

SISTEMA MÓVIL CON MODELO JERÁRQUICO HÍBRIDO PARA ESTIMAR LA SEVERIDAD FOLIAR DEL MILDIU EN QUINUA BAJO CAPTURA CONTROLADA

M.Sc. Edgar Jaldin Torrico

Posgrado SOE – UAGRM

<https://orcid.org/0009-0005-5329-4908>

Santa Cruz, Bolivia | edgarjaldintorrico@gmail.com



<https://doi.org/10.23670/FT.2026.1.27>

Recibido 20/04/2026 - Aceptado 13/05/2026

RESUMEN

La cuantificación visual de la severidad foliar del mildiu en quinua constituye un proceso laborioso y potencialmente variable entre observadores, lo que limita su estandarización en la evaluación fenotípica y el monitoreo fitopatológico. El presente estudio tuvo como objetivo evaluar experimentalmente un sistema móvil para apoyar la cuantificación automatizada de la severidad foliar del mildiu en quinua. Se desarrolló un sistema denominado QuinuApp, compuesto por una aplicación móvil, un dispositivo de captura estandarizada (BlackBox) y un estimador jerárquico híbrido basado en variables visuales interpretables. El estudio siguió un enfoque cuantitativo con diseño experimental, utilizando 772 imágenes RGB de hojas adaxiales individuales correspondientes a 92 accesiones. La severidad, expresada en una escala continua de 0 % a 100 %, fue estimada por una fitopatóloga experta y utilizada como referencia operativa. El pipeline incluyó normalización fotométrica, segmentación de la hoja principal, extracción de 55 variables clínico-texturales

derivadas de imagen, clasificación auxiliar por fases biológicas y regresión híbrida con calibración isotónica, incorporación de la variable contextual de esporulación, ajuste contextual por accesión y reglas biológicas. En el conjunto de prueba, el modelo obtuvo Pearson $r = 0.699$, CCC = 0.694, MAE = 17.17 % y RMSE = 24.75 %, alcanzando el mejor desempeño global frente a regresión lineal, Random Forest, XGBoost y SVR. El análisis de Bland–Altman mostró un sesgo medio de -0.58 %, con 50.78 % de predicciones dentro de ± 10 % y 69.43 % dentro de ± 20 %. El sistema propuesto mostró viabilidad técnica como herramienta de apoyo para la estimación automatizada de la severidad foliar del mildiu en quinua bajo condiciones experimentales controladas, con potencial de aplicación en fenotipado digital asistido.

Palabras clave: mildiu de la quinua, severidad foliar, fenotipado digital, visión computacional, estimación interpretable.

ABSTRACT

Visual quantification of downy mildew foliar severity in quinoa is a labor-intensive process and may vary across observers, thereby limiting its standardization in phenotypic assessment and phytopathological monitoring. The present study aimed to experimentally evaluate a mobile system designed to support the automated quantification of downy mildew foliar severity in quinoa. A system named QuinuApp was developed, comprising a mobile application, a standardized image acquisition device (BlackBox), and a hierarchical hybrid estimator based on interpretable visual features. The study followed a quantitative approach under an experimental design, using 772 RGB images of individual adaxial leaves from 92 accessions. Severity, expressed on a continuous scale from 0% to 100%, was assessed by an expert plant pathologist and used as the operational reference. The pipeline included photometric normalization, main leaf segmentation,

extraction of 55 image-derived clinical-textural features, auxiliary classification by biological phases, and hybrid regression with isotonic calibration, incorporation of sporulation as a contextual variable, accession-based contextual adjustment, and biologically informed rules. On the test set, the model achieved Pearson's $r = 0.699$, CCC = 0.694, MAE = 17.17%, and RMSE = 24.75%, attaining the best overall performance compared with linear regression, Random Forest, XGBoost, and SVR. Bland–Altman analysis showed a mean bias of -0.58 %, with 50.78% of predictions within ± 10 % and 69.43% within ± 20 %. The proposed system demonstrated technical feasibility as a supportive tool for the automated estimation of downy mildew foliar severity in quinoa under controlled experimental conditions, with potential applicability in assisted digital phenotyping.

Keywords: quinoa downy mildew, foliar severity, digital phenotyping, computer vision, interpretable estimation.

INTRODUCCIÓN

El mildiu veloso, causado por *Peronospora variabilis*, constituye una de las principales amenazas fitopatológicas para el cultivo de quinua (*Chenopodium quinoa* Willd.) en la región andina, especialmente bajo condiciones de alta humedad relativa y temperaturas moderadas (Colque-Little et al., 2021; Danielsen et al., 2003). La cuantificación de la severidad foliar es un componente central para la evaluación fenotípica, la selección de accesiones resistentes y el seguimiento epidemiológico de la enfermedad. Sin embargo, la estimación visual convencional y los enfoques basados en fotografía digital pueden presentar variabilidad entre observadores, diferencias asociadas a las condiciones de captura y limitaciones de reproducibilidad, afectando la comparabilidad de los resultados experimentales (Bock et al., 2010).

A diferencia de la abundante literatura orientada a la detección o clasificación categórica de enfermedades vegetales, los estudios que abordan la cuantificación continua de severidad foliar expresada como porcentaje de tejido afectado constituyen un cuerpo bibliográfico más restringido y metodológicamente heterogéneo. La estimación continua exige modelar relaciones no lineales entre coloración, textura, distribución espacial de los síntomas y proporción de tejido afectado, mientras que muchos enfoques recientes basados en aprendizaje profundo formulan el problema como clasificación binaria u ordinal, reduciendo la sensibilidad para diferenciar gradientes intermedios de daño (Barbedo, 2016; Bock et al., 2020; Li et al., 2021). Aunque existen antecedentes de estimación automatizada de severidad foliar en otros patosistemas mediante aplicaciones móviles RGB, sensado proximal, imágenes multiespectrales UAV o segmentación profunda, estos trabajos difieren sustancialmente en plataforma de captura, grado de automatización, escala de anotación, métricas reportadas y condiciones experimentales, lo que limita su comparabilidad directa (Chemura et al., 2018; Duarte-Carvajalino et al., 2018; Gao et al., 2021; Goncalves et al., 2021; Pethybridge & Nelson, 2015).

Aunque existen antecedentes de estimación automatizada de severidad foliar en otros patosistemas, la evidencia disponible para el mildiu en quinua se concentra principalmente en evaluaciones visuales, escalas ordinales, estimaciones agronómicas de severidad y estudios de resistencia genética o comportamiento varietal. En contraste, son escasos los trabajos que formulen la severidad foliar del mildiu en quinua como una variable continua de 0 % a 100 % estimada mediante modelos computacionales supervisados. Esta brecha metodológica es relevante porque la cuantificación continua exige no solo imágenes foliares trazables, sino también anotaciones expertas en escala porcentual, condiciones de captura consistentes y un modelo capaz de preservar gradientes intermedios de daño.

En este marco, QuinuApp se propone como una prueba de concepto experimental orientada a la estimación

continua, interpretable y computacionalmente asistida de la severidad foliar en el patosistema mildiu–quinua. Frente a esta brecha, se propone QuinuApp, un sistema móvil orientado a apoyar la cuantificación automatizada de la severidad foliar del mildiu en quinua bajo condiciones de captura controlada. El sistema integra una aplicación móvil, un diseño de dispositivo de adquisición estandarizada denominado BlackBox y un modelo predictivo basado en variables visuales interpretables. A diferencia de enfoques exclusivamente dependientes de representaciones latentes, la propuesta utiliza descriptores asociados a manifestaciones visibles de la enfermedad, como cambios cromáticos, patrones texturales y rasgos morfológicos de la hoja. Esta orientación busca favorecer la trazabilidad de las predicciones y su potencial aceptación en contextos aplicados, donde la confianza del usuario depende no solo del desempeño numérico, sino también de la posibilidad de relacionar la estimación con evidencias visuales observables (Kamilaris & Prenafeta-Boldu, 2018).

El presente estudio se diferencia de una investigación metodológica previa centrada en la estimación interpretable de severidad foliar continua en quinua mediante visión computacional, desarrollada sobre un subconjunto de imágenes y sin integración móvil. A diferencia de ese trabajo, este manuscrito evalúa un sistema ampliado denominado QuinuApp, incorpora un conjunto de datos mayor, variables contextuales de esporulación y accesión, comparación con modelos de referencia, análisis por rangos de severidad, validación complementaria por accesión y una arquitectura propuesta de despliegue móvil con BlackBox.

El objetivo de este estudio fue evaluar experimentalmente el desempeño de QuinuApp como sistema de apoyo para la cuantificación continua de la severidad foliar del mildiu en quinua. Para ello, se analizó un flujo de trabajo que combina procesamiento de imágenes, extracción de variables visuales interpretables y evaluación comparativa frente a modelos de referencia, considerando el desbalance natural de los datos entre fases de severidad.

METODOLOGÍA

La investigación se desarrolló bajo un enfoque cuantitativo y un diseño experimental aplicado, orientado al desarrollo y evaluación de un sistema computacional para estimar de forma continua la severidad foliar del mildiu en quinua (*Chenopodium quinoa* Willd.). El sistema integró adquisición estandarizada de imágenes, procesamiento computacional, modelado supervisado e inferencia móvil local bajo condiciones controladas.

Población y muestra

El material biológico provino de un ensayo controlado realizado en los invernaderos de la Universidad de Copenhague, Dinamarca, con 600 plantas de quinua correspondientes a 258 accesiones genéticas, sometidas a infección por *Peronospora variabilis* bajo condiciones reguladas de temperatura, humedad

y fotoperiodo. A partir del material generado, se conformó un conjunto final de 772 imágenes RGB de hojas adaxiales individuales, procedentes de 92 accesiones, cada una asociada a una anotación experta de severidad en escala continua de 0 % a 100 %, una accesión de pertenencia y un registro de esporulación. Los criterios de inclusión fueron: calidad visual suficiente para segmentación, correspondencia válida entre imagen y anotación, y cobertura de distintos niveles de severidad foliar.

Para fines analíticos, la severidad se organizó en tres fases biológicas: fase A o infección temprana, de 0–25 %, con 133 imágenes; fase B o expansión moderada, de 26–75 %, con 195 imágenes; y fase C o necrosis terminal, de 76–100 %, con 444 imágenes. Esta distribución evidenció un desbalance estructural hacia niveles altos de severidad.

El conjunto de datos se dividió mediante partición estratificada a nivel de imagen en 579 imágenes para entrenamiento y 193 para prueba. Dado que imágenes de una misma accesión podían estar presentes en ambos subconjuntos, los resultados se interpretan como desempeño bajo partición por imagen y no como validación estricta de generalización hacia accesiones no observadas.

Herramientas tecnológicas utilizadas

El pipeline fue desarrollado en Python 3.x, empleando OpenCV para procesamiento de imágenes, NumPy y pandas para manipulación de datos, scikit-learn para modelado supervisado, calibración y evaluación, SciPy para funciones estadísticas, joblib para serialización, matplotlib para visualización y XGBoost como uno de los modelos comparativos. El modelo entrenado fue exportado e integrado en una aplicación móvil Android denominada QuinuApp, diseñada para operar localmente sin conexión a internet.

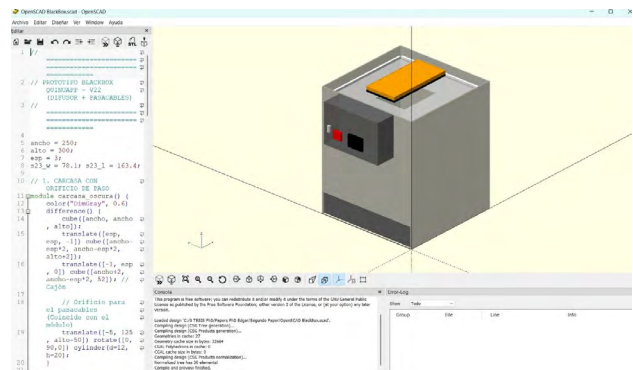
Para la fase de despliegue operativo se diseñó una arquitectura móvil compuesta por la aplicación Android QuinuApp, un teléfono Samsung Galaxy S23 Ultra como plataforma prevista de ejecución local y un dispositivo físico de captura estandarizada denominado BlackBox. Este dispositivo fue modelado paramétricamente en OpenSCAD e incorporó, a nivel de diseño, una carcasa opaca, iluminación perimetral, difusor interno de posicionamiento y soporte fijo para el teléfono móvil, como se muestra en la Figura 1.

La infraestructura fue concebida para reducir la variabilidad asociada a iluminación, fondo, distancia y ángulo de captura durante futuras adquisiciones controladas.

No obstante, en el presente estudio, la evaluación experimental se limitó al desempeño del pipeline computacional sobre el conjunto de imágenes disponible; por tanto, no se realizó una validación de campo del sistema integrado teléfono–BlackBox ni una evaluación operativa del algoritmo ejecutándose directamente en el dispositivo móvil.

Figura 1

Diseño paramétrico de la BlackBox propuesta para la implementación operativa de QuinuApp en condiciones de captura controlada



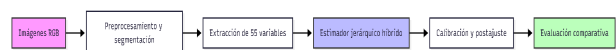
Nota. La figura muestra el diseño paramétrico de la BlackBox desarrollado en OpenSCAD como propuesta de infraestructura para captura controlada. Se representan la carcasa opaca, el sistema de iluminación perimetral superior, el difusor interno con abertura central, el módulo lateral de control, el cajón extraíble para posicionamiento de la muestra y el soporte superior para el teléfono móvil. El diseño fue concebido para estandarizar futuras capturas de imágenes y reducir la variabilidad asociada a iluminación, fondo, distancia y ángulo. En este estudio, la BlackBox se reporta como diseño tecnológico propuesto, no como dispositivo validado en campo.

Proceso metodológico

El flujo metodológico del sistema se estructuró en etapas secuenciales que abarcan la preparación del conjunto de datos, el preprocesamiento y segmentación de las imágenes, la extracción de variables visuales interpretables, el modelado predictivo, la calibración de las estimaciones y la evaluación comparativa frente a modelos de referencia. Una versión sintética de este flujo se presenta en la Figura 2.

Figura 2

Flujo metodológico del sistema propuesto para la estimación automatizada de severidad foliar en quinua



Nota. El diagrama sintetiza las etapas principales del pipeline computacional: entrada de imágenes RGB, preprocesamiento fotométrico, segmentación de la hoja principal, extracción de variables clínico-texturales, estimación predictiva, calibración y evaluación comparativa.

Procesamiento de imágenes y extracción de variables

Cada imagen fue sometida a normalización fotométrica mediante balance de blancos Gray-World y equalización adaptativa del contraste sobre el canal L^* del espacio CIELAB. Posteriormente, la hoja principal fue segmentada mediante una estrategia híbrida basada en exclusión de fondo, umbralización en espacios de color y selección geométrica del componente más próximo al centro de la imagen. En casos de segmentación inestable, se aplicó un procedimiento de respaldo basado en GrabCut. Sobre la región foliar segmentada se extrajeron 55 variables derivadas de imagen, agrupadas en cuatro bloques, fracciones de área sintomática, descriptores morfológicos, variables

cromáticas absolutas y descriptores texturales. Las fracciones sintomáticas representaron manifestaciones visibles de clorosis, necrosis, blanqueamiento o esporulación visible y necrosis marrón difusa. Los descriptores morfológicos resumieron propiedades espaciales de los síntomas; las variables cromáticas capturaron diferencias de color asociadas al daño foliar; y los descriptores texturales se calcularon mediante histogramas de Patrones Binarios Locales sobre la región segmentada.

Modelado predictivo

La severidad foliar fue tratada como una variable continua. No obstante, el modelo incorporó una discretización auxiliar en tres fases biológicas para guiar la especialización del estimador, infección temprana, expansión moderada y necrosis terminal. El núcleo predictivo se estructuró como un estimador jerárquico híbrido compuesto por un clasificador auxiliar de fase, regresores especializados por intervalo de severidad y un regresor global de respaldo. El clasificador auxiliar estimó la probabilidad de pertenencia de cada imagen a cada fase biológica. Estas probabilidades se utilizaron como mecanismo de ponderación para combinar las predicciones de los regresores especializados.

En paralelo, un regresor global estimó la severidad sobre todo el rango continuo de respuesta. La predicción final integró la salida local especializada y la salida global mediante una mezcla dinámica dependiente de la confianza del clasificador. Todas las variables numéricas fueron estandarizadas utilizando exclusivamente los datos de entrenamiento. Para compensar el desbalance de la muestra, se aplicaron pesos diferenciados durante el entrenamiento, asignando mayor relevancia a fases y rangos de severidad menos representados. Las predicciones fueron refinadas mediante calibración isotónica, ajuste contextual por accesión y reglas biológicas de postajuste. El ajuste por accesión se utilizó como información contextual histórica; por tanto, sus resultados no deben interpretarse como inferencia basada exclusivamente en la imagen.

Validación y modelos de referencia

El desempeño del sistema se evaluó sobre el conjunto de prueba mediante coeficiente de correlación de Pearson, coeficiente de concordancia de Lin, error absoluto medio, raíz del error cuadrático medio y proporción de predicciones dentro de ± 10 y ± 20 puntos porcentuales respecto al valor observado. La concordancia entre severidad observada y estimada se examinó adicionalmente mediante análisis de Bland–Altman, reportando sesgo sistemático y límites de acuerdo. El sistema propuesto se comparó con cuatro modelos de referencia entrenados sobre la misma matriz de predictores: regresión lineal múltiple, Random Forest, XGBoost y SVR con kernel RBF. Todos los modelos fueron evaluados bajo el mismo esquema de partición y con el mismo conjunto de métricas. Como referencia descriptiva adicional, se analizó la repetibilidad intraobservador en 24 imágenes independientes evaluadas en dos momentos por una especialista en fitopatología. Este análisis se consideró únicamente como evidencia de variabilidad humana en la estimación visual, no como comparación directa entre desempeño humano y algorítmico.

RESULTADOS

Desempeño comparativo del pipeline híbrido

El pipeline híbrido QuinuApp mostró el mejor desempeño global entre los modelos evaluados, al combinar el menor error absoluto medio, la mayor concordancia de Lin, el menor sesgo promedio y la mayor proporción de predicciones dentro de los márgenes de tolerancia de ± 10 % y ± 20 % sobre el conjunto de prueba. La Tabla 1 resume el desempeño comparativo del pipeline híbrido y de los modelos de referencia. QuinuApp alcanzó un MAE de 17.17 %, un CCC de 0.694 y un sesgo medio de -0.58 % en el análisis de Bland–Altman. Además, obtuvo el mayor porcentaje de predicciones dentro de ± 10 puntos porcentuales respecto al valor observado, con 50.78 %, y dentro de ± 20 puntos porcentuales, con 69.43 %. Estos resultados indican que el pipeline propuesto presentó la combinación más balanceada entre error absoluto, concordancia y bajo sesgo global.

Tabla 1

Comparación de las métricas de desempeño del pipeline híbrido QuinuApp y modelos de referencia en el conjunto de prueba

Modelo	Pearson r	CCC	MAE (%)	RMSE (%)	Sesgo BA (%)	Dentro $\pm 10\%$	Dentro $\pm 20\%$
Pipeline Híbrido	0.699	0.694	17.17	24.75	-0.58	50.78	69.43
Random Forest	0.707	0.561	19.34	24.71	0.25	33.16	63.73
XGBoost	0.692	0.614	17.95	24.34	1.52	41.97	65.80
SVR (RBF)	0.638	0.505	18.87	26.76	5.93	45.08	66.32
Regresión Lineal	0.622	0.568	19.56	26.21	0.26	40.93	61.14

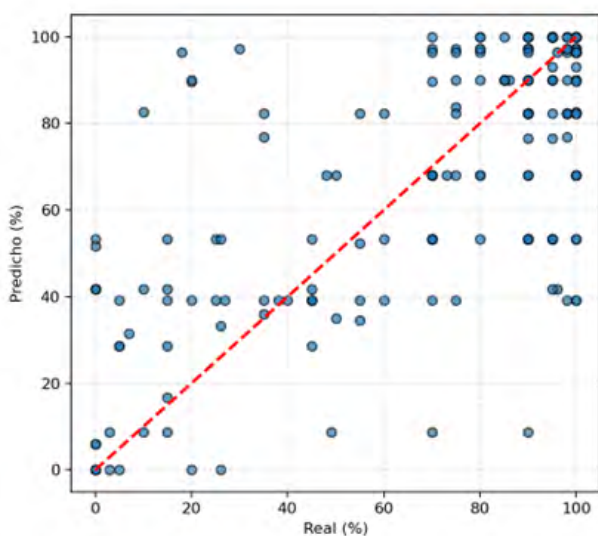
Nota. CCC = coeficiente de concordancia y correlación de Lin; MAE = error absoluto medio; RMSE = raíz del error cuadrático medio; Sesgo BA = sesgo medio estimado mediante el análisis de Bland–Altman; Dentro ± 10 % y Dentro ± 20 % representan la proporción de predicciones cuya diferencia absoluta respecto al valor observado fue ≤ 10 y ≤ 20 puntos porcentuales.

Aunque Random Forest obtuvo el mayor coeficiente de correlación de Pearson ($r = 0.707$), su concordancia global fue menor que la de QuinuApp (CCC = 0.561 frente a 0.694) y presentó una proporción sustancialmente inferior de predicciones dentro de $\pm 10\%$. XGBoost alcanzó el menor RMSE (24.34 %) y el segundo mejor MAE (17.95 %), pero no superó al pipeline híbrido en concordancia, sesgo ni proporción de predicciones dentro de los márgenes de tolerancia. En conjunto, estos resultados muestran que el desempeño de QuinuApp no depende de una única métrica aislada, sino de un equilibrio más consistente entre precisión, concordancia y estabilidad global.

La relación entre severidad observada y severidad predicha se presenta en la Figura 3. La distribución de los puntos muestra una asociación positiva general entre ambos valores, aunque con dispersión apreciable respecto de la diagonal de identidad. Esta dispersión fue más evidente en rangos intermedios de severidad, donde la transición entre fases biológicas y la superposición visual de síntomas dificultan la estimación precisa. Por tanto, el modelo preservó la tendencia global de la variable objetivo, pero mantuvo variabilidad residual a nivel de observaciones individuales.

Figura 3

Severidad observada frente a severidad predicha por el pipeline híbrido QuinuApp



Nota. La figura muestra la relación entre la severidad foliar observada y la severidad estimada por el pipeline híbrido QuinuApp sobre el conjunto de prueba. La línea discontinua representa la identidad $y = x$, correspondiente a concordancia perfecta entre valores observados y estimados. La dispersión de los puntos respecto a esta referencia permite visualizar la magnitud del error de predicción a lo largo del rango de severidad.

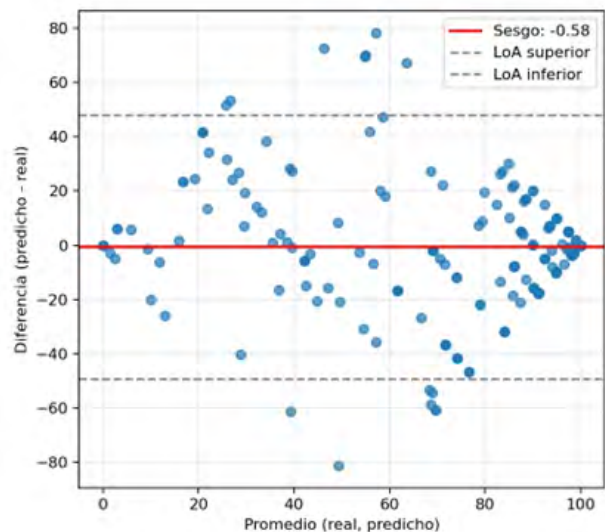
Análisis de concordancia Bland-Altman

El análisis de Bland-Altman permitió examinar la estructura del error del pipeline híbrido en función de la magnitud promedio de la severidad foliar (Bland & Altman, 1986). Como se observa en la Figura 4, el modelo presentó un sesgo medio de -0.58% , lo que indica ausencia de sobreestimación o subestimación sistemática relevante a nivel global. Sin embargo, los límites de acuerdo fueron amplios, entre -49.20% y

48.04% , evidenciando variabilidad considerable en la predicción individual.

Figura 4

Análisis de Bland-Altman entre severidad observada y severidad estimada por el pipeline híbrido QuinuApp



Nota. La figura presenta el análisis de Bland-Altman correspondiente a las predicciones del pipeline híbrido QuinuApp sobre el conjunto de prueba ($n = 193$). El eje horizontal representa el promedio entre severidad observada y severidad estimada, mientras que el eje vertical muestra la diferencia entre ambas. La línea central indica el sesgo medio (-0.58%) y las líneas discontinuas superiores e inferiores representan los límites de acuerdo (-49.20% a 48.04%).

La lectura conjunta de las Figuras 3 y 4 confirma que el pipeline híbrido reproduce adecuadamente la tendencia general de la severidad foliar y mantiene bajo sesgo global. No obstante, la amplitud de los límites de acuerdo indica que la precisión individual sigue siendo variable, especialmente en imágenes con severidad intermedia o con manifestaciones sintomáticas visualmente heterogéneas.

Este comportamiento es coherente con la complejidad de cuantificar severidad foliar continua a partir de imágenes RGB, donde la morfología de la hoja, la distribución espacial de los síntomas y la subjetividad de la anotación experta pueden introducir variabilidad adicional.

Una ampliación del dataset con mayor número de muestras reales y mejor equilibrio entre rangos podría mejorar la estabilidad del aprendizaje y reducir la incertidumbre en los intervalos menos representados, aunque dicha mejora debe verificarse experimentalmente y no asumirse como consecuencia automática del balanceo.

En síntesis, QuinuApp alcanzó el desempeño global más balanceado entre los modelos evaluados, con menor MAE, mayor CCC, bajo sesgo promedio y mayor proporción de predicciones dentro de márgenes de error clínicamente interpretables. Sin embargo, los resultados también evidencian que, aunque el modelo presenta bajo sesgo global, persiste variabilidad en la predicción individual, particularmente en rangos intermedios de severidad.

Desempeño por rango de severidad

El análisis estratificado por rangos mostró que el desempeño del pipeline híbrido QuinuApp no fue homogéneo a lo largo de la escala de severidad. Como se observa en la Tabla 2, el sistema presentó mejor comportamiento en el rango alto de severidad, con el menor MAE (14.15 %), el menor RMSE (21.44 %) y la mayor

proporción de predicciones dentro de ± 10 % (60.36 %) y ± 20 % (80.18 %). En contraste, el rango bajo de severidad mostró el desempeño más débil, con MAE = 26.60 %, RMSE = 35.12 % y solo 33.33 % de predicciones dentro de ± 10 %. El rango medio presentó un comportamiento intermedio, con MAE = 17.68 %, RMSE = 23.25 % y CCC = 0.401.

Tabla 2

Desempeño del pipeline híbrido QuinuApp por rango de severidad

Rango de severidad	n	MAE (%)	RMSE (%)	CCC	Dentro ± 10 (%)	Dentro ± 20 (%)	Dentro ± 20 %
Bajo, 0–25 %	33	26.603	35.115	0.138	33.333	45.455	69.43
Medio, 26–75 %	49	17.676	23.247	0.401	40.816	61.224	66.32
Alto, 76–100 %	111	14.150	21.440	0.014	40.816	80.180	61.14

Nota. La columna n indica el número de imágenes del conjunto de prueba incluidas en cada rango, definido según la severidad observada. Los rangos fueron establecidos como: bajo = 0–25 %, medio = 26–75 % y alto = 76–100 %. MAE = error absoluto medio; RMSE = raíz del error cuadrático medio; CCC = coeficiente de concordancia de Lin. Dado que el tamaño muestral difiere entre rangos especialmente entre el rango bajo (n = 33) y el rango alto (n = 111), las métricas deben interpretarse considerando la representatividad desigual de cada intervalo.

Estos resultados indican que la utilidad inmediata del sistema se concentra principalmente en escenarios de severidad alta, donde la presencia extensa de síntomas facilita la estimación del daño foliar. En cambio, el desempeño más débil en el rango de 0–25 % MAE = 26.60 % y RMSE = 35.12 % sugiere que el pipeline todavía presenta limitada sensibilidad para capturar manifestaciones tempranas de la enfermedad.

Este comportamiento debe interpretarse considerando el desbalance estructural del conjunto de datos, una condición frecuente en estudios fitopatológicos y biológicos, donde la progresión natural de la enfermedad no produce necesariamente una distribución uniforme de observaciones entre fases. En este estudio, el conjunto de prueba incluyó 33 imágenes en severidad baja, 49 en severidad media y 111 en severidad alta, lo que favorece una estimación más estable en el rango de daño avanzado y limita la robustez estadística de los rangos menos representados.

El rango medio presentó un comportamiento intermedio, con MAE = 17.68 %, RMSE = 23.25 % y CCC = 0.401, coherente con la dificultad de modelar transiciones graduales entre clorosis y esporulación visible. Debe señalarse que el CCC bajo en el rango alto no contradice el menor MAE observado en ese intervalo. Al restringir el análisis a severidades elevadas, la variabilidad interna de los valores observados puede disminuir, reduciendo la concordancia de Lin aun cuando el error absoluto sea menor.

En conjunto, el análisis por rangos evidencia que QuinuApp no debe interpretarse todavía como un estimador homogéneo en toda la escala de severidad, sino como un sistema con desempeño más estable

en daño avanzado y con necesidad de refinamiento específico para infecciones tempranas y transiciones intermedias.

Validación complementaria por accesión

Como análisis adicional de robustez metodológica, se repitió la evaluación del pipeline híbrido mediante partición estricta por accesión, en un esquema de tres conjuntos (entrenamiento / validación / prueba) con proporciones aproximadas 60/15/25. Bajo este diseño, las 92 accesiones se asignaron de forma exclusiva a una única partición (55 train, 14 val, 23 test), sin accesiones compartidas entre subconjuntos.

La estratificación se realizó sobre la clase de severidad dominante de cada accesión para garantizar representación de las tres fases biológicas en cada partición. El conjunto de validación cumplió una función específica: ajustar la calibración isotónica y los priors contextuales sin exponer el conjunto de prueba a ningún componente del pipeline.

Los resultados evidenciaron una degradación sustantiva del desempeño respecto a la partición por imagen Tabla 3. El MAE aumentó de 17.17% a 23.01% (+5.84 pp), el CCC descendió de 0.694 a 0.593, y las tasas de acuerdo dentro de ± 10 % y ± 20 % se redujeron a 25.82% y 42.86%, respectivamente, aproximadamente la mitad de los valores obtenidos con el esquema por imagen.

El sesgo Bland–Altman pasó de -0.58 pp a -9.66 pp, indicando una subestimación sistemática del modelo cuando las accesiones de prueba carecen de información contextual previa.

Tabla 3

Comparación entre partición por imagen y partición por accesión

Esquema de partición	Pearson r	CCC	MAE (%)	RMSE (%)	Sesgo BA (%)	Dentro $\pm 10\%$ (%)	Dentro $\pm 20\%$ (%)
Por imagen	0.699	0.694	17.17	24.75	-0.58	50.78	69.43
Por accesión	0.647	0.593	23.01	28.09	-9.66	25.82	42.86

Nota. MAE = error absoluto medio; RMSE = raíz del error cuadrático medio; CCC = coeficiente de concordancia de Lin; Sesgo BA = sesgo de Bland–Altman (pp); $\pm 10\%$ y $\pm 20\%$ = porcentaje de predicciones dentro de 10 y 20 puntos porcentuales del valor real. La partición por imagen emplea muestreo estratificado por clase (75/25, semilla 42). La partición por accesión asigna cada accesión exclusivamente a un conjunto (train 60 % / val 15 % / test 25 %), estratificando por clase dominante; 55 accesiones en entrenamiento, 14 en validación y 23 en prueba.

El análisis por rango de severidad en la validación complementaria por accesión reveló un patrón de sesgo asimétrico. En el rango bajo, de 0–25 %, el modelo tendió a sobreestimar la severidad, con un sesgo medio de +26.54 puntos porcentuales (n = 33).

En el rango medio, de 26–75 %, presentó una subestimación moderada, con un sesgo medio de -9.77 puntos porcentuales (n = 46; MAE = 18.65 %). En el rango alto, de 76–100 %, la subestimación fue más marcada, con un sesgo medio de -21.21 puntos porcentuales (n = 103). Este patrón sugiere una regresión de las predicciones hacia valores intermedios cuando el modelo se evalúa sobre accesiones no observadas durante el entrenamiento.

A diferencia de la partición principal por imagen, cuyo conjunto de prueba estuvo conformado por 193 imágenes, la validación por accesión generó un conjunto de prueba de 182 imágenes, debido a que la asignación se realizó por accesiones completas y no por imágenes individuales. En este escenario, las predicciones finales se concentraron en un número reducido de salidas, se registraron solo 14 valores predichos únicos en las 182 imágenes de prueba.

Este comportamiento evidencia una pérdida de resolución continua bajo el esquema inter-accesión y puede estar asociado a la ausencia de priors específicos para las accesiones de prueba reduce el efecto estabilizador del ajuste contextual por accesión.

DISCUSIÓN

El principal hallazgo de este estudio fue que el pipeline híbrido propuesto alcanzó el desempeño global más balanceado entre los modelos evaluados, registrando simultáneamente el menor error absoluto medio (MAE = 17.17 %), el mayor coeficiente de concordancia y correlación de Lin (CCC = 0.694) y el sesgo medio más cercano a cero en el análisis de Bland–Altman (-0.58 %).

En relación con el objetivo del estudio, estos hallazgos indican que la combinación de adquisición estandarizada, representación interpretable de síntomas y estimación jerárquica híbrida constituye una aproximación metodológicamente viable para apoyar la cuantificación continua de severidad foliar en quinua

bajo condiciones controladas.

Este desempeño (MAE = 17.17 %; CCC = 0.694) sugiere, además, que la especialización local por subrangos de la variable de salida puede mejorar la estabilidad de la predicción frente a distribuciones desbalanceadas y relaciones no lineales entre atributos y respuesta, en consonancia con principios de aprendizaje supervisado híbrido reportados en la literatura.

La relación observada entre severidad real y severidad predicha confirma que el sistema conserva la tendencia general de la variable objetivo, aunque con dispersión apreciable en distintos intervalos del rango. Esta característica se refuerza con el análisis de Bland–Altman, donde el sesgo global bajo contrasta con límites de acuerdo amplios. En términos prácticos, ello sugiere que el sistema reduce la desviación sistemática promedio, pero todavía mantiene variabilidad residual a nivel de observaciones individuales.

Este resultado es consistente con la complejidad del problema abordado, en el que convergen heterogeneidad morfológica entre accesiones, transición gradual entre fases biológicas y solapamiento visual entre síntomas. En consecuencia, QuinuApp debe interpretarse como una prueba de concepto experimental con desempeño prometedor en dominio controlado, pero aún insuficiente para uso autónomo en campo abierto sin validación externa adicional.

Comparación con antecedentes de cuantificación continua de severidad

Para contextualizar el desempeño relativo de QuinuApp, la Tabla 4 sintetiza antecedentes seleccionados de cuantificación continua o porcentual de severidad foliar mediante visión computacional o sensado proximal.

La comparación debe interpretarse con cautela, debido a diferencias entre cultivos, patosistemas, sensores, protocolos de captura, escalas de anotación, particiones experimentales y métricas reportadas.

En particular, no todos los estudios informan simultáneamente MAE, RMSE y CCC, por lo que la tabla no constituye un ranking directo de desempeño, sino una referencia metodológica para ubicar la contribución del presente trabajo.

Tabla 4

Antecedentes seleccionados de cuantificación continua de severidad foliar mediante visión computacional o sensado proximal en patosistemas comparables

Studio	Patosistema	Plataforma / sensor	Método	Métricas reportadas	Limitaciones principales
Pethybridge & Nelson (2015)	Seis enfermedades foliares	Smartphone, RGB	Umbralización de color asistida por usuario (Leaf Doctor)	$R^2 \geq 0.79$; Cb ≥ 0.959	Dependencia del usuario para selección de píxeles; sin inferencia automática completa
Chemura et al. (2018)	Roya del café	Espectrorradiómetro portátil	RBF-PLS sobre bandas espectrales	$R^2 = 0.92$; RMSE = 6.1 %	Condiciones controladas; tamaño muestral limitado; sensor especializado
Duarte-Carvajalino et al. (2018)	Tizón tardío de papa	UAV multiespectral	MLP, SVR, Random Forest, CNN	MAE = 13.63–19.00 %; RMSE = 18.00–28.03 %; $R^2 = 0.28–0.80$	Alta variabilidad entre modelos; requiere UAV y captura multiespectral
Gao et al. (2021)	Tizón tardío de papa	RGB de campo	Encoder–decoder para segmentación semántica	$R^2 = 0.655$; IoU lesión = 0.386	Dificultad en lesiones pequeñas; severidad evaluada en rango limitado
Gonçalves et al. (2021)	Tres patosistemas foliares	RGB, hojas individuales	FPN, U-Net, DeepLabv3+	Alto acuerdo con referencia; R^2 alto según arquitectura	Sobreestimación por clasificación errónea de píxeles sanos

Nota. La comparación realizada es contextual y no representa un ranking directo, debido a diferencias entre estudios en cultivo, patosistema, sensores, protocolos de captura, escalas de anotación, particiones experimentales y métricas reportadas. Además, no todos los trabajos informan simultáneamente MAE, RMSE y CCC. La inclusión de patosistemas distintos al mildiu en quinua responde a la escasa disponibilidad de estudios comparables sobre cuantificación continua de severidad foliar en este cultivo, situación asociada a la complejidad de generar bases de datos anotadas, condiciones controladas de captura e inoculación y disponibilidad de especialistas en fitopatología.

La inclusión de estudios en papa, café, trigo y otros patosistemas foliares responde a la limitada disponibilidad de antecedentes directamente comparables sobre cuantificación continua del mildiu en quinua. Esta limitación no es solo bibliográfica, sino también experimental y metodológica. Como señalan (Faye et al., 2023), la estimación automatizada de severidad vegetal constituye un campo heterogéneo, en el que coexisten enfoques basados en procesamiento de imágenes, aprendizaje automático clásico y aprendizaje profundo, con diferencias sustanciales en sensores, escalas de severidad, protocolos de captura y métricas de evaluación. Además, los autores identifican desafíos persistentes relacionados con la calidad de imagen, la segmentación de síntomas pequeños o tempranos y la dependencia de etiquetas generadas por especialistas. En este contexto, generar un conjunto de datos para severidad foliar continua en quinua exige ensayos controlados, inoculación o seguimiento fitopatológico consistente, imágenes con trazabilidad experimental y anotaciones expertas en escala porcentual. Por ello, la comparación presentada debe interpretarse como contextual y no como equivalencia directa de desempeño.

El aporte del presente estudio reside en proporcionar una base inicial, reproducible e interpretable para un patosistema aún poco representado en la literatura computacional. Frente a estos antecedentes, QuinuApp alcanzó MAE = 17.17 %, RMSE = 24.75 % y CCC = 0.694 bajo partición por imagen. Estos valores muestran una

viabilidad técnica inicial, pero también evidencian que el sistema todavía presenta error individual relevante, especialmente si se compara con estudios desarrollados bajo condiciones más controladas, con sensores especializados o con protocolos de segmentación asistida. Por tanto, la contribución principal del presente estudio no radica en afirmar superioridad métrica general frente a todos los antecedentes, sino en proponer una arquitectura específica para el patosistema mildiu–quinua, basada en variables clínico-texturales interpretables, modelado jerárquico híbrido, calibración postpredictiva y una propuesta de captura estandarizada. En quinua, la evaluación del mildiu causado por *Peronospora variabilis* se ha abordado principalmente desde enfoques fitopatológicos y agronómicos, mediante estimaciones visuales, escalas discretas de severidad, categorías ordinales o mediciones orientadas a caracterizar resistencia y comportamiento varietal. Este tipo de evaluación ha sido utilizado para describir germoplasma, comparar líneas de mejoramiento y estudiar la respuesta de accesiones frente al patógeno (Colque-Little et al., 2021; Danielsen et al., 2003; Stanschewski et al., 2021). En este contexto, el presente trabajo contribuye a trasladar la cuantificación de severidad hacia un enfoque computacional continuo, trazable y orientado al apoyo del fenotipado digital.

La orientación hacia variables interpretables constituye un componente metodológico relevante del sistema. La literatura sobre inteligencia artificial en agricultura

ha señalado que, aunque los modelos basados en aprendizaje profundo pueden alcanzar alto desempeño predictivo, su limitada transparencia puede dificultar la trazabilidad y la confianza en contextos aplicados (Kamilaris & Prenafeta-Boldu, 2018). En este sentido, el uso de fracciones sintomáticas, descriptores morfológicos, variables cromáticas y patrones texturales ofrece una ventaja operativa, ya que permite relacionar la predicción con manifestaciones visibles de la hoja y no únicamente con representaciones latentes de difícil interpretación. No obstante, esta característica debe entenderse como una forma de interpretabilidad basada en variables de entrada, no como una explicación causal completa del proceso patológico ni de la decisión del modelo. Estudios recientes evidencian una tendencia hacia sistemas móviles de fenotipado digital y diagnóstico asistido por inteligencia artificial, como StripeRust-Pocket y GranoScan, orientados a la detección de amenazas en trigo. En este contexto, QuinuApp se alinea con el desarrollo de herramientas móviles para apoyo fitopatológico; sin embargo, su validación operativa en campo aún no ha sido completada. Por ello, el estudio debe interpretarse como una prueba de concepto experimental en condiciones controladas y no como una validación definitiva en escenarios reales.

Entre las limitaciones del estudio debe señalarse, en primer lugar, el desbalance estructural del conjunto de datos, con predominio de observaciones en la fase alta severidad. Este comportamiento es frecuente en estudios biológicos y fitopatológicos, donde la progresión de la enfermedad no necesariamente produce una distribución uniforme entre fases. Aunque el efecto fue parcialmente mitigado mediante ponderación adaptativa, calibración y modelado jerárquico, la distribución de la muestra continúa condicionando la precisión en los rangos menos representados, especialmente en severidades bajas e intermedias. En segundo lugar, la evaluación principal se realizó mediante partición estratificada a nivel de imagen. Esto implica que imágenes procedentes de una misma accesión podían estar representadas tanto en entrenamiento como en prueba; por tanto, los resultados de este esquema deben interpretarse como evidencia de desempeño in-domain y no como validación estricta de generalización hacia accesiones no observadas. Para abordar esta limitación, se incorporó una validación complementaria por accesión, en la cual cada accesión fue asignada exclusivamente a entrenamiento, validación o prueba. Esta evaluación más exigente evidenció una disminución del desempeño, lo que indica que la generalización intergenotípica del sistema todavía requiere fortalecimiento mediante mayor volumen de datos, mejor balance entre rangos de severidad y validación externa con accesiones independientes.

En tercer lugar, el modelo incorporó información contextual mediante un prior por accesión y la variable de esporulación dentro de la matriz de entrada. Estos componentes pueden contribuir a estabilizar la predicción cuando existe información

histórica disponible de la accesión; sin embargo, limitan la interpretación del desempeño como resultado derivado exclusivamente de atributos visuales extraídos de la imagen RGB. En consecuencia, QuinuApp debe entenderse como un sistema híbrido que combina información visual y contextual, no como un estimador puramente visual. Además, bajo partición estricta por accesión, las accesiones del conjunto de prueba no disponen de priors históricos calculados desde entrenamiento, lo que contribuye a explicar la degradación observada en ese esquema de validación.

En cuarto lugar, el entrenamiento y la evaluación se realizaron con imágenes procedentes de un experimento controlado, no de una campaña amplia de captura en campo abierto. Asimismo, aunque se diseñó una arquitectura móvil compuesta por QuinuApp, un teléfono Android y la BlackBox, el presente estudio no incluyó validación operativa del algoritmo ejecutándose directamente en el dispositivo móvil ni validación de campo del sistema integrado teléfono-BlackBox. Por ello, la BlackBox debe interpretarse como una infraestructura tecnológica propuesta para futuras capturas estandarizadas, no como un componente experimental ya validado. La robustez del sistema bajo condiciones no controladas, con variación natural de iluminación, fondo, distancia y ángulo, requiere evaluación específica. Finalmente, el experimento intraobservador se realizó sobre un conjunto independiente de imágenes, por lo que sus resultados constituyen una referencia descriptiva de variabilidad humana, pero no una comparación estrictamente equivalente entre desempeño humano y desempeño algorítmico. Por esta razón, el estudio no permite afirmar que el sistema iguale o supere a una especialista en fitopatología; esa comparación requeriría un diseño humano-máquina sobre el mismo conjunto de prueba. Las implicaciones del estudio son relevantes tanto en el ámbito académico como en el profesional. Desde el plano académico, el trabajo aporta una arquitectura metodológica que articula visión computacional, aprendizaje supervisado híbrido, variables interpretables y calibración postpredictiva en un problema de cuantificación continua. Desde el plano aplicado, el sistema ofrece una base tecnológica para apoyar tareas de fenotipado digital y evaluación comparativa de accesiones bajo condiciones controladas. Sin embargo, su adopción operativa debe considerarse todavía preliminar, dado que requiere validación externa, evaluación inter-accesión, ampliación del conjunto de datos y pruebas directas del sistema móvil completo.

CONCLUSIONES

El principal hallazgo de este estudio fue que, bajo la evaluación principal con partición estratificada por imagen, el pipeline híbrido QuinuApp obtuvo el desempeño global más balanceado entre los modelos de referencia evaluados, alcanzando el menor error absoluto medio (MAE = 17.17 %), el mayor coeficiente de concordancia de Lin (CCC = 0.694) y un sesgo sistemático cercano a cero (-0.58 %) bajo condiciones

experimentales controladas. Estos resultados indican que la combinación de variables visuales interpretables, modelado jerárquico híbrido y calibración postpredictiva constituye una aproximación técnicamente viable para apoyar la estimación continua de la severidad foliar del mildiu en quinua. En relación con los objetivos planteados, se logró evaluar un pipeline computacional basado en 55 variables clínico-texturales derivadas de imagen, complementadas con información contextual de esporulación, para estimar la severidad foliar en una escala continua de 0 % a 100 %. Asimismo, se diseñó una arquitectura tecnológica de despliegue compuesta por QuinuApp y una BlackBox modelada paramétricamente como propuesta de captura estandarizada. En este sentido, el objetivo general del estudio se considera cumplido dentro del dominio experimental evaluado, centrado en el desempeño del modelo sobre el conjunto de imágenes disponible.

No obstante, los resultados deben interpretarse como una prueba de concepto experimental bajo condiciones controladas y no como una validación operativa definitiva para campo abierto. Aunque la evaluación principal por imagen mostró un desempeño favorable, la validación por accesión evidenció una reducción del rendimiento frente a accesiones no observadas, indicando limitaciones en la generalización intergenotípica. Asimismo, el modelo presentó un MAE de 17.17 % y variabilidad residual en las predicciones individuales, especialmente en severidades bajas e intermedias. Como líneas futuras de investigación, se recomienda priorizar la validación externa en campo abierto con accesiones no observadas y conjuntos de captura más diversos. Además, resulta necesario ampliar y equilibrar el dataset, particularmente en fases de baja e intermedia severidad, evaluar la arquitectura QuinuApp–teléfono–BlackBox en condiciones reales de uso y reducir la dispersión residual del error sin comprometer la interpretabilidad del modelo. Estos resultados posicionan a QuinuApp como una propuesta metodológicamente trazable y técnicamente prometedora para la estimación continua de la severidad foliar del mildiu en quinua. Su principal aporte radica en demostrar la viabilidad inicial de una arquitectura interpretable para fenotipado digital asistido, aunque su consolidación como herramienta operativa dependerá de validaciones externas, reducción del error individual y evaluación robusta en escenarios reales de campo.

BIBLIOGRAFÍA

- Barbedo, J. G. A. (2016). A review on the main challenges in automatic plant disease identification based on visible range images. *Biosystems Engineering*, 144, 52–60. <https://doi.org/10.1016/j.biosystemseng.2016.01.017>
- Bland, J. M., & Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet*, 327(8476), 307–310. [https://doi.org/10.1016/S0140-6736\(86\)90837-8](https://doi.org/10.1016/S0140-6736(86)90837-8)
- Bock, C. H., Poole, G. H., Parker, P. E., & Gottwald, T. R. (2010). Plant disease severity estimated visually, by digital photography and image analysis, and by hyperspectral imaging. *Critical Reviews in Plant Sciences*, 29(2), 59–107. <https://doi.org/10.1080/07352681003617285>
- Bock, C. H., Barbedo, J. G. A., Del Ponte, E. M., Bohnenkamp, D., & Mahlein, A.-K. (2020). From visual estimates to fully automated sensor-based measurements of plant disease severity: Status and challenges for improving accuracy. *Phytopathology Research*, 2(1), Artículo 9. <https://doi.org/10.1186/s42483-020-00049-8>
- Colque-Little, C., Amby, D. B., & Andreasen, C. (2021). A review of *Chenopodium quinoa* (Willd.) diseases—An updated perspective. *Plants*, 10(6), 1228. <https://doi.org/10.3390/plants10061228>
- Chemura, A., Mutanga, O., Sibanda, M., & Chidoko, P. (2018). Machine learning prediction of coffee rust severity on leaves using spectroradiometer data. *Tropical Plant Pathology*, 43(2), 117–127. <https://doi.org/10.1007/s40858-017-0187-8>
- Danielsen, S., Bonifacio, A., & Ames, T. (2003). Diseases of quinoa *Chenopodium quinoa*. *Food Reviews International*, 19(1–2), 43–59. <https://doi.org/10.1081/FRI-120018867>
- Dainelli, R., Bruno, A., Martinelli, M., Moroni, D., Rocchi, L., Morelli, S., Ferrari, E., Silvestri, M., Agostinelli, S., La Cava, P., Di Maggio, A., Folco, C., Mori, N., Stevanato, P., & Toscano, P. (2024). GranoScan: An AI-powered mobile app for in-field identification of biotic threats of wheat. *Frontiers in Plant Science*, 15, Article 1298791. <https://doi.org/10.3389/fpls.2024.1298791>
- Duarte-Carvajalino, J. M., Alzate, D. F., Ramirez, A. A., Santa-Sepulveda, J. D., Fajardo-Rojas, A. E., & Soto-Suárez, M. (2018). Evaluating late blight severity in potato crops using unmanned aerial vehicles and machine learning algorithms. *Remote Sensing*, 10(10), Article 1513. <https://doi.org/10.3390/rs10101513>
- Fondevilla, S., Calderón-González, Á., Rojas-Panadero, B., Cruz, V., & Matías, J. (2024). Genome-wide association study, combined with bulk segregant analysis, identify plant receptors and defense related genes as candidate genes for downy mildew resistance in quinoa. *BMC Plant Biology*, 24(1), 594. <https://doi.org/10.1186/s12870-024-05302-2>
- Faye, D., Diop, I., Mbaye, N., Dione, D., & Diedhiou, M. M. (2023). Plant disease severity assessment based on machine learning and deep learning: A survey. *Journal of Computer and Communications*, 11(9), 57–75. <https://doi.org/10.4236/jcc.2023.119004>
- Gao, J., Westergaard, J. C., Sundmark, E. H. R., Bagge, M., Liljeroth, E., & Alexandersson, E. (2021). Automatic late blight lesion recognition and severity quantification based on field imagery of diverse potato genotypes by deep learning. *Knowledge-Based Systems*, 214, Article 106723. <https://doi.org/10.1016/j.knsys.2020.106723>
- Gonçalves, J. P., Pinto, F. A. C., Queiroz, D. M., Villar, F. M. M., Barbedo, J. G. A., & Del Ponte, E. M. (2021). Deep learning architectures for semantic segmentation and automatic estimation of severity of foliar symptoms caused by diseases or pests. *Biosystems Engineering*, 210, 129–142. <https://doi.org/10.1016/j.biosystemseng.2021.08.011>
- Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture: A survey. *Computers and Electronics in Agriculture*, 147, 70–90. <https://doi.org/10.1016/j.compag.2018.02.016>
- Li, L., Zhang, S., & Wang, B. (2021). Plant disease detection and classification by deep learning—A review. *IEEE Access*, 9, 56683–56698. <https://doi.org/10.1109/ACCESS.2021.3069646>
- Liu, W., Chen, Y., Lu, Z., Lu, X., Wu, Z., Zheng, Z., Suo, Y., Lan, C., & Yuan, X. (2024). StripeRust-Pocket: A mobile-based deep learning application for efficient disease severity assessment of wheat stripe rust. *Plant Phenomics*, 6, Article 0201. <https://doi.org/10.34133/plantphenomics.0201>
- Pethybridge, S. J., & Nelson, S. C. (2015). Leaf Doctor: A new portable application for quantifying plant disease severity. *Plant Disease*, 99(10), 1310–1316. <https://doi.org/10.1094/PDIS-03-15-0319-RE>
- Stanschewski, C. S., Rey, E., Fiene, G., Wellman, G., Melino, V. J., & Tester, M. (2021). Quinoa phenotyping methodologies: An international consensus. *Plants*, 10(9), Article 1759. <https://doi.org/10.3390/plants10091759>