



**UNIVERSIDAD NACIONAL DE CHIMBORAZO
FACULTAD INGENIERIA
CARRERA ELECTRÓNICA Y TELECOMUNICACIONES**

**DETECCIÓN TEMPRANA DE MINADOR, MOSCA BLANCA Y
FUSARIUM EN EL TOMATE RIÑÓN, APLICANDO TÉCNICAS
DE VISIÓN ARTIFICIAL Y MACHINE LEARNING.**

**Trabajo de Titulación para optar al título de Ingeniero en
Electrónica y Telecomunicaciones**

Autor:

Jácome García Jaime Fernando

Tutor:

Mcs. José Luis Jinez Tapia.

Riobamba, Ecuador. 2022

DERECHOS DE AUTORÍA

Yo, Jaime Fernando Jácome García, con cédula de ciudadanía 060350201-4, autor (a) (s) del trabajo de investigación titulado: Detección temprana de minador, mosca blanca y fusarium en el tomate riñón, aplicando técnicas de visión artificial y machine learning., certifico que la producción, ideas, opiniones, criterios, contenidos y conclusiones expuestas son de mí exclusiva responsabilidad.

Asimismo, cedo a la Universidad Nacional de Chimborazo, en forma no exclusiva, los derechos para su uso, comunicación pública, distribución, divulgación y/o reproducción total o parcial, por medio físico o digital; en esta cesión se entiende que el cesionario no podrá obtener beneficios económicos. La posible reclamación de terceros respecto de los derechos de autor (a) de la obra referida, será de mi entera responsabilidad; librando a la Universidad Nacional de Chimborazo de posibles obligaciones.

En Riobamba, 16 de mayo del 2022.



Jaime Fernando Jácome García
C.I:060350201-4

DICTAMEN FAVORABLE DEL TUTOR Y MIEMBROS DE TRIBUNAL

Quienes suscribimos, catedráticos designados Tutor y Miembros del Tribunal de Grado para la evaluación del trabajo de investigación Detección temprana de minador, mosca blanca y fusarium en el tomate riñón, aplicando técnicas de visión artificial y machine learning., presentado por Jaime Fernando Jácome García, con cédula de identidad número 060350201-4, certificamos que recomendamos la APROBACIÓN de este con fines de titulación. Previamente se ha asesorado durante el desarrollo, revisado y evaluado el trabajo de investigación escrito y escuchada la sustentación por parte de su autor; no teniendo más nada que observar.

De conformidad a la normativa aplicable firmamos, en Riobamba 16 de mayo del 2022

PhD. Luis Tello Oquendo
PRESIDENTE DEL TRIBUNAL DE
GRADO

Firma

PhD. Marlon Basantes Valverde,
MIEMBRO DEL TRIBUNAL DE GRADO

Firma

Mgs. Edgar Giovanni Cuzco Silva
MIEMBRO DEL TRIBUNAL DE GRADO

Firma

Mcs. José Luis Jinez Tapia.
TUTOR

Firma

Jaime Fernando Jácome García
C.I.:060350201-4

CERTIFICADO DE LOS MIEMBROS DEL TRIBUNAL

Quienes suscribimos, catedráticos designados Miembros del Tribunal de Grado para la evaluación del trabajo de investigación Detección temprana de minador, mosca blanca y fusarium en el tomate riñón, aplicando técnicas de visión artificial y machine learning., presentado por Jaime Fernando Jácome García, con cédula de identidad número 060350201-4, bajo la tutoría de MEng. José Luis Jinez Tapia.; certificamos que recomendamos la APROBACIÓN de este con fines de titulación. Previamente se ha evaluado el trabajo de investigación y escuchada la sustentación por parte de su autor; no teniendo más nada que observar.

De conformidad a la normativa aplicable firmamos, en Riobamba 16 de mayo del 2022

Presidente del Tribunal de Grado
PhD. Luis Tello Oquendo



Firma

Miembro del Tribunal de Grado
PhD. Marlon Basantes Valverde,



Firma

Miembro del Tribunal de Grado
Mgs. Edgar Giovanny Cuzco Silva



Firma



CERTIFICADO ANTIPLAGIO

CERTIFICACIÓN

Que, **JAIME FERNANDO JÁCOME GARCÍA** con CC: **06033502014**, estudiante de la Carrera **ELECTRONICA Y TELECOMUNICACIONES, NO VIGENTE**, Facultad de **INGENIERIA**; ha trabajado bajo mi tutoría el trabajo de investigación titulado **“DETECCIÓN TEMPRANA DE MINADOR, MOSCA BLANCA Y FUSARIUM EN EL TOMATE RIÑÓN, APLICANDO TÉCNICAS DE VISIÓN ARTIFICIAL Y MACHINE LEARNING”** cumple con el 0%, de acuerdo al reporte del sistema Anti plagio **Urkund**, porcentaje aceptado de acuerdo a la reglamentación institucional, por consiguiente autorizo continuar con el proceso.

Riobamba, 10 de mayo de 2022

Mcs. José Luis Jinez Tapia
TUTOR

DEDICATORIA

Dedico al que es mi fuerza cuando estoy débil, al que cuando estoy solo es mi apoyo, al que cuando estoy perdido es mi sendero por ello este trabajo es para Dios.

De igual forma, dedico este trabajo de titulación a mis padres que a pesar de la distancia siempre estuvieron conmigo fomentándome el respeto, humildad, responsabilidad y disciplina, admirando su fortaleza y gran amor por todo lo que han hecho por mí.

A mis hermanos quienes son un pilar fundamental en mi vida.

AGRADECIMIENTO

Agradezco a Dios por guiarme y darme la sabiduría para superar obstáculos y debilidades, Ivon y Patricio mis padres les agradezco por siempre apoyarme, confiar y darme las herramientas necesarias para lograr lo que me proponga, Cristian mi primo el cual es mi ejemplo y mi apoyo incondicional en la vida, y a todas las personas que me ayudaron a cumplir una meta más en mi vida profesional.

ÍNDICE GENERAL

| | |
|---------------------------------------------------------------------------------------------------------------------------------------|----|
| DERECHOS DE AUTORÍA..... | 2 |
| DICTAMEN FAVORABLE DEL TUTOR Y MIEMBROS DE TRIBUNAL | 3 |
| CERTIFICADO DE LOS MIEMBROS DEL TRIBUNAL | 4 |
| CERTIFICADO ANTIPLAGIO..... | 5 |
| DEDICATORIA | 6 |
| AGRADECIMIENTO..... | 7 |
| ÍNDICE GENERAL | 8 |
| ÍNDICE DE TABLAS. | 11 |
| ÍNDICE DE FIGURAS | 11 |
| ÍNDICE DE ECUACIONES | 13 |
| RESUMEN..... | 14 |
| ABSTRACT | 15 |
| 1. CAPÍTULO I. INTRODUCCIÓN..... | 16 |
| 1.1 Antecedentes..... | 16 |
| 1.2 Planteamiento del Problema..... | 18 |
| 1.3 Justificación..... | 20 |
| 1.4 Objetivos..... | 21 |
| 1.4.1 General | 21 |
| 1.4.2 Específicos | 21 |
| 2. CAPÍTULO II. MARCO TEÓRICO | 22 |
| 2.1 Estado del Arte..... | 22 |
| 2.2 Marco Teórico..... | 23 |
| 2.2.1 Industria 4.0..... | 23 |
| 2.2.2 Trabajo colaborativo asistido por ordenador | 24 |
| 2.2.3 Sistemas físico-cibernéticos | 26 |
| 2.2.4 IoT | 30 |
| 2.2.5 Arquitecturas del internet de las cosas (IIoT) | 31 |
| 2.2.6 Datos Masivos (Big Data) | 32 |
| 2.2.7 Computación en la nube (Cloud computing), Computación en la niebla (Fog computing), Computación de borde (Edge computing) | 34 |

| | | |
|--------|--------------------------------------------------------------------------------------------|----|
| 2.2.8 | Aprendizaje automático (Machine Learning) en la industria | 36 |
| 2.2.9 | Madurez digital en la industria. | 37 |
| 2.2.10 | Plagas y Enfermedades | 38 |
| 3. | CAPÍTULO III. METODOLOGIA | 41 |
| 3.1 | Tipo de Investigación | 41 |
| 3.2 | Diseño de Investigación | 41 |
| 3.3 | Población de estudio y tamaño de muestra | 42 |
| 3.3.1 | Datos de la población..... | 42 |
| 3.3.2 | Muestra..... | 42 |
| 3.3.3 | El nivel de confianza | 43 |
| 3.3.4 | La estimación de la proporción | 43 |
| 3.3.5 | El margen de error deseado..... | 43 |
| 3.4 | Técnicas de recolección de Datos | 44 |
| 3.5 | Métodos de análisis, y procesamiento de datos. | 47 |
| 3.5.1 | Clasificación de las imágenes por carpetas: | 47 |
| 3.5.2 | Datos en la nube:..... | 48 |
| 3.5.3 | Entrenamiento del modelo | 48 |
| 3.5.4 | Carga y ejecución del modelo en la tarjeta | 53 |
| 4. | CAPÍTULO IV. RESULTADOS Y DISCUSIÓN | 55 |
| 4.1 | Resultados de la Madurez Digital de la industria agrícola en la comunidad Puctus. 55 | |
| 4.1.1 | Estadística de fiabilidad de la encuesta de madurez digital de la industria agrícola 56 | |
| 4.2 | Implementación en el prototipo | 57 |
| 4.3 | Resultados de los entrenamientos | 57 |
| 4.4 | Curvas de Aprendizaje | 59 |
| 4.5 | Pruebas Iniciales..... | 61 |
| 4.6 | Resultados del Prototipo | 62 |
| 4.6.1 | Estadística de fiabilidad de la encuesta de validación de prototipo | 63 |
| 4.7 | Revalidación del prototipo por agricultores expertos. | 64 |
| 4.8 | Google Coral Dev Board con conexión a Internet | 66 |
| 4.9 | Arquitectura General..... | 67 |

| | | |
|--------|------------------------------------------------------------------------------------------|----|
| 4.10 | Arquitectura del prototipo en su funcionamiento | 67 |
| 4.11 | Propuesta de arquitectura IoT para una industria agrícola de producción de tomate. 68 | |
| 4.11.1 | Propuesta propia | 68 |
| 5. | CAPÍTULO V. CONCLUSIONES y RECOMENDACIONES..... | 72 |
| 5.1 | Conclusiones..... | 72 |
| 5.2 | Recomendaciones..... | 73 |
| 6. | BIBLIOGRAFÍA..... | 74 |
| 7. | ANEXOS | 79 |
| 7.1 | ENCUESTA SOBRE LA MADUREZ DIGITAL DE LA INDUSTRIA AGRÍCOLA. | 79 |
| 7.2 | ENCUESTA VALIDACIÓN DE PROTOTIPO | 83 |
| 7.3 | Fotografías de la realización de encuesta y pruebas de campo | 85 |
| 7.4 | Ejemplos de imágenes usadas en el entrenamiento. | 87 |
| 7.5 | Presupuesto de Prototipo | 95 |

ÍNDICE DE TABLAS.

| | |
|-------------------------------------------------------------------------------------------------------------------------|----|
| Tabla 1. Madurez Digital – Variables | 19 |
| Tabla 2. Variables..... | 20 |
| Tabla 3. Datos de la población. | 42 |
| Tabla 4. Comparación de Tarjetas de Desarrollo | 45 |
| Tabla 5. Estadística de escala | 56 |
| Tabla 6. Resumen de configuraciones y resultados de los entrenamientos más notables. . | 58 |
| Tabla 7. Datos del proceso de aprendizaje en el Entrenamiento del modelo seleccionado | 58 |
| Tabla 8. Datos del proceso de aprendizaje del ajuste del modelo de entrenamiento seleccionado..... | 59 |
| Tabla 9. Estadística de escala de la encuesta de validación de prototipo..... | 63 |
| Tabla 10. Matriz de Confusión..... | 64 |
| Tabla 11. Comparación de las Exactitudes en el modelo TensorFlow original, modelo TensorFlow Lite y prueba real..... | 66 |
| Tabla 12. Presupuesto de Prototipo | 95 |

ÍNDICE DE FIGURAS

| | |
|---------------------------------------------------------------------------------------------------------------------------|----|
| Figura 1. Etapas en el desarrollo de sistemas de fabricación industrial adaptado de [20]. . | 24 |
| Figura 2. Métodos, temas y herramientas en CSCW adaptado de[21], [24],[16]. | 25 |
| Figura 3. Marco de resumen de los patrones de adopción de Industria 4.0 y sus tecnologías [16]..... | 26 |
| Figura 4. Distribuciones geográficas de publicaciones ciberfísicas industriales [16]. | 27 |
| Figura 5. Arquitectura 5C para la implementación de CPS adaptado de [26]..... | 28 |
| Figura 6. Arquitectura CPS del proyecto de Navantia [27]..... | 29 |
| Figura 7. Diagrama de interacción entre CPS y MES del proyecto de Navantia [27]. | 29 |
| Figura 8. Tipos de IoT. | 30 |
| Figura 9. Comparación de 3 arquitecturas base de IoT divididas en 3 capas..... | 32 |
| Figura 10. Las 5V de Big Data adaptado de [32],[33]. | 33 |
| Figura 11. Big Data dentro de la industria 4.0..... | 34 |
| Figura 12. Cloud – Fog – Edge Computing en IIoT (adaptado de [40], [41]). | 36 |
| Figura 13. Clasificación de los algoritmos de Machine Learning (Adaptado de [45]). | 37 |
| Figura 14. a) Enfermedad Fusarium vista desde microscopio [50]. b) Enfermedad fusarium en el tallo de una planta. | 38 |
| Figura 15. Ciclo de vida de plaga Mosca Blanca [51]..... | 39 |
| Figura 16. a) Plaga minador en fruto b) Plaga minador en follaje | 40 |
| Figura 17. Representación del nivel de confianza. | 43 |
| Figura 18. Procesos de la primera etapa de entrenamiento. | 46 |
| Figura 19. Procesos de la etapa de implementación del modelo entrenado. | 46 |
| Figura 20. Imágenes clasificadas en carpetas por categorías. | 47 |
| Figura 21. Datos en Google Cloud Storage en formato para desarrolladores. | 48 |
| Figura 22. Importación de los archivos y ejecución del modelo. | 54 |
| Figura 23. Madurez digital de la industria agrícola de la comunidad Puctus..... | 55 |

| | |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Figura 24. Alfa de Cronbach para encuesta de Madurez digital de la industria agrícola. ... | 56 |
| Figura 25. Prototipo montado. | 57 |
| Figura 26. Curvas de modelos entrenados. | 61 |
| Figura 27. Prototipo montado. | 62 |
| Figura 28. a) Prototipo en prueba final de funcionamiento con una sola predicción. b) Prototipo en prueba final de funcionamiento con predicción de varias clases. c) Prototipo en prueba final en invernadero con predicción de varias clases d) Prototipo en prueba final en invernadero con una sola predicción. | 63 |
| Figura 29. Alfa de Cronbach para encuesta de validación de prototipo. | 64 |
| Figura 30. Revalidación del prototipo por cada clase. | 66 |
| Figura 31. Coral Dev Board conectada a Internet. | 66 |
| Figura 32. Arquitectura del prototipo para el entrenamiento. | 67 |
| Figura 33. Arquitectura del Prototipo en funcionamiento. | 67 |
| Figura 34. Comparación de arquitecturas simplificadas Edge computing y Cloud computing. | 68 |
| Figura 35. Arquitectura implementada y propuesta de soluciones adicionales para una agroindustria. | 71 |
| Figura 36. Proceso de Encuesta a los agricultores en la comunidad Puctus. | 85 |
| Figura 37. Pruebas de campo del prototipo en el invernadero del Sr. Pedro Sucuy. | 85 |
| Figura 38. Pruebas de campo del prototipo en el invernadero del Sr. Ramon Morocho. | 85 |
| Figura 39. Pruebas de campo del prototipo en el invernadero de la Sra. María Quisay. | 86 |
| Figura 40. a), c) Muestra de Enfermedad Fusarium, obtenida en el invernadero del Sr. Pedro Sucuy en la etapa de floración del tomate riñón. b), d) Muestra de Enfermedad Fusarium, obtenida en el invernadero de la Sra. María Quisay en la etapa de Fructificación del tomate riñón. e) Muestra de Enfermedad Fusarium, obtenida en el invernadero del Sr. Ramon Morocho en la etapa de fructificación del tomate riñón. | 88 |
| Figura 41. a), b) Muestra de plaga Mosca Blanca, obtenida en el invernadero del Sr. Ramon Morocho en la etapa de fructificación del tomate riñón. c), d) Muestra de plaga Mosca Blanca, obtenida en el invernadero de la Sra. María Quisay en la etapa de Fructificación del tomate riñón. e) Muestra de plaga Mosca Blanca, obtenida en el invernadero del Sr. Ramon Morocho en la etapa de fructificación del tomate riñón. | 90 |
| Figura 42. a), e) Muestra de plaga Minador, obtenida en el invernadero de la Sra. María Quisay en la etapa de fructificación semana 6 del tomate riñón. b), d) Muestra de plaga Minador, obtenida en el invernadero del Sr. Pedro Sucuy en la etapa de Fructificación semana 8 del tomate riñón. c) Muestra de plaga Minador, obtenida en el invernadero del Sr. Ramon Morocho en la etapa de fructificación semana 7 del tomate riñón. | 92 |
| Figura 43. a), d) Muestra Planta Sana, obtenida en el invernadero del Sr. Pedro Sucuy en la etapa de fructificación semana 7 del tomate riñón. b), c) Muestra Planta Sana, obtenida en el invernadero de la Sra. María Quisay en la etapa de fructificación del tomate riñón. e) Muestra Planta Sana, obtenida en el invernadero del Sr. Ramon Morocho en la etapa de floración del tomate riñón. | 93 |
| Figura 44. a), e) Muestra Planta Enferma, obtenida en el invernadero del Sr. Ramon Morocho en la etapa de fructificación del tomate riñón. b) Muestra Planta Enferma, obtenida en el invernadero del Sr. Pedro Sucuy en la etapa de fructificación del tomate riñón. c), d) Muestra | |

Planta Enferma, obtenida en el invernadero de la Sra. María Quisay en la etapa de fructificación del tomate riñón. 95

ÍNDICE DE ECUACIONES

| | |
|---------------------------------------------------------------------------------------|----|
| Ecuación 1. Media Aritmética | 20 |
| Ecuación 2. Tamaño de la muestra..... | 42 |
| Ecuación 3. Probabilidad de que un miembro de la población no esté en la muestra..... | 42 |
| Ecuación 4. Estimación de la proporción | 43 |
| Ecuación 5. Peso por Invernadero | 44 |
| Ecuación 6. Accuracy de revalidación del prototipo | 65 |

RESUMEN

La industria agrícola en el Ecuador, en comparación con la realidad actual a nivel mundial, vive un retraso significativo en el uso de la tecnología, y aunque desde hace varias décadas atrás hasta la actualidad ha aumentado, lo ha hecho tan pausada y apocadamente que más parece estar estancada. El uso de la tecnología puede ser muy relevante, sobre todo en el sector agrario. Este trabajo de investigación propone un sistema portátil de identificación y clasificación de plagas y enfermedades en la planta de tomate riñón. Para conseguir este objetivo se utilizaron varias tecnologías disponibles, tanto maduras como emergentes, intentando ser parte de la concepción de la Industria 4.0 en aplicaciones agroindustriales. Inicialmente se realizó una evaluación de la madurez digital del sector agrícola de la localidad de Puctus, concluyendo que prácticamente no existe el uso de la tecnología en su producción. Posteriormente se tomaron suficientes muestras de las diferentes plagas y enfermedades para ser clasificadas manualmente con ayuda de expertos. Estos datos fueron almacenados en la nube apoyados del servicio de Google Cloud para desarrolladores y posteriormente utilizados en el entrenamiento de un algoritmo de clasificación usando la plataforma Google Colab y con librerías Aprendizaje Automático de TensorFlow. Luego de varios entrenamientos se obtuvo el modelo con mayor exactitud, el cual fue convertido a un formato de TensorFlow Lite, optimizado para ser ejecutado en tarjetas de desarrollo con unidades de procesamiento TPUs, en este caso la Google Coral Dev Board. La exactitud alcanzada en la validación del modelo original fue de 96.875 %, la del modelo LITE para la tarjeta de desarrollo fue de 98.438 %, mientras que, en las pruebas reales, validadas con la ayuda de expertos alcanzó una exactitud de 92.54%. El prototipo cumple con las expectativas de esta investigación, pero puede ser mejorado, para ello se incluye algunas recomendaciones basadas en la experiencia adquirida mientras se desarrollaban diferentes pruebas, como puede ser: el uso de cámaras multiespectrales en aplicaciones agroindustriales, el uso de bases de datos, o la implementación de este prototipo en sistemas de fumigación automática.

Palabras claves: Machine Learning, TPU, TensorFlow, Visión Artificial, IoT-IIoT, Edge-Computing, Industria 4.0, Tomate riñón.

ABSTRACT

The agricultural industry in Ecuador, in comparison with the current reality worldwide, is lagging significantly behind in the use of technology, and although it has been increasing for several decades, it has done so slowly and slowly that it seems to be stagnating. The use of technology can be very relevant, especially in the agricultural sector. This research work proposes a portable system for identification and classification of pests and diseases in the kidney tomato plant. To achieve this objective, several available technologies, both mature and emerging, were used, trying to be part of the industry 4.0 conception in agro-industrial applications. Initially, an assessment of the digital maturity of the agricultural sector in the town of Puctus was carried out, concluding that there is practically no use of technology in its production. Subsequently, sufficient samples of the different pests and diseases were taken to be classified manually with the help of experts. These data were stored in the cloud supported by the Google Cloud service for developers and subsequently used in the training of a classification algorithm using the Google Colab platform and with TensorFlow Machine Learning libraries. After several training sessions, the most accurate model was obtained, which was converted to a TensorFlow Lite format, optimized to be run on development boards with TPUs processing units, in this case the Google Coral Dev Board. The accuracy achieved in the validation of the original model was 96.875 %, that of the LITE model for the development board was 98.438 %, while in the real tests, validated with the help of experts, it reached an accuracy of 92.54%. The prototype meets the expectations of this research, but can be improved, for this we include some recommendations based on the experience gained while developing different tests, such as: the use of multispectral cameras in agro-industrial applications, the use of databases, or the implementation of this prototype in automatic spraying systems.

Keywords: Machine Learning, TPU, TensorFlow, Artificial Vision, , Edge-Computing, IoT-IIoT, Industry 4.0, Tomato.



Reviewed by:
Mgs. Geovanny Armas Pesántez
PROFESSOR OF ENGLISH
C.C. 0602773301

1. CAPÍTULO I. INTRODUCCIÓN

1.1 Antecedentes

En la actualidad, el estudio de las imágenes se ha incrementado debido a las necesidades de diferentes áreas. Por ejemplo, el análisis de imágenes es vital en la medicina, la agricultura, los deportes y el reconocimiento de objetos. En el ámbito médico, las imágenes revelan su importancia para mejorar los procesos de tratamiento de los pacientes. Por lo tanto, existe una amplia investigación para contribuir a la medicina mediante la segmentación y la identificación de áreas de interés (por ejemplo, órganos). La fase de preprocesamiento es crucial para conseguir mejores resultados. Se proponen varios métodos, como la reducción del ruido, la mejora de la iluminación o la detección de los límites. El entorno de las imágenes médicas es muy amplio e incluye imágenes de rayos X, resonancia magnética y tomografía computarizada.

Actualmente, las investigaciones se han centrado en las imágenes de tomografía computarizada (TC) para aplicar técnicas de procesamiento de imágenes debido a su papel esencial en la detección de enfermedades como en el hígado [1], los pulmones[2], el cerebro [3], el cráneo [4] y la enfermedad Covid-19 [5]. En consecuencia, para desarrollar el uso de las imágenes de TC, es necesario segmentar los órganos aplicando técnicas de procesamiento de imágenes y así realizar un tratamiento preventivo y un diagnóstico. Sin embargo, la segmentación de imágenes de TC es una tarea difícil debido a varios factores como el fondo complejo, el tamaño del conjunto de datos, las dimensiones, el ruido, la iluminación y las sombras.

El procesamiento de imágenes necesita mejorar sus técnicas utilizando métodos novedosos. Por ejemplo, la Inteligencia Artificial (IA) tiene diferentes formas de automatizar procesos como el Aprendizaje Automático (ML), las Redes Neuronales Artificiales (ANN) y las Redes Neuronales Convolucionales (CNN). Los algoritmos de aprendizaje automático funcionan con métodos basados en el aprendizaje supervisado, el aprendizaje no supervisado y el aprendizaje por refuerzo.

Su objetivo es aprender un modelo y hacer predicciones o reconocer patrones. En comparación, las ANN simulan el cerebro humano para aprender de los datos a través de las neuronas y finalmente extraer la información necesaria. Del mismo modo, las CNN también aprenden de los datos, pero utilizan capas más profundas para reconocer características, especialmente de las imágenes.

Los métodos que implican la segmentación de imágenes requieren un estudio en profundidad de los métodos tradicionales y no tradicionales. A lo largo de los años, los campos de la Inteligencia Artificial han ido revolucionando, desde las técnicas básicas de Aprendizaje Automático (Machine Learning) hasta los modelos de Deep Learning, aumentando la exactitud y exactitud de cada uno de los métodos.

En estos días, la nube manda, porque en gran medida es así. Una gran cantidad de los servicios de computación que hoy se prestan en el mundo se llevan a cabo en la nube, obviamente mediante internet. La propagación de la red y el aumento de la potencia de las conexiones vino a cambiarlo todo. Poco a poco, muchos servicios comenzaron a ampararse en la nube y se podría decir que es una vasta infraestructura, compuesta por miles de redes y servidores, que se dedican a llevar a cabo todo tipo de tareas de computación.

La computación de borde (Edge computing) supone, en cierto modo, una vuelta a los orígenes, sin perder por el camino todas las ventajas que Internet nos proporciona. En el futuro inmediato, se prevé que el Internet de las Cosas (IoT) crezca de una forma exponencial. Lo que ocurre con el IoT es que los dispositivos que forman parte de él suelen ser muy básicos y disponer de una capacidad de computación muy escasa; en muchas ocasiones se limitan a recoger datos para enviarlos a través de Internet, con el fin de que sean los equipos “en la nube” y mucho más potentes, los que hagan todo el trabajo de computación. Pero eso está a punto de cambiar. El edge computing, como su nombre indica, pretende poner la computación en el “borde”, es decir, quiere lograr que las tareas de computación o al menos una parte de ellas sea llevada a cabo por los propios dispositivos IoT.

En Ecuador, el tomate es el cuarto producto agrícola de mayor cultivo, en el país existen 3333 hectáreas de plantación, y la producción anual es de alrededor de 61426 toneladas por año. El tomate tiene una gran cantidad de vitaminas y sales minerales y es considerado un excelente elemento para añadir a la dieta [6] este trabajo se añadió por contener información bastante relevante pese a no cumplir un criterio de inclusión de información actualizada.

El proceso de comercialización del tomate depende directamente del tamaño y el color, estos factores intervienen en la calidad de sus productos derivados y posteriormente en el porcentaje de aceptación del consumidor; por ejemplo, una serie de pruebas sensoriales realizadas sobre la salsa de tomate han revelado que el brillo y el ángulo de tonalidad son los elementos responsables de la aceptación del consumidor. El pigmento que toman los tomates durante el proceso de maduración permite diferenciar los cambios de color en una progresión del tomate verde al rojo, así como la aparición de plagas y/o enfermedades [7]. Los agricultores deben tomar muchas decisiones sobre el proceso de plantación y recolección de los tomates basándose en el conocimiento empírico de las personas que recolectan los productos, y este método es uno de los más ineficientes [8]. En la actualidad, la mayoría de los agricultores clasifican los tomates manualmente, lo que significa que hay personas que, debido a su experiencia, identifican el color de cada tomate y detectan la presencia de cualquier defecto en la superficie del producto cosechado. El problema de este método de clasificación surge en la producción a gran escala, porque a medida que aumenta el número de tomates, se necesitan varios trabajadores experimentados para separar los buenos de los malos, y esto aumenta el coste de producción. Un sistema de control de calidad adecuado permitiría integrar programas o sistemas dentro del proceso de producción de tomates para mejorar la productividad y asegurar que la empresa busca la mejora continua de la calidad de sus productos [9]. La tecnología actual ofrece nuevos sistemas de visión por

ordenador que permiten la caracterización cuantitativa de diversas características físicas como el tamaño, la morfología, el color y las propiedades texturales mediante una imagen digital. Además, estos sistemas suelen ir acompañados de sistemas que pueden sustituir a la manipulación humana para que los procesos sean más eficaces.

La aplicación de esta tecnología en el campo tanto artesanal como industrial beneficiaría en calidad, costos y uso de recursos directamente a toda la cadena de valor del tomate riñón, es decir, desde su cultivo y sus agricultores, pasando por las fases intermedias de recolección y/o producción de derivados, hasta la entrega a los consumidores finales.

1.2 Planteamiento del Problema

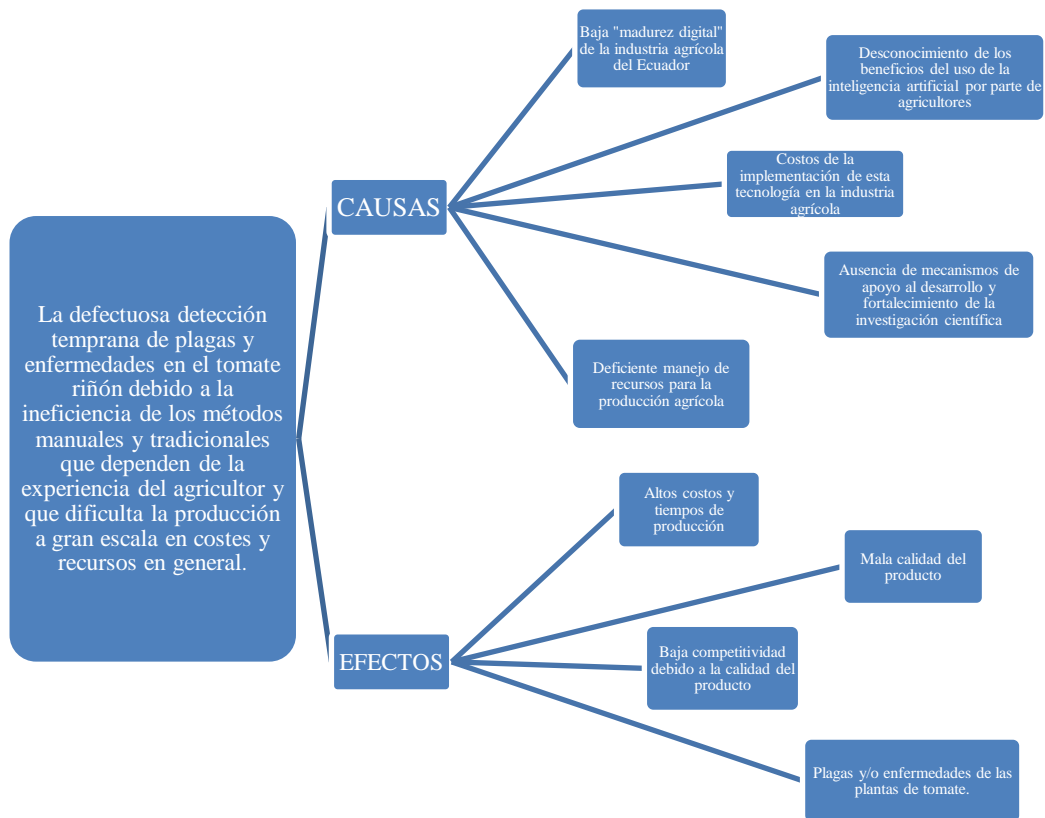


Tabla 1. Madurez Digital – Variables

| Madurez digital y variables de evaluación | | | | | |
|-------------------------------------------|----------------------------------------------------------------------|------------|---|---|---|
| Variables dependientes | Variables independientes | Valoración | | | |
| | | 1 | 2 | 3 | 4 |
| Tecnología | 1 Tecnología para el intercambio de información | | | | |
| | 2 Utilización de tecnología en la nube | | | | |
| | 3 Dispositivos móviles en planta | | | | |
| | 4 Almacenamiento de información descentralizado | | | | |
| | 5 Sensores para recopilación de datos | | | | |
| | 6 Computadoras integradas en máquinas | | | | |
| | 7 Computadoras integradas en herramientas | | | | |
| | 8 Fabricación aditiva | | | | |
| | 9 Utilización de robots | | | | |
| Productos | 10 Individualización de producto | | | | |
| | 11 Flexibilidad de las características del producto | | | | |
| | 12 Recopilación de información sobre el uso del producto | | | | |
| | 13 Componentes de procesamiento de datos en productos | | | | |
| | 14 Conexión a Internet de productos | | | | |
| | 15 Servicios de TI relacionados con productos físicos | | | | |
| Clientes y socios | 16 Apertura a las nuevas tecnologías | | | | |
| | 17 Competencia con las TIC modernas | | | | |
| | 18 Digitalización del contacto con el cliente | | | | |
| | 19 Utilización de datos relacionados con el cliente | | | | |
| Procesos de creación de valor | 20 Automatización de procesos de creación de valor | | | | |
| | 21 Autonomía de las máquinas de planta | | | | |
| | 22 Intercambio de información entre máquinas | | | | |
| | 23 Acceso remoto a las máquinas de planta | | | | |
| | 24 Control de calidad automatizado | | | | |
| | 25 Manejo de herramientas de automatización | | | | |
| | 26 Colaboración hombre - máquina (uso de HMLs) | | | | |
| Datos e información | 27 Procesos de información digital | | | | |
| | 28 Recopilación de datos automatizada | | | | |
| | 29 Análisis de datos recopilados | | | | |
| | 30 Toma de decisiones basada en datos | | | | |
| | 31 Suministro de información automatizado | | | | |
| | 32 Individualización de la información proporcionada | | | | |
| | 33 Visualización digital de los procesos | | | | |
| | 34 Simulación de escenarios futuros basada en datos | | | | |
| Estándares corporativos | 35 Contratación de servicios para la Industria 4.0 | | | | |
| | 36 Capacitaciones de empleados para competencias digitales | | | | |
| | 37 Nivel de seguridad cibernética | | | | |
| | 38 Reglas para empleados en entorno de trabajo digital | | | | |
| Empleados | 39 Abiertos a las nuevas tecnologías | | | | |
| | 40 Competentes con las TICs modernas | | | | |
| | 41 Experiencia con trabajo interdisciplinario | | | | |
| | 42 Voluntad de formación continua en el puesto de trabajo | | | | |
| Estrategia y liderazgo | 43 Hoja de ruta para la transformación hacia la Industria 4.0 | | | | |
| | 44 Recursos financieros para hacer realidad la Industria 4.0 | | | | |
| | 45 Voluntad de la gerencia para la transformación a la Industria 4.0 | | | | |
| | 46 Conocimiento o capacitación de la gerencia sobre Industria 4.0 | | | | |

En la Tabla 1, existen variables dependientes e independientes correspondientes específicamente al estudio de madurez digital. Las variables independientes tienen una ponderación de 1 a 4, aportando un valor numérico a las variables de evaluación, es decir que cada variable independiente aportará proporcionalmente a cada variable dependiente, la operación correspondiente es el promedio por cada variable dependiente, operación también conocida como la media aritmética, mediante la cual se obtienen los datos específicos para un análisis de la madurez digital de la industria agrícola, y posteriormente representado en un gráfico radial. Mientras que la Tabla 2 contiene las variables dependientes e independientes de la investigación y desarrollo del prototipo, en ella se detalla las operaciones respectivas.

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$$

Ecuación 1. Media Aritmética

Tabla 2. Variables

| Variables Dependientes | Variables Independientes | Operación |
|---------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------|
| Infección de las plantas | Muestras Fotográficas | Aprendizaje Automático (ML) |
| Costo de la implementación de la tecnología (Prototipo) | <ul style="list-style-type: none"> • Precio tarjeta de desarrollo • Precio de Cámara Google Coral • Precio de Memoria SD • Protector de Tarjeta de Desarrollo Google Coral Dev • Pantalla Digital | SUMA |

1.3 Justificación

La agricultura en el Ecuador viene relativamente precaria en comparación con la agricultura en Europa y Estados Unidos, donde la industrialización de sus procesos ha avanzado bastante, gracias al uso de las tecnologías. En gran parte del territorio ecuatoriano se maneja un modelo de agricultura de subsistencia o semisubsistencia.

La agricultura de subsistencia ocurre cuando los agricultores cultivan alimentos para satisfacer sus necesidades y las de sus familias en las pequeñas propiedades, y para satisfacer principalmente las necesidades locales, con poco o ningún excedente. A pesar de la primacía de la autosuficiencia en la agricultura de subsistencia, la mayoría de los agricultores de subsistencia también participan en el comercio hasta cierto punto. La mayoría de los agricultores de subsistencia operan hoy en países en desarrollo, como Ecuador. Aunque su volumen de comercio medido en efectivo es menor que el de los

consumidores en países con mercados modernos y complejos, muchos tienen importantes contactos y artículos comerciales que pueden producir debido a sus habilidades especiales o acceso especial a recursos valorados en el mercado. La agricultura de subsistencia generalmente presenta: necesidades de capital o financiamiento, cultivos mixtos, uso limitado de agroquímicos (por ejemplo, pesticidas y fertilizantes), variedades no mejoradas de cultivos, poco o ningún excedente de rendimiento para la venta, uso de herramientas crudas o tradicionales, realizada en pequeñas parcelas de tierra dispersas, la dependencia de mano de obra no calificada (a menudo miembros de la familia) y (generalmente) bajos rendimientos.

La falta de recursos viene a ser el principal inconveniente para migrar hacia una agricultura industrializada. Este proyecto de investigación pretende aportar a la producción local con un sistema que ayude con la producción de tomate riñón, otorgando soluciones gracias al uso de tecnologías que permiten tanto mejorar la producción como optimizar recursos. Adicionalmente el costo del prototipo inicial viene a ser económico en comparación con soluciones comerciales.

1.4 Objetivos

1.4.1 General

- Detectar plagas minador y mosca blanca, y la enfermedad fusarium en el tomate riñón de manera temprana mediante la aplicación de técnicas de visión artificial, machine learning y Edge-Computing, que optimicen el manejo de los recursos de producción.

1.4.2 Específicos

- Evaluar la madurez digital de la industria agrícola en la comunidad Puctus en la parroquia Yaruquies, cantón Riobamba a través de un método empírico propuesto por Fraunhofer Austria Research GmbH y la Universidad de Tecnologías de Viena, para conocer su realidad tecnológica.
- Obtener muestras (imágenes) válidas de las plantas de tomate riñón y sus productos sanos e infectados para su posterior tratamiento.
- Alcanzar modelos de clasificación optimizados que detecten y prevengan plagas de minador y mosca blanca, y la enfermedad Fusarium en el tomate riñón mediante la aplicación de técnicas de Machine Learning sobre las muestras obtenidas.
- Implementar el modelo entrenado en una tarjeta de desarrollo adecuada para procesamiento de imágenes y algoritmos de machine learning sobre Edge-Computing con el fin de ensamblar un prototipo funcional que detecte exitosamente las plagas y enfermedades objeto de este estudio.

2. CAPÍTULO II. MARCO TEÓRICO

2.1 Estado del Arte

El uso de la visión artificial y técnicas de aprendizaje automático (Machine Learning) no es relativamente nuevo en aplicaciones dentro de la industria agrícola, sin embargo, capacidades de procesamiento, técnicas de machine learning y específicamente hardware para aplicaciones embebidas han ido mejorando y aportando con el desarrollo de aplicaciones más eficientes. Además, nos ha permitido ampliar el uso de estas tecnologías en estudios que aún no han sido profundizados, como el caso específico de enfermedades en el tomate riñón. A continuación, se analizan algunos trabajos, reportados en bases científicas, y previamente seleccionados bajo criterios de inclusión del año de publicación no mayor a 5 años atrás y con resultados que aporten al entendimiento y desarrollo del trabajo de investigación.

Los autores del trabajo reportado en [10], proponen un sistema capaz de detectar enfermedades, pero en este caso en el pepino. Su trabajo está basado en redes neuronales convolucionales (CNN). Su sistema alcanza un promedio de exactitud de 94.9%. En este trabajo se han utilizado imágenes para el entrenamiento y el modelo de clasificación, y no utilizan otro tipo de datos de entrenamiento.

Otra técnica muy interesante que se ha encontrado en la búsqueda bibliográfica se menciona en [11]. Los autores proponen un sistema para la identificación automática de la enfermedad “tizón de la hoja del maíz del norte”. A diferencia del trabajo anterior, los autores utilizan técnicas redundantes, inicialmente entrenan el modelo con imágenes de plantas lesionadas y no lesionadas, posteriormente las predicciones fueron combinadas en mapas de calor. Los resultados alimentaron una red neuronal final consiguiendo una exactitud del 96.7% en sus pruebas.

En el trabajo [12] sus autores se enfocan en un método capaz de identificar enfermedades en el arroz que al igual de los métodos anteriores está basada en redes neuronales convolucionales, en este caso usando una arquitectura AlexNet. El aspecto concurrente en el uso de técnicas de machine learning con imágenes es la aplicación de las CNN. Los autores reportan que su modelo alcanzó el 95.48% de exactitud y para validar la fiabilidad de sus resultados su modelo fue sometido a comparaciones con otras técnicas como el SVM (support vector machine) que alcanzó el 91% de exactitud; con el BP (backpropagation) que llegó al 92% y finalmente con el método de optimización por enjambre de partículas con solamente con el 88%.

Un inconveniente general que se menciona en la mayoría de los trabajos encontrados es el sobreajuste, cuando no existe la cantidad de datos de entrenamiento suficientes. En este caso, en el trabajo [13], adicional al conjunto de datos de entrenamiento que fueron 2,430 imágenes, los autores analizaron el rendimiento de su modelo entrenado utilizando el método de transferencia de aprendizaje, es decir realizan un preentrenamiento con cierto conjunto de datos y posteriormente ajustan los parámetros con un nuevo entrenamiento logrando un 90.84% de exactitud. La conclusión principal del trabajo que aporta a esta investigación es

que la combinación de CNN con transferencia de aprendizaje resulta muy conveniente para la clasificación de enfermedades de plantas cuando se tiene datos a pequeña escala.

Un trabajo interesante, que se diferencia de los demás al no usar CNN se reporta en [14]. El trabajo compara y utiliza tres métodos de clasificación: K-NN (K-Nearest Neighbors), ANN MLP (Perceptron Multilayer Neural Network) y, K-means Clustering; La exactitud de cada uno fue superior al 90%, tanto para el clasificador K-NN como para el clasificador de red neuronal MLP, y el error medio (error de los clasificadores) fue igual o inferior al 6%, a diferencia del clasificador K-means, en el que la mayoría de los rendimientos fueron superiores al 80%, pero inferiores al 90%.

Dentro de la búsqueda bibliográfica, se trató de encontrar información científica particularmente del uso de Tarjetas de Google Coral en este tipo de aplicaciones, la computación de borde (Edge computing), sin embargo, es casi nula la información. Además, el uso de las librerías TensorFlow y TensorFlow Lite está muy poco reportada en información científica dado que fue liberada por Google hace apenas poquísimos años. Un sitio web de colaboradores y ayuda se puede encontrar en [15].

2.2 Marco Teórico

2.2.1 Industria 4.0

Las etapas en el desarrollo de sistemas de fabricación industrial, desde el trabajo manual hasta el concepto de Industria 4.0 se pueden presentar como un camino a través de las cuatro revoluciones industriales, esta ruta ha sido tratada en muchos documentos científicos [16]. El desarrollo se muestra en la Figura 1.

La Primera Revolución Industrial comenzó a fines del siglo XVIII y principios del XIX, que estuvo representada por la introducción de sistemas de fabricación mecánicos que utilizan agua y vapor. La Segunda Revolución Industrial comenzó a fines del siglo XIX, simbolizada por la producción en masa mediante el uso de energía eléctrica. La Tercera Revolución Industrial comenzó a mediados del siglo XX e introdujo la automatización y la tecnología microelectrónica en la fabricación. Hoy se encuentra en la cuarta revolución industrial impulsada por el desarrollo de las Tecnologías de la Información y las Comunicaciones (TIC). Su base tecnológica es la automatización inteligente de sistemas ciberfísicos con control descentralizado y conectividad avanzada (funcionalidades IoT) [17], [18]. La consecuencia de esta nueva tecnología para los sistemas de producción industrial es la reorganización de los sistemas de automatización jerárquica clásicos a un sistema de producción físico cibernético autoorganizado que permite una producción personalizada en masa flexible y flexibilidad en la cantidad de producción [19].



Figura 1. Etapas en el desarrollo de sistemas de fabricación industrial adaptado de [20].

2.2.2 Trabajo colaborativo asistido por ordenador

El factor humano es fundamental en la Industria 4.0 ya que la automatización industrial también da como resultado la "redefinición" de los puestos de trabajo, por lo que se recomienda el desarrollo educativo para los empleados de la organización. Evitar islas de conocimiento es uno de los objetivos de la interconectividad, y el intercambio de conocimiento requiere que no se quede dentro de los límites de la organización, sino que los supere [21], [22].

Por tanto, es necesario hablar de Trabajo Colaborativo Soportado por Computadora (CSCW - Computer-Supported Collaborative Work), para que incluya procesos y recursos que se están involucrando y ofrezca una integración de herramientas y métodos que apoyen la colaboración de los equipos de trabajo y potencialmente puedan mejorar la productividad y efectividad. de quienes trabajan en colaboración[23], [24]. Los temas relevantes relacionados con la Industria 4.0 se establecen en la Figura 2.

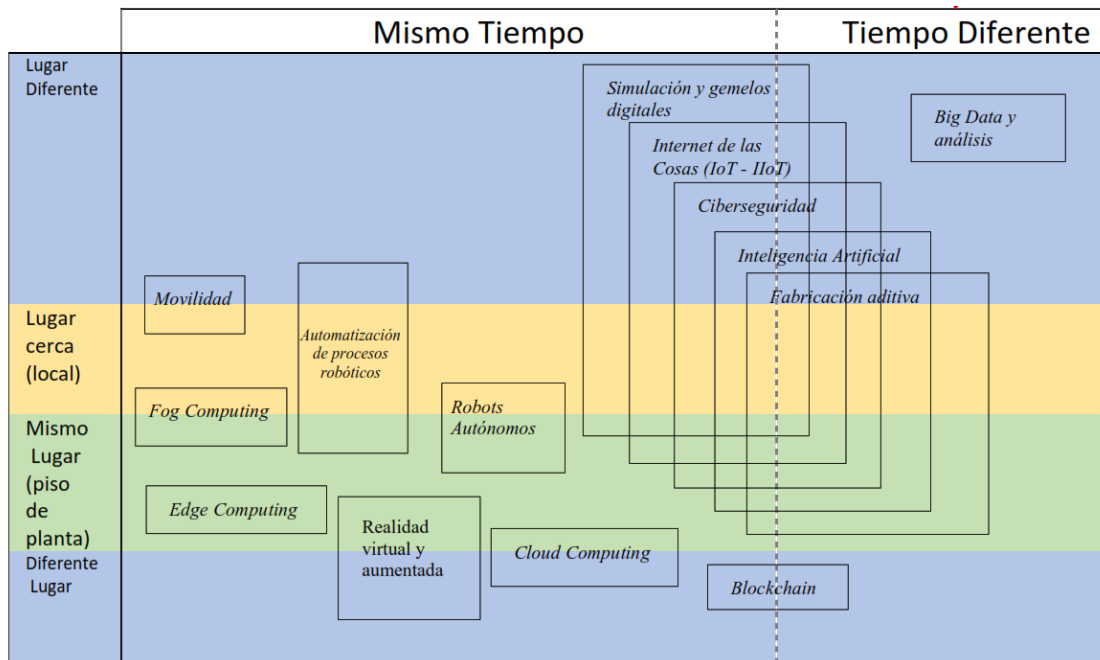


Figura 2. Métodos, temas y herramientas en CSCW adaptado de [21], [24], [16].

Estos son todos los factores por considerar en relación con la Industria 4.0 y lo que implica dentro de la Gestión del Conocimiento. Los requisitos, capacidades y restricciones deben tenerse en cuenta en lo que los expertos denominan la cuarta revolución industrial. En este caso, se centra en el aspecto tecnológico, tanto una revisión bibliográfica de tecnologías maduras como emergentes, que participan en el concepto de industria 4.0.

No puede hablar de tecnologías específicas, pues el concepto permite una constante evolución y continua integración de nuevas tecnologías según van madurando o adaptándose, en otras palabras, varias tecnologías emergentes están convergiendo para brindar soluciones digitales; pero si se menciona las principales, y su aplicación en este caso de estudio.

Pasar de la estrategia a la ejecución es la clave de un estudio de implantación de la Industria 4.0, sin embargo, al tener demasiada información sobre todas las tecnologías implicadas en un concepto tan amplio, esto se convierte en una tarea bastante complicada. Por lo tanto, una mala ejecución mostrará que la estrategia, por muy buena que sea, no tendrá éxito o no se alcanzarán los objetivos deseados. El esquema de la Figura 3 propuesto por los autores [16] puede ayudar a inspirar y contribuir al diseño de planes empresariales estratégicos que puedan ser implementados con éxito y alinear los recursos de TI, OT y humanos a la ejecución de estos planes y sobre todo entender en aspectos generales, el amplio concepto de la Industria 4.0.

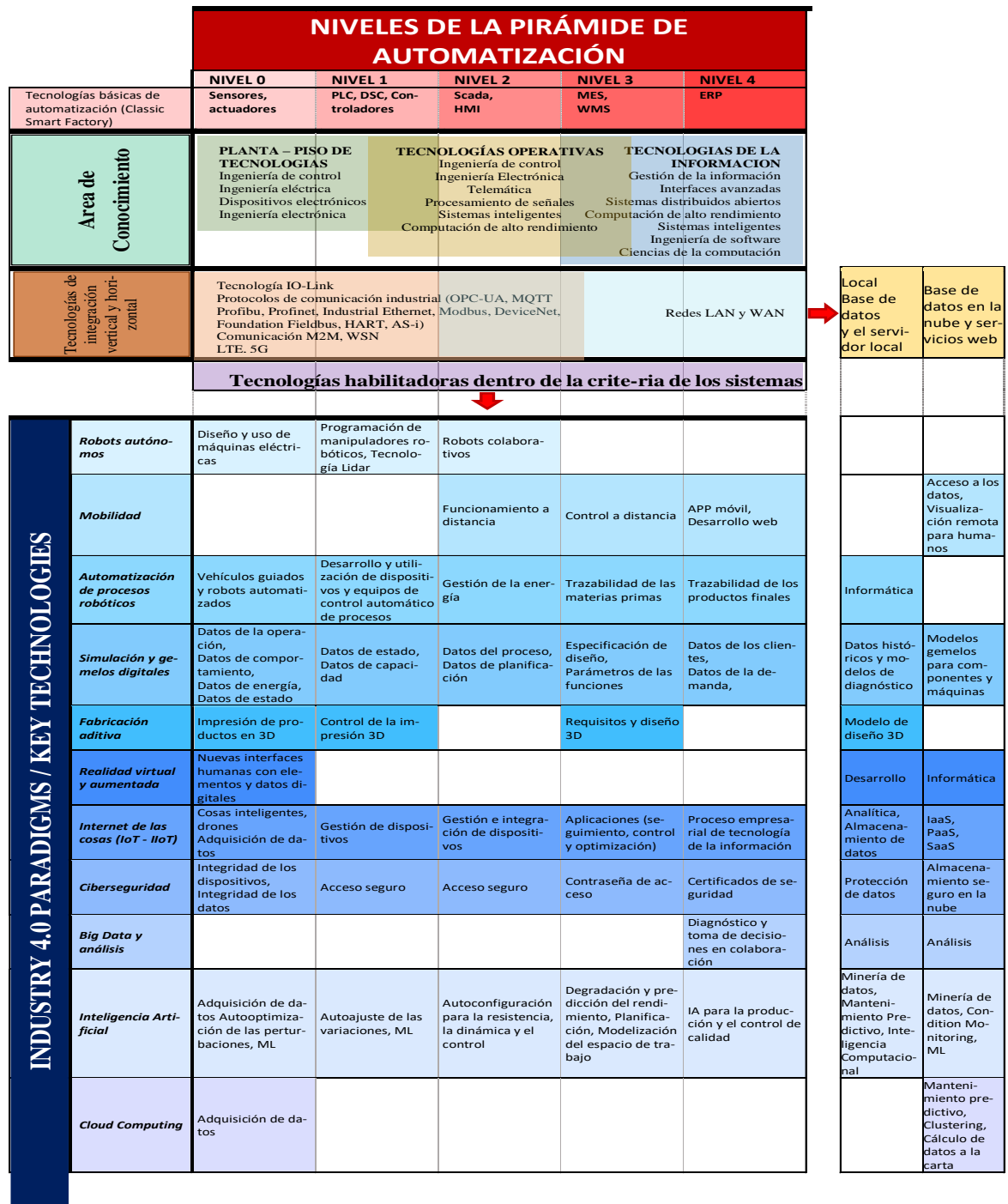


Figura 3. Marco de resumen de los patrones de adopción de Industria 4.0 y sus tecnologías [16].

2.2.3 Sistemas físico-cibernéticos

Este concepto no es tan nuevo, describe una amplia gama de sistemas de ingeniería de próxima generación complejos, multidisciplinarios y con conciencia física que integran

tecnologías informáticas integradas (partes cibernéticas) en el mundo físico. El progreso de diferentes tecnologías, entre ellas IoT, han hecho posible recopilar, acumular y analizar datos; para así proporcionar una retroalimentación de los resultados del análisis desde el ciberespacio de vuelta al mundo real, creando una relación recíproca [16]. Las investigaciones en este tema está en auge en el mundo científico y técnico, la Figura 4 ilustra los orígenes geográficos de publicaciones en este tema, siendo China y Alemania los países que más contribuyen[16].

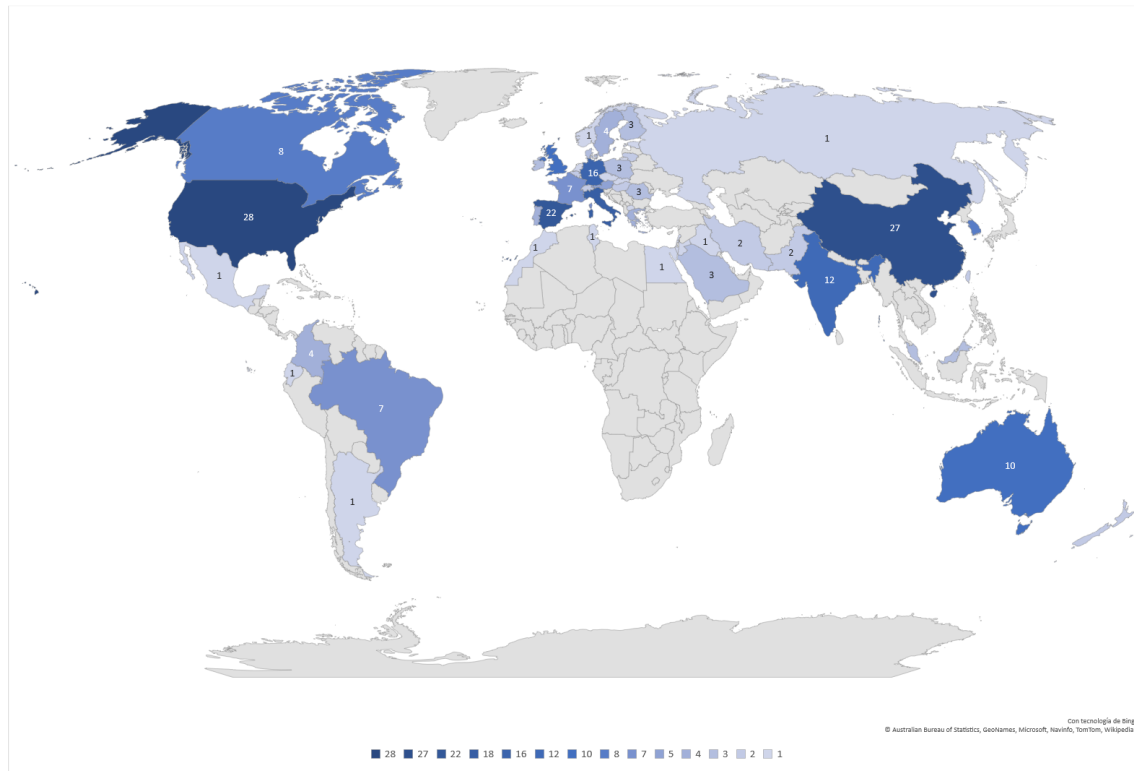


Figura 4. Distribuciones geográficas de publicaciones ciberfísicas industriales [16].

Desde que comenzaron a utilizarse dispositivos de TIC independientes, la tecnología ha progresado hasta el nivel de implementación de Cybert-physical systems (CPS) (nivel IV) a través de avances como la conexión en red de terminales individuales y el procesamiento de datos de alto rendimiento, y a medida que la realización de CPS toma forma concreta en varios sectores sociales, incluidos el sector industrial (manufactura), está provocando un cambio social radical más allá de la automatización y la dotación de inteligencia en desarrollo, la producción y la distribución [25].

Los sistemas ciberfísicos se han analizado y definido de varias maneras, y han existido diferentes propuestas de arquitecturas, una de ellas mostradas en [25], sin embargo existe una propuesta de arquitectura bastante completa reportada en [26] compuesta por 5 niveles y con una guía de implementación, sus autores la han denominado arquitectura 5C (connection, conversion, cyber, cognition, and configuration). La Figura 5 muestra cada nivel con sus funciones y atributos. Esto no quiere decir que la implementación deba ser

como dicha propuesta, pero será una referencia válida para un estudio más profundo de acuerdo con la naturaleza y necesidades de la empresa.

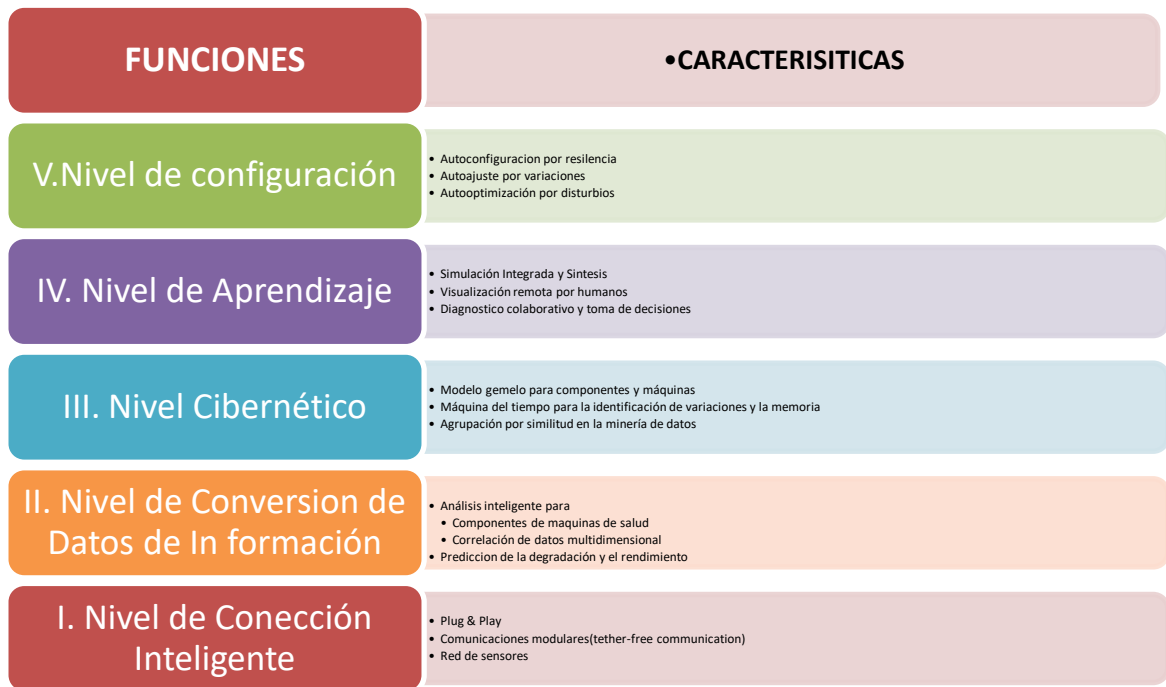


Figura 5. Arquitectura 5C para la implementación de CPS adaptado de [26].

Este concepto lleva a analizar diferentes tecnologías que se enmarcan en CPS, como son IoT, Fog computing, cloud computing, big data.

En 2018 se ha reportado un trabajo elaborado por la Universidad de Coruña (España), y Navantía (uno de los constructores navales más grandes del mundo). Se ha desarrollado a través de un proyecto que hace uso de algunas de las últimas tecnologías de la Industria 4.0, específicamente se describe un sistema ciberfísico (CPS), que utiliza etiquetas activas de identificación por radiofrecuencia (RFID) para rastrear tuberías y detectar eventos relevantes. Además, el CPS se ha integrado y probado junto con el sistema de ejecución de fabricación (MES) de Siemens (Simatic IT). Los experimentos que han realizado en el CPS muestran resultados interesantes que se podría considerar analizar, en los escenarios seleccionados del mundo real, las puertas de enlace de Fog responden más rápido que el servidor en la nube probado, siendo estas puertas de enlace también capaces de procesar con éxito más muestras en situaciones de alta carga, es decir, reportan la fiabilidad del uso de tecnología Fog computing en este tipo de aplicaciones y reaccionan más rápido que la alternativa cloud computing [27]. Su propuesta de arquitectura se puede observar en la Figura 6, se pueden identificar claramente tres capas.

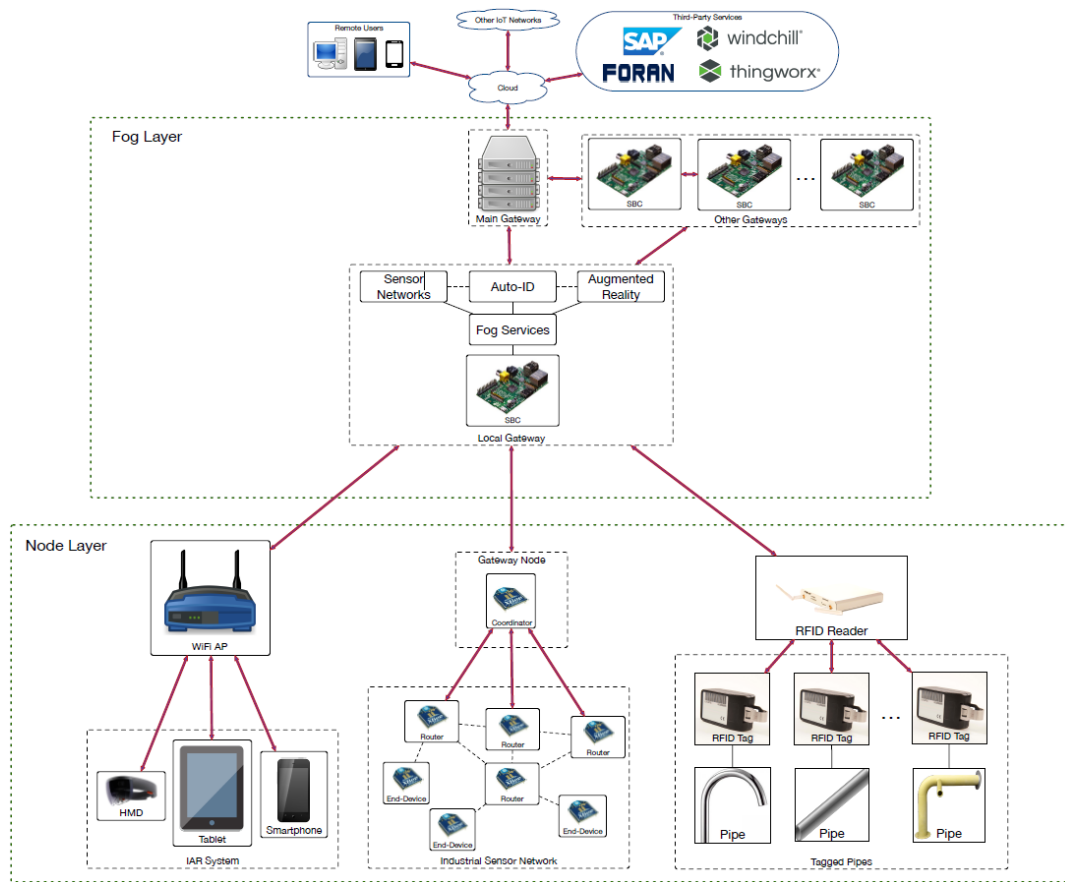


Figura 6. Arquitectura CPS del proyecto de Navantia [27].

La integración entre CPS y MES la representan en diagrama sencillo, acorde a la naturaleza de su industria. El diagrama resume el orden de fabricación y la interacción entre el sistema físico y el ciberespacio. La Figura 7 muestra un diagrama de secuencia simplificado que ilustra las interacciones CPS-MES.

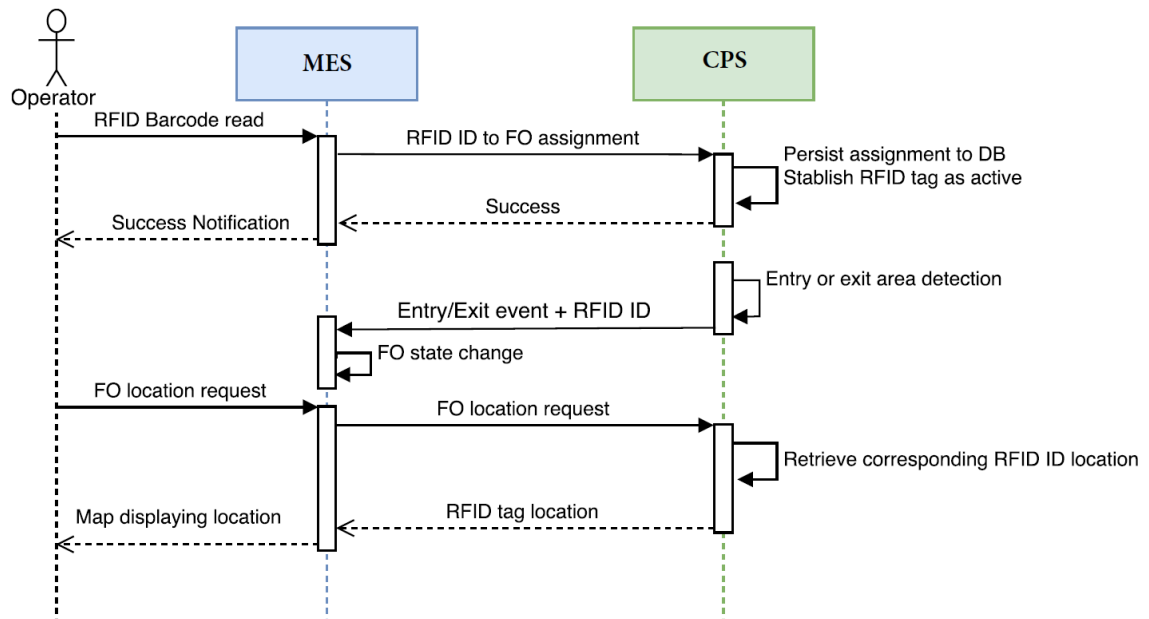


Figura 7. Diagrama de interacción entre CPS y MES del proyecto de Navantia [27].

En los trabajos [28], [29] se puede encontrar otras arquitecturas CPS en aplicaciones de visión artificial y sistemas embebidos respectivamente. Da a entender que existe cierta flexibilidad en su arquitectura.

2.2.4 IoT

Este concepto está muy relacionado al anterior (CPS), llegando al punto de ser interdependientes en un concepto industrial. El mundo físico se transforma en digitalizado y conecta todo. Una explosión de dispositivos y tecnologías inteligentes ha permitido a la humanidad estar en comunicación constante en cualquier lugar y en cualquier momento [30]. Industrial Internet of Things está impulsando la cuarta fase de la Revolución Industrial. Acelerará el viaje de las industrias a un nuevo nivel y transformará las economías, abriendo una nueva era de crecimiento y competitividad.

El Internet de las cosas (IoT) es una red de objetos físicos, dispositivos, instrumentos, vehículos, electrodomésticos, edificios y otros elementos integrados con componentes electrónicos, circuitos, software, sensores y conectividad de red que permite que estos objetos recopilen e intercambien datos. El Internet de las cosas permite que los objetos se detecten y controlen de forma remota a través de la infraestructura de red existente, creando oportunidades para una integración más directa del mundo físico en los sistemas basados en computadoras y dando como resultado una mayor eficiencia y exactitud [30]. En casi todos los ámbitos de la vida, el IoT parece estar cambiando el mundo.

Mientras que IoT afecta, entre otros, al transporte, la atención médica o los hogares inteligentes, el Internet industrial de las cosas (IIoT) se refiere en particular a los entornos industriales. En este contexto, Cyber Manufacturing Systems (CMS) evolucionó como un término significativo, entre ellos los conceptos de Fábrica Digital y los sistemas ciberfísicos [30]. Para identificar un poco más claro el marco del IoT, se lo puede clasificar en tres subcategorías según su uso y/o usuario: Consumer IoT, Commercial IoT, Industrial IoT. La Figura 8. representa un esquema del mundo del IoT, sin que esto limite su exclusividad de uso, ni norma.



Figura 8. Tipos de IoT.

2.2.5 Arquitecturas del internet de las cosas (IIoT)

El internet de las cosas (IoT) se perfila como una pieza fundamental para el desarrollo de la industria. Su rápido crecimiento y avance tecnológico se está convirtiendo en la base para un cambio de modelo de negocio para los diferentes interesados del sector. Sin embargo, es necesario definir una arquitectura y componentes necesarios.

Antes de proponer una arquitectura IoT posible, se debe saber que en temas de arquitecturas de IoT, no existe una única definición universalmente adoptada ni estandarizada, inclusive se han adaptado arquitecturas propias para aplicaciones específicas, la flexibilidad que esta tecnología presenta en este sentido se ha convertido en una de sus mayores ventajas al momento de adoptar por una transición digital de las industrias. Diferentes propuestas han surgido durante su desarrollo entre las que se puede encontrar: la arquitectura de 3 capas, la arquitectura de 5 capas, la arquitectura de nube (cloud architecture), la arquitectura de niebla (fog architecture) y la arquitectura de computación de Borde (edge architecture) entre muchas otras.

La literatura presenta un gran número de soluciones de arquitecturas IoT, pero en este trabajo de investigación se presenta un posible marco general para entender las principales capas que conforman IoT. La Figura 9 realiza una breve comparación entre tres de las arquitecturas más relevantes. A partir de ella, se realizará una propuesta de arquitectura simplificada para una posible implementación de un sistema que utilice el prototipo desarrollado.






























| | Cloud | Fog | Edge | |
|--------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| Número de dispositivos activos |  Computación y ML  Almacenamiento  Análisis  Intercambio de mensajes |  Intercambio de mensajes  Almacenamiento  Análisis |  Intercambio de mensajes  Almacenamiento  Análisis | Latencia |
| |  Fog nodes  Gateways |  Fog nodes  Gateways  Computación de niebla (Fog computing) |  Gateways | |
| |    Sensor  Actuadores |    Sensor  Actuadores |    Sensor  Actuadores  Computación de borde (Edge computing) | |

Figura 9. Comparación de 3 arquitecturas base de IoT divididas en 3 capas

2.2.6 Datos Masivos (Big Data)

Definición de Big Data

A diferencia de los datos tradicionales, el término Big Data se refiere a grandes conjuntos de datos en crecimiento que incluyen formatos heterogéneos: datos estructurados, no estructurados y semiestructurados. Big Data tiene una naturaleza compleja que requiere tecnologías potentes y algoritmos avanzados. Por tanto, las herramientas tradicionales de Business Intelligence estáticas ya no pueden ser eficientes en el caso de las aplicaciones de Big Data [31]. En resumen, Big Data, además de involucrar una gran masa de datos, incluye recursos que aseguran la velocidad de procesamiento y servicios que agregan valor. La Figura 10 muestra las dimensiones que diferencian el Big Data de los datos normales [32].

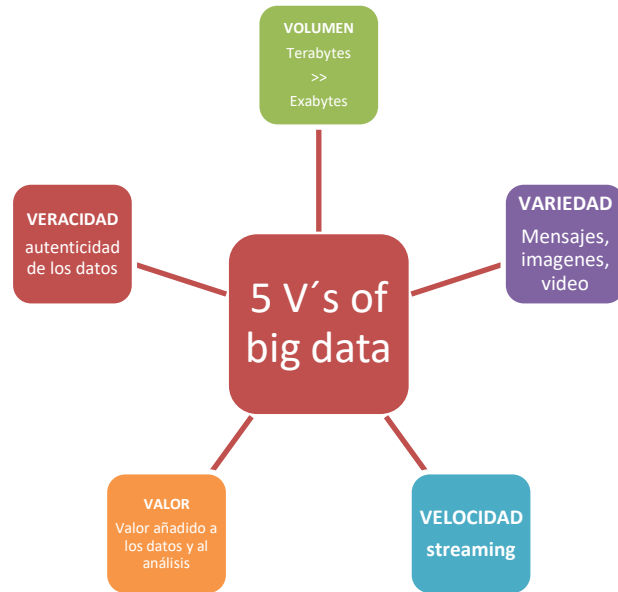


Figura 10. Las 5V de Big Data adaptado de [32],[33].

Las características de “5V” del big data industrial hacen posible proporcionar información completa y sistemática a través de la minería de datos y el descubrimiento de conocimientos [34]. Las tecnologías de minería de datos se utilizan para descubrir la siguiente información:

- Defectos de diseño de equipos
- Defectos de diseño del producto
- Condiciones de salud del equipo
- Defectos de procesamiento de producción
- Comportamiento y hábitos laborales del personal
- Comportamiento, hábitos y demandas del cliente

Big Data en la Industria 4.0

Con aparición del concepto de Industria 4.0, su desarrollo, y de Big Data en auge, la industria cuenta con grandes y únicas oportunidades en términos para impulsar su desempeño a un nuevo nivel [36]. El rendimiento se toma aquí en el sentido más amplio, desde los aspectos operativos, económicos y relacionados con el mercado hasta la seguridad del proceso y el medio ambiente [37].

La interdependencia entre la tecnología, los datos y su análisis es el ciclo principal (Figura 11), esta interrelación permite el surgimiento de técnicas, tecnologías, herramientas, que de alguna manera sostienen el complejo concepto de Industria 4.0.

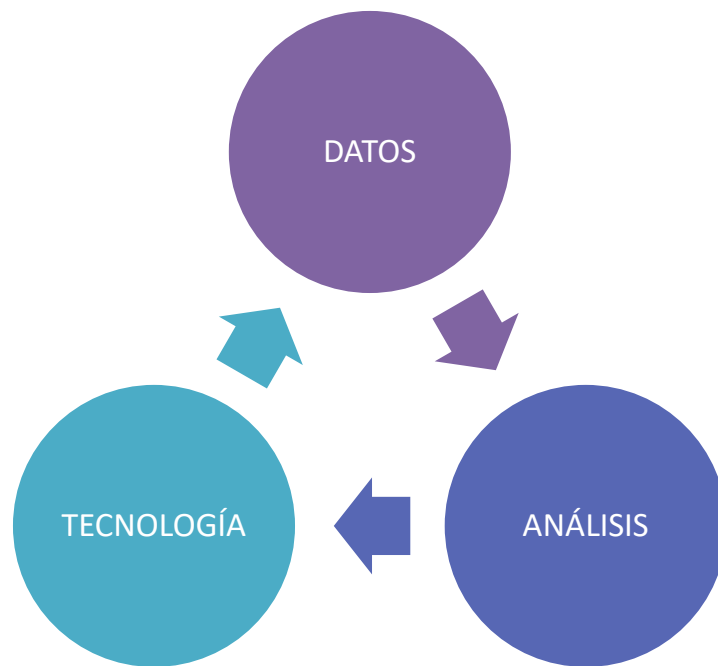


Figura 11. Big Data dentro de la industria 4.0.

2.2.7 Computación en la nube (Cloud computing), Computación en la niebla (Fog computing), Computación de borde (Edge computing)

"Cloud Computing" es un paradigma que permite "Utility Computing", es decir, el arrendamiento de recursos informáticos como: potencia de cálculo, almacenamiento y los relacionados a los recursos de red; en tiempo real, con una interacción mínima con el proveedor. De esta manera, Cloud Computing simplifica la operación, ya que no requiere un dimensionamiento y pronóstico cuidadosos de los recursos necesarios, lo que permite la facturación de pago por uso. Además, los clientes de la nube aprovechan recursos supuestamente infinitos bajo demanda y pueden aprovechar o entregar todo como un servicio. Los servicios más comunes son: Infraestructura, Plataforma o Software como servicio (IaaS, PaaS, y SaaS, respectivamente) [38]. La migración que se está produciendo a los servicios en la nube está impulsada por una tendencia emergente en la última década: es decir, la extensión de las funcionalidades integradas en los dispositivos de campo que les ha dotado de más inteligencia y flexibilidad, permitiendo así trasladar algunas funciones a la Nube, con la escalabilidad y la capacidad de respuesta relacionadas (y eventualmente, las ganancias económicas) [38].

Sin embargo, varias deficiencias del paradigma de la computación en la nube se han vuelto evidentes a lo largo de los años, principalmente relacionadas con la comunicación entre el dispositivo final y el centro de datos que aloja los servicios en la nube: la latencia, el ancho de banda, el costo y la disponibilidad de la conexión contribuyen a limitar un número de usos para Cloud Computing. La proliferación de dispositivos móviles generalizados

empeoró aún más este fenómeno, desafiando en gran medida el paradigma de la nube. De hecho, en varios contextos, la nube no puede cumplir con todos los requisitos de muchas aplicaciones, específicamente en aquellas que requieren tiempo de respuesta casi inmediata, esto hace surgir nuevas soluciones o arquitecturas diferentes. La computación en la niebla, más conocida como Fog Computing propone transferir algunos servicios de computación en la nube al borde de la red, cerca de los dispositivos del usuario, los nodos que realizan dicho procesamiento suelen estar ubicados en la misma red local que los dispositivos IoT y, por lo general, tienen capacidad de almacenamiento e informática [38], y posiblemente confiando parcialmente también en los recursos del dispositivo de los usuarios, distribuyendo así la carga entre los dispositivos finales y los centros de datos tradicionales en la nube y llevando, entre otros, seguridad a largo plazo, tasas de latencia bajas y una capacidad de respuesta más rápida, a la vez que ayuda a mejorar la escalabilidad del rendimiento en todo el sistema [39].

A través de IoT y el desarrollo de Smart Factory, la cantidad de datos generados en el borde está aumentando. Esto genera nuevos desafíos para las aplicaciones en tiempo real, porque el estado actual de las redes solo puede transmitir una cierta cantidad de datos y es el cuello de botella para su aplicación. Esta nueva necesidad, exige más a Fog computing, apareciendo el paradigma de Edge Computing. El esquema de arquitectura típico de Edge Computing consta de una capa de nodo, una capa de borde o niebla y una capa de nube. Con base en esta arquitectura, los beneficios de la computación en el borde son minimizar aún más la latencia, el consumo de energía, ahorrar ancho de banda y aumentar el nivel de seguridad y privacidad. La Figura 12 es un esquema referente de los 3 paradigmas de computación analizados previamente.

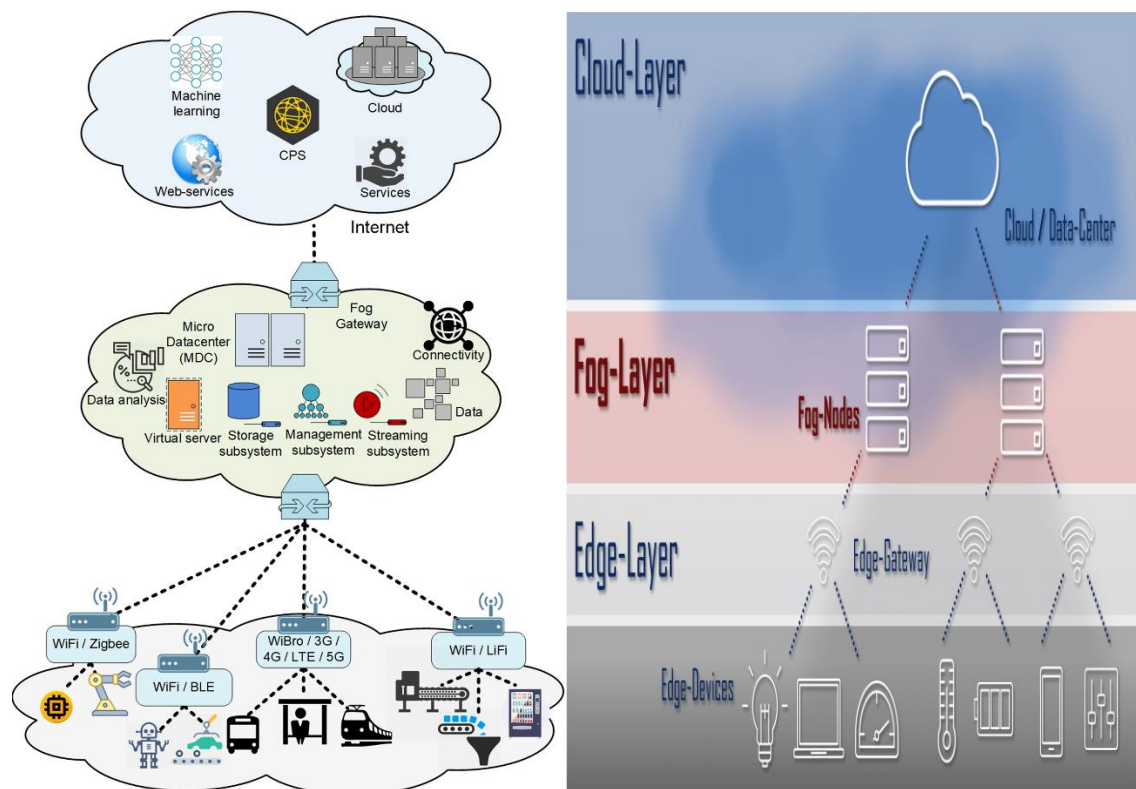


Figura 12. Cloud – Fog – Edge Computing en IIoT (adaptado de [40], [41]).

De esta revisión literaria se puede concluir considerar el uso de Fog y Edge Computing por ejemplo en aplicaciones de machine learning integradas en tiempo real para las operaciones de la Industria 4.0 [42].

2.2.8 Aprendizaje automático (Machine Learning) en la industria

El aprendizaje automático enseña a las computadoras a hacer lo que es natural para los humanos y los animales: aprender de la experiencia. Los algoritmos de aprendizaje automático utilizan métodos computacionales para "aprender" información directamente de los datos sin depender de una ecuación predeterminada como modelo. Los algoritmos mejoran su rendimiento de forma adaptativa a medida que aumenta el número de muestras disponibles para el aprendizaje. Los algoritmos de aprendizaje automático encuentran patrones naturales en los datos que generan información y lo ayudan a tomar mejores decisiones y predicciones.

Con el aumento del big data, el aprendizaje automático se ha vuelto particularmente importante para resolver problemas en el sector industrial, como estos: Procesamiento de imágenes y visión por computadora, para reconocimiento facial, detección de movimiento, detección de objetos, producción de energía, para previsión de precios y carga, automoción, aeroespacial, fabricación, mantenimiento predictivo, etc.

Elegir el algoritmo correcto puede parecer abrumador: existen bastantes algoritmos de aprendizaje automático supervisados y no supervisados, y cada uno tiene un enfoque diferente para el aprendizaje [43], [44].

No existe el mejor método. Encontrar el algoritmo correcto es, en parte, solo prueba y error; incluso los científicos de datos con mucha experiencia no pueden saber si un algoritmo funcionará sin probarlo. Pero la selección del algoritmo también depende del tamaño y el tipo de datos con los que está trabajando, los conocimientos que desea obtener de los datos y cómo se utilizarán esos conocimientos. La Figura 13 muestra los algoritmos disponibles para Machine Learning y su clasificación.

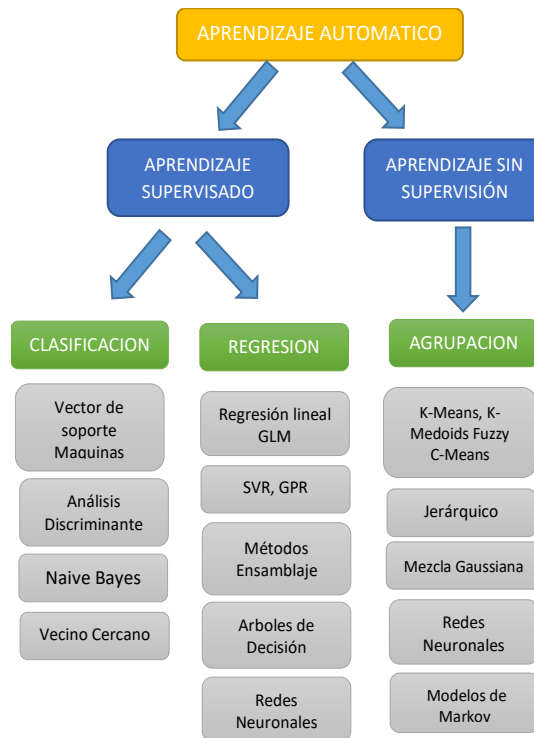


Figura 13. Clasificación de los algoritmos de Machine Learning (Adaptado de [45]).

2.2.9 Madurez digital en la industria.

La Industria 4.0 ofrece un sin límites de ventajas en la actualidad, y las empresas conocen claramente el potencial que ofrece su implementación, y están listas para invertir en su transformación. Sin embargo, a pesar de ser un concepto sencillo, y manejar tecnologías maduras en su mayoría, es muy importante manejar correctamente las estrategias, o al menos lo mejor posible, pues no existe un mapa de ruta definido claramente sin mencionar también la realidad de cada industria y su sector. Muchas empresas han reportado problemas en su proceso de digitalización cuando tratan de operar dentro de conceptos de integración horizontal y vertical. Esto puede ocasionar costes adicionales y muchas veces innecesarios, se debe entender el concepto de Industria 4.0 es aún abstracto y complejo a la vez, está aún en constante evolución y promete la integración posterior de tecnologías emergentes como 5G, siendo las principales barreras para su realización [45].

Existen muchos modelos de madurez en la literatura, propuestos tanto por la academia como por empresas de desarrollo y profesionales practicantes; todos ellos funcionales y efectivos. Estos modelos han sido desarrollados de manera empírica y de acuerdo con las necesidades y realidades locales o de cada empresa. Las estrategias para desarrollar estos modelos se han derivado, en la mayoría de los casos, de metodologías bien documentadas ya aplicadas [46]–[48]. Estos modelos conceptuales permiten recopilar datos sobre el estado actual de desarrollo de las industrias e identificar factores de éxito, necesidades y deficiencias. Las

diferentes empresas pueden utilizar los resultados de su autoevaluación como punto de partida para la implementación de las distintas medidas estratégicas. Para esta investigación, se ha tomado de referencia un modelo y se ha adaptado a la realidad de la industria local.

La mayoría de las propuestas de análisis de la madurez digital son empíricas y basadas en experiencias empresariales con el apoyo de la academia. Adicional a esto, cualquier propuesta puede ser adaptada a la realidad de la empresa o industria a analizar. El método seleccionado para esta tesis se ha considerado de los más generales y adaptables, por lo que facilita su análisis, ampliamente utilizado y siendo base teórica incluso de aplicaciones gubernamentales con ese propósito, como es el caso de HADA en España [49].

2.2.10 Plagas y Enfermedades

2.2.10.1 Fusarium

El tomate es una de las hortalizas de mayor importancia en varios países del mundo [50]. Entre las enfermedades más importantes que afectan a este cultivo se encuentra la marchitez vascular, cuyo agente causal es el Fusarium afectando también la calidad del producto. Esta enfermedad afecta a más de 30 países en gran diversidad de condiciones [50].

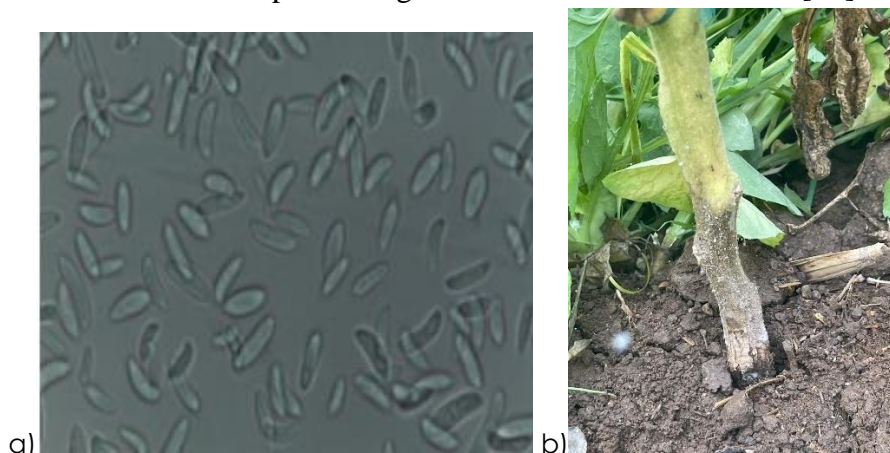


Figura 14. a) Enfermedad Fusarium vista desde microscopio [50]. b) Enfermedad fusarium en el tallo de una planta.

Características

Normalmente, el hongo infecta a las plantas de tomate riñón por las raíces, causando una obstrucción de los tejidos vasculares existe síntomas de la enfermedad como puede ser el amarillamiento del follaje, el cual puede ser confundida por falta de nitrógeno, la característica más notable en el fusarium se observa en el oscurecimiento vascular en los tallos, el tejido vascular de una planta enferma se torna de color pardo oscuro, siendo más notable en el punto de unión del peciolo con el tallo; este color es muy característico de la enfermedad y se utiliza para su correcta identificación, información base para aprendizaje automático (machine learning) dado que el algoritmo de entrenamiento se basa en la extracción de características de imágenes, color, forma, textura, perfiles, etc.. El

marchitamiento de las hojas, la pudrición de la raíz son señales de que la enfermedad se encuentra en el estado más avanzado [50].

2.2.10.2 Mosca Blanca

En el mundo de la agricultura en las plantas hortalizas la mosca blanca (*Trialeurodes vaporariorum*) es una plaga de alta dificultad de manejo. Por los daños económicos que estimula es considerada la segunda especie más importantes del mundo.[51]

Es polífaga esto quiere decir que se alimenta de varios huéspedes, causando daño a un gran número de especies. Los estadios ninfales y adultos succionan la savia de la planta, reduciendo los rendimientos principalmente en los cultivos hortícolas siendo el tomate uno de los cultivos más afectados.[51]

En países de América del Sur incluyendo Ecuador la producción hortícola tiene como desafío aumentar la productividad y mejorar aspectos para el manejo de las plagas.[51]



Figura 15. Ciclo de vida de plaga Mosca Blanca [51]

Características

El huevo de mosca blanca se establece en el envés de la hoja. Al inicio de su ciclo los huevos son blancos, luego toman un color amarillo y finalmente se torna café oscuro, el período de eclosión generalmente dura entre 6 y 10 días [51].

Al pasar este tiempo la ninfa recién emergida del huevo se mueve para localizar el lugar de alimentación hasta llegar al ciclo adulto donde su cuerpo toma un color amarillo limón, las alas son transparentes, angostas en la parte anterior, se ensanchan hacia atrás y están cubiertas por un polvo blanco lo cual ayuda para su identificación [51]. Información base para aprendizaje automático (machine learning) dado que el algoritmo de entrenamiento se basa en la extracción de características de imágenes, color, forma, textura, perfiles, etc...

2.2.10.3 Minador

El minador de tomate es un insecto-plaga de importancia económica en las zonas tomateras del país. Esta plaga conocida también como “gusano minador” representa una de las mayores limitantes para la producción rentable de esta hortaliza. [52]

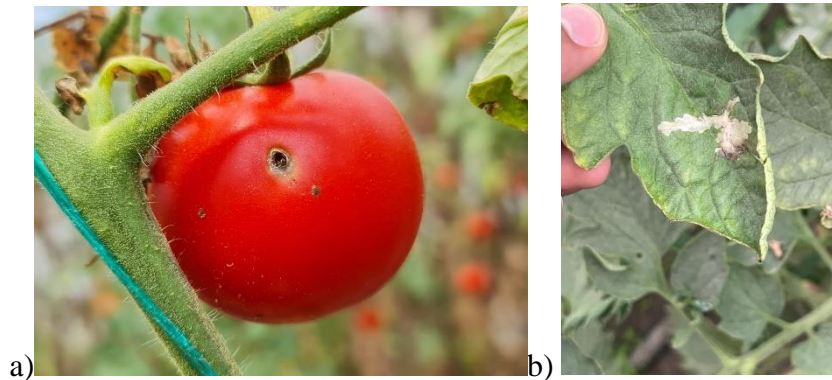


Figura 16. a) Plaga minador en fruto b) Plaga minador en follaje

Características

La larva afecta entre las dos epidermis de la hoja destrozando el mesófilo dejando áreas translúcidas necrosadas generalmente llamadas minas las cuales al unirse producen el quemado total de la hoja. De igual forma ataca tanto los frutos pequeños como los desarrollados. Lo cual es una característica muy principal para el reconocimiento de la plaga [52]. Información base para aprendizaje automático (machine learning) dado que el algoritmo de entrenamiento se basa en la extracción de características de imágenes, color, forma, textura, perfiles, etc.

3. CAPÍTULO III. METODOLOGIA

3.1 Tipo de Investigación

El análisis cuantitativo esta intrínseco en los estudios estadísticos del proyecto de investigación.

Así también el análisis cualitativo de la información de las fotografías con la ayuda de expertos para obtener las primeras clasificaciones. Estos resultados se han obtenido combinando una investigación descriptiva e interpretativa de los datos. Las principales fases de este análisis son las siguientes:

- Búsqueda y obtención de información (imágenes).
- La preparación, revisión y transcripción de la información.
- La organización de la información y los datos según criterios de expertos.
- La categorización, etiquetado y codificación de la información y los datos, que los prepara para el análisis.
- El análisis de los datos y la generación de proposiciones, utilidad, ejemplos y conclusiones.

3.2 Diseño de Investigación

Inicialmente, se realizó un estudio de la madurez digital de la industria agrícola en la comunidad de Puctus localizada en la Parroquia Yaruquies en el cantón Riobamba. Esto permitió recopilar datos sobre el estado actual de desarrollo de las industrias agrícolas e identificar factores de éxito, necesidades y deficiencias tecnológicas. Las diferentes industrias o artesanos productores podrán utilizar los resultados de las evaluaciones como punto de partida para la implementación de las distintas medidas estratégicas tecnológicas que ayuden a mejorar su producción; y también se puede tener una visión general de la realidad de la agroindustria y uso de nuevas tecnologías.

La metodología que se aplicó en esta investigación está dividida en dos etapas. En la primera etapa se realizó un estudio del caso, esto involucra el estudio detallado de un caso particular de detección de una plaga o enfermedad; para lo cual se recolectó la mayor cantidad de datos (en este caso imágenes) tanto de plantas sanas como infectadas de diferentes plagas o enfermedades, siempre cuidando de que las muestras de imágenes que se están tomando sean de buena calidad y adecuadas para su estudio, a continuación, se clasificó manualmente con el apoyo de expertos en el tema. Posterior a eso, se preprocesaron las imágenes aplicando diferentes filtros digitales. Este estudio abarcó a tres productores locales de tomate, Sr. Ramón Morocho, Sra. María Quisay, y Sr. Pedro Sucuy como propietarios. Ubicados en la Provincia de Chimborazo, Cantón Riobamba, Parroquia Yaruquies, Comunidad Puctus.

Correspondiente a las siguientes coordenadas EPSG:3857 -8758442.694772048,-189442.20281112657. A una altitud de 2.754 m.s.n.m.

3.3 Población de estudio y tamaño de muestra

3.3.1 Datos de la población

Tabla 3. Datos de la población.

| Productores Locales | Cantidad de plantas por invernadero |
|---------------------|-------------------------------------|
| Sr. Ramon Morocho | 2000 |
| Sra. María Quisay | 2000 |
| Sr. Pedro Sucuy | 2000 |
| TOTAL | 6000 |

3.3.2 Muestra

Calcular el tamaño de la muestra (o tamaño muestral) es fundamental. Una muestra más grande supone un desperdicio de recursos; una muestra más pequeña produce una pérdida en la calidad de los resultados. La ecuación a emplear depende del objetivo que se persiga (p.ej. una proporción, una media, etc.) y también depende del tamaño de la población, de si ésta es N es finita o infinita o muy grande. La elección de los parámetros intervinientes para el cálculo han de ser determinados con base experta expresada en la siguiente formula.

- **Fórmula**

$$n = \frac{z^2 * p * q * N}{z^2 * p * q + (N - 1) * \varepsilon^2}$$

Ecuación 2. Tamaño de la muestra

n= muestra

N= Población (6000)

p= probabilidad de que un miembro de la población esté en la muestra (0.5)

q= probabilidad de que un miembro de la población no esté en la muestra (0.5)

ε =error (0.03-0.05)

Z= valor normalizado con la confiabilidad (1.96)

$$q = 1 - p$$

Ecuación 3. Probabilidad de que un miembro de la población no esté en la muestra

$$q = 1 - 0,5$$

$$q = 0,5$$

$$n = \frac{z^2 * p * q * N}{z^2 * p * q + (N - 1) * \varepsilon^2}$$

$$n = \frac{1,96^2 * 0,5 * 0,5 * 6000}{1,96^2 * 0,5 * 0,5 + (6000 - 1) * 0,04^2}$$

$$n = 545.74$$

$$n \approx 546$$

3.3.3 El nivel de confianza

El nivel de confianza $(1-\alpha)$ se refiere a la probabilidad de que el dato deseado esté dentro del margen establecido. Este parámetro lo decide el investigador. Suele ser del 95%, $(\alpha = 0,05)$ al que se corresponde un coeficiente de confianza $Z = 1,96$, que es el que se pone en la fórmula. Es la semidistancia estandarizada en términos de desviaciones típicas que definen ambos extremos del intervalo.

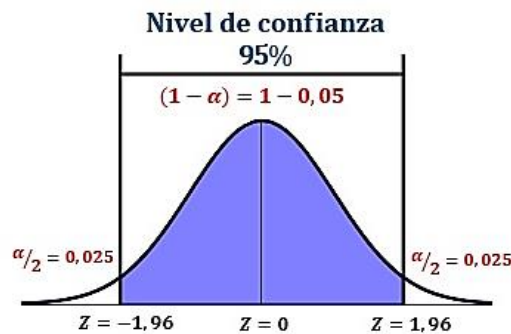


Figura 17. Representación del nivel de confianza.

3.3.4 La estimación de la proporción

La estimación de la proporción que se quiere medir es el tema clave. Se quiere estimar la proporción de los que cumplen la condición. El valor de esa estimación p la se obtendrá de estudios anteriores. En caso contrario, se considera que la condición la cumplan un 50% y, por tanto, que no la cumplan $(1 - p)$ el otro 50%. En ese caso, se pone en la fórmula:

- **Fórmula**

$$p \cdot (1 - p) = 0,5 \cdot 0,5 = 0,25$$

Ecuación 4. Estimación de la proporción

3.3.5 El margen de error deseado

El margen de error deseado, o exactitud, o margen de error admisible, se refiere a la diferencia entre la media muestral y la media poblacional. Desde luego, que no se pretende cometer errores. Se trata de un margen de error que se este dispuestos a tolerar.

Suele adoptarse $e = 4\%$ (0,04), aunque está entre:

- *Condición*

$$1\% \leq e \leq 10\% (0,01 \leq e \leq 0,10)$$

Error de $\epsilon=0.04$, en la cual se encuentra en el rango establecido.

PESOS-POR INVERNADERO

- **Fórmula**

$$\text{Peso por Invernadero} = \frac{\# \text{ de Plantas por Invernadero}}{\# \text{ Total de plantas}} * 100\%$$

Ecuación 5. Peso por Invernadero

$$\# \text{ de plantas por invernadero} = \frac{\text{Peso por invernadero} * n}{100\%}$$

$$\# \text{ de plantas por invernadero} = \frac{33.33\% * 546}{100\%}$$




$$\# \text{ de plantas por invernadero} = 181.98$$

$$\# \text{ de plantas por invernadero} \approx 182$$

3.4 Técnicas de recolección de Datos

Estos resultados estadísticos dan una noción de las cantidades de plantas necesarias para representar un invernadero, sin embargo, no da información de la cantidad de imágenes necesarias para entrenar un modelo eficiente. Esta información se la obtuvo empíricamente, según los resultados de cada entrenamiento que se realizó. Esta información se detallará en el capítulo IV de Resultados y Discusiones. La tarjeta de desarrollo que se consideró para esta investigación es la Coral Dev Board de Google, la cual tiene una unidad de procesamiento conocida como TPU (Unidad de Procesamiento de Tensores) y permitió ejecutar modelos de Aprendizaje Automático (Machine Learning) entrenados y desarrollados en la plataforma de Google Colab usando librerías de alto nivel de TensorFlow para procesamiento de imágenes y video en tiempo real conocido como Edge-Computing. Esta tarjeta fue considerada para esta investigación porque presentó mejores características en la mayoría de los criterios de selección, tanto técnicas como de costos, la Tabla 4. resume las características más relevantes de las tres opciones que fueron tomadas en cuenta. Adicional a estas características, un criterio muy importante fue la capacidad de la Google Coral Dev Board de ejecutar algoritmos que contengan librerías de TensorFlow optimizadas para Machine Learning.

Tabla 4. Comparación de Tarjetas de Desarrollo

| | Google Coral Dev Board | Jetson nano NVIDIA | Raspberry Pi 4 |
|----------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------|
| Foto |  <p>Fuente: Página oficial de Google Coral</p> |  <p>Fuente: Página oficial de nVIDIA Developer</p> |  <p>Fuente: Página oficial de Raspberry</p> |
| Procesador | Quad symmetric Cortex-A53 de of 1.5 GHz | ARM® A57 de cuatro núcleos a 1,43 GHz | Broadcom BCM2711, Quad core Cortex-A72 de 1.5GHz |
| Graphic Processing Unit (GPU) | Vivante GC7000Lite de 32 GFLOPs 32-bit or 64 GFLOPs 16-bit | NVIDIA Maxwell™ de 128 núcleos | NO |
| Acelerador (coprocesador) | Edge TPU coprocessor, ASIC diseñado por Google que provee alta eficiencia en modelos de machine learning de TensorFlow Lite. | NO | NO |
| Descodificación de video máxima | 4Kp60 HEVC/H.265 4Kp60 VP9 and 4Kp30 AVC/H.264 1080p60 MPEG-2, MPEG-4p2, VC-1, VP8, RV9, AVS, MJPEG, H.263 | 4Kp60, 2 de 4Kp30, 8 de 1080p30, 18 de 720p30 (H.264/H.265) | H.265 (4kp60 decode), H264 (1080p60 decode, 1080p30 encode) |
| Interfaz de cámara input | MIPI-CSI2 camera input (4-lane) | 1 conector CSI-2 MIPI | 2-lane MIPI CSI camera port |
| Pantalla | HDMI 2.0a MIPI-DSI 4 channels | HDMI | 2 × micro-HDMI 2-lane MIPI DSI |
| Interfaces | | Gigabit Ethernet, Wi-Fi, 1x USB 3.0 Type A, 2x USB 2.0 Type A, USB 2.0 Micro-B | 2.4 GHz and 5.0 GHz IEEE 802.11ac wireless, Bluetooth 5.0, BLE Gigabit Ethernet |

| | | | |
|---------------|-----------|-----------|-----------------------------------|
| | | | 2 USB 3.0 ports; 2 USB 2.0 ports. |
| Precio | \$ 180,00 | \$ 100,00 | \$ 90,00 |

Una vez obtenidas las muestras, clasificadas y ya adecuadas para su análisis, se aplicaron diferentes técnicas de procesamiento de imágenes para extracción de características [69], especialmente la convolución. Esta información extraída será la base para el entrenamiento de modelos de machine learning [54]. Se buscaron los algoritmos más adecuados para el objetivo [55]–[59]. El modelo obtenido fue optimizado y puesto a prueba con una parte de las muestras iniciales y con nuevas muestras, con el fin de verificar su efectividad en el reconocimiento y/o predicción, para lo cual se necesitará aún el apoyo de un experto en el tema de la agricultura del tomate. Finalmente, una vez verificado el funcionamiento del modelo, se puede dar por finalizada la primera etapa.

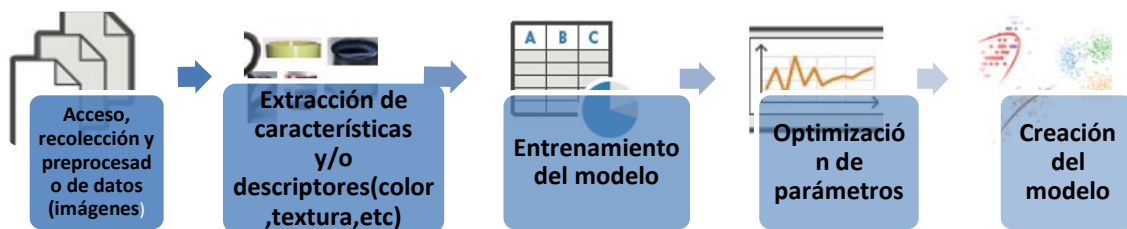


Figura 18. Procesos de la primera etapa de entrenamiento.

La segunda etapa consistió en la implementación, entendiendo que el modelo está entrenado y optimizado. El algoritmo con el modelo entrenado será implementado en una tarjeta de desarrollo para aplicaciones de machine learning, Coral Dev Board de Google. Esta tarjeta está diseñada para ejecutar eficientemente algoritmos de inteligencia artificial y procesamiento de imágenes en tiempo real y en aplicaciones embebidas, conocido como computación en la nube (Edge Computing) [60]–[63]. En esta tarjeta se procesó las imágenes, desde la extracción de características, la aplicación del modelo, y finalmente la detección y predicción final.



Figura 19. Procesos de la etapa de implementación del modelo entrenado.

La implementación de la tecnología Edge-Computing, en este tipo de aplicaciones tiene relevantes ventajas como una notable reducción de los costos de implementación y una rápida puesta en funcionamiento. Además de una respuesta con mínima latencia para aplicaciones críticas en tiempo real.

Para garantizar la eficiencia del modelo entrenado se utilizó una técnica redundante a la comprobación en el mismo modelo. Esta técnica consistió en realizar una encuesta a agricultores expertos acerca de los resultados de predicción del modelo con el fin de contrastar con los obtenidos. A estos datos se les realizó un análisis comparativo de medias para reforzar la fiabilidad del proyecto, así como para futuras estrategias de marketing y comercialización, entregando garantía y confianza a otros agricultores sobre el funcionamiento del sistema propuesto.

Finalmente, se implementó una interfaz de usuario final, de notificación sencilla en pantalla, la cual se encarga de informar en tiempo real el estado de la planta analizada, en formato video con texto informativo.

El prototipo resultante consiste en un dispositivo portátil capaz de analizar y detectar en tiempo real plagas, que a través una interfaz de usuario informa mediante video los resultados.

3.5 Métodos de análisis, y procesamiento de datos.

3.5.1 Clasificación de las imágenes por carpetas:

El primer análisis es manual, prácticamente las imágenes recolectadas se clasificaron en carpetas de acuerdo con su categoría, en este caso se seleccionaron 5 categorías: fusarium, minador, mosca blanca, planta sana y planta enferma. Este set de imágenes clasificadas servirá para el entrenamiento y validación del modelo, la Figura 20. muestra este sencillo procedimiento.

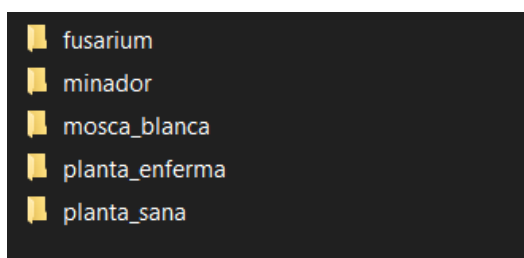
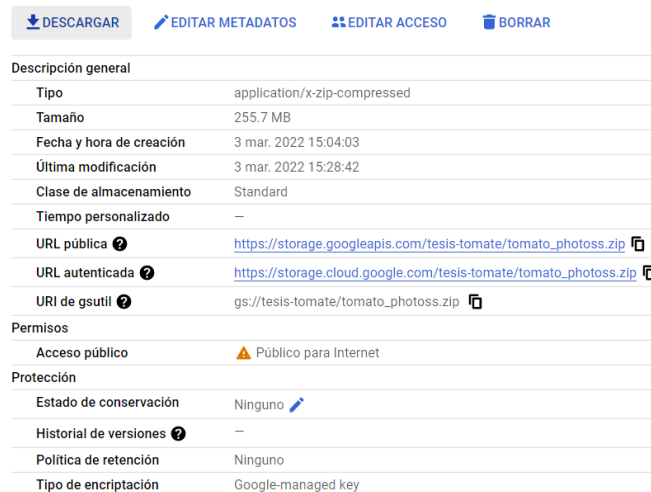


Figura 20. Imágenes clasificadas en carpetas por categorías.

3.5.2 Datos en la nube:

Las carpetas que contienen las categorías son subidas a la nube, en este caso se usa el servicio de “Google Cloud-Storage”. Los datos deben estar subidos en formato de desarrollador, para que se obtenga acceso directo de aplicaciones y APIs, eso es parte del servicio de Google Cloud Storage, en este caso, el algoritmo de entrenamiento del modelo accederá directamente a estos datos en la nube, en la Figura 21. Se observa cómo se representa los datos en la nube en formato para desarrolladores, conjuntamente con sus links de acceso directo para las aplicaciones.



The screenshot shows the details of a file in Google Cloud Storage. At the top, there are four action buttons: 'DESCARGAR' (Download), 'EDITAR METADATOS' (Edit metadata), 'EDITAR ACCESO' (Edit access), and 'BORRAR' (Delete). Below these is a table with the following data:

| Descripción general | |
|--------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------|
| Tipo | application/x-zip-compressed |
| Tamaño | 255.7 MB |
| Fecha y hora de creación | 3 mar. 2022 15:04:03 |
| Última modificación | 3 mar. 2022 15:28:42 |
| Clase de almacenamiento | Standard |
| Tiempo personalizado | — |
| URL pública | https://storage.googleapis.com/tesis-tomate/tomato_photoss.zip |
| URL autenticada | https://storage.cloud.google.com/tesis-tomate/tomato_photoss.zip |
| URI de gsutil | gs://tesis-tomate/tomato_photoss.zip |
| Permisos | |
| Acceso público | ⚠ Público para Internet |
| Protección | |
| Estado de conservación | Ninguno |
| Historial de versiones | — |
| Política de retención | Ninguno |
| Tipo de encriptación | Google-managed key |

Figura 21. Datos en Google Cloud Storage en formato para desarrolladores.

3.5.3 Entrenamiento del modelo

Este proceso consiste en la obtención de un modelo entrenado y optimizado de acuerdo con los datos y categorías que se ha asignado. La mayoría de guías disponibles utilizan un set de datos de internet con miles de datos todos convenientemente limpios y etiquetados, la verdad es que al hacer esto se está ahorrando la mayoría del trabajo y cuando se quiere aplicarlas con proyectos propios muy pronto se encontraran con dos problemas: conseguir suficientes datos para entrenar las redes no es fácil y entrenar redes neuronales con muchas capas y millones de parámetros se convierte en un procedimiento muy demoroso y muy costoso en términos de recursos de cálculo, por lo que se aplica una técnica conocida como “transferencia de aprendizaje” y que ayuda a obtener muy buenos resultados incluso cuando se tiene muy pocos datos de entrenamiento.

Para este caso se partirá del reentrenamiento de un modelo de clasificación, desarrollado por Google, este modelo se llama MobileNetV2., está previamente entrenado y optimizado para clasificar imágenes según ciertas categorías, se reentrenará este modelo, pero no todas las capas, sino las iniciales y algunas finales, con el fin de aprovechar el resto de información de la red neuronal.

Se inicia importando las librerías requeridas:


```

>>import tensorflow as tf
>>assert float(tf.__version__[:3]) >= 2.3

>>import os
>>import numpy as np
>>import matplotlib.pyplot as plt

```

A continuación, se descarga e importa el conjunto de datos que se encuentran almacenados en Google Cloud-Storage:

```

>>_URL = "https://storage.googleapis.com/tesis-
tomate/tomatos_photos.zip"
>>zip_file = tf.keras.utils.get_file(origin=_URL, fname="tomatos_phot
os.zip", extract=True)
>>tomato_dir = os.path.join(os.path.dirname(zip_file), 'tomato_photos
s')

```

A continuación, se utiliza ImageDataGenerator para reescalar los datos de la imagen en valores flotantes (dividir por 255 para que los valores del tensor estén entre 0 y 1), y se llama a flow_from_directory() para crear dos generadores: uno para el conjunto de datos de entrenamiento y otro para el conjunto de datos de validación.

La división es automática entre datos de entrenamiento y datos de validación, de acuerdo con la razón que se coloque en la variable “validation_split”, por ejemplo, si se coloca 0.2 en esta variable, dividirá el conjunto de datos en 80% para entrenamiento y 20% para validación. Así mismo, se ajusta el tamaño de la imagen a 224X224, en cada iteración, estos generadores proporcionan un lote de imágenes leyendo las imágenes almacenadas y procesándolas al tamaño de tensor adecuado (224 x 224). La salida es una tupla de (imágenes, etiquetas):.

```

>>IMAGE_SIZE = 224
>>BATCH_SIZE = 64
>>datagen = tf.keras.preprocessing.image.ImageDataGenerator(rescale=1
./255, validation_split=0.2)
>>train_generator = datagen.flow_from_directory(
    tomato_dir,
    target_size=(IMAGE_SIZE, IMAGE_SIZE),
    batch_size=BATCH_SIZE,
    subset='training')
>>val_generator = datagen.flow_from_directory(
    tomato_dir,
    target_size=(IMAGE_SIZE, IMAGE_SIZE),
    batch_size=BATCH_SIZE,
    subset='validation')

```

Ahora guarda las etiquetas de las clases en un archivo de texto, estas etiquetas servirán tanto para entrenar como para ejecutar el modelo:

```

>>print (train_generator.class_indices)
>>labels = '\n'.join(sorted(train_generator.class_indices.keys()))

```

```
>>with open('tomato_labels.txt', 'w') as f:
    f.write(labels)
>>!cat tomato_labels.txt
```

Ahora se crea un modelo que sea capaz de realizar transferencia de aprendizaje sólo en la última capa totalmente conectada. Se empieza con MobileNet V2 de Keras como modelo base, que está preentrenado con el conjunto de datos ImageNet (entrenado para reconocer 1.000 clases). Esto proporciona un gran extractor de características para la clasificación de imágenes y luego poder entrenar una nueva capa de clasificación con el conjunto de datos.

Al instanciar el modelo MobileNetV2, se especifica el argumento `include_top=False` para cargar la red sin las capas de clasificación en la parte superior. A continuación, se establece `trainable=False` para congelar todos los pesos del modelo base. Esto convierte efectivamente el modelo en un extractor de características porque todos los pesos y sesgos preentrenados se conservan en las capas inferiores cuando se comienza a entrenar para la clasificación:

```
>>IMG_SHAPE = (IMAGE_SIZE, IMAGE_SIZE, 3)
>>base_model = tf.keras.applications.MobileNetV2(input_shape=IMG_SHAPE,
include_top=False, weights='imagenet')
>>base_model.trainable = False
```

Ahora se crea un nuevo modelo secuencial y se pasa el modelo MobileNet congelado como modelo base, y añadir nuevas capas de clasificación para poder establecer la dimensión de salida final para que coincida con el número de clases de este conjunto de datos. Adicionalmente se agrega la convolución y filtrado a las imágenes:

```
>>model = tf.keras.Sequential([base_model,
tf.keras.layers.Conv2D(filters=32, kernel_size=3,
activation='relu'),
tf.keras.layers.Dropout(0.2),
tf.keras.layers.GlobalAveragePooling2D(),
tf.keras.layers.Dense(units=5, activation='softmax')])
```

Se finaliza la configuración del modelo. Aunque este método se llama `compile()`, es básicamente un paso de configuración que se requiere antes de poder empezar a entrenar. El optimizador Adam garantiza ajustar los pesos y sesgos para aprendizaje y descarta el aprendizaje que no mejore el modelo, es decir que el modelo no desaprenda. La métrica del modelo será la exactitud (`accuracy`):

```
>>model.compile(optimizer='adam', loss='categorical_crossentropy',
metrics=['accuracy'])
```

Ahora se realiza el proceso clave, entrenar el modelo utilizando los datos proporcionados por el `train_generator` y el `val_generator` que se crea al principio. La cantidad de épocas de entrenamiento será un valor empírico de acuerdo a los resultados de los entrenamientos que se realicen:

```
>>history = model.fit(train_generator,
                      steps_per_epoch=len(train_generator),
                      epochs=??,
                      validation_data=val_generator,
                      validation_steps=len(val_generator))
```

Hasta el proceso previo, sólo se ha entrenado las capas de clasificación, es decir la última capa; los pesos de la red preentrenada no se han modificado. Una forma de aumentar la exactitud es entrenar (o "afinar") más capas del modelo preentrenado. Es decir, liberar algunas capas del modelo base y ajustar esos pesos (que fueron entrenados originalmente con 1.000 clases de ImageNet) para que se ajusten mejor a las características encontradas en este conjunto de datos, se conoce que la totalidad de capas de esta red neuronal son 154. Así que en lugar de congelar todo el modelo base, se va a congelar capas individuales, se congela sólo las 100 capas inferiores:

```
>>base_model.trainable = True
>>fine_tune_at = 100
>>for layer in base_model.layers[:fine_tune_at]:
    layer.trainable = False
```

Ahora se configura de nuevo el modelo, pero esta vez con una tasa de aprendizaje más baja (la predeterminada es 0,001, con la que se entrenó previamente), la tasa de aprendizaje para este entrenamiento se ha establecido en 1e-5:

```
>>model.compile(optimizer=tf.keras.optimizers.Adam(1e-5),
                loss='categorical_crossentropy',
                metrics=['accuracy'])
```

Ahora se ajusta todas las capas entrenables nuevas que se han habilitado. Esto comienza con los pesos que ya entrenamos en las capas de clasificación, por lo que se necesita menos épocas:

```
>>history_fine = model.fit(train_generator,
                          steps_per_epoch=len(train_generator),
                          epochs=??,
                          validation_data=val_generator,
                          validation_steps=len(val_generator))
```

Con este proceso, se puede decir que se tiene un modelo entrenado y que puede ser utilizado para clasificar, sin embargo, no es compatible aún con la tarjeta de desarrollo Dev Board de Google Coral, las librerías para las que la tarjeta está desarrollada no son las de TensorFlow sino TensorFlow Lite, por lo que se tiene que hacer una conversión del modelo. Normalmente, la creación de un modelo TensorFlow Lite es sólo unas pocas líneas de código con TFLiteConverter. Por ejemplo, el siguiente código crea un modelo básico (sin cuantificar) de TensorFlow Lite del modelo entrenado:

```

>>converter = tf.lite.TFLiteConverter.from_keras_model(model)
>>tflite_model = converter.convert()
>>with open('mobilenet_v2_1.0_224.tflite', 'wb') as f:
    f.write(tflite_model)

```

Sin embargo, este archivo .tflite todavía utiliza valores de punto flotante para los datos de los parámetros, y se necesita cuantificar completamente el modelo al formato int8. Para cuantificar completamente el modelo, se necesita realizar una cuantificación posterior al entrenamiento con un conjunto de datos representativo, lo que requiere algunos argumentos más para el TFLiteConverter, y una función que construya un conjunto de datos representativo del conjunto de datos de entrenamiento. Así que se convierte el modelo de nuevo con cuantificación post-entrenamiento:

```

# Un generador que proporciona un conjunto de datos representativo
>>def representative_data_gen():
    dataset_list = tf.data.Dataset.list_files(tomato_dir + '/*/*')
    for i in range(100):
        image = next(iter(dataset_list))
        image = tf.io.read_file(image)
        image = tf.io.decode_jpeg(image, channels=3)
        image = tf.image.resize(image, [IMAGE_SIZE, IMAGE_SIZE])
        image = tf.cast(image / 255., tf.float32)
        image = tf.expand_dims(image, 0)
        yield [image]

>>converter = tf.lite.TFLiteConverter.from_keras_model(model)
# Esto permite la cuantificación
>>converter.optimizations = [tf.lite.Optimize.DEFAULT]
# Esto establece el conjunto de datos representativo para la
cuantificación
>>converter.representative_dataset = representative_data_gen
# Esto asegura que si alguna operación no puede ser cuantificada, el
convertidor lanza un error
>>converter.target_spec.supported_ops = [tf.lite.OpsSet.TFLITE_BUILTINS_INT8]
# Para la cuantificación completa de enteros, aunque los tipos
soportados son por defecto sólo int8, lo declaramos explícitamente para
mayor claridad.
>>converter.target_spec.supported_types = [tf.int8]
# Estos establecen los tensores de entrada y salida a uint8 (añadido
en r2.3)
>>converter.inference_input_type = tf.uint8
>>converter.inference_output_type = tf.uint8
>>tflite_model = converter.convert()
>> with open('mobilenet_v2_1.0_224_quant.tflite', 'wb') as f:
    f.write(tflite_model)

```

Finalmente, se está listo para compilar el modelo para el TPU Edge. Pero primero se descarga e instala el compilador del TPU Edge utilizando el siguiente código:

```
>>! curl https://packages.cloud.google.com/apt/doc/apt-key.gpg | sudo apt-key add -
>>! echo "deb https://packages.cloud.google.com/apt coral-edgetpu-stable main" | sudo tee /etc/apt/sources.list.d/coral-edgetpu.list
>>! sudo apt-get update
>>! sudo apt-get install edgetpu-compiler
```

A continuación, se compila el modelo con la siguiente línea de código:

```
>>! edgetpu_compiler mobilenet_v2_1.0_224_quant.tflite
```

Una vez compilado el modelo, está listo para ser ejecutado por la tarjeta Dev Board de Google coral, por lo que se tendrá que descargarlo. Los archivos que se descargan son el modelo entrenado y el archivo de texto que contiene las etiquetas. Estos archivos se deben pasar a la tarjeta y se puede hacer de diferentes modos, tanto de modo local por comandos o utilizando de igual manera el servicio de Google cloud-storage para descarga de archivos para desarrolladores directamente en la tarjeta ya configurada con acceso a internet. El código para descargar el modelo entrenado y las etiquetas es el siguiente:

```
>>from google.colab import files
>>files.download('mobilenet_v2_1.0_224_quant_edgetpu.tflite')
>>files.download('tomato_labels.txt')
```

3.5.4 Carga y ejecución del modelo en la tarjeta

En esta sección se discute brevemente la configuración de la tarjeta, debido a que esta información se le encuentra disponible en la página oficial de Google Coral y no vale ser detalla en el proyecto de investigación. Se inicia con la configuración de la tarjeta y la cámara, para poder ejecutar el modelo entrenado y convertido a formato TensorFlow Lite. El proceso de configuración se especifica a continuación:

- Revisión de requerimientos: Linux, Mac o Windows 10 con Python 3 o superior instalado.
- Instalación del sistema operativo de la tarjeta a través de una tarjeta microSD.
- Instalación de las líneas de comando MDT en el computador, para interactuar con el dispositivo.
- Conectar la tarjeta al computador por medio de USB y vía consola.
- Configurar la conexión de la tarjeta a internet.
- Actualizar el sistema operativo de la Tarjeta (Mendel Linux).

Ya configurada la tarjeta y una vez descargado el modelo, se procede a cargarlo en la tarjeta, existen diferentes métodos, en este caso se utiliza una descarga directa desde Google Cloud. La Figura 22. Muestra este proceso. Para la ejecución del modelo con imágenes de video en tiempo real, se utiliza un scrip de Google específico al cual simplemente se hace la transferencia de aprendizaje, es decir se adapta su funcionamiento al modelo entonado.

```
Símbolo del sistema - mdt shell
Microsoft Windows [Versión 10.0.19043.1526]
(c) Microsoft Corporation. Todos los derechos reservados.

C:\Users\JIMMY>mdt shell
Waiting for a device...
Connecting to deft-quill at 192.168.101.2
Linux deft-quill 4.14.98-imx #1 SMP PREEMPT Tue Nov 2 02:55:21 UTC 2021 aarch64

The programs included with the Mendel GNU/Linux system are free software;
the exact distribution terms for each program are described in the
individual files in /usr/share/doc/*/copyright.

Mendel GNU/Linux comes with ABSOLUTELY NO WARRANTY, to the extent
permitted by applicable law.
Last login: Sun Mar  6 21:34:21 2022 from 192.168.101.82
mendel@deft-quill:~$ edgetpu_classify \
> --model ${TESIS}/mobilenet_v2_1.0_224_quant_edgetpu.tflite \
> --labels ${TESIS}/tomato2_label.txt
```

Figura 22. Importación de los archivos y ejecución del modelo.

4. CAPÍTULO IV. RESULTADOS Y DISCUSIÓN

4.1 Resultados de la Madurez Digital de la industria agrícola en la comunidad Puctus.

La madurez tecnológica o madurez digital en la industria agrícola puede definirse como el nivel de implantación de las nuevas tecnologías dentro de esta industria de cara a mejorar la competitividad en todos sus factores. Existen diferentes modelos de análisis de la madurez digital, estos modelos conceptuales permiten recopilar datos sobre el estado actual de desarrollo de las industrias e identificar factores de éxito, necesidades y deficiencias. Las diferentes empresas o industrias agrícolas pueden utilizar los resultados de su autoevaluación como punto de partida para la implementación de las distintas medidas estratégicas en vía hacia una digitalización dentro de los conceptos de la industria 4.0.

En este trabajo de investigación se utilizó un modelo de evaluación a partir de una encuesta con apoyo de la experiencia de sectores industriales y la academia, propuesto por Fraunhofer Austria Research GmbH y la Universidad de Tecnologías de Viena recordando que existen varios modelos de madurez digital propuestos y estos modelos han sido desarrollados de manera empírica y de acuerdo con las necesidades y realidades locales o de cada empresa. Esta encuesta ha sido adaptada a la realidad socioeconómica local con el fin de obtener resultados más reales y objetivos, que ayuden a tener una visión más específica de la realidad tecnológica de este sector a nivel local.

En la Figura 23. se observa los resultados obtenidos. Han sido representados en un gráfico radial (gráfico de constelaciones), que se ha estandarizado para figurar los resultados de madurez digital.

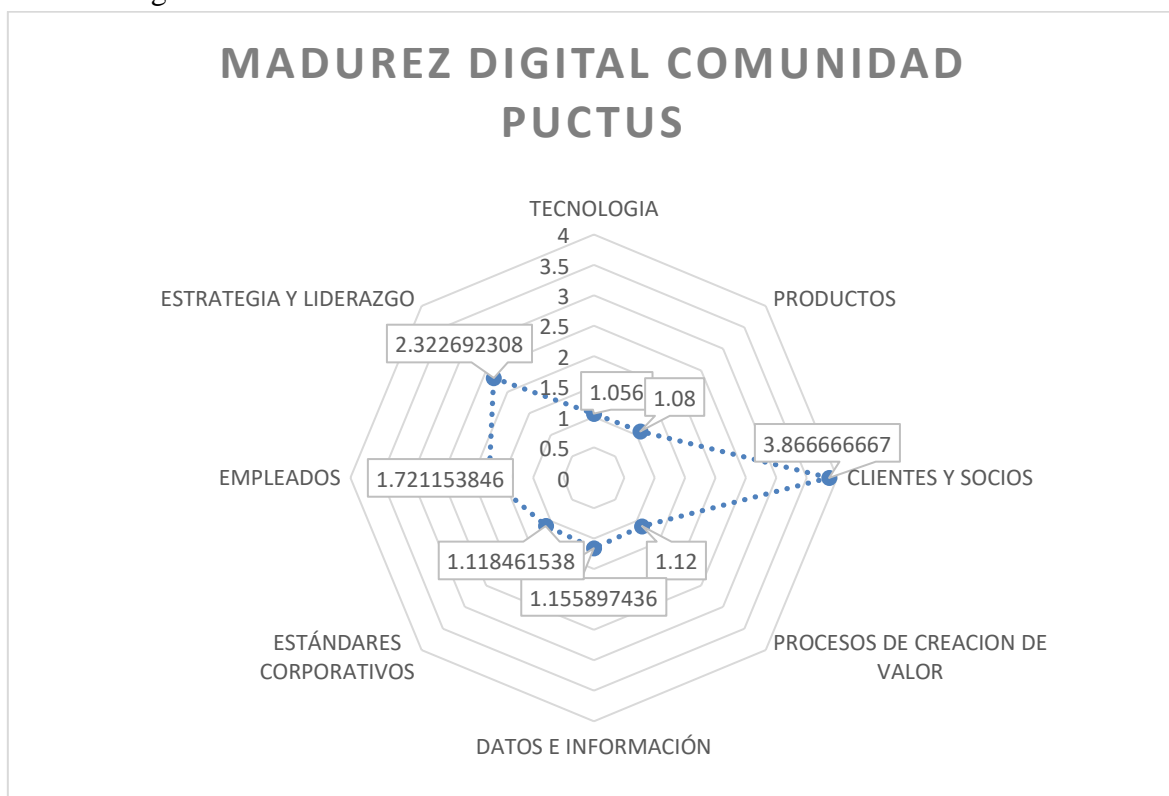


Figura 23. Madurez digital de la industria agrícola de la comunidad Puctus.

Si bien, el desarrollo tecnológico en general es bajo en el Ecuador, la madurez tecnológica en el agro también. El escaso dinamismo de la productividad en la región se debe principalmente a la poca diversificación productiva y la concentración de la producción en actividades intensivas en recursos naturales que no tienen efectos importantes en términos de derrames tecnológicos y creación de capacidades. Además, la inversión requerida resulta poco atractiva para quienes se han acostumbrado a una producción manual. Sin embargo, de los resultados se puede discutir la apertura y disponibilidad que existe por parte de clientes y socios, además de ciertos conocimientos en el uso de tecnología para comunicación. Por otra parte, de estos resultados también se concluye que prácticamente no existe digitalización de la industria agrícola en el sector de Puctus.

Así pues, se puede decir que la madurez digital en el agronegocio de la localidad no existe. El aumento de la competitividad y productividad agroindustrial dependerá de ello. Para lograr un desarrollo de agricultores, al tiempo que se crea valor añadido a sus productos y se impulsa la cualificación media, los productores deben adoptar soluciones innovadoras aprovechando el uso de la tecnología disponible.

A partir de esta evaluación, la industria agrícola local podría iniciar tu transición hacia la digitalización a través de dos fases complementarias: Digitalización de datos para toma de decisiones y la automatización de procesos y subprocesos.

4.1.1 Estadística de fiabilidad de la encuesta de madurez digital de la industria agrícola

El alfa de Cronbach no es un estadístico al uso, por lo que no viene acompañado de ningún p-valor que permita rechazar la hipótesis de fiabilidad en la escala. No obstante, cuanto más se aproxime a su valor máximo, 1, mayor es la fiabilidad de la escala. Además, en determinados contextos y por tácito convenio, se considera que valores del alfa superiores a 0,7 o 0,8 (dependiendo de la fuente) son suficientes para garantizar la fiabilidad de la escala.

Tabla 5. Estadística de escala

| Estadísticas de escala | | | |
|-------------------------------|----------|------------------|----------------|
| Media | Varianza | Desv. Desviación | N de elementos |
| 46,8400 | 10,307 | 3,21040 | 28 |

| Estadísticas de fiabilidad | |
|-----------------------------------|----------------|
| Alfa de Cronbach | N de elementos |
| ,717 | 28 |

Figura 24. Alfa de Cronbach para encuesta de Madurez digital de la industria agrícola.

Como se observa la encuesta o constructo de Madures Digital en la Industria Agrícola consta de 28 preguntas, el modelo arroja un valor de 0,717 para el Alfa de Cronbach lo cual dice

que la encuesta está bien realizada y todas las preguntas aportan positivamente con la investigación.

4.2 Implementación en el prototipo

El prototipo está conformado por 3 módulos. Como módulo principal está la tarjeta de desarrollo Google Coral, y los módulos complementarios: una pantalla LCD y la cámara coral de 5 Megapíxeles. El montaje es bastante sencillo. La Figura 25. representa el prototipo montado.

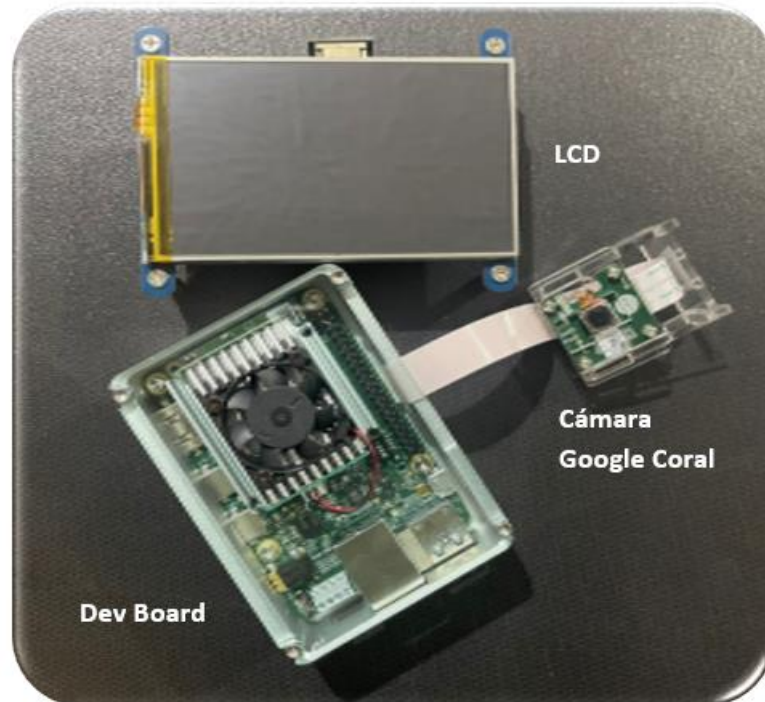


Figura 25. Prototipo montado.

4.3 Resultados de los entrenamientos

Durante el desarrollo de la investigación, como se conoce, las configuraciones iniciales del entrenamiento como el número de épocas, porcentaje de datos asignado a entrenamiento y validación, son ajustadas empíricamente de acuerdo con los resultados que se van obteniendo. Para este caso se realizaron varios entrenamientos hasta conseguir el modelo más adecuado y eficiente. Un sinnúmero de entrenamientos realizados no se reporta debido a que son irrelevantes, pero la Tabla 6 resume las configuraciones de los entrenamientos más notables que demuestran y justifican la elección por uno de ellos para ser implementado en la tarjeta.

Tabla 6. Resumen de configuraciones y resultados de los entrenamientos más notables.

| | Imágenes Training | Imágenes Validación | Split Validación | Épocas training 1 | Épocas training 2 | Accuracy Original | Accuracy Life |
|----------|--------------------------|----------------------------|-------------------------|--------------------------|--------------------------|--------------------------|----------------------|
| 1 | 350 | 87 | 0.2 | 10 | 5 | 91.667 % | 91.667 % |
| 2 | 699 | 174 | 0.2 | 10 | 5 | 90.625 % | 89.062 % |
| 3 | 1397 | 348 | 0.2 | 16 | 8 | 96.875 % | 98.438 % |
| 4 | 1224 | 521 | 0.3 | 16 | 8 | 92.188 % | 92.188 % |
| 5 | 1397 | 348 | 0.2 | 40 | 20 | 90.625 % | 90.625 % |
| 6 | 2794 | 696 | 0.2 | 16 | 8 | 92.188 % | 90.625 % |

De los resultados de la tabla 5, se ha seleccionado el modelo entrenado número 3, que es el modelo que mejor exactitud dio. El modelo seleccionado se entrenó con 16 épocas, las cuales fueron suficientes para alcanzar su exactitud, esto es gracias al uso de la transferencia de aprendizaje del modelo preentrenado especializado en extraer características de imágenes. En la Tabla 7 se observan los resultados del proceso de aprendizaje del modelo seleccionado, los datos que se muestran son tanto las pérdidas y la exactitud de la etapa de entrenamiento y de validación.

Tabla 7. Datos del proceso de aprendizaje en el Entrenamiento del modelo seleccionado

| Épocas | Perdidas de Entrenamiento | Exactitud del entrenamiento | Perdidas de validación | Exactitud de Validación |
|---------------|----------------------------------|------------------------------------|-------------------------------|--------------------------------|
| 1 | 0.6383 | 0.8239 | 0.3142 | 0.9224 |
| 2 | 0.0343 | 0.9893 | 0.3556 | 0.9138 |
| 3 | 0.0052 | 1 | 0.3730 | 0.9195 |
| 4 | 0.0036 | 1 | 0.3880 | 0.9195 |
| 5 | 0.0016 | 1 | 0.3790 | 0.9167 |
| 6 | 0.0012 | 1 | 0.3805 | 0.9339 |
| 7 | 9.9579e-04 | 1 | 0.3774 | 0.9339 |
| 8 | 6.8576e-04 | 1 | 0.3796 | 0.9368 |
| 9 | 6.4284e-04 | 1 | 0.3876 | 0.9368 |
| 10 | 4.7777e-04 | 1 | 0.3816 | 0.9339 |
| 11 | 4.5162e-04 | 1 | 0.3776 | 0.9339 |
| 12 | 4.5898e-04 | 1 | 0.3838 | 0.9368 |
| 13 | 4.2371e-04 | 1 | 0.3924 | 0.9398 |
| 14 | 2.6016e-04 | 1 | 0.3898 | 0.9339 |
| 15 | 3.3411e-04 | 1 | 0.3956 | 0.9310 |
| 16 | 2.8546e-04 | 1 | 0.3901 | 0.9339 |

La metodología de entrenamiento utilizada para esta investigación requirió un ajuste adicional al modelo entrenado. Este ajuste no requiere la misma cantidad de épocas de entrenamiento inicial, sin embargo, se liberan más neuronas para que sus pesos puedan modificarse con el nuevo ajuste. Esto hace que este modelo sea aún más preciso. Es necesario caer en cuenta que no solo no es necesario realizar un ajuste con muchas épocas sino también podría resultar contraproducente, haciendo que el modelo se sobreajuste debido a que para esta etapa se utilizan las mismas muestras del entrenamiento inicial, pero con el objetivo de modificar los pesos solamente de las neuronas nuevas que fueron liberadas. La Tabla 8

reporta los resultados del proceso de ajuste al modelo, el cual se realizó con 8 épocas solamente.

Tabla 8. Datos del proceso de aprendizaje del ajuste del modelo de entrenamiento seleccionado

| Épocas | Perdidas de Entrenamiento | Exactitud del entrenamiento | Perdidas de validación | Exactitud de Validación |
|--------|---------------------------|-----------------------------|------------------------|-------------------------|
| 1 | 0.1758 | 0.9520 | 0.3742 | 0.9167 |
| 2 | 0.0629 | 0.9821 | 0.3740 | 0.9138 |
| 3 | 0.0226 | 0.9950 | 0.3748 | 0.9167 |
| 4 | 0.0155 | 0.9957 | 0.3714 | 0.9195 |
| 5 | 0.0144 | 0.9971 | 0.3676 | 0.9224 |
| 6 | 0.0052 | 0.9993 | 0.3600 | 0.9195 |
| 7 | 0.0047 | 1 | 0.3567 | 0.9195 |
| 8 | 0.0062 | 0.9993 | 0.3560 | 0.9224 |

4.4 Curvas de Aprendizaje

Una curva de aprendizaje es una gráfica del rendimiento del aprendizaje del modelo sobre la experiencia o el tiempo, en este caso se analiza por épocas de entrenamiento.

Las curvas de aprendizaje son una herramienta de diagnóstico ampliamente utilizada en el aprendizaje automático para Algoritmos que aprenden de un conjunto de datos de entrenamiento de forma incremental. El modelo se puede evaluar en el conjunto de datos de entrenamiento y en un conjunto de datos de validación después de cada actualización durante el entrenamiento y se pueden crear gráficos del rendimiento medido para mostrar las curvas de aprendizaje.

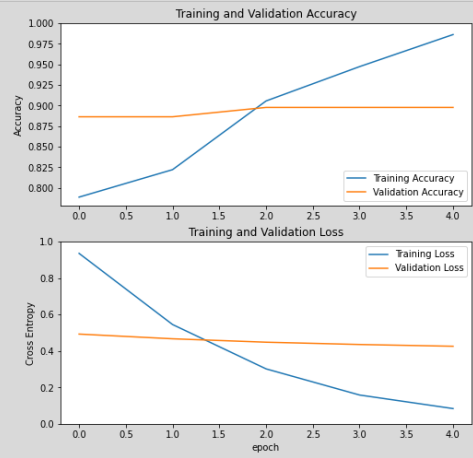
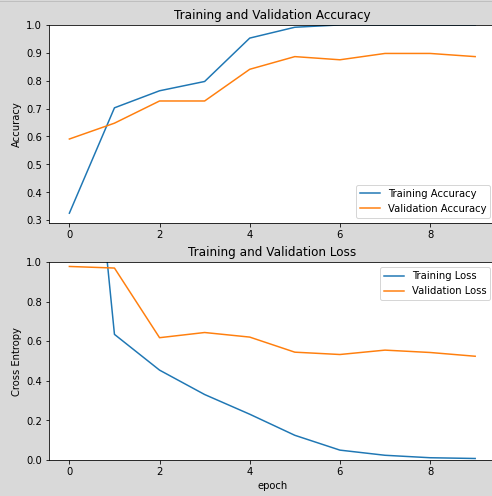
En la Figura 26. se observa las diferentes curvas de los entrenamientos realizados, el modelo con mejores resultados es el correspondiente al tres. La primera columna corresponde al entrenamiento inicial, es decir al que solo se entrenaron los pesos de la capa final. La segunda columna representa a las curvas de aprendizaje final de cada modelo, cuando se descongelaron más capas de entrenamiento para ajustar mejor al modelo. De cada entrenamiento se generan dos gráficas, la primera representa como mejora la exactitud (accuracy) por época, y la segunda es la curva de pérdidas por época.

En algunos entrenamientos, la pérdida de validación sigue siendo mayor que la pérdida de entrenamiento, por lo que podría haber un exceso de ajuste durante el entrenamiento. El sobreajuste también podría deberse a que el conjunto de datos de entrenamiento es relativamente pequeño y con menos varianza intraclase en algunas clases, en comparación con el conjunto de datos original de ImageNet utilizado para entrenar MobileNet V2. Sin embargo, en el entrenamiento 3, se acercan estas curvas, significa que no es un modelo perfecto, pero bastante preciso.

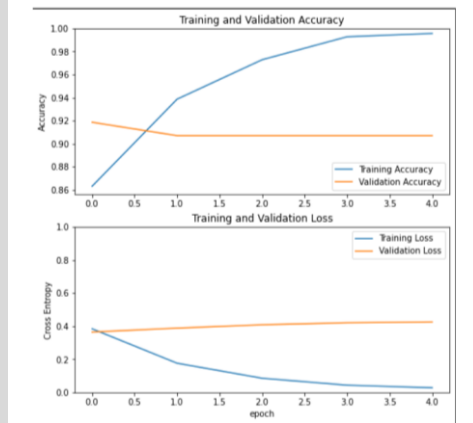
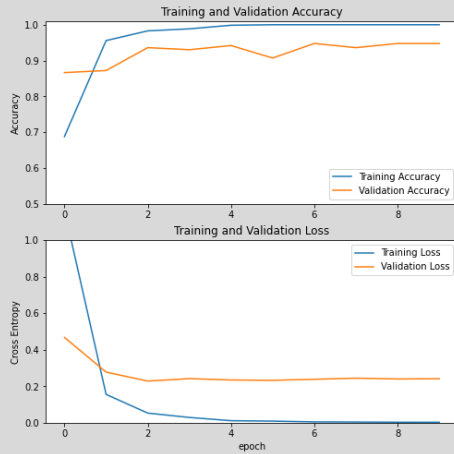
Entrenamiento Aprendizaje Inicial

Entrenamiento Aprendizaje final

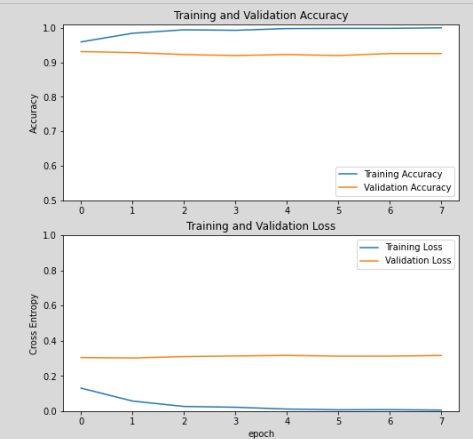
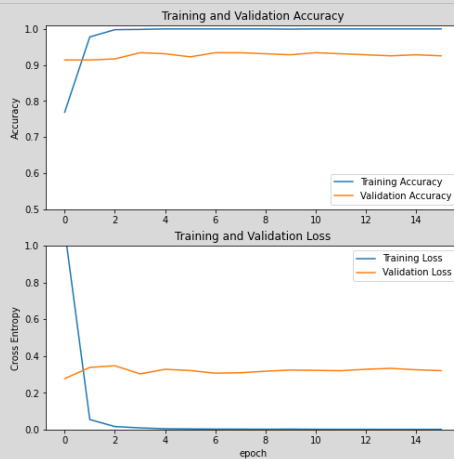
1



2



3



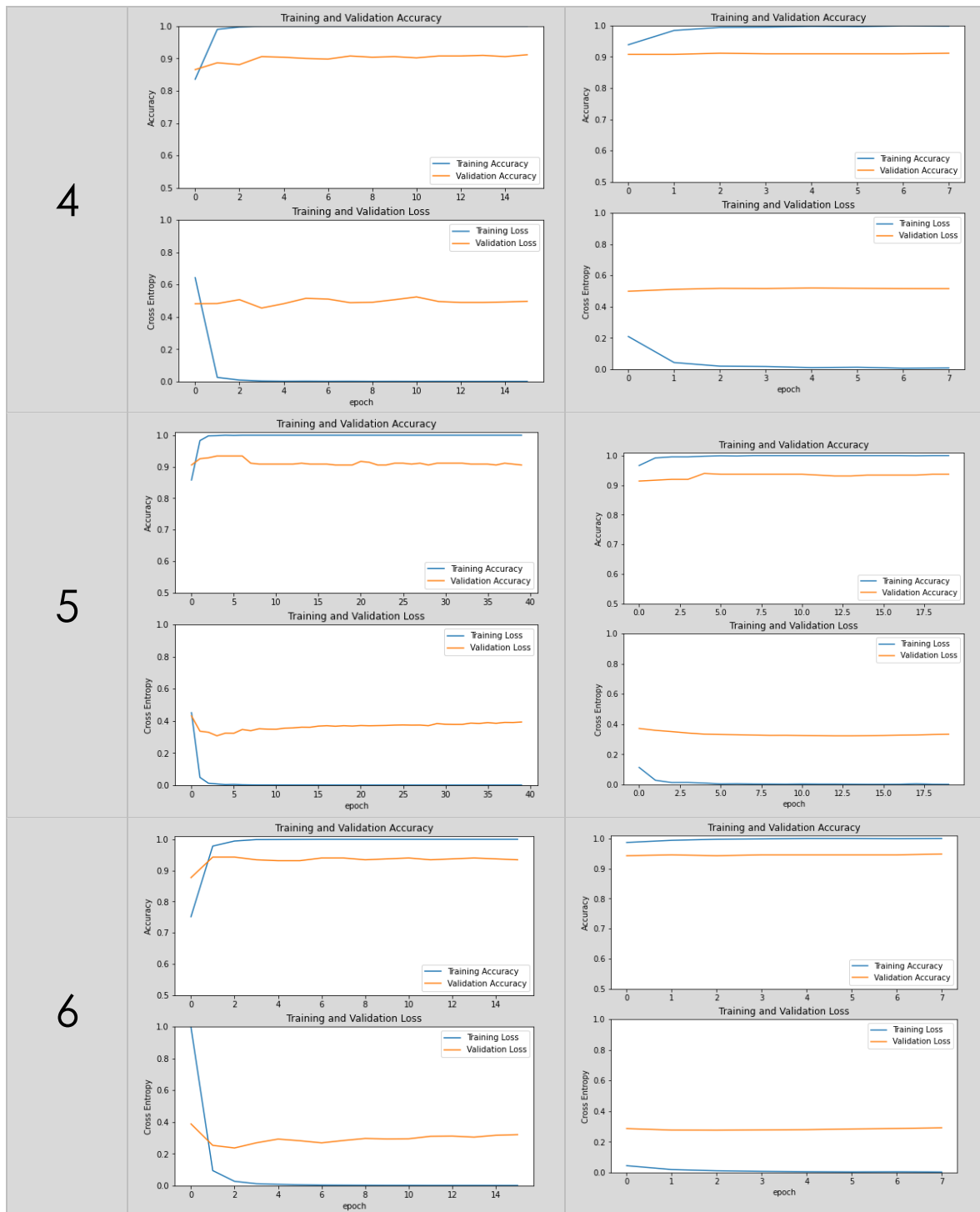


Figura 26. Curvas de modelos entrenados.

4.5 Pruebas Iniciales

Una vez montado el prototipo, cargado el modelo entrenado y configurada la tarjeta, se realizaron las pruebas iniciales a través de líneas de código por consola e imágenes cargadas directamente en la tarjeta, los resultados de estas comprobaciones iniciales son bastante alentadoras. En la Figura 27. se puede observar una captura de pantalla de estos resultados.

```

C:\Users\JIMMY>mdt devices
deft-quill (192.168.101.2)

C:\Users\JIMMY>mdt shell
Waiting for a device...
Connecting to deft-quill at 192.168.101.2
Linux deft-quill 4.14.98-imx #1 SMP PREEMPT Tue Nov 2 02:55:21 UTC 2021 aarch64

The programs included with the Mendel GNU/Linux system are free software;
the exact distribution terms for each program are described in the
individual files in /usr/share/doc/*/copyright.

Mendel GNU/Linux comes with ABSOLUTELY NO WARRANTY, to the extent
permitted by applicable law.
Last login: Sun Mar  6 21:36:48 2022
mendel@deft-quill:~$ python3 \home\mendel\algoritmo2\classify_image.py \
> --model \home\mendel\algoritmo2\mobilenet_v2_1.0_224_quant_edgetpu.tflite \
> --labels \home\mendel\algoritmo2\tomato_labels.txt \
> --input \home\mendel\algoritmo2\img26.jpg
----INFERENCE TIME----
Note: The first inference on Edge TPU is slow because it includes loading the model into Edge TPU memory.
11.9ms
3.8ms
3.0ms
2.6ms
2.6ms
-----RESULTS-----
planta_sana: 0.99609
mendel@deft-quill:~$

```

Figura 27. Prototipo montado.

4.6 Resultados del Prototipo

Posterior a realizar varias comprobaciones por consola, se configuró la tarjeta y se desarrolló el Scripts para acceder a la cámara y ejecutar el modelo sobre el vídeo en tiempo real. En este caso se usó nuevamente la técnica de Transferencia de Aprendizaje. Los resultados se observan en la Figura 28. En la parte inferior derecha de la pantalla, donde se proyecta el vídeo se visualiza el resultado de la predicción de la enfermedad, plaga o de ser el caso de una planta sana. En casos que el algoritmo tenga alta probabilidad de acertar la predicción y clasificar correctamente, imprime un solo resultado y se observa en la Figura 28 a.; en el caso de que el algoritmo detecte características que para asignar una categoría tenga como resultado altas probabilidades de pertenecer a varias, imprime todas esas posibilidades con su respectivo valor de exactitud (accuracy) y se observa en la Figura 28 b., para que el usuario final pueda verificar e identificar de las opciones que el algoritmo le propone, de ser el caso; por ejemplo, en casos que la imagen posea una enfermedad en cierta sección pero a la vez otra parte de la planta se encuentra sana, imprime ambas posibilidades con su respectivo valor de exactitud. Esto se da debido a que, en aplicaciones de agricultura, se complica identificar características tan específicas de diferentes categorías, esto se conoce como poca variación intraclase.



a)



Figura 28. a) Prototipo en prueba final de funcionamiento con una sola predicción. b) Prototipo en prueba final de funcionamiento con predicción de varias clases. c) Prototipo en prueba final en invernadero con predicción de varias clases d) Prototipo en prueba final en invernadero con una sola predicción.

4.6.1 Estadística de fiabilidad de la encuesta de validación de prototipo

Tabla 9. Estadística de escala de la encuesta de validación de prototipo

| Estadísticas de escala | | | |
|------------------------|----------|------------------|----------------|
| Media | Varianza | Desv. Desviación | N de elementos |
| 21,5000 | 3,364 | 1,83402 | 5 |

| Estadísticas de fiabilidad | |
|----------------------------|----------------|
| Alfa de Cronbach | N de elementos |
| ,839 | 5 |

Figura 29. Alfa de Cronbach para encuesta de validación de prototipo

Como se observa la encuesta o constructo de Validación del Prototipo consta de 5 preguntas, el modelo arrojo un valor de 0,839 para el Alfa de Cronbach lo cual dice que la encuesta está bien realizada y todas preguntas aportan positivamente con la investigación.

4.7 Revalidación del prototipo por agricultores expertos.

Este proceso de revalidación del prototipo se realizó mediante la puesta a prueba delante de un agricultor experto que identifique, ya sea la enfermedad o si está sana la planta, el resultado emitido por el prototipo será contrastado por el criterio del experto para verificar el funcionamiento efectivo y que la predicción sea correcta. Se realizaron diez pruebas por cada categoría en tres invernaderos, en cinco días diferentes todas durante el día en el que la iluminación se consideró casi constante; este tipo de muestreo se realizó de esta manera debido a que tanto las enfermedades como las plagas no aparecen o desarrollan un cambio relevante instantáneamente. Esto da un total de 150 pruebas por cada plaga se registraron los resultados, el promedio de aciertos de cada categoría se observa en la Figura 30. Estos resultados se compararon con la validación propia del prototipo realizado en el entrenamiento del algoritmo que se observa en la Tabla 11. La técnica utilizada para la revalidación de prototipo es la matriz de confusión.

Tabla 10. Matriz de Confusión

| | | Predicción | |
|-------------|-----------|----------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------|
| | | Positivos | Negativos |
| OBSERVACIÓN | Positivos | Verdadero Positivo (VP): Las veces que el prototipo predijo correctamente en cada plaga. | Falso Negativo (FN): Las veces que el prototipo no predijo correctamente en cada plaga. |
| | Negativos | Falso Positivo (FP): Las veces que el prototipo predice la plaga mientras se analizaba otra plaga. | Verdadero Negativo (VN): Las veces que el prototipo no predice a la plaga que no se analiza. |

El dato que se va a extraer de la matriz de confusión es la Exactitud (Accuracy).

$$Accuracy \% = \frac{VP + VN}{VP + FP + VN + FN} * 100\%$$

Ecuación 6. Accuracy de revalidación del prototipo

A continuación se reporta los resultados de cada clase. Los valores de los cuadrantes están de acuerdo con la matriz de confusión de la tabla 10.

- Minador:

| | |
|-----|-----|
| 130 | 20 |
| 21 | 579 |

$$Accuracy \text{ de Minador } \% = 94.5333$$

- Mosca Blanca:

| | |
|-----|-----|
| 127 | 23 |
| 24 | 576 |

$$Accuracy \text{ de Mosca Blanca } \% = 93.7333$$

- Fusarium:

| | |
|-----|-----|
| 129 | 21 |
| 27 | 573 |

$$Accuracy \text{ de Fusarium } \% = 93.6000$$

- Planta Sana:

| | |
|-----|-----|
| 124 | 26 |
| 29 | 571 |

$$Accuracy \text{ de Planta Sana } \% = 92.6666$$

- Planta Enferma:

| | |
|-----|-----|
| 114 | 36 |
| 47 | 553 |

$$Accuracy \text{ de Planta Enferma } \% = 88.9333$$

La exactitud promedio del prototipo es 92.693 %. Un dato menor al entregado por el entrenamiento y validación del modelo que es 98.438 % pero que fue realizado con pruebas de campo.

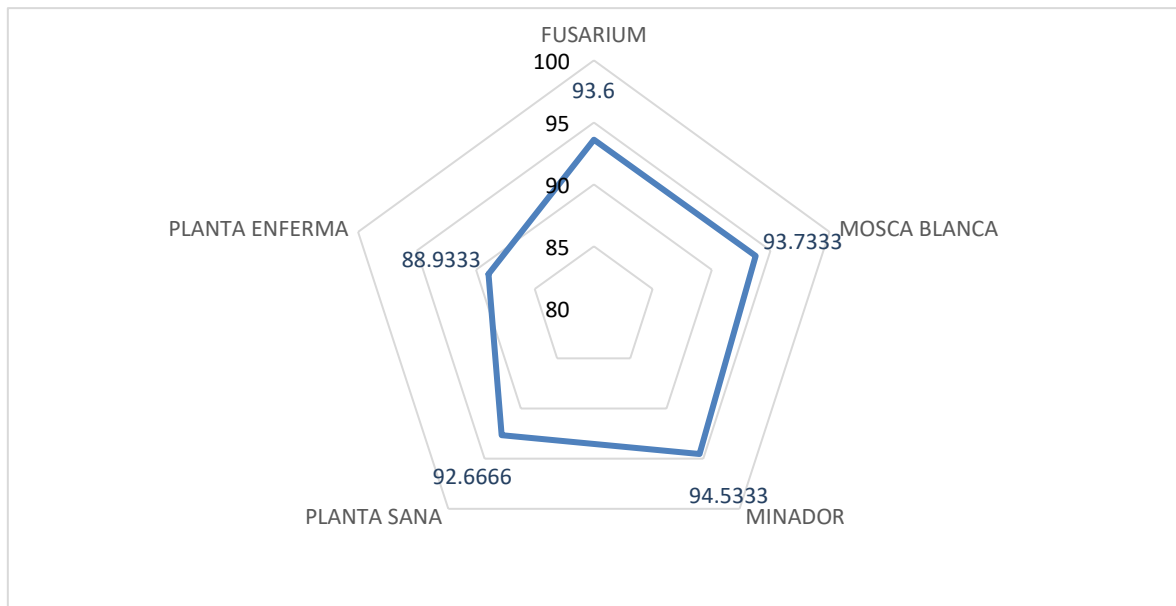


Figura 30. Revalidación del prototipo por cada clase.

Tabla 11. Comparación de las Exactitudes en el modelo TensorFlow original, modelo TensorFlow Lite y prueba real.

| | TensorFlow | TensorFlow Lite | Prueba real |
|-----------------------------|-------------------|------------------------|--------------------|
| <i>Exactitud (accuracy)</i> | 96.875 % | 98.438 % | 92.693% |

4.8 Google Coral Dev Board con conexión a Internet

La tarjeta Google Coral Dev Board, debe tener conexión a Internet, esto le permite acceder tanto a modelo preentrenados como a muestras en este caso estaban localizadas en la plataforma Google Cloud Storage

```

C:\Users\JIMMY>mdt shell
Waiting for a device...
Connecting to deft-quill at 192.168.101.2
Linux deft-quill 4.14.98-imx #1 SMP PREEMPT Tue Nov 2 02:55:21 UTC 2021 aarch64

The programs included with the Mendel GNU/Linux system are free software;
the exact distribution terms for each program are described in the
individual files in /usr/share/doc/*/copyright.

Mendel GNU/Linux comes with ABSOLUTELY NO WARRANTY, to the extent
permitted by applicable law.
Last login: Wed May 4 17:03:24 2022
mendel@deft-quill:~$ nmcli connection show
NAME                UUID                                  TYPE      DEVICE
Apple Network Jimmy  e1d8b8e7-f74c-451e-a3f5-4e40370687a3  wifi     wlan0

```

Figura 31. Coral Dev Board conectada a Internet

4.9 Arquitectura General

La arquitectura General puede acceder al almacenamiento de datos de Google Cloud Storage para poder realizar el entrenamiento de los modelos y así transferir el modelo óptimo a la tarjeta para su ejecución la cual mediante sus entradas, en este caso la cámara, realizará el procesamiento para ofrecer una respuesta.

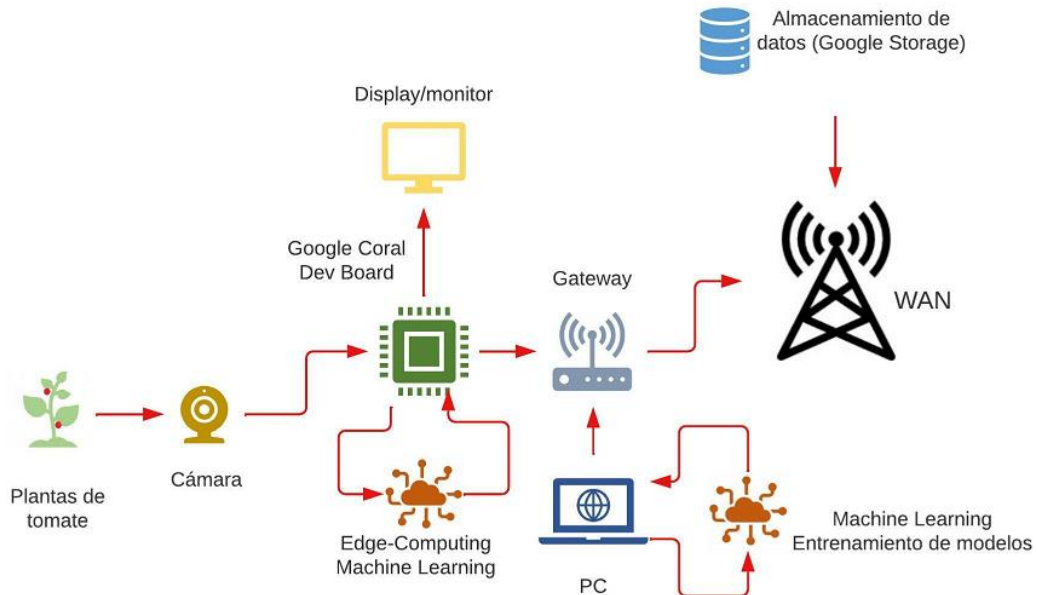


Figura 32. Arquitectura del prototipo para el entrenamiento

4.10 Arquitectura del prototipo en su funcionamiento

La arquitectura del prototipo en su funcionamiento efectivo no requiere una conexión permanente a internet, puesto que, posterior al entrenamiento fueron cargados los modelos a la tarjeta de desarrollo, la cual puede ejecutar los modelos de machine learning entrenados, sin complicaciones.

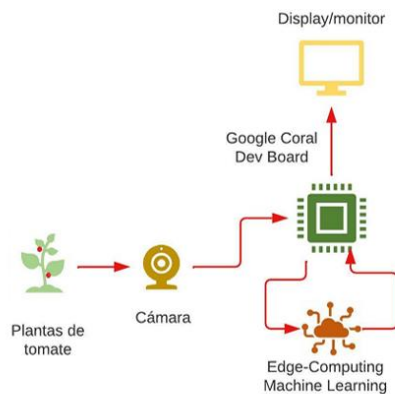


Figura 33. Arquitectura del Prototipo en funcionamiento

4.11 Propuesta de arquitectura IoT para una industria agrícola de producción de tomate.

4.11.1 Propuesta propia

Este trabajo de investigación propuso el uso del paradigma de computación Edge-Computing para procesar los datos basado en modelos entrenados de algoritmos de aprendizaje automático, por lo que la propuesta de una arquitectura IoT está basada en este modo de procesamiento de datos. La arquitectura simplificada del Edge-computing se observa en la Figura 34, en comparación con la que usa el paradigma de cloud-computing para procesar los datos. Sin embargo, la arquitectura Edge solamente muestra el bloque de procesamiento de datos, en una implementación completa, la arquitectura Edge es un complemento de la Cloud, mejorando y facilitando las acciones resultantes del procesamiento de datos, sobre todo en la latencia.

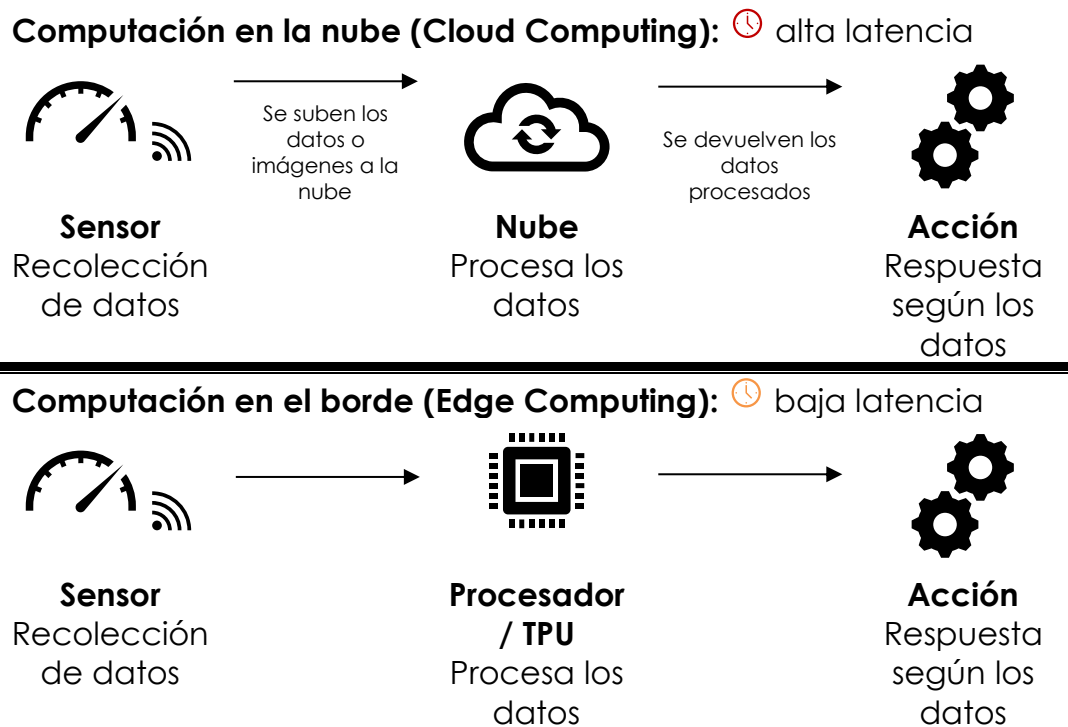


Figura 34. Comparación de arquitecturas simplificadas Edge computing y Cloud computing.

Una vez analizada brevemente las arquitecturas simplificadas, se procede a realizar una propuesta específica, para este caso se ha realizado en 4 fases:

FASE 1: Sensores y Actuadores

Las propuestas de arquitecturas comienzan con la identificación de sensores y actuadores que según la industria y sus especificaciones sean requeridos. En arquitecturas IoT, los sensores y actuadores son dispositivos conectados y/o

interconectados que monitorean (en el caso de los sensores) o controlan (en el caso de los actuadores) alguna "cosa" (término que se deriva del concepto del internet de las cosas) o proceso físico. En la agroindustria, los sensores más utilizados son aquellos que captan datos de una condición ambiental, como la temperatura, la humedad, la composición química, el flujo en una tubería, cámaras para aplicaciones de visión artificial.

En ciertas aplicaciones, el sensor puede detectar una condición o evento que necesite una respuesta casi instantánea para que un actuador pueda ejecutar acciones correctivas en tiempo real, como ajustar el caudal de un fluido, o los movimientos de un robot industrial. En estos casos, se requiere una latencia mínima entre el sensor y el actuador. Para evitar el retraso de un viaje de ida y vuelta de los datos a la nube para que sean procesados, analizados y el posterior envío del control a la "cosa", este procesamiento crítico se realiza muy cerca del proceso de monitoreo y control. Este procesamiento "en el borde" conocido como Edge-Computing puede realizarse mediante un dispositivo de sistema en módulo (SOM), como un Módulo Digi ConnectCore y/o un Router celular Digi con Python, Raspberry, y en este caso específico en una tarjeta de desarrollo Google Coral Dev Board.

FASE 2: Pasarelas de Internet y sistemas de adquisición de datos

Un sistema de adquisición de datos (DAS) recopila datos sin procesar de los sensores y convierte esos datos de un formato analógico a un formato digital. Luego, DAS recopila y formatea los datos antes de enviarlos por Internet a través de una WAN inalámbrica (como Wi-Fi o celular) o una WAN con cable para la siguiente etapa de procesamiento.

En este punto, el tamaño de los datos es el máximo. La cantidad de información puede ser bastante considerable, especialmente en industrias donde cientos de sensores pueden recopilar datos al mismo tiempo. Como resultado, los datos también se filtran y comprimen al tamaño óptimo para la transmisión. Los equipos que generalmente se utilizan en esta fase son los Gateway y los routers industriales.

FASE 3: Preprocesamiento: Análisis en la frontera

Una vez que los datos de IoT estén digitalizados y agregados, deberá procesarlos para reducir aún más la capacidad de datos antes de pasar al centro de datos o a la nube. Los dispositivos de borde pueden realizar una serie de análisis como parte del pretratamiento. El aprendizaje automático puede ser muy útil en este punto para ejecutar sistemas y mejorar continuamente las operaciones sin esperar instrucciones del centro de datos corporativo o la nube. Este tipo de procesamiento generalmente ocurre en un dispositivo ubicado cerca de donde se encuentran los sensores, como en un armario de cableado en el sitio.

Este trabajo de investigación se centra en esta fase, las operaciones de aprendizaje automático se realizan en dispositivos de borde sin la necesidad de ser procesados en

la nube. Los datos utilizados son imágenes, que en un procesamiento en la nube produciría mayor latencia.

FASE 4: Análisis en profundidad en la nube o en el centro de datos

En este trabajo de investigación, esta fase no ha sido estudiada a profundidad y tampoco utiliza computación en la nube, sin embargo, un servicio de la nube de Google ha sido utilizado para la etapa de entrenamiento.

Se puede utilizar potentes sistemas informáticos para analizar, gestionar y almacenar los datos de forma segura. Esto suele tener lugar en el centro de datos corporativo o en la nube, donde se puede combinar los datos de múltiples sitios/sensores de campo para proporcionar una imagen más amplia del sistema global de IoT y ofrecer información procesable tanto a los responsables de TI como a los de la empresa. Una empresa puede tener operaciones en diferentes zonas geográficas y los datos del IoT pueden analizarse para identificar tendencias y patrones clave, o para detectar anomalías.

A este nivel, se pueden utilizar aplicaciones específicas del sector y/o de la empresa para realizar análisis en profundidad y aplicar reglas de negocio para determinar si es necesario tomar medidas. Los datos entrantes pueden indicar cambios deseables en la configuración de los dispositivos u otras formas de optimizar el proceso, formando un bucle que facilita la mejora constante. La fase 4 también incluye el almacenamiento de datos, para el análisis y para una migración posterior.

Existen diferentes plataformas comerciales que ayudan a este tipo de análisis y operaciones en la nube que ya han sido optimizadas para este fin y se complementan con aplicaciones específicas de desarrolladores según los requerimientos de las industrias, por mencionar la más conocida y usada a nivel empresarial, SAP, que es una plataforma de desarrollo de programas de sistemas de análisis.

En la figura 35 se muestra la arquitectura IoT simplificada resultante del análisis de las cuatro fases precedentes.

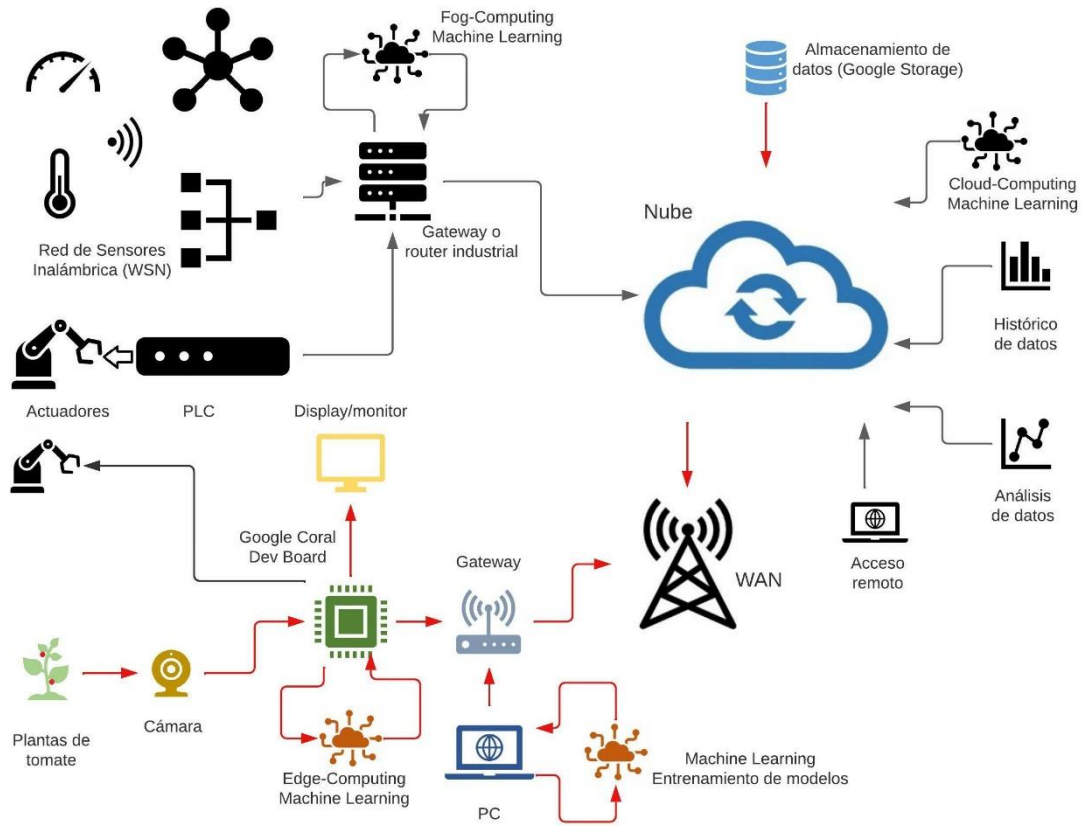


Figura 35. Arquitectura implementada y propuesta de soluciones adicionales para una agroindustria.

5. CAPÍTULO V. CONCLUSIONES y RECOMENDACIONES

5.1 Conclusiones.

Los niveles de digitalización de la industria agrícola son prácticamente nulos, por lo que existen muchos desafíos tecnológicos en este sector con el fin de mejorar su competitividad, incluso a nivel global.

Fueron necesarios muchos entrenamientos del modelo, sobre todo inicialmente, hasta encontrar los parámetros de configuración más adecuados y que den mejores resultados en su exactitud tanto en entrenamiento como en validación.

Las plagas “minador” y “mosca blanca” tienen poca variación interclase, es decir, las fotografías de sus hojas tienen características similares, lo que hacía que el algoritmo de clasificación tenga errores en la identificación entre ellas. Por lo que fue necesario buscar otra característica visual para identificar esta plaga, en este caso en el fruto.

Para el entrenamiento del modelo de clasificación de imágenes, se utilizó la técnica de “Transferencia de Aprendizaje”, es decir, se utiliza un modelo de clasificación previamente entrenado (en este caso MobilnetV2 de Google) que se ha especializado en la extracción de características de las imágenes, del cual se ha debido entrenar solamente los pesos de la capa final para definir las clases, y ciertos pesos de las capas centrales para la extracción de características específicas. Esto ha proporcionado una ventaja en recursos y tiempo.

Para esta investigación, no fue necesario entrenamientos con una gran cantidad de épocas, pues en una prueba realizada y reportada en este trabajo, una gran cantidad de épocas no mejoraba la exactitud del algoritmo y solamente era un consumo de recursos de procesamiento innecesario.

La calidad de los datos de entrenamiento fue muy importante, dado que, en pruebas iniciales, se tenía demasiada variación intraclase, lo que impedía al algoritmo generalizar y predecir correctamente. Esto se solucionó revisando y reclasificando manualmente cada fotografía con la ayuda de un experto, y revisando que la fotografía muestre claramente la plaga.

La unidad de procesamiento de la tarjeta Google Coral, conocido como TPU permitió ejecutar un modelo de clasificación en tiempo real, que pueden ser usadas en aplicaciones embebidas, por su alta capacidad de procesamiento. Su arquitectura está optimizada para procesamiento de imágenes y ejecución de modelos de machine learning.

Las curvas de aprendizaje finales, de modelo seleccionado, da a entender que el prototipo tendrá una alta exactitud, sin requerir demasiadas épocas de entrenamiento, sin embargo, en las pruebas reales, este valor de eficiencia se redujo un poco, pero sin alejarse de un valor aceptable.

Para usar adecuadamente los datos de entrenamiento, fue necesaria la carga de los datos en la nube, bajo los servicios de Google Cloud para desarrolladores. Lo que permite que estos datos sean directamente accesibles por aplicaciones, como es este caso. Este servicio tiene costo, pero este caso se utiliza la prueba de 30 días con un límite de capacidad de almacenamiento.

El modelo de entrenamiento propuesto en este trabajo de investigación puede ser adaptado a más datos, si es el caso del uso de cámaras multiespectrales, y ser reentrenado para mejorar su exactitud en la predicción.

Un aporte muy importante de este proyecto de investigación, tan importante como la misma solución a la producción agrícola de tomate, planteada como objetivo, es: El algoritmo para entrenamiento de modelos de clasificación de imágenes, la información para usarlo y modificar sus parámetros de entrenamiento, cómo convertir a un formato ejecutable en la tarjeta de desarrollo, cómo implementarlo en una aplicación de vídeo en tiempo real. Todo este aporte científico, usado en el desarrollo de este proyecto de investigación, es el resultado de la combinación ingeniosa de soluciones, la adaptación y modificación de varios algoritmos disponibles y varios métodos que son usados por separado en algunos casos; y sobre todo presenta una flexibilidad que se ha propuesto para que pueda ser utilizado en otros proyectos de investigación que requieran aplicaciones de inteligencia artificial para clasificación, en sistemas embebidos bajo el concepto de Edge-Computing.

5.2 Recomendaciones.

Dado que se identificó problemas en la diferenciación entre dos clases, minador y mosca blanca, el prototipo puede ser mejorado utilizando cámaras multiespectrales. Esta recomendación se basa en investigaciones actuales propiamente en aplicaciones agroindustriales, donde las cámaras convencionales no son suficientes para capturar características tan específicas, produciendo el problema de poca variación interclases.

Una continuación a esta investigación puede ser la adaptación de este prototipo a un sistema automático, capaz de tomar decisiones de fumigación de acuerdo con la identificación de las plagas, consiguiendo una optimización de recursos.

Este prototipo genera información continuamente, por lo que se recomienda una ampliación a la investigación mediante el uso de base de datos, que permita obtener información histórica del estado de la producción en el aspecto de plagas, y se pueda registrar y analizar dicha información en aplicaciones informáticas que monitoree y documente la producción, o incluso sea posible la aplicación de técnicas de inteligencia artificial que ayuden a la futura producción.

Se puede recomendar, por experiencia en esta investigación, usar el uso de la técnica de “Transferencia de Aprendizaje” porque permite aprovechar los conocimientos ya adquiridos de ciertos algoritmos previamente entrenados y especializados en ciertas tareas, como es el caso de extracción de características de imágenes. Esto permite optimizar bastante tiempo y recurso de procesamiento, en el caso que no estén disponibles para el uso.

6. BIBLIOGRAFÍA

- [1] D. Meng, L. Zhang, G. Cao, W. Cao, G. Zhang, and B. Hu, "Liver Fibrosis Classification Based on Transfer Learning and FCNet for Ultrasound Images," *IEEE Access*, vol. 5, no. 1, pp. 5804–5810, 2017, doi: 10.1109/ACCESS.2017.2689058.
- [2] P. Kavitha and S. Prabakaran, "A novel hybrid segmentation method with particle swarm optimization and fuzzy c-mean based on partitioning the image for detecting lung cancer," *Int. J. Eng. Adv. Technol.*, vol. 8, no. 5, pp. 1223–1227, 2019, doi: 10.20944/preprints201906.0195.v1.
- [3] H. M. Veit, "Processing Techniques," *Top. Mining, Metall. Mater. Eng.*, pp. 13–17, 2017, doi: 10.1007/978-3-319-15714-6_3.
- [4] J. J. Titano *et al.*, "Automated deep-neural-network surveillance of cranial images for acute neurologic events," *Nat. Med.*, vol. 24, no. 9, pp. 1337–1341, 2018, doi: 10.1038/s41591-018-0147-y.
- [5] Y. Zhang, M. Zhao, L. Yan, T. Gao, and J. Chen, "CNN-Based Anomaly Detection for Face Presentation Attack Detection with Multi-Channel Images," *2020 IEEE Int. Conf. Vis. Commun. Image Process. VCIP 2020*, no. December, pp. 189–192, 2020, doi: 10.1109/VCIP49819.2020.9301818.
- [6] C. Ausay, "RESPUESTA DE TOMATE RIÑÓN (*Lycopersicum esculentum* Mill) Cv DOMINIC BAJO INVERNADERO A DOS RELACIONES NITRATO/AMONIO MEDIANTE FERTIRIEGO POR GOTEO. AUSAY BASANTES ELVIA CRISTINA TESIS PRESENTADA COMO REQUISITO PARCIAL PARA OBTENER EL TÍTULO DE INGENIERO A," *Esc. Super. Politécnica Chimborazo*, 2015, [Online]. Available: [http://dspace.esPOCH.edu.ec/bitstream/123456789/4264/3/13T0808 .pdf](http://dspace.esPOCH.edu.ec/bitstream/123456789/4264/3/13T0808.pdf).
- [7] N. Petrellis, "A review of image processing techniques common in human and plant disease diagnosis," *Symmetry (Basel)*, vol. 10, no. 7, 2018, doi: 10.3390/sym10070270.
- [8] Á. Sola-Soria, "Modelo analítico de calidad de fruto en explotaciones agrícolas dedicadas al cultivo de tomate a partir de visión artificial.," Universidad de Alcalá, 2020.
- [9] "Control de calidad • gestiopolis." <https://www.gestiopolis.com/control-de-calidad/> (accessed Jul. 24, 2021).
- [10] Y. Kawasaki, H. Uga, S. Kagiwada, and H. Iyatomi, "Basic Study of Automated Diagnosis of Viral Plant Diseases Using Convolutional Neural Networks," doi: 10.1007/978-3-319-27863-6.
- [11] C. DeChant *et al.*, "Automated Identification of Northern Leaf Blight-Infected Maize Plants from Field Imagery Using Deep Learning," *Phytopathology*, vol. 107, no. 11, pp. 1426–1432, Nov. 2017, doi: 10.1094/PHYTO-11-16-0417-R.
- [12] Y. Lu, S. Yi, N. Zeng, Y. Liu, and Y. Zhang, "Identification of rice diseases using deep convolutional neural networks," *Neurocomputing*, vol. 267, pp. 378–384, Dec. 2017, doi: 10.1016/J.NEUCOM.2017.06.023.

- [13] J. Wang, L. Chen, J. Zhang, Y. Yuan, M. Li, and W. H. Zeng, "CNN transfer learning for automatic image-based classification of crop disease," *Commun. Comput. Inf. Sci.*, vol. 875, pp. 319–329, 2018, doi: 10.1007/978-981-13-1702-6_32.
- [14] W. D. N. Pacheco and F. R. J. López, "Tomato classification according to organoleptic maturity (coloration) using machine learning algorithms K-NN, MLP, and K-Means Clustering," *2019 22nd Symp. Image, Signal Process. Artif. Vision, STSIVA 2019 - Conf. Proc.*, 2019, doi: 10.1109/STSIVA.2019.8730232.
- [15] "google-coral." <https://github.com/google-coral> (accessed Jan. 31, 2022).
- [16] C. Rocha-j, F. M. Chavero, E. Guevara-cabezas, and E. H. Fort, "Industry 4 . 0 : A Proposal of Paradigm Organization Schemes from a Systematic Literature Review," 2022.
- [17] A. Moef, R. Pellerin, S. Lamouri, S. Tamayo-Giraldo, and R. Barbaray, "The industrial management of SMEs in the era of Industry 4.0," <https://doi.org/10.1080/00207543.2017.1372647>, vol. 56, no. 3, pp. 1118–1136, Feb. 2017, doi: 10.1080/00207543.2017.1372647.
- [18] W. V. B. de Souza *et al.*, "Using Crowdstorm to Prospect Innovations in Federal Institutions of Education in Brazil to Reduce Its Consumption of Electric Energy," *49th Hawaii Int. Conf. Syst. Sci. (HICSS 2016)*, pp. 2819–2828, 2016, doi: 10.1109/HICSS.2016.353.
- [19] A. Rojko, "Industry 4.0 Concept: Background and Overview," *Int. J. Interact. Mob. Technol.*, vol. 11, no. 5, pp. 77–90, Jul. 2017, doi: 10.3991/IJIM.V11I5.7072.
- [20] L. Da Xu, E. L. Xu, and L. Li, "Industry 4.0: state of the art and future trends," <https://doi.org/10.1080/00207543.2018.1444806>, vol. 56, no. 8, pp. 2941–2962, 2018, doi: 10.1080/00207543.2018.1444806.
- [21] J. I. P. Osmá, F. L. M. Salazar, and K. N. M. Gómez, "Knowledge Management and Industry 4.0 and Open Innovation," *Ing. Solidar.*, vol. 16, no. 2, May 2020, doi: 10.16925/2357-6014.2020.02.09.
- [22] K. Stachová, J. Papula, Z. Stacho, and L. Kohnová, "External Partnerships in Employee Education and Development as the Key to Facing Industry 4.0 Challenges," *Sustain. 2019, Vol. 11, Page 345*, vol. 11, no. 2, p. 345, Jan. 2019, doi: 10.3390/SU11020345.
- [23] J. H. J. Yun and Z. Liu, "Micro- and Macro-Dynamics of Open Innovation with a Quadruple-Helix Model," *Sustain. 2019, Vol. 11, Page 3301*, vol. 11, no. 12, p. 3301, Jun. 2019, doi: 10.3390/SU11123301.
- [24] C. Vila, D. Ugarte, J. Ríos, and J. V. Abellán, "Project-based collaborative engineering learning to develop Industry 4.0 skills within a PLM framework," *Procedia Manuf.*, vol. 13, pp. 1269–1276, Jan. 2017, doi: 10.1016/J.PROMFG.2017.09.050.
- [25] S. Kim and S. Park, "CPS(Cyber Physical System) based Manufacturing System Optimization," *Procedia Comput. Sci.*, vol. 122, pp. 518–524, Jan. 2017, doi: 10.1016/J.PROCS.2017.11.401.
- [26] J. Lee, B. Bagheri, and H. A. Kao, "A Cyber-Physical Systems architecture for Industry 4.0-based manufacturing systems," *Manuf. Lett.*, vol. 3, pp. 18–23, Jan. 2015, doi: 10.1016/J.MFGLET.2014.12.001.

- [27] T. M. Fernández-Caramés, P. Fraga-Lamas, M. Suárez-Albela, and M. A. Díaz-Bouza, “A fog computing based cyber-physical system for the automation of pipe-related tasks in the industry 4.0 shipyard,” *Sensors (Switzerland)*, vol. 18, no. 6, 2018, doi: 10.3390/s18061961.
- [28] P. O’Donovan, C. Gallagher, K. Bruton, and D. T. J. O’Sullivan, “A fog computing industrial cyber-physical system for embedded low-latency machine learning Industry 4.0 applications,” *Manuf. Lett.*, vol. 15, pp. 139–142, Jan. 2018, doi: 10.1016/J.MFGLET.2018.01.005.
- [29] F. I. J. Ramirez and J. M. J. Barrionuevo, “Cyber-physical system for quality control of spur gears through artificial vision techniques,” *2019 IEEE 4th Ecuador Tech. Chapters Meet. ETCM 2019*, Nov. 2019, doi: 10.1109/ETCM48019.2019.9014881.
- [30] S. Munirathinam, “Industry 4.0: Industrial Internet of Things (IIOT),” *Adv. Comput.*, vol. 117, no. 1, pp. 129–164, Jan. 2020, doi: 10.1016/BS.ADCOM.2019.10.010.
- [31] A. Oussous, F. Z. Benjelloun, A. Ait Lahcen, and S. Belfkih, “Big Data technologies: A survey,” *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 30, no. 4, pp. 431–448, Oct. 2018, doi: 10.1016/J.JKSUCI.2017.06.001.
- [32] F. A. Coda, R. M. D. Salles, F. Junqueira, D. J. S. Filho, J. R. Silva, and P. E. Miyagi, “Big data systems requirements for Industry 4.0,” *2018 13th IEEE Int. Conf. Ind. Appl. INDUSCON 2018 - Proc.*, pp. 1230–1236, Jan. 2019, doi: 10.1109/INDUSCON.2018.8627269.
- [33] S. Yin and O. Kaynak, “Big Data for Modern Industry: Challenges and Trends,” *Proc. IEEE*, vol. 103, no. 2, pp. 143–146, Feb. 2015, doi: 10.1109/JPROC.2015.2388958.
- [34] J. Yan, Y. Meng, L. Lu, and L. Li, “Industrial Big Data in an Industry 4.0 Environment: Challenges, Schemes, and Applications for Predictive Maintenance,” *IEEE Access*, vol. 5, pp. 23484–23491, Oct. 2017, doi: 10.1109/ACCESS.2017.2765544.
- [35] C. L. Philip Chen and C. Y. Zhang, “Data-intensive applications, challenges, techniques and technologies: A survey on Big Data,” *Inf. Sci. (Ny)*, vol. 275, pp. 314–347, Aug. 2014, doi: 10.1016/J.INS.2014.01.015.
- [36] V. Pilloni, “How Data Will Transform Industrial Processes: Crowdsensing, Crowdsourcing and Big Data as Pillars of Industry 4.0,” *Futur. Internet 2018, Vol. 10, Page 24*, vol. 10, no. 3, p. 24, Mar. 2018, doi: 10.3390/FI10030024.
- [37] M. S. Reis and G. Gins, “Industrial process monitoring in the big data/industry 4.0 era: From detection, to diagnosis, to prognosis,” *Processes*, vol. 5, no. 3, 2017, doi: 10.3390/pr5030035.
- [38] T. M. Fernández-Caramés, P. Fraga-Lamas, M. Suárez-Albela, and M. Vilar-Montesinos, “A Fog Computing and Cloudlet Based Augmented Reality System for the Industry 4.0 Shipyard,” *Sensors 2018, Vol. 18, Page 1798*, vol. 18, no. 6, p. 1798, Jun. 2018, doi: 10.3390/S18061798.
- [39] G. Aceto, V. Persico, and A. Pescapé, “Industry 4.0 and Health: Internet of Things, Big Data, and Cloud Computing for Healthcare 4.0,” *J. Ind. Inf. Integr.*, vol. 18, p. 100129, Jun. 2020, doi: 10.1016/J.JII.2020.100129.
- [40] M. Aazam, S. Zeadally, and K. A. Harras, “Deploying Fog Computing in Industrial Internet of Things and Industry 4.0,” *undefined*, vol. 14, no. 10, pp. 4674–4682, Oct.

- 2018, doi: 10.1109/TII.2018.2855198.
- [41] “Fog computing: la nube se prepara para el Internet de las cosas - IONOS.” <https://www.ionos.es/digitalguide/servidores/know-how/fog-computing/> (accessed Jan. 02, 2022).
- [42] P. O’Donovan, C. Gallagher, K. Bruton, and D. T. J. O’Sullivan, “A fog computing industrial cyber-physical system for embedded low-latency machine learning Industry 4.0 applications,” *undefined*, vol. 15, pp. 139–142, Jan. 2018, doi: 10.1016/J.MFGLET.2018.01.005.
- [43] D. P. Kroese, Z. I. Botev, T. Taimre, and R. Vaisman, “Data Science and Machine Learning: Mathematical and Statistical Methods,” *Data Sci. Mach. Learn.*, Nov. 2019, doi: 10.1201/9780367816971.
- [44] “Data Clustering Algorithms - k-means clustering algorithm.” <https://sites.google.com/site/dataclusteringalgorithms/k-means-clustering-algorithm> (accessed Jan. 02, 2022).
- [45] “MathWorks - Creadores de MATLAB y Simulink - MATLAB y Simulink - MATLAB & Simulink.” <https://la.mathworks.com/> (accessed Jan. 02, 2022).
- [46] O. Lorenzo Ochoa, “Modelos de madurez digital: ¿En qué consisten y qué podemos aprender de ellos?,” *Boletín Estud. económicos*, ISSN 0006-6249, Vol. 71, N° 219, 2016 (*Ejemplar Dedic. a Transform. Digit. págs. 573-590*, vol. 71, no. 219, pp. 573–590, 2016, Accessed: Jan. 02, 2022. [Online]. Available: <https://dialnet.unirioja.es/servlet/articulo?codigo=5808328&info=resumen&idioma=ENG>.
- [47] C. Cimini, R. Pinto, and S. Cavalieri, “The business transformation towards smart manufacturing: a literature overview about reference models and research agenda,” vol. 50, no. 1, pp. 14952–14957, Jul. 2017, doi: 10.1016/J.IFACOL.2017.08.2548.
- [48] L. Canetta, A. Barni, and E. Montini, “Development of a Digitalization Maturity Model for the Manufacturing Sector,” *2018 IEEE Int. Conf. Eng. Technol. Innov. ICE/ITMC 2018 - Proc.*, Aug. 2018, doi: 10.1109/ICE.2018.8436292.
- [49] “Industria Conectada 4.0 - HADA.” <https://www.industriaconectada40.gob.es/programas-apoyo/Paginas/HADA.aspx> (accessed Jan. 07, 2022).
- [50] L. Fernanda Cardona-Piedrahíta and J. Castaño-Zapata, “Comparación de métodos de inoculación de *Fusarium oxysporum* f. sp. *lycopersici* (Sacc.) Snyder & Hansen, causante del marchitamiento vascular del tomate,” 2019, doi: 10.18257/raccefy.854.
- [51] B. Ricardo Suazo, “"Identificación de defensas inducibles a mosca blanca."”
- [52] J. Salas, “Biología de *Phthorimaea operculella* (Lepidoptera: Gelechiidae) el minador grande de la hoja del tomate,” Accessed: Apr. 11, 2022. [Online]. Available: <https://www.researchgate.net/publication/347928244>.
- [53] Z. Iqbal, M. A. Khan, M. Sharif, J. H. Shah, M. H. ur Rehman, and K. Javed, “An automated detection and classification of citrus plant diseases using image processing techniques: A review,” *Comput. Electron. Agric.*, vol. 153, no. September 2017, pp. 12–32, 2018, doi: 10.1016/j.compag.2018.07.032.
- [54] The Mathworks, “Introducing Machine Learning What is Machine,” *Perspect. Ontol.*

- Learn.*, no. January 2014, 2016.
- [55] P. Ambika, “Machine learning and deep learning algorithms on the Industrial Internet of Things (IIoT),” in *Advances in Computers*, 1st ed., vol. 117, no. 1, Elsevier Inc., 2020, pp. 321–338.
 - [56] M. S. Farooq, S. Riaz, A. Abid, T. Umer, and Y. Bin Zikria, “Role of iot technology in agriculture: A systematic literature review,” *Electron.*, vol. 9, no. 2, 2020, doi: 10.3390/electronics9020319.
 - [57] Matlab Book Editors, *Mastering Machine Learning A Step-by-Step Guide with MATLAB*. 2018.
 - [58] “MathWorks - Creadores de MATLAB y Simulink - MATLAB y Simulink - MATLAB & Simulink.” <https://es.mathworks.com/> (accessed Nov. 16, 2020).
 - [59] MathWorks, S. M. Chelly, C. Denis, and MathWorks, “Getting Started with Machine Learning,” *What is Mach. Learn.*, vol. 33, no. 2, pp. 326–333, 2016, [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/11224825>.
 - [60] I. Sittón-Candanedo, R. S. Alonso, S. Rodríguez-González, J. A. García Coria, and F. De La Prieta, “Edge Computing Architectures in Industry 4.0: A General Survey and Comparison,” in *14th International Conference on Soft Computing Models in Industrial and Environmental Applications (SOCO 2019)*, 2020, pp. 121–131.
 - [61] P. O’Donovan, C. Gallagher, K. Leahy, and D. T. J. O’Sullivan, “A comparison of fog and cloud computing cyber-physical interfaces for Industry 4.0 real-time embedded machine learning engineering applications,” *Comput. Ind.*, vol. 110, pp. 12–35, 2019, doi: 10.1016/j.compind.2019.04.016.
 - [62] P. O’Donovan, C. Gallagher, K. Bruton, and D. T. J. O’Sullivan, “A fog computing industrial cyber-physical system for embedded low-latency machine learning Industry 4.0 applications,” *Manuf. Lett.*, vol. 15, pp. 139–142, 2018, doi: 10.1016/j.mfglet.2018.01.005.
 - [63] C. Yang, S. Lan, W. Shen, L. Wang, and G. Q. Huang, “Software-defined Cloud Manufacturing with Edge Computing for Industry 4.0,” *2020 Int. Wirel. Commun. Mob. Comput. IWCMC 2020*, pp. 1618–1623, 2020, doi: 10.1109/IWCMC48107.2020.9148467.

7. ANEXOS

7.1 ENCUESTA SOBRE LA MADUREZ DIGITAL DE LA INDUSTRIA AGRÍCOLA.

Estimados encuestados estoy analizando la información acerca de la madurez digital de la industria agrícola. Por favor, contesta la siguiente encuesta, la información proporcionada será utilizada para conocer la situación actual de la industria agrícola en la comunidad Puctus en la parroquia Yaruquies, cantón Riobamba.

- **TECNOLOGÍA**

Esta sección nos permite conocer cómo se encuentra la tecnológicamente la industria agrícola

¿Utiliza usted tecnología para el intercambio de información?

| | 1 | 2 | 3 | 4 | |
|-------------------------------------------------------------------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|----------------------------------------------------------------------------------------|
| Toda la información obtenida o intercambiada lo hace usted sin apoyarse de la tecnología. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | Toda la información obtenida o intercambiada lo hace usted con apoyo de la tecnología. |

¿Utiliza usted tecnología en la nube?

| | 1 | 2 | 3 | 4 | |
|-----------------------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|----------------------------------------------------------|
| Sin uso de servicios e información de la nube | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | Todos los servicios e información disponibles de la nube |

¿Almacena información de forma descentralizada?

| | 1 | 2 | 3 | 4 | |
|---------------------------------------------------------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------------------------------------------------------------------------------|
| No utiliza o al menos utiliza un dispositivo para almacenamiento de información | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | Utiliza varios dispositivos o se ayuda de la tecnología de la nube para almacenar información |

¿Utiliza sensores para recopilación de datos en el invernadero?

| | 1 | 2 | 3 | 4 | |
|-------------------------------------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|---------------------------------------------------------------------------|
| No utiliza ningún tipo de sensor para recopilar información | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | Utiliza todos los sensores que sean necesarios para recopilar información |

¿Utiliza o se ayuda de algún tipo de robot?

| | 1 | 2 | 3 | 4 | |
|-------------------------|-----------------------|-----------------------|-----------------------|-----------------------|-------------------------------------------------|
| No utiliza ningún robot | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | Todas sus actividades están apoyadas por robots |

- **PRODUCTOS**

Esta sección nos permite conocer específicamente el uso de la tecnología sobre el producto en la industria agrícola

¿Utiliza usted componentes para el procesamiento de datos del producto?

| | | | | | |
|---------|-----------------------|-----------------------|-----------------------|-----------------------|---------------------------------------------------------------|
| | 1 | 2 | 3 | 4 | |
| Ninguna | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | Totalmente toda la información de los productos se digitaliza |

¿Utiliza usted TI (tecnologías de la información) relacionado con sus productos?

| | | | | | |
|---------|-----------------------|-----------------------|-----------------------|-----------------------|---------------------------------------------|
| | 1 | 2 | 3 | 4 | |
| Ninguna | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | Totalmente la información esta digitalizada |

• **CLIENTES Y SOCIOS**

Esta sección nos permite conocer a los clientes y socios en la industria agrícola y el uso de la tecnología para interactuar con ellos.

¿Qué tan dispuesto estaría a la apertura de las nuevas tecnologías?

¿Qué tan dispuesto estaría a la apertura de las nuevas tecnologías?

| | | | | | |
|--------------------------------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------------------------|
| | 1 | 2 | 3 | 4 | |
| No, Me mantengo de la forma tradicional sin tecnología | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | Totalmente me apoyaría de la tecnología |

¿Qué tan dispuesto estaría a digitalizar el contacto que tiene con el cliente?

| | | | | | |
|--------------------------------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------------------------|
| | 1 | 2 | 3 | 4 | |
| No, Me mantengo de la forma tradicional sin tecnología | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | Totalmente me apoyaría de la tecnología |

¿Usted estaría dispuesto que la tecnología utilice los datos relacionados con el cliente para su beneficio?

| | | | | | |
|--------------------------------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------------------------|
| | 1 | 2 | 3 | 4 | |
| No, Me mantengo de la forma tradicional sin tecnología | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | Totalmente me apoyaría de la tecnología |

• **PROCESOS DE CREACION DE VALOR**

Esta sección nos permite saber el proceso en el que se encuentra la industria agrícola para aportar valor a la materia prima hasta su producto final.

¿Posee una automatización de máquinas, sensores, actuadores en el invernadero?

| | | | | | |
|----------------------------------------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|----------------------------------------------|
| | 1 | 2 | 3 | 4 | |
| No poseo de ninguna automatización todo se realiza manualmente | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | Totalmente, mi invernadero esta automatizado |

¿Posee usted la información generada de las máquinas, sensores, actuadores en el invernadero?

| | 1 | 2 | 3 | 4 | |
|---------------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|-------------------------------------------------------|
| No poseo ninguna información generada | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | Totalmente conozco la información de los dispositivos |

¿Posee usted acceso remoto a las máquinas, sensores, actuadores en el invernadero?

| | 1 | 2 | 3 | 4 | |
|--------------------------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|--------------------------------------------|
| No poseo ningún acceso remoto a los dispositivos | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | Si, poseo acceso remoto a los dispositivos |

- **DATOS E INFORMACIÓN**

Esta sección nos permitirá saber cómo los datos son evaluados en la industria agrícola

¿La recopilación de datos se ejecuta de manera automatizada?

| | 1 | 2 | 3 | 4 | |
|------------------------------------------------------------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|---------------------------------------------------------------|
| No, la recopilación de datos se hace manual, o no se realiza recopilación de datos | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | Totalmente la recopilación de datos es de manera automatizada |

¿Realiza un análisis de los datos recopilados?

| | 1 | 2 | 3 | 4 | |
|--------------------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|-------------------------------------------------------|
| No se realiza ningún análisis de los datos | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | Se realiza un análisis de todos los datos recopilados |

¿Usted toma decisiones basadas en los datos obtenidos?

| | 1 | 2 | 3 | 4 | |
|------------------------------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|-------------------------------------------------------------------|
| No tomo decisiones basadas en datos o no poseo datos | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | Totalmente mis decisiones son tomadas basadas los datos obtenidos |

- **ESTÁNDARES CORPORATIVOS**

Esta sección nos permitirá conocer a nivel empresarial superior (de ser el caso) los niveles de digitalización en la industria agrícola.

¿Tiene contratado servicios para la Industria 4.0?

| | | | | | |
|-------------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|--------------------------------------------------------------------------------------|
| | 1 | 2 | 3 | 4 | |
| No poseo servicios de Industria 4.0 | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | Tengo contratado servicios de industria 4.0 para todo el requerimiento de la empresa |

¿Sus empleados se encuentran capacitados en competencias digitales?

| | | | | | |
|------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|-------------------------------------------------------------------|
| | 1 | 2 | 3 | 4 | |
| No se encuentran capacitados | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | Todos mis empleados están capacitados para competencias digitales |

¿La información corporativa está protegida con algún nivel de seguridad cibernética?

| | | | | | |
|----------------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|--------------------------------------------|
| | 1 | 2 | 3 | 4 | |
| No posee ninguna seguridad cibernética | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | Toda la información se encuentra protegida |

- EMPLEADOS

Esta sección nos permite conocer el nivel de conocimiento de los empleados

¿Los empleados están abiertos a nuevas tecnologías?

| | | | | | |
|---------------------------------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|-------------------------------------------------------------------------------------|
| | 1 | 2 | 3 | 4 | |
| No están dispuestos a aprender sobre nuevas tecnologías | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | Totalmente todos los empleados están dispuestos a aprender sobre nuevas tecnologías |

¿Los empleados tienen competencia en TICs modernas?

| | | | | | |
|----------------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|---------------------------------------------------------|
| | 1 | 2 | 3 | 4 | |
| No tienen competencia en TICs modernas | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | Todos los empleados tienen competencia en TICs modernas |

¿Los empleados tienen experiencia en trabajos interdisciplinarios (conocimiento en varias actividades de trabajo)?

| | | | | | |
|-------------------------------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------------------------------------------------------|
| | 1 | 2 | 3 | 4 | |
| No poseen experiencia en trabajos interdisciplinarios | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | Todos los empleados tiene experiencia en trabajos interdisciplinarios |

¿Los empleados poseen la voluntad de formación continua en el puesto de trabajo?

| | | | | | |
|----------------------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|-------------------------------------------------------------------|
| | 1 | 2 | 3 | 4 | |
| No poseen voluntad de una formación continua | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | Todos los empleados poseen una voluntad de una formación continua |

- **ESTRATEGIA Y LIDERAZGO**

Esta sección nos permite conocer la estrategia y liderazgo que posee la empresa y propietarios dentro de la industria 4.0

¿Posee una planificación u hoja de ruta para la transformación hacia la Industria 4.0?

| | | | | | |
|--------------------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|----------------------------------------|
| | 1 | 2 | 3 | 4 | |
| No poseo una planificación ni hoja de ruta | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | Poseo una planificación y hoja de ruta |

¿Posee los recursos financieros para hacer realidad la industria 4.0?

| | | | | | |
|-------------------|-----------------------|-----------------------|-----------------------|-----------------------|--------------------------------------------------------------------|
| | 1 | 2 | 3 | 4 | |
| No poseo recursos | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | Poseo recursos para una transformación completa a la industria 4.0 |

¿Posee voluntad para la transformación a la industria 4.0?

| | | | | | |
|-------------------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------------------------------------------------------------------------|
| | 1 | 2 | 3 | 4 | |
| No me interesa una transformación digital | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | Estoy totalmente dispuesto a una transformación digital y al total uso de la tecnología |

¿Posee el conocimiento o esta dispuesto a una capacitación sobre la Industria 4.0?

| | | | | | |
|----------------------------------------------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|--------------------------------------------------------------------------------------------------|
| | 1 | 2 | 3 | 4 | |
| No me interesa la Industria 4.0 ni ninguna transformación tecnología | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | Poseo conocimiento y estoy dispuesto a tener capacitación si es necesaria sobre la industria 4.0 |

7.2 ENCUESTA VALIDACIÓN DE PROTOTIPO

Estimados encuestados estoy realizando varias pruebas para detección de plagas y enfermedades. Por favor conteste la encuesta la cual me ayudara a dar validez del algoritmo entrenado con base a sus conocimientos

¿De las 10 pruebas realizadas cuantas clasificó de manera correcta el prototipo en la enfermedad Fusarium?

0 1 2 3 4 5 6 7 8 9 10

¿De las 10 pruebas realizadas cuantas clasificó de manera correcta el prototipo en la plaga mosca blanca?

0 1 2 3 4 5 6 7 8 9 10

¿De las 10 pruebas realizadas cuantas clasificó de manera correcta el prototipo en la plaga minador?

0 1 2 3 4 5 6 7 8 9 10

¿De las 10 pruebas realizadas cuantas clasificó de manera correcta el prototipo en la planta sana?

0 1 2 3 4 5 6 7 8 9 10

¿De las 10 pruebas realizadas cuantas clasificó de manera correcta el prototipo en la planta enferma?

0 1 2 3 4 5 6 7 8 9 10

7.3 Fotografías de la realización de encuesta y pruebas de campo



Figura 36. Proceso de Encuesta a los agricultores en la comunidad Puctus.



Figura 37. Pruebas de campo del prototipo en el invernadero del Sr. Pedro Sucuy



Figura 38. Pruebas de campo del prototipo en el invernadero del Sr. Ramon Morocho



Figura 39. Pruebas de campo del prototipo en el invernadero de la Sra. María Quisay

7.4 Ejemplos de imágenes usadas en el entrenamiento.





| Enfermedad o Plaga | Ejemplo de imágenes de entrenamiento |
|--------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Fusarium | <p data-bbox="742 846 769 878">a)</p>  <p data-bbox="671 1281 699 1312">b)</p>  <p data-bbox="869 1621 896 1653">c)</p>  |



Figura 40. a), c) Muestra de Enfermedad Fusarium, obtenida en el invernadero del Sr. Pedro Sucuy en la etapa de floración del tomate riñón. b), d) Muestra de Enfermedad Fusarium, obtenida en el invernadero de la Sra. María Quisay en la etapa de Fructificación del tomate riñón. e) Muestra de Enfermedad Fusarium, obtenida en el invernadero del Sr. Ramon Morocho en la etapa de fructificación del tomate riñón.

| Enfermedad o Plaga | Ejemplo de imágenes de entrenamiento |
|-------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p data-bbox="427 779 616 815">Mosca Blanca</p> |  <p data-bbox="842 607 868 642">a)</p> <p data-bbox="842 1093 868 1128">b)</p> <p data-bbox="842 1608 868 1644">c)</p> <p>The table contains three photographs of whiteflies on green leaves. Image (a) shows a close-up of a leaf with several whiteflies. Image (b) shows a leaf with a few whiteflies. Image (c) shows a leaf with a large number of whiteflies.</p> |

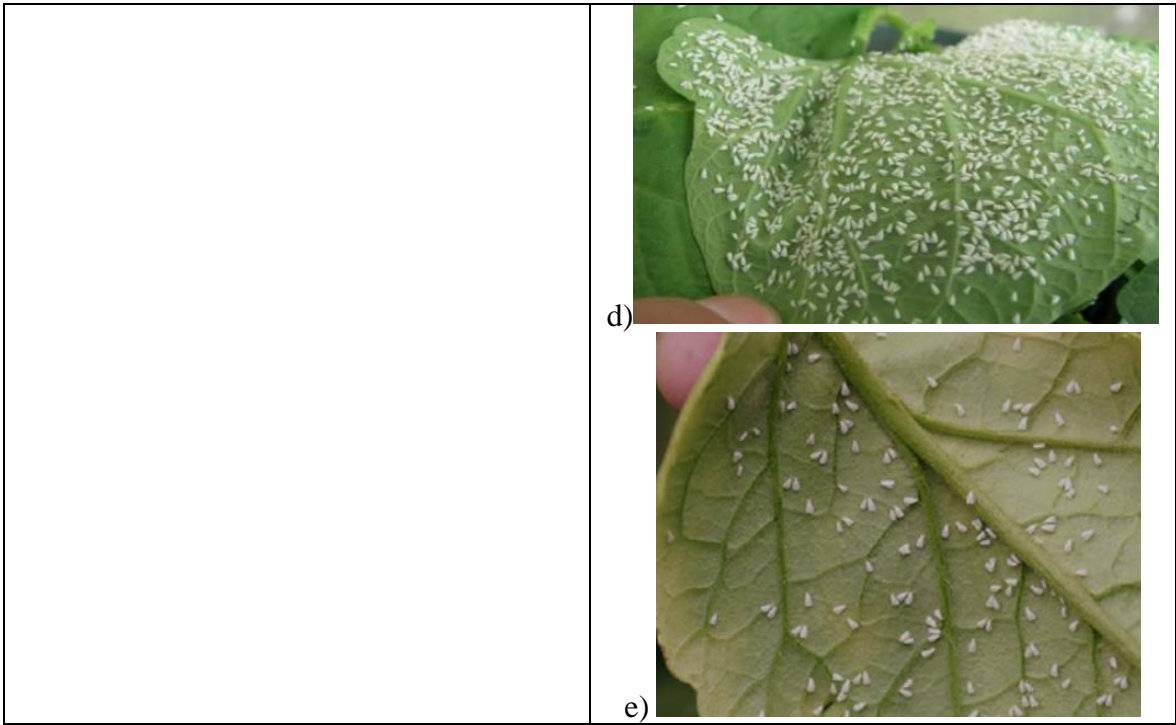



Figura 41. a), b) Muestra de plaga Mosca Blanca, obtenida en el invernadero del Sr. Ramon Morocho en la etapa de fructificación del tomate riñón. c), d) Muestra de plaga Mosca Blanca, obtenida en el invernadero de la Sra. María Quisay en la etapa de Fructificación del tomate riñón. e) Muestra de plaga Mosca Blanca, obtenida en el invernadero del Sr. Ramon Morocho en la etapa de fructificación del tomate riñón.

| Enfermedad o Plaga | Ejemplo de imágenes de entrenamiento |
|--------------------|------------------------------------------------------------------------------------------------|
| <p>Minador</p> |  <p>a)</p> |



b)



c)



d)

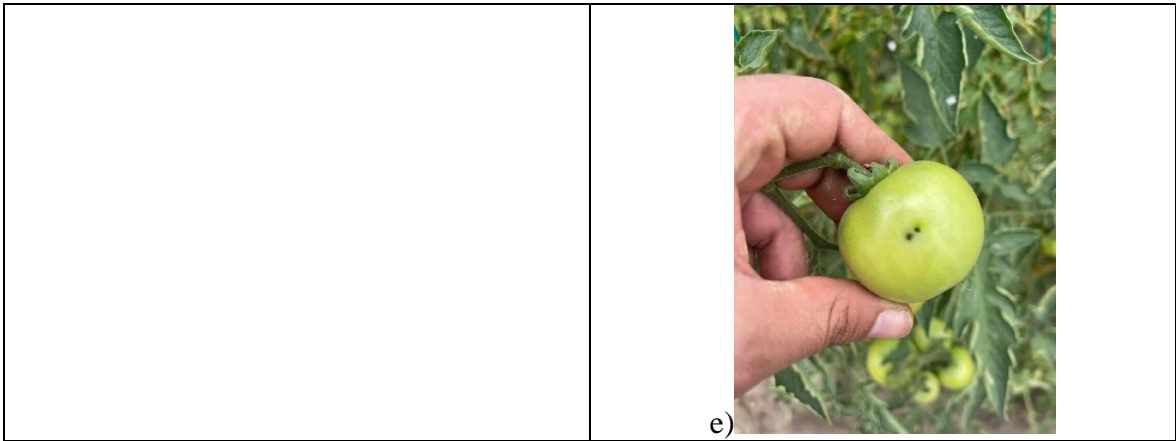


Figura 42. a), e) Muestra de plaga Minador, obtenida en el invernadero de la Sra. María Quisay en la etapa de fructificación semana 6 del tomate riñón. b), d) Muestra de plaga Minador, obtenida en el invernadero del Sr. Pedro Sucuy en la etapa de Fructificación semana 8 del tomate riñón. c) Muestra de plaga Minador, obtenida en el invernadero del Sr. Ramon Morocho en la etapa de fructificación semana 7 del tomate riñón.


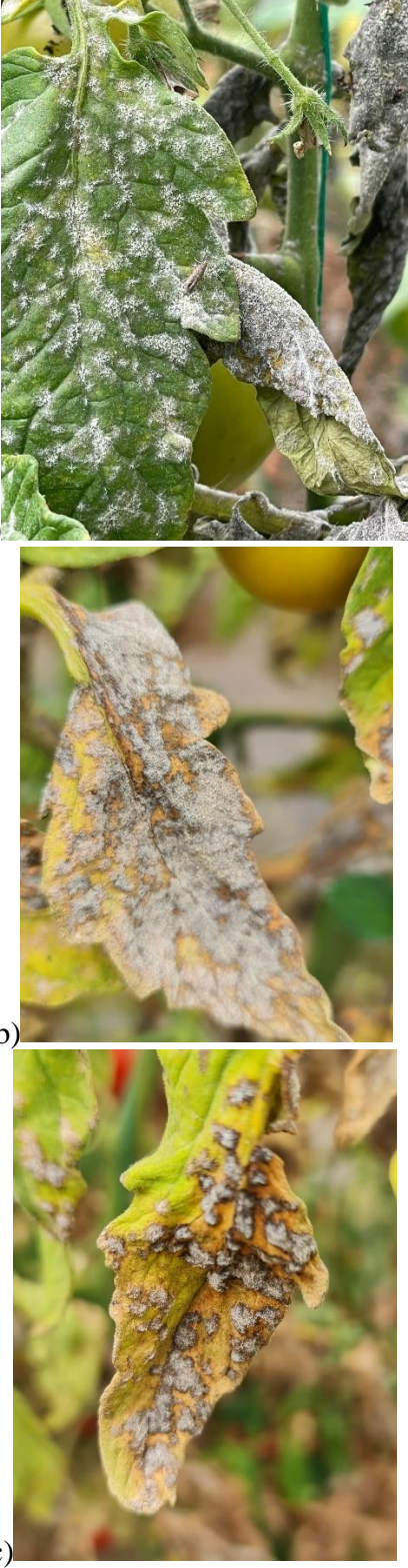
| Enfermedad o Plaga | Ejemplo de imágenes de entrenamiento |
|--------------------|---------------------------------------------------------------------------------------------------------|
| <p>Planta Sana</p> |  <p>a)</p> <p>b)</p> |



Figura 43. a), d) Muestra Planta Sana, obtenida en el invernadero del Sr. Pedro Sucuy en la etapa de fructificación semana 7 del tomate riñón. b), c) Muestra Planta Sana, obtenida en el invernadero de la Sra. María Quisay en la etapa de fructificación del tomate riñón. e) Muestra Planta Sana, obtenida en el invernadero del Sr. Ramon Morocho en la etapa de floración del tomate riñón.

| Enfermedad o Plaga | Ejemplo de imágenes de entrenamiento |
|---------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p data-bbox="347 633 692 674">Planta Enferma (cenicilla)</p> |  <p data-bbox="890 745 917 779">a)</p> <p data-bbox="906 1249 933 1283">b)</p> <p data-bbox="906 1765 933 1798">c)</p> |

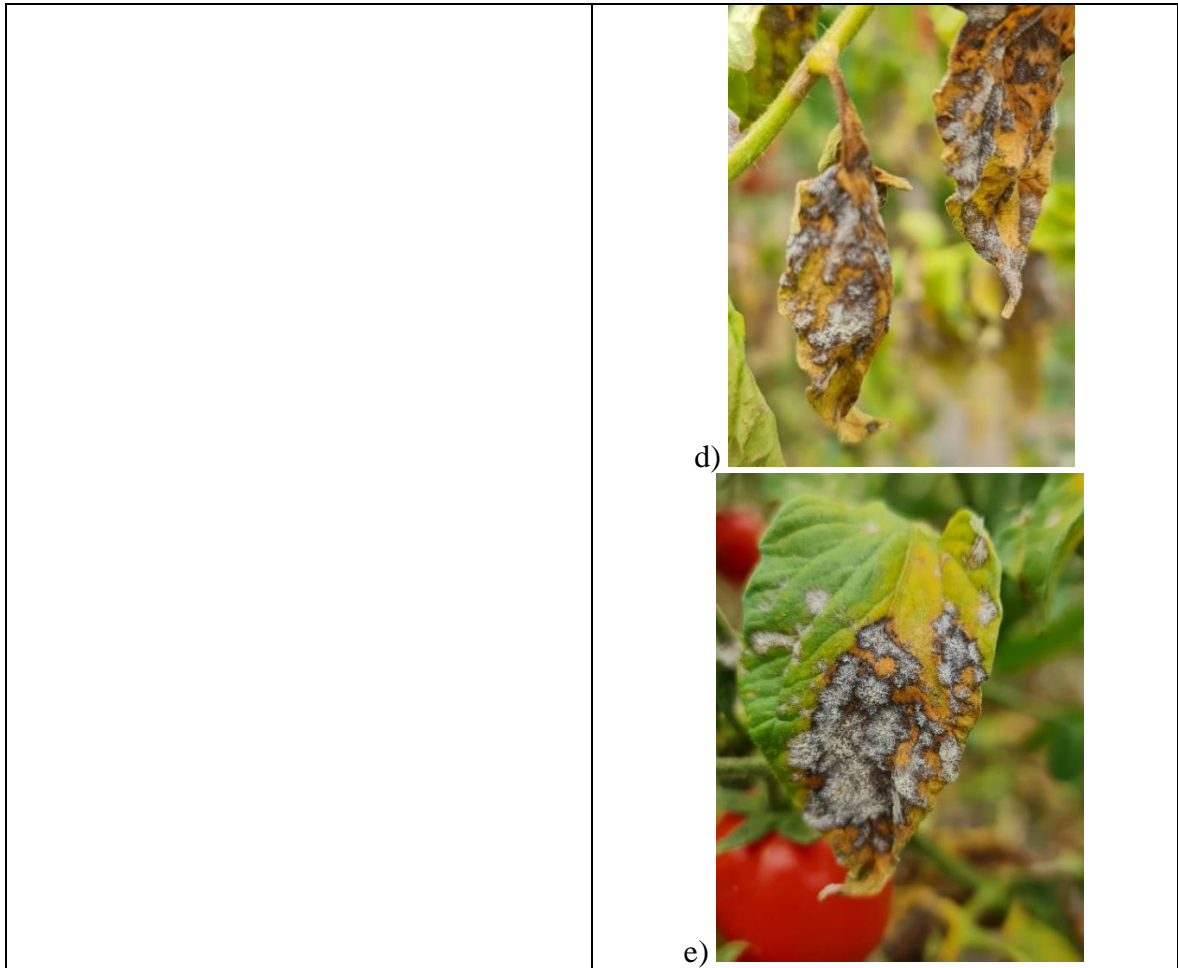


Figura 44. a), e) Muestra Planta Enferma, obtenida en el invernadero del Sr. Ramon Morocho en la etapa de fructificación del tomate riñón. b) Muestra Planta Enferma, obtenida en el invernadero del Sr. Pedro Sucuy en la etapa de fructificación del tomate riñón. c), d) Muestra Planta Enferma, obtenida en el invernadero de la Sra. María Quisay en la etapa de fructificación del tomate riñón.

7.5 Presupuesto de Prototipo

Tabla 12. Presupuesto de Prototipo

| ITEM | COSTOS |
|-----------------------------------------------------|------------------|
| Tarjeta de Desarrollo Google Coral Dev | 130.00 \$ |
| Tarjeta Micro SD de 64Gb | 20.00 \$ |
| Google Coral Cámara | 30.00 \$ |
| Protector de Tarjeta de Desarrollo Google Coral Dev | 50.00 \$ |
| Pantalla Digital 4 pulgadas | 50\$ |
| Software | 0 \$ |
| TOTAL | 780.00 \$ |