Tarea 2 (tantas h)
- 3.3. Tarea 3 (tantas h)
- 3.4. Tarea 4 (tantas h)
- 3.5. Tarea 5 (tantas h)

Cantidad total de horas: (tantas h)

Se recomienda que no haya ninguna tarea que lleve más de 40 h.

10. Diagrama de Activity On Node

Armar el AoN a partir del WBS definido en la etapa anterior.

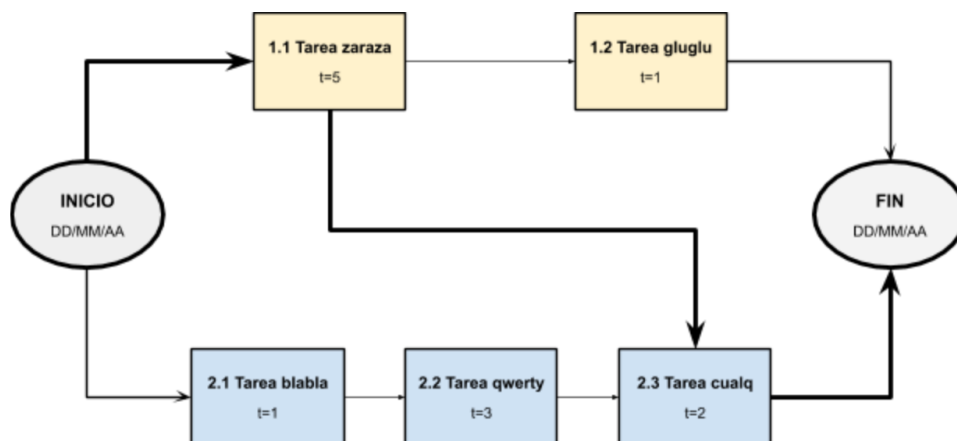


Figura 2. Diagrama de *Activity on Node*.

Indicar claramente en qué unidades están expresados los tiempos. De ser necesario indicar los caminos semicríticos y analizar sus tiempos mediante un cuadro. Es recomendable usar colores y un cuadro indicativo describiendo qué representa cada color, como se muestra en el siguiente ejemplo:

11. Diagrama de Gantt

Existen muchos programas y recursos *online* para hacer diagramas de Gantt, entre los cuales destacamos:

- Planner
- GanttProject

- Trello + *plugins*. En el siguiente link hay un tutorial oficial:
<https://blog.trello.com/es/diagrama-de-gantt-de-un-proyecto>
- Creately, herramienta online colaborativa.
<https://creately.com/diagram/example/ieb3p3ml/LaTeX>
- Se puede hacer en latex con el paquete *pgfgantt*
<http://ctan.dcc.uchile.cl/graphics/pgf/contrib/pgfgantt/pgfgantt.pdf>

Pegar acá una captura de pantalla del diagrama de Gantt, cuidando que la letra sea suficientemente grande como para ser legible. Si el diagrama queda demasiado ancho, se puede pegar primero la “tabla” del Gantt y luego pegar la parte del diagrama de barras del diagrama de Gantt.

Configurar el software para que en la parte de la tabla muestre los códigos del EDT (WBS).
Configurar el software para que al lado de cada barra muestre el nombre de cada tarea.
Revisar que la fecha de finalización coincida con lo indicado en el Acta Constitutiva.

En la figura 3, se muestra un ejemplo de diagrama de Gantt realizado con el paquete de *pgfgantt*. En la plantilla pueden ver el código que lo genera y usarlo de base para construir el propio.

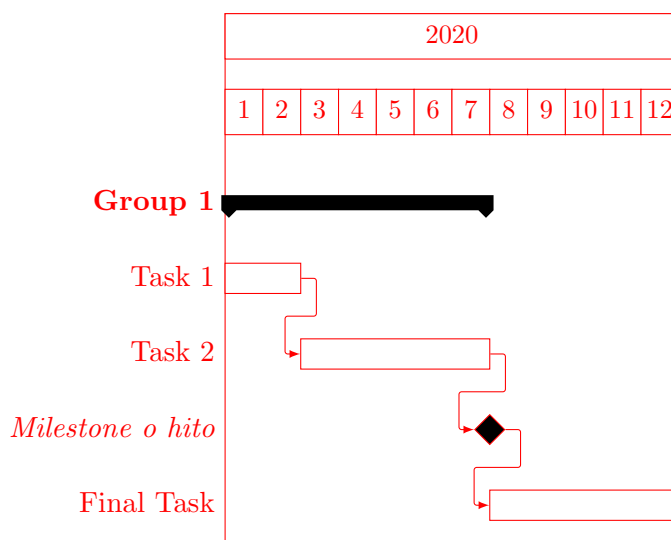


Figura 3. Diagrama de Gantt de ejemplo



Figura 4. Ejemplo de diagrama de Gantt rotado

12. Presupuesto detallado del proyecto

Si el proyecto es complejo entonces separarlo en partes:

- Un total global, indicando el subtotal acumulado por cada una de las áreas.
- El desglose detallado del subtotal de cada una de las áreas.

IMPORTANTE: No olvidarse de considerar los **COSTOS INDIRECTOS**.

COSTOS DIRECTOS			
Descripción	Cantidad	Valor unitario	Valor total
SUBTOTAL			
COSTOS INDIRECTOS			
Descripción	Cantidad	Valor unitario	Valor total
SUBTOTAL			
TOTAL			

13. Gestión de riesgos

a) Identificación de los riesgos (al menos cinco) y estimación de sus consecuencias:

Riesgo 1: detallar el riesgo (riesgo es algo que si ocurre altera los planes previstos de forma negativa)

- Severidad (S): mientras más severo, más alto es el número (usar números del 1 al 10). Justificar el motivo por el cual se asigna determinado número de severidad (S).
- Probabilidad de ocurrencia (O): mientras más probable, más alto es el número (usar del 1 al 10). Justificar el motivo por el cual se asigna determinado número de (O).

Riesgo 2:

- Severidad (S):
- Ocurrencia (O):

Riesgo 3:

- Severidad (S):

■ Ocurrecia (O):

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

Riesgo	S	O	RPN	S*	O*	RPN*

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

Nota: los valores marcados con (*) en la tabla corresponden luego de haber aplicado la mitigación.

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

Riesgo 1: plan de mitigación (si por el RPN fuera necesario elaborar un plan de mitigación). Nueva asignación de S y O, con su respectiva justificación: - Severidad (S): mientras más severo, más alto es el número (usar números del 1 al 10). Justificar el motivo por el cual se asigna determinado número de severidad (S). - Probabilidad de ocurrencia (O): mientras más probable, más alto es el número (usar del 1 al 10). Justificar el motivo por el cual se asigna determinado número de (O).

Riesgo 2: plan de mitigación (si por el RPN fuera necesario elaborar un plan de mitigación).

Riesgo 3: plan de mitigación (si por el RPN fuera necesario elaborar un plan de mitigación).

14. Gestión de la calidad

Elija al menos diez requerimientos que a su criterio sean los más importantes/críticos/que aportan más valor y para cada uno de ellos indique las acciones de verificación y validación que permitan asegurar su cumplimiento.

- Req #1: copiar acá el requerimiento.
 - Verificación para confirmar si se cumplió con lo requerido antes de mostrar el sistema al cliente. Detallar
 - Validación con el cliente para confirmar que está de acuerdo en que se cumplió con lo requerido. Detallar

Tener en cuenta que en este contexto se pueden mencionar simulaciones, cálculos, revisión de hojas de datos, consulta con expertos, mediciones, etc. Las acciones de verificación suelen considerar al entregable como “caja blanca”, es decir se conoce en profundidad su funcionamiento interno. En cambio, las acciones de validación suelen considerar al entregable como “caja negra”, es decir, que no se conocen los detalles de su funcionamiento interno.

15. Procesos de cierre

Establecer las pautas de trabajo para realizar una reunión final de evaluación del proyecto, tal que contemple las siguientes actividades:

- Pautas de trabajo que se seguirán para analizar si se respetó el Plan de Proyecto original:
- Indicar quién se ocupará de hacer esto y cuál será el procedimiento a aplicar.
- Identificación de las técnicas y procedimientos útiles e inútiles que se emplearon, y los problemas que surgieron y cómo se solucionaron: - Indicar quién se ocupará de hacer esto y cuál será el procedimiento para dejar registro.
- Indicar quién organizará el acto de agradecimiento a todos los interesados, y en especial al equipo de trabajo y colaboradores: - Indicar esto y quién financiará los gastos correspondientes.